Image segmentation through metaheuristics optimization : application to brain magnetic resonance images

Thuy Xuan Pham

Université Paris-Est
ÉCOLE DOCTORALE MATHEMATIQUES ET STIC (MSTIC, E.D. 532)

THÈSE

présentée et soutenue publiquement le 9 décembre 2019

pour l'obtention du

Doctorat de l'Université Paris-Est

DOMAINE DE RECHERCHE : SIGNAL, IMAGE, ET AUTOMATIQUE (CNU 61)

Thèse présentée par

Thuy Xuan PHAM

Image segmentation through metaheuristics optimization: application to brain magnetic resonance images

Directeur de thèse: **Pr. Patrick Siarry**

Composition du jury

C. FERNANDEZ-MALOIGNE	Université de Poitiers	Rapporteur
J-K. HAO	Université d'Angers	Rapporteur
L. DUMAS	Université de Versailles Saint Quentin en Yvelines	Examinateur
M. GARREAU	Université de Rennes 1	Examinateur
H. OULHADJ	Université Paris-Est Créteil	Examinateur
P. SIARRY	Université Paris-Est Créteil	Directeur

Acknowledgments

First of all, I would like to express my special appreciation and thanks to my supervisor, Professor Patrick SIARRY, Université Paris-Est Créteil (UPEC), France. He introduced me to the fields of Image Processing and Metaheuristics, educated me with the principles and methods of research, and guided me patiently throughout this work. I am highly indebted and express my deep sense of gratitude to him for his invaluable guidance, constant inspiration and motivation with enormous moral support. Working with him, a person with strong commitment and positive attitude is a highly rewarding experience of mine. I would also like to provide my special thanks to Dr. Hamouche OULHADJ, my co-supervisor. I was very pleased by his suggestions and contributions in this research project.

I would like to thank the members of my dissertation committee, Prof. Christine FERNANDEZ-MALOIGNE and Prof. Jin-Kao HAO, for having accepted to serve as referees, Prof. Mireille GARREAU and Prof. Laurent DUMAS, for having accepted to examine this work.

Many thanks to all people whose friendship and companionship has given me a lot of encouragement during my research work. I want to express my sincere appreciation to Ms. Thi Thanh Tam LE for extending her help and support in editing manuscripts. I also want to extend my profound appreciation to my relatives who have supported and encouraged me over the past years.

Personally, I owe my largest debt to my family. I am highly indebted to my parents, parents-in-laws, brothers, sisters and my beloved kids for their love, affection, constant encouragement and invaluable support throughout this particular period. I also want to express my gratitude to my wife Thu Huong for motivating me, daily taking care of and educating our daughters, Bao Chi and Bao Linh. I would never have completed this work without their love, encouragement and support. I dedicate this work to all of them.

Finally, I gratefully acknowledge the financial support from the Government of Vietnam (Project 911). Also, I would like to thank all my friends and colleagues at LiSSi Lab and doctoral school MSTIC for their support during the course of this doctoral project.

Abstract

Image segmentation is the process of partitioning an image into smaller non-overlapped and meaningful regions based in part on some homogeneity characteristics. Many high-level processing tasks such as feature extraction, object recognition and medical diagnosis depend heavily on the quality of solutions. In medical image analysis, images usually contain some artifacts such as noise, image volume and bias field effects due to various factors, for instance, environment and acquisition devices, and have complex structures. Therefore, image segmentation remains a difficult task even if various techniques and methods of different accuracy and degree of complexity have been introduced in the literature. Several approaches, such as fuzzy clustering, region-based active contour and Markov random field, have been found that can produce promising results; however, still many key open issues remain to be investigated. Up to now, there is no gold standard method and segmentation procedures still need a significant amount of expert intervention to achieve acceptable performance.

Metaheuristics are high-level procedures designed to solve optimization problems by the process of searching acceptable suboptimal solutions to a particular problem of interest. This family of algorithms is generally applied to problems for which there is no satisfactory algorithm able to solve them effectively. Therefore, they are widely used to solve complex problems and have proven to be successful in many fields of application with varying degrees of success. Considering the image segmentation problem as one of the optimization problems solved by metaheuristics, image segmentation has attracted many researchers in recent years. In many successful applications, it can be seen that the traditional approaches for image segmentation have been combined with metaheuristics in different perspectives in order to improve their performance.

Bearing those in mind, we propose in this work three image segmentation methods for magnetic resonance (MR) brain images based on mono-objective, multi-objective and hybrid metaheuristic optimization techniques. In each method, first, the basic model for the image segmentation problem is extended to incorporate more image information (spatial or spectral) such that more and better characteristics in segmented image can be achieved. Then, metaheuristic algorithms are adapted or developed to take place in optimization step. The proposed methods were evaluated on both simulated MR images and real MR images and compared with a set of recent methods in the literature. The obtained results show clearly the efficiency of the proposed approaches.

Keywords: Image segmentation, fuzzy clustering, region-based active contour, Markov random field, metaheuristics, multi-objective optimization, hybrid metaheuristic, MRI.

.

Résumé

La segmentation d'image est le processus de partitionnement d'une image en régions plus petites, non superposées et de taille significative, reposant en partie sur certaines caractéristiques d'homogénéité. De nombreuses tâches de traitement de haut niveau, telles que l'extraction de caractéristiques, la reconnaissance d'objets et le diagnostic médical, dépendent fortement de la qualité de la segmentation. En analyse d'images médicales, les images contiennent généralement des artefacts tels que le bruit, l'effet de volume d'image et l'effet de champ de polarisation dus à divers facteurs, comme l'environnement et les dispositifs d'acquisition; en outre, ces images possèdent des structures complexes. Par conséquent, la segmentation des images reste une tâche difficile, même si diverses techniques et méthodes de précision et de complexité diverses ont été proposées dans la littérature. Plusieurs approches, telles que le regroupement flou des pixels de l'image, les contours actifs basés basés régions, ainsi que les champs aléatoires de Markov, peuvent produire des résultats prometteurs. Cependant, de nombreuses questions restent à ce jour en attente de réponse. Jusqu'à maintenant, il n'existe pas de méthode de référence et les procédures de segmentation nécessitent toujours l'intervention d'un nombre important d'experts pour améliorer les performances.

Les métaheuristiques sont des procédures de haut niveau conçues pour résoudre les problèmes d'optimisation en cherchant des solutions sous-optimales acceptables à un problème complexe donné. Les métaheuristiques sont généralement appliquées à des problèmes pour lesquels il n'existe pas d'algorithme satisfaisant capable de les résoudre de manière efficace. Par conséquent, elles sont largement utilisées pour résoudre des problèmes complexes et elles ont fait leurs preuves dans de nombreux domaines d'application, avec plus ou moins de succès. Ces dernières années, de nombreux chercheurs se sont intéressés à la segmentation d'images en la considérant comme un problème d'optimisation, que l'on peut résoudre de manière approchée au moyen des métaheuristiques. Dans de nombreuses applications, les approches traditionnelles de la segmentation d'images ont été ainsi combinées avec des métaheuristiques, afin d'améliorer leurs performances.

Poursuivant cet objectif, nous proposons dans ce travail trois méthodes de segmentation dédiées aux images cérébrales par résonance magnétique (IRM), basées sur des techniques d'optimisation mono-objectif, multi-objectif et hybride, mettant en oeuvre des. Pour chaque méthode, le modèle de base du problème de segmentation d'image est d'abord étendu en incorporant davantage d'informations (de nature spatiale ou spectrale): de manière à obtenir des résultats de segmentation pertinents, qui conservent les caractéristiques initiales des images d'entrée. Ensuite, des algorithmes à base de métaheuristiques sont développés ou adaptés, afin de conduire la phase d'optimisation. Les méthodes proposées ont été évaluées à la fois sur des images IRM simulées et sur des images IRM réelles, en les comparant avec un ensemble de méthodes récentes de la littérature. Les résultats confirment sans ambiguïté la pertinence et l'efficacité des méthodes développées.

Mots-clés: Segmentation d'images, classification floue, contours actifs basés régions, champs aléatoires de Markov, métaheuristiques, optimisation multi-objectif, métaheuristiques hybrides, MRI.

Introduction générale

La segmentation d'images est le processus de partitionnement d'une image en zones ou régions connexes significatives, au sens d'un critère quantitatif donné, tel que le niveau de gris, la couleur, la texture ou une combinaison de ces derniers. Deux caractéristiques essentielles de la segmentation d'image sont l'homogénéité d'une région et la discontinuité marquant les régions adjacentes disjointes. En imagerie médicale, la segmentation revêt une importance considérable pour fournir des informations non invasives sur les structures du corps humain. Du point de vue de ce champ d'application, le besoin d'une segmentation robuste et précise des images médicales constitue une étape importante pour une large gamme de problèmes, notamment l'identification des organes du corps, à des fins de mise au point d'atlas de patients, d'étiquetage rapide en tomodensitométrie à rayons X, ou d'analyse d'informations issues de l'imagerie par résonance magnétique (IRM). D'autres applications médicales concernent l'imagerie cérébrale (pour guider les procédures d'intervention), l'analyse des poumons et des voies respiratoires environnantes (pour surveiller et traiter les maladies pulmonaires), la distinction des vaisseaux, des valves et des cavités dans les images de l'appareil cardiovasculaire (afin de quantifier l'étendue de la maladie), ou encore la segmentation des polypes observés en coloscopie par tomodensitométrie (afin de guider les procédures de traitement). La liste des applications potentielles peut à l'évidence être encore plus longue, compte tenu de la variété des problèmes traités.

Cependant, la segmentation des images médicales se heurte à de nombreux défis, qui doivent être surmontés, pour obtenir des résultats fiables et exploitables en milieu clinique. Le principal défi est que les différentes parties anatomiques du corps humain présentent des variations significatives de forme et d'apparence, causées par une multitude de facteurs: (1) bruits et artefacts des capteurs (inhérents à leur principe physique ou au processus de formation de l'image); (2) diversité morphologique des patients et variété de leurs mouvements (respiration, cycle cardiaque, circulation sanguine et cérébrospinale, péristaltisme et déglutition, et bien sûr les mouvements volontaires); (3) pathologie, chirurgie et agents de contraste (l'apparence des images sous différentes phases de contraste est variable); (4) balayage partiel et champ de vision; et (5) tissus mous (frontières floues entre les organes internes). Un autre défi réside dans la précision des traitements, la robustesse et les exigences en vitesse de calcul, découlant d'applications réelles en milieu clinique.

La résolution des problèmes de segmentation des images médicales a longtemps été considérée comme un problème fondamental, qui demeure une pierre d'achoppement de la recherche en analyse d'images médicales. Les premiers travaux dans ce domaine ont débuté il y a plusieurs décennies; ils ont souvent consisté à adapter à des problèmes d'imagerie médicale des techniques développées dans la vision par ordinateur et la reconnaissance de formes. Cependant, au cours des vingt dernières années, des chercheurs spécialisés dans l'analyse d'images médicales ont pu identifier de nombreux problèmes spécifiques à ce domaine en général, et à la segmentation d'images médicales en particulier. Les difficultés rencontrées incluent notamment la nature déformable des structures ou des régions sous-jacentes à segmenter, la variation statistique naturelle de ces structures ou des régions elles-mêmes, les paramètres qui pourraient en être dérivés, ainsi que le problème singulier de la segmentation des images médicales, qui implique souvent la connaissance de plusieurs structures, régions ou informations contextuelles, inhérentes aux images traitées.

Au fil du temps, de nombreuses techniques de segmentation, décrites dans la littérature, ont été développées. De manière générale, aucune technique de segmentation n'est universelle pour fonctionner

correctement avec toutes les applications, et diverses approches de la précision, de la rapidité et du degré de complexité ont été explorées, pour différents problèmes pratiques. Leur catégorisation est souvent basée sur des objectifs et aspects spécifiques du traitement des données impliquées. Parmi ces approches, il y a celles qui considèrent le problème de segmentation des images comme un problème d'optimisation, où la segmentation désirée minimise ou maximise une ou plusieurs fonctions d'énergie (ou de coût) définies pour le problème particulier. La résolution du problème d'optimisation, grâce aux métaheuristiques, a intéressé de nombreux chercheurs ces dernières années. Une métaheuristique est une procédure de haut niveau, conçue pour résoudre des problèmes d'optimisation réputés difficiles. En effet, les métaheuristiques sont en général conçues pour résoudre des problèmes complexes, pour lesquels il n'existe pas d'algorithmes spécifiques capables de les résoudre de manière satisfaisante. Divers algorithmes ont été proposés pour résoudre de cette façon le problème de la segmentation d'images. Ces algorithmes ont été expérimentés dans de nombreuses applications, avec des degrés divers de réussite, et de nombreuses questions restent à ce jour en attente de réponse. En particulier, les questions encore ouvertes sont: (1) comment construire des fonctions objectifs optimisées, reflétant de manière satisfaisante les caractéristiques des images traitées?; (2) quel est l'algorithme efficace pour trouver le meilleur résultat, les fonctions objectifs étant généralement non-convexes, par nature non uniques, et susceptibles de comporter de nombreux minimums locaux?

Grâce au développement rapide des technologies d'imagerie, il est désormais possible d'explorer de manière non invasive des organes in vivo du corps humain. En particulier, pour explorer l'anatomie du cerveau et étudier son fonctionnement. Dans le diagnostic des troubles cérébraux, l'IRM est couramment utilisée: cette technique peut fournir une grande quantité de données, avec un niveau de qualité de plus en plus élevé. Cependant, les images d'IRM du cerveau contiennent encore des artefacts, dus au bruit, la résolution spatiale (taille du plus petit détail détectable) et des biais liés à divers facteurs, tels que l'environnement et les dispositifs d'acquisition; en outre, ces images peuvent avoir des structures complexes. Leur analyse constitue une tâche fastidieuse et complexe pour les cliniciens. De ce fait, il existe un grand besoin de méthodes capables d'interpréter les données d'IRM avec précision. Pour contribuer à ce domaine, nous proposons dans cette thèse trois méthodes de segmentation dédiées aux images d'IRM du cerveau.

Cette thèse a été préparée au sein du Laboratoire Images, Signaux et Systèmes Intelligents (LiSSi, E.A. 3956) de l'Université Paris-Est-Créteil (UPEC). Le travail de recherche a été réalisé sous la direction du professeur Patrick SIARRY et du Dr. Hamouche OULHADJ, tous deux membres permanents du groupe SIMO (Signal, Image et Optimisation) du LiSSi, qui s'intéresse au développement du traitement d'image et des techniques d'optimisation.

Objectifs de l'étude et contributions majeures

L'objectif principal de ce travail de recherche est de développer de nouvelles méthodes de segmentation d'images, de grande précision et dédiées spécifiquement aux images d'IRM du cerveau, en s'appuyant sur la mise en oeuvre d'algorithmes robustes à base de métaheuristiques d'optimisation. Nous nous attachons à démontrer l'efficacité des méthodes développées en les confrontant à un ensemble de techniques de segmentation concurrentes, disponibles dans la littérature.

Les objectifs de base de ce travail sont les suivants:

- Proposer de nouvelles solutions de segmentation d'images, en exploitant de nouvelles informations (de nature spatiale ou spectrale) dans les images à traiter.
- Adapter ou développer des algorithmes s'appuyant sur des métaheuristiques d'optimisation, dédiés aux images d'IRM du cerveau, de manière à permettre une interprétation plus précise et

plus fiable des résultats.

- Pour chaque méthode, analyser qualitativement et quantitativement les résultats produits, en impliquant des bases de données standard de la littérature, afin de mieux évaluer leur performance et leur portée.

- Confronter les méthodes proposées à un ensemble de méthodes de segmentation concurrentes, afin de démontrer leur capacité à surmonter, tout au moins partiellement, les limites des approches existantes.

Les principales contributions de cette thèse peuvent être résumées par les points qui suivent:

- Dans le cadre de l'optimisation mono-objectif, nous avons développé une nouvelle méthode de segmentation des images d'IRM du cerveau, en mettant en oeuvre un algorithme d'optimisation par essaims de particules (PSO) et une approche de regroupement des pixels de l'image (*clustering*, en anglais) par le biais du concept d'entropie floue. La méthode que nous avons mise au point surpasse un ensemble de méthodes récentes de la littérature. L'algorithme proposé a été validé en impliquant à la fois des images simulées d'IRM du cerveau et des images réelles.

- Dans le cadre de l'optimisation multi-objectif, nous avons mis au point une nouvelle méthode de segmentation, qui exploite simultanément deux critères complémentaires, à savoir la compacité et la séparation. Cette méthode de segmentation est organisée autour d'un *clustering* à noyau à base d'entropie floue (avec une information spatiale locale et une correction de biais) et un poids d'énergie adaptatif combiné au modèle *Global Energy Fit*. En outre, une nouvelle variante de PSO multi-objectif a été développée pour déterminer l'estimateur optimal. La méthode proposée a été également validée sur la base, à la fois, d'images simulées d'IRM du cerveau et d'images réelles.

- Enfin, en matière d'hybridation de métaheuristiques, une nouvelle méthode de segmentation a été élaborée, en combinant une variante du modèle des champs aléatoires de Markov cachés et une nouvelle métaheuristique hybride. Celle-ci combine deux autres métaheuristiques classiques, qui sont PSO et *Cuckoo Search*. La méthode proposée est validée en impliquant des images simulées d'IRM du cerveau et des images réelles. L'étude expérimentale a montré que notre approche produit des résultats plus rapides, avec une segmentation de meilleure qualité.

Organisation de la thèse

La thèse est structurée en quatre chapitres. L'objectif du premier chapitre est d'introduire les concepts exploités dans nos travaux de recherche; chacun des chapitres suivants détaille les méthodes que nous avons développées et leur validation expérimentale. L'organisation générale et le contenu de ces différents chapitres sont donnés dans la suite de ce rapport.

Le chapitre 1 traite de la description formelle du problème de segmentation des images d'IRM du cerveau, dans le cadre général d'une approche reposant sur l'optimisation. Nous présentons en outre un état de l'art des métaheuristiques d'optimisation existantes, des approches de segmentation d'image couramment utilisées, ainsi que de leurs améliorations au moyen de métaheuristiques.

Le chapitre 2 expose notre première contribution, qui concerne l'amélioration de la classification floue pour la segmentation des images d'IRM du cerveau, à l'aide de métaheuristiques d'optimisation. Nous améliorons d'abord la fonction de coût, reposant sur le modèle de classification à base du concept d'entropie floue, puis nous développons une variante de PSO, pour déterminer l'estimateur optimal.

Le chapitre 3 décrit une extension de la première contribution à une approche d'optimisation multi-objectif, qui exploite un critère complémentaire, à savoir un contour actif basé-région. Les deux critères sont optimisés simultanément, au moyen d'un algorithme impliquant un PSO multi-objectif amélioré. Ainsi, deux caractéristiques indépendantes, qui sont la compacité et la séparation, sont prises en considération dans le processus de segmentation des images.

Le chapitre 4 est consacré à une amélioration ultime, obtenue en exploitant les résultats de la contribution précédente et une nouvelle variante du modèle des champs aléatoires de Markov cachés. Une métaheuristique hybride est mise en oeuvre dans l'étape d'optimisation.

Enfin, ce rapport de thèse se termine par une conclusion générale, dans laquelle nous récapitulons nos contributions et nous proposons des perspectives de développement du présent travail, compte tenu des résultats actuels obtenus.

General introduction

Image segmentation is the process of partitioning an image space into non-overlapped meaningful homogeneous regions or objects, according to given quantitative criteria: gray level, color, texture or combination of them. Two main properties that segmentation methods are usually based on are: the homogeneity of a region and the discontinuities characterising adjacent disjoint regions. In medical imaging, segmentation is of considerable importance in providing non-invasive information about human body structures. From this application field standpoint, the need for robust and accurate medical image segmentation is an important step for a wide range of problems. They could be: (1) the identification of the body organs for use in the development of patient atlases for rapid labelling of X-ray Computed Tomography (CT) or Magnetic Resonance Imaging (MRI) derived information; (2) finding brain neuroanatomy for guiding interventional procedures; (3) parsing out the lungs and surrounding airways to monitor and treat lung disease; or (4) isolating vessels, valves and chambers in the cardiovascular system for quantifying the extent of disease and polyp segmentation from CT colonoscopies to guide treatment procedures. This list is obviously much longer, and includes problems from a variety of imaging scales to visualizing and capturing the signature of individual cells.

However, medical image segmentation confronts several challenges to obtain results that can be used in clinical applications. The main challenge is that the different anatomical parts of the human body exhibit significant shape and appearance variations caused by a multitude of factors: (1) sensor noise/artifact (inherent to its own physical sensor, image formation process, modality and imaging configuration); (2) diversity of patients and variety of their movements (build forms; motions from respiration, cardiac cycle, blood and cerebrospinal fluid flow, peristalsis and swallowing, and voluntary movement); (3) pathology, surgery, and contrast agents (image appearances under different contrast phases are different); (4) partial scan and field of view; and (5) soft tissue (weak boundary between internal organs). Another challenge lies in stringent accuracy, robustness and speed requirements arising from real clinical applications.

Solving the problems of medical image segmentation has long been thought of as a basic touchstone of research in medical image analysis. The early work in this area, that began several decades ago, often consisted of adapting techniques and strategies developed in the computer vision and pattern recognition communities to problems in medical imaging. However, over the past 20 years, researchers dedicated to working specifically in the area of medical image analysis have been able to articulate and identify many of the issues specific to this field in general, and medical image segmentation in particular. Difficulties to overcome include noting the deformable nature of the underlying structures or regions to be segmented, the natural statistical variation of these structures or regions themselves and any parameters that might be derived from them, as well as the singular problem of the segmentation of medical images, which often involves the knowledge of several structures, regions or contextual information inherent to the processed images.

Throughout the years, many segmentation techniques have been developed and reported in the literature. Generally, no segmentation technique works well for all the applications, and various approaches of different accuracy, speed, and degree of complexity have been explored for different practical problems. Their categorization is often based on goals and specific aspects of data processing involved. Among those, the approach which considers the image segmentation problem as one of the optimization problems, where the desired segmentation minimizes or maximizes one or some

energy (or cost) functions defined by the particular application. Solving the optimization problem, using metaheuristics, has attracted many researchers in recent years. Metaheuristics are high-level procedures designed to solve optimization problems by the process of searching acceptable suboptimal solutions to a particular problem of interest. This family of algorithms is generally applied to problems for which there is no satisfactory algorithm able to solve them effectively. Even though various techniques and methods for the image segmentation problem have been proposed and proven to be successful in many applications with different degrees of success, many key open issues still remain to be investigated. Such issues might be, for instance, the following problems: (1) how to construct objective functions being optimized that reflect satisfactory characteristics of segmented results; (2) what is the efficient algorithm for finding the optimal estimator since the objective functions are generally non-convex, non-unique in nature, and may have several local minimum points.

Thanks to the rapid development of imaging technologies, now living organs and organisms of human body can be explored non-invasively. Among those, to analyse and explore brain anatomy and function in diagnosis of brain disorders, MRI is commonly used since it can provide a large amount of data with an increasingly high level of quality. However, brain MR images still contain some artifacts such as noise, image volume effect and bias field effect due to various factors, for instance, the environment and the acquisition devices and have complex structures. Analysing of these images becomes a tedious and complex task for clinicians; therefore, there is a great need for methods that process with the interpretation of the data with high accuracy. To alleviate these difficulties, in the present research work, we propose three segmentation methods for brain MR images.

This thesis was prepared within the Laboratoire Images, Signaux et Systèmes Intelligents (LiSSi, E.A. 3956) from the Université Paris-Est Créteil (UPEC). This research work has been proposed and carried out under the supervision of Professor Patrick SIARRY and co-supervised by Dr. Hamouche OULHADJ in the SIMO group (Signal, Image and Optimization) of LiSSi. This team is interested in the development of both image processing and optimization techniques.

Studied objectives and major contributions

The core objective of this work is to create new variants of image segmentation methods for brain MR images with high accuracy, thanks to the contributions of metaheuristics. To verify the effectiveness of the methods developed, we compared them with a set of recent segmentation techniques available in the literature.

The basic objectives pursued by this research work are:

- To propose new variants of image segmentation criteria (energy or cost functions) that better reflect the characteristics of the input images by incorporating additional information (spatial or spectral) in the segmentation process.
- To adapt or develop segmentation algorithms based on optimization metaheuristics, and dedicated to MR images of the brain, with more reliable results, easier to analyze and interpret.
- To examine both qualitatively and quantitatively the segmented results, using the standard databases of the literature to better evaluate the scope and the performance of the methods developed.
- To confront the proposed methods with a set of competing segmentation methods of the literature to demonstrate their ability to overcome, at least partially, the limitations of existing approaches.

The major contributions of the thesis can be summarized as follows.

- In mono-objective metaheuristic optimization approach, a new segmentation method for brain MR images using improved Particle Swarm Optimization (PSO) and a new variant of fuzzy entropy clustering model is proposed. The performance is found better than a set of recent methods in the literature. The proposed algorithm is validated by using both simulated and real brain MR images.

- In multi-objective metaheuristic optimization approach, a new segmentation method that simultaneously optimizes two complementary properties, namely compactness and separation, in the segmentation process, is proposed. The two criteria are based on kernelized fuzzy entropy clustering with local spatial information and bias correction, and adaptive energy weight combined with global and local fitting energy active contour model. In addition, a new variant of multi-objective particle swarm optimization algorithm is proposed for finding the optimal estimator. The proposed method is also validated by using both simulated and real brain MR images.

- In hybrid metaheuristic approach, a new segmentation method, that combines a new variant of hidden Markov random field model and a novel hybrid metaheuristic method, is proposed. The hybrid metaheuristic algorithm developed is based on two well-known metaheuristic algorithms, namely PSO and Cuckoo search. The proposed method is validated by using both simulated and real brain MR images. Experiments show that faster and better segmented results can be obtained.

Thesis organization

The thesis is structured into four chapters. The first chapter is devoted to information in support of this research. The next chapters explain in detail each of the proposed methods. We detail below the general organization of these chapters:

Chapter 1 deals with the formal description of the problem of segmentation, brain MR image segmentation, in the optimization-based framework. Furthermore, a bibliographic study is presented in which there are various metaheuristic optimization techniques, major segmentation approaches commonly used as well as their improvements by combination with metaheuristics.

Chapter 2 presents our first contribution, which is the improvement of fuzzy clustering approach for brain MR image segmentation, using optimization metaheuristics. Particularly, first, fuzzy entropy clustering model is improved and considered as the cost function, then an improved PSO is adapted to find the optimal estimator.

Chapter 3 describes an extension of the first contribution to a multi-objective optimization approach, where a complementary criterion, that implements a region-based active contour, is introduced. Both segmentation criteria are simultaneously optimized by using an improved multi-objective particle swarm optimization algorithm such that two independent characteristics in segmented results, namely compactness and separation, can be achieved.

Chapter 4 is dedicated to a further enhancement of the segmentation, using the results of the second contribution and a new variant of hidden Markov random ?eld model. Hybrid metaheuristic approach is used in the optimization step.

Finally, the thesis ends with a general conclusion, in which we summarize our contributions and propose perspectives, taking into account the current results obtained.

Chapter I

State of the art: image segmentation and optimization metaheuristics

I.1 Introduction

Image segmentation is the process of partitioning an image space into non-overlapped meaningful homogeneous regions or objects, according to given quantitative criteria: gray level, color, texture or combination of them. For medical image analysis, the success of an image analysis system depends heavily on the quality of the segmentation process. Among different medical image modalities such as X-ray, Ultrasound (US), Computed Tomography (CT) and Positron Emission Tomography (PET), Magnetic Resonance Imaging (MRI) has become a leading technique, widely used for imaging human soft tissue anatomy. Its applications are extended over all parts of the human body, and it presents the most common method of human brain exam. However, the existence of noise, low contrast, intensity non-uniformity, and the complexity of objects' structure are critical obstacles that stand in the way of achieving an ideal segmentation system. Over the past decades, several approaches have been proposed to obtain desirable segmentation results. A class of rather promising approaches reformulates the problem of image segmentation into an optimization problem, such as energy minimization or maximum-a-posteriori estimation, mainly because of: (1) rigorous and formal mathematical formulation; (2) availability of mathematical tools for optimization; (3) capability to incorporate several criteria as terms in the objective function; (4) ability to examine the relative performance of different solutions; and (5) availability of quantitative metrics to measure the extent by which a method satisfies different criteria [Nosrati and Hamarneh, 2016]. Since the energy fitting functions, which contain a large amount of information or image characteristics, are non-convex, non-unique in nature and may have several local minimum points, using optimization metaheuristic techniques is a promising approach to solve efficiently the problem.

In this chapter, we present the basic information of MRI technique in Section I.2. Then, the formulation of the image segmentation problem is given in Section I.3. Without claiming completeness, we provide the major concepts related to optimization metaheuristics in Section I.4. In Section I.5, different methods for the image segmentation problem, as well as their improvements by combination with optimization metaheuristics, are described. Finally, Section I.6 presents a conclusion summarizing what we wish to develop to achieve a more efficient segmentation.

I.2 Magnetic resonance imaging

I.2.1 MRI principles

The principles of magnetic resonance (MR) imaging is based on the absorption and emission of energy in the radio frequency range of the electromagnetic spectrum. The basic objective is to map

the spatial location and associated properties of specific nuclei or protons present in the object being imaged.

In the human body, hydrogen atoms, which have a fundamental property of odd atomic weight and/or odd atomic numbers with the possession of an angular moment called spin, contain a MR signal. These protons carry an electric charge and spins around their axes. The charged protons create a magnetic field around them and thus they act as tiny magnets, possessing both an angular moment and a magnetic moment. The magnetic moment is proportional to the spin angular moment and is related through a constant, called gyro-magnetic ratio, a quantum property of the proton. The relationship is given by:

$$m = \xi \mathrm{J} \tag{I.1}$$

where m and J are the magnetic moment and the spin angular moment, respectively; and ξ is gyro-magnetic ratio defined in MHz/T.

The atomic numbers have their orientations corresponding to energy levels with or without the present of an external magnetic field. Without an external magnetic field, the direction of the magnetic moments of the spinning protons is completely random. In contrast, if an external magnetic field exists, the magnetic moments of nuclei result in a nuclear paramagnetic polarization with specific orientations (either aligned along the external magnetic field or against the field) and energy levels as characterized by their spin quantum number.

Using the principle of classical mechanics, the torque generated by the interaction of the magnetic moments and the external magnetic field is equal to the rate of change of angular momentum, which is given by:

$$\frac{d\overrightarrow{\mathrm{J}}}{dt} = m \times \overrightarrow{\mathrm{H}}_0 = m \times \mathrm{H}_0 \overrightarrow{k} \tag{I.2}$$

where H_0 is the strength of the external magnetic field and $\overrightarrow{k}$ is the unit vector along the z-direction. The solutions of Eqs. (I.1) and (I.2) lead to an important relationship that provides the angular frequency, ω_0:

$$\omega_0 = \xi \mathrm{H}_0 \tag{I.3}$$

The relationship provided in Eq. (I.3) is known as the Larmor equation. From this equation, we can state that the possession frequency, ω_0, depends on the type of nuclei (determined by ξ) and the intensity of the external magnetic field (H_0). This frequency is the frequency at which the protons can receive the Radio Frequency (RF) energy to change their states for exhibiting magnetic resonance. The energy is provided by a RF electromagnetic coil that transmits an oscillating RF wave at the Larmor frequency to cause nuclear excitation. After the RF pulse is turn off, the excited nuclei go through a relaxation phase emitting the absorbed energy at the same Larmor frequency that can be detected as an electrical signal, called the Free Induction Decay (FID). FID is the basic signal that is used to create MR images and can be acquired through the same RF coil tuned at the Larmor frequency.

Assuming that $\mathcal{N}$ is the total number of spinning nuclei in the object being imaged, a stationary magnetization vector $\overrightarrow{\mathrm{M}}_s$ can be defined as follows:

$$\overrightarrow{\mathrm{M}}_s = \sum_{i=1}^{\mathcal{N}} \overrightarrow{m}_i \tag{I.4}$$

with the stationary magnetization vector, $\overrightarrow{\mathrm{M}}_s$, and a rotating magnetization vector, $\overrightarrow{\mathrm{M}}_r$ defined

as follows:

$$\begin{aligned}\overrightarrow{\mathrm{M}}_s &= \mathrm{M}_x\overrightarrow{i} + \mathrm{M}_y\overrightarrow{j} + \mathrm{M}_z\overrightarrow{k} \\ \overrightarrow{\mathrm{M}}_r &= \mathrm{M}'_x\overrightarrow{i}' + \mathrm{M}'_y\overrightarrow{j}' + \mathrm{M}'_z\overrightarrow{k}'\end{aligned} \tag{I.5}$$

the relationship between them can be expressed as follows:

$$\frac{d\overrightarrow{\mathrm{M}}_s}{dt} = \frac{\partial\overrightarrow{\mathrm{M}}_r}{\partial t} + \omega \times \overrightarrow{\mathrm{M}}_r \tag{I.6}$$

where ω is the angular frequency at which the transverse plane rotates. The rate of change in the net stationary magnetization vector can be described as follows (Bloch equation):

$$\frac{d\overrightarrow{\mathrm{M}}_s}{dt} = \xi\overrightarrow{\mathrm{M}}_s \times \overrightarrow{\mathrm{H}} \tag{I.7}$$

where $\overrightarrow{\mathrm{H}}$ is the net effective magnetic field. Considering the total respond of the spin system in the presence of an external magnetic field along with the RF pulse for nuclei excitation followed by the nuclear relaxation phase, the change of the net magnetization vector can be expressed as follows:

$$\frac{d\overrightarrow{\mathrm{M}}_s}{dt} = \xi\overrightarrow{\mathrm{M}}_s \times \overrightarrow{\mathrm{H}} - \left(\frac{\mathrm{M}_z\overrightarrow{i} - \mathrm{M}_y\overrightarrow{j}}{T2}\right) - \frac{(\mathrm{M}_z - \mathrm{M}_z^0)\overrightarrow{k}}{T1} \tag{I.8}$$

where M_z^0 is the net magnetization vector in thermal equilibrium in the presence of H_0. $T1$ and $T2$ are the longitudinal (spin-lattice) and the transverse (spin-spin) relaxation times, respectively. The magnetic flux $\phi(t)$ through the RF coil can be given as follows:

$$\phi(t) = \int_{\text{object}} \overrightarrow{\mathrm{H}}(s)\overrightarrow{\mathrm{M}}_s(s,t)ds \tag{I.9}$$

where s is a spatial location vector in the spinning nuclei system. Using the Faraday's law of electromagnetic induction, the voltage, $v(t)$, induced in the RF coil is the raw MR signal, which can be expressed as follows:

$$v(t) = -\frac{\partial\phi(t)}{\partial t} = -\frac{\partial}{\partial t}\int_{\text{object}} \overrightarrow{\mathrm{H}}(s)\overrightarrow{\mathrm{M}}_s(s,t)ds \tag{I.10}$$

The realization of spatial location, s, is responsible to create a MR image that maps the magnetic resonance response of the spinning nuclei available in that location. At last, generating MR signal is obtained through encoding techniques.

I.2.2 Effecting factors on MR signal

There are several factors that affect the MR signal acquisition causing artifacts or degradation in the reconstructed image. These factors include noise, field inhomogeneities, flow of nuclei and change in resonance parameters due to chemical shifts within the object.

- **Noise**: various sources, which include many external causes in the patient (Brownian motion of ions in bodily fluids), transmission system and environmental factors (medical devices located anywhere in the MR procedure room), can influence on MR signal. They result in several types of noise such as Gaussian, Poisson, blurred, speckle and salt-and-pepper noise existing in MR image [Kumar and Nachamai, 2012].

- **Magnetic field inhomogeneities**: field inhomogeneities and gradient fields also cause a direct dephasing effect to the transverse relaxation process [Simmons et al., 1994].
- **Magnetic susceptibility**: the presence of other substances in the imaging medium is another factor influencing the relaxation times.
- **Chemical shifts**: the presence of a molecular or chemical environment, which can change the magnetic characteristics influence on the protons, can cause a deviation in the Larmor frequency of spinning protons.

Though there are some artifacts existing in MR image, MRI technique has some excellent characteristics such as high soft tissue contrast, non-invasive character, high partial resolution and easy slice selection in any orientation. As a result, it is used for many applications such as Alzheimer disease, Parkinson or Parkinson related syndrome, congenital brain malformations or perinatal brain damage, post-traumatic syndrome, etc.

Note that, among three types of MR images, named T1-weighted, T2-weighted and Photon Density (PD), T1-weighted MR images are used for the validation purpose in this study. The main reason is that T1-weighted images offer the highest contrast between the brain soft tissues.

I.3 Formulation of the image segmentation problem

Formally, the segmentation of an image can be reduced to an optimization problem, which can be stated as follows. Given a feature image $\mathbf{F}$ and an uniformity criterion Γ, define predicate:

$$P(\mathbf{F}) = \text{True}, \quad \text{if} \quad \exists \varepsilon \ni |\Gamma(i,j) - \varepsilon| < \xi, \forall (i,j) \in \mathbf{F} \tag{I.11}$$

Partition image into C subsets $\mathbf{Y} = (\Omega_i)_C$ such that

- Completeness: $\mathbf{F} = \cup_{i=1}^{C} \Omega_i$
- Disjointness: $\Omega_i \cap \Omega_j = \oslash$ for $i \neq j$
- Satisfiability: $P(\Omega_i) = \text{True}, \forall i$
- Segmentability: $P(\Omega_i \cup \Omega_j) = \text{False}, \forall i \neq j$

We also observe that the uniqueness of the segmentation is not guaranteed by these four conditions, because the segmentation results depend not only on the information contained in the image, but also the method used to process that information. Generally, to reduce the problem of non uniqueness of the solution, the segmentation problem is regularized by an optimization constraint of a function $\mathcal{H}$, characterizing the quality of a good segmentation. As a result, a fifth condition is added [Capelle et al., 2003]:

$$\mathcal{H}(\mathbf{F}, \mathbf{Y}^*) = \min_{\mathbf{Y} \in \mathbf{\Omega}_\mathrm{Y}} / \max_{\mathbf{Y} \in \mathbf{\Omega}_\mathrm{Y}} \mathcal{H}(\mathbf{F}, \mathbf{Y}) \tag{I.12}$$

where $\mathcal{H}$ is a function and $\mathbf{\Omega}_\mathrm{Y}$ is the set of possible partitions.

It is clear that such condition (I.12) does not entirely solve the problem of uniqueness of segmentation. There are still cases where several segmentations can have the same optimal value [Sulaiman et al., 2014]. Hence, to obtain an acceptable segmentation of the image, several criteria need to be satisfied, such as compactness, separation, and non-overlapping. In other words, there is no single and sufficient criterion for optimally segmenting all images.

There are two major approaches for formulating the image segmentation problem as an energy optimization problem: *regularization* and *Bayes' theorem*. In the regularization approach, an energy function, $\mathcal{J}_{\mathrm{R}} : \mathbf{Y} \longrightarrow \Re$, usually consists of several objectives that are divided into two main categories: *regularization terms*, $\mathcal{R}_{\mathrm{i}} : \mathbf{Y} \longrightarrow \Re$, and *data terms*, $\mathcal{D}_{\mathrm{i}} : \mathbf{Y} \longrightarrow \Re$. The regularization terms correspond to priors on the space of feasible solutions and penalize any deviation from the enforced prior such as shape, length, etc. The data terms measure the strength with which a pixel should be associated with a specific label/segment. These objectives can be derived as:

$$\mathcal{J}_{\mathrm{R}}(\mathbf{F}, \mathbf{Y}) = \lambda \sum_{i} \mathcal{R}_{\mathrm{i}}(\mathbf{Y}) + \eta \sum_{i} \mathcal{D}_{\mathrm{i}}(\mathbf{F}, \mathbf{Y}) \tag{I.13}$$

where λ and η are constants that balance the contribution/importance of the regularization term and the data term in the problem.

The image segmentation problem is then formulated as follows:

$$\mathbf{Y}^{*} = \underset{\mathbf{Y} \in \mathbf{\Omega}_{\mathrm{Y}}}{\operatorname{argmin}} \,/\, \underset{\mathbf{Y} \in \mathbf{\Omega}_{\mathrm{Y}}}{\operatorname{argmax}} \mathcal{J}_{\mathrm{R}}(\mathbf{F}, \mathbf{Y}) \tag{I.14}$$

An optimization-based image segmentation problem can also be formulated as a maximization problem using Bayes' theorem.

$$\mathbf{Y}^{*} = \underset{\mathbf{Y} \in \mathbf{\Omega}_{\mathrm{Y}}}{\operatorname{argmax}} P(\mathbf{Y}|\mathbf{F}) = \underset{\mathbf{Y} \in \mathbf{\Omega}_{\mathrm{Y}}}{\operatorname{argmax}} P(\mathbf{F}|\mathbf{Y}) P(\mathbf{Y}) \tag{I.15}$$

where $P(\mathbf{Y}|\mathbf{F})$ is the *posteriori probability* that defines the degree of belief in $\mathbf{Y}$ given the evidence $\mathbf{F}$. $P(\mathbf{F}|\mathbf{Y})$ is the image *likelihood* measuring the probability of the evidence in $\mathbf{F}$ given the segmentation $\mathbf{Y}$, and $P(\mathbf{Y})$ is the *priori probability* that indicates the initial (prior to observing $\mathbf{F}$) degree of belief in $\mathbf{Y}$. For more details, we refer the reader to Section IV.2.1

Thus, the image segmentation problem can be mathematically formulated in different ways depending on the approach and image information used. Generally, the objective functions are complex, non-convex, non-unique in nature and may have several local minimum points. Consequently, two major problems have arisen: (1) How to obtain appropriate models as well as to benefit image information for the problem? (2) What is the efficient algorithm for finding the optimal estimator to avoid stagnation at local minima? These problems will be partially solved in this study.

I.4 Optimization metaheuristics

Optimization can exist in almost every aspects of life including engineering, industry, business or even social science. Metaheuristics are widely recognized as efficient approaches for optimization problems of such areas. Because of simplicity and robustness of produced results, a great interest has been devoted to metaheuristics. In what follows, first, we present a short introduction related to optimization metaheuristics; then, we briefly describe the major areas of study in this field.

I.4.1 Metaheuristic algorithms

Metaheuristics are generally applied to problems for which there is no satisfactory problem-specific algorithm to solve them. In order to find desirable solutions, for the multi-population metaheuristics, the optimization process begins with the creation of an initial group of random solutions (population) that satisfy the restriction of the problem to be solved. Then, the set containing all the

solutions is iteratively evaluated by one or more target functions associated with the problem and one iterates over generations to minimize or maximize the objective(s). These iterations run until the solution found meets some pre-defined criteria. This final solution (near optimal solution) is said to be an "optimal" solution and the system reaches a converged state [Hussain et al., 2018]. Note that, though this procedure is quite simple, finding solutions for real-life problems requires considering and addressing several issues, from which the most important are: local optimal avoidance, computational cost of function evaluation, constraints handling, multiplicity of objective functions, and uncertainties.

Basically, to search for an optimal solution of a given problem, a metaheuristic algorithm can be formulated as:

$$\min_{\mathbf{X}\in\mathbf{\Omega}_{\mathcal{X}}} / \max_{\mathbf{X}\in\mathbf{\Omega}_{\mathcal{X}}} \mathcal{Q}(\mathbf{X}), \quad \mathbf{X} = (x_1, \ldots, x_C) \in \Re^C \tag{I.16}$$

where $\mathbf{X}$ is the design vector that encodes C decision variables of the problem, usually, a design vector is also called candidate solution. $\mathbf{\Omega}_{\mathcal{X}}$ is the feasible search space or solution space limited by the lower and upper bounds.

A metaheuristic will be successful on a given optimization problem if it can provide a balance between two cornerstones features: the exploration and the exploitation (also referred to as diversification and intensification, respectively). Exploration is the ability to expand search in wide spread domain to explore unvisited areas such that parts of search space with high-quality solutions are identified. Exploitation, via accumulated search experience, is important to intensify the search in some promising areas. The main differences among the existing metaheuristics concern the particular way in which they try to achieve this balance [Birattari et al., 2001]. Some promising approaches, which can be used to maintain the trade-off balance, are: parameter tuning, population size control, and diversity maintenance through deterministic, adaptive, and self-adaptive techniques.

Almost all metaheuristics share the following characteristics: they are nature-inspired (based on principles of physics, biology or ethology); they make use of stochastic elements (random variables); they do not use the gradient or Hessian matrix of the objective function; existing several parameters need to be configured to the problem at hand [BoussaïD et al., 2013]. The classification of metaheuristics can be performed with respect to different aspects concerning the search path they follow, the use of memory, the kind of neighbourhood exploration used or the number of solutions maintained during iterations. For a formal classification of metaheuristics, we refer the reader to [Talbi, 2009].

In order to analyse the performance of metaheuristic algorithms, researchers have commonly used benchmark test functions such as those used in [Héliodore et al., 2017]. Then, some validation criteria such as best, worst, mean and standard deviation of objective function values obtained over specific number of runs are used to make a comparison. For specific engineering problems, which are also solved by using metaheuristics (for instance the image segmentation problem), the quality of the final solution is evaluated by using validation criteria in those domains (such as Dice coefficient, Hausdorff distance, etc.).

The idea of solving optimization problems through heuristic approaches was envisioned more than forty years ago when Operations Research was in its infancy during World War II. The formal kick off of metaheuristic research took place when initial metaheuristic methods like Simulated Annealing and Tabu Search were introduced in 1980s. However, the boom of this field of research was witnessed in 1990s after the wider applications of Genetic algorithms (GA), Ant colony optimization (ACO) and Particle swarm optimization (PSO). Despite the success of metaheuristic methods on diversified areas of science, engineering and technology, there remains a sufficient gap that needs to be filled in order to reach maturity level as compared to other established fields of research.

Now that the above description has already established preliminary knowledge about meta-

heuristics, the upcoming sections explore more about some attractive categories of metaheuristics.

I.4.2 Single-solution based metaheuristics

Single-solution based metaheuristics, also called *trajectory methods*, are based on the evolution of a single solution during the search process. Typically, these methods start with a single initial solution and move away from it, describing a trajectory in the search space. Usually, basic single-solution based algorithms are more exploitation oriented. Many methods in this category can be found in the literature; however, in this section we present two common ones used not only in image segmentation but also in other applications, named Simulated Annealing (SA) and Tabu search (TS).

I.4.2.1 Simulated annealing

Simulated annealing (SA) is a stochastic optimization technique, first introduced by Kirkpatrick [Kirkpatrick et al., 1983] and independently by Cerny [Černỳ, 1985]. SA is inspired by the annealing technique in which high temperature metal if cooled at an appropriate cooling rate will reach an absolute minimum energy state related to complete atomic ordering of metal. If the high temperature metal is cooled at a fast rate, the atoms will reach a sub-optimal energy state. The hypothesis of this method is that: system energy at higher temperature (T) is allowed to move uphill as well as downhill, but as temperature (T) goes down gradually, the energy is allowed to move downhill only. Thus, evolution of states is sensitive to coarser energy variation when (T) is large and to finer variation when (T) is small.

Transposing the process of annealing to the optimization process is based on the following analogies: the objective function to be optimized is similar to the energy of a material, and the temperature is represented by a controllable parameter defining the cooling scheme.

The algorithm begins with selecting an initial solution and later generating a new state, randomly generating a new solution ($\mathbf{X}' = \mathbf{X} + \triangle\mathbf{X}$) in the neighbourhood of the current solution ($\mathbf{X}$); this is called a neighbour solution. This new state is evaluated and compared with the previous solution in terms of fitness value ($\mathfrak{f}$). If the solution of the new state is better than the previous one, it is accepted; but if it is not, it is accepted with a probability, $P(T, \mathfrak{f}_{\mathbf{X}'}, \mathfrak{f}_{\mathbf{X}}) = \exp(-\frac{\mathfrak{f}_{\mathbf{X}'} - \mathfrak{f}_{\mathbf{X}}}{T})$. The temperature ($T$) is gradually decreased during the search process. By repeatedly following this Metropolis rule of acceptance, a sequence of solutions is generated, which constitutes a Markov chain (in the sense that each solution depends on only that one which immediately precedes it). With this formalism in place, it is possible to show that, when the chain is of infinite length, the system can reach (in practice, can approach) an equilibrium point at the temperature considered. The process is terminated when the system is "solidified" (either the temperature has reached zero or no more moves causing an increase in energy have been accepted).

It is clear that, at high temperature, $\exp(-\frac{\mathfrak{f}_{\mathbf{X}'} - \mathfrak{f}_{\mathbf{X}}}{T})$ is close to 1, and therefore the majority of the moves are accepted and the algorithm becomes equivalent to a simple random walk in the configuration space. However, at low temperature, $\exp(-\frac{\mathfrak{f}_{\mathbf{X}'} - \mathfrak{f}_{\mathbf{X}}}{T})$ is close to 0, and therefore the majority of the moves that increase the energy are rejected. At an intermediate temperature, the algorithm intermittently allows transformations that degrade the objective function: hence it leaves a chance for the system to be pulled out of a local minimum.

Note that the cooling schedule for T is critical to the efficiency of SA. If T is reduced too rapidly, a premature convergence to a local minimum may occur. In contrast, if it is too slow, the algorithm is very slow to converge. Given a "sufficiently large" number of iterations at each temperature, SA is proved to converge almost surely to the global optimum [Hajek, 1988]. The Simulated Annealing

method is presented in **Algorithm** I.1.

Algorithm I.1: The general SA algorithm

Initialization: Initialize randomly a solution $\mathbf{X}$ for the system to be optimized; initialize the temperature T
Results : The optimal solution $\mathbf{X}^*$
$k \leftarrow 1$
repeat
1 Apply random perturbations to the state: $\mathbf{X}' \leftarrow (\mathbf{X} + \triangle\mathbf{X})$
2 Evaluate changes in the energy: $\triangle \mathfrak{f}_{\mathbf{X}} \leftarrow (\mathfrak{f}_{\mathbf{X}'} - \mathfrak{f}_{\mathbf{X}})$
3 **if** $\triangle \mathfrak{f}_{\mathbf{X}} < 0$ **then**
3.1 Keep the new state: $\mathbf{X} \leftarrow \mathbf{X}'$
else
3.2 Accept the new state with probability: $P(T, \mathfrak{f}_{\mathbf{X}'}, \mathfrak{f}_{\mathbf{X}}) = \exp(-\frac{\mathfrak{f}_{\mathbf{X}'} - \mathfrak{f}_{\mathbf{X}}}{T})$
4 Decrease T: $T \leftarrow (T - \triangle T)$
5 $k \leftarrow k + 1$
until *termination conditions are met*

SA has been successfully applied to many engineering problems [Chibante, 2010], though it has been found too slow to converge to the global optimum. For the image segmentation problem, several works [Cook et al., 1996, Liu et al., 2007, Aguilera et al., 2012, Zhang et al., 2016b] have been found in the literature. To accelerate SA search, some variants of SA have been proposed, and we refer the reader to [Suman and Kumar, 2006, Gendreau et al., 2018] for more details.

I.4.2.2 Tabu search

Tabu search (TS) is also a single-solution based metaheuristic, which was formalized in 1986 by Glover [Glover, 1986]. Originally, the method was developed for very large combinatorial optimization and was later extended to continuous optimization [Cvijović and Klinowski, 1995]. Tabu search uses a set of strategies and learned information to mimic human insights for problem-solving.

Essentially, TS is a greedy local search (also known as hill-climbing) method that explores the solution space beyond local optimality and adopts a memory structure that imitates human behaviour, and uses past experiences to improve decision-making.

In TS, a neighbourhood structure is introduced to the solution space $\mathcal{A}(\mathbf{X})$ in the following way: Each solution $\mathbf{X} \in \mathcal{A}(\mathbf{X})$ has an associated set $\mathcal{N}(\mathbf{X}) \subset \mathcal{A}(\mathbf{X})$, with $\mathbf{X} \notin \mathcal{N}(\mathbf{X})$, called the *neighbourhood* of $\mathbf{X}$. $\mathcal{N}(\mathbf{X})$ is defined as the set of all $\mathbf{X}' \in \mathcal{A}(\mathbf{X})$ that can be obtained directly from $\mathbf{X}$ by a modification called a move $m(\mathbf{X}, \mathbf{X}')$ from $\mathbf{X}$ to $\mathbf{X}'$. The number of solutions in a neighbourhood may be very large and the quality of these solutions may vary a lot. Usually, the sizes of neighbourhoods are much larger than can be evaluated by the algorithm, and thus only the most attractive part of a neighbourhood is actually explored.

In order to guide the search process in an intelligent manner, TS procedure incorporates a flexible memory structure as its essential component. Various types of memory structures can be used to remember specific properties of the trajectory through the search space that the algorithm has undertaken. The memory structure, $\mathcal{D}$, is formally defined using so-called *tabu lists*: $T_i, i \in \{1, \ldots, p\}$ with lengths $L_i, i \in \{1, \ldots, p\}$, and $\mathcal{D} = T_1 \cup \cdots \cup T_p$. The main role of this structure in TS is to prevent cycling back to some previously generated solutions and to diversify the search process, i.e. induce the search of new subregions of the solution space $\mathcal{A}(\mathbf{X})$. However, the disadvantage of the use

of a tabu list is that a forbidden attribute may be part of an attractive solution of a neighbourhood that has not been visited so far. To cope with this problem, an aspiration criterion, involving a set of rules, is used to override tabu restrictions (if the aspiration criterion is satisfied, the move is allowed). A typical aspiration criterion is to keep a solution that is better than the best solution found so far.

It should be noted that the length of the tabu list controls the memory of the search process. If the length of the list is low, the search will concentrate on small areas of the search space (intensification). On the contrary, a high length forces the search process to explore larger regions (diversification), because it forbids revisiting a higher number of solutions.

The TS procedure starts from an initial feasible solution and at each step moves from the current solution to the best one in its neighbourhood, trying to reach an optimal solution. At step (k), a subset $\mathcal{N}'(\mathbf{X}^{(k)})$ of the modified neighbourhood $\mathcal{N}(\mathbf{X}^{(k)}, \mathcal{D})$, which is defined according to $\mathcal{D}$, is constructed and the best solution in $\mathcal{N}'(\mathbf{X}^{(k)})$ is chosen as the next solution $\mathbf{X}^{(k+1)}$ (even if it brings no improvement in terms of fitness value). Then, this solution $\mathbf{X}^{(k+1)}$ will replace the previous one $\mathbf{X}^*$ (the best solution discovered in the neighbourhood so far) if it is found to be better in terms of fitness value. At the end of the step (k), the memory structure $\mathcal{D}$ and the aspiration criterion are updated, preparing for next iteration $(k+1)$. If the stopping criteria are met, the TS algorithm terminates and the best solution found is returned. The TS procedure can be expressed in the most general way as in **Algorithm** I.2:

Algorithm I.2: The general TS algorithm

Initialization: Generate an initial solution $\mathbf{X}^{(0)}$; create the memory structure $\mathcal{D}$
Results : The optimal solution $\mathbf{X}^*$
$k \leftarrow 1$
repeat
1 Define $\mathcal{N}(\mathbf{X}^{(k)}, \mathcal{D})$ and evaluate fitness values, $\mathfrak{f}_{\mathcal{N}(\mathbf{X}^{(k)}, \mathcal{D})}$
2 Generate a set $\mathcal{N}'(\mathbf{X}^{(k)})$ as a subset of $\mathcal{N}(\mathbf{X}^{(k)}, \mathcal{D})$
3 Determine $\mathbf{X}^{(k+1)}$ by optimizing the objective function over $\mathcal{N}'(\mathbf{X}^{(k)})$
4 **if** $\mathfrak{f}_{\mathbf{X}^{(k+1)}}$ *better than* $\mathfrak{f}_{\mathbf{X}^*}$ **then**
4.1 $\mathbf{X}^* \leftarrow \mathbf{X}^{(k+1)}$
5 Update the memory structure $\mathcal{D}$, and the aspiration criteria
6 $k \leftarrow k + 1$
until *the stopping criteria are satisfied*

The stopping criteria can have the following forms. First, the TS procedure is terminated if the number of consecutive iterations, performed without any improvement of the currently best objective function value, is greater than a specified number. Second, if the optimal fitness value of the objective function is known in advance, then the process can be interrupted as soon as this value is reached.

