Towards the prediction and understanding of tribological effects on system performance through artificial intelligence

Nikzad Motamedi

THÈSE

Pour l'obtention du titre de :

Docteur de l'Université de Lille

Spécialité : Mécanique, Génie Mécanique, Génie Civil

Ecole doctorale : ENGSYS

Présentée par :
Nikzad MOTAMEDI

Vers la prédiction et la compréhension des effets tribologiques sur les performances systèmes par l'intelligence artificielle

Towards the prediction and understanding of tribological effects on system performance by artificial intelligence

Thèse soutenue publiquement, mardi le 04 avril 2023 à 10h devant le jury composé de :

Guilhem MOLLON	Rapporteur	Maître de conférences - HDR INSA Lyon
Hassen DRIRA	Rapporteur	Professeur Université de Strasbourg
Céline BOUBY	Examinatrice	Maître de conférences Université de Lorraine
Faten CHAIEB	Examinatrice	Maître de conférences - HDR Université Panthéon-Assas-Paris
Philippe DUFRENOY	Examinateur	Professeur Université de Lille
Frédéric LEBON	Président du jury	Professeur Université de Marseille
Hazem WANNOUS	Co-directeur de thèse	Professeur Université IMT Nord-Europe, Lille
Vincent MAGNIER	Directeur de thèse	Maître de conférences - HDR Université de Lille

Remerciement

Ce travail de thèse a été réalisé au sein de l'équipe µFrein du Laboratoire de Mécanique Multiphysique et Multiéchelle (LaMcube) de l'Université de Lille.

Mes sincères remerciements vont à la Région Hauts-de-France pour avoir financé cette thèse, ce qui a indéniablement contribué à son aboutissement.

Je souhaite exprimer ma sincère gratitude envers mon directeur de thèse, Vincent MAGNIER, pour sa passion, son dévouement et sa vision inspirante dans la recherche. Sa disponibilité, ses conseils pertinents et sa capacité à susciter ma curiosité ont grandement contribué à la réussite de cette thèse. Travailler sous sa direction a été un honneur et une expérience enrichissante. Je le remercie chaleureusement pour son soutien inconditionnel et son mentorat précieux.

Je souhaite également exprimer ma profonde gratitude envers mon co-directeur de thèse, Hazem WANNOUS, pour ses conseils éclairés, son soutien constant et sa contribution précieuse à mes travaux de recherche. Son expertise et son engagement ont été d'une importance capitale dans la réalisation de cette thèse et je le remercie sincèrement pour son dévouement et sa collaboration précieuse.

Je tiens à remercier Frédéric Lebon d'avoir accepté d'être le président de mon jury de thèse, et Guilhem Mollon et Hassen Drira d'avoir accepté d'être les rapporteurs de ma thèse. Leur présence et leurs évaluations ont enrichi mes travaux et je suis honoré d'avoir bénéficié de leur expertise.

J'adresse mes remerciements sincères à tous les membres de l'équipe µFrein pour leur collaboration, leur esprit d'équipe et les moments partagés ensemble. Je souhaite remercier tout particulièrement Philippe Dufrénoy, J-F Brunel, Esso, Van, Dimitri, Maël, Florent, pour leur contribution significative à ce travail collectif.

Mes remerciements également au personnel administratif du LaMcube, Jean-Baptiste Colliat, Nathalie Labaeye, Ana-Maria Moudart pour leur gentillesse et leur accueil chaleureux.

Un grand merci à tous mes amis (Thierry, Mehrnoush, Saied, Ali (Tommy), Adrien, Victor, Nagesh, Valentin, Mathis, Edouard, Yue, Sanna, Arnaud, Annie, Guillaume) pour leur gentillesse, les moments de divertissement partagés et leur précieuse aide. Votre présence et votre soutien ont rendu cette expérience encore plus enrichissante et inoubliable.

Je souhaite exprimer ma gratitude envers mon père, ma mère, ma sœur Firouzeh, mon frère Farzad, ma sœur Soulmaz et ma sœur Fatemeh pour leur soutien constant et leur amour inconditionnel. Leurs encouragements ont été d'une importance capitale tout au long de cette thèse.

Je tiens à remercier chaleureusement mon épouse Mehrnaz pour son soutien indéfectible, sa patience et ses encouragements tout au long de cette aventure. Sa présence a été une source d'inspiration et de motivation pour moi.

Je souhaite également remercier toutes les personnes qui, de près ou de loin, ont contribué à la réalisation de cette thèse. Leur soutien et leur collaboration ont été précieux.

Abstract

The contact between two parts, especially when induced, as in braking systems, needs to be improved and better understood. The phenomena involved are complex as they encompass multi-scale and multi-physics concepts within a context of constant evolution. An additional challenge arises from the closed nature of the contact, making it difficult to explicitly observe the interface phenomena that play a crucial role in the targeted performance such as noise pollution, emission of fine particles, and wear, among others. Considering the increasingly stringent European environmental standards, it is imperative to establish new strategies for comprehensively addressing this problem. The team aims to bridge the gap between numerical simulations and experimental methods. The "experimental" component relies on dedicated test benches equipped with advanced and precise instrumentation. The "numerical" aspect involves multi-scale and multi-physics modeling, striving to account for tribological mechanisms within a complete system. One challenge lies in comparing and aligning these two components.

Therefore, the objective of this PhD thesis is to propose predictive models that connect the contact interface with the complete system using artificial intelligence. Initially, we will attempt to determine the natural frequencies of a pin-on-disk system by considering various surface types for the interface. More specifically, the interface will exhibit a numerically generated roughness field. Subsequently, an AI model will be developed to predict the contact distribution during a test. This part relies on measurements obtained from a thermocouple array embedded near the surface of the friction material during the tests. Lastly, based on experimental acquisitions of the surface profile at different time intervals, a model will be proposed to determine the wear evolution. The AI models specifically developed for these three components employ algorithms such as CNN, GAN, RNN, etc. As these concepts may not be familiar to the mechanical engineering community, they will be illustrated using a simple example of behavior identification in the introduction of this manuscript. In terms of results, the obtained outcomes demonstrate a high level of satisfaction when compared to simulation and/or experimental data. This confirms the value of utilizing AI to advance the prediction capabilities of these models. Additionally, AI enables understanding the significance of input parameters, potentially facilitating system optimization and test control in the medium term.

Keywords: Tribology, Deep Learning, Contact, Multi-Physical, Multi-Scale, Evolutionary

Resumé

Le contact entre deux pièces, notamment quand il est provoqué comme sur les systèmes de freinage, mérite d'être amélioré et donc d'être mieux compris. Les phénomènes mis en jeu sont complexes car ils font appel à des notions multi-échelle, multi-physique dans un contexte d'évolution constante. Une difficulté supplémentaire est que le contact est fermé, et donc il est par conséquent difficile d'observer explicitement les phénomènes d'interface qui joue un rôle prépondérant dans les performances visés (pollution sonore, émission des particules fines, usure etc.). Au regard des normes environnementales européennes qui seront de plus en plus sévères, il est indispensable d'établir de nouvelles stratégies permettant de mieux appréhender le problème dans sa globalité. L'équipe souhaite tendre vers le couplage numérique et expérience. La partie « expérience » s'appuie sur des bancs d'essais dédiés avec une instrumentation riche et fine. La partie « numérique » s'appuie sur des modélisations multi-échelles et multi-physiques tentant de considérer des mécanismes tribologiques au sein d'un système complet. Une difficulté est de vouloir comparer (recaler) ces deux parties.

Ainsi, l'objectif de cette thèse est de proposer des modèles prédictifs liant l'interface de contact avec le système complet par l'**intelligence artificielle**. Dans un premier temps, on tentera de déterminer les fréquences propres d'un système pion-disque en prenant en compte n'importe quelles surfaces pour l'interface. Plus précisément, cette interface présentera un champ de rugosité qui sera généré numériquement. Dans un second temps, un modèle IA est développé pour prédire la distribution de contact pendant un essai. Cette partie s'appuie sur des mesures d'une nappe de thermocouples noyée dans le matériau de friction en proche surface pendant les essais. Dans un troisième temps, en s'appuyant sur des acquisitions expérimentales du profil de surface à différent instant, un modèle est proposé pour déterminer les évolutions de l'usure. Les modèles IA spécifiquement développés pour ces trois parties font appel à des algorithmes de type CNN, GAN, RNN etc. Ces notions n'étant pas commune dans la communauté mécanicienne, elles sont illustrées sur un exemple simple d'identification de comportement en préambule de ce manuscrit. En termes de bilan, les résultats obtenus sont très satisfaisants au regard de la comparaison aux données de simulation et/ou expérimentales. Cela confirme l'intérêt de l'utilisation de l'IA afin de passer un cap dans la prédiction des modèles. De plus, l'IA permet aussi la compréhension et l'importance des paramètres d'entrée qui pourrait servir à moyen terme à optimiser le système ou de piloter les essais.

Mots-clés : Tribologie, Deep Learning, Contact, Multi-Physique, Multi-Echelle, Évolutif

Table of Contents

An Overview of the Study

1. INTRODUCTION

The contact mechanism is a tribological issue and one of the most complex problems that engineers might face. As it is a multi-scale, multi-physical and evolutionary problem, a correct understanding of its physical characteristics leads to make optimal designs in the industrial sector and prevents the loss of energy, materials, and the occurrence of many disturbances in which this mechanism plays a major role. One of the most common applications in which the contact mechanism has a great impact is friction brakes. This type of braking is widely used in the automotive and railway industries, but there are two major problems associated with friction brakes. First, the squealing of the brake pads, and second, the release of fine particles during the braking process, which cause many diseases, including mental [1][2], respiratory [3], skin diseases [4] and visual disturbances [5].

But diseases are not the only aspect of this problem because the occurrence of these disorders causes high costs for after-sales service companies and the replacement of these parts causes the production of large amounts of waste, which has many destructive environmental effects [6][7].

As the contact issue is multi-scale, multi-physical, and evolutionary, and the contact surfaces evolve under the effects of sliding, loads, and different mechanical properties of materials, they can lead to different behaviors of the system, which can cause problems. In order to prevent these problems to occur, different aspects of the contact mechanism must be identified.

For this reason, in order to clarify the aspect of this issue, each of these aspects will be described separately, and by reviewing the valuable research that has been done before, we will have an overview of the road map.

2. MULTI-SCALE POINT OF VIEW

To analyze the problem from the multi scale point of view, different scales which are illustrated in Figure 1 are presented by [8]. The contact process is basically multi-scale, from the atomic up to the system scale. Generally, three scales are considered in tribological problems. The first scale is the macro-scale models where the physical quantities are influenced by small scales. The second scale is associated to the different components of the global system and their surface topography (roughness, plateaus, etc.). The third scale is related to the wear debris, which results from the initial contact between the bodies.

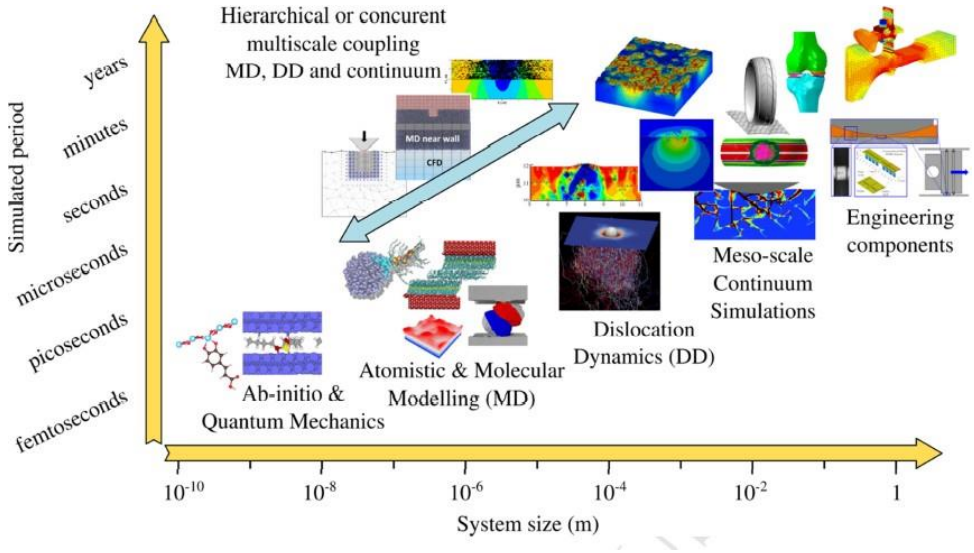


Figure 1: A time-vs scales amount diagram of models developed in tribology [8].

The figure highlights a remarkable challenge: as the scale increases, the computation time also increases dramatically. This can result in calculations for large-scale structures taking years to complete. As a consequence, the application of these models becomes highly impractical, making it nearly impossible to accurately predict complete systems that incorporate all relevant parameters.

One of the methods that can help to understand the process of contact is the use of analytical or modeling methods. In this regard, many researchers have tried to use these methods to predict contact behavior from different aspects.

Analytical methods: Hertz is the pioneer in the study of contact mechanics and tribology. Hertz theory tries to solve this problem by using two non-conforming elastic objects that are exposed to frictionless contact mechanics [8][9]. It should be noted that the initial models used to investigate the problem were based on Hertz theory and they have taken adhesion into consideration and have helped to understand the mechanics of contact [10][11]. But these models have many limitations and are only valid for very small contact surfaces [8]. For this reason, in recent years, other methods such as semi-analytical [12], multilayered elastic coating or the introduction of anisotropy [13], the use of heterogeneous materials [14][15] and the use of viscoelastic properties [16] were used to make these methods closer to reality. But it must be admitted that in a real contact phenomenon, the contact surfaces are rough, so the models need to be realistic and close to reality, and it is necessary to consider roughness as a parameter in models. In fact, this roughness makes the actual contact surface much smaller than a perfectly smooth contact surface, which leads to a completely different behavior.

The effects of surface roughness are highlighted by [17], which investigates the relationship between electrical conductivity and surface roughness across a wide range of charge values. The

results confirm that the actual contact area is much smaller than the nominal contact area. [18] Provides a valid model for modeling a contact surface that is close to rough surfaces in reality.

Braking scale: In recent years, the advanced IT science, programming, hardware, and computing modeling methods such as finite elements method (FEM) has helped to model the contact mechanism and a complete braking system in its whole complexity [19][20] and it is mostly used to solve these problems leading to a high number of degree of freedom (DOF) models. Figure 2 presents the common braking system mostly used in academic and industrial domains [19].

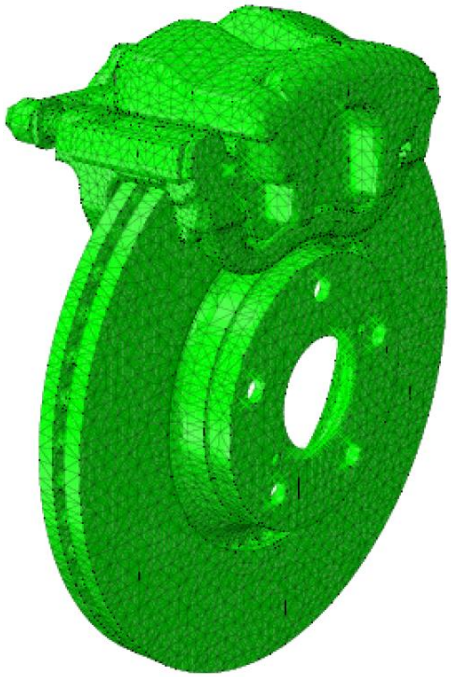


Figure 2: Complete FEM disc brake model [19].

Generally, in these models, the contact surface is considered smooth and the roughness or heterogeneity of the materials are not considered in the calculations, which reduces the accuracy. Also the contact modeling is often simplified by considering only flat surfaces [21][22][23][24]. Furthermore, in studying the dynamic behavior of the system, other minimal approaches have also been proposed alongside the reduction of degrees of freedom. For instance, with mode Lock-in theory, the stability of frictional systems has been analyzed using simplified mass-spring systems [25][26].

Surface scale: Since the state of the contact surface plays a significant role in the behavior of tribological systems, models have been presented to consider the contact surface state. For example,[27] analyzes and models the effect of roughness on the mechanical responses of the system, including stress fields [28]. Additionally, [29] investigates the thermomechanical responses

of the system to predict the resulting dynamic aspects, including vibrations. But in these models, the heterogeneity of materials and the evolution of the state of the contact surface are not considered. So, there is a need to analyze specially the surface state evolution in a third body scale.

Lately, the simplified models that consider roughness were born as the role of surface roughness in the occurrence of disturbances in the contact process was inevitable [30]. Obviously, the computation time in such models is much less than a complete braking system model, and by changing the parameters in the surface scale, the responses at the system scale are examined. But these models also could not provide a completely appropriate results, because many unknown phenomena are involved in the contact problem or to have a complete model, many parameters must be entered into the models, which will firstly increase the computation time and secondly working with such models would be very difficult and impractical with such models.

Third body scale: As the current approaches used for third body models differ from static rough modeling, most of these models are based on particle interactions, which contribute to the evolution of the contact surface.

As an example, the third body is considered as a fluid granular medium made of rigid or deformable particles in the discrete element method (DEM) [31][32].

As the large-scale interactions between contact objects cannot be considered in this approach, an improved approach called "Non Smooth Contact Dynamics" [33] has been used in [34] to combine FEM to deform contact objects and DEM for third body particles.

The approaches used for the third body issue are various, such as the cellular automaton [35]. In this method, the medium is seen as a discrete set of cells having different states and linked to each other by many rules controlling the evolution of the third body.

The neighboring cells define the state of each specific cell. The movable cellular automaton approach is the enhancing factor of this method [36][37] which consists of pairing two cells by a state corresponding to their chemical bonds. Thereafter, a general equation is obtained to describe the motion of the automaton considering the links between the cells and the force of inter-cell interaction. To solve the problem, a response function is introduced to consider the material behavior (elastic, plastic, etc.) by linking the cell deformation to stress.

3. MULTI-PHYSICS POINT OF VIEW

When two solids come into contact with each other, various tribological phenomena occur at the contact interface (Figure 3), and their interaction may give rise to different physical phenomena [8][38]. Among these various phenomena, mechanical (fatigue, deformations, micro-cracks), thermomechanical, and vibration aspects could be mentioned.

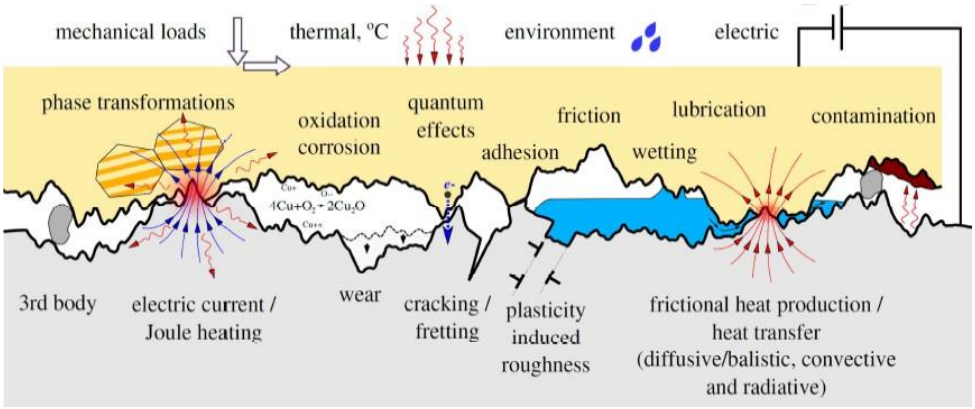


Figure 3 : multi-physics phenomena associated to contact [8]**.**

Various mechanisms can typically occur in contact surfaces either individually or in interaction with each other, leading to disruptions in contact systems. These mechanisms include fatigue, elastic and plastic deformations, as well as the expansion of micro-cracks under loads.

Fatigue: The Finite Element Method (FEM) has been utilized to investigate the fatigue aspect in contact problems within brake systems, which can lead to system disturbances under cyclic loads. The modeling approach aims to comprehend and predict the effects of fatigue on the system's behavior [39].

Deformations: Two kinds of deformation, elastic and plastic could happen due to loading materials during the contact process. To understand this mechanism by using a sphere on a flat surface, a new method was presented to calculate the elastic deformation. This model considers the influence of short-range as well as long-range attractive forces both inside and outside the actual contact area [40]. The influence of strain hardening on the cumulative plastic deformation, also known as ratcheting, occurring in repeated rolling and sliding contacts, has been assessed through the implementation of a non-linear kinematic hardening law. This assessment considers plastic deformations and stress levels on the contact surface [41].

Thermo-mechanical: Heat generation is another mechanism that occurs at the contact surface. Indeed the studies show that 95% of the mechanical energy is transformed into heat [42] [43][44]

and this temperature increase can spread through the contact surface or in the depth of the solids [45]. Changes in the mechanical properties of materials can be caused by high temperature [46] or expansion [47]. Since these changes in contact surface are heterogeneous, they lead to completely different contact localization [48], and as a result, different responses could be received from the system.

Therefore, initially, experimental models [49][50] were presented for the analysis of the contact surface, but since the contact surface is closed during contact and it is not possible to obtain much information about the contact state, these models could not provide a complete view of this phenomenon.

For this reason, in an effort to understand the temperature evolution on the contact surface, various models [51] [52] have been developed. However, due to their tendency to oversimplify the problem and not account for critical factors such as material heterogeneity, the accuracy of these models is compromised.

Therefore, methods such as the inverse method [48] by placing thermocouples under the contact surface were created so that they could analyze the behavior of the system by locating the contact zones, but these models could not predict the contact location, which is a crucial issue in the contact problem.

Vibrational aspects (friction-induced vibration): Many forms of waves and oscillations within solids might be created by the friction in the sliding contact, they can cause noise production and its propagation to the environment [53]. These instabilities caused by friction, are generally called friction-induced vibrations [38].

One of the major issues of interest in the transportation industry is the vibration due to friction on the brakes. As the friction makes the frequency higher than 1 kHz, it would cause a great discomfort for human ear [53]. Furthermore, the consequences of these issues can include safety concerns, inconvenient driving experiences, and increased maintenance requirements. As it is known, brakes are one of the essential components in terms of performance and safety in the field of transportation. Therefore, car manufacturers are looking for the ways to improve the design of the braking system to increase its performance quality, efficiency, and durability [38][54].

It must be noted that, the surface topography and pressure distribution are the major causes of surface instability and then the noise [30][55]. When the contact surface changes due to separation or combination of the third body, the friction surface will be changed and consequently the system behavior changes [55] (Figure 4), therefore studying and analyzing this issue would be beneficial.

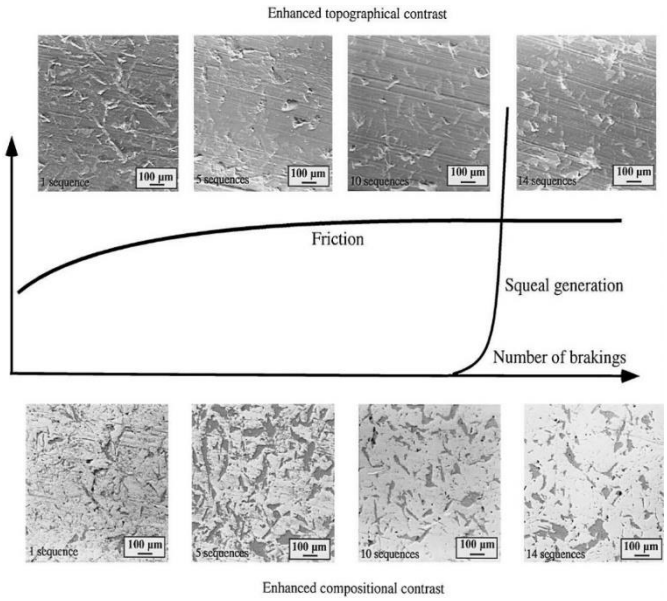


Figure 4: Surface evolution and corresponding squeal generation curves [55]

4. EVOLUTIVE POINT OF VIEW

When two solids come into contact with each other due to the presence of surface roughness, speed, and load, the contact surface undergoes evolution [56]. These evolutions can be caused by the separation of small parts from the contact surface, which either come out of the system or stick to the contact surfaces under the effect of high temperature and force during the contact process [57], which in any case causes the creation of a new topography generation and, as a result, different behavior. Various mechanisms cause the evolution of the contact surface, including Adhesive wear, Abrasive wear, Corrosive wear, Fatigue wear, but due to the complexity and smallness of these mechanisms, a model that can predict these changes well has not been presented.

Adhesive wear is due to the transfer of material from one surface to another surface by shearing of solid welded junctions of asperities. It leaves pits, voids, cavities or valley on the surface [58][59]. This wear occurs because of the adhesive bond. At the contact points, the adhesive bond is stronger than the cohesive bond of the weaker material of the Pair. Normally, adhesion occurs when two similar chemical composition metals are in contact or contact surface are free from oxide layer (vacuum or an inert atmosphere).

Abrasive wear occurs due to hard particles or protuberances sliding along a soft solid surface. It results from ploughing, wedging and cutting phenomena [59] [60]. In ploughing (also called ridge

formation) process, material is displaced at both the sides and forms a groove with or without removal of material.

Erosive wear is due to mechanical interaction between solid surface and fluid, or impinging liquid or solid particles [59] [61]. When particles with some velocity are impacted on the surface of metal, the pits and large scale subsurface deformation occur on the metal surface.

Fretting wear is due to small amplitude of oscillatory or reciprocating movement between two surfaces [59] [62]. It is a two-step mechanism. Initially, the adhesive wear occurs due to rubbing of two surfaces and then they become oxidized due to large quantity of energy stored in wear particles.

Fatigue wear is caused by fracture arising from surface fatigue due to cyclic loading [59] [63]. It results in a series of pits or voids. It usually occurs in rolling or sliding contact bodies such as bearings, roads, etc. After repeated cyclic loading, a crack is observed on the subsurface or the surface. The subsurface cracks propagate, connect with other cracks, reach the surface and generate wear particles.

Corrosive wear occurs when sliding takes place in corrosive or oxidative environment. During dry sliding also, the oxygen from the normal environment or other gases present in the environment can react with the solid surface [59] [64]. The excessive presence of anti-wear additives or other chemical agents also can bring corrosive wear.

Due to the complexities in the wear problem, many models have been presented to help to understand the mechanism, for example Adhesive wear [65][66] Abrasive Wear [67][68] Erosive wear [69][70] Fretting wear [71] [72] Fatigue/ Delamination wear [73] [74] Corrosive wear [75] but most of these models are either simplified or do not consider the interaction between the mechanisms.

5. AI POINT OF VIEW

The advancements in computer science, particularly in artificial intelligence and hardware such as Central Processing Units (CPUs), Graphics Processing Units (GPUs), and Tensor Processing Units (TPUs), have opened up new opportunities for mechanical engineers to tackle complex problems. However, despite the long history of AI research and development, the limited capabilities of earlier hardware prevented its widespread use in engineering. With the advent of modern technology and powerful computing devices such as GPUs and TPUs, AI has become a valuable tool for solving complex tasks in science. While AI is widely used in fields such as economics, health, and marketing, it is not as commonly used in engineering due to several factors. Engineers often view AI as a "black box" and have doubts about its applicability, despite the fact

that AI is based on simple mathematical equations that have been automated through programming. Additionally, the lack of proper knowledge and understanding of AI algorithms and data structure among engineers has led to limited use of AI in engineering. Furthermore, the lack of a consistent framework for storing data in engineering sciences has hindered the use of AI in these fields. Finally, AI algorithms have primarily been developed for purposes other than engineering and may require modification to be effectively used in engineering applications. To fully understand and utilize AI in engineering, it is necessary to introduce and discuss the potential uses of AI in different areas of engineering.

Previously, deep learning based methods have been presented to predict system-scale responses by examining the system's behavior [76]. In these methods, global system responses are analyzed and predicted by computer vision [77] while [78] analyzes the system's responses by considering the contact history by using recurrent neural network algorithms. And [79] analyzes and predict NVH (noise, vibration, and harshness) brake pad modal characteristics by using artificial neural network and machine learning. Or [80] they tried by image processing to predict the dry friction at the contact scale. Although these researches claim to have obtained good results, most of them analyze the responses of the system and predict them at the system scale, but the origin and root of these responses are not investigated.

Therefore, the main objective of this thesis is to conduct a comprehensive examination of the fundamental causes of mechanical problems through the utilization of suitable Machine Learning (ML) and Deep Learning (DL) algorithms. The research aims to delve into the root of these problems by analyzing the underlying mechanisms and identifying key factors that contribute to their occurrence. To achieve this goal, various algorithms will be employed to analyze and interpret data, as well as simulate and model the mechanical systems in question. The results of this analysis will be used to generate insights and recommendations for addressing the problems under investigation. The hope is that the research will contribute to the development of more effective solutions for mechanical problems and improve the overall efficiency and performance of mechanical systems.

6. MAIN SKETCH OF THIS THESIS

So far, a summary of tribology and contact issues has been provided, along with methods for understanding and predicting these problems. However, due to the inherent complexity of contact problems, the methods presented thus far have not fully elucidated many aspects of this scientific domain. As a result, a paradigm shift in the analytical approach and philosophy is necessary to

address the contact mechanism problem more effectively. In this study, the application AI, specifically DL, is explored to provide a more accurate and comprehensive analysis of contact mechanisms.

Despite the significant potential of deep learning techniques in addressing problems related to mechanical science, their application in this field remains moderate, as per literature review. To fully harness the capabilities of this method and uncover hidden aspects related to mechanical problems, it is essential to first evaluate the effectiveness of deep learning techniques in tackling complex mechanical issues. Subsequently, knowledge and experience can be applied to the specific area of contact mechanics.

First, a comprehensible example by using deep learning in a problem related to mechanical and civil engineers is presented. In this example, a database is firstly generated using finite element modeling. Then it is analyzed by using machine learning and deep learning algorithms and considering the physics of the problem. After explaining this example and gaining sufficient view about deep learning algorithms and understanding of mechanical problems, we will enter the main part of the thesis, which is the problem of contact mechanism.

The main body of this thesis consists of three parts. In the first part, using modeling data, the effect of surface topography on the occurrence of system instability is investigated to determine a criterion for these disorders, and the surfaces that have a higher risk of instability are analyzed.

After understanding the surface physics and examining the modeled data, it is time to analyze the experimental data to predict the temperature evolution (another physical aspect) and the performance of algorithms in facing experimental data that many parameters are involved in the occurrence of system behaviors, and the behavior of the system is studied by predicting the evolution temperature of system.

In the final chapter, an analysis of surface changes, one of the most complex issues in tribological studies, is conducted. Based on the surface condition, predictions are made regarding these evolutions.

