



Deep learning for carbonate rocks petrography

Axel Ransinangue

► To cite this version:

Axel Ransinangue. Deep learning for carbonate rocks petrography. Environmental Sciences. Université de Bordeaux, 2025. English. ⟨NNT : 2025BORD0262⟩. ⟨tel-05523669⟩

HAL Id: tel-05523669

<https://theses.hal.science/tel-05523669v1>

Submitted on 23 Feb 2026

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



HAL Authorization

THÈSE PRÉSENTÉE
POUR OBTENIR LE GRADE DE

DOCTEUR DE
L'UNIVERSITÉ DE BORDEAUX

ÉCOLE DOCTORALE : SCIENCES ET ENVIRONNEMENTS
SPÉCIALITÉ : GÉOLOGIE SÉDIMENTAIRE, PALÉOCÉANOGRAPHIE ET CLIMATOLOGIE

Par Axel RANSINANGUE

Apprentissage Profond pour la Pétrographie des Roches Carbonatés

Sous la direction de : Pr. Raphael BOURILLOT, Dr. Richard LABOURDETTE, Pr. Nesrine CHEHATA
(Co-encadrants: Dr. Emmanuel DUJONCQUOY, Dr. Sebastien GUILLON, M. Erwann HOUZAY)

Soutenue le 10/11/2025

Membres du jury :

M. Raphaël BOURILLOT	Professeur	Université de Bordeaux	Directeur de thèse
M. Cédric JOHN	Professeur	Queen Mary University of London (GB)	Rapporteur
M. Stephen LOKIER	Professeur associé	University of Derby (GB)	Rapporteur
M. Yannick BERTHOUMIEU	Professeur	Université de Bordeaux	Examineur
Mme Anneleen FOUBERT	Professeure	Université de Fribourg (CH)	Examinatrice
Mme Emmanuelle VENIN	Professeure	Université de Bourgogne (Dijon)	Examinatrice
M. Richard LABOURDETTE	Docteur	TotalEnergies (Pau)	Co-Directeur
Mme Nesrine CHEHATA	Professeure	Université de Bordeaux	Co-Directeur

Membres invités :

Mme FRANCO Fabienne	Ingénieure	TotalEnergies (Pau)	Invité
M. RAZIN Philippe	Professeur	Université de Bordeaux	Invité

RÉSUMÉ

L'objectif de ce projet est d'étudier la capacité de l'apprentissage profond à améliorer l'identification des facteurs pouvant impacter le comportement dynamique des réservoirs carbonatés. Ceci par l'intégration d'un ensemble complet de données d'images acquises au fil des décennies sur des réservoirs carbonatés d'hydrocarbures et analogues d'affleurements identifiés.

Les réseaux de neurones convolutifs constituent des premiers candidats naturels pour analyser les milliers de lames minces disponibles. Cette thèse a développé une suite de méthodes complémentaires pour leur description quantitative et automatisée.

Une stratégie de subdivision optimale des images haute résolution a pu être établie pour une caractérisation multi-échelle de ce type de données permettant de se rapprocher des méthodes traditionnellement utilisées en analyse microscopique. L'approche combine l'apprentissage de représentation par un autoencoder variationnel invariant aux rotations, clustering automatique, permettant une accélération double du temps de labellisation tout en formant des groupes géologiquement cohérents basés majoritairement sur les classifications de R. Dunham pour les calcaires, R. Folk pour les dolomites et l'incorporation de minéraux secondaires tels que les faciès à anhydrites. Le modèle peut ensuite être raffiné sur ces tâches de classification subsidiaires permettant une description sémantique des échantillons étudiés (>88 % mAP).

Pour pallier l'hétérogénéité des descriptions existantes en pétrographie et le manque de quantification précise sur des données réelles, deux méthodes ont été développées : SynSection génère procéduralement des données d'entraînement en simulant les processus de dépôt (+20% mAP) dans les calcaires à grains, tandis que SynSection2 unifie transformation synthétique-réaliste et segmentation (+5% mAP supplémentaire). Ces approches permettent conjointement la segmentation automatique des éléments (bioclastes, intraclastes, oïdes et péloïdes) et l'obtention d'une base de données quantifiée homogène.

La segmentation sémantique profonde appliquée aux calcaires granulaires rend accessible la quantification des proportions relatives et l'analyse des distributions granulométriques. Les métriques développées atteignent des coefficients de détermination R^2 de 0,78 pour la taille moyenne des grains et 0,72 pour le tri.

Cette méthodologie d'ensemble permet une description standardisée et reproductible suivant les nomenclatures établies en pétrographie et pouvant se superposer à l'analyse du réseau poreux. La quantification automatisée des types de pores révèle des hétérogénéités facilitant l'analyse comparative entre les échantillons.

Cette approche constitue un outil puissant pour l'analyse pétrographique quantitative, pouvant intégrer des raffinements locaux afin d'exploiter le plein potentiel de l'apprentissage profond dans la caractérisation des systèmes carbonatés et la compréhension des réponses pétrophysiques associées.

Mots clés: Apprentissage Profond, Vision par Ordinateur, Pétrographie, Lame Mince, Carbonate

ABSTRACT

This project aims to investigate deep learning capacity to enhance the identification of factors that may impact the dynamic behavior of carbonate reservoirs. This is achieved through the integration of comprehensive datasets comprising images acquired over decades from hydrocarbon carbonate reservoirs and identified outcrop analogues.

Convolutional neural networks represent natural candidates for analyzing the thousands of available thin sections. This thesis developed a suite of complementary methods for their quantitative and automated description.

An optimal subdivision strategy for high-resolution images was established to enable multi-scale characterization of this data type, closely approximating traditional microscopic analysis methods. The approach combines representation learning through a rotation-invariant variational autoencoder with automatic clustering, achieving a two-fold acceleration in labeling time while forming geologically coherent groups based primarily on R. Dunham's classifications for limestones, R. Folk's classifications for dolomites, and the incorporation of secondary minerals such as anhydrite facies. The model can then be refined for these subsidiary classification tasks, enabling semantic description of the samples studied (>88% mAP).

To address the heterogeneity of existing petrographic descriptions and the lack of precise quantification on real data, two methods were developed: SynSection procedurally generates training data by simulating depositional processes (+20% mAP) in grain limestones, while SynSection2 unifies synthetic-to-realistic transformation with segmentation (+5% additional mAP). These approaches jointly enable automatic segmentation of elements (bioclasts, intraclasts, ooids, and peloids) and the creation of a homogeneous quantified database.

Deep semantic segmentation applied to granular limestones provides access to quantification of relative proportions and grain size distribution analysis. The developed metrics achieve determination coefficients (R^2) of 0.78 for mean grain size and 0.72 for sorting.

This integrated methodology enables standardized and reproducible description following established petrographic nomenclatures and can be superimposed on pore network analysis. Automated quantification of pore types reveals heterogeneities that facilitate comparative analysis between samples.

This approach constitutes a powerful tool for quantitative petrographic analysis, capable of incorporating local refinements to fully exploit deep learning's potential in characterizing carbonate systems and understanding associated petrophysical responses.

Keywords: Deep Learning, Computer Vision, Petrography, Thin Section, Carbonate

UMR EPOC

Environnements et Paléoenvironnements Océaniques et Continentaux, n°5805

TotalEnergies

OT/TL/G&R/IS/TR – OT/R&D/SUB/FRES

Phoenix.AI

Conditions optionnelles de réutilisation :



REMERCIEMENTS

Je souhaite adresser mes remerciements les plus profonds aux personnes m'ayant accompagné tout au long de ce parcours scientifique.

Mes premières pensées s'adressent à mon directeur de thèse Richard LABOURDETTE. Merci de m'avoir accordé cette confiance et autonomie qui m'ont permis d'aboutir à ce résultat. Ton soutien constant et ta compréhension à chaque étape m'ont permis de naviguer sereinement dans cette aventure. Tous ces moments partagés resteront précieux et n'auraient pu être meilleurs. J'espère sincèrement que ce travail n'est que le début de quelque chose de plus ambitieux, et que tu auras la satisfaction de voir se dessiner la vision d'ensemble de tout ce que tu as initié une fois toutes les pièces du puzzle assemblées.

Je remercie Fabienne FRANCO notre ancienne responsable d'équipe, merci pour ton soutien, les discussions et l'ensemble des attentions que tu as pu avoir envers chacun de nous quand tu dirigeais le projet. Dès mon arrivée, j'ai eu la chance d'intégrer une équipe soudée, complémentaire et bienveillante. Je mesure aujourd'hui la rareté d'un tel environnement de travail et te suis reconnaissant d'avoir pu évoluer dans ce cadre que tu sais si bien créer. Je te souhaite tout le meilleur pour la suite.

Je remercie chaleureusement Emmanuel DUNJONCQUOY de m'avoir accueilli lors de mes stages qui ont mené à cette thèse. Merci pour ton accompagnement, d'avoir su me guider et m'encourager à explorer de nouveaux horizons avec la même motivation que tu sais montrer dans tes propres projets. Mes remerciements vont également à Julie CHAMPAGNE, pour votre accueil bienveillant à tous les deux chez AGEOS, durant ces quelques mois d'attente avant le début de la thèse.

Je remercie également Erwann HOUZAY et Sebastien GUILLON, qui ont été une source d'inspiration constante tout au long de cette thèse. Ils m'ont ouvert les portes du traitement d'image et de l'apprentissage numérique, disciplines qui m'étaient alors inconnues. Merci d'avoir partagé votre expertise et votre passion avec tant de générosité. Chacun de nos échanges a enrichi ma réflexion, m'a appris à aborder les problèmes sous de nouveaux angles et à conserver un esprit critique nécessaire à toute avancée. J'espère avoir l'opportunité de continuer à échanger avec vous à l'avenir et de progresser dans ces domaines.

Je tiens à exprimer ma gratitude envers mes encadrants universitaires, Nesrine CHEHATA et Raphael BOURILLOT. Merci pour votre accompagnement et la confiance que vous m'avez témoignée en me laissant toute l'autonomie possible dans mes pistes d'investigation. Je vous souhaite le meilleur dans vos projets personnels et scientifiques.

Je souhaite remercier tous les membres de l'équipe au côté de laquelle j'ai pu travailler au cours de ces trois dernières années et qui m'ont permis de me sentir pleinement intégré. Un grand merci à toi Pauline BRAUN pour tous nos échanges et ta bonne humeur communicative qui ont su me réconforter lors des moments de doute. Merci à Maddalen LEPPHAILLE et Skah Geordy BISSAVETSY, vous êtes des personnes formidables et j'espère de tout cœur que nous aurons l'occasion de nous retrouver dans l'avenir. Merci à Claire CIESLAR, Dirk GELDERBLOM, Lucio DI-PETRO, Sebastien CAMBON, Philippe AUGÉ, Maxime CADIAT, Jeremie FERNAGU et Alexandre LAJOINIE. Mes meilleurs vœux de réussite vous accompagnent pour vos prochaines aventures.

J'adresse mes remerciements à toute l'équipe de l'ancien service carbonate qui m'a ouvert ses portes à mon arrivé dans l'entreprise, et plus particulièrement Arnaud MEYER, Vincent MARLOT et amis Jules GRELLIER, Pierre POLARD-TAINE, Cyprien LANTEAUME, Pierre-Alexandre TEBOUL, Ali MOHAND-SAID et Alexandre LETTTERON pour tous ces moments de complicité et de rire partagés.

Un grand merci à Lionel MARTINEZ et l'ensemble des membres de la carothèque qui ont passé un temps considérable et précieux à acquérir les scans de lames minces sur lesquels j'ai pu travailler. Je remercie également Bruno CONCHE pour m'avoir donné accès aux infrastructures de calcul qui ont rendu possible la réalisation de mes travaux.

Je souhaite plein de réussite à mes amis doctorants Rihame NBRI et Yanis CHERIFI. Votre aide, en plus des liens que nous partageons aujourd'hui m'ont énormément apporté. Beaucoup de choses n'auraient pu être rendues possible à ce jour si vous n'aviez pas été là. Je suis certain que vous enrichirez considérablement vos domaines de recherche et que vos futures découvertes marqueront des avancées significatives. J'ai hâte de voir tout ce que vous accomplirez. Mes encouragements accompagnent également Alexandre THOUVENOT et Julien LALANNE pour leur thèse. Je remercie aussi Lionel BOILLOT pour les échanges enrichissants qu'il a su initier autour de la recherche et de l'image.

J'adresse une pensée toute particulière à mes enseignants du primaire Arnaud DULUC, Laetitia LAHET et son compagnon Sylvain, ainsi qu'à ceux du collège Maïté CAVERIVIÈRE et Christian ESPESO, qui m'ont accompagné durant une période importante de ma vie. Je remercie également Philippe RAZIN et Bertrand FASENTIEUX de m'avoir fait découvrir des domaines passionnants tout au long de ma formation.

A mes parents, toi maman pour tes encouragements et réconfort inestimable, toi papa pour ton soutien constant, je suis heureux d'en être arrivé là grâce vous. A ma sœur avec qui je n'ai pu partager que trop peu de choses depuis le début de cette thèse j'ai hâte que ce nouveau chapitre nous permette de passer tout le temps possible ensemble. Ces derniers mots pour toi Eléa qui a été chaque jour présente et d'un soutien sans égal, je vous aime.

TABLE OF CONTENTS

CHAPTER 1: GENERAL INTRODUCTION	1
1.1 Context and Challenges	1
1.2 Conventional Workflows for 2G&R Synthesis	1
1.2.1 Sedimentological Characterization	3
1.2.2 Petrophysical Characterization	4
1.2.3 Integration and Reservoir Modeling	5
1.3 Phoenix.AI for AI-Enhanced 2G&R Synthesis	6
1.4 Research Framework	8
1.5 Thesis Question and Objectives	9
REFERENCES	11
CHAPTER 2: STATE OF THE ART	14
2.1 Carbonate Rock Analysis	14
2.1.1 Limestone Classification	14
2.1.2 Dolomite Classification	16
2.1.3 Methodologies Discussion	18
2.1.4 Numerical Methods Integration	20
2.2 Grain Size Analysis	23
2.2.1 Grain Size Classification and Statistical Analysis	24
2.2.2 Methodologies Discussion	28
2.2.3 Numerical Methods Integration	29
2.3 Porous Network Analysis	31
2.3.1 Pores Classification	31
2.3.2 Methodologies Discussion	35
2.3.3 Numerical Methods Integration	36
2.4 Synthesis	37
2.4.1 Limitations of Traditional Methods	37
2.4.2 Numerical Development and Persistent Gaps	37
2.5 Contributions	38
2.5.1 Carbonate Rock Analysis	38
2.5.2 Grain Size Analysis	39
2.5.3 Porous Network Analysis	39
REFERENCES	40
CHAPTER 3: FROM UNSUPERVISED TO SEMI-SUPERVISED CLASSIFICATION IN CARBONATES PETROGRAPHY USING A ROTATION INVARIANT VAE	44
Abstract	44
3.1 Introduction	45

3.2 Data	47
3.2.1 Raw Data	47
3.2.2 Working Data	47
3.3 Methods	49
3.3.1 Variational Autoencoder	49
3.3.2 Rotation Invariance	50
3.3.3 Model Architecture and Hyperparameters	51
3.3.4 Model Evaluation	52
3.3.5 Representation Clustering	52
3.3.6 Semi-Supervised Fine-Tuning	53
3.4 Results	53
3.4.1 Model Training	53
3.4.2 Latent Space Distribution	55
3.4.3 Representation Clustering	56
3.4.4 Semi-Supervised Fine-Tuning	57
3.4.5 Full Thin Section Characterization	58
3.5 Discussion and Recommendations	59
3.6 Conclusion	60
REFERENCES	62
CHAPTER 4: SYNSECTION – SEDIMENTOLOGY-DRIVEN DATA GENERATION FOR DEEP LEARNING APPLICATIONS IN CARBONATES PETROGRAPHY	65
<i>Abstract</i>	65
4.1 Introduction	66
4.2 Data	69
4.2.1 Raw Data	69
4.2.2 Working Data	70
4.3 Methods	71
4.3.1 Replication of Carbonate Rock Formation Process	73
4.3.2 2D Object Packing	74
4.3.3 Texture Synthesis	77
4.3.4 Image Blending	80
4.3.5 SynSection Dataset Construction	82
4.3.6 Image Classification	83
4.3.7 Image Segmentation	85
4.4 Results	88
4.4.1 Image Classification	88
4.4.2 Image Segmentation	90
4.4.3 Description and Quantification at Full Thin Section Scale	94
4.5 Discussions	95
4.5.1 Synthetic Data Benefits and Limits	95
4.5.2 Model Performance Analysis in Classification and Segmentation	97

4.5.3	Practical Workflow Implications and Future Recommendations	98
4.6	Conclusion	99
REFERENCES		100
CHAPTER 5: SYNSECTION 2 – ADAPTIVE SYNTHETIC DATA AUGMENTATION WITH UNIFIED DOMAIN TRANSFER AND SEGMENTATION IN CARBONATES PETROGRAPHY		108
<i>Abstract</i>		108
5.1	Introduction	109
5.2	Data	111
5.2.1	Raw Data	111
5.2.2	Working Data	111
5.2.3	Data Preprocessing	111
5.2.4	Data Labeling	111
5.2.5	Synthetic Dataset Generation	112
5.3	Methods	112
5.3.1	Unsupervised Domain Adaptation	113
5.3.2	From Unsupervised Domain Adaptation to Semi-Supervised Image Translation	113
5.3.3	Semi-Supervised Adaptive Image Translation with Semantic Consistency	114
5.4	Results	125
5.4.1	Image Quality and Features Analysis	125
5.4.2	Image Segmentation and Practical Petrographic Performance	126
5.5	Discussions	131
5.5.1	Effectiveness of Domain Translation for Petrographic Analysis	131
5.5.2	Grain Size Analysis Using Local Thickness	131
5.5.3	Architectural Considerations and Training Strategy	132
5.5.4	Limitations and Challenges	132
5.5.5	Recommendations	133
5.6	Conclusion	133
REFERENCES		135
CHAPTER 6: GENERAL SYNTHESIS, CONCLUSION AND WAY FORWARD		140
6.1	Deep Quantitative Petrography	140
6.2	Methodology Overview	140
6.2.1	Image Preprocessing and Tiling Strategy	140
6.2.2	Litho-Textural Analysis	140
6.2.3	Granular Analysis	143
6.3	Limitations and Recommendations	146
6.3.1	Manufacturing Impacts	146
6.3.2	Classification Coverage Limitations	146
6.3.3	Dolomite Classification Challenges	146
6.3.4	Groundmass Material Recognition Challenges	146

6.3.5	3D Grain Representation	147
6.3.6	Multi-Scale Segmentation	147
6.3.7	Architecture Updates	147
6.3.8	Foundation Model Integration	147
6.4	Conclusion	148
6.5	Way Forward	149
REFERENCES		151

LIST OF FIGURES

FIGURE 1.1	CONVENTIONAL WORKFLOW FOR RESERVOIR CHARACTERIZATION SYNTHESIS INVOLVING THREE MAJOR DISCIPLINES: SEDIMENTOLOGY, PETROPHYSICS, AND GEOMODELLING.	2
FIGURE 1.2	THE PHOENIX.AI PROJECT FOR DEEP LEARNING APPROACHES INTEGRATION INTO GEOLOGICAL RESERVOIR CHARACTERIZATION.	7
FIGURE 1.3	PHD TIMELINE SHOWING THE PERIODS DEVOTED TO EACH PHASE OF THE PROJECT.	10
FIGURE 2.1	TYPICAL COMPONENTS AND CORRESPONDING CHARACTERISTICS OBSERVED IN LIMESTONES THIN SECTIONS (MICROSCOPIC SCALE). ALLOCHEMICAL CONSTITUENTS ARE DIVIDED IN VARIOUS SUBTYPES (FOSSILS, INTRACALASTS, OIDS AND PELLETS) AS WELL AS ORTHOCHEMICAL CONSTITUENTS (MICRITE AND SPARRY CALCITE).	14
FIGURE 2.2	FOLK'S LIMESTONE CLASSIFICATION FRAMEWORK. THE SYSTEM USES VISUAL PROPORTION ESTIMATION OF ALLOCHEMICAL TYPES AND OTHOCHEMICAL CONTENT FOR CATEGORIZING SAMPLES.	15
FIGURE 2.3	DUNHAM'S LIMESTONE CLASSIFICATION FRAMEWORK. THE SYSTEM ESTABLISHES THE PRINCIPLE OF DEPOSITIONAL TEXTURE USING PROPORTION OF GRAINS AND GROUNDMASS MATERIAL FOR PRINCIPAL CATEGORIZATION. SUBCLASSES WERE LATER INTRODUCED FOR SAMPLES PRESENTING PARTICLES > 2MM OR ROCKS ORGANICALLY BUILT.	16
FIGURE 2.4	FOLK'S DOLOMITE CLASSIFICATION FRAMEWORK. THE SYSTEM RELIES ON ORIGINAL GHOST-PRESERVED GRAIN STRUCTURES AND CRYSTAL SIZE CRITERIA.	17
FIGURE 2.5	SIBLEY AND GREGG'S DOLOMITE CLASSIFICATION FRAMEWORK. THE SYSTEM REFERS TO CRYSTAL SIZE MODALITY DISTRIBUTION, CRYSTALLIZATION GEOMETRY AND REPLACED MATERIAL TO PROVIDE VARIOUS DESCRIPTION COMBINATION POSSIBILITIES.	17
FIGURE 2.6	EXAMPLES OF COGNITIVE BIASES OBSERVED IN THIN SECTION IMAGE ANALYSIS. THE TWO CASES HIGHLIGHT THE DIFFICULTY OF EVALUATING THE RELATIVE GRAIN PROPORTION IN TERMS OF OCCUPIED SURFACE. SAMPLE A ALLOWS A GOOD PROPORTION ESTIMATION RELATED TO THE 1-DIMENSIONAL PREFERENTIAL DISTRIBUTION (VERTICAL). IN SAMPLE B, WHERE GRAIN CLASSES SPATIAL DISTRIBUTION IS MIXED, OIDS AND INTRACLASTS SEEMS TO BE REPRESENTED IN SIMILAR PROPORTION, HOWEVER SINCE INTRACLASTS SHAPE IS IRREGULAR THEY DO NOT MAXIMIZE THE	

OCCUPIED SURFACE SUCH AS OOLITHS AND ARE LESS REPRESENTED. SAMPLE C AND SAMPLE D HIGHLIGHT A BIAS RELATED TO THE NUMBER OF GRAINS AND THEIR SPACING WHERE PERCENTAGE OF GRAINS IN SAMPLE C SEEMS INFERIOR TO SAMPLE D BUT IS EQUAL. ____	18
FIGURE 2.7 EXAMPLES SHOWING PROCESSING DIFFICULTIES IN LIMESTONE THIN SECTION CLASSIFICATION TASK. SAMPLE A PRESENTS MULTIPLE ROCK TEXTURES DEPENDING ON THE SCALE OF OBSERVATION. SAMPLE B AND SAMPLE C, WHICH SIMPLY DIFFER FROM A $\pm 2\%$, MUST RESULT IN DIFFERENT NAMING BASED ON DUNHAM CLASSIFICATION RULES, HOWEVER IT IS CURRENTLY ALMOST VISUALLY IMPOSSIBLE TO CLEARLY DISCRIMINATE. 19	
FIGURE 2.8 HETEROGENEITIES OF CLASSIFICATION HIGHLIGHTED IN CARBONATE THIN SECTION ANALYSIS BY LOKIER'S STUDY. _____	19
FIGURE 2.9 PRINCIPAL PARADIGMS AND NEURAL NETWORK DESIGNS INVESTIGATED IN THE STATE-OF-THE-ART NUMERICAL CARBONATE PETROGRAPHY APPLIED TO THIN SECTIONS. IMAGE CLASSIFICATION ARCHITECTURE OUTPUT IS A DISCRETE VALUE CORRESPONDING TO MOST PROBABLE CLASS. OBJECT DETECTION MODEL PROVIDES A LIST OF BOXES COORDINATES AND MOST PROBABLE CLASS PER OBJECT. THE IMAGE SEGMENTATION FRAMEWORK PRODUCES A MASK WHERE THE MOST PROBABLE CLASS IS AFFECTED PER PIXEL. _____	23
FIGURE 2.10 WENTWORTH'S GRAIN SIZE CLASSIFICATION SYSTEMS, COMPLEMENTED WITH KRUMBEIN'S CONTRIBUTIONS WITH PHI SCALE DEVELOPMENT AND FOLK CARBONATES CONSTITUENT INTEGRATION. _____	24
FIGURE 2.11 SHEPARD'S TERNARY DIAGRAM SYSTEM FOR MIXED GRAIN SIZE CHARACTERIZATION COMPLEMENTED BY SCHLEE FOR GRAVEL SIZE FRACTIONS (UPPER) AND FOLK'S NOMENCLATURE USING SIMILAR STRUCTURE WITH VARYING CLASSES AND BOUNDARIES (LOWER). _____	26
FIGURE 2.12 FOLK'S GRAIN SIZE STATISTICAL CLASSIFICATION USING GRAPHIC MEAN, INCLUSIVE GRAPHIC STANDARD DEVIATION, INCLUSIVE GRAPHIC SKEWNESS AND GRAPHIC KURTOSIS DERIVED FROM CUMULATIVE FREQUENCY DESCRIPTORS (EQUATION 1 TO 4 USING QUANTILES 5, 16, 50, 84 AND 95). VISUAL IMPACT OF THE DIFFERENT STATISTICAL PARAMETERS IS HIGHLIGHTED THROUGH THE PROBABILITY DENSITY FUNCTION. _____	27
FIGURE 2.13 TYPICAL GRAIN SIZE ANALYSIS CHARTS USED FOR MEAN GRAIN SIZE AND GRAIN SORTING CLASSES ASSIGNMENTS. _____	29
FIGURE 2.14 PRINCIPAL PARADIGMS AND NEURAL NETWORK DESIGNS INVESTIGATED IN THE STATE-OF-THE-ART NUMERICAL GRAIN SIZE ANALYSIS. DIRECT PARTICLE SIZE DISTRIBUTION USES IMAGE ENCODING FOR PRE-DEFINED QUANTILES PREDICTION. INTERMEDIATE SEGMENTATION REPLICATES HUMAN PROCESSING TASKS BY PREDICTING THE BINARY PARTICLE MASK AND AUTOMATING MEASUREMENTS. _____	30
FIGURE 2.15 CHOQUETTE AND PRAY'S POROSITY CLASSIFICATION FRAMEWORK. PORES ARE CATEGORIZED USING INTRICATE RELATIONSHIPS WITH SEDIMENTOLOGICAL INTERPRETATIONS AND VOIDS MORPHOLOGY FOR GENETIC SIGNIFICANCE. _____	33
FIGURE 2.16 LUCIA'S POROSITY CLASSIFICATION FRAMEWORK. PORE CLASSES ARE SEPARATED BETWEEN INTERPARTICLE AND VUGGY POROSITY. INTERPARTICLE POROSITY IS SUBDIVIDED INTO DUNHAM TEXTURE INTEGRATING PARTICLE SIZE AND SORTING IN THE CLASSIFICATION. VUGGY POROSITY USES CHOQUETTE AND PRAY'S NOMENCLATURE WITH THE ADDITION OF CONNECTIVITY PATTERN QUALIFICATION. _____	34

FIGURE 2.17 INSIGHTS IN CARBONATE POROSITY-PERMEABILITY RELATIONSHIPS ESTABLISHED IN LUCIA'S AND AKBAR'S STUDIES WHERE SEDIMENTOLOGICAL AND DIAGENETIC FEATURES ARE SUPPOSED TO BE RELATED TO PETROPHYSICAL SIGNATURES. _____	35
FIGURE 2.18 PRINCIPAL PARADIGMS AND APPROACHES IN THE STATE-OF-THE-ART NUMERICAL POROSITY ANALYSIS OF CARBONATE ROCKS. MORPHOMETRICAL DESCRIPTORS ARE EXTRACTED AFTER POROSITY THRESHOLDING AND ARE THEN STATISTICALLY ANALYZED OR RE-USED FOR A DOWNSTREAM CLASSIFICATION OR REGRESSION TASK PERFORMED BY A MACHINE LEARNING MODEL. _____	36
FIGURE 3.1 PETROGRAPHIC CLASSIFICATION AND COMPARATIVE CHARACTERISTICS OF CARBONATE (LIMESTONE AND DOLOMITE) AND EVAPORITIC ROCKS (ANHYDRITE) LEADING TO COMPOSITIONAL AND TEXTURES VARIABILITY SEPARATED IN DIFFERENT CLASSES. CORRESPONDING REPRESENTATIVE THIN SECTION IMAGES ILLUSTRATE EACH CATEGORY. _____	48
FIGURE 3.2 LABELED DATA DISTRIBUTION FOR THE MANUALLY CONSTRUCTED THIN SECTION IMAGE DATASET. _____	48
FIGURE 3.3 ROTATION INVARIANCE LOSS MINIMIZES IN THIS EXAMPLE THE DIFFERENCE BETWEEN $Dec(Enc(\tau_0(X))), Dec(Enc(\tau_{90}(X))), Dec(Enc(\tau_{180}(X)))$ AND $Dec(Enc(\tau_{270}(X)))$ THAT ARE ALL RECONSTRUCTED TO THE CANONICAL ORIENTATION $\tau_{270}(X)$. THE RECONSTRUCTION RELAXED TERM MINIMIZES THE DIFFERENCE $Dec(Enc(\tau_0(X)))$ AND $\tau_{270}(X)$ WHICH IS THE ROTATED INPUT CORRESPONDING TO THE MATCHING RECONSTRUCTED IMAGE. _____	50
FIGURE 3.4 RESIDUAL VAE NETWORK ARCHITECTURE MADE OF 4 RESIDUAL CONVOLUTION DOWNSAMPLING AND UPSAMPLING STAGES. LAST CONVOLUTIONAL FEATURE MAPS GENERATED BY THE ENCODER GO THROUGH AN INTERMEDIATE FULLY CONNECTED LAYER BEFORE GOING TO THE LATENT MEAN AND VARIANCE WITH THE REPARAMETERIZATION TRICK ($z = \mu + \sigma \odot \varepsilon$) ALLOWING GRADIENT BACK-PROPAGATION. THE PATCH DISCRIMINATOR SERVES AS A TEXTURE CLASSIFIER, DETERMINING THE REALISM OF LOCALLY GENERATED PATTERNS WITHIN IMAGE PATCHES. _____	52
FIGURE 3.5 EVOLUTION OF VAE TRAINING LOSS COMPONENTS ACROSS 100 EPOCHS. _____	54
FIGURE 3.6 IMAGE RECONSTRUCTION ANALYSIS SHOWING INPUT IMAGES AT 60° INCREMENTAL ROTATIONS WHERE ILLUSTRATION INCLUDES LATENT SPACE MAPPING AND RECONSTRUCTIONS, WITH THE FINAL OUTPUT ROTATED BY OPERATOR τ_{120} RELATIVE TO THE REFERENCE IMAGE. _____	54
FIGURE 3.7 T-SNE MAPPING OF LATENT VECTORS FOR THE SET OF LABELED IMAGES. _____	55
FIGURE 3.8 DISTRIBUTION OF LITHO-TEXTURE ROCK TYPES ACROSS FORMED CLUSTERS. _____	56
FIGURE 3.9 CLASSIFICATION METRICS AND CONFUSION MATRICES FOR THE BEST-PERFORMING SUPERVISED AND SEMI-SUPERVISED MODELS, EVALUATED ON VALIDATION DATA. _____	57
FIGURE 3.10 CLASSIFICATION MAPS EXAMPLES OF FULL THIN SECTION SAMPLES HIGHLIGHTING VARIOUS LITHO-TEXTURE TYPES. _____	58
FIGURE 4.1 INTEGRATION OF ARTIFICIAL INTELLIGENCE AND MODERN IMAGING TECHNIQUES TO ADDRESS TRADITIONAL CHALLENGES IN PETROGRAPHY: A COMPREHENSIVE FRAMEWORK OF PROBLEMS AND SOLUTIONS FOR ENHANCED GEOLOGICAL ANALYSIS. _____	67

FIGURE 4.2 PETROGRAPHIC ANALYSIS OF A CARBONATE THIN SECTION: A SELECTED REGION OF INTEREST IS LABELED AT THE TILE SCALE (GRAINSTONE), WITH GRAINS ANNOTATED BY CLASS (SKELETAL, OOID, INTRACLAST, PELOID) AND BACKGROUND MARKED USING BOUNDING BOXES. ANNOTATED GRAINS (THUMBNAIL IMAGES AND MASKS) ARE EXTRACTED AND CATEGORIZED BY CLASS.	69
FIGURE 4.3 DATA DISTRIBUTION FOR CLASSIFICATION AND SEGMENTATION TASKS.	70
FIGURE 4.4 DUNHAM CLASSIFICATION USED FOR MODEL TRAINING: SYNTHETIC EXAMPLES A, B, C, AND D ARE GENERATED USING THE SYNSECTION METHOD.	71
FIGURE 4.5 PROPOSED METHOD FOR SYNTHETIC IMAGE TRAINING DATASET GENERATION (SYNSECTION): ANNOTATED GRAINS ARE PACKED TO INTRODUCE SPATIAL VARIABILITY, AND DIVERSE TEXTURED BACKGROUNDS ARE SYNTHESIZED USING A STYLEGAN-2 GENERATOR. BACKGROUNDS AND GRAIN COMPOSITIONS ARE BLENDED USING PIXEL DISTRIBUTION ADAPTATION COMBINED WITH GAUSSIAN AND LAPLACIAN PYRAMIDS. ANNOTATION MASKS ARE CREATED DURING THE PACKING PROCESS, WITH LABELS DERIVED FROM INPUT GENERATION PARAMETERS. THE ENTIRE WORKFLOW IS STREAMLINED INTO A UNIFIED PROCESSING CHAIN FOR USER CONVENIENCE.	72
FIGURE 4.6 GRAIN COMPOSITIONS GENERATED WITH CONTROLLED SEDIMENTOLOGICAL PACKING CONSTRAINTS: MIMICKING GEOLOGICALLY PLAUSIBLE ARRANGEMENTS BY REGULATING CLASS DISTRIBUTION (P), GRAIN DENSITY (P), LOCAL PROXIMITY (T), AND ORIENTATION (Θ).	75
FIGURE 4.7 EFFICIENCY OF GRAIN PACKING COMPUTATION TIME ACROSS 5 REALIZATIONS PER STEP: RANGES INCLUDE [0–100] WITH A STEP SIZE OF 10, [100–1,000] WITH A STEP SIZE OF 100, AND TWO ADDITIONAL STEPS AT 2,000 AND 3,000 PACKED OBJECTS. EXAMPLES A, B, AND C DEPICT INSTANCE MASK REPRESENTATIONS OF A COMPOSITE INSTANCE MASK AT DIFFERENT STEPS, CONTAINING 115, 546, AND 2,754 PACKED OBJECTS, RESPECTIVELY.	75
FIGURE 4.8 STYLEGAN 2 MODEL ARCHITECTURE USED FOR BACKGROUND TEXTURE SYNTHESIS.	78
FIGURE 4.9 COMPARISON OF HARALICK FEATURE DISTRIBUTIONS BETWEEN GENERATED AND REAL TEXTURES.	79
FIGURE 4.10 LARGE TEXTURED BACKGROUNDS GENERATED USING THE STYLEGAN2 GENERATOR (512×512 PIXELS): EXAMPLES A, B, AND C REPRESENT SPAR CEMENT TEXTURES, WHILE D, E, AND F DEPICT MICRITIC MATRIXES. THESE TEXTURES WERE CLASSIFIED USING A KNN ALGORITHM.	80
FIGURE 4.11 COMPARISON OF IMAGE COMPOSITION METHODS: ALPHA BLENDING, LAPLACIAN BLENDING, AND PIXEL DISTRIBUTION ADAPTATION COMBINED WITH LAPLACIAN BLENDING.	81
FIGURE 4.12 IMPLEMENTED IMAGE CLASSIFICATION MODEL ARCHITECTURES USED FOR ROCK TEXTURE RECOGNITION.	84
FIGURE 4.13 IMPLEMENTED IMAGE SEGMENTATION MODEL ARCHITECTURES USED FOR GRAIN IDENTIFICATION.	86
FIGURE 4.14 MODELS PERFORMANCE METRICS COMPARISON FOR EACH TRAINING STRATEGY: HIGHLIGHTING THE IMPACT OF SYNTHETIC DATA INTEGRATION IN THE IMAGE	

CLASSIFICATION TRAINING PROCEDURE THROUGH PERFORMANCE IMPROVEMENT MONITORING. _____	89
FIGURE 4.15 CONFUSION MATRICES COMPARING CLASSICAL AND TRANSFER LEARNING FROM SYNTHETIC TO REAL IMAGE TRAINING: SAMPLES A, B, C, AND D REPRESENT SELECTED GRAIN-ANNOTATED DATASET EXAMPLES FOR QUANTIFYING ELEMENT ABUNDANCE FROM BINARY MASKS A, B, C, D. IMAGES ARE OVERLAID WITH GRAD-CAM. _____	89
FIGURE 4.16 MODELS PERFORMANCE METRICS COMPARISON FOR EACH TRAINING STRATEGY: COMPARING BINARY AND MULTICLASS SEGMENTATION WITH THE IMPACT OF SYNTHETIC DATA INTEGRATION IN THE TRAINING PROCEDURE THROUGH PERFORMANCE IMPROVEMENT MONITORING. _____	91
FIGURE 4.17 COMPARISON OF BINARY AND MULTICLASS SEGMENTATION ACROSS VARIOUS GRAIN ARRANGEMENTS AND ROCK TEXTURES: SAMPLE A (MUDSTONE WITH SKELETAL AND PELOIDAL ELEMENTS), SAMPLE B (GRAINSTONE WITH PREDOMINANT OIDS AND PELOIDS), SAMPLE C (GRAINSTONE WITH SKELETAL, OOID, AND PELOID GRAINS), SAMPLE D (GRAINSTONE WITH INTRACLASTS, SKELETALS, AND PELOIDS), AND SAMPLE E (MIXED GRAINSTONE). _____	93
FIGURE 4.18 QUANTITATIVE ANALYSIS OF THIN SECTIONS AT FULL IMAGE SCALE: FROM ROCK TEXTURE TO GRAIN IDENTIFICATION. SAMPLE A (HETEROGENEOUS ROCK TEXTURE PRIMARILY COMPOSED OF MUDSTONE AND WACKESTONE), SAMPLE B (RELATIVE PROPORTIONS OF INTRA-GRANULAR AND INTER-GRANULAR POROSITY), SAMPLE C (COARSE INTRACLAST LAMINATIONS WITHIN AN OOLITIC-DOMINATED DEPOSIT). _____	94
FIGURE 4.19 SYNTHETIC IMAGES SHOWING COLORS DISMATCH BETWEEN ALLOCHEMS AND THE SELECTED BACKGROUNDS HIGHLIGHTING THE INTRODUCTION OF POTENTIAL NON REALISTIC EXAMPLES IN THE DATABASE. _____	96
FIGURE 5.1 SIMPLIFIED REPRESENTATION OF SYNSECTION2 FRAMEWORK TRANSLATING IMAGES FROM SYNTHETIC TO REAL DOMAIN WITH SEMANTIC CONSISTENCY AND REFERENCE STYLE CONSTRAINTS. _____	110
FIGURE 5.2 CONVENTIONAL METRICS USED IN COMPUTER VISION BOTH FOR IMAGE GENERATION AND IMAGE SEGMENTATION COMPLETED WITH SPECIFIC PETROGRAPHIC METRICS. _____	110
FIGURE 5.3 GRAIN DATA PROPERTIES SHOWING THE DISTRIBUTION OF LABELED INSTANCES ACROSS SELECTED IMAGE TILES AND AVERAGE SURFACE AREA PER GRAIN TYPE. _____	112
FIGURE 5.4 VISUALIZATION OF CONSTRAINT EFFECTS IN IMAGE TRANSLATION NETWORKS FOR DEEP LEARNING APPLICATIONS IN CARBONATE PETROGRAPHY. _____	117
FIGURE 5.5 NETWORK ARCHITECTURE FOR SYNTHETIC-TO-REAL DOMAIN TRANSLATION OF CARBONATE THIN SECTION IMAGES WITH STYLISTIC AND SEMANTIC CONSISTENCY. _____	118
FIGURE 5.6 SENSITIVITY ANALYSIS OF GRAIN SIZE STATISTICS TO SEGMENTATION ERRORS: COMPARISON OF TRADITIONAL DIAMETER MEASUREMENTS VERSUS LOCAL THICKNESS METHODS _____	124
FIGURE 5.7 LATENT SPACE ANALYSIS COMPARING SYNTHETIC, DOMAIN-ADAPTED, AND REAL IMAGE DISTRIBUTIONS. FID SCORES BETWEEN NON-ADAPTED AND ADAPTED IMAGES ARE SHOWN VIA HISTOGRAM. THE PRIOR DISTRIBUTION AND PRINCIPAL REFERENCE STYLE USED AS DATA PARTITIONING ARE ALSO DISPLAYED. _____	126

FIGURE 5.8 QUALITATIVE COMPARISON OF SYNTHETIC IMAGE ADAPTATION USING REAL IMAGES AS STYLE REFERENCES. _____	126
FIGURE 5.9 COMPARATIVE SEGMENTATION RESULTS BETWEEN CONVENTIONAL TRANSFER LEARNING USING SYNTHETIC IMAGES DURING PRE-TRAINING THEN REAL DATA FOR FINE-TUNING AND THE PROPOSED END-TO-END TRAINING ON DOMAIN-ADAPTED IMAGES. _____	129
FIGURE 5.10 FULL THIN SECTION SEGMENTATION EXAMPLES SHOWING GRAIN COMPOSITION DISTRIBUTIONS AND GRAIN SIZE ANALYSIS WITH PROBABILITY DENSITY AND CUMULATIVE DISTRIBUTION FUNCTION. _____	130
FIGURE 6.1 EXAMPLE OF MULTI-SCALE CARBONATE THIN SECTION ANALYSIS RESULTING FROM THE COMBINED DEEP APPROACHES OUTPUTS. _____	141
FIGURE 6.2 FIRST ORDER THIN SECTION CHARACTERIZATION USING DEEP CLASSIFICATION APPROACHES FOR QUANTITATIVE ANALYSIS. _____	142
FIGURE 6.3 SECOND ORDER THIN SECTION CHARACTERIZATION IN CARBONATE ROCKS USING DEEP SEGMENTATION APPROACHES FOR QUANTITATIVE ANALYSIS. _____	145
FIGURE 6.4 MICP-BASED ROCK TYPING SHOWING SEDIMENTOLOGICAL AND PETROPHYSICAL MISMATCHES THAT EXPOSE INADEQUATE DATA INTEGRATION IN THE CLUSTERING METHODOLOGY, LEADING TO POORLY DEFINED PETROPHYSICAL GROUPS (PG). _____	150

LIST OF TABLES

TABLE 2.1 COMPARISON OF MANUAL VERSUS AI-ASSISTED LABELLING APPROACHES FOR CARBONATE THIN SECTION ANALYSIS, PRESENTING PROCESSING OBJECTIVES, EMPLOYED METHODS, AVERAGE PROCESSING RATES OBSERVED DURING THE STUDY PERIOD, AND QUALITY-EFFICIENCY ASSESSMENTS. AI-ASSISTED IMAGE CLASSIFICATION IMPLEMENTS THE METHODOLOGY DEVELOPED IN PAPER 1, WHILE ALLOCHEM ANNOTATION UTILIZES A SPECIALIZED LABELLING INTERFACE INTEGRATED WITH META'S SEGMENT ANYTHING MODEL (SAM). _____	20
TABLE 2.2 STATE-OF-THE-ART MACHINE AND DEEP LEARNING APPROACHES FOR CARBONATE THIN SECTION CLASSIFICATION. (ACC. : ACCURACY) _____	21

TABLE 2.3 STATE-OF-THE-ART MACHINE LEARNING APPROACHES FOR CARBONATE THIN SECTION OBJECT IDENTIFICATION AND CLASSIFICATION. (ACC. : ACCURACY, IOU: INTERSECTION OVER UNION, MAP: MEAN AVERAGE PRECISION)	21
TABLE 2.4 STATE-OF-THE-ART DEEP LEARNING APPROACHES FOR GRAIN SIZE ANALYSIS. (MAE: MEAN ABSOLUTE ERROR, MAP: MEAN AVERAGE PRECISION, MAPE: MEAN AVERAGE PERCENTAGE ERROR, RMSE: ROOT MEAN SQUARE ERROR, R^2 : COEFFICIENT OF DETERMINATION, WASSERSTEIN: DISTRIBUTION SIMILARITY)	30
TABLE 2.5 ARCHIE'S POROSITY CLASSIFICATION FRAMEWORK. PORE TYPES ARE DEFINED ACCORDING TO ROCK COMPOSITIONAL TEXTURE IN WHICH THEY ARE OBSERVED, VOID SIZES AND THE SURFACE THEY COVER.	32
TABLE 3.1 CLASSIFICATION METRICS FOR SUPERVISED AND SEMI-SUPERVISED APPROACHES, AVERAGED ACROSS 3 TRAINING REPETITIONS.	57
TABLE 4.1 MODELING AND SYNTHESIS OF SEDIMENTARY CARBONATES ROCK TEXTURES THROUGH NUMERICAL APPROACHES: KEY PROCESSES, CONSTRAINTS, AND COMPUTATIONAL SOLUTIONS.	72
TABLE 4.2 CLASSIFICATION MODEL CONFIGURATIONS AND PERFORMANCE METRICS: COMPARISON OF RESULTS BETWEEN TRAINING ON A SMALL REAL IMAGE DATASET AND TRANSFER LEARNING FROM SYNTHETICS TO REAL IMAGES.	88
TABLE 4.3 SEGMENTATION MODEL CONFIGURATIONS AND PERFORMANCE METRICS FOR BINARY TASKS: COMPARISON OF RESULTS BETWEEN TRAINING ON A SMALL REAL IMAGE DATASET AND TRANSFER LEARNING FROM SYNTHETIC TO REAL IMAGES.	91
TABLE 4.4 SEGMENTATION MODEL PERFORMANCE METRICS FOR MULTICLASS TASKS: COMPARISON OF RESULTS BETWEEN TRAINING ON A SMALL REAL IMAGE DATASET AND TRANSFER LEARNING FROM SYNTHETIC TO REAL IMAGES.	92
TABLE 5.1 GRAIN SIZE WITH CORRESPONDING SORTING CLASSIFICATION AND STATISTICAL PARAMETERS FOR SEDIMENTARY ANALYSIS.	123
TABLE 5.2 PERFORMANCE COMPARISON OF BASELINE SEGMENTATION APPROACH AND DOMAIN ADAPTATION METHOD ACROSS SEGMENTATION AND PETROGRAPHIC MEASUREMENT METRICS	127
TABLE 5.3 CLASS-WISE SEGMENTATION PERFORMANCE COMPARISON BETWEEN BASELINE AND DOMAIN ADAPTATION APPROACHES	127

GLOSSARY – PETROGRAPHY

Allochem

A broad term referring to identifiable grains within carbonate rocks, typically formed from biotic or abiotic carbonate material. Common examples include bioclasts, ooids, oncoids, intraclasts or peloids.

Anhydrite

A calcium sulfate mineral (CaSO_4) that forms through the dehydration of gypsum or direct precipitation from highly saline waters. Commonly found in evaporite sequences and locally in carbonate systems, it appears as white to gray crystalline masses.

Bioclast

Fragments of skeletal material, such as shells, foraminifers or corals that contribute to the composition of a carbonate rock.

Cement

Crystalline carbonate material precipitated from pore fluids that binds allochems together, forming one of the potential rock's frameworks.

Diagenesis

Post-depositional processes, such as cementation, recrystallization, and dolomitization, that alter the texture, composition, and porosity of carbonate rocks.

Dolomite

A carbonate mineral ($\text{CaMg}(\text{CO}_3)_2$) and the rock composed predominantly of this mineral. Dolomite typically forms through the replacement of limestone during diagenesis, where magnesium-rich fluids alter the original calcite. It exhibits distinctive rhombohedral crystals.

Grainstone

A type of carbonate rock that is grain-supported with a sparry cement.

Intraclast

Fragments of pre-existing carbonate rock reworked into a new depositional environment.

Limestone

A sedimentary rock composed primarily of calcium carbonate (CaCO_3), typically in the form of calcite or aragonite. Limestones form through biological processes (accumulation of skeletal fragments), chemical precipitation, or mechanical accumulation of carbonate grains. They are characterized by their textures ranging from fine-grained micrite to coarse-grained deposits.

Micrite

Microcrystalline calcite or lime mud forming the fine-grained matrix in carbonate rocks. It can be a primary deposit or a product of diagenetic recrystallization.

Ooid

A small, spherical to ellipsoidal carbonate grain formed by concentric precipitation of calcium carbonate around a nucleus, typically in shallow, high-energy environments.

Packstone

A type of carbonate rock that is grain-supported with a micritic matrix.

Peloid

A fine-grained, structureless carbonate grain generally formed by the fecal pellets of organisms or the micritization of fragmented allochems.

Porosity

The volume of voids or pores in a rock, useful for assessing reservoir quality in carbonate rocks. Principal types of porosity include moldic, intergranular, vuggy, and fracture porosity.

Recrystallization

A diagenetic process in which original carbonate minerals are replaced or restructured into new crystal forms, often altering the original textures.

Sparite

Coarse-crystalline calcite cements that precipitate in pore spaces during diagenesis, contributing to the infill of voids.

Stylolite

Irregular, serrated surfaces formed by pressure dissolution during burial, generally marked by insoluble residue.

Wackestone

A carbonate rock type that is matrix-supported with more than 10% of allochems and a micritic matrix.

GLOSSARY – DEEP LEARNING

Activation Function

A mathematical function applied to the output of a neuron to introduce non-linearity into the model. Common examples include ReLU (Rectified Linear Unit), sigmoid, softmax, tanh, and leaky ReLU.

Artificial Neural Network

A computing system inspired by biological neural networks, consisting of interconnected nodes organized in layers.

Backpropagation

A training algorithm for neural networks that computes the gradients of the loss function with respect to each weight by propagating errors backward through the network.

Batch

A subset of the training data processed together in one forward and backward pass. Processing data in batches improves computational efficiency.

Convolutional Neural Network (CNN)

A type of neural network designed to process structured grid data, such as images. It uses convolutional layers to capture spatial hierarchies in the input.

Deep Learning

A subset of machine learning using neural networks with multiple layers.

Epoch

One complete pass through the entire training dataset during the training process.

Feature Map

The output of applying a convolutional filter to an input, capturing specific patterns or features.

Fine-tuning

The process of taking a pre-trained model and further re-training it on a specific task or dataset after decreasing the initial learning rate.

Graphics Processing Unit (GPU)

Hardware specially designed for parallel processing, commonly used to accelerate deep learning computations.

Gradient Descent

An optimization algorithm that updates the model's weights by minimizing the loss function. Variants include Stochastic Gradient Descent (SGD), RMSprop, and Adam.

Gradient Vanishing/Exploding Problem

Issues in deep networks where gradients become too small or too large during backpropagation, hindering effective learning.

Ground Truth

Accurate and objective data or labels that serve as the standard against which a model's predictions are evaluated. It is collected through reliable methods such as human annotation, sensor readings, or experimental measurements and is used to train and validate models.

Hyperparameter

A configuration variable set before training begins, such as learning rate or batch size.

Latent Space

The compressed representation of input data in a lower-dimensional space learned by the model, often used in autoencoders and VAEs.

Learning Rate

A hyperparameter that controls the step size of weight updates during gradient descent and affects the convergence speed of the model.

Loss Function

A mathematical function that measures the discrepancy between the model's predictions and the ground truth values. This function is differentiable with respect to the model's parameters, enabling the propagation of errors through the network using an optimizer. Common examples include Mean Squared Error (MSE) for regression tasks and Cross-Entropy Loss for classification problems.

Model Architecture

The specific arrangement of layers, nodes, and connections in a neural network.

Normalization

The process of adjusting values to a common scale, generally to improve training stability.

Overfitting

A scenario where the model learns to fit the training data too closely, failing to generalize to unseen data.

Regularization

Techniques to reduce overfitting by adding constraints to the model, such as L1 and L2 regularization, data augmentation and specific design of loss terms.

Training Set

The portion of data that is used to train the model.

Transfer Learning

A technique where a pre-trained model on a large dataset is fine-tuned for a related task on a smaller dataset.

Validation Set

The portion of data that is used to evaluate model performance during training.

Weights

Learnable parameters in a neural network determining the strength of connections between neurons.

CHAPTER 1: GENERAL INTRODUCTION

1.1 Context and Challenges

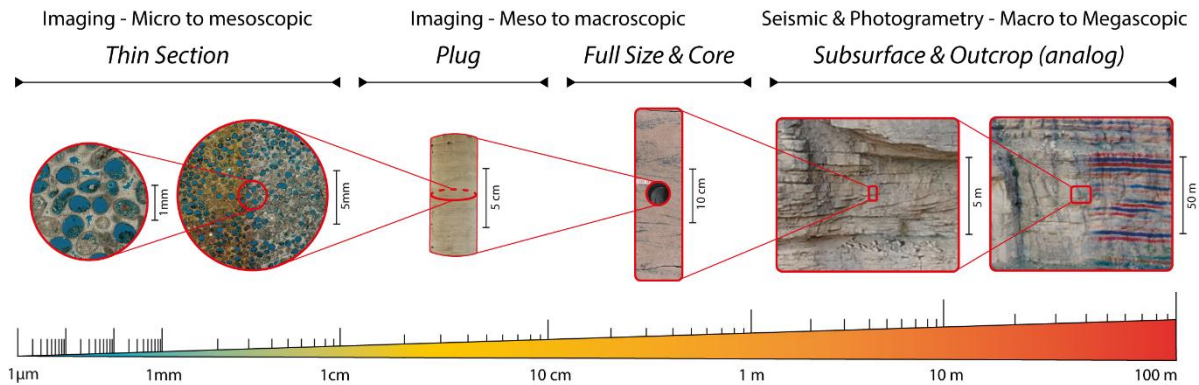
This thesis was conducted in collaboration with the University of Bordeaux (UMR EPOC – ENSEGID) and TotalEnergies, forming part of the Phoenix.AI research project hosted within the company. The objective of this larger-scale project is to support Geological, Geophysical and Reservoir (2G&R) synthesis in carbonate systems through the application of AI algorithms. The workflow focuses on multiscale data description and integration, facilitating homogenization, storage, and understanding of geological reservoir behaviors.

1.2 Conventional Workflows for 2G&R Synthesis

A 2G&R synthesis represents an advanced stage in a geological reservoir characterization process. It typically follows preliminary studies conducted in specific domains such as sedimentology and petrophysics before the final geomodelling step (Figure 1.1).

Conceptual Scales and Processes for 2G&R Synthesis

Acquisitions:

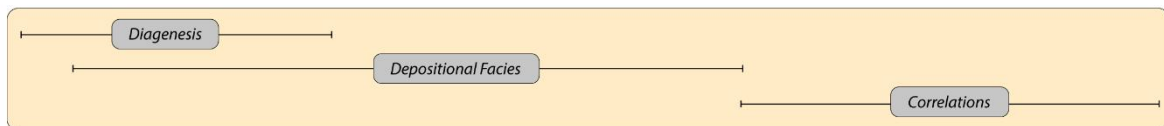


Characterization:

• Sedimentological



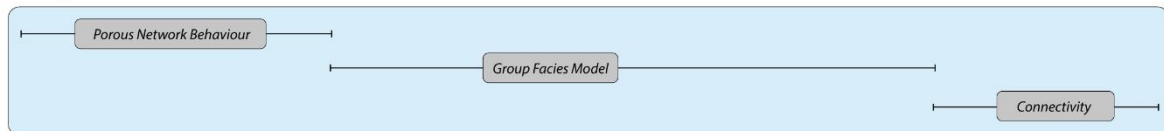
• Interpretations



• Petrophysical



• Interpretations



• Reservoir

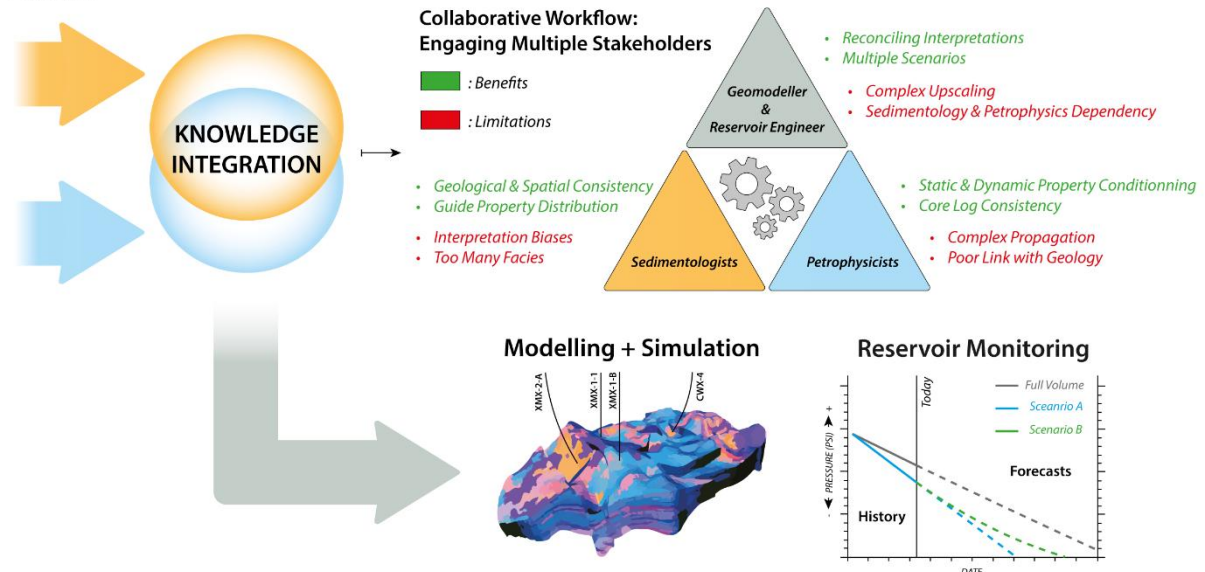


Figure 1.1 Conventional workflow for reservoir characterization synthesis involving three major disciplines: sedimentology, petrophysics, and geomodelling.

1.2.1 Sedimentological Characterization

Sedimentology studies sedimentary rocks and interprets their formation mechanisms by analyzing the transport, deposition, and precipitation processes that control elemental composition, from initial particle accumulation through consolidation and dissolution events that may occur during diagenesis. In the studied systems, regional architecture results from complex interactions among these processes across scales ranging from individual grains to entire basins, creating multiple levels of sedimentary heterogeneity (Cross & Burchette, 2024; Mohammed & Velledits, 2024; Tavakoli, 2019).

1.2.1.1 Microscale: Thin Section and Plug Analysis

Original sediments composed of calcium carbonate grains are predominantly precipitated from biological sources (shells, coral reefs, foraminifera) and through saltation of particles in shallow high-energy marine environments (Flügel, 2004; Tucker, 2009). Once sealed together through muddy deposits or interstitial cement (Harris et al., 1985), grains form micro- to centimeter-scale arrangements that are analyzed through the observation of thin sections (30 µm thickness) and core plugs. At this scale, the nature and distribution of elements and porosity within a sample can be characterized to provide initial indicators for interpreting both depositional conditions and environments, as well as the diagenetic processes that shaped the rock's current state.

1.2.1.2 Mesoscale: Sedimentary Structures and Features on Cores

Information extracted from thin sections and core plugs represents very local compositional information and does not allow deduction of spatial distribution laws for constructing a conceptual depositional model (Pomar & Kendall, 2008). Marine sediments, generally transported during sedimentary cycles, provide clues about the nature and intensity of past currents through alternating grain sizes that reveal structural patterns such as ripples or laminations (Demicco & Hardie, 1995; Yilmaz & Altiner, 2001).

These sedimentary structures, visible in cores at centimeter to multi-decimeter scales, enhance thin section interpretations to provide a comprehensive understanding of how sedimentary records evolved vertically. Since core plugs and thin sections are typically sampled at one-foot intervals in the oil and gas industry, the compositional data they provide can be extrapolated across the broader visual textures observed between sample points in the cores. Over recent years, especially in field studies, sampling has become more precise and targeted, focusing on well-defined litho-facies.

This integration of microscopic and mesoscopic data forms the foundation for defining depositional facies that will feed the future geological model.

1.2.1.3 Macroscale: Sedimentary Bodies and Stratigraphic Architecture

The succession and boundaries between different structural patterns can lead to identification of sedimentary bodies, such as dunes or channels; whose envelopes and connectivity appear to play a key role in resource recovery or trapping (Fogg & Lucia, 1990; Herbert et al., 2015; Tatum et al., 2012). However, the lateral extension of sedimentary bodies is not directly observable through core analysis.

Empirical laws from field analogues or current systems are necessary to define geometric relationships between thickness and width of sandbodies (Harris, 2019; Harris et al., 2011),

enabling geostatistical modeling of their distribution (Burchette et al., 1990; Palermo et al., 2010). These relationships constrain three-dimensional sedimentary body architecture from one-dimensional well observation.

1.2.1.4 Regional Scale: Stratigraphic Sequences and Seismic Correlations

After defining facies and identifying sedimentary bodies within formations, stacking patterns are hierarchized to establish stratigraphic sequences that reflect lateral migration of depositional environments (Ravenne, 2002). This sequential analysis integrates higher-order cycles controlled by eustatic, tectonic, and flexural variations.

These processes can lead to low-frequency topographic (slope) changes and modify structural unit geometries at hectometer to kilometer scales. Seismic sections or photogrammetric outcrop models track this evolution and enable stratigraphic correlations between wells or sections by defining temporal boundaries. This data type thus helps in constraining unit geometry and lateral continuity at basin scale.

1.2.2 Petrophysical Characterization

Petrophysics studies the physical properties of reservoir rocks and analyzes their capacity to store and transmit fluids through analysis of static properties (porosity, permeability, saturation) to their dynamic characterization during fluid flow test through porous media. In carbonate reservoirs, petrophysical properties result from complex interactions between original rock texture and subsequent diagenetic modifications, creating multiscale heterogeneities from pore scale to reservoir scale.

1.2.2.1 Microscale: Laboratory Analysis on Plugs

Measurements qualifying fundamental petrophysical properties are primarily acquired at micro- to centimeter scales through Conventional Core Analysis (CCA). These measurement methods, such as porosity through fluid saturation or helium pycnometry, gas permeability, and grain density, quantify the storage and transmission capacity of reservoirs (McLallen, 1959).

At this scale, original pore distribution or that resulting from dissolution and precipitation processes can be characterized through Special Core Analysis (SCAL). Mercury Injection Capillary Pressure (MICP) techniques reveal pore size distribution and pore network connectivity, while NMR (Nuclear Magnetic Resonance) provides signatures of fluids in place and their mobility in different pore environments (Gu et al., 2024). Dynamic flow simulations on plugs, including relative permeability curves and displacement tests, reconstruct the history of fluid-rock interactions and understand flow behavior under near-reservoir conditions (Hu et al., 2022).

1.2.2.2 Mesoscale: Well Logging and Facies Modeling

Information extracted from plugs, as in sedimentology, typically represents punctual petrophysical properties and does not allow deduction of extended distribution for constructing a petrophysical model.

Conventional logs (gamma-ray, neutron, density, resistivity) recorded continuously along wells provide indications of lithological nature, porosity volume, nature of inner fluids, and invasion of drilling fluids in a close radius around wells (Chombart, 1960). Variation in recorded signals

generally reveals characteristic patterns for temporally correlating different intervals across multiple wells or identifying source, reservoir, or seal units.

Borehole Images (BHI) can be acquired over the entire drilled section and complement core descriptions by observing and picking planes highlighted by contrast variations, revealing structural directions not accurately interpretable from core images (Yousef & Morozov, 2022; Zou et al., 2020).

Production Logging Tools (PLT) complete this approach by identifying contributing zones and their relative flow rates (Al-Khalifa et al., 2019; Hamoudi et al., 2019).

These quantitative approaches typically employ algorithmic clustering to integrate point data from core plugs and well logs, defining petrophysical facies groups for reservoir property modeling (Aliouane et al., 2013; Aliouane & Ouadfeul, 2018; Stadtmüller & Jarzyna, 2023).