TS has been applied to many optimization problems: vehicle routing [Toth and Vigo, 2003], continuous optimization [Battiti and Tecchiolli, 1996, Chelouah and Siarry, 2000], multi-criteria optimization [Jaeggi et al., 2008], stochastic programming [Løkketangen and Woodruff, 1996] and real-time decision problems [Gendreau et al., 2006]. For the image segmentation problem, several works [Jiang and Yang, 2002, Nanda et al., 2004, Jiang et al., 2015] have been found in the literature. A good review of the method is provided in [Gendreau, 2003], and we refer the reader to it for more details.

I.4.3 Population-based metaheuristics

Population-based metaheuristics deal with a set of solutions rather than with a single solution. In this section, we present the most studied population-based methods related to Evolutionary Computation (EC) and Swarm Intelligence (SI). EC algorithms are inspired by Darwin's evolutionary theory, where a population of individuals is modified through recombination and mutation operators. SI is a branch of biologically inspired algorithms which is focused on the collective behaviour of swarms in order to develop metaheuristics which mimic the swarm's problem solution abilities.

I.4.3.1 Evolutionary computation

From a conventional point of view, an EC method is an algorithm that simulates at some level of abstraction a biological, natural or social system. Evolutionary computation (EC) [Eiben et al., 2003, Simon, 2013] methods are derivative-free procedures, which do not require that the objective function must be two-times differentiable or uni-modal. Therefore, EC methods as global optimization algorithms can deal with non-convex, non-linear, and multimodal problems subject to linear or non-linear constraints with continuous or discrete decision variables. Despite the existing differences in these methods, they all share a common underlying idea of simulating the evolution of individual structures via processes of evolutionary operators (*selection, recombination, mutation*). The general scheme of an evolutionary algorithm can be given in **Algorithm** I.3 in a pseudocode form.

Algorithm I.3: The general scheme of an evolutionary algorithm

Initialization: Initialize the population $\mathbf{X} = \{\mathbf{X}_1, \ldots, \mathbf{X}_P\}$ randomly; evaluate their fitness values

Results : The optimal solution $\mathbf{X}^*$

$k \leftarrow 1$

repeat

1. *Select* parents
2. *Recombine* pairs of parents
3. *Mutate* a few individuals
4. *Evaluate* new individuals
5. *Select* individuals for the next generation
6. $k \leftarrow k + 1$

until *the stopping criteria are satisfied*

EC methods do not require hypotheses on the optimization problem nor any kind of prior knowledge on the objective function. They obtain knowledge about the structure of an optimization problem by utilizing information obtained from the possible solutions (i.e., candidate solutions) evaluated in the past. This knowledge is used to construct new candidate solutions which are likely to have a better quality. Figure I.1 presents a graphical representation of a basic cycle of an EC method.

In order to solve an optimization problem by using evolutionary computation method, a population $\mathbf{X} = \{\mathbf{X}_1, \ldots, \mathbf{X}_P\}$ with P candidate solutions (individuals) evolves from the initial point ($k = 1$) to a total number of iterations, N_{iter}. In its initial point, the algorithm begins by initializing the set of P candidate solutions with values that are randomly and uniformly distributed between the pre-specified lower (x_{min}) and upper (x_{max}) limits. In each iteration, a set of evolutionary operators are applied over the population $\mathbf{X}^{(k)}$ to build the new population $\mathbf{X}^{(k+1)}$. The quality of each candidate solution $\mathbf{X}_i^{(k)}$ is evaluated by using an objective function representing the fitness value, $\mathfrak{f}_{\mathbf{X}_i^{(k)}}$. During the evolution process, the best candidate solution $\mathbf{X}^*$ seen so far is preserved since it represents the

best available solution.

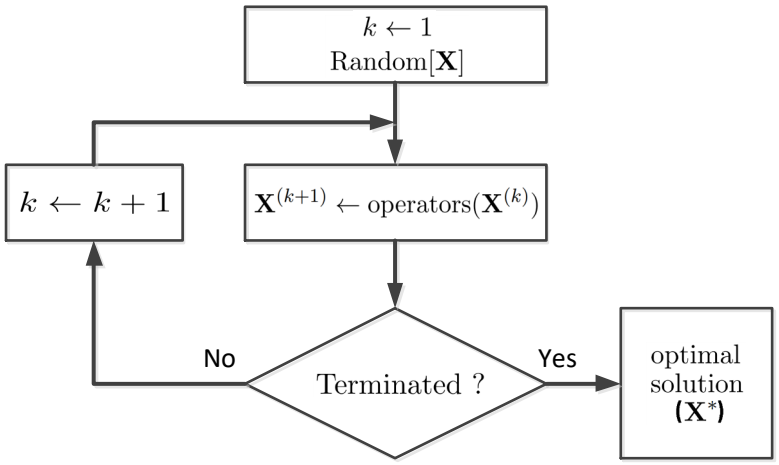

Figure I.1: The basic cycle of an EC method

Over the years, Evolutionary algorithms have been applied with a good measure of success to many areas such as combinatorial optimization problems [Blum and Roli, 2003], constrained optimization problems [Coello, 2002], Data mining and Knowledge discovery [Freitas, 2003], etc. For the image segmentation problem, many works have been also proposed [Bhandarkar and Zhang, 1999, Ho and Lee, 2003, Talbi et al., 2007, Maulik, 2009, Bhandari et al., 2015]. For more details, valuable sources for current issues, algorithms, and existing systems in this area are presented in [Coello et al., 2007, Zhou et al., 2011].

I.4.3.2 Swarm intelligence

Swarm Intelligence (SI) is an innovative distributed intelligent paradigm for solving optimization problems that takes inspiration from the collective intelligence behaviour of self-organized and decentralized systems, e.g., artificial groups of simple agents. The members of a swarm must be active, dynamic and simple (with no or very little inherent knowledge of the surroundings). The algorithms in the class of swarm intelligence primarily consist of two phases, namely the *variation phase* and the *selection phase*. These phases are responsible for maintaining the balance between exploration and exploitation and forcing the entire swarm, i.e., the set of potential solutions, to update their positions. While the variation phase explores different areas of the search space, the selection phase works for the exploitation of the previous experiences [Bansal et al., 2019].

A group of homogeneous agents exhibits the swarm intelligence if and only if it follows two conditions: self-organization and division of labour. According to Bonabeau et al. [Bonabeau et al., 1999], self-organization is categorized into four strategies: positive feedback, negative feedback, fluctuations and multiple interactions. While positive feedback is revealed to the input system to promote formation of appropriate structures (provides diversity), negative feedback balances the positive feedback and provides stabilization to the collective pattern (refers to exploitation). Fluctuations meanwhile provide new situations in the process and help to get rid of stagnation. Multiple interactions improve the overall intelligence of the swarm by sharing information among individuals within their searching area. The second condition, division of labour, is defined as the simultaneous execution of various simple and feasible tasks by individuals. This division allows the swarm to be capable of handling changed conditions in the search space.

In this section, we briefly present the most notable swarm intelligence techniques for obtaining approximate solutions to optimization problems: Ant colony optimization (ACO), Particle swarm optimization (PSO) and Cuckoo search (CS). These optimization methods will be explained below.

Ant colony optimization

Ant colony optimization (ACO) was introduced by M. Dorigo and colleagues [Dorigo, 1992, Drigo, 1996] for the solution of hard combinatorial optimization problems. ACO is inspired from the way how the ant colonies find the shortest route between the food source and their nest. When searching for food, these ants initially explore the area surrounding their nest by performing a randomized walk. Along their path between food and nest, ants deposit a chemical pheromone trail on the ground that guides other ants to the food source [Blum, 2005]. During the return trip, the quantity of pheromone that an ant leaves on the ground may depend on the quantity and quality of the food. After some time, the shortest path presents a higher concentration of pheromones, and therefore attracts more ants. Ant system [Drigo, 1996] exploited this characteristic of real ant colonies to build solutions for an optimization problem and exchange information on their quality through a communication scheme that is reminiscent of the one adopted by real ants.

ACO involves solution construction on a graph. Many ants travel through the solution space adding solution components to partial solutions until they reach a complete solution. The selection of the components depends on the pheromone content of the paths and a heuristic evaluation [Dorigo and Birattari, 2010]. At each step of construction, the m^{th} ant selects the next node using a probabilistic action selection rule, which dictates the probability at which the m^{th} ant will choose to go from the current node (i) to next node (j). At the k^{th} generation (or iteration), the probability is defined as follows:

$$p_{ij}^{m}(k) = \begin{cases} \frac{[\tau_{ij}(k)]^{\alpha}[\eta_{ij}]^{\beta}}{\sum_{l \in \mathcal{N}_i^m}[\tau_{il}(k)]^{\alpha}[\eta_{il}]^{\beta}}, & \text{if} \quad j \in \mathcal{N}_i^m \\ 0, & \text{otherwise} \end{cases} \tag{I.17}$$

where τ_{ij} is the pheromone content of the arc from node (i) to node (j), $\mathcal{N}_i^m$ is the set of nodes that remain to be visited by the m^{th} ant positioned at node (i) to make the solution feasible. η_{ij} is the heuristic information for going from node (i) to node (j). The heuristic information is a measure of the cost of extending the current partial solution. The constants α and β represent the influence of pheromone content and heuristic, respectively.

Once a solution is built, it is evaluated and amount of pheromone is relatively deposited to the quality of the solution. The ants deposit pheromone on the arcs they visited as follows:

$$\tau_{ij}(k+1) = \tau_{ij}(k) + \sum_{m=1}^{P} \Delta\tau_{ij}^{m}(k) \tag{I.18}$$

where $\Delta\tau_{ij}^{m}$ is the amount of pheromone ant m will add to the arc going from node (i) to node (j), and P is the total number of ants. The amount of pheromone added is defined by:

$$\Delta\tau_{ij}^{m}(k) = \begin{cases} \frac{Q}{L_m(k)}, & \text{if arc is in the path of ant } m \\ 0, & \text{otherwise} \end{cases} \tag{I.19}$$

where Q is a constant and $L_m(k)$ is the total cost of the path solution (the penalized objective function value for ant m) at the k^{th} iteration. All arcs in the same path will have the same cost value.

Once all ants of the colony have completed the construction of their solution, pheromone evaporation, which provides an effective strategy to avoid rapid convergence to local optima and to favour the exploration of new regions of the search space, is performed as follows:

$$\tau_{ij}(k+1) = (1-\rho)\tau_{ij}(k) + \rho\tau_{ij}(0) \tag{I.20}$$

where $\rho \in (0,1)$ is a pheromone decay parameter. The basic ACO algorithm can be summarized as in **Algorithm** I.4.

Algorithm I.4: The ACO algorithm

Initialization: Initialize the pheromone matrix $\mathcal{T} = [\tau_{ij}]$ and the number of ants P; and set: $\mathbf{X}^* \leftarrow$ Null
Results : The optimal solution, $\mathbf{X}^*$
$k \leftarrow 1$
repeat
1 Initialize the set of solutions obtained by ants
`/* Solution construction */`
2 **for** *each ant* **do**
2.1 Choose next node by applying the state transition rule given by Eq. (I.17)
2.2 Update pheromone on ant's path
2.3 Build solution by the selected items
2.4 Evaluate solution and update the best one, $\mathbf{X}^*$
`/* Update pheromones */`
3 Update and evaporate the pheromone matrix, $\mathcal{T}$, according to solutions and $\mathbf{X}^*$
4 $k \leftarrow k+1$
until *the stopping criteria are satisfied*

ACO has several advantages including offering positive feedback resulting in rapid solution finding, and having distributed computation which avoids premature convergence. However, ACO has some drawbacks such as slower convergence compared with other heuristic-based methods and lack of a centralized processor to guide it towards good solutions. Even though the convergence is frequently obtained, the time for convergence is uncertain. In addition, ACO shows its poor performance within problems with large search spaces [Selvi and Umarani, 2010].

Over the years, a number of ACO variants, which share the same characteristic idea, have been created with the aim to improve overall performance and many successful applications are now available [Mohan and Baskaran, 2012]. For the image segmentation problem, many works [Ouadfel and Batouche, 2003, Malisia and Tizhoosh, 2006, Wang et al., 2006, Han and Shi, 2007a, Huang et al., 2008, Lü et al., 2015, Rogai et al., 2016, Khorram and Yazdi, 2019] are available in the literature.

Particle swarm optimization

Particle swarm optimization (PSO) was initially introduced in 1995 by James Kennedy and Russell Eberhart as a global optimization technique [Kennedy and Eberhart, 1995]. The initial intent of the particle swarm concept was to graphically simulate the graceful and unpredictable choreography of a bird flock, with the aim of discovering patterns that govern the ability of birds to fly synchronously, and to suddenly change direction with a regrouping in an optimal formation.

In PSO algorithm, each particle is a candidate solution to the problem, and is represented by a velocity, a location in the search space and has a memory which helps it in remembering its previous

best solution. To find the best solution, PSO simply adjusts the moving vector of each particle according to its personal best, $pBest$ (cognition aspect) and the global best, $gBest$ (social aspect) positions of particles in the entire swarm at each iteration. A detailed description of PSO algorithm is given in Section II.2.2 and Section IV.3.2.

PSO has been successfully applied to many applications [Poli, 2008, Parsopoulos and Vrahatis, 2010], such as power systems [AlRashidi and El-Hawary, 2008, Del Valle et al., 2008], networkings [Kulkarni and Venayagamoorthy, 2010, Elbes et al., 2019], data clustering [Rana et al., 2011, Alam et al., 2014, Esmin et al., 2015]. For the image segmentation problem, many works have been also found in the literature [Omran et al., 2005, Li and Li, 2008, Ghamisi et al., 2013, Li et al., 2015, Pham et al., 2018, Tan et al., 2019]. However, there are some issues associated with PSO, which are: the stagnation in some points in the search space, inability to change the value of one or more decision variables, poor performance in case of a small swarm or large number of dimensions, and sensitivity to the rotation of the search space [Du and Swamy, 2016]. Therefore, a great effort has been deployed to get rid of these drawbacks, and then the performance of PSO is improved. For more details, readers are referred to [Banks et al., 2007, 2008, Zhang et al., 2015a] for a comprehensive review.

Cuckoo search

The Cuckoo search (CS) algorithm is one of the latest metaheuristic approaches introduced by Yang and Deb in 2009 [Yang and Deb, 2009]. This algorithm is inspired by the behaviour of cuckoo species, such as brood parasites, in combination of the Levy flight behaviour of some birds and fruit flies. CS employs three basic rules or operations in its implementation. First, each cuckoo is only allowed to lay one egg in each iteration, and the nest is chosen randomly by the cuckoo to lay its egg in. Second, the eggs and nests with high quality are carried forward to the next generation. Finally, the number of available host nests is fixed and the egg laid by a cuckoo is discovered by a host bird using a probability of $p_a \in [0, 1]$. In other words, the host can choose whether to throw the egg away or abandon the nest and build a new nest completely.

The main merits of the CS over other optimization algorithms are as follows: the number of parameters needed to be configured in the initial search is very little, and the inexperienced user can easily interact with it. CS has the strength points of TS in exploitation through random walk and of EC in exploration through Levy flights. It is an efficient metaheuristic algorithm that balances between the local search strategy (exploitation) and the whole space (exploration) [Roy and Chaudhuri, 2013]. However, CS also comprises several limitations such as low performance of local search, easily trapping into local optimum and lack of diversity of the local search and global search [Nguyen et al., 2018a]. A detailed description of CS algorithm is given in Section IV.3.1.

Cuckoo search has been applied in many areas of optimization and computational intelligence with promising efficiency. Comprehensive reviews about developments and applications of CS can be found in [Fister Jr et al., 2013, Shehab et al., 2017, Chiroma et al., 2017]. For the image segmentation problem, several works have been found in the literature [Agrawal et al., 2013, Bhandari et al., 2014, Ilunga-Mbuyamba et al., 2016, Suresh and Lal, 2016, Wang et al., 2018, Sumathi et al., 2018].

I.4.4 Multi-objective metaheuristics

Many sectors of industry (mechanics, chemistry, telecommunication, environment, transport, etc.) are concerned with complex problems of large dimensions that must be optimized. These optimization problems are seldom mono-objective; on the contrary, they frequently have several contradictory criteria or objectives that must be satisfied simultaneously. Multi-objective optimization (MOO) is a discipline centered on the resolution of these kinds of problems. For them, instead of

seeking for a single solution, we seek for a set of non-dominated solutions representing the compromise solutions between different conflicting objectives [Talbi, 2009], called the Pareto optimal set. A solution that belongs to this set is said to be a Pareto optimum and, when the solutions of this set are plotted in the objective space, they are collectively known as the *Pareto front*. A short description of MOO is illustrated in Section III.2.2.

Furthermore, many real-world multi-objective optimization problems (MOPs) typically need computationally expensive methods for computing the objective functions and constraints. In this context, deterministic techniques are generally not applicable, which leads therefore to using approximate methods [Gendreau et al., 2010]. Among them, metaheuristics are nowadays used extensively to deal with MOPs.

Algorithm I.5: The general scheme of a basic MOO evolutionary algorithm

Initialization: Initialize population and controlling parameters; evaluate fitness values of all individuals; build an initial Pareto set approximation
Results : Pareto set approximation
$k \leftarrow 1$
repeat
1 Evolve the population using evolutionary operators (crossover, mutation, and selection) to generate a new population
2 Evaluate the individuals of the created population in terms of fitness value
3 Build a new Pareto set approximation using dominance concept
4 Maintain non-dominated solutions (external archives, crowding distance, adaptive grid, density estimation, ranking relations,...)
5 $k \leftarrow k + 1$
until *the stopping criteria are satisfied*

The success of a MOO metaheuristic algorithm depends mainly on the definition of some issues: (1) solution representation; (2) choice of objective function; and (3) design of the operators [Mukhopadhyay et al., 2015]. To assess its performance, three items are interesting to measure: (1) the number of elements of the Pareto optimal set found; (2) the distance of the Pareto front produced by the algorithm with respect to the theoretical Pareto front (assuming we know it); and (3) the spread of solutions found [Coello et al., 2006]. The main steps of a basic MOO evolutionary algorithm can be summarized as in **Algorithm** I.5.

Over the years, many methods have been proposed including both evolutionary approaches, MOEAs, (mostly) and non-evolutionary approaches [Talbi et al., 2012]. The most popular techniques include: SPEA2 [Zitzler et al., 2001], PAES [Knowles and Corne, 2000], NSGA-II [Deb et al., 2002], MOEA/D [Zhang and Li, 2007], MOPSO [Coello and Lechuga, 2002], AMOSA [Bandyopadhyay et al., 2008]. An excellent repository (`(http://delta.cs.cinvestav.mx/~ccoello/EMOO/)`) in which the-state-of-the-art MOO algorithms can be found is available for research community. Nowadays, many MOO metaheuristic algorithms have been successfully used in different areas, for instance, scheduling, data mining, circuits and communications, control systems and robotics, manufacturing, and image processing [Zhou et al., 2011]. For the image segmentation problem, several works have been also found in the literature [Mukhopadhyay and Maulik, 2011, Saha and Bandyopadhyay, 2011, Zhang et al., 2016a, Zhao et al., 2015, Benaichouche et al., 2016, Pham et al., 2019b].

I.4.5 Hybrid metaheuristics

Recently, quite an impressive number of algorithms have been reported that do not purely follow the paradigm of a single traditional metaheuristic. On the contrary, they combine various algorithmic components, often originating from algorithms of various research areas on optimization. These approaches are commonly referred to as *hybrid metaheuristics*. The main motivation behind the hybridization of different algorithms is to exploit the complementary character of different optimization strategies, while simultaneously trying to minimize any substantial disadvantages. It is mostly due to the *no free lunch* theorems [Wolpert et al., 1997] that there cannot exist a general optimization strategy which is globally better than any other. In fact, to solve a problem at hand most effectively, one almost always needs a specialized algorithm that includes several mechanisms.

All the existing metaheuristics share some ideas and differ among each other by certain characteristic *key* components; making a toolbox of these components, from which we can pick in the design of an optimization algorithm (hybrid algorithm), is the most appropriate approach tailored to the specific characteristics of one problem at hand [Raidl, 2006]. Unfortunately, developing an effective hybrid approach is in general a difficult task which requires expertise from different areas of optimization. Moreover, the literature shows that it is non-trivial to generalize, because the hybridization of algorithms might work well for specific problems, but it might perform poorly for others.

The work on hybrid algorithms is relatively recent and can be subdivided into two different categories: collaborative hybrids and integrative hybrids [Ting et al., 2015]. In collaborative hybrids, two or more algorithms are combined in manner of running either in sequential or parallel. On the other hand, in integrative hybrids, one algorithm is regarded as a subordinate, embedded in a master metaheuristic. Good resources for studying on hybrid metaheuristics can be found in [Blum et al., 2008, Talbi et al., 2013, Blum and Raidl, 2016].

It should be noted that though hybrid approach offers a great advantage of increasing the diversity in a population and hence enhancing the search capability of the developed algorithm, there are some drawbacks. First, the hybridization process usually creates extra components, and hence the complexity of the hybrid method is increased. Second, the developed algorithm usually uses a higher number of (internal or implicit) iterations. In addition, most hybrid algorithms require an increasing number of tuning parameters, and due to its complicated structure, a hybrid algorithm is harder to be analysed.

However, in reality, hybrid algorithms have been proved successful in solving a wide range of applications, such as power systems [Katsigiannis et al., 2012, Peres et al., 2018], scheduling [Behnamian et al., 2009], telecommunications [Shankar et al., 2016, Kaur and Mahajan, 2018], data clustering [García et al., 2014], image processing [Hoseini and Shayesteh, 2013, de Paiva et al., 2016], and many others. For the image segmentation problem, some interesting works can be found in the literature [Siddhartha Bhattacharyya, 2016, Djemame and Batouche, 2018, Oliva et al., 2019, Ramadas and Abraham, 2019, Ewees et al., 2018, Jia et al., 2019].

I.5 Image segmentation using optimization metaheuristics

In the past few decades, various segmentation techniques of different accuracy and degree of complexity have been developed and reported in the literature. In this section, we present the main approaches commonly used and their improvements using optimization metaheuristics. For each method, first, an overview of how to implement it is provided, from which its capacities, advantages and disadvantages are highlighted. Then, the improvements made by optimization metaheuristics are pointed out.

I.5.1 Thresholding

Thresholding is the simplest image segmentation method. In essence, the method uses the intensity histogram to select threshold values, $\boldsymbol{\tau}$, that are intended to separate classes in the image histogram. The segmentation is then achieved by grouping all pixels between thresholds into one class.

For thresholding purpose, the pixels of a gray image $\mathbf{F} = (f_1, f_2, \ldots, f_N)$ having N pixels with L gray levels $\mathbf{L} = \{0, 1, \ldots, L-1\}$, are classified into C regions: $\{\Omega_1, \ldots, \Omega_C\}$ using a set of $(C-1)$ thresholds $\boldsymbol{\tau} = \{\tau_1, \ldots, \tau_{C-1}\}$ such that $\tau_1 < \tau_2 < \cdots < \tau_{C-1}$. For convenience, we assume two extreme thresholds $\tau_0 = f_{min}$ and $\tau_C = f_{max}$, where f_{min} and f_{max} are the lower and higher gray level in the image, respectively. A pixel with gray level f_j is assigned to region Ω_i if $\tau_{i-1} < f_j < \tau_i$, $i = \{1, 2, \ldots, C\}$.

The thresholding problem consists in selecting the set of thresholds $\boldsymbol{\tau}^*$ which optimizes an objective function $\mathcal{J}(\mathbf{F}, \boldsymbol{\tau})$, such that:

$$\boldsymbol{\tau}^* = \underset{\boldsymbol{\tau} \in \mathbf{L}}{\operatorname{argmin}} \,/\, \underset{\boldsymbol{\tau} \in \mathbf{L}}{\operatorname{argmax}} \, \mathcal{J}(\mathbf{F}, \boldsymbol{\tau}) \tag{I.21}$$

Segmentation methods based on thresholding can be divided into parametric and nonparametric [Sezgin and Sankur, 2004]. Parametric approaches estimate the parameters of a probability density function to describe each class (assuming that the intensity distribution of each class obeys a given distribution). In contrast, non-parametric approaches use criteria such as class variance, entropy, and error rate to find the best thresholds. These criteria, which commonly include: Otsu criterion [Otsu, 1979], Kapur entropy [Kapur et al., 1985], Tsallis entropy [Tsallis, 1988, De Albuquerque et al., 2004], Renyi entropy [Sahoo et al., 1997] and cross entropy [Li and Lee, 1993], are optimized to find the optimal threshold values, providing robust and accurate methods.

Thresholding is a fast and computationally efficient method. It works well in cases where there is a large variation in the intensity between the pixel values [Hiralal and Menon, 2016]. However, the method does not take into account the spatial characteristics of an image (neighbourhood information); as a result, it is sensitive to noise and intensity in-homogeneities. In low-contrast images, thresholding techniques tend to produce scattered groups of pixels rather than connected regions and require connectivity algorithms as a post-processing step. In general, this approach is not suitable for textured images because the perceptual qualities of this type of image is based on higher order interactions between image elements or objects in the scene.

In both parametric and non-parametric approaches, the image segmentation problem typically leads to an optimization problem (Eq. (I.21)) of which solution is computationally expensive and time-consuming. For solving such problem, while enumerative search methods are sensitive to initial values of thresholds, may converge to a local optimum, and deteriorate their performance with the complexity of the problem related to the number of thresholds, the metaheuristic techniques are able to escape from local optima with a reasonable time. As a result, these techniques are preferred in finding the optimal thresholds.

Over the years, many methods have been proposed to use optimization metaheuristics for the thresholding problem. Single-solution based metaheuristic approach was used in [Karasulu and Korukoglu, 2011, Jiang et al., 2012, 2015]. Population-based metaheuristic approach was used in [Hammouche et al., 2008, Zhang et al., 2014, Manikandan et al., 2014, Bhandari, 2018, Raj et al., 2019, Khorram and Yazdi, 2019, Qin et al., 2019, Borjigin and Sahoo, 2019, Suresh and Lal, 2016, Pare et al., 2017, Bhakat and Periannan, 2019]. Multi-objective metaheuristics were used in [Nakib et al., 2007, 2010, Yin and Wu, 2017, Sarkar et al., 2017, El Aziz et al., 2018, Elaziz et al., 2019]. Hybrid metaheuristics were used in [Zahara et al., 2005, Fan and Lin, 2007, Sun et al., 2016, Mlakar et al.,

2016, Dehshibi et al., 2017, Lang and Jia, 2019, Agrawal et al., 2020].

I.5.2 Region-based methods

In this kind of image segmentation methods, the images are split into a different number of regions that are constructed by dissociating or associating their neighbour pixels based on predefined criteria of similarities [Freixenet et al., 2002, Cufi et al., 2003]. Specifically, the neighbouring pixels inside a region have the same characteristics (gray level, shape, texture and/or colour), and they are dissimilar with the pixels of other regions. There are several region-based segmentation techniques that have been introduced in the literature [Pavlidis and Liow, 1990, Chen et al., 1991, Beucher and Meyer, 1993]. In this section, we consider two popular methods, namely region growing and region split and merge.

I.5.2.1 Region growing

Region growing [Zucker, 1976] is one of the simplest and most popular algorithms for region-based segmentation. The algorithm starts by choosing or automatically generating a starting point called *seed* pixel. Then, the region grows by adding similar neighbouring pixels according to a certain homogeneity criterion, increasing step by step the size of the region. So, the homogeneity criterion has the function of deciding whether a pixel belongs to the growing region or not. The region growing methods can be classified into two categories: Seeded Region Growing (SRG) [Adams and Bischof, 1994] and Unseeded Region Growing (UsRG) [Lin et al., 2000].

The advantage of region growing is that it is capable of correctly segmenting regions that have similar properties and generating connected regions. Some researchers have proved that the region growing is an effective approach with less computation effort than other non-region-based methods for segmenting MR images of brain tumors, especially for the homogeneous tissues and regions. However, the drawbacks of these techniques are its sensitivity to noise and a requirement for manual interaction to obtain the seed point which has the largest effect on the quality of the segmented image. In the presence of noise, segmented regions are disconnected or have holes. In addition, if the seed point and homogeneity criterion are not properly defined, the growing region can leak out and merge with the regions that do not belong to the object of interest.

Note that, region growing is not often used alone but within a set of image processing operations, particularly for the delineation of small, simple structures such as tumours and lesions.

I.5.2.2 Region split and merge

Typically, split and merge techniques [Horowitz, 1974] consist of two basic steps. First, the whole image is considered as one region. If this region does not satisfy a homogeneity criterion that usually takes into account edge and intensity characteristics, the region is split into four quadrants (subregions) and each quadrant is tested in the same way; this process is recursively repeated until every square region created in this way contains homogeneous pixels. However, such process of splitting is not sufficient for image segmentation since it limits the segmented shapes. Therefore, in the second step, all adjacent regions with similar attributes may be merged by using other (or the same) criteria. This process (splitting and merging) is continued recursively so that no further splits or merges are possible.

The main advantages of this method are: (1) the image could be split progressively according to our demanded resolution because the number of splitting levels can be determined beforehand;

(2) choosing criteria for splitting and merging is flexible. However, the method may produce blocky segments or more regions (over-segmentation) by splitting [Sharma et al., 2012]. Though the method could be improved by splitting in higher level, the trade off is that the computation time will arise. In addition, it can fail to provide a unique solution [Cheevasuvit et al., 1986].

To deal with the limitations of region-based methods, the metaheuristic algorithms can be used to improve their performance. For instance, several works can be found in the literature [Yang et al., 2008, Mirghasemi et al., 2013, Zanaty and Ghiduk, 2013, Al-Faris et al., 2014, Preetha et al., 2015, Manousakas et al., 1998].

I.5.3 Clustering

Clustering is a process of organizing the objects into groups based on their attributes. A cluster is a collection of objects which are similar between them and are dissimilar to the objects belonging to other clusters. Given a vector of N measurements describing each feature pixel or group of feature pixels (i.e., region) in an image, a similarity of the measurement vectors and their clustering in the N-dimensional measurement space implies similarity of the corresponding pixels or pixel groups. In other words, clustering in measurement space can be considered as an indicator of similarity of image regions and may be used for segmentation purposes. Thus, clustering can be defined as the process of identifying groups of similar image primitives in terms of image segmentation [Puzicha et al., 1999].

Clustering techniques can be classified into supervised clustering (requires the interaction of an expert to guide and verify if the separation of the data is correct) and the unsupervised clustering (decides the clustering criteria by itself). A variety of clustering techniques have been introduced to solve the image segmentation problem such as Hierarchical clustering [Johnson, 1967], Partitional clustering [Jin and Han, 2010b], K-means clustering [MacQueen et al., 1967], Fuzzy clustering (developed by Dunn in 1973 [Dunn, 1973] and improved by Bezdek in 1981 [Bezdek, 1981]) and Expectation-maximization [Jin and Han, 2010a]. Among those, Fuzzy C-means (FCM) introduced by Bezdek [Bezdek, 1981] is one of the most popular methods that is widely used in the literature.

I.5.3.1 Fuzzy C-means clustering

In essence, FCM tries to minimize the intra-cluster variation through iterations. The unlabelled pixels are assigned to the nearest clusters basing on their distances to the cluster centroids, then the cluster centroid is updated and the pixels are re-assigned. The algorithm ends until all the pixels have fixed labels. Let $\mathbf{F} = (f_1, f_2, \ldots, f_N)$ denote an image with N pixels to be partitioned into C clusters, where f_j represents multispectral (features) data. The algorithm is an iterative optimization that minimizes the cost function defined as follows:

$$\mathcal{J}_{\mathrm{FCM}}(\mathbf{F}, \mathbf{C}, \mathbf{U}) = \sum_{i=1}^{C} \sum_{j=1}^{N} u_{ij}^{p} \mathrm{D}_{ij} \tag{I.22}$$

where p controls the fuzziness degree of clustering. Particularly, it indicates that how much the clusters can overlap with each other. The bigger the value of the p is, the more the clusters overlap with each other. u_{ij} is fuzzy membership of data f_j to cluster with center c_i, and D_{ij} is distance between data f_j and center of the cluster c_i. The membership values $\mathbf{U} = (u_{ij})_{C \times N}$ comply with the following conditions:

$$\mathbf{U} \in \left\{ u_{ij} \in [0,1] \,\middle|\, \sum\nolimits_{i=1}^{C} u_{ij} = 1, \quad \forall j \quad \text{and} \quad 0 < \sum\nolimits_{j=1}^{N} u_{ij}, \quad \forall i \right\} \tag{I.23}$$

Since FCM is an iterative process, the membership function and center of each cluster are obtained as follows:

$$u_{ij} = \left[\sum_{r=1}^{C} \left(\frac{\mathrm{D}_{ij}}{\mathrm{D}_{rj}} \right)^{\frac{1}{p-1}} \right]^{-1} \tag{I.24}$$

$$c_i = \frac{\sum_{j=1}^{N} u_{ij}^{p} f_j}{\sum_{j=1}^{N} u_{ij}^{p}} \tag{I.25}$$

The FCM algorithm iteratively optimizes $\mathcal{J}_{\mathrm{FCM}}$ by evaluating Eqs. (I.24) and (I.25) until stopping criteria are satisfied. The procedure of the FCM algorithm for the image segmentation problem can be summarised as the pseudo code shown in **Algorithm** I.6. Once a membership value u_{ij} for class i is assigned to a pixel j, a defuzzification process takes place so as to convert the partition matrix $\mathbf{U}$ to a crisp partition. Commonly, the maximum membership procedure method, which assigns the pixel j to the cluster i with the highest membership, is adopted as follows:

$$\Omega_i = \operatorname{argmax} \{u_{ij}\}, \text{with } i = 1, 2, \cdots, C \tag{I.26}$$

Algorithm I.6: The FCM algorithm

Initialization: Read the input image; fix the number of clusters, C; randomly initialize $\mathbf{U} = (u_{ij})_{C \times N}$ satisfying Eq. (I.23); set termination criterion ϵ, maximum number of iterations N_{iter} and the parameter p; compute the cluster centers $\mathbf{C} = (c_i)_C$ using $\mathbf{U}$ and Eq. (I.25)

Results : Cluster centroids $\mathbf{C}$ and membership matrix $\mathbf{U}$

$k \leftarrow 1$

repeat

1. Calculate similarity between cluster centers and image elements, $\mathbf{D} = (\mathrm{D}_{ij})_{C \times N}$
2. Update the partition matrix $\mathbf{U} = (u_{ij})_{C \times N}$ using Eq. (I.24)
3. Update the cluster centers $\mathbf{C} = (c_i)_C$ using Eq. (I.25)
4. $k \leftarrow k + 1$

until $(\max \{|\mathbf{U}_{new} - \mathbf{U}_{old}|\} < \epsilon)$ **or** $(k > N_{iter})$

The main advantages of using fuzzy clustering for the image segmentation problem are robust characteristics for ambiguity. However, the two major drawbacks of these methods are: (1) the sensitivity to noise and INU artifact, since no local spatial information in the image is considered; (2) the high vulnerability of the algorithms to trapping into local minima, depending on the choice of the initial clustering centroids.

To deal with such problems, many works, which improve either the objective function or searching method or both, have been proposed. For instance, to overcome the first problem, some interesting works with different directions (adding feature information of the neighbourhood pixels into the objective function or into a similarity measure between cluster centers and image elements, transferring non-linear problem to linear problem, using effective metric for measuring similarity, etc) are [Ahmed et al., 2002, Zhang and Chen, 2004, Chuang et al., 2006, Krinidis and Chatzis, 2010, Chen et al., 2011, Adhikari et al., 2015, Verma et al., 2016, Lei et al., 2018, Singh and Bala, 2019, Memon et al., 2019]. To address the second problem, the metaheuristic optimization algorithms have been introduced. Many works in this category can be found in the literature [Han and Shi, 2007b, Izakian and Abraham, 2011, Ding and Fu, 2015, Maulik and Saha, 2010, Indrajit Saha and Plewczynski, 2011, Chaoshun Li and Xiao, 2012, Alsmadi, 2014, Silva Filho et al., 2015, Pham et al., 2018, Zhao and Fan, 2019].

I.5.4 Deformable models

Deformable model is defined as surfaces or curves in image domain, which are influenced by the internal and external forces in deformation of the image. In the process of deformation, an internal force is to maintain the smoothness and same features of the curves, whereas external forces are defined such that the model is attracted towards an object or other features of interest within the image. Deformable models extract elements with similar structure and integrate the boundaries to be consistent and produce coherent structures.

There are basically two main types of deformable models: *parametric/explicit* and *geometric/implicit*. The former represents curves and surfaces explicitly in their parametric forms during deformation, allowing direct interaction with the model and leading to a compact representation for implementation. On the other hand, the later handles topological changes naturally since these models are based on the theory of curve evolution [Adalsteinsson and Sethian, 1995], and they represent curves and surfaces implicitly, as a level set of a higher-dimensional scalar function.

I.5.4.1 Parametric deformable models

Parametric deformable models which are well-known as snakes, explicit deformable models and active contour models, were first introduced by Kass et al. [Michael Kass and Terzopoulos, 1988]. Starting from an initial closed configuration representing a rough approximation of the shape to be segmented, the deformation procedure is driven by the minimization of an energy function until the deformable model coincides with the object boundary. Let $s \in [0,1]$ be the parametric domain and $\mathbf{v}(s)$ be a representation of the contour, then the energy of $\mathbf{v}(s)$ is given by:

$$\mathcal{J}_{\mathrm{PDM}}(\mathbf{v}) = \underbrace{\alpha \int_0^1 \left|\frac{\partial \mathbf{v}}{\partial s}\right|^2 ds + \beta \int_0^1 \left|\frac{\partial^2 \mathbf{v}}{\partial s^2}\right|^2 ds}_{\text{interal energy}} + \underbrace{\int_0^1 \mathrm{P}(\mathbf{v}) ds}_{\text{external energy}} \tag{I.27}$$

where $\mathrm{P}(\mathbf{v})$ denotes a scalar potential function defined in the image plane; α, and β are the system's parameters, which dictate the simulated physical characteristics of the contour; α controls the "tension" of the contour (more or less resistant to stretching); β controls its "rigidity" (more or less resistant to bending).

The main advantages of these models are their ability to directly generate closed parametric curves or surfaces and their incorporation of a smoothness constraint. In addition, the external forces act in a quite intuitive manner and they can be adapted to track dynamic objects. However, the basic method also presents some limitations, namely the sensitivity to local minima, the dependency on initialization, the absence of prior shape knowledge and impossibility of managing topological changes [Mesejo and Cagnoni, 2016].

I.5.4.2 Geometric deformable models

Geometric deformable models [Malladi and Vemuri, 1995, Caselles, 1995] are based on the curve evolution theory [Sapiro and Tannenbaum, 1993] and the level set method [Sethian, 1999]. Since the evolution of curves and surfaces do not depend on the particular way the curve has been parameterized, the evolving curves can be represented implicitly as a level set of a higher dimensional function. Amongst geometric models, the level set method is one of the most popular methods relying on an evolving closed surface defined by a moving interface which expands outwards until it reaches the desired boundary.

Given an interface $\Gamma(t)$ that is characterized by a Lipschitz continuous function:

$$\begin{cases} \phi(t,\mathbf{x}) > 0, & \text{for } \mathbf{x} \text{ inside } \Gamma(t) \\ \phi(t,\mathbf{x}) < 0, & \text{for } \mathbf{x} \text{ outside } \Gamma(t) \\ \phi(t,\mathbf{x}) = 0, & \text{for } \mathbf{x} \text{ on } \Gamma(t) \end{cases} \tag{I.28}$$

where $\phi(t,\mathbf{x})$ is a level set function at time t. The evolution of ϕ is commonly described by the following differential equation:

$$\frac{\partial \phi}{\partial t} + \mathbf{V}\,|\bigtriangledown \phi| = 0 \tag{I.29}$$

where $\mathbf{V}$ is the speed function normal to the curve, which can depend on position, time or geometry of the interface, and $\bigtriangledown$ is the spatial gradient operator. Importantly, $\mathbf{V}$ describes the local movement of the contour and is analogue to the energy function used in parametric deformable models.

One of the main advantages of the level set-based methods is the natural ability of a single surface to seemingly split apart and merge without losing its identity. However, there are also two main disadvantages, namely computationally demanding and requiring considerable design effort to construct appropriate velocity functions, $\mathbf{V}$, for adapting the level set function.

To partially solve problems mentioned above, metaheuristics have been involved. Many works can be found in the literature. For instance, in parametric approach, they can be used to set the initial location of the snakes and the appropriate number of control points [Mobahi et al., 2006, Chih-Yu Hsu and Chen, 2008], to search the best set of values for the parameters that control the evolution of the active contour model [Caselles, 2000, Teixeira et al., 2008], to evolve each control point of the model and search for best location in a given search space [MacEachern and Manku, 1998, Covavisaruch and Tanatipanond, 1999, Tang et al., 2007]. In geometric approach, metaheuristics have been used to initialize the contour and/or extract the prior information which is to be used by the level set method [Xiao et al., 2006, Mesejo et al., 2015] or to directly guide the optimization process avoiding local minima [Kan et al., 2007, Ghosh et al., 2010, Law et al., 2008].

I.5.5 Markov random field models

Digital images consisting of pixels and being defined on a bounded 2-D lattice have two important attributes. First, pixel intensities of a nearly homogenous object will follow a certain distribution, called the conditional intensity distribution of the pixel intensities. Second, the pixels close together or lying in a neighbourhood will tend to have similar intensity values, known as contextual information. Markov random field (MRF) modelling itself is not a segmentation method but a statistical model which can be used within segmentation methods. MRFs model spatial interactions between neighbouring or nearby pixels. Combining MRFs with the conditional intensity distribution under a Bayesian framework, we can estimate the true intensities of the image much more accurately than those based on only the information derived from the image intensities. A detailed description of MRFs framework for the image segmentation problem is given in Section IV.2.

There are two main advantages of using MRF models for the image segmentation problem [Huawu Deng and Clausi, 2005]: (1) the spatial relationship can be seamlessly integrated into a segmentation procedure through contextual constraints of neighbouring pixels; (2) different types of image features can be utilized in the MRF-based segmentation model via the Bayesian framework. However, there are still several problems which limit its performance [Kato et al., 1999, Marroquin et al., 2002]: (1) How to provide reasonably good initial values for mean and variance of each class (assuming that the classes are represented by Gaussian distributions)?; (2) What is the efficient

algorithm for finding the optimal estimator to avoid the convergence to the first encountered minimum when including spatial coherence assumptions?; (3) The basic model is inaccurate in nature.

To address the aforementioned problems, there are many approaches in which metaheuristics have been used. For instance, for the first problem, metaheuristics have been used to find the initialization of MRFs [Cuadra et al., 2005, T Krishnan et al., 2016]. In order to seek the global extremum solution in the optimization step, several works in which metaheuristics are used as searching engines can be found [Yousefi et al., 2012, Ben George et al., 2015, Guerrout et al., 2017]. For the last problem, which is a key in image segmentation, two main directions have been exploited. While one is to achieve a proper balance between the two components [Deng and Clausi, 2004, Huawu Deng and Clausi, 2005] in the standard model [Li, 2001], the other is to use spatial context or neighbourhood information efficiently [Yang et al., 2015, Ahmadvand et al., 2017, Chen et al., 2017a].

I.5.6 Hybrid techniques

Since the selection of the most appropriate technique for a given application is often a difficult task, a combination of several techniques may be necessary to obtain the segmentation goal. Therefore, hybrid or combined segmentation methods have been used extensively in different segmentation applications [Li et al., 2011a, Ortiz et al., 2013, Abdel-Maksoud et al., 2015, Gupta et al., 2015]. The main idea is to combine different complementary segmentation methods into a hybrid approach to avoid many of the disadvantages of each method alone and improve segmentation accuracy.

However, the traditional segmentation methods commonly share the same drawbacks such as the dependency on the initialization and a high possibility of being trapped into local minima, that they cannot deal with by themselves. As a result, metaheuristics are a good approach for such problems. Several works in this direction can be found in the literature [Kaur et al., 2012, Al-Faris et al., 2014, Mekhmoukh and Mokrani, 2015].

Note that, combining different segmentation methods often leads to increased complexity. This induces a higher computational time and a higher number of different parameters that need to be tuned for specific application. Therefore, designing a hybrid segmentation method should be done carefully.

I.6 Conclusion

In this chapter, we presented the basic information of MRI technique as well as the formulation of the image segmentation problem. In addition, the main metaheuristics and segmentation methods were also provided. In each category, we have cited several works in which metaheuristics have been taken into account to improve the segmentation performance. Besides, current limitations of the segmentation methods have been pointed out, from which we can have a clear view to go further. Even though there is a growing number of works in this field, using metaheuristics for solving the image segmentation problem has been proven to be successful. By considering the limitations and taking advantages of current works, we propose in the following chapters some methods that are contributions in this area.

Chapter II

Mono-objective optimization metaheuristic approach for image segmentation using fuzzy entropy clustering

II.1 Introduction

In this chapter, we present a new method based on fuzzy entropy clustering (FEC) for segmentation of brain magnetic resonance (MR) images. Because of simplicity and applicability, the FEC algorithm is widely used in medical image segmentation. The FEC method is stand on the principles of fuzzy logic and entropy theory, which are not only able to tolerate inaccurate, incomplete or uncertain information but also save information in the original image. However, there are two major problems associated with this method, which are: (1) the sensitivity to noise and intensity non-uniformity (INU) artifact and (2) the trapping into local minima and dependency on initial clustering centroids. The method we propose in this chapter makes it possible to overcome these drawbacks.

This chapter is organized as follows. In Section II.2, we present the related works on which the proposed method is based. Section II.3 introduces the proposed method: segmentation of brain MR images using particle swarm optimization (PSO) with a novel objective function based on fuzzy entropy clustering. Evaluation of the proposed method and comparison with other methods in the literature are shown in Section II.4. Finally, in Section II.5, we conclude with some thoughts on the strengths of the presented method and shortcomings to improve to go further in the results.

II.2 Related works

II.2.1 Image segmentation using fuzzy entropy clustering

The fuzzy entropy clustering (FEC) algorithm proposed by Tran and Wagner [Tran and Wagner, 2000] is an alternative generalization of hard C-means (HCM) clustering algorithm. The objective of this method is finding both the centers of classes and the degrees of data belonging to different classes. For a given feature image $\mathbf{F} = (f_1, f_2, \ldots, f_N)$ with N pixels, where f_j represents a feature of the j^{th} pixel, the algorithm allows partitioning of the image, by minimizing an objective function with respect to a partition matrix $\mathbf{U} = (u_{ij})_{C\times N}$ and a set of C cluster prototypes $\mathbf{C} = (c_i)_C$ with $(1 < C < N)$. The standard FEC objective function is given as follows:

$$\mathcal{J}_{\text{FEC}}(\mathbf{F}, \mathbf{C}, \mathbf{U}) = \underbrace{\sum_{i=1}^{C}\sum_{j=1}^{N} u_{ij} d^2(f_j, c_i)}_{\text{shape and size}} + \underbrace{\frac{1}{n}\sum_{i=1}^{C}\sum_{j=1}^{N} u_{ij}\log(u_{ij})}_{\text{degree of non-membership}} \tag{II.1}$$

It should satisfy the following conditions:

$$\mathbf{U} \in \left\{ u_{ij} \in [0,1] \middle| \sum_{i=1}^{C} u_{ij} = 1, \quad \forall j \quad \text{and} \quad 0 < \sum_{j=1}^{N} u_{ij}, \quad \forall i \right\} \tag{II.2}$$

where f_j is the feature of the j^{th} pixel in the image which can be intensity value or gray value; $d^2(f_j, c_i)$ is the Euclidean distance between the feature f_j and the cluster center c_i, and $1/n$ denotes the degree of fuzzy entropy.

Since u_{ij} and c_i cannot be determined simultaneously, optimizing the objective function Eq. (II.1) under the constraints Eq. (II.1) is an ill-posed problem. In order to solve this constrained problem, first, the problem is transformed into an unconstrained problem by using the Lagrange multipliers, then the Picard iteration is applied. Particularly, the FEC objective function, $\mathcal{J}_{\text{FEC}}$, is iteratively minimized by using the following update equations:

$$u_{ij} = \left\{ \sum_{r=1}^{C} \left\{ \exp[d^2(f_j, c_i) - d^2(f_j, c_r)] \right\}^n \right\}^{-1} \tag{II.3}$$

$$c_i = \frac{\sum_{j=1}^{N} u_{ij} f_j}{\sum_{j=1}^{N} u_{ij}} \tag{II.4}$$

When the algorithm has converged, a defuzzification process takes place so as to convert the partition matrix $\mathbf{U}$ to a crisp partition. Commonly, the maximum membership procedure method, which assigns the j^{th} pixel to the i^{th} region, Ω_i, with the highest membership, is adopted as follows:

$$\Omega_i = \text{argmax}\{u_{ij}\}, \text{with } i = 1, 2, \cdots, C \tag{II.5}$$

The procedure of the FEC algorithm can be described as the pseudo code shown in **Algorithm** II.1

Algorithm II.1: The FEC algorithm

Initialization: Read the input image; fix the number C of cluster centers; randomly initialize $\mathbf{U} = (u_{ij})_{C \times N}$ satisfying Eq. (II.2); set termination criterion ϵ, maximum number of iterations N_{iter} and the parameter n; compute the cluster centers $\mathbf{C} = (c_i)_C$ using $\mathbf{U}$ and Eq. (II.4)

Results : Cluster centroids $\mathbf{C}$ and membership matrix $\mathbf{U}$

$k \leftarrow 1$

repeat

1 Update the partition matrix $\mathbf{U} = (u_{ij})_{C \times N}$ using Eq. (II.3)

2 Update the cluster centers $\mathbf{C} = (c_i)_C$ using Eq. (II.4)

3 $k \leftarrow k + 1$

until $(\max\{|\mathbf{U}_{new} - \mathbf{U}_{old}|\} < \epsilon)$ **or** $(k > N_{iter})$

Note that, in the objective function, $\mathcal{J}_{\text{FEC}}$, while the first term on the right-hand side is the sum-of-squared-errors function controlling the shape and size of the clusters, the second term expresses the average degree of non-membership of members. Minimizing $\mathcal{J}_{\text{FEC}}$ means simultaneously minimizing the dispersion within clusters and maximizing the degree of membership of members. However, this energy function is non-convex, non-unique in nature and may have several local minimum points, hence optimizing it by using the gradient descent technique faces the problem of getting stuck in local

minima. In addition, $\mathcal{J}_{\text{FEC}}$ does not take into account any local spatial and bias information, and the metric is the Euclidean distance $d^2(f_j, c_i)$. This metric assumes that each feature of data points is equally important, independent from others, and belonging to a cluster with spherical shape. This assumption may not be always satisfied in real applications, especially in medical images. Furthermore, the parameter $1/n$, which controls the influence of the degree of fuzzy entropy, needs to be carefully chosen. If $1/n$ is relatively small, the FEC clustering reduces to HCM clustering. In contrast, if $1/n$ is large enough, u_{ij} in Eq. (II.3) tends towards $1/C$, it means we only obtain a single cluster.

II.2.2 Low-discrepancy sequence initialized particle swarm optimization

Particle swarm optimization (PSO) is a population-based stochastic optimization technique regarded as a global search strategy, originally introduced by Kennedy and Eberhart [Kennedy and Eberhart, 1995]. In the PSO algorithm, each member of the population, called *particle*, represents a potential solution to the optimization problem; and the population, called *swarm*, is evolved through successive iterations. The quality of a candidate solution is evaluated by the fitness value, associated with each particle. Each particle in the swarm with size of P, denoted by i, has a position vector $\mathbf{X}_i = (x_{ir})_C$, a velocity vector $\mathbf{V}_i = (v_{ir})_C$, its own best position $pBest$ found so far, and interacts with neighbouring particles through the best position $gBest$ discovered in the neighbourhood so far. At the k^{th} iteration, each particle is moved according to equations (II.6) and (II.7):

$$\mathbf{V}_i^{(k+1)} = w^{(k)}\mathbf{V}_i^{(k)} + c_1 r_1 \left[pBest^{(k)} - \mathbf{X}_i^{(k)}\right] + c_2 r_2 \left[gBest^{(k)} - \mathbf{X}_i^{(k)}\right] \tag{II.6}$$

$$\mathbf{X}_i^{(k+1)} = \mathbf{X}_i^{(k)} + \mathbf{V}_i^{(k+1)} \tag{II.7}$$

where c_1 and c_2 are acceleration coefficients that scale the influence of the *cognitive* and *social* components; r_1 and r_2 are two random values, uniformly distributed in $[0, 1]$; and $w^{(k)}$ is inertia weight at the k^{th} iteration. The higher $w^{(k)}$ is, the higher the possibility of searching in the global solution space is, and the smaller $w^{(k)}$ is, the higher the possibility of searching in the local solution space is.

There are two basic criteria for assessing performance of the PSO algorithm, named the convergence speed and the ability to find global optima. To optimize both criteria, keeping balance between global exploration and local exploitation is crucial. From Eqs. (II.6) and (II.7), it is clear that the performance depends not only on the controlling parameters $\{w^{(k)}, c_1, c_2\}$, but also on the size and structure of the neighbourhood.

In this work, the population initialization and the tuning of parameters $\{w^{(k)}, c_1, c_2\}$ are adopted from the low-discrepancy sequence initialized particle swarm optimization (LHNPSO) algorithm designed by Yang et al. [Yang et al., 2015]. Particularly, the initial population of particles is generated by using the Halton sequence to cover the search space effectively, $\{c_1, c_2\}$ are set to constants, equal to 2, and the inertia weight is updated as follows:

$$w^{(k)} = w_{max} - (w_{max} - w_{min}) \cdot \left(\frac{k}{N_{iter}}\right)^{\frac{1}{\pi^2}} \tag{II.8}$$

where $[w_{min}, w_{max}]$ is a range of the inertia weight, with $w_{min} = 0.4$ and $w_{max} = 0.9$. k and N_{iter} are the iteration numbers starting from iteration one and a maximum number of allowable iterations, respectively.

This law of variation of $w^{(k)}$ increases the exploration of the search space in the beginning stage of iterations of the algorithm, and the exploitation of the best solutions found so far towards the end of the algorithm. This constitutes a reasonable balance between the phases of exploration of the search space and the phases of exploitation of the solutions.

Note that, current research [Harrison et al., 2016] reported that this variant outperformed the other variants of PSO with two main advantages, namely, (1) a faster converging process with more accurate final solution and (2) an easy implementation.

The procedure of the LHNPSO algorithm is summarized in **Algorithm** II.2:

Algorithm II.2: The LHNPSO algorithm

Initialization: Set maximum number of iterations N_{iter}; initialize particles using the Halton sequence; set c_1, c_2 and the range $[w_{min}, w_{max}]$; determine $pBest$ and $gBest$.
Results : The optimal solution $gBest$
$k \leftarrow 1$
repeat
`/* Flight particles */`
1 Update the inertia weight $w^{(k)}$ according to Eq. (II.8)
2 Update $\mathbf{X}_i$ and $\mathbf{V}_i$ according to Eqs. (II.6) and (II.7)
`/* Evaluate solutions */`
3 Evaluate fitness values, ƒ
4 Update $pBest$ and $gBest$
5 $k \leftarrow k + 1$
until *the stopping criteria are met*

II.3 Proposed method

The proposed segmentation method for MR images is introduced based on three advanced concepts:

1. An appropriate image model is applied by considering both noise and intensity non-uniformity (INU) artifact.
2. Kernel method is used for measuring the similarity to overcome problems of complex structure and non-spherical shape tissues, which are common in MR images.
3. The LHNPSO algorithm is utilised to solve the segmentation problem with a novel objective function.

These three concepts are detailed in the following sections.