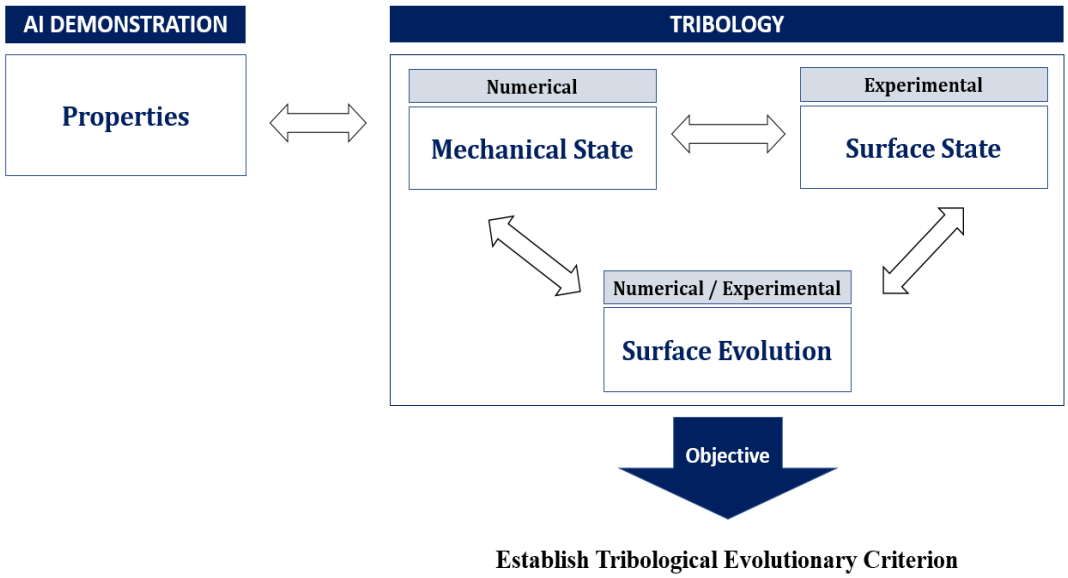
As the outline of this thesis:

- In Chapter I, a numerical modeling dataset is used to classify the mechanical properties of materials using a neural network architecture. Then, using a 2D-CNN architecture, the parameters of the material constitutive law are predicted by analyzing the displacement field.
- In Chapter II, a finite element model dataset is used to investigate the risk of instability in a pin-on-disc system based on the roughness of the contact surface. First, a 2D-CNN

architecture is used to classify the contact surfaces as stable or unstable, then another 2D-CNN is used to predict instability frequencies for the unstable surfaces.

- In Chapter III, the contact surface temperature evolution of a pin-on-disc system is studied using an experimental trial dataset and 2D-CNN and RNN architectures to analyze and predict the system responses then.
- In Chapter IV, the evolution of contact surface roughness is generated using a dataset obtained from experimental tests (pin-on-disc system) using GAN algorithms.

The general scenario of this thesis is described below.



The general scenario of the thesis, which in this chapter artificial intelligence algorithms are introduced through an example.

Transition

One of the main needs in mechanical engineering, in order to carry out optimal designs, is knowing the materials constitutive law and its parameters. However, due to the varying behavior of materials, such as anisotropy and the presence of numerous constitutive parameters, identifying their behavior becomes a challenging and laborious process. Consequently, in order to prevent energy and material wastage, it is essential to acquire knowledge about these parameters.

Therefore, in the first chapter of this manuscript (indicated with a blue box in the main outline (Figure I) through a mechanical example), two main goals are followed. The first is to classify the constitutive law by analyzing the received responses from a mechanical system, and the second is to determine the parameters of the constitutive law.

However, the primary objective of this thesis is to establish a connection with the field of tribology. Nonetheless, through this mechanical example, artificial intelligence algorithms are introduced to serve as an interface between the mechanical and computer engineering communities. This integration aims to provide a comprehensive perspective on the challenges that lie ahead.

One of the challenges in solving mechanical problems is selecting appropriate algorithms. So, the needs, geometry, physics, and the type of data, etc. all must be considered.

After providing a general point of view of artificial intelligence algorithms for mechanical engineers and mechanical problems for computer engineers, the main topic of the thesis, which is the use of artificial intelligence in tribology problems, will be discussed.

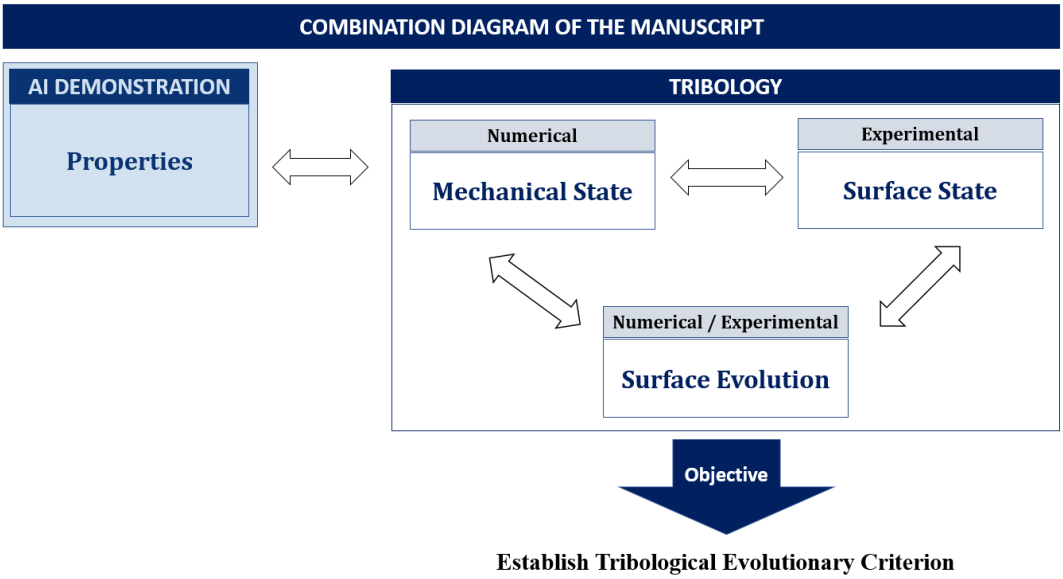


Figure I: The general scenario of the thesis, which in this chapter artificial intelligence algorithms are introduced through an example.

I. AI and Mechanics: Demonstrator on Classification and Regression Applied to the Constitutive Laws

Table of Contents

Abstract

The determination of the stress-strain relationship is a challenge especially when the material presents complex behaviors such as anisotropy. Here, the developed method is based on the use of Artificial Intelligence (AI) via deep learning (DL) techniques. The input data is the displacement field which could, from an experimental point of view, be obtained by image correlation. Here, the kinematics field is generated by Finite Element Method (FEM) calculations. The main purpose of this section is to classify and identify the constitutive law parameters, for these materials. By using the proposed methods, very promising results are obtained. The advantage of the proposed method is the quasi-instantaneousness of solution. Also, the obtained average errors are less than 3% on all the identified parameters. It should be added that the architecture of the AI is of paramount importance where the multi-scale notion must be considered, for this reason, an encapsulated architecture is used to determine the mechanical properties.

Keywords: Inverse identification, Deep Learning, different constitutive laws

I.1. INTRODUCTION

Understanding the mechanical properties of materials is not an easy task and requires a coupling between experiment and analytical/numerical tools. In order to determine the mechanical properties of materials, aid of relationship between stress and strain is required. Indeed, in the simplified case (1D), where the width and height of the sample are significantly smaller compared to its length, stress can be defined as the force per unit area (σ=F/A). In this case, stress (σ) is calculated as the ratio of the force (F) to the surface area (A). Consequently, when a body is subjected to an external force, its shape changes, leading to a strain. Strain (ε) is determined by dividing the change in length (ΔL) by the initial length (L) of the body (Figure 1a) [81].

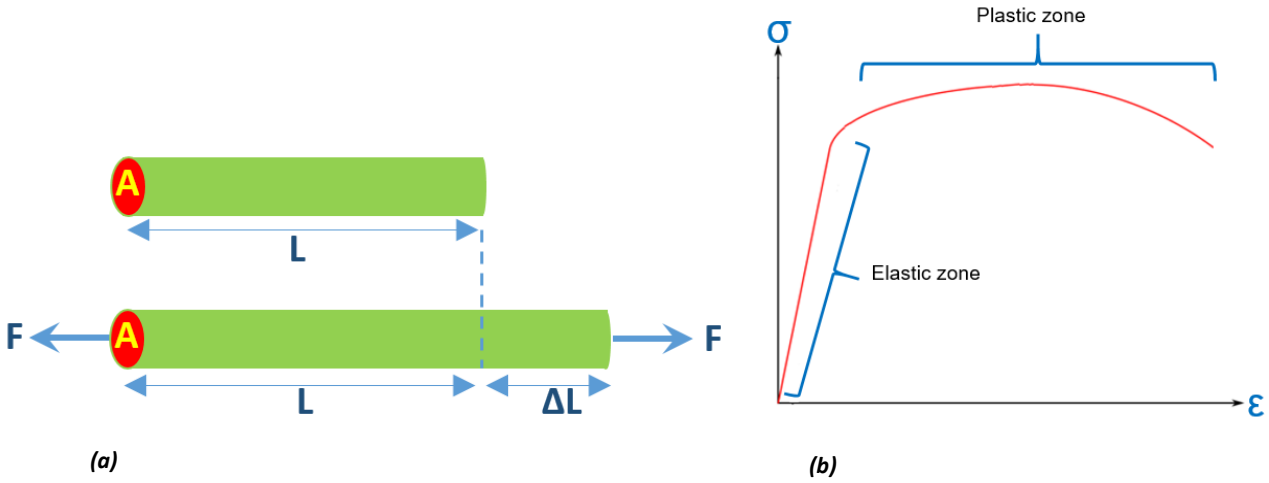


Figure 1a-b:

a) The change in the length of the object due to the external load applied to the longitudinal axis of the object in 1D case.
b) Strain vs stress diagram for elastoplastic materials which initially has a linear behavior, then a non-linear behaviour.

In Figure 1b, the relationship between stress and strain is shown. As it shows initially the strain increases linearly with the increase of the stress, but by the increase of the stress and passing through linear stage, the relationship turns nonlinear.

But in reality, not only these deformations happen in the direction of the force, but also transverse deformations occur (Figure 2), therefore the complexity of the problem increases where according to the direction of the applied force, the equations change according to the deformations of the object.

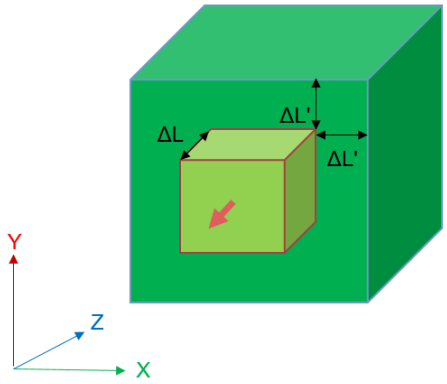


Figure 2: The change in the length of the object due to the external load applied to the longitudinal axis of the object in 3D case.

The equations corresponding to the deformations are shown in Figure 3, where σ represents stresses, ε represents strains, E represents Young's modulus, ν represents Poisson's ratio, τ represents shear stress, and γ represents shear strain.

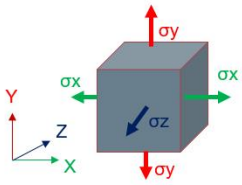
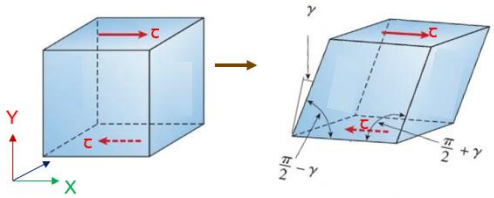
The strain caused by the applied load	The shear strain caused by an applied shear load
	
εx = 1/E[σx-ν (σy+σz)] εy = 1/E[σy-ν (σx+σz)] εz = 1/E[σz-ν (σx+σy)]	γ = τ [(2(1+ν)/E)]

Figure 3: The equations of the changes in shapes according to the applied force direction.

In addition, determining the stress-strain relationship is a challenge especially when the material exhibits complex behavior (anisotropy is the property of a material that allows it to change or assume different properties in different orientations, heterogeneity, damage, asymmetry of tensile/compression behavior, etc.). Awareness of this relationship is the key to understand the behavior of materials at the structural scale in order to obtain the best possible designs. This is in a context where environmental standards are increasingly restrictive and the need to lighten the structure has become a priority to save material and energy. To meet these needs, it is essential to better understand the new materials which are used in present day (composites, 3D printing, etc.) with the aim of being able to optimize them on both their composition and their shape.

From a bibliographical point of view, the current approach to characterize the mechanical properties of these materials, consists mainly in taking up the standardized tests, simplified

tests [82], or reverse engineering methods to determine the mechanical properties from global information [83][84][85][86][87]. In these methods the effort is mainly focused on the preparation of the test in order to apply only one loading mode in particular (traction, compression, shear, etc.). Although these approaches make it possible to estimate a constitutive law and its associated parameters, they are less accurate when dealing with complex or heterogeneous materials.

Recently, other methods have tried to study the behavior of materials by the help of receiving more information from the experimental test process. For example, when the test is heavily instrumented, in particular by means of cameras, more advanced studies have made it possible to develop strategies based on a numerical- experimental dialogue [88][89] such as the Finite Element Model Updating (FEMU) [90] methods and its variations [91]. Nevertheless, the use of these numerical methods does not give full satisfaction when the number of parameters to be determined becomes important. The difficulty lies in the minimization of a single cost function that can generate several pairs of "admissible" solutions.

In recent years, several methods such as point clouds obtained from simple or complex tests and the constitutive law have been proposed to show the relationship between stresses and strains to determine the parameters of the constitutive law. But the relationship could only be shown when they were calculated by tensile tests or isostatic homogeneous tests, analytical formulae and a simple geometry [92][93][94][95][96][97][98].

Recently, new methods have been developed such as "Data-Driven Computational Mechanics" (DDCM) [99][100][101][102] and in continuity, Data-Driven Identification (DDI) [103][104]. These methods are based on the construction of statically admissible field from field measurements. Therefore, these methods allow the identification of certain parameters [103][104][105]. But the complexity of these methods and the high computational time when the number of calculated parameters makes them less efficient. Since different materials behave differently and the parameters of the constitutive law are various, it is better to first classify them according to the constitutive law, then determine the parameters of the constitutive law. There are methods [106] which perform the classification of the behavior of materials, but are not accurate enough due to the complex material behavior and which take a lot of time to perform [99][100]. In this section, a new way is proposed to first, classify the constitutive law and second identify the parameters of the constitutive law.

In the presented method, the displacement, strain, and stress obtained from a modeled tensile test are utilized to meet the objectives of the problem. From an experimental test point of view, the system responses can be calculated using the image correlation method. This method is

solely based on the measurement of kinematic fields from a characterization trial. The used methodology utilizes artificial intelligence (AI) on a dataset generated by numerical models. This numerical data is exploited from finite element method (FEM). In fact, in the first part, the classification between three materials classes (perfect plasticity, bilinear plasticity and transversely isotropic) is done. In the second part, by using a supervised learning algorithm, the feasibility and capacity of AI is demonstrated to determine the mechanical properties. A specific example on a transversely isotropic material is carried out. These material categories were chosen because by changing the orientation of mechanical properties and having more parameters than perfect plasticity and bilinear plasticity, they can well challenge the ability of AI to deal with complex mechanical problems. At the end, the objectives will be to determine the Young modulus in the longitudinal and transverse directions of the isotropy axis, the Poisson's ratio, the shear modulus, as well as the orientation angle of the transversely isotropic material. In this study, we propose a global approach that incorporates the characterization of this constitutive law.

I.2. AI: WHAT IS IT REALLY?

By development of computer science field, especially in the artificial intelligence and hardware such as central processing unit (CPUs), graphics processing unit (GPUs), tensor processing unit (TPUs) and the increase of computing power, a new prospect has opened up for mechanical engineers to be able to understand and solve complex problems.

Although the AI is not a new subject [107], and since many efforts have been made to develop algorithms in this science in the late 20th century, the low configuration hardware did not allow to use this powerful tool for a long time. By using the modern technology as well as computers and the availability of (GPUs) and (TPUs), a powerful assistant was born to use in the complex tasks of science.

Although the artificial intelligence is widely used in economics, health, marketing, etc. it is less used in engineering, because of several reasons.

First of all, engineers often perceive AI as a black box, and since the primary objective of research for this group is to understand and subsequently address problems, they often harbor doubts about using AI. However, it is important to emphasize that artificial intelligence is not an incomprehensible black box. It is rooted in simple mathematical equations that have been automated through programming science to expedite the solution of these equations.

Secondly, although mechanical and civil engineers have a good view of programming, the lack of proper knowledge of complex artificial intelligence algorithms and lack of proper view of

data structure and database, has made them reluctant to use this method, so they use it in very minimal sections.

Thirdly, unlike other sciences that have established well-structured frameworks for data storage over many years, the field of engineering lacks a standardized approach to data storage. As a result, each researcher tends to archive and analyze data in their own individual manner, leading to a lack of coherence in engineering data repositories.

Fourthly, the algorithms used in artificial intelligence have been developed for purposes other than engineering, and in order to use them, they need to be slightly modified or innovated to meet the needs well.

According to the mentioned points and in order to understand the algorithms in artificial intelligence properly, we first introduce them and then discuss the potential of using each of them in different sections.

Artificial intelligence: AI is a branch of computer science that refers to the general ability of computers to mimic human thinking and perform tasks in real environments, while machine learning (ML) is a subset of artificial intelligence and refers to techniques that enable systems to learn from examples. Deep learning (DL) is a subfield of machine learning, where neural networks form the backbone of its models. Figure 4 shows an overview of artificial intelligence and its scope.

In general, ML algorithms can be categorized according to the algorithms used to predict or categorize (ML and DL), and the learning methods (supervised learning, semi-supervisory learning and non-supervised learning) [108][109].

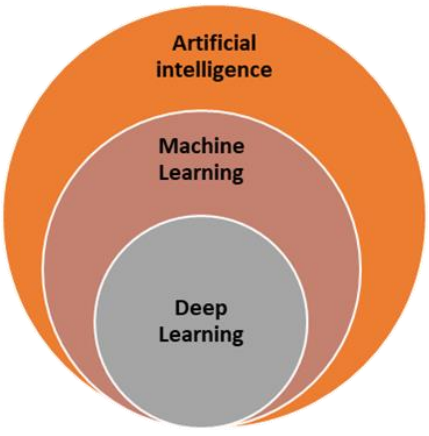


Figure 4: Overview of AI, machine learning and deep learning

Machine learning: (ML) finds a logical connection between inputs and responses using the examples given and non-complex algorithms, without explicit planning, human dictation and intervention. In machine learning, instead of planning everything, data is given to an algorithm, so it is the algorithm that builds its logic based on the given data [109].

Some of the most important machine learning algorithms are:

- Regression Algorithms
- Instance-based Algorithms
- Regularization Algorithms
- Decision Tree Algorithms
- Bayesian Algorithms
- Clustering Algorithms
- Association Rule Learning Algorithms
- Artificial Neural Network Algorithms (ANN)

These algorithms can be used according to problem needs which contain prediction or classification.

Deep Learning: (DL) is a subset of machine learning in which, inspired by human brain algorithms, a large amount of data is presented to the model, and by analyzing this data, the model can learn and analyze a wide range of problems. Deep learning algorithms, such as the human brain, can learn from experience, run a program repeatedly, and at each step, the algorithm improves itself to make the answers more accurate [110].

Some of the most important machine learning algorithms are:

- Convolutional Neural Network (CNN)
- Recurrent Neural Networks (RNNs)
- Generative Adversarial Networks (GANs)
- Auto-Encoders
- Deep Boltzmann Machine (DBM)
- Deep Belief Networks (DBN)

These algorithms also can be used according to problem needs which contain prediction or classification and can classify them into three general categories in terms of learning methods.

Supervised learning: Supervised algorithms are rhythms that attempt to find logic between input and output data using the examples given to them. In this method, all data is labeled and the output is labeled according to each input. In this way, the model takes the input data and tries to find the outputs. If the model responses were significantly different from the actual

responses, the algorithm attempts to correct the model in order to minimize the discrepancy between the predicted and actual answers.

Unsupervised learning: In unsupervised learning there is no label for the input data and the model tries to categorize the data according to the similarity of the data or mathematical rules.

Semi-supervised learning: In this method, some data is labeled, and some is not. By looking at the labeled data, the model tries to find a logic between input and output data, then develops its logic on the unlabeled data.

It is not possible to say which algorithm is the best to use in mechanical engineering or civil engineering, because each of these algorithms, depending on their architecture, can be very suitable for a series of problems and very unsuitable for a number of problems. For this reason, before using any of them, two aspects of the problem must be considered. Firstly, the physics of the problem should be known well, and secondly, by studying the existing algorithms, the most appropriate algorithm for solving the problem could be selected. After introducing AI, it is time to solve the defined problem. For this purpose, a scenario has been defined that is described in the next part.

I.3. GLOBAL VIEW OF THE PROBLEM RESOLUTION

In fact, the goal of the proposed scenario is: first classify different materials based on their behavior, then predict the parameters of the constitutive law according to the mechanical responses of the system. A scenario overview is shown in Figure 5. This process consists of different steps:

- Step 1: various mechanical properties are incorporated into the numerical model, leading to distinct behavioral characteristics exhibited by the system.
- Step 2: the responses gained from the model include displacements, deformations, stresses, and reaction forces, for each configuration are prepared and recorded in the database.
- Step 3. this data is used for training the AI model in order to classify the mechanical properties of materials.
- Step 4. after classifying the mechanical properties, this data is used to determine the mechanical properties.

- Step 5. the model performance is analyzed facing new data. Each of the mentioned steps is explained in detail in pursuit.

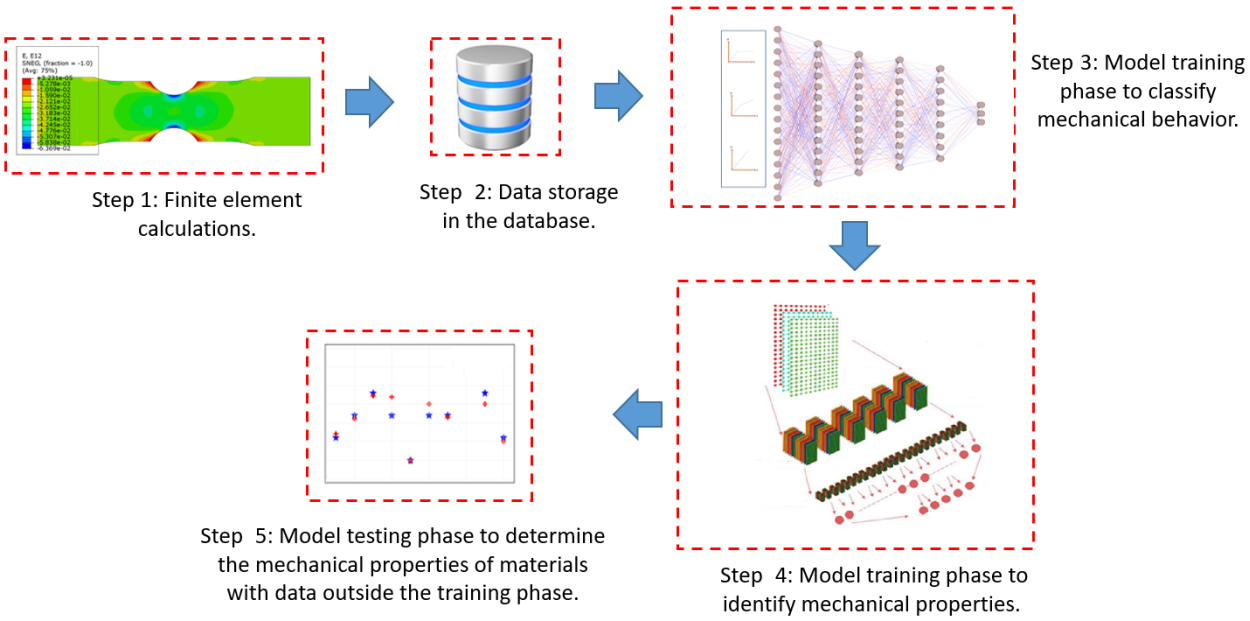


Figure 5: Strategy global to link system responses with mechanical properties

I.3.1. The database generation process

The database for the training algorithm is generated artificially using a finite element software. In order to demonstrate the approach, a tensile test on a standardized specimen is proposed. The geometry is detailed in Figure 6 where we find a section reduction in the center of the specimen. The dimensions are specified in millimeter and allow being in conformity with the standardized tests of the ISO-6892-1 standard. The database contains the whole kinematic field of the reduced section (stress, displacements, strains), as it could be done from the measurement of fields by cameras from an experimental point of view. A set of 2544 (consists of 48 rows and 53 columns) nodes are present as shown in Figure 7. The length of the arch in the area in question is 41.5 mm, which consists of 52 elements and the height of this area varies between 24 to 44 mm, which consists of 47 elements. Average size of the elements in the longitudinal direction is 0.743 mm and in the transverse direction is 0.66 mm.

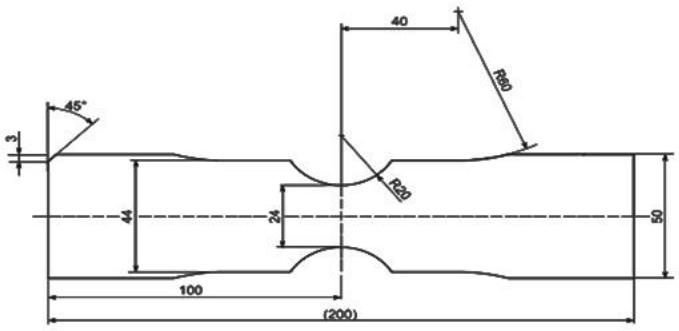


Figure 6: The geometry of model (Dimensions are in millimeter and the thickness of the sample is 1 mm)

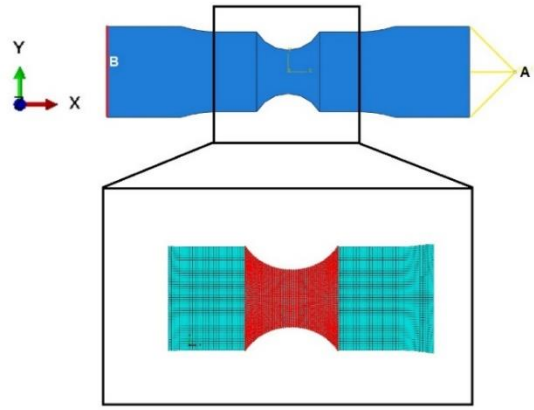


Figure 7: Boundary conditions: displacements= 0 mm on the left edge(B), and the displacement applied to point A= 5 mm

Concerning the boundary conditions, a displacement of 5 mm in the X direction is imposed on point A which is coupled with all the displacements of the right edge of the specimen. On the left edge (noted B in Figure 7), an embedding is imposed on all the nodes in order to comply with an experimental tensile test. Then, the stress, displacements and strains are recorded on all the points located on the central part of the test specimen and reactions force for the left edge. An illustration of the results obtained from the model is presented in Figure 8 where the displacement and strain fields are shown. The objective is also to demonstrate the variations in displacement and strain gradients, which can serve as signatures of the properties.

All these calculations are done sequentially on a server running at 2.50Ghz in a single processor. The calculation time to generate the whole database is approximately one day.

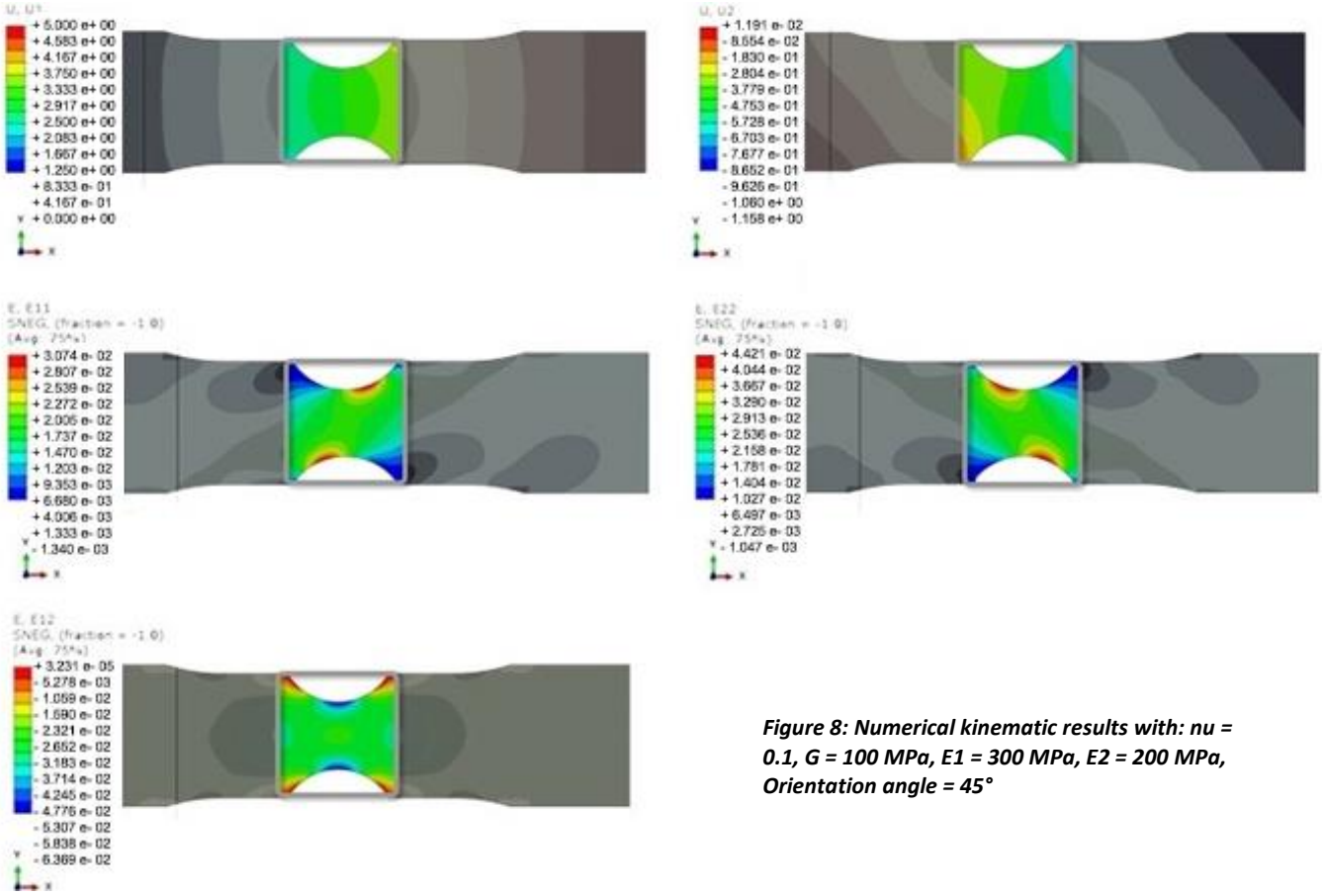


Figure 8: Numerical kinematic results with: nu = 0.1, G = 100 MPa, E1 = 300 MPa, E2 = 200 MPa, Orientation angle = 45°

After obtaining all the results, the data is normalized and then split as training, validation and test sets. This normalization is an essential step, so that there is no higher weighting for a family of data.

Normalizing a set of data transforms the set of data to be on a similar scale. The goal is usually to recenter and rescale the data to be arranged between 0 and 1 or -1 and 1. Normalization can help the training phase, as the different features are on a similar scale, which helps to converge faster the training phase [111].

Here the used normalization method is finding the absolute maximum value of each feature, then diving all values by this value, the scale of all data will be the same between 0 and 1.

I.3.2. Classification of the material behavior

According to the mentioned goals, the behavior of materials will be classified by using the artificial neural networks algorithm. For this purpose, strain-stress data of three classes of materials, including “perfect plasticity” (Figure 9a), “bilinear plasticity” materials (Figure 9b), and “transversely isotropic material” (Figure 9c), were stored in the database and the model was asked to classify them according to the different behavior of each material.