1.2.2.3 Macroscale: Well Tests and Reservoir Connectivity

Petrophysical facies boundaries and variations create heterogeneities that either enhance or impede fluid flow. However, well logs cannot reveal effective lateral connectivity between units.

Well tests such as DST (Drill Stem Tests) are typically required to evaluate connectivity around wells over large radii and define flow parameters at drainage scale (Timmerman & van Poolen, 1972). These tests determine effective permeability, permeability heterogeneities within the investigation radius, and potential flow barriers. Pressure transient curve analysis reveals flow geometry and permeable body extent.

1.2.3 Integration and Reservoir Modeling

Once petrophysical facies are defined and nearby flow units identified, it is necessary to hierarchize heterogeneities to establish a reservoir model. The spatial distribution of observed reservoir behaviors is guided by conceptual laws from geological interpretation jointly linking evolution of sedimentary facies to petrophysical facies. Although this approach offers explicability beyond what petrophysics alone can provide through mathematical property propagation laws, the fusion of sedimentary and petrophysical facies via rock-typing processes presents significant challenges (Bize-Forest et al., 2014; Hollis et al., 2010; Krivoshchekov et al., 2023; Raznitsyn et al., 2022; Rebelle, 2014; Wibowo & Permadi, 2013). These difficulties emerge when sedimentological facies outnumber petrophysical facies or when fundamental incompatibilities exist between the two classification systems.

Geological studies typically yield qualitative to semi-quantitative interpretations, as precise measurements cannot be obtained within reasonable timeframes. Multiple biases and uncertainties are consequently often difficult to assess. Information reconciliation generally occurs during rock-type construction and reservoir modeling, where multiple scenarios are developed to identify the most appropriate representation. Fine subdivision of locally assigned model properties requires cell upscaling using specific laws that are often limited by global averaging, which can obscure previously identified fine-scale heterogeneities (Madalimov et al., 2024; Thomas et al., 2023).

Once reservoir models are constructed, flow simulations can be performed to monitor field resource recovery and make forecasts to anticipate potential investments for maintaining production plateaus until reaching depletion levels that no longer allow recovery.

1.3 Phoenix.AI for AI-Enhanced 2G&R Synthesis

Since the massive popularization of AI over the past decade, deep learning models have demonstrated significant efficiency in information processing for vastly diverse applications and modalities (structured data, images, texts, etc.) (Lecun et al., 2015).

Current reservoir characterization methods encounter specific limitations that could potentially be overcome through the integration of AI approaches to address these challenges.

Geoscience datasets from thin sections, cores, and outcrops contain rich information but remain underexploited because traditional descriptive analysis requires extensive human time investment. While these materials can now be digitized, only a few computational tools exist specifically designed for processing and analyzing this type of data.

The Phoenix.AI project focuses on developing machine learning algorithms designed to create enhanced descriptions of imaging data across multiple scales for geological synthesis and reservoir characterization (Figure 1.2). An aspiration is to build explainable methods that align with established geoscience practices, thereby building user confidence. The expected high-resolution descriptions would enable better understanding of data organization for filtering and detailed distribution analysis while integrating petrophysical measurements acquired at corresponding image locations. This approach would identify and hierarchize heterogeneity classes more consistently across different resolution levels. Using description proxies to find image-petrophysics correlation laws is necessary since image-measurement pairs are generally rare and therefore difficult to approximate using AI methods requiring large example quantities. Beyond predictive capability, explicability is primarily sought in image descriptions that are themselves relatable to original rock formation processes.

Constituting a common encoding framework for sedimentological descriptions and petrophysical measurements appears to be the next critical step for developing successive upscaling laws that enable progression across scales, ultimately leading to robust rock-type model construction. This methodology would guarantee that well-characterized heterogeneities are faithfully captured in reservoir models while maintaining interpretability through the synthesis of various underlying processes.

The Phoenix.AI project for 2G&R Synthesis

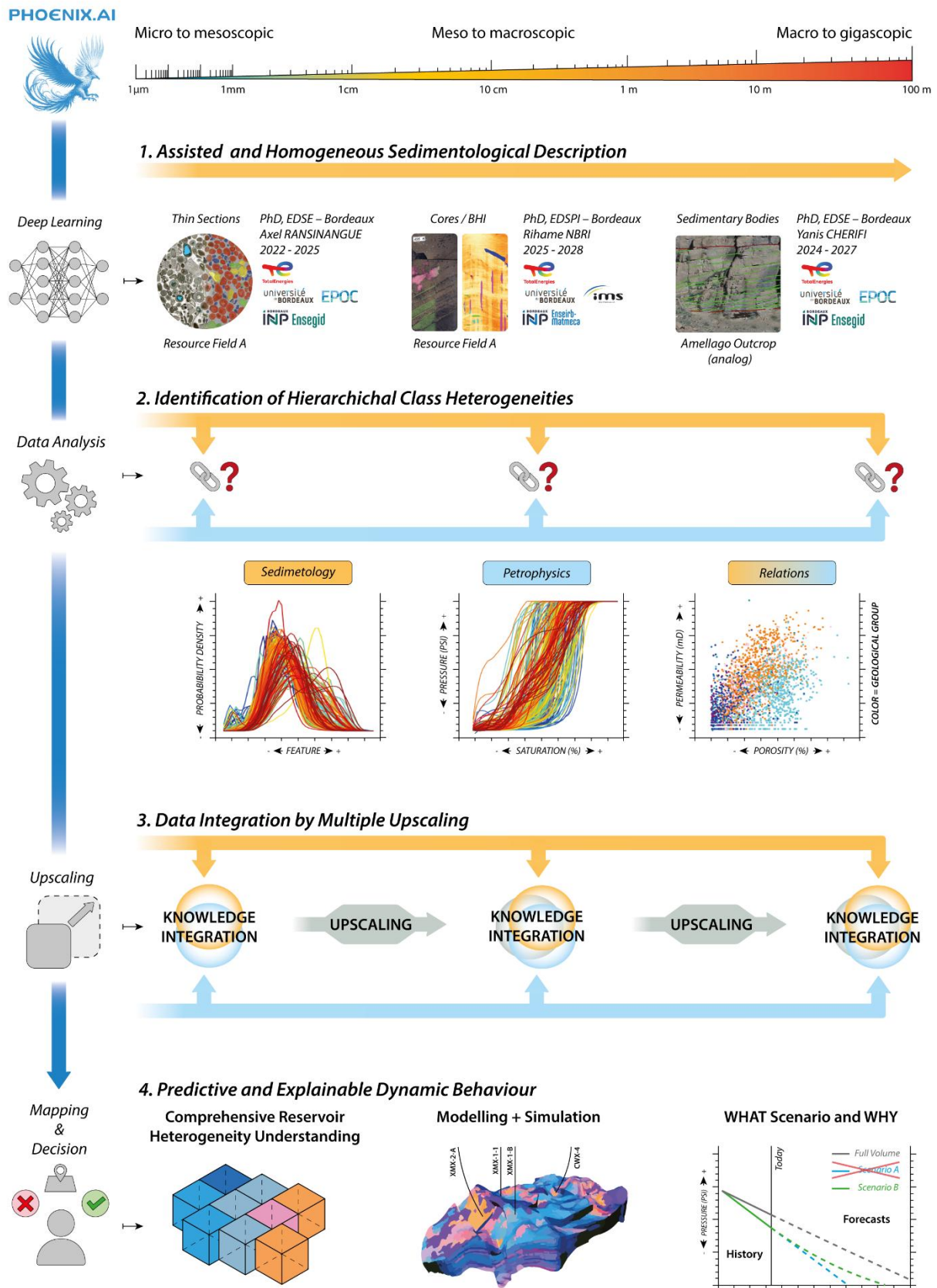


Figure 1.2 The Phoenix.AI project for deep learning approaches integration into geological reservoir characterization.

1.4 Research Framework

This thesis is the first in a three-part series investigating geological information extraction from ooid-dominated carbonate systems.

Theses Timeline:

1. **Axel RANSINANGUE (2022-2025)**: Deep learning for carbonate rocks petrography.
2. **Rihame NBRI (2025-2028)**: Automatic extraction of facies and sedimentary bodies from core photographs and borehole imagery in oolitic carbonate systems.
3. **Yanis CHERIFI (2024-2027)**: Qualitative and quantitative analysis of carbonate sand bodies in stratigraphic records from 3D outcrops, multiscale index extraction for characterizing underground reservoir heterogeneities.

While the complete series will address different scales of image acquisition, this initial study focuses specifically on thin section images, which serve as the primary data source for microscale rock characterization.

Petrographic analysis of carbonate thin sections relies on extensive knowledge bases covering general concepts and standardized nomenclatures to extract relevant parameters for various analytical tasks. While these traditionally qualitative to semi-quantitative methods support descriptive work, they become limited when precise sample differentiation is needed based on characteristics that could theoretically be quantified.

Computer vision and machine learning models offer promising solutions for accelerating image recognition and quantification tasks. This would allow geologists to shift their focus from time-intensive manual sample description to higher-level interpretation and contextual analysis of samples within sedimentary frameworks. However, training such models requires large, annotated image databases, and no comprehensive reference database currently exists in the published literature.

1.5 Question and Objectives

The various reflections on the subject led to the formulation of the following problem :

How can deep learning-based computer vision enable quantitative petrographic analysis of carbonate rock thin sections using initial limited annotated databases?

To answer this question, a work plan was formulated addressing different thesis objectives and stages:

- State of the art of traditional and digital methods used in carbonate thin section petrography.
- Data collection and preparation.
- Construction of initial learning datasets for classification and segmentation of thin section images.
- Development of data augmentation methods through procedural synthetic image generation for classification and segmentation.
- Investigation and implementation of diverse architectures for image classification and segmentation.
- Domain adaptation for optimizing image segmentation methods through enhanced and adaptive synthetic data generation.
- Concatenation of developed methods for multiscale quantitative thin section characterization approaches.

This work follows a 3-year research period marked by annual progress committees, participation in conferences and scientific exchanges, as well as the production of papers related to the contributions made to the field of study (Figure 1.3).

PhD Timeline

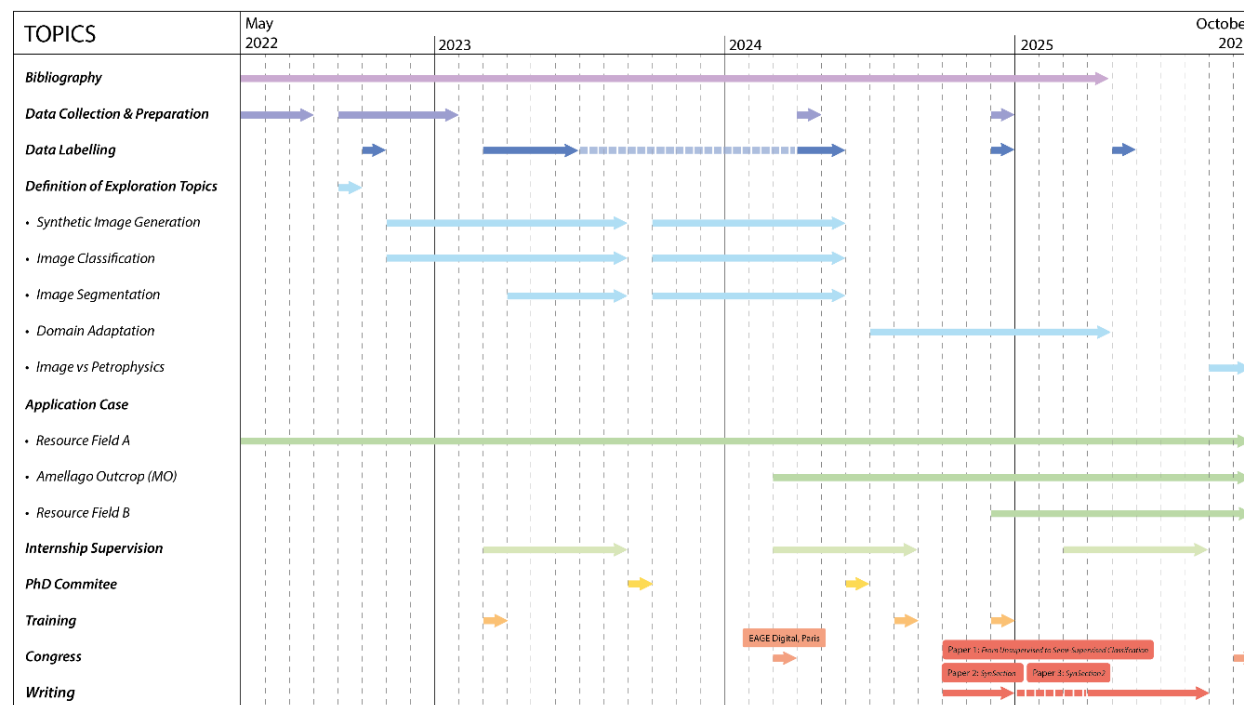


Figure 1.3 PhD timeline showing the periods devoted to each phase of the project.

REFERENCES

- Aliouane, L., & Ouadfeul, S.-A. (2018). Automatic lithofacies classification from well-logs data using the Walsh transform combined with the self-organizing map. *ECMOR XVI - 16th European Conference on the Mathematics of Oil Recovery*, Barcelona, Spain. <https://doi.org/10.3997/2214-4609.201802174>
- Aliouane, L., Ouadfeul, S.-A., Djarfour, N., & Boudella, A. (2013). Lithofacies prediction from well logs data using different neural network models. *International Conference on Pattern Recognition Applications and Methods (ICPRAM)*, Barcelona, Spain, 702-706. <https://doi.org/10.5220/0004265407020706>
- Al-Khalifa, N. F., Nugroho, C., Memiche, L., Al-Blooshi, A., Tiwary, A., Al-Moraikhi, H., & Joshi, D. (2019). Conforming production logging tool with high-resolution electrical imager in carbonate reservoir. *SPE Middle East Oil and Gas Show and Conference*. <https://doi.org/10.2118/194738-MS>
- Bize-Forest, N., Baines, V., Boyd, A., Moss, A., & Oliveira, R. (2014). Carbonate reservoir rock typing and the link between routine core analysis and special core analysis. *SCA International Symposium*, Avignon, France, SCA2014-003.
- Burchette, T. P., Paul Wright, V., & Faulkner, T. J. (1990). Oolitic sandbody depositional models and geometries, Mississippian of southwest Britain: Implications for petroleum exploration in carbonate ramp settings. *Sedimentary Geology*, 68(1), 87–115. [https://doi.org/10.1016/0037-0738\(90\)90121-9](https://doi.org/10.1016/0037-0738(90)90121-9)
- Chombart, L. G. (1960). Well logs in carbonate reservoirs. *Geophysics*, 25(4), 779–853. <https://doi.org/10.1190/1.1438766>
- Cross, N. E., & Burchette, T. (2024). The impact of carbonate reservoir heterogeneity on the development of Middle Eastern oilfields. *First EAGE Workshop on Advances in Carbonate Reservoirs from Prospects to Development*. <https://doi.org/10.3997/2214-4609.202412109>
- Demicco, R. V., & Hardie, L. A. (1995). *Sedimentary structures and early diagenetic features of shallow marine carbonate deposits*. SEPM Society for Sedimentary Geology. <https://doi.org/10.2110/sepmatl.01>
- Flügel, E. A. (2004). *Microfacies of carbonate rocks: Analysis, interpretation and application*. Springer-Verlag, Berlin.
- Fogg, G. E., & Lucia, F. J. (1990). *Reservoir modeling of restricted platform carbonates: Geologic/geostatistical characterization of interwell-scale reservoir heterogeneity, Dune Field, Crane County, Texas*. Bureau of Economic Geology, University of Texas at Austin, Report of Investigations No. 194.
- Gu, Z., Wang, S., Guo, P., & Zhao, W. (2024). Utilizing differences in mercury injection capillary pressure and nuclear magnetic resonance pore size distributions for enhanced rock quality evaluation: A Winland-style approach with physical meaning. *Applied Sciences*, 14(7), 2759. <https://doi.org/10.3390/app14072759>
- Hamoudi, M., Humoodi, A., & Mohammed, B. (2019). Oil well testing using production logging tool in Khurmala field in Kurdistan Region-Iraq. *UKH Journal of Science and Engineering*, 3(2), 1-10. <https://doi.org/10.25079/ukhjse.v3n2y2019.pp1-10>

- Harris, P. M. (2019). Lessons from a modern carbonate sandbody – A personal experience of comparative sedimentology. *The Depositional Record*, 5(3), 438–450. <https://doi.org/10.1002/dep2.82>
- Harris, P. M., Purkis, S. J., & Ellis, J. (2011). Analyzing spatial patterns in modern carbonate sand bodies from Great Bahama Bank. *Journal of Sedimentary Research*, 81(3), 185–206. <https://doi.org/10.2110/jsr.2011.21>
- Harris, P. M., St. C. Kendall, C. G., & Lerche, I. (1985). Carbonate cementation - A brief review. In *Carbonate Cements* (pp. 79–95). SEPM (Society for Sedimentary Geology). <https://doi.org/10.2110/pec.85.36.0079>
- Herbert, C., Alexander, J., & de Álvaro, M. J. M. (2015). Reservoir heterogeneity resulting from bottomset deposition in the lee of dunes and unit bars. *AAPG Bulletin*, 99(8), 1437–1462. <https://doi.org/10.1306/02241514209>
- Hollis, C., Vahrenkamp, V. C., Tull, S. J., Mookerjee, A., Taberner, C., & Huang, Y. (2010). Pore system characterisation in heterogeneous carbonates: An alternative approach to widely-used rock-typing methodologies. *Marine and Petroleum Geology*, 27(4), 772–793. <https://doi.org/10.1016/j.marpetgeo.2009.12.002>
- Hu, Y., Wei, J., Li, T., Zhu, W., Gong, W., Hui, D., & Wang, M. (2022). Numerical simulation of fluid flow in carbonate rocks based on digital rock technology. *Energies*, 15(10), 3748. <https://doi.org/10.3390/en15103748>
- Krivoshchekov, S. N., Kochnev, A., Kozyrev, N., Botalov, A., Kochneva, O., & Ozhgibesov, E. (2023). Rock typing approaches for effective complex carbonate reservoir characterization. *Energies*, 16(4), 1840. <https://doi.org/10.3390/en16041840>
- LeCun, Y., Bengio, Y., & Hinton, G. (2015). Deep learning. *Nature*, 521(7553), 436–444. <https://doi.org/10.1038/nature14539>
- Madalimov, M., Li, D., Haynes, B., Dair, Y., Ehighebolo, I. T., & Ionescu, C. L. (2024). Heterogeneity modeling and heterogeneity-based upscaling for reservoir characterization and simulation. *SPE Journal*, 29(5), 2358–2375. <https://doi.org/10.2118/218383-PA>
- McLallen, D. L. (1959). *The characterization of porous media by porosity, permeability, formation factor, and mercury injection capillary*. PhD dissertation, University of Texas at Austin.
- Mohammed, A., & Velledits, F. (2024). Microfacies impacts on reservoir heterogeneity of early Cretaceous Yamama carbonate reservoir in South Iraq. *Scientific Reports*, 14(1), 6847. <https://doi.org/10.1038/s41598-024-57370-4>
- Palermo, D., Aigner, T., Nardon, S., & Blendinger, W. (2010). Three-dimensional facies modeling of carbonate sand bodies: Outcrop analog study in an epicontinental basin (Triassic, southwest Germany). *AAPG Bulletin*, 94(4), 475–512. <https://doi.org/10.1306/08180908168>
- Pomar, L., & Kendall, C. C. G. (2008). Architecture of carbonate platforms: A response to hydrodynamics and evolving ecology. In *Controls on Carbonate Platform and Reef Development* (pp. 187–216). SEPM Special Publication No. 89. <https://doi.org/10.2110/pec.08.89.0187>
- Ravenne, C. (2002). Sequence stratigraphy evolution since 1970. *Comptes Rendus Palevol*, 1(6), 415–438. [https://doi.org/10.1016/S1631-0683\(02\)00068-4](https://doi.org/10.1016/S1631-0683(02)00068-4)

- Raznitsyn, A. V., Putilov, I. S., & Melkishev, O. (2022). Comprehensive approach to carbonate rocks typing based on core laboratory experiments data. *IOP Conference Series: Earth and Environmental Science*, 1021, 012047. <https://doi.org/10.1088/1755-1315/1021/1/012047>
- Rebelle, M. (2014). Rock-typing in carbonates: A critical review of clustering methods. *SPE Reservoir Characterisation and Simulation Conference and Exhibition*, Abu Dhabi, UAE. <https://doi.org/10.2118/171759-MS>
- Stadtmüller, M., & Jarzyna, J. A. (2023). Estimation of petrophysical parameters of carbonates based on well logs and laboratory measurements, a review. *Energies*, 16(4), 1787. <https://doi.org/10.3390/en16041787>
- Tatum, D., Nursaidova, R., Hern, C. Y., Westerman, A. R., Francke, J., & Gardiner, A. R. (2012). Reservoir impact of small-scale aeolian dune architecture - From acquisition to simulation in the Wahiba Sands, Oman. *AAPG Annual Convention and Exhibition*, Long Beach, California, USA.
- Tavakoli, V. (2019). *Carbonate reservoir heterogeneity*. Springer. <https://doi.org/10.1007/978-3-030-35392-2>
- Thomas, H., Brigaud, B., Blaise, T., Zordan, E., Zeyen, H., Catinat, M., Andrieu, S., Mouche, E., & Fleury, M. (2023). Upscaling of geological properties in a world-class carbonate geothermal system in France: From core scale to 3D regional reservoir dimensions. *SSRN Electronic Journal*. <https://doi.org/10.2139/ssrn.4480152>
- Timmerman, E. H., & van Poolen, H. K. (1972). Practical use of drill-stem tests. *Journal of Canadian Petroleum Technology*, 11(4), 187-193. <https://doi.org/10.2118/72-04-04>
- Tucker, M. (2009). *Carbonate sedimentology*. John Wiley & Sons. <https://doi.org/10.1002/9781444314175.ch1>
- Wibowo, A. S., & Permadi, P. (2013). A type curve for carbonates rock typing. *SPE/IATMI Asia Pacific Oil & Gas Conference and Exhibition*, Jakarta, Indonesia. <https://doi.org/10.2118/165495-MS>
- Yılmaz, İ. Ö., & Altın, D. (2001). Use of sedimentary structures in the recognition of sequence boundaries in Upper Jurassic–Upper Cretaceous peritidal carbonates of the Western Taurides, Turkey. *International Geology Review*, 43(8), 736–753. <https://doi.org/10.1080/00206810109465040>
- Yousef, I., & Morozov, V. (2022). Fractures characterization of carbonate reservoir using core analyses and borehole image logs. *Petroleum Science and Technology*, 40(14), 1772–1797. <https://doi.org/10.1080/10916466.2022.2042725>
- Zou, X., Wang, C., Wang, Y., & Song, H. (2020). Morphological feature description method of structural surface in borehole image during in-situ instrumentation. *Rock Mechanics and Rock Engineering*, 53(7), 2947–2956. <https://doi.org/10.1007/s00603-020-02078-0>

CHAPTER 2: STATE OF THE ART

2.1 Carbonate Rock Analysis

Carbonate rock classification systems were developed to convey comprehensive geological information through concise terminology, eliminating the need for lengthy detailed descriptions. These frameworks have remained relatively stable since their inception between the late 1950s and late 1980s (Dunham, 1962; Folk, 1959; Sibley & Gregg, 1987). The foundational schemes established by R. Folk and R. Dunham, which primarily addressed limestones, complemented by D. Sibley and J. Gregg's detailed approach to dolomites, continue to serve as the cornerstone of modern carbonate petrography, with recent refinements targeting specialized applications in sedimentology, or reservoir characterization. Contemporary advances have increasingly incorporated numerical methods and machine learning approaches to enhance classification accuracy and migrated towards assisted processes.

2.1.1 Limestone Classification

2.1.1.1 Folk's Framework

R. Folk introduced a classification system proposing that most limestone rocks could be categorized based on their constituent particles (allochems) and enclosing material (orthochems) (Figure 2.1), analogous to siliciclastic rocks (Folk, 1959).

Primary Limestone Constituents

Allochems	Fossils (bioclasts)	Intraclasts	Ooids	Pellets (peloids)	Orthochems	Micrite	Sparry Calcite (spar)
Schematic Illustration					Schematic Illustration		
Size Range	10s μm (microscopic foraminifera) – few cm (macroscopic shells)	100s μm – few cm	250 μm – 2 mm diameter	< 250 μm diameter	Crystal Size	< 4 μm	> 10 μm
Shape	Convex to non-convex with rounded to sharp edges	Angular to subrounded	Spherical	Ellipsoidal to spherical	Color	Dark to light brown	Light grey to cream
Internal Structure	Complex geometries and patterns	Homogeneous fine-grained to heterogeneous with multi-components	Concentric radial, tangential, or mixed fabric arrangements	Homogeneous fine-grained without obvious lamination	Texture	Massive microcrystalline	Coarse crystalline calcite cement
Origin	Skeletal remains of carbonate-secreting organisms	Reworked fragments of previously lithified carbonate sediment	Formed by precipitation around a nucleus in agitated, shallow and warm-water	Fecal pellets, micritized grains, or precipitates	Formation Conditions	Quiet water settings with carbonate supersaturation	Early marine to late burial diagenetic phases where fabric types (equant, bladed, fibrous, or syntaxial overgrowth) indicate specific diagenetic conditions
Environment Indicator	Taxonomic diversity may reflect paleoenvironmental conditions (water depth, temperature, etc.) and age	Indicates erosion of nearby carbonates substrates	High-energy, shallow, warm-water conditions	Strong biological activity and moderate energy			

Figure 2.1 Typical components and corresponding characteristics observed in limestones thin sections (microscopic scale). Allochemical constituents are divided in various subtypes (fossils, intraclasts, ooids and pellets) as well as orthochemical constituents (micrite and sparry calcite).

The systematic estimation of allochemical and orthochemical components proportion notably allowed to establish relationships with formation mechanisms and sediment transport history (Flügel, 2004; Tucker, 2009). This framework provided foundational rules for detailed description at the thin section scale by grouping samples according to shared characteristics (Figure 2.2).

*Original Folk's Limestone Classification
(1959)*

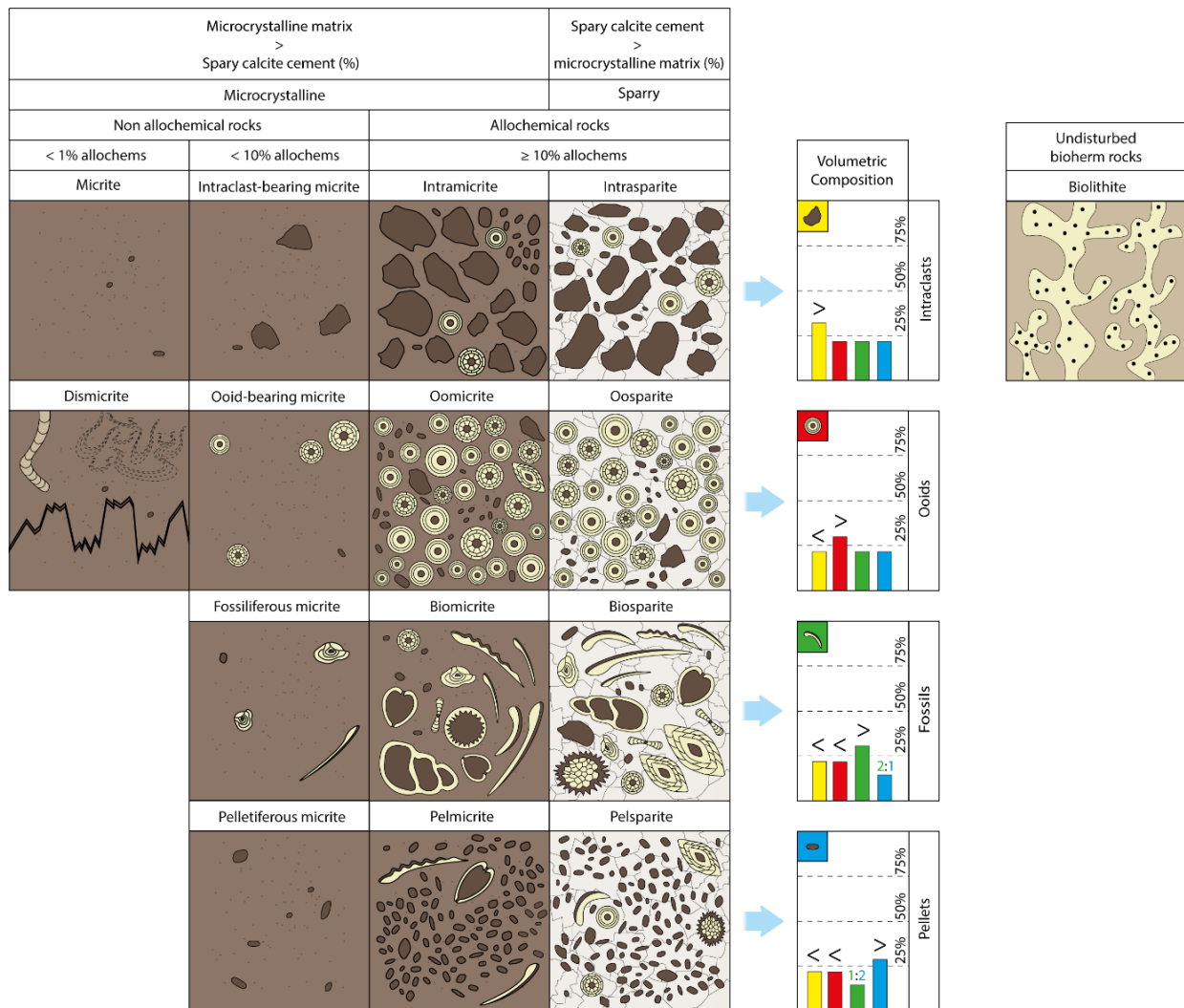


Figure 2.2 Folk's limestone classification framework. The system uses visual proportion estimation of allochemical types and othochemical content for categorizing samples.

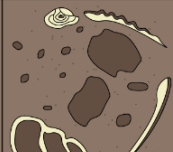
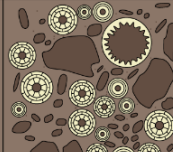
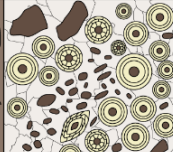
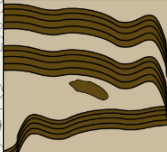
Nevertheless, its complexity prevented widespread adoption and led to the development of a simplified system with R. Dunham's approach.

2.1.1.2 Dunham's Framework

R. Dunham's classification framework also relies on the principle that carbonate components provide essential insights into energy conditions and environmental settings at multiple scales of observation (Dunham, 1962). The main advancements proposed through this system lie in reducing the number of classes and descriptive parameters, particularly regarding proportion estimation for each allochem class, which was simplified for estimating the general support of allochems regardless of their origin. The distinction between orthochem types is preserved, and a grain size boundary (2 mm) is also used to refine the classes, as well as fabric types for components bound during deposition in its extended version (Embry & Klovan, 1971). Consequently, terminology has evolved with the notion of depositional textures being introduced to refer to specific ranges of characteristics enumerated for each defined class (Figure 2.3). This

system was specifically designed to provide a framework that could be used both in the field and under microscopic analysis.

Original Dunham's Limestone Classification (1962)

Recognizable depositional texture				
Components not bound together organically during deposition				Components bound together organically during deposition
Mud-supported < 50% grains		Grain-supported ≥ 50% grains		
< 10% grains	≥ 10% grains	Contains mud	Lacks mud	
Mudstone	Wackestone	Packstone	Grainstone	Boundstone
				

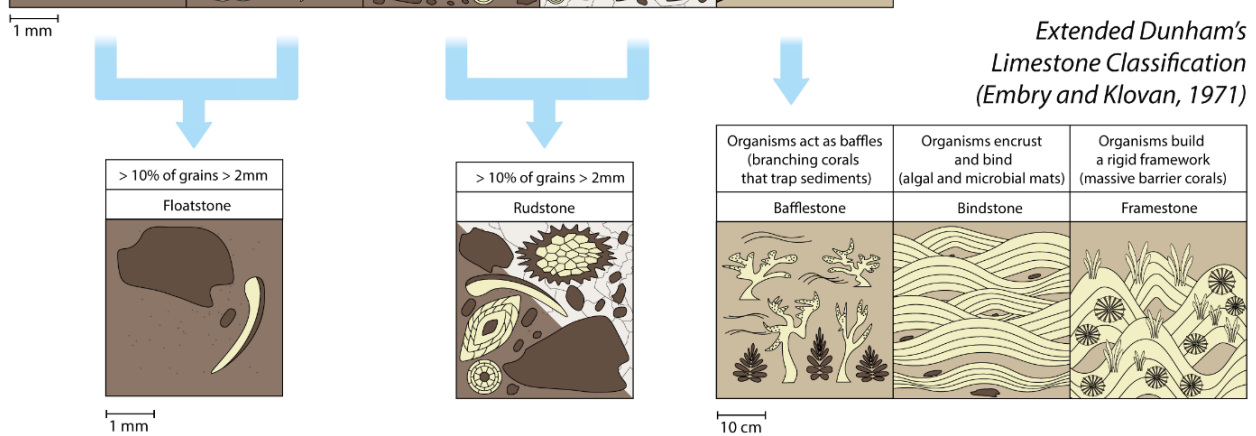


Figure 2.3 Dunham's limestone classification framework. The system establishes the principle of depositional texture using proportion of grains and groundmass material for principal categorization. Subclasses were later introduced for samples presenting particles > 2mm or rocks organically built.

The classification methods described were initially developed for limestone rocks, though the sedimentary systems in which they form undergo diagenetic processes that can dissolve and reprecipitate minerals such as dolomite, which necessitated the development of dedicated categorization systems.

2.1.2 Dolomite Classification

2.1.2.1 Folk's Framework

Dolomite classification, introduced by R. Folk, is primarily based on the preservation of allochem ghosts and crystal size criteria (Folk, 1962) (Figure 2.4). These features prove essential for gaining insight into the conditions to which these rocks were exposed (Flügel, 1982).

Original Folk's Dolomite Classification
(1962)

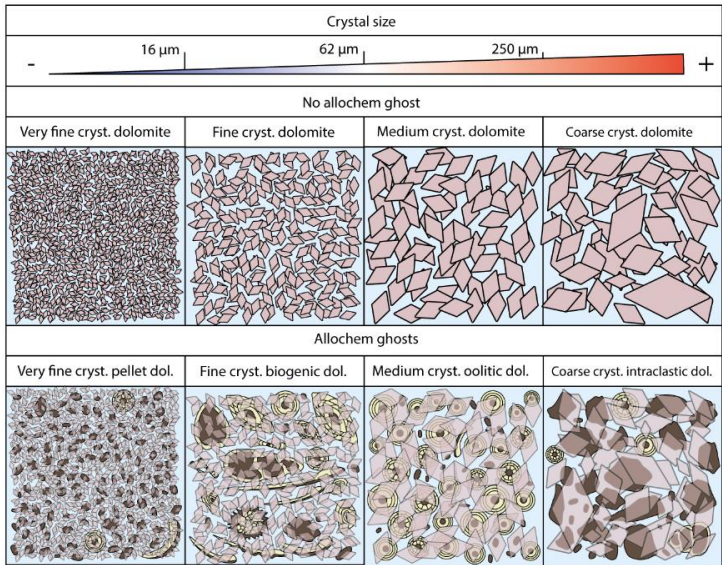


Figure 2.4 Folk's dolomite classification framework. The system relies on original ghost-preserved grain structures and crystal size criteria.

2.1.2.2 Sibley and Gregg's Framework

To improve reconstruction of historical nucleation, temperature, and fluid composition evolution, D. Sibley and J. Gregg reinterpret the classification framework of R. Folk by incorporating crystal size modality, crystal morphology, type and volume of replaced material patterns (Figure 2.5), establishing correlations between these parameters and physicochemical events (Sibley & Gregg, 1987).

Original Sibley and Gregg's Dolomite Classification
(1987)

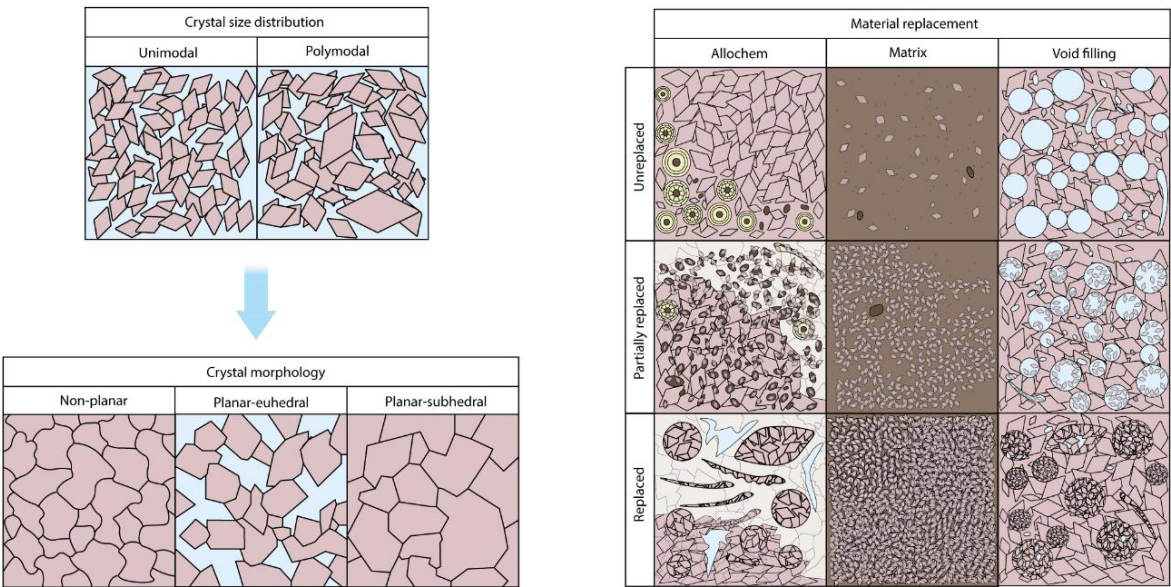


Figure 2.5 Sibley and Gregg's dolomite classification framework. The system refers to crystal size modality distribution, crystallization geometry and replaced material to provide various description combination possibilities.

2.1.3 Methodologies Discussion

Although the classification tools built over recent decades have provided a structured framework for understanding and characterizing carbonate rock formation mechanisms, these methods are subject to certain cognitive biases.

Visual estimations form the basis for most analytical methods used in petrographic imaging. While graphical charts sometimes aid, these estimations are generally imprecise and create challenges when assessing material proportions and uncertainties, particularly when comparing samples with varying compositional or spatial distribution patterns (Figure 2.6) (Bercean et al., 2023).

Cognitive Biases in Carbonate Thin Section Image Analysis

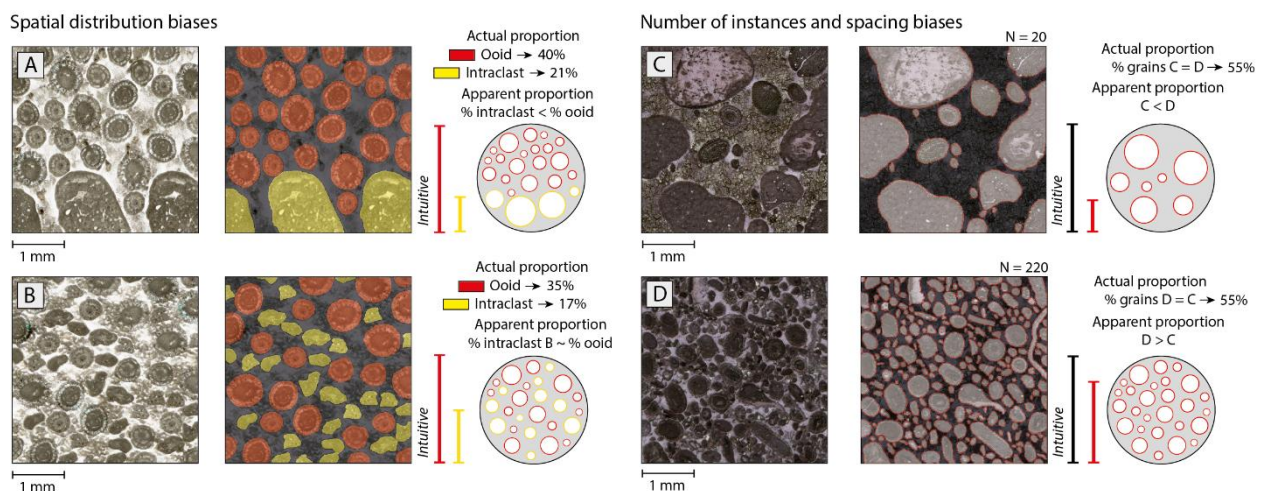


Figure 2.6 Examples of cognitive biases observed in thin section image analysis. The two cases highlight the difficulty of evaluating the relative grain proportion in terms of occupied surface. Sample A allows a good proportion estimation related to the 1-dimensional preferential distribution (vertical). In sample B, where grain classes spatial distribution is mixed, ooids and intraclasts seems to be represented in similar proportion, however since intraclasts shape is irregular they do not maximize the occupied surface such as ooids and are less represented. Sample C and sample D highlight a bias related to the number of grains and their spacing where percentage of grains in sample C seems inferior to sample D but is equal.

R. Folk's classification can therefore be defined as impractical when allochem proportion estimations must be made for multiple element classes, potentially introducing cumulative sources of error. Conversely, R. Dunham's classification has maintained its dominance in carbonate geology due to its simplicity and practical applicability. Nevertheless, this method is not without limitation. Like any classification system, processing difficulties arise when handling examples with inherent complexity, including shared characteristics across different classes or proximity to class boundaries, which also makes it sometimes difficult to identify an optimal scale of observation for comprehensive analysis (Figure 2.7).

Maintaining consistent and accurate descriptions then becomes challenging over time and studies. This issue has been particularly evident in carbonate thin section analysis, with description heterogeneities observed across various levels of operator expertise (from bachelor's degree to specialist) (Figure 2.8) (Lokier & Al Junaibi, 2016).

Encountered Classification Complexities for Limestone Rocks

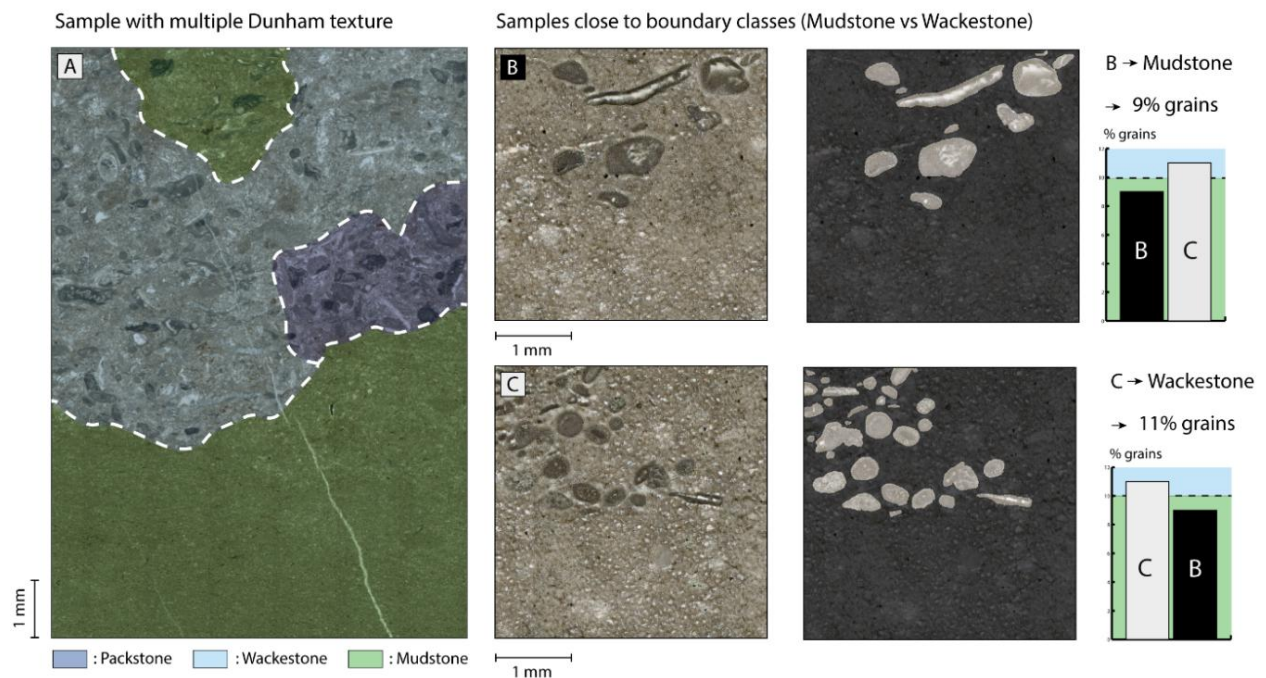


Figure 2.7 Examples showing processing difficulties in limestone thin section classification task. Sample A presents multiple rock textures depending on the scale of observation. Sample B and sample C, which simply differ from a $\pm 2\%$, must result in different naming based on Dunham classification rules, however it is currently almost visually impossible to clearly discriminate.

Classification Heterogeneities in Carbonate Thin Section Analysis (Lokier et al., 2016)

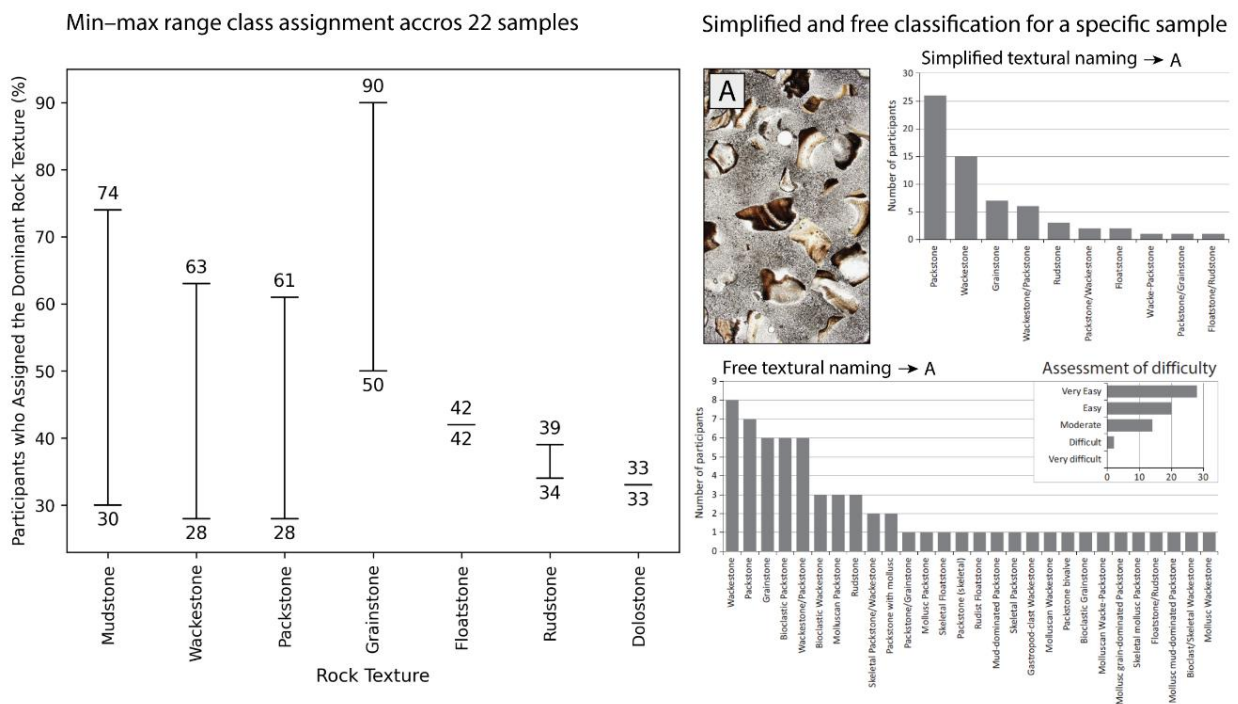


Figure 2.8 Heterogeneities of classification highlighted in carbonate thin section analysis by Lokier's study.

Despite methodological advances in distinguishing mineral formation processes, dolomite classification rules encounter limitations like those found in limestone classification. The microscopic heterogeneity commonly present in these rocks creates multiple textures within individual samples, complicating reliable crystal size measurements. Consequently, ghost material preservation criteria remain the only readily observable and robust diagnostic feature.

2.1.4 Numerical Methods Integration

The exponential evolution of digital methods for image characterization has benefited numerous fields, including geosciences and carbonate rocks petrography. Early applications employed feature extraction methods (pixel distribution, edge detection, convexity analysis, etc.) for classifying thin sections (Marmo et al., 2005). Subsequently, the complexities associated with feature engineering led to renewed focus on data-driven deep neural network solutions that are built to handle this step on their own. The progressive adoption of deep learning methods over the past decade has enabled new applications, addressing diverse research challenges.

The primary obstacle faced with implementing neural network models lies in constructing annotated datasets for supervised or transfer learning. Annotating thin sections, whether through global image tagging or precise grain boundary delineation, is time-intensive and costly (Table 2.1), yet remains an essential step for validating investigated approaches. The annotation process has been the subject of extensive optimization efforts to maximize efficiency given its significant cost. Moreover, once obtained, these datasets unlock numerous analytical possibilities.

Labelling Methodology	Traditional Manual Approach		AI-Assisted Process	
	Standard Image Classification using Expert Knowledge	Detailed Allochems Annotation with Manual Mask Creation	Image Classification Enhanced with Pre-clustering Algorithms	Assisted Allochems Annotation using Segment Anything Model (SAM)
Processing Rate (images per day)	250 to 500 images	1 to 2 images	500 to 1000 images	10 to 15 images
Quality & Efficiency Assessment	High precision with sometimes too many created classes. Time-intensive sequential processing and potential annotation drift	Comprehensive geological feature identification varying on expertise-level and thin section quality. Generally lack precision for grain contours	Enables batch classification with standardized criteria that reduce annotation variability	Superior mask precision with automated boundary detection allowing focus on grain classification

Table 2.1 Comparison of manual versus AI-assisted labelling approaches for carbonate thin section analysis, presenting processing objectives, employed methods, average processing rates observed during the study period, and quality-efficiency assessments. AI-assisted image classification implements the methodology developed in *Paper 1* (chapter 3), while allochem annotation utilizes a specialized labelling interface integrated with Meta's Segment Anything Model (SAM).

Author(s) and Year	Classification	Classes	Images	Method	Architecture	Result
Marmo et al., 2005	Dunham	4	400	Feature-based	MLP	94% (Acc.)
Budenny et al., 2017	Lithology (lime./dol./sand.)	3	–	Feature-based	Decision Tree	96% (Acc.)
Jobe et al., 2018	Dunham	6	9,000	Transfer Learning	Inception-v3	83% (Acc.)
Pires de Lima et al., 2019	Micro fossils	7	1,850	Transfer Learning	Inception-v3 MobileNetV2	100% (Acc.)
Pires de Lima et al., 2019	Reservoir quality	3	161	Transfer Learning	Inception-v3 MobileNetV2	81% (Acc.)
Pires de Lima et al., 2020	Mudstone subtypes	5	5,515	Transfer Learning	VGG-19	93% (Acc.)
Koeshidayatullah et al., 2020	Allochems/orthochems	6	13,000	Transfer Learning	VGG-16	92% (F1)
Shebl et al., 2021	Dunham	6	1,658	Transfer Learning	GoogLeNet	96% (Acc.)
Mosser et al., 2022	Dunham	6	66	Transfer Learning	ResNet-18	59% (F1)
Liu et al., 2023	Dunham	5	758	Transfer Learning	DenseNet-121	89% (Acc.)

Table 2.2 State-of-the-art machine and deep learning approaches for carbonate thin section classification. (Acc. : accuracy)

Given that the primary challenge in carbonate thin section characterization stems from data categorization, research has naturally gravitated toward image classification models (Table 2.2). However, most studies are limited to R. Dunham's classification system and rarely integrate other nomenclatures or lithologies such as dolomites or secondary mineralization like anhydrite, features that are commonly observed in carbonate rocks. Omitting these components at this stage implies that a preliminary sorting of major lithologies has already been performed, though this entire process could potentially be accomplished in a single, integrated step.

While results from these studies show promise, they generally fail to address how bias could be mitigated in training datasets, particularly regarding granular limestone classifications. Since training datasets are typically constructed manually, no established methods exist to verify accurate labeling with certainty, especially near end member classes. The existence of gaps in learned representations remains unverifiable, potentially leading to model confusion when the full range and balance of class possibilities are not covered during training.

Author(s) and Year	Detection/ Segmentation	Images	Objects	Classes	Architecture	Result
Koeshidayatullah et al., 2020	Object detection	4,000	13,000	9	Faster RCNN (Inception-ResNet-v2)	84% (mAP)
Nanjo et al., 2020	Segmentation	306	–	4	U-Net (Adversarial)	84% (Acc.)
Koeshidayatullah et al., 2022	Segmentation (ooids)	–	–	2	Autoencoder	–
Liu et al., 2023	Segmentation (oncoids)	17	–	2	U-Net	70% (Acc.) 84% (IoU)
Wang et al., 2023	Object detection (ooids)	103	–	2	Faster RCNN (ResNet50)	92% (mAP)

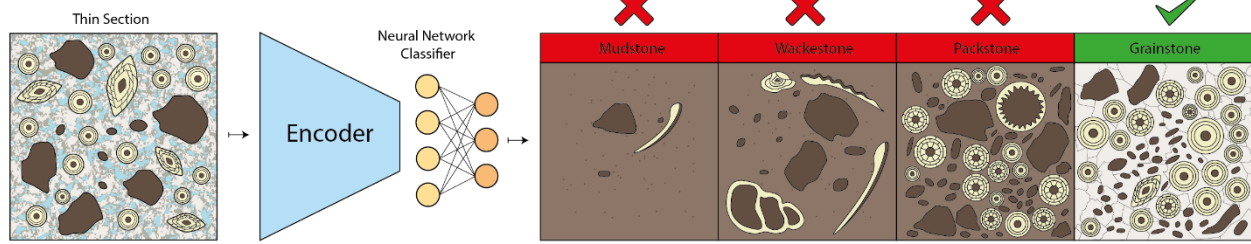
Table 2.3 State-of-the-art machine learning approaches for carbonate thin section object identification and classification. (Acc. : accuracy, IoU: intersection over union, mAP: mean average precision)

Sedimentological analysis involves interpreting information across multiple scales, making sophisticated classification systems like R. Folk's framework crucial for distinguishing between sample contents and understanding their formative processes. Recent advances in object detection methods have emerged as the most documented approach for limestone petrography (Table 2.3), offering the ability to count and classify individual grains. However, these methods face significant scalability challenges when applied to complete thin section images. The increasing resolution and size of modern microscopy images have created computational bottlenecks that are less problematic with cropped images on zoomed regions. Current object detection frameworks typically require image decomposition, where large images must be divided into smaller tiles for training and inference. This approach introduces complex technical challenges including the management of overlapping bounding boxes across tile boundaries, the merging of detection results from adjacent tiles, and the accurate reprojection of findings across the full image area. These computational complexities highlight the interconnected evolution of deep-learning-based technologies and data capabilities, where advances in imaging hardware have simultaneously created new opportunities and new technical obstacles for automated petrographic analysis.

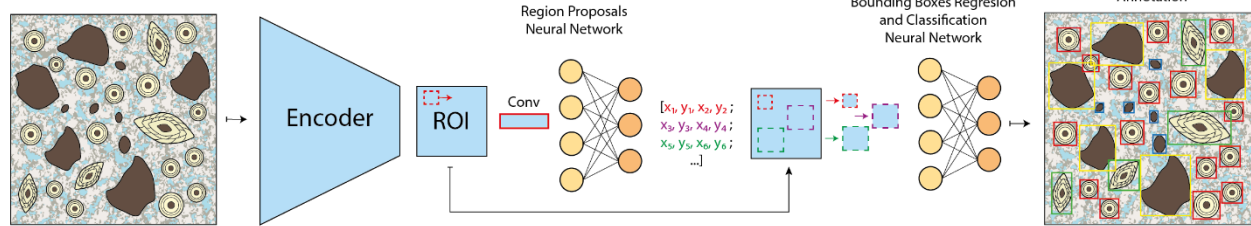
Segmentation methods based on pixel-level classification rules present a promising alternative that remains underexplored in current research. These approaches offer several distinct advantages over object detection algorithms. First, they provide a richer scene representation by identifying not only "what" is present but also "where exactly" it is located, enabling more sophisticated spatial analysis of object relationships. Additionally, grains observed in thin sections are typically in close contact or touching, and while detection methods may struggle to separate adjacent objects, segmentation can precisely distinguish pixels belonging to each individual instance. Furthermore, the extraction of morphological parameters from detected objects, which typically align with ideal petrographic descriptions, was not accessible in the object detection model implementations used in previous studies.

Principal Image-based Methodologies for Carbonate Petrography

Image Classification



Object Detection (Two-Stages)



Semantic Segmentation

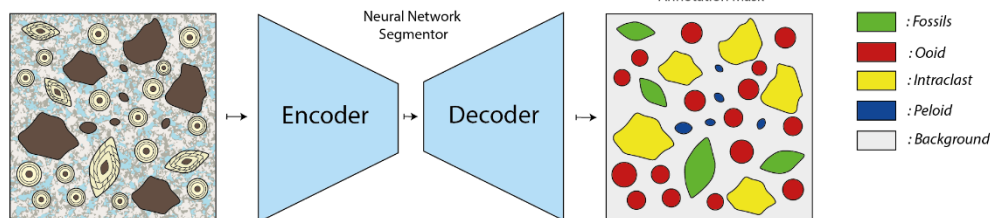


Figure 2.9 Principal paradigms and neural network designs investigated in the state-of-the-art numerical carbonate petrography applied to thin sections. Image classification architecture output is a discrete value corresponding to most probable class. Object detection model provides a list of boxes coordinates and most probable class per object. The image segmentation framework produces a mask where the most probable class is affected per pixel.

The transition from traditional feature extraction to deep learning methods represents a significant advancement in petrographic analysis (Figure 2.9). However, addressing dataset bias, expanding classification systems beyond traditional frameworks, and implementing robust segmentation approaches remain crucial for advancing the field's analytical capabilities with the most valuable outputs.

2.2 Grain Size Analysis

Grain size analysis and classification systems emerged as the earliest standardized methods for characterizing sedimentary particles, establishing consistent terminology among geologists and enabling quantitative assessment of depositional processes. Initially developed as simple threshold-based descriptive approaches (Folk, 1974; Wentworth, 1922), these frameworks subsequently evolved into more sophisticated methodologies incorporating concrete diagram structures or statistical parameters from particle size distributions (Folk, 1954, 1966, 1974; Krumbein, 1934; Shepard, 1954) to support genetic interpretations of sedimentary environments. Modern sedimentary analysis remains anchored in these foundational principles, with numerical tools assisting these tedious tasks and ongoing modifications addressing specialized research applications.

2.2.1 Grain Size Classification and Statistical Analysis

2.2.1.1 Wentworth's Framework

C. Wentworth introduced a systematic grain size classification for siliciclastic rocks, initially based on a geometric progression of mean particle diameters (Figure 2.10) (Wentworth, 1922). This grade scale uses 1 mm as its reference point, where each size class is bounded by minimum and maximum values following a 2^n step expressed in mm with n an integer, typically representing successive doubling or halving of average particle diameters.

Building upon earlier work (Udden, 1914), the system was specifically designed to reflect natural breaks in sediment size distributions observed in field and laboratory studies, making it a first practical approach for understanding sediment transport mechanisms and depositional energy conditions.

Original Wentworth's Grain Size Classification

Krumbein's Phi Scale and Folk's Extension to Carbonate Rocks

(Wentworth, 1922; Krumbein, 1934; Folk 1974)

Millimeters (mm)	Micrometers (μm)	Phi (Φ) (Krumbein)	Siliciclastic (Wentworth)		Carbonates Transported (Folk)	Carbonates Authigenic (Folk)
4096.00		-12.0	Boulder	GRAVEL		Extremely Coarsely Crystalline
256.00		-8.0	Cobble		Very Coarse Calcirudite	
64.00		-6.0	Very Coarse Pebble		Coarse Calcirudite	
32.00		-5.0	Coarse Pebble		Medium Calcirudite	
16.00		-4.0	Medium Pebble			
8.00		-3.0	Fine Pebble			
4.00		-2.0	Granule		Fine Calcirudite	Very Coarsely Crystalline
2.00		-1.0	Very Coarse Sand	SAND	Coarse Calcarenite	Coarsely Crystalline
1.00		0.0	Coarse Sand		Medium Calcarenite	Medium Crystalline
1/2	0.50	1.0	Medium Sand		Fine Calcarenite	
1/4	0.25	2.0	Fine Sand		Very Fine Calcarenite	
1/8	0.125	3.0	Very Fine Sand			
1/16	0.0625	4.0	Coarse Silt	SILT	Coarse Calcilutite	Finely Crystalline
1/32	0.031	5.0	Medium Silt		Medium Calcilutite	
1/64	0.0156	6.0	Fine Silt		Fine Calcilutite	Very Finely Crystalline
1/128	0.0078	7.0	Very Fine Silt			
1/256	0.0039	8.0		CLAY		Aphanocrystalline
1/1000	0.001	10.0	Clay			Cryptocrystalline

Figure 2.10 Wentworth's grain size classification systems, complemented with Krumbein's contributions with phi scale development and Folk carbonates constituent integration.

Framework Extensions:

Recognizing the need for simplification, finer subdivisions or extended applications, several updates have been proposed. First, the initial grade scale was transformed into a phi (ϕ) scale (Figure 2.10), where $\phi = -\log_2(d)$ with d the average grain size diameter in mm (Krumbein, 1934). This mathematical simplification principally allowed to convert the geometric progression into an arithmetic scale.

Detailed terminology aligning with established boundaries was introduced for carbonates rocks, enabling characterization of both transported and authigenic constituents (Folk, 1974) (Figure 2.10).

Additionally, the classification scheme was extended to accommodate particles exceeding 4096 mm, facilitating identification of transport mechanisms specific to these materials (Blair & McPherson, 1999).

2.2.1.2 Shepard and Folk's Framework

F. Shepard and R. Folk developed classification systems specifically tailored for mixed terrigenous sediments with variable compositional proportions. Since natural sediments typically consist of multiple size fractions, these frameworks were designed to convey complete textural characteristics rather than merely average particle diameter. Such approaches notably participated in improving sedimentological description of heterogeneous deposits.

Both systems introduced ternary diagram concepts initially based on relative proportions of three end-member components: sand, silt, and clay (Figure 2.11 left), utilizing particle size boundaries from the C. Wentworth's system. Despite not being explicitly proposed, the transposition to carbonates can be applied with the same logic between calcilutite, calcarenite and calcirudite fractions.

These original frameworks underwent temporal refinements and stabilized with the incorporation of intermediate classes (e.g. mud, a mix of silt and clay in relative equal proportions) or additional diagrams for gravels characterization (Folk, 1974; Schlee, 1973) (Figure 2.11 right).