II.3.1 Image model

Different models of MR image formation have been proposed in the literature, depending on how the true image $\left(f_j^t\right)_N$, the intensity inhomogeneity field $(b_j)_N$, and the noise $(n_j)_N$ interact [Vovk et al., 2007]. In this work, the most common model of MR image formation, which assumes that the noise is approximated by Gaussian probability distribution and is independent of the intensity inhomogeneity field [Pham and Prince, 1999], is used. In addition, the bias field is typically modelled as a low-frequency multiplicative field. Accordingly, the acquired image $(f_j)_N$ is obtained as:

$$(f_j)_N = \left(f_j^t\right)_N (b_j)_N + (n_j)_N \tag{II.9}$$

where f_j is the measured intensity of the j^{th} MRI pixel; f_j^t is the true intensity value; b_j is the unknown smoothly varying bias field caused by imperfections of imaging devices, particularly, due to the magnetic field, n_j is an additive zero-mean Gaussian noise with relatively small variance at the j^{th} pixel, and N is the number of pixels in the MR image. As an example, Figure II.1 shows an example of estimated bias field and noise when segmenting a simulated MR image with characteristics of 9% noise and 40% INU (slice 80 from the brainWeb dataset). Thus, the accurate segmentation of a MR image involves an accurate estimation of the unknown bias field $(b_j)_N$ and removal of both this bias field and noise from the measured MR signal.

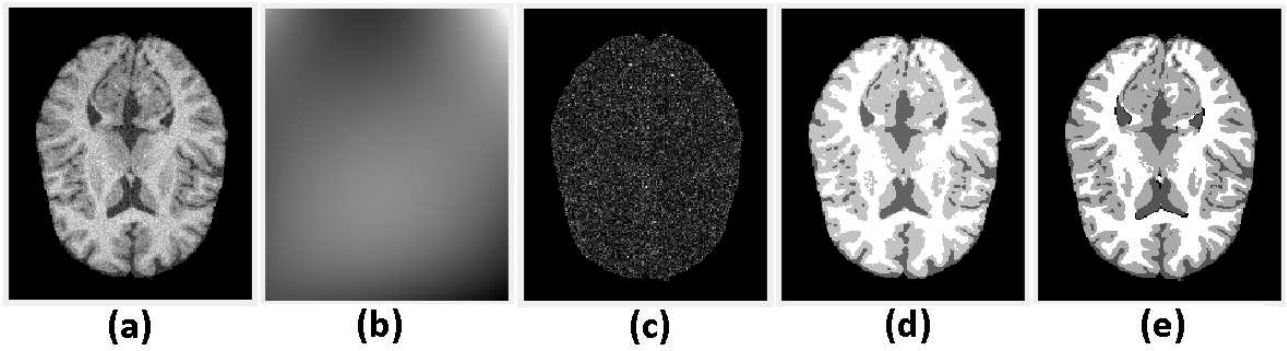

Figure II.1: An example of estimating intensity non-uniformity and noise for a simulated MR image: (a) original image; (b) bias field estimation; (c) noise estimation; (d) segmentation result; (e) ground truth.

II.3.2 Kernel method for similarity measurement

Recently, "kernel method" has been widely applied to fuzzy clustering, which is referred to as kernel-based fuzzy clustering [Graves and Pedrycz, 2010]. The essence of the kernel method is to perform a non-linear transformation $\Phi : X^P \to H$, where X^P is an input data space with a low dimension and H is a high dimensional feature space. With this transformation, the structure of input data points, which may be inadequate for the analysis in the original space, but can be now analysed in space H. More specifically, in the segmentation of brain MR images, input images which have complex structure, non-spherical shape tissues, can more likely be linearly segmented in the feature space according to the well-known Cover's theorem [Cover, 1965].

It is well known that every linear algorithm that uses inner products can be easily extended to a non-linear version through the kernels that satisfy the Mercer's conditions [Muller et al., 2001]. A kernel in the feature space calculating the inner product can be represented as function K below:

$$K\left(f_j, c_i\right) = \left\langle \Phi\left(f_j\right), \Phi\left(c_i\right) \right\rangle \tag{II.10}$$

where f_j and c_i are a feature of j^{th} pixel and an i^{th} cluster center in the image, respectively; $\left\langle \Phi\left(f_j\right), \Phi\left(c_i\right) \right\rangle$ denotes the inner product operation. In this approach, Gaussian kernel is adopted as:

$$K(f_j, c_i) = \langle \Phi(f_j), \Phi(c_i) \rangle = \exp\left(-\frac{\|f_j - c_i\|^2}{\sigma^2}\right) \tag{II.11}$$

where σ^2 is a Gaussian kernel parameter.

It is worth to note that the performance of kernel-based fuzzy clustering depends greatly on the selection of kernel parameter [Graves and Pedrycz, 2010]. Particularly, if σ^2 is small enough, $K\left(\cdot, \cdot\right)$

is close to zero, it means that any two mapped data samples are approximately orthogonal with no similarity. On the other hand, if σ^2 is relatively large, $K(\cdot,\cdot)$ is approximately one, which means the samples are roughly overlapped. Both situations make data clustering infeasible. Despite the fact that some works have attempted to provide an optimal value of the parameter [Wang et al., 2003], the estimation of σ^2 remains as an open problem [Liao and Zhang, 2011]. Hence, in this work σ^2 is estimated by generalizing the result of Yang and Tsai's work [Yang and Tsai, 2008], which is defined as follows:

$$\sigma^2 = \frac{\lambda}{N}\sum_{j=1}^{N}\left\|f_j - \overline{f}\right\|^2, \text{ with } \overline{f} = \sum_{j=1}^{N} f_j \tag{II.12}$$

where $\overline{f}$ is the mean intensity value of the entire image, and λ is a constant defined by experiments.

Without explicitly using the transformation Φ, the Euclidean distance $d^2(f_j, c_i) = \|f_j - c_i\|^2$, which measures the similarity between two data samples $\{f_j, c_i\}$, can be executed by substitution of $\|\Phi(f_j) - \Phi(c_i)\|^2$ in the feature space. It is defined by using kernel function K as below:

$$\begin{aligned}\|\Phi(f_j) - \Phi(c_i)\|^2 &= K(f_j, f_j) + K(c_i, c_i) - 2K(f_j, c_i)\\ &= 2\left[1 - \exp\left(-\frac{\|f_j - c_i\|^2}{\sigma^2}\right)\right]\end{aligned} \tag{II.13}$$

II.3.3 A novel image segmentation model based on fuzzy entropy clustering and kernel method

Let Ω be the image domain, and assume that the true image $\left(f_j^t\right)_N$ is segmented into C disjoint regions, $\{\Omega_1, \Omega_2, \cdots, \Omega_C\}$, with distinct cluster centers, $\{c_1, c_2, \cdots, c_C\}$, respectively, where $\Omega = \cup_{i=1}^{C}\Omega_i$ and $\Omega_i \cap \Omega_j = \oslash$ for $i \neq j$. Similarly, the segmentation of the acquired image $(f_j)_N$ totally includes C segments which have cluster centers $\{m_1, m_2, \cdots, m_C\}$. Then, in view of the image model in Eq. (II.9) and because of $(b_j)_N$ varying slowly, the acquired image $(f_j)_N$ can be approximated as follows:

$$f_j \approx (m_i + n_j) \approx (c_i b_j + n_j) \quad \text{for } f_j \in \Omega_i \tag{II.14}$$

For the objective function, $\mathcal{J}_{\text{FEC}}(\mathbf{F}, \mathbf{C}, \mathbf{U})$ in Eq. (II.1), considering bias field estimation $(b_j)_N$, the modified objective function could be given as follows:

$$\mathcal{J}_{\text{FECB}}(\mathbf{F}, \mathbf{C}, \mathbf{U}, \mathbf{B}) = \underbrace{\sum_{i=1}^{C}\sum_{j=1}^{N} u_{ij} d^2(f_j, c_i b_j)}_{\text{shape and size}} + \underbrace{\frac{1}{n}\sum_{i=1}^{C}\sum_{j=1}^{N} u_{ij}\log(u_{ij})}_{\text{degree of non-membership}} \tag{II.15}$$

Incorporating this criterion with a novel local spatial constraint term $\sum_{i=1}^{C}\sum_{j=1}^{N} u_{ij} d^2\left(\overline{f}_j, c_i b_j\right)$, then kernelizing it through the newly induced distance measure substitution (II.13), a new objective function for image clustering segmentation is obtained as follows:

$$\begin{aligned}\mathcal{J}_{\text{KFECSB}}\left(\mathbf{F},\mathbf{C},\mathbf{U},\mathbf{B}\right) = & \underbrace{\sum_{i=1}^{C}\sum_{j=1}^{N} u_{ij}\left[1-K(f_j,c_i b_j)\right]}_{\text{shape and size}} + \underbrace{\frac{1}{n}\sum_{i=1}^{C}\sum_{j=1}^{N} u_{ij}\log\left(u_{ij}\right)}_{\text{degree of non-membership}} \\ & + \underbrace{\eta\sum_{i=1}^{C}\sum_{j=1}^{N} u_{ij}\left[1-K(\overline{f}_j,c_i b_j)\right]}_{\text{artifact in-sensitiveness and detail preservation}}\end{aligned} \tag{II.16}$$

where the constraints in Eq. (II.2) should be satisfied.

In this new objective function (II.16), the first term is actually a transformation of the first term in Eq. (II.1) thanks to the kernel substitution trick. The third term is the local spatial constraint term, in which the parameter η controls the effect of the penalty. In essence, this term, equivalently, aims at guaranteeing noise and INU artifact in-sensitiveness, and image detail preservation. The median of the neighbours within a window of size 3×3 around the j^{th} pixel is used to represent $\overline{f}_j$, which can be computed in advance. In addition, in this work, the bias field estimation of $\mathbf{B} = (b_j)_N$ based on Li et al.'s work [Li et al., 2014] is computed by using partition matrix $\mathbf{U} = (u_{ij})_{C\times N}$ and cluster centers $\mathbf{C} = (c_i)_C$ as follows:

$$\mathbf{B} = \left[\left(\sum_{i=1}^{C}\mathbf{G}\times\mathbf{F}\sum_{j=1}^{N} c_i u_{ij}\right)^{-1}\left(\mathbf{G}\times\mathbf{G}^T\sum_{j=1}^{N} c_i^2 u_{ij}\right)\right]^T\mathbf{G} \tag{II.17}$$

where $\mathbf{G} = (g_1, g_2, \ldots, g_H)^T$ is a set of basic functions. It is known that theoretically, any function can be approximated by a linear combination of a set of basis functions up to arbitrary accuracy [Powell, 1981], given a sufficient large number of basis functions. In this implementation, $2D$ orthogonal polynomials are used, specifically the Legendre polynomials [Kim and Park, 1998], as the basis functions.

Furthermore, considering the segmentation of brain MR images as an optimization problem, minimizing $\mathcal{J}_{\text{KFECSB}}$ means that the optimal regions $(\hat{\Omega}_i)_C$, with cluster centers $(\hat{c}_i)_C$ and the bias field $(\hat{b}_j)_N$, are simultaneously accomplished. By using the Lagrange multiplier method, the necessary conditions for the minimization of the $\mathcal{J}_{\text{KFECSB}}$ with the constraints in in Eq. (II.2) can be found. Specifically, taking the first derivatives of the $\mathcal{J}_{\text{KFECSB}}$ with respect to u_{ij} and c_i, and zeroing them respectively, the two necessary but not sufficient conditions for $\mathcal{J}_{\text{KFECSB}}$ to be a local optimal solution will be obtained as follows:

$$u_{ij}^{-1} = \sum_{r=1}^{C}\frac{\exp\left\{n\left[\left(1-K(f_j,c_i b_j)\right)+\eta\left(1-K(\overline{f}_j,c_i b_j)\right)\right]\right\}}{\exp\left\{n\left[\left(1-K(f_j,c_r b_j)\right)+\eta\left(1-K(\overline{f}_j,c_r b_j)\right)\right]\right\}} \tag{II.18}$$

$$c_i = \frac{\sum_{j=1}^{N} u_{ij}\left[f_j K(f_j,c_i b_j)+\eta\overline{f}_j K(\overline{f}_j,c_r b_j)\right]}{\sum_{j=1}^{N} b_j u_{ij}\left[K(f_j,c_i b_j)+\eta K(\overline{f}_j,c_r b_j)\right]} \tag{II.19}$$

Thus, the proposed objective function $\mathcal{J}_{\text{KFECSB}}$ deals with some limits of the original one by including local spatial information, bias correction, and deploying kernel-induced distance. However, Eqs. (II.18) and (II.19) are only necessary conditions for minimizing the objective function $\mathcal{J}_{\text{KFECSB}}$; as a result, the obtained clustering solutions may be local optima if the gradient descent technique is applied. To overcome this drawback, the metaheuristic optimization-based approach is used, as explained in Section II.3.4 below.

II.3.4 A new method for segmentation of brain MR images

To solve the segmentation problem by optimizing the objective function described in Eq. (II.16) efficiently, the LHNPSO algorithm and guidances given by Eqs. (II.18) and (II.19) are used. With this approach, not only trapping into local minima can be avoided but also faster convergence and more accurate solutions can be achieved. Particularly, while the LHNPSO algorithm, which is a global optimization method, can provide global optimal solution, the new objective function guides the search process faster and more precisely. As such, a new MR image segmentation algorithm, called PSO-KFECSB algorithm, is proposed. The details of this algorithm are provided in the next section.

II.3.4.1 Particle representations

Applying the LHNPSO algorithm to optimize the proposed objective function, $\mathcal{J}_{\text{KFECSB}}$, cluster prototypes, $\mathbf{C} = (c_i)_C$, are chosen to be optimization variables and encoded as the positions of particles. For P particles or solutions, there are in total $(C \cdot P)$ optimization variables that need to be encoded. The position of the i^{th} particle can be described as: $\mathbf{X_i} = (x_{i1}, x_{i2}, \ldots, x_{iC})$. Here, x_{ij} $(j = 1, \cdots, C)$ represents the j^{th} cluster center among C centers of the i^{th} solution. In this way, the cluster centers are encoded in position vector $\mathbf{X_i}$ and then $\mathbf{C}$ can be obtained by decoding $\mathbf{X_i}$.

II.3.4.2 Fitness function

After decoding $\mathbf{X_i}$ to obtain cluster centers $\mathbf{C}$ and calculating fuzzy partition matrix $\mathbf{U}$ according to Eq. (II.18), the value of the fitness function, $\mathfrak{f}_\mathfrak{i}$, corresponding to the i^{th} particle, is calculated by evaluating $\mathcal{J}_{\text{KFECSB}}$ according to Eq. (II.16).

$$\mathfrak{f}_\mathfrak{i} = \mathcal{J}_{\text{KFECSB}}\left(\mathbf{F}, \mathbf{X_i}, \mathbf{U_i}, \mathbf{B_i}\right) \tag{II.20}$$

The minimization of $\mathfrak{f}_\mathfrak{i}$ is the same as the minimization of the objective function, $\mathcal{J}_{\text{KFECSB}}$, which leads to an optimal partitioning of the MR image.

II.3.4.3 The PSO-KFECSB algorithm

The proposed PSO-KFECSB algorithm takes advantages of both the excellent feature of LHNPSO algorithm in optimizing the objective function of kernelized fuzzy entropy clustering with spatial information and bias correction, and the guidances provided by the gradient method in speeding up convergence.

To make sure that all particles are moving within the search space and avoiding divergent behavior, their positions and velocities are limited as follows:

$$v_{ij}^{(k)} = \begin{cases} +\,\text{rand}\,() \cdot v_{max}, & \text{if } v_{ij}^{(k)} > v_{max} \\ -\,\text{rand}\,() \cdot v_{max}, & \text{if } v_{ij}^{(k)} < -v_{max} \\ +\,v_{ij}^{(k)}, & \text{otherwise} \end{cases} \tag{II.21}$$

$$x_{ij}^{(k)} = \begin{cases} \frac{1}{2} \cdot \text{rand}\,() \cdot (x_{max} - x_{min}), & \text{if} \quad (x_{ij}^{(k)} < x_{min}) \quad \text{or} \quad (x_{ij}^{(k)} > x_{max}) \\ x_{ij}^{(k)}, & \text{otherwise} \end{cases} \tag{II.22}$$

where v_{max} is the largest allowable step size in any dimension; and $\{x_{min}, x_{max}\}$ are the bounds of the search space in each dimension. In image clustering, commonly, v_{max} is set to 1 and $\{x_{min}, x_{max}\}$

Algorithm II.3: The PSO-KFECSB algorithm

Initialization: Read the input image; initialize all parameters for the LHNPSO algorithm: $\{c_1, c_2, P, \ldots\}$, the objective function, and generate particles using the Halton sequence; randomly set up $\mathbf{U}$ satisfying Eq. (II.2) and set $\mathbf{B} = 1$; determine $pBest$ and $gBest$

Results : The cluster centers $\mathbf{C}$ ($gBest$) and the partition matrix $\mathbf{U}$

$k \leftarrow 1$

repeat

`/* Flight particles */`

1 Update the inertia weight w using Eq. (II.8)

2 Update the positions $\mathbf{X_i}$ and the velocities $\mathbf{V_i}$ using Eqs. (II.6) and (II.7)

3 Map $\mathbf{X_i}$ and $\mathbf{V_i}$ into search space using Eqs. (II.21) and (II.22)

`/* Evaluate solutions */`

4 **for** *each particle (or solution)* **do**

4.1 Calculate kernel distance K using Eq. (II.11)

4.2 Update the partition matrix $\mathbf{U}$ using Eq. (II.18)

4.3 Evaluate the fitness value $\mathfrak{f}_\mathbf{i}$ using Eq. (II.20)

5 Update the $pBest$ and $gBest$, and $\mathbf{B}$

6 $k \leftarrow k + 1$

until *the stopping criteria are met*

are the minimum and maximum of the feature (intensity or gray value) of the image. Putting all of the above issues together, the PSO-KFECSB algorithm is described in **Algorithm** II.3.

In this work, the number of non-significant improvements of the partition matrix ($N_{non-imp}$) and the maximum number of iterations (N_{iter}) are used as the stopping criteria of the algorithm. Particularly, if ($\max\left\{\left|\mathbf{U}_{new}^{gBest} - \mathbf{U}_{old}^{gBest}\right|\right\} < \epsilon$) is completed $N_{non-imp}$ times or the condition ($k > N_{iter}$) is reached, the algorithm is immediately stopped.

II.4 Experimental results

In this section, the performance of the proposed algorithm is evaluated both qualitatively and quantitatively. In the comparative study, the algorithm is confronted with four well-known pixel-based classification methods and one region-based method in the literature: the standard fuzzy entropy clustering (FEC) [Tran and Wagner, 2000], the second version of fuzzy c-means with spatial constraints based on kernel-induced distance (KFCMS2) proposed by Chen and Zhang [Chen and Zhang, 2004], the fuzzy local information c-means (FLICM) [Krinidis and Chatzis, 2010], the conditional spatial fuzzy c-means (csFCM) [Adhikari et al., 2015], and the multiplicative intrinsic component optimization (MICO) [Li et al., 2014].

All of the algorithms are implemented in MATLAB 2014b and executed with a computer with Intel Core i3 1.5 GHz CPU, 4GB RAM using Microsoft Windows 7.

II.4.1 Experimental setup

To perform experiments, the parameters of the proposed algorithm are set as follows: population size $P = 40$, maximum number of allowable iterations $N_{iter} = 100$, number of non-significant improvements $N_{non-imp} = 5$, and termination criteria parameter $\epsilon = 0.0001$. Note that, the kernel parameter

σ^2 in Gaussian kernel has a very important effect on the performances of kernel-based algorithms. However, how to choose an appropriate value is still an open problem. In this work, we adopt the "trial-and-error" technique and set the parameter λ in Eq. (II.12) to compute σ^2 equal to 1/9. We also found that for a range of λ around 1/9 (e.g. from 1/12 to 1/6), there appears no apparent changes in results. Thus, we use this constant value in all experimental evaluations of the proposed algorithm. All of the other parameters for the other algorithms are set to the default values.

II.4.2 Datasets

The MR images used in this study include both T1-weighted simulated and real 2D brain MR images. For simulated MR images, they are downloaded from a well-known database: the brainWeb from a McConnell Brain Imaging Center [Kwan et al., 1999], which can be reached in (https://brainweb.bic.mni.mcgill.ca/brainweb/). This dataset includes six different noise levels (eg. from 0 to 9%), and two different INU levels: 20%, and 40%, and five different slice thicknesses. Images with size of 181×217 and thickness of 1 mm are used in this work. On the other hand, real MR images are taken in the 20-normal MR brain data sets, which contain manual segmentation by an expert technician, provided by the Center for Morphometric Analysis at Massachusetts General Hospital. The data sets are available at (http://www.nitrc.org/projects/ibsr/). Images with characteristics of size 135×142 and 1.171751 mm thickness, are used in our experiments.

II.4.3 Performance measures

Since the ground truth images are available in the datasets, for quantitatively comparing the performance, two criteria are involved, which are the Jaccard index (JAC) [Jaccard, 1912] and the Hausdorff distance (HD) [Beauchemin et al., 1998]. Current research [Taha and Hanbury, 2015] reports that they are appropriate metrics for the evaluation of different segmentation methods when there exist outliers with or without small segments, complex boundaries as well as low densities in the image. These metrics are defined below.

II.4.3.1 Jaccard index

The Jaccard index is an overlap-based metric which directly compares a segmented image ($\mathbf{F_s}$) with a ground truth image ($\mathbf{F_t}$) by measuring similarity between them. A higher value indicates a better result. Given an input image with N pixels $\mathbf{F} = (f_1, f_2, \ldots, f_N)$, and its two partitions, $\mathbf{F}_\mathrm{s} = (F_{\mathrm{s}1}, F_{\mathrm{s}2}, \cdots, F_{\mathrm{s}N})$ (the segmented image) and $\mathbf{F}_\mathrm{t} = (f_{\mathrm{t}1}, f_{\mathrm{t}2}, \cdots, f_{\mathrm{t}N})$ (the ground truth image), there are four common cardinalities that reflect the overlap between the two partitions, namely the true positives (TP), the false positives (FP), the true negatives (TN), and the false negatives (FN). Then, the Jaccard index is defined as follows:

$$\mathrm{JAC}\left(\mathbf{F_s}, \mathbf{F_t}\right) = \frac{\mathrm{TP}}{\mathrm{TP} + \mathrm{FP} + \mathrm{FN}} \tag{II.23}$$

II.4.3.2 Hausdorff distance

The Hausdorff distance is a distance-based metric which measures the dissimilarity between the segmented image ($\mathbf{F}_\mathrm{s}$) and the ground truth image ($\mathbf{F}_\mathrm{t}$). A lower value indicates a better result. This is also a widely used metric, defined as follows:

$$\mathrm{HD}\left(\mathbf{F}_\mathrm{s}, \mathbf{F}_\mathrm{t}\right) = \max\left\{h\left(\mathbf{F}_\mathrm{s}, \mathbf{F}_\mathrm{t}\right), h\left(\mathbf{F}_\mathrm{t}, \mathbf{F}_\mathrm{s}\right)\right\} \tag{II.24}$$

where $h(\mathbf{F}_s, \mathbf{F}_t)$ is called the directed Hausdorff distance given by:

$$h(\mathbf{F}_s, \mathbf{F}_t) = \max_{f_{si} \in \mathbf{F}_s} \min_{f_{ti} \in \mathbf{F}_t} \|f_{si} - f_{ti}\| \tag{II.25}$$

where $\|f_{si} - f_{ti}\|$ is the Euclidean distance between the intensity values of the f_{si} pixel and the f_{ti} pixel in the segmented and ground truth images, respectively.

II.4.4 Parameters setting

There are two parameters n and η of the fitness function $\mathfrak{f}_i$, which have significant influence on the partition matrix $\mathbf{U} = (u_{ij})_{C \times N}$ and the cluster prototypes $\mathbf{C} = (c_i)_C$, according to Eqs. (II.18) and (II.19), and thereby effecting the final segmentation results. To determine these parameter values, a 3-step procedure has been carried out. First, estimating performance in terms of Jaccard index (JAC), according to Eq. (II.23), of the proposed algorithm is done by fixing η and modifying n. Then, the same evaluation of the algorithm is estimated by fixing n and modifying η. Finally, a set of (n, η) values in their ranges which have strong effects on the performance is examined so that the nearly optimal (n, η) values can be found. Note that the (n, η) values cannot be high because the membership degree u_{ij} and the cluster centers c_i in Eqs. (II.18) and (II.19) are the exponential functions of these parameters, and high values of these parameters will lead to u_{ij}, c_i becoming sensitive to the noise. Hence, in our work, the ranges of (n, η) values are set to $[1, 20]$.

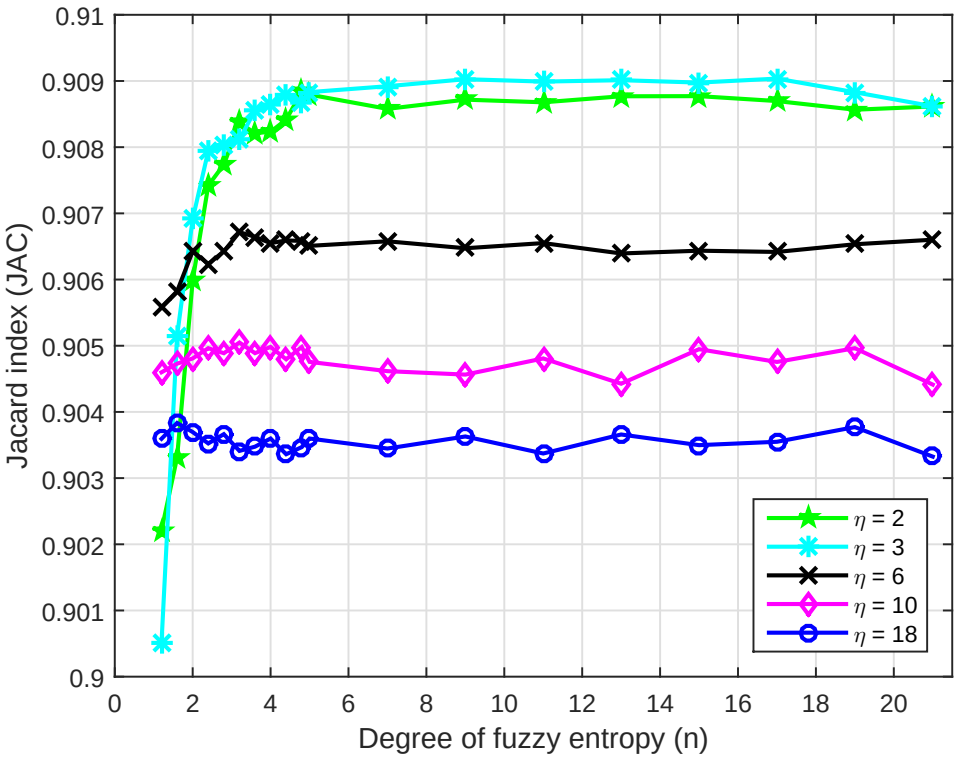

Figure II.2: Performance of the PSO-KFECSB algorithm in terms of Jaccard index by fixing η and modifying n

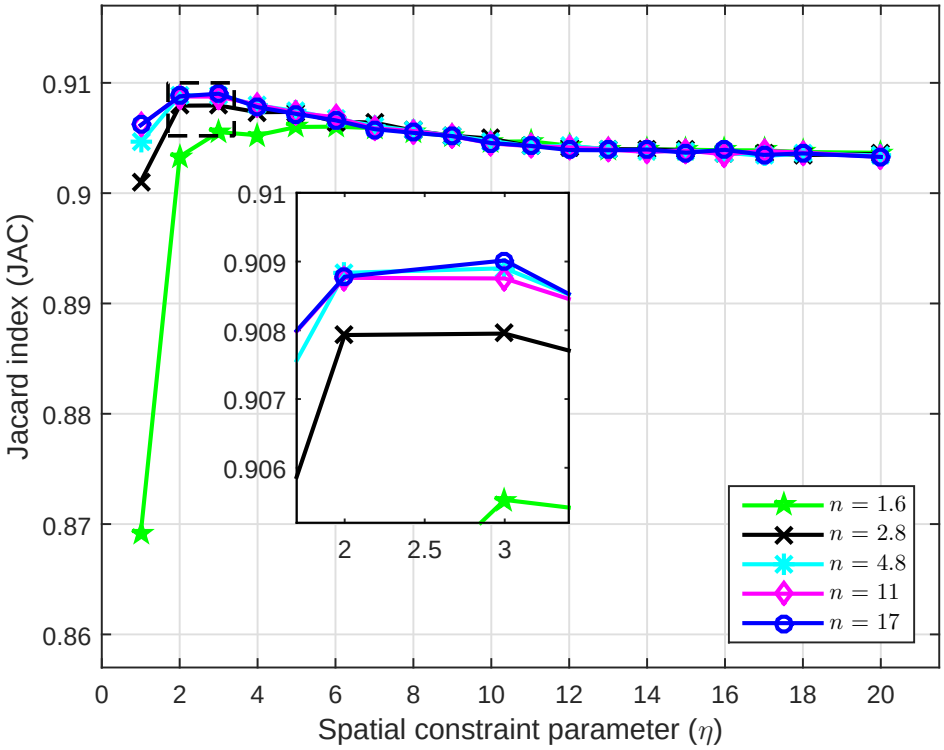

Figure II.3: Performance of the PSO-KFECSB algorithm in terms of Jaccard index by fixing n and modifying η

Table II.1: Performance of the PSO-KFECSB algorithm in terms of Jaccard index with different simulated MR images and different (n, η) values

$(\mathbf{n}, \eta)$	**Z planes with 7% noise and 20% INU**										**Average**
	50	**55**	**60**	**65**	**70**	**75**	**80**	**85**	**90**	**95**	
$(4, 1)$	0.8440	0.7855	0.7964	0.9089	0.9061	0.9030	0.9039	0.9115	0.9215	0.9291	0.8810
$(17, 1)$	0.8900	0.8790	0.9140	0.9118	0.9079	0.9043	0.9062	0.9127	0.9224	0.9302	0.9078
$(6, 2)$	0.8922	0.8813	0.9162	0.9145	0.9120	0.9093	0.9086	0.9149	0.9234	0.9316	0.9104
$(10, 2)$	0.8933	0.9016	0.9165	0.9142	0.9119	0.9091	0.9086	0.9151	0.9233	0.9316	0.9125
(10, 2.5)*	0.8933	0.9035	0.9173	0.9141	0.9117	0.9089	0.9088	0.9150	0.9223	0.9317	**0.9126**
$(10, 3)$	0.8926	0.9040	0.9160	0.9134	0.9116	0.9082	0.9090	0.9143	0.9222	0.9315	0.9123
$(17, 3)$	0.8927	0.9041	0.9164	0.9131	0.9112	0.9084	0.9090	0.9143	0.9222	0.9314	0.9123
$(4, 6)$	0.8896	0.8973	0.9141	0.9108	0.9098	0.9065	0.9067	0.9123	0.9203	0.9300	0.9098
$(17, 6)$	0.8894	0.9024	0.9142	0.9114	0.9097	0.9064	0.9064	0.9121	0.9206	0.9297	0.9102
$(20, 6)$	0.8916	0.8963	0.9146	0.9112	0.9099	0.9063	0.9067	0.9121	0.9206	0.9298	0.9099

*The values on gray background indicate the best performance.

Figure II.2 illustrates the performance of the PSO-KFECSB algorithm in terms of Jaccard index by fixing η and modifying n. From the figure, it is obvious that the PSO-KFECSB algorithm provides superior results using $\eta = 2$ and $\eta = 3$. In addition, the results also indicate that the performance is quite constant when modifying n in the interval $[6, 18]$.

Figure II.3 illustrates the performance of the PSO-KFECSB algorithm in terms of Jaccard index

by fixing n and modifying η. It confirms more properly that the PSO-KFECSB algorithm provides superior results when η value is in the interval $[2, 3]$ and changing n in the interval $[6, 18]$ has a minor effect.

Table II.1 gives the average Jaccard index scores of the proposed algorithm on different simulated MR images with different (n, η) values which are selected from the conclusions of two steps above. Note that each provided result is the average of 10 independent program runs throughout the paper. It is clearly illustrated that the PSO-KFECSB algorithm with the values of $(10, 2.5)$ for the parameters provides better performance than the others. Therefore, the subsequent experiments are performed by using $n = 10$ and $\eta = 2.5$.

II.4.5 Results on simulated MR images

In this section, simulated MR brain images from the brainWeb are used for the purpose of performance evaluation. The experiment is conducted on a data set of 30 images. These images having characteristics of Mobility T1-weighted, slice thickness 1mm, with three different levels of noise $(5\%, 7\%, 9\%)$, and two different levels of non-uniformity $(20\%, 40\%)$, are segmented into 4 regions: background, cerebral spinal fluid (CSF), gray matter (GM), and white matter (WM).

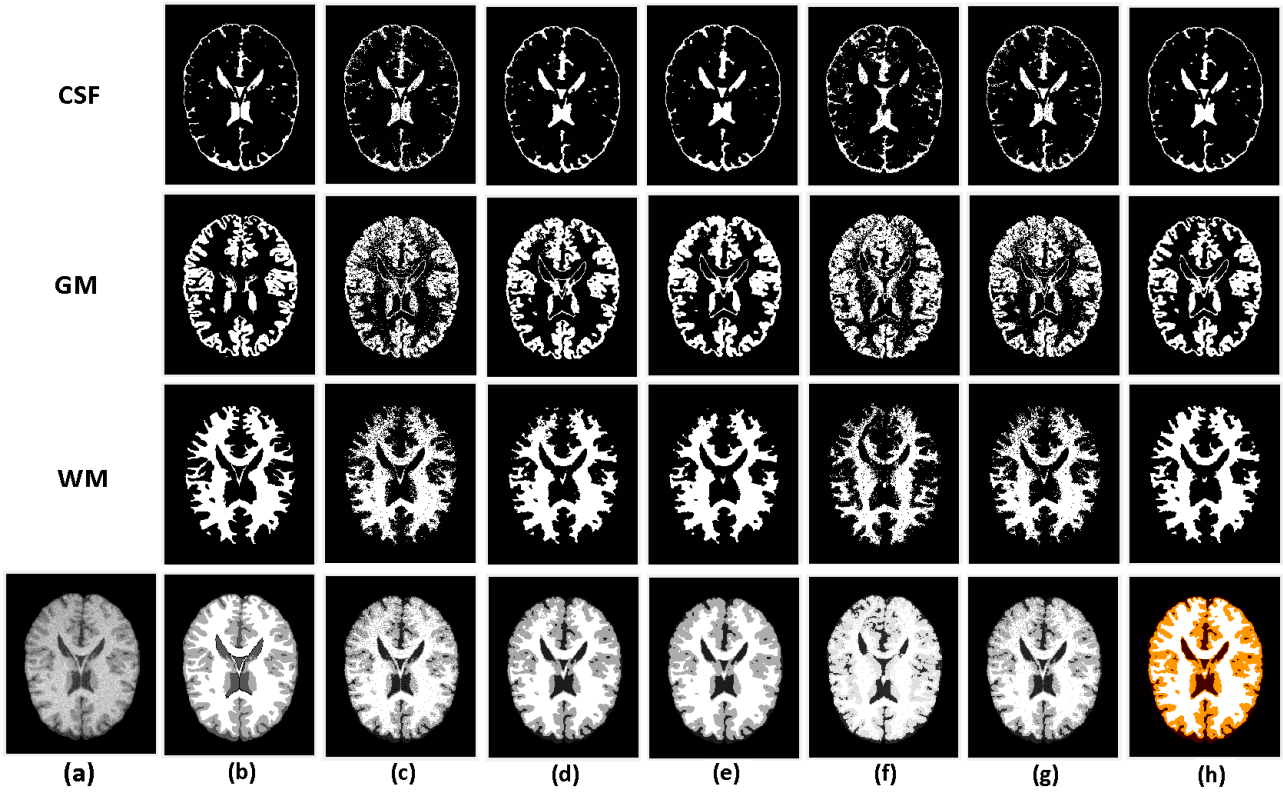

Figure II.4: Qualitative results of segmentation of simulated brain MR image (slice 90) with 9% noise and 40% INU provided by the different algorithms:(a) original image; (b) ground truth; (c) FEC results; (d) KFCMS2 results; (e) FLICM results; (f) MICO results; (g) csFCM results; (h) PSO-KFECSB results.

Figure II.4 shows the qualitative results of the segmentation of T1-weighted image (slice 90) with 9% noise and 40% INU by the different competing algorithms. This figure reveals that though the KFCMS2, FLICM and the proposed PSO-KFECSB algorithm, yield positive outcomes, those

from the proposed algorithm are the best ones, particularly, the GM segment. Figure II.5 shows the qualitative results of the segmentation of 7 T1-weighted images (slices 55, 60, 65, 70, 75, 80 and 85) with 7% noise and 20% inhomogeneity. It can be seen that the proposed algorithm achieves an impressive performance when segmenting simulated brain MR images. To obtain the results, the PSO-KFECSB algorithm spends approximately 300 seconds to converge. Thus, it can be concluded that our proposed algorithm qualitatively provides satisfactory results with better performance than the other algorithms.

Table II.2: Average values of the total segmented results in terms of Jaccard index and Hausdorff distance on T1-weighted brain MR images using different algorithms.

INU	Noise	Metrics	FEC	KFCMS2	FLICM	MICO	csFCM	Proposed
20%	5%	JAC	0.9096	0.9192	0.9081	0.7402	0.9183	**0.9267***
		HD	057.00	061.20	065.20	082.20	**043.80**	058.80
	7%	JAC	0.8739	0.9087	0.9020	0.7114	0.8902	**0.9145**
		HD	123.00	049.60	064.20	121.60	095.00	**038.00**
	9%	JAC	0.8308	0.8962	0.8937	0.6710	0.8476	**0.9004**
		HD	122.20	055.00	067.40	127.20	118.60	**037.60**
40%	5%	JAC	0.8949	0.9040	0.8878	0.7439	0.9017	**0.9290**
		HD	082.00	067.60	069.20	083.00	082.00	**040.20**
	7%	JAC	0.8604	0.8908	0.8812	0.7174	0.8735	**0.9170**
		HD	108.20	061.40	063.00	109.20	098.40	**054.00**
	9%	JAC	0.8154	0.8778	0.8738	0.6877	0.8365	**0.9004**
		HD	114.00	055.20	060.80	125.40	114.20	**044.40**

*The values in bold indicate the best performance.

In order to compare more clearly the performance of different segmentation algorithms, the quantitative evaluation is considered. Tables II.2 - II.4 show the average values of Jaccard index and Hausdorff distance metrics of six competing algorithms: FEC, KFCMS2, FLICM, MICO, csFCM, and the proposed PSO-KFECSB. Table II.2, which summarizes the average values of the overall results of the segmentation, is a striking illustration of the fact that the PSO-KFECSB algorithm significantly outperforms the existing techniques. In addition, as can be seen from Table II.3, when segmenting images with various added noise levels in the images (the same INU artifact level), the proposed method generally gives the best scores, except for CSF segmenting scores in terms of Hausdorff distance. Similarly, when segmenting images with different levels of INU artifact in the images (the same added noise level), the data from Table II.4 demonstrate that our proposed method performs more efficiently on simulated brain MR images compared to the others. Thus, the proposed algorithm provides more accurate and stable segmentation results compared to its competitors.

II.4.6 Results on real MR images

We have also examined the performance of the proposed method on real MR brain images. The number of tissue regions in the segmentation is set to three, which corresponds to the cerebrospinal fluid (CSF), gray matter (GM), and white matter (WM). The background pixels are ignored in the computation.

Table II.3: Average values on T1-weighted brain MR images with the same level of INU artifact and different levels of noise using different algorithms.

Regions	INU	Metrics	FEC	KFCMS2	FLICM	MICO	csFCM	proposed
CSF	**20%**	JAC	0.9655	0.9706	0.9644	0.9661	0.9662	**0.9727***
		HD	173.00	185.20	247.27	211.00	**165.47**	189.33
	40%	JAC	0.9578	0.9672	0.9605	0.9668	0.9600	**0.9735**
		HD	194.27	183.80	251.53	**171.1**	193.93	192.20
GM	**20%**	JAC	0.8550	0.8905	0.8847	0.8075	0.8645	**0.8965**
		HD	107.47	094.00	097.60	120.80	103.07	**089.40**
	40%	JAC	0.8238	0.8669	0.8583	0.8068	0.8371	**0.8992**
		HD	121.73	089.67	096.67	123.47	116.27	**086.07**
WM	**20%**	JAC	0.8929	0.9288	0.9276	0.8401	0.9043	**0.9332**
		HD	130.13	062.93	081.60	184.40	109.40	**062.80**
	40%	JAC	0.8669	0.9074	0.9038	0.8390	0.8819	**0.9349**
		HD	144.87	076.73	103.80	183.60	122.07	**071.20**

*The values in bold indicate the best performance.

Table II.4: Average values on T1-weighted brain MR images with the same level of noise and different INU artifact using different algorithms.

Regions	Noise	Metrics	FEC	KFCMS2	FLICM	MICO	csFCM	Proposed
CSF	5%	JAC	0.9736	0.9731	0.9652	0.9765	0.9748	**0.9780***
		HD	**148.80**	186.70	263.80	162.50	176.80	198.40
	7%	JAC	0.9660	0.9699	0.9629	0.9686	0.9686	**0.9736**
		HD	190.70	181.90	261.00	193.40	**143.60**	196.00
	9%	JAC	0.9454	0.9637	0.9593	0.9544	0.9459	**0.9678**
		HD	211.40	184.90	223.40	217.30	218.70	**177.90**
GM	5%	JAC	0.8743	0.8914	0.8788	0.8451	0.8847	**0.9138**
		HD	106.20	093.70	098.70	106.10	093.20	**088.70**
	7%	JAC	0.8450	0.8806	0.8731	0.8121	0.8614	**0.8991**
		HD	118.20	089.10	095.20	128.50	110.90	**089.00**
	9%	JAC	0.8257	0.8860	0.8863	0.7835	0.8345	**0.9003**
		HD	119.40	092.70	097.50	131.80	124.90	**085.20**
WM	5%	JAC	0.9078	0.9280	0.9214	0.8708	0.9185	**0.9466**
		HD	105.10	071.00	090.20	173.30	067.40	**062.70**
	7%	JAC	0.8831	0.9193	0.9169	0.8430	0.8997	**0.9350**
		HD	161.50	073.10	096.60	188.40	119.90	**056.30**
	9%	JAC	0.8489	0.9070	0.9088	0.8049	0.8610	**0.9207**
		HD	145.90	**065.40**	091.30	190.30	159.90	082.00

*The values in bold indicate the best performance.

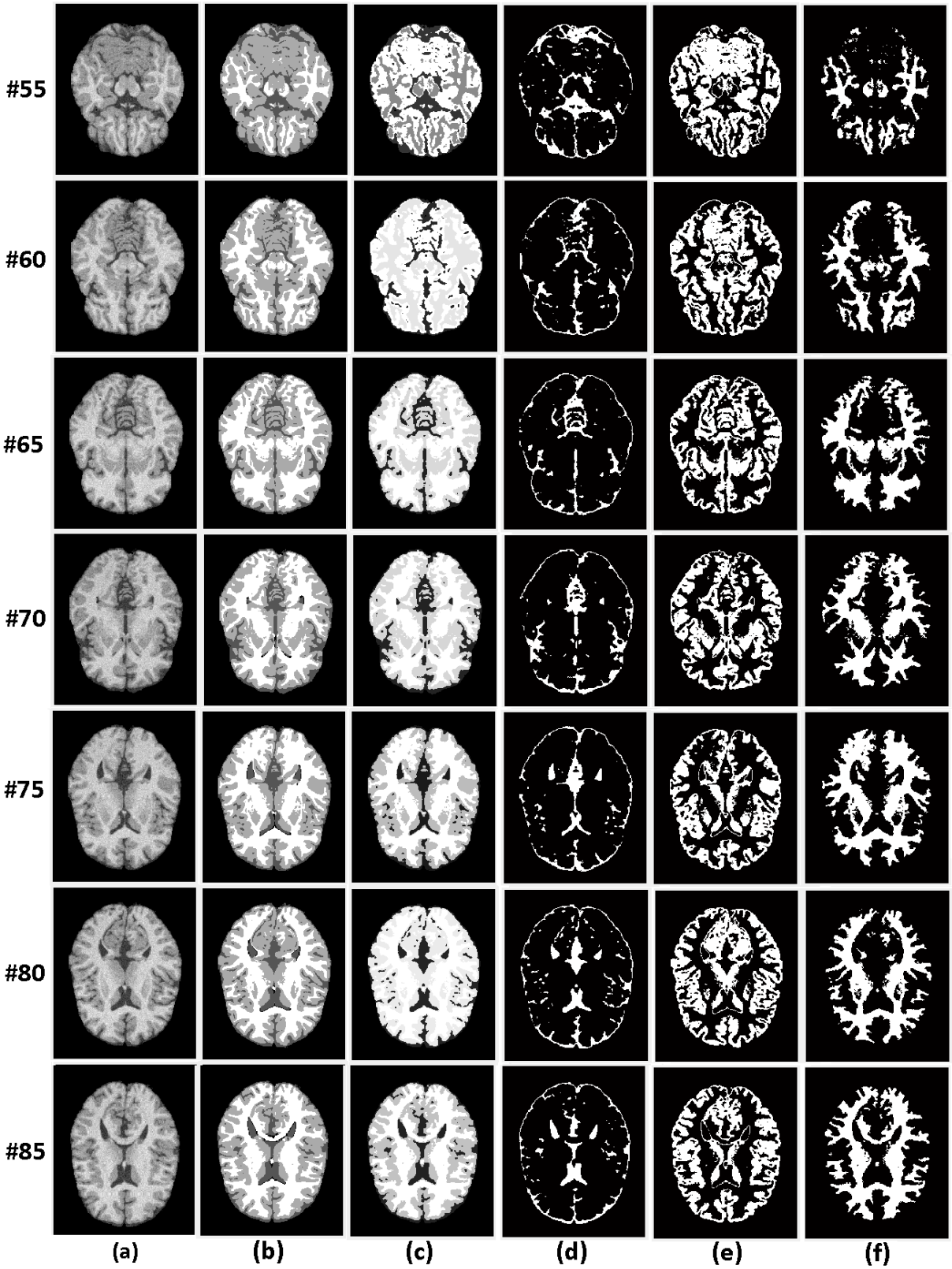

Figure II.5: Qualitative results of segmentation of simulated brain MR images (slice 55, 60, 65, 70, 75, 80, and 85) with 9% noise and 40% INU provided by the proposed algorithm versus the ground truth images: (a) original images; (b) ground truth images; (c) segmentation results; (d) CSFs; (e) GMs; (f) WMs.

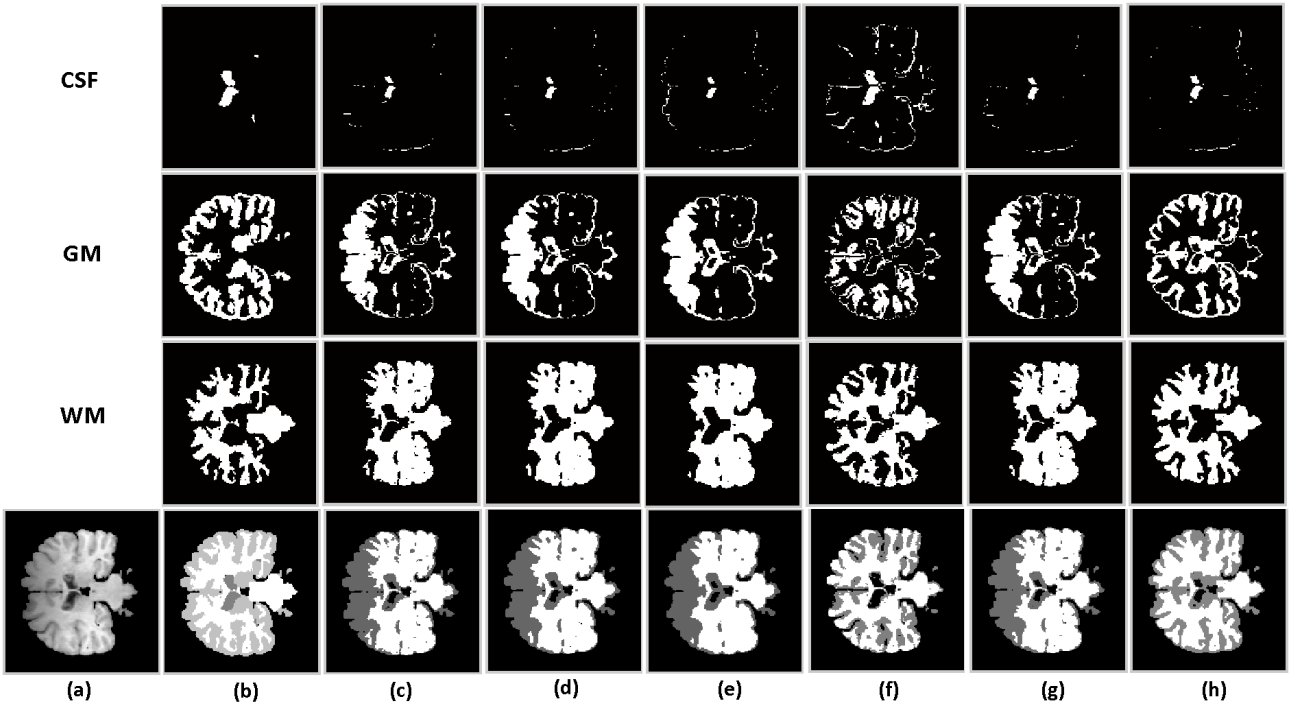

Figure II.6: Qualitative results of segmentation of real brain MR image (Z plane 25 in the 4^{th} volume) provided by the different algorithms:(a) original image; (b) ground truth; (c) FEC results; (d) KFCMS2 results; (e) FLICM results; (f) MICO results; (g) csFCM results; (h) PSO-KFECSB results.

Figure II.6 shows the qualitative results of the segmentation of T1-weighted image (Z plane 25) by the competing algorithms. Through a careful look at the results presented in this figure, it can be seen that our proposed algorithm provides superior results as compared to other five different algorithms. Among the rest, even though the MICO algorithm produces a satisfactory WM segment, the other results (CSF and GM segments) from this algorithm are not acceptable.

Table II.5: Average values in terms of Jaccard index and Hausdorff distance using real brain MR images

Regions	Metrics	FEC	KFCMS2	FLICM	MICO	csFCM	Proposed
CSF	JAC	0.9756	0.9777	0.9758	0.9480	0.9756	**0.9835***
	HD	1635.8	1827.2	1776.0	1821.0	1636.0	**1627.2**
GM	JAC	0.7449	0.7460	0.7361	0.7593	0.7476	**0.7985**
	HD	187.80	186.60	184.80	182.60	187.20	**181.60**
WM	JAC	0.7539	0.7570	0.7529	0.8011	0.7590	**0.8160**
	HD	147.80	169.60	175.60	119.00	166.20	**082.20**
Total	JAC	0.8610	0.8654	0.8736	0.8911	0.8615	**0.9019**
	HD	076.40	079.80	091.60	054.60	079.00	**046.80**

*The values in bold indicate the best performance.

Figure II.7 shows qualitative results of the segmentation of 5 normal T1-weighted images (Z planes 25, 27, 30, 35, and 37 in the 4^{th} volume) by the PSO-KFECSB algorithm. This figure reveals that the proposed algorithm accomplishes appropriately the segmentation of real brain MR images. It may be worth mentioning here that though the KFCMS2 and FLICM algorithms work quite well on simulated brain MR images, as shown in Fig. II.4, these algorithms provide unsatisfactory results on real MRI brain images. On the other hand, the PSO-KFECSB algorithm does well in both cases.

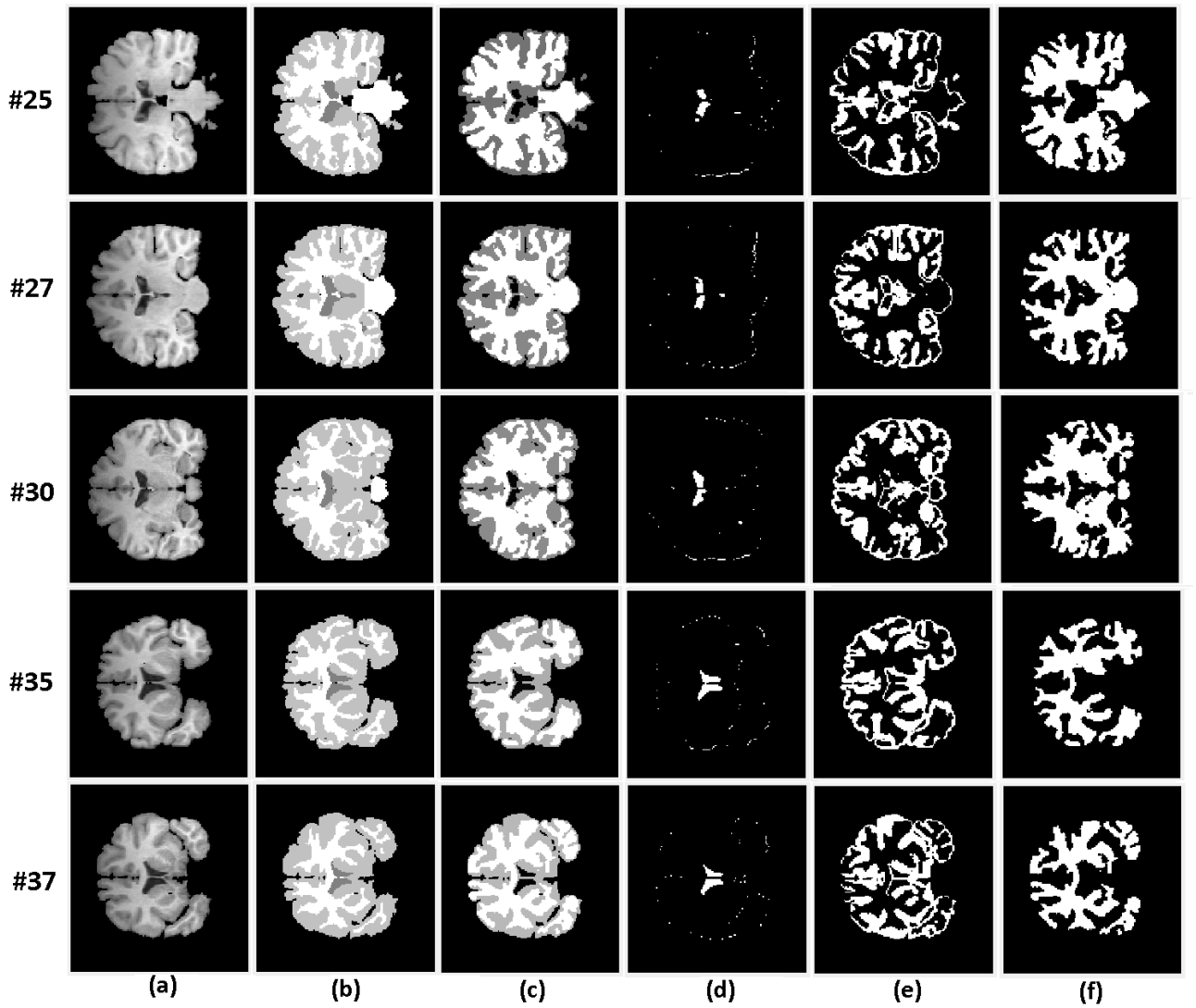

Figure II.7: Qualitative results of segmentation of real brain MR images (Z planes 25, 27, 30, 35, 37 in the 4^{th} volume) provided by the PSO-KFECSB algorithm versus the ground truth images: (a) original images; (b) ground truth images; (c) segmentation results; (d) CSFs; (e) GMs; (f) WMs.

Table II.5 shows the average values of the Jaccard index and the Hausdorff distance metrics using real brain MR images (Z planes 25, 27, 30, 35, and 37 in the 4^{th} volume) for six competing algorithms: FEC, KFCMS2, FLICM, MICO, csFCM, and the proposed PSO-KFECSB. Note that to archive the results, the PSO-KFECSB algorithm spends around 150 seconds to converge. These results show that the PSO-KFECSB algorithm outperforms the FEC, KFCMS2, FLICM, MICO, csFCM algorithms for segmenting the images. Again, the results presented here show the efficiency of the PSO-KFECSB algorithm and also demonstrate its superiority over its competitors.