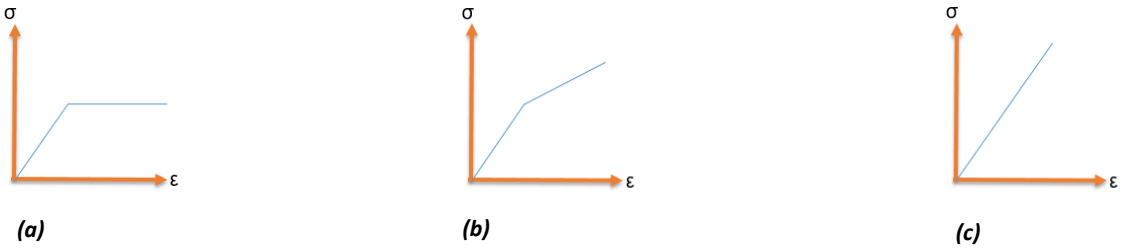


Figure 9a-c: Schematic of strain stress curves for different materials

a) **Schematic of strain stress curve for perfect plasticity (Class 1)**
b) **Schematic of strain stress curve for bilinear plasticity (Class 2)**
c) **Schematic of strain stress curve for transversely isotropic elastic (Class 3)**

In fact, by changing the mechanical properties of each category, the system responses, including the stresses and strains associated with each configuration, are recorded.

As, for the perfect plasticity, parameters including Young modulus (between 50 and 500 MPa), the elastic limit (between 25 and 80 MPa), and Poisson's ratio (between 0.1 and 0.4) are variable.

For the bilinear plasticity model, parameters including Young modulus (between 50 and 500 MPa), the elastic limit (between 25 and 80 MPa), Poisson's ratio (between 0.1 and 0.4), and plastic strain (between 0.05 and 0.25) are variable.

For transverse isotropic elastic materials, Young's modulus varied (between 50 and 500 MPa) in the longitudinal and transverse directions, shear modulus (between 50 and 200 MPa), transverse angle (between 0 and 90), and Poisson's ratio (between 0.1 and 0.4) are variable.

Each material shows different behavior due to their mechanical properties, which causes them to have different strain-stress curve. Strain-stress data for each of these configurations are stored in the database, and a label is assigned to each of these behaviors according to the different received responses.

2000 samples were made for each type of material, and 1700 cases were used in the training phase and 300 in the test phase. Then, after the data has been prepared for the first part of modeling, which is classification, the model training phase begins.

From this point of view, neural network algorithms were used to classify mechanical behavior, because in the presented example, the purpose of classification is 3 different classes of materials, and in machine learning algorithms, when the number of classes increases, the accuracy of the model decreases, because the algorithms have to analyze more features, but neural network models have a good performance in multi-class classification [112].

The schematic of the artificial network model used is shown in the Figure 10.

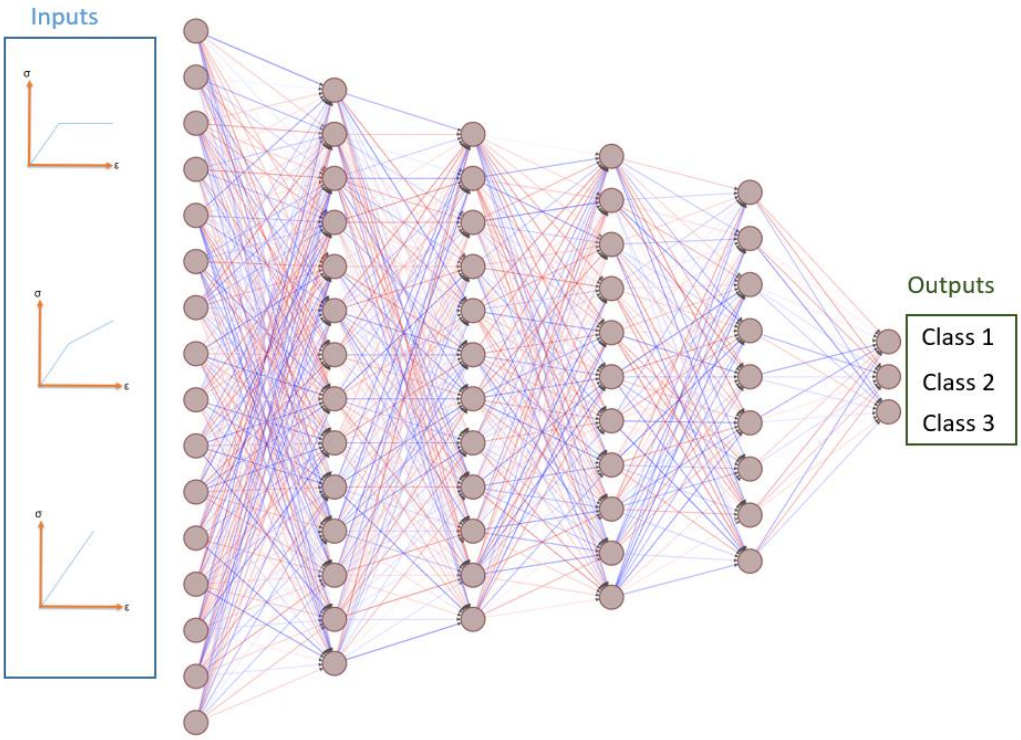


Figure 10: The schematic of the used (ANN) model to determine the mechanical properties.

In the following, neural networks are introduced and then they are classified into different categories using a mechanical behavior neural network.

I.3.3. Use of Artificial neural networks (ANN) to classify system responses

An artificial neural network (or simply neural network) consists of an input layer of neurons (or nodes, units, perceptron), several intermediate layers of neurons, and a final layer of output neurons [113].

First, we will give a summary of neural networks, then, it will be applied to the problem.

A model made of ANN algorithms generally consists of the following parts:

- Neurons
- Weights
- Activations
- Optimization
- Loss function

Multilayer perceptron: Neural networks or multilayered perceptron, are made up of a certain number of neurons [114]. Neural networks are designed to model the structure of the human brain in order to address challenging computational tasks, including prediction. The goal of

these algorithms is not to create real models of the brain, but to create robust algorithms and data structures that we can use to model difficult problems. A row of neurons is called a layer, and a network can have multiple layers (Figure 11). The architecture of network neurons is often called the network topology.

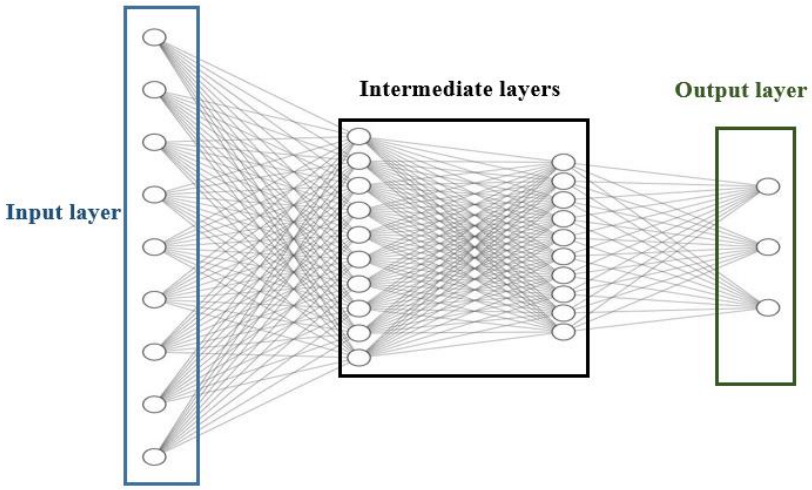


Figure 11: Schematic of an artificial neural network

Neurons: A block of neural networks is made of artificial neurons [111]. They are simple computational units that generate the output signals by giving weight to the input signals. (Figure 12)

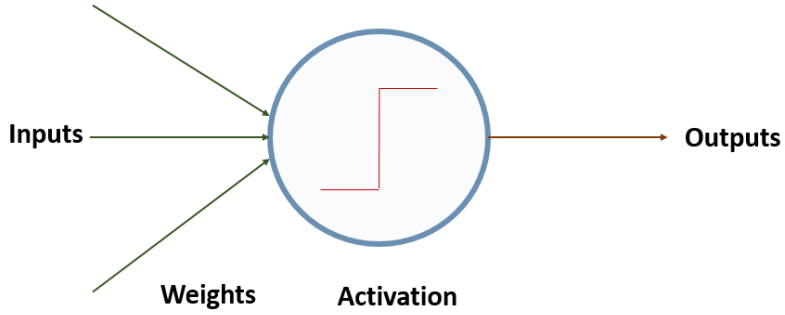


Figure 12: Schematic of an artificial neuron

After creating the database and choosing the algorithm that suits the needs of the problem, the model training phase begins.

The training phase of models made of neural networks consists of two parts. In the first part, which is called training, the data of the training phase is entered to the model and the weights are updated according to the input data and forecasts. Then in the next part, the validation data is entered to the model and the model tries to edit the weights by comparing the responses of the validation data to the real data so that the model has a good generalizability and has an acceptable performance in the test phase. One of the methods that can explain the model

behavior in the training and validation phase is the loss function curve, which measures the error or mismatch between predicted and actual values, helping evaluate the performance of the model. For having an acceptable performance in the training and validation phase, two points must be considered. The first is that the loss curves tend to zero, indicating a decreasing error between predicted and actual values. The second is that both the training and validation loss curves converge, suggesting that the model is effectively learning and generalizing across both datasets. An artificial intelligence model can perform well when its hyperparameters are well adjusted.

In order to set the best hyperparameters for model training, an iterative strategy is adopted. Hyperparameters are adjustable parameters that are not learned by the model itself, but rather set by the user before the training process. They control the behavior and performance of the model, such as the learning rate, batch size, number of hidden layers, and activation functions. In order to decrease calculations time, two models were considered in parallel, in each the loss functions were constant. And each model was run with different configurations to find the best hyperparameters. Finally, the average error percentage was recorded for each hyperparameter. In this strategy, to reduce the calculation time, two models were built in parallel, in each, the loss function, the number of epochs (An epoch is a complete learning cycle involving the forward and backward propagation using all the data, and the model tries to optimize the results by examining the error rate relative to the actual values.) (70 epochs), and the number of batch size (32 batch), which refers to the number of samples processed together during each iteration of training, were considered as fixed parameters. Then, in each model, the optimizer and metric were changed to find the best ones. In the following, after finding the best optimizer, metric and loss function, these hyper parameters are considered constant, and in the next step, the best batch size is found by keeping the number of epochs constant. Then the number of data required for training the model was found. Then best value of epoch is determined by changing the number of the epochs. The results are shown in (Table 1). The computation time for setting hyper parameters and complete training phase of the classification model by a server computer (Configuration: Memory: 128 GB, Processor: Intel® Xeon® Gold 5215 CPU @ 2.50GHz *40, Graphics: NVIDIA Corporation GV100GL [Tesla V100SPCIe 32GB] and a Samsung SSD 2 TB M.2 NVMe) is nearly 72 hours.

Table 1: Received responses from the classification model for different configurations

Loss Function	categorical_crossentropy			sparse_categorical_crossentropy	
Error %	17.62			27.4721	
Optimizer	Adagrad	RMSprop	adam	nadam	adamax
Error %	11.80	15.280	8.75	11.65	24.35
Metrics	Accuracy			categorical_accuracy	
Error %	8.0478			6.3715	
Batch Size	32		64	128	
Error %	8.4701		6.1174	7.4795	
Number of data	1400	1500	1600	1700	
Error %	6.942	6.2508	2.97	2.83	
Epoch	70		80	90	100
Error %	7.51		6.234	2.99	2.63

By using the hyperparameters obtained from the previous step, an artificial neural network model is built to find the best number of model layers that has the lowest error percentage. For this purpose, first the model was trained with 4 layers with a percentage error of 3.01%, then with 5 layers with a percentage error of 2.86%, 6 layers with a percentage error of 2.63%, 7 layers with a percentage error of 2.71%. and finally, 8 layers with a percentage error of 2.69%. The results obtained by the number of layers compared to the error percentage are shown in Figure 13.

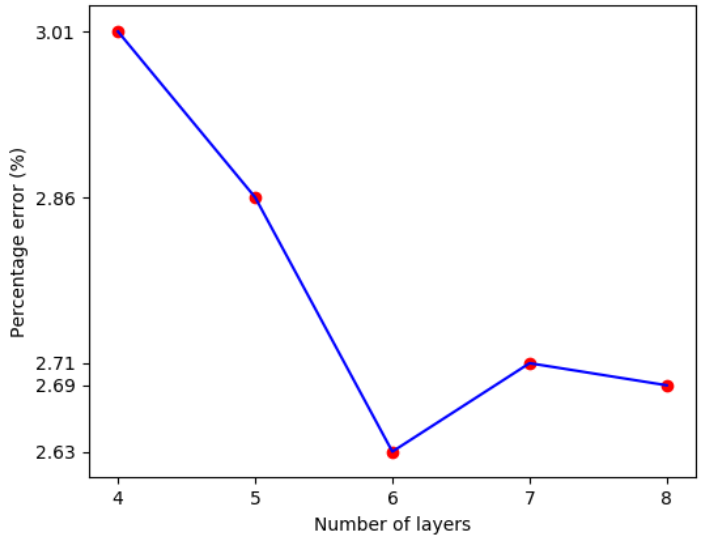


Figure 13: Error percentage vs. the number of layers

By looking at the model, it can be seen that initially the error percentage of the model decreases as the number of layers increases, but after 6 layers, there is no significant change in the performance of the model. For this reason, the best number of layers was 6. Table 2 shows the final specifications of the used model.

Table 2: Model specifications for classification

Model specifications	
Learning method:	Supervised learning
Algorithm:	Classification via "neural network"
Aim:	Classify the Constitutive Law
Number and type of layers:	6 Dense layers
Loss function:	Categorical-Crossentropy
Activation function:	sigmoid
Optimization function:	adam
Metric:	Categorical Accuracy

The loss function relative to the epoch curve can represent the behavior of the model in this phase.

The loss function-epochs curves and the accuracy evolution for this classification is shown in Figure 14. The blue curve represents the loss function in the training phase and the orange curve represents the loss function of the model in the validation part. In fact, the learning process consists of two parts, training and validation phases, and the model seeks to minimize the value of loss function in these phases, also tending this trend to zero indicates a good training phase for the model. It is clear that the model has a good training phase so that the loss function curves in the training and validation phases tended to zero and both of these curves were converged, and also, the accuracy of the model tends to one.

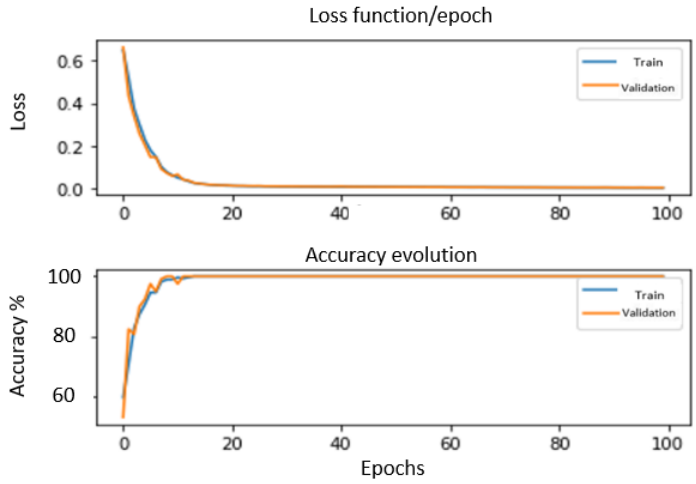


Figure 14: Loss function curves and accuracy evolution of model percentage

It is important because if the model does not perform well during training and has ***over-fit*** *or* ***under-fit***, measures should be taken to guide the model to better training.

The over-fitting is one of biggest problems in training neural networks. It means that the neural network at the certain time during the training period does not improve its ability to solve problem anymore. But just starts to learn some random regularity contained in the set of training patterns [115][116][117]. Over-fitting occurs when astatically model describes random error or noise instead of the underlying relationship [116][118][119].

Under-fitting is the opposite of Over-fitting. This occurs when the model is incapable of capturing the variability of the data [116][120]. Under-fitting, on the other hand, means the model has not captured the underlying logic of the data. It doesn't know what to do with the given task, therefore, it provides an answer that is far from reality.

In order to test the generalization performance of the model facing new data, a test phase is performed on 300 samples of test data not encountered in the training phase. The result is presented in the confusion matrix (Table3). The classification time in the test phase is nearly 10ms for each case.

Table 3: Confusion matrix to classify the behavior of the material

Prediction / Actual	Perfect plasticity	Bilinear plasticity	Transverse isotropic (elastic)
Perfect plasticity	292	7	1
Bilinear plasticity	11	289	0
Transverse isotropic (elastic)	2	5	293

Looking at the confusion matrix, it can be seen that the model can classify well, so it recognized 292 cases out of 300 (perfect plasticity materials) well the rest 7 cases misidentified bilinear plasticity and 1 case transverse isotropic (elastic). In the case of bilinear plasticity, it categorized 289 cases well and incorrectly identified 11 cases as perfect plasticity. In the case of transverse isotropic (elastic), it categorized 293 cases correctly and 7 cases incorrectly, also 5 cases were bilinear plasticity, and 2 cases were perfect plasticity. The error percentage of model in detecting perfect plasticity is 2.6%, bilinear plasticity is 3.6% and transverse isotropic (elastic) is 2.3%.

Two examples of misclassified are shown in following. In the Figure 15, the bilinear plastic material is miss-classified as perfect plasticity by the model. It can be seen that the material behavior is very similar to perfect plasticity and the curve has entered the plastic phase with a very small angle that is why the model mistook it as perfect plasticity. Or in Figure 16, the model is identified a bilinear plastic material as elastic material, which can be due to the angle

of the diagram between the elastic and plastic phases. In short, the problem comes from a bad discretization in the finite element method that should be refined.

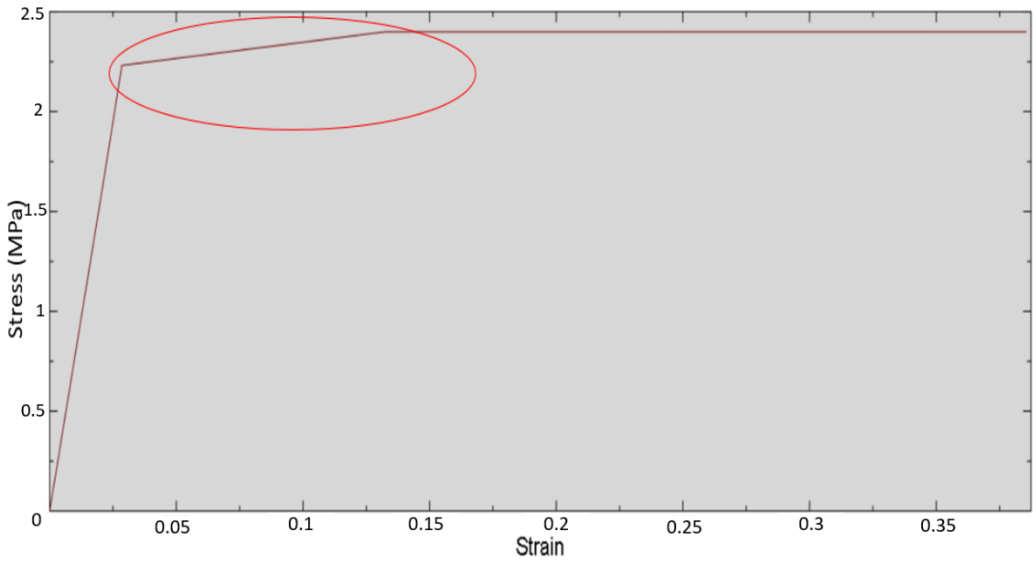


Figure 15: A bilinear plastic material which is misclassified as perfect plasticity

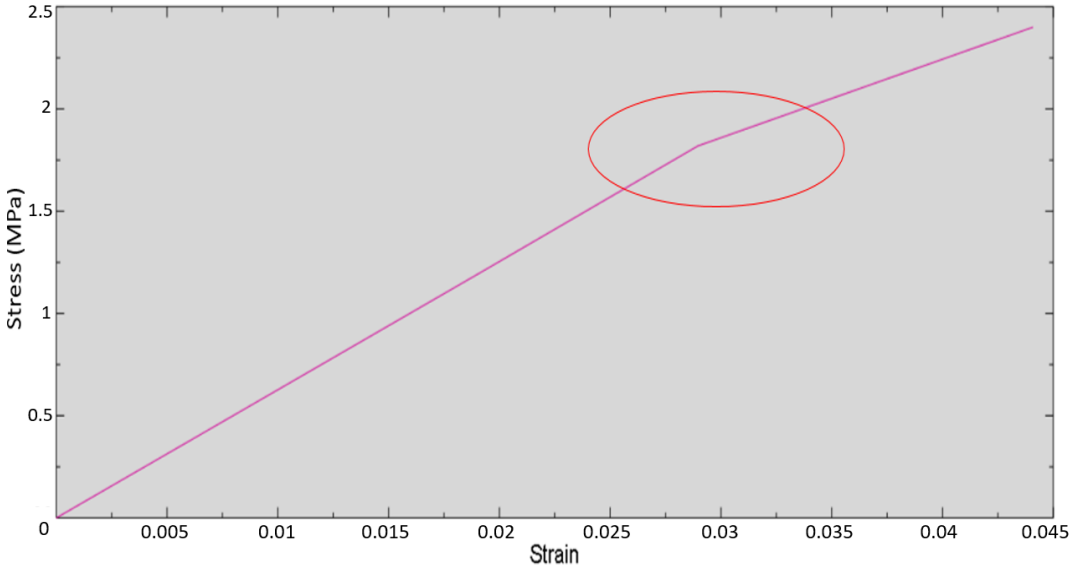


Figure 16: A bilinear plastic material which is misclassified as elastic material

After having done the classification model, now is the time to determine the mechanical properties of the materials according to the displacements.

In the next part, prediction of the mechanical properties of transverse isotropic materials is presented. These materials were selected because they have a larger number of parameters and more complex behavior than other materials in the database.

I.3.4. Identify the parameters of the transversely isotropic material

The class of transversely isotropic elastic materials will be used because its constitutive law consists of numerous parameters (5 parameters) and the ability of artificial intelligence algorithms can be challenged facing a relatively complex behavior. In a plane, this class of material is characterized by an axis that allows to orient longitudinal and transverse Young's modulus, the Poisson's ratio as well as the shear modulus. An overview of these materials is shown in Figure 17. These parameters will be the input of the finite element model. The value ranges for these parameters are globally inspired by commonly used materials where Young's modulus for both directions will be between 50 and 500 MPa, the shear modulus will vary between 50 and 200 MPa, and finally the values of the Poisson's ratio will be in the interval [0.06-0.4]. The angle of the axis of the transverse isotropy can be arbitrary, thus ranging from 0 to 180°. The distribution of mechanical properties in the database are not homogeneous and the data are denser in the range of materials that are mostly found in nature, the data distribution is shown in Figure 18.

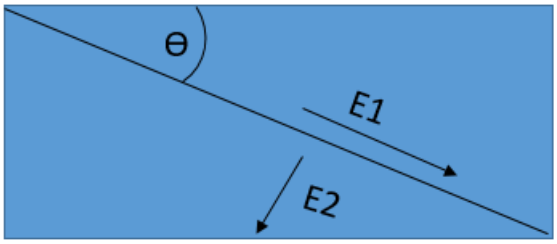


Figure 17: transversely isotropic elastic materials schematic

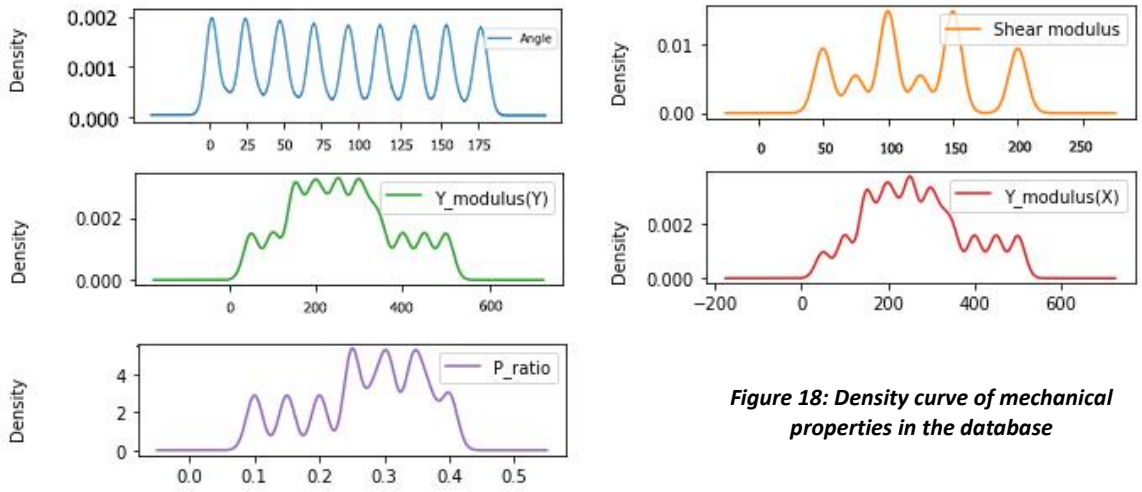


Figure 18: Density curve of mechanical properties in the database

To summarize, 17680 samples with different properties are generated from the finite element analysis. Each case is unique by combining the dispersion of all the parameters. This data number is the minimum data by which the model can have proper training phase, and the reason is shown in Table 4. With this mapping, a wide range of data is generated that helps to build a coherent database, also the generalization capacity of the model built with a complete database is very high.

I.3.5. First proposal: Machine Learning exploiting the displacement field

Several strategies were considered to determine the mechanical properties of transverse isotropic materials. These strategies include machine learning and deep learning which are described below.

Indeed, the goal is to use exclusively the displacement field and the reaction force (obtained on the point A Figure 7) without any particular treatment. From an experimental point of view, these displacements could be derived from image correlation for example, and the reaction force can be deduced by the information given by the machine test. In order to train the model, 70% of the data is allocated to the training phase, 20% of the data to the validation phase, and 10% to the model test.

In the first scenario, the data such as displacements and reaction force are entered to the model in form of vector in order to determine the mechanical properties of the materials including, the Poisson's ratio, Young modulus in the longitudinal and transverse directions of the isotropy axis, the shear modulus as well as the orientation angle of the transversely isotropic material. The model scheme is shown in Figure 19. An iterative strategy is used to adjust the hyperparameters so that to reduce the calculation time. Four models are built in parallel, in each, the loss function, the number of epochs (100 epochs), and the number of batch size (32 batch) are considered as fixed parameters. Then, in each model, the optimizer is changed to find the best ones. In the following, after finding the best optimizer and loss function, these hyper parameters are considered constant, and in the next step, the best batch size is found by keeping the number of epochs constant. Then the number of data required for training the model is found. Then best value of epoch is determined by changing the number of the epochs. The results are shown in Table 4 and the model specifications are shown in Table 5. The computation time for setting hyperparameters and completing the training phase of the identification model by the server computer (explained in the classification of the material behavior part) is nearly 45 hours.

The loss function curves in the training phase are shown in Figure 20, which shows that the training and validation curves have converged but did not tend to zero, indicating that the model could not find a good correlation between input and output data.

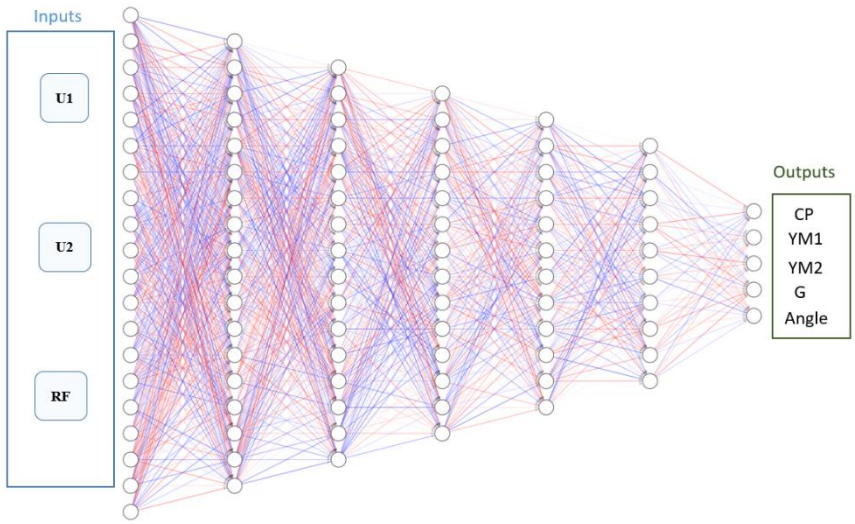


Figure 19: The schematic of the used (ANN) model to determine the mechanical properties

Table 4: Received responses from the identification model for different configurations

Loss_function:		mean_squared_error	mean_absolute_error	huber_loss	log_cosh
Metrics:	Training MSE	3.41	5.11	7.21	9.21
	Validation MSE	4.27	8.63	5.09	10.4

Optimizer:		Adagrad	RMSprop	adam	nadam	adamax	adadelta
Metrics:	Training MSE	7.063	6.18	3.12	2.76	6.37	6.24
	Validation MSE	6.75	8.15	2.87	3.532	6.48	5.92

Batch Size:		16	32	64	128	256
Metrics:	Training MSE	2.7	2.92	2.14	1.52	1.72
	Validation MSE	2.37	2.26	1.85	1.32	1.44

Epochs:		25	50	100	150	200	250	300
Metrics:	Training MSE	1.24	1.22	1.22	1.21	1.2	1.2	1.119
	Validation MSE	1.24	1.21	1.22	1.21	1.19	1.2	1.18

Number of data		2000	4000	8000	16000	17680
Metrics:	Training MSE	9.38	8.61	5.11	2.3	1.24
	Validation MSE	8.9	8.43	5.28	2.45	1.24

Table 5: (ANN) model specifications to determine the mechanical properties.

Model specifications	
Learning method:	Supervised learning
Algorithm:	Regression via "neural network"
Aim:	Identify the material properties
Number and type of layers:	7 Dense layers
Loss function:	mean_squared_error
Activation function:	relu
Optimization function:	adam
Metric:	mean_squared_error

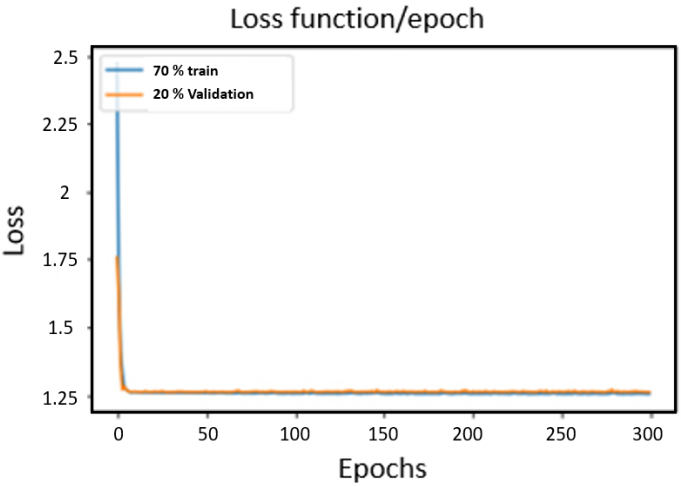
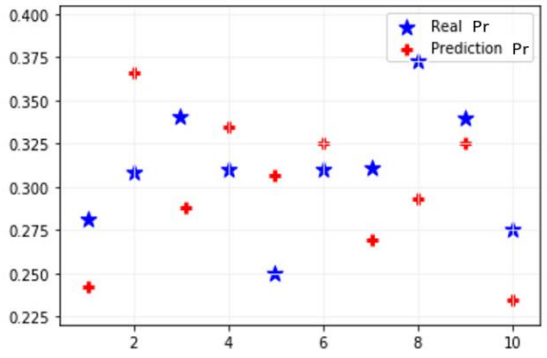


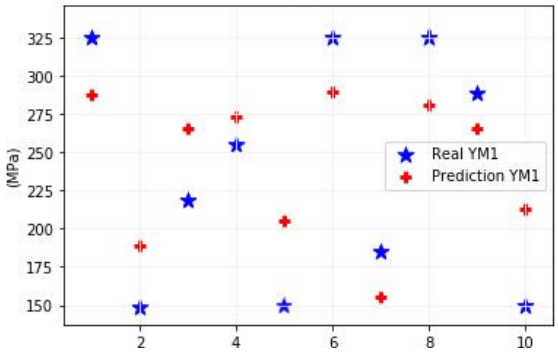
Figure 20: Loss function curves of (ANN) to determine the mechanical properties.