Original Ternary Diagrams for Grain Size Classification
(Shepard and Folk, 1954; Schlee, 1973; Folk, 1974)

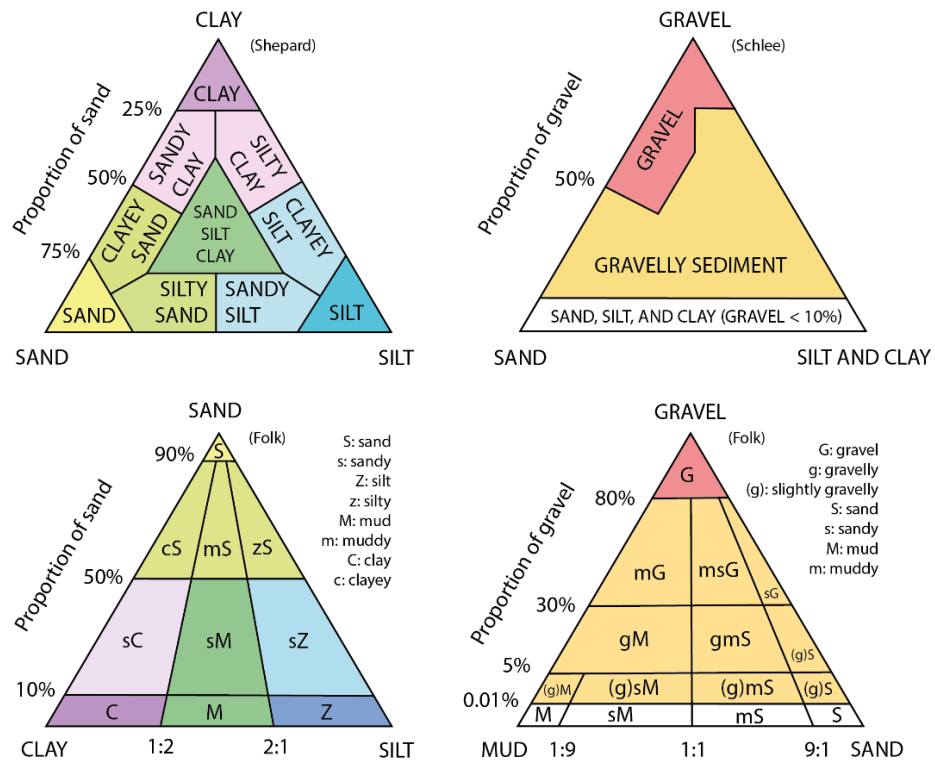


Figure 2.11 Shepard's ternary diagram system for mixed grain size characterization complemented by Schlee for gravel size fractions (upper) and Folk's nomenclature using similar structure with varying classes and boundaries (lower).

Framework Extensions:

R. Folk later enhanced grain size classifications by integrating statistical parameters to provide quantitative parameters extracted from particle size distributions (Equation 2.1 to 2.4), building upon W. Krumbein's earlier method. This approach employs cumulative frequency curves of grain diameters converted to the phi (ϕ) scale, enabling systematic comparison between sample populations based on their derived statistical properties characterizing the geometric signature of curve distributions (Figure 2.12).

- Graphic Mean (M_z) indicates average grain size:

$$M_z = \frac{\varphi_{16} + \varphi_{50} + \varphi_{84}}{3} \quad (2.1)$$

- Inclusive Graphic Standard Deviation (σ_I) reflects the grain sorting:

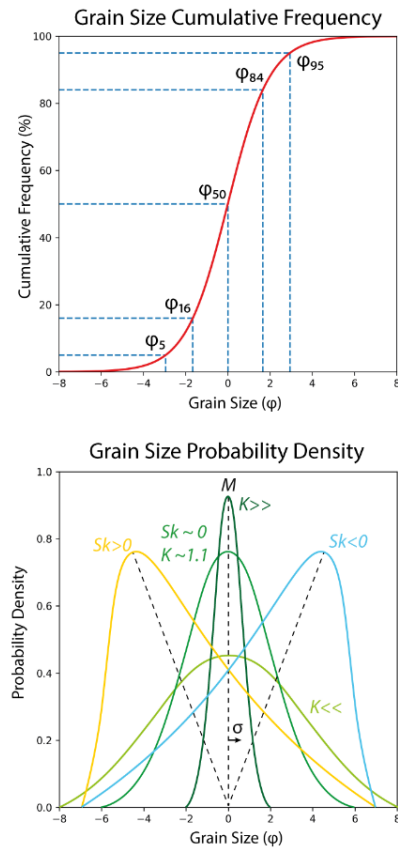
$$\sigma_I = \frac{(\varphi_{84} - \varphi_{16})}{4} + \frac{(\varphi_{95} - \varphi_5)}{6.6} \quad (2.2)$$

- Inclusive Graphic Skewness (Sk_I) denotes the symmetry of the distribution:

$$Sk_I = \frac{\varphi_{16} + \varphi_{84} - 2\varphi_{50}}{2(\varphi_{84} - \varphi_{16})} + \frac{\varphi_5 + \varphi_{95} - 2\varphi_{50}}{2(\varphi_{95} - \varphi_5)} \quad (2.3)$$

- Graphic Kurtosis (K_G) specifies the distribution peakedness:

$$K_G = \frac{\varphi_{95} - \varphi_5}{2.44(\varphi_{75} - \varphi_{25})} \quad (2.4)$$



Statistical Parameter	Category	Range Values
Graphic Mean (M_z)	Use Wentworth classification boundaries (Indicates average grain size)	
Inclusive Graphic Standard Deviation (σ_I)	Very well sorted	$\sigma_I < 0.35$
	Well sorted	$0.35 \leq \sigma_I < 0.50$
	Moderately well sorted	$0.50 \leq \sigma_I < 0.70$
	Moderately sorted	$0.70 \leq \sigma_I < 1.00$
	Poorly sorted	$1.00 \leq \sigma_I < 2.00$
	Very poorly sorted	$2.00 \leq \sigma_I < 4.00$
Inclusive Graphic Skewness (Sk_I)	Very coarse-skewed	$-1.0 < Sk_I < -0.3$
	Coarse-skewed	$-0.3 \leq Sk_I < -0.1$
	Symmetrical	$-0.1 \leq Sk_I < +0.1$
	Fine-skewed	$+0.1 \leq Sk_I < +0.3$
	Very fine-skewed	$+0.3 \leq Sk_I < +1.0$
Graphic Kurtosis (K_G)	Very platykurtic	$K_G < 0.67$
	Platykurtic	$0.67 \leq K_G < 0.90$
	Mesokurtic	$0.90 \leq K_G < 1.11$
	Leptokurtic	$1.11 \leq K_G < 1.50$
	Very leptokurtic	$1.50 \leq K_G < 3.00$
	Extremely leptokurtic	$K_G \geq 3.00$

Figure 2.12 Folk's grain size statistical classification using Graphic Mean, Inclusive Graphic Standard Deviation, Inclusive Graphic Skewness and Graphic Kurtosis derived from cumulative frequency descriptors (Equation 1 to 4 using quantiles 5, 16, 50, 84 and 95). Visual impact of the different statistical parameters is highlighted through the probability density function.

Moreover, one frequently questioned aspect of this method concerns the quantification of particle size based on the number of instances, or whether weighting by mass or occupied surface should be applied during two-dimensional observation. This consideration is important given that within a defined volume or surface area; there cannot necessarily be more large particles than small ones. Therefore, it is common practice to apply this weighting to achieve better dispersion of values within the calculated distribution. The statistical parameters are then decomposed using specific thresholds for practical applications (Figure 2.12).

2.2.2 Methodologies Discussion

Rather than competing approaches, the different classification systems served complementary roles, each addressing specific observational needs while collectively providing comprehensive tools for sedimentological analysis.

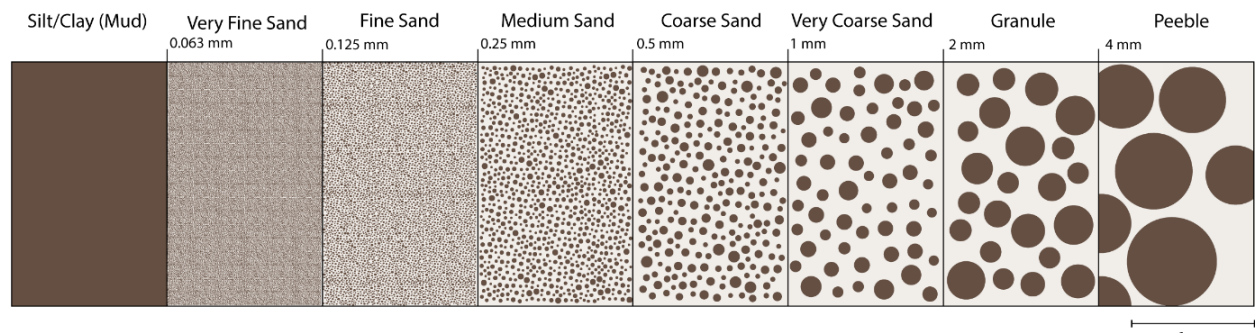
The C. Wentworth classification's enduring significance derives from its establishment of size-based boundaries that correlate with depositional energy and transport distance (Chambers & Upchurch, 1979; Folk, 1974). The fundamental principle that grain size decreases with increased transport distance is widely accepted for siliciclastic systems where terrigenous sediments result primarily from mechanical erosion processes. However, this relationship remains debated for carbonate sediments (de Kruijf et al., 2021), where constituents form through complex mechanisms that depend on the type of carbonate factory (Schlager, 2000). Analyzing carbonate particle sizes introduces additional complexity since grain growth may reflect various environmental conditions or dynamic evolutionary processes (Trower et al., 2015; Védérine et al., 2007). The mathematical refinements introduced through W. Krumbein's phi scale, along with specialized extensions for coarse-grained materials and integrated carbonate terminology, demonstrate the system's adaptability while preserving its fundamental simplicity.

The Shepard classification shifted focus from average grain sizes to textural relationships within mixed sediments. The ternary approach provides a practical framework that simplifies description through visual assessment of material proportions across specified grain-size fractions.

R. Folk's system combined the precision of C. Wentworth boundaries with F. Shepard system logic, enhanced through quantitative statistical parameters. This integration significantly improved analytical rigor and provided quantitative indicators for sedimentary process analysis. However, the semantic complexity and numerous classes in these frameworks limit their practical application. Furthermore, across the different parameters most have not achieved equal popularity, with recent studies typically concentrating on average grain size and grain sorting (An et al., 2017; Javanbakht et al., 2022; J. Liu et al., 2021; Recking et al., 2022). The development of graphical charts for field sedimentological description has particularly encouraged the preferential use of these two key measurements (Figure 2.13).

Grain Size Analysis Charts

Mean Grain Size



Grain Sorting

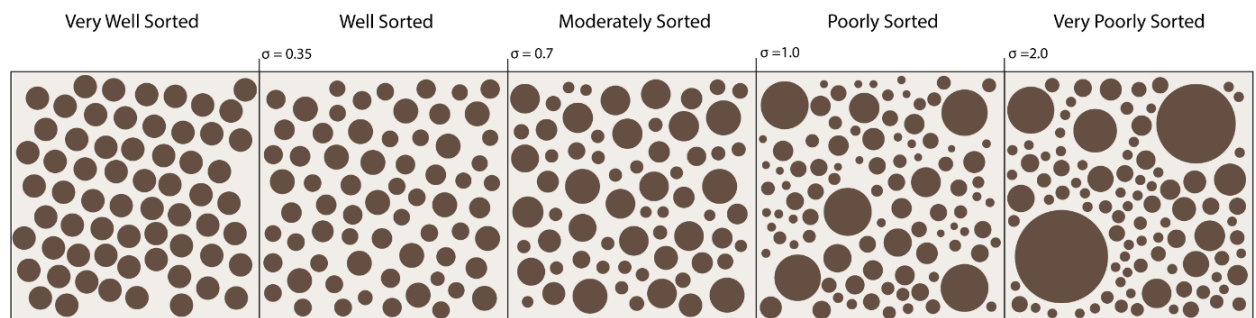


Figure 2.13 Typical grain size analysis charts used for mean grain size and grain sorting classes assignments.

Contemporary sedimentological practice increasingly relies on integrating multiple parameters to differentiate samples into families with shared characteristics (Benavente et al., 2021). Classification rules can be either concept-driven or established through arbitrary threshold definitions. In the case of grain size characterization and distribution analysis, this essentially represents a regression problem where the resolution matters more than the specific classes or boundaries themselves.

2.2.3 Numerical Methods Integration

Recent advances in machine learning and numerical methods have enhanced particle size analysis across different applications and scales, moving beyond traditional petrographic techniques. These developments align with broader goals in geosciences and porous media characterization, where understanding how particle arrangement affects pore network structure can be investigated to characterize key petrophysical properties and heterogeneities in porosity or permeability (Graczyk & Matyka, 2020).

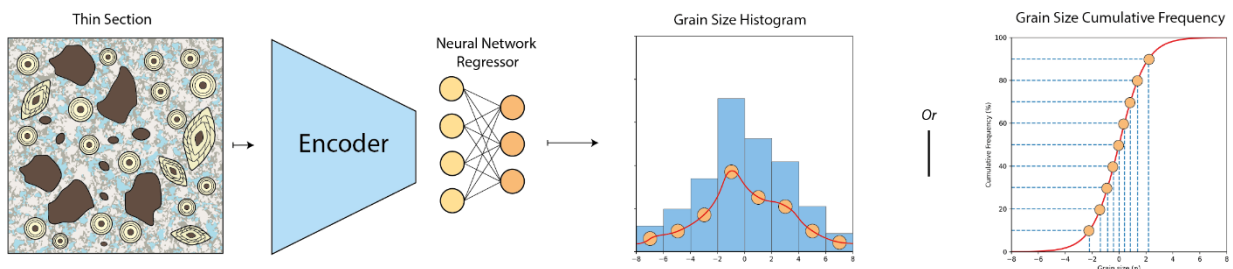
The studies are generally confronted initially with difficulties related to the acquisition of a large database of measurements. The most studied cases involve unconsolidated sediments where samples containing known size proportion of material are generated after sieving, synthetic image generation processes are also used, or even particle diameter labeling is performed on real images (Javad et al., 2023; Lang et al., 2021; Olivier et al., 2020). Another general observation is the inconsistent use of evaluation metrics, which often express errors that are difficult to interpret when assessing model performance. In contrast, a coefficient of determination (R^2) would provide more informative results (Table 2.4).

Author(s) and Year	Database	Images	Particles	Method	Architecture	Result
Olivier et al., 2020	Real Stitched Images with known PSD	781	–	Cumulative Frequency (Continuous)	VGG-16	75% (R^2)
Soloy et al., 2020	Real Labelled Images	46	3,598	Segmentation + Cumulative Frequency (Continuous)	Mask R-CNN	45–75% (R^2)
Lang et al., 2021	Real Labelled Images	1,491	> 180,000	Histogram (Discrete)	GRAINet (Residual)	0.42 (Wasserstein)
				Weighted Volume (Discrete)		0.65 (Wasserstein)
				Mean Diameter (Continuous)		18% (Normalized Error)
Javad et al., 2023	Synthetic Images	53,003	–	Cumulative Frequency (Continuous)	PSDNet	2.8% (RMSE)
Warren et al., 2023	Real Labelled and Synthetic Images	640	–	Segmentation + Mean Grain Size	U-Net	8.88% (MAE)
Mair et al., 2024	Real Labelled Images	56	< 1,000	Segmentation + Diameter Cumulative Frequency	Cellpose (U-Net)	57.5% (mAP)
Wei et al., 2024	Real SSE Images	500	–	Cumulative Frequency (Continuous)	ResNet18	10.91% (MAPE), 13.11% (NRMSE)

Table 2.4 State-of-the-art deep learning approaches for grain size analysis. (MAE: mean absolute error, mAP: mean average precision, MAPE: mean average percentage error, RMSE: root mean square error, R^2 : coefficient of determination, Wasserstein: distribution similarity)

Principal Image-based Methodologies for Particle Size Distribution Estimation

Direct Particle Size Distribution Prediction



Intermediate Segmentation then Particle Size Distribution from Measurements

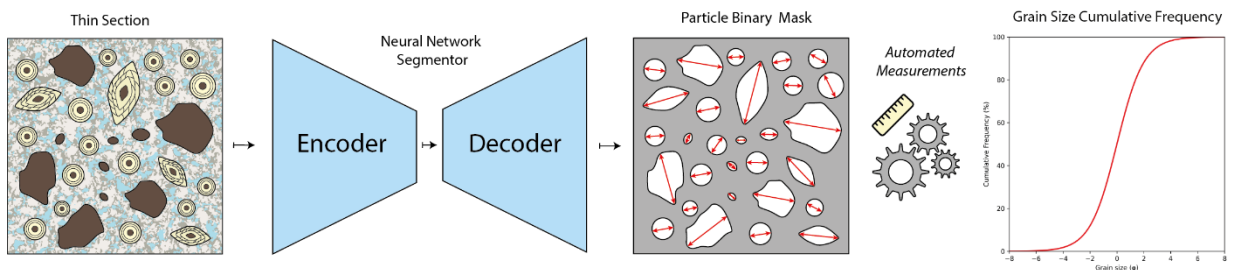


Figure 2.14 Principal paradigms and neural network designs investigated in the state-of-the-art numerical grain size analysis. Direct particle size distribution uses image encoding for pre-defined quantiles prediction. Intermediate segmentation replicates human processing tasks by predicting the binary particle mask and automating measurements.

The characterization of particle sizes distribution is then generally addressed through predicting the cumulative frequency curve of particle sizes (Javad et al., 2023; Olivier et al., 2020) or through discretization of size intervals and prediction of the histogram for each proportion according to pre-defined intervals directly from the image (Lang et al., 2021) (Figure 2.14, upper). Other methods refer to an approach divided into two processes: first, particles segmentation, and second, automated measurements to provide similar analytical curves (Mair et al., 2024; Soloy et al., 2020) (Figure 2.14, lower). It is also possible that certain specific parameters are directly predicted, such as the average grain size or other statistical descriptors proposed by R. Folk (Lang et al., 2021; P. Warren et al., 2024).

Each of these methods has its advantages and limitations. Direct approaches allow for easier use but lack explainability in case of model drift, and predictions will generally be limited to the interval of minimum and maximum particle sizes seen during training. Methods based on measurements from segmentation masks allow for better understanding of the process and errors potentially made with different grain size intervals, however the resulting measurements will be sensitive to the quality of this same segmentation.

2.3 Porous Network Analysis

Complementary to rock composition and grain size analysis, porosity is studied through thin sections to characterize the network that controls fluid flow and storage capacity in carbonate rocks. Originally developed as qualitative or semi-quantitative descriptive methods (Archie, 1952; Choquette & Pray, 1970; Lucia, 1995, 1999), petrographic pore analysis has evolved into quantitative frameworks incorporating digital image analysis, using descriptors such as morphological parameters, complemented with various sources of acquisition to support detailed interpretations of reservoir quality and diagenetic processes (Anselmetti & Eberli, 1993; White et al., 1998). Contemporary carbonate petrography integrates these established approaches with growing prominence referring to imaging techniques, enabling enhanced characterization of pore networks for both fundamental research and practical applications.

2.3.1 Pores Classification

2.3.1.1 Archie's Framework

One of the earliest classifications of porous space was proposed by G. Archie to investigate direct relationships with petrophysical parameters (Archie, 1952). This classification system is structured around three primary characteristics: texture of matrix, visible pore size and frequency of occurrence (Table 2.5).

Texture Type	Hand Sample Appearance	Microscopic Appearance (10× to 15× ⇔ few μm/pixel)
Type I: Compact Crystalline	Crystalline, hard, dense, with sharp edges and smooth faces on breaking. Resinous luster.	Matrix composed of tightly interlocking crystals with no visible pore space. Commonly produces “feather edge” on breaking due to fracturing at crystal edges in thin flakes.
Type II: Chalky	Dull, earthy or chalky appearance. Crystalline structure not apparent due to small, loosely interlocked crystals that reflect light in multiple directions, or composed of extremely fine granules. May be siliceous or argillaceous.	Crystals less effectively interlocked than Type I, joining at various angles. Extremely fine textures may still appear “chalky” under this magnification, while others begin to show crystalline structure.
Type III: Granular or Saccharoidal	Sandy or sugar-like appearance (sucrose texture). Crystal/granule sizes: Fine = 0.05 mm; Medium = 0.1 mm; Coarse = 0.4 mm.	Crystals interlocking at various angles, generally allowing considerable inter-crystal porosity. Includes oolitic and other granular textures.

Porosity Class	Description
Class A	No visible porosity under 10× magnification; pore diameter < 0.01 mm
Class B	Visible porosity with pore diameter between 0.01 mm and 0.1 mm
Class C	Visible porosity with pore diameter > 0.1 mm but smaller than cutting size
Class D	Visible porosity evidenced by secondary crystal growth on cutting faces or weathered surfaces showing fracturing/solution channels; pore size greater than cutting size

Porosity Quality	Surface Coverage (%)
Excellent	20
Good	15
Fair	10
Poor	5

Table 2.5 Archie's porosity classification framework. Pore types are defined according to rock compositional texture in which they are observed, void sizes and the surface they cover.

Each class is defined through a combination of the identified features (e.g. I-C-Good) admitting the inverse relationship between pore size and frequency, as pore dimensions increase across the classes, their occurrence frequency typically decreases.

2.3.1.2 Choquette and Pray's Framework

The P. Choquette and L. Pray's porosity classification represents the most frequently used framework in carbonate reservoir characterization (Choquette & Pray, 1970). This system constructs a descriptive and genetic structure around the relationship between porosity and the rock's original fabric elements, establishing three primary divisions based on this fabric dependency (Figure 2.15).

Original Choquette and Pray's Porosity Classification for Carbonate Rocks (1970)

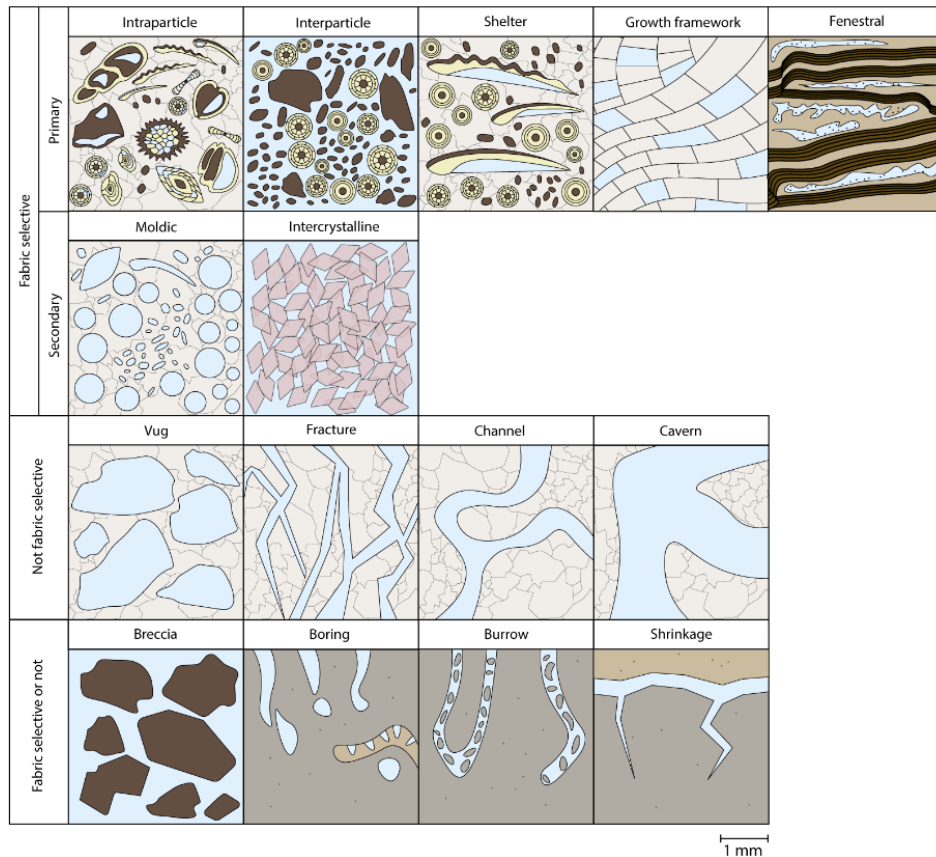


Figure 2.15 Choquette and Pray's porosity classification framework. Pores are categorized using intricate relationships with sedimentological interpretations and voids morphology for genetic significance.

The first major category, fabric-selective porosity is directly controlled by the original depositional and diagenetic rock fabric. The shape and location of the pores are determined by the rock's components, such as grains, fossils, or crystals. This class is subdivided into two subsections with primary and secondary porosity. Primary porosity forms during the original deposition of the rock with residual voids trapped in sediments while secondary porosity develops after deposition through dissolution processes resulting from physical-chemical condition modifications.

The second category, non-fabric-selective porosity, groups pore types that develop independently of the original rock fabric and cutting across the depositional and diagenetic components of the rock. These pores are typically formed by larger-scale processes that are not tied to the rock's specific texture.

A third category is defined when the dependency between porosity and original rock fabrics can be selective or not selective. This can be specifically related to mechanical fragmentation, bioturbation or desiccation events.

The methodology systematically identifies fifteen distinct porosity types based on the enumerated characteristics and geometric features across various depositional and diagenetic environments.

2.3.1.3 Lucia's Framework

F. Lucia's classification follows G. Archie's original approach by incorporating petrophysical principles to establish porosity–permeability laws (Lucia, 1995). This framework modifies the original P. Choquette and L. Pray's classification by restructuring terminology and integrating quantitative parameters (Figure 2.16).

Original Lucia's Porosity Classification in Carbonates (1995)

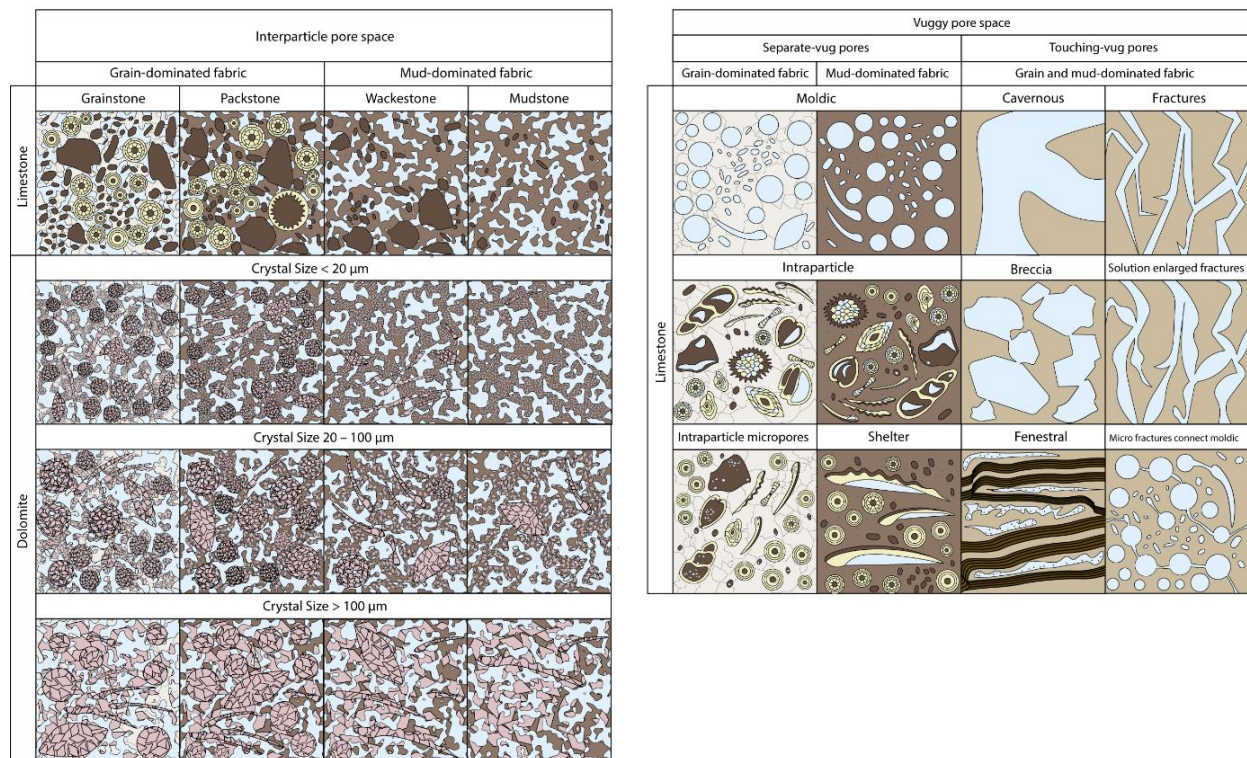


Figure 2.16 Lucia's porosity classification framework. Pore classes are separated between interparticle and vuggy porosity. Interparticle porosity is subdivided into Dunham texture integrating particle size and sorting in the classification. Vuggy porosity uses Choquette and Pray's nomenclature with the addition of connectivity pattern qualification.

The core of the system lies in its focus on two main pore types: interparticle porosity and vuggy porosity.

Interparticle porosity is split according to the original Dunham texture, average grain size and sorting of the rock's particles. Dolomitized rocks with intercrystalline voids follow the same principle replacing particles by crystal size features.

Vuggy porosity encompasses pore spaces grouped according to their dimensional relationship with constituent particles using Choquette and Pray's original nomenclature. This category includes both intraparticle voids (within individual grains or crystals) and significantly larger spaces that exceed particle dimensions, representing various dissolution and structural porosity types. The framework further subdivides vuggy porosity based on connectivity patterns that are supposed to directly influence permeability. Separate-vug pores generally rely exclusively on interparticle networks for fluid communication, while touching-vug pores form interconnected networks among themselves.

2.3.2 Methodologies Discussion

The classification systems that have been developed over time for carbonates have allowed enhanced understanding of porous networks, from their genesis and evolution to their influence on fluid circulation within these specific rocks. However, inherent limitations that affect practical application and analytical precision are exhibited by these methods. These constraints are primarily derived from subjective assessment requirements, scale-dependent observations, and the complex nature of carbonate pore systems, making consistent analytical protocols difficult to establish (Bust et al., 2011).

All the classification approaches rely heavily on visual interpretation of pore characteristics, introducing significant operator-dependent variability. G. Archie's system requires visual identification of pore presence and size estimation, while Choquette and Pray's framework demands recognition of fabric relationships that may be ambiguous in complex diagenetic environments where a single sample may span multiple pore categories. Similarly, Lucia's approach makes connectivity assessments between separate-vug and touching-vug categories dependent on interpretation of pore network geometry from two-dimensional observations. Substantial challenges are generated by these visual estimations, which are generally rendered imprecise because of the numerous parameters that are integrated into these nomenclatures. Thus, error propagation is inevitably introduced by the cumulative effect of these estimations, complicating replicable studies on large datasets.

Despite these identified limitations, certain promising characteristics should be retained for petrophysical characterization to be conducted on carbonate rocks. An essential parameter for relationships with permeability to be investigated is represented by the differentiation between intergranular and intragranular porosity in limestones, which is often characterized as intricately correlated in these rock types. Pore network geometry can be controlled by particle size distribution, which also appears to be a feature that can be related to fluid displacement capacity within the network (Figure 2.17) (Akbar et al., 1995; Lucia, 1995).

Porosity–Permeability Diagram related to Particle Size and Diagenesis in Carbonate (Lucia, 1995; Akbar, 1995)

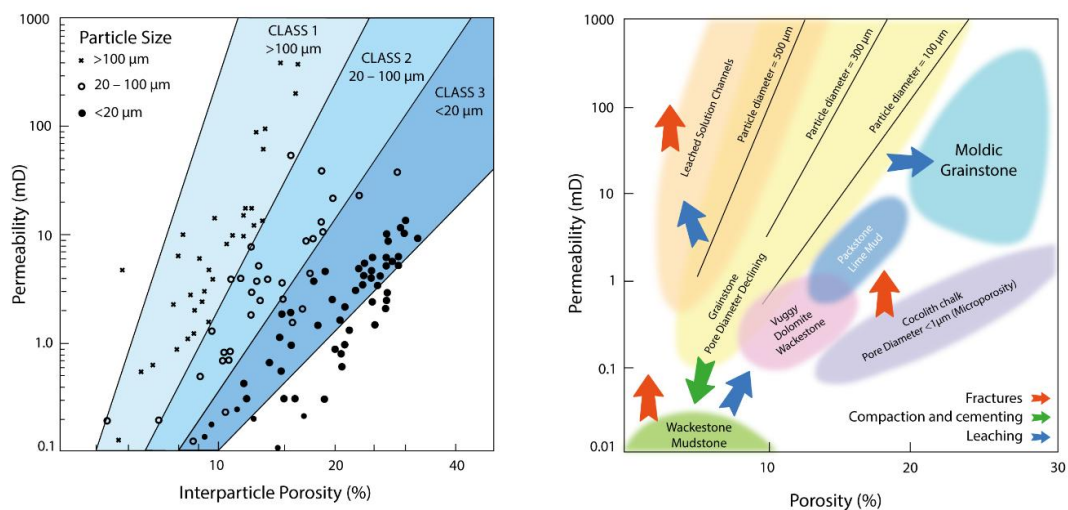


Figure 2.17 Insights in carbonate porosity-permeability relationships established in Lucia's and Akbar's studies where sedimentological and diagenetic features are supposed to be related to petrophysical signatures.

Furthermore, better understanding of the origin and functioning of dissolution processes is enabled by the interaction and favorable development of porosity in certain facies types, allochems, or orthochems, allowing improved integration into reservoir modeling workflows.

2.3.3 Numerical Methods Integration

The introduction of numerical methods and machine learning approaches into porosity classification represents an emerging frontier that addresses limitations inherent to traditional visual assessment techniques by providing quantitative analysis of observed elements. However, applying these methods to porosity characterization presents distinct challenges compared to their more established use in carbonate texture classification and grain size analysis.

Unlike carbonate classification systems that have benefited from several deep learning implementations, porosity classification methods remain largely dependent on manual interpretation or basic image processing techniques. Early digital approaches focused primarily on pore detection and measurements, including threshold-based segmentation and morphometrical analysis (area, circularity, aspect ratio, tortuosity, shape factor, etc.) (Figure 2.18) (Akbar et al., 1995; Karsanina, Gerke, & Skvortsova, 2015). These methods successfully automated pore space quantification but struggled to integrate the numerous extracted features with sedimentological interpretations to provide genetic relationships for better and explainable property propagation in reservoir modeling, which was the goal of classifications like those proposed by P. Choquette and L. Pray or F. Lucia.

Principal Image-based Methodologies for Porosity Analysis

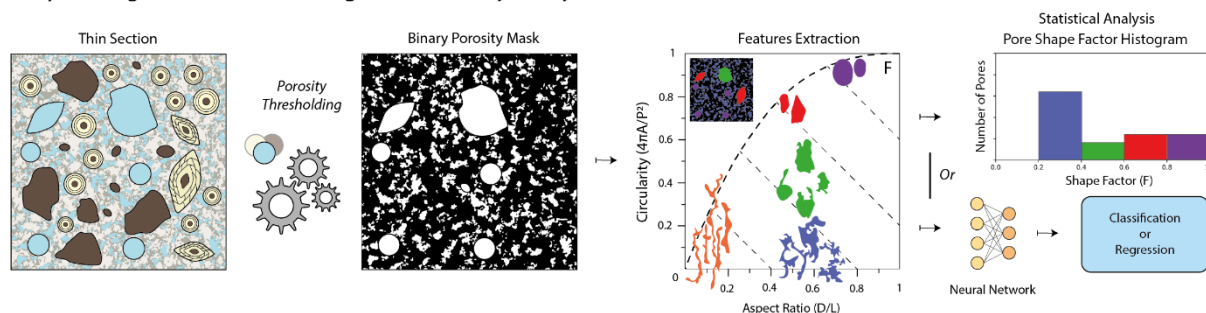


Figure 2.18 Principal paradigms and approaches in the state-of-the-art numerical porosity analysis of carbonate rocks. Morphometrical descriptors are extracted after porosity thresholding and are then statistically analyzed or re-used for a downstream classification or regression task performed by a machine learning model.

The transition toward machine learning applications in porosity analysis has explored predicting porosity classes from P. Choquette and L. Pray's framework using morphological attributes (Figure 2.18) (Abedini et al., 2018) or even estimating permeability using pore size distribution (S. Peng et al., 2016). However, the first application serves as a proxy for converting void shapes into interpreted geological information, which proves challenging when limited to simple image descriptors that lack integration with rock component information. Therefore, it would be first appropriate to process sedimentological descriptions and morphometric feature extraction of porosity separately, then combine them in a subsequent step to establish pore classes with geological significance derived from sedimentary or diagenetic processes. This approach would also avoid complex image annotation, allowing information integration to occur in a more straightforward manner.

2.4 Synthesis

Carbonate rock petrography has evolved from foundational visual classification systems into growing and integrated numerical approaches. The traditional frameworks of R. Folk, R. Dunham, and D. Sibley & J. Gregg for carbonate classification, alongside C. Wentworth's, F. Shepard's and R. Folk's grain size analysis, then P. Choquette & L. Pray's and F. Lucia's porosity classification, established fundamental descriptive standards that remain relevant today. However, these methods face significant limitations that modern digital approaches are beginning to address.

2.4.1 Limitations of Traditional Methods

The primary constraint across all traditional classification systems lies in their reliance on visual estimation and subjective interpretation. In carbonate texture analysis, operators must estimate allochemical and orthochemical proportions, leading to cumulative errors and inconsistencies across different expertise levels. R. Folk's complex framework proved impractical for widespread adoption, while R. Dunham's simplified approach, though more accessible, struggles with samples near class boundaries where distinguishing between categories becomes nearly impossible with visual methods alone.

Grain size analysis faces similar challenges when simple charts support only a weak representation of grain composition possibilities. The statistical parameters introduced by R. Folk for quantitative grain size description have seen limited adoption, with most studies focusing only on mean grain size and sorting due to analytical complexity.

Porosity classification encounters perhaps the greatest interpretive challenges, as systems like P. Choquette & L. Pray's framework require recognition of fabric relationships that may be ambiguous in complex diagenetic environments.

2.4.2 Numerical Development and Persistent Gaps

Recent advances in computer science have demonstrated significant potential for overcoming these limitations, with deep learning approaches showing improved performance compared to traditional feature extraction methods. However, current numerical implementations reveal several gaps that may limit their practical application.

Most studies concentrate exclusively on R. Dunham's limestone classification using supervised learning methods, with limited incorporation of other lithologies such as dolomite or anhydrite that are essential to describing carbonate sedimentary systems. Additionally, alternative training strategies have been rarely investigated beyond conventional supervised approaches. A concerning limitation is also the systematic absence of methods to mitigate training dataset bias, a critical oversight given that these datasets are constructed using the same manual interpretation methods that automated systems are designed to replace.

Object detection methods for granular analysis face substantial scalability challenges when applied to complete thin section images, primarily due to computational complexity in managing numerous overlapping bounding boxes. Segmentation approaches present promising alternatives by enabling pixel-level classification that produces detailed semantic outputs, yet this methodology remains significantly underexplored in petrographic applications because of labels acquisition costs.

In porosity analysis, the transition from traditional methods to digital imaging has been notably limited, with most applications focusing on basic morphometric measurements that fail to capture geologically meaningful distinctions. Current automated approaches overlook critical sedimentological classifications such as intraparticle versus interparticle porosity, or the identification of materials intersected by pore networks. These distinctions represent a rich source of information that aligns closely with established petrophysical relationships and correlations that remain largely unexploited in digital analysis workflows.

All these fundamental gaps between quantitative measurements and geological significance represent a key challenge in effectively bridging traditional petrographic expertise with emerging automated analytical capabilities.

2.5 Contributions

The evolution toward numerical methods in carbonate petrography represents both significant opportunity and methodological challenge, particularly when addressing the limitations of current approaches that have been critically examined across traditional thin section analysis. While deep learning approaches can potentially limit operator-dependent variability and provide consistent, reproducible results, they must address the fundamental issue of training data quality and comprehensiveness alongside the essential parameters useful for understanding sedimentary systems and reservoir levels.

The success of these approaches depends critically on developing robust and as unbiased as possible datasets that capture the full range of carbonate rock variability across different classification systems. This challenge is fundamentally addressed by reproducing naturalist approaches used in geosciences through a digital workflow, where each step progresses from high-level to low-level characterization, adding detail to descriptions and transitioning from qualitative approaches to quantitative methods.

The most promising direction involves integrated approaches that combine multiple analytical scales, from global image classification to detailed grain-level segmentation, enabling comprehensive characterization that mirrors the multi-parameter assessment traditionally conducted by expert petrographers. These proposals offer contributions to address the primary challenges encountered when implementing both traditional and digital characterization methods. However, realizing this potential requires systematic attention to dataset construction and careful validation against established petrographic expertise to ensure geological relevance of automated classifications, thereby bridging the gap between historical petrographic methods and modern numerical approaches in carbonate rock analysis.

2.5.1 Carbonate Rock Analysis

Paper 1 (chapter 3) presents a rotation-invariant representation learning approach using a variational autoencoder to perform clustering in the latent space of carbonate thin section images using traditional machine learning algorithms. The objective is to accelerate the construction of a supervised training dataset with limited examples, which subsequently serves to refine the previously trained model for classifying large numbers of images. This approach

demonstrates improved results compared to purely supervised training methods on the same architecture initialized with random weights on a small set of manually classified images.

This study defines an appropriate scale and patch size for images adapted to the studied dataset, enabling observation of local homogeneity in rock textures within a constrained observation field. This allows for subsequent visualization of global sample heterogeneity by making predictions using a sliding average over overlapping patch areas. The constructed dataset incorporates a broader range of classes than previously proposed in the literature, notably adding dolomitic rock classes (sucrosic and with preserved ghost-allochems) and anhydrite classes (homogeneous, breccias, and mixed with dolomites) that can be observed in specifically studied systems.

Paper 2 (chapter 4) develops a method for generating synthetic image training datasets used for classification and segmentation of granular limestone thin sections to address the typical lack of annotated data volume encountered in petrography. This approach combines procedural methods with texture generation via a generative model. The developed method reproduces geologically plausible images by integrating spatial distribution constraints and object classes, rapidly generating large numbers of examples with variable complexity.

Using these controlled data during model training enables the creation of balanced datasets in terms of class and content, something practically impossible in classical studies while respecting laws established by traditional classification nomenclatures such as R. Folk and R. Dunham, thus theoretically mitigating classification bias. The introduction of segmentation as a complementary descriptive approach in this study provides geologists with spatial image descriptions linked to R. Folk classification for more detailed characterization of sample granular content.

Paper 3 (chapter 5) introduces a method for adapting synthetic images created using the approach from paper 2 to make training data more realistic and achieve better segmentation results for improved exploitation of generated masks for quantifying detected elements.

2.5.2 Grain Size Analysis

Paper 3 introduces specific metrics for characterizing grain sizes observed in thin sections from produced segmentation masks. Since average grain size and sorting parameters have been identified as essential from both sedimentological perspectives (for understanding depositional conditions) and diagenetic perspectives (related to porosity classification and links to petrophysical properties), these two continuous values are derived from segmentation masks using a surface quantification method rather than instance-based measurement methods, which are shown to produce more errors in automated measurements.

2.5.3 Porous Network Analysis

Paper 1 presents how granular segmentation and porosity segmentation can be combined to discriminate and quantify intragranular and intergranular porosity. The synthesis section (chapter 6) then shows the integration of morphometric descriptors conventionally used for each identified pore type according to which component type it intersects, using semantic segmentation masks produced via methods from papers 1 and 2. This aims to enable fine data filtering for more detailed understanding of diagenetic processes.

REFERENCES

- Abedini, M., Ziaii, M., Negahdarzadeh, Y., & Ghiasi-Freez, J. (2018). Porosity classification from thin sections using image analysis and neural networks including shallow and deep learning in Jahrum formation. *Journal of Mining and Environment*, 9(2), 513–525. <https://doi.org/10.22044/jme.2018.6386.1458>
- Akbar, M., Petricola, M. J. C., Watfa, M. K., & Dhabi, A. (1995). Classic interpretation problems: Evaluating carbonates. *Oilfield Review*, 7(1), 38-47.
- An, C., Fu, X., Wang, G., & Parker, G. (2017). Effect of grain sorting on gravel bed river evolution subject to cycled hydrographs: Bed load sheets and breakdown of the hydrograph boundary layer. *Journal of Geophysical Research: Earth Surface*, 122(6), 1513–1533. <https://doi.org/10.1002/2016JF004146>
- Anselmetti, F. S., & Eberli, G. P. (1993). Controls on sonic velocity in carbonates. *Pure and Applied Geophysics*, 141(2), 287–323. <https://doi.org/10.1007/BF00998333>
- Archie, G. E. (1952). Classification of carbonate reservoir rocks and petrophysical considerations. *AAPG Bulletin*, 36(2), 278–298. <https://doi.org/10.1306/3D9343F7-16B1-11D7-8645000102C1865D>
- Benavente, C. A., Mancuso, A. C., & Bohacs, K. M. (2021). Reconstructing paleoenvironmental conditions through integration of paleogeography, stratigraphy, sedimentology, mineralogy and stable isotope data of lacustrine carbonates: An example from Early Middle Triassic strata of southwest Gondwana, Cuyana Rift, Argentina. In *Limnogeology: Progress, Challenges and Opportunities* (pp. 341-388). Springer. https://doi.org/10.1007/978-3-030-66576-0_12
- Bercean, B. A., Birhala, A., Ardelean, P. G., Barbulescu, I., Benta, M. M., Rasadean, C. D., Costachescu, D., Avramescu, C., Tenescu, A., Iarca, S., Buburuzan, A. S., Marcu, M., & Birsasteanu, F. (2023). Evidence of a cognitive bias in the quantification of COVID-19 with CT: An artificial intelligence randomised clinical trial. *Scientific Reports*, 13(1), 4887. <https://doi.org/10.1038/s41598-023-31910-3>
- Blair, T. C., & McPherson, J. G. (1999). Grain-size and textural classification of coarse sedimentary particles. *Journal of Sedimentary Research*, 69(1), 6–19. <https://doi.org/10.2110/jsr.69.6>
- Bust, V., Oletu, J., & Worthington, P. (2011). The challenges for carbonate petrophysics in petroleum resource estimation. *SPE Reservoir Evaluation & Engineering*, 14(1), 25–34. <https://doi.org/10.2118/142819-PA>
- Chambers, R. L., & Upchurch, S. B. (1979). Multivariate analysis of sedimentary environments using grain-size frequency distributions. *Mathematical Geology*, 11(1), 27–43. <https://doi.org/10.1007/BF01043244>
- Choquette, P. W., & Pray, L. C. (1970). Geologic nomenclature and classification of porosity in sedimentary carbonates. *AAPG Bulletin*, 54(2), 207–250. <https://doi.org/10.1306/5D25C98B-16C1-11D7-8645000102C1865D>
- Dunham, R. J. (1962). Classification of carbonate rocks according to depositional textures. In W. E. Ham (Ed.), *Classification of Carbonate Rocks: A Symposium* (pp. 108-121). *AAPG Memoir*, 1. <https://doi.org/10.1306/M1357>

- Embry, A. F., & Klovan, J. E. (1971). A late Devonian reef tract on northeastern Banks Island, N.W.T. *Bulletin of Canadian Petroleum Geology*, 19(4), 730–781. <https://doi.org/10.35767/gscpgbull.19.4.730>
- Flügel, E. (1982). Classifications of carbonate rocks. In *Microfacies Analysis of Limestones* (pp. 366–382). Springer Berlin Heidelberg. https://doi.org/10.1007/978-3-642-68423-4_6
- Flügel, E. A. (2004). *Microfacies of carbonate rocks: Analysis, interpretation and application*. Springer-Verlag, Berlin. <https://doi.org/10.1007/978-3-662-08726-8>
- Folk, R. L. (1954). The distinction between grain size and mineral composition in sedimentary-rock nomenclature. *The Journal of Geology*, 62(4), 344–359. <https://doi.org/10.1086/626171>
- Folk, R. L. (1959). Practical petrographic classification of limestones. *AAPG Bulletin*, 43(1), 1–38. <https://doi.org/10.1306/0BDA5C36-16BD-11D7-8645000102C1865D>
- Folk, R. L. (1962). Spectral subdivision of limestone types. In W. E. Ham (Ed.), *Classification of Carbonate Rocks: A Symposium* (pp. 62–84). *AAPG Memoir*, 1. <https://doi.org/10.1306/M1357>
- Folk, R. L. (1966). A review of grain-size parameters. *Sedimentology*, 6(2), 73–93. <https://doi.org/10.1111/j.1365-3091.1966.tb01572.x>
- Folk, R. L. (1974). *Petrology of sedimentary rocks*. Hemphill Publishing Company, Austin, Texas.
- Graczyk, K. M., & Matyka, M. (2020). Predicting porosity, permeability, and tortuosity of porous media from images by deep learning. *Scientific Reports*, 10(1), 21488. <https://doi.org/10.1038/s41598-020-78415-x>
- Javad, M., Pirnia, P., Manashty, A., Ujan, S., Toews, M., & Duhaime, F. (2023). PSDNet: Determination of particle size distributions using synthetic soil images and convolutional neural networks. *ArXiv preprint*, arXiv:2303.04269. <https://doi.org/10.48550/arXiv.2303.04269>
- Javanbakht, M., Beheshtipur, M. R., & Farimani, S. R. (2022). Factors affecting average grain size changes in rivers of a catchment area (Ardak catchment area, northeast Iran). *Arabian Journal of Geosciences*, 15(8), 736. <https://doi.org/10.1007/s12517-022-09994-0>
- Karsanina, M. V., Gerke, K. M., & Skvortsova, E. B. (2015). Universal spatial correlation functions for describing and reconstructing soil microstructure. *PLOS ONE*, 10(5), e0126515. <https://doi.org/10.1371/journal.pone.0126515>
- Krumbein, W. C. (1934). Size frequency distributions of sediments. *Journal of Sedimentary Research*, 4(2), 65–77. <https://doi.org/10.1306/D4268EB9-2B26-11D7-8648000102C1865D>
- Lang, N., Irniger, A., Rozniak, A., Hunziker, R., Wegner, J. D., & Schindler, K. (2021). GRAINet: Mapping grain size distributions in river beds from UAV images with convolutional neural networks. *Hydrology and Earth System Sciences*, 25(5), 2567–2597. <https://doi.org/10.5194/hess-25-2567-2021>
- Liu, J., Chen, X., Yin, P., Cao, K., Gao, F., Meng, Y., Qiu, J., & Li, M. (2021). Sediment characteristics, sources, and transport patterns in Kompong Som Bay, Cambodia: Indications from grain size and heavy minerals. *Journal of Ocean University of China*, 20(2), 329–339. <https://doi.org/10.1007/s11802-021-4532-5>
- Lokier, S. W., & Al Junaibi, M. (2016). The petrographic description of carbonate facies: Are we all speaking the same language? *Sedimentology*, 63(7), 1843–1885. <https://doi.org/10.1111/sed.12293>

- Lucia, F. J. (1995). Rock-fabric/petrophysical classification of carbonate pore space for reservoir characterization. *AAPG Bulletin*, 79(9), 1275–1300. <https://doi.org/10.1306/7834D4A4-1721-11D7-8645000102C1865D>
- Lucia, F. J. (1999). Characterization of petrophysical flow units in carbonate reservoirs: Discussion. *AAPG Bulletin*, 83(7), 1161–1163. <https://doi.org/10.1306/E4FD2E29-1732-11D7-8645000102C1865D>
- Mair, D., Witz, G., Do Prado, A. H., Garefalakis, P., & Schlunegger, F. (2024). Automated detecting, segmenting and measuring of grains in images of fluvial sediments: The potential for large and precise data from specialist deep learning models and transfer learning. *Earth Surface Processes and Landforms*, 49(3), 1099–1116. <https://doi.org/10.1002/esp.5755>
- Marmo, R., Amodio, S., Tagliaferri, R., Ferreri, V., & Longo, G. (2005). Textural identification of carbonate rocks by image processing and neural network: Methodology proposal and examples. *Computers & Geosciences*, 31(5), 649–659. <https://doi.org/10.1016/j.cageo.2004.11.016>
- Olivier, L. E., Maritz, M. G., & Craig, I. K. (2020). Estimating ore particle size distribution using a deep convolutional neural network. *IFAC-PapersOnLine*, 53(2), 12038–12043. <https://doi.org/10.1016/j.ifacol.2020.12.740>
- Peng, S., Hassan, A., & Loucks, R. G. (2016). Permeability estimation based on thin-section image analysis and 2D flow modeling in grain-dominated carbonates. *Marine and Petroleum Geology*, 77, 763–775. <https://doi.org/10.1016/j.marpetgeo.2016.07.024>
- Recking, A., Vazquez, D. O., & Piton, G. (2022). The contribution of grain sorting to the dynamics of the bedload active layer. *Earth Surface Processes and Landforms*, 47(4), 979–996. <https://doi.org/10.1002/esp.5304>
- Schlager, W. (2000). Sedimentation rates and growth potential of tropical, cool-water and mud-mound carbonate systems. *Geological Society, London, Special Publications*, 178(1), 217–227. <https://doi.org/10.1144/GSL.SP.2000.178.01.14>
- Schlee, J. S. (1973). *Atlantic continental shelf and slope of the United States; sediment texture of the northeastern part*. U.S. Geological Survey Professional Paper 529-L. <https://doi.org/10.3133/pp529L>
- Shepard, F. P. (1954). Nomenclature based on sand-silt-clay ratios. *Journal of Sedimentary Research*, 24(3), 151–158. <https://doi.org/10.1306/D4269774-2B26-11D7-8648000102C1865D>
- Sibley, D. F., & Gregg, J. M. (1987). Classification of dolomite rock textures. *Journal of Sedimentary Research*, 57(6), 967–975. <https://doi.org/10.1306/212F8CBA-2B24-11D7-8648000102C1865D>
- Soloy, A., Turki, I., Fournier, M., Costa, S., Peuziat, B., & Lecoq, N. (2020). A deep learning-based method for quantifying and mapping the grain size on pebble beaches. *Remote Sensing*, 12(21), 3659. <https://doi.org/10.3390/rs12213659>
- Trower, L., Lamb, M. P., & Fischer, W. W. (2015). What controls ooid grain size? *Geological Society of America Abstracts with Programs*, 47(7), 855.
- Tucker, M. (2009). *Carbonate sedimentology*. John Wiley & Sons. <https://doi.org/10.1002/9781444314175.ch1>

Udden, J. A. (1914). Mechanical composition of clastic sediments. *Geological Society of America Bulletin*, 25(1), 655–744. <https://doi.org/10.1130/GSAB-25-655>

Védrine, S., Strasser, A., & Hug, W. (2007). Oncoid growth and distribution controlled by sea-level fluctuations and climate (Late Oxfordian, Swiss Jura Mountains). *Facies*, 53(4), 535–552. <https://doi.org/10.1007/s10347-007-0114-4>

Warren, P., Raju, N., Prasad, A., Hossain, M. S., Subramanian, R., Kapat, J., Manjooran, N., & Ghosh, R. (2024). Grain and grain boundary segmentation using machine learning with real and generated datasets. *Computational Materials Science*, 233, 112739. <https://doi.org/10.1016/j.commatsci.2023.112739>

Wentworth, C. K. (1922). A scale of grade and class terms for clastic sediments. *The Journal of Geology*, 30(5), 377–392. <https://doi.org/10.1086/623003>

White, J. V., Kirkland, B. L., & Gournay, J. P. (1998). Quantitative porosity determination of thin sections using digitized images. *Journal of Sedimentary Research*, 68(2), 220–222. <https://doi.org/10.1306/D4268724-2B26-11D7-8648000102C1865D>

CHAPTER 3: FROM UNSUPERVISED TO SEMI-SUPERVISED CLASSIFICATION IN CARBONATES PETROGRAPHY USING A ROTATION INVARIANT VAE

Paper 1: In Preparation

Abstract

This study introduces a novel method for analyzing carbonate rock images using a rotation-invariant variational autoencoder combined with unsupervised clustering and supervised classification techniques. The approach tackles a key challenge in petrography by analyzing large collections of unlabeled thin section images to accelerate the creation of supervised datasets for further classifier training. Working with a dataset of 5,110 high-resolution carbonate and evaporitic (anhydrite) thin section scans that yielded 169,166 image tiles, the study demonstrates both the capabilities and limitations of this method. The model successfully captured meaningful geological features within its latent space, validated through t-SNE visualization and Gaussian Mixture Model clustering of 45,699 manually labeled samples across 10 litho-textural classes. The results reveal distinct clustering patterns, with strong performance for anhydrite samples and reasonable separation between dolomites, and limestone varieties. The encoder component of the model, initially trained on the reconstruction pre-text task and subsequently fine-tuned on a small set of labeled images, demonstrates superior performance compared to traditional supervised approaches, achieving a +3.5% improvement in mean average precision. Moreover, when applied to complete thin section characterization, the method effectively identified major lithological and textural regions and their boundaries. This approach represents a step forward in automated petrographic analysis, providing a scalable solution for processing large datasets of thin sections where fully manual classification would be impractical while preserving geological accuracy and relevance.

Keywords: Carbonates, Petrography, Deep Learning, Classification, Supervised, Unsupervised, VAE, Orientation Invariance.

3.1 Introduction

Petrography is a fundamental technique used in sedimentary studies, essential for understanding natural rock formation processes and the complex arrangement of microstructures (Boggs, 2009; Flugel, 2004). This method presents significant challenges due to intricate overlapping features that require considerable expertise to extract information useful for developing conceptual geological models. Petrographic analysis typically involves examining high-resolution scanned thin sections at various magnification levels, with scale selection driven by the size of the features being studied. When describing a thin section, it becomes difficult to comprehensively capture all available information while maintaining consistency and minimizing interpretational bias (Lokier & Al Junaibi, 2016). Although initial broad screening of data is typically performed to understand the main dataset heterogeneities and facilitate preliminary data sorting, this process can be time-consuming when dealing with large sample sets (>100 – 1,000). Moreover, as more samples are analyzed, previously identified features may be overlooked, or new characteristics might go unnoticed. Following initial data screening, thin section characterization typically involves selecting multiple regions of interest (ROIs) that exhibit local homogeneity to simplify their description. However, this selective approach may lead to sampling bias and the potential loss of important information in data description.

In recent years, Deep Learning (DL) has experienced unprecedented growth and development. These advances have proven effective across various fields, including computer vision applications within both general geoscience research (Bergen et al., 2019; Reichstein et al., 2019) and specific petrographic challenges, primarily using Deep Convolutional Neural Networks (DCNNs) (Koeshidayatullah et al., 2020; Liu et al., 2023; Pires de Lima et al., 2020). While this technology has shown promise in simplifying image description through supervised classification, a major limitation of these approaches remains the scarcity of large, labeled datasets. DCNNs are data-driven models that typically require extensive training examples to optimize their parameters for effective generalization. Furthermore, creating comprehensive datasets is time-consuming and can present significant challenges due to the large variety and heterogeneity of sedimentary rocks. Their analysis involves varying compositional characteristics that sometimes require different descriptive criteria that are not easily transferable between lithologies and textures (Dott, 1964; Dunham, 1962; Folk, 1959, 1974; Sibley & Gregg, 1987; Warren, 2010).

More recently, self-supervised learning has emerged as a powerful paradigm in computer vision (Gui et al., 2023), allowing models to learn representations from unlabeled images. This approach leverages the inherent structure and relationships within visual data by using auxiliary tasks to generate latent features that can later be used in clustering frameworks or to train projection head on downstream tasks (e.g., classification, regression, etc.) with minimal labeled data. While pre-text tasks, such as image inpainting, puzzle solving, and learning transformation-invariant features, have proven effective in various domains (Burgess et al., 2024; He et al., 2021; Noroozi & Favaro, 2016), their application to petrographic analysis remains rarely explored (Mamode et al., 2025) and represents a promising approach to sort data, accelerate the construction of supervised datasets, and train subsequent classifiers.

Among these approaches, methods such as contrastive learning (CL) have shown remarkable success, particularly SimCLR (Chen et al., 2020), which introduced the triplet learning framework by minimizing the distance between an image and its augmented version (e.g., flips, rotations, brightness and contrast shifts, etc.) while simultaneously maximizing the distance to randomly sampled images from the dataset. Although this approach has proven effective, it may suffer from the absence of robust priors for identifying appropriate contrastive pairs. This limitation can result in the inadvertent selection of similar images as negative pairs, forcing the model to separate examples from the same semantic class, potentially compromising the learning process.

Image reconstruction using Autoencoders (AE) has also proven promising for learning latent representations (Hinton & Salakhutdinov, 2006). However, traditional AEs suffer from non-continuous mapping of latent space, leading to potential holes in the learned representations, which can further compromise their efficiency. To address this issue, Variational Autoencoders (VAEs) were introduced. Unlike traditional autoencoders that encode inputs to fixed points, VAEs learn to map inputs to probability distributions in the latent space (Kingma & Welling, 2013). This approach gives VAEs several advantages over deterministic autoencoders, including better generalization and continuous latent representations that enable the generation of new samples. Combining VAEs with discriminators (Munjal et al., 2019; Rivera, 2023), inspired by Generative Adversarial Networks (GANs) (Goodfellow et al., 2014), also addresses part of the blurry reconstruction gap observed with standard VAEs, where textural details in thin section images might be important for discriminating between specific rock classes. Thus, this approach was identified as relevant for obtaining meaningful compressed information vectors for thin section ROIs.

Training a VAE to extract features from thin section thumbnail images can be somewhat restrictive, which has motivated the investigation of invariant characteristics relevant to the petrographic domain. During traditional thin section description processes, image orientation is often considered as secondary information or may be entirely omitted. Moreover, lithology characterization, depositional texture analysis, or rock facies identification are typically not impacted by this missing information. Therefore, a method that enforces latent space invariance to arbitrary rotation was investigated and was found to be relatively efficient for subsequent clustering using traditional machine learning algorithms. This approach allows batches containing numerous images of interest within clusters to be selected for further development of labeled datasets. These newly created datasets can then be used to train a classification head on top of the trained encoder, with superior performance being demonstrated compared to traditional supervised training. The classification method can subsequently be extended to complete thin section samples, and compositional variability across spatial dimensions can be highlighted.

3.2 Data

3.2.1 Raw Data

In this study, a collection of 5,110 high-resolution images of carbonate (limestone and dolomite) and evaporitic thin sections from various private sources were assembled. These images were captured under plain polarized light at a resolution of $8.8 \times 10^{-1} \mu\text{m}$ per pixel, resulting in image dimensions of approximately $50,000 \times 25,000$ pixels.

3.2.2 Working Data

3.2.2.1 Data Preprocessing

The raw data underwent preprocessing to convert images into a format suitable for both human operators and computational analysis. Initially, images were downsampled to a working scale of $28 \mu\text{m}$ per pixel, a resolution determined by human interpreters as necessary for identifying both lithology and rock texture in limestones. To facilitate annotation work and closely replicate the visual experience of describing a region of interest (ROI), images were divided into tiles of 128×128 pixels. This tile size ensures that observable features remain adequate for litho-textural identification while making the images computationally manageable for algorithms such as DCNNs. This resulted in the generation of 169,166 tiles for the 5,110 thin sections.

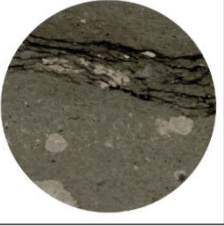
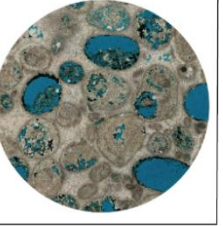
3.2.2.2 Data Sub-Sampling

A batch of 46,000 tiles was initially selected randomly from the complete set of generated tiled images. This sample size corresponds to the number of tiles that could feasibly be classified during the study, a limitation primarily imposed by human capacity constraints. This dataset was first used as a validation set to analyze class distribution within the latent space of the trained model. A subset of this data was subsequently used to train a supervised classifier.

3.2.2.3 Data Labeling

Out of the 46,000 tiles, 45,699 were manually categorized into 10 classes based on lithology, depositional textures for limestone samples (Dunham, 1962), and specific features regrouped under the term *disturbed* as presented in Figure 3.1. This classification resulted in the following data distribution: 4,145 mudstones, 3,381 disturbed mudstones, 4,091 wackestones, 4,500 packstones, 5,174 grainstones, 3,362 dolo-grainstones, 4,509 dolomites (sucrosic), 6,600 anhydrites, 5,687 disturbed anhydrites, and 4,250 anhydrites with ghost-preserved dolomitized grains (Figure 3.2). The remaining 301 tiles were excluded for two reasons: first, some of them belonged to limestone textures: floatstones, rudstones, or boundstones (Embry & Klovan, 1971) ; for which insufficient examples were available in the dataset to enable reliable class discrimination; second, others exhibited excessive complexity with multiple litho-textures or heavy diagenetic alteration in the same field of view, which were deemed too ambiguous for classification.