II.5 Conclusion

In this chapter, a new clustering method, which takes both advantages of improved PSO algorithm, LHNPSO algorithm, and kernelized fuzzy entropy clustering with spatial information and bias correction, called PSO-KFECSB algorithm, is presented. This method is used to segment brain MR images. There are three major improvements, which have been done here. First, kernel method has been applied to measure the similarity between data points (intensity of pixels) such that the problem of complex structure of brain MR images can be overcome. Second, a novel objective function, which takes into account both spatial information and bias field existing in input image, has been proposed. Finally, the LHNPSO algorithm is used to deal with the problem of getting stuck in local minima. Therefore, the two major drawbacks of fuzzy clustering algorithms, which are the sensitivity to noise and INU artifact and the trapping of the solution into local minima, have been partially overcome; thereby providing superior segmentation results. The proposed algorithm has been tested on both simulated and real brain MR images. The experiments are examined qualitatively as well as quantitatively, using the Jaccard index and the Hausdorff distance. The experimental results show that the proposed algorithm is more effective in comparison with five powerful states of the art methods in the literature. Particularly, on the one hand, the visual segmentation results of proposed algorithm show more compact and with less artifacts than its competitors; on the other hand, the quantitative results of analysing segmentation also prove the outstanding of PSO-KKFECSB algorithm compared to its competitors. However, when high level of noise along with INU artifact are added into MRI data, the performance of the PSO-KFECSB algorithm may decrease. In addition, only one criterion $\mathcal{J}_{\text{KFECSB}}$ is used to guide the solution-searching process, which may lead to the situation where the solution is a global optimum specific to the criterion used, but may not be the best optimum for the segmentation. To produce better segmentation results, we propose to use multi-objective optimization approach in which the complementary strengths of other criteria can be exploited. This new method is presented in detail in the next chapter.

Note that the research reported in this chapter gave rise to our publications [Pham et al., 2017a,b, 2018].

Chapter III

Multi-objective optimization metaheuristic approach for image segmentation using fuzzy entropy clustering and region-based active contour

III.1 Introduction

In the previous chapter, we proposed an improvement of fuzzy entropy clustering (FEC) for segmentation of brain magnetic resonance (MR) images. Despite good results obtained, we noticed that segmentation errors still remain. In order to further improve performance, we propose in this chapter to add a novel segmentation criterion based on region-based active contour. Both criteria are optimized simultaneously in multi-objective optimization approach via an improved multi-objective particle swarm optimization algorithm. From that, a decision maker is used to select the best trade-off solution among the solutions of the Pareto front. As a result, higher accurate image segmentation, with two independent characteristics, named compactness and separation, that are derived from the fuzzy entropy clustering and the region-based active contour approaches, can be achieved.

This chapter is organized as follows. In Section III.2, we present the background information on which the proposed method is based. Section III.3 introduces the proposed method: segmentation of brain MR images using multi-objective optimization approach. Evaluation and comparison with a set recent methods in the literature are shown in Section III.4. The results of the comparative evaluation are described and discussed in Section III.5. Finally, conclusion is drawn in Section III.6.

III.2 Multi-objective optimization approach for image segmentation

In this section, we present different aspects of multi-objective optimization approach for the image segmentation problem. The details are described in the following sections.

III.2.1 Problem formulation and proposed solution

In the previous chapter, the model of MR image formation described in Eq. (II.9) is based on assumptions, that intensity inhomogeneity varies slowly over the entire image domain; otherwise, the intensity would be constant for the same tissue type. Besides, the intensity inhomogeneity is in multiplicative type since it is consistent with the inhomogeneous sensitivity of the reception coil. Noise in the image is approximated by Gaussian probability distribution and is independent of the intensity inhomogeneity field. To model the input image more accurately, we use the same approach proposed by Huang and Zeng [Huang and Zeng, 2015], who assume that there is a difference between

the measured image and the traditionally approximated models in the local region. The difference is also independent of the intensity inhomogeneity and noise. Accordingly, an acquired MR image can be modelled as:

$$(f_j)_N = \left(f_j^{\mathrm{t}}\right)_N (b_j)_N + (d_j)_N + (n_j)_N \tag{III.1}$$

where f_j is the measured intensity of the j^{th} MR pixel; f_j^{t} is the true intensity value approximated by a constant c_i^{t} for all pixels in the i^{th} tissue; b_j is the unknown smoothly varying bias field, d_j is the local difference between the measured image and the traditionally approximated model, n_j is an additive zero-mean Gaussian noise with relatively small variance at the j^{th} pixel, and N is the number of pixels in the MR image.

Considering an input image modelled by Eq. (III.1) as a decomposition of four components: $(f_j^{\mathrm{t}})_N$, $(b_j)_N$, $(d_j)_N$, and $(n_j)_N$, we formulate the image segmentation problem as a multi-objective optimization (MOO) problem and a decision-making process of finding optimal values of these components with a view of an image as a function $\Gamma : \Omega \to \Re$ on a continuous domain Ω. The MOO approach to the problem is motivated by seeking two important properties in segmented images: compactness and separation, from two complementary approaches named fuzzy entropy clustering and region-based active contour. It is well known that fuzzy entropy clustering is to reduce the variation or scattering of the data within a particular tissue (compactness), and region-based active contour attempts to isolate the clusters/regions from each other (separation). Note that, the local spatial information and bias correction are also considered to deal with noise and INU artifact. After a set of non-dominated Pareto optimal solutions is produced, a decision maker will take place to find the best trade-off solution between the two characteristics.

III.2.2 Multi-objective optimization

From a mathematical viewpoint, a general MOO problem can be formulated as follows:

$$\begin{aligned} &\text{minimize: } \mathcal{Q}(\boldsymbol{x}) = (q_1(\boldsymbol{x}), q_2(\boldsymbol{x}), \ldots, q_m(\boldsymbol{x}))^T \\ &\text{subject to } \begin{cases} g_j(\boldsymbol{x}) \geq 0, & j = 1, \cdots, J \\ h_k(\boldsymbol{x}) = 0, & k = 1, \cdots, K \end{cases} \end{aligned} \tag{III.2}$$

where J and K are the numbers of inequality and equality constraints, respectively. $\boldsymbol{x} = (x_1, x_2, \ldots, x_L)$ is a L-dimensional candidate solution in search space Ω. The mapping function $\Delta : \Omega \to \Re^m$ defines m objective functions $(q_i(\boldsymbol{x}), i = 1, ..., m)$ and $\Re^m$ is called the objective space.

The solving of a MOO problem relies on a crucial concept, which is that of dominance. A decision vector $\boldsymbol{x_p} = (x_{p1}, x_{p2}, \ldots, x_{pL})$ is said to strictly dominate another decision vector $\boldsymbol{x_r} = (x_{r1}, x_{r2}, \ldots, x_{rL})$, denoted by $\boldsymbol{x_p} \prec \boldsymbol{x_r}$. If the following conditions are met, one can say that $\boldsymbol{x_p}$ dominates $\boldsymbol{x_r}$ or $\boldsymbol{x_p}$ is better than $\boldsymbol{x_r}$:

$$\forall i : q_i(\boldsymbol{x_p}) \leq q_i(\boldsymbol{x_r}) \quad \text{and} \quad \exists j : q_j(\boldsymbol{x_p}) < q_j(\boldsymbol{x_r}) \tag{III.3}$$

where $i = 1, ..., m$; $j = 1, ..., m$. In MOO, when a solution is not dominated by any others, it is referred to as a Pareto optimal solution. The latter is said to be non-dominated, and the set of all non-dominated solutions forms the *Pareto front* of optimal solutions, a consequence of not being able to minimize or maximize all the objective functions simultaneously.

Note that, in simultaneous multi-objective optimization of criteria, it is possible that a conflict appears. It is more often that one criterion cannot be improved without damaging at least one other. In addition, it is rare to be in a situation where a single solution $\hat{\boldsymbol{x}}$ represents the optimal solution

for all objective functions. In the literature, there are several approaches for solving multi-objective optimization problems [Deb, 2001]. In this work, we will benefit from the optimization metaheuristic approach by using multi-objective particle swarm optimization (MOPSO) described in the next section.

III.2.3 Multi-objective particle swarm optimization algorithm

In order to solve multi-objective optimization problems, the original PSO algorithm needs to be adapted. In terms of multi-objective optimization, each particle can have several non-dominant *pBest*. The swarm can also have several non-dominant *gBest*. As a result, extending PSO to MOO problems raises three fundamental questions [Reyes-Sierra and Coello, 2006]:

(i_1) How to select the *pBest* for each particle and the *gBest* for the swarm to give preference to non-dominated solutions ?

(i_2) How to retain the non-dominated solutions found during the search process so that the solutions are non-dominated with respect to all the past populations, not only to the current one ?

(i_3) How to maintain diversity in the swarm in order to avoid convergence towards a single solution ?

Algorithm III.1: The improved MOPSO algorithm

Initialization: Set the maximum number of iterations N_{iter}; set $\{w, c_1, c_2\}$; initialize population P; determine *pBest* and *gBest* and store them in external archives pA(t) and gA(t), respectively; generate hypercubes as coordinate systems

Results : Report results in the gA(t)

$k \leftarrow 1$

repeat

`/* Flight particles */`

1 Update the flight parameters using Eq. (III.4)

2 Update $\mathbf{X_i}$ and $\mathbf{V_i}$ according to Eqs. (II.6) and (II.7)

3 Do mutation as detailed in [Sierra and Coello, 2005]

4 Check boundary conditions for both $\mathbf{X_i}$ and $\mathbf{V_i}$

`/* Evaluate and maintain solutions */`

5 **for** *each particle (or solution)* **do**

Evaluate fitness values $\{q_1(\mathbf{X_i}), q_2(\mathbf{X_i}), \ldots, q_m(\mathbf{X_i})\}$

6 Determine the *pBest* using the strategy described in Section III.2.3.2

7 Determine the *gBest* using the strategy described in Section III.2.3.3

8 Update and store non-dominated solutions in pA(t) and gA(t)

9 Maintain gA(t) using crowding distance technique

10 Update the contents in the hypercubes as detailed in [Sierra and Coello, 2005]

11 $k \leftarrow k + 1$

until *the stopping criteria are met*

In this work, the algorithm proposed by Sierra and Coello [Sierra and Coello, 2005] is considered as a basic MOPSO algorithm to develop for solving the image segmentation problem. To deal with the questions mentioned above, we utilize the state of the art in MOPSO developments, particularly in the selection of *pBest*, *gBest*, and updating flight parameters ($w, c_1,$ and c_2). As a result, the convergence

speed is increased and the diversity of non-dominated solutions is maintained. The procedure of developed MOPSO algorithm is summarized in **Algorithm** III.1.

Note that, the crowding distance is calculated by using the procedure proposed by Kalyanmoy Deb et al. [Deb et al., 2002], which is described in **Algorithm** III.2.

Algorithm III.2: Crowding distance procedure

```
1 Procedure crowdingDist()
    Inputs : The non-dominated solutions with size l in the gA(t)
    Output: The distance corresponded to each solution, Dist
    /* Initialization                                          */
2   for each particle (or solution i) do
3     Dist[i] ← 0
    /* Calculation                                             */
4   for each objective (m) do
5     Sort fitness values
6     Dist[1] ← ∞ and Dist[l] ← ∞
7     for i = 2   to   (l − 1) do
8       Dist[i] ← Dist[i] + [ (gA(i+1)_m − gA(i−1)_m) / (q_m^max − q_m^min) ]
```

Here, $gA(i)_m$ refers to the m^{th} fitness value of the i^{th} individual in the global archive $gA(t)$ and parameters $\mathfrak{q}_m^{max}$ and $\mathfrak{q}_m^{min}$ are the maximum and minimum values of the m^{th} objective function, respectively.

Furthermore, the hypercubes in this work are currently built using an adaptive grid [Knowles and Corne, 2000]. At each iteration, the objective space is divided using a grid so that the crowding of the solutions is measured using the crowding distance in the objective space within the grid. As a result, solutions at the highly populated cells can be removed or replaced to reduce computational overhead and overcome the problem of size limitation of the storing archives.

III.2.3.1 Flight parameter mechanism

To have a better global exploration of the search space, most of existing works on MOPSO show that (w, c_1) parameters should be large and c_2 should be small at the beginning stage of iterations. Then, towards the end of the algorithm, to promote a better local exploitation, (w, c_1) parameters should be small and c_2 should be large. Hence, the law of updating flight parameters $(w, c_1, \text{and } c_2)$ in this work is adopted from Zhang et al.'s work [Zhang et al., 2017], in which the w is fixed (equal to 0.4) and the (c_1, c_2) are modified as follows:

$$c_1^{(k)} = 2.5 - 2.\,(k/N_{iter})\,,\;\; c_2^{(k)} = 0.5 + 2.\,(k/N_{iter}) \tag{III.4}$$

where k and N_{iter} are, respectively, the iteration number starting from 1 and maximum number of allowable iterations.

III.2.3.2 Personal best selection mechanism

In this work, the "Diversity" strategy proposed by Branke and Mostaghim [Branke and Mostaghim, 2006] is used because of its improvement ability in maintaining diversity for MOPSO algorithm. Par-

ticularly, the *pBest* is stored in a personal archive and updated in each generation. Among all the *pBest* solutions visited in the past, the one which has the largest minimum crowding distance to any others, is selected. In other words, one selects the most isolated solution from the others as the current *pBest* to force the particle to explore the regions which are far from those in the population.

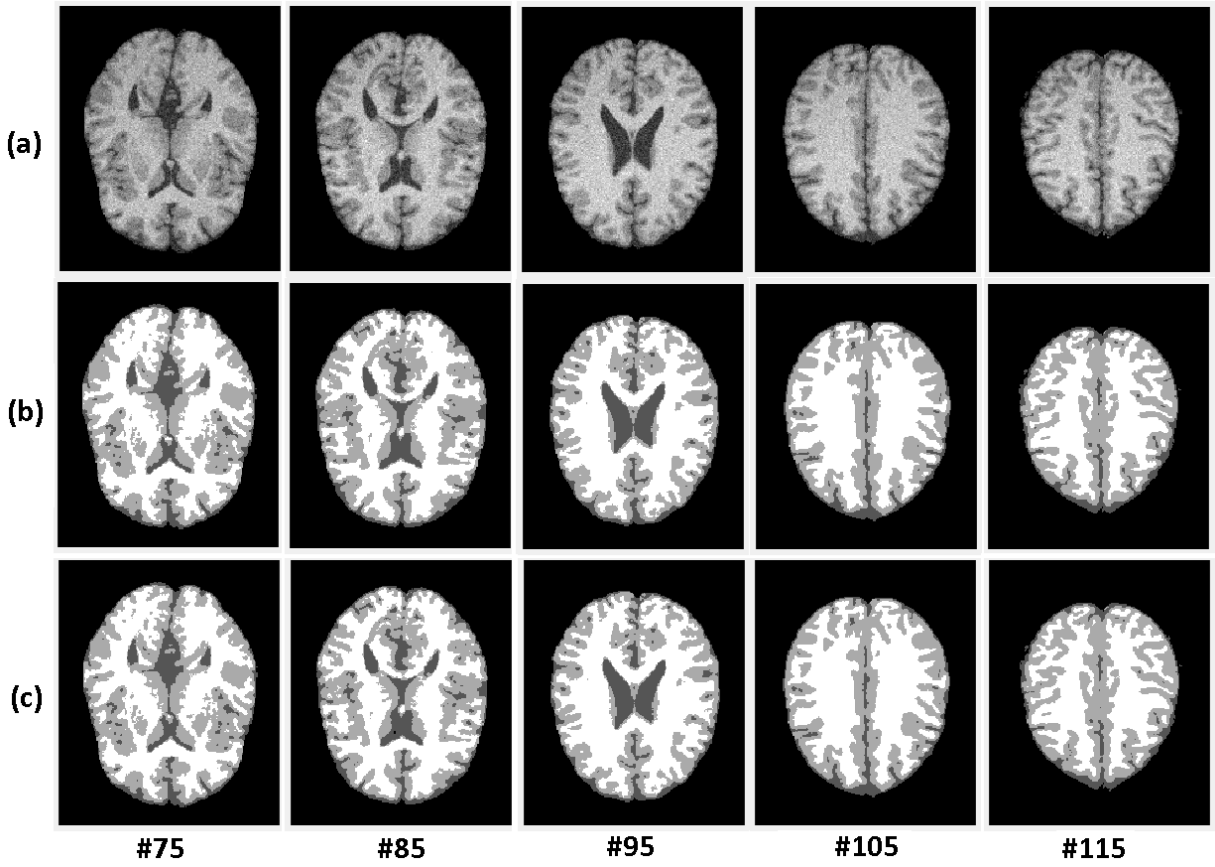

Figure III.1: Qualitative segmentation results obtained from the proposed strategy and the original one using T1-weighted brain MR images: (a) input images; (b) results using the proposed strategy; (c) results using the roulette wheel selection strategy.

Table III.1: The average values in terms of DICE coefficient obtained when doing simulated MR image segmentation by using different strategies for selecting the *gBest*.

Strategies	**Slices**					**Average**
	75	85	95	105	115	
The original	0.9425	0.9410	0.9583	0.9684	0.9688	0.9558
Proposed	0.9459	0.9479	0.9610	0.9725	0.9714	**0.9597**

*The value in bold indicates the best performance.

III.2.3.3 Global best selection mechanism

The selection of *gBest* is adapted from the leader selection strategy of whole process proposed by Gong et al. [Gong et al., 2017], which makes MOPSO algorithm avoid trapping in a local optimum

by considering the distribution of the Pareto set. $gBest$ is selected among the leftmost, rightmost, and middle solutions in the current Pareto front with certain selection probability (higher for the middle one, and smaller for the others). Specifically, the middle solution is selected among all current Pareto solutions with the minimum $\mathfrak{q}_i$.

$$\mathfrak{q}_i = \sqrt{\sum_{j=1}^{m} \overline{\mathfrak{q}}_{ij}^2} \tag{III.5}$$

where m is the number of objective functions. $\overline{\mathfrak{q}}_{ij}$ is the fitness value of the j^{th} objective function of the i^{th} solution after normalization. Obviously, the middle particle is a good trade-off solution of all objective functions.

Note that, for the problem with two fitness functions, the middle solution, actually, is the closest one to the ideal solution which simultaneously has minimum fitness value.

To examine the impact and efficiency of this mechanism, we did a comparison with the original one (applying roulette wheel selection to choose one hypercube among those having more than one particle, then, randomly selecting a particle in the hypercube and considering it as the $gBest$) in the MOPSO algorithm proposed by Coello et al [Sierra and Coello, 2005] in terms of DICE coefficient. The two mechanisms were tested on simulated MR images (slices 75, 85, 95, 105, 115 with 9% noise and 40% INU artifact downloaded from the brainWeb). From Table III.1 and Figure III.1, it can be seen that the proposed strategy generally provides better segmentation results, which are about 0.39% in terms of DICE coefficient in average.

III.2.4 Decision making

When a set of trade-off solutions is obtained, a decision point needs to be chosen to proceed further. Consequently, we are facing a multi-criteria decision making (MCDM) problem. Fortunately, there are several methods which have been proposed in the literature for this non-trivial task [Padhye and Deb, 2011, Wang and Rangaiah, 2017]. These methods can be divided into two main classes: requiring prior information from users and without requiring any information from users.

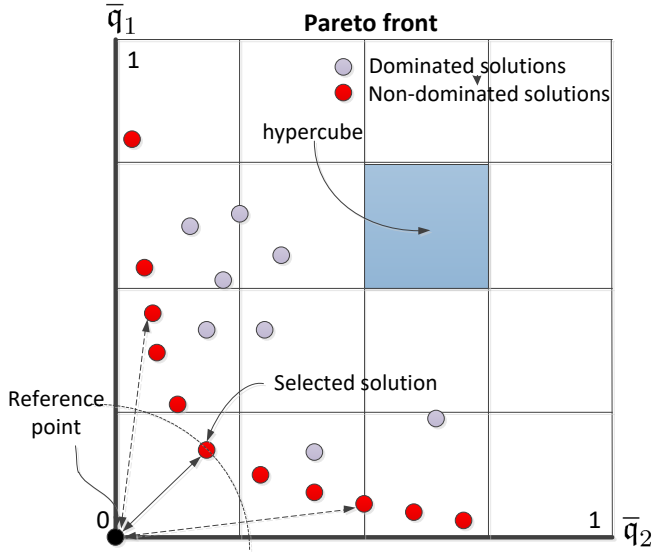

Figure III.2: $\mathcal{L}_2$-metric method for selecting final solution in a Pareto front.

In this work, we use a distance-based technique, called $\mathcal{L}_2$-metric method, proposed by Padhye et al. [Padhye and Deb, 2011]. This is a straightforward method to select one solution out of many non-dominated solutions without requiring any information from users. Specifically, first, each objective is normalized within $[0, 1]$. Then, the ideal point (solution) is constructed, which is origin in case of normalized space, and taken as the reference point (Figure III.2). Euclidean distance ($\mathcal{L}_2$) of each point in the non-dominated set is calculated from the reference point and the one which has the smallest Euclidean distance will be chosen.

III.3 Proposed method

This section describes the proposed image segmentation method, which is based on the improved MOPSO algorithm with two criteria and the $\mathcal{L}_2$-metric method used to choose the best solution on the Pareto front. The first criterion is that of the improved fuzzy entropy clustering ($\mathcal{J}_{\text{KFECSB}}$), which has been illustrated in Section II.3.3 in the previous chapter. Minimizing $\mathcal{J}_{\text{KFECSB}}$ leads to a more compact segmentation result. The second criterion takes advantages of the region-based active contours such that a higher separation between regions in segmented images can be achieved. Both criteria are optimized simultaneously using the improved MOPSO algorithm. The result is a set of distributed segmentation solutions in which one cannot be estimated as better compared to the others. To choose the final solution, the $\mathcal{L}_2$-metric method is used as a decision-making process based on the sorting of the solutions in the Pareto set. This method is presented in detail below.

III.3.1 Particle representation

In this work, the particles are made up of real numbers which represent the coordinates of the cluster/region centers in the range of pixel values in the image. For P solutions with C distinct centers, there are in total $(P \times C)$ optimization variables that need to be encoded. For instance, the position of i^{th} solution in the population is encoded as: $\mathbf{X_i} = (x_{i1}, x_{i2}, \ldots, x_{iC})$. Here, x_{ir} $(r = 1, \cdots, C)$ represents the r^{th} cluster centers among C centers of the i^{th} solution. In this way, the estimated cluster centers, $\mathbf{C} = (c_r)_C$, are obtained by decoding $\mathbf{X_i}$.

III.3.2 Segmentation criteria

The performance of a segmentation algorithm critically depends upon the criteria it tries to optimize simultaneously. We propose to jointly optimize two independent criteria to obtain both the complementary properties of segmented images: compactness and separation. In addition, the local spatial information and bias correction are also considered to deal with noise and INU artifact existing in input images.

III.3.2.1 Criterion based on fuzzy entropy clustering

The first objective function is the kernelized fuzzy entropy clustering with local spatial information and bias correction, $\mathcal{J}_{\text{KFECSB}}$. Note that, brain MR images may be corrupted by different artifacts. Therefore, a pixel at the boundary of two or more regions can easily induce classification errors under the grayness ambiguity or geometric blur that reduces the sharpness of the objects contours in the image. By using the fuzzy concept to consider each pixel belonging to several regions with different degrees, we can manage and overcome these difficulties more efficiently. This criterion

is calculated using the following equations (for more details, refer to Section II.3.3):

$$K(f_j, c_i) = \exp\left(-\frac{\|f_j - c_i\|^2}{\sigma^2}\right) \tag{III.6}$$

$$u_{ij}^{-1} = \sum_{r=1}^{C} \frac{\exp\left\{n\left[\left(1 - K(f_j, c_i b_j^{fe})\right) + \eta\left(1 - K(\overline{f}_j, c_i b_j^{fe})\right)\right]\right\}}{\exp\left\{n\left[\left(1 - K(f_j, c_r b_j^{fe})\right) + \eta\left(1 - K(\overline{f}_j, c_r b_j^{fe})\right)\right]\right\}} \tag{III.7}$$

$$\mathbf{B^{fe}} = \left[\left(\sum_{i=1}^{C} \mathbf{G} \times \mathbf{F} \sum_{j=1}^{N} c_i u_{ij}\right)^{-1} \left(\mathbf{G} \times \mathbf{G}^T \sum_{j=1}^{N} c_i^2 u_{ij}\right)\right]^T \mathbf{G} \tag{III.8}$$

$$\begin{aligned} \mathcal{J}_{\text{KFECSB}}\left(\mathbf{F}, \mathbf{C}, \mathbf{U}, \mathbf{B^{fe}}\right) = & \underbrace{\sum_{i=1}^{C}\sum_{j=1}^{N} u_{ij}\left[1 - K(f_j, c_i b_j^{fe})\right]}_{\text{shape and size}} + \underbrace{\frac{1}{n}\sum_{i=1}^{C}\sum_{j=1}^{N} u_{ij}\log\left(u_{ij}\right)}_{\text{degree of non-membership}} \\ & + \underbrace{\eta\sum_{i=1}^{C}\sum_{j=1}^{N} u_{ij}\left[1 - K(\overline{f}_j, c_i b_j^{fe})\right]}_{\text{artifact in-sensitiveness and detail preservation}} \end{aligned} \tag{III.9}$$

The objective function, $\mathcal{J}_{\text{KFECSB}}$, described in Eq. (III.9) should satisfy the following conditions:

$$\mathbf{U} \in \left\{ u_{ij} \in [0,1] \middle| \sum\nolimits_{i=1}^{C} u_{ij} = 1, \quad \forall j \quad \text{and} \quad 0 < \sum\nolimits_{j=1}^{N} u_{ij}, \quad \forall i \right\} \tag{III.10}$$

III.3.2.2 A novel criterion based on region-based active contour

The second objective function is the adaptive energy weight combined with global and local fitting energy active contour models, $\mathcal{J}_{\text{AWGLAC}}$. As there are more than two regions in a single slice in MR brain images, hence, we focus on the multiphase approach (3 or 4 regions) in this work.

Chan and Vese proposed a multiphase model (MC-V) [Vese and Chan, 2002], called *global fitting energy* (GFE), which is used to extract 2^n objects with n level set functions. For $C = 4$ phases, let $\mathbf{\Phi} = (\phi_1, \phi_2)$ be a vector level set function, and $\mathbf{\Theta} = (\Theta_1, \Theta_2)$ be a set of closed curves expressed by the zero level set functions ϕ_i. $\{\Theta_1 : \phi_1 = 0\} \cup \{\Theta_2 : \phi_2 = 0\}$ partitions an image into four regions $\mathbf{\Omega} = (\Omega_1, \Omega_2, \Omega_3, \Omega_4)$, where the image domain $\Omega = \cup_{i=1}^{C} \Omega_i$ and $\Omega_i \cap \Omega_j = \oslash$ for $i \neq j$, with their mean values $\mathbf{C} = (c_i)_C$. In the level set framework [Sethian, 1999], for the given image $\mathbf{F}(x)$ with point $x \in \Omega$, the energy functional of $\mathbf{\Phi}$ can be formulated as:

$$\mathcal{J}^{\text{GFE}}\left(\mathbf{F}, \mathbf{\Phi}, \mathbf{C}\right) = \sum_{i=1}^{C} \lambda_i \int_{\Omega} |\mathbf{F}(x) - c_i|^2 M_i(\mathbf{\Phi}(x)) dx + \nu \mathcal{L}(\mathbf{\Phi}) \tag{III.11}$$

where λ_i is a positive constant which defines the weight of a term in the energy functional. ν is a positive constant fixed empirically. $M_i(\mathbf{\Phi})$ is the membership function representing the region Ω_i: $M_1(\mathbf{\Phi}) = H_\epsilon(\phi_1).H_\epsilon(\phi_2), M_2(\mathbf{\Phi}) = H_\epsilon(\phi_1).(1 - H_\epsilon(\phi_2)), M_3(\mathbf{\Phi}) = (1 - H_\epsilon(\phi_1)).H_\epsilon(\phi_2)$, and $M_4(\mathbf{\Phi}) = (1 - H_\epsilon(\phi_1)).(1 - H_\epsilon(\phi_2))$, where $H_\epsilon(s) = \frac{1}{2}(1 + \frac{2}{\pi} arctan(\frac{s}{\epsilon}))$ with a positive constant ϵ, is a

smooth version of Heaviside function. $\mathcal{L}(\mathbf{\Phi})$ is a regularization term that computes the arc length of the zero level set contours of $\mathbf{\Phi}$ and serves to smooth them during evolution, defined as follows:

$$\mathcal{L}(\mathbf{\Phi}) = \int_{\Omega} |\nabla H_{\epsilon}(\phi_1(x))| dx + \int_{\Omega} |\nabla H_{\epsilon}(\phi_2(x))| dx \tag{III.12}$$

By minimizing Eq. (III.11) with respect to $\mathbf{\Phi}$ and $\mathbf{C}$, the segmented image is obtained as follows:

$$\mathbf{F_s} = \sum_{i=1}^{C} c_i M_i(\mathbf{\Phi}) \tag{III.13}$$

Li et al. [Li et al., 2018] pointed out that the energy weight parameters (λ_i) in the energy functional (III.11) are important and impact on the segmentation accuracy, especially when the areas of the objects and background are significantly different. They argue that each object should have the same contribution to the energy functional (III.11) and it would not be dependent on the others. Hence, to cope with this problem, the authors proposed adaptive energy weight functions to configure (λ_i) as illustrated below:

$$\lambda_i^{(k)} = \frac{\text{size}(\mathbf{F})}{\text{Area}(\Omega_i^{(k)})} = \frac{N}{\text{Area}(\Omega_i^{(k)})}, \quad \text{with} \quad \text{Area}(\Omega_i^{(k)}) = \int_{\Omega} M_i(\mathbf{\Phi}^{(k)}(x)) dx \tag{III.14}$$

where $\lambda_i^{(k)}$, $\text{Area}(\Omega_i^{(k)})$, and $\mathbf{\Phi}^{(k)}(x)$ are the energy weight, the area, and level set functions in the k^{th} iteration of the i^{th} region, respectively.

It is clear that the energy functional (III.11) does not take into account local partial information and thus is unable to deal with INU artifact. Inspired from the work of Li et al. [Li et al., 2011b], Huang and Zeng [Huang and Zeng, 2015] proposed the model which is better in approximating the measured image and simultaneously solves the problem of intensity inhomogeneity by incorporating the local difference information between the acquired image and Li et al.'s estimate. The model based on local partial information of the image (called *local fitting energy* (LFE)) is given by:

$$\begin{aligned} \mathcal{J}^{\text{LFE}}(\mathbf{F}, \mathbf{\Phi}, \mathbf{C}, \mathbf{B^{ac}}, \mathbf{D}) &= \nu \mathcal{L}(\mathbf{\Phi}) + \mu \mathcal{P}(\mathbf{\Phi}) \\ &+ \int_{\Omega} \left(\sum_{i=1}^{C} \beta_i \int_{\Omega} K_{\sigma}(y-x) |\mathbf{F}(x) - b^{ac}(y) c_i - d(y)|^2 M_i(\mathbf{\Phi}(x)) dx \right) dy \end{aligned} \tag{III.15}$$

where $\mathbf{C} = (c_i)_C$, $\mathbf{B^{ac}} = \left(b_j^{ac}\right)_N$, and $\mathbf{D} = (d_j)_N$ are the set of C mean intensity values, the bias field estimation and the local difference, respectively. β_i is a positive constant which defines the weight of different terms in the local energy functional. $K_{\sigma}(y-x)$ is a Gaussian kernel function (viewed as the weight at point y in local region of x) with a window of size $(4\text{w}+1) \times (4\text{w}+1)$ and standard deviation σ, which is larger than w. μ is a positive constant fixed empirically. $\mathcal{P}(\mathbf{\Phi})$ is a regularization term proposed by Li et al. [Li et al., 2005] to eliminate the re-initialization step, which is defined as follows:

$$\mathcal{P}(\mathbf{\Phi}) = \frac{1}{2} \int_{\Omega} (|\nabla \phi_1(x)| - 1)^2 dx + \frac{1}{2} \int_{\Omega} (|\nabla \phi_2(x)| - 1)^2 dx \tag{III.16}$$

Minimizing Eq. (III.15) with respect to $\mathbf{\Phi}$, $\mathbf{C}$, $\mathbf{B^{ac}}$, and $\mathbf{D}$, the segmented image is obtained by using Eq. (III.13)

By using local partial information, the work of Huang and Zeng is able to provide desirable segmentation results even in the presence of INU artifact. However, to some extent, the model is sensitive to initialization and hard to handle with the presence of noise. To solve this issue, several works [Liu et al., 2013, Zhao et al., 2017] have been proposed by using the approach that consists in combining the global and local fitting energy functions. While the global information is used to improve the robustness with respect to noise and initialization (sometimes, it also maintains the boundary details [Wang et al., 2010]), the local information is considered to deal with INU artifact occurred in the given image. However, even most of the previous works assume that the energy weights associated with each region are fixed (equal to 1), this is not true as pointed out by Li et al. [Li et al., 2018]. Moreover, tuning simultaneously several parameters in the combining models is a tedious work.

In this work, with the same approach, which uses both the global and local fitting energy functions, we propose a novel region-based active contour model with adaptive weights described as follows:

$$\mathcal{J}_{\text{AWGLAC}}(\mathbf{F},\mathbf{\Phi},\mathbf{C},\mathbf{B^{ac}},\mathbf{D}) = \underbrace{\omega.\sum_{i=1}^{C}\lambda_i\int_\Omega |\mathbf{F}(x)-c_i|^2 M_i(\mathbf{\Phi}(x))dx}_{\text{The global fitting energy (GFE)}} + \underbrace{\mathcal{R}(\mathbf{\Phi})}_{\text{regularization and arc length}}$$
$$+ \underbrace{(1-\omega).\int_\Omega\left(\sum_{i=1}^{C}\beta_i\int_\Omega K_\sigma(y-x)|\mathbf{F}(x)-b^{ac}(y)c_i-d(y)|^2 M_i(\mathbf{\Phi}(x))dx\right)dy}_{\text{The local fitting energy (LFE)}} \tag{III.17}$$

where $\omega = (\frac{1}{2\pi})^2.(1-h_{\text{mean}})$ is a constant defined by the input image, which controls the influence of the GFE force and LFE force. Here, h_{mean} is the average of the local entropy with window of size 9×9 centered on each pixel in the whole image, which reflects the degree of the intensity inhomogeneity [Zhao et al., 2017]. It is clear from the energy functional (III.17) that, when the presence of INU artifact (h_{mean}) in the given image is higher, the LFE force will be increased to deal with it. $\mathcal{R}(\mathbf{\Phi}) = \nu\mathcal{L}(\mathbf{\Phi})+\mu\mathcal{P}(\mathbf{\Phi})$ is the combination of the regularization term and arc length term. The energy weight parameters (λ_i,β_i) are adaptively adjusted as follows:

$$\lambda_i^{(k)} = \rho.\frac{\min\{\text{Area}(\Omega_i^{(k)})\}}{\text{Area}(\Omega_i^{(k)})} \tag{III.18}$$

$$\beta_i^{(k)} = \zeta.\frac{\text{Area}(\Omega_i^{(k)})}{\min\{\text{Area}(\Omega_i^{(k)})\}}, \quad \text{with } i=1,\cdots,C. \tag{III.19}$$

where (ρ,ζ) are scale parameters determined by experiments and $\text{Area}(\Omega_i^{(k)})$ is defined as in Eq. (III.14).

With the definition of the energy weight parameters as in Eqs. (III.18) and (III.19), it is obvious that each object fairly contributes in the energy functional (controlled by $\lambda_i^{(k)}$) in the whole image. On the other hand, in local area the difference among regions is enhanced (controlled by $\beta_i^{(k)}$). Hence, in this way, by minimizing Eq. (III.17), we can obtain a better result of image segmentation.

Using the gradient descent method, the solution of $(\mathbf{\Phi},\mathbf{C},\mathbf{B^{ac}},\mathbf{D})$ when we minimize Eq. (III.17) with respect to each of its variables, given the others in previous iteration, is achieved as follows.

For fixed $\mathbf{\Phi},\mathbf{B^{ac}}$, and $\mathbf{D}$, the optimal $\mathbf{C}=(c_i)_C$ that minimizes (III.17), is given by:

$$c_i = \frac{\omega.\lambda_i\int_\Omega \mathbf{F}(x)M_i(\mathbf{\Phi})dx + (1-\omega).\beta_i\int_\Omega (K_\sigma * \mathbf{B^{ac}})(\mathbf{F}(x)-\mathbf{D})M_i(\mathbf{\Phi})dx}{\omega.\lambda_i\int_\Omega M_i(\mathbf{\Phi})dx + (1-\omega).\beta_i\int_\Omega (K_\sigma * (\mathbf{B^{ac}})^{\mathbf{2}})M_i(\mathbf{\Phi})dx} \tag{III.20}$$

Keeping $\mathbf{C}, \mathbf{B^{ac}}$, and $\mathbf{D}$ fixed, and minimizing the energy functional $\mathcal{J}_{\text{AWGLAC}}$ in (III.17) with respect to $\mathbf{\Phi}$, we can deduce the associated Euler-Lagrange equations for $\mathbf{\Phi}$. By introducing an artificial time variable $t \geq 0$, the gradient flow equation for updating $\mathbf{\Phi}$ is as follows:

$$\frac{\partial \phi_l^{(k)}}{\partial t} = -\sum_{i=1}^{L} \frac{\partial M_i(\mathbf{\Phi}^{(k)})}{\partial \phi_l^{(k)}} . e_i^{(k)} + \nu . \delta_\epsilon(\phi_l^{(k)}) \text{div}\left(\frac{\nabla \phi_l^{(k)}}{|\nabla \phi_l^{(k)}|}\right) + \mu . \left(\nabla^2 \phi_l^{(k)} - \text{div}\frac{\nabla \phi_l^{(k)}}{|\nabla \phi_l^{(k)}|}\right) \tag{III.21}$$

$$\phi_l^{(k+1)} = \phi_l^{(k)} + \triangle t \frac{\partial \phi_l^{(k)}}{\partial t} \tag{III.22}$$

with the initial and Neumann boundary, conditions are defined by:

$$\phi_l(x, 0) = \phi_l^{(0)}(x), \quad x \in \Omega \tag{III.23}$$

$$\frac{\partial \phi_l^{(k)}}{\partial \mathbf{n}} = 0, \quad x \in \partial\Omega \tag{III.24}$$

where $l = 1, 2$, and $\phi_l^{(0)}(x)$ is the initial level set l^{th} function. $\mathbf{n}$ denotes the exterior normal to the boundary $\partial\Omega$. $\triangle t$ is time step. $\delta_\epsilon(\phi_l^{(k)})$ is the Dirac delta function, which is the derivative of $H_\epsilon(\phi_l^{(k)})$: $\delta_\epsilon(s) = H_\epsilon{}'(s) = \frac{1}{\pi}\frac{\epsilon}{\epsilon^2 + s^2}$. And, $e_i^{(k)}$ is a function defined as follows:

$$e_i^{(k)} = (1 - \omega) \int_\Omega \beta_i^{(k)} K_\sigma(y - x) |\mathbf{F}(x) - b^{ac}(y) x_i^{\text{c}} - d(y)|^2 dy + \omega \lambda_i^{(k)} |\mathbf{F}(x) - x_i^{\text{c}}|^2 \tag{III.25}$$

Similarly, we obtain the optimal $\mathbf{B}$ and $\mathbf{D}$ as follows:

$$\mathbf{B^{ac}} = \frac{K_\sigma * ((\mathbf{F} - \mathbf{D}) . \sum_{i=1}^{C} c_i . M_i(\mathbf{\Phi}))}{K_\sigma * \sum_{i=1}^{C} (c_i)^2 . M_i(\mathbf{\Phi})} \tag{III.26}$$

$$\mathbf{D} = \frac{K_\sigma * \sum_{i=1}^{C} (\mathbf{F} - c_i . \mathbf{B}) M_i(\mathbf{\Phi})}{K_\sigma * \sum_{i=1}^{C} M_i(\mathbf{\Phi})} \tag{III.27}$$

where ' $*$ ' is the convolution operator.

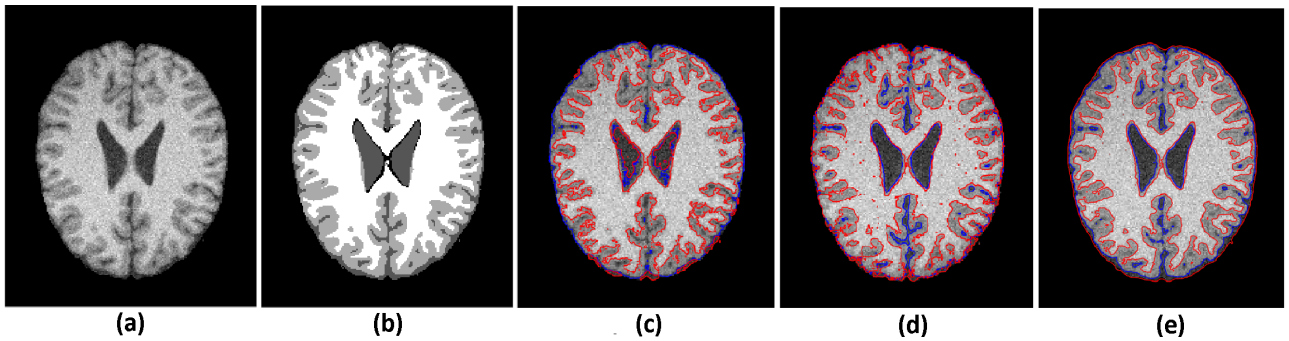

Figure III.3: Segmentation results of a simulated MR brain image by different objective functions: (a) the original; (b) ground truth; (c) result with the energy functional (III.11); (d) result with the energy functional (III.15); (e) result with the energy functional (III.17)

To verify the effectiveness of the proposed model (III.17), we segment MR images using the PSO-based framework proposed by Mandal et al. [Mandal et al., 2014], that we extended for multiphase case with the same termination conditions introduced in [Wang et al., 2010], for three fitness functions: (III.11), (III.15), and (III.17), with default settings as in the literature. As a simple example, Figure III.3 shows the segmentation results for the different fitness functions when segmenting T1-weighted image (slice 97 from the brainWeb dataset) with noise level of 9% and intensity inhomogeneity of 40%. It is clear that when using the energy functional (III.17) as a fitness function of the image segmentation problem, we can obtain superior results compared to the counterparts.

III.3.3 Optimum search process

In this section, the improved MOPSO algorithm illustrated in Section III.2.3 is used as a searching engine to find a set of Pareto solutions by minimizing the two fitness functions, $\mathcal{J}_{\text{KFECSB}}$ and $\mathcal{J}_{\text{AWGLAC}}$. The MOO framework for the image segmentation problem is summarized in **Algorithm III.3**.

In the initialization, the population $(\mathbf{X_i})_P$ is generated randomly. Each individual $\mathbf{X_i} = (x_{i1}, x_{i2}, \ldots, x_{iC})$ is a C-dimension decision vector (with $C = 3$ or 4), where C is the number of cluster centroids. And then, the initial level set function associated with the solution $\mathbf{\Phi}^{(0)}$ is configured through a succession of steps. First, the elements of the decision vector are sorted in ascending order, $(x'_{i1}, x'_{i2}, \ldots, x'_{i4})$. Next, the $\mathbf{\Phi}^{(0)}$ is determined as explained bellow

If the number of regions is three ($C = 3$), where $M_1(\mathbf{\Phi}) = H_\epsilon(\phi_1).H_\epsilon(\phi_2)$, $M_2(\mathbf{\Phi}) = H_\epsilon(\phi_1).(1 - H_\epsilon(\phi_2))$, and $M_3(\mathbf{\Phi}) = 1 - H_\epsilon(\phi_1)$, the $\mathbf{\Phi}^{(0)}$ are initialized as:

$$\phi_1^{(0)} = \alpha.\text{sign}(\mathbf{F} > x'_{i1}) \quad \text{and} \quad \phi_2^{(0)} = \alpha.\text{sign}(\mathbf{F} > x'_{i2}) \tag{III.28}$$

If the number of regions is four ($C = 4$), the $\mathbf{\Phi}^{(0)}$ are initialized as:

$$\phi_1^{(0)} = \alpha.\text{sign}(\mathbf{F} < x'_{i3}) \quad \text{and} \quad \phi_2^{(0)} = \alpha.\text{sign}((\mathbf{F} > x'_{i2}) \cup (\mathbf{F} < x'_{i4})) \tag{III.29}$$

where α is a constant, set equal to 2, and $\mathbf{F}$ is the input image.

To make sure that all particles are moving within the search space, the boundary conditions for their velocities and positions ($\mathbf{V_i}$, $\mathbf{X_i}$) are limited as follows:

$$v_{ij}^{(k)} = \begin{cases} v_{max}, & \text{if } v_{ij}^{(k)} > v_{max} \\ v_{min}, & \text{if } v_{ij}^{(k)} < v_{min} \\ +\, v_{ij}^{(k)}, & \text{otherwise} \end{cases} \tag{III.30}$$

$$x_{ij}^{(k)} = \begin{cases} x_{max}, & \text{if } x_{ij}^{(k)} > x_{max} \\ x_{min}, & \text{if } x_{ij}^{(k)} < x_{min} \\ +\, x_{ij}^{(k)}, & \text{otherwise} \end{cases} \tag{III.31}$$

where v_{min} and v_{max} are the smallest and largest allowable step sizes in any dimension ($v_{min} = -v_{max} = -1$ is set in this paper); and $\{x_{min}, x_{max}\}$ are the bounds of the search space in each dimension. Actually, they are the minimum and maximum of the intensity value of the input image.

Furthermore, to provide a flexible choice for the user and a resource for further post-processing stage, maintaining the external archive gA(t), where are stored the non-dominated solutions, includes both output information for fuzzy entropy clustering ($\mathbf{U}$, $\mathbf{B^{fe}}$) and region-based active contour ($\mathbf{\Phi}$,

$\mathbf{B^{ac}}$) methods. However, in this work we get segmented images by using the output of fuzzy entropy clustering method. In addition, the terminating condition in **Algorithm** III.3 is the maximum number of iterations. When ($k > N_{iter}$) is reached, the algorithm is immediately stopped.

Algorithm III.3: MOPSO-based image segmentation algorithm

Initialization: Read input image and determine the range of search space $[x_{min}, x_{max}]$; set the number of regions: C; set the maximum number of iterations N_{iter}; initialize population: $(\mathbf{X_i})_P$; initialize level set functions: $(\phi_1^{(0)}, \phi_2^{(0)})$; initialize flight parameters: $\{w, c_1, c_2\}$; set parameters for the two fitness functions ($\mathcal{J}_{\text{KFECSB}}$ and $\mathcal{J}_{\text{AWGLAC}}$); calculate fitness values; determine $pBest$, $gBest$ and store them in two external archives pA(t) and gA(t), respectively; generate hypercubes as coordinate systems

Results : A set of non-dominated solutions in the gA(t)

$k \leftarrow 1$

repeat

for *each particle (say* $\mathbf{X_i}$*)* **do**

`/* Flight particles */`

Update flight parameters $\{w, c_1, c_2\}$ using Eq. (III.4)

Update $\mathbf{X_i}$ and $\mathbf{V_i}$ according to Eqs. (II.6) and (II.7)

Generate new particle (say $\mathbf{X_i'}$) by using mutation operators

Check boundary conditions in search space for $\mathbf{X_i'}$ and $\mathbf{V_i'}$

`/* Evaluate solution (`$\mathbf{X_i'}$`) in terms of fitness values */`

Evaluate fitness value for $\mathcal{J}_{\text{KFECSB}}$ as following steps:

- Calculate kernel distance K by using Eq. (III.6)
- Update the partition matrix $\mathbf{U}$ by using Eq. (III.7)
- Estimate the bias correction $\mathbf{B^{fe}}$ by using Eq. (III.7)
- Evaluate the fitness value by using Eq. (III.9)

Evaluate fitness value for $\mathcal{J}_{\text{AWGLAC}}$ according to following steps:

- Evolve the level set functions $\mathbf{\Phi}$ by using Eq. (III.22)
- Update energy weights (λ_i, β_i) by using Eqs. (III.18) and (III.19)
- Estimate the bias correction $\mathbf{B^{ac}}$ by using Eq. (III.26)
- Estimate the difference $\mathbf{D}$ by using Eq. (III.27)
- Evaluate the fitness value by using Eq. (III.17)

`/* Maintain non-dominated solutions */`

Update $pBest$ and $gBest$ using the the proposed strategies descibed in Section III.2.3.2 and Section III.2.3.3

Maintain gA(t) using crowding distance technique

Update the contents in the hypercubes

$k \leftarrow k + 1$

until *the stopping criteria are met*

III.3.4 Solution selection

After getting a Pareto set from the multi-objective optimization algorithm, the $\mathcal{L}_2$-metric method will take place to select the final solution. Note that, the reference point (the ideal solution) is the intersection point of lines passing through the top right and bottom left solutions of the Pareto front. Euclidean distance ($\mathcal{L}_2$) from each point in non-dominated set to the ideal point is calculated as

follows:

$$\mathcal{L}_2 = \sqrt{w_{fe}.\text{norm}(\mathcal{J}_{\text{KFECSB}})^2 + w_{ac}.\text{norm}(\mathcal{J}_{\text{AWGLAC}})^2} \tag{III.32}$$

where w_{fe} and w_{ac} define the importance of the two objective functions. These constants are set equal to 1 in all experiments.

The whole algorithm for the image segmentation problem is illustrated in Figure III.4.

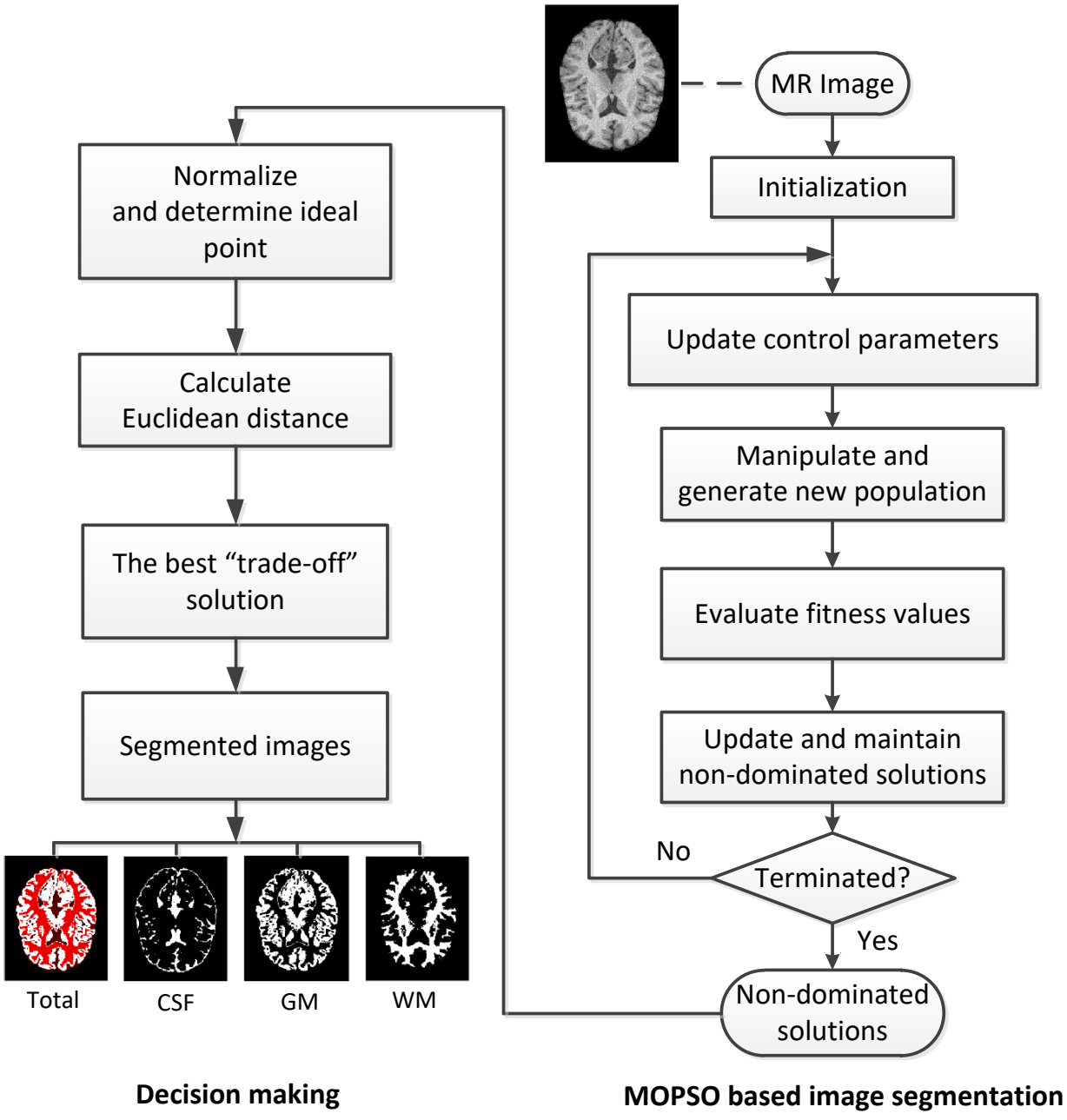

Figure III.4: Block diagram of the proposed method

III.4 Experimental results

In this section, we empirically evaluate the performance of the proposed method. Particularly, the aim is to address the possibility of the proposed algorithm to effectively improve the performance in segmentation of MR brain images. To validate the effectiveness of our framework, the results are examined both qualitatively and quantitatively, and compared with those obtained from six state-of-the-art algorithms and a PSO based algorithm with the second objective function (III.17), called PSO-AWGLAC algorithm. They are current variants of FCM, named adaptively regularized kernel-based fuzzy C-means clustering (ARKFCM) [Elazab et al., 2015] and FCM algorithm based on morphological reconstruction and membership filtering (FRFCM) [Lei et al., 2018], the multiplicative intrinsic component optimization (MICO) [Li et al., 2014] and the well-known region-based level set method, called LIC [Li et al., 2011b]. Additionally, the metaheuristic multilevel threshold with Markov random field (PSO-MRF) [T Krishnan et al., 2016] and our PSO-KFECSB algorithm [Pham et al., 2018], are used.

All algorithms are implemented in MATLAB 2014b and executed with a computer with Intel Core i7 1.8 GHz CPU, 8GB RAM using Microsoft Windows 10.

III.4.1 Experimental setup

To perform experiments, the parameters of the proposed algorithm are set as follows: population size $P = 30$, external archive of non-dominated solutions size $N_r = 30$, number of grids per dimension $N_g = 10$, maximum number of iterations $N_{iter} = 200$ (180 for real MR images), and selection probabilities for the leftmost, rightmost, and middle solutions: $15\%, 15\%$, and 70%, respectively. For the first fitness function (III.9), $\mathcal{J}_{\text{KFECSB}}$, the values of n, η, and λ are 10, 2.5, and 1/9 [Pham et al., 2018], respectively. For the second fitness function (III.17), $\mathcal{J}_{\text{AWGLAC}}$, unless otherwise specified, we set $\epsilon = 3, \sigma = 5, \alpha = 2, \triangle t = 1$, and $\mu = 1$, as common values in the literature. Note that, due to the nature of active contour models, the weighting coefficients ν, ρ, and ζ are differently set for each type of image. The effect of the arc length parameter ν was useful and briefly discussed in [Li et al., 2008, Lankton and Tannenbaum, 2008]. In addition, depending on the level of noise and INU artifact, one can increase the value of ρ (for high level of intensity inhomogeneity), or ζ (for high level of noise), and vice versa. The algorithms' parameters setting is specified in Table III.2.