After the training phase, although, the model does not have a proper evolution in the training phase, it is tested by the test phase data to evaluate its performance facing new data. 10 tests were randomly selected from the test database, but it did not yield acceptable results.

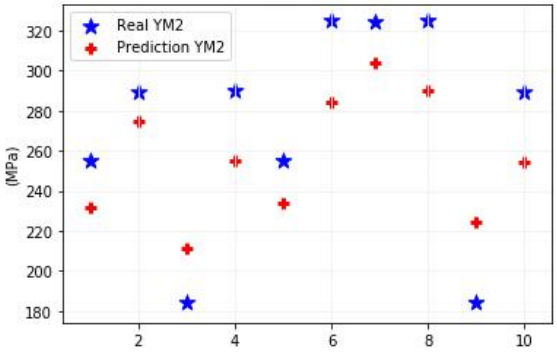
These results are shown for each of the mechanical properties in Figures 21 a to 1e.



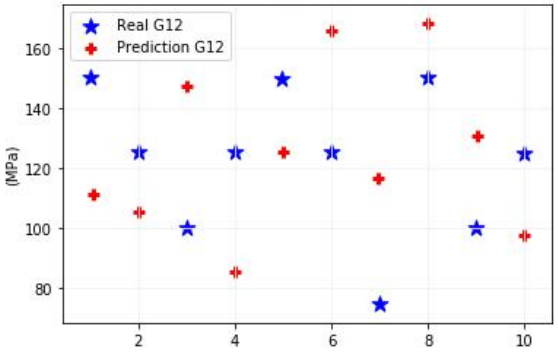
(a)



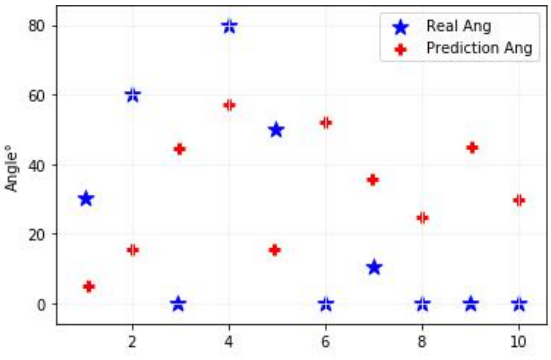
(b)



(c)



(d)



(e)

Figure 21a-e: Results of the test phase to determine the material properties.

- ***a) Results of the test of the model to determine the Poisson's ratio.***
- ***b) Results of the test of the model to determine the Young's modulus (in longitudinal direction).***
- ***c) c: Results of the test of the model to determine the Young modulus (in transverse direction).***
- ***d) d: Results of the test of the model to determine the Young modulus (in transverse direction).***
- ***e) e: Results of the test phase to determine the shear modulus.***

In these figures, the blue stars show the actual results, and the red stars the predicted values in the test phase.

The error rate for Poisson's ratio is 10.3%, Young modulus in the longitudinal direction is 9.5%, 11.5% in the transverse direction, the transverse angle is 12.7%, and the shear modulus is 14%. In fact, it can be said that since the spatial form of the input data is not preserved, the model could not find a relationship between displacements and mechanical properties. For this reason, another method must be used to deal with this issue.

I.3.6. Second proposal: deep learning exploiting displacement fields - 2D Convolutional Neural Networks (2D-CNN)

In fact, since acceptable results were not obtained in the previous section, it was decided to preserve the spatial form of the input data, including the longitudinal and transverse displacement fields in form of matrix in different layers, that is why 2D-CNN is used, given that in the previous scenario, the data was entered into the model in rows form.

Convolutional Neural Networks are a powerful artificial neural network technique. They are popular because researchers are achieving state-of-the-art results on difficult computer vision [77][121][122].

In fact, in terms of the structure, 2-dimensional convolutional networks are composed of perceptron that are placed in a two-dimension plan.

Convolutional neural networks, not only maintain spatial shape of the input matrix in the network, but also establish correlation between the input matrix arrays. Here are some benefits of using convolutional neural networks:

- They use fewer parameters (weights) to learn than a fully connected network.
- They are designed to be invariant to object position and distortion in the scene.

- They automatically learn and generalize features from the inputs.

A model made of (CNN) generally consists of the following parts:

- Convolutional Layers
- Filters
- feature maps
- Pooling layers
- Fully connected Layers

According to the needs of the problem and the described algorithms, now is the time to build a model to assess the ability of these algorithms to deal with a mechanical problem.

The AI scheme for the second strategy is summarized in Figure 22. It consists of six 2D-convolutional layers and various filters that are applied to make the input data clear for the model. Since the goal is prediction, the activation function "linear" are used in the first two layers and the "relu" activation function is used in the other layers. Then, seven "dense" layers with "relu" activation function are used. The used optimizer is "adam" and the loss function used is "mean_squared_error". The strategy of the previous part is used to adjust the hyperparameters so that to reduce the calculation time, again four models are built in parallel, in each, the loss function, the number of epochs (100 epochs), and the number of batch size (32 batch) are considered as fixed parameters. Then, in each model, the optimizer is changed to find the best ones. In the following, after finding the best optimizer and loss function, these hyper parameters are considered constant, and in the next step, the best batch size is found by keeping the number of epochs constant. Then best value of epoch is determined by changing the number of the epochs. The results are shown in Table 6 and the model specifications are shown in Table 7. The computation time for setting hyperparameters and completing the training phase of the identification model by the server computer (explained in the classification of the material behavior part) is nearly 80 hours.

In order to clarify the behavior of the model, the loss function-epoch curves during the learning phase are shown in Figure 23, for each target (Poisson's ratio, Angle, Young modulus 1, Young modulus 2, Shear modulus). But Figure 23 shows that the trend of the model is not good and indicates the lack of correct training of the model. It is evident that the loss function curves have not converged. However, the rate of decrease in the loss function is closer to zero compared to the previous architecture. This divergence shows that the model cannot find a

correlation between the inputs and the outputs. This means that the model does not have enough information to find the mechanical properties. Other works has been done on the AI scheme by increasing the number of layers in particular, which did not solve the problem.

Table 6: Received responses from the identification model for different configurations

Loss_function:		mean_squared_error	mean_absolute_error	huber_loss	log_cosh
Metrics:	Training MSE	0.17	0.73	1.49	1.56
	Validation MSE	0.35	1.07	1.8	1.71

Optimizer:		Adagrad	RMSprop	adam	nadam	adamax	adadelta
Metrics:	Training MSE	1.55	2.09	0.09	0.27	2.28	3.91
	Validation MSE	1.89	2.32	0.14	0.68	3.18	4.78

Batch Size:		16	32	64
Metrics:	Training MSE	0.04	0.025	0.034
	Validation MSE	0.081	0.04	0.047

Epochs:		25	50	100	150	200	250
Metrics:	Training MSE	0.008	0.004	0.002	0.001	0.0008	0.0007
	Validation MSE	0.028	0.021	0.019	0.019	0.024	0.018

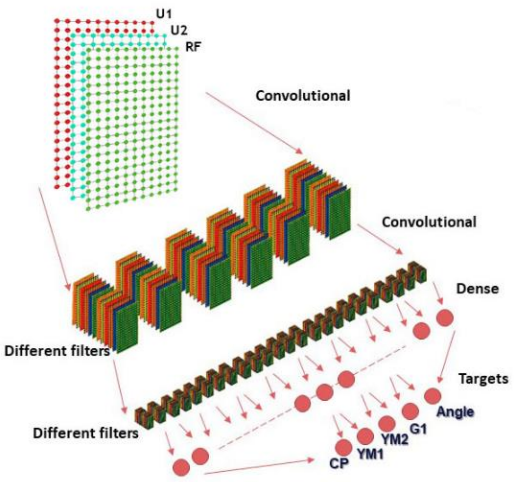
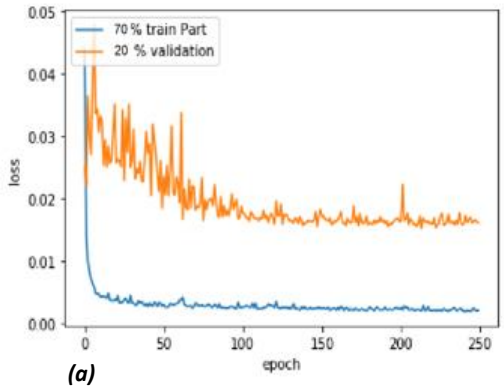


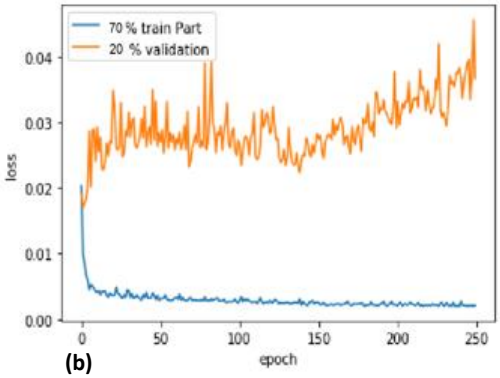
Figure 22: The 2D-CNN schema of the used meta-model to determine the mechanical properties

Table 7: (2D-CNN) model specifications to determine the mechanical properties

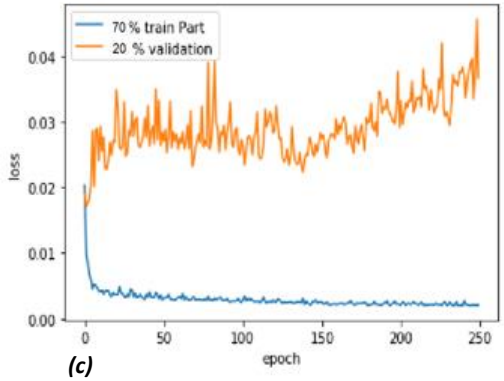
Model specifications	
Learning method:	Supervised learning
Algorithm:	Regression via "CNN"
Aim:	Identify the material properties
Number and type of layers:	6 CNN and 7 Dense layers
Loss function:	mean_squared_error
Activation function:	linear, relu
Optimization function:	adam
Metric:	mean_squared_error



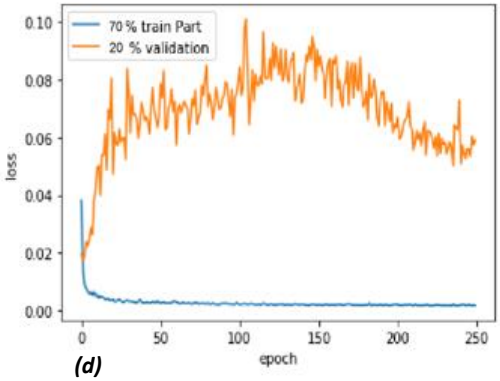
(a)



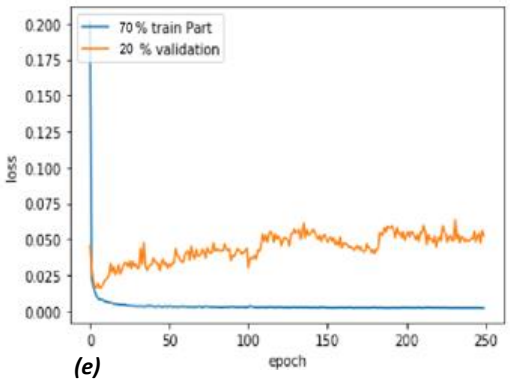
(b)



(c)



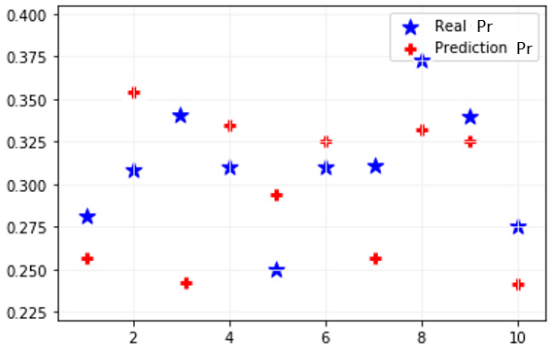
(d)



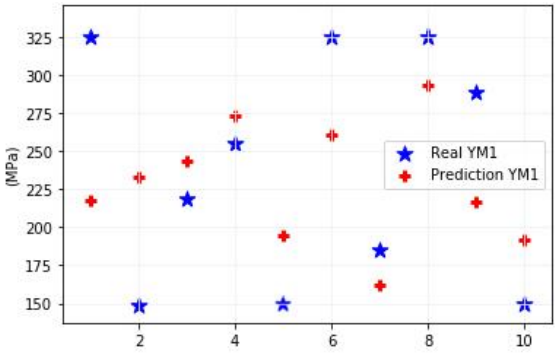
(e)

Figure 23a-e: Loss functions to determine mechanical properties using displacements and reaction force at point A.

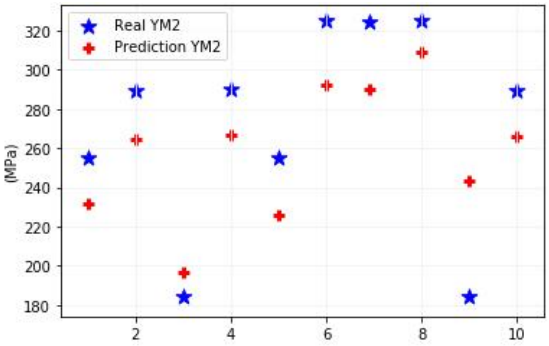
- **a)** ***Loss function curve to predict Poisson's ratio***
- **b)** ***Loss function curve to predict Young's modulus (in longitudinal direction)***
- **c)** ***Loss function curve to predict Young's modulus (in transverse direction)***
- **d)** ***Loss function curve to predict the shear modulus***
- ***e)*** ***Loss function curve to predict the orientation angle***



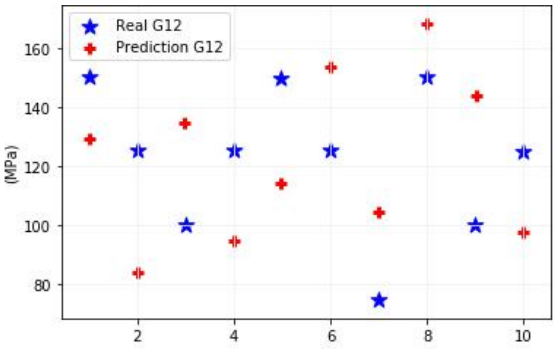
(a)



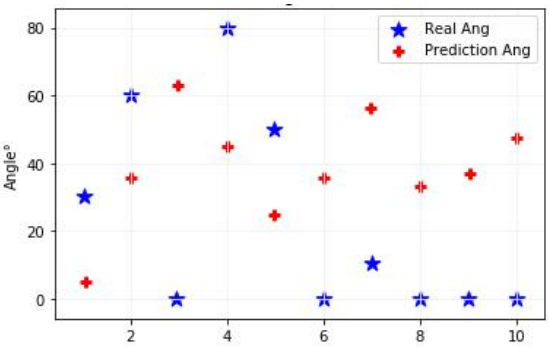
(b)



(c)



(d)



(e)

Figure 24a-e: Results of the test phase to determine the material properties

- ***a) Results of the test of the model to determine the Poisson's ratio***
- ***b) Results of the test of the model to determine the Young's modulus (in longitudinal direction)***
- ***c) Results of the test of the model to determine the Young modulus (in transverse direction)***
- ***d) Results of the test phase to determine the shear modulus***
- ***e) Results of the model in test phase to determine the orientation angle***

Although the rate of decrease in the loss function during the training phase may be close to zero, the divergence of the loss function curves between the training and validation phases indicates that the model did not generalize well during training. This discrepancy suggests that the model may have overfit the training data, resulting in poor performance on unseen data. In addition, in the test phase, the model cannot provide good results, and the error rate for the Poisson's ratio is 8.3%, Young modulus in the longitudinal direction is 9.7%, 11.2% in the transverse direction, the transverse angle is 12.6%, and the shear modulus is 11.8% which confirms that the model is not able to solve this problem (Figure 24 a-e shows the results).

I.3.7. Third proposal: deep learning exploiting displacement field and strain field

As the previous results are not completely conclusive, it is then essential to physically enrich the initial model by drawing inspired "Physics Informed Neural Networks" (PINNs) [123][124], where the strain field would be introduced as an input data like the displacement fields. This strain field is deduced from another AI scheme which had an input data in the displacement fields. The last schema is also derived from a 2D-CNN model. The global scheme is presented in Figure 25. Here the purpose is to improve the performance by dividing the learning process into two separate parts. In the first part, the model predicts the strains by analyzing the displacements in the longitudinal and transverse directions.

Then the outputs of this step enter into the second part with the inputs of the first step to predict the mechanical properties. The first structure of model consists of five 2D-convolutional layers. The activation function used in the first layer is "linear" and for the other layers, the activation function "relu" is used. After the 2D-convolutional layers, four "dense" layers are applied, and the found hyperparameters for the model in (I.3.6) are used. In the second step of the model, the same model as in (I.3.6) is used, with the difference that instead of using seven dense layers, six layers are used. The computation time for setting hyperparameters and completing the training phase of the identification model by the server computer (explained in the classification of the material behavior part) is nearly 80 hours. (Table 8 shows the model specifications)

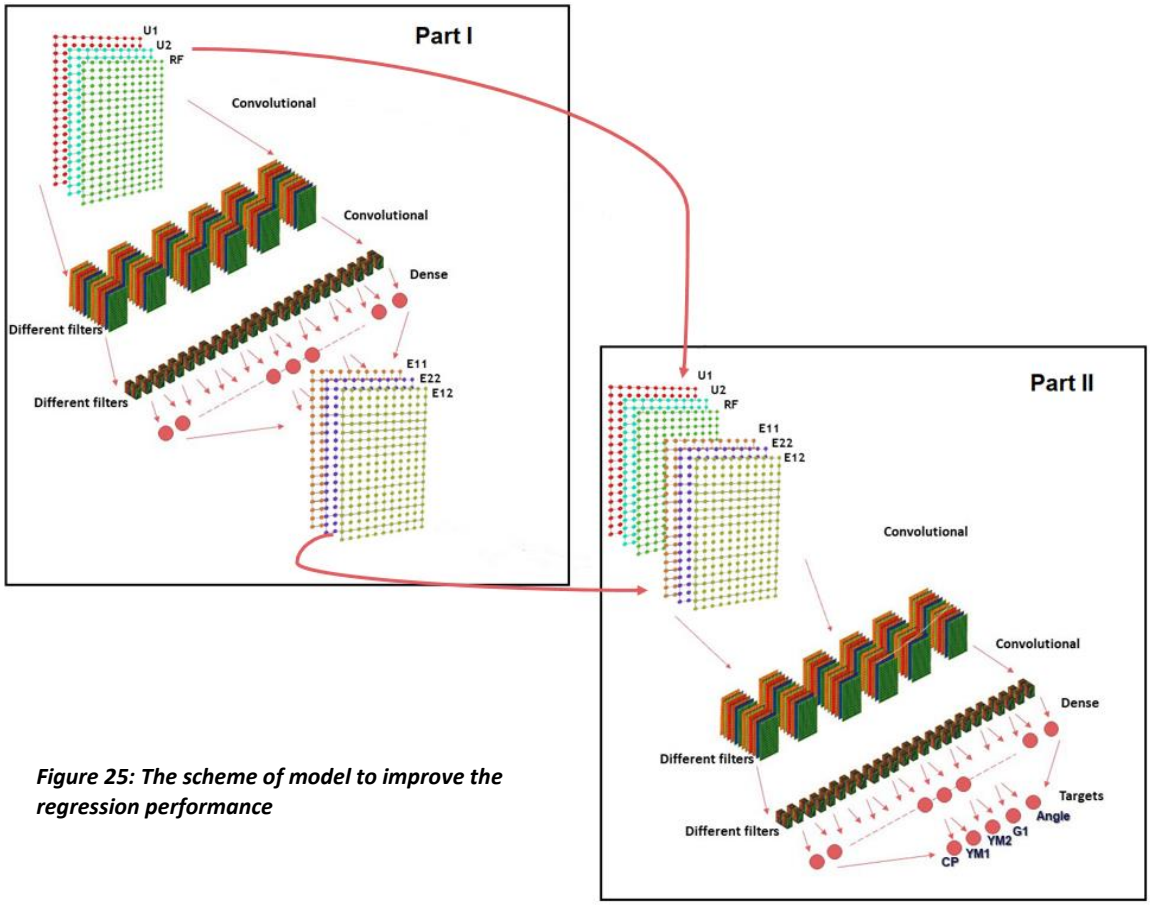


Figure 25: The scheme of model to improve the regression performance

Table 8: (2D-CNN) model specifications to determine the mechanical properties

Model specifications part I	
Learning method:	Supervised learning
Algorithm:	Regression via "2D-CNN"
Aim:	Identify the material properties
Number and type of layers:	5 CNN and 4 Dense layers
Loss function:	mean_squared_error
Activation function:	linear, relu
Optimization function:	adam
Metric:	mean_squared_error

Model specifications part II	
Learning method:	Supervised learning
Algorithm:	Regression via "2D-CNN"
Aim:	Identify the material properties
Number and type of layers:	6 CNN and 6 Dense layers
Loss function:	mean_squared_error
Activation function:	linear, relu
Optimization function:	adam
Metric:	mean_squared_error

In the following, the behavior of the model is verified in two parts, first with the analysis of the loss function and then with the test data. In fact, in this section, the displacements of the test data samples are given to the artificial intelligence model to check the generalizability of the model. The first parameter to analyze is the Poisson ratio, and loss function curves are shown in Figure 26. As shown, the curves tend to zero and indicates a good learning operation of the

model. After completing the training phase, the effectiveness of the model is verified with data outside the original database.

Figure 27 compares Poisson ratio results across 10 tests randomly selected from the test data. Blue stars represent the values given to the (FEM) and red stars represent values predicted by AI. Figure 27 shows the results obtained from the model and indicates that the model is working well, also the maximum relative error for predicting this parameter is 8.5%. The model error rate for the Poisson's ratio in prediction phase is 5.37%.

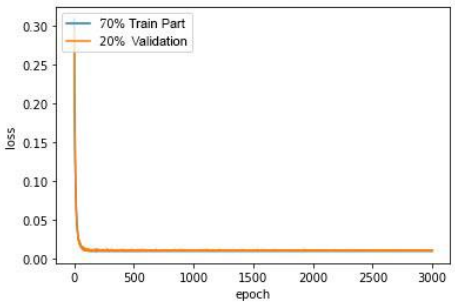


Figure 26: Loss function curve to predict Poisson's ratio

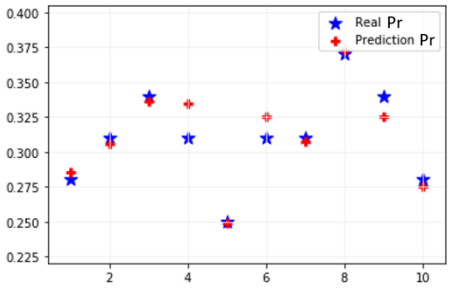


Figure 27: Results of the test of the model to determine the Poisson's ratio

Young's modulus in the longitudinal direction is the second mechanical property to predict. The loss function curve is shown in Figure 28. The training and validation curves show that the training process is well done, then the model should be evaluated with test data to ensure the accuracy of the results. The maximum error rate for this parameter in the test section of the model is 3.48%, and the average error associated with the Young's modulus in the longitudinal direction is 2.67 %, which is an acceptable percentage. Figure 29 shows the results of real and predicted values for several samples. The red stars are the results of the AI model, and the blue stars are the actual values, used for this parameter in the (FEM) model.

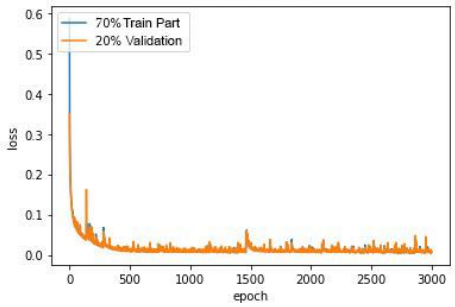


Figure 28: Loss function curve to predict Young's modulus (in longitudinal direction)

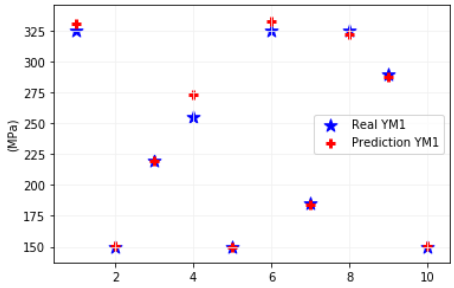


Figure 29: Results of the test of the model to determine the Young's modulus (in longitudinal direction)

The above trend for finding the Young's modulus in the longitudinal direction is repeated for finding the Young's modulus in the transverse direction. Figure 30 shows the training and

validation curves that indicate a proper performance of the model in the training phase. In Figure 31 the test results are shown, and it should be noted that the maximum error rate for this target is 4.79%. Also, the average error rate of the model is 3.47% which provides very good responses from the model.

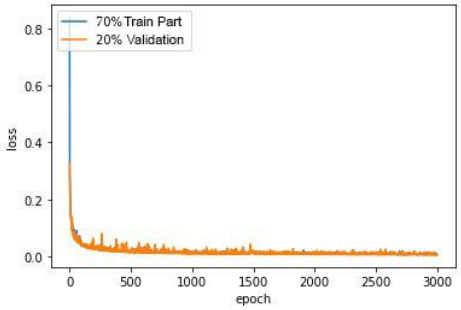


Figure 30: Loss function curve to predict Young's modulus (in transverse direction)

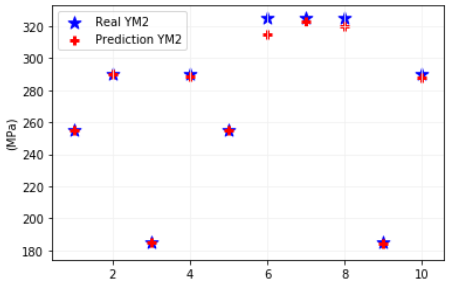


Figure 31: Results of the test of the model to determine the Young modulus (in transverse direction)

The model training process and the loss function for the prediction the shear modulus is shown in Figure 32 and the model test results are shown in Figure 33. The error rate is 1.98%, which is a good indication for the efficiency of the model. The performance of the model in this section is quite well as the maximum value is 3.5%.

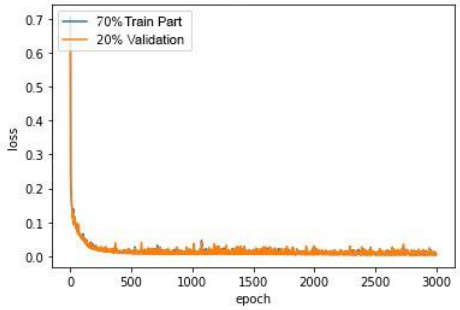


Figure 32: Loss function curve to predict the shear modulus

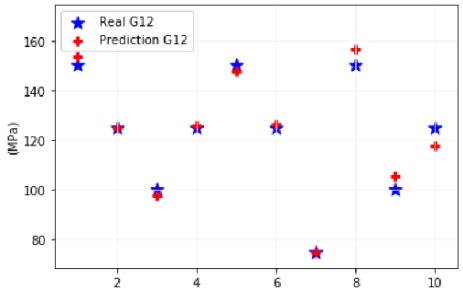


Figure 33: Results of the test phase to determine the shear modulus

The last parameter that is examined in this part is determining the angle of the transverse isotropy. The process of training and validation to predict this parameter is shown in Figure 34. The model then is evaluated by test data which these results are shown in Figure 35. The maximum error value for the model-testing phase to find this parameter is 3.2%. The average error of the model for predicting the transverse angle is 1.35% and results indicate the high accuracy of the model to predict this parameter.

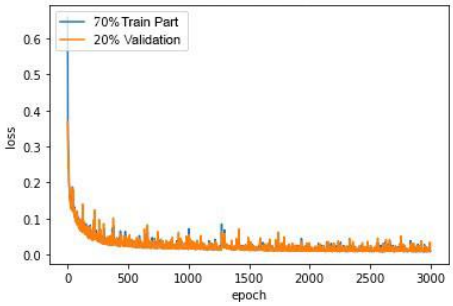


Figure 34: Loss function curve to predict the orientation angle

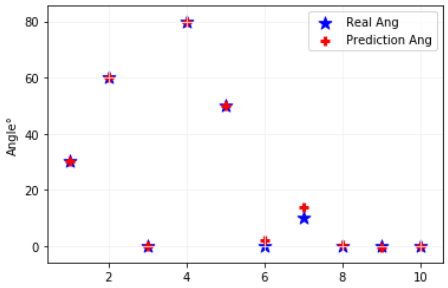


Figure 35: Results of the model in test phase to determine the orientation angle

In conclusion, quantitatively, the results are quite convincing regardless of the parameter to be identified. and the classification time in the test phase is 7ms for each case. Therefore, the step of introducing the strain field in the AI algorithm proved to be crucial.

I.4. IDENTIFICATION OF INFLUENTIAL AREAS FOR PARAMETER DETERMINATION

After making predictions in order to find the mechanical properties with respect to displacements and strains, it is important to visualize the zone on the surface that are more important for the algorithm and in order to find a logical correlation between inputs and outputs can lead to a better understanding of the physics of the problem. In fact, the goal is to understand the performance of the algorithm by visualizing the zones that have more weight on the structure of model. In order to determine and illustrate these zones, it is necessary to find the weight of each input elements of the model. For this reason, a machine learning algorithm (feature sensitivity also known as feature contribution in XGBoost which is an implementation of gradient boosting machines [125]) is used. In fact, in this process, each of the constitutive law parameters is analyzed separately. For this purpose, displacement and deformation matrices are converted into vector form and are considered as the input for the feature sensitivity algorithm, then each of the parameters is considered individually as a target so that the feature sensitivity is determined. Then the results were converted into matrix form to show their spatial form. Finally, after determining the weight of each element, these weights are plotted by assigning color as the weight.

I.4.1. Influential zones for finding Poisson's ratio

First, the zones that are most influential in order to find the responses for prediction the Poisson's ratio is investigated. The results are shown in Figure 36 by associating a dark color to the places where the zone is influential with respect to the search for the value of the Poisson's ratio. As it can be seen, the algorithm used a special pattern to determine the Poisson's ratio. The edges and a pattern of oblique lines in the central areas have the most presence.

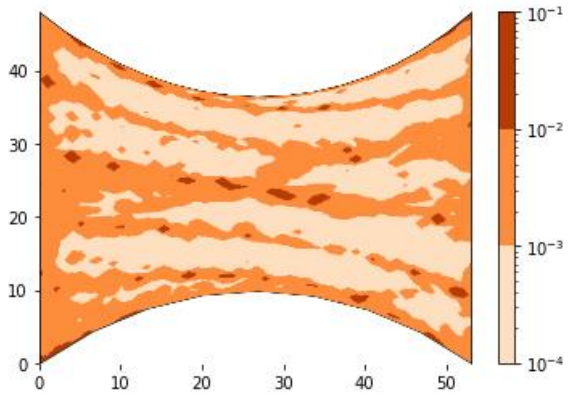


Figure 36: The influential zone for the algorithm to predict the Poisson's ratio

I.4.2. Influential zones for finding Young's modulus in the longitudinal direction

The Young's modulus in the longitudinal direction is the second target to be analyzed. The strategy used in the previous section is also used for this section. The aim is to better understand the performance of the algorithm to make clearer the learning process of the deep learning model. In Figure 37 more important zones for the algorithm are illustrated. The pattern of diagonal lines that start from the top left and continue to the right is shown in the Figure 37, in addition, a cloud of points in the center of the image tilted to the left have more effects in this analysis.