Litho-Texture Classification

LIMESTONE				
Mudstone	Mudstone Disturbed	Wackestone	Packstone	Grainstone
- Micritic matrix - Mud-supported < 10% grains - No to low porosity	- Micritic matrix - Mud-supported < 10% grains - Stylolites - Convolute laminations - Anhydrite nodules - No to low porosity	- Micritic matrix - Mud-supported > 10% grains - No to low porosity	- Micritic matrix - Grain-supported - No to low porosity	- Sparry cement - Grain-supported - Medium to high porosity
				
1 mm				

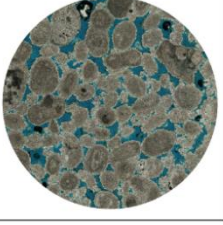
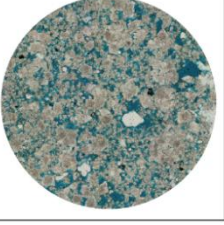
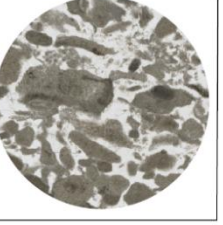
DOLOMITE		ANHYDRITE		
Dolo-grainstone	Dolomite	Anhydrite	Anhydrite Disturbed	Anhydrite Grainy
- Preserved or dissolved matrix - Preserved allochems - Medium to fine crystals - Texture alteration - Low to high porosity - Local anhydrite patches	- Sucrosic crystals - Medium to large crystals - High porosity	- Homogeneous to fibrous texture - No porosity	- Breccia - Mixed with residual dolomite - No to low porosity	- Anhydrite matrix - Preserved allochems - No to low porosity
				
1 mm				

Figure 3.1 Petrographic classification and comparative characteristics of carbonate (limestone and dolomite) and evaporitic rocks (anhydrite) leading to compositional and textures variability separated in different classes. Corresponding representative thin section images illustrate each category.

Labeled Data Distribution

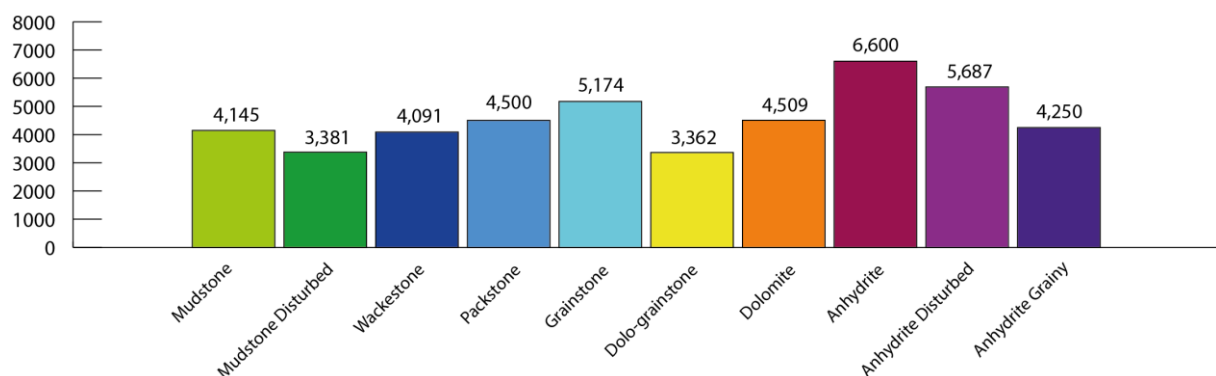


Figure 3.2 Labeled data distribution for the manually constructed thin section image dataset.

3.3 Methods

This study utilizes a large collection of unlabeled thin section tiles (169,166) combined with a smaller subset of labeled examples (45,699) to develop a VAE framework that learns orientation-invariant representations. The learned embeddings are analyzed to examine class distributions within the latent space and evaluate whether unsupervised clustering can effectively isolate distinct geological groups. Subsequently, the model encoder is fine-tuned on a subset of classified images to adapt it for downstream classification tasks that align with established geological classes. The approach is further validated by applying the classification model to complete thin section samples to assess its spatial characterization capabilities.

3.3.1 Variational Autoencoder

VAEs are a type of neural network that aims to learn a compressed, probabilistic representation of data in a continuous latent space. In this specific case it consists of three main components: an encoder network (Enc), parameterized by ϕ , that maps input data x into the latent space z by approximating the posterior distribution $q_\phi(z|x)$; a decoder network (Dec), parameterized by θ , that reconstructs the input x by modeling the likelihood $p_\theta(x|z)$; and a discriminator network, parameterized by ψ , that distinguishes between real and generated samples to improve reconstruction quality. The training objective minimizes two key components: a reconstruction loss (Equation 3.1), which ensures the model can recreate inputs and a regularization of the latent space (Equation 3.2) through Kullback-Leibler (KL) divergence that constrains the learned latent distribution $q_\phi(z|x)$ to match a prior distribution $p_\theta(z)$ (typically standard Gaussian). To enhance the detailed characteristics of generated examples, adversarial training components that include a generator loss (Equation 3.3) and a discriminator loss (Equation 3.4) can be added to the training framework. The generator loss trains both encoder and decoder components based on discriminator feedback, where E represents the expected value computed over all possible values of x sampled from the data distribution $p(x)$.

$$\mathcal{L}_R = E_{q_\phi(z|x)}[||x - \hat{x}||_2^2] \quad (3.1)$$

$$\mathcal{L}_{KL,z} = D_{KL}(q_\phi(z|x)|p_\theta(z)) \quad (3.2)$$

$$\mathcal{L}_{gen} = -E_{x \sim p(x)} \left[\log D_\psi \left(Dec_\theta \left(Enc_\phi(x) \right) \right) \right] \quad (3.3)$$

$$\mathcal{L}_{dis} = -E_{x \sim p(x)} [\log D_\psi(x)] - E_{x \sim p(x)} \left[\log \left(1 - D_\psi \left(Dec_\theta \left(Enc_\phi(x) \right) \right) \right) \right] \quad (3.4)$$

This probabilistic framework enables VAEs to generate new data samples, interpolate in the latent space, and capture patterns in the data representations.

3.3.2 Rotation Invariance

VAEs designed with rotation invariance can encode input images into latent representations that capture essential features regardless of how the image is rotated. In contrast, the convolutional operations used in Deep Convolutional Neural Networks (DCNNs) lack this rotation invariance and produce different outputs when the same image is rotated to different orientations. To achieve it, several approaches can be implemented: incorporating rotational data augmentation during training, using specialized architectures with built-in rotational equivariance (Cohen & Welling, 2016; Mo & Zhao, 2024) , or designing custom functions that penalize orientation-dependent representations (Kurihana, Moyer, et al., 2021; Matsuo et al., 2017). Methods employing rotational equivariant convolutions majorly depend on spatial reprojections, which add complexity to architecture. In contrast, incorporating specific terms in the loss function streamlines the implementation and is aligned with the study's objectives (Kurihana, Moyer et al., 2021). Figure 3.3, Equation 3.6 and Equation 3.7 illustrate the introduced functions with $\mathcal{L}_{Inv}$ the rotational invariance term and $\mathcal{L}_{Relaxed}$ the relaxed reconstruction.

Core Principles for Analyzing Loss Components

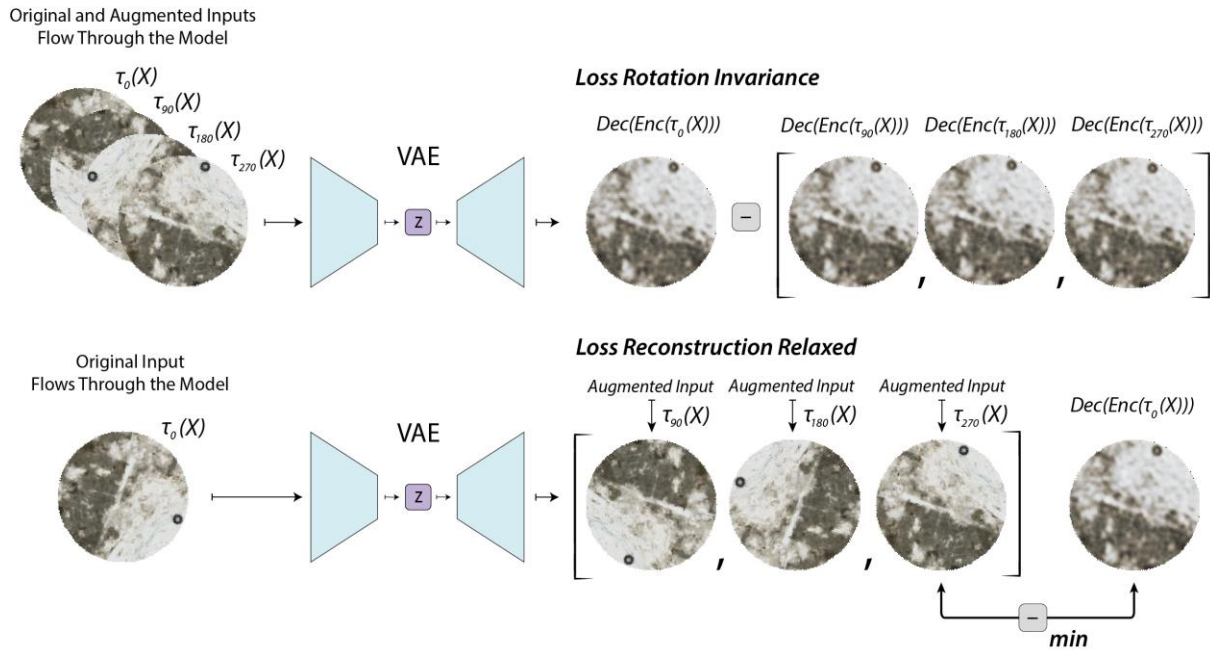


Figure 3.3 Rotation invariance loss minimizes in this example the difference between $Dec(Enc(\tau_0(X)))$, $Dec(Enc(\tau_{90}(X)))$, $Dec(Enc(\tau_{180}(X)))$ and $Dec(Enc(\tau_{270}(X)))$ that are all reconstructed to the canonical orientation $\tau_{270}(X)$. The reconstruction relaxed term minimizes the difference $Dec(Enc(\tau_0(X)))$ and $\tau_{270}(X)$ which is the rotated input corresponding to the matching reconstructed image.

Formally, for each image x in a minibatch S , the model restores both the original image and N versions that have been rotated using a unique scalar rotation operator τ from a set of

rotations R . The invariance loss in Equation 3.6 reduces the difference between the restored original output and all restored augmented outputs.

Then, to enable the network to learn spatial substructures in images, the restoration term is relaxed in Equation 3.7 to minimize the difference between the restoration $\hat{x}$ of an original image x , and the sample that produce the minimum error among all the versions of original input when rotated by an operator $\tau \in R$, so that spatial structure can be preserved in the input while restoration can be performed to a new canonical orientation.

$$\mathcal{L}_{Inv} = \frac{1}{N} \sum_{x \in S} \sum_{\tau \in R} ||Dec_{\theta}(Enc_{\phi}(x)) - Dec_{\theta}(Enc_{\phi}(\tau(x)))||_2^2 \quad (3.6)$$

$$\mathcal{L}_{Relaxed} = \sum_{x \in S} \min_{\tau \in R} ||Dec_{\theta}(Enc_{\phi}(x)) - \tau(x)||_2^2 \quad (3.7)$$

3.3.3 Model Architecture and Hyperparameters

The model architecture (Figure 3.4) was designed for processing training images using residual convolution blocks (He et al., 2015a) with downsampling and upsampling convolutional layers combined with a patch-discriminator. The encoder consists of four convolutional stages with [64, 128, 256, 256] filters of kernel size 3×3 progressively downsampling the input spatial dimensions from 128×128 to 16×16. Each residual block contains two-unit convolution layers followed by instance normalizations across spatial dimensions, and LeakyReLU activations with 0.2 negative slope. A skip connection adds each block's input to its output. These connections were used to improve gradient flow and feature reuse. The final latent block is flattened and passed through a fully connected layer with 196 neurons and linear activation, to then compute the latent distribution's mean (μ) and log-variance (σ) with dimension $z = 196$. The decoder mirrors the encoder part, using upsampled convolutions instead of stride convolutions. The rotation mechanism consists of 5 operators, each separated by 60° increments to complete a full 360° rotation. This design choice was implemented due to memory constraints. The discriminator consists of 4 convolutional layers with respectively [64, 128, 256, 256] filters and strides from the second to penultimate layer. Each layer employs instance normalization and LeakyReLU activation with 0.2 negative slope, except for the final layer that uses sigmoid activation to perform patch-wise classification (Demir & Unal, 2018; Zhu et al., 2017).

Global Network Architecture

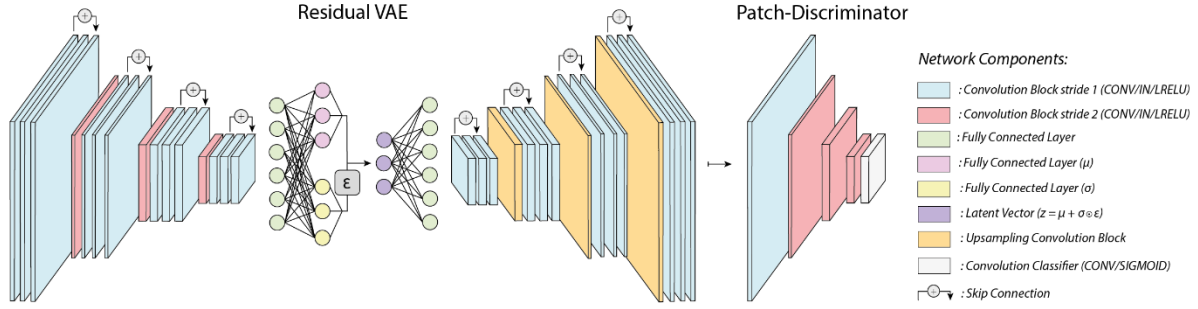


Figure 3.4 Residual VAE network architecture made of 4 residual convolution downsampling and upsampling stages. Last convolutional feature maps generated by the encoder go through an intermediate fully connected layer before going to the latent mean and variance with the reparameterization trick ($z = \mu + \sigma \odot \epsilon$) allowing gradient back-propagation. The patch discriminator serves as a texture classifier, determining the realism of locally generated patterns within image patches.

The model was optimized using Equation 3.8 (with $\alpha = 2 \times 10^{-2}$, $\gamma = 1 \times 10^{-2}$, $\beta = 1 \times 10^{-1}$) on 21,145 batches, each containing 8 images. This training involved 100 iterations (epochs) with constant index shuffling and no data augmentation. The Adam optimizer was selected with an exponential decay learning rate strategy going from 1×10^{-4} to 5×10^{-4} .

$$\min (\mathcal{L}_{Global} = \alpha \cdot \mathcal{L}_{Relaxed} + \gamma \cdot \mathcal{L}_{Inv} + \mathcal{L}_{dis} + \mathcal{L}_{gen} + \beta \cdot \mathcal{L}_{KL,z}) \quad (3.8)$$

3.3.4 Model Evaluation

The performance of the network was first evaluated through continuous monitoring of component losses variation and the assessment of reconstruction quality. Following the training process, an analysis of the representations between the original images and their rotated augmentations was conducted to ensure that the model effectively captured this feature invariance. Then, latent vectors of pre-classified samples were extracted from the encoder network, followed by the application of t-SNE (t-distributed Stochastic Neighbor Embedding) (Van Der Maaten & Hinton, 2008), a technique that reduces high-dimensional data to smaller dimensions by converting distances between points into probabilities and creating a representation that preserves local relationships between points, making it easier for visualizing clusters in data. The number of components chosen for the t-SNE representation is 2 with a perplexity of 30.

3.3.5 Representation Clustering

Representation clustering analysis was performed using Gaussian Mixture Models (GMM) (McLachlan & Peel, 2000) to group latent vectors learned by the neural architecture. GMM represents a suitable choice given its natural compatibility with the Gaussian prior inherent in VAE latent spaces. Thus, embeddings from the labelled dataset were extracted and used to train a 12-component clustering model optimized using the Bayesian Information Criterion (BIC). This approach aims to identify inherent clusters within the latent space, revealing distinct subgroups that correspond to either meaningful or negligible variations in input features.

3.3.6 Semi-Supervised Fine-Tuning

While learning representations from images and performing successive clustering often yields less robust performance compared to traditional supervised training, this approach can still provide valuable insights when properly refined. To enhance the model capabilities, the trained VAE encoder was augmented using a projection head consisting of successive layers comprising 32 and 10 fully connected neurons respectively with ReLU and softmax activations. This adapted architecture was fine-tuned on the labeled set of images for 100 epochs, using a weighted cross-entropy loss addressing classes imbalances and a learning rate that decreased progressively from 5×10^{-4} to 1×10^{-5} . The data was split equally between training and validation to better test model generalization (ratio = 0.5).

Two approaches were compared, the proposed method using pre-trained VAE features against the equivalent model initialized with random weights. The effectiveness of learning representations from the full dataset of images was assessed to determine if classification performance could be improved on the validation set using standard classification metrics including accuracy (Equation 3.9), precision (Equation 3.10), recall (Equation 3.11), and F1 score (Equation 3.12), all computed from the underlying true positives (TP), true negatives (TN), false positives (FP), and false negatives (FN). Confusion matrices were also used and analyzed during inference (Fawcett, 2006).

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \quad (3.9)$$

$$Precision = \frac{TP}{TP + FP} \quad (3.10)$$

$$Recall = \frac{TP}{TP + FN} \quad (3.11)$$

$$F1 = 2 \times \frac{Precision \times Recall}{Precision + Recall} \quad (3.12)$$

The training procedure was replicated using 3-fold cross-validation to ensure representative evaluation of the tested approaches

3.4 Results

3.4.1 Model Training

The different training losses evolved from 18.54, 2.39, 2.67, 8.75 and 3.75 to 8.05, 0.53, 3.63, 4.27 then 0.59 respectively for $\mathcal{L}_{R_{relaxed}}$, $\mathcal{L}_{KL,z}$, $\mathcal{L}_{Inv}$, $\mathcal{L}_{Gen}$ and $\mathcal{L}_{Disc}$. As exhibited in Figure 3.5, the training loss $\mathcal{L}_{R_{relaxed}}$ and $\mathcal{L}_{KL,z}$ globally show a decreasing trend while $\mathcal{L}_{Inv}$ slightly increased over early epochs and decreased lately. The adversarial losses $\mathcal{L}_{Gen}$ and $\mathcal{L}_{Disc}$ decreased for early epochs and oscillated over stable values until training ends. The total average loss started from 9.11 and ended down to 4.19.



Figure 3.5 Evolution of VAE training loss components across 100 epochs.

Figure 3.6 represents an input sample and its 5 rotated versions. The encoded vectors of dimensions 196 were reshaped to size 14×14 for visualization. These latent representations show a similar structure even if not perfectly identic. The decoded vectors led to the reconstruction of the original input image rotated by the operator τ_{120} .

Image Reconstruction and Rotation Invariance

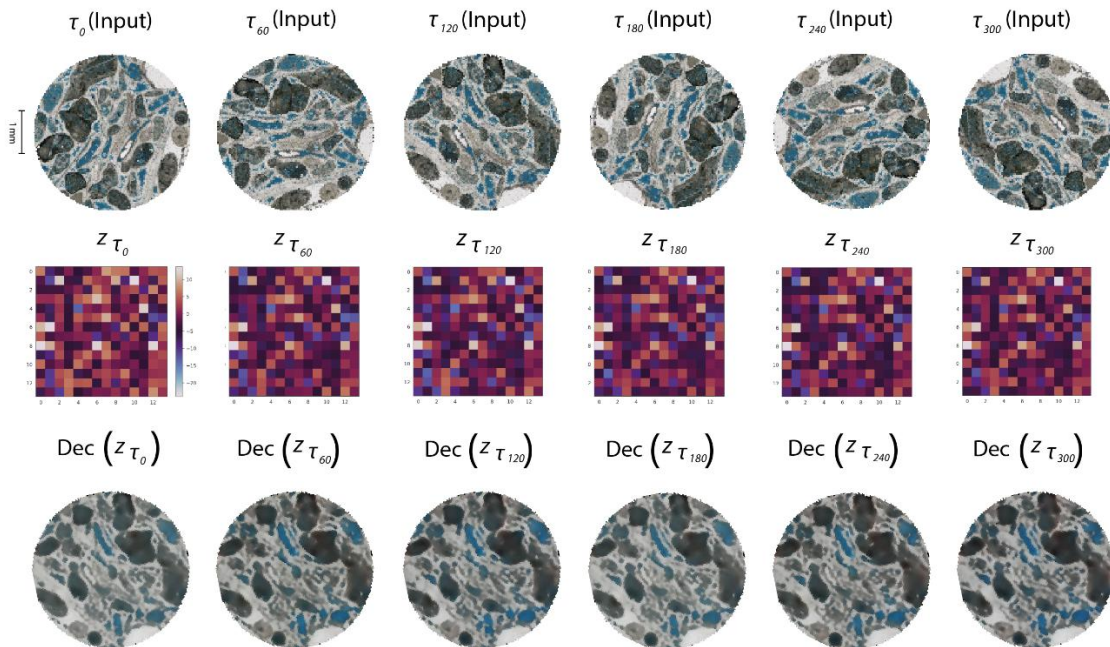


Figure 3.6 Image reconstruction analysis showing input images at 60° incremental rotations where illustration includes latent space mapping and reconstructions, with the final output rotated by operator τ_{120} relative to the reference image.

3.4.2 Latent Space Distribution

The analysis was conducted on the labeled set of images, generating 45,699 latent vectors with 196 dimensions. These vectors were then projected into a probabilistic space using t-SNE, as illustrated in Figure 3.7.

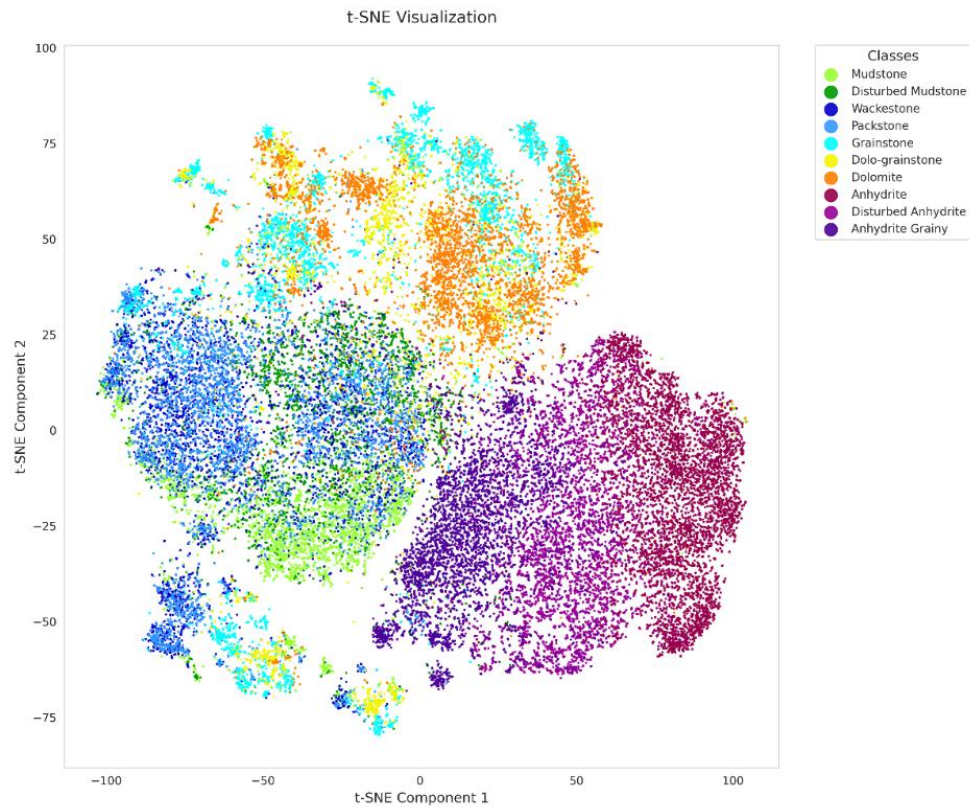


Figure 3.7 t-SNE mapping of latent vectors for the set of labeled images.

The visualization reveals distinct clustering patterns among various litho-texture classes that effectively discriminate between principal rock types (anhydrites, dolomite and limestones).

Anhydrite samples (shown in magenta to purple hues) form the most cohesive and well-separated clusters in the lower-right portion of the plot. These samples demonstrate tight grouping with a clear compositional gradient observable between the three anhydrite sub-classes (anhydrite, disturbed anhydrite, and anhydrite grainy), suggesting systematic variations in crystallinity patterns or structural disturbance.

Dolomite and dolo-grainstone samples (yellow and orange) are preferentially located in the upper region of the visualization space and show local mixing with grainstone samples (light blue). This spatial arrangement reflects their original grain-supported characteristics and similar visual appearance within the dataset, with grainstones that were transformed into dolomites.

Higher-order divisions in limestone classes, notably composed of muddy matrix, are distinguishable as separate groups but exhibit overlaps between different sub-classes. This

overlap pattern captures the natural textural continuum present in carbonate rocks, where grain-to-matrix ratios transition gradually rather than forming sharp boundaries. The mudstone classes (light and dark green) occupy intermediate positions and show distinguishable clustering, though with some transitional overlap.

Isolated clusters are also evident in the lower region of the plot, consisting primarily of grain-supported limestones (with mud or spar cement) that have undergone local dolomitization. These distinctive groupings potentially highlight shared characteristics that set them apart from other carbonate groups.

3.4.3 Representation Clustering

Clustering the encoder representations enables detailed analysis of class distributions within each group, revealing distinct trends in the data. Figure 3.8 demonstrates the class associations within each cluster, highlighting groups with varying label compositions and local affinities for specific litho-textures.

Image Representations Clustering

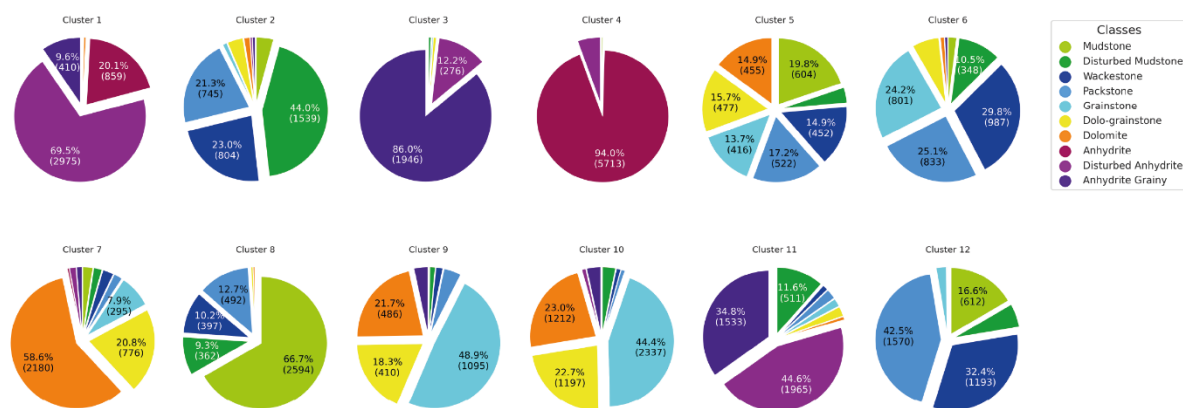


Figure 3.8 Distribution of litho-texture rock types across formed clusters.

Clusters 1, 3, and 4 are predominantly composed of anhydrite variations (>95%). Cluster 4 exhibits particularly high exclusivity for homogeneous facies (94%), while clusters 3 and 1 are primarily characterized by grainy (86%) and disturbed textures (69.5%), respectively. Cluster 11 also consists mainly of anhydrite, with disturbed and grainy sub-textures occurring in relatively close proportions (44.6% and 34.8%).

Dolomite is most prominently represented in cluster 7 (58.6%), where it groups with dolo-grainstones. It also appears in clusters 5, 9, and 10, though in lower proportions (<23%).

The grainstone class dominates clusters 9 (48.9%) and 10 (44.4%), appearing alongside dolo-grainstones and dolomites. This class is also present in clusters 5 and 6, but in lower abundances (<25%).

Muddy limestones containing significant grain proportions such as wackestone and packstone are primarily found in cluster 12 (42.5% and 32.4%, respectively), cluster 6 (25.1% and 24.2%), and cluster 2 (21.3% and 23%).

Clusters 2 and 8 consist mainly of mud-dominated limestones, with cluster 2 being primarily associated with disturbed facies (44%) and cluster 8 with homogeneous forms (66.7%).

Cluster 5 stands apart as the only group without a clearly predominant component, making it more challenging to characterize and utilize effectively.

3.4.4 Semi-Supervised Fine-Tuning

Models initialized with pre-trained weights consistently outperform those trained from scratch, demonstrating the value of leveraging prior knowledge of the target pretext task (Figure 3.9 and Table 3.1).

Comparative Analysis of Top-Performing Model

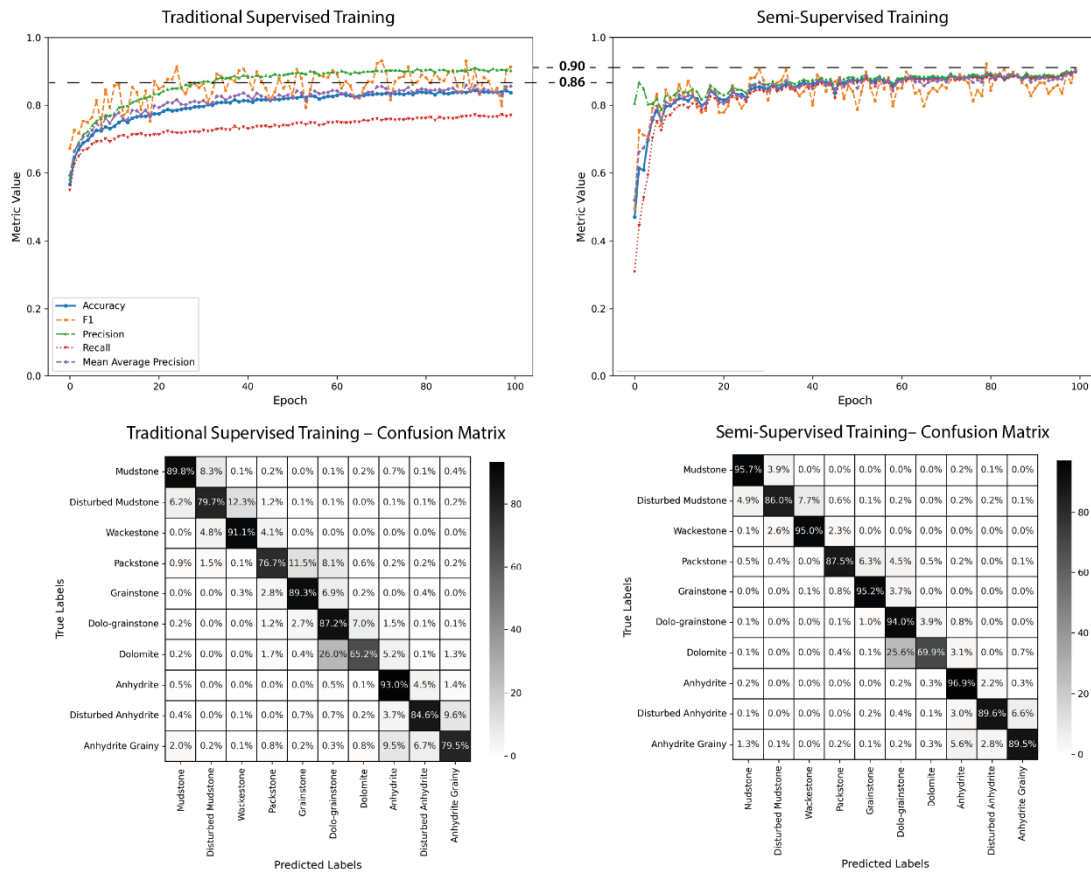


Figure 3.9 Classification metrics and confusion matrices for the best-performing supervised and semi-supervised models, evaluated on validation data.

Metric	Training Methods		Δ
	Traditional	Semi-Supervised	
Accuracy	82.8%	86.7%	+3.9%
F1	88.6%	88.9%	+0.3%
Precision	89.3%	89.5%	+0.2%
Recall	79.9%	83.7%	+3.8%
mAP	85.1%	88.6%	+3.5%

Table 3.1 Classification metrics for supervised and semi-supervised approaches, averaged across 3 training repetitions.

The semi-supervised approach produces several key advantages with faster convergence, more stable training and more uniform metric distributions. Performance gains range from +0.3% to +3.9% when averaged across the three-time cross validation training configuration.

Analysis of the confusion matrices reveals that the pre-trained model makes fewer classification errors overall. However, both approaches still struggle with dolomite and dolo-grainstone differentiation, indicating this remains a challenging classification boundary regardless of the initialization strategy.

3.4.5 Full Thin Section Characterization

Carbonate rocks display significant heterogeneity at the microscopic scale, making it valuable to classify regions of interest (ROI) in thin section images to reveal the lithological and textural variations within samples. To accomplish this, the model fine-tuned for litho-texture classification was employed to analyze patches of complete thin sections. A sliding window approach with 80% overlap was used to generate multiple predictions for each area, and the results were averaged at patch intersections, ultimately producing detailed classification maps that capture the spatial distribution of different rock types and textures. Figure 3.10 presents five samples arranged from left to right in order of increasing complexity.

Global Thin Section Characterization using Semi-Supervised Classification Network

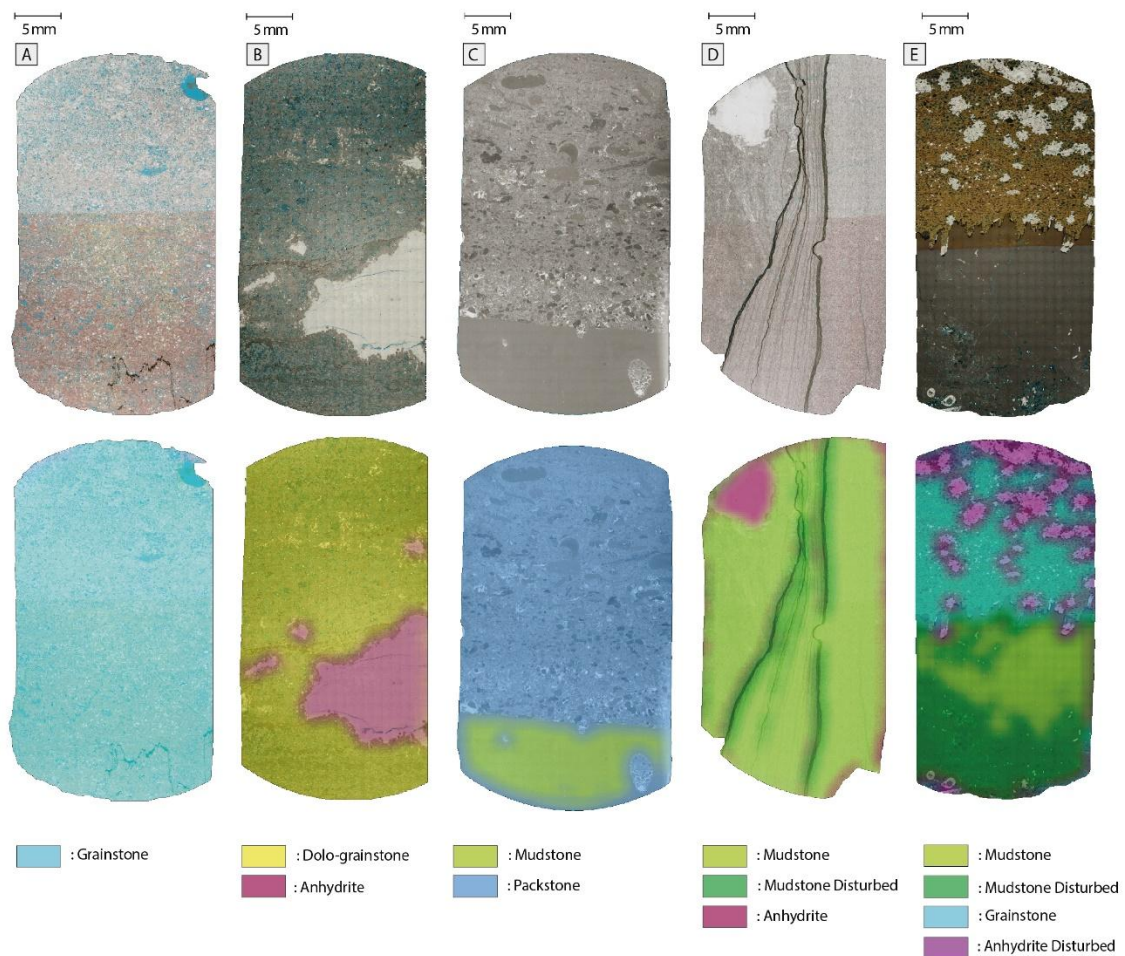


Figure 3.10 Classification maps examples of full thin section samples highlighting various litho-texture types.

Figure 3.10A shows a grainstone with a globally homogeneous texture throughout the section, excluding minor grain size variation and stylolites that were outside the scope of this classification.

Figure 3.10B is primarily composed of two distinct litho-textures: a dolo-grainstone section featuring ghost-preserved allochems and a dissolved matrix, accompanied by localized anhydrite nodules.

Figure 3.10C exhibits two principal limestone classifications: an upper section displaying a grain-supported framework within a muddy matrix, and a lower section marked by an abrupt transition to mud deposits.

The predicted classifications in Figure 3.10D delineate a sample predominantly consisting of carbonate mud with stylolitized regions in the central area and an anhydrite nodule positioned in the upper left corner.

The sample illustrated in Figure 3.10E reveals an upper grainstone section containing small, heterogeneous anhydrite patches. A stylolite demarcates the transition to the lower section, which comprises homogeneous to heterogeneous mud containing sparse grains.

The patch classification performed across complete thin section areas demonstrated globally accurate identification of litho-texture types present within the samples, although localized border effects remain observable, as evidenced in Figures 3.10E and 3.10D.

3.5 Discussion and Recommendations

The rotational-invariant network architecture achieved global convergence despite low or oscillating minimization in some training loss components. This behavior is typical for adversarial losses, and the initial increase in rotational invariance can be attributed to early reconstructions being noisy and structurally incoherent. Hyperparameter tuning proved sensitive for optimal convergence, suggesting that additional testing could further improve results. While increasing latent space dimensionality might enhance both rotational invariance and reconstruction quality, this must be carefully balanced against clustering complexity and potential overfitting in downstream classification tasks.

Testing confirmed the model's ability to maintain structural consistency in latent representations across rotational transformations. Encoded vectors for rotated versions of the same image exhibited similar latent space structures, though total invariance is still complex to be perfectly achieved. The reconstructed output consistently oriented to canonical rotations, demonstrating that the model successfully learned to decode features while preserving common spatial substructures.

t-SNE visualization revealed distinct and meaningful clustering patterns among different lithologies and textures. Anhydrite samples formed the most well-separated clusters probably due to their simpler visual properties. Dolomite and dolo-grainstone samples showed preferential clustering with partial grainstone samples mixing, reflecting their shared characteristics related to allochem presence. Limestone classes displayed separated yet overlapping groups, potentially capturing the gradual transitions characteristic of muddy carbonate rocks rather than clearly indicating model limitations.

GMM clustering confirmed these results, showing clearly distinct clusters for anhydrite and its variations. However, other fine-grained rock textures were intermixed with geologically similar classes, making it difficult to reliably separate the subclasses found in limestones and dolomites. This observation justifies fine-tuning the model on labeled data to achieve the precision needed for accurate rock litho-texture identification. Other methods could be explored for future experiments, including integrating Gaussian Mixture Models within the VAE learning process to learn probabilistic latent representations where data points are directly assigned to different Gaussian clusters (Dilokthanakul et al., 2016). VQ-VAE offers another promising approach through discrete latent representations by quantizing encoder outputs to a learned codebook of discrete vectors (Oord et al., 2017), potentially creating more structured and interpretable latent spaces suitable for clustering frameworks.

Using the pre-trained VAE encoder outperformed random initialization, delivering better results on all metrics while converging faster and maintaining more stable training. This highlights the relevance of semi-supervised methods that learn representations from unlabeled data during pre-training (Nellas et al., 2022). Nevertheless, both methods had difficulty distinguishing between dolomite and dolo-grainstone, suggesting this classification challenge stems potentially from fundamental issues in how the dataset was constructed rather than the modeling approach itself. Indeed, ghost-preserved allochems and authigenic dolomite often display indistinguishable microstructures, making their petrographic distinction uncertain.

The patch-classification approach using a sliding window with overlap successfully demonstrated practical applicability for spatial heterogeneity analysis. The method effectively highlighted lithological and textural boundaries across complete samples and accurately identified major litho-textural regions in samples of varying complexity. Localized border effects were observed in some samples and could potentially be mitigated through proper translation augmentation techniques. Moreover, this study did not specifically target stylolites and grain component heterogeneities, which would be better addressed using dedicated segmentation methods.

3.6 Conclusion

This study demonstrates the usefulness of a rotational-invariant VAE framework for carbonate petrographic analysis, offering a practical solution for processing initial large unlabeled and limited annotated datasets. The method effectively learns meaningful representations of carbonate rock litho-textures while maintaining rotational invariance, a relevant feature in petrographic analysis where image orientation varies arbitrarily.

The approach achieves a first unsupervised clustering that reveals major lithological groupings, though fine-textural differentiation required semi-supervised fine-tuning with minimal labeled data. This hybrid methodology outperformed traditional supervised approaches and demonstrated practical value through full thin section characterization that accurately identifies spatial heterogeneity.

Most significantly, semi-supervised approaches that rely on meaningful data encoding are considered the right direction, given both the difficulty of manual classification at scale and

the potential for classification nomenclatures to change based on specific needs. These new approaches for efficient preliminary data sorting and identification of specific or isolated groups provide a foundation for future developments in assisted geological image analysis systems.

REFERENCES

- Bergen, K. J., Johnson, P. A., De Hoop, M. V., & Beroza, G. C. (2019). Machine learning for data-driven discovery in solid Earth geoscience. *Science*, 363(6433), eaau0323. <https://doi.org/10.1126/science.aau0323>
- Boggs, S. J. (2009). *Petrology of sedimentary rocks* (2nd ed.). Cambridge University Press. <https://doi.org/10.1017/CBO9780511626487>
- Burgess, J., Nirschl, J. J., Zanellati, M. C., Lozano, A., Cohen, S., & Yeung-Levy, S. (2024). Orientation-invariant autoencoders learn robust representations for shape profiling of cells and organelles. *Nature Communications*, 15(1), 1758. <https://doi.org/10.1038/s41467-024-45362-4>
- Chen, T., Kornblith, S., Norouzi, M., & Hinton, G. (2020). A simple framework for contrastive learning of visual representations. *Proceedings of the 37th International Conference on Machine Learning*, 119, 1597-1607. arXiv:2002.05709
- Cohen, T. S., & Welling, M. (2016). Group equivariant convolutional networks. *Proceedings of the 33rd International Conference on Machine Learning*, 48, 2990-2999. arXiv:1602.07576
- Demir, U., & Unal, G. (2018). Patch-based image inpainting with generative adversarial networks. arXiv:1803.07422 <https://doi.org/10.48550/arXiv.1803.07422>
- Dilokthanakul, N., Mediano, P. A. M., Garnelo, M., Lee, M. C. H., Salimbeni, H., Arulkumaran, K., & Shanahan, M. (2016). Deep unsupervised clustering with Gaussian mixture variational autoencoders. arXiv:1611.02648
- Dott, R. H. (1964). Wacke, graywacke and matrix; what approach to immature sandstone classification? *Journal of Sedimentary Petrology*, 34(3), 625-632. <https://doi.org/10.1306/74D71109-2B21-11D7-8648000102C1865D>
- Dunham, R. J. (1962). Classification of carbonate rocks according to depositional textures. In W. E. Ham (Ed.), *Classification of carbonate rocks—A symposium* (pp. 108-121). American Association of Petroleum Geologists Memoir 1.
- Embry, A. F., & Klován, J. E. (1971). A late Devonian reef tract on northeastern Banks Island, N.W.T. *Bulletin of Canadian Petroleum Geology*, 19(4), 730-781.
- Fawcett, T. (2006). An introduction to ROC analysis. *Pattern Recognition Letters*, 27(8), 861-874. <https://doi.org/10.1016/j.patrec.2005.10.010>
- Flügel, E. (2004). *Microfacies of carbonate rocks: Analysis, interpretation and application* (2nd ed.). Springer. <https://doi.org/10.1007/978-3-662-08726-8>
- Folk, R. L. (1959). Practical petrographic classification of limestones. *AAPG Bulletin*, 43(1), 1-38.
- Folk, R. L. (1974). *Petrology of sedimentary rocks* (2nd ed.). Hemphill Publishing Company.
- Goodfellow, I., Pouget-Abadie, J., Mirza, M., Xu, B., Warde-Farley, D., Ozair, S., Courville, A., & Bengio, Y. (2014). Generative adversarial networks. *Advances in Neural Information Processing Systems*, 27, 2672-2680.

- Gui, J., Chen, T., Zhang, J., Cao, Q., Sun, Z., Luo, H., & Tao, D. (2023). A survey on self-supervised learning: Algorithms, applications, and future trends. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 45(7), 8927-8949. arXiv:2301.05712
- He, K., Chen, X., Xie, S., Li, Y., Dollár, P., & Girshick, R. (2021). Masked autoencoders are scalable vision learners. arXiv:2111.06377
- He, K., Zhang, X., Ren, S., & Sun, J. (2015). Deep residual learning for image recognition. arXiv:1512.03385
- Hinton, G. E., & Salakhutdinov, R. R. (2006). Reducing the dimensionality of data with neural networks. *Science*, 313(5786), 504-507. <https://doi.org/10.1126/science.1127647>
- Kingma, D. P., & Welling, M. (2013). Auto-encoding variational Bayes. arXiv:1312.6114
- Koeshidayatullah, A., Morsilli, M., Lehrmann, D. J., Al-Ramadan, K., & Payne, J. L. (2020). Fully automated carbonate petrography using deep convolutional neural networks. *Marine and Petroleum Geology*, 122, 104687. <https://doi.org/10.1016/j.marpetgeo.2020.104687>
- Kurihana, T., Moyer, E., Willett, R., Gilton, D., & Foster, I. (2021). Data-driven cloud clustering via a rotationally invariant autoencoder. *IEEE Transactions on Geoscience and Remote Sensing*, 60, 1-15. <https://doi.org/10.1109/TGRS.2021.3098008>
- Liu, X., Chandra, V., Ramdani, A. I., Zuhlke, R., & Vahrenkamp, V. (2023). Using deep-learning to predict Dunham textures and depositional facies of carbonate rocks from thin sections. *Geoenery Science and Engineering*, 227, 211906. <https://doi.org/10.1016/j.geoen.2023.211906>
- Lokier, S. W., & Al Junaibi, M. (2016). The petrographic description of carbonate facies: Are we all speaking the same language? *Sedimentology*, 63(7), 1843-1885. <https://doi.org/10.1111/sed.12293>
- Mamode, H. I., Hampson, G. J., & John, C. M. (2025). Do more with less: Exploring semi-supervised learning for geological image classification. *Applied Computing and Geosciences*, 25, 100216. <https://doi.org/10.1016/j.acags.2024.100216>
- Matsuo, T., Fukuhara, H., & Shimada, N. (2017). Transform invariant auto-encoder. In *2017 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)* (pp. 2359-2364). IEEE. <https://doi.org/10.1109/IROS.2017.8206047>
- McLachlan, G. J., & Peel, D. (2000). *Finite mixture models*. John Wiley & Sons.
- Mo, H., & Zhao, G. (2024). RIC-CNN: Rotation-invariant coordinate convolutional neural network. *Pattern Recognition*, 146, 109994. <https://doi.org/10.1016/j.patcog.2023.109994>
- Munjal, P., Paul, A., & Krishnan, N. C. (2019). Implicit discriminator in variational autoencoder. In *2020 International Joint Conference on Neural Networks (IJCNN)* (pp. 1-8). IEEE.
- Nellas, I. A., Tasoulis, S., Plagianakos, V. P., & Georgakopoulos, S. V. (2022). Supervised dimensionality reduction and image classification utilizing convolutional autoencoders. *SN Computer Science*, 3(6), 455.

Noroozi, M., & Favaro, P. (2016). Unsupervised learning of visual representations by solving jigsaw puzzles. In *European Conference on Computer Vision* (pp. 69-84). Springer. arXiv:1603.09246

Oord, A. van den, Vinyals, O., & Kavukcuoglu, K. (2017). Neural discrete representation learning. *Advances in Neural Information Processing Systems*, 30, 6306-6315. arXiv:1711.00937

Pires de Lima, R., Duarte, D., Nicholson, C., Slatt, R., & Marfurt, K. J. (2020). Petrographic microfacies classification with deep convolutional neural networks. *Computers and Geosciences*, 142, 104481. <https://doi.org/10.1016/j.cageo.2020.104481>

Reichstein, M., Camps-Valls, G., Stevens, B., Jung, M., Denzler, J., Carvalhais, N., & Prabhat. (2019). Deep learning and process understanding for data-driven Earth system science. *Nature*, 566(7743), 195-204. <https://doi.org/10.1038/s41586-019-0912-1>

Rivera, M. (2023). How to train your VAE. In *2023 International Conference on Information Photonics (ICIP)* (pp. 1-6). IEEE.

Sibley, D. F., & Gregg, J. M. (1987). Classification of dolomite rock textures. *Journal of Sedimentary Research*, 57(6), 967-975.

Van Der Maaten, L., & Hinton, G. (2008). Visualizing data using t-SNE. *Journal of Machine Learning Research*, 9, 2579-2605.

Warren, J. K. (2010). Evaporites through time: Tectonic, climatic and eustatic controls in marine and nonmarine deposits. *Earth-Science Reviews*, 98(3-4), 217-268. <https://doi.org/10.1016/j.earscirev.2009.11.004>

Zhu, J.-Y., Park, T., Isola, P., & Efros, A. A. (2017). Unpaired image-to-image translation using cycle-consistent adversarial networks. In *Proceedings of the IEEE International Conference on Computer Vision* (pp. 2223-2232). IEEE. arXiv:1703.10593

CHAPTER 4: SYNSECTION – SEDIMENTOLOGY-DRIVEN DATA GENERATION FOR DEEP LEARNING APPLICATIONS IN CARBONATES PETROGRAPHY

Abstract

Paper 2: Published in Marine and Petroleum Geology

doi.org/10.1016/j.marpetgeo.2025.107490

Patent INV23-205

The analysis of carbonate rocks through petrographic methods has long posed significant challenges in geological sciences, particularly regarding the systematic description and quantification of thin sections. Recent developments in artificial intelligence suggest promising avenues for automation; however, the field remains constrained by limited training data availability, as manual annotation requires considerable time investment and can result in variable interpretations. This study addresses these challenges by introducing SynSection, a method for generating synthetic pairs of carbonate thin section images with corresponding labels for both classification and segmentation tasks. This approach has been carefully designed to incorporate sedimentological heuristics, ensuring the geological validity of generated samples through the integration of three primary components: 2D grain packing, groundmass generation using texture synthesis, and image blending for image composition. The advantage of this methodology lies in its ability to dramatically expand a limited set of manually labeled images, thereby reducing the resources required for dataset creation while maintaining consistency in annotation. The method was evaluated through the generation of 100,000 synthetic image-annotation pairs, implementing a transfer learning strategy in which models were initially trained on synthetic data before being fine-tuned on real thin section images. This approach demonstrated significant improvements in both classification and segmentation tasks compared to conventional training methods that rely solely on limited real data. In classification tasks, the transfer learning strategy enhanced performance metrics by 7.2%, achieving a top accuracy score of 94.9%. The impact of the method was particularly notable in segmentation tasks, where the binary grain segmentation model achieved an Intersection over Union (IoU) of 77.2% and a point counting determination score of 93.2%. The multiclass segmentation achieved a mean Intersection over Union (mIoU) of 55.3% and a point counting score of 89.6%, reflecting the inherent complexity of grain type differentiation. These results, which are consistent with medical imaging applications, demonstrate the method's potential to optimize carbonate thin section analysis by reducing time allocated to descriptive tasks. The approach generates diverse training data that maintains geological validity, establishing a quantitative framework for automated petrographic analysis that aligns with standardized classification criteria and expert-defined features.

Keywords: Carbonates, Petrography, Deep Learning, Synthetic Data, Image Classification, Image Segmentation.

4.1 Introduction

The petrographic analysis of carbonate rocks represents one of the major challenges in sedimentary geology. The complexity of this analysis lies in the multiplicity of processes involved in the formation and evolution of these rocks, from their initial deposition to their present state. Carbonates may contain diverse components: skeletal grains, non-skeletal grains, matrix, cements and pore spaces, resulting from the interaction between biological, physico-chemical, and diagenetic factors (Flügel, 2004). Adding to this, carbonates often show significant variations in both textures and grain distribution within the same rock body (Scholle & Ulmer-Scholle, 2003). The analysis of these rocks can be conducted through the examination of high-resolution thin section images at multiple scales, with the choice of scale typically determined by the size of the features to be characterized.

When visually analyzing a thin section at its full extent, it becomes challenging to capture all the visible information while maintaining consistency and avoiding bias in the description process due to various possible scales of analysis (Figure 4.1 challenge A.1). Furthermore, the final interpretation of depositional sequences is typically reduced to a facies classification system that groups major observational categories but often fails to fully represent the heterogeneities observed within the samples.

The characterization of thin sections typically involves selecting multiple regions of interest (ROI) that present some degree of local homogeneity, which simplifies the analysis process. Initial examination of the selected regions is usually performed qualitatively through visual description to determine the rock texture. This textural analysis typically follows the Dunham classification, which remains the traditional method to categorize carbonate rocks (Dunham, 1962; Embry & Klovan, 1971). Although this approach is probably the most widely used to describe carbonates, it is inherently subjective as it relies on visual estimation of quantitative parameters (Figure 4.1 challenge A.2). These parameters are mainly based on the relative proportion and size of grains, as well as the type of groundmass material surrounding the elements within the observed field of view. This subjectivity can lead to biases, errors, or the creation of intermediate classes, particularly when samples display characteristics that fall between two adjacent or multiple categories (Lokier & Al Junaibi, 2016).

Quantitative analyses provide additional insights to help interpret the depositional environment of sediment accumulation. Grain identification and quantification are typically measured in several ROIs and then qualitatively extrapolated and reported at full image scale. This is necessary given that a thin section may contain thousands of grains, making manual analysis impractical (Figure 4.1 challenge A.3). Therefore, this method emphasizes its limitations when the selected regions are not representative of the overall sample. The orientation of the sectioning plane within a complex three-dimensional grain framework or the presence of overlying diagenetic imprints may also introduce further ambiguity by partially or completely obscuring the sediment's primary characteristics.

Limitations of Traditional Methods Addressed Through Modern Numerical Approaches:

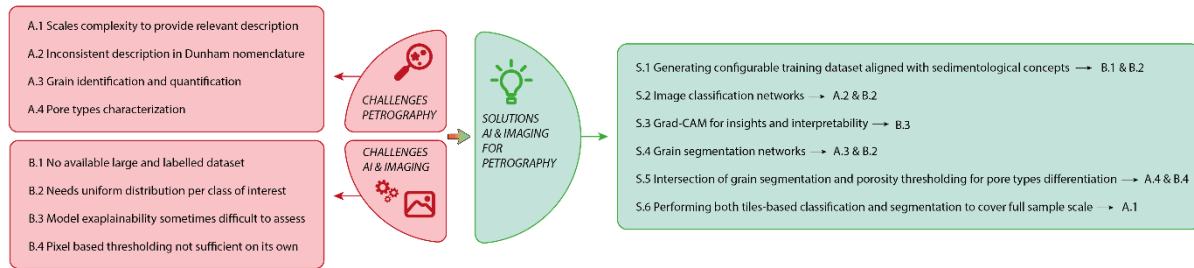


Figure 4.1 Integration of artificial intelligence and modern imaging techniques to address traditional challenges in petrography: a comprehensive framework of problems and solutions for enhanced geological analysis.

Diagenetic processes drive the types of mineralogical transformations and dissolution that can affect grains, cement, or matrix. Identifying the cement or pore types is generally valuable for understanding the conditions the rock experienced to reconstruct its history (Harris et al., 1985). However, accurately delineating and classifying these features can be challenging due to indistinct boundaries between cement phases or the complex arrangement of pore structures (Figure 4.1 challenge A.4), making this analytical step one of the most detailed in carbonates petrography.

Over the past decade, the field of Deep Learning (DL) has experienced a remarkable surge in interest and development. These advancements have demonstrated effectiveness across various domains, including computer vision applications for both broad geoscience research (Bergen et al., 2019; Reichstein et al., 2019) and specific carbonates related challenges (Anjos et al., 2020; Dawson et al., 2023; Jarvis et al., 2024; Maas et al., 2023; Zheng et al., 2021), primarily using Deep Convolutional Neural Networks (DCNNs). However, the application of advanced techniques to the deep analysis of carbonate thin section images remains limited (Koeshidayatullah et al., 2020, 2022; Liu et al., 2023). The use of such technology could significantly simplify the tasks of describing and quantifying these images, which are otherwise tedious, complex, and time-consuming when manually processed. While a petrographer must carefully select relevant ROIs for analysis, a deep learning model can process numerous images in seconds, maximizing the value of the original data and ensuring consistency over time. Nevertheless, a major limitation in applying DL to carbonate studies is the lack of available and properly labeled datasets (Figure 4.1 challenge B.1). DCNNs are data-driven models that conventionally require large quantities of training examples to optimize the millions of parameters they contain, enabling task generalization across diverse new data (Lecun et al., 2015). To address this issue, data augmentation techniques are generally employed to increase the variability of the training images but are typically limited to simple transformations (e.g. rotations, flips, brightness and contrast adjustments, colorimetric changes, distortions, etc.) (Krizhevsky et al., 2012). While these methods can be useful, they may not be sufficient when working with small training databases. Therefore, this study proposes SynSection, a method to create a synthetic training dataset using a limited initial database. This approach aims to better support various DL approaches, aligning with the characterization process conventionally followed by specialists. The generation technique is driven by sedimentological concepts and is configured to produce geologically plausible

examples, along with realistically rendered images and labels (Figure 4.1 solution S.1 addressing challenge B.1).

To date, most studies applying DL to carbonate petrography have focused on image classification tasks (Liu et al., 2023; Mosser et al., 2022; Pires de Lima et al., 2020; Shebl et al., 2021). Although these approaches have demonstrated high accuracy (> 95%) and work well for qualitative analysis, their robustness in categorizing extreme cases of Dunham classes is generally not assessed due to unknown grain density distribution within the training datasets (Figure 4.1 challenge B.2). Moreover, raw classification networks generally lack interpretability for highlighting what contributes the most for decision making (Figure 4.1 challenge B.3). This work investigates various DCNNs architectures using a transfer learning approach (Zhuang et al., 2019), first training randomly initialized models on a large amount of synthetic data before fine-tuning them with a small number of real images. While transfer learning has been successfully applied to various size image-based geoscientific datasets (Dawson et al., 2023), this approach has not yet been evaluated using synthetic data for a preliminary training stage, despite showing promising results in fields such as medical imaging (Ding et al., 2023; Pezoulas et al., 2024). The developed method showed robust learning of essential Dunham classification parameters, even for complex examples by incorporating balanced distribution of grain abundance per synthetic image classes (Figure 4.1 solutions S.1 and S.2 addressing challenges A.2 and B.2). It was also observed that incorporating the generated synthetic images during training has more positive impact rather than model architecture type. Moreover, a method for model explainability is proposed to ensure it focuses on relevant image features for classification and avoids overfitting (Figure 4.1 solution S.3 addressing challenge B.3).

When it comes to quantifying grains within thin sections, classification models reach their limits, as they only provide a probability vector for the defined classes. To handle this, object detection approaches have been explored to locate and classify grains within carbonate images (Koeshidayatullah et al., 2020; Wang et al., 2023). However, these more complex architectures do not allow for extracting directly the precise contours of grains, which is necessary for subsequent size or morphological analysis (Azzam et al., 2024; Portugal et al., 2024).

Image semantic segmentation, which performs pixel-level classification, has been identified as a method to precisely delineate and classify grains simultaneously. While these techniques have been investigated in clastic petrography (Saxena et al., 2021; Vellappally et al., 2024), they are still relatively rare for carbonates thin section analysis (Koeshidayatullah et al., 2022; Liu et al., 2023; Nanjo, 2020). In this study several DCNNs models with U-Net-like architectures (Ronneberger et al., 2015) were trained with the same transfer learning strategy proposed for image classification allowing control over grain class distribution within synthetic images (Figure 4.1 solutions S.1 and S.4 addressing challenges A.3 and B.2). The use of synthetic data significantly improved the results compared to a traditional training approach, while also minimizing the impact of model architectures. This work found that a binary grain segmentation model was more effective for precisely identifying grains, regardless of their class, compared to a multiclass model, which made the task more complex but produced satisfactory results for further geological interpretations.

Porosity in thin sections can be easily identified by adding blue dye to the epoxy resin that fills the pore spaces. This simplifies the extraction and allows its characterization using dedicated size and morphological features. However, this method does not allow for determining the specific type of pores (Figure 4.1 challenges A.4 and B.4). To overcome this, the combination of porosity thresholding and grain segmentation was used to distinguish between intragranular and intergranular porosity, which could then be quantified (Figure 4.1 solution S.5 addressing challenges A.4 and B.4). The analysis of cement types was not covered in this study, despite its potential to provide valuable petrographic information.

Finally, this investigation explored the use of DL approaches for characterizing carbonate thin section images at multiple scales. Previously proposed methods were combined to extend local-scale descriptions to entire images, demonstrating how these tools effectively facilitate analysis and maximize the data's value by capturing its heterogeneity (Figure 4.1 solution S.6 addressing challenge A.1).

4.2 Data

4.2.1 Raw Data

In this study, 14,679 high-resolution images of carbonate rock thin sections from various private sources were compiled (Figure 4.2A). These images were acquired under plain polarized light at a resolution of $8.8 \times 10^{-1} \mu\text{m}$ per pixel, resulting in image sizes around $50,000 \times 25,000$ pixels. Using raw data is currently limited for several reasons. The most significant limitation is the human labeling factor. Manually labeling a small number of complete images is not feasible in a short timeframe and risks being unrepresentative of the entire dataset. Another reason is the computational constraints. DCNNs are trained using Graphics Processing Units (GPU) with limited memory, which can make the training process impossible if the model size and the batch of images with such dimensions exceed the available resource.

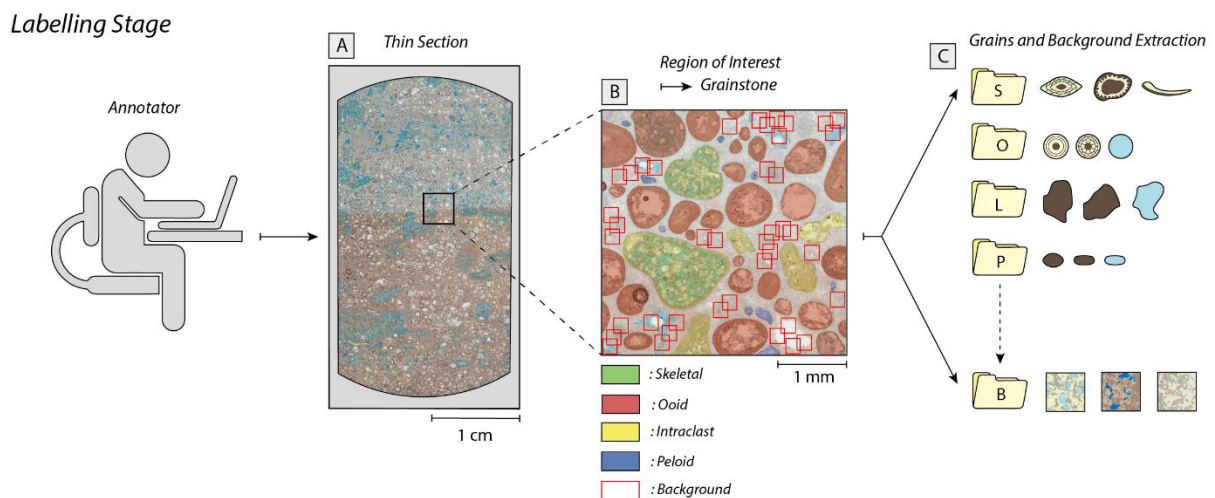


Figure 4.2 Petrographic analysis of a carbonate thin section: a selected region of interest is labeled at the tile scale (grainstone), with grains annotated by class (skeletal, ooid, intraclast, peloid) and background marked using bounding boxes. Annotated grains (thumbnail images and masks) are extracted and categorized by class.