III.4.2 Datasets

The MR images used in this study include both T1-weighted simulated and real 2D brain MR images. For simulated MR images, they are downloaded from a well-known database: the brainWeb from a McConnell Brain Imaging Center [Kwan et al., 1999], which can be reached in (https://brainweb.bic.mni.mcgill.ca/brainweb/). Images with size of 181×217 and thickness of 1 mm are used in this work. On the other hand, real images are taken in the 20 normal MR brain data sets provided by the Center for Morphometric Analysis at Massachusetts General Hospital, which are available at (http://www.nitrc.org/projects/ibsr/). Images with characteristics of size 135×142 and 1.171751 mm thickness, are used in our experiments.

Note that several image modalities used in medical imaging, such as Roentgen rays (X-Rays), Magnetic Resonance Imaging (MRI), Ultrasound (US), Computed Tomography (CT) and Positron Emission Tomography (PET), each with its own advantages and disadvantages, have appeared over the years. This study focused on images produced by structural imaging techniques and satisfied assumptions mentioned in Section III.2.1. As a result, images produced by the other techniques,

Table III.2: Parameters setting used for the compared algorithms

Algorithms	Parameters setting
ARKFCM	number of clusters 4 (or 3 for real MR images), window size of 3×3 pixels, maximum number of iterations $t = 100$, median filter, and $\varepsilon = 0.001$.
FRFCM	number of clusters 4 (or 3 for real MR images), window size of 3×3 pixels, square structuring element of size 3×3, maximum number of iterations $t = 100$, minimum amount of improvement $\eta = 0.001$, and exponent for the partition matrix $m = 2$.
MICO	multiphase $L = 3$, 15 polynomials of the first four orders as the basis functions, maximum number of iterations $t = 15$, exponent for the partition matrix $q = 1$, and minimum amount of improvement $\varepsilon = 0.001$.
LIC	multiphase $L = 3$, the convolution kernel is constructed as a 17×17 mask (scale parameter $\sigma = 4$), the time step $\Delta t = 0.1$, $\mu = 1$, weight of length term $\nu = 0.001 * 255^2$, $\varepsilon = 3$ for Heaviside function, number of inner iterations $N_{iter} = 10$, and number of outer iterations $N_{outer} = 100$.
PSO-MRF	number of clusters 4, number of EM iterations and MAP iterations: $N_{em} = 3$ and $N_{map} = 3$.
PSO-KFECSB	number of clusters 4 (or 3 for real MR images), the kernel parameter $\lambda = 1/9$, the population size $P = 40$, constraints of the velocities $[-1, 1]$, maximum number of iterations $N_{iter} = 100$, termination criteria parameter $\epsilon = 0.0001$, the number of non-significant improvements $N_{non-imp} = 5$, and model parameters $n = 10$, $\eta = 2.5$.
PSO-AWGLAC	number of clusters 4 (or 3 for real MR images), the population size $P = 40$, constraints of the velocities $[-1, 1]$, maximum number of iterations $N_{iter} = 100$, the number of non-significant improvements $N_{non-imp} = 5$, the time step $\Delta t = 0.1$, $\mu = 1$, weight of length term $\nu = 0.003 * 255^2$, $\varepsilon = 3$, Gaussian parameter $\sigma = 4$, and scale parameters (ρ, ζ) are $(1/1.9, 1.9)$ (or $(1/0.9, 0.9)$ for real MR images).

functional imaging [Foster et al., 2014], such as PET, may not be suitable for our method.

III.4.3 Performance measures

Since the ground truth images are available in the datasets, for comparing quantitatively the performance, five performance measures are considered, which are: the Dice Similarity Coefficient (DICE), the Hausdorff distance (HD), the Jaccard index (JAC), the Accuracy (AC) and the Sensitivity (SI) [Taha and Hanbury, 2015]. While JAC and HD are presented in Section II.4.3, the remain ones are defined below.

III.4.3.1 Dice coefficient

The Dice coefficient [Dice, 1945] (DICE) is an overlap-based metric which directly compares a segmented image ($\mathbf{F}_\mathrm{s}$) with a ground truth image ($\mathbf{F}_\mathrm{t}$) by measuring similarity between them. This metric is one of the most used measures in validating medical volume segmentations. Given an input image with N pixels $\mathbf{F} = (f_1, f_2, \ldots, f_N)$, and its two partitions, $\mathbf{F}_\mathrm{s} = (F_{\mathrm{s}1}, F_{\mathrm{s}2}, \cdots, F_{\mathrm{s}N})$ (the segmented image) and $\mathbf{F}_\mathrm{t} = (f_{\mathrm{t}1}, f_{\mathrm{t}2}, \cdots, f_{\mathrm{t}N})$ (the ground truth image), there are four common cardinalities that reflect the overlap between the two partitions, namely the true positives (TP), the false positives (FP), the true negatives (TN), and the false negatives(FN). Then, the pair-wise overlap of the repeated segmentations is calculated using the DICE, which is defined by:

$$\mathrm{DICE}\,(\mathbf{F}_\mathrm{s}, \mathbf{F}_\mathrm{t}) = \frac{2.|\mathbf{F}_\mathrm{s} \bigcap \mathbf{F}_\mathrm{t}|}{|\mathbf{F}_\mathrm{s}| + |\mathbf{F}_\mathrm{t}|} = \frac{2.\mathrm{TP}}{2.\mathrm{TP} + \mathrm{FP} + \mathrm{FN}} \tag{III.33}$$

From Eqs. (II.23) and (III.33), we can see that: DICE = 2.JAC/(1 + JAC) and JAC = DICE/(2 - DICE) for any input. That means that both metrics measure the same aspects and provide the same system ranking.

III.4.3.2 Accuracy

This criterion determines how much the segmentation algorithm results match with the ground truth. It is defined as below:

$$\mathrm{AC}\left(\mathbf{F}_{\mathrm{s}}, \mathbf{F}_{\mathrm{t}}\right) = \frac{\mathrm{TP} + \mathrm{TN}}{\mathrm{TP} + \mathrm{TN} + \mathrm{FP} + \mathrm{FN}} \tag{III.34}$$

III.4.3.3 Sensitivity

This criterion indicates true positivity and it is the probability that a pixel in segmented image belongs to a particular region. Definition of the criterion is given below:

$$\mathrm{SI}\left(\mathbf{F}_{\mathrm{s}}, \mathbf{F}_{\mathrm{t}}\right) = \frac{\mathrm{TP}}{\mathrm{TP} + \mathrm{FN}} \tag{III.35}$$

III.4.4 Results on simulated MR images

In this section, simulated MR brain images from the brainWeb are used for the purpose of performance evaluation. Each image is segmented into four regions: cerebrospinal fluid (CSF), gray matter (GM), white matter (WM), and the background. The rest of parameters is a set of the weighting coefficients ($\nu = 0.00015 * 255^2, \rho = 9$, and $\zeta = 0.1$).

To establish the performance of the proposed algorithm with respect to noise and INU artifact, experiments have been performed by considering images containing high level of noise and intensity inhomogeneity. Particularly, those are T1-weighted images with 9% noise and 40% INU artifact.

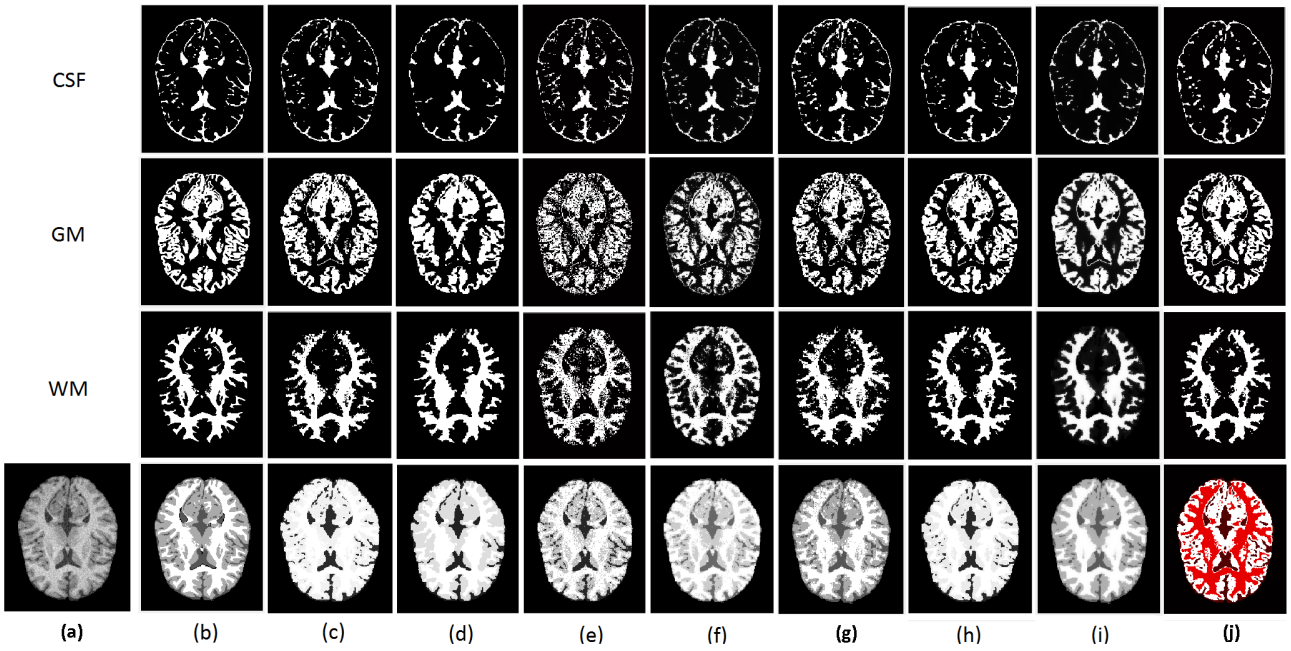

Figure III.5: Qualitative segmentation results of a simulated brain MR image provided by the competing algorithms: (a) the original; (b) ground truth images; (c) ARKFCM results; (d) FRFCM results; (e) MICO results; (f) LIC results; (g) PSO-MRF results; (h) PSO-KFECSB results; (i) PSO-AWGLAC results; (j) proposed method results.

Figure III.5 shows the qualitative results of the segmentation of a T1-weighted image (slice 80) provided by the competing algorithms. This figure reveals that though the PSO-KFECSB, FRFCM and the proposed algorithm, among the existing methods, yield positive outcomes, those from the multi-objective approach are the best ones, preserving the correct details of the image in specific. Figure III.6 shows the qualitative results of the segmentation of five T1-weighted images (slices: 75, 80, 85, 95, 105, 115). In spite of artifact existence, it can be seen that the proposed algorithm achieves a high performance when segmenting simulated MR brain images. Hence, it can be concluded that our proposed algorithm provides qualitatively satisfactory results.

Table III.3: Average values in terms of five criteria (DICE, HD, JAC, AC, SI) on simulated brain MR images using the competing algorithms.

Methods	**Regions**	**Performance Criteria**				
		DICE	**HD**	**JAC**	**AC**	**SI**
ARKFCM	CSF	0.9242	202.3333	0.8714	0.8770	0.9407
	GM	0.8842	126.5000	0.7993	0.8389	0.8722
	WM	0.9451	108.3333	0.8963	0.9251	0.9299
	Total	0.9062	080.5000	0.8373	0.9322	0.8932
FRFCM	CSF	0.9776	479.8333	0.9563	0.9607	0.9741
	GM	0.9241	124.6667	0.8589	0.8974	0.9289
	WM	0.9493	210.1667	0.9038	0.9324	0.9495
	Total	0.9415	081.0000	0.8901	0.9582	0.9402
MICO	CSF	0.9792	**184.5000**	0.9593	0.9638	**0.9826***
	GM	0.8828	180.1667	0.7904	0.8392	0.8723
	WM	0.9041	187.6667	0.8250	0.8729	0.9135
	Total	0.9187	162.5000	0.8502	0.9420	0.9227
LIC	CSF	**0.9828**	195.1667	**0.9661**	**0.9697**	0.9748
	GM	0.8875	172.6667	0.7983	0.8443	0.8658
	WM	0.9018	182.6667	0.8214	0.8728	0.9324
	Total	0.9201	151.6667	0.8529	0.9428	0.9161
PSO-MRF	CSF	0.9658	214.0000	0.9350	0.9403	0.9800
	GM	0.9163	145.8333	0.8460	0.8863	0.9092
	WM	0.9425	136.8333	0.8915	0.9231	0.9409
	Total	0.9242	122.0000	0.8601	0.9438	0.9346
PSO-KFECSB	CSF	0.9797	224.0000	0.9634	0.9635	0.9691
	GM	**0.9414**	114.0000	**0.8924**	0.9203	0.9424
	WM	0.9641	078.1667	0.9308	0.9522	0.9694
	Total	0.9543	064.6667	0.9171	0.9664	0.9544
PSO-AWGLAC	CSF	0.9545	380.1667	0.9141	0.9157	0.9161
	GM	0.9112	180.1667	0.8384	0.8762	0.9068
	WM	0.9360	204.5000	0.8819	0.9159	0.9571
	Total	0.9147	167.0000	0.8442	0.9351	0.8826
Proposed	CSF	0.9819	234.5000	0.9645	0.9678	0.9719
	GM	0.9412	**105.5000**	0.8890	**0.9208**	**0.9467**
	WM	**0.9654**	**076.0000**	**0.9331**	**0.9539**	**0.9704**
	Total	**0.9575**	**053.3333**	**0.9187**	**0.9695**	**0.9551**

*The values in bold indicate the best performance.

Table III.3 shows the average values of five metrics of eight competing algorithms: ARKFCM, FRFCM, MICO, LIC, PSO-MRF, PSO-KFECSB, PSO-AWGLAC, and the proposed algorithm. As can be seen from the table, when segmenting images with high levels of noise and INU artifact, the proposed method generally gives the best scores. Even though the scores for CSF are not the best, its

values are comparable to the best one.

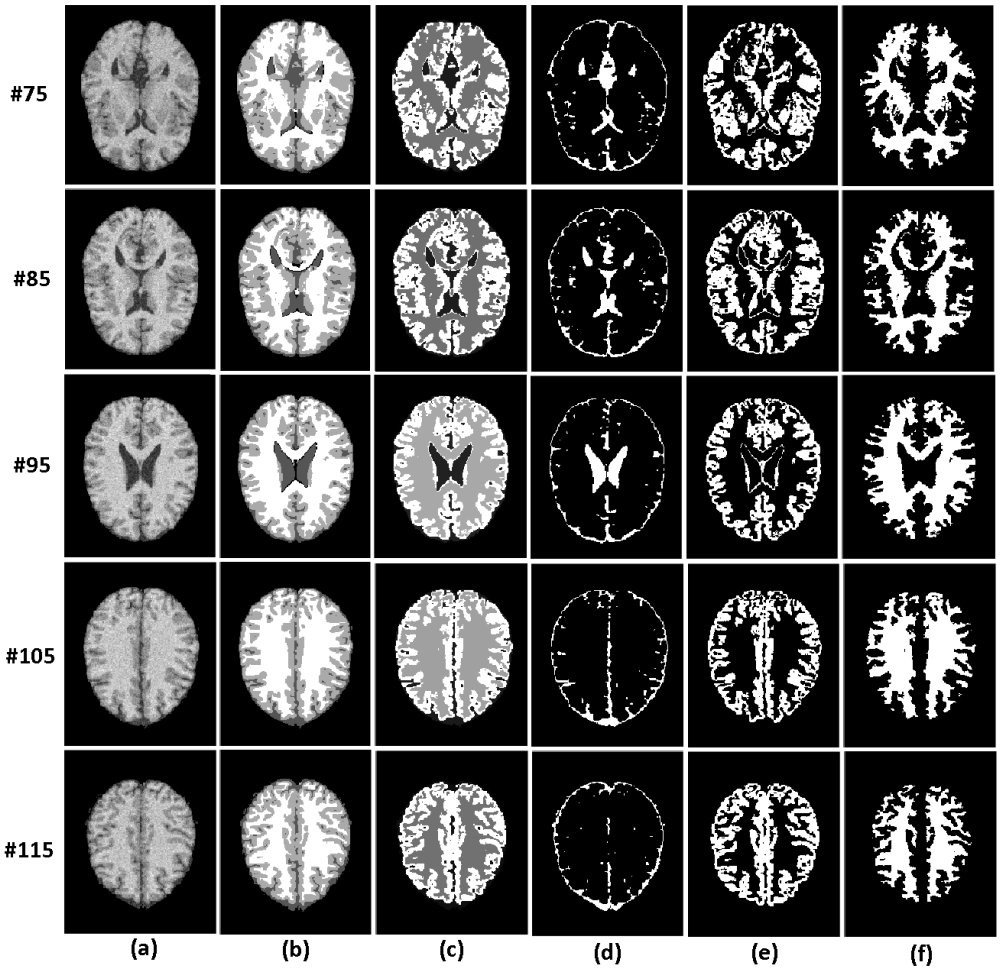

Figure III.6: Qualitative segmentation results of simulated brain MR images (slices 75, 85, 95, 105, 115) provided by the proposed algorithm versus the ground truth images: (a) the original; (b) ground truth images; (c) segmented results; (d) CSFs; (e) GMs; (f) WMs.

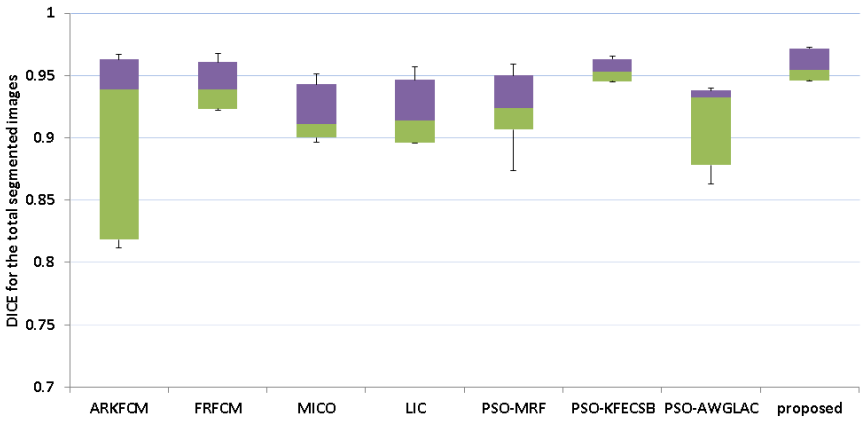

Figure III.7: Average values in terms of DICE coefficient on simulated brain MR images (slices 75, 80, 85, 95, 105, 115) using the competing algorithms.

Furthermore, Figures III.7 and III.8 show the values of Dice coefficient (DICE) and Hausdorff distance (HD) in range of the obtained values for total segmented image using eight competing algorithms, respectively. It can be seen that our method performs better than the other ones for simulated brain MRI segmentation. Thus, the proposed algorithm provides more accurate and stable segmentation results than its competitors when dealing with noise and intensity inhomogeneity problems.

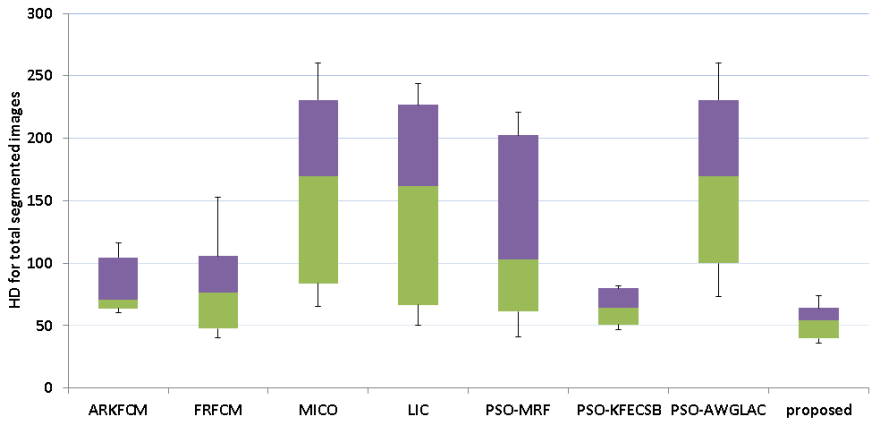

Figure III.8: Average values in terms of HD distance on simulated brain MR images (slices 75, 80, 85, 95, 105, 115) using the competing algorithms.

III.4.5 Results on real MR images

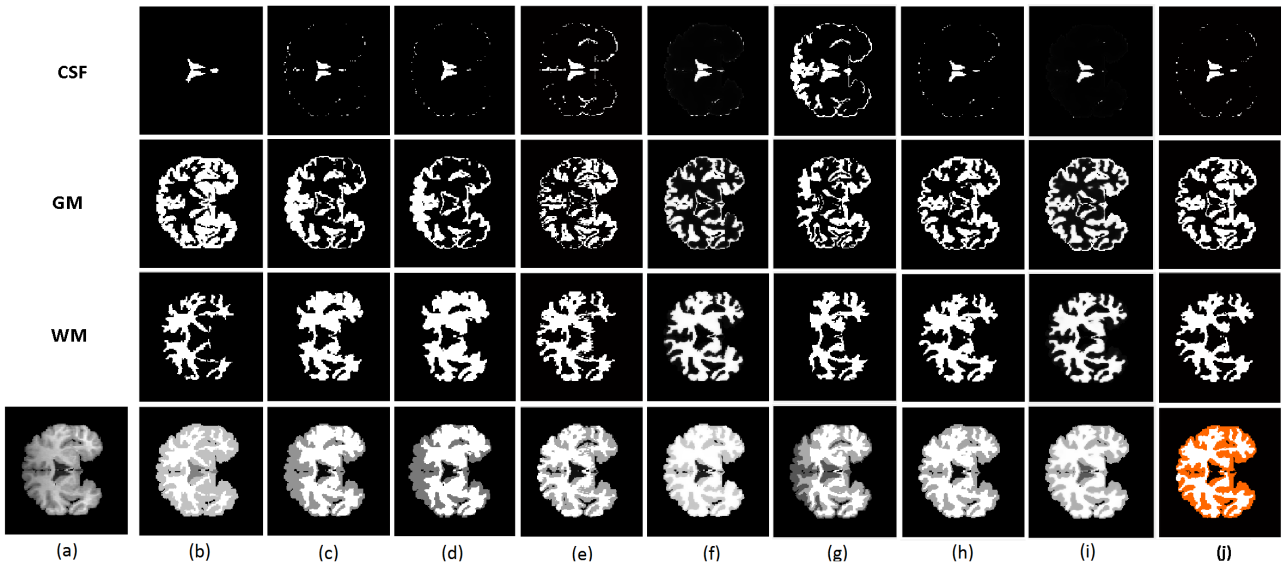

Figure III.9: Qualitative segmentation results of a real brain MR image (Z plane 34 in the 4^{th} volume) provided by the competing algorithms:(a) the original; (b) ground truth images; (c) ARKFCM results; (d) FRFCM results; (e) MICO results; (f) LIC results; (g) PSO-MRF results; (h) PSO-KFECSB results; (i) PSO-AWGLAC results; (j) proposed method results.

We have also examined the ability of the proposed method when segmenting real MR images from Internet Brain Segmentation Repository (IBSR) database (the 20 normal T1-weighted MR brain images in the 4^{th} volume). Each image is segmented into three regions: cerebrospinal fluid (CSF),

gray matter (GM) and white matter (WM). The background pixels are ignored in the computation. The rest of parameters is a set of the weighting coefficients ($\nu = 0.003 * 255^2, \rho = 0.4$, and $\zeta = 1.65$).

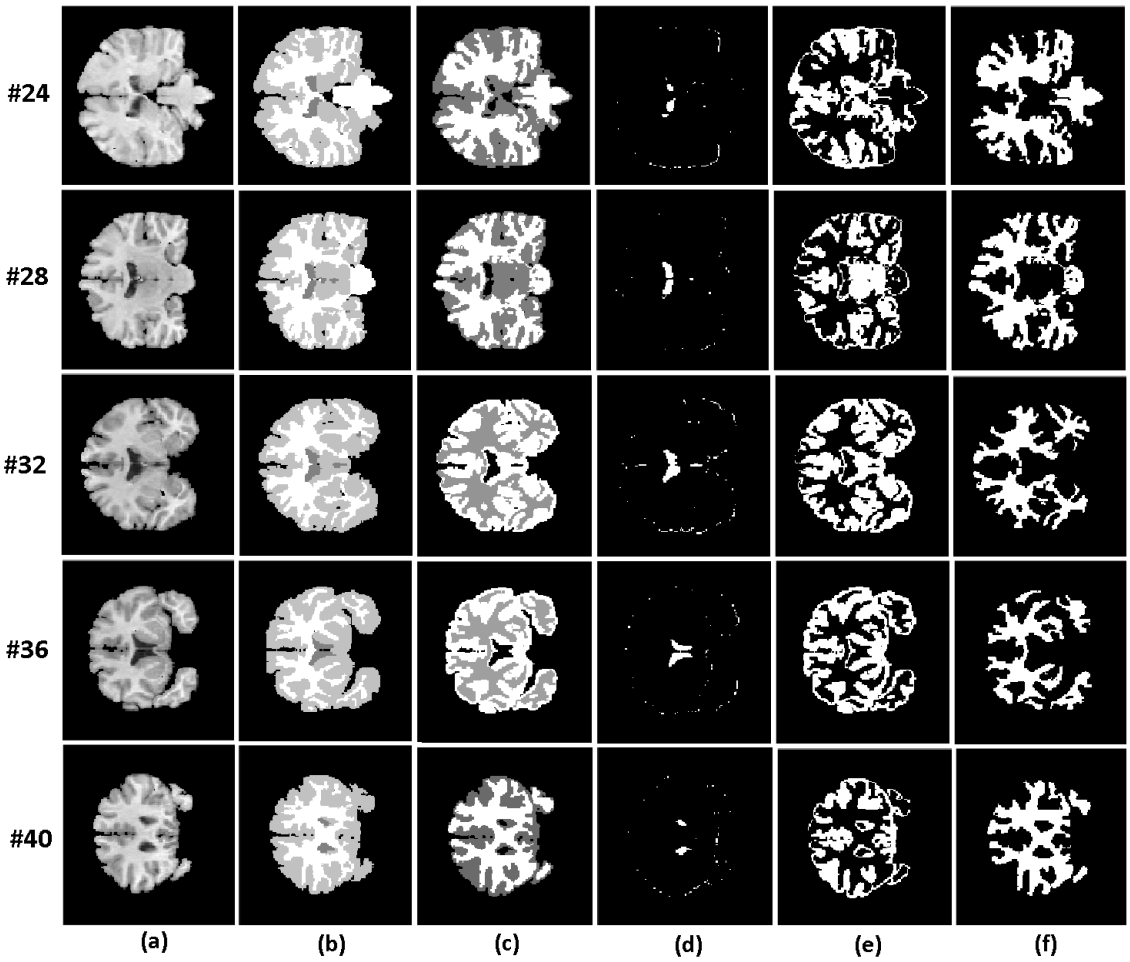

Figure III.10: Qualitative segmentation results of real brain MR images (Z planes 24, 28, 32, 36, 40 in the 4^{th} volume) by the proposed algorithm versus the ground truth images: (a) the original; (b) ground truth images; (c) segmented results; (d) CSFs; (e) GMs; (f) WMs.

Figure III.9 shows the qualitative results of the segmentation by the competing algorithms. Taking a look at the results in this figure, it is clear that our proposed algorithm provides superior results, compared to others. Even though the PSO-KFECSB, PSO-AWGLAC, MICO, and LIC algorithms can provide a reasonable WM segmentation, they have difficulty to cope with smooth transient areas in the image. Figure III.10 shows qualitative results of the segmentation of five images (Z planes 24, 28, 32, 36, and 40) provided by the proposed algorithm. This figure shows that the proposed method has identified the tissues of the real brain MR images reasonably well.

Table III.4 shows the average values of five metrics when segmenting real MR brain images (Z planes 24, 28, 32, 34, 36, and 40) by the competing algorithms. These results show that the proposed method outperforms the ARKFCM, FRFCM, MICO, LIC, PSO-MRF, PSO-KFECSB, and PSO-AWGLAC algorithms. In addition, Figure III.11 and III.12 show the values of Dice coefficient (DICE) and Hausdorff distance (HD), respectively, in range of the obtained values for total segmented image using eight competing algorithms. Again, the results presented here show the efficiency of the proposed approach and also demonstrate its superiority over its competitors.

Table III.4: Average values in terms of five criteria (DICE, HD, JAC, AC, SI) on real MR brain images using the competing algorithms.

Methods	Regions	Performance Criteria				
		DICE	HD	JAC	AC	SI
ARKFCM	CSF	0.9574	1878.5000	0.9206	0.9221	0.9938
	GM	0.8496	0166.6667	0.7425	0.7879	0.8059
	WM	0.8915	0315.1667	0.8077	0.8437	0.9122
	Total	0.9387	0059.6667	0.8888	0.9467	0.9510
FRFCM	CSF	0.9921	1877.8333	0.9843	0.9845	0.9904
	GM	0.8594	0156.6667	0.7559	0.8008	0.8094
	WM	0.8646	0175.0000	0.7649	0.8114	0.9228
	Total	0.9383	0098.3333	0.8865	0.9446	0.9378
MICO	CSF	0.9715	1977.0000	0.9447	0.9457	**0.9974***
	GM	0.8709	**0143.1667**	0.7721	0.8160	0.8163
	WM	0.9014	0129.3333	0.8217	0.8586	0.9429
	Total	0.9454	0049.1667	0.8978	0.9508	0.9549
LIC	CSF	**0.9936**	**1745.5000**	**0.9878**	**0.9884**	0.9931
	GM	0.8760	0147.3333	0.7805	0.8215	0.8159
	WM	0.8775	0143.5000	0.7837	0.8300	0.9456
	Total	0.9442	0069.3333	0.8953	0.9488	0.9403
PSO-MRF	CSF	0.9147	2011.6667	0.8431	0.8460	0.9970
	GM	0.8216	0177.6667	0.6986	0.7450	0.7669
	WM	0.9026	0423.0000	0.8254	0.8550	0.8991
	Total	0.9376	0046.8333	0.8843	0.9449	0.9586
PSO-KFECSB	CSF	0.9932	1873.1667	0.9865	0.9867	0.9927
	GM	0.8902	0146.0000	0.8027	0.8453	0.8428
	WM	0.8982	0121.1667	0.8161	0.8557	0.9519
	Total	0.9499	0050.3333	0.9054	0.9542	0.9504
PSO-AWGLAC	CSF	0.9931	2011.6667	0.9857	0.9862	0.9915
	GM	0.8853	0147.8333	0.7961	0.8411	0.8397
	WM	0.8938	0168.1667	0.8086	0.8476	0.9486
	Total	0.9485	0059.1667	0.9025	0.9526	0.9495
Proposed	CSF	0.9931	1873.1667	0.9863	0.9865	0.9931
	GM	**0.8996**	0145.1667	**0.8181**	**0.8609**	**0.8620**
	WM	**0.9103**	**0119.1667**	**0.8362**	**0.8713**	**0.9530**
	Total	**0.9540**	**0048.1667**	**0.9126**	**0.9580**	**0.9560**

*The bold numerical values indicate the best performance.

III.5 Discussion

III.5.1 Effects of skull-stripped removal

Skull stripping is crucial to study neuroimaging data. Several existing techniques, such as Brain Extraction Tool (BET), Brain Surface Extractor (BSE), ANALYZE 4.0 (The Biomedical Imaging Resource at Mayo Clinic, Rochester, MN) and modified Region Growing tool (mRG), have been proposed to remove the skull and the background region from MRI [Lee et al., 2003]. In this work, we have also developed a skull stripping algorithm (Figure III.13), which is based on brain anatomy and image intensity, to examine the segmentation ability of the proposed method. The algorithm is demonstrated as follows. First, the histogram of a given image is calculated, from which a proper threshold is selected to get a binary image. Then, **imfill** operation is used to clean the binary image. By utilizing the solidity property and using morphological operations, the skull and brain regions are

identified. Finally, the skull-stripped image is obtained by masking the original image with the brain region mask above.

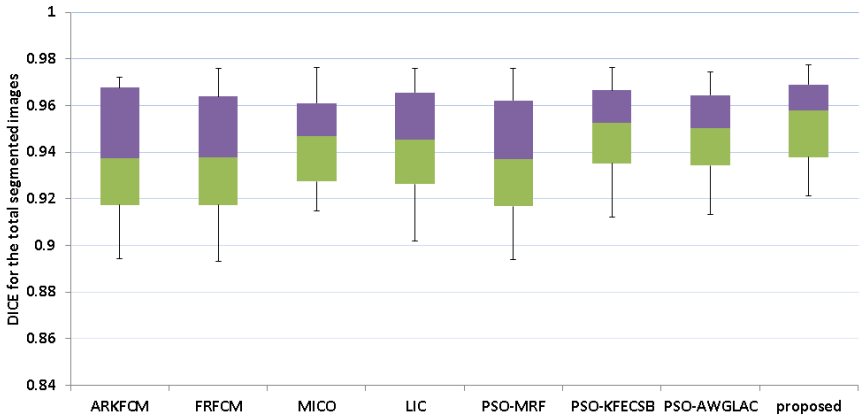

Figure III.11: Average values in terms of DICE coefficient on real brain MR images (slices 24, 28, 32, 34, 36, and 40) using the competing algorithms.

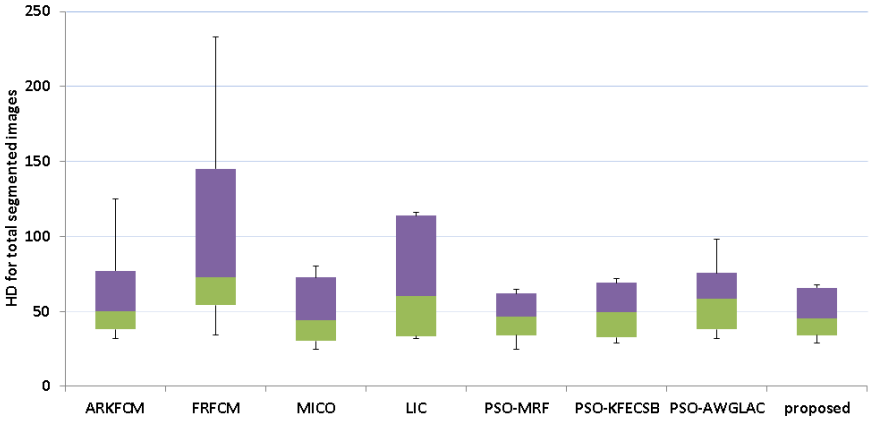

Figure III.12: Average values in terms of HD distance on real MR brain images (slices 24, 28, 32, 34, 36, and 40) using different algorithms.

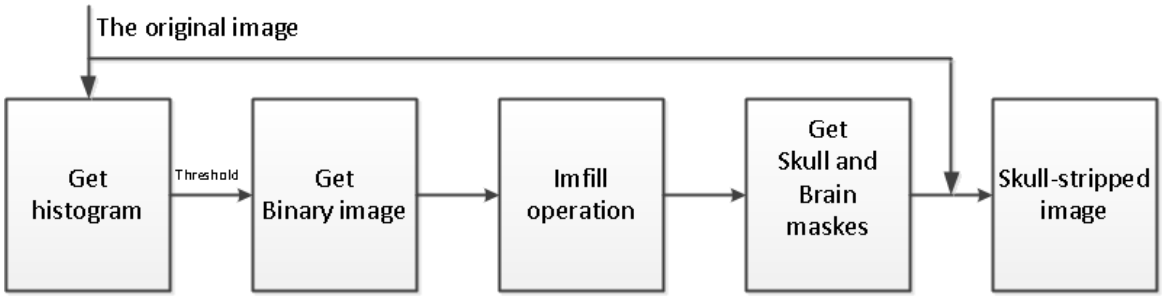

Figure III.13: Skull stripping procedure.

By using this skull stripping algorithm, we are able to obtain skull-stripped images from original images with skull. From that, the proposed method can be evaluated on both types of brain MRI image. Figure III.14 shows an example of segmentation results when applying the proposed method

to segment both types of MRI brain images (slice 80 in the brainWeb dataset with 9% noise and 40% INU artifact). It should be noted that the proposed method provides better qualitative segmentation results in case of the skull-stripped image. To clarify this issue, three different unsupervised criteria: Levine Nazif, Rosenberger Type I and Borsotti Q [Chabrier et al., 2006], were used for a quantitative performance evaluation. Table III.5 also demonstrates that better segmentation results can be achieved with input image without skull.

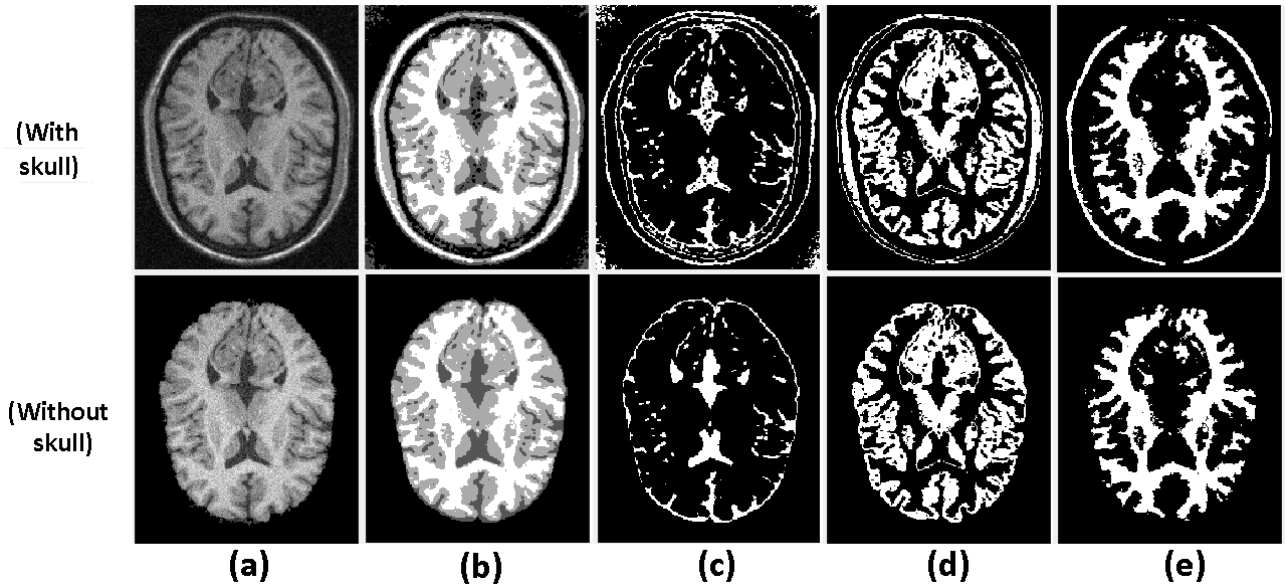

Figure III.14: Qualitative segmentation results by using the proposed method with both T1-weighted skull brain image and skull-stripped brain image: (a) input images; (b) segmentation results; (c) CSFs; (d) GMs; (e) WMs.

Table III.5: Quantitative segmentation results in terms of unsupervised criteria of the proposed method on slice 80 in simulated MR image dataset.

Types of images	Unsupervised criteria		
	Levine Nazif	Rosenberger Type I	Borsotti Q
With skull	0.5643	0.5728	0.0504
Without skull	**0.5911**	**0.6010**	**0.0448**

*The values in bold indicate the best performance.

III.5.2 Parameter settings

There are two constant parameters (ρ, ζ) which control the effect of the GFE force and LFE force in the second objective function, $\mathcal{J}_{\text{AWGLAC}}$. These scale weights directly determine the accuracy of the proposed model, hence they indirectly define the quality of the proposed method. However, as pointed out in Section III.3.2.2, when the intensity inhomogeneity in the input image is severe, such as for real MR images in the IBSR database, the correctness of the proposed region-based active contour model relies on the LFE force. In such case, we should choose a higher value for (ζ); otherwise, the GFE force may cause the deformable curves falling into fake edges. For images with minor inhomogeneity

and high level of noise, such as simulated MR images examined in Section III.4.4, the influence of the GFE force in the model should be increased. In this situation, we can use a smaller value for (ζ) and a higher value for (ρ). In the experiment, we need to select appropriate values for (ρ, ζ) according to the degree of INU artifact and noise existing in input images.

III.5.3 Computational analysis

Table III.6: Mean ± standard deviation of time cost of the proposed algorithm.

Dimension	Dataset	Number of images	Image size	Time/iteration (s)
2D	brainWeb	10	181 x 217	4.76 ± 0.15
	IBSR	10	135 x 142	2.12 ± 0.06

The proposed method benefits both advantages of fuzzy entropy clustering and region-based active contour approaches to satisfy multiple requirements of image segmentation problem. In addition, using multi-objective particle swarm optimization algorithm to avoid the dependence on initial labelling is also taken into account. To analyse the computational cost of the proposed algorithm, we calculated the running time when segmenting brain MR images. Both simulated and real brain MR image datasets are involved in experiments. Since the eventual computational cost will be the multiplication of cost for each iteration and the number of iterations for convergence, the average time cost of each iteration is recorded. The mean and standard deviation of the time cost of our algorithm (Intel Core i7 1.8 GHz CPU, 8GB RAM and MATLAB 2014b) are listed in Table III.6.

III.6 Conclusion

In this chapter, we have proposed a new method, which is based on multi-objective optimization metaheuristic approach, for the segmentation of MR brain images. In the proposed method, an improved MOPSO algorithm, which takes advantages of the states-of-the-art developments in multi-objective clustering by using MOPSO engine, is used to optimize two independent and complementary properties (compactness and separation) in segmented images. While the compactness is characterized by the first fitness function, called kernelized fuzzy entropy clustering with local spatial information and bias correction, the second property is represented by a novel region-based active contour model, called adaptive energy weight global and local fitting energy active contour model. The result of the optimization process is a set of Pareto-optimal solutions. By using the $\mathcal{L}_2$-metric method, the best trade-off solution is found and considered as the final result. In this way, the major drawbacks in fuzzy clustering and active contour methods (such as sensitivity to noise and INU artifact, possibility to be trapped into the first local minimum they encounter when using the gradient descent search method), have been partially solved. Therefore, the algorithm can produce better segmentation results. To confirm the effectiveness of the proposed method, it has been examined on both simulated and real MR images, then compared to six recent segmentation methods in the literature. The experimental results show that our method can produce better segmentation results and is more robust against high levels of noise and INU artifact contained in input images. In particular, not only the visual segmentation is more accurate than others but also the quantitative results of segmentation show the better scores compared to the counterparts. However, when segmenting images with high level of noise and severely inhomogeneous intensity such as real MR brain images in the IBSR dataset, the proposed method still shows its limitations. To improve the accuracy of segmentation results, we propose to use hybrid metaheuristic optimization approach in which another segmentation criterion will be further

optimized. In such way, not only more information in the input image can be utilised, but also other properties in segmented images can be achieved. This new method is presented in detail in the next chapter.

Note that the research reported in this chapter gave rise to our publication [Pham et al., 2019b].

Chapter IV

Hybrid metaheuristic for image segmentation using Markov random field

IV.1 Introduction

In this chapter, we present a new segmentation method combining a new Hidden Markov Random Field (HMRF) model and a novel hybrid metaheuristic method based on Cuckoo search (CS) and Particle swarm optimization (PSO) algorithms. Even though the method proposed in the previous chapter can produce promising segmentation results, its performance still shows some limitations, especially when being applied to segment complex brain MR images as in the IBSR database. To improve the quality of segmented images, hence, an additional criterion based on Markov random field (MRF) model for the image segmentation problem is optimized by using a hybrid metaheuristic. Particularly, in the new segmentation criterion, not only local spatial interactions between pixels in the input image is taken into account, but also balancing energy contributions of each segmented region is addressed. Furthermore, the hybrid metaheuristic, which takes into account the advantages of both developed CS and PSO algorithms, is designed such that fast and accurate solutions can be obtained. As a result, two major problems of the MRF approach: (1) the inaccuracy of the basic model in nature; (2) the trapping into the first encountered minimum when including spatial coherence assumptions and dependency on initially segmented image, can be solved, and thus better segmentation results can be achieved.

This chapter is organized as follows. The general theory and development used in this work for MR image segmentation using the MRF models are presented in Section IV.2. The hybrid metaheuristic based on CS and PSO to enhance the searching ability and the quality of solutions is presented in Section IV.3. The description of the proposed method, using the developed MRF model and the hybrid optimization algorithm, is given in Section IV.4. Experimental results and discussions are given in Section IV.5. Finally, we reach a conclusion in Section IV.6.

IV.2 MRF-based segmentation model

In the literature, Hidden Markov Random Field (HMRF) models, a generalization of Hidden Markov Model (HMM), are widely used as a probabilistic robust-to-noise approach to model the joint probability distribution of the image pixels in terms of local spatial interactions [Geman and Geman, 1987]. Many methods based on HMRF have been proposed for brain MR segmentation [Marroquin et al., 2002, Guerrout et al., 2016, Ahmadvand et al., 2017]. The main idea underlying the segmentation process in this approach is that the image to segment (the observed image) and the segmented image (the hidden image) are considered as MRF. The segmented image is computed sequentially according to the MAP (Maximum A Posteriori) criterion, it leads to the minimization of an energy function [Zhang et al., 2001]. There are two main advantages when using the MRF

models for image segmentation [Huawu Deng and Clausi, 2005]: (1) The spatial relationship can be seamlessly integrated into a segmentation procedure through contextual constraints of neighbouring pixels; (2) Different types of image features can be utilized in the MRF-based segmentation model via the Bayesian framework. However, the basic model is inaccurate in nature. As a result, unexpected artifacts appear in segmentation results. This section describes a novel MRF model to deal with this problem.

IV.2.1 Image segmentation problem in Bayesian framework

Let $\Omega = \{s = (i, j) | 1 \leqslant i \leqslant H, 1 \leqslant j \leqslant W\}$ be the set of image lattice sites, where H and W are the image height and width in pixels. In the two-dimensional image lattice Ω, the pixel values $\mathbf{I} = \{\mathrm{I}_s | s \in \Omega\}$ are a realization of random variables $\mathbf{I^r} = \{\mathrm{I^r}_s | s \in \Omega\}$.

The segmentation problem can be expressed in the Bayesian framework. Let us denote the observed image extracted from a random image $\mathbf{I^r} = \mathbf{I}$ by $\mathbf{F^r} = \mathbf{F}$, where $\mathbf{F^r}$ and $\mathbf{F}$ are a random variable and its instance, respectively. $\mathbf{Y^\Omega} = \mathbf{Y}$ stands for a segmented image based on the vector $\mathbf{F^r} = \mathbf{F}$, with the set of all possible configurations on Ω, $\Omega_\mathbf{Y}$. For the gray image, $\mathbf{F}$ takes its values in the space $\mathcal{E}_\mathbf{F} = \{0, \ldots, 255\}$ and $\mathbf{Y}$ takes its values in the discrete space $\mathcal{E}_\mathbf{Y} = \{1, \ldots, C\}$, where C is the number of classes or homogeneous regions in the image.

According to the Bayes rule, the segmentation problem is formulated as follows:

$$P\left(\mathbf{Y^\Omega} = \mathbf{Y} \mid \mathbf{F^r} = \mathbf{F}\right) = \frac{P\left(\mathbf{F^r} = \mathbf{F} \mid \mathbf{Y^\Omega} = \mathbf{Y}\right) P\left(\mathbf{Y^\Omega} = \mathbf{Y}\right)}{P\left(\mathbf{F^r} = \mathbf{F}\right)} \tag{IV.1}$$

where $P\left(\mathbf{Y^\Omega} = \mathbf{Y} \mid \mathbf{F^r} = \mathbf{F}\right)$ is the *posteriori* probability of $\mathbf{Y^\Omega} = \mathbf{Y}$ conditioned on $\mathbf{F^r} = \mathbf{F}$. $P\left(\mathbf{F^r} = \mathbf{F} \mid \mathbf{Y^\Omega} = \mathbf{Y}\right)$ denotes the probability distribution of $\mathbf{F^r} = \mathbf{F}$ conditioned on $\mathbf{Y^\Omega} = \mathbf{Y}$, which is referred to as the feature modelling component. $P\left(\mathbf{Y^\Omega} = \mathbf{Y}\right)$ is a *priori* probability of $\mathbf{Y^\Omega} = \mathbf{Y}$ that describes the label distribution of a segmented result. In the literature, $P\left(\mathbf{Y^\Omega} = \mathbf{Y}\right)$ is referred to as the region labelling component. $P\left(\mathbf{F^r} = \mathbf{F}\right)$ is the probability distribution of $\mathbf{F^r} = \mathbf{F}$.

Segmentation of an image can be considered as seeking the best realization $\mathbf{Y}^*$ by maximizing the probability $P\left(\mathbf{Y^\Omega} = \mathbf{Y} \mid \mathbf{F^r} = \mathbf{F}\right)$, that is the *Maximum a posteriori* (MAP) estimate. As $\mathbf{F^r} = \mathbf{F}$ is known, $P\left(\mathbf{F^r} = \mathbf{F}\right)$ does not vary with respect to any solution $\mathbf{Y^\Omega} = \mathbf{Y}$ and hence can be disregarded.

$$\begin{aligned} \mathbf{Y}^* &= \underset{\mathbf{Y} \in \Omega_\mathbf{Y}}{\operatorname{argmax}} \left\{ P\left(\mathbf{Y^\Omega} = \mathbf{Y} \mid \mathbf{F^r} = \mathbf{F}\right) \right\} \\ &= \underset{\mathbf{Y} \in \Omega_\mathbf{Y}}{\operatorname{argmax}} \left\{ P\left(\mathbf{F^r} = \mathbf{F} \mid \mathbf{Y^\Omega} = \mathbf{Y}\right) P\left(\mathbf{Y^\Omega} = \mathbf{Y}\right) \right\} \end{aligned} \tag{IV.2}$$

We assume that each component of $\mathbf{F^r} = \mathbf{F}$ be independent on the other components with respect to $\mathbf{Y^\Omega} = \mathbf{Y}$ (conditional independence). Then, with M components in the feature vector $\mathbf{F} = \{f^m \mid m = 1, \ldots, M\}$, $P\left(\mathbf{F^r} = \mathbf{F} \mid \mathbf{Y^\Omega} = \mathbf{Y}\right)$ is transformed into:

$$P\left(\mathbf{F^r} = \mathbf{F} \mid \mathbf{Y^\Omega} = \mathbf{Y}\right) = \prod_{m=1}^{M} \left[P\left(f^m \mid \mathbf{Y^\Omega} = \mathbf{Y}\right)\right] \tag{IV.3}$$

where $P\left(f^m \mid \mathbf{Y^\Omega} = \mathbf{Y}\right)$ is the probability distribution of the extracted data component f^m conditioned on the segmented image $\mathbf{Y^\Omega} = \mathbf{Y}$.

From Eqs. (IV.2) and (IV.3), the image segmentation problem becomes:

$$\mathbf{Y}^* = \underset{\mathbf{Y} \in \Omega_\mathbf{Y}}{\operatorname{argmax}} \left\{ \prod_{m=1}^{M} \left[P\left(f^m \mid \mathbf{Y^\Omega} = \mathbf{Y}\right)\right] P\left(\mathbf{Y^\Omega} = \mathbf{Y}\right) \right\} \tag{IV.4}$$

Now, depending on how to model the feature modelling component and the region labelling component, we can get exact formulas for the image segmentation problem as described in the following sections.

IV.2.2 Intensity distribution model

The form of $P\left(f^m \mid \mathbf{Y}^{\mathbf{\Omega}} = \mathbf{Y}\right)$ may be different depending on the features used. We assume that the distribution of all feature data is a Gaussian function with different means μ_i^m and standard deviations σ_i^m. That is:

$$\begin{aligned} P\left(f_s^m \mid \mathrm{Y}_\mathrm{s}^{\Omega} = i\right) &= \frac{1}{\sqrt{2\pi(\sigma_i^m)^2}} \exp\left[-\frac{(f_s^m - \mu_i^m)^2}{2(\sigma_i^m)^2}\right] \\ &= \exp\left[-\frac{(f_s^m - \mu_i^m)^2}{2(\sigma_i^m)^2} - \log(\sqrt{2\pi(\sigma_i^m)^2})\right] \end{aligned} \tag{IV.5}$$

where μ_i^m and σ_i^m are the mean and standard deviation for the i^{th} class in the m^{th} feature component.

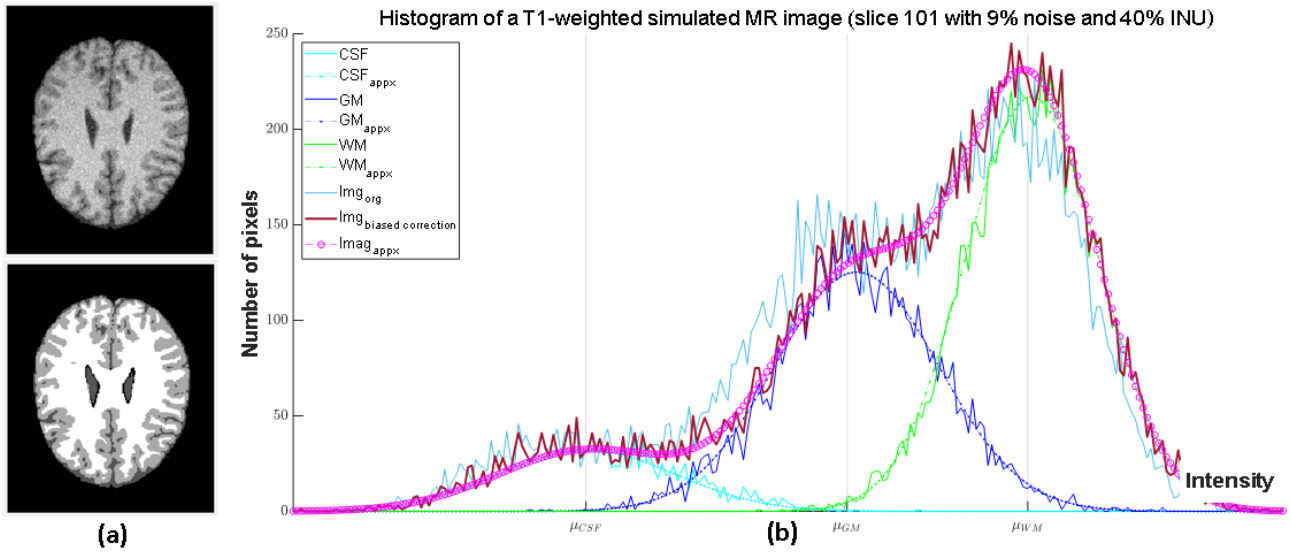

Figure IV.1: An example of fitting Gaussian distribution to a simulated brain MR image: (a) slice 101 (up) and the truth (down); (b) histograms of tissues based on the truth.

The feature modelling component, which is the product of all $P\left(f_s^m \mid \mathrm{Y}_\mathrm{s}^{\Omega} = i\right)$, has the form:

$$P\left(\mathbf{F}^{\mathbf{r}} = \mathbf{F} \mid \mathbf{Y}^{\mathbf{\Omega}} = \mathbf{Y}\right) = \exp\left\{-\sum_{s,\mathrm{Y}_\mathrm{s}^{\Omega}=i}\sum_{m=1}^{M}\left[\frac{(f_s^m - \mu_i^m)^2}{2(\sigma_i^m)^2} + \log(\sqrt{2\pi(\sigma_i^m)^2})\right]\right\} \tag{IV.6}$$

In this work, for the task of partitioning brain MR images, the intensity feature is used as the only image feature to extract different tissues ($M = 1$). Eq. (IV.6) can be rewritten as follows:

$$P\left(\mathbf{F}^{\mathbf{r}} = \mathbf{F} \mid \mathbf{Y}^{\mathbf{\Omega}} = \mathbf{Y}\right) = \exp\left\{-\sum_{i=1}^{C}\sum_{s\in\Omega_i}\left[\frac{(f_s - \mu_i)^2}{2\sigma_i^2} + \log(\sqrt{2\pi\sigma_i^2})\right]\right\} \tag{IV.7}$$

where μ_i and σ_i are the mean and standard deviation in terms of intensity for the i^{th} region, Ω_i, of the image. They are defined as follows:

$$\begin{aligned} \mu_i &= \frac{1}{|\Omega_i|} \sum_{s \in \Omega_i} f_s \\ \sigma_i &= \sqrt{\frac{1}{|\Omega_i|} \sum_{s \in \Omega_i} (f_s - \mu_i)^2} \\ \Omega_i &= \left\{ s | \mathrm{Y}_s^{\Omega} = i \right\} \end{aligned} \tag{IV.8}$$

As an example, Figure IV.1 shows how the Gaussian distribution fits in with the brain MR image. It is clear that after doing bias correction, the Gaussian distribution fits really well, hence it can be used to approximate different classes effectively.