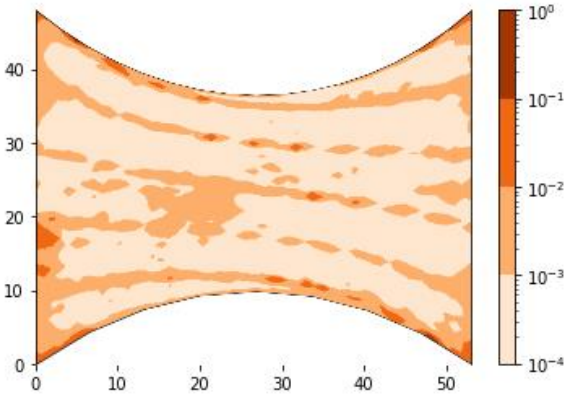


Figure 37: The influential zone for the algorithm to predict the Young's modulus in the longitudinal direction

I.4.3. Effective zones for finding Young's modulus in the transverse direction

The next goal is to find more significant areas for the Young's modulus in the transverse direction. It can be derived from Figure 38 that the density of the lines is higher on the top and bottom edge and less density towards the center.

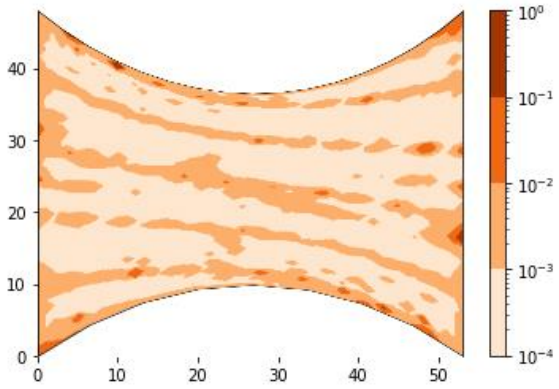


Figure 38: The zone importance for the algorithm to predict the Young's modulus in the transversal direction

I.4.4. Effective zones for finding Shear modulus

In this part, the effective zones for the algorithm in order to predict the shear modulus are studied and analyzed. Figure 39 shows the active points are more in the center and corners of the surface. The point to be made is that the more effective areas are along the load applied to the geometry.

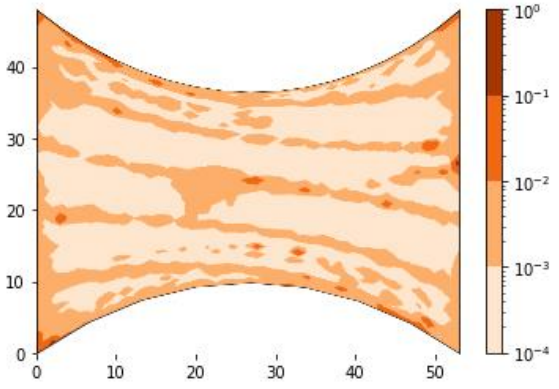


Figure 39: The zone importance for the algorithm to predict the Shear modulus

I.4.5. Effective zones for finding transversal angle

The last parameter to be examined is the transversal angle. The diagonal lines at on the top and bottom of the surface indicate the most important zones for the algorithm in order to find a link between inputs and transversal angle. Also, a horizontal line that is propagated along the surface is more important for the algorithm, this behavior was not observed in other parameters. (Figure 40)

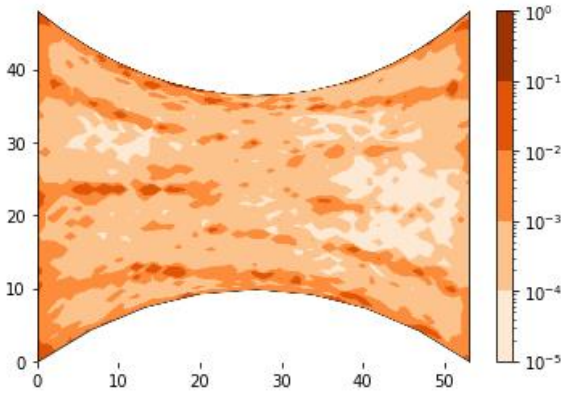


Figure 40: The zone importance for the algorithm to predict the Transversal angle

By looking at all the influence maps (Figures 36-40) on the different parameters, a tendency seems to show that the important zones to consider are almost the same. In particular, through the lines that cross the central zone of the specimen which are systematically present on all the figures. In fact, by looking at the results, it can be inferred that the model has been able to define the mechanical properties by using some specific points and not by using all displacements so that the multi-scale nature of the approach is considered. Also, according to the results of the important zones, some limited zones control the behavior of system, and it can be stated that the deep learning algorithm finds the answers in a larger scale by examining and considering the behavior of the system in the smaller scale. Therefore, the position of each point plays a significant role in finding answers for the model.

I.5. CONCLUSION

In this chapter, the capabilities of Artificial Intelligence and Machine Learning in mechanical engineering were demonstrated through the example of material behavior understanding. Indeed, a method based on artificial intelligence is proposed for two goals, first to classify the material behavior and second to identify the parameters of an elastic type behavior law in a transverse isotropic material framework. The study is conducted in a plane

where the objective was to determine the constitutive parameters and the main angle of the material in the manner of an experimental test. From the obtained results, we highlight several points.

- First, it is necessary to optimize the input data. By taking only the displacement field and even by adding layers in the schema or by modifying the hyper-parameters, the meta-model does not work as expected. It was necessary to guide the meta-model by adding a mechanical field: the deformations. This last one, even if it is artificially determined as it is the case here, allowed to obtain convincing results with an average error of 3%. This observation shows that it is essential to think and to put as much physics as possible in the construction of the meta-model.
- Second, the multi-dimensional nature of the resolution. The convolutional aspect of the 2D resolution is judicious. Other machine learning type schemes (not presented here) have been developed and have shown insufficient results. Moreover, the multi-scale nature of the resolution through the choice of the convolution data sizes is essential. By its structure, the scheme proposed here establishes a link coupling information at a fine scale with other more global information in a multi-level manner. This link has been highlighted and allows a better understanding of the crucial areas for obtaining good results.
- Third, the generalization of the proposed scheme for other forms of specimens, and other classes of materials. Studies are carried out with hyper-elastic materials or with materials showing plasticity. The results are equally convincing with multi-parameter behavior laws.
- Fourth, the execution cost. The constitution or generation of the database is the point that takes the most time. For the case of the transverse isotropic elastic material where 5 parameters of different nature had to be identified, nearly 18000 cases were generated. In a non-optimized way, both on the number and on the resolution, which was done sequentially, the simulation times took almost 24h. Even if this time may seem long, from an experimental point of view, the machines or the tests performed often remain unchanged. Therefore, the proposed base does not have to be renewed continuously. On the other hand, once the meta-model is obtained, the results are almost instantaneous, allowing for quick and efficient analysis without the need for extensive computational time. This remark allows for the quasi-direct treatment of parameter identification in the constitutive law, leading to smart testing possibilities.

Transition

After presenting a demonstrator in Chapter I and framing a mutual understanding between mechanical engineers and computer engineers, the main axis of this research which is tribology is being proposed.

Tribology is one of the main issues in the industry since it concerns all problems where there is contact. It depends on the system, but it can cause problems with wear, noise, etc.

Most of these disturbances are caused by the state of the contact surface during the contact process. One of the main factors that has a direct effect on system efficiency and system responses is the state of the contact surface, which can be caused by the surface roughness.

These conditions can be risky for the environment and human health. For this reason, it can be useful to predict the behavior of tribological systems in order to prevent disturbances. But due to multi-mechanism and multi-scale tribological systems, modeling these systems is very complicated. Due to the complexity of the problem, it can be studied from different aspects in order to understand it better. Therefore, in this research, the problem of the contact process is investigated in three phases, from three aspects.

In this chapter, by following the contact surface roughness and the mechanical state of the contact surface, the responses of the system are studied. As it was mentioned, the tribology issue is a multi-mechanical and multi-physical problem that can be analyzed from different aspects. Therefore, in order to provide a comprehensive tribology model, first, the surface mechanical state and its effect on the system's responses are investigated. Then, the contact surface state and its behavior during the contact process, and finally, the surface contact evolution is analyzed (the box highlighted in blue in Figure II is analyzed in this chapter).

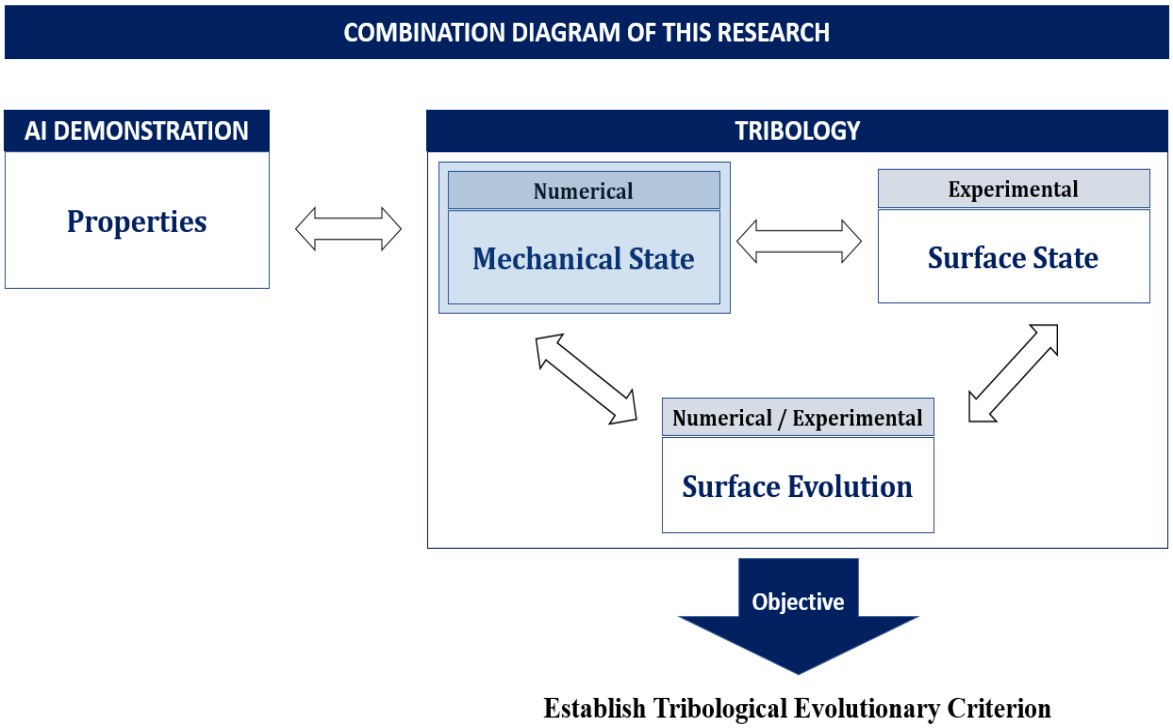


Figure II: The general scenario of the thesis, which in this chapter artificial intelligence algorithms are introduced through an example.

Here, two main goals are pursued. In the first part, by analyzing the surface roughness, the system instability risk is studied, and in the second part, the probable frequencies which can be produced by the risky surfaces would be foreseen. Finally, by analyzing the deep learning model behavior, the zones that might have a greater impact on the system disorders are determined.

II. Factors of Understanding Between the Surface State and the Vibratory Behavior

Table of Contents

Abstract

The real surfaces in the contact process are generally non-regular at different scales, which results in a complex and non-predictable behavior of the vibratory response of the system due to the presence of roughness, contact plateaus, bumps, and other factors. Moreover, considering that friction-induced vibration is a multi-scale problem, the localization of contact plays a crucial factor in determining the behavior. Thus, in order to avoid squeal problems in designing process, the engineer often starts a trial-and-error approach which limits the understanding of the phenomena and does not allow to clearly identify the relevant criteria for the "ideal surfaces" to be defined. A complete approach is proposed where, through a database generated from a multi-scale contact simulation, a link is established between the asperity field generated through the contact and the obtained frequencies. In fact, this analysis in the first part deals with classifying the roughness of the surfaces that lead to system vibration, and in the second part by analyzing the surface roughness which is at risk of vibration, the probable generated frequencies produced by this roughness are predicted.

In order to investigate this problem, deep learning algorithms and more precisely 2D-convolutional neural networks (2D-CNN) are used. To train the deep learning model, a database using finite element modeling is used, which models a pin-on-disk system.

Very satisfactory results were obtained. An accuracy of 99% in the classification section and accurate results, less than 4% error rate in the frequency prediction section, is reached by the model. Finally, a discussion on the areas of influence and the importance of the multi-scale aspect on vibratory behavior is addressed to establish a criterion.

Keywords: Multi-scale contact, Deep learning, Pin-on-disc configuration, Squeal prediction

II.1. INTRODUCTION

One of the main problems faced by the transport industries is the undesirable squeal during braking phases. Although the disc brake system has been used for more than a century, the brake squeal still remains an unsolved problem [126][127]. Squeal is due to system instability that occurs during braking [128]. This instability is caused by friction between the contact surfaces. Since mechanical interactions in the braking system are overly complex and depend on many parameters, this issue cannot be easily solved [126]. This noise does not technically disturb the performance of the braking system, however, consumers have been complaining about these annoying noises and tend to replace braking pads, which imposes huge costs on after-sales service companies [53]. In recent years, efforts have been made to investigate the effects associated with this phenomenon [129], and many methods have been used to solve it [128]. A lot of researchers have tried to study the instability of the system through various methods, such as complex eigenvalue analysis and calculating the friction effects between contact surfaces using linear elements in an asymmetric stiffness matrix [130][131] . A simplified linear mathematical model of a disc brake is presented, which investigates the possible occurrence of unstable states under an oscillating motion [132] . Other approaches include building a physical model in the laboratory [133][134], among others.

However, in the contact issue, there are strong nonlinearities due to the presence of the surface defects which are generally not taken into account [128]. Indeed, in the majority of cases, surface evolution and contact localization is globally not considered, while it is a key factor in the prediction of squeal [135]. From the “system” point of view, different research teams have been looking at the problem and working on the system responses [136][137].

There are also many methods to analyze the contact process such as**:** numerical methods [138][139]; investigating the effect of geometrical parameters [140]; a method based on fuzzy arithmetic [141]; a variance-based global sensitivity applied to the oscillator parameters [142]; predicting unstable frequencies, trying to be similar to the behavior of the experimental model [143]. But, the results are still not satisfactory due to the complexity of the problem [144]. Afterwards some researches started using finite elements method in this field by considering the contact process as a multi-scale problem and entering the contact stiffness heterogeneity to the calculations [26], a Hertz theory based model is presented, it considers the contact surfaces as ellipsoidal roughness and analyzes the frictionless rough contact surfaces [28]. For better understanding this phenomenon, porosity, material heterogeneity and contact surface heterogeneity is taken into consideration [145][146], but they were not able to meet the reality

factors as it is a multi-scale and multi-mechanical issue (tribology, system scale, etc.) which has a crucial impact on squealing. It can also be added that these methods are time consuming, and the complexity of these methods makes them not easy to use and requires a lot of specialized skills. Recently the methods for predicting brake squeal by DL have been presented [144], however they are not strong in two aspects. First, they are based on a limited set of empirical data and cannot be generalized because they are mostly case studies. Second, these models are based on observation and are not able to investigate the origin of the problems like surface topography to provide an indicator for understanding the origin of the problem. Considering these points, to have a high performance DL model, a combination of experimental data and modeled data can be a good solution. For this reason, a numerically efficient and realistic numerical model could be utilized. By incorporating these validated models, a comprehensive braking system can be accurately modeled [139]. Additionally, models that simulate pin-on-disk systems [26][28][146] contribute to the development of a reliable database. Although these methods are time consuming, but when the database is created and the DL model training is completed, using the DL model would be immediate. In short, time will be consumed only once to build a database and the DL model training but the proficiency of the built model is instantaneous.

The motivation of building up a database using the pin-on-disk system is that it is a simplified system and shows the contact problem. In addition, there are many physical models that model a pin-on-disk system which can easily be a triangle of the AI, numerical modeling and experimental modeling to solve the complex problem of contact after years of study and research.

On one hand, built models using DL methods are more flexible and can easily extract maximum information from data obtained from experiments or modeling. On the other hand, DL models have the ability to predict the behavior of the system, which makes it possible to prevent the occurrence of such disturbances. Regarding the main problem, which is surface vibration, using 2D-Convolutional Neural Network (2D-CNN) (explained in I.3.6) can be a good idea to investigate the issue of squeal. This approach is therefore used to establish a link between the defects of contact surfaces and the natural frequencies. Due to the nature of the data and the team's know-how on multi-scale numerical developments, a numerical database is generated and a “supervised learning” algorithm is used because the used data is labeled. Moreover, the powerful architecture of 2D-CNN is used due to the input dimensions (2D) which is the surface roughness.

In this research, the modeling phase consists of two main parts. In the first part, by using classification algorithms based on the 2-DCNN architecture, the surfaces that have the risk of making vibration on the system scale are identified.

In the second part, by using the 2-DCNN, the model predicts the disturbing generated frequency by analyzing the roughness of the surfaces that increases the risk of system vibration. Usually, in deep learning models that use 2-DCNN, the input data is an image, which usually does not consider the multi-scale effect, but here, the model input is the surface roughness matrix, which considers the roughness as one of the main parameters in the occurrence of disturbances. Not only using an image as input cannot be a good idea in this issue, but also using the roughness matrix allows the multi-scale effect to enter the deep learning model.

It should be considered that the evaluation of the robustness of the developed meta-model will be based on known criteria, such as the Receiver Operating Characteristic (ROC) and the confusion matrix.

Once the relevance of the meta-model (a meta-model is a model that is trained to make predictions about other models.) is ensured, it will be used to propose a selection of criterion for the crunchy configurations and thus show the multi-scale dimension of the problem.

II.2. GLOBAL VIEW OF SOLVING THE PROBLEM

A scenario overview is shown in Figure 1. This process consists of different steps:

- Step 1: the surface roughness is produced by numerical methods.
- Step 2: this data enters the numerical model, which according to their asperities may cause instability of the model.
- Step 3: the received responses from the model include stability or instability, also system frequencies for each surface are recorded in the database.
- Step 4: the data is used for training the AI model in classification part.
- Step 5: the data is used for training the AI model in regression part.
- Step 6: Finally, by analyzing the results received from the AI model, the mechanism of this model is interpreted. Each of the mentioned steps is explained in detail in pursuit.

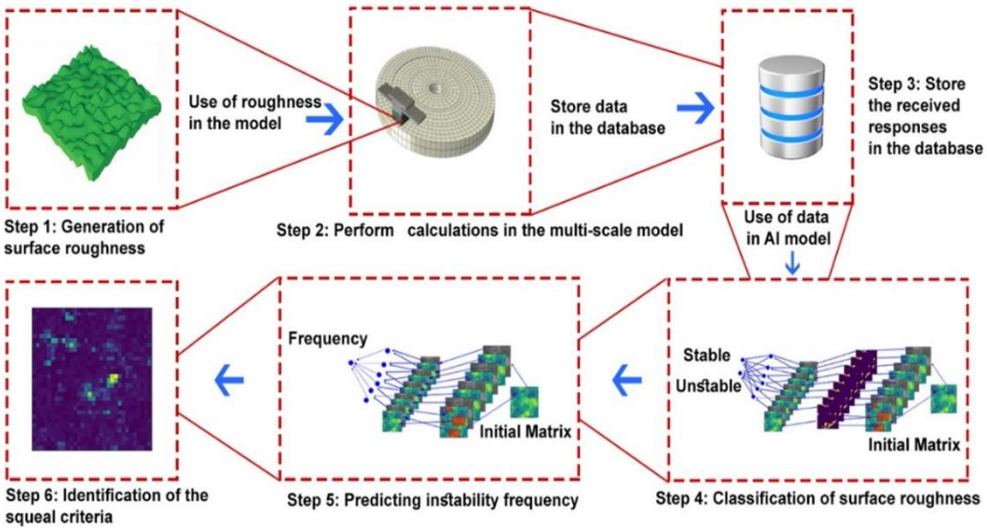


Figure 2: Global Strategy to link surface defects to vibratory behaviors

II.2.1. Step 1: Generation of the roughness surface: creation of the database

The input parameters of the global strategy are the topographies of the contact surface. These topographies are generated numerically inspired by experimental measurements. Indeed, depending on different factors such as manufacturing processes or wear during its use, the surfaces can be different. These surfaces have been generated numerically using fractal techniques [147]. Indeed, a self-affine surface is generated using a power-law spectral density.

$$S(\mathbf{k}) = S_0 \, |k|^{-2(H+1)} \qquad \textbf{(1)}$$

S_0 is a constant and H is the Hurst exponent which is related to the fractal dimension $Df = 3-H$. A randomly rough surface can be generated with any given spectral density.

$$h(x) = \sum_k B(k) e^{i.(\frac{2\pi}{L} k.x + \emptyset(k))} \qquad \textbf{(2)}$$

$B(k) = 2\pi(\frac{s(k)}{A_0})^{1/2}$, A_0 is the surface area, L is the square root of A_0 and ∅(k) are independent random variables which are uniformly distributed in the interval [0, 2π]. In order to have a better view, two of these surfaces are illustrated in (Figure 2). The fractal dimension is 0.59 (deduced by experimental data from [148]), the cutoff wave vectors are 2 and 10 and the surface is Gaussian with a root mean square which is a parameter. The amplitude of roughness can also be modified with a range from a few micrometers to a few hundred micrometers.

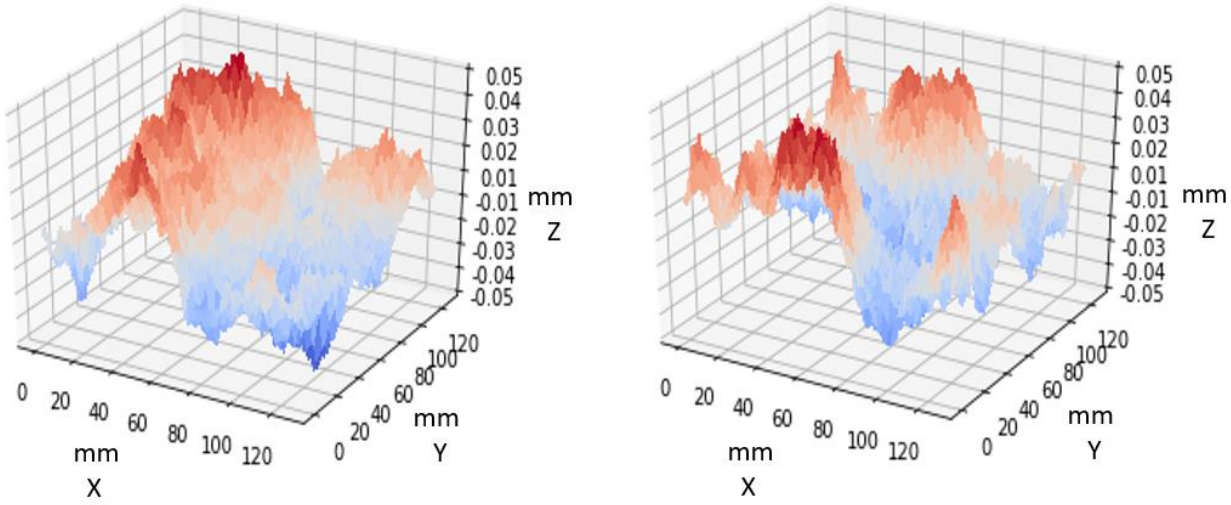


Figure 2: Two examples of numerically generated surface roughness.

II. 2.2. Step 2: Presentation of the multi-scale model (Finite element model for database generation)

A multi-scale model is used which considers surface defects within a pin-on-disc [28] [146]. It consists of a disc, topped with a pad that is held to a flexible plate (Figure 3). A force (300N) is applied to the ends of the plate while the disc is rotating around its axis. The disc and the plate are made of steel while the pad has properties of a classical friction material (E=3000MPa and nu=0.2). A rotation of the disc around its axis is imposed. The overall numerical model is explained in full in [149].

A method of enrichment is being taken into account, completing the mentioned models (other enrichments are possible such as contact plates [28], the undulation of the disc [150] etc.). The disc will be considered perfectly flat. This method is based on numerical homogenization [28][151] through a local evaluative contact stiffness. To provide information, a calculation lasts ~10min (per each roughness) on a (high performance computing) HPC cluster [152].

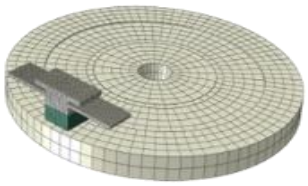


Figure 3: The contact model a disk-on-pin system.

To better understand the used multi-scale contact model, Figure 4 shows the kind of state of surfaces that are used as input data. Figure 4a shows a smooth surface while (Figure 4b) shows a rough one. The leading edge is localized at the bottom edge of both figures. Figure 4c and 4d shows the pressure fields associated with Figure 4a and 4b respectively. In the perfect case (Figure 4a), an overpressure (Figure 4c) occurs at the contact entrance, while in the case with roughness (Figure 4b), the pressure (Figure 4d) is slightly more distributed over the whole surface, mainly on the highest roughness.

In the used model, the surface of the disc is considered completely smooth, and after adding roughness on the pad surface, the results obtained from the system are recorded. For example, in Figure 4a, the surface of the pad is considered smooth and the responses show that in this case the stresses are localized at the contact entrance.

But when roughness enters the model, stresses are distributed on the contact surface, which reduces stress concentration and temperature rise in limited points that disrupts the system's efficiency [153].

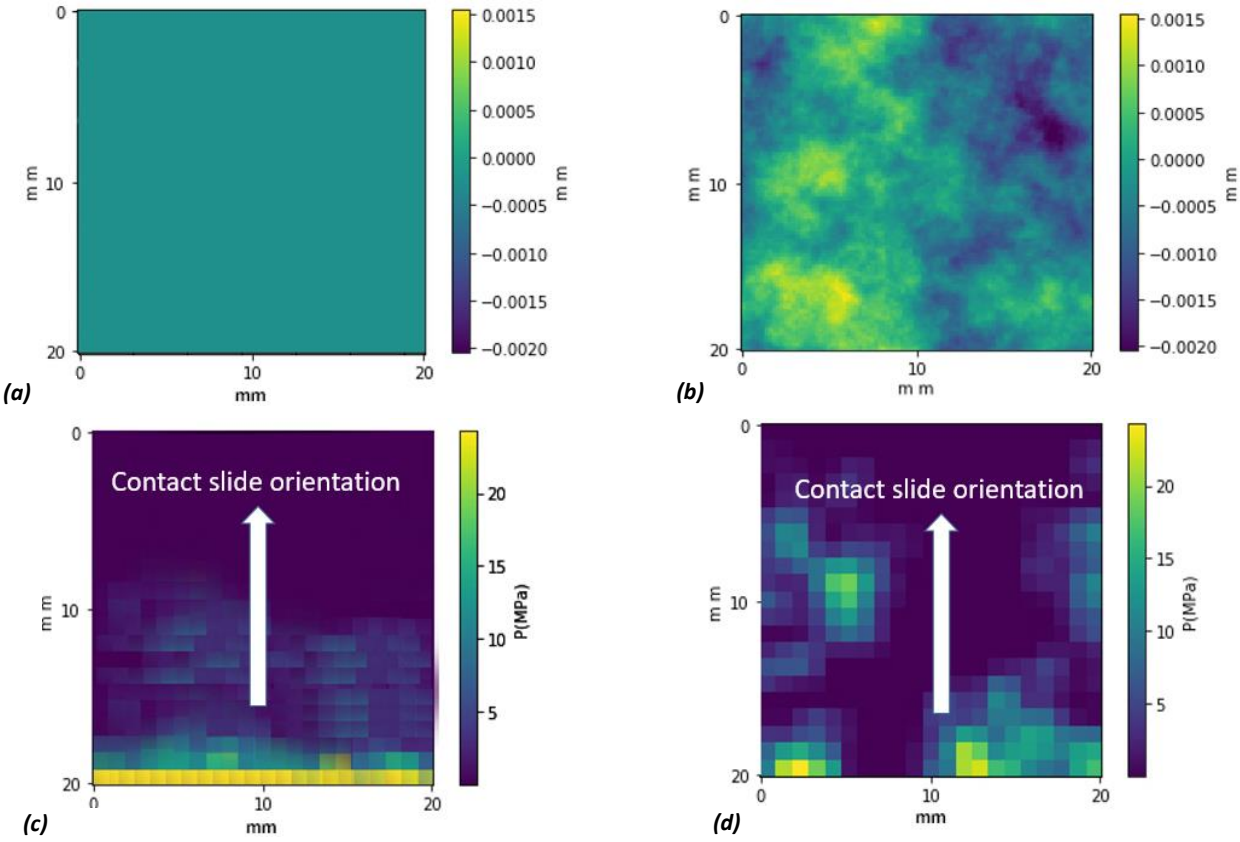


Figure 4a-d: Surface of friction pad with associated pressure distribution

- *a) Perfectly flat surface*
- *b) Rough surface*
- *c) Stress distribution for a flat surface*
- *d) Stress distribution for a heterogeneous surface*

Due to the nature of the contact problem, different surfaces generate different frequencies during the contact. For example, perfectly flat surfaces produce unstable frequencies, close to 7.5 kHz during the contact, but rough surfaces produce different frequencies (between 2.1kHz and 3.5kHz) which are sometimes stable or unstable, also unstable frequencies could be changed with the evolution of the surface. These complex behaviors during the contact make it difficult to understand this problem. Figure 5 shows some examples of frequencies generated by rough surfaces and their effects on the system behavior.

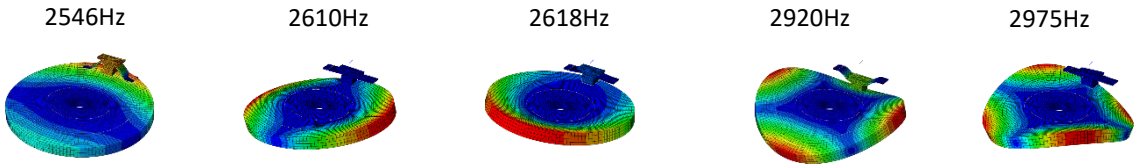


Figure 5: Different frequencies produced by different rough surfaces

II.2.3. Step 3: Data preparation and storing responses received from the finite element model

In step 3, the data is prepared and stored. In the stored database, 10,500 examples were used to train and test the model. In fact, for each roughness surface (in a matrix form, including 128 rows and 128 columns, which elements of these matrices are the surface roughness altitude for each point), the responses of the system, including the system stability and instability, and the received frequencies are stored. How to produce different roughness with different configurations is explained in [26][28][146]. The data distribution is shown in the Figure (6).