Global Labeled Data Distribution

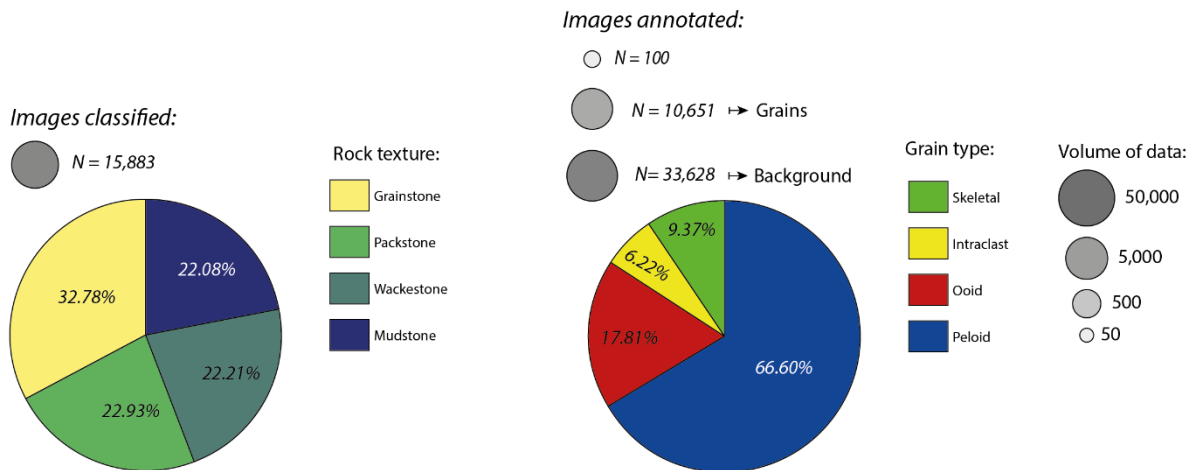


Figure 4.3 Data distribution for classification and segmentation tasks.

4.2.2 Working Data

4.2.2.1 Data Preprocessing

Raw data were preprocessed to convert the images into a format suitable for both human operators and computational resources. First, the images had to be downsampled to a working scale deemed necessary by a human interpreter to identify the rock texture and the grains comprising the thin section for this dataset. This resolution was set to 7 μm per pixel. To facilitate the annotator's work and closely approximate what is observed when describing a ROI, the images were divided into non-overlapping tiles in a way that ensured grains are contained within each tile, with a reasonable number of observed grains, typically between a few dozen and one or two hundred at maximum (Figure 4.2B). This tile size was determined to 512×512 pixels, making it manageable for algorithms like DCNNs.

4.2.2.2 Data Sub-Sampling

A batch of 16,000 tiles was initially randomly selected out of the whole generated images. This choice is justified by a human-limiting factor and corresponds to the number of images that were attempted to be classified during the study.

4.2.2.3 Data Labeling

Out of the 16,000 tiles, 15,883 tiles were manually categorized into four classes according to the Dunham classification leading to the following data distribution: 3,507 mudstones, 3,528 wackestones, 3,642 packstones and 5,206 grainstones (Figure 4.3). The binary grain segmentation InceptionResU-Net model described in section 3.7.1 was used to cross-validate this qualitative classification by quantifying grain proportions in images, confirming alignment with the grain abundance parameters for each class (Figure 4.4). The other 117 tiles were set aside as they belonged to floatstones, rudstones, or boundstones, which lacked sufficient data for accurate identification. Some of these tiles also presented excessive complexity showing heavy diagenetic alterations or poor image qualities.

Dunham's Classification

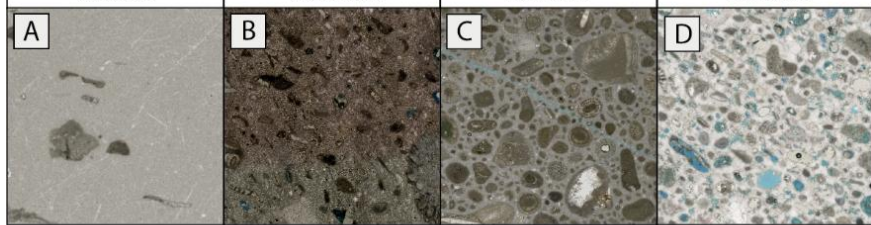
Generated Depositional Texture			
Less than 10% components $\geq 2\text{mm}$ & not bound together organically during deposition			
Mud-supported < 50% grains		Grain-supported $\geq 50\%$ grains	
< 10% grains	$\geq 10\%$ grains	Contains mud	Lacks mud
Mudstone	Wackestone	Packstone	Grainstone
A	B	C	D
			
1 mm			

Figure 4.4 Dunham classification used for model training: synthetic examples A, B, C, and D are generated using the SynSection method.

Then, 150 tiles within the set of classified tiles were selected for grain segmentation purpose and were labeled using an open-source software, Computer Vision Annotation Tool (CVAT) (Sekachev, et al., 2020). Grains that were not intersecting image borders were extracted based on their minimum enclosed area making pairs of thumbnail images and masks organized into dedicated directories per class (Figure 2C). This process resulted in 32,695 extracted objects distributed across 4 categories: 5,669 skeletal grains, 3,228 intraclasts, 6,671 ooids, 17,127 peloids (Figure 4.3). Background cement and matrix were also annotated using bounding boxes, enabling the extraction of 53,917 patches of textures of size 96×96 pixels. The training dataset is impacted by class imbalances among labeled grains. This can be attributed to the distribution of the original data, which is primarily composed of grainstones dominated by ooids and peloids. Additionally, large objects (e.g., shells, algae, bryozoans, etc.) covering a more significant area of the image than small grains are consequently labeled less extensively for a given thin section.

4.3 Methods

A set of around ten thousand classified tiles (15,883) and only a hundred (150) with precise grain annotations respectively constitute small and very small datasets that are difficult to exploit in a DL framework without resorting to data augmentation. Indeed, the 150 tiles annotated for grain segmentation, when scaled back to an entire thin section, would only be equivalent to the annotation of one to two full images. However, the number of labeled individual objects (32,695) and patches of texture background (53,917) is significant (Figure 4.3). Using annotated elements, a complementary data augmentation method called SynSection is then proposed to create a synthetic training dataset for image classification and segmentation purposes. The approach was developed and broken down into multiple steps, specifically designed to mimic and respect the physical constraints of the natural formation process of carbonate rocks (Figure 4.5).

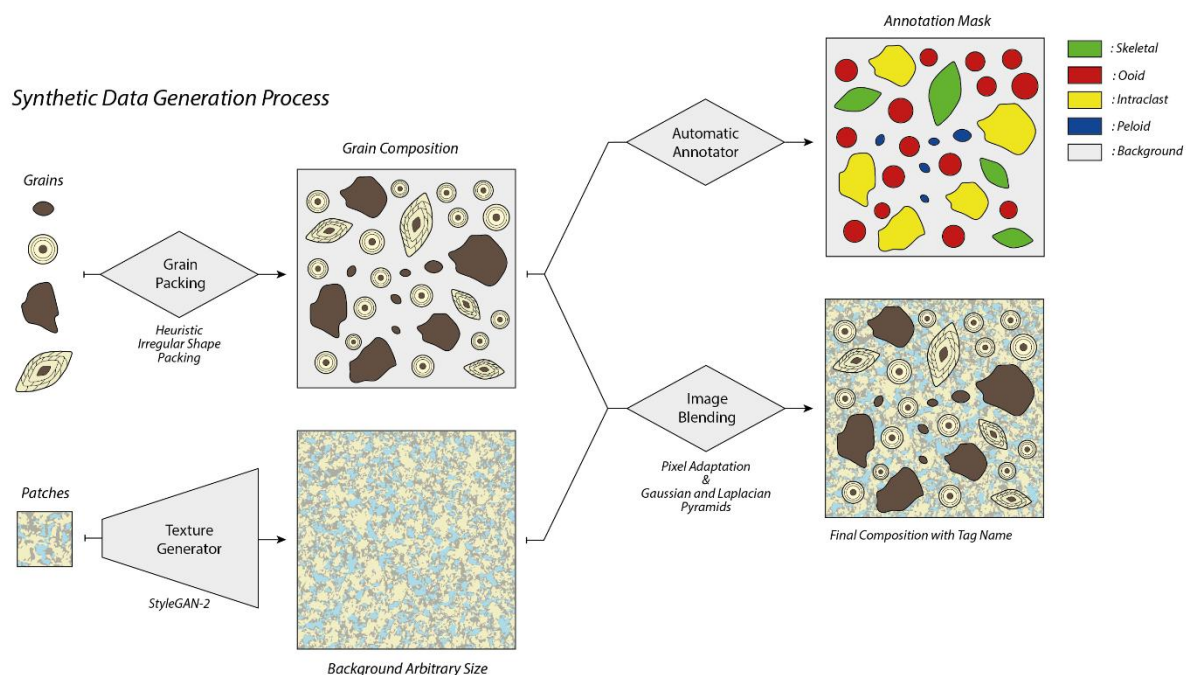


Figure 4.5 Proposed method for synthetic image training dataset generation (SynSection): annotated grains are packed to introduce spatial variability, and diverse textured backgrounds are synthesized using a StyleGAN-2 generator. Backgrounds and grain compositions are blended using pixel distribution adaptation combined with gaussian and laplacian pyramids. Annotation masks are created during the packing process, with labels derived from input generation parameters. The entire workflow is streamlined into a unified processing chain for user convenience.

Process	Description	Constraints	Solution
A. Grain Deposition	Initial deposition of grains, influenced by environmental dynamics.	A.1: 2D non-overlapping grains. A.2: Diverse grain typologies. A.3: Localized spatial relationships. A.4: Nuanced density variations.	Modeled as a distance-based packing problem using heuristics to address constraints A.1-A.4.
B. Cementation	Diagenetic process binding grains via groundmass formation (micritic to calcite).	B.1: Groundmass varies (muddy micritic to coarse crystalline calcite).	Generated using StyleGAN-2 for texture synthesis, producing scalable, diverse groundmass images matching cement/matrix.
C. Dissolution	Circulating fluids dissolve existing grains or groundmass.	Implicitly handled via input data.	Porosity distribution implicitly induced via extracted grains/textures already affected by dissolution in input data.
D. Integration	Combining grains and groundmass into a cohesive synthetic image.	D.1: Seamless blending required. D.2: Packing constraints (A.1-A.4) depend on background texture.	Background texture generated <i>before</i> grain placement. Multi-scale image blending (Gaussian/Laplacian pyramids) used for seamless composition.

Table 4.1 Modeling and synthesis of sedimentary carbonates rock textures through numerical approaches: key processes, constraints, and computational solutions.

4.3.1 Replication of Carbonate Rock Formation Process

Carbonate rock formation initiates with grain deposition (Table 4.1 process A), a process governed by environmental dynamics that determines the three-dimensional spatial configuration and characteristics of accumulated grains (Tucker, 2009). When examined through thin sections, carbonates reveal a two-dimensional arrangement of non-overlapping grains (Table 4.1 constraint A.1), characterized by diverse typologies (Table 4.1 constraint A.2), localized spatial relationships (Table 4.1 constraint A.3), and nuanced density variations (Table 4.1 constraint A.4). Following grain deposition, diagenesis begins with a cementation process (Table 4.1 process B) that binds elements together (Harris et al., 1985). The groundmass ranging from muddy micritic to coarse crystalline calcite (Table 4.1 constraint B.1), fills pore spaces and interconnects grains while enabling fluid circulation through residual porosity. During this process, circulating fluids can potentially dissolve existing structural components, including grains and groundmass (Table 4.1 process C).

The processes described are proposed to be numerically addressed through SynSection in a three-step pipeline (Figure 4.5). The first step involves modeling grain deposition as the placement of one or multiple objects on a restricted surface, analogous to a packing problem categorized under combinatorial optimization (Du et al., 2023; Guo et al., 2022). A distance-based packing method, incorporating heuristics, serves as the foundation for the image generation process (Table 4.1 solution process A). The second step focuses on generating groundmass images of arbitrary size from annotated background cement and matrix. This material, composed of fine crystals at varying scales, aligns with the principles of texture synthesis domain, which aims to create repeating patterns infused with a degree of randomness. Recent studies have successfully employed Generative Adversarial Networks (GANs) (Goodfellow et al., 2014) for texture synthesis (Bergmann et al., 2017; Jetchev et al., 2016). These methods were deemed appropriate for this study due to their ability to generate a diverse and extensive collection of high-quality images. StyleGAN-2 (Karras et al., 2019) was chosen for its advanced and optimized architectural improvements (Table 4.1 solution process B). The final step in the image generation process completes the integration of packed grain images with the generated background textures (Table 4.1 process D), ensuring a seamless result (Table 4.1 constraint D.1). Image composition is an essential aspect for creating realistic visual content and has been extensively studied (Prados & Garcia, 2014). This integration is achieved through a conventional approach that combines statistical pixel distribution adaptation with Gaussian and Laplacian pyramid-based image blending (Burt & Adelson, 1983a; Burt & Adelson, 1983b) (Table 4.1 solution process D). For consistency, the background texture is generated and selected before object placement, since the packing constraints vary depending on the specific rock texture being created (Table 4.1 constraint D.2). Since the extracted grains and patches of background texture are already affected by dissolution (Table 4.1 process C), the distribution of the generated porosity is implicitly induced in all the previous detailed generation steps.

4.3.2 2D Object Packing

Object packing is a challenging problem encountered in various application domains. It involves arranging a set of items within a confined area to maximize space utilization. Formally, the problem can be defined as given a surface S with dimensions W (width) and H (height) and a set of objects $\{O_1, \dots, O_n\} \in S$ with respective areas $\{A_1, \dots, A_n\}$. Each object O_i can be rotated by an operator $R_{\theta_i}(O_i)$, with $\theta_i \in [\theta_{min}, \theta_{max}]$ a range of allowed rotation angles. The objective is to minimize function F as defined in Equation 4.1 under the constraints in Equation 4.2.

$$\min (F = W \cdot H - \sum_{i=1}^n A_i) \quad (4.1)$$

$$R_{\theta_i}(O_i) \cap R_{\theta_j}(O_j) = \emptyset, \quad \forall i \neq j \quad (4.2)$$

Several approaches to particle placement have been developed, primarily using heuristic and dynamic methods. These techniques rely on experience and optimization strategies to find near-optimal solutions. A common strategy involves simplifying complex shapes into basic geometric forms like circles or rectangles (Hinrichsen et al., 1990; Jylänki, 2010; Wang et al., 2006). However, this approach can be problematic as they generally produce unrealistic compositions with a well-organized structure. Alternatively, geometrical methods, particularly those based on constructive approaches, have gained attention but seem to present similar limits (Zhao et al., 2023).

Particle packing can be achieved using a weighted Voronoi tessellation to generate particle convexity. This method involves creating polygons with specific size and orientation distributions by iteratively adjusting seed positions and radii. The packing process then involves fitting these particles using a cubic-polynomial curve (Wang et al., 2019). While this approach demonstrates innovative potential, it is currently limited to convex objects and reveals several computational complexity and implementation challenges.

4.3.2.1 2D Grain Packing with Sedimentological Heuristics

In this study, the proposed method aims to pack grains with irregular shapes by iteratively positioning them on a blank composition. A distance map is computed at each iteration to identify available locations for subsequent element placement, guided by geological heuristics that are subsequently developed (Table 4.1 constraints A.1 to A.4).

Configured Grain Compositions

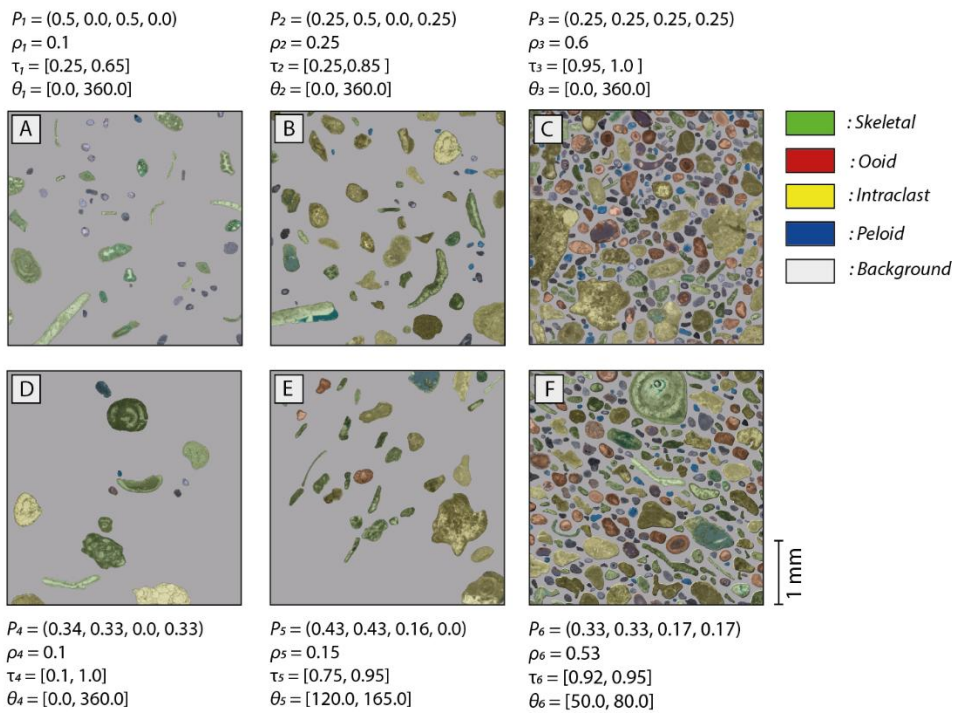


Figure 4.6 Grain compositions generated with controlled sedimentological packing constraints: Mimicking geologically plausible arrangements by regulating class distribution (P), grain density (ρ), local proximity (τ), and orientation (θ).

Grain Packing Evaluation over Time

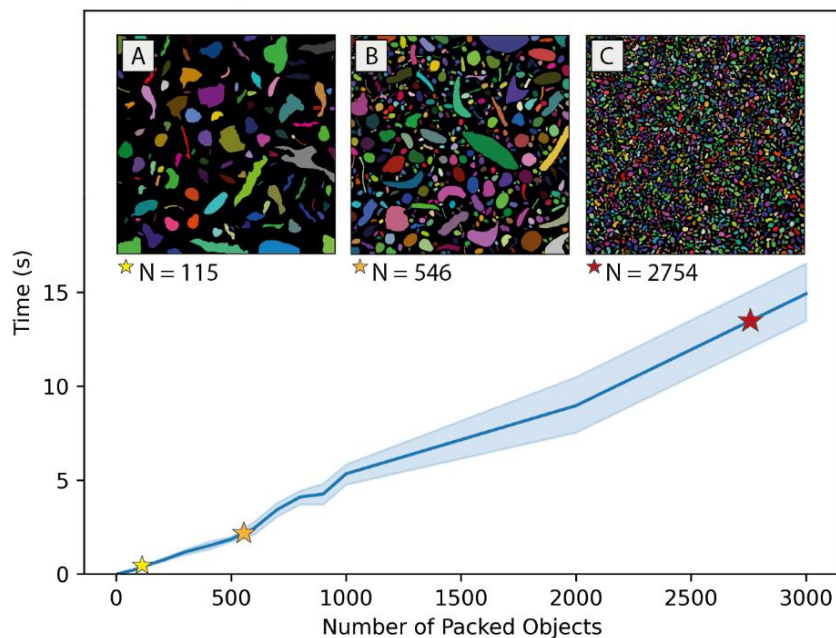


Figure 4.7 Efficiency of grain packing computation time across 5 realizations per step: Ranges include [0–100] with a step size of 10, [100–1,000] with a step size of 100, and two additional steps at 2,000 and 3,000 packed objects. Examples A, B, and C depict instance mask representations of a composite instance mask at different steps, containing 115, 546, and 2,754 packed objects, respectively.

To be detected using a DL model, each grain should be observed in all potential contexts of observation. Thus, an initial consideration is managing the distribution of grain classes within the image composition. Classes distribution may have affinities with specific depositional environments and help in defining sedimentary facies. Consequently, controlling sampling during the packing process becomes essential by introducing a parameter P that determines the probability of sampling each class. As demonstrated in Figure 4.6A, the composition is predominantly comprised of skeletal and peloidal grains while Figure 4.6C shows equivalent class distribution. The global grain density, defined as the ratio of grain-occupied area to total image area, is a critical parameter for characterizing rock texture. This density is controlled by a parameter ρ that determines the termination point of the packing process when a specified limit is reached. Figure 4.6 also illustrates various grain density values: $\rho_{1 \rightarrow 6} = \{0.10, 0.25, 0.60, 0.10, 0.15, 0.53\}$. During deposition, grains can be positioned at varying distances. A local proximity operator, defined within range $\tau \in [0, 1]$ is then used to control their spacing. As the value approaches 1, objects become closer, and conversely. Figures 4.6B and 4.6D illustrate this parameter's impact, showing a more homogeneous grain spatial distribution in Figure 4.6B compared to Figure 4.6D. Marine currents (e.g., tidal currents) influence grain arrangement on the seabed, creating directional markers through laminated structures (laminations). Thus, grains may respond to specific orientation constraints applied via a rotational operator, as evidenced in Figure 4.6E and Figure 4.6F. A final feature prevents object placement within the background texture's porosity, simulating a visible intergranular porosity. The process and implementation are complementary detailed in Algorithm 4.1.

Algorithm 1: 2D Grain Packing with Sedimentological Heuristics

```

Inputs :  $B \in [W, H]$  /* Background texture of size (W, H) */
            $O = \{o_1, \dots, o_n\}$  /* X objects belonging to N classes */
            $P(C) = \{P(c_1), \dots, P(c_k)\} \mid k = N$  /* Probabilities per class */
            $\rho, \tau \in [0, 1]$  /* Object density and local proximity */
            $\theta \in [\theta_{min}, \theta_{max}]$  /* Rotation angle range */

Outputs:  $ObjectComp$  /* Object composition image */
            $O_{packed} \subseteq O$  /* Packed objects */
            $Coords = \{(x_1, y_1), \dots, (x_k, y_k)\}$  /* Object coordinates */

Select a background  $B$ 
 $ObjectComp \leftarrow \text{Binarize}(B)$  /* Apply porosity thresholding */
 $O_{packed} \leftarrow \emptyset, Coords \leftarrow \emptyset$ 
while true do
    Select an object  $o_i$  according to  $P(C)$ 
     $o_i \leftarrow R_{\theta_i}(o_i)$  /* Apply rotation */
     $r \leftarrow \text{radius}(o_i) \times \frac{1}{\tau}$  /* Dilate radius */
     $d \leftarrow \text{EuclideanDistanceTransform}(ObjectComp)$ 
     $AvailableCoords \leftarrow \{(x, y) \in [W, H] \mid d(x, y) - r \geq 0\}$ 
    if  $AvailableCoords \neq \emptyset$  then
        Select a random  $(x, y)$  from  $AvailableCoords$ 
         $ObjectComp(x, y) \leftarrow o_i$ 
         $O_{packed} \leftarrow O_{packed} \cup \{o_i\}$ 
         $Coords \leftarrow Coords \cup \{(x, y)\}$ 
        if  $\frac{1}{WH} \sum_{n=1}^N (ObjectComp - \text{Binarize}(B)) \geq \rho$  /* Density limit */
            then
                break
            else
                end
        else
            break
        end
    end
end
return  $ObjectComp, O_{packed}, Coords$ 

```

Algorithm 4.1 Grain packing algorithm with integrated sedimentological priors.

4.3.2.2 Packing Performance Evaluation

The packing performance was evaluated by tracking the number of placed objects over time (Figure 4.7), which revealed a sub-linear packing rate of 214 objects per second, that proved efficient enough to produce a large volume of compositions (100,000) within a reasonable timeframe.

4.3.3 Texture Synthesis

Traditional methods for generating textures can be categorized into two primary categories. Parametric techniques rely on matching statistical properties of images, such as histograms derived from a steerable decomposition (Heeger & Bergen, 1995) or various sub-band decompositions to synthesize a diverse range of textures (Portilla & Simoncelli, 1999). These methods fundamentally focus on solving an optimization problem within the space of image pixels. Non-parametric techniques constitute the second category, which diverge significantly from parametric approaches (Efros & Leung, 1999; Paget & Longsta, 1995). These methods resample pixels and patches from example textures, strategically randomizing these inputs while rigorously preserving their perceptual properties. While these techniques can produce high-quality images, they essentially reorder the input texture using local similarity.

More recently, the advent of DL contributed to the development of texture synthesis across a wide variety of natural images using a pre-trained VGG-19 network (Simonyan & Zisserman, 2015) to extract features from a specified texture. Subsequently, starting from a noised input image, the network synthesizes another texture that aligns with the Gram-matrix representation of the original texture (Gatys et al., 2015b). Despite yielding a powerful technique that adeptly captures expressive image statistics, the creation of a single output texture necessitates solving an optimization problem through iterative backpropagation, involving both time and memory costs. The introduction of GANs (Goodfellow et al., 2014) marked a significant breakthrough in image generation techniques. These innovative dual-network models consist of a generator that synthesizes images from random noise and a discriminator network that evaluates whether the generated images match real data distribution. The application of GANs in texture synthesis was initially introduced with the Spatial-GAN (Jetchev et al., 2016) for creating textures of arbitrary size. A notable advancement in the generator architecture came with the StyleGAN (Karras et al., 2018) which introduced an intermediate latent space to adapt the style at each convolution layer using a statistical alignment method. This architecture also incorporates explicit noise injections, effectively separating global features from local stochastic variations. Building upon this foundation, StyleGAN-2 (Karras et al., 2019) addressed the visual artifacts present in the original StyleGAN by redesigning normalization techniques and eliminating progressive training. This refined architecture has already demonstrated its potential for generating petrographic images (Ferreira et al., 2022), highlighting its significance for this study.

4.3.3.1 Groundmass Generation using StyleGAN-2

A lightweight variant of StyleGAN-2 (Figure 4.8) was implemented with a generator comprising 5 convolution blocks paired with 5 learnable input style vectors. The discriminator features a symmetric structure with residual blocks, concluding with a single-unit fully connected layer after the final convolution block. The Wasserstein loss function was employed to address two critical challenges: improving training stability and getting meaningful metric that correlates with perceptual image quality (Arjovsky et al., 2017). Under this framework, the discriminator is conceptualized as a critical model that assigns scalar scores to images, measuring how closely they match the statistical properties of the training examples, rather than outputting probabilities in the classical GAN sense. Two complementary regularization techniques were also implemented: a gradient penalty applied to the discriminator (Gulrajani et al., 2017) and a path length regularization term imposed on the generator (Karras et al., 2019). The gradient penalty ensures the discriminator maintains Lipschitz continuity, while the path length regularization promotes smoother generator mapping between the latent space and generated images. The model was trained for 100 epochs using the Adam optimizer (Kingma & Ba, 2014) with an exponential learning rate decay schedule going from 1×10^{-4} to 1×10^{-5} . To generate larger textures during inference, the fully convolutional nature of the generator was exploited by increasing the spatial dimensions of the input noise. This straightforward approach allows the network to maintain its learned patterns while producing outputs of arbitrary sizes.

Texture Synthesis Architecture

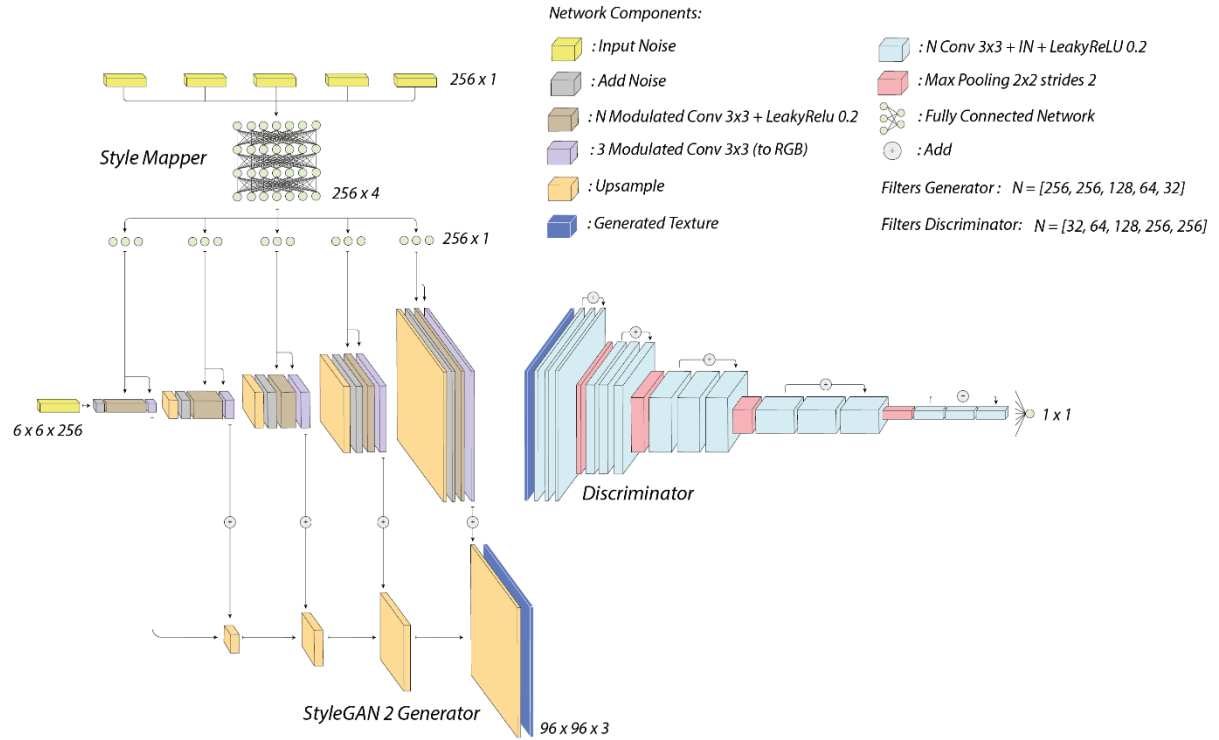


Figure 4.8 StyleGAN 2 model architecture used for background texture synthesis.

4.3.3.2 Model Evaluation

A common metric for evaluating the image quality generated by GANs is the Frechet Inception Distance (FID) (Heusel et al., 2017). This score quantifies the similarity between real and synthetic images using a pre-trained Inception-v3 model (Szegedy et al., 2015). A lower FID indicates better image generation and greater diversity. For petrographic applications, this score has already been measured at 12.49 (Ferreira et al., 2022), with values below 10 representing very good image quality. In this study, the FID was calculated at various training stages, achieving an optimal score of 13.52. To further assess the model's textural variability during inference, Haralick features distribution were compared between real and generated images (Haralick, 1973). Haralick features are derived from the Gray Level Co-occurrence Matrix (GLCM), which records the frequency of adjacent gray-level pixels in an image. From this matrix, 13 texture-quantifying values can be extracted. These features were calculated for 50,000 real and synthetic images. As illustrated in Figure 4.9, the model generally captured and reproduced the training data distribution during inference. However, slight variations (around a quartile) were observed in specific features: Angular Second Moment, Contrast, Sum of Squares, Inverse Difference Moment, Sum Average, Sum Variance, Sum Entropy, Entropy, Difference Variance, Difference Entropy, Measure of Correlation f12, and Measure of Correlation f13. Some examples of generated large textures are presented in Figure 4.10.

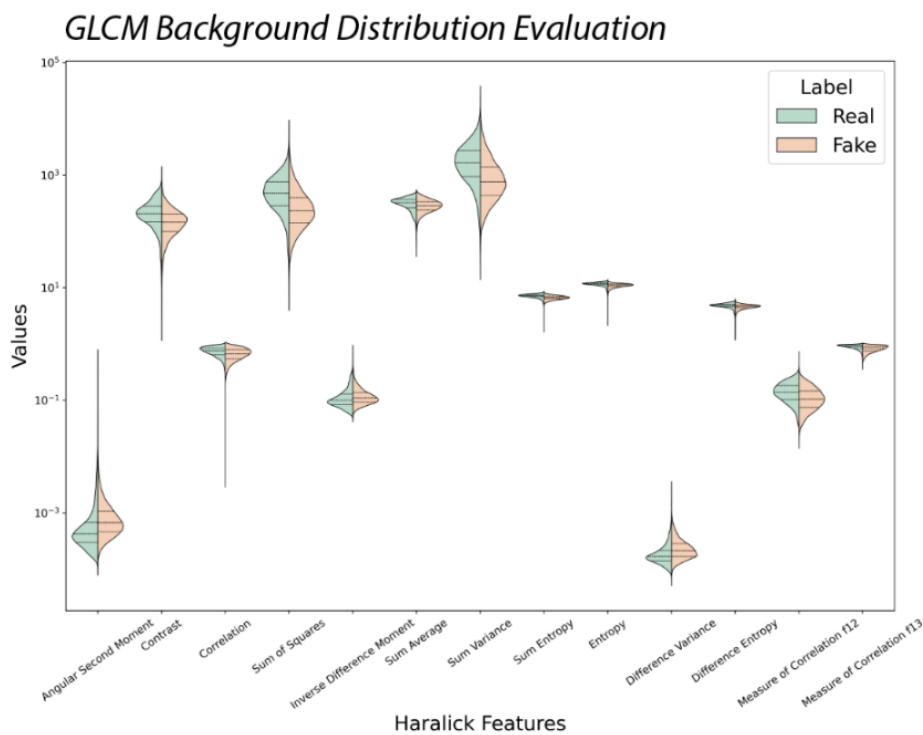


Figure 4.9 Comparison of Haralick feature distributions between generated and real textures.

Generated Large Textures

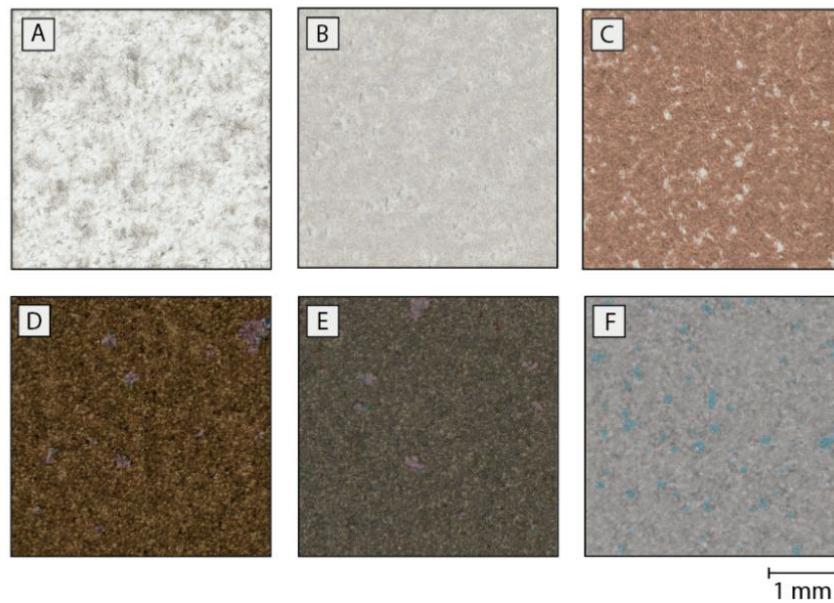


Figure 4.10 Large textured backgrounds generated using the StyleGAN2 generator (512×512 pixels): examples A, B, and C represent spar cement textures, while D, E, and F depict micritic matrixes. These textures were classified using a KNN algorithm.

4.3.4 Image Blending

In computer vision, image blending is a technique that seamlessly integrates multiple images or layers to create visually compelling composites. This process goes beyond simple merging, emphasizing smooth transitions while preserving critical visual details to maintain realistic representations.

Traditional approaches like alpha blending utilize a weighted pixel contribution, where each pixel's intensity is modulated by an alpha value. Although efficient and widely adopted for its simplicity, this technique can struggle with complex images featuring intricate textures or sharp edges (Figure 4.11A). To address these limitations, advanced methods have emerged. Notably, Poisson image editing (Pérez et al., 2003) represents a significant advancement and reimagines blending as a partial differential equation problem. By enforcing gradient continuity, this method achieves seamless merging of images. Despite its impressive capabilities, the technique can generate undesirable ghost effects when applying a precisely defined mask around source and target images, and when the pixel distributions of the images demonstrate significant discontinuity.

Emerging DL methods have been explored for image blending. However, these approaches may require paired training data, including source and reference images alongside their corresponding well-blended counterparts (Wu et al., 2019). Another proposed method involves optimizing a Poisson blending loss, along with style and content loss, and iteratively updating pixels using a solver to reconstruct the blending region (Zhang et al., 2019). While this approach demonstrates promising potential, the computational complexity of pixel-by-pixel optimization renders it impractical for generalized application, as each image sample necessitates individual, resource-intensive processing.

4.3.4.1 Grain composition and groundmass fusion

In this study the image blending process has been addressed using a two-step approach. First, a pixel distribution adaptation was performed between source objects (grains) and reference textured background (cement or matrix) to homogenize their visual aspects. This process involves transforming an original background B and a set of objects $\{O_1, \dots, O_n\}$ into a standard distribution space using a selected transform T (Equation 4.3). Subsequently, the object normalized images are individually mapped into the background image's distribution space by applying the inverse of the transformation fitted on the reference image (Equation 4). This approach improves grain integration into the background texture, as evidenced by the visual differences between Figure 4.11B and Figure 4.11C, and the similarity with a real example in Figure 4.11D.

$$T_X(x) = \frac{x - \mu_X}{\sigma_X} \quad (4.3)$$

$$O_{i_{adapted}} = T_B^{-1} \circ T_{O_i}(O_i) \quad (4.4)$$

In the second step, a Gaussian and Laplacian pyramid blending was employed to create the seamless composition (Burt & Adelson, 1983a; Burt & Adelson, 1983b). This method involves constructing K multi-scale Gaussian and Laplacian pyramids representation for both the grain composition and the background image, followed by alpha blending across a selected mask at each level of its Gaussian pyramid decomposition. The final output is then synthesized by summing the blended results from the different scales, ensuring a smooth and natural integration of the grain composition into the background image (Algorithm 4.2).

Image Composition

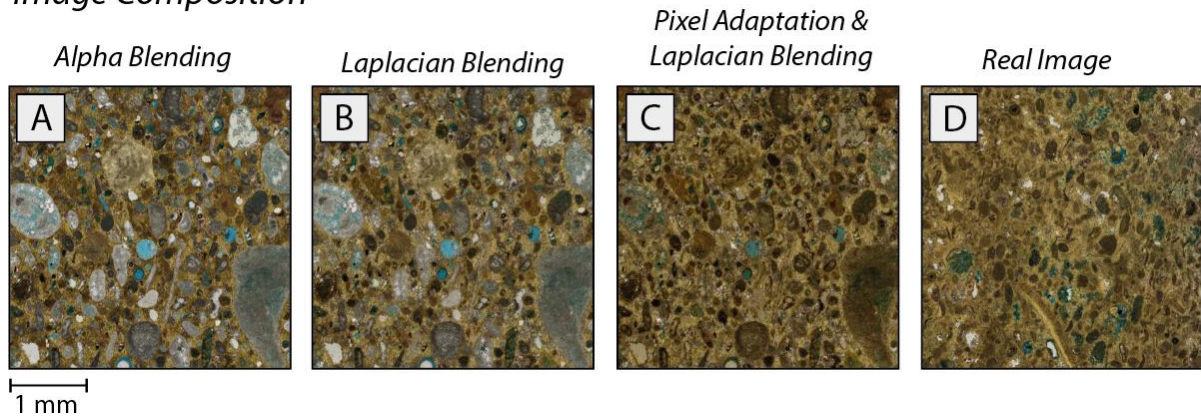


Figure 4.11 Comparison of image composition methods: alpha blending, laplacian blending, and pixel distribution adaptation combined with laplacian blending.

Algorithm 2: Gaussian and Laplacian Pyramid Blending

```

Inputs :  $ObjectComp_A \in [W, H]$  /* Object composition with adaptation */
            $B \in [W, H]$  /* Background texture */
            $M \in [W, H]$  /* Binary object composition mask */
            $K = \max\{K \in \mathbb{N}^+ \mid 2^K \mid \min(W, H)\}$  /* Pyramid levels */

Outputs:  $I_{blend}$  /* Final blended image */

Step 1: Construct Gaussian Pyramids
 $G_{ObjectComp_A}[0] \leftarrow ObjectComp_A$ ,  $G_B[0] \leftarrow B$ ,  $G_M[0] \leftarrow M$ 
for  $k = 1$  to  $K$  do
     $G_{ObjectComp_A}[k] \leftarrow \text{Blur}(\text{Downscale}(G_{ObjectComp_A}[k-1]))$ 
     $G_B[k] \leftarrow \text{Blur}(\text{Downscale}(G_B[k-1]))$ 
     $G_M[k] \leftarrow \text{Blur}(\text{Downscale}(G_M[k-1]))$ 

Step 2: Construct Laplacian Pyramids
for  $k = 0$  to  $K - 1$  do
     $L_{ObjectComp_A}[k] \leftarrow G_{ObjectComp_A}[k] - \text{Upscale}(G_{ObjectComp_A}[k+1])$ 
     $L_B[k] \leftarrow G_B[k] - \text{Upscale}(G_B[k+1])$ 
 $L_{ObjectComp_A}[K] \leftarrow G_{ObjectComp_A}[K]$ ,  $L_B[K] \leftarrow G_B[K]$  /* Last pyramid levels */

Step 3: Blend Laplacian Pyramids
for  $k = 0$  to  $K$  do
     $L_{blend}[k] \leftarrow G_M[k] \cdot L_{ObjectComp_A}[k] + (1 - G_M[k]) \cdot L_B[k]$ 

Step 4: Reconstruct the Final Image
 $I_{blend} \leftarrow L_{blend}[K]$ 
for  $k = K - 1$  down to  $0$  do
     $I_{blend} \leftarrow \text{Upscale}(I_{blend}) + L_{blend}[k]$ 
return  $I_{blend}$ 

```

Algorithm 4.2 Image blending using gaussian and laplacian pyramids techniques.

4.3.5 SynSection Dataset Construction

To carefully attribute corresponding cement in grainstone and matrix in mudstone, wackestone or packstone textures, 100,000 large background images with a size of 512×512 pixels were generated and subsequently classified using a K-Nearest Neighbors (KNN) algorithm with 5 neighbors (Guo et al., 2003). This proximity-based classifier was trained on colorimetric attributes (average hue, saturation, and value) and Haralick features extracted from a representative subset of the generated textures. The inclusion of a color indicator proved valuable, particularly since some original patches were stained with alizarine ferricyanide as illustrated on Figure 4.10C to Figure 4.10E. Illustrations from Figure 4.10A to Figure 4.10C highlight examples of generated and classified cement, while Figure 4.10D to Figure 4.10F display matrix textures.

Once the groundmass images were classified, 25,000 synthetic pairs of synthetic images and annotation masks were generated for each carbonate texture (mudstone, wackestone, packstone, and grainstone) to ensure data balancing (Figure 4.4). For each rock texture, heuristic parameters were randomly assigned within their defined ranges using a uniform distribution. This approach introduced high compositional variability while ensuring a balanced distribution for further image classification and segmentation tasks.

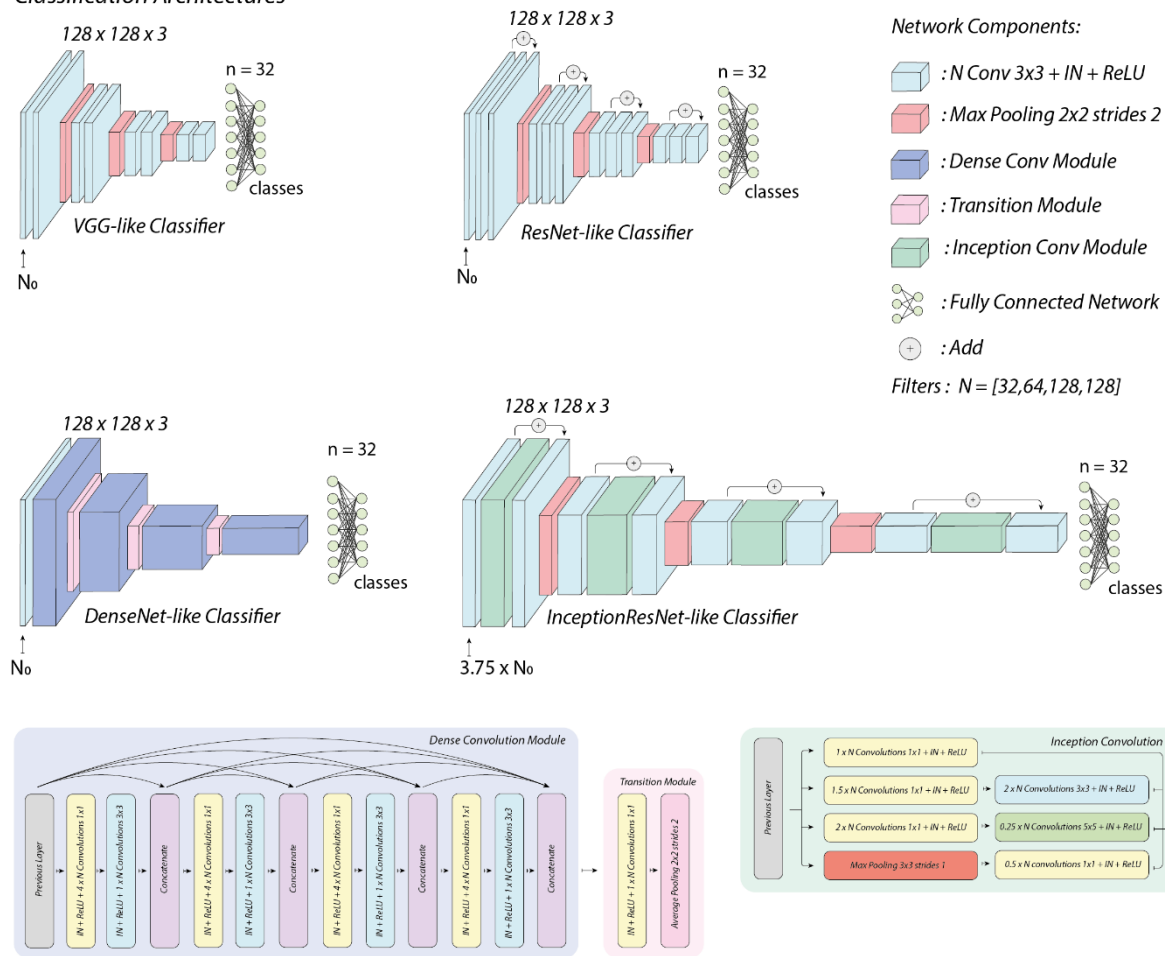
4.3.6 Image Classification

Image classification has been a major challenge in computer vision for decades. Early approaches focused on simple template matching and geometric methods, then machine learning methods became dominant, particularly with the use Support Vector Machines (SVM) combined with hand-crafted features (Evgeniou & Pontil, 2001). However, these approaches were limited by the need for manual feature engineering.

The arrival of data-driven models like DCNNs has opened a way to solve more and more complex image classification problems by increasing their depths and the size of the training datasets. In the context of rock texture classification, supervised models can be trained on datasets sorted by multiple individuals with varying levels of expertise. This inevitably introduces biases during the learning process. Moreover, the classification rules used for carbonate rock textures are based on quantitative parameters, which are challenging to visually assess when examples fall at the boundary between two classes or multiple classes. The image generation method presented in this study enables precise control over the final rock texture composition by incorporating the key parameters of Dunham's classification system. By adjusting variables such as groundmass type, grain density, grain type and spatial distribution, this method can produce a broader spectrum of rock textures. This comprehensive approach facilitates a deeper understanding of how different carbonate rock classifications transition between their endmember states.

4.3.6.1 Model Architectures and Training

Four lightweight models reusing essential modules proposed in the following architectures: VGG; Residual (ResNet); Dense (DenseNet) and Inception-Residual (InceptionResNet), have been implemented following a sequence of 4 convolutional blocks separated by max-pooling layers (He et al., 2015b; G. Huang et al., 2017; Simonyan & Zisserman, 2014; Szegedy et al., 2014). The classification part is made of a fully connected network with successively 32 units with ReLU activation and 4 units with a softmax activation corresponding to the classes of interest. The different architectures are detailed in Figure 4.12.



The training protocol followed a two-stage process, with results averaged across 5 cross-validation cycles (Kohavi, 1995). In the initial training phase, models were trained on 3,125 batches, each containing 32 synthetic images (100,000 training examples). This training involved 10 iterations with standard data augmentation (rotations, flips, slight brightness and contrast shifts) and dynamic image resizing to 128×128 pixels. During this stage, Adam optimizer was selected with an exponential decay learning rate strategy going from 1×10^{-4} to 0.5×10^{-4} . Following initial training, the models underwent fine-tuning using 338 batches of 32 real images (10,816 fine-tuning examples) over 10 additional iterations with same standard augmentations. During this phase, the convolutional layers were frozen, allowing only the classification layers to be adjusted, with a fixed learning rate of 1×10^{-5} . The training employed a categorical cross-entropy loss function with applied class weights to address data imbalances in the original dataset (Figure 4.3). The different models were then evaluated on a set of blind test images comprising 5,067 examples.

4.3.6.2 Model Evaluation

The performance of each network architecture was evaluated through continuous monitoring of loss variations and a comprehensive set of metrics. The evaluation framework incorporated standard classification metrics including accuracy (Equation 4.5), precision (Equation 4.6), recall (Equation 4.7), and F1 score (Equation 4.8), all computed from the underlying true positives (TP), true negatives (TN), false positives (FP), and false negatives (FN). To provide additional information into classification behavior, confusion matrices were analyzed during inference on the test dataset (Fawcett, 2006). The evaluation also included measuring how long image processing took during inference to check model's efficiency.

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \quad (4.5)$$

$$Precision = \frac{TP}{TP + FP} \quad (4.6)$$

$$Recall = \frac{TP}{TP + FN} \quad (4.7)$$

$$F1 = 2 \times \frac{Precision \times Recall}{Precision + Recall} \quad (4.8)$$

Finally, Grad-CAM was employed to gain further insight into the model's decision-making process (Selvaraju et al., 2016). This explainability technique, computes the gradient of the predicted class score with respect to the feature maps of the final convolutional layer, revealing which regions most strongly participate to the classification decision.

4.3.7 Image Segmentation

A traditional approach for image segmentation is thresholding, where pixels are classified based on intensity values. While simple, this approach is effective for certain applications, especially when the objects of interest have distinct intensity characteristics. However, it often falls short when dealing with complex images with overlapping intensity distributions. Region-based segmentation methods were introduced to aggregate pixels sharing common features. The region-growing algorithm exemplifies this approach by expanding from seed points to incorporate neighboring pixels that meet defined criteria (Asmussen et al., 2015). While computationally efficient, this method's effectiveness depends heavily on initial seed point selection and growth criteria. Then, edge-based techniques focus on boundary detection between objects, with the Canny edge detector (Canny, 1986) being a notable example that utilizes gradient information. However, these methods often encounter difficulties with noisy images and struggle to differentiate between adjacent objects displaying similar intensity gradients, a limitation particularly evident in this dataset.

Growing in prominence, DCNNs have emerged as powerful tools for image segmentation. The U-Net architecture (Ronneberger et al., 2015) is a major advancement through its distinctive design: a contracting path captures contextual information, while a symmetric expanding path enables precise localization. The architecture's effectiveness stems from its integration of skip

connections between corresponding downsampling and upsampling layers, preserving both contextual and fine-grained spatial information.

4.3.7.1 Model Architectures and Training

Four custom U-Net-like architectures were developed, incorporating transfer learning strategies like those applied in the classification stage. The segmentation dataset was limited to 150 annotated images (Figure 4.3), significantly fewer than those available for classification.

Two distinct approaches were implemented to assess task feasibility. The first approach employed binary segmentation, treating all grain types as a single class to facilitate subsequent grain quantification analysis. The second approach incorporated grain class information to generate both spatial distribution and relative proportions of different elements within thin sections.

The implemented models follow similar architectures to U-Net, Residual U-Net (ResU-Net), Residual with Attention Gate U-Net (AttentionResU-Net), Inception-Residual U-Net (InceptionResU-Net) (Oktay et al., 2018; Punj & Agarwal, 2020; Ronneberger et al., 2015; Zhang et al., 2017). Each model features symmetrical convolution blocks in both encoder and decoder sections, with max-pooling layers for downsampling in the encoder and upsampled convolutions in the decoder. The final convolutional layer is task-specific: binary segmentation utilizes a single filter with sigmoid activation, while multiclass segmentation employs six filters with softmax activation to differentiate between the defined grain types and background. Detailed architectural representations are provided in Figure 4.13.

Segmentation Architectures

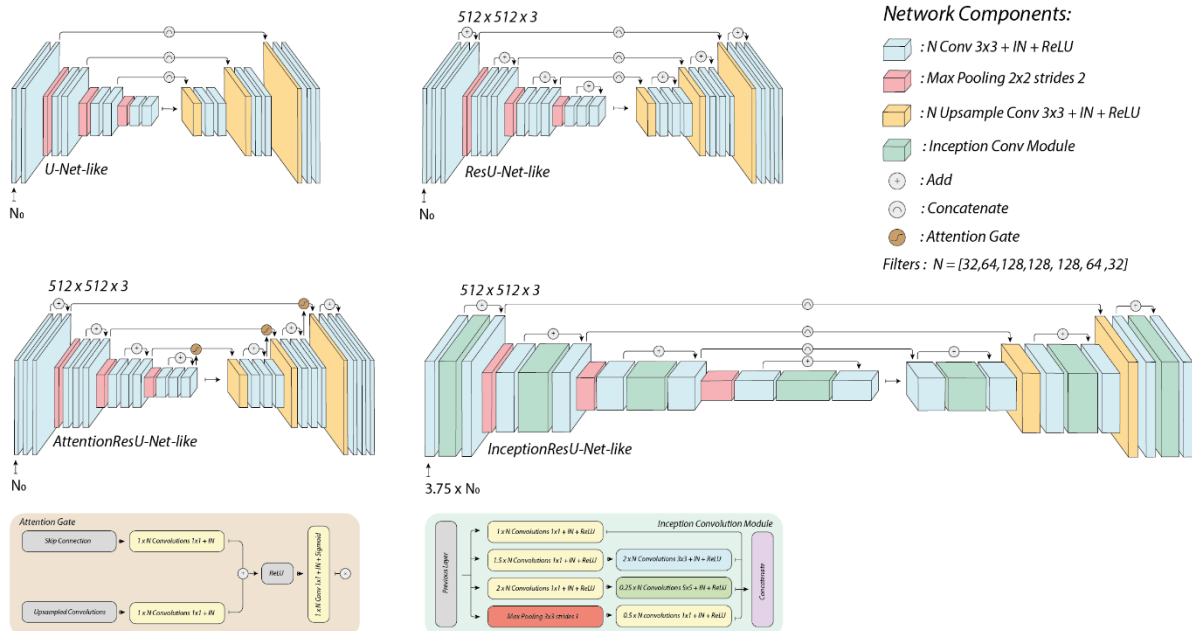


Figure 4.13 Implemented image segmentation model architectures used for grain identification.

Initially, each model was trained on 12,500 batches of 8 synthetic images (100,000 training examples) over 10 iteration loops with the same augmentation parameters used during classification. Subsequently, the models underwent fine-tuning using a validation set of 13 batches containing 8 real images (104 fine-tuning examples), processed through 100 additional iteration loops with augmentations. A separate set of 46 images was preserved exclusively for testing purposes.

The study acknowledged a fundamental challenge in image segmentation: accurately distinguishing adjacent objects within compact and dense visual compositions. To address this limitation, a penalty map that strategically assigned substantial weight to background pixels separating touching objects has been applied to the cross-entropy loss as described in the original U-Net paper (Ronneberger et al., 2015). This approach was augmented by combining this term with an additional Dice loss (Sudre et al., 2017), creating a hybrid loss mechanism that simultaneously optimizes pixel-level accuracy and overall segmentation quality. The refined loss enables more precise isolation of individual grains in close spatial proximity, enhancing the model's ability to discern subtle boundaries between nearly overlapping elements. Adam optimizer was selected for the training process, incorporating an exponential learning rate decay that gradually reduced from 1×10^{-4} to 0.5×10^{-4} . Following the initial training phase, all convolutional layers were unfrozen during fine-tuning, utilizing a fixed learning rate of 1×10^{-5} .

To assess the benefits of incorporating synthetic data as a complementary augmentation, separate models were trained from scratch exclusively on a set of real images (same 104 fine-tuning examples) for 250 iterative loops, enabling a comparative analysis of performance metrics on a blind test set (same 46 examples).

4.3.7.2 Model Evaluation

Each network architecture's performance was assessed through comprehensive loss monitoring and a targeted set of performance metrics: accuracy, Intersection over Union (IoU) (Equation 4.9), Dice Coefficient (DC) (Equation 4.10), and point counting determination score (R^2) (Equation 4.11).

$$IoU = \frac{\text{Area of Overlap}}{\text{Area of Union}} \quad (4.9)$$

$$\text{Dice Coefficient} = \frac{2 \times \text{Area of Overlap}}{\text{Total Area}} \quad (4.10)$$

$$\text{Point Counting } R^2 = 1 - \frac{\sum_{i=1}^n (y_i - \hat{y}_i)^2}{\sum_{i=1}^n (y_i - \bar{y})^2} \quad (4.11)$$

Accuracy provides an overall segmentation correctness measure, while IoU and DC offer more nuanced assessments of class imbalance and prediction-ground truth overlap. The point counting determination score evaluates the model's ability to identify individual objects through a two-step process: thresholding probability masks using Otsu's method (Otsu, 1979) and applying a connected components algorithm (He et al., 2017) to isolate individual object.

This approach enabled a thorough performance assessment of each architecture for first-scale order petrographic analysis.

4.4 Results

The different trainings were conducted until the loss plateaued. Performance was evaluated on the real-image test set, with additional analysis performed on a subset of challenging cases that had both complexities and available grain abundancy quantitative parameters.

4.4.1 Image Classification

In all training scenarios, the ResNet model consistently demonstrated superior performance in both conventional training and synthetic data-augmented approaches where it respectively achieved average scores of 87.56% and 95.05% across all metrics (Table 4.2 and Figure 4.14). Models trained with synthetic images before being fine-tuned on real images all showed improved results, with a global average increase of +7.23%. InceptionResNet model emerged as another strong performer, delivering a F1 score of 86.75% with traditional training and 94.92% through transfer learning which is generally a reference metric for unbalanced datasets. However, this model presents the disadvantage of being four times heavier in memory than ResNet, demonstrating that effective rock texture classification can be successfully accomplished with lightweight architectural designs. While the VGG and DenseNet models produced comparatively lower results, they remained acceptable for the study's objectives. All models exhibited fast inference capabilities, consistently generating more than 15 predictions per second.

Model	Image Input Size	Depth	Parameters	Size
VGG	128 × 128 × 3	9	2,679,652	6,425 MB
ResNet	128 × 128 × 3	13	2,706,724	6,553 MB
DenseNet	128 × 128 × 3	14	6,841,332	26,905 MB
InceptionResNet	128 × 128 × 3	33	6,933,636	27,291 MB

Training Real Images Only: Metrics on Test Data					
Model	Accuracy	Precision	Recall	F1	Images/s
VGG	79.92	84.24	75.36	79.38	17.61
ResNet	86.19	89.92	87.90	86.22	17.44
DenseNet	84.55	88.19	80.86	84.27	16.52
InceptionResNet	89.61	90.13	83.74	86.75	16.31

Training Synthetic and Fine-Tuning Real Images: Metrics on Test Data					
Model	Accuracy	Precision	Recall	F1	Images/s
VGG	86.61	90.13	83.74	86.75	—"
ResNet	94.86	96.54	93.93	94.85	—"
DenseNet	91.33	92.59	90.47	91.50	—"
InceptionResNet	94.88	95.61	94.17	94.92	—"

Table 4.2 Classification model configurations and performance metrics: comparison of results between training on a small real image dataset and transfer learning from synthetics to real images.

Contribution of Synthetic Data Integration for Image Classification

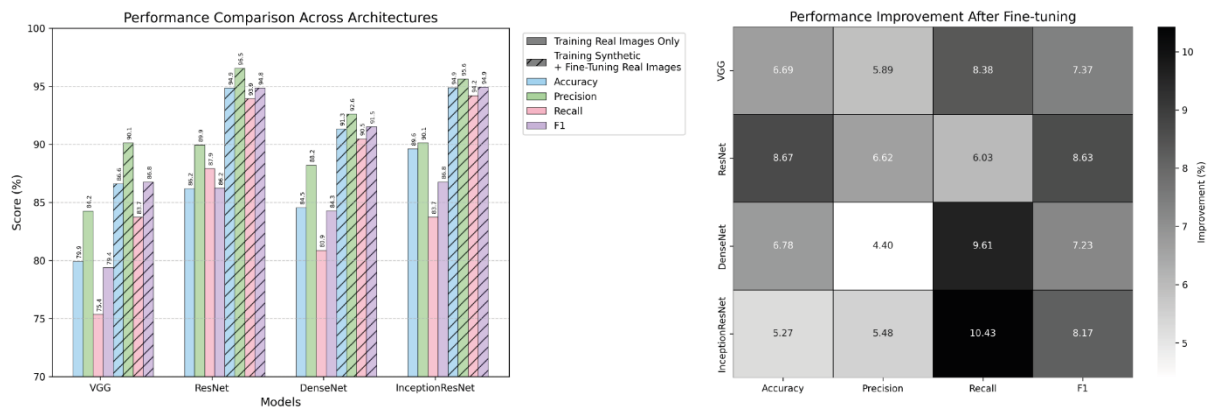


Figure 4.14 Models performance metrics comparison for each training strategy: highlighting the impact of synthetic data integration in the image classification training procedure through performance improvement monitoring.

Image Classification Results

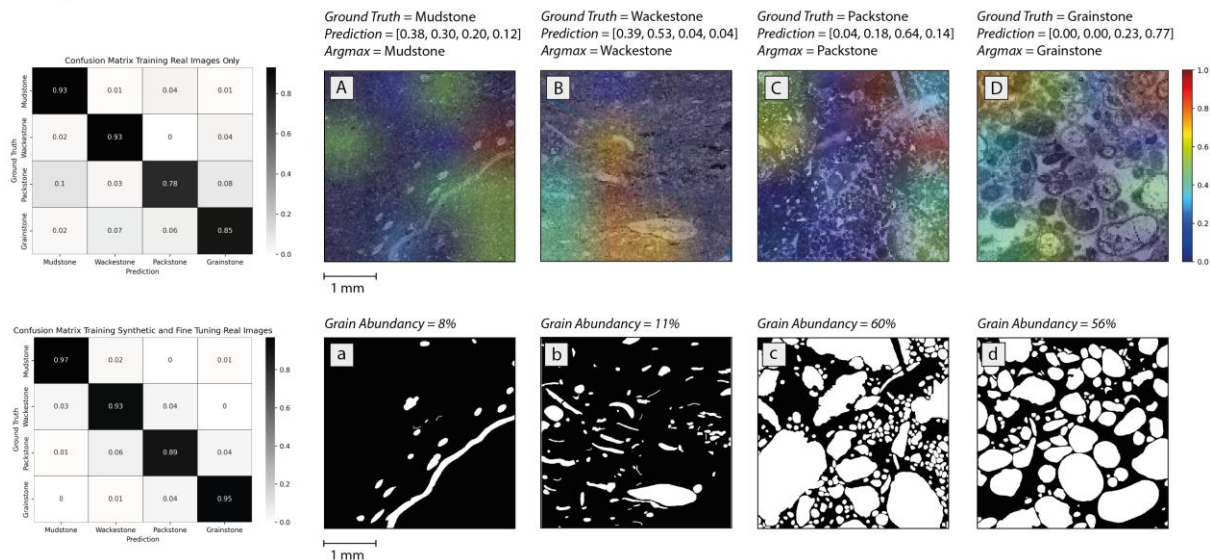


Figure 4.15 Confusion matrices comparing classical and transfer learning from synthetic to real image training: samples A, B, C, and D represent selected grain-annotated dataset examples for quantifying element abundance from binary masks a, b, c, d. Images are overlaid with Grad-CAM.