IV.2.3 Spatial distribution model

The other term in Eq. (IV.4) is $P\left(\mathbf{Y}^{\mathbf{\Omega}} = \mathbf{Y}\right)$, which describes the prior knowledge about the spatial distribution of brain tissues in the image. For a given tissue class, one can consider the prior probability unchanged over the image. Other authors suppose that the prior probability is varying at a given location, depending on the tissues found at the neighbouring locations. This can be done by using a MRF to model spatial interactions among tissue classes [Zhang, 1992].

Definition 4.1 A random field $\mathbf{Y}^{\mathbf{\Omega}}$ is a MRF with respect to the neighbourhood system $N = \{\mathcal{N}_s, s \in \Omega\}$, if and only if:

$$P\left(\mathbf{Y}^{\mathbf{\Omega}} = \mathbf{Y}\right) > 0, \ \forall \mathbf{Y} \in \Omega_{\mathbf{Y}} \tag{IV.9}$$

and

$$P\left(\mathbf{Y}^{\mathbf{\Omega}} = \mathrm{Y}_s \mid \mathbf{Y}^{\mathbf{\Omega}} = \mathrm{Y}_t, t \neq s\right) = P\left(\mathbf{Y}^{\mathbf{\Omega}} = \mathrm{Y}_s \mid \mathbf{Y}^{\mathbf{\Omega}} = \mathrm{Y}_t, t \in \mathcal{N}_s\right) \tag{IV.10}$$

According to the Hammersley-Clifford theorem [Besag, 1974], a MRF can be equivalently characterized by a Gibbs distribution:

$$P\left(\mathbf{Y}^{\mathbf{\Omega}} = \mathbf{Y}\right) = Z^{-1} \exp\left\{-U(\mathbf{Y}, \beta)\right\} \tag{IV.11}$$

where Z is a normalizing constant and $U(\mathbf{Y}, \beta)$ is the energy function. β is a positive constant that controls the size of partitioning or interaction between the sites.

Note that, the normalization factor of the Gibbs distribution is theoretically well-defined as follows:

$$Z(U) = \sum_{\mathbf{Y}} \exp\left\{-U(\mathbf{Y}, \beta)\right\} \tag{IV.12}$$

where the sum runs over all possible configurations of $\mathbf{Y}$, which is usually unknown [Geman and Geman, 1987]. Hence, Z is said to be computationally intractable.

On the other hand, the choice of the energy function $U(\mathbf{Y}, \beta)$ is arbitrary and there are several definitions of it in the framework of image segmentation [Yang et al., 2014, Ahmadvand et al., 2017]. Using the multi-level logistic (MLL) model, a general expression of the energy function for pairwise interactions can be formulated by:

$$U(\mathbf{Y}, \beta) = \sum_{\forall s \in \Omega} \left[V_1\left(\mathrm{Y}_s\right) + \beta \sum_{t \in \mathcal{N}_s, t \neq s} V_2\left(\mathrm{Y}_s, \mathrm{Y}_t\right) \right] \tag{IV.13}$$

where $V_1(\mathrm{Y}_s)$ is an external field that weights the relative importance of the different classes present in the image. This is usually unknown, hence, $V_1(\mathrm{Y}_s) = 0$ is used. $V_2(\mathrm{Y}_s, \mathrm{Y}_t)$ models the interactions between neighbours. Figure IV.2 illustrates the second-order neighbourhood system with all possible pairwise cliques.

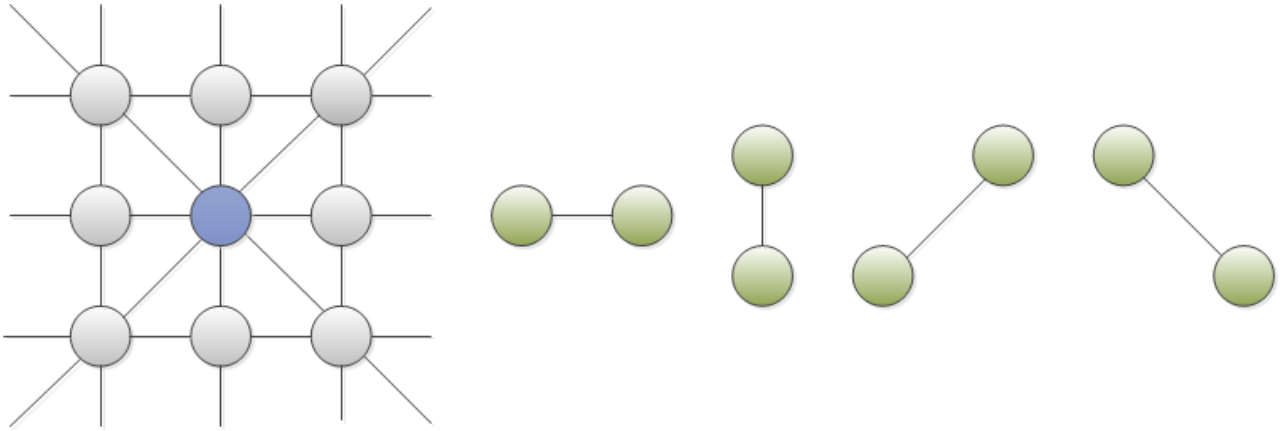

Figure IV.2: The second-order neighbourhood system with all two-points cliques.

Here, we define $V_2(\mathrm{Y}_s, \mathrm{Y}_t)$ as follows:

$$V_2(\mathrm{Y}_s, \mathrm{Y}_t) = \begin{cases} -(\exp(-|\mathrm{I}_s - \mathrm{I}_t|) + \frac{1}{2}) \times \frac{1}{d(s,t)}, & \text{if} \quad \mathrm{Y}_s = \mathrm{Y}_t \\ +(\frac{2}{2-\exp(-|\mathrm{I}_s - \mathrm{I}_t|)} - \frac{1}{2}) \times \frac{1}{d(s,t)}, & \text{otherwise} \end{cases} \tag{IV.14}$$

where $|\mathrm{I}_s - \mathrm{I}_t|$ and $d(s,t)$ are the absolute difference of intensities and the Euclidean distance between the center pixel and one of its eight neighbours ($\mathcal{N}_s = 8$), respectively.

From Eqs. (IV.11), (IV.13) and (IV.14), we can see that a pixel is classified like the majority of its neighbours. The more neighbours that have the same label as the center pixel, the more the $P\left(\mathbf{Y}^{\mathbf{\Omega}} = \mathbf{Y}\right)$ is increased. In addition, the proposed clique potential $V_2(\mathrm{Y}_s, \mathrm{Y}_t)$ involves the intensity information of the image, which makes the prior segmentation probability $P\left(\mathbf{Y}^{\mathbf{\Omega}} = \mathbf{Y}\right)$ adaptive to the local intensity. Particularly, if two pixels have the same label, the probability will be increased when the difference of intensity is small. In contrast, if two pixels have different labels, the probability will be increased when the difference of intensity is large.

IV.2.4 A novel MRF-based image segmentation model

From Eqs. (IV.1), (IV.7), (IV.11) and (IV.13), we obtain:

$$P\left(\mathbf{Y}^{\mathbf{\Omega}} = \mathbf{Y} \mid \mathbf{F}^{\mathbf{r}} = \mathbf{F}\right) = Z' \exp\left(-\Psi\left(\mathbf{Y}, \mathbf{F}\right)\right) \tag{IV.15}$$

where Z' is a constant and $\Psi\left(\mathbf{Y}, \mathbf{F}\right)$ is defined as follows:

$$\Psi\left(\mathbf{Y}, \mathbf{F}\right) = \sum_{i=1}^{C} \sum_{s \in \Omega_i} \left[\frac{(f_s - \mu_i)^2}{2\sigma_i^2} + \log(\sqrt{2\pi\sigma_i^2}) + \beta \sum_{t \in \mathcal{N}_s, t \neq s} V_2(\mathrm{Y}_s, \mathrm{Y}_t) \right] \tag{IV.16}$$

Thus, maximizing the probability $P\left(\mathbf{Y}^{\mathbf{\Omega}} = \mathbf{Y} \mid \mathbf{F}^{\mathbf{r}} = \mathbf{F}\right)$ is equivalent to minimizing the function $\Psi\left(\mathbf{Y}, \mathbf{F}\right)$. For that, we seek the segmentation result $\mathbf{Y}^*$ as follows:

$$\mathbf{Y}^* = \underset{\mathbf{Y} \in \Omega_{\mathbf{Y}}}{\operatorname{argmin}} \left\{\Psi\left(\mathbf{Y}, \mathbf{F}\right)\right\} \tag{IV.17}$$

The energy function $\Psi(\mathbf{Y}, \mathbf{F})$ is then derived:

$$\begin{aligned}\Psi(\mathbf{Y}, \mathbf{F}) &= \sum_{i=1}^{C} \lambda(i) \sum_{s \in \Omega_i} \left[\frac{(f_s - \mu_i)^2}{2\sigma_i^2} + \log(\sqrt{2\pi\sigma_i^2}) + \beta \sum_{t \in \mathcal{N}_s, t \neq s} V_2(\mathrm{Y}_s, \mathrm{Y}_t) \right] \\ &= \sum_{i=1}^{C} \lambda(i) \mathcal{J}_{\mathrm{MRF}}(i)\end{aligned} \tag{IV.18}$$

where $\mathcal{J}_{\mathrm{MRF}}(i)$ is the energy of the i^{th} region. $\lambda(i)$ is a weighting parameter corresponding to the i^{th} region, which determines how much the region contributes to the entire energy.

In general, the energy weight parameter $\lambda(i)$ in the energy function $\Psi(\mathbf{Y}, \mathbf{F})$ is set to be 1.0 for all classes. However, the energy functional is affected by the energy weight parameters, which could reduce the segmentation accuracy. It is clear that $\mathcal{J}_{\mathrm{MRF}}(i)$ is related to the number of pixels in associated region Ω_i. Hence, it could be said that the energy term $\mathcal{J}_{\mathrm{MRF}}(i)$ is positively correlated with the associated area of Ω_i. As a result, the convergence of the energy functional would be controlled by the larger regions (GM and WM) and the smaller regions (CSF) would be covered.

To overcome this problem, adaptive weight functions are proposed to configure $\lambda(i)$ and adjust the contribution of each region. The weight functions are set as follows:

$$\lambda(i) = 0.5 \frac{\min\{\mathrm{Area}(\Omega_i^{(k)})\}}{\mathrm{Area}(\Omega_i^{(k)})} \tag{IV.19}$$

where $\mathrm{Area}(\Omega_i^{(k)})$ is the number of pixels in the k^{th} iteration of the i^{th} region.

Furthermore, we can always compute $\mathbf{Y}$ through $\boldsymbol{\mu} = (\mu_1, \ldots, \mu_C)$ by classifying f_s into the nearest mean μ_j i.e. $\mathrm{Y}_s = j$ if the nearest mean to f_s is μ_j. Thus, with the same approach as in [Guerrout et al., 2018], we seek for $\boldsymbol{\mu}^*$ instead of $\mathbf{Y}^*$. Then, the image segmentation problem becomes:

$$\boldsymbol{\mu}^* = \underset{\boldsymbol{\mu} \in \Omega_\mu}{\operatorname{argmin}} \{\Psi(\boldsymbol{\mu})\} \tag{IV.20}$$

where $\Psi(\boldsymbol{\mu})$ is defined:

$$\Psi(\boldsymbol{\mu}) = \sum_{i=1}^{C} \lambda(i) \sum_{s \in \Omega_i} \left[\underbrace{\frac{(f_s - \mu_i)^2}{2\sigma_i^2} + \log(\sqrt{2\pi\sigma_i^2})}_{\text{feature modelling}} + \beta \underbrace{\sum_{t \in \mathcal{N}_s, t \neq s} V_2(\mathrm{Y}_s, \mathrm{Y}_t)}_{\text{region modelling}} \right] \tag{IV.21}$$

where $V_2(\mathrm{Y}_s, \mathrm{Y}_t)$ and $\lambda(i)$ are defined in Eqs. (IV.14) and (IV.19), respectively.

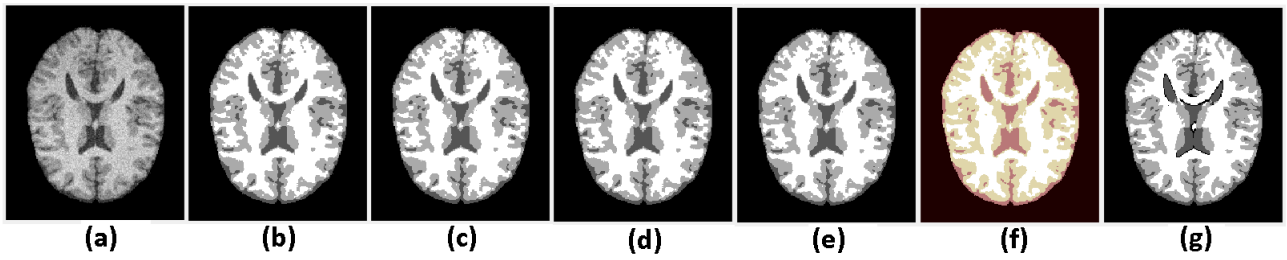

Figure IV.3: Segmentation results of a simulated brain MR image associated with different configurations of $\lambda(i)$: (a) original image; (b) result with Opt 1; (c) result with Opt 2; (d) result with Opt 3; (e) result with Opt 4; (f) result with the proposed strategy; (g) ground truth.

Table IV.1: DICE coefficient obtained from segmenting a simulated MR image (slice 87, with 9% noise and 40% INU, from the BrainWeb) by using different strategies for selecting weight parameters.

Metric	Regions	Configurations of $\{\lambda(1),\ldots,\lambda(C)\}, C=4$				
		Opt 1	**Opt 2**	**Opt 3**	**Opt 4**	**Proposed**
		{1,0.5,0.5,0.5}	{1,1,0.5,0.5}	{1,1,1,0.5}	{1,1,1,1}	adaptive
DICE	CSF	0.9827	0.9838	0.9831	0.9845	**0.9853**
	GM	0.9371	0.9361	0.9381	0.9372	**0.9418**
	WM	0.9601	0.9591	0.9589	0.9591	**0.9654**

*The values in bold indicate the best performance.

To examine the impact and efficiency of the adaptive weighting parameters on MRF-based segmentation model, we have conducted some experiments on simulated MR images with different configurations of $\{\lambda(1),\ldots,\lambda(C)\}$. Here, the simulated MR images, which are taken from the brain-Web, have characteristics of 9% noise and 40% intensity non-uniformity. The images are segmented into four regions ($C=4$): background, CSF, GM and WM corresponding to ($i=1,\ldots,C$), respectively. In addition, to analyse the performance quantitatively, DICE coefficient is used as an evaluating criterion. As an example, Table IV.1 and Figure IV.3 show quantitative and qualitative results when segmenting a simulated brain MR image (slice 87, with 9% noise and 40% INU) with different configurations of weighting parameters (other settings are the same for all experiments). It can be seen that the proposed strategy provides better segmentation results than the competing ones.

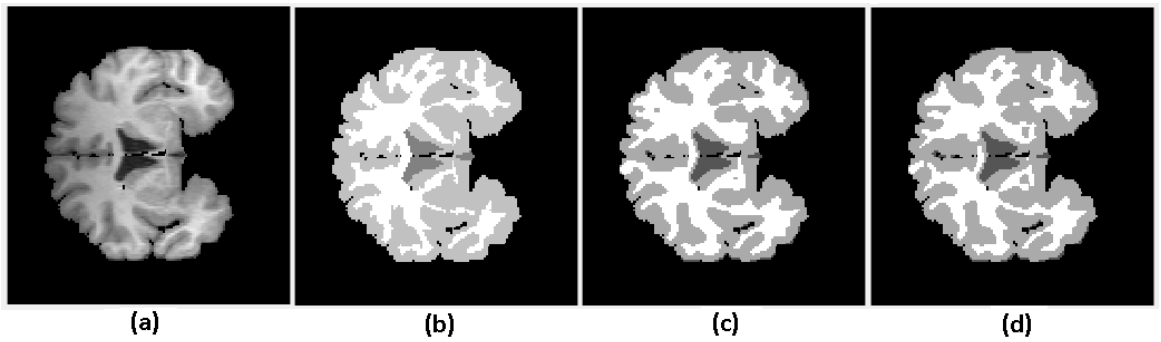

Figure IV.4: Segmentation results of a real brain MR image by using different clique potential functions: (a) original image; (b) ground truth; (c) result with Potts model; (d) result with the proposed model.

Table IV.2: DICE coefficient obtained from segmenting a real MR image by using different clique potential models.

Metric	Models of clique potential function	Regions		
		CSF	GM	WM
DICE	Potts model	0.9920	0.9269	0.9441
	Proposed	**0.9924**	**0.9335**	**0.9532**

*The values in bold indicate the best performance.

Furthermore, the influence of the proposed clique potential model (IV.14) is also examined. To do this, we have conducted some experiments on real MR images using two models: Potts model [Guerrout et al., 2018] and the proposed one. As an example, Table IV.2 and Figure IV.4 show quantitative and qualitative results when segmenting a real brain MR image (slice 34 in the 20-normal brain MR data sets, from the IBSR repository) for the two models (other settings are the same for all experiments). It can be seen that the proposed approach leads to a higher quality of segmentation results.

IV.3 Hybrid metaheuristic

To solve the image segmentation problem by using optimization techniques with the objective function given in Eq. (IV.21), several approaches can be used such as direct search methods [Guerrout et al., 2016] or metaheuristic algorithms [Yousefi et al., 2012]. Here, we exploit the second approach to cope with the second problem mentioned in Section IV.1 by introducing a new hybrid metaheuristic which benefits advantages of both the two well-known algorithms, named CS and PSO. The proposed algorithm is described below.

IV.3.1 An improved Cuckoo search algorithm

The Cuckoo Search (CS) algorithm, which is a population-based stochastic global search algorithm, is based on two behaviours of birds: the breeding behaviour of some cuckoo species and the Levy flight behaviour of some bird species [Yang and Deb, 2009]. Using CS has two main advantages: (1) The number of parameters which have to be configured in the initial search is very small; (2) The inexperienced user can easily interact with it. However, it is common to find that CS shows relatively slow convergence speed and low search accuracy because of the loss of diversity in the population [Abed-alguni and Alkhateeb, 2017].

There are three rules that idealize behaviour of cuckoos in order to become appropriate for implementation as an optimization tool: (1) Each cuckoo lays one egg at a time and dumps it in a randomly chosen nest; (2) The best nests with high-quality of eggs will be carried further to the next generation; (3) The number of available host nests is fixed, and the egg laid by a cuckoo is discovered by the host bird with a probability of $p_a \in [0, 1]$. In this case, the host bird can either get rid of the egg or simply abandon the nest and build a completely new nest.

Basically, in the CS algorithm, each egg in a nest corresponds to a potential solution and each Cuckoo's egg corresponds to a new solution. The CS algorithm attempts to iteratively improve the candidate solutions (eggs in the nests) by replacing them with better generated solutions (Cuckoo's eggs) based on the fitness values. The algorithm consists of two different phases: a global Levy flight random walk (exploration) and a local random walk (exploitation), which are controlled by the switching parameter p_a, from which the population is updated during the whole process.

In the first phase, the global Levy flight random walk is used to generate new solutions around the best nest, $gBest$, in the current generation. Assuming that, the CS algorithm processes with a population of P eggs, $\mathbf{X} = \{\mathbf{X}_1, \ldots, \mathbf{X}_P\}$, where each egg is composed of C decision variables, $\mathbf{X}_i = \{x_{i1}, \ldots, x_{iC}\}$. The step with the Levy flight is calculated based on Mantegna's algorithm [Mantegna, 1994], as below:

$$stepsize_i^{(k)} = 0.05 \times \frac{u_k}{|v_k|^{\frac{1}{\alpha}}} \times \left(\mathbf{X}_i^{(k)} - gBest\right) \tag{IV.22}$$

where $\mathbf{X}_i^{(k)}$ is the i^{th} egg of the population in the k^{th} iteration. In our work, α is set to 1.5. u and v are

Algorithm IV.1: The ICS algorithm

Initialization: Randomly initialize a population $\mathbf{X}$ of P host nests $\mathbf{X}_i$; determine $gBest$ (the global best nest)
Results : The optimal solution $gBest$
$k \leftarrow 1$
repeat
/* Exploration phase */
for *each nest* **do**
Generate a cuckoo (say $\mathbf{X}_i$) randomly by Levy flights, Eq. (IV.24)
Evaluate the fitness value, $\mathfrak{f}_\mathfrak{i}$
Randomly choose a nest among P (say $\mathbf{X}_j$)
if $\mathfrak{f}_\mathfrak{i}$ *is better than* $\mathfrak{f}_\mathfrak{j}$ **then**
Replace nest $\mathbf{X}_j$ by the new one, $\mathbf{X}_i$
else
Create a cuckoo (say $\mathbf{X}'_i$) by using Eq. (IV.25)
Replace nest $\mathbf{X}_j$ if an improvement is found
/* Exploitation phase */
Abandon a fraction p_a of worse nests and build new ones, Eq. (IV.27)
Update the $gBest$ nest and pass to the next generation
$k \leftarrow k + 1$
until *the stopping criteria are met*

normally distributed stochastic variables, $u \sim N\left(0, \sigma_u^2\right)$ and $v \sim N\left(0, \sigma_v^2\right)$. The standard deviation of the random matrix generated is defined as:

$$\sigma_u\left(\alpha\right) = \left[\frac{\Gamma\left(1+\alpha\right)\sin\left(\frac{\pi\alpha}{2}\right)}{\Gamma\left(\frac{(1+\alpha)}{2}\right)\alpha 2^{\frac{(\alpha-1)}{2}}}\right]^{\frac{1}{\alpha}} \quad \text{and} \quad \sigma_v = 1 \tag{IV.23}$$

where Γ corresponds to the standard gamma function.

Then, the new egg, $\mathbf{X}_i^{(k+1)}$, can be obtained by the following equation:

$$\mathbf{X}_i^{(k+1)} = \mathbf{X}_i^{(k)} + stepsize_i^{(k)} \times randn[C] \tag{IV.24}$$

where $randn[C]$ represents random scalars drawn from the standard normal distribution.

In order to improve the global exploration abilities within a relatively small number of generations by discovering more untouched areas in the search space (enhance the diversity), we propose to add a further step of generating new eggs if the better ones cannot be found. Particularly, if a generated egg is not better than the old one in terms of fitness value, a Differential Evolution (DE) scheme proposed by Mohamed et al. [Mohamed et al., 2019], is used to lay another egg, $\mathbf{X}_i^{\prime(k+1)}$. The scheme is expressed as follows:

$$\mathbf{X}_i^{\prime(k+1)} = \mathbf{X}_r^{(k)} + F_m \times \left(\mathbf{X}_{best}^{(k)} - \mathbf{X}_{worst}^{(k)}\right) \tag{IV.25}$$

where $\mathbf{X}_{best}^{(k)}$ and $\mathbf{X}_{worst}^{(k)}$ are the best and the worst solutions in terms of fitness value in the current population. $\mathbf{X}_r^{(k)}$ is a solution which has a rank of its fitness value in range of $[\text{round}(\varepsilon P), P -$

round(εP)] with $\varepsilon \in (0, 0.5)$. In our work, ε is set to 0.15 and F_m is generated uniformly between (0.1,1).

Using this scheme, there are two main benefits, which can be observed. First, the target vector is attracted towards good solutions, since it will always follow the same direction of the better ones. Second, the direction of the worst solution is avoided, which will force the direction of the search process towards the promising regions in the search space.

In the second phase, CS continues to generate new eggs in terms of biased/selective random walk. By considering the probability of cuckoos of being discovered, a crossover operator is used to construct a new egg as follows:

$$\mathbf{X}_i^{\prime\prime(k+1)} = \begin{cases} \mathbf{X}_i^{(k)} + F_c \times \left(\mathbf{X}_{r1}^{(k)} - \mathbf{X}_{r2}^{(k)}\right), & \text{if} \quad \text{rand}[0,1] > p_a \\ \mathbf{X}_i^{(k)}, & \text{otherwise} \end{cases} \tag{IV.26}$$

where $r1$ and $r2$ are mutually different random integers; F_c denotes the scaling factor which is a uniformly distributed random number in the interval $[0, 1]$.

Actually, the operator, $\mathbf{X}_{jump}^k = \mathbf{X}_i^{(k)} + F_c \times \left(\mathbf{X}_{r1}^{(k)} - \mathbf{X}_{r2}^{(k)}\right)$, aims to create a jump to avoid the local trap. However, if $\mathbf{X}_{r1}^{(k)}$ and $\mathbf{X}_{r2}^{(k)}$ are relatively close to each other, it is unable to explore any new prospective zone. To overcome this problem, in this work, instead of using $\mathbf{X}_{jump}^k = \mathbf{X}_i^{(k)} + F_c \times \left(\mathbf{X}_{r1}^{(k)} - \mathbf{X}_{r2}^{(k)}\right)$, we deploy an operator proposed by Nguyen et al. [Nguyen et al., 2018b], which is defined as follows:

$$\mathbf{X}_{jump}^k = \begin{cases} \mathbf{X}_i^{(k)} + F_c \times \left(\mathbf{X}_{r1}^{(k)} - \mathbf{X}_{r2}^{(k)}\right), & \text{if} \quad FDR_i > 10^{-3} \\ \mathbf{X}_i^{(k)} + F_c \times \left[(\mathbf{X}_{r1}^{(k)} - \mathbf{X}_{r2}^{(k)}) + (\mathbf{X}_{r3}^{(k)} - \mathbf{X}_{r4}^{(k)})\right], & \text{otherwise} \end{cases} \tag{IV.27}$$

where $r1$, $r2$, $r3$ and $r4$ are mutually different random integers; FDR_i is the fitness difference ratio (FDR) of the i^{th} solution in the current population, which is defined as follows:

$$FDR_i = \left| \frac{\mathfrak{f}_i - \mathfrak{f}_{gBest}}{\mathfrak{f}_{gBest}} \right| \tag{IV.28}$$

where $\mathfrak{f}_i$ and $\mathfrak{f}_{gBest}$ are the fitness values of the i^{th} solution and the best solution found so far, respectively.

Furthermore, to increase the exploration of the search space in the beginning stage and the exploitation of the best solutions found so far towards the end of the algorithm, p_a is updated as follows:

$$p_a^{(k)} = p_{a_{max}} - (p_{a_{max}} - p_{a_{min}}) \cdot (k/N_{iter}) \tag{IV.29}$$

where $p_a^{(k)}$ is the switching parameter p_a at the k^{th} iteration; $[p_{a_{min}}, p_{a_{max}}]$ is the range of p_a, with $p_{a_{min}} = 0.01$ and $p_{a_{max}} = 0.5$; N_{iter} is the maximum number of allowable iterations.

Finally, by using the greedy strategy, the next generation solution is created. At the end of each iteration process, the best solution obtained so far is updated. The procedures of improved Cuckoo search (ICS) algorithm can be described as the pseudo code shown in **Algorithm** IV.1

IV.3.2 An improved particle swarm optimization algorithm

The particle swarm optimization (PSO) algorithm is also a population-based stochastic optimization algorithm and regarded as a global search strategy. This algorithm is inspired from the

cooperation and communication of a swarm of birds [Eberhart and Kennedy, 1995]. Due to the simple representation and relatively low number of adjustable parameters, PSO has become one of the most popular choices and is efficiently applicable to optimization problems [Zhang et al., 2015b]. However, the major drawback of the PSO is that it may be trapped into a local optimal solution region.

In PSO, each individual (called *particle*), which represents a potential solution to the optimization, of a given population (called *swarm*), is updated according to its own experience and that of its neighbours. The quality of a candidate solution is evaluated by the fitness value associated with it. Let us consider a swarm of P particles, where each particle has a position vector, $\mathbf{X}_i = \{x_{i1}, \ldots, x_{iC}\}$, a velocity vector, $\mathbf{V}_i = \{v_{i1}, \ldots, v_{iC}\}$, its own best position $pBest$ found so far, and interacts with neighbouring particles via the best position $gBest$ discovered in the neighbourhood so far. At iteration k^{th} in the search process, particles are moved according to the following equations:

$$\mathbf{V}_i^{(k+1)} = w^{(k)}\mathbf{V}_i^{(k)} + c_1 r_1 \left[pBest^{(k)} - \mathbf{X}_i^{(k)}\right] + c_2 r_2 \left[gBest^{(k)} - \mathbf{X}_i^{(k)}\right] \tag{IV.30}$$

$$\mathbf{X}_i^{(k+1)} = \mathbf{X}_i^{(k)} + \mathbf{V}_i^{(k+1)} \tag{IV.31}$$

where r_1 and r_2 are random variables, uniformly distributed in $[0, 1]$ to provide stochastic weighting of the different components participating in the velocity. c_1 and c_2 are acceleration coefficients that scale the influence of the *cognitive* and *social* components, respectively, and w is an inertia weight. In addition, the flying velocity is limited to a reasonable range $[v_{min}, v_{max}]$. In this work, $v_{min} = -v_{max} = -3$ is set to constrain particle movement.

In order to improve the performance of the PSO algorithm, in this work, we use the same approach as in [Chen et al., 2017b] that creates a high diversity population to provide a good guidance for particles. As a result, the problem of premature convergence can be partially avoided and the exploitation ability of the algorithm can be improved. Here, we employ a differential mutation scheme proposed by Mohamed [Mohamed, 2018] for this purpose. The scheme is expressed as follows:

$$\begin{aligned}\mathbf{X}_i^{(k+1)} = \overline{\mathbf{X}}_i^{(k)} &+ F_{p1} \times \left(\mathbf{X}_{best}^{(k)} - \mathbf{X}_{better}^{(k)}\right) \\ &+ F_{p2} \times \left(\mathbf{X}_{best}^{(k)} - \mathbf{X}_{worst}^{(k)}\right) + F_{p3} \times \left(\mathbf{X}_{better}^{(k)} - \mathbf{X}_{worst}^{(k)}\right)\end{aligned} \tag{IV.32}$$

where $\mathbf{X}_{best}^{(k)}$, $\mathbf{X}_{better}^{(k)}$ and $\mathbf{X}_{worst}^{(k)}$ are the tournament best, better and worst three randomly selected particles in the current population, respectively. F_{p1} , F_{p2} and F_{p3} are the mutation factors generated independently from a uniform distribution in $(0, 1)$. And, $\overline{\mathbf{X}}_i^{(k)}$ is a convex combination vector of the triangle, which is defined as follows:

$$\overline{\mathbf{X}}_i^{(k)} = \delta_1 \mathbf{X}_{best}^{(k)} + \delta_2 \mathbf{X}_{better}^{(k)} + \delta_3 \mathbf{X}_{worst}^{(k)} \tag{IV.33}$$

where the real weights $\delta_i, (i = 1, 2, 3)$ are given by $\delta_i = p_i / \sum_{i=1}^{3} p_i$. Here, p_1, p_2 and p_3 are set representatively to 1, $rand(0.75, 1)$ and $rand(p_2, 1)$, in which $rand(a, b)$ returns a real number between a and b.

From Eqs. (IV.32) and (IV.33), it can be observed that there are two main benefits. First, the landscape of different sub-regions around the best vectors can be explored by forming many different sizes and shapes of triangles through the optimization process. Second, the global solution can be easily reached if all vectors follow the direction of the best vectors since the convex combination $\overline{\mathbf{X}}_i^{(k)}$ consists of the best vector with higher weight.

Furthermore, to have a reasonable balance between exploration and exploitation during the optimization process (increase exploration ability in the beginning stage and increase exploitation

Algorithm IV.2: The IPSO algorithm

Initialization: Initialize randomly a population $\mathbf{X}$ of size P; set up the flight control parameters: $\{c_1, c_2\}$ and w; determine $pBest$ and $gBest$
Results : The optimal solution $gBest$
$k \leftarrow 1$
repeat
/* Provide new population with high diversity */
1 **for** *each particle (say* $\mathbf{X}_i$*)* **do**
1.1 Generate a new particle (say $\mathbf{X}'_i$) using Eqs. (IV.32) and (IV.33)
1.2 Evaluate fitness value, $\mathfrak{f}'_i$
1.3 **if** $\mathfrak{f}'_i$ *is better* **then**
Replace $\mathbf{X}_i$ by the new one, X'_i
/* Flight particles by using PSO operators */
2 Update the inertia weight w using Eq. (IV.34)
3 **for** *each particle (say* $\mathbf{X}_j$*)* **do**
3.1 Update the velocity $\mathbf{V}_j$ using Eq. (IV.30)
3.2 Update the position $\mathbf{X}_j$ using Eq. (IV.31)
/* Evaluate solutions */
4 Evaluate fitness values, $\mathfrak{f}$
5 Update the $pBest$ and the $gBest$
6 $k \leftarrow k + 1$
until *the stopping criteria are met*

towards the end), here, a parameter control strategy proposed by Yang et al. [Yang et al., 2015] is used. In the strategy, $c_1 = c_2 = 2$ and the inertia weight, w, updating scheme is expressed as follows:

$$w^{(k)} = w_{max} - (w_{max} - w_{min}) \cdot (k/N_{iter})^{\frac{1}{\pi^2}} \tag{IV.34}$$

where $w^{(k)}$ is the inertia weight at the k^{th} iteration. $[w_{min}, w_{max}]$ is the range of inertia weight, with $w_{min} = 0.4$ and $w_{max} = 0.9$. The procedures of improved PSO algorithm can be described as the pseudo code shown in **Algorithm** IV.2.

IV.3.3 Proposed hybrid metaheuristic

There are no theoretical or experimental guarantees that any optimization algorithm can avoid getting stuck in suboptimal solutions. In order to increase the accuracy of the results and decrease the likelihood of trapping into local solution regions, the common approach consists in hybridizing or combining different techniques [Fister Jr et al., 2013, Sengupta et al., 2018].

In this work, an effective hybrid optimization algorithm, called hybrid ICS/IPSO, is proposed based on the strategies described above. The proposed algorithm is built based on two main steps. The first step is used for finding different promising regions in the search space by using both the improved CS and improved PSO algorithms. The second step is used for selecting and updating solutions by maintaining the merits of both algorithms. Hence, we can achieve fast convergence and partially avoid the problem of getting stuck in suboptimal solutions. These two steps are repeated alternately until the termination criteria are satisfied. The main procedure of hybrid ICS/IPSO is given in Algorithm IV.3.

There are several works, which hybridize CS and PSO algorithms [Wang et al., 2015, Bouyer and

Hatamlou, 2018], for solving the global optimization problem. However, the proposed algorithm has its own specific characteristics, which make it different from others used in the literature. First, both CS and PSO algorithms are improved to explore more untouched areas in the search space. Second, since the ICS and IPSO are independently operated during the optimization process, the merits of both are maintained. Third, by using a simple effective mechanism of selection, potential solutions not only follow the direction of the best one, but also avoid the direction of the worst. In addition, sharing solutions can help both ICS and IPSO algorithms to compensate for their weaknesses.

Algorithm IV.3: The hybrid ICS/IPSO algorithm

Initialization: Initialize randomly a population $\mathbf{X}$ and two external archives: pA_{ic} and pA_{ip}, of size P; set up parameters for both ICS and IPSO algorithms; evaluate the population with both ICS and IPSO algorithms to determine $gBest_{ic}$, $pBest_{ip}$, $gBest_{ip}$ and $gBest$

Results : The optimal solution $gBest$

$k \leftarrow 1$

repeat

`/* ICS phase, operating on X */`

1 **for** *each potential solution in* $\mathbf{X}$ *(say* $\mathbf{X}_i$*)* **do**

Perform from step (1.1) to step (1.4.3) in Algorithm IV.1

2 Perform step (2) in Algorithm IV.1

3 Sort and store new solutions in pA_{ic}

`/* IPSO phase, operating on X */`

4 Perform from step (1) to step (4) in Algorithm IV.2

5 Update $pBest_{ip}$, sort and store new solutions in pA_{ip}

`/* Selection phase */`

6 Determine the current best by:

$$\underset{\mathbf{X}_j \in (pA_{ic} \cup pA_{ip})}{\operatorname{argmin}} \mathfrak{f}(\mathbf{X}_j) \quad \textbf{or} \quad \underset{\mathbf{X}_j \in (pA_{ic} \cup pA_{ip})}{\operatorname{argmax}} \mathfrak{f}(X_j)$$

7 Update $gBest$, then assign: $(gBest_{ic} \leftarrow gBest)$ and $(gBest_{ip} \leftarrow gBest)$

8 Update $\mathbf{X}$ by selecting the best half of solutions in pA_{ic} and pA_{ip}

9 $k \leftarrow k + 1$

until *the stopping criteria are met*

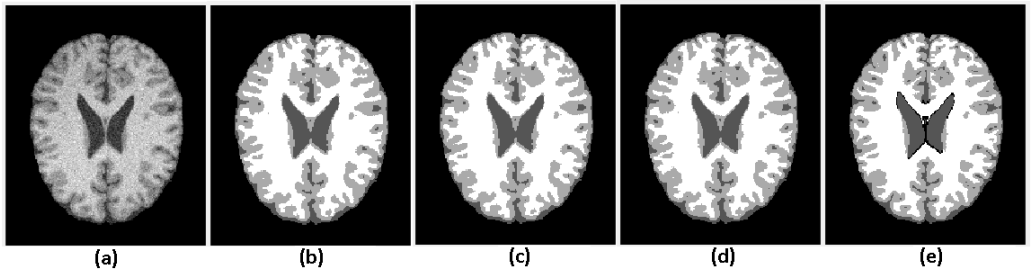

Figure IV.5: Segmentation results of a simulated brain MR image produced by different algorithms: (a) original image; (b) result with ICS algorithm; (c) result with IPSO algorithm; (d) result with hybrid ICS/IPSO algorithm; (e) ground truth.

To examine the efficiency of the proposed algorithm, we conducted some experiments by applying it to solve the MR image segmentation problem. Here, the fitness value is calculated by using Eq.

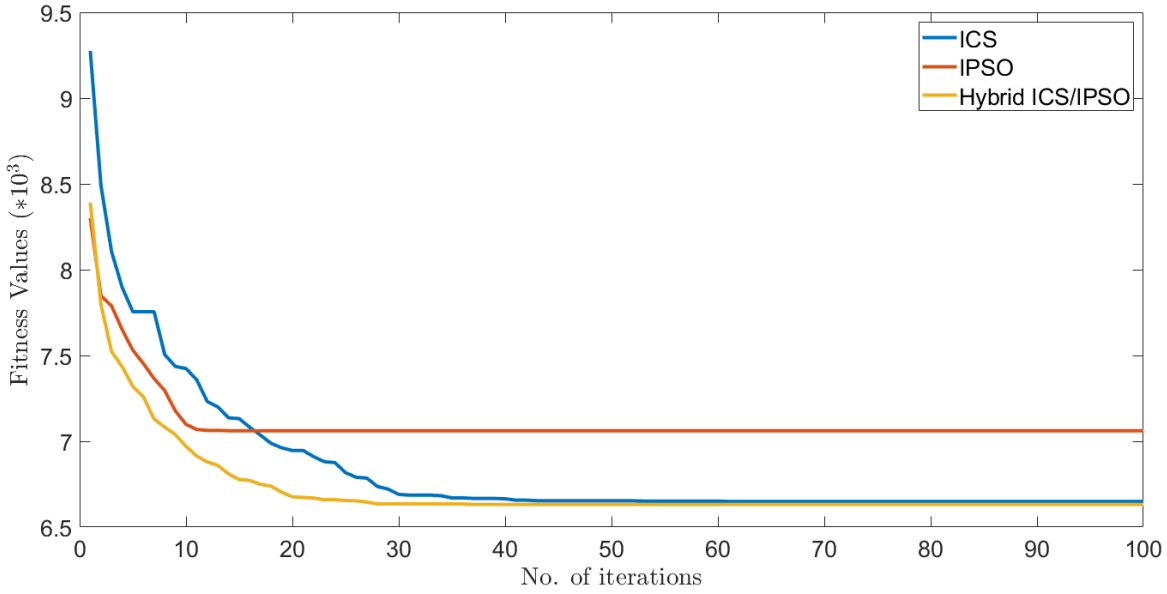

Figure IV.6: Convergence characteristics of segmenting a simulated brain MR image using ICS, IPSO and Hybrid ICS/IPSO algorithms.

(IV.21). The simulated MR images from the brainWeb, which have characteristics of 9% noise and 40% intensity non-uniformity, are used. To evaluate the segmentation results quantitatively, we use DICE coefficient as a metric.

As an example, Figure. IV.6 depicts the convergence process of the objective function values when segmenting a simulated brain MR image (slice 95). As shown in this figure, it is obvious that the proposed algorithm can converge to the optimal solution more rapidly and accurately than the other algorithms. In addition, Figure. IV.5 and Table IV.3 also reveal that better segmentation results can be achieved when using the proposed algorithm.

Table IV.3: DICE coefficients obtained from segmenting a simulated MR image by using different algorithms.

Metric	**Regions**	**Algorithms**		
		Improved CS	**Improved PSO**	**Hybrid ICS/IPSO**
DICE	CSF	0.9875	0.9875	**0.9875**
	GM	0.9524	0.9513	**0.9525**
	WM	0.9685	0.9669	**0.9692**

*The values in bold indicate the best performance.

IV.4 Proposed method

This section describes the proposed brain MR segmentation method, which is based on the proposed MRF-based segmentation model and the hybrid metaheuristic algorithm (hybrid ICS/IPSO). The fitness function is calculated by using Eq. (IV.21). To obtain the segmented images after achieving the optimal mean intensity of each region, two different steps need to be done: (1) Labelling each pixel in the image by classifying it to the nearest mean; (2) Filtering the labelled image by using Median filter with the structure of $[3 \times 3]$. The details of this algorithm are illustrated in the following sections.

IV.4.1 Solution representation

In this study, the potential solutions are made up of real numbers, which are *nests* in the ICS or *positions* in the IPSO. The potential solutions represent the mean intensities of various regions in the image. For P solutions with C distinct elements, there are in total $(P.C)$ optimization variables that need to be encoded. For example, in the population $\mathbf{X} = \{\mathbf{X}_1, \ldots, \mathbf{X}_P\}$, the i^{th} potential solution can be described as: $\mathbf{X}_i = \{x_{i1}, \ldots, x_{iC}\}$, where x_{ij} (with $j = 1, \ldots, C$) represents the j^{th} mean intensity of the j^{th} region. In this way, the mean intensity of each region, $\boldsymbol{\mu} = \{\mu_1, \ldots, \mu_C\}$, can be obtained by decoding $\mathbf{X}_i$.

IV.4.2 Segmentation criterion

In this research, the value of the fitness function, $\mathfrak{f}_i$, corresponding to the i^{th} solution, is calculated by evaluating $\Psi(\mathbf{X}_i)$ according to Eq. (IV.21).

$$\mathfrak{f}_i = \Psi(\mathbf{X}_i) \tag{IV.35}$$

Algorithm IV.4: The ICS/IPSO-based image segmentation algorithm

Initialization: Read the input image; set the number of regions C; set maximum number of allowable iterations N_{iter}; do bias correction and initialize population $\mathbf{X}$; initialize parameters for ICS, IPSO and MRF-based segmentation model ($\beta = 0.95$); determine $gBest_{ic}$, $pBest_{ip}$, $gBest_{ip}$ and $gBest$; initialize two external archives: pA_{gBest} and pA_{CJV}.

Results : The optimal region centers, $(\boldsymbol{\mu}^* = gBest^*)$

for $(t = 1;\ t \leq T;\ t = t + 1)$ **do**

`/* Hybrid ICS/IPSO algorithm phase */`

1 **repeat**

`/* Computing the fitness values by using Eq. (IV.35) */`

`/* argmin is used in step (6) */`

Perform from step (1) to step (8) in Algorithm IV.3

until *the stopping criteria are met*

`/* Get solution corresponding to each value of` β `*/`

2 Store $gBest$ in pA_{gBest}

3 Calculate CJV criterion using Eq. (IV.36) and store it in pA_{CJV}

`/* Increase` β `with step size of 0.1 */`

4 $\beta = \beta + 0.1 \times t$

5 Get the optimal region centers by:

$$gBest^* = \underset{gBest \in pA_{gBest}}{\operatorname{argmin}} \{pA_{CJV}(gBest)\}$$

The minimization of $\mathfrak{f}_i$ is the same as the minimization of the objective function, $\Psi(\mathbf{X}_i)$, which can lead to an optimal partitioning of the MR image.

IV.4.3 Optimal search process

The hybrid ICS/IPSO algorithm described in Section IV.3.3 is used as a searching engine to do the optimization step in the MRF segmentation approach. By taking advantages of both the proposed

MRF-based segmentation model and the hybrid ICS/IPSO algorithm, the optimal solution for the MR image segmentation problem can be found. The framework for the problem is summarized in Algorithm IV.4.

Note that, we assume that initialization and bias correction of segmented images are well established by using results from the previous chapters (from the PSO-KFECSB algorithm (**Algortihm** II.3) or the MOPSO-based algorithm (**Algortihm** III.3)). Particularly, a good approximation solution, called *starting point*, needs to be provided such that the proposed algorithm can converge quickly and be able to obtain the global optimal solution. Here, we use the results produced by the MOPSO-based algorithm. The main reason is that the MOPSO-based method provides a set of solutions instead of one solution only. As a result, we can have a flexible choice for setting up the starting point of the proposed algorithm. In addition, the coefficient of joint variation (CJV) between WM and GM regions [Chua et al., 2009], which is a well-known criterion to evaluate bias field correction methods, is used as a metric to determine the point. From that, the initial population is created by randomly generating solutions around it.

The CJV is defined as follows:

$$CJV = \frac{\sigma(GM) + \sigma(WM)}{|\mu(GM) - \mu(WM)|} \tag{IV.36}$$

where $\sigma(.)$ and $\mu(.)$ denote the standard deviation and the mean intensity. A smaller CJV value corresponds to the better performance.

As an example, Figure. IV.7 shows the bias correction and the initial labelling image ($CJV = 0.4501; \boldsymbol{\mu}_0^* = (0.4502, 86.0597, 114.1679)$) selected by using CJV criterion (slice 24 in the 20-normal brain MR datasets). As can be seen in the figure, the results are satisfactory for the initial conditions.

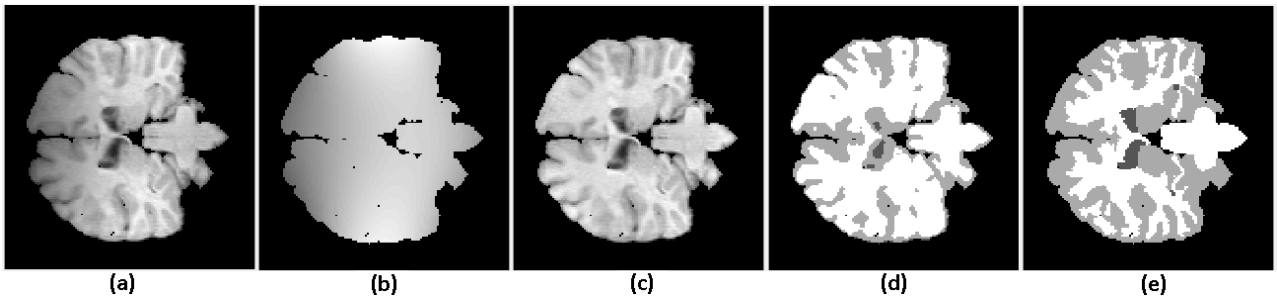

Figure IV.7: Initialization and bias correction of a real brain MR image by using the MOPSO-based method: (a) original image; (b) bias field; (c) bias correction image; (d) initial labelling image; (e) ground truth.

In addition, to ensure that all solutions are moving within the search space and avoiding divergent behaviour, the boundary conditions for the i^{th} potential solution are limited as follows:

$$x_{ij}^{(k)} = \begin{cases} x_{max}, & \text{if } x_{ij}^{(k)} > x_{max} \\ x_{min}, & \text{if } x_{ij}^{(k)} < x_{min} \\ +x_{ij}^{(k)}, & \text{otherwise} \end{cases} \tag{IV.37}$$

$$v_{ij}^{(k)} = \begin{cases} v_{max}, & \text{if } v_{ij}^{(k)} > v_{max} \\ v_{min}, & \text{if } v_{ij}^{(k)} < v_{min} \\ +v_{ij}^{(k)}, & \text{otherwise} \end{cases} \tag{IV.38}$$

where v_{min} and v_{max} are the smallest and largest allowable step sizes in any dimension ($v_{min} = -v_{max} = -3$ is set in this paper); and $\{x_{min}, x_{max}\}$ are the bounds of the search space in each dimension. Actually, they are the minimum and maximum of the intensity of the input image.

Furthermore, to stop the phase of hybrid ICS/IPSO algorithm efficiently, two criteria are set: the maximum number of the allowable iterations and the maximum number of non-significant improvements of the fitness value, $\mathfrak{f}_{gBest}$. Particularly, if $(k > N_{iter})$ is reached or $(|\mathfrak{f}_{gBest,\text{new}} - \mathfrak{f}_{gBest,\text{old}}| < 10^{-4})$ is completed $(0.1 \times N_{iter})$ times, this phase is immediately stopped.

Finally, to get rid of the problem of selecting parameter β, which is described in Section IV.5.1, we use CJV criterion to estimate the quality of solutions [Xie et al., 2015] so that the final solution can be found. The whole algorithm for the image segmentation problem is illustrated in Figure IV.8.

IV.5 Results and discussions

In this section, we empirically evaluate the performance of the proposed method. Both qualitative and quantitative evaluations are involved in this study.

IV.5.1 Experimental setup

To validate the effectiveness, the performance metric values of the proposed method have been evaluated and compared with four state-of-the-art algorithms in the literature. These algorithms are: FCM algorithm based on morphological reconstruction and membership filtering (FRFCM) [Lei et al., 2018], the multiplicative intrinsic component optimization (MICO) [Li et al., 2014], the integrating metaheuristic multilevel threshold with Markov random field (PSO-MRF) [T Krishnan et al., 2016] and the MOPSO-based method (iMOPSO) [Pham et al., 2019b]. The parameter settings of the competing algorithms are specified in Table IV.4.

To perform experiments, the parameters of the proposed algorithm are set as follows: population size $P = 30$, maximum number of allowable iterations $N_{iter} = 100$. Note that, the parameter β in Eq. (IV.21), which controls the balance between two components: feature modelling and region labelling, has an important effect on the performance of the algorithm. Particularly, it tunes the degree of homogeneity of each region in the segmented image. If β makes the feature modelling component dominant (small value of β), spatial relationship information will be ignored. On the other hand, if β makes the region labelling component dominant (large value of β), the values of estimated solutions may deviate considerably and the segmented result is not consistent (excessive smoothing of boundaries may occur). Unfortunately, there is no closed-form definition for β as the normalizing constant $Z(U)$ is intractable (Eqs. (IV.12), (IV.13) and (IV.14)). In addition, different input images can have different spatial organizations. However, from experiments we found that $\beta \in [0.95, 1.45]$ can produce satisfactory results in most cases. Hence, in this work, β is determined empirically as proposed in [Besag, 1986] by gradually increasing its value, from 0.95 to 1.45 with step of 0.1 ($T = 5$), through the algorithm loops. All the other parameters are set as in the previous Sections.

All algorithms are implemented in MATLAB 2014b and executed with a computer with Intel Core i7 1.8 GHz CPU, 8GB RAM using Microsoft Windows 10.

IV.5.2 Datasets

The MR images used in this study include both T1-weighted simulated and real 2D brain MR images. For simulated MR images, they are downloaded from a well-known database: the brainWeb

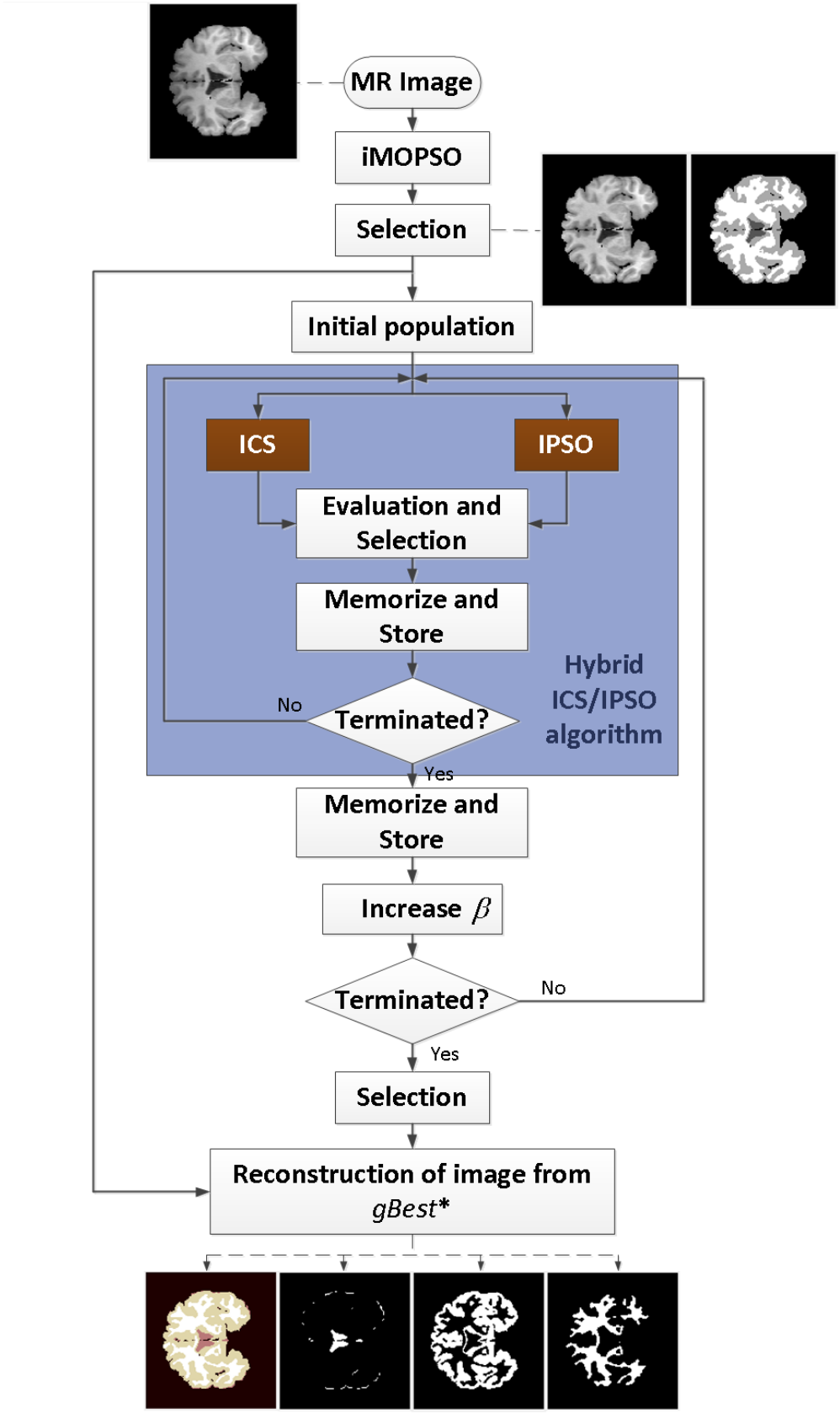

Figure IV.8: Block diagram of the proposed method.