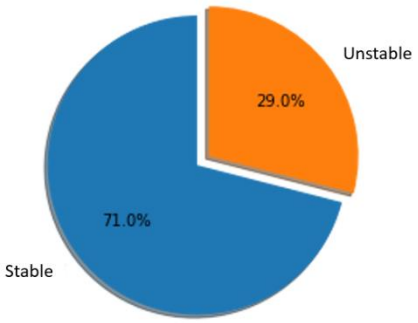


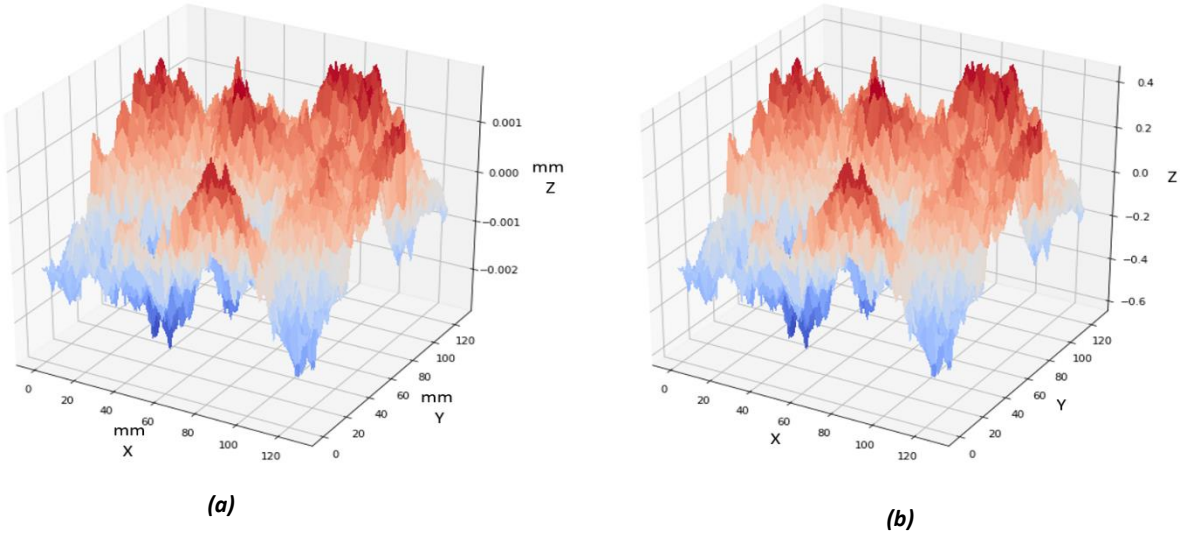
Figure 6: The percentage of stable cases compared to the unstable cases

One of the major challenges in this issue is the imbalanced data, because the number of unstable cases in the database is much less than the stable ones, 29 percent (3045 rough surfaces) unstable and 71 percent (7455 rough surfaces) stable, which caused the database to be imbalanced. In order to solve this problem, several techniques can be used.

One of the possible solutions that has been used here in facing imbalanced databases is the technique of double copying the less repeated data in the database [32].

After accomplishing this process, before entering the data into the model, a normalization step is performed which led to the data set being included in the interval [1, -1].

In fact, since the roughness data is randomly generated, each surface has different maximum and minimum values for surface roughness. In order to make the roughness be on the same scale, their minimum and maximum value from all roughness were found, then the minimum value was set to -1 and the maximum value was set to +1. The surface roughness data before normalization is shown in figure 7a and after normalization in 7b.



(a) *(b)*

Figure 7a-b: Surface roughness before and after applying normalization

a) Before applying normalization
b) After applying normalization

II.2.4. Step 4: Development of an AI model - 2D CNN model for classification: mode lock-in or not?

The aim of this step is to provide a model for classifying surface roughness, therefore, by using a sequential model, the structure of the classification model is constructed.

Due to the nature of the data, which is surface topography, 2D-Convolutional Neural Network (2D-CNN) is used. A key advantage of 2D-CNNs is their ability to preserve the spatial structure of data, as referenced in studies [144] and [154] By keeping this structure, the algorithm can accurately interpret inputs and produce the desired outputs. The model architecture for classification is shown in Figure 8.

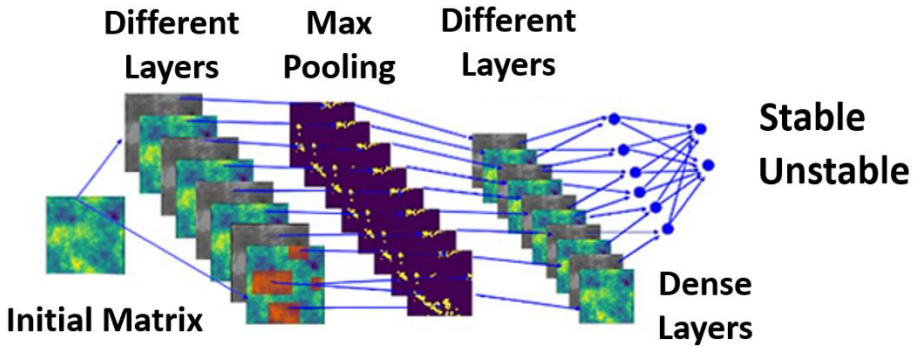


Figure 8: 2D-CNN Diagram for Classification

In this model the input data is the surface roughness represented as a matrix. In fact, the data is presented in a matrix format, with 128 rows and 128 columns. The elements within these matrices represent the surface roughness altitude for each point. In the layers of the model, several filters are applied to the initial matrix, which help to find the features helping to understanding and describing the system behavior. For example, as the higher points play the major role when contact occurs, applying MaxPooling layers can preserve information that is more useful and remove unnecessary information in the training process. Then, by flattening the feature maps, the shape becomes one-dimensional before applying Dense layers and the model tries to analyze the information obtained from (2D-CNN) layers in the best way and classifies the contact surfaces into two categories: stable and unstable.

In order to provide an optimal model, the hyper parameters of the model, such as “Loss function”, “Activation function”, “Number of epochs”, “Optimization function”, “Metrics”, etc. are adjusted using an iterative strategy which is described below.

In this strategy, to reduce the calculation time, four models were built in parallel, in each, the loss function, the number of epochs (64 epochs), and the batch size (32 batch) were considered as fixed parameters. Then, in each model, the optimizer and metric were changed to find the best ones. In the following, after finding the best optimizer, metric and loss function, these hyper parameters are fixed, and in the next step, the best batch size is found by keeping the number of epochs fixe. The best value of epoch is determined by changing the number of the epochs. The results are shown in Table 1.

Table 1: Received responses from the classification model for different configurations

Loss Function	categorical_ crossentropy	kullback_leible_ divergence	binary_ crossentropy	sparse_categorical_ crossentropy
Error %	10.01	30.2	11.11	37.2

Optimizer	Adagrad	RMSprop	adam	nadam	adamax	adadelta
Error %	13.80	20.07	8.75	11.65	28.35	22.22

Metrics	Accuracy	categorical_accuracy
Error %	7.47	5.93

Batch Size	4	16	32	128	256	450	800	900
Error %	13.85	11.83	10.9	4.93	6.86	8.47	7.72	12.54

Epoch	16	32	64	128	256	450	650	750
Error %	10.34	11.26	11.7	7.9	6.19	4.83	2.67	2.97

According to figure 9, it can be seen that the model has not been able to analyze the relationship between the surface roughness and the behavior of the system properly at the beginning when there is not a sufficient amount of data for training phase, so that the error percentage of the model is very high. The higher number of data, the lower error percentage. In other words, in order to understand the relationship between the surface roughness and the behavior of the system, the model needs to have enough examples. It is necessary to provide enough data to the model to achieve the desired error percentage which is below one percent.

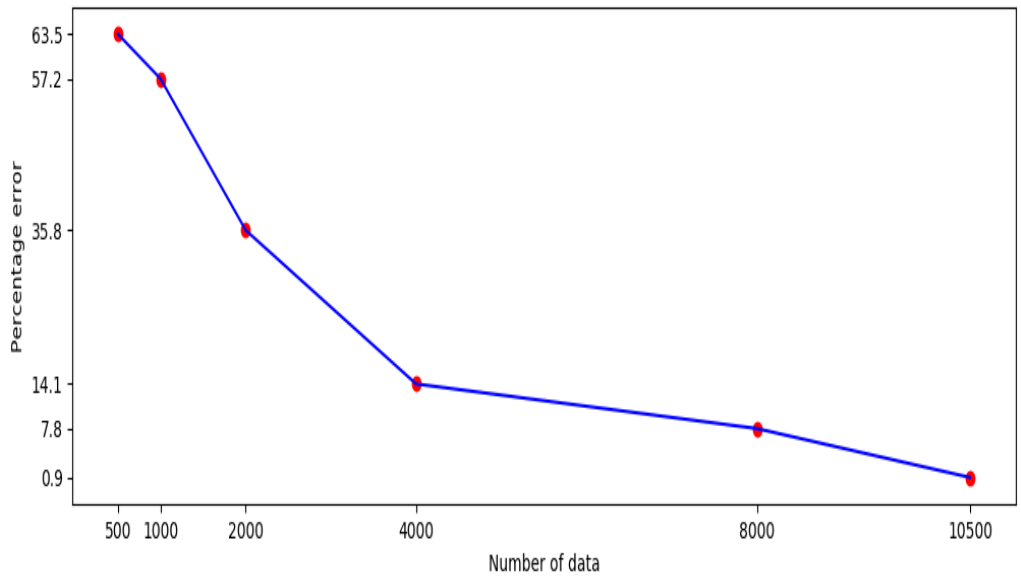


Figure 9: Number of data vs error rate

To analyze the data in order to train the meta-model, data has been divided into three parts, sixty percent (6300 rough surfaces) for training, about twenty percent (2101 rough surfaces) for validation and around twenty percent (2099 rough surfaces) for testing.

The findings indicate that the model failed to establish a strong correlation between the input and output data during the initial stage, which could be from two aspects, either the model does not fit the needs of the problem, or the provided data to the model is not sufficient. For this reason, first the number of data was increased, so the model obtained better results, and the model was tested for each number of data with different configurations, then when the model showed the best configuration with a small number of data, the number of data was increased to reach the percentage of the acceptable error which is below one percent (which is 10500 cases). Hyperparameters and final specifications of the model are shown in Table 2.

Table 2: Model specifications for classification

Model specifications	
Learning method:	Supervised learning
Algorithm:	Classification via "2D-CNN"
Aim:	Classification of contact surfaces
Number and type of layers:	6 (2D-CNN), 3 Max Pooling and 6 Dense layers
Loss function:	Categorical-Crossentropy
Activation function:	sigmoid
Optimization function:	adam
Metric:	Categorical Accuracy

The computation time for setting hyper parameters and complete training phase of the classification model by a server computer (Configuration: Memory: 128 GB, Processor: Intel® Xeon® Gold 5215 CPU @ 2.50GHz *40, Graphics: NVIDIA Corporation GV100GL [Tesla V100SPCIe 32GB] and a Samsung SSD 2 TB M.2 NVMe) is less than 40 hours.

Results of the classification model

After completing the training phase of the model (consists of training and validation), its performance is evaluated with test data. The main purpose of the classification task is to find the squeal risk of the surfaces.

Indeed, by analyzing the surface which includes the asperities and their position at the micro scale, the model seeks to find a logical relationship between the topography and the risk of system squeal at the macro scale. To evaluate the performance of the classification model, the Receiver Operating Characteristic (ROC) curve is used [155][156]. To draw the ROC curve, "True Positive Rate" (TPR) and "False Positive Rate" (FPR) are required. Also, to calculate

FPR and TPR, parameters like (“True Positive (TP)”, “False positive (FP)”, “True Negative (TN)” and “False Negative (FN)”) are needed and they can be obtained from the confusion matrix (Table 3).

Table 3: Confusion Matrix

Actual class / Predicted	**Stable**	**Unstable**
Positive	1448(TP)	0(FP)
Negative	5 (FN)	646(TN)

The confusion matrix shows that the model is able to correctly identify 1448 out of 1448 silent cases without any misidentified cases with vibration risk. On the other hand, among the 651 cases with squeal risk, 646 cases are detected correctly by the model but 5 cases are misdiagnosed.

According to the values obtained from the confusion matrix the TPR (True Positive Rate) and FPR (False Positive Rate) can be calculated to draw the ROC curve:

$$TPR = \frac{TP}{TP + FN} = 0.996558844 \quad \textbf{(3)}$$

$$FPR = \frac{FP}{FP + TN} = 0 \quad \textbf{(4)}$$

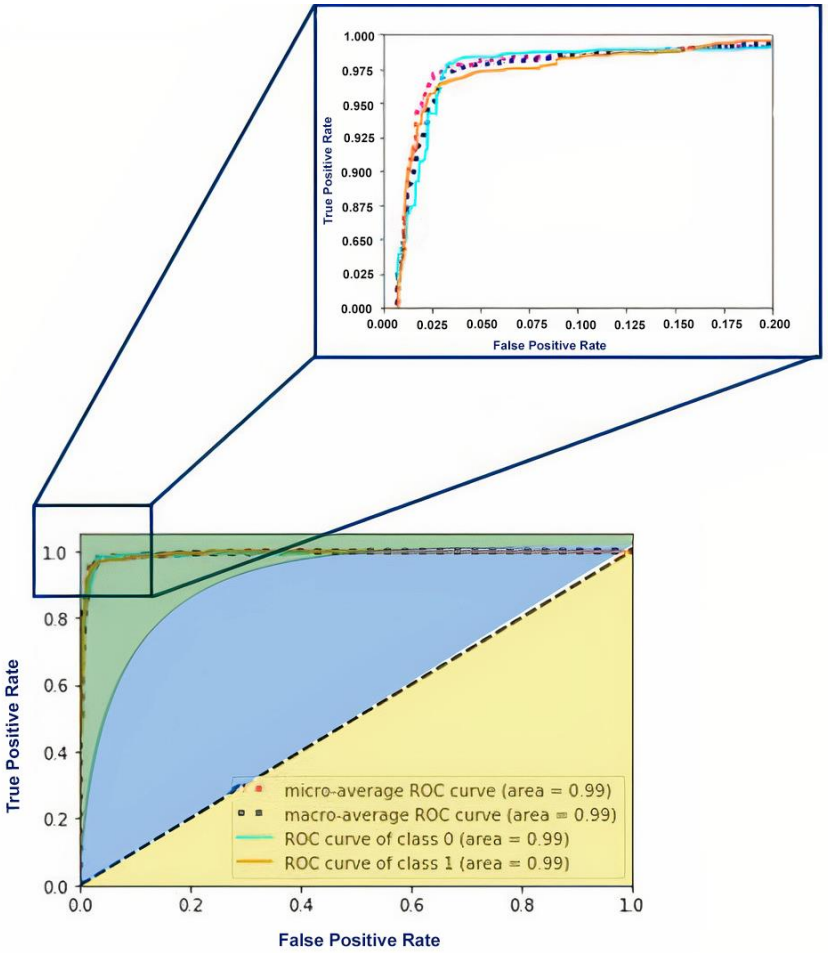


Figure 10: ROC results in categorical classification case.

As the TPR value is close to 1 and the FPR value is 0, it can be inferred that the model is efficient.

Figure 10 shows the TPR vs FPR curves which is divided to three areas: beige, blue and green, if the curves are in the beige area or below the diagonal line, the model has not been able to perform properly, and it has actually done the classification randomly.

If the curves are placed in the blue area, the model has been able to perform a relatively good classification, and if the graphs are in the green area, it means that the classification model is performing very well. As it can be seen in Figure (10), the ROC curve converges well to the upper left corner and is located in the green area. It indicates that the performance of the model is particularly good. In addition, the "Area Under the Curve" (AUC) value is close to 1, which is extremely high performance for a classification model. By calculating the "Accuracy" (ACC) of the model, the efficiency of the model can be examined from another aspect. The accuracy value is calculated below:

$$ACC = \frac{TP + TN}{TP + TN + FP + FN} * 100 = \frac{1448 + 646}{1448 + 646 + 0 + 5} * 100 = 0.997617913 \quad \textbf{(5)}$$

Given that the database under study is imbalanced, to ensure the performance of the model, "Balanced Accuracy" (BA) variable should be calculated and (BA) gives a good view of the actual performance of the model. To achieve (BA) calculating "True Negative Rate" (TNR) is needed first.

$$TNR = \frac{TN}{TN + FN} = \frac{646}{646 + 5} = 0.99231950 \quad \textbf{(6)}$$

$$BA = \frac{TPR + TNR}{2} = \frac{0.964985994 + 0.99231950}{2} = 0.998279422 \quad \textbf{(7)}$$

The (BA) which is closer to 1, indicates that the model is so capable in classification process with an imbalanced dataset.

Another practical variable in such problems, is the "F1-score" or "F-measure" variable to check the performance of the model. In fact, this variable is a kind of average between TPR and positive predictive value (PPV).

$$PPV = \frac{TP}{TP + FP} = \frac{1448}{1448 + 0} = 1 \quad \textbf{(8)}$$

F1 is calculated through the following formula:

$$F1\ Score = 2 * \frac{PPV * TPR}{PPV + TPR} = 2 * \frac{1 * 0.964985994}{1 + 0.964985994} = 0.982181040 \quad \textbf{(9)}$$

The value of F1-score which is close to 1 indicates that the model is appropriate to the problem, and it can do the classification properly.

Here is a case in which the AI model mistakenly classified 5 out of 651 surfaces as non-risky, when in fact they were risky. These 5 cases have one factor in common which is the contact entrance. Although these surfaces are rough and must be considered no risk, the contact entrances are flat which can cause vibration risk and make it risky but the AI model assumed them no risk.

Figure 11 shows a no risk surface, Figure 12 is a risky one and Figures 13 are the risky surfaces which were considered no risk by the AI model. It should be mentioned that all of the testing procedure happened in 7 sec. and each one needed 35 ms.

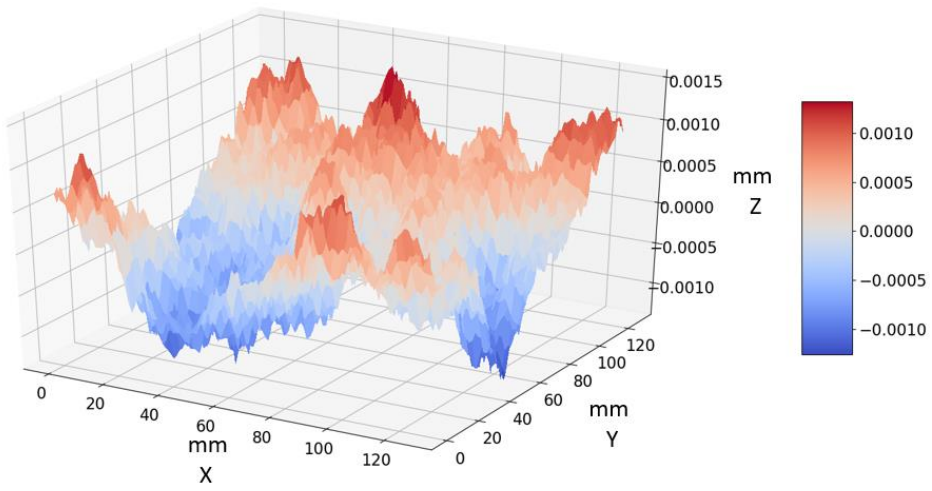


Figure 11: A no risky surface

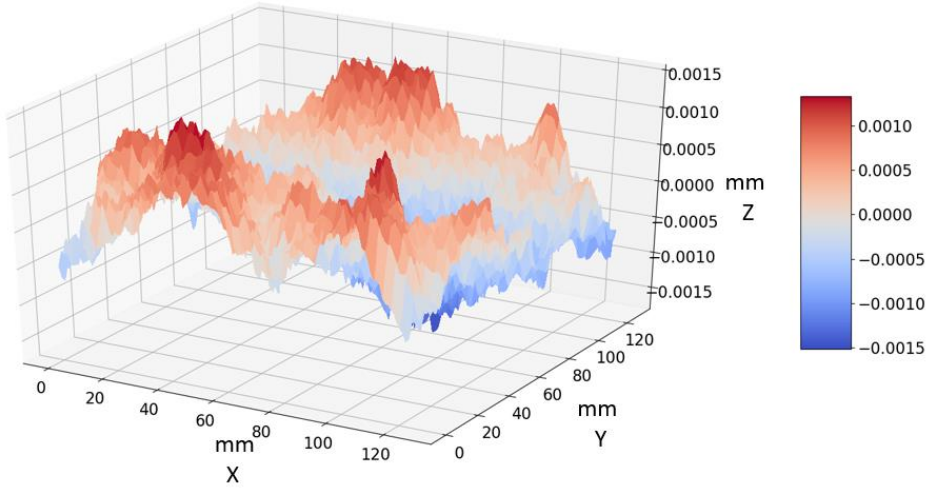


Figure 12: A risky surface

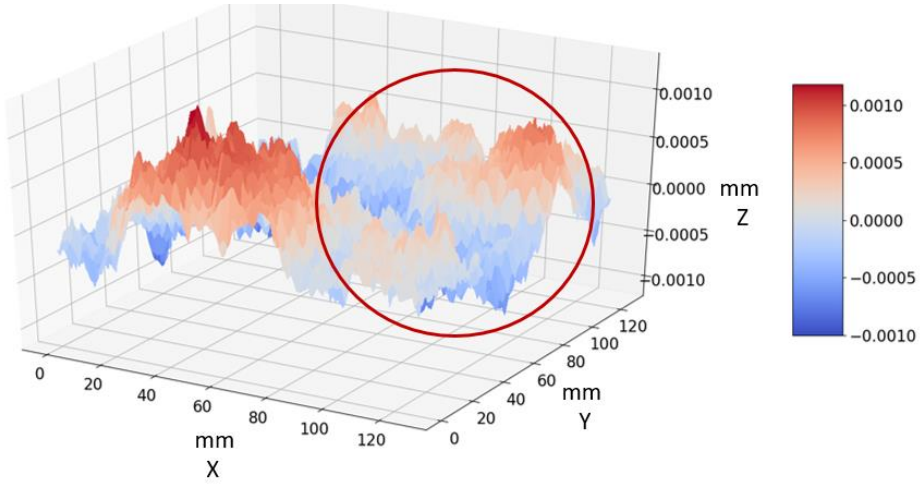


Figure 13: A risky surface which is considered no risk

To explain the cause of these mistakes, it must be clarified that they happen in the cases where the roughness is mostly the same and the model has to consider a borderline to separate them. The model tries to make the borderline not so complicated in order to avoid over-fitting. Hence, some of the factors will remain on the wrong side, which leads the model to make such mistakes. Figure 14 shows a schematic for this explanation.

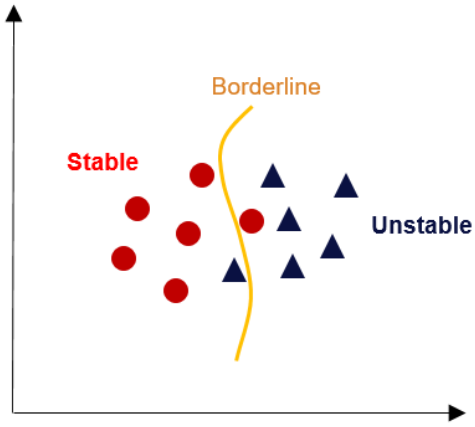


Figure 14: A simplified schematic of borderline assessment by the model, which causes the mistake

II.2.5. Step 5: 2D-CNN model for prediction: value of the mode lock-in frequency

After finding the surfaces that have the risk of squeal, the next goal is applying deep learning to determine the instability frequencies values. For this purpose, regression algorithms are used, and according to the data structure, which is the surface topography, again (2D-CNN) architecture is used to preserve the spatial form of the inputs. The inputs of this approach are similar to the previous one, but the training process and target output is different in comparison with the classification approach, because after finding the risk of instability by the classification algorithm, the regression model predicts the instability frequencies value. In fact, the approach by analyzing the surface roughness, seeks to find a correlation between asperities of the contact surface and the value of vibration frequency, then it predicts these frequencies. This mechanism makes the loss function, activation function, metric and other hyper parameters act differently from the classification models and require the use of functions appropriate to the prediction problems. Obviously, the roughness matrix forms the model input, where we target to extract the maximum information from the roughness by applying different filters. Then, by applying consecutively “Convolutional” layers, the model is allowed to investigate all points of the surface and improve its performance. In order to present an optimal architecture, the same

strategy of the previous step is used (Table 4). The architecture of this model is shown in Figure 15.

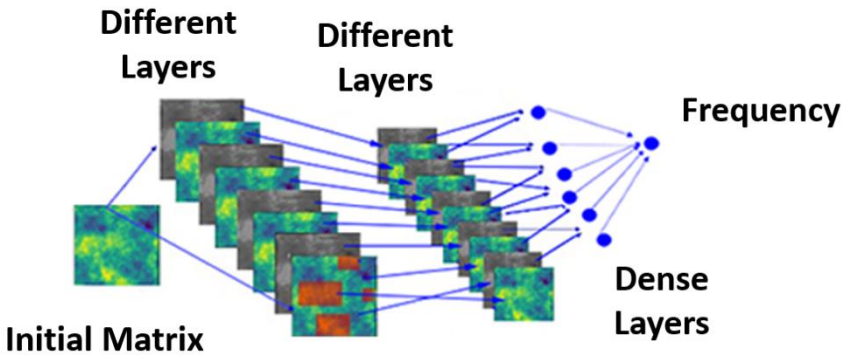


Figure 15: 2D CNN Diagram for Prediction

Table 4: Received responses from the regression model for different configurations

Loss_function :		mean_squared_error	mean_absolute_error	huber_loss	log_cosh
Metrics :	Training MSE	0.0070	0.0008933	0.0075	0.0074
	Validation MSE	0.0072	0.0019	0.0073	0.0073

Optimizer :		Adagrad	RMSprop	adam	nadam	adamax	adadelta
Metrics :	Training MSE	0.0022	0.000453	0.001977	0.000918	0.00088172	0.0176
	Validation MSE	0.0032	0.025	0.0021	0.0028	0.0025	0.0154

Batch Size :		16	32	64	128	256
Metrics :	Training MSE	0.0015	0.0014	0.0012	0.0011	0.000985
	Validation MSE	0.0031	0.0028	0.0023	0.0013	0.0025

Epochs :		16	32	64	128	256	450	650
Metrics :	Training MSE	0.00019	0.000185	0.000182	0.000180	0.000181	0.000182	0.000182
	Validation MSE	0.0002	0.000188	0.000183	0.000180	0.000182	0.000182	0.000182

The strategy begins with selecting four models in parallel and in each one, the loss function, the number of epochs and the number of batch sizes are considered constant. After determining

the best loss function, the optimizer is adjusted and the general structure of the model is found, now the batch size and the number of epochs are going to be determined. In this process, two goals are followed to find the best hyperparameters. First, the MSE value in the training and validation trend must be minimum. Second, the MSE value in the training and validation phases must be close to each other. For example, in the optimization functions, the minimum value in the validation phase is very low, but it is far from the minimum value in the training phase, that is why the "adam" optimization function is selected. Table 5 shows the final specifications of the model.

Table 5: Model specifications for classification

Model specifications	
Learning method:	Supervised learning
Algorithm:	Regression via "2D-CNN"
Aim:	Prediction of instability frequencies
Number and type of layers:	5 (2D-CNN) and 8 Dense layers
Loss function:	mean_squared_error
Activation function:	relu
Optimization function:	adam
Metric:	mean_squared_error

The "Mean Squared Error" (MSE) metric is used to evaluate the performance of the model in the regression phase, so that the performance of each hyper parameter is checked by this indicator.

Loss function is the first parameter that is set to build the model. There are different loss functions for regression problems, which the best loss function can be selected by considering the physics of the problem and its performance. By looking at the results, it can be understood that the "Mean Squared Error" loss function has the best performance. The next hyper parameter that has been analyzed is the optimizer and the results indicate that "Adam" Optimizer has the best performance.

Batch size analysis shows that the optimal performance of the model belongs to the value of 128, it also helps to converge and reduce the value of MSE in the training and validation phases. Epoch is the last hyper parameter that is analyzed, and the results show that the performance of the model with 128 epochs is acceptable.

The computation time of setting hyper parameters and complete training phase (consists of training and validation) of the regression model with a server computer (Configuration: "Memory: 128 GB, Processor: Intel® Xeon® Gold 5215 CPU @ 2.50GHz *40, Graphics:

NVIDIA Corporation GV100GL [Tesla V100SPCIe 32GB] and a Samsung SSD 2 TB M.2 NVMe) was less than 28 hours.

Results of the regression model

In this phase, the goal is to predict the vibration frequency. The model training process is shown in Figure 16. This curve shows the loss function changes over epochs. In regression issues, in order to have an acceptable performance for the model, two indicators must be considered. First, the loss function must be reduced during the training phase, second the training and validation curves must be converged.

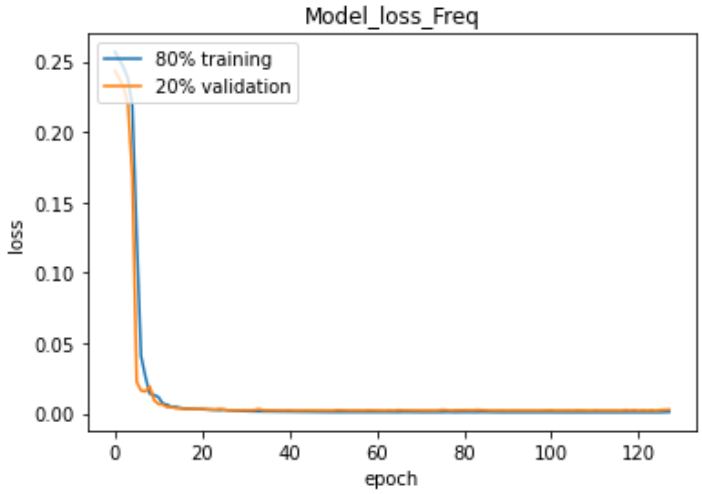


Figure 16: Loss function curve that shows the model training process

As it can be seen in Figure 16, the loss function is reduced during the training phase, also, the validation and training curves are well converged. In addition, during the training and validation, the behavior of the model is stable and there is not any noise between epochs 40 and 128, which indicates the model is able to be generalized. In order to determine and evaluate the ability of the model, it is tested by out of the database. In Figure 17, the blue circles indicate the value of expected frequencies, and the red circles are values that have been predicted by deep learning. It shows that the prediction results in most cases are close to the real cases, the error rate for all test data is 3.7%, which indicates the proper performance of the model in comparison to the out of database data, this indicates the model ability to generalize.

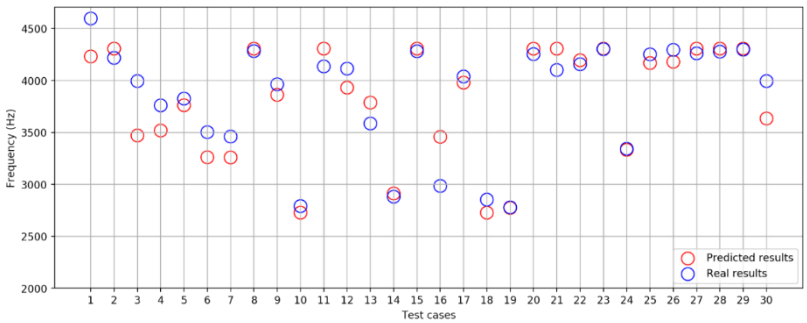


Figure 17: Results from model testing

Two examples of surface roughness are shown below: Figure 18 displays a surface roughness for which the model successfully predicts the frequency of instability, while Figure 19 depicts a contact surface roughness where the model struggles to accurately detect the frequency of instability. By referring to the roughness of the surface, no difference can be observed between these surfaces. Therefore, it can be inferred that the lack of variety surface roughness in the database is the cause. In other words, the roughness heights provided during the training phase were different from those encountered in the test phase, despite having the same shape. As a result, the model made predictions based on the training phase data and faced difficulties when presented with unfamiliar roughness heights during testing. In order to have higher accuracy, more roughness and the maximum possible state should be presented to the model in the training phase so that the model has good generalizability.