The confusion matrices revealed distinct performances between the two model training strategies (Figure 4.15). When trained exclusively on real images, the model exhibited some classification challenges. Specifically, packstone and grainstone demonstrated the highest rates of misclassification, with packstone occasionally being mistaken for mudstone (10%) and grainstone being confused with wackestone (7%) with some potential samples with fine micro-spar cement and low lightness of images explaining this confusion. The transfer learning method proved more effective in mitigating these classification errors. By initially training on synthetic data and subsequently adjusting model weights on real images, the models better converged to the target distribution, thereby reducing misclassification rates and improving overall performance. The packstone stood out as the most challenging rock texture class where it is still rarely confused with wackestone (6%).

Complementary tests were conducted on specific samples (Figure 4.15A to Figure 4.15D) to evaluate whether the model effectively captured the quantitative principles of Dunham's classification and identified relevant regions for decision-making using Grad-CAM visualizations. These samples, selected from the segmentation dataset, allowed precise quantification of grain density through binary masks (Figure 4.15a to Figure 4.15d) and were chosen for their inherent ambiguity, with Figure 4.15A to Figure 4.15C positioned near the boundaries between classes.

In Figure 4.15A, the sample was classified as mudstone, containing just under 10% grains, but its proximity to the wackestone class was evident, with a 30% probability assigned to this alternative. Similarly, Figure 4.15B, a wackestone with 11% grains, was correctly identified but showed a 39% likelihood for mudstone as the second most probable class. The sample in Figure 4.15C, primarily a packstone, exhibited areas of cement precipitation between grains that could potentially lead to its misclassification as a grainstone. Probabilities for wackestone and grainstone classes were 18% and 14%, respectively. Despite these challenges, the model accurately classified all three ambiguous cases, with uncertainties aligning with the most plausible alternative classes.

The Grad-CAM analysis revealed that for Figure 4.15A and Figure 4.15B, the model focused on both grainy and muddy regions, closely replicating the interpretative approach a specialist might adopt. For the more complex samples in Figure 4.15C and Figure 4.15D, the model's focus spanned multiple regions, reflecting the importance of the overall image. While this weighting of regions was more difficult to interpret, it was consistent with the model's reliance on global features for accurate predictions. These results highlight the model's ability to navigate ambiguous scenarios effectively and demonstrates its robust application of Dunham's classification principles.

4.4.2 Image Segmentation

4.4.2.1 Binary Segmentation

Training solely on a limited set of real images constrained the models' ability to learn from diverse examples. Among traditional approaches, the AttentionResU-Net model demonstrated the most robust performance, with an accuracy of 87.50%, an IoU of 67.46%, and a point counting R^2 score of 80.06%. ResU-Net architecture, however, achieved the highest DC of 77.71%.

When trained on synthetic data and fine-tuned with real images, all models achieved comparable performance metrics with values fluctuating around $\pm 1.76\%$ (Figure 4.16A). AttentionResU-Net achieved the highest accuracy at 89.31%, while the highest IoU score of 77.17% was achieved by InceptionResU-Net. This model also showed the best performance in the point counting task, reaching a R^2 score of 93.20%. Meanwhile, a top DC of 87.61% was recorded for the ResU-Net architecture (Table 4.3).

Model	Input Size	Depth	Parameters	Size
U-Net	$512 \times 512 \times 3$	14	6,481,256	25,360 MB
ResU-Net	$512 \times 512 \times 3$	21	6,568,072	25,712 MB
AttentionResU-Net	$512 \times 512 \times 3$	30	6,674,152	26,097 MB
InceptionResU-Net	$512 \times 512 \times 3$	56	11,692,924	45,788 MB

Training Real Images Only: Metrics on Test Data

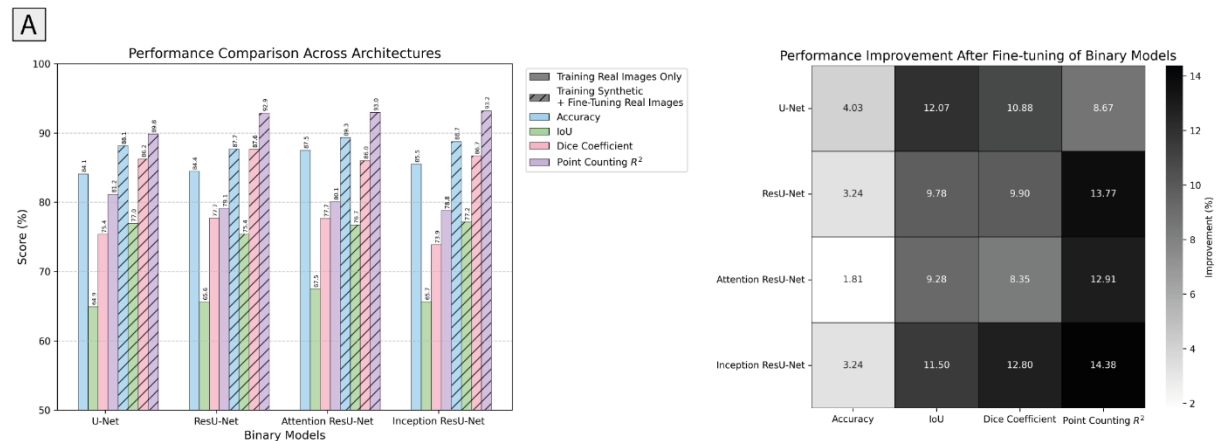
Binary Model	Accuracy	IoU	Dice	Point Counting R^2	Images/s
U-Net	84.08	64.91	75.36	81.16	18.48
ResU-Net	84.44	65.63	77.71	79.12	18.12
AttentionResU-Net	87.50	67.46	77.68	80.06	17.92
InceptionResU-Net	85.50	65.67	73.87	78.82	17.86

Training Synthetic and Fine-Tuning Real Images: Metrics on Test Data

Binary Model	Accuracy	IoU	Dice	Point Counting R^2	Images/s
U-Net	88.11	76.98	86.24	89.83	—” —
ResU-Net	87.68	75.41	87.61	92.89	—” —
AttentionResU-Net	89.31	76.74	86.03	92.97	—” —
InceptionResU-Net	88.74	77.17	86.67	93.20	—” —

Table 4.3 Segmentation model configurations and performance metrics for binary tasks: comparison of results between training on a small real image dataset and transfer learning from synthetic to real images.

Contribution of Synthetic Data Integration for Binary Image Segmentation



Contribution of Synthetic Data Integration for Multiclass Image Segmentation

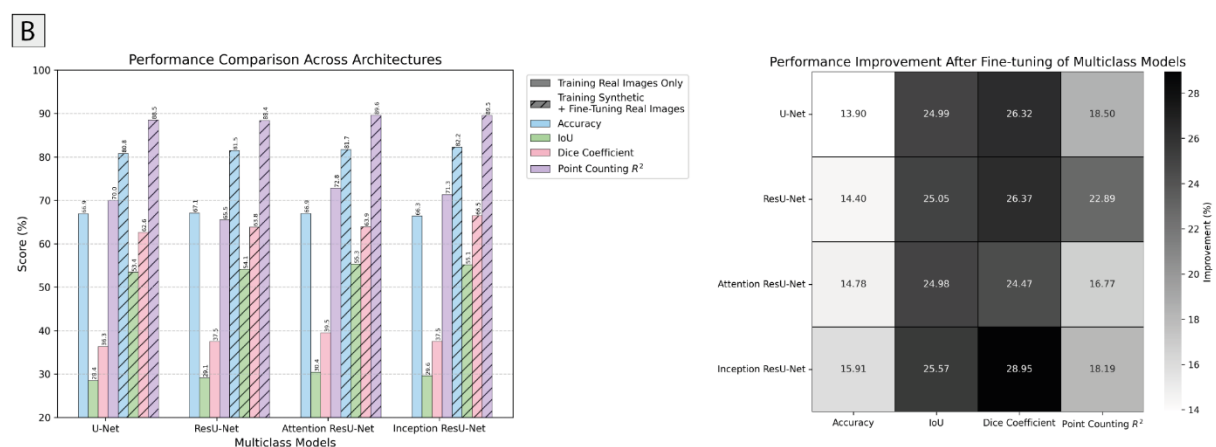


Figure 4.16 Models performance metrics comparison for each training strategy: comparing binary and multiclass segmentation with the impact of synthetic data integration in the training procedure through performance improvement monitoring.

The incorporation of synthetic data in the training process enhanced overall performance, yielding an average improvement of +9.04% across all performance metrics. The image processing speed matched that of classification models (Table 2 and Table 3), making this approach particularly valuable given the time-intensive nature of manual detouring.

4.4.2.2 Multiclass Segmentation

The multiclass grain segmentation, like the binary segmentation, was evaluated using the two developed training approaches: with synthetic data augmentation and without (Figure 4.16B). The same observation as for binary segmentation is made regarding the negligible impact of model architecture. The IoU metric, calculated for each object class, served as an additional evaluation criterion.

When using only real images during training, AttentionResU-Net slightly emerged as the overall top performer, achieving an IoU of 30.35%, a DC of 39.45%, and a point counting R^2 score of 72.80%. The ResU-Net model achieved the highest accuracy at 67.06%. Class-specific IoU scores resulting from this training method were generally poor to modest. Only the intraclast category in the conventional U-Net architecture exceeded 30%, while skeletal elements showed the lowest performance at 6.87% (Table 4.4).

Multiclass Model	Accuracy	IoU	Dice	Point Counting R^2
U-Net	66.88	28.44	36.29	70.01
ResU-Net	67.06	29.06	37.46	65.48
AttentionResU-Net	66.93	30.35	39.45	72.80
InceptionResU-Net	66.34	29.56	37.50	71.33

Training Real Images Only: Detailed IoUs on Test Data

Multiclass Model	Background	Skeletal	Intraclast	Ooid	Peloid
U-Net	66.64	6.87	30.22	20.86	17.60
ResU-Net	67.34	9.30	28.35	22.21	18.09
AttentionResU-Net	68.62	12.24	27.31	24.57	18.99
InceptionResU-Net	68.26	11.09	26.98	23.11	18.36

Training Real Images Only: Metrics on Test Data

Multiclass Model	Accuracy	IoU	Dice	Point Counting R^2
U-Net	80.78	53.43	62.61	88.51
ResU-Net	81.46	54.11	63.83	88.37
AttentionResU-Net	81.71	55.33	63.92	89.57
InceptionResU-Net	82.25	55.13	66.45	89.52

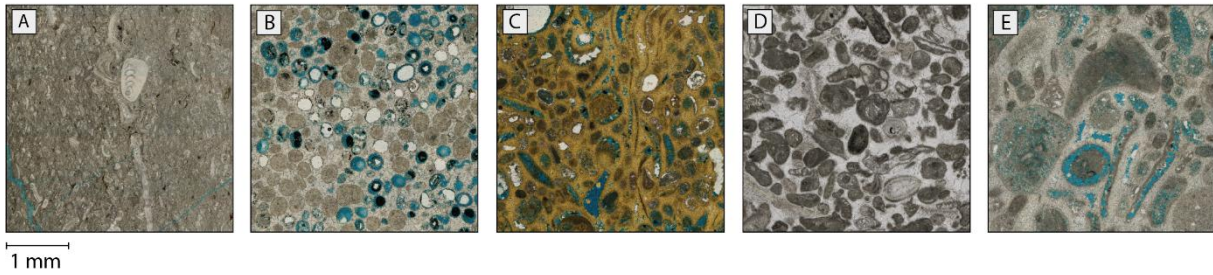
Training Synthetic and Fine-Tuning Real Images: Detailed IoUs on Test Data

Multiclass Model	Background	Skeletal	Intraclast	Ooid	Peloid
U-Net	79.01	47.46	47.53	47.85	45.29
ResU-Net	77.42	51.20	49.13	48.04	46.78
AttentionResU-Net	80.56	49.97	50.18	48.85	47.06
InceptionResU-Net	79.54	50.06	50.02	49.04	47.01

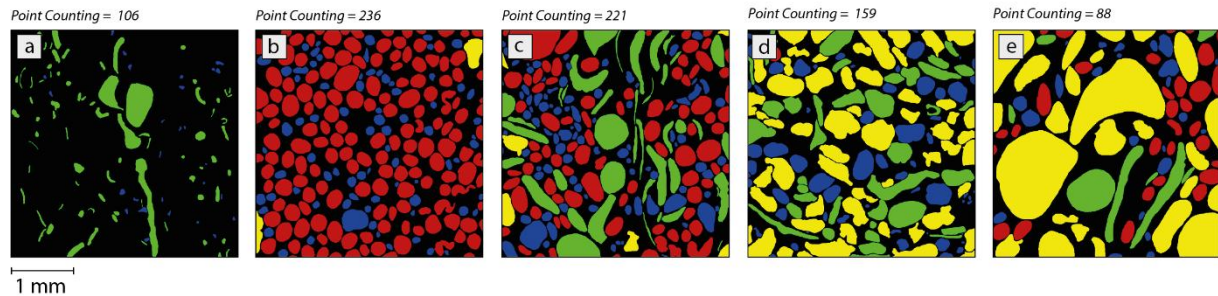
Table 4.4 Segmentation model performance metrics for multiclass tasks: comparison of results between training on a small real image dataset and transfer learning from synthetic to real images.

Image Segmentation Results

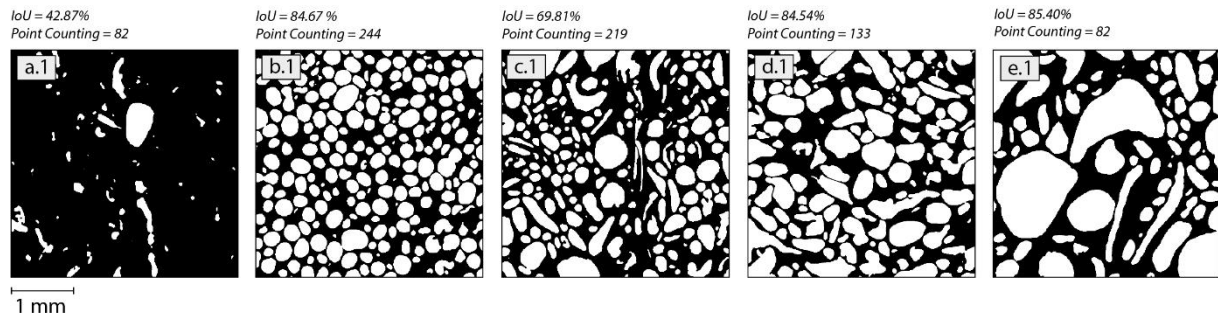
Test Images



Ground Truth



Binary Prediction



Multiclass Prediction

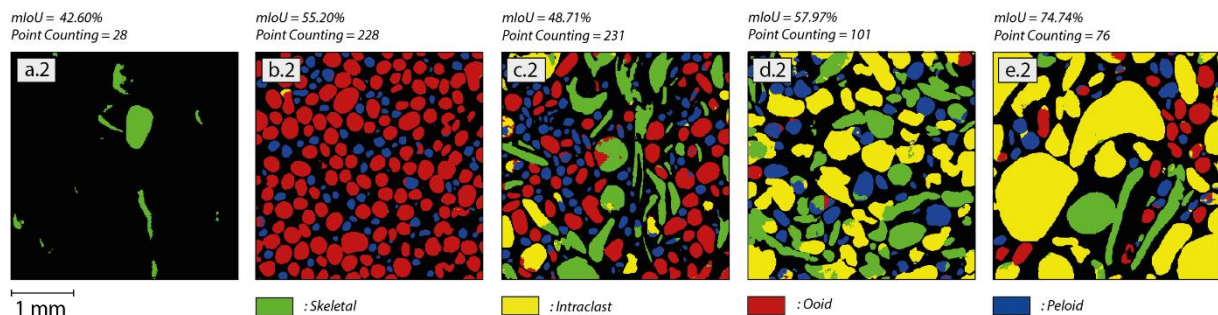


Figure 4.17 Comparison of binary and multiclass segmentation across various grain arrangements and rock textures: sample A (mudstone with skeletal and peloidal elements), sample B (grainstone with predominant ooids and peloids), sample C (grainstone with skeletal, ooid, and peloid grains), sample D (grainstone with intraclasts, skeletal, and peloids), and sample E (mixed grainstone).

Pre-training with synthetic data led to substantial global improvements. The InceptionResU-Net model achieved overall best performance with an accuracy 82.25% and a DC of 66.45%, while AttentionResU-Net topped an IoU score of 55.33% and a point counting R^2 of 89.57%. However, the binary model outperformed the multiclass model in the point counting task,

with a +3.63% difference between the two top-performing models. This can be pointed out in Figure 4.17d.1 and Figure 4.17d.2, where the boundaries between touching grains are better predicted in the binary mask. The multiclass model also tends to under-segment grains, as seen in Figure 4.17a.1 and Figure 4.17a.2. The integration of synthetic data also yielded more balanced and consistent class-specific IoU results (Table 4.4). For grain segmentation, scores improved from a range of 6.87% – 30.22% to 45.29% – 51.20%. Background segmentation also saw enhancement, with scores increasing from 66.64% – 68.62% to 77.42% – 80.56%.

Finally, the use of synthetic data as a first training stage in a multiclass segmentation framework delivered the most significant improvement among all the tasks investigated, with an average increase of + 20.9% across all metrics.

4.4.3 Description and Quantification at Full Thin Section Scale

The description of entire thin sections in petrographic analysis is complex and can vary significantly between interpreters. This variability stems primarily from the need to extrapolate local-scale observations to broader areas. The challenge is even more pronounced when analyzing samples with spatial heterogeneities in rock textures, grain distribution, or porosity types. However, classification and segmentation performed at the tile scale can be effectively replicated across a global image at consistent resolution using a sliding window prediction with overlapping tiles.

From Image Classification to Grain Segmentation

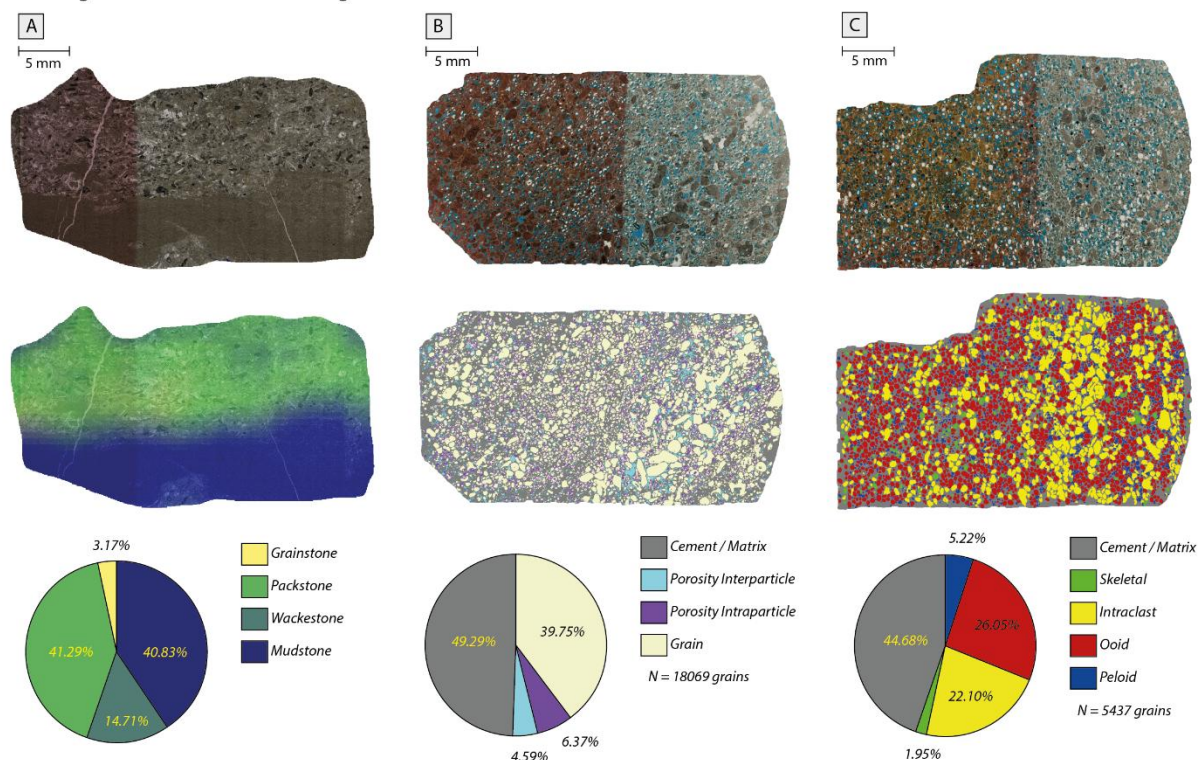


Figure 4.18 Quantitative analysis of thin sections at full image scale: from rock texture to grain identification. Sample A (heterogeneous rock texture primarily composed of mudstone and packstone), sample B (relative proportions of intra-granular and inter-granular porosity), sample C (coarse intraclast laminations within an oolitic-dominated deposit).

Figure 4.18A illustrates a multi-textured sample that presents some of these descriptive issues, featuring an upper region with dense grain concentration and a lower region dominated by mud material. The rock texture classification model was applied using prediction tiles matching the training stage's size and resolution, with 85% of overlap, assigning specific colors to the identified rock textures. The resulting image was produced in 35.76 seconds. Moreover, the substantial overlap enabled precise delineation of the horizontal boundary between the packstone region (41.29% of the total thin section) and the mudstone region (40.83%). The analysis also identified wackestone textures (14.71%) in less grain-supported areas and grainstone textures (3.17%) in grainy and cemented locations. This approach enables consistent regional characterization based on learned representations, facilitating comprehensive sample description while preserving heterogeneity information.

The grain segmentation task employed a similar overlapping tiles approach but with a reduced overlap of 25%. Predictions at patch intersections were averaged, and multi-inference was implemented using traditional augmentations (rotations, brightness and contrast adjustments). Figure 4.18B shows the binary model's segmentation results, which proved more effective for grain counting compared to the multiclass model (Table 3 and Table 4). During inference, this grain segmentation was generated in 39.43 seconds. Furthermore, it was combined with blue color thresholding in the HSV color space of the original images to identify stained epoxy resin that fills the porous spaces. This integrated approach enabled the identification and quantification of two distinct types of visible porosity: intra-granular porosity (occurring within grains) at 6.37%, and inter-granular porosity (occurring between grains) at 4.59%.

Figure 4.18C presents the multiclass grain segmentation across an entire thin section. A post-processing addressed multiple pixel class assignments within single grains by applying the predominant label to each grain. Indeed, U-Net based architectures are prone to this phenomenon (Figure 4.17b.2 to Figure 4.17e.2) since they are basically pixel classifiers. The resulting semantic mask was here generated in 37.64 seconds and enabled the quantification of relative grain distribution with 26.05% of ooids, 22.10% of intraclasts, 5.22% of peloids, and 1.95% of skeletal grains. Note that the background category encompasses both cement and matrix, as distinguishing between these components was beyond these study's objectives.

Finally, this method demonstrates how classification and segmentation techniques can complement each other in a multiscale approach for thin section characterization.

4.5 Discussions

4.5.1 Synthetic Data Benefits and Limits

The synthetic image generation method proposed with SynSection presents both promising opportunities and limitations for carbonate thin section analysis. While no previous studies have attempted a similar approach in petrography, synthetic data has been successfully incorporated in training pipelines for other fields such as medical imaging and autonomous driving (Pezoulas et al., 2024; Song et al., 2023), enabling representation of scenarios that would otherwise be underrepresented or entirely absent from datasets.

A significant advantage of the proposed approach is its ability to efficiently generate large and annotated datasets, addressing the fundamental challenge of time-consuming manual labeling in petrographic analysis. Precise control over both grain abundance, spatial distribution and class selection, as well as background types, is provided by the configurable nature of SynSection, which directly aligns with geological classification like the Dunham nomenclature. This controllability enables the generation of examples across the parameter range of each rock texture type, exposing models to a diverse spectrum of characteristics beyond what can be reproduced by conventional data augmentation.

The benefits of this data generation approach are evident in the experimental results. While exclusive use of synthetic data proved insufficient for optimal model training, substantial performance improvements were yielded by its integration with real data in a transfer learning strategy. This finding aligns with observations in Shorten & Khoshgoftaar (2019), who emphasized the complementary role of synthetic data in domains with limited labeled examples. The effectiveness of the method in enhancing DL model training outcomes is demonstrated by the improvement in performance metrics, particularly for overcoming the limitations of small datasets that cannot capture the full spectrum of carbonate facies variability.

However, several limitations are identified in the generation process. The primary challenge lies in capturing the full variability of grain sectioning planes in carbonate thin sections, particularly in skeletal elements that present an inherent limitation during data acquisition. Additionally, composite grains and important diagenetic features such as compaction figures, fractures, stylolites, and dolomitization patterns remain unaddressed in the current study.

The generation of background cement and matrix textures represents an area with potential for further exploration with conditional (amount of porosity, cement or matrix types, etc.). Diverse textures can be effectively produced by StyleGAN-2, although certain cement types were likely missed during the labeling process. A sensitivity to hyperparameter tuning is also observed in this architecture, which can lead to model collapse and produce unrealistic or low-variability results (Kossale et al., 2022). Less spatial heterogeneities are generally displayed in the generated textures compared to natural counterparts. Limitations are also revealed in the image blending process when handling significant disparities between grain and background color distributions, suggesting the need for more sophisticated techniques. Examples of these observations are presented in Figure 4.19.

Synthetic Examples with Color Blending Problem

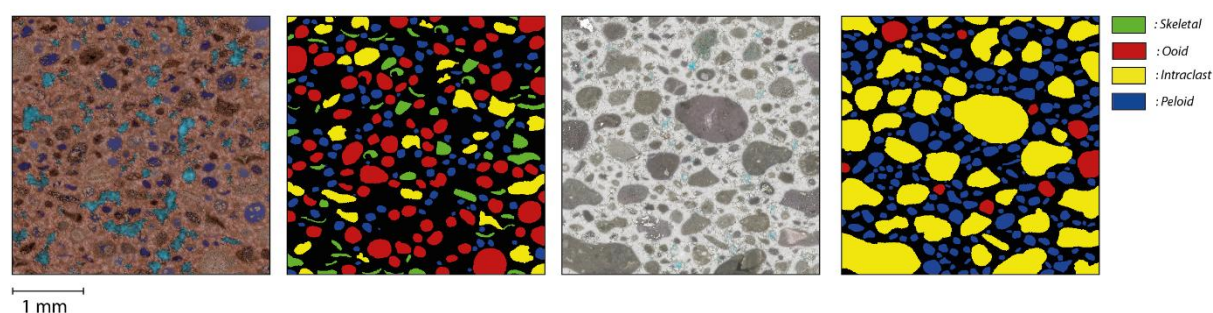


Figure 4.19 Synthetic images showing colors mismatch between allochems and the selected backgrounds highlighting the introduction of potential non realistic examples in the database.

These limitations and observations provide valuable direction for future research in synthetic image generation applied to carbonates petrography. Addressing these challenges could further enhance the utility and performance of DL in geological image analysis applications.

4.5.2 Model Performance Analysis in Classification and Segmentation

Among the tested classification architectures, superior performance was demonstrated by ResNet and InceptionResNet models. This performance difference can likely be attributed to residual layers capabilities and inception multiscale kernel approach, by which the hierarchical organization of features in carbonate thin sections may be probably better captured. The findings in Dawson et al. (2023) for core image analysis are corroborated with this observation as well as the training instabilities encountered with the VGG-like model. Despite these performance differences between architectures, it is clearly demonstrated by the results that the major factor driving metric improvements was the addition of synthetic images, highlighting that diversity of examples remains the foundation of effective deep learning approaches.

Valuable insights into model behavior were provided using Grad-CAM, whereby alignment with expert-defined key features was ensured, addressing concerns raised by Dramsch (2020) regarding the geological relevance of features extracted by deep learning models. This method was also particularly informative in detecting overfitting, as weighted regions were primarily located on image borders when this occurred. The development of specialized network architectures demonstrates that lightweight models can achieve state-of-the-art results. These results are comparable to pre-built and pre-trained models originally designed for open-source natural image datasets such as ImageNet. The key advantage of the proposed approach is that it extracts features specifically tailored to the target task, whereas conventional transfer learning relies on more generalized feature extraction that may not be perfectly optimized for the specific application.

To date, segmentation in carbonate thin sections is still rarely investigated (Koeshidayatullah et al., 2022; Liu et al., 2023) especially for multiple grain identification, making these results a potential starting point for future work. The different U-Net like implementations did not show significant differences in performance. However, the metrics obtained are comparable to results reported in nuclei segmentation studies in medical imaging (Caicedo et al., 2019; Lagree et al., 2021). This performance level is particularly noteworthy given the shared characteristics between these two types of source data containing compact objects that can be either closely spaced or dispersed. Higher performance in grain counting was demonstrated by the binary segmentation approach compared to multiclass segmentation, likely due to its more focused objective. Moreover, it was observed that the weight maps applied to the cross-entropy loss function were critical for delineating closely spaced grains in both segmentation tasks, addressing challenge similarly encountered and addressed in Ronneberger et al. (2015). Although lower accuracy was exhibited by multiclass segmentation, reflecting the challenges of grain type classification where diagnostic features can be obscured by interpretation, diagenetic alterations and sectioning effects, valuable compositional estimates are still provided. Furthermore, as highlighted in Mun et al. (2023) study, IoU scores

can easily decrease with a shift of a few pixels when working with small objects as observed in carbonates thin sections. The semantic quality of the predictions and the geological interpretability of the results were enhanced by the implementation of post-processing techniques, specifically the assignment of dominant labels to isolated grains.

4.5.3 Practical Workflow Implications and Future Recommendations

The most immediate practical benefit of the approach is the substantial reduction in analysis time. Traditional manual methods required around 500 person-hours to classify 15,883 examples and annotate 150 tiles with precise grain boundaries. In contrast, the trained models could process equivalent volumes in just a few minutes for image classification and a few seconds for segmentation. A critical bottleneck in petrographic workflows where traditional visual frequency estimation may be difficult to assess and biased, is here addressed by migrating towards quantification on predicted outputs at full sample scale. This approach can also help to highlight sedimentological features such as laminations formed by specific grain types, as shown in Figure 4.18C.

Promise for further characterizing and quantifying porous network attributes in carbonates is shown by combining the output segmentation mask with porosity thresholding. As demonstrated in Azzam et al. (2024) and Portugal et al. (2024) diagenetic processes can be better comprehended through pore type and geometry. More nuanced understanding of structure-property relationships in reservoir rocks can be developed by analysts through the integration of grain segmentation and classification with pore network attributes. This offers a potential way to refine geological characterization that may constitute input information for enhanced rock-type processing (Rebelle & Lalanne, 2014).

Despite these advances, several challenges remain for future research. Only a subset of the Dunham classification system is addressed by the current implementation, and further investigation is required for its extension to unaddressed classes. Current grain identification processes face challenges due to single-person annotation, which can lead to interpretation biases. Establishing a reference dataset through multiple annotators examining the same images and determining class attribution through consensus would enhance reliability.

The image generation method proposed through SynSection could be improved in several areas. Color harmonization challenges previously evocated could be addressed using deep learning methods incorporating self-supervised image harmonization (Jiang et al., 2021), while the packing method could benefit from more advanced geometric constraints or reinforcement learning approaches (Du et al., 2023; Fang et al., 2023). Furthermore, the groundmass texture synthesis model could be enhanced with refined style layers to produce realizations with greater spatial heterogeneity. All these improvements aim at addressing a certain domain gap that exists between synthetic and real images, which would become even more pronounced when applying transfer learning methods to thin sections from different locations and stratigraphic contexts. To overcome this challenge, exploring various techniques for adapting images from synthetic to real domains also represents a promising direction (Cherian & Sullivan, 2019; Peng et al., 2018; Toldo et al., 2020).

Image segmentation presents potential for further investigation since the infinite variability of grains, particularly in skeletal elements across geological timescales, poses a significant challenge. The current practice of grouping these elements into a single class, while efficient for training purposes, currently limits their utility in paleoecological and biostratigraphic applications. A more nuanced classification system would therefore be valuable with sufficient annotations. Although incorporating more real-world data would potentially improve model accuracy, the costs of data collection and annotation remain a significant barrier. An alternative approach would involve first developing robust binary segmentation models to identify grains, then implementing self-supervised learning strategies for pseudo grain classification based on similarity measures (Burgess et al., 2024). Finally, the multi-scale nature of carbonate petrography presents unique challenges, as some grains span multiple orders of magnitude in size, potentially exceeding single tile dimensions. An ensemble learning approach utilizing dual models would be recommended: one for detecting large grains at the full thin section scale, and another for segmenting smaller grains at higher resolution (Podda et al., 2024a).

4.6 Conclusion

This study introduced SynSection, a method addressing the data limitation challenge in deep learning approaches for carbonates petrography through conditional synthetic image generation. This technique enables deep learning applications in a domain traditionally constrained by limited labeled examples. The performance improvements achieved through transfer learning from synthetic to real domain, suggest a viable solution for other studies facing data scarcity. The automation of routine descriptive tasks through these methods could reshape how geologists conduct thin section analysis, allowing more time for interpretation and advanced analysis. In both research and industrial settings, this capability for systematic quantitative analysis has practical applications in reservoir characterization and resource assessment workflows where traditional approaches are limited to frequency estimations.

The integration of sedimentological principles into the synthetic data generation process demonstrates the value of incorporating domain expertise into machine learning solutions. This strategy of constraining synthetic data generation with field-specific knowledge could be adapted for other specialized scientific domains prone to description biases and where training data must reflect natural constraints. Development of explainability techniques aligned with geological insights may also facilitate adoption among practitioners.

Several research directions emerge from this work. Investigation into minimizing the domain gap between synthetic and real images could enhance performance, while expansion to other clastic rocks or core data could broaden applicability.

REFERENCES

- Anjos, C. E. M. dos, Avila, M. R. V., Vasconcelos, A. G. P., Neta, A. M. P., Medeiros, L. C., Evsukoff, A. G., & Surmas, R. (2021). Deep learning for lithological classification of carbonate rock micro-CT images. *Computational Geosciences*, 25(3), 971–983. <https://doi.org/10.1007/s10596-021-10033-6>
- Arjovsky, M., Chintala, S., & Bottou, L. (2017). Wasserstein GAN. *arXiv preprint*, arXiv:1701.07875. <https://doi.org/10.48550/arXiv.1701.07875>
- Asmussen, P., Conrad, O., Günther, A., Kirsch, M., & Riller, U. (2015). Semi-automatic segmentation of petrographic thin section images using a "seeded-region growing algorithm" with an application to characterize weathered subarkose sandstone. *Computers & Geosciences*, 83, 89–99. <https://doi.org/10.1016/j.cageo.2015.05.001>
- Azzam, F., Blaise, T., & Brigaud, B. (2024). Automated petrographic image analysis by supervised and unsupervised machine learning methods. *Sedimentologica*, 2(1), e22.1594. <https://doi.org/10.57035/journals/sdk.2024.e22.1594>
- Bergen, K. J., Johnson, P. A., De Hoop, M. V., & Beroza, G. C. (2019). Machine learning for data-driven discovery in solid Earth geoscience. *Science*, 363(6433), eaau0323. <https://doi.org/10.1126/science.aau0323>
- Bergmann, U., Jetchev, N., & Vollgraf, R. (2017). Learning texture manifolds with the periodic spatial GAN. *arXiv preprint*, arXiv:1705.06566. <https://doi.org/10.48550/arXiv.1705.06566>
- Burgess, J., Nirschl, J. J., Zanellati, M. C., Lozano, A., Cohen, S., & Yeung-Levy, S. (2024). Orientation-invariant autoencoders learn robust representations for shape profiling of cells and organelles. *Nature Communications*, 15(1), 1758. <https://doi.org/10.1038/s41467-024-45362-4>
- Burt, P., & Adelson, E. (1983). The Laplacian Pyramid as a compact image code. *IEEE Transactions on Communications*, 31(4), 532–540. <https://doi.org/10.1109/TCOM.1983.1095851>
- Burt, P. J., & Adelson, E. H. (1983). A multiresolution spline with application to image mosaics. *ACM Transactions on Graphics*, 2(4), 217–236. <https://doi.org/10.1145/245.247>
- Caicedo, J. C., Goodman, A., Karhohs, K. W., Cimini, B. A., Ackerman, J., Haghighi, M., Heng, C. K., Becker, T., Doan, M., McQuin, C., Rohban, M., Singh, S., & Carpenter, A. E. (2019). Nucleus segmentation across imaging experiments: The 2018 Data Science Bowl. *Nature Methods*, 16(12), 1247–1253. <https://doi.org/10.1038/s41592-019-0612-7>
- Canny, J. (1986). A computational approach to edge detection. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 8(6), 679–698. <https://doi.org/10.1109/TPAMI.1986.4767851>
- Cherian, A., & Sullivan, A. (2019). Sem-GAN: Semantically-consistent image-to-image translation. In *2019 IEEE Winter Conference on Applications of Computer Vision (WACV)* (pp. 1797–1806). IEEE. <https://doi.org/10.1109/WACV.2019.00196>

- Dawson, H. L., Dubrule, O., & John, C. M. (2023). Impact of dataset size and convolutional neural network architecture on transfer learning for carbonate rock classification. *Computers & Geosciences*, 171, 105284. <https://doi.org/10.1016/j.cageo.2022.105284>
- Ding, K., Zhou, M., Wang, H., Gevaert, O., Metaxas, D., & Zhang, S. (2023). A large-scale synthetic pathological dataset for deep learning-enabled segmentation of breast cancer. *Scientific Data*, 10(1), 231. <https://doi.org/10.1038/s41597-023-02125-y>
- Dramsch, J. S. (2020). Chapter One - 70 years of machine learning in geoscience in review. In B. Moseley & L. Krischer (Eds.), *Machine Learning in Geosciences, Advances in Geophysics* (Vol. 61, pp. 1–55). Elsevier. <https://doi.org/10.1016/bs.agph.2020.08.002>
- Du, L., Liu, X., Han, Y., & Deng, Z. (2023). Generation of irregular particle packing with prescribed statistical distribution, spatial arrangement, and volume fraction. *Journal of Rock Mechanics and Geotechnical Engineering*, 15(2), 375–394. <https://doi.org/10.1016/j.jrmge.2022.03.009>
- Dunham, R. J. (1962). Classification of carbonate rocks according to depositional textures. In W. E. Ham (Ed.), *Classification of Carbonate Rocks: A Symposium* (pp. 108–121). AAPG Memoir, 1. <https://doi.org/10.1306/M1357>
- Efros, A. A., & Leung, T. K. (1999). Texture synthesis by non-parametric sampling. In *Proceedings of the Seventh IEEE International Conference on Computer Vision* (Vol. 2, pp. 1033–1038). IEEE. <https://doi.org/10.1109/ICCV.1999.790383>
- Embry, A. F., & Klován, J. E. (1971). A late Devonian reef tract on northeastern Banks Island, N.W.T. *Bulletin of Canadian Petroleum Geology*, 19(4), 730–781. <https://doi.org/10.35767/gscpgbull.19.4.730>
- Evgeniou, T., & Pontil, M. (2001). Support vector machines: Theory and applications. In *Lecture Notes in Computer Science* (Vol. 2049, pp. 249–257). Springer. https://doi.org/10.1007/3-540-44673-7_12
- Fang, J., Rao, Y., Zhao, X., & Du, B. (2023). A hybrid reinforcement learning algorithm for 2D irregular packing problems. *Mathematics*, 11(2), 327. <https://doi.org/10.3390/math11020327>
- Fawcett, T. (2006). An introduction to ROC analysis. *Pattern Recognition Letters*, 27(8), 861–874. <https://doi.org/10.1016/j.patrec.2005.10.010>
- Ferreira, I., Ochoa, L., & Koeshidayatullah, A. (2022). On the generation of realistic synthetic petrographic datasets using a style-based GAN. *Scientific Reports*, 12(1), 8172. <https://doi.org/10.1038/s41598-022-16034-4>
- Flügel, E. A. (2004). *Microfacies of carbonate rocks: Analysis, interpretation and application*. Springer, Berlin. 976 pp. <https://doi.org/10.1007/978-3-662-08726-8>
- Gatys, L. A., Ecker, A. S., & Bethge, M. (2015). Texture synthesis using convolutional neural networks. *Advances in Neural Information Processing Systems*, 28, 262–270.
- Goodfellow, I., Pouget-Abadie, J., Mirza, M., Xu, B., Warde-Farley, D., Ozair, S., Courville, A., & Bengio, Y. (2014). Generative adversarial networks. *Advances in Neural Information Processing Systems*, 27, 2672–2680. <https://doi.org/10.1145/3422622>

- Gulrajani, I., Ahmed, F., Arjovsky, M., Dumoulin, V., & Courville, A. (2017). Improved training of Wasserstein GANs. In *Proceedings of the 31st International Conference on Neural Information Processing Systems* (pp. 5769–5779). Curran Associates Inc.
- Guo, B., Zhang, Y., Hu, J., Li, J., Wu, F., Peng, Q., & Zhang, Q. (2022). Two-dimensional irregular packing problems: A review. *Frontiers in Mechanical Engineering*, 8, 966691. <https://doi.org/10.3389/fmech.2022.966691>
- Guo, G., Wang, H., Bell, D., Bi, Y., & Greer, K. (2003). KNN model-based approach in classification. In R. Meersman, Z. Tari, & D. C. Schmidt (Eds.), *On The Move to Meaningful Internet Systems 2003: CoopIS, DOA, and ODBASE* (pp. 986–996). Springer. https://doi.org/10.1007/978-3-540-39964-3_62
- Haralick, R. M. (1973). Textural features for image classification. *IEEE Transactions on Systems, Man, and Cybernetics*, SMC-3(6), 610–621. <https://doi.org/10.1109/TSMC.1973.4309314>
- Harris, P. M., St. C. Kendall, C. G., & Lerche, I. (1985). Carbonate cementation - A brief review. In *Carbonate Cements* (pp. 79–95). SEPM (Society for Sedimentary Geology). <https://doi.org/10.2110/pec.85.36.0079>
- He, K., Zhang, X., Ren, S., & Sun, J. (2016). Deep residual learning for image recognition. In *2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)* (pp. 770–778). IEEE. <https://doi.org/10.1109/CVPR.2016.90>
- He, L., Ren, X., Gao, Q., Zhao, X., Yao, B., & Chao, Y. (2017). The connected-component labeling problem: A review of state-of-the-art algorithms. *Pattern Recognition*, 70, 25–43. <https://doi.org/10.1016/j.patcog.2017.04.018>
- Heeger, D. J., & Bergen, J. R. (1995). Pyramid-based texture analysis/synthesis. In *Proceedings of the 22nd Annual Conference on Computer Graphics and Interactive Techniques* (pp. 229–238). ACM. <https://doi.org/10.1145/218380.218446>
- Heusel, M., Ramsauer, H., Unterthiner, T., Nessler, B., & Hochreiter, S. (2017). GANs trained by a two time-scale update rule converge to a local Nash equilibrium. *Advances in Neural Information Processing Systems*, 30, 6626–6637.
- Hinrichsen, E. L., Feder, J., & Jøssang, T. (1990). Random packing of disks in two dimensions. *Physical Review A*, 41(8), 4199–4209. <https://doi.org/10.1103/PhysRevA.41.4199>
- Huang, G., Liu, Z., Van Der Maaten, L., & Weinberger, K. Q. (2017). Densely connected convolutional networks. In *2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)* (pp. 2261–2269). IEEE. <https://doi.org/10.1109/CVPR.2017.243>
- Jarvis, E., Wang, H., Dietz, J., & Van Der Looven, T. (2024). Large volume analysis of core and thin section images in the assessment of Brazil pre-salt reservoir distribution. *First Break*, 42(3), 79–87. <https://doi.org/10.3997/1365-2397.fb2024018>
- Jetchev, N., Bergmann, U., & Vollgraf, R. (2016). Texture synthesis with spatial generative adversarial networks. *arXiv preprint*, arXiv:1611.08207. <https://doi.org/10.48550/arXiv.1611.08207>

- Jiang, Y., Zhang, H., Zhang, J., Wang, Y., Lin, Z., Sunkavalli, K., Chen, S., Amirghodsi, S., Kong, S., & Wang, Z. (2021). SSH: A self-supervised framework for image harmonization. In *2021 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)* (pp. 4832–4841). IEEE. <https://doi.org/10.1109/CVPR46437.2021.00479>
- Jylänki, J. (2010). *A thousand ways to pack the bin - A practical approach to two-dimensional rectangle bin packing*. Retrieved from <http://pds25.egloos.com/pds/201504/21/98/RectangleBinPack.pdf>
- Karras, T., Laine, S., & Aila, T. (2019). A style-based generator architecture for generative adversarial networks. In *2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)* (pp. 4401–4410). IEEE. <https://doi.org/10.1109/CVPR.2019.00453>
- Karras, T., Laine, S., Aittala, M., Hellsten, J., Lehtinen, J., & Aila, T. (2020). Analyzing and improving the image quality of StyleGAN. In *2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)* (pp. 8107–8116). IEEE. <https://doi.org/10.1109/CVPR42600.2020.00813>
- Kingma, D. P., & Ba, J. (2014). Adam: A method for stochastic optimization. *arXiv preprint*, arXiv:1412.6980. <https://doi.org/10.48550/arXiv.1412.6980>
- Koeshidayatullah, A., Morsilli, M., Lehrmann, D. J., Al-Ramadan, K., & Payne, J. L. (2020). Fully automated carbonate petrography using deep convolutional neural networks. *Marine and Petroleum Geology*, 122, 104687. <https://doi.org/10.1016/j.marpetgeo.2020.104687>
- Koeshidayatullah, A., Trower, E. J., Li, X., Mukerji, T., Lehrmann, D. J., Morsilli, M., Al-Ramadan, K., & Payne, J. L. (2022). Quantitative evaluation of the roles of ocean chemistry and climate on ooid size across the Phanerozoic: Global versus local controls. *Sedimentology*, 69(6), 2486–2506. <https://doi.org/10.1111/sed.12998>
- Kohavi, R. (1995). A study of cross-validation and bootstrap for accuracy estimation and model selection. In *Proceedings of the 14th International Joint Conference on Artificial Intelligence* (Vol. 2, pp. 1137–1143). Morgan Kaufmann.
- Kossale, Y., Airaj, M., & Darouichi, A. (2022). Mode collapse in generative adversarial networks: An overview. In *2022 8th International Conference on Optimization and Applications (ICOA)* (pp. 1–6). IEEE. <https://doi.org/10.1109/ICOA55659.2022.9934291>
- Krizhevsky, A., Sutskever, I., & Hinton, G. E. (2012). ImageNet classification with deep convolutional neural networks. In F. Pereira, C. J. Burges, L. Bottou, & K. Q. Weinberger (Eds.), *Advances in Neural Information Processing Systems* (Vol. 25, pp. 1097–1105). Curran Associates, Inc.
- Lagree, A., Mohebpour, M., Meti, N., Saednia, K., Lu, F. I., Slodkowska, E., Gandhi, S., Rakovitch, E., Shenfield, A., Sadeghi-Naini, A., & Tran, W. T. (2021). A review and comparison of breast tumor cell nuclei segmentation performances using deep convolutional neural networks. *Scientific Reports*, 11(1), 8025. <https://doi.org/10.1038/s41598-021-87496-1>
- LeCun, Y., Bengio, Y., & Hinton, G. (2015). Deep learning. *Nature*, 521(7553), 436–444. <https://doi.org/10.1038/nature14539>

- Liu, X., Chandra, V., Ramdani, A. I., Zuhlke, R., & Vahrenkamp, V. (2023). Using deep-learning to predict Dunham textures and depositional facies of carbonate rocks from thin sections. *Geoenergy Science and Engineering*, 227, 211906. <https://doi.org/10.1016/j.geoen.2023.211906>
- Lokier, S. W., & Al Junaibi, M. (2016). The petrographic description of carbonate facies: Are we all speaking the same language? *Sedimentology*, 63(7), 1843–1885. <https://doi.org/10.1111/sed.12293>
- Maas, M. V. R., Bedle, H., & de Matos, M. C. (2023). Seismic identification of carbonate reservoir sweet spots using unsupervised machine learning: A case study from Brazil deep water Aptian pre-salt data. *Marine and Petroleum Geology*, 151, 106199. <https://doi.org/10.1016/j.marpetgeo.2023.106199>
- Mosser, L., Ghon, G., & Baechele, G. (2022). Interpretation of deep neural networks for carbonate thin section classification. In *SEG Technical Program Expanded Abstracts* (pp. 371–375). Society of Exploration Geophysicists. <https://doi.org/10.1190/image2022-3748082.1>
- Mun, C., Lee, S., Uh, Y., Choe, J., & Byun, H. (2023). Small objects matters in weakly-supervised semantic segmentation. In *2023 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)* (pp. 8425–8434). IEEE. <https://doi.org/10.1109/CVPR52729.2023.00815>
- Nanjo, T. (2020). Carbonate lithology identification with generative adversarial networks. *International Petroleum Technology Conference*, Dhahran, Saudi Arabia. <https://doi.org/10.2523/IPTC-20226-ABSTRACT>
- Oktay, O., Schlemper, J., Folgoc, L. Le, Lee, M. J., Heinrich, M. P., Misawa, K., Mori, K., McDonagh, S. G., Hammerla, N. Y., Kainz, B., Glocker, B., & Rueckert, D. (2018). Attention U-Net: Learning where to look for the pancreas. *arXiv preprint*, arXiv:1804.03999. <https://doi.org/10.48550/arXiv.1804.03999>
- Otsu, N. (1979). A threshold selection method from gray-level histograms. *IEEE Transactions on Systems, Man, and Cybernetics*, 9(1), 62–66. <https://doi.org/10.1109/TSMC.1979.4310076>
- Paget, R., & Longstaff, I. D. (1995). Texture synthesis via a non-parametric Markov random field. In *Proceedings of Digital Image Computing: Techniques and Applications* (pp. 547–552).
- Peng, X., Usman, B., Saito, K., Kaushik, N., Hoffman, J., & Saenko, K. (2018). Syn2Real: A new benchmark for synthetic-to-real visual domain adaptation. In *2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops (CVPRW)* (pp. 856–8565). IEEE. <https://doi.org/10.1109/CVPRW.2018.00143>
- Pérez, P., Gangnet, M., & Blake, A. (2003). Poisson image editing. *ACM SIGGRAPH 2003 Papers* (pp. 313–318). ACM. <https://doi.org/10.1145/1201775.882269>
- Pezoulas, V. C., Zaridis, D. I., Mylona, E., Androustos, C., Apostolidis, K., Tachos, N. S., & Fotiadis, D. I. (2024). Synthetic data generation methods in healthcare: A review on open-source tools and methods. *Computational and Structural Biotechnology Journal*, 23, 2892–2910. <https://doi.org/10.1016/j.csbj.2024.07.005>

- Pires de Lima, R., Duarte, D., Nicholson, C., Slatt, R., & Marfurt, K. J. (2020). Petrographic microfacies classification with deep convolutional neural networks. *Computers & Geosciences*, 142, 104481. <https://doi.org/10.1016/j.cageo.2020.104481>
- Podda, A. S., Balia, R., Barra, S., Carta, S., Neri, M., Guerriero, S., & Piano, L. (2024). Multi-scale deep learning ensemble for segmentation of endometriotic lesions. *Neural Computing and Applications*, 36(16), 14895–14908. <https://doi.org/10.1007/s00521-024-09828-2>
- Portilla, J., & Simoncelli, E. P. (2000). A parametric texture model based on joint statistics of complex wavelet coefficients. *International Journal of Computer Vision*, 40(1), 49–70. <https://doi.org/10.1023/A:1026553619983>
- Portugal, D., Souza, J. P., Kuroda, M., Vidal, A. C., Mello, A., Resende, M., & Silva, Y. M. (2024). Quantification of mineral grains and pores in thin sections using mathematical morphology. In *85th EAGE Annual Conference & Exhibition* (pp. 1–5). EAGE Publications. <https://doi.org/10.3997/2214-4609.2024101299>
- Prados, R., & Garcia, R. (2014). State of the art in image blending techniques. In *SpringerBriefs in Computer Science* (pp. 35–60). Springer. https://doi.org/10.1007/978-3-319-05558-9_3
- Punn, N. S., & Agarwal, S. (2020). Inception U-Net architecture for semantic segmentation to identify nuclei in microscopy cell images. *ACM Transactions on Multimedia Computing, Communications and Applications*, 16(1), 15. <https://doi.org/10.1145/3376922>
- Rebelle, M., & Lalanne, B. (2014). Rock-typing in carbonates: A critical review of clustering methods. In *SPE Abu Dhabi International Petroleum Exhibition and Conference* (pp. 792–805). Society of Petroleum Engineers. <https://doi.org/10.2118/171759-MS>
- Reichstein, M., Camps-Valls, G., Stevens, B., Jung, M., Denzler, J., Carvalhais, N., & Prabhat. (2019). Deep learning and process understanding for data-driven Earth system science. *Nature*, 566(7743), 195–204. <https://doi.org/10.1038/s41586-019-0912-1>
- Ronneberger, O., Fischer, P., & Brox, T. (2015). U-Net: Convolutional networks for biomedical image segmentation. In *Medical Image Computing and Computer-Assisted Intervention – MICCAI 2015* (pp. 234–241). Springer. https://doi.org/10.1007/978-3-319-24574-4_28
- Saxena, N., Day-Stirrat, R. J., Hows, A., & Hofmann, R. (2021). Application of deep learning for semantic segmentation of sandstone thin sections. *Computers & Geosciences*, 152, 104778. <https://doi.org/10.1016/j.cageo.2021.104778>
- Scholle, P. A., & Ulmer-Scholle, D. S. (2003). *A color guide to the petrography of carbonate rocks: Grains, textures, porosity, diagenesis*. American Association of Petroleum Geologists. <https://doi.org/10.1306/M77973>
- Sekachev, B., Manovich, N., Zhiltsov, M., Zhavoronkov, A., Kalinin, D., Hoff, B., Tosmanov, Kruchinin, D., Zankevich, A., DmitriySidnev, Markelov, M., Johannes, Chenuet, M., Andre, A., Telenachos, Melnikov, A., Kim, J., Ilouz, L., Glazov, N., Priya, Tehrani, R., Jeong, S., Skubriev, V., Yonekura, S., Vugia, T., Zliang, Lihming, & Truong, T. (2020). OpenCV/CVAT: v1.1.0. Zenodo. <https://doi.org/10.5281/zenodo.4009388>

- Selvaraju, R. R., Cogswell, M., Das, A., Vedantam, R., Parikh, D., & Batra, D. (2017). Grad-CAM: Visual explanations from deep networks via gradient-based localization. In *2017 IEEE International Conference on Computer Vision (ICCV)* (pp. 618–626). IEEE. <https://doi.org/10.1007/s11263-019-01228-7>
- Shebl, H. T., Al Tamimi, M. A., Boyd, D. A., & Nehaid, H. A. (2021). Automation of carbonate rock thin section description using cognitive image recognition. In *Abu Dhabi International Petroleum Exhibition & Conference*. Society of Petroleum Engineers. <https://doi.org/10.2118/208149-MS>
- Shorten, C., & Khoshgoftaar, T. M. (2019). A survey on image data augmentation for deep learning. *Journal of Big Data*, 6(1), 60. <https://doi.org/10.1186/s40537-019-0197-0>
- Simonyan, K., & Zisserman, A. (2014). Very deep convolutional networks for large-scale image recognition. *arXiv preprint*, arXiv:1409.1556. <https://doi.org/10.48550/arXiv.1409.1556>
- Song, Z., He, Z., Li, X., Ma, Q., Ming, R., Mao, Z., Pei, H., Peng, L., Hu, J., Yao, D., & Zhang, Y. (2023). Synthetic datasets for autonomous driving: A survey. *IEEE Transactions on Intelligent Vehicles*, 8(11), 4748–4768. <https://doi.org/10.1109/TIV.2023.3331024>
- Sudre, C. H., Li, W., Vercauteren, T., Ourselin, S., & Cardoso, M. J. (2017). Generalised dice overlap as a deep learning loss function for highly unbalanced segmentations. In *Deep Learning in Medical Image Analysis and Multimodal Learning for Clinical Decision Support* (pp. 240–248). Springer. https://doi.org/10.1007/978-3-319-67558-9_28
- Szegedy, C., Liu, W., Jia, Y., Sermanet, P., Reed, S., Anguelov, D., Erhan, D., Vanhoucke, V., & Rabinovich, A. (2015). Going deeper with convolutions. In *2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)* (pp. 1–9). IEEE. <https://doi.org/10.1109/CVPR.2015.7298594>
- Szegedy, C., Vanhoucke, V., Ioffe, S., Shlens, J., & Wojna, Z. (2016). Rethinking the inception architecture for computer vision. In *2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)* (pp. 2818–2826). IEEE. <https://doi.org/10.1109/CVPR.2016.308>
- Toldo, M., Maracani, A., Michieli, U., & Zanuttigh, P. (2020). Unsupervised domain adaptation in semantic segmentation: A review. *Technologies*, 8(2), 35. <https://doi.org/10.3390/technologies8020035>
- Tucker, M. (2009). *Carbonate sedimentology*. Blackwell Publishing Ltd. 482 pp. <https://doi.org/10.1002/9781444314175.ch1>
- Vellappally, A., Hou, S., & Emmings, J. (2024). Automated grain segmentation and mineral classification in rock thin sections. In *85th EAGE Annual Conference & Exhibition* (pp. 1–5). EAGE Publications. <https://doi.org/10.3997/2214-4609.2024101405>
- Wang, H., Cao, W., Zhou, Y., Yu, P., & Yang, W. (2023). Multitarget intelligent recognition of petrographic thin section images based on Faster RCNN. *Minerals*, 13(7), 872. <https://doi.org/10.3390/min13070872>

- Wang, W., Wang, H., Dai, G., & Wang, H. (2006). Visualization of large hierarchical data by circle packing. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (pp. 517–520). ACM. <https://doi.org/10.1145/1124772.1124851>
- Wang, X., Gong, J., An, A., Zhang, K., & Nie, Z. (2019). Random generation of convex granule packing based on weighted Voronoi tessellation and cubic-polynomial-curve fitting. *Computers and Geotechnics*, 113, 103088. <https://doi.org/10.1016/j.compgeo.2019.05.003>
- Wu, H., Zhang, J., Zheng, S., & Huang, K. (2019). GP-GaN: Towards realistic high-resolution image blending. In *MM 2019 - Proceedings of the 27th ACM International Conference on Multimedia* (pp. 2487–2495). Association for Computing Machinery. <https://doi.org/10.1145/3343031.3350944>
- Zhang, L., Wen, T., & Shi, J. (2019). Deep image blending. *arXiv preprint*, arXiv:1910.11495. <https://doi.org/10.48550/arXiv.1910.11495>
- Zhang, Z., Liu, Q., & Wang, Y. (2018). Road extraction by deep residual U-Net. *IEEE Geoscience and Remote Sensing Letters*, 15(5), 749–753. <https://doi.org/10.1109/LGRS.2018.2802944>
- Zhao, X., Rao, Y., Qi, P., Lyu, Q., Yang, P., & Yu, S. (2023). An efficient constructive heuristic for the rectangular packing problem with rotations. *PLOS ONE*, 18(11), e0295206. <https://doi.org/10.1371/journal.pone.0295206>
- Zheng, W., Tian, F., Di, Q., Xin, W., Cheng, F., & Shan, X. (2021). Electrofacies classification of deeply buried carbonate strata using machine learning methods: A case study on Ordovician paleokarst reservoirs in Tarim Basin. *Marine and Petroleum Geology*, 123, 104720. <https://doi.org/10.1016/j.marpetgeo.2020.104720>
- Zhuang, F., Qi, Z., Duan, K., Xi, D., Zhu, Y., Zhu, H., Xiong, H., & He, Q. (2021). A comprehensive survey on transfer learning. *Proceedings of the IEEE*, 109(1), 43–76. <https://doi.org/10.1109/JPROC.2020.3004555>

CHAPTER 5: SYNSECTION 2 – ADAPTIVE SYNTHETIC DATA AUGMENTATION WITH UNIFIED DOMAIN TRANSFER AND SEGMENTATION IN CARBONATES PETROGRAPHY

Paper 3: In Preparation

Abstract

Petrographic analysis of carbonate rocks faces limitations due to time-intensive manual grain identification and the scarcity of properly labeled training datasets for deep learning applications. While synthetic data generation through SynSection has shown promise for training segmentation models, a persistent domain shift between synthetic and real petrographic images compromises both training efficiency and validation accuracy. This study introduces SynSection2, a unified framework that simultaneously performs adaptive image-to-image translation and semantic segmentation to bridge the synthetic-real domain gap while maintaining geological consistency.

Inspired by CycleGAN approaches, the proposed method integrates a shared encoder-decoder architecture with domain-specific decoders, adversarial training, and Adaptive Instance Normalization (AdaIN) for style transfer. Crucially, the framework employs masked cycle consistency loss applied exclusively to grain-containing regions, enabling enhanced textural variation in backgrounds while preserving fine-grained details within objects. A segmentation constraint prevents artifact generation while directly training a target-domain segmentation model within the same network.

The method demonstrates substantial improvements over the previously established synthetic data training and real data fine-tuning approach. Domain adaptation achieved a 67-point reduction in Fréchet Inception Distance (from 86.12 to 18.69), indicating enhanced image quality and distributional alignment. Segmentation performance improved consistently across all metrics: accuracy (+4%), Dice Coefficient (+5%), and Intersection over Union (+4%). Most notably, derived petrographic regression tasks showed marked enhancement, with mean grain size estimation improving by +12% ($R^2 = 0.78$) and grain sorting analysis by +6% ($R^2 = 0.72$).

The study introduces local thickness measurements as a robust alternative to traditional grain diameter calculations, demonstrating superior resilience to segmentation noise while aligning with visual petrographic interpretations. Full thin section characterization examples showcase the framework's practical utility for quantitative geological analysis, enabling objective compositional assessment and grain size distribution analysis that complements traditional expert interpretation.

Keywords: Carbonates, Petrography, Deep Learning, Synthetic Data, Image Segmentation, Domain Adaptation.

5.1 Introduction

Petrographic analysis of carbonates is a cornerstone methodology for understanding the depositional environments and diagenetic evolution of these sedimentary rocks. This traditional approach allows geologists to unravel complex geological histories through detailed examination of grains, groundmass material and mineral compositions in thin sections. However, several inherent challenges limit its effectiveness in modern studies. The intricate nature of carbonate textures, coupled with time-consuming manual grain identification where expertise is required for their specific classification, often restricts the acquisition of statistically robust quantitative data across large sample sets.

The emergence of digital imaging technologies and Deep Learning (DL) offer unprecedented opportunities to enhance and streamline the study of carbonate rocks (Dawson et al., 2023; Koeshidayatullah et al., 2020, 2022; Liu et al., 2023; Pires de Lima et al., 2020; Shebl et al., 2021). While human expertise continues to be essential for interpreting nuanced textural relationships and complex diagenetic features, DL models can provide rapid and homogeneous characterization of fundamental sedimentological parameters through classification, detection or segmentation algorithms (Azzam et al., 2024; Dawson & John, 2024; Mamode et al., 2025). This technological breakthrough addresses several long-standing challenges in traditional petrography, including operator subjectivity, time limitations, and the statistical constraints of conventional limited grain characterization techniques.

The implementation of DL methods also faces significant challenges, primarily due to the scarcity of properly labeled training datasets. In geological contexts, this problem of data acquisition is particularly related to the substantial time investment and specialized expertise required for accurate grain labeling in petrographic images. To overcome these constraints, data augmentation strategies have been developed, notably SynSection, which generates conditional synthetic training datasets from sparse annotated samples for both image classification and segmentation purposes (Ransinangue et al., 2025). These methodological advances have facilitated the creation, and the training of deep learning models specifically designed for carbonate petrographic analysis applications.

Although synthetic data generation has been shown to improve model performance compared to traditional training approaches using limited datasets, a domain shift between synthetic and real petrographic images persists, compromising both training efficiency and validation accuracy. To address this issue, this study proposes a dual task learning approach that simultaneously performs stylized image-to-image translation with semantic consistency and segmentation (Figure 5.1). This technique aims to transform synthetic images of a source domain to match the visual characteristics of target real petrographic data while conducting grain identification, thereby enhancing downstream segmentation performance in the target domain used during inference.