Table IV.4: Parameter settings of the competing algorithms.

Algorithms	Parameter settings
FRFCM	number of clusters 4 (or 3 for real MR images), window size of 3×3 pixels, square structuring element of size 3×3, maximum number of iterations $t = 100$, minimum amount of improvement $\eta = 0.001$, and exponent for the partition matrix $m = 2$.
MICO	multiphase $L = 3$, 15 polynomials of the first four orders as the basis functions, maximum number of iterations $t = 15$, exponent for the partition matrix $q = 1$, and minimum amount of improvement $\varepsilon = 0.001$.
PSO-MRF	number of clusters 4, number of EM iterations and MAP iterations: $N_{em} = 3$ and $N_{map} = 3$.
iMOPSO	$P = 30$, $N_r = 30$, $N_g = 10$, $N_{iter} = 200$ (180 for real MR images). The values of (n, η, and λ) are 10, 2.5, and 1/9. $\epsilon = 3, \sigma = 5, \alpha = 2, \triangle t = 1$, and $\mu = 1$, are set as common values. $\nu = 0.00015 * 255^2, \rho = 9$ and $\zeta = 0.1$ for simulated MR images and $\nu = 0.003 * 255^2, \rho = 0.4$, and $\zeta = 1.65$ for real MR images.

from a McConnell Brain Imaging Center [Kwan et al., 1999], which can be reached in (https://brainweb.bic.mni.mcgill.ca/brainweb/). This dataset includes different noise levels: from 0 to 9%, and different INU levels: 0, 20%, and 40%, and five different slice thicknesses. Images with size of 181×217 and thickness of 1 mm are used in this work. On the other hand, real MR images are taken in the 20-normal brain MR data sets, which contain manual segmentation by an expert technician, provided by the Center for Morphometric Analysis at Massachusetts General Hospital. The data sets are available at (http://www.nitrc.org/projects/ibsr/). Images with characteristics of size 135×142 and 1.171751 mm thickness, are used in our experiments.

IV.5.3 Performance measures

Since the ground truth images are available in the datasets, for quantitatively comparing the performance, three criteria are involved, which are the Dice Similarity Coefficient (DICE), the Hausdorff distance (HD), and the Accuracy (AC). These metrics are presented in details in the previous chapters (Section II.4.3 and Section III.4.3).

IV.5.4 Results on simulated MR images

In this section, simulated brain MR images from the brainWeb are used for the purpose of performance evaluation. The experiment is conducted on a set of 9 images (slices: 85, 87, 89, 91, 93, 95, 97, 99 and 101) representing the "worst case" (low contrast and relatively large spatial inhomogeneities); the images have the characteristics of 9% noise and 40% INU artifact. Each image is segmented into four classes: cerebrospinal fluid (CSF), gray matter (GM), white matter (WM) and the background.

Figure IV.9 shows the qualitative results of the segmentation of a T1-weighted image (slice 101) provided by the considered algorithms. This figure reveals that though the MOPSO-based method and the proposed algorithm produce better results among those, the proposed method reserves the best in details of the image. Figure IV.10 shows the qualitative results of the segmentation of 8 T1-weighted images (slices: 85, 87, 89, 91, 93, 95, 97, and 99) produced by the proposed algorithm. In spite of existing artifact, as can be seen from the figure, the segmented images are almost close to the ground reference data, hence, our method achieves a high performance when segmenting brain MR images. Thus, it can be concluded that the proposed method qualitatively provides satisfactory results.

In order to compare more clearly the performance of the considered methods, quantitative evaluation is also taken into account. Two supervised metrics, named DICE and AC, are involved. Note that a higher value indicates a better correspondence to the ground-truth. Figure IV.11, IV.12

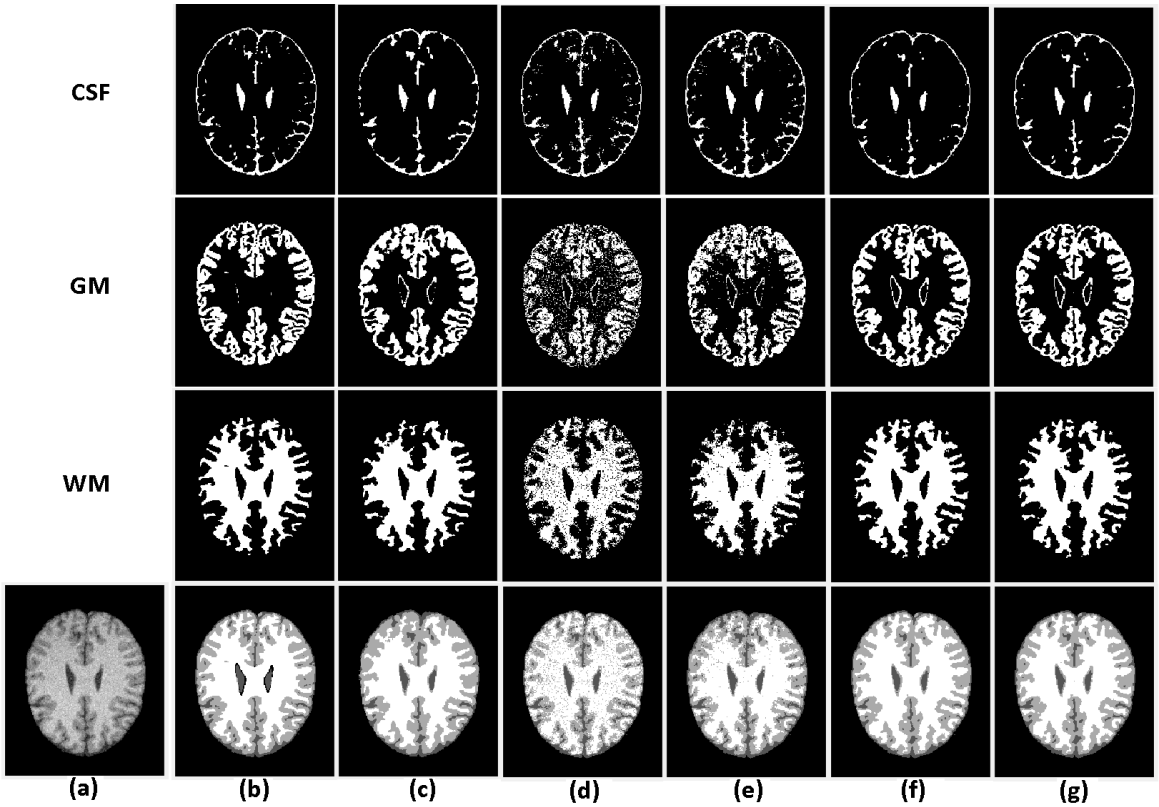

Figure IV.9: Qualitative segmentation results of a simulated brain MR image provided by the considered algorithms: (a) original image; (b) ground truth; (c) FRFCM results; (d) MICO results; (e) PSO-MRF results; (f) iMOPSO results; (g) the proposed method results.

Table IV.5: Mean and standard deviation of DICE and AC results for the considered algorithms calculated over the set of 9 simulated brain MR images from the brainWeb.

Methods	**Metrics**	**Regions**					
		CSF		**GM**		**WM**	
		Mean	Std. Dev.	Mean	Std. Dev.	Mean	Std. Dev.
FRFCM	DICE	0.9791	0.0019	0.9333	0.0081	0.9532	0.0047
	AC	0.9635	0.0033	0.9073	0.0081	0.9395	0.0067
MICO	DICE	0.9815	0.0022	0.8896	0.0073	0.9040	0.0049
	AC	0.9679	0.0036	0.8444	0.0072	0.8772	0.0061
PSO-MRF	DICE	0.9790	0.0031	0.9219	0.0102	0.9451	0.0050
	AC	0.9639	0.0053	0.8905	0.0107	0.9288	0.0072
iMOPSO	DICE	0.9847	0.0015	0.9436	**0.0060**	0.9633	**0.0023**
	AC	0.9739	0.0028	0.9222	**0.0062**	0.9530	**0.0035**
Proposed	DICE	**0.9865**	**0.0015**	**0.9469**	0.0071	**0.9661**	0.0033
	AC	**0.9764**	**0.0026**	**0.9261**	0.0074	**0.9566**	0.0045

*The values in bold indicate the best performance.

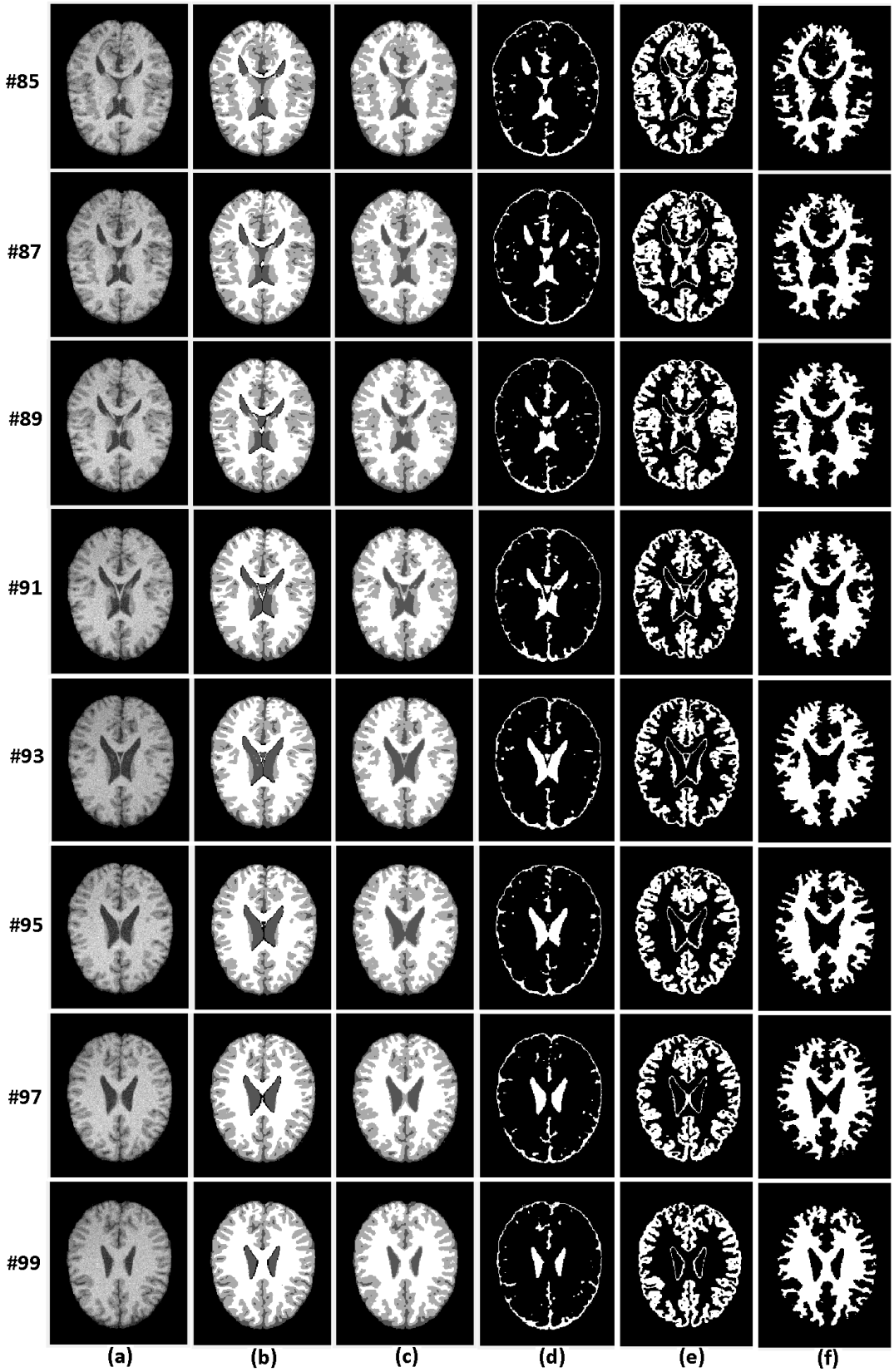

Figure IV.10: Qualitative segmentation results of simulated brain MR images provided by the proposed algorithm versus the ground truth images: (a) original images; (b) ground truth images; (c) segmentation results; (d) CSFs; (e) GMs; (f) WMs.

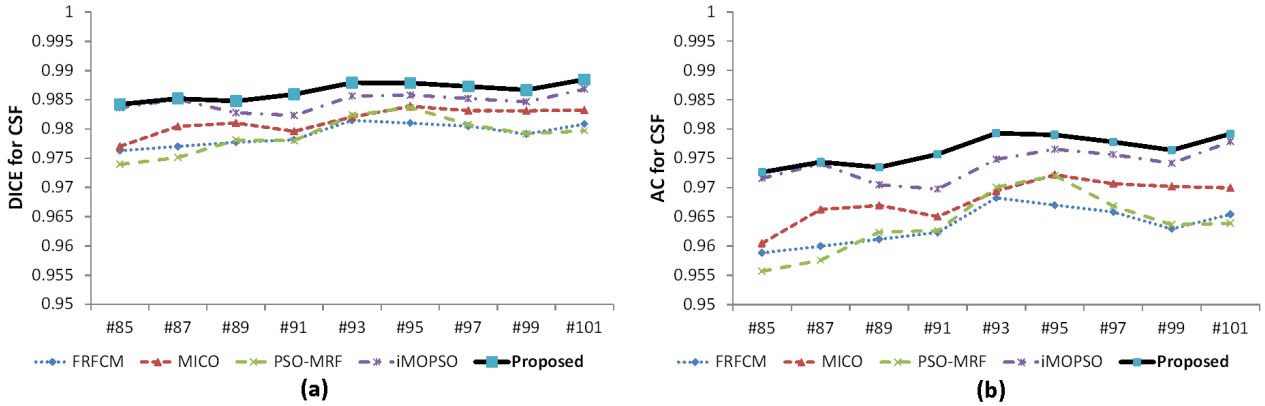

Figure IV.11: Dice coefficient and accuracy associated with CSF tissue for the competing segmentation algorithms on different simulated brain MR scans with 9% noise and 40% INU artifact from the brainWeb datasets: (a) DICE; (b) AC.

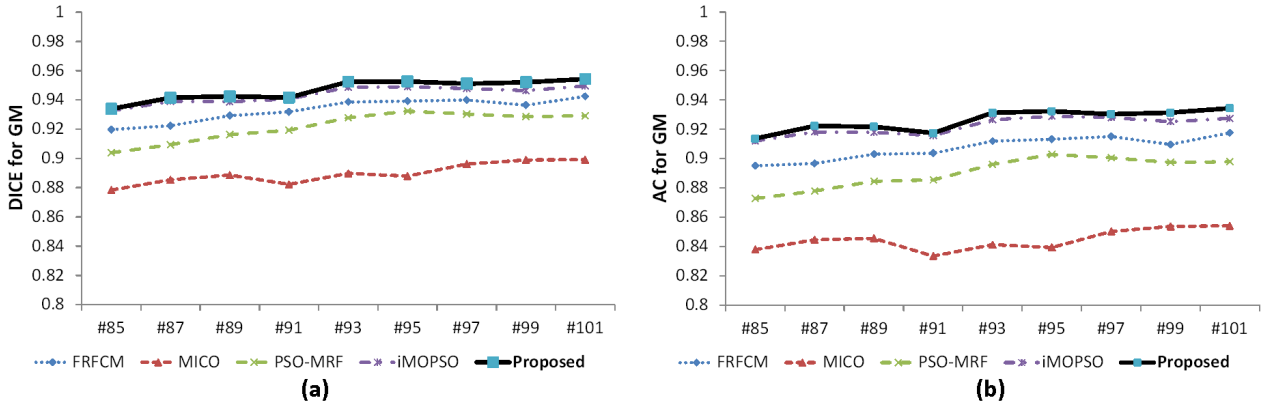

Figure IV.12: Dice coefficient and accuracy associated with GM tissue for the competing segmentation algorithms on different simulated brain MR scans with 9% noise and 40% INU artifact from the brainWeb datasets: (a) DICE; (b) AC.

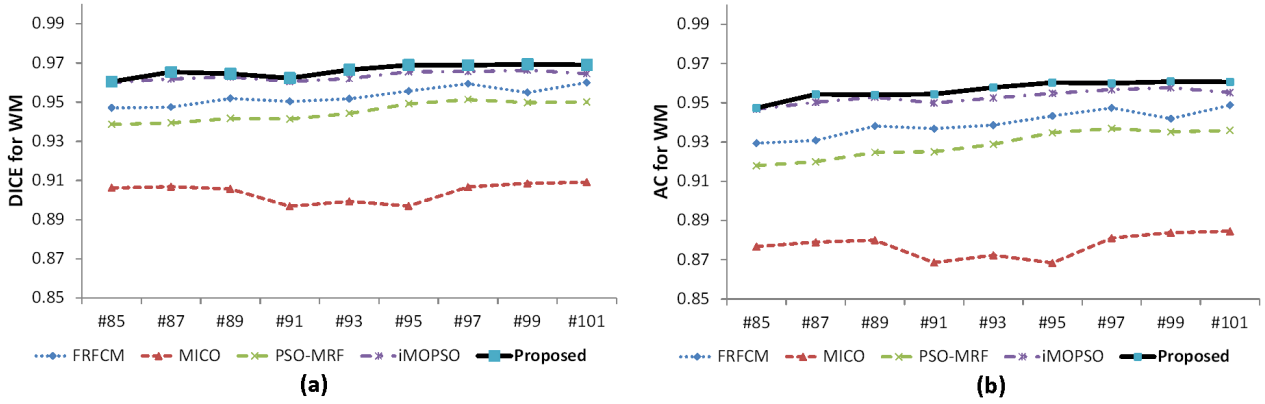

Figure IV.13: Dice coefficient and accuracy associated with WM tissue for different segmentation algorithms on the competing simulated brain MR scans with 9% noise and 40% INU artifact from the brainWeb datasets: (a) DICE; (b) AC.

and IV.13 show the comparison results. The results are summarized in Table IV.5.

As can be seen from the figures and table, the proposed method generally gives the best scores. Even though the MOPSO-based method provides a little more consistent GM and WM results, the proposed method achieves better performance. This confirms that the proposed method performs more efficiently on simulated brain MR images compared to its competitors.

IV.5.5 Results on real MR images

We have also examined the performance of our method on real brain MR images. The experiment is conducted on a set of 9 images in the 20-normal T1-weighted real brain MR dataset (slices: 24, 26, 28, 30, 32, 34, 36, 38 and 40). It has been used in a variety of volumetric studies in the literature as it contains varying levels of difficulty, with the worst scans consisting of low contrast and relatively large spatial inhomogeneities. The number of regions in the segmentation process is set to three: cerebrospinal fluid (CSF), gray matter (GM), white matter (WM). The background pixels are ignored in the computation.

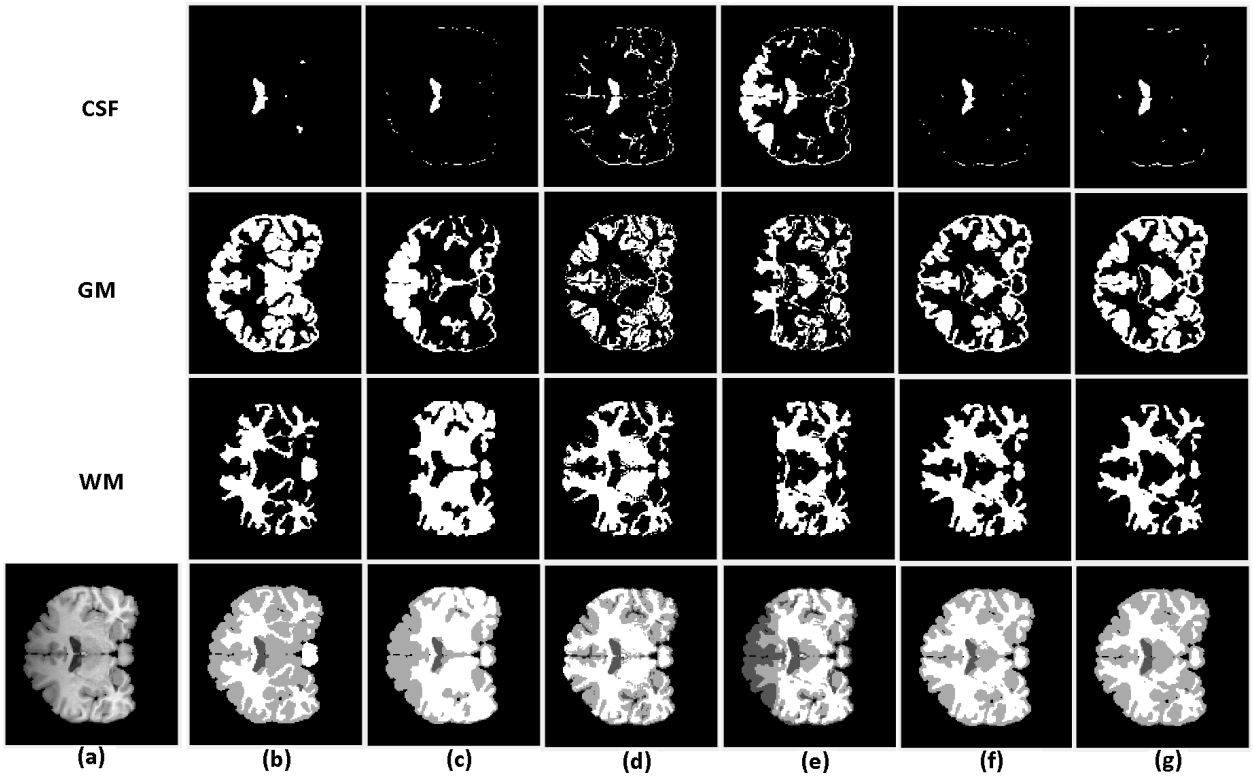

Figure IV.14: Qualitative segmentation results of a real brain MR image provided by the considered algorithms:(a) original image; (b) ground truth; (c) FRFCM results; (d) MICO results; (e) PSO-MRF results; (f) iMOPSO results; (g) the proposed method results.

Figure IV.14 shows the qualitative results of the segmentation of a T1-weighted image (slice 30) provided by the considered algorithms. This figure reveals that the proposed method provides superior results. It may be worth mentioning here that though the MOPSO-based method is comparable with the proposed method on simulated MR images, the current method achieves much higher performance on real brain MR images. Figure IV.15 shows the qualitative results of the segmentation of 8 T1-weighted images. As can be seen from this figure, the proposed method provides an appropriate segmentation of real brain MR images.

To evaluate quantitatively the segmentation of the selected images, in this experiment, the surface distance-based metric, Hausdorff distance, is also calculated along with DICE and AC. Figure IV.16, IV.17 and IV.18 show the comparison results. The results are summarized in Table IV.6. From the figures and the table, it can be seen that the proposed method outperforms its competitors. Again, the results presented here confirm the efficiency of the proposed method and demonstrate its superiority over the others.

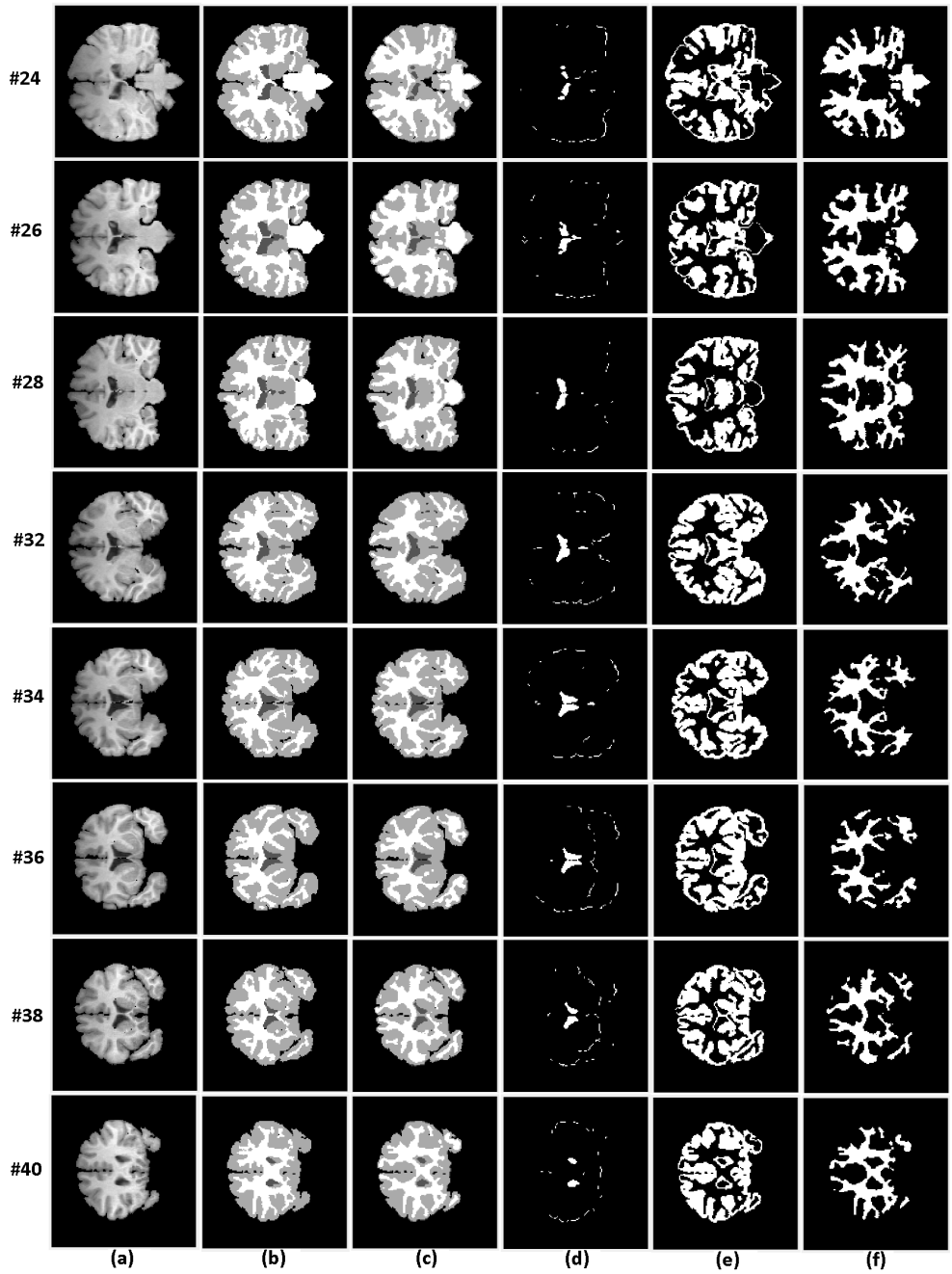

Figure IV.15: Qualitative segmentation results of real brain MR images provided by the proposed algorithm versus the ground truth images: (a) original images; (b) ground truth images; (c) segmentation results; (d) CSFs; (e) GMs; (f) WMs.

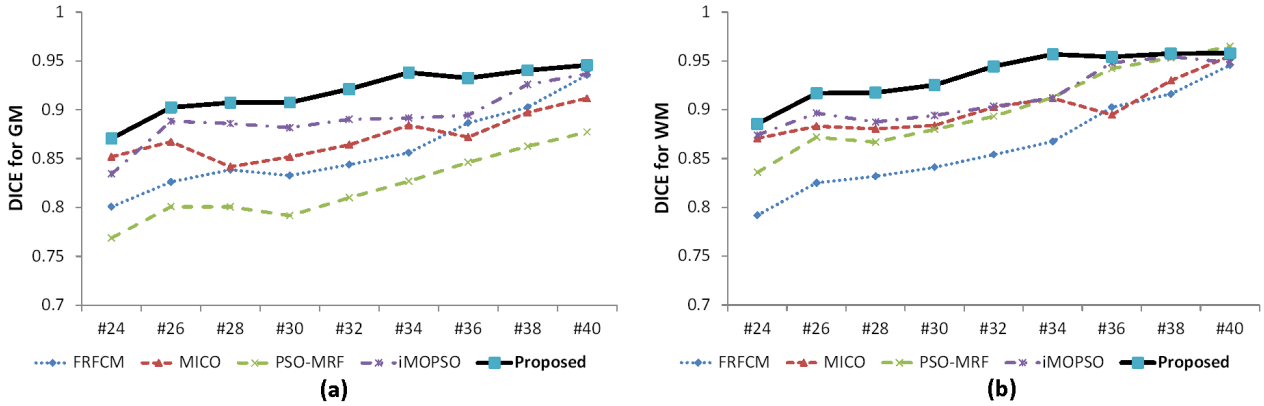

Figure IV.16: Dice coefficient associated with GM and WM tissues for the competing segmentation algorithms on different real brain MR scans: (a) GM; (b) WM.

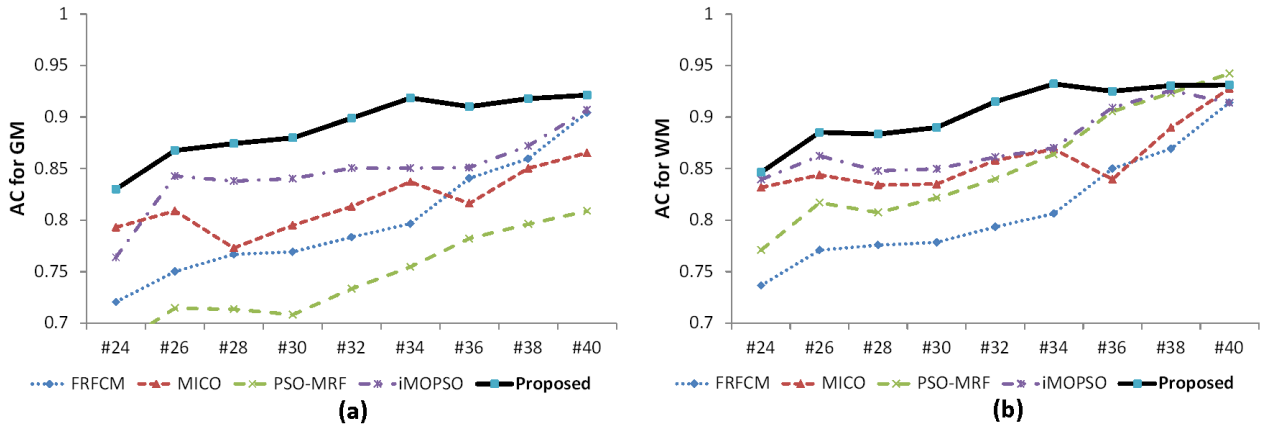

Figure IV.17: Accuracy associated with GM and WM tissues for the competing segmentation algorithms on different real brain MR scans: (a) GM; (b) WM.

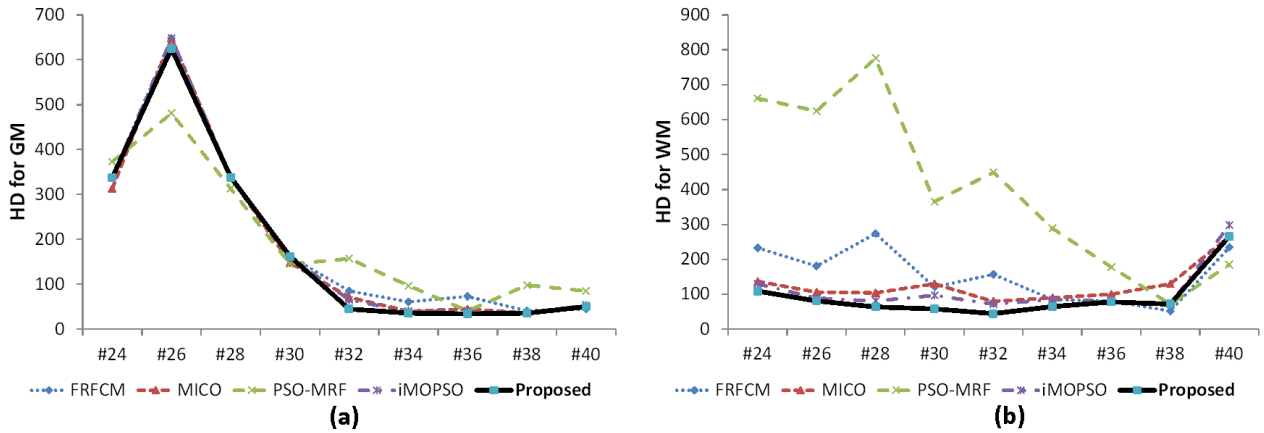

Figure IV.18: Hausdorff distance associated with GM and WM tissues for the competing segmentation algorithms on different real brain MR scans: (a) GM; (b) WM.

Table IV.6: Mean and standard deviation of DICE, AC and HD results for the considered algorithms calculated over the set of 9 real brain MR images from the IBSR repository.

Regions	Metrics	Statistic	Methods				
			FRFCM	MICO	PSO-MRF	iMOPSO	Proposed
GM	DICE	Mean	0.8581	0.8713	0.8205	0.8921	**0.9184**
		Std. Dev.	0.0423	**0.0228**	0.0355	0.0288	0.0240
	AC	Mean	0.7990	0.8169	0.7432	0.8462	**0.8909**
		Std. Dev.	0.0583	**0.0294**	0.0449	0.0376	0.0306
	HD	Mean	199.00	188.00	198.89	191.11	**184.89**
		Std. Dev.	205.61	208.77	**152.50**	211.38	207.96
WM	DICE	Mean	0.8640	0.9014	0.9023	0.9132	**0.9352**
		Std. Dev.	0.0491	0.0270	0.0438	0.0298	**0.0253**
	AC	Mean	0.8105	0.8587	0.8547	0.8755	**0.9044**
		Std. Dev.	0.0563	0.0323	0.0582	0.0322	**0.0298**
	HD	Mean	157.56	126.89	400.00	111.56	**093.00**
		Std. Dev.	078.46	**055.27**	244.74	072.10	066.90

*The values in bold indicate the best performance.

IV.5.6 Computational analysis

The proposed method benefits advantages of both the Markov random field based model and the hybrid metaheuristic algorithm to satisfy requirements of image segmentation problem. However, to solve the problem of selecting appropriate value of β, the core of the proposed method has to be repeated several times (T), hence the total time required to end increases in proportion to (T). To analyse the computational complexity of the proposed algorithm, we calculated the running time when segmenting brain MR images. Both simulated and real brain MR image datasets are involved in our experiments. Since the eventual computational cost will be the multiplication of cost for each for-loop iteration and the number of for-loop iterations (see Algorithm IV.4), the average time cost of each for-loop iteration is recorded. The mean and standard deviation of the time cost (Intel Core i7 1.8 GHz CPU, 8GB RAM and MATLAB 2014b) are listed in Table IV.7.

Table IV.7: Mean ± standard deviation of time cost for each for-loop iteration.

Dimension	Datasets	Number of images	Image size	Time/loop (s)
2D	BrainWeb	9	181 x 217	302.72 ± 45.06
	IBSR	9	135 x 142	90.46 ± 15.43

IV.6 Conclusion

In this chapter, a new method, which takes advantages of the hidden Markov random field approach and searching ability of a hybrid metaheuristic, has been proposed for the segmentation of brain MR images. To achieve satisfactory segmentation results, first, a novel Markov random field model is derived. By using an adaptive weight mechanism and a new potential function, not only balancing contributions of components in the optimization process is achieved but also both spatial and spectral information are utilized to deal with artifacts existing in images. Secondly, a new hybrid metaheuristic optimization algorithm, which is based on two well-known algorithms, Cuckoo search (CS) and Particle swarm optimization (PSO), is also proposed. These two algorithms cooperate together by working on the same population in a parallel way and with a solution selection mechanism. Therefore, this operating scenario can produce better results in terms of quality in a shorter time. Furthermore, to enhance the efficiency of searching solutions, the results from the previous work (the MOPSO-based method) are utilized to provide a starting point for generating initial population. To confirm the effectiveness of the proposed method, it has been examined on both simulated and real MR images, then compared to four recent segmentation methods in the literature. The experimental results show that our method can produce better segmentation results and is able to handle high levels of noise and INU artifact contained in input images. In particular, both qualitative and quantitative results of segmentation show the better performance compared to the competing algorithms. However, in this method, the computational cost is high because of the problem of selecting appropriate value of the parameter β.

Note that the research reported in this chapter gave rise to our publication [Pham et al., 2019a] and a paper under submission [Pham et al., 2019c].

Conclusions and future work

General conclusion

This thesis addresses the problem of automatic segmentation of brain magnetic resonance (MR) images. The objective of the project is to develop new methods combining classical segmentation methods with metaheuristics in order to achieve the desired segmentation performance. This work is stimulated by the following observations: (1) the energy/cost functions of the classical segmentation methods are incorrect in nature since brain MR images usually contain some artifacts such as noise and bias field effect and have complex structures; (2) the need for efficient optimization algorithms is to avoid trapping into local minima since the objective functions are generally non-convex, non-unique in nature and may have several local minimum points; (3) metaheuristics are excellent tools to solve complex problems without having to deeply adapt to each of problem. Therefore, this work attempts to devise and develop brain MR image segmentation methods using metaheuristics that would augment the automatic vision system and thus will facilitate the clinical experts for more accurate diagnosis.

The first proposed method is directed towards developing an unsupervised image segmentation algorithm based on fuzzy clustering approach. The main advantage of using the fuzzy clustering approach for the image segmentation problem lies in its robustness for handling uncertain data, such as region boundaries of cerebral MRI images, which can not be defined in a clear and precise manner. However, two major drawbacks of these methods remain to be overcome: (1) the sensitivity to noise and INU artifact, since no local spatial information in the image is considered; (2) the high vulnerability of the algorithms to trapping into local minima, dependence on the choice of the initial clustering centroids. To overcome these defects, we have developed a new method that has a new cost function taking into account both spatial information and bias field correction, and adapts an improved PSO algorithm to avoid trapping into local minima and dependence on the initialization. The algorithm is found to provide better segmented results compared to a set of recent methods in the literature both qualitatively and quantitatively. However, when high level of noise along with INU artifact is added into MRI data, the performance of the proposed algorithm may decrease.

The second method is an extension of the previous one, since no single criterion is sufficient for dealing with the different properties of the images. In other words, the use of a single objective is only dedicated to exploring a subset of search space and cannot model all the geometric properties of segmentation solution. Hence, we have introduced an additional criterion based on region-based active contour. These two criteria are simultaneously optimized in a multi-objective metaheuristic optimization framework, using a new variant of the MOPSO algorithm. As the output is a set of solutions in which no solution is better than another, we choose the best trade-off solution using a distance-based technique called $\mathcal{L}_2$-metric method. The experimental results show that the proposed method can produce better segmented results, which confirm its robustness against high levels of noise and INU artifact contained in input images compared to a set of recent segmentation methods in the literature. However, when segmenting images with severely inhomogeneous intensity such as real MR brain images in the Internet Brain Segmentation Repository (IBSR) dataset, the proposed method still shows its limitations.

To further improve the quality of segmented images, a third method is proposed. In this method,

an additional criterion, which is based on MRF model for the image segmentation problem, is developed and optimized using a hybrid metaheuristic algorithm. The new model not only takes into account spatial and spectral features of the input image but also balances energy contributions of each segmented region. The hybrid metaheuristic algorithm, which takes into account the advantages of both developed Cuckoo search and particle swarm optimization algorithms, is designed such that fast and accurate solutions can be obtained. To confirm the effectiveness of this method, it has been examined on both simulated and real MR images, then compared to four recent segmentation methods in the literature. The experimental results show that our method can produce better segmentation results and is able to handle high levels of noise and INU artifact contained in input images.

Future work

Although the provided methodologies and results within this dissertation are quite good and constitute a set of powerful tools for image segmentation, it is necessary to mention their perspectives. Here, we briefly describe some interesting research topics that should be further investigated and addressed:

- This work is mainly focused on images produced by structural imaging techniques, MRI modality in particular. It could be interesting to extend the proposed methods to other modalities used in medical imaging, such as Roentgen rays (X-Rays), Ultrasound (US), Computed Tomography (CT) and Positron Emission Tomography (PET).
- In addition, from this work, it can be seen that the more segmentation criteria are optimized, the more and better characteristics of segmented images can be achieved. But how? And in which way? So, it is suggested to go further by examining the performance of combinations of the cost functions from other approaches as well as the capability of other metaheuristic algorithms in order to improve the performance.
- In terms of the correctness of the image segmentation models, optimizing their control parameters could be studied. In addition, incorporating various types of prior information such as appearance prior, shape prior, topological prior, boundary information, etc, into the image segmentation models could be further explored.
- Using different kinds of techniques for post-processing steps in order to improve the overall performance of the proposed methods should be also exploited. Correcting (detecting and reclassifying) potentially misclassified pixels is usually more convenient and simpler than designing a complex method.
- Finally, it would also be interesting to extend the methods for segmentation of a sequence of images, in the context of optimization in dynamic environments, by taking into account the evolution of the objective function over time.

Glossary

clique A clique of a graph G is a fully connected subgraph of G, in which every vertex is a neighbour of all others.

clustering Grouping together image regions or pixels into larger homogeneous regions sharing some property.

compactness A scale, translation and rotation invariant descriptor based on the ratio $\frac{perimeter^2}{area}$.

entropy A measure of the average information content of a random variable X.

feature A numerical property generally used in a classifier.

fuzzy logic A form of logic that allows a range of possibilities between true and false.

gray scale image A monochrome image in which pixels typically represents brightness values ranging from a minimum value (V_{min}) to a maximum value (V_{max}).

ground truth It is the true value, or the most accurate value achievable, of the output of a specific instrument under analysis.

intensity histogram A data structure that records the number of pixels of each intensity value.

kernel function The function in some sense quantifies the similarity of inputs x and y.

labeling problem Given a set S of image structures and a set of labels L, the labeling problem is the question of how to assign a label $l \in L$ for each image structure $s \in S$.

level set The set of data points $\mathbf{x}$ that satisfy a given equation of the form: $f(\mathbf{x}) = c$.

Pareto front Pareto font denotes a set of solutions, none of which performs better than another on all objectives.

posteriori probability The probability $p(s|e)$ that some situation s holds after some evidence e has been observed.

priori probability The probability $p(s)$ corresponding to the probability of the situation s before the observation of any evidence.

segmentation problem The problem of dividing a data set into parts according to a given set of rules.

separability It refers to whether the data is capable of being split into distinct subclasses.

under-segmentation This occurs when one or more segmented regions form the union of many desired regions.

References

[Abdel-Maksoud et al., 2015] Abdel-Maksoud, E., Elmogy, M., and Al-Awadi, R. (2015). Brain tumor segmentation based on a hybrid clustering technique. *Egyptian Informatics Journal*, 16(1):71–81.

[Abed-alguni and Alkhateeb, 2017] Abed-alguni, B. H. and Alkhateeb, F. (2017). Novel selection schemes for Cuckoo search. *Arabian J. Sci. Eng.*, 42(8):3635–3654.

[Adalsteinsson and Sethian, 1995] Adalsteinsson, D. and Sethian, J. A. (1995). A fast level set method for propagating interfaces. *Journal of Computational Physics*, 118.

[Adams and Bischof, 1994] Adams, R. and Bischof, L. (1994). Seeded region growing. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 16(6):641–647.

[Adhikari et al., 2015] Adhikari, S. K., Sing, J. K., Basu, D. K., and Nasipuri, M. (2015). Conditional spatial fuzzy C-means clustering algorithm for segmentation of MRI images. *Applied Soft Computing*, 34:758–769.

[Agrawal et al., 2013] Agrawal, S., Panda, R., Bhuyan, S., and Panigrahi, B. K. (2013). Tsallis entropy based optimal multilevel thresholding using Cuckoo search algorithm. *Swarm and Evolutionary Computation*, 11:16–30.

[Agrawal et al., 2020] Agrawal, S., Samantaray, L., Panda, R., and Dora, L. (2020). A new hybrid adaptive Cuckoo search-squirrel search algorithm for brain MR image analysis. In *Hybrid Machine Intelligence for Medical Image Analysis*, pages 85–117. Springer.

[Aguilera et al., 2012] Aguilera, C. A., Ramos, M. A., and Sappa, A. D. (2012). Simulated annealing: A novel application of image processing in the wood area. In *Simulated Annealing-Advances, Applications and Hybridizations*. IntechOpen.

[Ahmadvand et al., 2017] Ahmadvand, A., Yousefi, S., and Manzuri Shalmani, M. (2017). A novel Markov random field model based on region adjacency graph for T1 magnetic resonance imaging brain segmentation. *International Journal of Imaging Systems and Technology*, 27(1):78–88.

[Ahmed et al., 2002] Ahmed, M. N., Yamany, S. M., Mohamed, N., Farag, A. A., and Moriarty, T. (2002). A modified fuzzy C-means algorithm for bias field estimation and segmentation of MRI data. *IEEE Transactions on Medical Imaging*, 21(3):193–199.

[Al-Faris et al., 2014] Al-Faris, A. Q., Ngah, U. K., Isa, N. A. M., and Shuaib, I. L. (2014). Breast MRI tumour segmentation using modified automatic seeded region growing based on particle swarm optimization image clustering. In *Soft Computing in Industrial Applications*, pages 49–60. Springer.

[Alam et al., 2014] Alam, S., Dobbie, G., Koh, Y. S., Riddle, P., and Rehman, S. U. (2014). Research on particle swarm optimization based clustering: a systematic review of literature and techniques. *Swarm and Evolutionary Computation*, 17:1–13.

[AlRashidi and El-Hawary, 2008] AlRashidi, M. R. and El-Hawary, M. E. (2008). A survey of particle swarm optimization applications in electric power systems. *IEEE Transactions on Evolutionary Computation*, 13(4):913–918.

[Alsmadi, 2014] Alsmadi, M. K. (2014). A hybrid firefly algorithm with fuzzy C-mean algorithm for mri brain segmentation. *American Journal of Applied Sciences*, 11.

[Bandyopadhyay et al., 2008] Bandyopadhyay, S., Saha, S., Maulik, U., and Deb, K. (2008). A simulated annealing-based multiobjective optimization algorithm: AMOSA. *IEEE Transactions on Evolutionary Computation*, 12(3):269–283.

[Banks et al., 2007] Banks, A., Vincent, J., and Anyakoha, C. (2007). A review of particle swarm optimization. Part I: background and development. *Natural Computing*, 6(4):467–484.

[Banks et al., 2008] Banks, A., Vincent, J., and Anyakoha, C. (2008). A review of particle swarm optimization. Part II: hybridisation, combinatorial, multicriteria and constrained optimization, and indicative applications. *Natural Computing*, 7(1):109–124.

[Bansal et al., 2019] Bansal, J. C., Singh, P. K., and Pal, N. R. (2019). *Evolutionary and swarm intelligence algorithms*. Springer.

[Battiti and Tecchiolli, 1996] Battiti, R. and Tecchiolli, G. (1996). The continuous reactive Tabu search: blending combinatorial optimization and stochastic search for global optimization. *Annals of Operations Research*, 63(2):151–188.

[Beauchemin et al., 1998] Beauchemin, M., Thomson, K. P., and Edwards, G. (1998). On the Hausdorff distance used for the evaluation of segmentation results. *Canadian Journal of Remote Sensing*, 24(1):3–8.

[Behnamian et al., 2009] Behnamian, J., Ghomi, S. F., and Zandieh, M. (2009). A multi-phase covering Pareto-optimal front method to multi-objective scheduling in a realistic hybrid flowshop using a hybrid metaheuristic. *Expert Systems with Applications*, 36(8):11057–11069.

[Ben George et al., 2015] Ben George, E., Rosline, G. J., and Rajesh, D. G. (2015). Brain tumor segmentation using Cuckoo search optimization for magnetic resonance images. In *2015 IEEE 8th GCC Conference Exhibition*, pages 1–6.

[Benaichouche et al., 2016] Benaichouche, A. N., Oulhadj, H., and Siarry, P. (2016). Multiobjective improved spatial fuzzy C-means clustering for image segmentation combining Pareto-optimal clusters. *J Heuristics*, 22(4):383–404.

[Besag, 1974] Besag, J. (1974). Spatial interaction and the statistical analysis of lattice systems. *J R Stat Soc Series B Stat Methodol*, 36(2):192–225.

[Besag, 1986] Besag, J. (1986). On the statistical analysis of dirty pictures. *J R Stat Soc Series B Stat Methodol*, 48(3):259–279.

[Beucher and Meyer, 1993] Beucher, S. and Meyer, F. (1993). The morphological approach to segmentation: the watershed transformation. *Mathematical Morphology in Image Processing*, 34:433–481.

[Bezdek, 1981] Bezdek, J. C. (1981). *Pattern recognition with fuzzy objective function algorithms*. Springer.

[Bhakat and Periannan, 2019] Bhakat, S. and Periannan, S. (2019). Brain tumor detection using Cuckoo search algorithm and histogram thresholding for MR images. In *Smart Innovations in Communication and Computational Sciences*, pages 85–95. Springer.

[Bhandari, 2018] Bhandari, A. K. (2018). A novel beta differential evolution algorithm-based fast multilevel thresholding for color image segmentation. *Neural Computing and Applications*, pages 1–31.

[Bhandari et al., 2015] Bhandari, A. K., Kumar, A., and Singh, G. K. (2015). Tsallis entropy based multilevel thresholding for colored satellite image segmentation using evolutionary algorithms. *Expert Systems with Applications*, 42(22):8707–8730.

[Bhandari et al., 2014] Bhandari, A. K., Singh, V. K., Kumar, A., and Singh, G. K. (2014). Cuckoo search algorithm and wind driven optimization based study of satellite image segmentation for multilevel thresholding using Kapurs entropy. *Expert Systems with Applications*, 41(7):3538–3560.

[Bhandarkar and Zhang, 1999] Bhandarkar, S. M. and Zhang, H. (1999). Image segmentation using evolutionary computation. *IEEE Transactions on Evolutionary Computation*, 3(1):1–21.

[Birattari et al., 2001] Birattari, M., Paquete, L., Strutzle, T., and Varrentrapp, K. (2001). Classification of metaheuristics and design of experiments for the analysis of components.

[Blum, 2005] Blum, C. (2005). Ant colony optimization: Introduction and recent trends. *Physics of Life Reviews*, 2(4):353–373.

[Blum and Raidl, 2016] Blum, C. and Raidl, G. R. (2016). *Hybrid Metaheuristics: Powerful Tools for Optimization*. Springer.

[Blum and Roli, 2003] Blum, C. and Roli, A. (2003). Metaheuristics in combinatorial optimization: Overview and conceptual comparison. *ACM Computing Surveys (CSUR)*, 35(3):268–308.

[Blum et al., 2008] Blum, C., Roli, A., and Sampels, M. (2008). *Hybrid metaheuristics: an emerging approach to optimization*, volume 114. Springer.

[Bonabeau et al., 1999] Bonabeau, E., Marco, D. d. R. D. F., Dorigo, M., Theraulaz, G., et al. (1999). *Swarm intelligence: from natural to artificial systems*. Number 1. Oxford University Press.

[Borjigin and Sahoo, 2019] Borjigin, S. and Sahoo, P. K. (2019). Color image segmentation based on multi-level Tsallis–Havrda–Charvát entropy and 2D histogram using PSO algorithms. *Pattern Recognition*, 92:107–118.

[BoussaïD et al., 2013] BoussaïD, I., Lepagnot, J., and Siarry, P. (2013). A survey on optimization metaheuristics. *Information sciences*, 237:82–117.

[Bouyer and Hatamlou, 2018] Bouyer, A. and Hatamlou, A. (2018). An efficient hybrid clustering method based on improved Cuckoo optimization and modified particle swarm optimization algorithms. *Appl Soft Comput*, 67:172 – 182.

[Branke and Mostaghim, 2006] Branke, J. and Mostaghim, S. (2006). About selecting the personal best in multi-objective particle swarm optimization. In *Proc PPSN IX*, pages 523–532. Springer.

[Capelle et al., 2003] Capelle, A.-S., Colot, O., and Fernandez-Maloigne, C. (2003). 3D segmentation of MR brain images into white matter, gray matter and cerebrospinal fluid by means of evidence theory. In *Conference on Artificial Intelligence in Medicine in Europe*, pages 112–116. Springer.

[Caselles, 1995] Caselles, V. (1995). Geometric models for active contours. In *Proceedings of Comput. Soc. Press International Conference on Image Processing*, volume 3. IEEE.

[Caselles, 2000] Caselles, V. (2000). Automatic parameter setting for balloon models. In *SPIE Proceedings of Medical Imaging 2000 - San Diego, CA*, volume 3979.

[Černỳ, 1985] Černỳ, V. (1985). Thermodynamical approach to the traveling salesman problem: An efficient simulation algorithm. *Journal of Optimization Theory and Applications*, 45(1):41–51.

[Chabrier et al., 2006] Chabrier, S., Emile, B., Rosenberger, C., and Laurent, H. (2006). Unsupervised performance evaluation of image segmentation. *EURASIP Journal on Applied Signal Processing*, 2006:217–217.

[Chaoshun Li and Xiao, 2012] Chaoshun Li, Jianzhong Zhou, P. K. and Xiao, J. (2012). A novel chaotic particle swarm optimization based fuzzy clustering algorithm. *Neurocomputing*, 83.

[Cheevasuvit et al., 1986] Cheevasuvit, F., Maitre, H., and Vidal-Madjar, D. (1986). A robust method for picture segmentation based on a split-and-merge procedure. *Computer Vision, Graphics, and Image Processing*, 34(3):268–281.

[Chelouah and Siarry, 2000] Chelouah, R. and Siarry, P. (2000). Tabu search applied to global optimization. *European Journal of Operational Research*, 123(2):256–270.

[Chen et al., 2011] Chen, L., Chen, C. P., and Lu, M. (2011). A multiple-kernel fuzzy C-means algorithm for image segmentation. *IEEE Transactions on Systems, Man, and Cybernetics, Part B (Cybernetics)*, 41(5):1263–1274.

[Chen et al., 2017a] Chen, M., Yan, Q., and Qin, M. (2017a). A segmentation of brain MRI images utilizing intensity and contextual information by Markov random field. *Computer Assisted Surgery*, 22(sup1):200–211.

[Chen and Zhang, 2004] Chen, S. and Zhang, D. (2004). Robust image segmentation using FCM with spatial constraints based on new kernel-induced distance measure. *IEEE Transactions on Systems, Man, and Cybernetics, Part B (Cybernetics)*, 34(4):1907–1916.

[Chen et al., 1991] Chen, S.-Y., Lin, W.-C., and Chen, C.-T. (1991). Split-and-merge image segmentation based on localized feature analysis and statistical tests. *CVGIP: Graphical Models and Image Processing*, 53(5):457–475.

[Chen et al., 2017b] Chen, Y., Li, L., Peng, H., Xiao, J., Yang, Y., and Shi, Y. (2017b). Particle swarm optimizer with two differential mutation. *Appl Soft Comput*, 61:314 – 330.

[Chibante, 2010] Chibante, R. (2010). *Simulated Annealing: Theory with Applications*. BoD–Books on Demand.

[Chih-Yu Hsu and Chen, 2008] Chih-Yu Hsu, C.-Y. L. and Chen, C.-M. (2008). Automatic segmentation of liver PET images. *Computerized Medical Imaging and Graphics*, 32.

[Chiroma et al., 2017] Chiroma, H., Herawan, T., Fister Jr, I., Fister, I., Abdulkareem, S., Shuib, L., Hamza, M. F., Saadi, Y., and Abubakar, A. (2017). Bio-inspired computation: Recent development on the modifications of the Cuckoo search algorithm. *Applied Soft Computing*, 61:149–173.

[Chua et al., 2009] Chua, Z. Y., Zheng, W., Chee, M. W., and Zagorodnov, V. (2009). Evaluation of performance metrics for bias field correction in MR brain images. *JMRI-J Magn Reson Im*, 29(6):1271–1279.

[Chuang et al., 2006] Chuang, K.-S., Tzeng, H.-L., Chen, S., Wu, J., and Chen, T.-J. (2006). Fuzzy C-means clustering with spatial information for image segmentation. *Computerized Medical Imaging and Graphics*, 30(1):9–15.

[Coello et al., 2006] Coello, C., De Computación, S., and Zacatenco, C. (2006). Twenty years of evolutionary multi-objective optimization: A historical view of the field. *IEEE Computational Intelligence Magazine*, 1(1):28–36.