For example, in the figure below, two surface roughness are shown, which at first glance are very similar to each other, but during the contact process, surface A produces a frequency of 4148 Hz, and the second surface produces 2900 Hz. In fact, very small changes in the contact level can cause different responses (pay attention to the different scales of the following figures which makes them different).

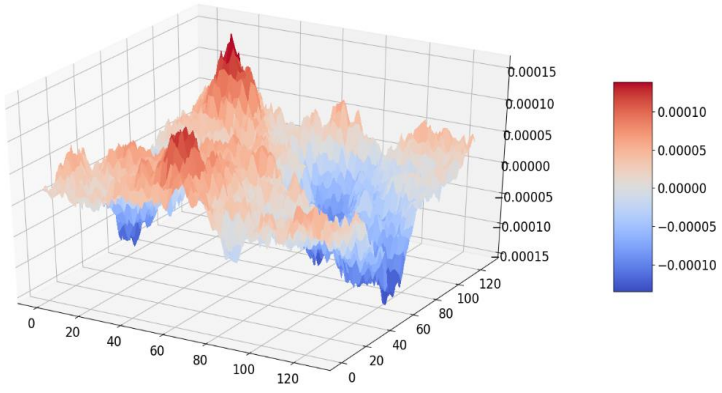


Figure 18: Results from model testing

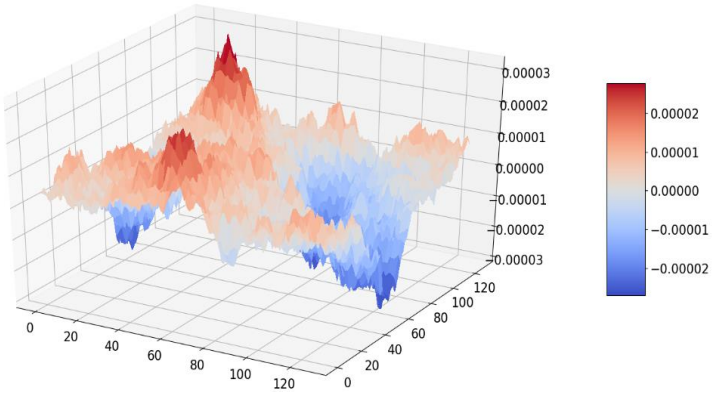


Figure 19: Results from model testing

II.3. DISCUSSION OF RESULTS

II.3.1. Influential zones

According to the results of the previous sections that small changes can lead to different responses, this idea can be raised that maybe some surface areas have a more significant role in the occurrence of system responses. For this reason, by using XGBoost [125] (explained in Chapter I, Section I.4) and Random Forest algorithms [157] (which is an extension of bootstrap aggregation (bagging) of decision trees and can be used for classification and regression problems) , the location of points that have a greater influence on the behavior of different responses of the system was discussed. The surface roughness matrices are transformed into vector form and used as input for the feature sensitivity algorithm to determine system stability or instability. The results are then converted back into matrix form to display their spatial distribution. The weight of each roughness point is determined and visualized by assigning color to the weight. This process is repeated for different roughness scales by reducing the matrix size from 128*128 to 64*64, 32*32, 16*16, and finally 8*8 to observe the impact of surface roughness scale on the localization of critical points. The results for Random Forest and XGBoost are presented in figures 20a-e and 21a-e respectively.

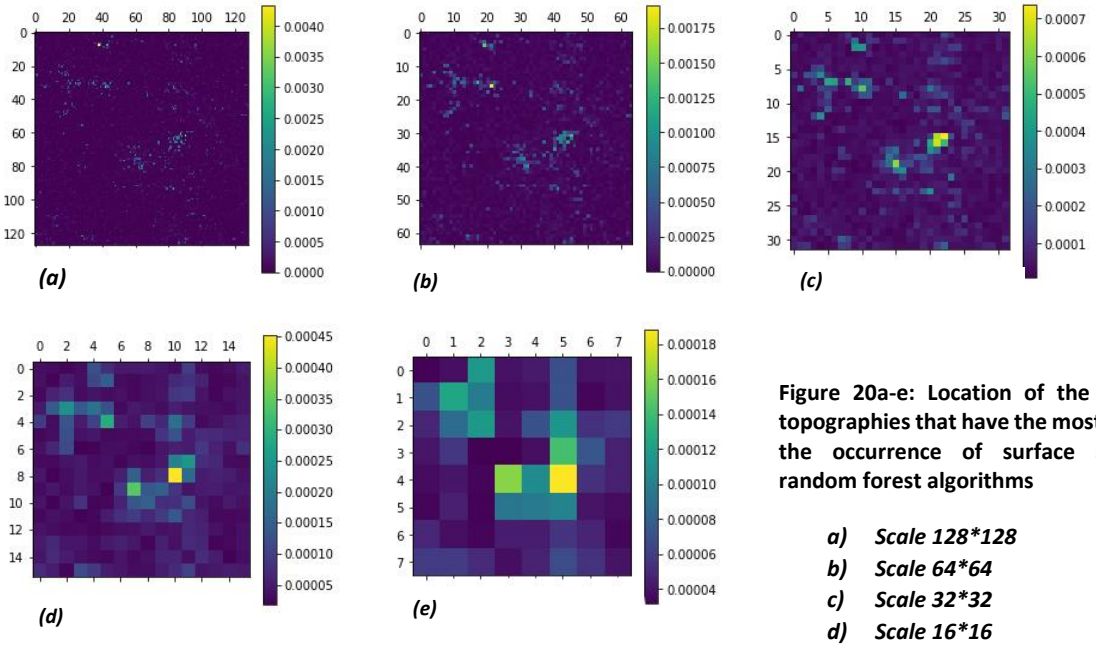


Figure 20a-e: Location of the roughness topographies that have the most effects on the occurrence of surface squeal by random forest algorithms

- ***a) Scale 128*128***
- ***b) Scale 64*64***
- ***c) Scale 32*32***
- ***d) Scale 16*16***
- ***e) Scale 8*8***

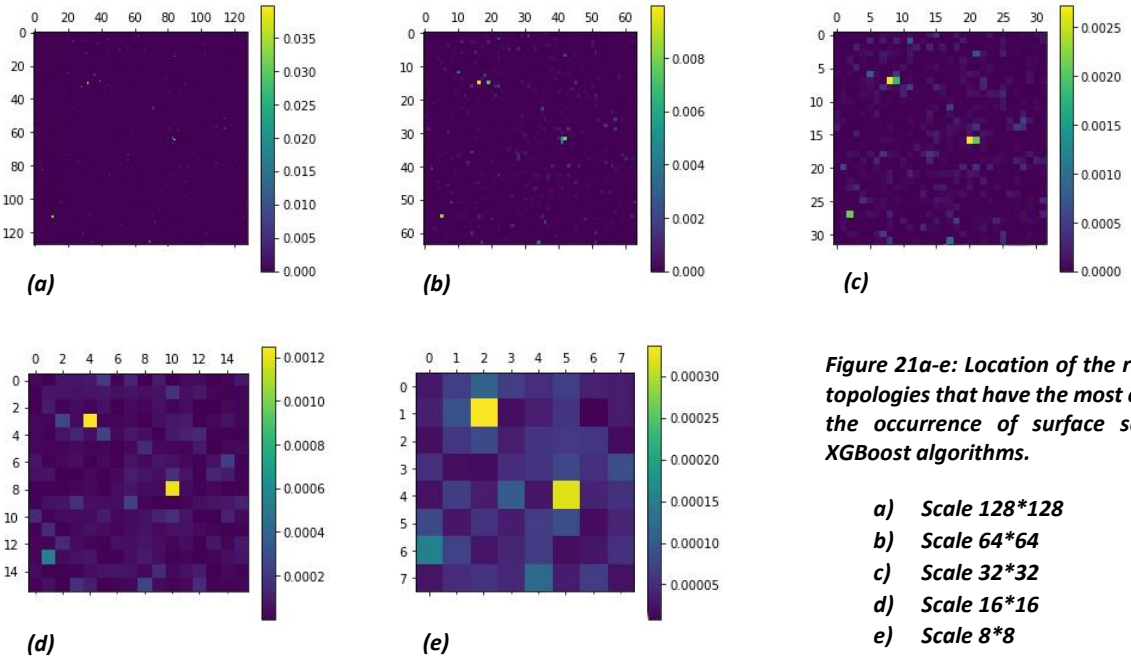


Figure 21a-e: Location of the roughness topologies that have the most effects on the occurrence of surface squeal by XGBoost algorithms.

a) Scale 128*128
b) Scale 64*64
c) Scale 32*32
d) Scale 16*16
e) Scale 8*8

It must be clarified that the output matrix in scale of 128*128 is 16 times finer than 8*8 and these 2 algorithms had almost the same responses and they both assumed the center and the up left of the surface as the important parts.

Figures 20 and 21 show that the brighter points have the most effects on the occurrence of vibration and both algorithms show almost the same location for important zones. It is also well visible that in the multi-scales, the location of the points of more importance are different. In other words, different responses are received from the system at different scales. That is why finding correct areas that play the main role in squeal must be considered in a small scale.

II.3.2. Influence of the roughness scale in the learning phase

In this subsection, the objective is to clarify the needs for considering the multiscale dimension to predict system squeal. To do so, the training phase is carried out by the surfaces with the large roughness height (-0.0002 mm to +0.0002 mm intervals). Then, the meta-model is tested by ultrafine roughness height (-0.0001 mm to +0.0001 mm) and tries to classify the instability risk for these asperities.

As Figure 22 shows, the results are not satisfactory, since the meta-model is trained with larger roughness height, it cannot perform well on a smaller scale. In order to prove this theory, in the next step, the database is inversed, using asperities with small size. In this step, model training is performed by small scale roughness height then the meta-model is tested by data in large scale. The ROC curve for testing the model is shown in Figure 23.

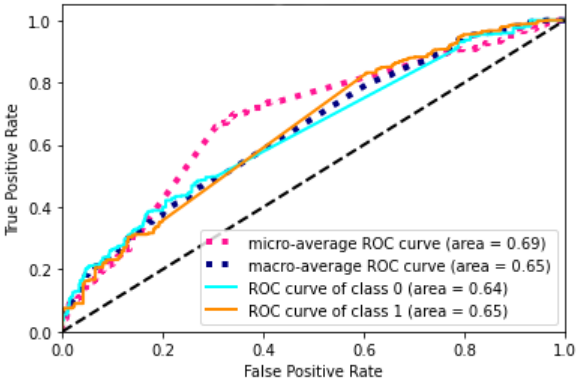


Figure 22: ROC curve of the model which has been trained by large roughness height and tested by small roughness height data

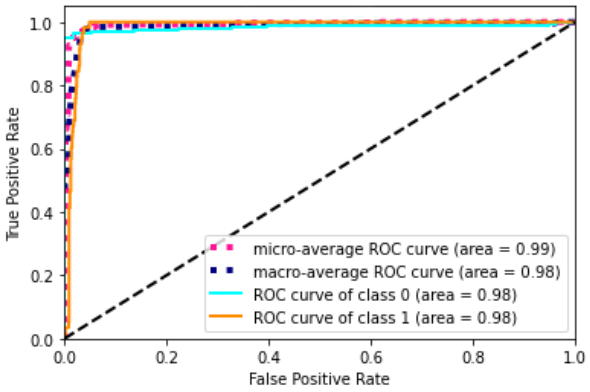


Figure 23: ROC curve of the model which has been trained by small roughness height and tested by large roughness height data

Figure 23 shows that when the model is trained by finer roughness height, it can be well generalized to the other surfaces. The confusion matrix in the model testing phase is shown in (Table 6). The confusion matrix shows that the model is able to correctly identify 1436 out of 1448 silent cases and 9 misidentified cases with vibration risk. On the other hand, among the 651 cases with squeal risk, 636 cases are detected correctly by the model, but 15 cases are misdiagnosed.

Table 6: Confusion Matrix

Actual class / **Predicted**	**Stable**	**Unstable**
Positive	1439 (TP)	9(FP)
Negative	15 (FN)	636 (TN)

A summary of the predicted outputs the test phase is shown in Table 7.

The results show that the accuracy obtained by this approach is significant and the model has an acceptable generalizability.

By comparing the results in Table 3 to the Table 6, it is seen that the results in Table 3 are more accurate. By analyzing the input data, it was clarified that the inputs of II.3.2. model was much

more various so it helped the models generalizability but as the inputs in table 6 were more limited, the model was in trouble to generalize and extend the data.

Table 7: Model performance analysis

Indicator	**Value**
TPR True Positive Rate	0.9896
FPR False Positive Rate	0.0139
TNR True Negative Rate	0.978
ACC Accuracy	0.988
BA Balanced Accuracy	0.984
PPV positive predictive value	0.994
F1 Score	0.992

The model's behavior during training and the results obtained during testing suggest that surface data scales play a significant role and that the DL model employs a multiscale approach to understand the correlation between surface roughness height and system instability. In other words, the strategy can be summarized as exploiting only data with the finest possible roughness height.

II.4. CONCLUSION

In this chapter, contact surface roughness and its effect on system responses are investigated. For this purpose, two-dimensional convolutional networks are used according to the shape the inputs, which are surface roughness and in matrix form. Surface roughness is analyzed from two aspects. First, by using a classification model, surfaces with high vibration risk are found, then by a prediction model, the instability frequencies related to each roughness are predicted. The error of the model is less than 1% in the classification section and less than 3% in the prediction section. The results demonstrate that minor variations in surface roughness yield different system responses, resulting in an increased error rate. To improve the model's performance, it is crucial to generate additional data for the training phase, addressing the need for an increased data volume. Even though it may take some time, having a comprehensive and representative dataset will be beneficial in the long term.

In order to build the dataset introduced in this section, a finite element model is used. Parameters such as the friction coefficient are considered constant, as changing any of these

parameters would significantly increase the calculation time for different configurations. However, in experimental tests, these parameters exhibit high variability and interdependence with each other. They have many interactions and affect the behavior of the system. For this reason, according to the acceptable results obtained in this section, it is appropriate to create a dataset of experimental trial, and by analyzing these system responses, the behavior of the system can be checked and predicted.

It should be added that in the approach used in this section, the surface of the disk is considered completely smooth, although the roughness of the disc surface can also be the reason for the different behavior of the system in experimental tests.

Transition

In Chapter II, the dynamic state system in squealing process was studied. In order to analyze contact process, surfaces were generated using function based on experimental measurements [158]. Despite the function's ability to work with real experimental data, it is evident that there are differences present in actual contact surfaces. For example, in reality, we encounter surfaces that exhibit multi-scale characteristics, whereas the function focuses on a specific scale. Also, in reality during the contact process, surface evolution happens which causes the primary assumptions do not work properly as the surface has changed.

On the other hand, during the contact process, obtaining information is not feasible as the contact process is closed. However, it is evident that in the contact process between two surfaces, the highest points make contact first. So, the stress and consequently temperature raise in these zones due to the friction. So, by studying the temperature in different zones, it will be helpful to detect the contact zones state.

Finally, measuring the contact surface temperature evolution can lead to get information about the contact surface during the contact process. In this chapter, the surface state is investigated by analyzing the temperature evolution, which is depicted in a blue box in the main thesis diagram (Figure III).

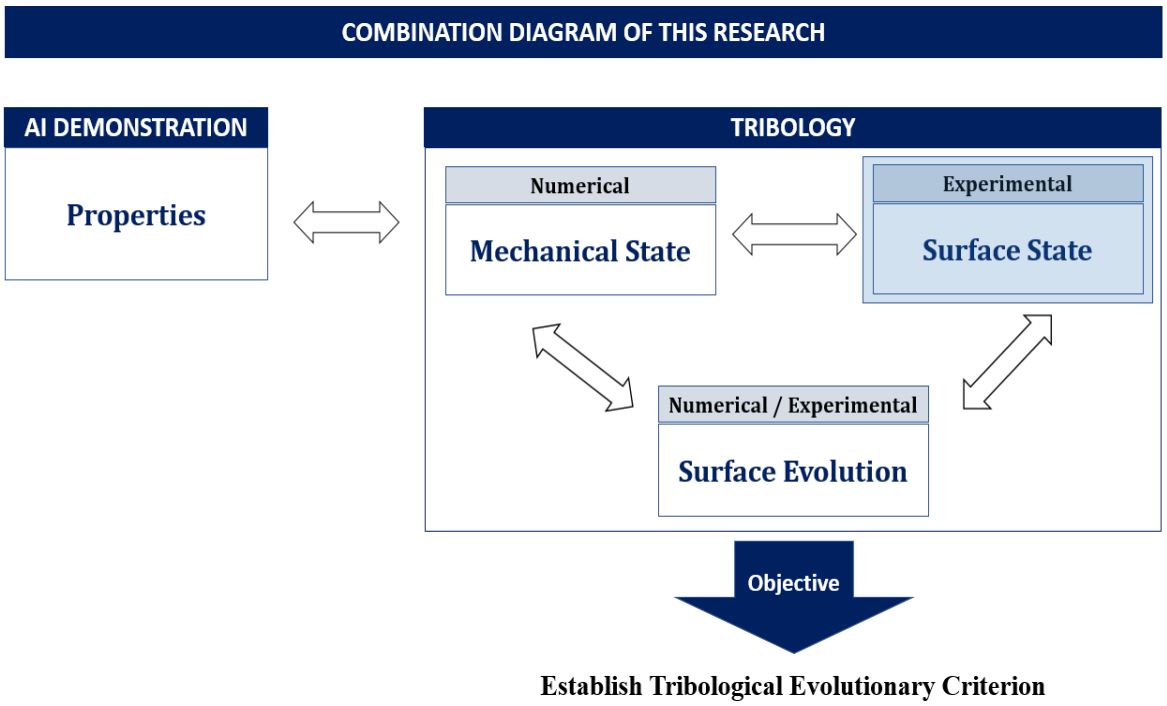


Figure III: The general scenario of the thesis, which in this chapter the surface state, during the contact process is analyzed.

III. Toward Predicting Contact Localization in a Pin-on-disc System: Deep Learning Approach

Table of Contents

Abstract

From the tribological point of view, it is well known that the frictional contact between two parts is, on the one hand, localized only on some contact areas and, on the other hand, evolving. To try to predict these localizations over time, an approximate coupling is employed between a pin-on-disc system and a proposed AI hyper-model, which refers to an advanced model utilizing artificial intelligence techniques. The contact locations are "captured" through a sheet of thermocouples embedded in the pin. Nearly 600 tests are performed with different boundary conditions (contact force, disc speed and holding brake time) to obtain variability. Based on a large amount of data, a hyper-model based on an innovative 2-D recurrent convolutional neural networks (2D-RCNN) architecture is built to predict the evolution of temperature. This hyper-model appears to be very efficient as it presents an error lower than 4% in the test phase. An intelligent control system of the applying contact force is proposed, with the aim of achieving contact homogeneity. This, in turn, suggests the potential for "smart braking".

Keywords: Tribology, contact localization, experimental-numerical coupling, deep learning, heavily instrumented trials

III.1. INTRODUCTION

The contact mechanism is a tribology issue, and due to the presence of roughness in the contact interface, the actual contact surface is less than the nominal contact surface [158], which causes the concentration of stress, and as a result, the temperature rises in these zones [159][30]. So the behavior of contact interface has a great effect on the system performance [160]. Experimental studies show that the propensity to squeal in the contact process directly depends on the tribological parameters such as surface temperature, humidity, local pressure, and loading history [55][161][162]. To gain a clear understanding of tribological systems is important because, with a proper understanding of this mechanism, disorders such as friction-induced vibration [163][164] which is the main cause of noise pollution [165][166] during the contact process, can be prevented. Since the contact process is a complicated multi-scale [158][146] tribological issue, which is also multi-physical [167] and evolutionary [30][168]. Moreover, various factors including contact localization [169][170], third body [171][172], contact materials [146][168], and rheology are involved in this problem. These complexities cannot be adequately understood through simplified models [173] or classical experimental tests.

In addition, as the contact interface is closed during the contact process, it is no longer possible to easily obtain information about the state of the surface. Therefore, one of the primary and critical parameters which helps to understand the contact surface behavior in the closed condition, is the evolution of temperature during the contact process in different contact zones. In recent years, various methods [169][170][174] have been proposed to locate and predict these evolutions. However, as per the aforementioned points, the contact problem is highly complex, making it impractical to develop a comprehensive model that encompasses all aspects of this problem. For these reasons, there are still many unsolved problems in this field. Recently, some methods for understanding system behavior using Deep Learning approaches have been proposed [144], but these models are based on system-scale responses and do not analyze the cause and origin of these behaviors. This is a key aspect because by a correct understanding of the contact process and by finding the origin of these disorders, some actions can be taken during the contact process to control the occurrence of annoying frequencies generated by the system. They could reduce the costs of after-sales services, increase quality, and reduce noise and environmental pollution. Hence, by using DL methods including Recurrent Neural Network (RNN) and 2D-CNN, a method is presented and applied on an experimental database to study and predict the surface temperature evolution during the contact

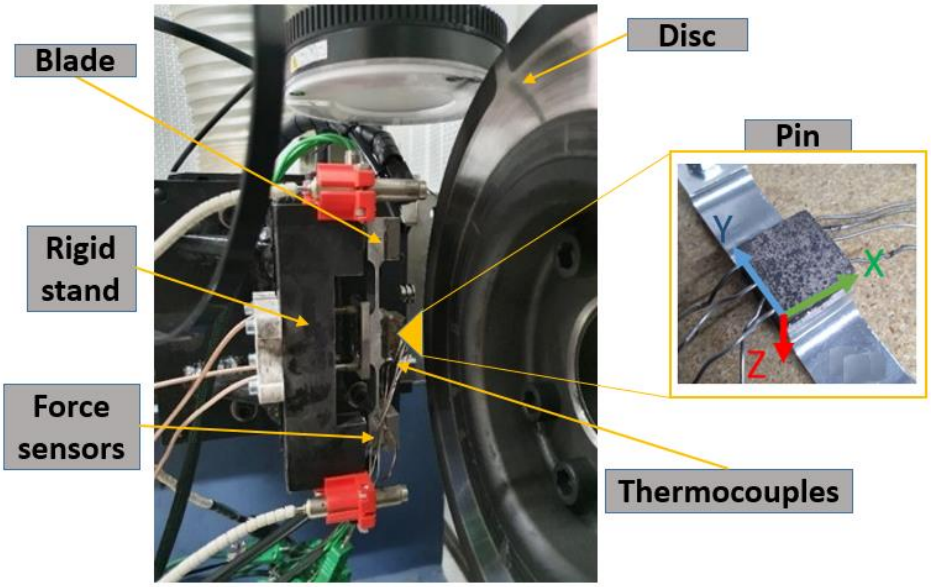
process. The combination of two types of architecture, CNN to capture the spatial features, and RNN to model the relationship between them across the temporal domain, could offer an interesting property to understanding the frictional contact behavior.

III.2. THE BENCH-TEST

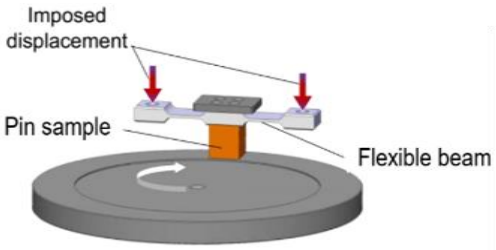
III.2.1. Presentation

The experimental set-up that is being used in this study is based on the design previously established in [175], and it has been utilized in [48]. It is based on a pin-on-disc configuration (Figure 1). A thin plate is fixed at its extremities on a rigid stand. A 20 mm x 20 mm square face pin (extracted from a brake pad) is glued at the center of the blade. The rigid stand moves along the axial disc direction and pushes the pin against the disc, the normal load is obtained with the bending of the plate. The disc is designed in a solid automotive shape, and it is made of 15CrMoV6 steel material. The average radius of the friction track is 100 mm. The pin is made of a sintered composite material with a metal matrix (copper and iron) graphite particles and abrasives

Since it is impossible to obtain much information about the contact interface when two bodies come into contact and the surfaces are closed, hence the surface temperature changes can indicate the contact state during the process. Therefore, materials with high thermal conductivity are used to transfer temperature changes quickly to thermocouples. Indeed, the used materials have transversal isotropic properties and the heat conduction in the direction perpendicular to the contact surface is much higher than in other directions, therefore, by placing thermocouples in the depth of the pad near the contact surface and recording the contact changes during the contact process it is possible to get information about the contact surface temperature during the process.



(a)



(b)

Figure 1a-b: Equipment installed on the experimental test and test schematic.

a) Equipment installed in situ to record system responses.
b) The physical model schematic includes a pin on the disc.

III.2.2. Instrumentations

According to the objectives of the study, several measurements are performed:

- **Acoustic measurements**: A microphone placed along the axis of rotation of the disc at the distance of 1 m from the pin-disc contact surface, is used for the determination of the squealing frequencies. They were deduced from the raw acoustic pressure using short-time Fourier Transform. A peak is the spectrogram is considered as a squealing frequency if its sound pressure level is greater than 80 dB and a duration greater than 1s.

- **Force measurements:** Two 3D piezoelectric sensors are inserted between the extremities of the thin plate and the rigid stand. Normal and tangential contact forces can be estimated from these measurements.

- **Thermal measurements:** 8 thermocouples are inserted inside the pin at 3 mm depth from the contact surface. The position of the thermocouples is shown in Figure 2, and during the

process, two thermocouples (no.7 and no.8) were detached from their positions, while the remaining six thermocouples continued to be used for the tests. The used instrumentation has already been in the same as [48] to track the load bearing area at a macroscopic scale.
The sampling frequencies for the first two measurements are 50 kHz and for the thermocouples 90 kHz.

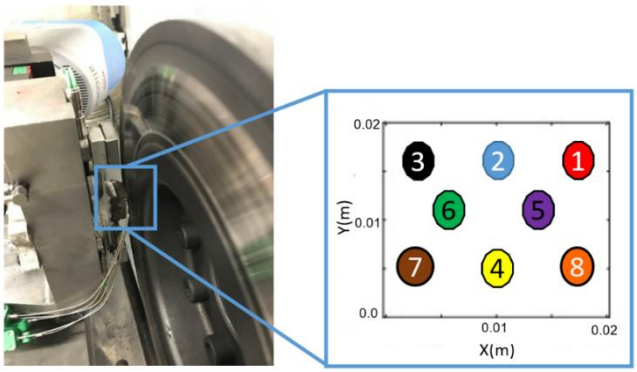
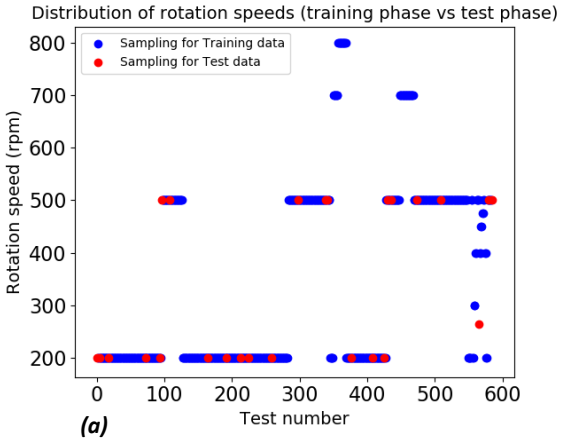


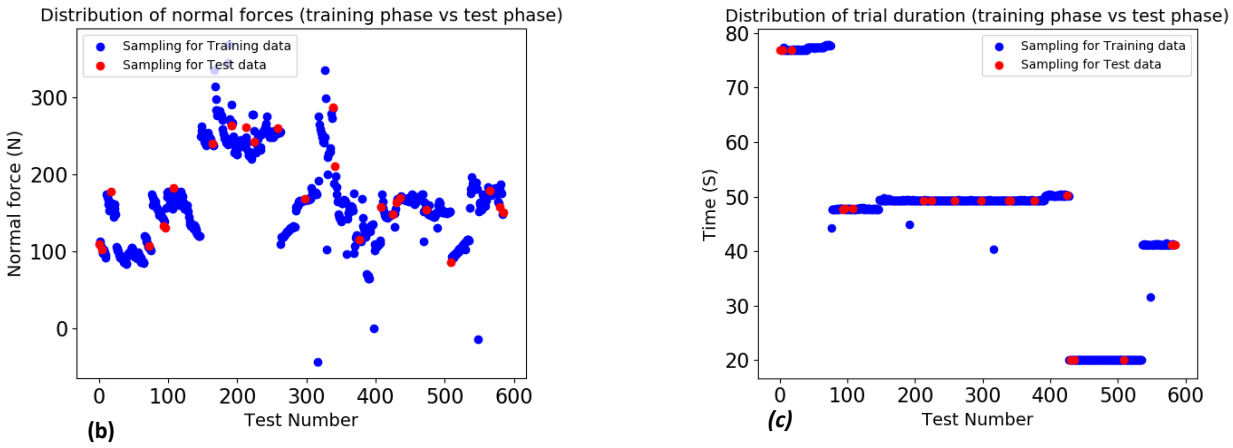
Figure 2: Position of thermocouples relative to the contact surface.

III.2.3. Protocol and some illustrations of experimental configuration

The trial sequence is decomposed into series of friction tests (typically around 10 tests) including contact and no-contact conditions. The rotation speed remains constant (drag conditions) for each trial (Figure 3a), the normal forces (Figure 3b) and the trial duration (Figure 3c) are different from each other. The complete sequence is finally composed of 584 friction tests with different speed rotations, normal force, and duration of contact.



(a)



Figures 3a-c: The disparity in the input parameters in all the trials

- ***a) Distribution of disc rotation speed in the test series.***
- ***b) Distribution of normal force in the test series.***
- ***c) Distribution of tests duration in the test series***

The point marked with blue points are used data in the training phase of the deep learning model, and the data marked with red points are used from experimental data into the model evaluation phase, which will be described later. In order to have a homogeneous distribution from all the test phase, the data selection strategy for the model training phase and the model testing phase have been completely random.

III.3. AI MODEL FOR LOCAL TEMPERATURE PREDICTION

III.3.1. General presentation of the model

As the performance of a tribology system is highly dependent on the behavior of the contact interface [158][55][161][146], the interface analysis can provide useful information to predict system performance, and makes it possible to prevent system-scale disturbances. Due to the complexity and existence of multi-physical interaction in the contact mechanism, current experimental or numerical methods cannot cover all aspects of this process to provide a complete model with high generalizability. Here, using deep learning (DL) methods, the contact interface is analyzed in order to predict the behavior of the system.

Since the evolution of surface temperature and contact localization are believed to be the main factors influencing the behavior of a tribological system, the aim is to predict this evolution during the contact process, in order to accurately anticipate the system's behavior.

As the main method, a deep learning type scheme and more precisely, a 2D-recurrent convolutional neural networks (2D-RCNN) is used. This choice is motivated because of two

main reasons. First, the history effects must be considered in the problem (time series data). Second, the temperature of different thermocouples can have a direct or indirect effect on each other (spatial features).

For a better understanding the used algorithm, first recurrent algorithm is introduced, then different scenarios are presented to solve the problem.

Recurrent neural networks: A type of artificial neural networks which is very suitable for analyzing, processing and forecasting sequences and time series [176].

One of the advantages of recurrent neural networks is that they can process and predict sequences with different lengths, which helps the generalization of models built with this algorithm [177].

Recurrent neural networks are different because they are trained by taking the history of a sequence into account. It lets the algorithm in this structure receive information from the previous data and use them to influence the next outputs and inputs [178]. In fact, recurrent neural networks are a type of neural networks that their performance makes deep learning models dynamic. Figure 4 shows the schematic of the operation of recurrent neural networks.