SynSection2 Framework

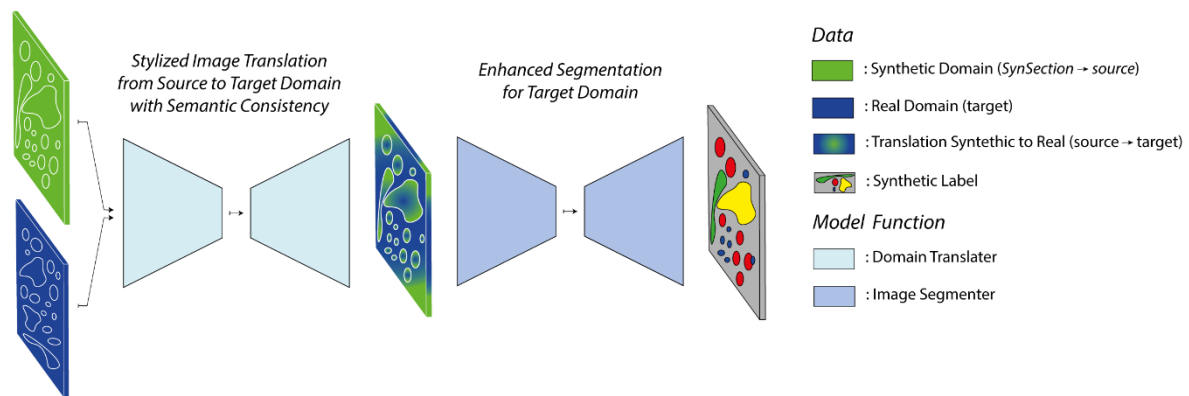


Figure 5.1 Simplified representation of SynSection2 framework translating images from synthetic to real domain with semantic consistency and reference style constraints.

SynSection2 Evaluation Performance

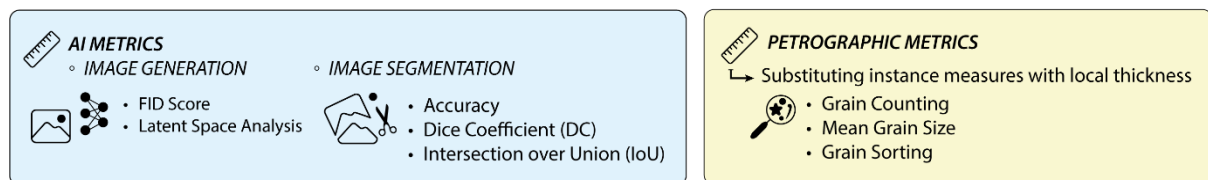


Figure 5.2 Conventional metrics used in computer vision both for image generation and image segmentation completed with specific petrographic metrics.

Another challenge identified in applying DL to geological studies is the development of appropriate evaluation metrics. Standard deep learning metrics used for image segmentation often fail to capture the specific requirements of geological analysis effectively. In petrography, one essential parameter is the characterization of grain size distribution, which provides information about hydrodynamic conditions of depositional environments. To address this need, this study introduces novel metrics specifically designed to evaluate grain size distribution parameters, assessing both the mean grain size and the grain sorting derived from predicted segmentation masks (Blott & Pye, 2001). The traditional grain size measurement method is proposed to be replaced by the local thickness approach (Dahl & Dahl, 2023; Gostick et al., 2019), which is observed to be less sensitive to segmentation noise and better aligned with visual petrographic interpretations (Figure 5.2).

The improvements observed in synthetic data generation, image segmentation, and evaluation metrics represent new steps toward making DL-assisted petrographic analysis more practical and reliable. These advancements could enable more efficient and accurate characterization of carbonate rocks, potentially leading to better understanding of geological processes and more informed decision-making in various applications, such as resource exploration and development.

5.2 Data

5.2.1 Raw Data

In this study, 199 high-resolution images of carbonate thin sections from Amellago outcrop, located in the central High Atlas Mountains of Morocco were compiled. The studied thin sections represent a shallow marine environment, characterized by limestones deposited in environments ranging from high to low energy alternating with low energy marly intervals. These images were acquired under plain polarized light at a resolution of $8.8 \times 10^{-1} \mu\text{m}$ per pixel, resulting in image sizes around $50,000 \times 25,000$ pixels. This database was initially unlabeled at the start of the study.

The use of raw data faces several constraints. Manual annotation presents the primary bottleneck, as selecting and labeling even a small subset of complete images requires extensive time investment while potentially failing to capture the full dataset's diversity. Additionally, computational limitations impose restrictions, as GPU capacity constraints can render training infeasible when model complexity and high-dimensional image batches exceed available memory.

5.2.2 Working Data

5.2.3 Data Preprocessing

Raw data underwent preprocessing to optimize images for both manual annotation and computational handling. Initially, images were downsampled to $7 \mu\text{m}$ per pixel resolution, which was found the minimum scale required for human interpreters to effectively identify constituent grains within thin sections for this dataset. Subsequently, images were partitioned into non-overlapping 512×512 tiles designed to preserve grain integrity while maintaining a manageable grain count per image (less than a few hundred grains) to facilitate annotator workflows and approximate typical Region of Interest (ROI) analysis conditions. Moreover, the predetermined tile dimensions facilitate straightforward manipulation and efficient processing by DCNNs algorithms.

5.2.4 Data Labeling

From the 18,452 generated tiles, 100 were manually selected for grain annotation and subsequently labeled using the open-source Computer Vision Annotation Tool (CVAT) (Sekachev et al., 2020). Grains not intersecting image borders were extracted based on their minimum enclosed area, creating pairs of thumbnail images and masks, which were then organized by class into dedicated directories. This process yielded 10,651 extracted objects distributed across four categories: 998 skeletal grains, 663 intraclasts, 1,897 ooids, 7,093 peloids (Figure 5.3A). Background components, such as cement and matrix, were also annotated using bounding boxes, enabling the extraction of 33,628 texture patches (128×128 pixels). The resulting instances exhibit class imbalance among labeled grains. This imbalance is attributed to the original data distribution, which consists primarily of grainstones dominated by peloids and ooids. Additionally, larger objects (e.g., intraclasts, shells) that occupy more significant image areas than smaller grains consequently receive less extensive labeling per image (Figure 5.3B).

Grain Characteristics

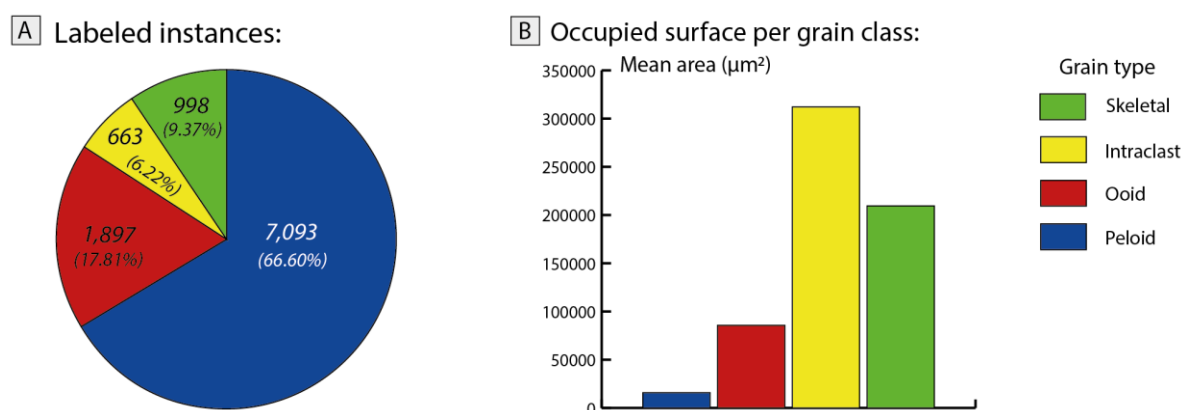


Figure 5.3 Grain data properties showing the distribution of labeled instances across selected image tiles and average surface area per grain type.

5.2.5 Synthetic Dataset Generation

A synthetic training dataset consisting of 50,000 image-mask pairs was constructed using SynSection approach, which leverages annotated grains and background patches. This method integrates three key processes: positioning grains according to the depositional heuristics, generating cement or matrix backgrounds through a generative network, and seamlessly blending grains into the generated textures while creating corresponding annotation masks.

The depositional heuristics provide precise control over the listed parameters: background selection, grain class distribution, grain density, local proximity relationships, and grain orientation. For dataset generation, grain class distributions were randomly sampled for each synthetic composition, with grain density varying between 0% and around +65%. Background selection followed geological principles: matrix backgrounds were used exclusively for compositions with grain densities < 50% (representing mud-supported classes of the Dunham classification), while both matrix and cement backgrounds were available for higher densities above this threshold (representing grain-supported classes). Local proximity parameters were randomly assigned, incorporating both preferential and fully random grain orientations throughout the dataset construction process.

5.3 Methods

Transfer learning represents the most conventional approach for training DL models when annotated datasets are limited (Zhuang et al., 2019). This technique leverages knowledge from a source domain by adapting pre-trained models to target distributions that are assumed to be closely aligned with the intended task. In practice, this may involve fine-tuning models initially trained on synthetic data using real-world examples (Ding et al., 2023; Peng et al., 2018; Ros et al., 2016; Xiao et al., 2021). However, synthetic data frequently fails to capture the complete complexity of real-world distributions, as certain realistic content remains challenging to reproduce accurately, ultimately limiting the model's generalization capability.

To address these limitations, domain adaptation techniques have emerged to align features and input data from different distributions into a shared representation space (Farahani et al., 2020).

5.3.1 Unsupervised Domain Adaptation

Unsupervised Domain Adaptation (UDA) addresses the fundamental challenge of transferring knowledge from a source to a target domain to maintain model performance (Liu et al., 2022). In computer vision applications, this approach aims to bridge variations in image characteristics without requiring new labels, making it a highly practical approach for real-world scenarios where annotation is expensive or unfeasible to acquire. UDA research in computer vision encompasses various DL approaches that can be categorized into several paradigms based on their adaptation mechanisms.

Adversarial learning-based methods have emerged as a dominant approach, inspired by Generative Adversarial Networks (GANs) (Goodfellow et al., 2014). These methods typically employ a domain discriminator to achieve feature-alignment across various input domains (Ganin et al., 2015; Tzeng et al., 2017). Similar objectives can also be achieved through contrastive methods that maximize similarity between representation of semantically similar inputs from different domains while minimizing similarity between cross-domain inputs (Thota & Leontidis, 2021). While these approaches could potentially generate similar feature distributions between synthetic and real images for further semantic segmentation tasks, they remain unexplored in this study. Indeed, adversarial approaches that achieve similar feature distributions across domains do not explicitly guarantee improved performance on the primary task (e.g., segmentation or classification). Furthermore, contrastive methods would require establishing similar and dissimilar pairs between synthetic and real images, which would be impractical to construct within reasonable timeframes.

5.3.2 From Unsupervised Domain Adaptation to Semi-Supervised Image Translation

Moving towards alternative paradigms, generative approaches have proposed adapting the input data itself rather than modifying learned features or parameters.

Style transfer represents a fundamental approach that preserves image content while adopting the style characteristics of another image (Gatys et al., 2015a). This method separates and recombines content and style features using a pre-trained VGG-19 network (Simonyan & Zisserman, 2014), minimizing a loss function that balances content reconstruction with style representation through Gram matrices that capture texture and style information. While producing visually compelling results, this method has limitations as it requires individual optimization processes for each image pair to be made up. Despite these constraints, style transfer remains promising, with subsequent approaches directly modifying source content features to match target style statistics using Adaptive Instance Normalization (AdaIN) (Karras et al., 2018). This advancement has notably allowed arbitrary style transfer between pairs of images coming from different domains (Huang & Belongie, 2017) and guided a part of the proposed approach.

CycleGAN represents another significant method, employing image-to-image translation to generate target domain images resembling the source domain without requiring paired training data (Zhu et al., 2017). The framework consists of two generators and two discriminators. The first generator maps images from one domain to another, while the second generator performs the inverse operation. The discriminators are trained to distinguish between real images from their respective domains and fake images generated by the opposing generator. Built on a straightforward concept, this method offers fully unsupervised learning using two sets of unlabeled images with shared characteristics. Although this architecture provides advantages through a cycle-consistency loss ensuring that translating an image to another domain and back reconstructs the original, semantic consistency lacks explicit formulation and may result in undesired texture mapping in certain regions of transformed images (Cherian & Sullivan, 2019).

Recent studies addressed the problem of semantic consistency in CycleGAN by incorporating a classification or segmentation tasks to preserve label information during translation with Cycle-Consistent Adversarial Domain Adaptation (CyCADA) (Hoffman et al., 2017). This approach enables effective adaptation to new domains without requiring labeled target data, making it useful for image classification and segmentation in diverse environments. However, this method is constrained by generating only a single adapted version per image learned during training. Consequently, producing multiple variations from individual synthetic images offers opportunities to enhance output diversity.

5.3.3 Semi-Supervised Adaptive Image Translation with Semantic Consistency

This study explores the potential of image-to-image translation from synthetic to real domain for downstream real image segmentation. In relation to the previously mentioned challenges, this approach is primarily built to preserve semantic content while incorporating diverse visual styles during training to achieve enhanced segmentation.

5.3.3.1 Network Components and Formal Training Process

Given $\{X_s, Y_s\}$ pairs of examples with corresponding labels from the source domain (synthetic) and X_t examples from the target domain (real), an autoencoder composed of an encoder E and decoder D is first trained on both domains (Equation 5.1).

$$\hat{X} = D(E(X)) , X \in \{X_s, X_t\} \text{ minimizing } L = \|\hat{X} - X\|^2 \quad (5.1)$$

Once trained, the encoder E is frozen and integrated into the global framework to ensure a common latent representation of both source and target domains.

The unified network comprises two generators, each consisting of the shared encoder E , paired with domain-specific decoders D_{st} (source to target) and D_{ts} (target to source). Following CycleGAN principles, the first generator learns to produce images from synthetic to real domain (Equation 5.2).

$$X_{st} = D_{st}(E(X_s)) \quad (5.2)$$

The second generator performs the inverse transformation to return to the source domain (Equation 5.3).

$$X'_s = D_{ts}(E(X_{st})) \quad (5.3)$$

Similarly to GANs, a discriminator $Disc_t$ is used to distinguish whether images X_{st} produced by the decoder D_{st} belong to the target distribution $p_t(X_t)$ (Equation 5.4).

$$p_{st}(X_{st}) = p_t(X_t) \quad (5.4)$$

Inspired by style transfer approach, Adaptive Instance Normalization (AdaIN) is integrated into the source-to-target transformation (Equation 5.5). This module accepts source content and target style inputs, aligning the channel-wise mean and standard deviation of the source domain to match the target domain (Equation 5.6), facilitating translation by providing target distribution indicators.

$$X_{st} = D_{st}[AdaIN(E(X_s), E(X_t))] \quad (5.5)$$

Where the AdaIN function is expressed as:

$$AdaIN(E(X_s), E(X_t)) = \sigma(E(X_t)) \cdot \left(\frac{E(X_s) - \mu(E(X_s))}{\sigma(E(X_s))} \right) + \mu(E(X_t)) \quad (5.6)$$

This method offers the opportunity to generate diverse adaptations for each synthetic image by using various real images as style inputs through training-time shuffling.

Ultimately, a segmentation model S is added to segment X_{st} , the transformed synthetic images, into proper labels (Equation 5.7).

$$Y'_s = S(X_{st}) \quad (5.7)$$

This approach directly benefits from producing a target domain segmentation model within the same compact network.

Except for the shared encoder E that has been frozen, all the new components are trained simultaneously combining multiple losses.

A cycle training policy is used to maintain generators consistency between image-to-image translations. Nevertheless, cycle consistency that is typically computed across entire images to preserve original input details can be applied to grain-containing regions only. Given that X_s , the synthetic image created by SynSection, which is a combination of grains and background with M the associated binary grain mask (Equation 5.8):

$$X_s = Grain \cdot M + Background \cdot (1 - M) \quad (5.8)$$

The grain consistency loss (Equation 5.9) regardless of background is formulated as:

$$\mathcal{L}_{grain} = \left\| (X_s - D_{st} \left(E \left(D_{ts} (E(X_s)) \right) \right)) \cdot M \right\|_2^2 \quad (5.9)$$

The implementation of this modification enables the model to generate enhanced textural variations in backgrounds, which typically demonstrate important variability in thin sections. Furthermore, this approach results in improved preservation of fine-grained details within the objects. Figure 5.4A provides direct illustration of these observations, with example A.3 displaying an image generated by a model trained with cycle consistency loss calculated across the entire image, while example A.4 presents an image produced by a model employing the same loss function applied exclusively to the grain-containing regions.

Since the framework uses adversarial training, the transformed examples X_{st} generated by the decoder D_{st} try to fool the discriminator $Disc_t$ leading to specific generator and discriminator losses (Equation 5.10 and Equation 5.11).

$$\mathcal{L}_{Generator} = E_{X_s \sim p_s(X_s)} [\log(1 - Disc_t(X_{st}))] \quad (5.10)$$

$$\mathcal{L}_{Discriminator} = E_{X_t \sim p_t(X_t)} [\log Disc_t(X_t)] + E_{X_s \sim p_s(X_s)} [\log(1 - Disc_t(X_{st}))] \quad (5.11)$$

The above E denotes the expectation operator which calculates the average value of a function across the entire data distribution and $p_x(X_x)$ the specific domain distribution.

To ensure informative target features for the decoder D_{st} , transforming synthetic-to-real images, generated images X_{st} are passed through the shared encoder E whose resulting features must match the statistics (mean and standard deviation) of conditioning target features, inspired from the method introduced with arbitrary style transfer (Equation 5.12).

$$\mathcal{L}_{style} = \|Style(X_t) - Style(X_{st})\|_2^2 \quad (5.12)$$

Where the style function (Equation 5.13) is defined as:

$$Style(X) = [\mu(E(X)), \sigma^2(E(X))] \quad (5.13)$$

Ultimately, the segmentation loss (Equation 5.14) measures the difference between the true probability Y_{s_i} and predicted probability distributions X_{st_i} of k multiple classes, where i represents the i -th class:

$$\mathcal{L}_{Seg} = - \sum_{i=1}^k Y_{s_i} \cdot \log(S(X_{st_i})) \quad (5.14)$$

This approach, in addition to directly providing a segmentation model in the target domain, formulates a constraint that prevents the generation of undesired artifacts. Figure 5.4B demonstrates results obtained without the segmentation constraint, where example B.3 shows newly generated grains within the cement, whereas example B.4 illustrates muddy heterogeneities incorporated within the cement without the creation of new grains.

Impact of Training

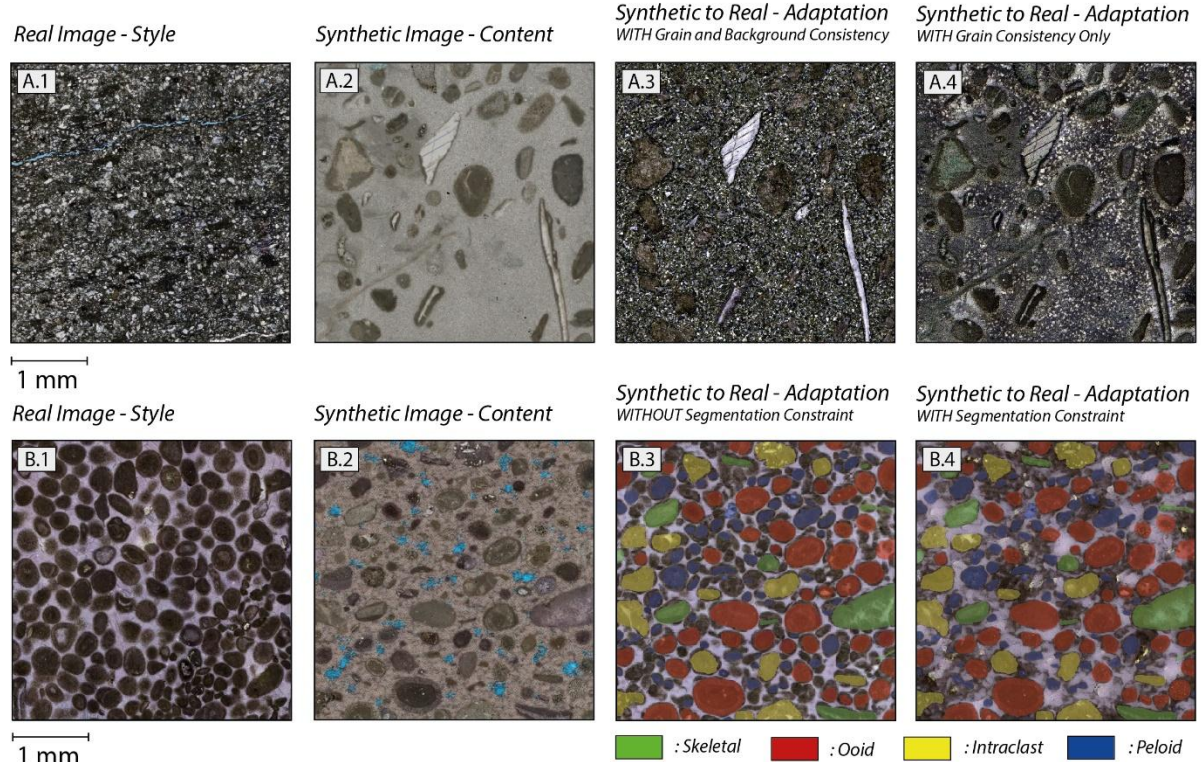


Figure 5.4 Visualization of constraint effects in image translation networks for deep learning applications in carbonate petrography.

5.3.3.2 Model Architecture and Training Hyperparameters

The global architecture presented in Figure 5.5 first presents the shared encoder built from five sequential downsampling convolutional layers. These layers use progressively increasing filter counts of [64, 128, 128, 256, 256] channels, with corresponding kernel sizes [7×7, 5×5, 5×5, 3×3, 3×3], and strides [1, 2, 1, 2, 1] respectively. To minimize undesired border effects that typically arise from zero padding, reflection padding is applied throughout the encoder. Between each convolutional layer, Leaky ReLU activation functions are employed with an alpha parameter of 0.2 allowing small negative gradients. The feature extractor is completed by two additional residual blocks, each containing 256 filters with 3×3 kernels and unit stride. These blocks maintain the same padding and activation function as the previous layers.

Global Model Architecture

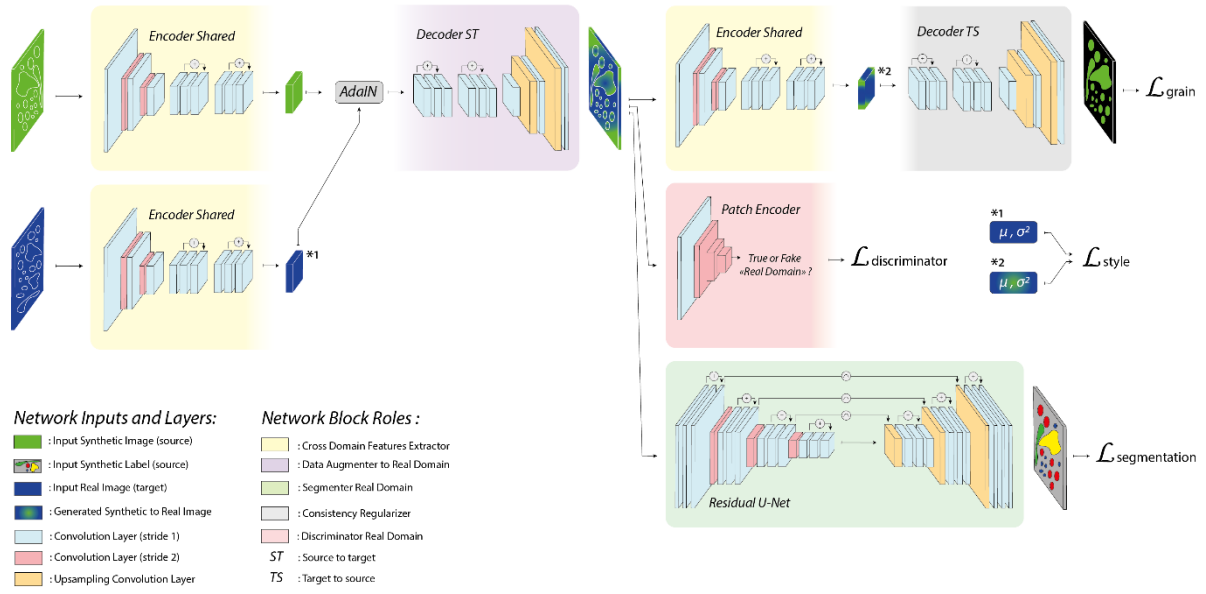


Figure 5.5 Network architecture for synthetic-to-real domain translation of carbonate thin section images with stylistic and semantic consistency.

The decoder sections mirror the encoder architecture but replace downsampling operations with upsampling convolutions. This design choice specifically addresses the checkerboard artifact problem commonly associated with transposed convolutions (Odena et al., 2016). Notably, normalization layers are deliberately omitted from the encoder-decoder architecture since previous work has demonstrated that such layers can significantly impair style transfer performance (Huang & Belongie, 2017). The final layer of each decoder consists of a single convolutional layer with three output filters, unit kernel size and stride, followed by sigmoid activation to produce valid RGB image outputs.

The discriminator network implements a fully convolutional design that progressively increases feature depth through layers containing [64, 128, 256, 256] filters respectively. All convolutional operations use 3×3 kernels with stride-2 downsampling, while no padding operations are performed. LeakyReLU activations with alpha 0.2 are used between each convolutional layer. The final classification layer employs a single filter with unit kernel size and stride, followed by sigmoid activation for binary classification output. This architectural approach significantly reduces the number of trainable parameters compared to traditional discriminators that rely on fully connected final layers. More importantly, the design functions as a texture discriminator rather than evaluating global image realism (Isola et al., 2016), with overall consistency maintained through cycle loss functions detailed in the training methodology.

The segmentation model utilizes standard residual U-Net architecture (Zhang et al., 2017), chosen based on findings from SynSection instigations indicating that network architecture plays a secondary role compared to input data distribution quality. The encoder pathway consists of convolutional blocks with a sequence filters [64, 128, 256, 256], all using 3×3 kernels. Each block incorporates LeakyReLU activations and intermediate instance

normalization layers to stabilize training dynamics. Spatial downsampling between encoder blocks is achieved through max-pooling operations, while the decoder pathway employs upsampled convolutions for feature map expansion. Critical to the U-Net design, skip connections directly link corresponding feature block levels between encoder and decoder part, enabling the network to preserve fine-grained spatial information.

To address the imbalance between real and synthetic dataset sizes, dynamic duplication of real images is implemented alongside traditional data augmentation techniques. The augmentation pipeline includes random horizontal and vertical flips, 90-degree rotations, and subtle modifications to brightness and contrast parameters to increase data diversity.

The complete model training process spans 12,500 batches over 50 iterations, with each batch containing 4 pairs of synthetic images with their corresponding annotation masks and four real images. The Adam optimizer is employed across all network components, though with different learning rate schedules. For the decoders and the discriminator, an exponential decay strategy reduces the learning rate from 5×10^{-5} to 1×10^{-5} . The segmentation network uses the same optimizer but with a higher initial learning rate of 1×10^{-4} , decaying to the same final value of 1×10^{-5} .

The training process employs task-specific loss functions optimized for each network component's objectives. Cycle grain consistency is enforced through Mean Absolute Error (MAE) and Structural SIMilarity (SSIM) (Nilsson & Akenine-Möller, 2020), while style consistency relies on Mean Squared Error (MSE). The segmentation task utilizes weighted Categorical Cross-Entropy (CCE) combined with Dice loss (Sudre et al., 2017), where the weighting incorporates a distance map that specifically penalizes regions containing closely positioned objects. This approach encourages individual grain separation during training (Ronneberger et al., 2015). The generator and discriminator training employ least squares loss, following the original CycleGAN idea to reduce vanishing gradient risks compared to standard cross-entropy approaches (Zhu et al., 2017). During backpropagation, discriminator weights are updated using only the discriminator-specific loss term while generator, grain consistency, style, and segmentation losses are accumulated across the decoders and the segmentation network before weight updates. Indeed, the decoders also receive updates based on the segmentation network for maintaining semantic consistency objective throughout the training process. The different loss terms are balanced using various weighting parameters λ that have been identified optimal for this use case. These lasts are either fixed or dynamically adjusted from start to end of the training: $\lambda_{\text{generator}} = 2.5$; $\lambda_{\text{grain}} = 7.5$; $\lambda_{\text{style}} = 10^2 \rightarrow 10^3$; $\lambda_{\text{segmentation}} = 1.0 \rightarrow 3.0$; $\lambda_{\text{discriminator}} = 3.5$. The detailed training loop is presented in Algorithm 5.1.

Algorithm 1: Adaptive Image Translation via Adversarial Learning and Segmentation for Semantic Consistency

Inputs : X_s /* Synthetic domain images */
 Y_s /* Synthetic domain segmentation masks */
 M /* Synthetic domain binary grain masks */
 X_t /* Real domain images */
 w_{edt} /* Euclidian Distance Transform penalty maps */

Outputs: X_{st} /* Synthetic images translated to real domain */
 Y'_s /* Cycled segmentation masks */

Initialize networks:

$E \leftarrow$ Encoder weights from pre-trained autoencoder

$D_{st}, D_{ts}, Disc_t, S \leftarrow$ Random weights

Initilize weighting terms:

$\lambda_1 = 2.5$; $\lambda_2 = 7.5$; $\lambda_3 = 10^2 \rightarrow 10^3$; $\lambda_4 = 1.0 \rightarrow 3.0$; $\lambda_5 = 3.5$

while not converged do

Step 1: Image Translation

$X_{st} \leftarrow D_{st}(\text{AdaIN}(E(X_s), E(X_t)))$ /* Transform synthetic images */

Step 2: Consistency Cycle

$X'_s \leftarrow D_{ts}(E(X_{st}))$ /* Transform back */

Step 3: Segmentation Cycle

$Y'_s \leftarrow S_R(X_{st})$ /* Predict masks from transformed images */

Step 4: Discriminator Evaluation

$Z_{fake.t} \leftarrow Disc_t(X_{st})$

$Z_{true.t} \leftarrow Disc_t(X_t)$

Step 5: Discriminator Loss

 Compute loss then update discriminator weights

$Loss_{discriminator} \leftarrow \frac{1}{2} \left((Z_{true.t} - 1)^2 + Z_{fake.t}^2 \right) \cdot \lambda_5$

Step 6: Generator Losses

 Compute losses then update decoders and segmenter weights

$\mathcal{L}_{generator} \leftarrow \frac{1}{2} (Z_{fake.r} - 1)^2$

$\mathcal{L}_{grain} \leftarrow \text{MAE}(X_s, X'_s) \cdot M + \frac{1}{2} \text{SSIM}(X_s, X'_s) \cdot M$

$\mathcal{L}_{style} \leftarrow \text{MSE}((\mu(E(X_s)), \sigma^2(E(X_s))), (\mu(E(X_{st})), \sigma^2(E(X_{st}))))$

$\mathcal{L}_{segmentation} \leftarrow w_{edt} \cdot \text{CCE}(Y_s, Y'_s) + \text{DICE}(Y_s, Y'_s)$

$\mathcal{L}_{total} \leftarrow \lambda_1 \cdot \mathcal{L}_{generator} + \lambda_2 \cdot \mathcal{L}_{content} + \lambda_3 \cdot \mathcal{L}_{style} + \lambda_4 \cdot \mathcal{L}_{segmentation}$

end

return X_{st}, Y'_s

Algorithm 5.1 Unified adversarial training loop combining domain adaptation, cycle consistency, stye conditioning and segmentation for synthetic-to-real carbonate thin section image characterization.

5.3.3.3 Model Evaluation

The network architecture's performance was evaluated by continuously monitoring losses variations and using comprehensive metrics tailored to each objective task, as detailed in the following sections. The primary goal of this study is to determine whether the proposed method achieves superior petrographic characterization compared to the baseline SynSection approach, which involves initial training on synthetic data followed by fine-tuning on limited

real examples. To ensure fair comparison, the baseline was reproduced using the same segmentation architecture as that employed in the current unified framework. Given the small validation dataset of 100 examples, 5-fold cross-validation was performed on the baseline method, allocating 70% of each fold for fine-tuning and 30% for testing, with performance metrics averaged across all folds. In contrast, the segmentation component of the proposed model was evaluated directly on all 100 validation images.

5.3.3.3.1 Image Quality and Features Distribution

The evaluation framework employs the Frechet Inception Distance (FID) (Heusel et al., 2017) to assess image quality by measuring the distributional similarity between real and generated images using features extracted from a pre-trained Inception-v3 model (Szegedy et al., 2015). Lower FID scores indicate better distributional alignment between real and generated images. To further analyze domain adaptation effectiveness, the latent space distributions of three image sets are examined: domain-translated synthetic images, original synthetic images, and original real images. Through this analysis, it is determined whether better alignment with real image distributions is achieved by the adapted synthetic images. Features are extracted from the pre-trained Inception-v3 model at the first fully connected layer, then t-SNE dimensionality reduction is applied to visualize the feature distributions in a low-dimensional probabilistic space. A qualitative assessment is also conducted to evaluate the realism of the translated images and verify that the conditioning input style is effectively transferred to the output.

5.3.3.3.2 Image Segmentation and Petrographic Metrics

Conventional accuracy (Equation 5.15), Intersection over Union (IoU) (Equation 5.16), and Dice Coefficient (DC) (Equation 5.17) were calculated and completed with specific petrographic metrics such as point counting, mean grain size and grain sorting determination score (R^2) that are directly derived from the segmentation mask. TP, TN, FP, FN respectively refers to True Positive, True Negative, False Positive and False Negative in Equation 5.15, while y indicates the ground truth petrographic measurements and $\hat{y}$ the values extracted from the predictions in Equation 5.18.

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \quad (5.15)$$

$$IoU = \frac{Area\ of\ Overlap}{Area\ of\ Union} \quad (5.16)$$

$$Dice\ Coefficient = \frac{2 \times Area\ of\ Overlap}{Total\ Area} \quad (5.17)$$

$$R^2 = 1 - \frac{\sum_{i=1}^n (y_i - \hat{y}_i)^2}{\sum_{i=1}^n (y_i - \bar{y})^2} \quad (5.18)$$

The point counting metric assesses the model's ability to identify individual objects through a two-step process: grain probability masks are thresholded using Otsu's method (Otsu, 1979), followed by application of a connected components algorithm (He et al., 2017) to isolate individual objects. This approach enables thorough performance assessment for first-scale petrographic analysis.

Complementary information extracted from this process includes evaluation of mean grain size and sorting (Blott & Pye, 2001; Folk, 1954, 1966), both derived from grain size distribution and serving as key indicators for sedimentological interpretations (e.g., hydrodynamic depositional conditions). These measures provide relevant feedback regarding practical applicability for specialist interpreters. Traditionally, the mean grain size is defined as the sum of maximum grain diameters D weighted by their mass w , over the total mass of all grains (Equation 5.19), while sorting is calculated as a sum of weighted differences between distribution quantiles where diameters in millimeters (also weighted by mass) are transformed to $\log_2$ base (ϕ scale) (Equation 5.20). Corresponding descriptive terminology resorting from these measurements are provided with Table 5.1. The previous equations were adapted for image-based grain size analysis using pixel-level operations (Equation 5.21 and Equation 5.22). Since grain diameters are typically weighted by mass in conventional analysis, pixel area is proposed as a proxy, with area values A being assigned to each pixel p within isolated grain regions G . Statistical characterization can thus be performed directly on the image's scalar values, with background pixels (zero values) being excluded from the analysis.

$$\bar{D} = \frac{\sum_{i=1}^N D_i \cdot w_i}{\sum_{i=1}^N w_i} \quad (5.19)$$

$$\text{Sorting} = \frac{\phi_{84} - \phi_{16}}{4} + \frac{\phi_{95} - \phi_5}{6.6} \quad (5.20)$$

$$\overline{D}_{\text{image}} = \frac{\sum_{p \in G} D_p \cdot A_p}{\sum_{p \in G} A_p} \quad (5.21)$$

$$\text{Sorting}_{\text{image}} = \frac{\phi_{84}^A(G) - \phi_{16}^A(G)}{4} + \frac{\phi_{95}^A(G) - \phi_5^A(G)}{6.6} \quad (5.22)$$

$$\phi_q^A = -\log_2(D_q^A) \quad (5.23)$$

With:

D_p = grain diameter at pixel p

A_p = pixel area (used as weight, analogous to mass in traditional method)

G = set of all grain pixels (excluding background pixels with zero values)

$q \in \{5, 16, 84, 95\}$ for the respective percentiles

$\phi_q^A(G)$ = area-weighted q -th percentile of grain diameters in ϕ scale within grain pixel set G

$D_q^A(G)$ = quantile $_q(\{D_p: p \in G, A_p > 0\}, \text{weights} = \{A_p: p \in G, A_p > 0\})$

Grain Size Statistics			
Grain Type	mGrain Size (μm)	Sorting Type	Grain Sorting
Medium to very coarse calcirudite	> 4,000	Very well sorted	0.35
Fine calcirudite	1000–4,000	Well sorted	0.35–0.50
Coarse calcarenite	500–1,000	Moderately well sorted	0.50–0.70
Medium calcarenite	250–500	Moderately sorted	0.70–1.00
Fine calcarenite	125–250	Poorly sorted	1.00–2.00
Very fine calcarenite	63–125	Very poorly sorted	2.00–4.00
Calcilutite	< 63	Extremely poorly sorted	> 4.0

Table 5.1 Grain size with corresponding sorting classification and statistical parameters for sedimentary analysis.

However, image segmentation using deep learning is subject to inherent errors, and its application to carbonate petrography presents specific challenges. Grains observed within thin sections are often close or in contact with each other. Size measurements are therefore directly impacted when predicted masks lack well-defined boundaries between originally separated elements, or even when over detection of non-existing grains as well as misdetection are noticed.

To address this issue, local thickness (LT) that is conventionally used in porous media analysis (Gostick et al., 2019) is proposed as an alternative for estimating mean grain size and grain sorting, since carbonate elements of this dataset are mainly compact. Local thickness is defined at any point of a particle as the radius of the largest circle that fits within the particle and contains that point. Formally, for an object $O \subset \mathbb{R}^2$ and a point $x \in O$ the local thickness (LT) is given by Equation 5.24 where $S(c, r)$ is a closed circle centered at c with radius r . It is common to extend LT by admitting that $LT(x) = 0$ for x outside O (Dahl & Dahl, 2023).

$$LT(x) = \max\{r \in \mathbb{R} : x \in S(c, r) \subseteq O, c \in \mathbb{R}^2\} \quad (5.24)$$

To assess the benefits of using local thickness instead of traditional measurements, a theoretical application is conducted on the set of real annotated masks (100). First, the mean grain size and the grain sorting are computed for each example using both traditional and local thickness methods. The operation is then repeated with random transformations applied to annotated masks to simulate commonly observed prediction errors. These augmentations comprise positive noise additions, morphological operations (dilation, erosion), and slight distortions. The computed features are then compared between methods to evaluate sensitivity to errors.

The results presented in Figure 5.6 demonstrate that traditional methods are significantly affected by applied transformations on particle masks. Connected grains and newly appeared small objects introduce measurement bias where the mean grain size evolved from 985.11 μm to 1141.7 μm and grain sorting increased from 1.27 to 2.64 (Figure 5.6A and 5.6B). Conversely, local thickness measurements show consistent results between original and transformed masks: the mean grain size altered minimally from 755.27 μm to 734.21 μm and sorting changed from 1.49 to 1.53 (Figure 5.6C and 5.6D).

When extended to the entire validation set, similar observations are made. Mean grain size and sorting show poor correlation between original and augmented masks under transformation, with R^2 scores dropping below 0.5 (Figures 5.6E and 5.6G). In contrast, local thickness measurements remain more stable across transformations, preserving determination scores above the 0.5 reliability benchmark (Figures 5.6F and 5.6H).

This experiment demonstrates that grain size measurements using local thickness are less susceptible to segmentation noise due to the method's tendency to create distinct value gaps at grain boundaries. Consequently, local thickness measurement is adopted for all subsequent analyses in this study.

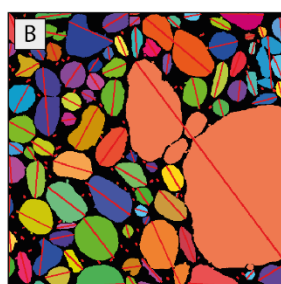
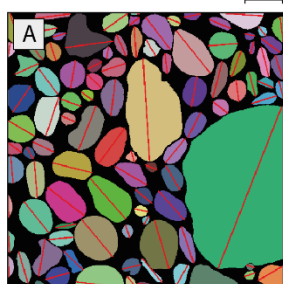
Traditional Grain Measurements

Single-sample comparison

Original Mask

500 μm

Transformed Mask



mGrain Size = 985.11 μm

Grain Sorting = 1.27

mGrain Size = 1141.7 μm

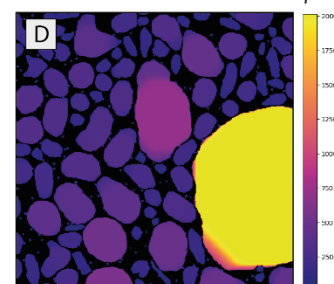
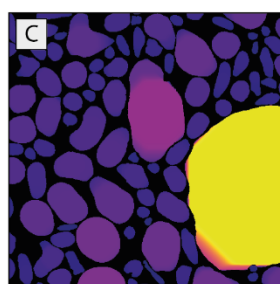
Grain Sorting = 2.64

Local Thickness Measurements

Original Mask

Transformed Mask

μm



mGrain Size = 755.27 μm

Grain Sorting = 1.49

mGrain Size = 734.21 μm

Grain Sorting = 1.53

Comparison across 100 samples

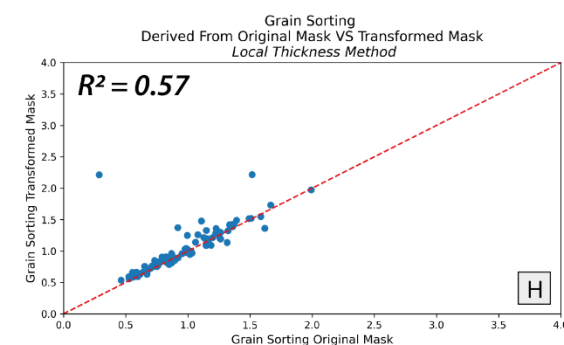
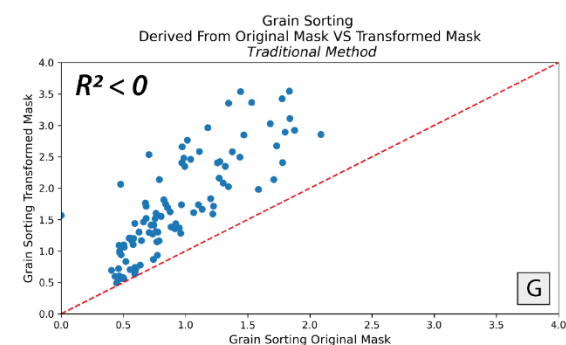
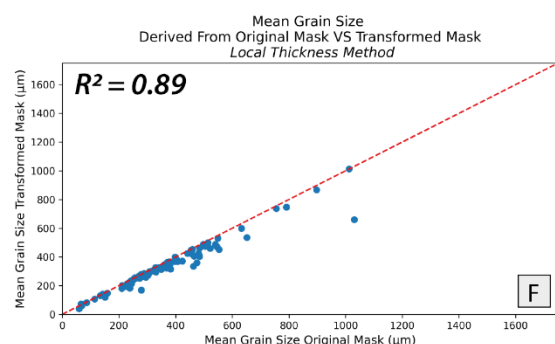
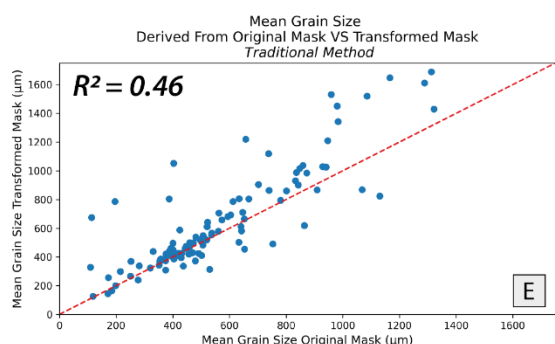


Figure 5.6 Sensitivity analysis of grain size statistics to segmentation errors: comparison of traditional diameter measurements versus local thickness methods.

To fully evaluate segmentation performance, the local image characterization method is proposed to be applied to complete thin section samples. This approach aims to provide better insights into the practical potential of image-assisted characterization techniques.

5.4 Results

The training was conducted until the losses plateaued with complementary qualitative evaluation of translated image quality.

5.4.1 Image Quality and Features Analysis

A FID of 86.12 was obtained for synthetic images without domain adaptation, which can be compared to the improved score of 18.69 achieved for images that underwent adaptation. The latent space distribution resulting from the Inception-v3 model of the utilized datasets is illustrated through various t-SNE visualizations (Figure 5.7). In Figure 5.7A, the distribution of real images (green) is contrasted with conventional synthetic images (red), where it can be observed that synthetic images are predominantly positioned at the periphery of the real image distribution, with only minimal overlap between synthetic examples and the real domain.

When translated synthetic images using random real samples as style inputs were incorporated, as shown in Figure 5.7B, the real domain was found to be extensively covered by the newly generated images (orange), with only limited regions remaining uncovered.

Furthermore, a visual prior is known for the real image dataset, which consists primarily of three recurring example types: alizarine ferricyanide-stained grainstone (Figure 5.7C.1), unstained grainstone (Figure 5.7C.2) and muddy limestones (Figure 5.7C.3). A final case is demonstrated in Figure 5.7C, where fewer than ten images of these three types were selected as reference style inputs. Geological consistency was maintained by applying real muddy limestone styles to synthetic muddy limestones, with the same methodology being reproduced for the grainstones. As a result of this approach, it is observed that the three defined input styles, when applied to the original synthetic images, enabled the new synthetic distribution to achieve substantial overlap with the real image distribution.

Complementary to this quantitative and probabilistic analysis, some examples of translated images are presented in Figure 5.8 with reference style and source content inputs.

t-SNE Representations

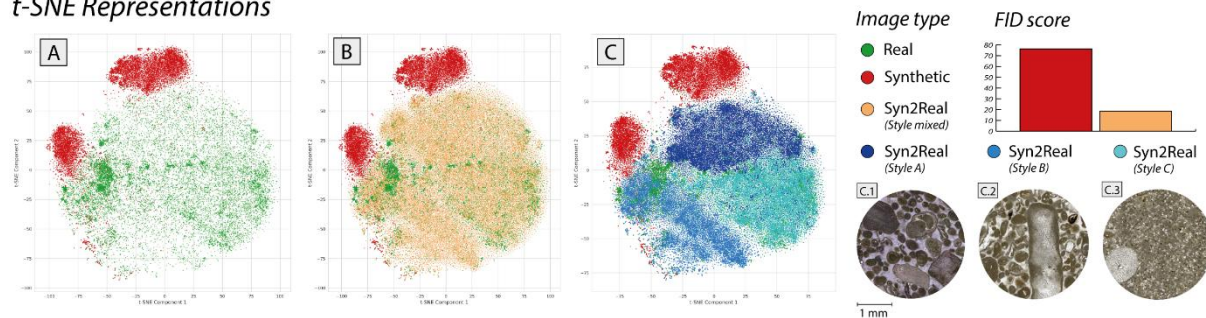
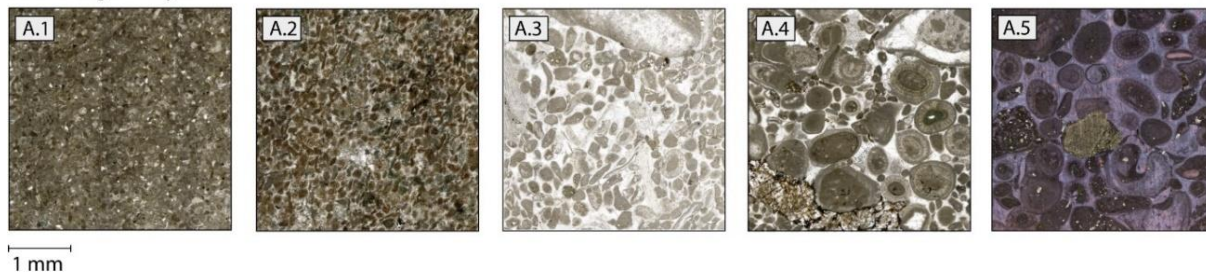


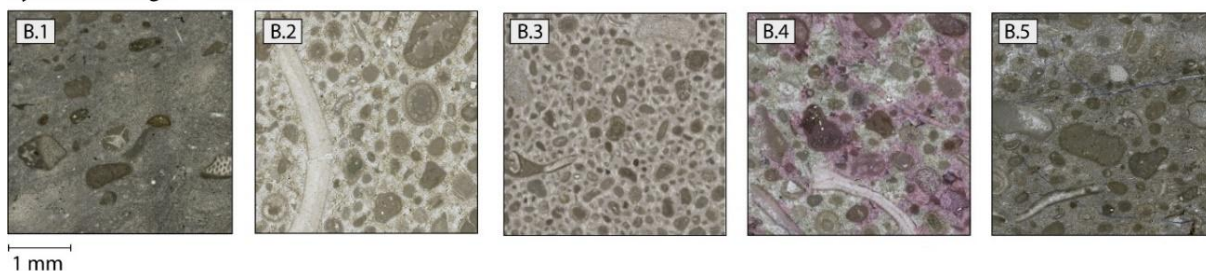
Figure 5.7 Latent space analysis comparing synthetic, domain-adapted, and real image distributions. FID scores between non-adapted and adapted images are shown via histogram. The prior distribution and principal reference style used as data partitioning are also displayed.

Image Translation Results

Real Images - Style



Synthetic Images - Content



Synthetic to Real Images - Translation

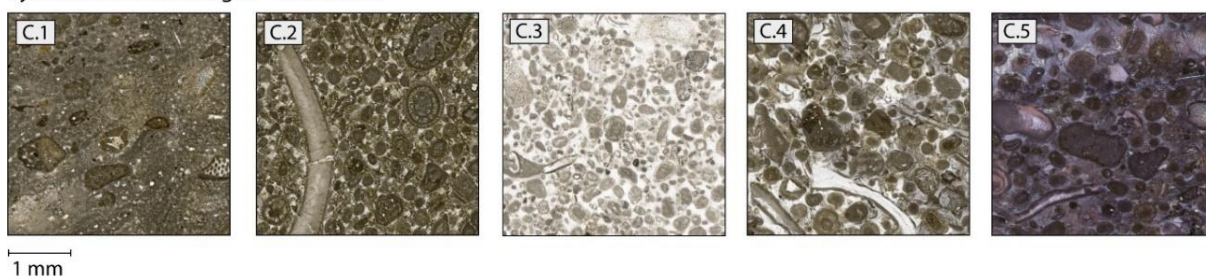


Figure 5.8 Qualitative comparison of synthetic image adaptation using real images as style references.

5.4.2 Image Segmentation and Practical Petrographic Performance

The quantitative results demonstrate the effectiveness of the unified network in performing segmentation within target domain across multiple evaluation metrics. Table 5.2 presents a comprehensive comparison between the baseline method and the newly proposed approach.

Method	mAccuracy	mDC	mIoU	Point Counting R ²	mGrain Size R ²	Grain Sorting R ²
Baseline	0.64	0.63	0.50	0.80	0.66	0.66
With Domain Adaptation	0.68	0.68	0.55	0.85	0.78	0.72
Improvement	+0.04	+0.05	+0.04	+0.05	+0.12	+0.06

Table 5.2 Performance comparison of baseline segmentation approach and domain adaptation method across segmentation and petrographic measurement metrics.

Class	Baseline				With Domain Adaptation		
	Acc	DC	IoU	Δ	Acc	DC	IoU
Background	0.89	0.85	0.75	↑	0.94	0.87	0.77
Grain	0.80	0.80	0.71	↑	0.82	0.83	0.73
Skeletal	0.57	0.57	0.45	↑	0.64	0.63	0.49
Intraclast	0.61	0.55	0.42	↑	0.67	0.66	0.53
Ooid	0.64	0.63	0.51	↑	0.68	0.70	0.57
Peloid	0.49	0.54	0.39	≈	0.48	0.53	0.40

Table 5.3 Class-wise segmentation performance comparison between baseline and domain adaptation approaches.

The baseline method achieved moderate performance with an accuracy of 0.64, DC of 0.63, and IoU of 0.50. For the regression tasks, the baseline showed reasonable performance with R² values of 0.80 for point counting, 0.66 for grain size estimation, and 0.67 for grain sorting analysis.

The incorporation of domain adaptation led to consistent improvements across all evaluated metrics. The enhanced model achieved an improvement in accuracy to 0.68 (+0.04). Similarly, the DC increased to 0.68 (+0.05) and IoU improved to 0.55 (+0.04), indicating better segmentation quality and spatial overlap with ground truth annotations.

The most substantial improvements were observed in the petrographic regression tasks derived from predicted masks. Point counting performance increased from R² = 0.80 to R² = 0.85 (+0.05). Mean grain size estimation demonstrated the largest improvement, with R² increasing from 0.66 to 0.78 (+0.12), suggesting that domain adaptation particularly benefits fine-grained morpho-metrical analysis. Grain sorting analysis also showed improvement, with R² increasing from 0.66 to 0.72 (+0.06).

Table 5.3 compares the performance of the baseline method and domain adaptation approach across six distinct classes: background, grain, skeletal, intraclast, ooid, and peloid. The grain category encompasses all granular components regardless of their specific classification, providing a comprehensive evaluation of general grain segmentation performance. This class was excluded from the previously presented average metrics. Moreover, this approach proves particularly valuable because grain boundary delineation forms the foundation for morphometric characterization analyses, where accurate segmentation takes precedence over precise class attribution. Evaluating grain segmentation independently of class-specific identification allows for better isolation and understanding of segmentation errors, which primarily stem from inter-class confusion among different grain types rather than fundamental boundary detection failures. This strategy directly connects

segmentation metrics to subsequent quantitative analyses of grain counting, size distribution, and sorting parameters.

The results reveal varying degrees of improvement across the different categories, highlighting the selective benefits of domain adaptation.

The background class demonstrated strong baseline performance with accuracy, DC, and IoU values of 0.89, 0.85, and 0.75, respectively. Domain adaptation further enhanced these metrics to 0.94 (+0.05), 0.87 (+0.02), and 0.77(+0.02). This indicates that the method successfully maintains and enhances performance on the dominant background class.

The generic grain class showed robust baseline performance with accuracy of 0.80, DC of 0.80, and IoU of 0.71, while domain adaptation yielded consistent improvements across all metrics with resulting accuracy of 0.82 (+0.02), DC of 0.83 (+0.03), and IoU of 0.73 (+0.02).

The most substantial improvements were observed in classes that exhibited lower baseline performance.

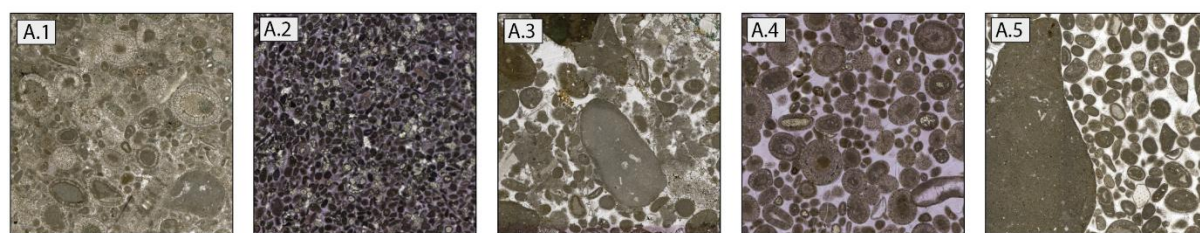
The skeletal class, initially achieving 0.57 in accuracy, 0.57 in DC, and 0.45 in IoU, showed marked improvement by training the segmentation network on adapted images reaching 0.64, 0.63, and 0.49 respectively (+0.07, +0.06, +0.04). Similarly, the intraclast category demonstrated significant enhancement from baseline values of 0.61 in accuracy, 0.55 in DC, and 0.42 IoU to 0.67 (+0.06), 0.66 (+0.11), and 0.53 (+0.11) respectively. The ooids identification also benefited from domain adaptation, with performance increasing from 0.64, 0.63, and 0.51 to 0.68 (+0.04), 0.70 (+0.07), and 0.57 (+0.06), showing respective improvements across accuracy, DC, and IoU metrics. However, the peloid class was observed to be the only instance where slight performance degradation occurred, with accuracy being minimally reduced from 0.49 to 0.48 (-0.01) and DC showing decline from 0.54 to 0.53 (-0.01). IoU was found to be marginally improved from 0.39 to 0.40 (+0.01).

Overall, domain adaptation demonstrated positive effects for five out of six classes. The minimal magnitude of the changes observed for the peloid category indicates that the domain adaptation method maintained relatively stable performance for this challenging class.

Figure 5.9 illustrates qualitative examples where the highest amplitude improvements in segmentation performance were observed between the baseline and current implementation. The validation images are presented in Figure 5.9A1 to 5.9A5, while the corresponding ground truth segmentation masks are shown in Figure 5.9B1 to 5.9B5. The baseline segmentation results are displayed in Figure 5.9C1 to 5.9C5, and the output masks obtained with the domain-adapted segmentation network are presented in Figure 5.9D1 to 5.9D5. The first example (first column) demonstrates enhanced overall segmentation where ooids that were previously misdetected by the baseline approach are now better identified through the new methodology. The second sample reveals improved results where peloids that had been incorrectly classified as ooids by the baseline method are now properly detected. The third and fourth examples exhibit more refined results where contours were better delineated, and grain boundaries were more effectively separated. The final validation image demonstrates superior segmentation performance where holes observed within grains are filled, contrasting with the incomplete results observed in the baseline approach.

Image Segmentation Results

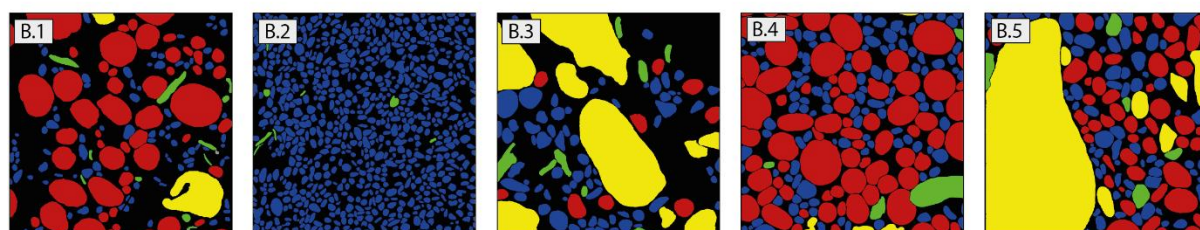
Real Validation Images



1 mm

Ground Truth

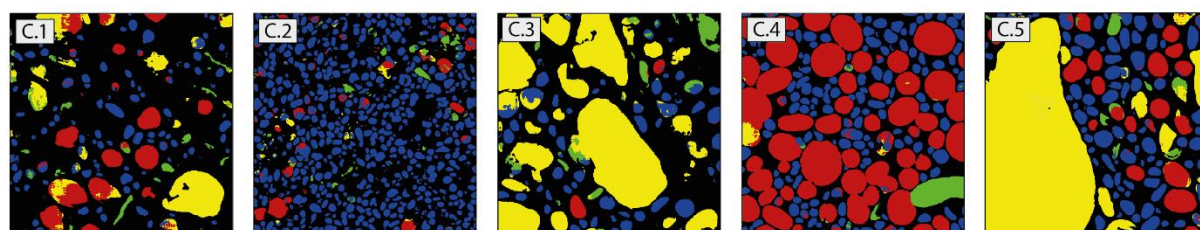
Grain Counting = 146 mGrain Size = 349.04 μm Grain Sorting = 1.12	Grain Counting = 635 mGrain Size = 73.70 μm Grain Sorting = 0.61	Grain Counting = 83 mGrain Size = 561.94 μm Grain Sorting = 1.25	Grain Counting = 126 mGrain Size = 312.20 μm Grain Sorting = 0.89	Grain Counting = 107 mGrain Size = 887.64 μm Grain Sorting = 1.47
--	---	---	--	--



Legend: : Skeletal (green), : Ooid (red), : Intraclast (yellow), : Peloid (blue)

Baseline

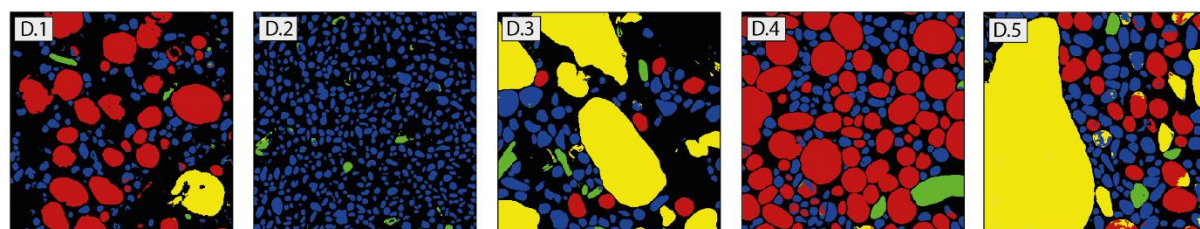
Grain Counting = 187 mGrain Size = 245.18 μm Grain Sorting = 1.21	Grain Counting = 654 mGrain Size = 82.8 μm Grain Sorting = 0.71	Grain Counting = 128 mGrain Size = 481.72 μm Grain Sorting = 1.30	Grain Counting = 183 mGrain Size = 318.50 μm Grain Sorting = 0.96	Grain Counting = 131 mGrain Size = 1465.22 μm Grain Sorting = 1.87
--	--	--	--	---



mAccuracy = 0.70 mDC = 0.69 mIoU = 0.53	mAccuracy = 0.64 mDC = 0.64 mIoU = 0.53	mAccuracy = 0.84 mDC = 0.84 mIoU = 0.73	mAccuracy = 0.94 mDC = 0.92 mIoU = 0.86	mAccuracy = 0.92 mDC = 0.89 mIoU = 0.8
---	---	---	---	--

With Domain Adaptation

Grain Counting = 174 mGrain Size = 277.58 μm Grain Sorting = 1.18	Grain Counting = 632 mGrain Size = 72.62 μm Grain Sorting = 0.69	Grain Counting = 126 mGrain Size = 546.12 μm Grain Sorting = 1.28	Grain Counting = 170 mGrain Size = 314.56 μm Grain Sorting = 0.94	Grain Counting = 126 mGrain Size = 1271.94 μm Grain Sorting = 1.38
--	---	--	--	---



mAccuracy = 0.85 mDC = 0.86 mIoU = 0.75	mAccuracy = 0.77 mDC = 0.77 mIoU = 0.63	mAccuracy = 0.95 mDC = 0.95 mIoU = 0.91	mAccuracy = 0.96 mDC = 0.95 mIoU = 0.91	mAccuracy = 0.95 mDC = 0.93 mIoU = 0.87
---	---	---	---	---

Figure 5.9 Comparative segmentation results between conventional transfer learning using synthetic images during pre-training then real data for fine-tuning and the proposed end-to-end training on domain-adapted images.