[Coello, 2002] Coello, C. A. C. (2002). Theoretical and numerical constraint-handling techniques used with evolutionary algorithms: a survey of the state of the art. *Computer Methods in Applied Mechanics and Engineering*, 191(11-12):1245–1287.

[Coello et al., 2007] Coello, C. A. C., Lamont, G. B., Van Veldhuizen, D. A., et al. (2007). *Evolutionary algorithms for solving multi-objective problems*, volume 5. Springer.

[Coello and Lechuga, 2002] Coello, C. C. and Lechuga, M. S. (2002). MOPSO: A proposal for multiple objective particle swarm optimization. In *Proceedings of the 2002 Congress on Evolutionary Computation. CEC'02 (Cat. No. 02TH8600)*, volume 2, pages 1051–1056. IEEE.

[Cook et al., 1996] Cook, R., McConnell, I., Stewart, D., and Oliver, C. J. (1996). Segmentation and simulated annealing. In *Microwave Sensing and Synthetic Aperture Radar*, volume 2958, pages 30–37. International Society for Optics and Photonics.

[Covavisaruch and Tanatipanond, 1999] Covavisaruch, N. and Tanatipanond, T. (1999). Deformable contour for brain MR images by genetic algorithm: from rigid to training approaches. In *Proceedings, Image and Vision Computing New Zealand (IVCNZ'99)*.

[Cover, 1965] Cover, T. M. (1965). Geometrical and statistical properties of systems of linear inequalities with applications in pattern recognition. *IEEE Transactions on Electronic Computers*, (3):326–334.

[Cuadra et al., 2005] Cuadra, M. B., Cammoun, L., Butz, T., Cuisenaire, O., and Thiran, J. . (2005). Comparison and validation of tissue modelization and statistical classification methods in T1-weighted MR brain images. *IEEE Trans. Med. Imag.*, 24(12):1548–1565.

[Cufi et al., 2003] Cufi, X., Munoz, X., Freixenet, J., and Marti, J. (2003). A review of image segmentation techniques integrating region and boundary information. *Advances in Imaging and Electron Physics*, 120:1–39.

[Cvijović and Klinowski, 1995] Cvijović, D. and Klinowski, J. (1995). Taboo search: an approach to the multiple minima problem. *Science*, 267(5198):664–666.

[De Albuquerque et al., 2004] De Albuquerque, M. P., Esquef, I. A., and Mello, A. G. (2004). Image thresholding using Tsallis entropy. *Pattern Recognition Letters*, 25(9):1059–1065.

[de Paiva et al., 2016] de Paiva, J. L., Toledo, C. F., and Pedrini, H. (2016). An approach based on hybrid genetic algorithm applied to image denoising problem. *Applied Soft Computing*, 46:778–791.

[Deb, 2001] Deb, K. (2001). *Multi-objective optimization using evolutionary algorithms*, volume 16. John Wiley & Sons.

[Deb et al., 2002] Deb, K., Pratap, A., Agarwal, S., and Meyarivan, T. (2002). A fast and elitist multiobjective genetic algorithm: NSGA-II. *IEEE Transactions on Evolutionary Computation*, 6(2):182–197.

[Dehshibi et al., 2017] Dehshibi, M. M., Sourizaei, M., Fazlali, M., Talaee, O., Samadyar, H., and Shanbehzadeh, J. (2017). A hybrid bio-inspired learning algorithm for image segmentation using multilevel thresholding. *Multimedia Tools and Applications*, 76(14):15951–15986.

[Del Valle et al., 2008] Del Valle, Y., Venayagamoorthy, G. K., Mohagheghi, S., Hernandez, J.-C., and Harley, R. G. (2008). Particle swarm optimization: basic concepts, variants and applications in power systems. *IEEE Transactions on Evolutionary Computation*, 12(2):171–195.

[Deng and Clausi, 2004] Deng, H. and Clausi, D. A. (2004). Unsupervised image segmentation using a simple MRF model with a new implementation scheme. *Pattern Recognit*, 37(12):2323–2335.

[Dice, 1945] Dice, L. R. (1945). Measures of the amount of ecologic association between species. *Ecology*, 26(3):297–302.

[Ding and Fu, 2015] Ding, Y. and Fu, X. (2015). Kernel-based fuzzy C-means clustering algorithm based on genetic algorithm. *Neurocomputing*.

[Djemame and Batouche, 2018] Djemame, S. and Batouche, M. (2018). A hybrid metaheuristic algorithm based on quantum genetic computing for image segmentation. In *Hybrid Metaheuristics for Image Analysis*, pages 33–48. Springer.

[Dorigo, 1992] Dorigo, M. (1992). Optimization, learning and natural algorithms. *PhD Thesis, Politecnico di Milano*.

[Dorigo and Birattari, 2010] Dorigo, M. and Birattari, M. (2010). *Ant colony optimization*. Springer.

[Drigo, 1996] Drigo, M. (1996). The ant system: optimization by a colony of cooperating agents. *IEEE Transactions on Systems, Man, and Cybernetics-Part B*, 26(1):1–13.

[Du and Swamy, 2016] Du, K.-L. and Swamy, M. (2016). Search and optimization by metaheuristics. *Techniques and Algorithms Inspired by Nature; Birkhauser: Basel, Switzerland*.

[Dunn, 1973] Dunn, J. C. (1973). A fuzzy relative of the ISODATA process and its use in detecting compact well-separated clusters.

[Eberhart and Kennedy, 1995] Eberhart, R. and Kennedy, J. (1995). Particle swarm optimization. In *Proceedings of the IEEE International Conference on Neural networks*, volume 4, pages 1942–1948. Citeseer.

[Eiben et al., 2003] Eiben, A. E., Smith, J. E., et al. (2003). *Introduction to evolutionary computing*, volume 53. Springer.

[El Aziz et al., 2018] El Aziz, M. A., Ewees, A. A., Hassanien, A. E., Mudhsh, M., and Xiong, S. (2018). Multi-objective Whale optimization algorithm for multilevel thresholding segmentation. In *Advances in Soft Computing and Machine Learning in Image Processing*, pages 23–39. Springer.

[Elazab et al., 2015] Elazab, A., Wang, C., Jia, F., Wu, J., Li, G., and Hu, Q. (2015). Segmentation of brain tissues from magnetic resonance images using adaptively regularized kernel-based fuzzy C-means clustering. *Comput Math Method M*, 2015:12 pages.

[Elaziz et al., 2019] Elaziz, M. A., Oliva, D., Ewees, A. A., and Xiong, S. (2019). Multi-level thresholding-based grey scale image segmentation using multi-objective multi-verse optimizer. *Expert Systems with Applications*, 125:112–129.

[Elbes et al., 2019] Elbes, M., Alzubi, S., Kanan, T., Al-Fuqaha, A., and Hawashin, B. (2019). A survey on particle swarm optimization with emphasis on engineering and network applications. *Evolutionary Intelligence*, pages 1–17.

[Esmin et al., 2015] Esmin, A. A., Coelho, R. A., and Matwin, S. (2015). A review on particle swarm optimization algorithm and its variants to clustering high-dimensional data. *Artificial Intelligence Review*, 44(1):23–45.

[Ewees et al., 2018] Ewees, A. A., Elaziz, M. A., and Oliva, D. (2018). Image segmentation via multilevel thresholding using hybrid optimization algorithms. *Journal of Electronic Imaging*, 27(6):063008.

[Fan and Lin, 2007] Fan, S.-K. S. and Lin, Y. (2007). A multi-level thresholding approach using a hybrid optimal estimation algorithm. *Pattern Recognition Letters*, 28(5):662–669.

[Fister Jr et al., 2013] Fister Jr, I., Fister, D., and Fister, I. (2013). A comprehensive review of Cuckoo search: variants and hybrids. *International Journal of Mathematical Modelling and Numerical Optimisation*, 4(4):387–409.

[Foster et al., 2014] Foster, B., Bagci, U., Mansoor, A., Xu, Z., and Mollura, D. J. (2014). A review on segmentation of positron emission tomography images. *Computers in Biology and Medicine*, 50:76–96.

[Freitas, 2003] Freitas, A. A. (2003). A survey of evolutionary algorithms for data mining and knowledge discovery. In *Advances in Evolutionary Computing*, pages 819–845. Springer.

[Freixenet et al., 2002] Freixenet, J., Muñoz, X., Raba, D., Martí, J., and Cufí, X. (2002). Yet another survey on image segmentation: Region and boundary information integration. In *European Conference on Computer Vision*, pages 408–422. Springer.

[García et al., 2014] García, M. L. L., García-Ródenas, R., and Gómez, A. G. (2014). Hybrid metaheuristic optimization algorithms for time-domain-constrained data clustering. *Applied Soft Computing*, 23:319–332.

[Geman and Geman, 1987] Geman, S. and Geman, D. (1987). Stochastic relaxation, Gibbs distributions, and the Bayesian restoration of images. In *Readings in Computer Vision: Issues, Problem, Principles, and Paradigms*, pages 564 – 584. Morgan Kaufmann.

[Gendreau, 2003] Gendreau, M. (2003). An introduction to Tabu search. In *Handbook of metaheuristics*, pages 37–54. Springer.

[Gendreau et al., 2006] Gendreau, M., Guertin, F., Potvin, J.-Y., and Séguin, R. (2006). Neighborhood search heuristics for a dynamic vehicle dispatching problem with pick-ups and deliveries. *Transportation Research Part C: Emerging Technologies*, 14(3):157–174.

[Gendreau et al., 2010] Gendreau, M., Potvin, J.-Y., et al. (2010). *Handbook of metaheuristics*, volume 2. Springer.

[Gendreau et al., 2018] Gendreau, M., Potvin, J.-Y., et al. (2018). *Handbook of metaheuristics*, volume 272, chapter Simulated Annealing: From Basics to Applications. Springer.

[Ghamisi et al., 2013] Ghamisi, P., Couceiro, M. S., Martins, F. M., and Benediktsson, J. A. (2013). Multilevel image segmentation based on fractional-order Darwinian particle swarm optimization. *IEEE Transactions on Geoscience and Remote sensing*, 52(5):2382–2394.

[Ghosh et al., 2010] Ghosh, P., Mitchell, M., Tanyi, J. A., and Hung, A. (2010). A genetic algorithm-based level set curve evolution for prostate segmentation on pelvic CT and MRI images. In *Biomedical Image Analysis and Machine Learning Technologies: Applications and Techniques*, pages 127–149. IGI Global.

[Glover, 1986] Glover, F. (1986). Future paths for integer programming and links to artificial intelligence. *Computers and Operations Research*, 13(5):533–549.

[Gong et al., 2017] Gong, C., Chen, H., He, W., and Zhang, Z. (2017). Improved multi-objective clustering algorithm using particle swarm optimization. *PloS one*, 12(12):e0188815.

[Graves and Pedrycz, 2010] Graves, D. and Pedrycz, W. (2010). Kernel-based fuzzy clustering and fuzzy clustering: A comparative experimental study. *Fuzzy sets and Systems*, 161(4):522–543.

[Guerrout et al., 2016] Guerrout, E., Ait-Aoudia, S., Michelucci, D., and Mahiou, R. (2016). Hidden Markov random fields and direct search methods for medical image segmentation. In *ICPRAM*, pages 154–161.

[Guerrout et al., 2018] Guerrout, E.-H., Ait-Aoudia, S., Michelucci, D., and Mahiou, R. (2018). Hidden Markov random field model and Broyden–Fletcher–Goldfarb–Shanno algorithm for brain image segmentation. *J Exp Theor Artif Intell*, 30(3):415–427.

[Guerrout et al., 2017] Guerrout, E.-H., Mahiou, R., and Ait-Aoudia, S. (2017). Hidden Markov random fields and particle swarm combination for brain magnetic resonance image segmentation. *Int Arab J Inf Techn*, 15(3).

[Gupta et al., 2015] Gupta, D., Anand, R., and Tyagi, B. (2015). A hybrid segmentation method based on Gaussian kernel fuzzy clustering and region based active contour model for ultrasound medical images. *Biomedical Signal Processing and Control*, 16:98–112.

[Hajek, 1988] Hajek, B. (1988). Cooling schedules for optimal annealing. *Mathematics of Operations Research*, 13(2):311–329.

[Hammouche et al., 2008] Hammouche, K., Diaf, M., and Siarry, P. (2008). A multilevel automatic thresholding method based on a genetic algorithm for a fast image segmentation. *Computer Vision and Image Understanding*, 109(2):163–175.

[Han and Shi, 2007a] Han, Y. and Shi, P. (2007a). An improved ant colony algorithm for fuzzy clustering in image segmentation. *Neurocomputing*, 70(4-6):665–671.

[Han and Shi, 2007b] Han, Y. and Shi, P. (2007b). An improved ant colony algorithm for fuzzy clustering in image segmentation. *Neurocomputing*, 70.

[Harrison et al., 2016] Harrison, K. R., Engelbrecht, A. P., and Ombuki-Berman, B. M. (2016). Inertia weight control strategies for particle swarm optimization. *Swarm Intelligence*, 10(4):267–305.

[Héliodore et al., 2017] Héliodore, F., Nakib, A., Ismail, B., Ouchraa, S., and Schmitt, L. (2017). *Metaheuristics for intelligent electrical networks*, chapter Performance Evaluation of Metaheuristics. Wiley Online Library.

[Hiralal and Menon, 2016] Hiralal, R. and Menon, H. P. (2016). A survey of brain MRI image segmentation methods and the issues involved. In *The International Symposium on Intelligent Systems Technologies and Applications*, pages 245–259. Springer.

[Ho and Lee, 2003] Ho, S.-Y. and Lee, K.-Z. (2003). Design and analysis of an efficient evolutionary image segmentation algorithm. *Journal of VLSI Signal Processing Systems for Signal, Image and Video Technology*, 35(1):29–42.

[Horowitz, 1974] Horowitz, S. L. (1974). Picture segmentation by a directed split-and-merge procedure. In *IJCPR*, pages 424–433.

[Hoseini and Shayesteh, 2013] Hoseini, P. and Shayesteh, M. G. (2013). Efficient contrast enhancement of images using hybrid ant colony optimisation, genetic algorithm, and simulated annealing. *Digital Signal Processing*, 23(3):879–893.

[Huang and Zeng, 2015] Huang, C. and Zeng, L. (2015). An active contour model for the segmentation of images with intensity inhomogeneities and bias field estimation. *PloS one*, 10(4):e0120399.

[Huang et al., 2008] Huang, P., Cao, H., and Luo, S. (2008). An artificial ant colonies approach to medical image segmentation. *Computer Methods and Programs in Biomedicine*, 92(3):267–273.

[Huawu Deng and Clausi, 2005] Huawu Deng and Clausi, D. A. (2005). Unsupervised segmentation of synthetic aperture radar sea ice imagery using a novel Markov random field model. *IEEE Trans. Geosci. Remote Sens*, 43(3):528–538.

[Hussain et al., 2018] Hussain, K., Salleh, M. N. M., Cheng, S., and Shi, Y. (2018). Metaheuristic research: a comprehensive survey. *Artificial Intelligence Review*, pages 1–43.

[Ilunga-Mbuyamba et al., 2016] Ilunga-Mbuyamba, E., Cruz-Duarte, J. M., Avina-Cervantes, J. G., Correa-Cely, C. R., Lindner, D., and Chalopin, C. (2016). Active contours driven by Cuckoo search strategy for brain tumour images segmentation. *Expert Systems with Applications*, 56:59–68.

[Indrajit Saha and Plewczynski, 2011] Indrajit Saha, U. M. and Plewczynski, D. (2011). A new multi-objective technique for differential fuzzy clustering. *Applied Soft Computing*, 11.

[Izakian and Abraham, 2011] Izakian, H. and Abraham, A. (2011). Fuzzy C-means and fuzzy swarm for fuzzy clustering problem. *Expert Systems with Applications*, 38.

[Jaccard, 1912] Jaccard, P. (1912). The distribution of the flora in the alpine zone. *New Phytologist*, 11(2):37–50.

[Jaeggi et al., 2008] Jaeggi, D. M., Parks, G. T., Kipouros, T., and Clarkson, P. J. (2008). The development of a multi-objective Tabu search algorithm for continuous optimisation problems. *European Journal of Operational Research*, 185(3):1192–1212.

[Jia et al., 2019] Jia, H., Lang, C., Oliva, D., Song, W., and Peng, X. (2019). Hybrid grasshopper optimization algorithm and differential evolution for multilevel satellite image segmentation. *Remote Sensing*, 11(9):1134.

[Jiang and Yang, 2002] Jiang, T. and Yang, F. (2002). An evolutionary Tabu search for cell image segmentation. *IEEE Transactions on Systems, Man, and Cybernetics, Part B (Cybernetics)*, 32(5):675–678.

[Jiang et al., 2012] Jiang, Y., Tsai, P., Hao, Z., and Cao, L. (2012). A novel auto-parameters selection process for image segmentation. In *2012 IEEE Congress on Evolutionary Computation*, pages 1–7. IEEE.

[Jiang et al., 2015] Jiang, Y., Tsai, P., Hao, Z., and Cao, L. (2015). Automatic multilevel thresholding for image segmentation using stratified sampling and Tabu search. *Soft Computing*, 19(9):2605–2617.

[Jin and Han, 2010a] Jin, X. and Han, J. (2010a). Expectation maximization clustering. *Encyclopedia of Machine Learning*, pages 382–383.

[Jin and Han, 2010b] Jin, X. and Han, J. (2010b). Partitional clustering. *Encyclopedia of Machine Learning*, pages 766–766.

[Johnson, 1967] Johnson, S. C. (1967). Hierarchical clustering schemes. *Psychometrika*, 32(3):241–254.

[Kan et al., 2007] Kan, J., Li, H., and Li, W. (2007). Level set method in standing tree image segmentation based on particle swarm optimization. In *MIPPR 2007: Automatic Target Recognition and Image Analysis; and Multispectral Image Acquisition*, volume 6786, page 67864V. International Society for Optics and Photonics.

[Kapur et al., 1985] Kapur, J. N., Sahoo, P. K., and Wong, A. K. (1985). A new method for gray-level picture thresholding using the entropy of the histogram. *Computer Vision, Graphics, and Image Processing*, 29(3):273–285.

[Karasulu and Korukoglu, 2011] Karasulu, B. and Korukoglu, S. (2011). A simulated annealing-based optimal threshold determining method in edge-based segmentation of grayscale images. *Applied Soft Computing*, 11(2):2246–2259.

[Kato et al., 1999] Kato, Z., Zerubia, J., and Berthod, M. (1999). Unsupervised parallel image classification using Markovian models. *Pattern Recognit*, 32(4):591 – 604.

[Katsigiannis et al., 2012] Katsigiannis, Y. A., Georgilakis, P. S., and Karapidakis, E. S. (2012). Hybrid simulated annealing–Tabu search method for optimal sizing of autonomous power systems with renewables. *IEEE Transactions on Sustainable Energy*, 3(3):330–338.

[Kaur et al., 2012] Kaur, J., Agrawal, S., and Vig, R. (2012). Integration of clustering, optimization and partial differential equation method for improved image segmentation. *International Journal of Image, Graphics & Signal Processing*, 4(11).

[Kaur and Mahajan, 2018] Kaur, S. and Mahajan, R. (2018). Hybrid metaheuristic optimization based energy efficient protocol for wireless sensor networks. *Egyptian Informatics Journal*, 19(3):145–150.

[Kennedy and Eberhart, 1995] Kennedy, J. and Eberhart, R. (1995). Particle swarm optimization. In *Proceedings of IEEE International Conference on Neural Networks IV*, volume 1000, pages 1942–1948.

[Khorram and Yazdi, 2019] Khorram, B. and Yazdi, M. (2019). A new optimized thresholding method using ant colony algorithm for MR brain image segmentation. *Journal of Digital Imaging*, 32(1):162–174.

[Kim and Park, 1998] Kim, B.-H. and Park, R.-H. (1998). Multi-image photometric stereo using surface approximation by Legendre polynomials. *Pattern recognition*, 31(8):1033–1047.

[Kirkpatrick et al., 1983] Kirkpatrick, S., Gelatt, C. D., and Vecchi, M. P. (1983). Optimization by simulated annealing. *Science*, 220(4598):671–680.

[Knowles and Corne, 2000] Knowles, J. D. and Corne, D. W. (2000). Approximating the non-dominated front using the Pareto archived evolution strategy. *Evolutionary Computation*, 8(2):149–172.

[Krinidis and Chatzis, 2010] Krinidis, S. and Chatzis, V. (2010). A robust fuzzy local information C-means clustering algorithm. *IEEE Transactions on Image Processing*, 19(5):1328–1337.

[Kulkarni and Venayagamoorthy, 2010] Kulkarni, R. V. and Venayagamoorthy, G. K. (2010). Particle swarm optimization in wireless-sensor networks: A brief survey. *IEEE Transactions on Systems, Man, and Cybernetics, Part C (Applications and Reviews)*, 41(2):262–267.

[Kumar and Nachamai, 2012] Kumar, N. and Nachamai, M. (2012). Noise removal and filtering techniques used in medical images. *Indian Journal of Computer Science and Engineering*, 3(1):146–153.

[Kwan et al., 1999] Kwan, R.-S., Evans, A. C., and Pike, G. B. (1999). MRI simulation-based evaluation of image-processing and classification methods. *IEEE Transactions on Medical Imaging*, 18(11):1085–1097.

[Lang and Jia, 2019] Lang, C. and Jia, H. (2019). Kapur entropy for color image segmentation based on a hybrid Whale optimization algorithm. *Entropy*, 21(3):318.

[Lankton and Tannenbaum, 2008] Lankton, S. and Tannenbaum, A. (2008). Localizing region-based active contours. *IEEE T Image Process*, 17(11):2029–2039.

[Law et al., 2008] Law, Y. N., Lee, H. K., and Yip, A. M. (2008). A multiresolution stochastic level set method for Mumford-Shah image segmentation. *IEEE Transactions on Image Processing*, 17(12):2289–2300.

[Lee et al., 2003] Lee, J.-M., Yoon, U., Nam, S. H., Kim, J.-H., Kim, I.-Y., and Kim, S. I. (2003). Evaluation of automated and semi-automated skull-stripping algorithms using similarity index and segmentation error. *Computers in Biology and Medicine*, 33(6):495–507.

[Lei et al., 2018] Lei, T., Jia, X., Zhang, Y., He, L., Meng, H., and Nandi, A. K. (2018). Significantly fast and robust fuzzy C-means clustering algorithm based on morphological reconstruction and membership filtering. *IEEE T Fuzzy Syst*, 26(5):3027–3041.

[Li et al., 2011a] Li, B. N., Chui, C. K., Chang, S., and Ong, S. H. (2011a). Integrating spatial fuzzy clustering with level set methods for automated medical image segmentation. *Computers in Biology and Medicine*, 41(1):1–10.

[Li et al., 2014] Li, C., Gore, J. C., and Davatzikos, C. (2014). Multiplicative intrinsic component optimization (MICO) for MRI bias field estimation and tissue segmentation. *Magnetic Resonance Imaging*, 32(7):913–923.

[Li et al., 2011b] Li, C., Huang, R., Ding, Z., Gatenby, J. C., Metaxas, D. N., and Gore, J. C. (2011b). A level set method for image segmentation in the presence of intensity inhomogeneities with application to MRI. *IEEE T Image Process*, 20(7):2007–2016.

[Li et al., 2008] Li, C., Kao, C.-Y., Gore, J. C., and Ding, Z. (2008). Minimization of region-scalable fitting energy for image segmentation. *IEEE T Image Process*, 17(10):1940–1949.

[Li et al., 2005] Li, C., Xu, C., Gui, C., and Fox, M. D. (2005). Level set evolution without re-initialization: a new variational formulation. In *Proc IEEE Conf Comput Vis Pattern Recognit (CVPR)*, volume 1, pages 430–436. IEEE.

[Li and Lee, 1993] Li, C. H. and Lee, C. (1993). Minimum cross entropy thresholding. *Pattern Recognition*, 26(4):617–625.

[Li et al., 2015] Li, H., He, H., and Wen, Y. (2015). Dynamic particle swarm optimization and K-means clustering algorithm for image segmentation. *Optik*, 126(24):4817–4822.

[Li and Li, 2008] Li, L. and Li, D. (2008). Fuzzy entropy image segmentation based on particle swarm optimization. *Progress in Natural Science*, 18(9):1167–1171.

[Li, 2001] Li, S. Z. (2001). *Markov Random Field Modeling in Computer Vision*. New York: Springer-Verlag.

[Li et al., 2018] Li, X., Wang, X., and Dai, Y. (2018). Adaptive energy weight based active contour model for robust medical image segmentation. *J Signal Process Syst*, 90(3):449–465.

[Liao and Zhang, 2011] Liao, L. and Zhang, Y. (2011). MRI image segmentation based on fast kernel clustering analysis. *Frontiers of Electrical and Electronic Engineering in China*, 6(2):363–373.

[Lin et al., 2000] Lin, Z., Jin, J., and Talbot, H. (2000). Unseeded region growing for 3D image segmentation. In *Selected papers from the Pan-Sydney Workshop on Visualisation-Volume 2*, pages 31–37. Australian Computer Society, Inc.

[Liu et al., 2013] Liu, L., Zhang, Q., Wu, M., Li, W., and Shang, F. (2013). Adaptive segmentation of magnetic resonance images with intensity inhomogeneity using level set method. *Magn Reson Imaging*, 31(4):567–574.

[Liu et al., 2007] Liu, X. L., Zhang, Y. S., and Xie, Y. (2007). Image segmentation algorithm based on simulated annealing and fuzzy c-means clustering. *Journal of Engineering Graphics*, 1:89–93.

[Løkketangen and Woodruff, 1996] Løkketangen, A. and Woodruff, D. L. (1996). Progressive hedging and Tabu search applied to mixed integer (0, 1) multistage stochastic programming. *Journal of Heuristics*, 2(2):111–128.

[Lü et al., 2015] Lü, C., Yang, X., and Qi, S. (2015). Color image segmentation based on the ant colony algorithm. In *2015 8th International Congress on Image and Signal Processing (CISP)*, pages 438–442. IEEE.

[MacEachern and Manku, 1998] MacEachern, L. A. and Manku, T. (1998). Genetic algorithms for active contour optimization. In *ISCAS'98. Proceedings of the 1998 IEEE International Symposium on Circuits and Systems (Cat. No. 98CH36187)*, volume 4, pages 229–232. IEEE.

[MacQueen et al., 1967] MacQueen, J. et al. (1967). Some methods for classification and analysis of multivariate observations. In *Proceedings of the fifth Berkeley Symposium on Mathematical Statistics and Probability*, volume 1, pages 281–297. Oakland, CA, USA.

[Malisia and Tizhoosh, 2006] Malisia, A. R. and Tizhoosh, H. R. (2006). Image thresholding using ant colony optimization. In *The 3rd Canadian Conference on Computer and Robot Vision (CRV'06)*, pages 26–26. IEEE.

[Malladi and Vemuri, 1995] Malladi, R., S. J. and Vemuri, B. (1995). Shape modeling with front propagation: a level set approach. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 17.

[Mandal et al., 2014] Mandal, D., Chatterjee, A., and Maitra, M. (2014). Robust medical image segmentation using particle swarm optimization aided level set based global fitting energy active contour approach. *Eng Appl Artif Intel*, 35:199–214.

[Manikandan et al., 2014] Manikandan, S., Ramar, K., Iruthayarajan, M. W., and Srinivasagan, K. (2014). Multilevel thresholding for segmentation of medical brain images using real coded genetic algorithm. *Measurement*, 47:558–568.

[Manousakas et al., 1998] Manousakas, I., Undrill, P., Cameron, G., and Redpath, T. (1998). Split-and-merge segmentation of magnetic resonance medical images: performance evaluation and extension to three dimensions. *Computers and Biomedical Research*, 31(6):393–412.

[Mantegna, 1994] Mantegna, R. N. (1994). Fast, accurate algorithm for numerical simulation of Levy stable stochastic processes. *Physical Review E*, 49(5):4677.

[Marroquin et al., 2002] Marroquin, J. L., Vemuri, B. C., Botello, S., Calderon, E., and Fernandez-Bouzas, A. (2002). An accurate and efficient Bayesian method for automatic segmentation of brain MRI. *IEEE Trans. Med. Imag.*, 21(8):934–945.

[Maulik, 2009] Maulik, U. (2009). Medical image segmentation using genetic algorithms. *IEEE Transactions on Information Technology in Biomedicine*, 13(2):166–173.

[Maulik and Saha, 2010] Maulik, U. and Saha, I. (2010). Automatic fuzzy clustering using modified differential evolution for image classification. *IEEE Transactions on Geoscience and Remote Sensing*, 48.

[Mekhmoukh and Mokrani, 2015] Mekhmoukh, A. and Mokrani, K. (2015). Improved fuzzy C-means based particle swarm optimization (pso) initialization and outlier rejection with level set methods for MR brain image segmentation. *Computer Methods and Programs in Biomedicine*, 122(2):266–281.

[Memon et al., 2019] Memon, K. H., Memon, S., Qureshi, M. A., Alvi, M. B., Kumar, D., and Shah, R. A. (2019). Kernel possibilistic fuzzy C-means clustering with local information for image segmentation. *International Journal of Fuzzy Systems*, 21(1):321–332.

[Mesejo and Cagnoni, 2016] Mesejo, Pablo, I. O.-C. O. and Cagnoni, S. (2016). A survey on image segmentation using metaheuristic-based deformable models: State of the art and critical analysis. *Applied Soft Computing*.

[Mesejo et al., 2015] Mesejo, P., Valsecchi, A., Marrakchi-Kacem, L., Cagnoni, S., and Damas, S. (2015). Biomedical image segmentation using geometric deformable models and metaheuristics. *Computerized Medical Imaging and Graphics*, 43:167–178.

[Michael Kass and Terzopoulos, 1988] Michael Kass, A. W. and Terzopoulos, D. (1988). Snakes: Active contour models. *International Journal of Computer Vision*, 1.

[Mirghasemi et al., 2013] Mirghasemi, S., Rayudu, R., and Zhang, M. (2013). A new image segmentation algorithm based on modified seeded region growing and particle swarm optimization. In *2013 28th International Conference on Image and Vision Computing New Zealand (IVCNZ 2013)*, pages 382–387. IEEE.

[Mlakar et al., 2016] Mlakar, U., Potočnik, B., and Brest, J. (2016). A hybrid differential evolution for optimal multilevel image thresholding. *Expert Systems with Applications*, 65:221–232.

[Mobahi et al., 2006] Mobahi, H., Ahmadabadi, M. N., and Araabi, B. N. (2006). Swarm contours: A fast self-organization approach for snake initialization. *Complexity*, 12.

[Mohamed, 2018] Mohamed, A. W. (2018). A novel differential evolution algorithm for solving constrained engineering optimization problems. *J Intell Manuf*, 29(3):659–692.

[Mohamed et al., 2019] Mohamed, A. W., Mohamed, A. K., Elfeky, E. Z., and Saleh, M. (2019). Enhanced directed differential evolution algorithm for solving constrained engineering optimization problems. *International Journal of Applied Metaheuristic Computing (IJAMC)*, 10(1):1–28.

[Mohan and Baskaran, 2012] Mohan, B. C. and Baskaran, R. (2012). A survey: Ant colony optimization based recent research and implementation on several engineering domain. *Expert Systems with Applications*, 39(4):4618–4627.

[Mukhopadhyay and Maulik, 2011] Mukhopadhyay, A. and Maulik, U. (2011). A multiobjective approach to MR brain image segmentation. *Appl Soft Comput*, 11(1):872–880.

[Mukhopadhyay et al., 2015] Mukhopadhyay, A., Maulik, U., and Bandyopadhyay, S. (2015). A survey of multiobjective evolutionary clustering. *ACM Computing Surveys (CSUR)*, 47(4):61.

[Muller et al., 2001] Muller, K.-R., Mika, S., Ratsch, G., Tsuda, K., and Scholkopf, B. (2001). An introduction to kernel-based learning algorithms. *IEEE Transactions on Neural Networks*, 12(2):181–201.

[Nakib et al., 2007] Nakib, A., Oulhadj, H., and Siarry, P. (2007). Image histogram thresholding based on multiobjective optimization. *Signal Processing*, 87(11):2516–2534.

[Nakib et al., 2010] Nakib, A., Oulhadj, H., and Siarry, P. (2010). Image thresholding based on Pareto multiobjective optimization. *Engineering Applications of Artificial Intelligence*, 23(3):313–320.

[Nanda et al., 2004] Nanda, P., Patra, D., and Pradhan, A. (2004). Unsupervised image segmentation using Tabu search and hidden Markov random field model and hidden Markov random field model.

[Nguyen et al., 2018a] Nguyen, T., Vo, D., Vu Quynh, N., Van Dai, L., et al. (2018a). Modified Cuckoo search algorithm: a novel method to minimize the fuel cost. *Energies*, 11(6):1328.

[Nguyen et al., 2018b] Nguyen, T. T., Vo, D. N., and Dinh, B. H. (2018b). An effectively adaptive selective cuckoo search algorithm for solving three complicated short-term hydrothermal scheduling problems. *Energy*, 155:930–956.

[Nosrati and Hamarneh, 2016] Nosrati, M. S. and Hamarneh, G. (2016). Incorporating prior knowledge in medical image segmentation: a survey. [Online].

[Oliva et al., 2019] Oliva, D., Elaziz, M. A., and Hinojosa, S. (2019). *Metaheuristic Algorithms for Image Segmentation: Theory and Applications*, volume 825. Springer.

[Omran et al., 2005] Omran, M., Engelbrecht, A. P., and Salman, A. (2005). Particle swarm optimization method for image clustering. *International Journal of Pattern Recognition and Artificial Intelligence*, 19(03):297–321.

[Ortiz et al., 2013] Ortiz, A., Palacio, A. A., Górriz, J. M., Ramírez, J., and Salas-González, D. (2013). Segmentation of brain MRI using SOM-FCM-based method and 3D statistical descriptors. *Computational and Mathematical Methods in Medicine*, 2013.

[Otsu, 1979] Otsu, N. (1979). A threshold selection method from gray-level histograms. *IEEE Transactions on Systems, Man, and Cybernetics*, 9(1):62–66.

[Ouadfel and Batouche, 2003] Ouadfel, S. and Batouche, M. (2003). MRF-based image segmentation using ant colony system. *ELCVIA Electronic Letters on Computer Vision and Image Analysis*, 2(1):12–24.

[Padhye and Deb, 2011] Padhye, N. and Deb, K. (2011). Multi-objective optimisation and multi-criteria decision making in SLS using evolutionary approaches. *Rapid Prototyping J*, 17(6):458–478.

[Pare et al., 2017] Pare, S., Kumar, A., Bajaj, V., and Singh, G. K. (2017). An efficient method for multilevel color image thresholding using Cuckoo search algorithm based on minimum cross entropy. *Applied Soft Computing*, 61:570–592.

[Parsopoulos and Vrahatis, 2010] Parsopoulos, K. E. and Vrahatis, M. N. (2010). Particle swarm optimization and intelligence: advances and applications.

[Pavlidis and Liow, 1990] Pavlidis, T. and Liow, Y.-T. (1990). Integrating region growing and edge detection. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 12(3):225–233.

[Peres et al., 2018] Peres, W., Júnior, I. C. S., and Passos Filho, J. A. (2018). Gradient based hybrid metaheuristics for robust tuning of power system stabilizers. *International Journal of Electrical Power & Energy Systems*, 95:47–72.

[Pham and Prince, 1999] Pham, D. L. and Prince, J. L. (1999). Adaptive fuzzy segmentation of magnetic resonance images. *IEEE Transactions on Medical Imaging*, 18(9):737–752.

[Pham et al., 2017a] Pham, T. X., Siarry, P., and Oulhadj, H. (2017a). Image clustering using improved particle swarm optimization. In *International Conference on Industrial Networks and Intelligent Systems*, pages 359–373. Springer.

[Pham et al., 2017b] Pham, T. X., Siarry, P., and Oulhadj, H. (2017b). An improved particle swarm optimization algorithm for MRI image segmentation. In *13th Biennal International Conference on Artificial Evolution, Artificial Evolution*, pages 277–284.

[Pham et al., 2018] Pham, T. X., Siarry, P., and Oulhadj, H. (2018). Integrating fuzzy entropy clustering with an improved PSO for MRI brain image segmentation. *Appl Soft Comput*, 65:230–242.

[Pham et al., 2019a] Pham, T. X., Siarry, P., and Oulhadj, H. (2019a). A cooperating metaheuristic approach for MR image segmentation. In *2019 3rd International Conference on Bio-engineering for Smart Technologies (BioSMART)*, pages 1–5. IEEE.

[Pham et al., 2019b] Pham, T. X., Siarry, P., and Oulhadj, H. (2019b). A multi-objective optimization approach for brain MRI segmentation using fuzzy entropy clustering and region-based active contour methods. *Magn Reson Imaging*, 61:41–65.

[Pham et al., 2019c] Pham, T. X., Siarry, P., and Oulhadj, H. (2019c). Segmentation of MR brain images through hidden Markov random field and hybrid metaheuristic algorithm. *IEEE Transactions on Image Processing*, (Under reviewed).

[Poli, 2008] Poli, R. (2008). Analysis of the publications on the applications of particle swarm optimisation. *Journal of Artificial Evolution and Applications*, 2008.

[Powell, 1981] Powell, M. J. D. (1981). *Approximation theory and methods*. Cambridge university press.

[Preetha et al., 2015] Preetha, M. M. S. J., Suresh, L. P., and Bosco, M. J. (2015). Cuckoo search based color image segmentation using seeded region growing. In *Power Electronics and Renewable Energy Systems*, pages 1573–1583. Springer.

[Puzicha et al., 1999] Puzicha, J., Hofmann, T., and Buhmann, J. M. (1999). Histogram clustering for unsupervised image segmentation. In *Proceedings. 1999 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (Cat. No PR00149)*, volume 2, pages 602–608. IEEE.

[Qin et al., 2019] Qin, J., Shen, X., Mei, F., and Fang, Z. (2019). An Otsu multi-thresholds segmentation algorithm based on improved ACO. *The Journal of Supercomputing*, 75(2):955–967.

[Raidl, 2006] Raidl, G. R. (2006). A unified view on hybrid metaheuristics. In *International Workshop on Hybrid Metaheuristics*, pages 1–12. Springer.

[Raj et al., 2019] Raj, A., Gautam, G., Abdullah, S. N. H. S., Zaini, A. S., and Mukhopadhyay, S. (2019). Multi-level thresholding based on differential evolution and Tsallis fuzzy entropy. *Image and Vision Computing*.

[Ramadas and Abraham, 2019] Ramadas, M. and Abraham, A. (2019). *Metaheuristics for data clustering and image segmentation*. Springer.

[Rana et al., 2011] Rana, S., Jasola, S., and Kumar, R. (2011). A review on particle swarm optimization algorithms and their applications to data clustering. *Artificial Intelligence Review*, 35(3):211–222.

[Reyes-Sierra and Coello, 2006] Reyes-Sierra, M. and Coello, C. A. C. (2006). Multi-objective particle swarm optimizers: A survey of the state-of-the-art. *IJCIR*, 2(3):287–308.

[Rogai et al., 2016] Rogai, F., Manfredi, C., and Bocchi, L. (2016). Metaheuristics for specialization of a segmentation algorithm for ultrasound images. *IEEE Transactions on Evolutionary Computation*, 20(5):730–741.

[Roy and Chaudhuri, 2013] Roy, S. and Chaudhuri, S. S. (2013). Cuckoo search algorithm using lévy flight: a review. *International Journal of Modern Education and Computer Science*, 5(12):10.

[Saha and Bandyopadhyay, 2011] Saha, S. and Bandyopadhyay, S. (2011). Automatic MR brain image segmentation using a multiseed based multiobjective clustering approach. *Appl Intell*, 35(3):411–427.

[Sahoo et al., 1997] Sahoo, P., Wilkins, C., and Yeager, J. (1997). Threshold selection using Renyi's entropy. *Pattern Recognition*, 30(1):71–84.

[Sapiro and Tannenbaum, 1993] Sapiro, G. and Tannenbaum, A. (1993). Affine invariant scale-space. *International Journal of Computer Vision*, 11.

[Sarkar et al., 2017] Sarkar, S., Das, S., and Chaudhuri, S. S. (2017). Multi-level thresholding with a decomposition-based multi-objective evolutionary algorithm for segmenting natural and medical images. *Applied Soft Computing*, 50:142–157.

[Selvi and Umarani, 2010] Selvi, V. and Umarani, R. (2010). Comparative analysis of ant colony and particle swarm optimization techniques. *International Journal of Computer Applications*, 5(4):1–6.

[Sengupta et al., 2018] Sengupta, S., Basak, S., and Peters, R. (2018). Particle swarm optimization: A survey of historical and recent developments with hybridization perspectives. *Mach Learn Knowl Extr*, 1(1):157–191.

[Sethian, 1999] Sethian, J. (June 1999). *Level Set Methods and Fast Marching Methods*. Cambridge University Press, Cambridge.

[Sezgin and Sankur, 2004] Sezgin, M. and Sankur, B. (2004). Survey over image thresholding techniques and quantitative performance evaluation. *Journal of Electronic Imaging*, 13(1):146–166.

[Shankar et al., 2016] Shankar, T., Shanmugavel, S., and Rajesh, A. (2016). Hybrid HSA and PSO algorithm for energy efficient cluster head selection in wireless sensor networks. *Swarm and Evolutionary Computation*, 30:1–10.

[Sharma et al., 2012] Sharma, N., Mishra, M., and Shrivastava, M. (2012). Colour image segmentation techniques and issues: an approach. *International Journal of Scientific and Technology Research*, 1(4):9–12.

[Shehab et al., 2017] Shehab, M., Khader, A. T., and Al-Betar, M. A. (2017). A survey on applications and variants of the Cuckoo

search algorithm. *Applied Soft Computing*, 61:1041–1059.

[Siddhartha Bhattacharyya, 2016] Siddhartha Bhattacharyya, Paramartha Dutta, S. D. G. K. e. (2016). *Hybrid Soft Computing for Image Segmentation*. Springer International Publishing.

[Sierra and Coello, 2005] Sierra, M. R. and Coello, C. A. C. (2005). Improving PSO-based multi-objective optimization using crowding, mutation and ϵ-dominance. In *Proc EMO'05*, pages 505–519. Springer.

[Silva Filho et al., 2015] Silva Filho, T. M., Pimentel, B. A., Souza, R. M., and Oliveira, A. L. (2015). Hybrid methods for fuzzy clustering based on fuzzy C-means and improved particle swarm optimization. *Expert Systems with Applications*, 42(17):6315–6328.

[Simmons et al., 1994] Simmons, A., Tofts, P. S., Barker, G. J., and Arridge, S. R. (1994). Sources of intensity nonuniformity in spin echo images at 1.5 T. *Magn Reson Med*, 32(1):121–128.

[Simon, 2013] Simon, D. (2013). *Evolutionary optimization algorithms*. John Wiley & Sons.

[Singh and Bala, 2019] Singh, C. and Bala, A. (2019). A transform-based fast fuzzy C-means approach for high brain MRI segmentation accuracy. *Applied Soft Computing*, 76:156–173.

[Sulaiman et al., 2014] Sulaiman, S. N., Non, N. A., Isa, I. S., and Hamzah, N. (2014). Segmentation of brain MRI image based on clustering algorithm. In *2014 IEEE Symposium on Industrial Electronics and Applications (ISIEA)*, pages 60–65. IEEE.

[Suman and Kumar, 2006] Suman, B. and Kumar, P. (2006). A survey of simulated annealing as a tool for single and multiobjective optimization. *Journal of The Operational Research Society*, 57(10):1143–1160.

[Sumathi et al., 2018] Sumathi, R., Venkatesulu, M., and Arjunan, S. P. (2018). Extracting tumor in MR brain and breast image with Kapurs entropy based Cuckoo search optimization and morphological reconstruction filters. *Biocybernetics and Biomedical Engineering*, 38(4):918–930.

[Sun et al., 2016] Sun, G., Zhang, A., Yao, Y., and Wang, Z. (2016). A novel hybrid algorithm of gravitational search algorithm with genetic algorithm for multi-level thresholding. *Applied Soft Computing*, 46:703–730.

[Suresh and Lal, 2016] Suresh, S. and Lal, S. (2016). An efficient Cuckoo search algorithm based multilevel thresholding for segmentation of satellite images using different objective functions. *Expert Systems with Applications*, 58:184–209.

[T Krishnan et al., 2016] T Krishnan, P., Balasubramanian, P., and Krishnan, C. (2016). Segmentation of brain regions by integrating meta heuristic multilevel threshold with Markov random field. *Current Medical Imaging Reviews*, 12(1):4–12.

[Taha and Hanbury, 2015] Taha, A. A. and Hanbury, A. (2015). Metrics for evaluating 3D medical image segmentation: analysis, selection, and tool. *BMC Medical Imaging*, 15(1):29.

[Talbi, 2009] Talbi, E.-G. (2009). *Metaheuristics: from design to implementation*, volume 74. John Wiley & Sons.

[Talbi et al., 2012] Talbi, E.-G., Basseur, M., Nebro, A. J., and Alba, E. (2012). Multi-objective optimization using metaheuristics: non-standard algorithms. *International Transactions in Operational Research*, 19(1-2):283–305.

[Talbi et al., 2013] Talbi, E.-G. et al. (2013). *Hybrid metaheuristics*, volume 166. Springer.

[Talbi et al., 2007] Talbi, H., Batouche, M., and Draa, A. (2007). A quantum-inspired evolutionary algorithm for multiobjective image segmentation. *International Journal of Mathematical, Physical and Engineering Sciences*, 1(2):109–114.

[Tan et al., 2019] Tan, T. Y., Zhang, L., Lim, C. P., Fielding, B., Yu, Y., and Anderson, E. (2019). Evolving ensemble models for image segmentation using enhanced particle swarm optimization. *IEEE Access*, 7:34004–34019.

[Tang et al., 2007] Tang, L., Wang, K., Feng, G., and Li, Y. (2007). An image segmentation algorithm based on the simulated annealing and improved snake model. In *2007 International Conference on Mechatronics and Automation*, pages 3876–3881. IEEE.

[Teixeira et al., 2008] Teixeira, G. M., Pommeranzembaum, I. R., de Oliveira, B. L., Lobosco, M., and Dos Santos, R. W. (2008). Automatic segmentation of cardiac MRI using snakes and genetic algorithms. In *International Conference on Computational Science*, pages 168–177. Springer.

[Ting et al., 2015] Ting, T., Yang, X.-S., Cheng, S., and Huang, K. (2015). Hybrid metaheuristic algorithms: past, present, and future. In *Recent Advances in Swarm Intelligence and Evolutionary Computation*, pages 71–83. Springer.

[Toth and Vigo, 2003] Toth, P. and Vigo, D. (2003). The granular Tabu search and its application to the vehicle routing problem. *Informs Journal on Computing*, 15(4):333–346.

[Tran and Wagner, 2000] Tran, D. and Wagner, M. (2000). Fuzzy entropy clustering. In *Fuzzy Systems, 2000. FUZZ IEEE 2000. The Ninth IEEE International Conference on*, volume 1, pages 152–157. IEEE.

[Tsallis, 1988] Tsallis, C. (1988). Possible generalization of boltzmann-gibbs statistics. *Journal of Statistical Physics*, 52(1-2):479–487.

[Verma et al., 2016] Verma, H., Agrawal, R., and Sharan, A. (2016). An improved intuitionistic fuzzy C-means clustering algorithm incorporating local information for brain image segmentation. *Applied Soft Computing*, 46:543–557.

[Vese and Chan, 2002] Vese, L. A. and Chan, T. F. (2002). A multiphase level set framework for image segmentation using the Mumford and Shah model. *Int J Comput Vision*, 50(3):271–293.

[Vovk et al., 2007] Vovk, U., Pernus, F., and Likar, B. (2007). A review of methods for correction of intensity inhomogeneity in MRI. *IEEE Transactions on Medical Imaging*, 26(3):405–421.

[Wang et al., 2015] Wang, L., Yiwen, Z., and Yilong, Y. (2015). A hybrid cooperative Cuckoo search algorithm with particle swarm optimisation. *International Journal of Computing Science and Mathematics*, 6(1):18–29.

[Wang et al., 2018] Wang, M., Wan, Y., Gao, X., Ye, Z., and Chen, M. (2018). An image segmentation method based on fuzzy C-means clustering and Cuckoo search algorithm. In *Ninth International Conference on Graphic and Image Processing (ICGIP 2017)*, volume 10615, page 1061525. International Society for Optics and Photonics.

[Wang et al., 2003] Wang, W., Xu, Z., Lu, W., and Zhang, X. (2003). Determination of the spread parameter in the Gaussian kernel for classification and regression. *Neurocomputing*, 55(3):643–663.

[Wang et al., 2010] Wang, X.-F., Huang, D.-S., and Xu, H. (2010). An efficient local Chan-Vese model for image segmentation.

Pattern Recogn, 43(3):603–618.

[Wang et al., 2006] Wang, X.-N., Feng, Y.-J., and Feng, Z.-R. (2006). Ant colony optimization with active contour models for image segmentation. *Control Theory and Applications*, 23(4):515–522.

[Wang and Rangaiah, 2017] Wang, Z. and Rangaiah, G. P. (2017). Application and analysis of methods for selecting an optimal solution from the Pareto-optimal front obtained by multiobjective optimization. *Ind Eng Chem Res*, 56(2):560–574.

[Wolpert et al., 1997] Wolpert, D. H., Macready, W. G., et al. (1997). No free lunch theorems for optimization. *IEEE Transactions on Evolutionary Computation*, 1(1):67–82.

[Xiao et al., 2006] Xiao, M., Xia, S., and Wang, S. (2006). Geometric active contour model with color and intensity priors for medical image segmentation. In *2005 IEEE Engineering in Medicine and Biology 27th Annual Conference*, pages 6496–6499. IEEE.

[Xie et al., 2015] Xie, M., Gao, J., Zhu, C., and Zhou, Y. (2015). A modified method for MRF segmentation and bias correction of MR image with intensity inhomogeneity. *Med Biol Eng Comput*, 53(1):23–35.

[Yang et al., 2015] Yang, C., Gao, W., Liu, N., and Song, C. (2015). Low-discrepancy sequence initialized particle swarm optimization algorithm with high-order nonlinear time-varying inertia weight. *Applied Soft Computing*, 29:386–394.

[Yang and Tsai, 2008] Yang, M.-S. and Tsai, H.-S. (2008). A Gaussian kernel-based fuzzy c-means algorithm with a spatial bias correction. *Pattern recognition Letters*, 29(12):1713–1725.

[Yang et al., 2008] Yang, W.-l., Guo, L., XU, Z., XIAO, G.-c., and ZHAO, T.-y. (2008). Image segmentation method based on region growing and ant colony clustering. *Application Research of Computers*, 5.

[Yang et al., 2014] Yang, X., Gao, X., Tao, D., Li, X., and Li, J. (2014). An efficient MRF embedded level set method for image segmentation. *IEEE Trans. Image Process.*, 24(1):9–21.

[Yang et al., 2015] Yang, X., Gao, X., Tao, D., Li, X., and Li, J. (2015). An efficient MRF embedded level set method for image segmentation. *IEEE Trans. Image Process.*, 24(1):9–21.

[Yang and Deb, 2009] Yang, X.-S. and Deb, S. (2009). Cuckoo search via lévy flights. In *2009 World Congress on Nature & Biologically Inspired Computing (NaBIC)*, pages 210–214. IEEE.

[Yin and Wu, 2017] Yin, P.-Y. and Wu, T.-H. (2017). Multi-objective and multi-level image thresholding based on dominance and diversity criteria. *Applied Soft Computing*, 54:62–73.

[Yousefi et al., 2012] Yousefi, S., Azmi, R., and Zahedi, M. (2012). Brain tissue segmentation in MR images based on a hybrid of MRF and social algorithms. *Med Image Anal*, 16(4):840–848.

[Zahara et al., 2005] Zahara, E., Fan, S.-K. S., and Tsai, D.-M. (2005). Optimal multi-thresholding using a hybrid optimization approach. *Pattern Recognition Letters*, 26(8):1082–1095.

[Zanaty and Ghiduk, 2013] Zanaty, E. A. and Ghiduk, A. S. (2013). A novel approach based on genetic algorithms and region growing for magnetic resonance image (mri) segmentation. *Comput. Sci. Inf. Syst.*, 10(3):1319–1342.

[Zhang and Chen, 2004] Zhang, D.-Q. and Chen, S.-C. (2004). A novel kernelized fuzzy C-means algorithm with application in medical image segmentation. *Artificial Intelligence in Medicine*, 32(1):37–50.

[Zhang, 1992] Zhang, J. (1992). The mean field theory in EM procedures for Markov random fields. *IEEE Trans. Image Process.*, 40(10):2570–2583.

[Zhang et al., 2014] Zhang, J., Li, H., Tang, Z., Lu, Q., Zheng, X., and Zhou, J. (2014). An improved quantum-inspired genetic algorithm for image multilevel thresholding segmentation. *Mathematical Problems in Engineering*, 2014.

[Zhang et al., 2016a] Zhang, M., Jiao, L., Ma, W., Ma, J., and Gong, M. (2016a). Multi-objective evolutionary fuzzy clustering for image segmentation with MOEA/D. *Appl Soft Comput*, 48:621–637.

[Zhang and Li, 2007] Zhang, Q. and Li, H. (2007). MOEA/D: A multiobjective evolutionary algorithm based on decomposition. *IEEE Transactions on Evolutionary Computation*, 11(6):712–731.

[Zhang et al., 2001] Zhang, Y., Brady, M., and Smith, S. (2001). Segmentation of brain MR images through a hidden Markov random field model and the Expectation-Maximization algorithm. *IEEE Trans. Med. Imag.*, 20(1):45–57.

[Zhang et al., 2017] Zhang, Y., Gong, D.-w., and Cheng, J. (2017). Multi-objective particle swarm optimization approach for cost-based feature selection in classification. *IEEE ACM T Comput Bi*, 14(1):64–75.

[Zhang et al., 2015a] Zhang, Y., Wang, S., and Ji, G. (2015a). A comprehensive survey on particle swarm optimization algorithm and its applications. *Mathematical Problems in Engineering*, 2015.

[Zhang et al., 2015b] Zhang, Y., Wang, S., and Ji, G. (2015b). A comprehensive survey on particle swarm optimization algorithm and its applications. *Math Probl Eng*, 2015:38 pages.

[Zhang et al., 2016b] Zhang, Y., Yan, H., Zou, X., Tao, F., and Zhang, L. (2016b). Image threshold processing based on simulated annealing and Otsu method. In *Proceedings of The 2015 Chinese Intelligent Systems Conference*, pages 223–231. Springer.

[Zhao and Fan, 2019] Zhao, Feng, C. Y. L. H. and Fan, J. (2019). Alternate PSO-based adaptive interval type-2 intuitionistic fuzzy C-means clustering algorithm for color image segmentation. *IEEE Access*.

[Zhao et al., 2015] Zhao, F., Liu, H., and Fan, J. (2015). A multiobjective spatial fuzzy clustering algorithm for image segmentation. *Appl Soft Comput*, 30:48–57.

[Zhao et al., 2017] Zhao, L., Zheng, S., Wei, H., and Gui, L. (2017). Adaptive active contour model driven by global and local intensity fitting energy for image segmentation. *Optik*, 140:908–920.

[Zhou et al., 2011] Zhou, A., Qu, B.-Y., Li, H., Zhao, S.-Z., Suganthan, P. N., and Zhang, Q. (2011). Multiobjective evolutionary algorithms: A survey of the state of the art. *Swarm and Evolutionary Computation*, 1(1):32–49.

[Zitzler et al., 2001] Zitzler, E., Laumanns, M., and Thiele, L. (2001). SPEA2: Improving the strength Pareto evolutionary algorithm. *TIK-report*, 103.

[Zucker, 1976] Zucker, S. W. (1976). Region growing: Childhood and adolescence. *Computer Graphics and Image Processing*, 5(3):382–399.