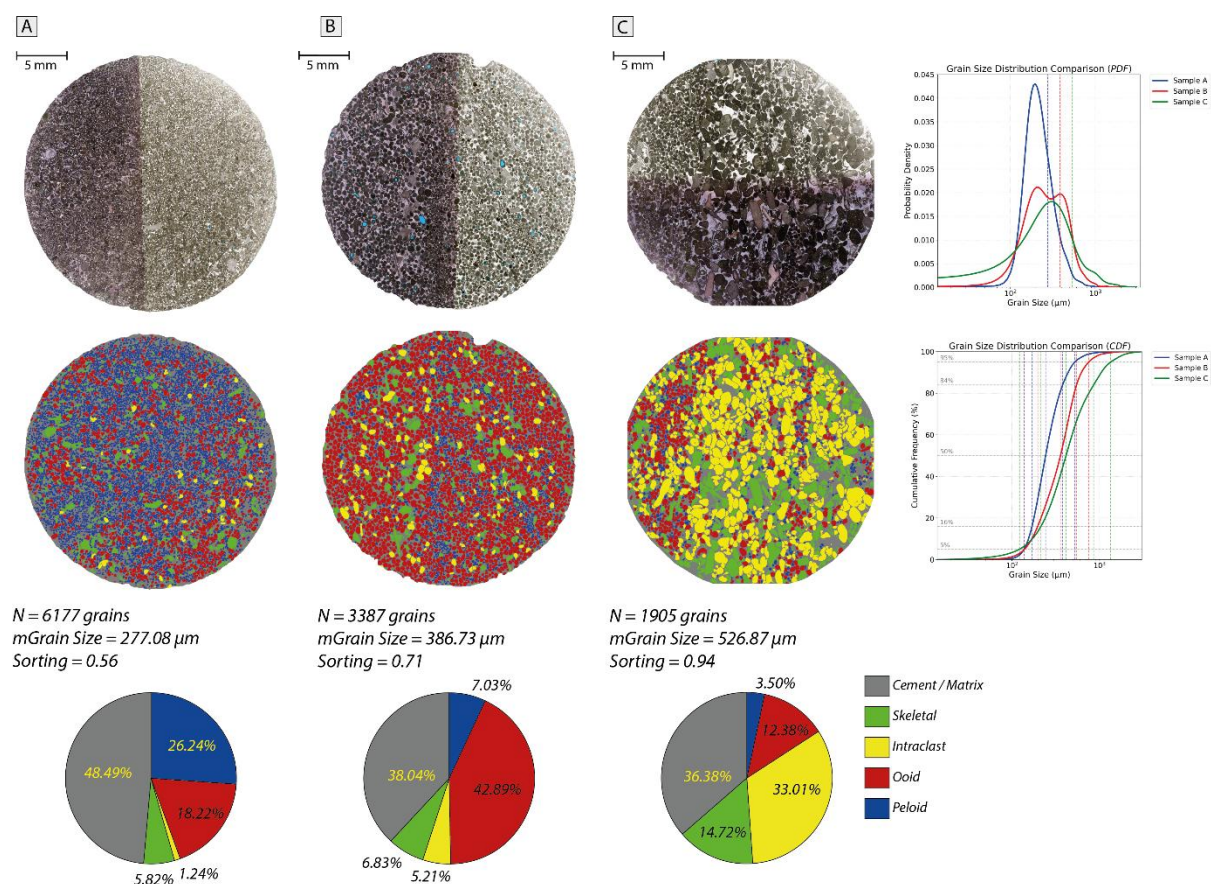


Figure 5.10 Full thin section segmentation examples showing grain composition distributions and grain size analysis with probability density and cumulative distribution function.

The extension of semantic segmentation to full thin sections for practical reliability and demonstration is shown in Figure 5.10, where three samples with distinct grain component distributions are characterized. This segmentation task utilized an overlapping tiles strategy with 25% of overlap. Predictions at patch intersections were averaged, and multi-inference was conducted using traditional augmentations (rotations, brightness and contrast adjustments). A post-processing addressed instances of multiple pixel class assignments within individual grains by applying the predominant label to each grain. Sample A exhibits a grain-supported framework (51.51%) predominantly composed of peloids (26.24%) and ooids (18.22%), with minor skeletal grains (5.82%) and rare intraclasts (1.24%). The mean grain size is classified as medium calcarenite (277.08 μm), with a moderately well-sorting (0.56). Sample B displays a grain-supported structure (61.96%) dominated by ooids (42.89%), accompanied by subordinate peloids (7.03%), skeletal elements (6.83%), and intraclasts (5.21%). The average particle size falls within the medium calcarenite category (386.73 μm), exhibiting moderately sorted characteristics (0.71). Sample C presents a grain-supported framework (63.62%) primarily constituted of intraclasts (33.01%), with moderate skeletal elements (14.72%) and ooids (12.38%), plus minor peloids (3.50%). The mean grain size is classified as coarse calcarenite (526.87 μm) with moderately sorted distribution (0.94). The probability density and cumulative density function of grain size parameters are additionally presented, enabling detailed comparative analysis between samples.

5.5 Discussions

5.5.1 Effectiveness of Domain Translation for Petrographic Analysis

The results demonstrate that the proposed method for synthetic-to-real domain adaptation enhances petrographic characterization compared to the baseline approach. The substantial improvement in FID scores indicates that the domain adaptation successfully participates in reducing the distributional gap between synthetic and real petrographic images. This improvement is useful in the context of carbonate thin section analysis, where visual realism proved to impact the reliability of assisted characterization.

The t-SNE visualization results provide evidence for the effectiveness of the domain adaptation strategy. The initial segregation between synthetic and real image distributions, with synthetic samples positioned at the limit of the real domain, highlights a fundamental challenge in applying synthetic training data to real-world scenarios (Tsai et al., 2018). The comprehensive coverage of the real domain mostly achieved by translated synthetic images demonstrates that the method successfully learns to map synthetic features into the target distribution space with limited remaining uncovered areas. Moreover, it was evidenced that specific regions of the target latent distribution can be reached when applying specific style inputs. This suggests that the approach preserves meaningful petrographic characteristics during translation and specific samples sets could be generated using prior criterion.

The segmentation performance improvements represent meaningful advances for practical applications. The +4 to +6% differential improvements in accuracy, DC, and IoU metrics translate into more precise grain class identification, which is fundamental for subsequent quantification. Similar performance improvements have been noticed in medical research using image translation techniques for nuclei segmentation where the shared problematic of detecting small compact object in heterogeneous textural environments is encountered (Haq et al., 2023; Takato, 2020; Trullo et al., 2022; Utz et al., 2023). The substantial enhancements observed for *skeletal* (+7% accuracy) and *intraclast* (+6% accuracy) categories suggest that domain adaptation particularly benefits to these classes that exhibit important morphological complexity or variability.

5.5.2 Grain Size Analysis Using Local Thickness

The adoption of LT measurements represents a promising and meaningful finding for assisted petrographic analysis as its use in porous media analysis is already known for pore size and geometry characterization (Ávila et al., 2022; Chiang et al., 2009; Gostick et al., 2019). The comparative analysis demonstrating LT's robustness against noisy segmentation addresses a limitation in deep learning-based grain diameter measurements derived from segmentation masks (Nichols et al., 2024). This benefit shows particularly valuable given the inherent challenges in carbonate petrography, where grain boundaries are often ambiguous and contact relationships complex. The natural weighting of LT measurements by grain surface area also aligns with traditional methods and visual interpretation practices, where larger particles better influence overall semi-quantitative assessments.

The most substantial improvements in petrographic regression tasks, particularly noticed for mean grain size estimation, demonstrate the practical value of combining domain adaptation

with robust measurement techniques. This improvement in characterization could therefore directly benefit downstream sedimentological interpretations.

5.5.3 Architectural Considerations and Training Strategy

The unified network architecture effectively balances multiple competing objectives through careful design choices. The shared encoder approach with domain-specific decoders enables efficient feature extraction while maintaining computational efficiency (Huang & Belongie, 2017; Y. Zhang et al., 2022). The use of masked cycle consistency loss represents an adaptation of the traditional CycleGAN (Zhu et al., 2017) approach designed for petrographic applications. By restricting consistency calculations to grain-containing regions, the method achieves enhanced textural variation in backgrounds while preserving fine-grained details within objects. Research related to cycle consistency properties have also proved their impact on provided contextual information and ways of regularization (Kasaraneni & Mishra, 2020; T. Wang & Lin, 2024). This modification here addresses the heterogeneous nature of thin section backgrounds, where significant variability is both expected and geologically meaningful. The identity cycle loss used in conventional image translation approaches for training stability, that consists of ensuring that a source image when translated to the same domain remain unaltered was identified to not be mandatory when initial learning rate was set to a sufficient initial low value ($\sim 5 \times 10^{-5}$). The same observation was done for migrating from a two-cycle training strategy to a unique-cycle. Segmentation constraint integration prevents undesired generation artifacts while simultaneously training a target-domain segmentation model. This dual-purpose approach maximizes training efficiency and ensures that translated images maintain semantic consistency with their synthetic counterparts (Cherian & Sullivan, 2019; Hoffman et al., 2017). Ultimately, the reference style mapping ability of the proposed framework allows to get control over the final appearance of adapted images and might be beneficial to generate a larger diversity of input data for the previously mentioned segmentation task.

5.5.4 Limitations and Challenges

Despite the progress achieved, several limitations persist. The current segmentation performance, although improved, remains below high-performance thresholds for certain classes. Nevertheless, relative abundance assessments which are key to evaluate supporting materials in sedimentary deposits, were found to be visually respected and metrics are still in the state-of-the-art range of medical nuclei segmentation (Basu et al., 2024; Caicedo et al., 2019; Lagree et al., 2021).

Performance variations across grain classes, particularly the lower metrics for peloids, underscore persistent difficulties in characterizing small particles. This limitation is partially attributed to the sensitivity of IoU scores to minor misalignments in small-object detection (Mun et al., 2023). Furthermore, the relatively small validation set (100 examples) constrains the statistical robustness of the evaluation.

Style transfer methods, while effective, may introduce subtle artifacts that are visually undetectable but could impact downstream analyses. Additionally, the evaluation focused

primarily on segmentation and grain size, potentially overlooking geological subtleties considered by petrographers.

Although LT provided robust measurements for compact grains, its utility may be reduced in samples dominated by elongated bioclasts such as bivalve or brachiopod fragments. Given that LT reflects the largest inscribed circle, it does not capture grain anisotropy. Moreover, the inherent limitations of analyzing 2D sections of 3D objects must be acknowledged, as observed shapes vary with orientation.

5.5.5 Recommendations

Future work should prioritize the development of larger, more diverse real image datasets to better evaluate the method's generalization capabilities. Collaborative efforts to compile reference datasets would provide valuable insights into the method's performance across different imaging conditions, staining protocols or geological contexts.

The framework could be extended to incorporate multi-scale analysis capabilities that can simultaneously characterize fine-grained matrix components and larger framework elements. This enhancement would provide more comprehensive textural characterizations and better support for complex carbonate facies interpretations.

Beyond local thickness measurements, investigation of additional morphometric parameters such as grain shape factors, roundness indices, and orientation statistics could provide richer characterizations for paleoenvironmental reconstructions (Karsanina, Gerke, Skvortsova, et al., 2015). The integration of these parameters with existing workflows would enhance the geological relevance of assisted analyses.

Development of active learning strategies could optimize the selection of real images for style transfer, potentially achieving similar performance improvements with reduced data requirements (Mittal et al., 2023). This approach would be particularly valuable for specialized applications where real data acquisition is challenging or expensive.

Development of class-specific adaptation strategies could address the variable performance observed across different grain types (Wang et al., 2023). Targeted approaches for challenging categories might yield more consistent improvements across all components.

Systematic comparison with expert petrographer interpretations would provide crucial validation of the method's geological relevance and identify areas where automated approaches diverge from established practices.

5.6 Conclusion

A unified framework has been presented in this study for synthetic-to-real domain adaptation and segmentation in assisted petrographic analysis of carbonate thin sections. This method completes the original sedimentology driven workflow of SynSection for conditional synthetic training dataset generation. Through the integration of image translation with semantic segmentation regularization, measurable improvements have been observed in bridging the gap between low-level features of synthetic training data and real-world geological examples.

Grain size metrics have been introduced for morphometric analysis, providing meaningful information for further sedimentological interpretations. The implementation of local thickness as a measurement method addresses boundary ambiguities encountered in segmentation applications to petrographic field. The combination of reference style image adaptation with adapted measurement techniques has resulted in quantifiable improvements in petrographic grain segmentation and regression tasks.

This work establishes a foundation for transforming petrographic workflows within the geological community. The framework enables full thin section characterization through quantitative and objective analyses that complement traditional expert interpretations. By maintaining geological consistency while achieving enhanced performance, this approach opens new possibilities for valuable throughput petrographic studies, standardized characterization protocols, and improved reproducibility in carbonate sedimentology research.

The integration of synthetic data generation with real-domain analysis offers the geological community a scalable solution to the persistent challenge of limited annotated datasets. Using synthetic data enables transparency over training datasets, as generation parameters can be directly expressed to align with sedimentological constraints. This work demonstrates that synthetic training data can be effectively adapted for real-world applications, providing a pathway for developing petrographic tools across diverse geological contexts that support both research applications and industrial geological characterization needs.

REFERENCES

- Ávila, J., Pagalo, J., & Espinoza-Andaluz, M. (2022). Evaluation of geometric tortuosity for 3D digitally generated porous media considering the pore size distribution and the A-star algorithm. *Scientific Reports*, 12(1), 19463. <https://doi.org/10.1038/s41598-022-23643-6>
- Azzam, F., Blaise, T., & Brigaud, B. (2024). Automated petrographic image analysis by supervised and unsupervised machine learning methods. *Sedimentologica*, 2(2), e22.1594. <https://doi.org/10.57035/journals/sdk.2024.e22.1594>
- Blott, S. J., & Pye, K. (2001). GRADISTAT: A grain size distribution and statistics package for the analysis of unconsolidated sediments. *Earth Surface Processes and Landforms*, 26(11), 1237–1248. <https://doi.org/10.1002/esp.261>
- Cherian, A., & Sullivan, A. (2019). Sem-GAN: Semantically-consistent image-to-image translation. In *2019 IEEE Winter Conference on Applications of Computer Vision (WACV)* (pp. 1797–1806). IEEE. <https://doi.org/10.1109/WACV.2019.00196>
- Chiang, M. Y. M., Landis, F., Wang, X., Smith, J. R., Cicerone, M. T., Dunkers, J. P., & Luo, Y.-K. (2009). Local thickness and anisotropy approaches to characterize pore size distribution of three-dimensional porous networks. *Tissue Engineering Part C: Methods*, 15(1), 65–76. <https://doi.org/10.1089/ten.tec.2008.0204>
- Dahl, V. A., & Dahl, A. B. (2023). Fast local thickness. In *2023 IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops (CVPRW)* (pp. 4336–4344). IEEE. <https://doi.org/10.1109/CVPRW59228.2023.00458>
- Dawson, H. L., Dubrule, O., & John, C. M. (2023). Impact of dataset size and convolutional neural network architecture on transfer learning for carbonate rock classification. *Computers & Geosciences*, 171, 105284. <https://doi.org/10.1016/j.cageo.2022.105284>
- Dawson, H. L., & John, C. M. (2024). Object detection algorithms to identify skeletal components in carbonate cores. *Marine and Petroleum Geology*, 167, 106965. <https://doi.org/10.1016/j.marpetgeo.2024.106965>
- Ding, K., Zhou, M., Wang, H., Gevaert, O., Metaxas, D., & Zhang, S. (2023). A large-scale synthetic pathological dataset for deep learning-enabled segmentation of breast cancer. *Scientific Data*, 10(1), 231. <https://doi.org/10.1038/s41597-023-02125-y>
- Farahani, A., Voghoei, S., Rasheed, K. M., & Arabnia, H. R. (2021). A brief review of domain adaptation. In *Advances in Data Science and Information Engineering* (pp. 877–894). Springer. https://doi.org/10.1007/978-3-030-71704-9_65
- Ganin, Y., Ustinova, E., Ajakan, H., Germain, P., Larochelle, H., Laviolette, F., Marchand, M., & Lempitsky, V. (2016). Domain-adversarial training of neural networks. *Journal of Machine Learning Research*, 17(59), 1–35. <https://doi.org/10.48550/arXiv.1505.07818>
- Gatys, L. A., Ecker, A. S., & Bethge, M. (2016). Image style transfer using convolutional neural networks. In *2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)* (pp. 2414–2423). IEEE. <https://doi.org/10.1109/CVPR.2016.265>
- Goodfellow, I., Pouget-Abadie, J., Mirza, M., Xu, B., Warde-Farley, D., Ozair, S., Courville, A., & Bengio, Y. (2014). Generative adversarial networks. *Advances in Neural Information Processing Systems*, 27, 2672–2680. <https://doi.org/10.1145/3422622>

- Gostick, J., Khan, Z., Tranter, T., Kok, M., Agnaou, M., Sadeghi, M., & Jervis, R. (2019). PoreSpy: A Python toolkit for quantitative analysis of porous media images. *Journal of Open Source Software*, 4(37), 1296. <https://doi.org/10.21105/joss.01296>
- Haq, M. M., Ma, H., & Huang, J. (2023). NuSegDA: Domain adaptation for nuclei segmentation. *Frontiers in Big Data*, 6, 1082779. <https://doi.org/10.3389/fdata.2023.1082779>
- He, L., Ren, X., Gao, Q., Zhao, X., Yao, B., & Chao, Y. (2017). The connected-component labeling problem: A review of state-of-the-art algorithms. *Pattern Recognition*, 70, 25–43. <https://doi.org/10.1016/j.patcog.2017.04.018>
- Heusel, M., Ramsauer, H., Unterthiner, T., Nessler, B., & Hochreiter, S. (2017). GANs trained by a two time-scale update rule converge to a local Nash equilibrium. *Advances in Neural Information Processing Systems*, 30, 6626–6637. <https://doi.org/10.48550/arXiv.1706.08500>
- Hoffman, J., Tzeng, E., Park, T., Zhu, J.-Y., Isola, P., Saenko, K., Efros, A. A., & Darrell, T. (2018). CyCADA: Cycle-consistent adversarial domain adaptation. In *35th International Conference on Machine Learning* (pp. 1989–1998). PMLR. <https://doi.org/10.48550/arXiv.1711.03213>
- Huang, X., & Belongie, S. (2017). Arbitrary style transfer in real-time with adaptive instance normalization. In *2017 IEEE International Conference on Computer Vision (ICCV)* (pp. 1510–1519). IEEE. <https://doi.org/10.1109/ICCV.2017.167>
- Isola, P., Zhu, J.-Y., Zhou, T., & Efros, A. A. (2017). Image-to-image translation with conditional adversarial networks. In *2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)* (pp. 5967–5976). IEEE. <https://doi.org/10.1109/CVPR.2017.632>
- Karras, T., Laine, S., & Aila, T. (2019). A style-based generator architecture for generative adversarial networks. In *2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)* (pp. 4401–4410). IEEE. <https://doi.org/10.1109/CVPR.2019.00453>
- Karsanina, M. V., Gerke, K. M., Skvortsova, E. B., & Mallants, D. (2015). Universal spatial correlation functions for describing and reconstructing soil microstructure. *PLOS ONE*, 10(5), e0126515. <https://doi.org/10.1371/journal.pone.0126515>
- Kasaraneni, S. H., & Mishra, A. (2020). Image completion and extrapolation with contextual cycle consistency. In *2020 IEEE International Conference on Image Processing (ICIP)* (pp. 1901–1905). IEEE. <https://doi.org/10.1109/ICIP40778.2020.9191115>
- Koeshidayatullah, A., Morsilli, M., Lehrmann, D. J., Al-Ramadan, K., & Payne, J. L. (2020). Fully automated carbonate petrography using deep convolutional neural networks. *Marine and Petroleum Geology*, 122, 104687. <https://doi.org/10.1016/j.marpetgeo.2020.104687>
- Koeshidayatullah, A., Trower, E. J., Li, X., Mukerji, T., Lehrmann, D. J., Morsilli, M., Al-Ramadan, K., & Payne, J. L. (2022). Quantitative evaluation of the roles of ocean chemistry and climate on ooid size across the Phanerozoic: Global versus local controls. *Sedimentology*, 69(6), 2486–2506. <https://doi.org/10.1111/sed.12998>
- Liu, X., Chandra, V., Ramdani, A. I., Zuhlke, R., & Vahrenkamp, V. (2023). Using deep-learning to predict Dunham textures and depositional facies of carbonate rocks from thin sections. *Geoenergy Science and Engineering*, 227, 211906. <https://doi.org/10.1016/j.geoen.2023.211906>

- Liu, X., Yoo, C., Xing, F., Oh, H., El Fakhri, G., Kang, J.-W., & Woo, J. (2022). Deep unsupervised domain adaptation: A review of recent advances and perspectives. *APSIPA Transactions on Signal and Information Processing*, 11(1), e25. <https://doi.org/10.1561/116.00000192>
- Mamode, H. I., Hampson, G. J., & John, C. M. (2025). Do more with less: Exploring semi-supervised learning for geological image classification. *Applied Computing and Geosciences*, 25, 100216. <https://doi.org/10.1016/j.acags.2024.100216>
- Mittal, S., Niemeijer, J., Schäfer, J. P., & Brox, T. (2023). Best practices in active learning for semantic segmentation. In *DAGM German Conference on Pattern Recognition* (pp. 364–378). Springer. https://doi.org/10.1007/978-3-031-54605-1_24
- Nichols, K., Hauwiler, M. R., Propes, N., Wu, S., Hernandez, S., & Kautzky, M. (2024). Segment Anything Model for grain characterization in hard drive design. *arXiv preprint*, arXiv:2408.12732. <https://doi.org/10.48550/arXiv.2408.12732>
- Nilsson, J., & Akenine-Möller, T. (2020). Understanding SSIM. *arXiv preprint*, arXiv:2006.13846. <https://doi.org/10.48550/arXiv.2006.13846>
- Odena, A., Dumoulin, V., & Olah, C. (2016). Deconvolution and checkerboard artifacts. *Distill*, 1(10), e3. <https://doi.org/10.23915/distill.00003>
- Otsu, N. (1979). A threshold selection method from gray-level histograms. *IEEE Transactions on Systems, Man, and Cybernetics*, 9(1), 62–66. <https://doi.org/10.1109/TSMC.1979.4310076>
- Peng, X., Usman, B., Saito, K., Kaushik, N., Hoffman, J., & Saenko, K. (2018). Syn2Real: A new benchmark for synthetic-to-real visual domain adaptation. In *2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops (CVPRW)* (pp. 856–8565). IEEE. <https://doi.org/10.1109/CVPRW.2018.00143>
- Pires de Lima, R., Duarte, D., Nicholson, C., Slatt, R., & Marfurt, K. J. (2020). Petrographic microfacies classification with deep convolutional neural networks. *Computers & Geosciences*, 142, 104481. <https://doi.org/10.1016/j.cageo.2020.104481>
- Ronneberger, O., Fischer, P., & Brox, T. (2015). U-Net: Convolutional networks for biomedical image segmentation. In *Medical Image Computing and Computer-Assisted Intervention – MICCAI 2015* (pp. 234–241). Springer. https://doi.org/10.1007/978-3-319-24574-4_28
- Ros, G., Sellart, L., Materzynska, J., Vazquez, D., & Lopez, A. M. (2016). The SYNTHIA dataset: A large collection of synthetic images for semantic segmentation of urban scenes. In *2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)* (pp. 3234–3243). IEEE. <https://doi.org/10.1109/CVPR.2016.352>
- Sekachev, B., Manovich, N., Zhiltsov, M., Zhavoronkov, A., Kalinin, D., Hoff, B., Tosmanov, Kruchinin, D., Zankevich, A., DmitriySidnev, Markelov, M., Johannes222, Chenuet, M., a-andre, telenachos, Melnikov, A., Kim, J., Ilouz, L., Glazov, N., Priya, Tehrani, R., Jeong, S., Skubriev, V., Yonekura, S., Vugia, T., Zliang, Lizhming, & Truong, T. (2020). OpenCV/CVAT: v1.1.0. Zenodo. <https://doi.org/10.5281/zenodo.4009388>
- Shebl, H. T., Al Tamimi, M. A., Boyd, D. A., & Nehaid, H. A. (2021). Automation of carbonate rock thin section description using cognitive image recognition. In *Abu Dhabi International Petroleum Exhibition & Conference*. Society of Petroleum Engineers. <https://doi.org/10.2118/208149-MS>

- Simonyan, K., & Zisserman, A. (2014). Very deep convolutional networks for large-scale image recognition. *arXiv preprint*, arXiv:1409.1556. <https://doi.org/10.48550/arXiv.1409.1556>
- Sudre, C. H., Li, W., Vercauteren, T., Ourselin, S., & Cardoso, M. J. (2017). Generalised dice overlap as a deep learning loss function for highly unbalanced segmentations. In *Deep Learning in Medical Image Analysis and Multimodal Learning for Clinical Decision Support* (pp. 240–248). Springer. https://doi.org/10.1007/978-3-319-67558-9_28
- Szegedy, C., Vanhoucke, V., Ioffe, S., Shlens, J., & Wojna, Z. (2016). Rethinking the inception architecture for computer vision. In *2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)* (pp. 2818–2826). IEEE. <https://doi.org/10.1109/CVPR.2016.308>
- Takato, Y. (2021). Generative synthetic augmentation using label-to-image translation for nuclei image segmentation. In *2021 IEEE 18th International Symposium on Biomedical Imaging (ISBI)* (pp. 1561–1565). IEEE. <https://doi.org/10.1109/ISBI48211.2021.9434130>
- Thota, M., & Leontidis, G. (2021). Contrastive domain adaptation. In *2021 IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops (CVPRW)* (pp. 2209–2218). IEEE. <https://doi.org/10.1109/CVPRW53098.2021.00249>
- Trullo, R., Bui, Q. A., Tang, Q., & Olfati-Saber, R. (2022). Image translation based nuclei segmentation for immunohistochemistry images. In *Deep Generative Models, and Data Augmentation, Labelling, and Imperfections* (pp. 133–142). Springer. https://doi.org/10.1007/978-3-031-18576-2_13
- Tsai, Y.-H., Hung, W.-C., Schuler, S., Sohn, K., Yang, M.-H., & Chandraker, M. (2018). Learning to adapt structured output space for semantic segmentation. In *2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition* (pp. 7472–7481). IEEE. <https://doi.org/10.1109/CVPR.2018.00780>
- Tzeng, E., Hoffman, J., Saenko, K., & Darrell, T. (2017). Adversarial discriminative domain adaptation. In *2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)* (pp. 2962–2971). IEEE. <https://doi.org/10.1109/CVPR.2017.316>
- Utz, J., Weise, T., Schlereth, M., Wagner, F., Thies, M., Gu, M., Uderhardt, S., & Breininger, K. (2023). Focus on content not noise: Improving image generation for nuclei segmentation by suppressing steganography in CycleGAN. In *Medical Image Computing and Computer Assisted Intervention – MICCAI 2023* (pp. 212–222). Springer. https://doi.org/10.1007/978-3-031-43993-3_21
- Wang, T., & Lin, Y. (2024). CycleGAN with better cycles. *arXiv preprint*, arXiv:2408.15374. <https://doi.org/10.48550/arXiv.2408.15374>
- Wang, W., Wang, M., Dong, X., Lan, L., Zu, Q., Zhang, X., & Wang, C. (2023). Class-specific and self-learning local manifold structure for domain adaptation. *Pattern Recognition*, 142, 109654. <https://doi.org/10.1016/j.patcog.2023.109654>
- Xiao, A., Huang, J., Guan, D., Zhan, F., & Lu, S. (2021). Transfer learning from synthetic to real LiDAR point cloud for semantic segmentation. In *Proceedings of the AAAI Conference on Artificial Intelligence* (Vol. 36, pp. 2795–2803). AAAI Press. <https://doi.org/10.1609/aaai.v36i3.20181>

- Zhang, Y., Tang, F., Dong, W., Huang, H., Ma, C., Lee, T.-Y., & Xu, C. (2022). Domain enhanced arbitrary image style transfer via contrastive learning. In *ACM SIGGRAPH 2022 Conference Proceedings* (pp. 1–8). ACM. <https://doi.org/10.1145/3528233.3530736>
- Zhang, Z., Liu, Q., & Wang, Y. (2018). Road extraction by deep residual U-Net. *IEEE Geoscience and Remote Sensing Letters*, 15(5), 749–753. <https://doi.org/10.1109/LGRS.2018.2802944>
- Zhu, J.-Y., Park, T., Isola, P., & Efros, A. A. (2017). Unpaired image-to-image translation using cycle-consistent adversarial networks. In *2017 IEEE International Conference on Computer Vision (ICCV)* (pp. 2242–2251). IEEE. <https://doi.org/10.1109/ICCV.2017.244>
- Zhuang, F., Qi, Z., Duan, K., Xi, D., Zhu, Y., Zhu, H., Xiong, H., & He, Q. (2021). A comprehensive survey on transfer learning. *Proceedings of the IEEE*, 109(1), 43–76. <https://doi.org/10.1109/JPROC.2020.3004555>

CHAPTER 6: GENERAL SYNTHESIS, CONCLUSION AND WAY FORWARD

6.1 Deep Quantitative Petrography

This study presented a comprehensive suite of complementary methods for quantitative, computer-assisted and multi-scale description of thin section images from carbonate rocks. The extracted parameters from these descriptions can be integrated into various established nomenclature systems, providing standardized and reproducible petrographic analysis (Figure 6.1).

6.2 Methodology Overview

The methodology unfolds from the broad sample characterization (Figure 6.2) to granular implementation details (Figure 6.3).

6.2.1 Image Preprocessing and Tiling Strategy

Modern petrographic images can achieve very high resolutions ($<1\mu\text{m}/\text{pixel}$) with correspondingly large dimensions ($>50,000\times 25,000$ pixels). A critical first step involves defining an appropriate image subdivision strategy that balances computational feasibility with sufficient quality for identifying desired descriptive criteria (litho-texture, allochems, pore types, cement, and matrix).

Based on analysis of oil and gas field samples and outcrop data, it was determined that a resolution of approximately $7\mu\text{m}/\text{pixel}$ with tile sizes of 512×512 pixels provide optimal results for both image labeling and grain type annotation.

6.2.2 Litho-Textural Analysis

6.2.2.1 Representation Learning and Raw Data Clustering

The tiling process generates substantial datasets ($>100,000$ tiles from $\sim 15,000$ large images). Even basic sorting of this data represents a significant bottleneck in the classification process using standard carbonate petrography nomenclatures.

A representation learning approach was employed for dimensionality reduction of the input image space, followed by clustering in the reduced latent space. Chapter 3 explored this approach using a VAE-GAN that learned rotation-invariant representations, mimicking the rotating stage commonly used in microscopy.

This method successfully formed geologically coherent groups based primarily on R. Dunham's classification of limestones and R. Folk's allochem preservation criteria in dolomites, plus certain anhydrite types. While imperfect, this approach accelerated the creation of training datasets through clustering and manual sampling, reducing tile dimensions to 128×128 pixels and achieving a twofold acceleration in traditional labeling time.

Deep Learning for Carbonate Rocks Petrography – General Synthesis, Conclusion and Way Forward

Quantitative Carbonates Petrography using Deep Learning

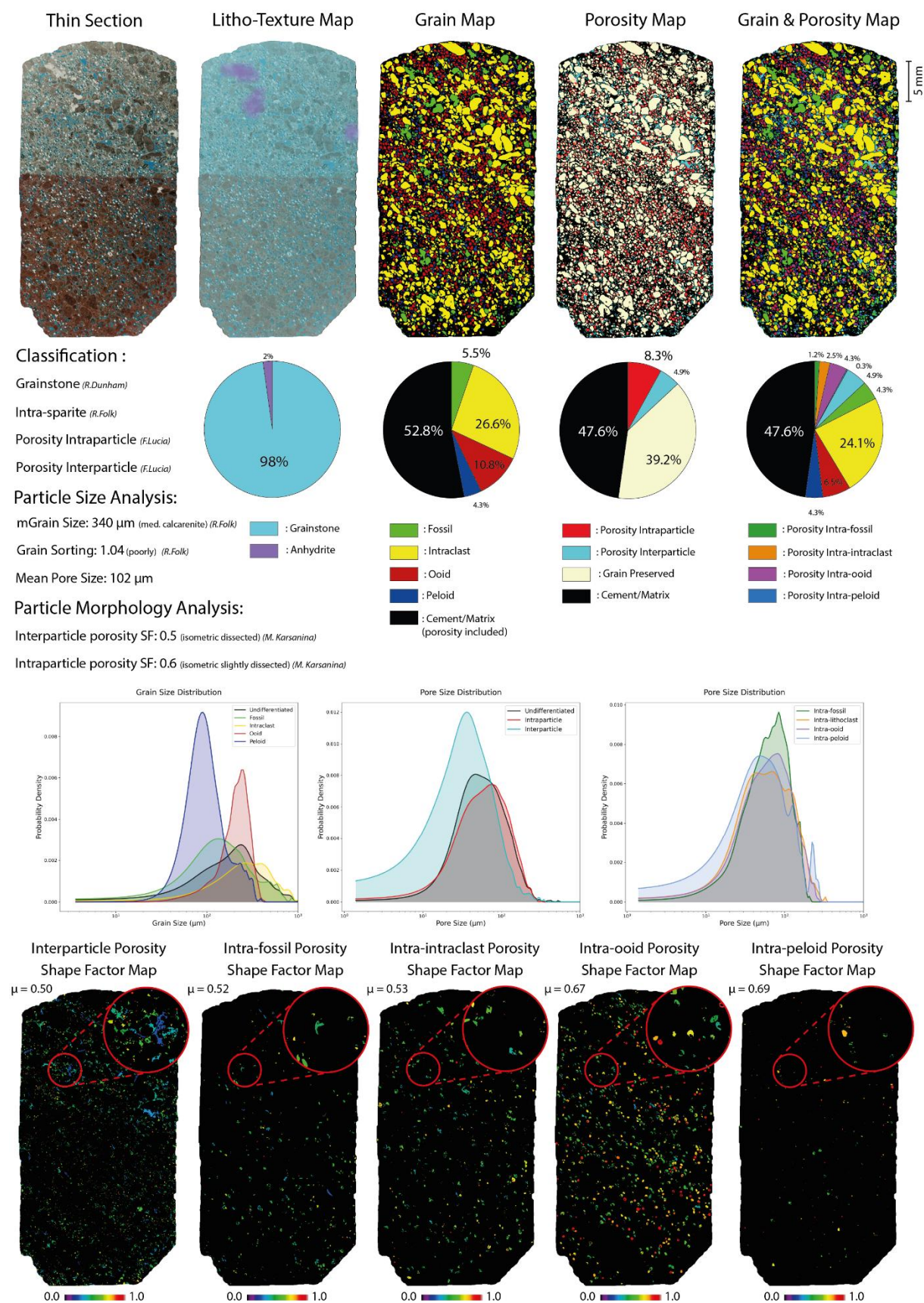


Figure 6.1 Example of multi-scale carbonate thin section analysis resulting from the combined deep approaches outputs.

Global Synthesis: Combined Deep Learning Approaches for Assisted Carbonate Petrography (Classification)

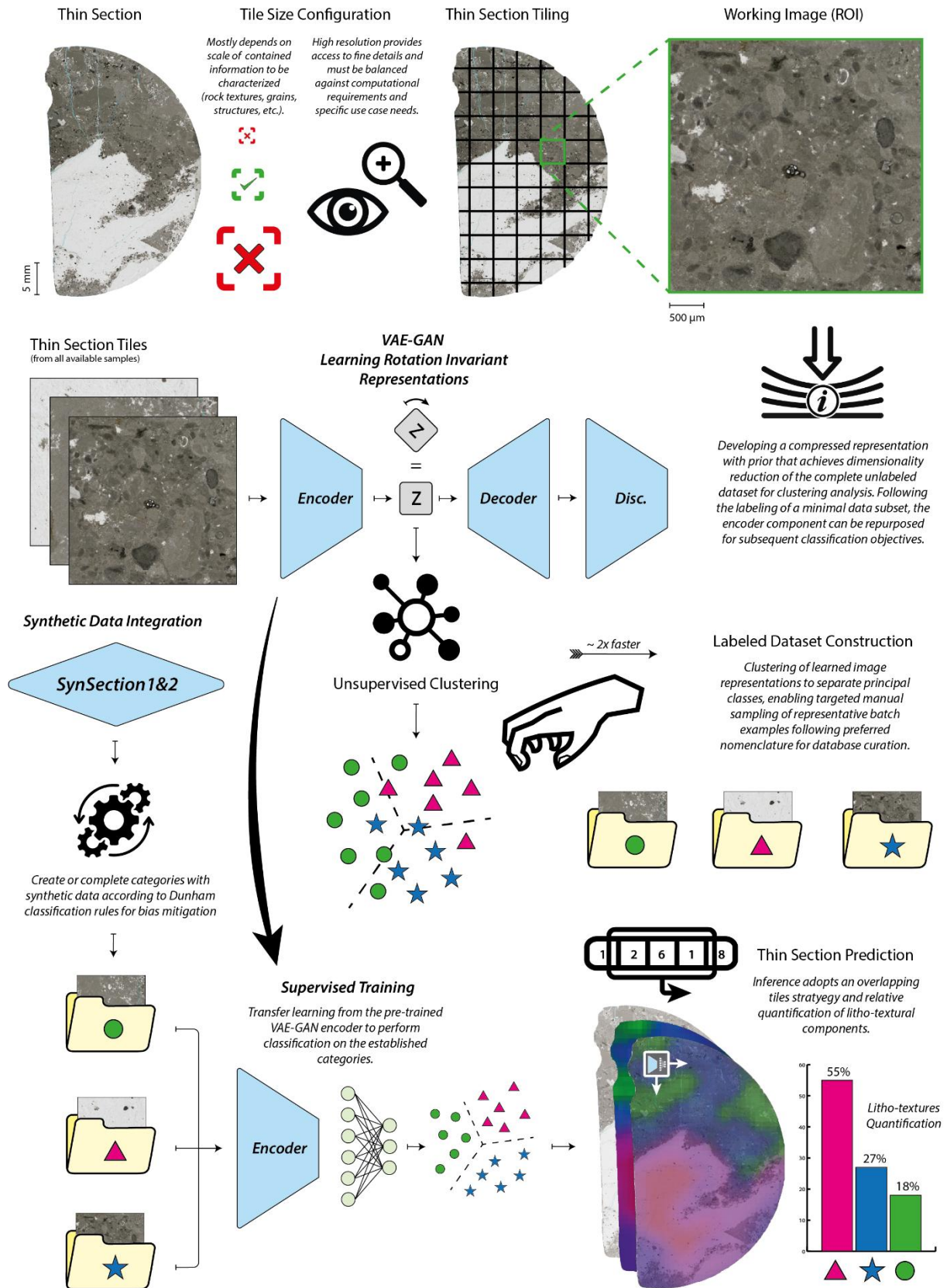


Figure 6.2 First order thin section characterization using deep classification approaches for quantitative analysis.

6.2.2.2 Supervised Transfer Learning Implementation for Image Classification

The encoder from the previously trained VAE-GAN was retrained using transfer learning techniques. The model, having already learned representations from unlabeled data, demonstrated improved performance (+3.5% mAP) on test data compared to randomly initialized models, highlighting the utility and effectiveness of this approach.

Addressing Classification Bias with Synthetic Data:

Accelerated labeling, while beneficial, does not resolve identified biases in visual differentiation between classes following R. Dunham or R. Folk principles, particularly around class boundaries (mudstone vs. wackestone, wackestone vs. packstone, etc.).

Methods to generate synthetic data that limit interpretative biases were developed in chapter 4 and 5 by creating configurable image examples that respect criteria for each granular limestone class while remaining realistic and domain appropriate. The integration of such data in a representation learning process may benefit preliminary clustering and more advanced supervised classification stage for ensuring both data balancing and variability.

6.2.2.3 Global-Scale Application

Thin sections exhibit spatial heterogeneities at their global scale, necessitating predictions using sliding window methods with overlap across the entire sample to obtain litho-textural quantifications. This approach provides first-order descriptions particularly useful for mixed facies that were observed in muddy limestones, dolomites and anhydrite.

6.2.3 Granular Analysis

6.2.3.1 Data Labeling and Segmentation Approaches

To reveal heterogeneity in granular limestones (grainstones to packstones), more detailed descriptions involving grain type identification within these rocks became necessary.

Grain-supported limestone description represents a meticulous operation in traditional quantitative petrography. These types of thin sections typically contain thousands of grains, making complete manual annotation impractical (requiring ~2-3 months for complete manual labeling versus ~1 week using segmentation assistance tools like SAM). Although labeling an entire thin section is not necessarily required for all petrographic analyses, depending on the objective, this type of quantitative study remains a relevant approach for heterogeneous samples where a few restricted fields of view would limit the global context of the analysis.

The grain quantification task was investigated using deep semantic segmentation approaches, which align well with desired parameters such as relative proportions of grain types per surface unit, grain size distribution parameters, combination with porosity thresholding for semantic pore type classification.

6.2.3.2 Synthetic Data Generation using SynSection

Like image classification, creating labeled training datasets is crucial for segmentation model development. A procedural method (SynSection) for generating triplets of images, labels, and annotation masks were proposed in chapter 4. This approach requires relatively small initial annotation quantities (~100 images) of grains, cement, and matrix compared to ideal

requirements (10,000 to 100,000+ images). SynSection simulates depositional processes using packing methods with spatial distribution rules, controlling relative grain proportions in images, grain types, local proximity and orientation. The grain composition is then merged with textured backgrounds generated by models trained on annotated backgrounds.

SynSection method rapidly generates complete training datasets (100,000+ images and masks in 1-2 days) for pre-training U-Net segmentation models before fine-tuning on real labeled data. The approach showed significant performance improvement compared to simple training on augmented real data (+20% mAP).

6.2.3.3 Advanced Synthetic Data Adaptation using SynSection2

While SynSection effectively introduces high-level characteristic variability in low-frequency image content, it struggles with low-level characteristics like textures, color associations between grains and backgrounds, and sometimes overly smooth generated images.

A model that learns both synthetic-to-realistic image transformation and segmentation within a single unified framework was implemented. This approach offers several benefits that simplifies the two-stage process (separate adaptation and segmentation models), preserves semantic content during image adaptation through segmentation constraints, and allows greater textural diversity in the transformed background through consistency penalties between detransformed and original synthetic images that are applied only in grain-containing regions. This method achieved additional improvements over chapter 4 approaches (+5% mAP), with the internal segmentation model progressively learning to segment synthetic images in the real domain.

Beyond pure image-based segmentation quality indicators, metrics related to downstream petrographic processes were developed focusing on grain size distribution parameter estimation, including mean grain size and sorting derived from segmentation masks. These metrics achieved determination coefficients (R^2) of 0.78 and 0.72 respectively, surpassing baseline results using non-adapted synthetic images for training and real images for model fine-tuning (+12% and +6%). The measurements were obtained using a local thickness approach quantifying surface area of detected particles through spherical filling simplification. This method proved less sensitive to noise in generated masks and works well for spherical particle compositions, though maximum diameter measurements should complement these measures for predominantly elongated particles.

6.2.3.4 Global Scale Application

Similar to classification processes, segmentation models trained on tiles can detect grain types at the global thin section scale using overlapping sliding window predictions, enabling quantification of elements in previously uncharacterized areas. Fine-grained description of granular content opens multiple possibilities such as sample categorization following R. Folk nomenclature (previously considered too complex), combination of semantic masks with porosity thresholding (blue dyed epoxy) for differentiating intra-granular and inter-granular porosities. Nomenclatures defined by Choquette & Pray (1970) or Lucia (1995, 1999) are also accessible by introducing morphometric descriptors of detected and classified pores based on intersected material.

Deep Learning for Carbonate Rocks Petrography – General Synthesis, Conclusion and Way Forward

Global Synthesis: Combined Deep Learning Approaches for Assisted Carbonate Petrography (Segmentation)

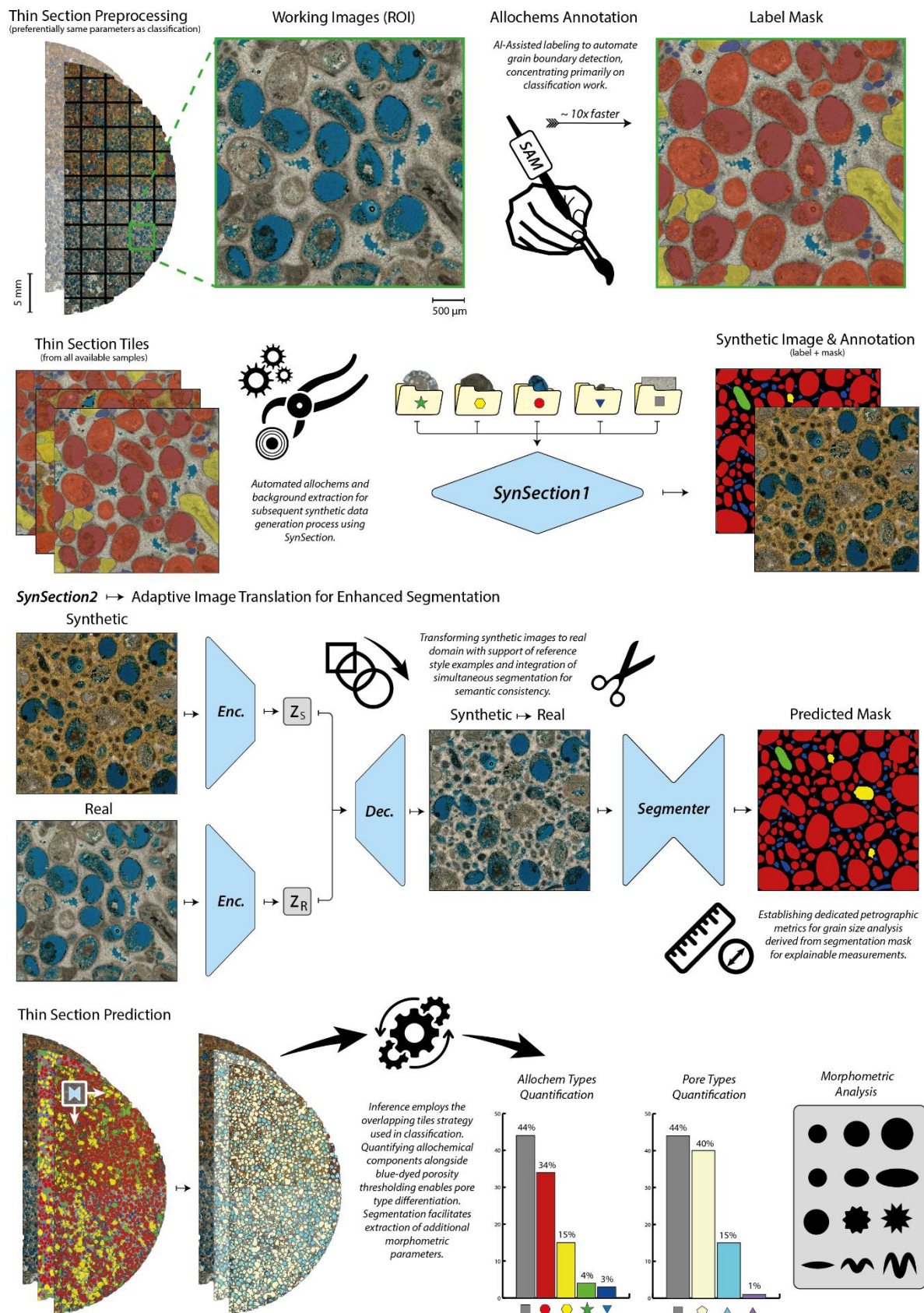


Figure 6.3 Second order thin section characterization in carbonate rocks using deep segmentation approaches for quantitative analysis.

6.3 Limitations and Recommendations

Quantitative approaches for carbonate rock petrography have been developed and decomposed into two main tasks for multi-scale characterization.

A first local classification step allows defining the lithology and texture of a thin section across its entire spatial coverage, with the goal of obtaining proportion maps reflecting millimetric-order heterogeneities. Moreover, the cases examined in this study primarily involved grain-supported limestones, which required enhanced descriptive resolution by transitioning from image classification to grain segmentation tasks for revealing the semantic heterogeneity of the samples. This method notably enabled the quantification of main allochem class proportions. Through analysis of grain and pore size distributions along with their specific relationships, the level of detail provided was increased. This approach focused on certain key parameters to simplify the task for operators who would otherwise be required to reference numerous qualification systems. These methods, although aligned with the initially set objectives, have certain limitations.

6.3.1 Manufacturing Impacts

Thin section thicknesses and manufacturing methods directly impact quality. Very thin sections tend to eliminate essential characteristics for clear litho-texture identification, causing confusion even for trained experts or expert systems. Very thick sections darken the observed field but generally have less negative impact than very thin sections. Recommendations would favor manufacturing standardization procedures over data augmentation approaches for training deep learning models, which can be complex to implement effectively in these cases.

6.3.2 Classification Coverage Limitations

Several facies potentially observed in carbonate systems (microbial or biological constructions, corals) were not included due to sample scarcity. Future systems should try to incorporate these data types for complete carbonate rock classification coverage (Dawson et al., 2023).

6.3.3 Dolomite Classification Challenges

Dolomite characterization employed simple binary classification (sucrosic vs preserved allochems) adapted to the used case studies, though these rocks can show much greater complexity as demonstrated by various nomenclatures established by R. Folk and Gregg and Sibleys. Differentiating dolomitized muddy facies is also visually complex, typically requiring sample density measurement (limestone $\rho=2.75 \text{ g/cm}^3$, dolomite $\rho=2.85 \text{ g/cm}^3$) for definitive lithological determination. Given that differentiation criteria are primarily based on crystal size and morphology, controlled synthetic data generation methods similar to SynSection appear essential for training classification models, given time constraints for rigorous real data labeling.

6.3.4 Groundmass Material Recognition Challenges

Small-scale (pixel) recognition and differentiation of cements and matrix were not addressed in this study (Bathurst, 1976; Flügel, 2004). A similar approach to litho-texture classification

could be developed for this problem using higher resolution image versions and potentially smaller tile sizes.

6.3.5 3D Grain Representation

Grains observed in thin sections represent only 2D cuts of potentially complex 3D objects. Ideally, 3D digital models should be created (Eber, 2023; Sun et al., 2024; Zuber et al., 2017) from which random sections could be extracted to enrich synthetic image generation databases, enabling finer grain class subdivision for specific biostratigraphic needs.

6.3.6 Multi-Scale Segmentation

Current segmentation approaches are limited by object sizes seen during training. Thin sections contain objects ranging from millimeters to centimeters without major textural variation from backgrounds (e.g. large intraclasts), leading sometimes to poor segmentation and unusable masks for quantification. A recommendation to address this issue might employ ensemble approaches working in parallel on large tiles or entire thin section images for large object segmentation and smaller tiles for small grain detection, with subsequent recombination for robust global segmentation (Podda et al., 2024b).

6.3.7 Architecture Updates

Incorporating Vision Transformer (ViT) architectures for classification and segmentation (Dosovitskiy et al., 2020; Strudel et al., 2021; Sylvester et al., 2025), alongside diffusion-based approaches for texture generation and image translation (Arslan et al., 2024; Ho et al., 2020; Tumanyan et al., 2022), represents a promising direction for unifying methodologies and establishing robust performance benchmarks. Preliminary evaluation of ViTs showed no performance gains over conventional methods, which is attributed to the limited size of available training datasets rather than fundamental architectural limitations.

6.3.8 Foundation Model Integration

Models developed in this study are custom architectures. The popularization of foundation models (SAM, DINO) (Caron et al., 2021; Kirillov et al., 2023; Oquab et al., 2023; Sim'eoni et al., 2025) demonstrates multi-domain information extraction capabilities for various downstream tasks, warranting investigation given their massive training datasets.

Given that the petrographic domain will likely never benefit from large-scale manually annotated and expert-verified data (>10,000), leveraging transverse research appears to be a sound approach for contributing to general understanding efforts.

6.4 Conclusion

Deep learning applied to computer vision proves to be a powerful tool for general or detailed image description when all necessary material for training architectures is assembled. This typically requires substantial annotated databases, though approaches like self-supervision now help limit initially massive label collection requirements.

This study's success resulted from combining different image processing and deep learning approaches. No single paradigm or specific approach appears superior to others; rather, dedicated and adapted use of specific architectures for identified problem types, requiring domain knowledge to understand extent and boundaries of traditionally used frameworks, enabled these results.

For supervised approaches with limited labeled image databases in carbonate petrography, representation learning approaches followed by clustering or downstream classification task refinement proved beneficial for exploring data variability before obtaining quantitative parameters at global sample scales aligned with modular descriptions following geological nomenclatures and highlighting possible textural heterogeneities. Most classification systems used rely on quantitative parameters derived from images, making it necessary to augment possible examples of certain classes (granular limestones) using configurable synthetic image generation following these same parameters to limit biases or potential gaps in learned distributions.

Complex image types that cannot be adequately characterized through single labels or that necessitate extensive adjectival sequences for proper definition and differentiation were addressed through enhanced description resolution. This enhancement was achieved by transitioning from global classification methodologies to pixel-level classification approaches, specifically semantic segmentation techniques. Focus was thus directed toward describing simplified granular limestone content to derive grain type quantifications and parameters from element size distributions.

Since obtaining labels for utilized images remained the primary limiting factor, synthetic data generation for model pre-training prior to refinement on real data was demonstrated to be both useful and necessary for this approach. However, realistic image generation through traditional processing methods can introduce unrealistic examples from low-level characteristic perspectives, including contours, colors, and textures artifacts during training phases. This limitation was substantially addressed through the implementation of deep generative domain adaptation approaches, where the desired final segmentation task was integrated within the constructed model itself, thereby contributing to semantic consistency during synthetic image transformation into realistic representations.

The incorporation of dedicated metrics for grain size parameter characterization further enhanced confidence in results generated during element segmentation. Additionally, the combination of semantic masks with observed porosity enabled the identification of distinct pore classes, making them fully accessible for various morphometric characterizations and facilitating improved understanding of their spatial arrangements for subsequent study of formation mechanisms.

The standardized integration of deep learning models for thin section characterization in petrographic image analysis represents a tool that requires considerable refinement to fully realize its potential capabilities. While traditional description tasks are increasingly being abandoned or accelerated in geological studies, such data is considered fundamental and should be exploited more systematically to understand certain petrophysical responses in carbonate systems. The homogeneity of rapid descriptions obtainable from large data volumes through these methods enables the establishment of constrained frameworks that limit potential variations naturally occurring across multiple studies of identical datasets. However, requirements may evolve alongside general subject knowledge advancement. Description nomenclatures employed for specific cases may prove inadequate for others; therefore, movement toward foundational systems developed with domain experts, supplemented by modular branches capable of addressing diverse possibilities and specialized disciplinary nuances, is considered essential.

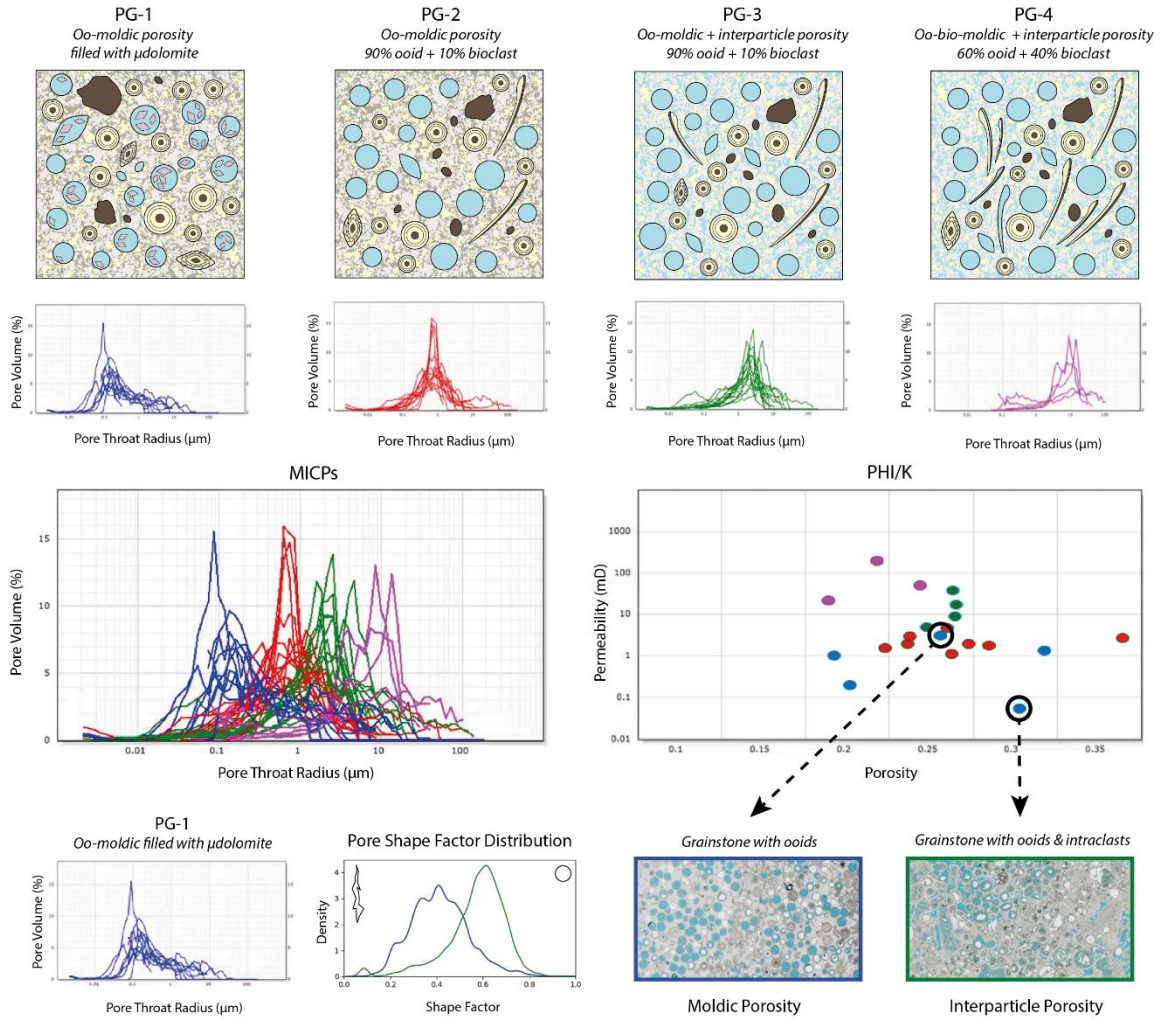
6.5 Way Forward

A multitude of possibilities and questions can emerge from this study. The focus on implications of parameters extracted from images when compared with static and dynamic petrophysical measurements acquired on core plugs appears to be the most promising direction for exploring and evaluating potential links between these two domains, sedimentology and petrophysics, unified through rock typing studies.

Micro-scale rock typing methods are diverse (Rebelle, 2014), and one approach based on Mercury Injection Capillary Pressure clustering, which has been historically used in studies, already shows its limitations when comparing morphological parameters derived from thin section image porosity with these same petrophysical measurements (Figure 6.4).

The integration of semantic quantifications from thin section images therefore represents a major asset for standardizing the groupings made in rock typing before moving to a higher scale of characterization using logs and core images.

Micro Scale Rock-Typing based on MICPs Clustering and Inconsistencies



« PG-1 mixing different facies with different petrophysic responses and pore shape factor »

Figure 6.4 MICP-based rock typing showing sedimentological and petrophysical mismatches that expose inadequate data integration in the clustering methodology, leading to poorly defined petrophysical groups (PG).

REFERENCES

- Arslan, F., Kabas, B., Dalmaz, O., Özbey, M., & Cukur, T. (2024). Self-consistent recursive diffusion bridge for medical image translation. *Medical Image Analysis*, 96, 103199. <https://doi.org/10.1016/j.media.2024.103199>
- Bathurst, R. G. C. (1976). Essai de caractérisation sédimentologique des dépôts carbonatés. Eléments d'analyse. (An attempt at sedimentological characterisation of carbonate deposits. Analytic elements): ELF-Aquitaine, Centre de Recherches de Boussens et de Pau, 1975, 173 pp. £7.00 [Book review]. *Sedimentary Geology*, 16(4), 319-320.
- Caron, M., Touvron, H., Misra, I., Jégou, H., Mairal, J., Bojanowski, P., & Joulin, A. (2021). Emerging properties in self-supervised vision transformers. In *Proceedings of the IEEE/CVF International Conference on Computer Vision* (pp. 9650-9660). IEEE. arXiv:2104.14294
- Dawson, H. L., Dubrule, O., & John, C. M. (2023). Impact of dataset size and convolutional neural network architecture on transfer learning for carbonate rock classification. *Computers and Geosciences*, 171, 105284. <https://doi.org/10.1016/j.cageo.2022.105284>
- Dosovitskiy, A., Beyer, L., Kolesnikov, A., Weissenborn, D., Zhai, X., Unterthiner, T., Dehghani, M., Minderer, M., Heigold, G., Gelly, S., Uszkoreit, J., & Houlsby, N. (2020). An image is worth 16x16 words: Transformers for image recognition at scale. arXiv:2010.11929
- Eber, A. (2023). Fossils in a digital age. *Pathfinder: A Canadian Journal for Information Science Students and Early Career Professionals*, 4(2), 45-62.
- Flügel, E. A. (2004). *Microfacies of carbonate rocks: Analysis, interpretation and application* (2nd ed.). Springer. <https://doi.org/10.1007/978-3-662-08726-8>
- Ho, J., Jain, A., & Abbeel, P. (2020). Denoising diffusion probabilistic models. In *Advances in Neural Information Processing Systems* (Vol. 33, pp. 6840-6851). arXiv:2006.11239
- Kirillov, A., Mintun, E., Ravi, N., Mao, H., Rolland, C., Gustafson, L., Xiao, T., Whitehead, S., Berg, A. C., Lo, W.-Y., Dollár, P., & Girshick, R. (2023). Segment anything. In *Proceedings of the IEEE/CVF International Conference on Computer Vision* (pp. 4015-4026). IEEE. arXiv:2304.02643
- Oquab, M., Darcet, T., Moutakanni, T., Vo, H., Szafraniec, M., Khalidov, V., Fernandez, P., Haziza, D., Massa, F., El-Nouby, A., Assran, M., Ballas, N., Galuba, W., Howes, R., Huang, P.-Y., Li, S.-W., Misra, I., Rabbat, M., Sharma, V., Szegedy, C., Bojanowski, P., & Joulin, A. (2023). DINOv2: Learning robust visual features without supervision. arXiv:2304.07193
- Podda, A. S., Balia, R., Barra, S., Carta, S., Neri, M., Guerriero, S., & Piano, L. (2024). Multi-scale deep learning ensemble for segmentation of endometriotic lesions. *Neural Computing and Applications*, 36(24), 14895-14908. <https://doi.org/10.1007/s00521-024-09828-2>
- Rebelle, M. (2014, April). Rock-typing in carbonates: A critical review of clustering methods. In *SPE Reservoir Characterization and Simulation Conference and Exhibition* (pp. 1-15). Society of Petroleum Engineers. <https://doi.org/10.2118/171759-MS>

- Siméoni, O., Vo, H. V., Seitzer, M., Baldassarre, F., Oquab, M., José, C., Khalidov, V., Szafraniec, M., Yi, S., Ramamonjisoa, M., Massa, F., Haziza, D., Wehrstedt, L., Wang, J., Darcet, T., Moutakanni, T., Sentana, L., Roberts, C., Vedaldi, A., Larlus, D., & Bojanowski, P. (2025). DINOv3: Scaling self-supervised vision transformers. *arXiv preprint*. arXiv:2501.03179
- Strudel, R., Garcia, R., Laptev, I., & Schmid, C. (2021). Segmenter: Transformer for semantic segmentation. In *Proceedings of the IEEE/CVF International Conference on Computer Vision* (pp. 7262-7272). IEEE. arXiv:2105.05633
- Sun, M., Zhong, Y., Qin, Z., Liu, Z., Gao, C., & Gong, Y. (2024). Research on 3D reconstruction of paleontological fossils based on photoacoustic microscopy. *Journal of Physics: Conference Series*, 2822(1), 012068. <https://doi.org/10.1088/1742-6596/2822/1/012068>
- Sylvester, Z., Stockli, D. F., Howes, N., Roberts, K., Malkowski, M. A., Poros, Z., Martindale, R. C., & Bai, W. (2025). Segmentever grain: A Python module for segmentation of grains in images. *Journal of Open Source Software*, 10(105), 7953. <https://doi.org/10.21105/joss.07953>
- Tumanyan, N., Geyer, M., Bagon, S., & Dekel, T. (2022). Plug-and-play diffusion features for text-driven image-to-image translation. In *2023 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)* (pp. 1921-1930). IEEE. <https://doi.org/10.1109/CVPR52729.2023.00190>
- Zuber, M., Laaß, M., Hamann, E., Kretschmer, S., Hauschke, N., van de Kamp, T., Baumbach, T., & König, T. (2017). Augmented laminography, a correlative 3D imaging method for revealing the inner structure of compressed fossils. *Scientific Reports*, 7(1), 41413. <https://doi.org/10.1038/srep41413>