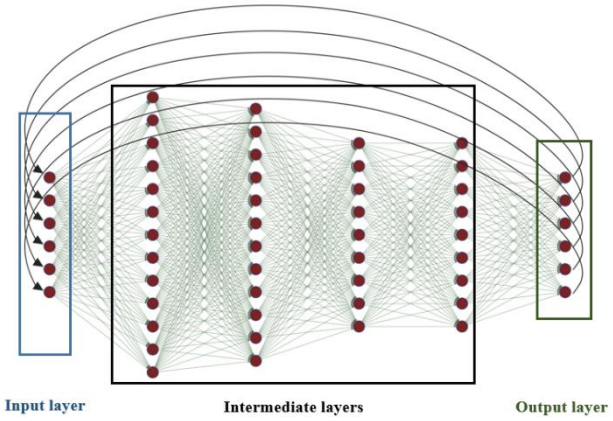


Figure 4: Schematic of the recurrent neural network algorithms mechanism.

After creating the dataset and choosing a suitable algorithm to the needs of the issue, it is time to adopt a scenario facing the problem.

By looking at the experimental results which are the temperature changes compared to the initial contact surface temperature, in the first-time no-good signals would be received. Because it can be clearly seen that they are completely heterogeneous from one to another.

Figure 5 shows three of these tests randomly selected from the database (each thermocouple is indicated by a specific number and color specified in Figure 2).

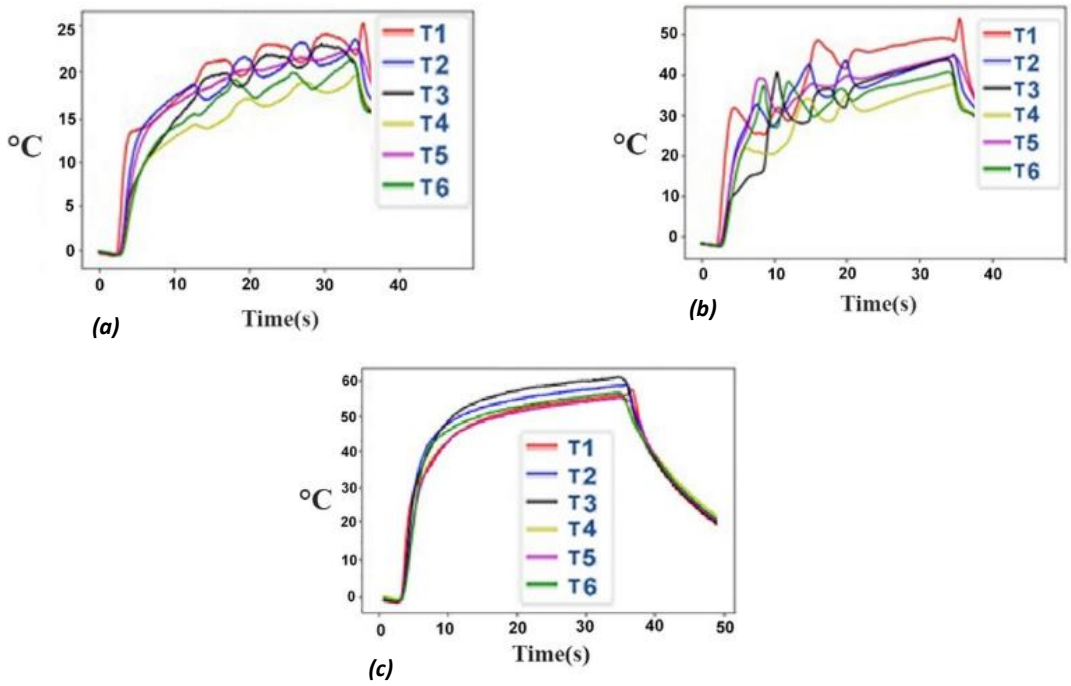


Figure 5a-c: Three different tests show the system behavior heterogeneity.

a) Test no. 96
b) Test no. 108
c) Test no. 225

The evolution recorded by each thermocouple during each test is also different, for example, test no. 96 in Figure 6 shows this heterogeneity, and during each test at different times, a different behavior of each thermocouple is obtained.

As a result, this heterogeneity with respect to time and location of each thermocouple provides us with an extremely complex problem. In fact, this inhomogeneity in the evolution of the surface temperature in different zones and over time is caused by tribological phenomena, which causes the contact surface and contact behavior to evolve. For this reason, a conventional understanding of this phenomenon cannot be obtained with classical research methods.

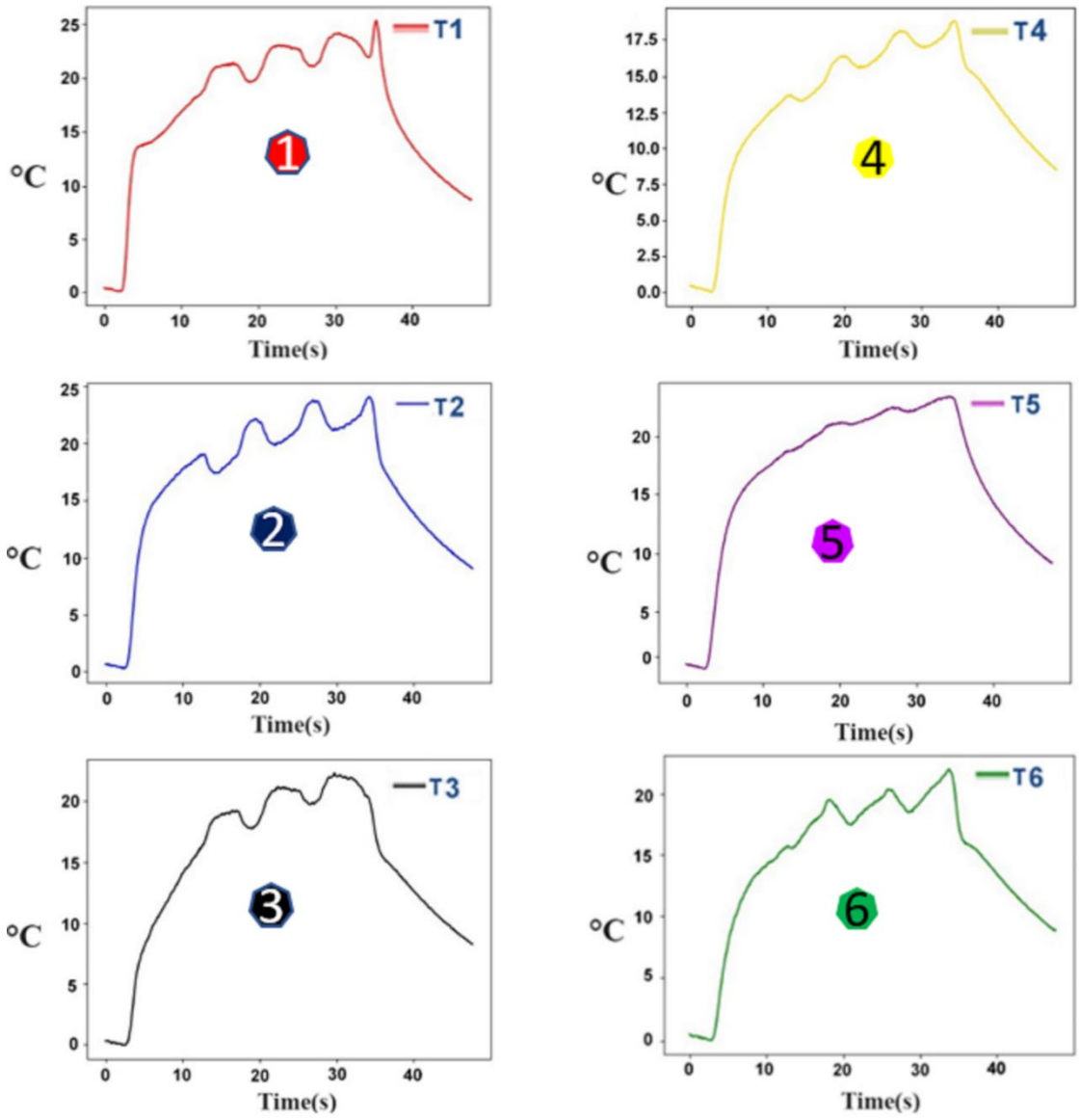


Figure 6: Temperature evolution heterogeneity in different surface zones (test no. 96).

According to the mentioned points and in order to use deep learning, an appropriate scenario for this problem must be used. The global scenario facing this issue is shown in Figure 7.

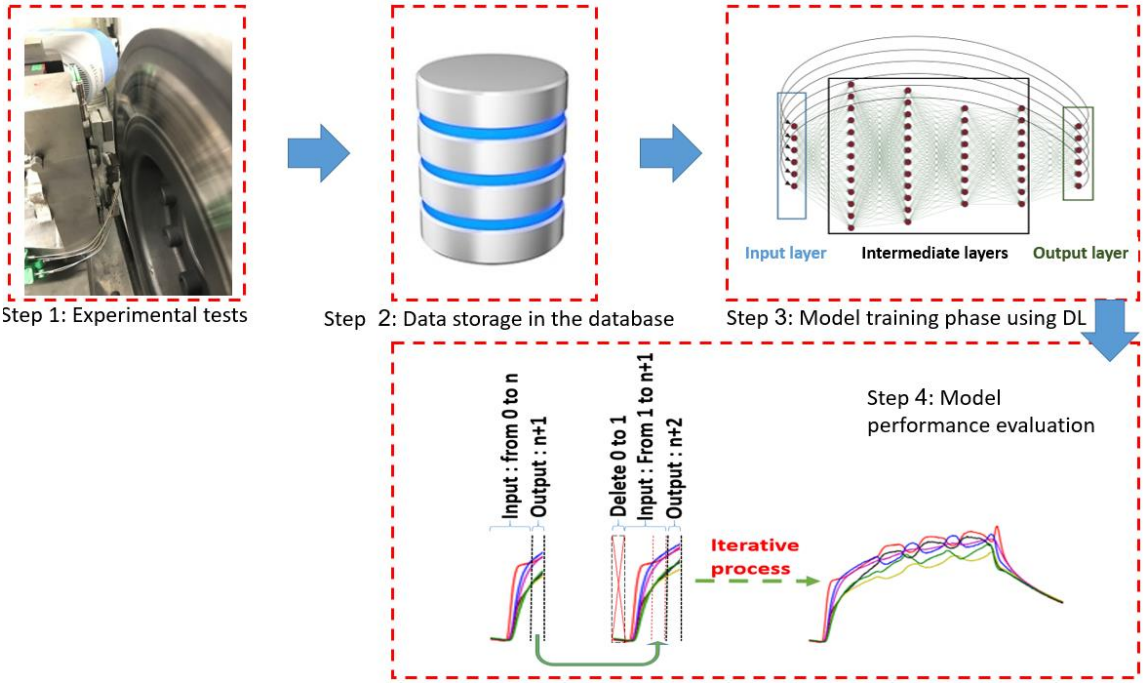


Figure 7: The scenario chosen to solve this issue.

This scenario consists of several steps, in the first step, experimental tests are performed, in the second step, data including the disc rotation speed, the force applied from the pad to the disc, the time and the temperature evolution during each trial are stored in the database. In the third step, after some data preparation such as data normalization (for example, a normalization is carried out so that all the values are located between 0 and 1, implying a homotheticity compared to the extreme values experimentally recorded), the model training phase begins, and the model tries to predict the temperature evolution by analyzing the data for a few seconds from the beginning of the test. Finally, in the fourth step, the built model is evaluated with data outside the training phase to verify its performance.

To solve the problem, two different strategies are considered to find the best, most appropriate and optimal strategy which will be described in the following.

III.3.2. Different scenarios to solve the problem

- First strategy: An LSTM (Long Short-Term Memory) architecture is used to build a model that considers the temperature evolution, normal force, and rotation speed for a certain time as known data and the training phase is done with this data. Then, in the test phase, by providing data such as rotation speed and normal force as known data, the model is asked to predict the evolution of the temperature of the contact surface. In fact, the data from an experiment is given to the model from the second 0 to 20 in the training phase, each second contains 10 frames so

20 seconds equals to 200 frames. Then the model predicts the extension of the experiment. Figure 8 shows the general architecture of the model.

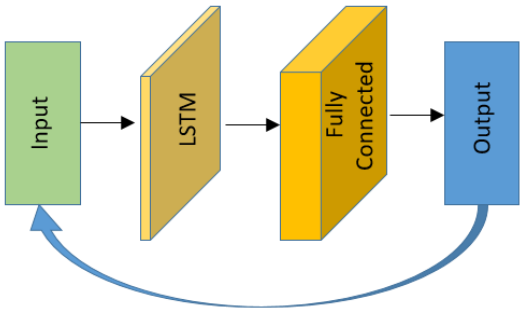


Figure 8: The model architecture used to predict surface temperature evolution.

In order to set and find the best hyperparameters of the model, an iterative strategy is used. These results are shown in Table 1.

Table 1: Model results for tuning hyperparameters (first strategy)

Loss_function :		mean_squared error	mean_absolute_error	huber_loss	log_cosh
Metrics :	Training MSE	0.002	0.8	0.0475	0.024
	Validation MSE	0.0034	0.045	0.0973	0.143

Optimizer :		Adagrad	RMSprop	adam	nadam	adamax	adadelta
Metrics :	Training MSE	0.0252	0.43	0.009	0.000918	0.0008817	0.0176
	Validation MSE	0.0588	0.925	0.0011	0.0028	0.0025	0.0154

Batch Size :		4	8	16	32	64
Metrics :	Training MSE	0.0025	0.074	0.02	0.081	0.985
	Validation MSE	0.0028	0.088	0.04	0.13	0. 25

Epochs :		16	32	64	128	256	450	650
Metrics :	Training MSE	0.017	0.028	0.0001	0.032	0.0058	0.0061	0.0052
	Validation MSE	0.08	0.0147	0.0003	0.0452	0.079	0.0095	0.007

LSTM layers :		2	4	6	8	10	12
Metrics :	Training MSE	0.00012	0.00143	0.0039	0.012	0.059	0.1
	Validation MSE	0.00007	0.0025	0.0097	0.035	0.083	0.17

Dense layers :		2	4	6	8	10	12
Metrics :	Training MSE	0.011	0.0001	0.094	0.22	0.207	0.27
	Validation MSE	0.068	0.0007	0.14	0.31	0.4201	0.394

The final specifications of the model are shown in Table 2.

Table 2: Specifications of the model used for model training and testing (first strategy)

Model specifications	
Learning method:	Supervised learning
Algorithm:	Regression via "RNN"
Aim:	Predicting surface temperature evolution
Number and type of layers:	2 LSTM and 4 Dense layers
Loss function:	mean_squared_error
Activation function:	relu
Optimization function:	adam
Metric:	mean_squared_error

After setting the hyperparameters, to find the minimum desired frame so that the model has a correct understanding of the history of temperature evolution, the different frames number is presented to the model and the error percentage corresponding to each of the configurations is recorded. The model error percentage for each interval is shown in Figure 9.

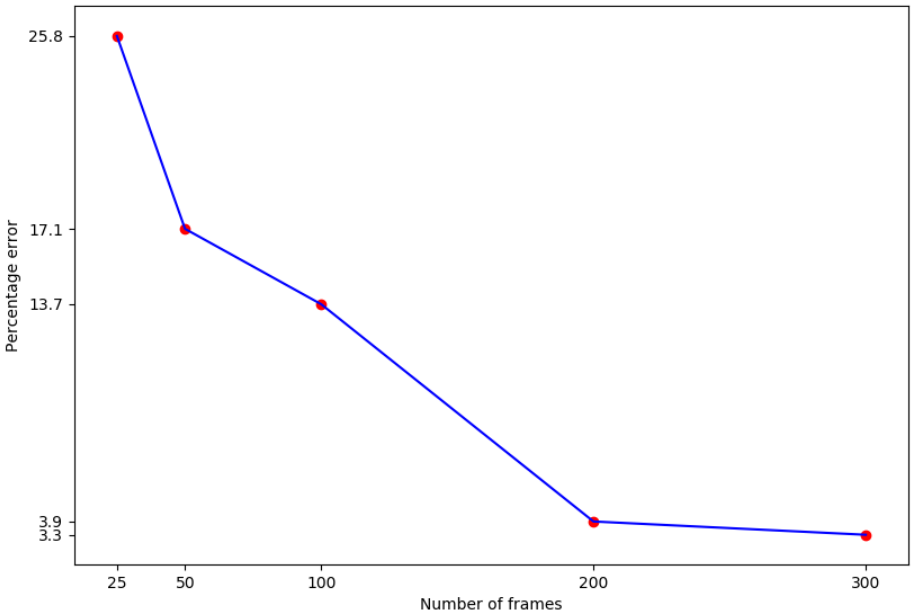


Figure 9: Model error percentage vs the number of used frames

Although, the results of the model considering 300 frames (30 sec.) are more accurate than 200 frames (20 sec.) but considering that the duration of the used test when the pad is in contact with the disc is approximately 35 seconds, choosing 300 frames (30 Sec.) is not a suitable choice for predicting temperature evolution. For this reason, 200 frames (20 Sec.) are used for the model training phase. Figure 10 shows the training process of the model.

In this process, the training phase is done from second 0 to 20, marked by the vertical red line in the Figure 10a, then the model predicts the temperature evolution of the contact surface from second 20 to 35, according to the history and the trend of temperature evolution during the contact process.

In this strategy, the disc rotation speed, normal forces, and temperature changes for a specific interval are considered known data, and the model predicts the temperature changes in the next step.

By comparing the predicted results by the model in Figure 10a after the red vertical line and comparing it with the real data in Figure 10b, it can be seen that the performance of the model is acceptable.

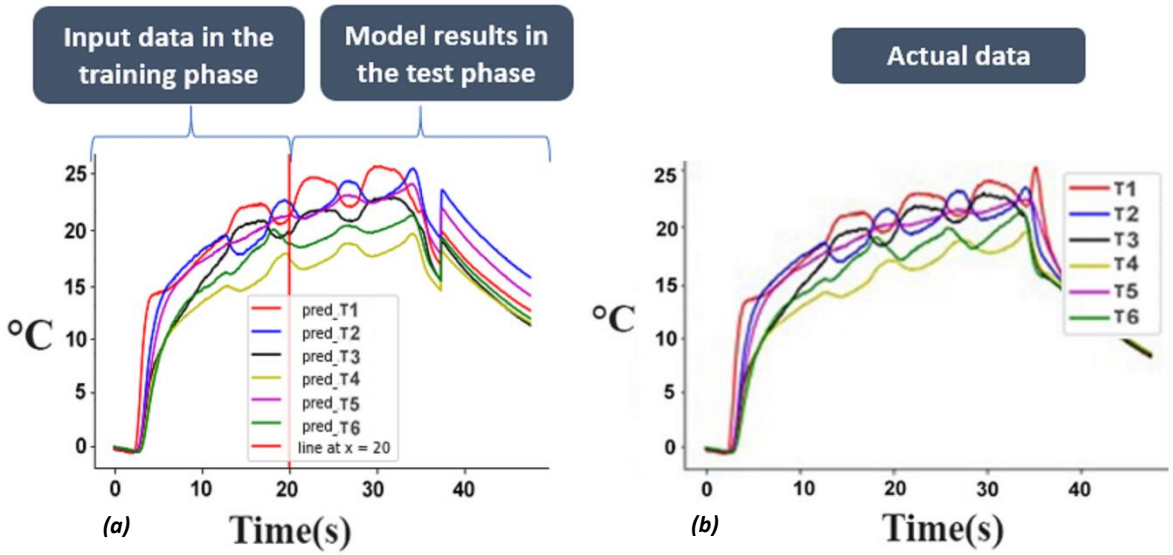


Figure 10a-b: The model performance using the first twenty seconds data.

- ***a) Model training phase and test results (test no. 96).***
- ***b) Temperature evolution in the trial no. 96.***

The results of this scenario are satisfactory so that the error percentage of the model is 3.9 %, and it can predict the continuation of the surface temperature evolution. In order to evaluate the performance of the model, then it is tested by data outside the training phase. The results are shown in Figure 11.

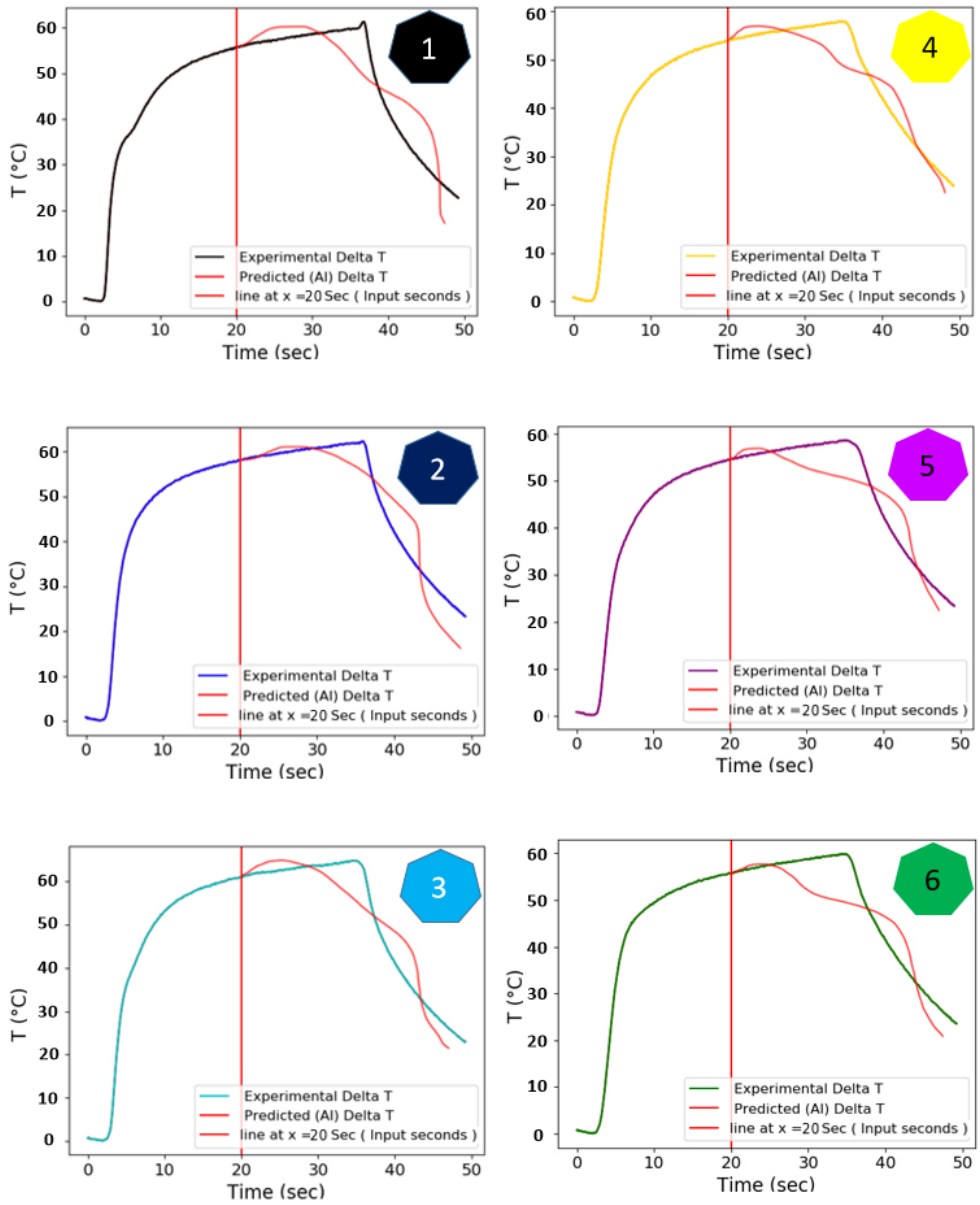


Figure 11: Model performance facing new data.

As can be seen, the approach made with this method is inefficient because if applied to another experiment with different behavior, it predicts completely incorrect results (8.9% error). Even though the data is given to the model with a large interval, the model has not been able to predict the phenomena well. The prediction results and the real data respectively are shown in Figure 11. By comparing tests no. 96 and no.225, it can be seen that two completely different prediction results have been received from the system, that's why the model has not been able to make a good prediction. According to the obtained results, another strategy is presented that is more accurate and generalizable.

- Second strategy: Due to the heterogeneity of system responses in the experimental test series, the training process should change entirely to provide a model with a proper performance. Therefore, the training phase is done discontinuously, contrary to the philosophy of recurrent

neural networks. In this strategy, the data to perform the training phase is divided into time intervals, including **n** frames as input and the **n+1st** frame as output, then data is entered into the model randomly from all parts of the training phase even in a single test and in a discrete way (Figure 12a).

In this strategy, the disc rotation speed, normal forces, and temperature changes for a specific interval are considered as known data, and the model predicts the temperature changes in the next step. But the testing phase is done in a continuous manner, in such a way that the data for the first **n** frames are entered into the model as input, and the model predicts the temperature changes for the **n+1st** frame, then the predicted data is entered into the model to predict the temperature changes for the **n+2nd** frame and this trend continues until the end of the test. The mentioned process is shown in Figure 12b.

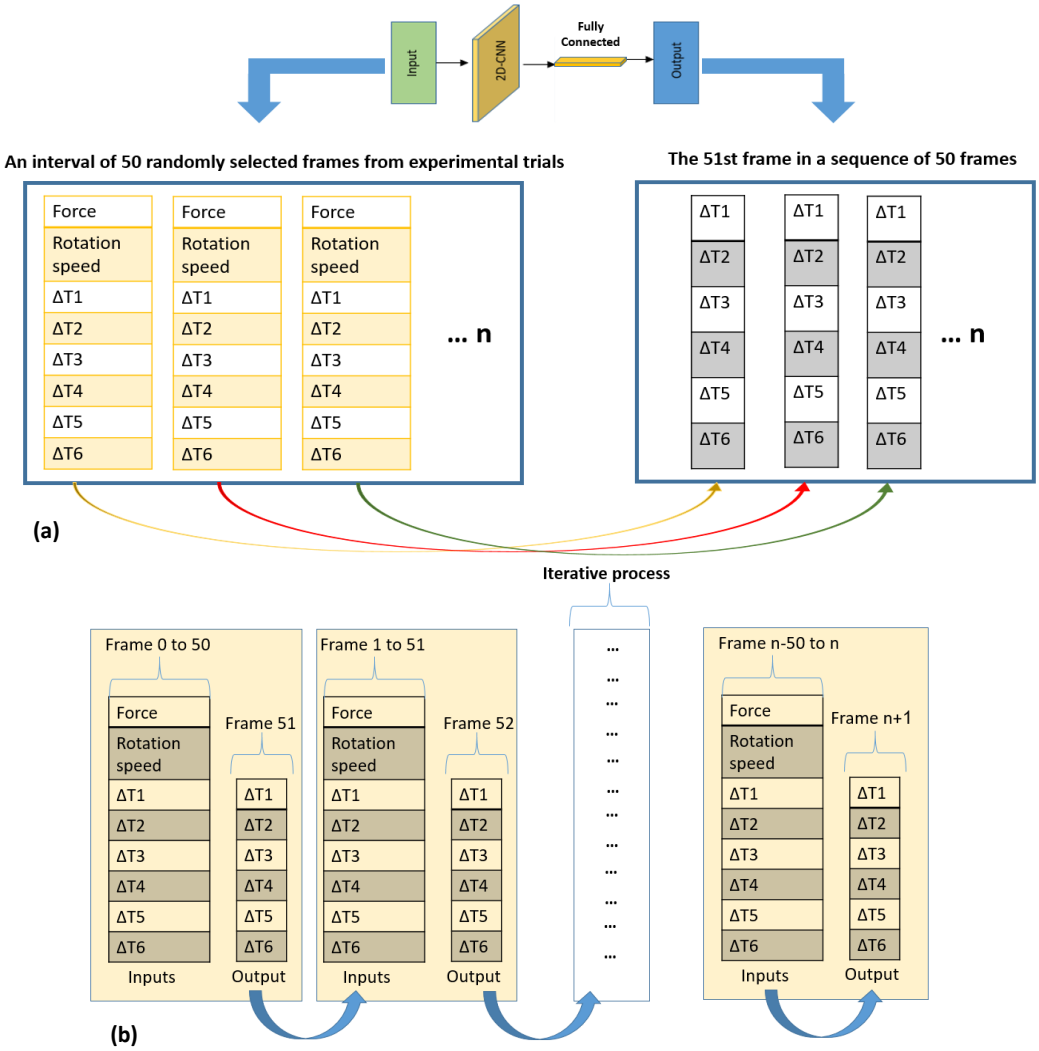


Figure 12a-b: **Enhancing model performance through discontinuous training and iterative testing for heterogeneous system responses.**

a) Discontinuous training approach for improving model performance in heterogeneous system responses
b) Model testing process that predicts the entire temperature using an iterative process

In order to apply DL, 85% of the data from all the trials described above are used in the training and 10% for the validation phases. Once the DL model is established, the 5% remaining is used for the testing phase. In order to adjust the hyperparameters of the model, an iterative strategy is used, the results for each hyperparameter are shown in the Table 3.

Table 3: Model results for tuning hyperparameters (second strategy)

Loss_function :		mean_squared_error	mean_absolute_error	huber_loss	log_cosh
Metrics :	Training MSE	0.0192	0.310	0.212	0.024
	Validation MSE	0.014	0.541	0.0279	0.130

Optimizer :		Adagrad	RMSprop	adam	nadam	adamax	adadelta
Metrics :	Training MSE	0.171	0.759	0.001	0. 18	0.009	0.25
	Validation MSE	0.111	0.652	0.0021	0.037	0.0045	0.34

Batch Size :		8	16	32	64	128
Metrics :	Training MSE	0.0804	0.0415	0.0045	0.00047	0.0042
	Validation MSE	0.012	0.057	0.00063	0.00022	0. 0052

Epochs :		32	64	128	512	750	850	950
Metrics :	Training MSE	0.75	0.176	0.0125	0.0074	0.002508	0.00019	0.0001
	Validation MSE	0.532	0.328	0.0257	0.00571	0.003801	0.00095	0.00056

2D-CNN layers :		2	4	6	8	10	12
Metrics :	Training MSE	0.000042	0.00257	0.003587	0.0019	0.075	0.079
	Validation MSE	0.0000397	0.00428	0.00467	0.078	0.12	0.047

Dense layers :		2	4	6	8	10	12
Metrics :	Training MSE	0.2514	0.0241	0.0098	0.00092	0.01201	0. 0183
	Validation MSE	0.17503	0.0197	0.0014	0.001	0.00694	0.0179

The final specifications of the model are shown in Table 4.

Table 4: Specifications of the model used for model training and testing (second strategy)

Model specifications	
Learning method:	Supervised learning
Algorithm:	Regression via "RNN"
Aim:	Predicting surface temperature evolution
Number and type of layers:	2 (2D-CNN) and 8 Dense layers
Loss function:	mean_squared_error
Activation function:	relu
Optimization function:	adam
Metric:	mean_squared_error

After adjusting the hyperparameters, selecting the optimal number of frames as the input for the DL model is the next goal. So that it has the highest accuracy and the least number of frames as input in order to have the best prediction in the shortest possible time. To find the best

interval for the model training, in the first step, the first 10 frames (1 sec.) are taken. The error rate is more than 17.5 %. Then the model is trained by 25 frames (2.5 sec.) as an interval, which significantly reduced the error rate to 8%. The next try is 50 frames (5 sec.), which reduced the model error percentage to less than 4%. 3.2 % for a 75 frame (7.5 sec.) interval and 2.6% for a 100 frame (10 sec.) interval. It can be seen that as the number of frames increases, the accuracy of the model would be increased, but the goal in this analysis is to get an error less than 5% with the minimum input frame so that the 50 frames as input has a good performance. The model error percentage for each of the listed intervals is shown in Figure 13.

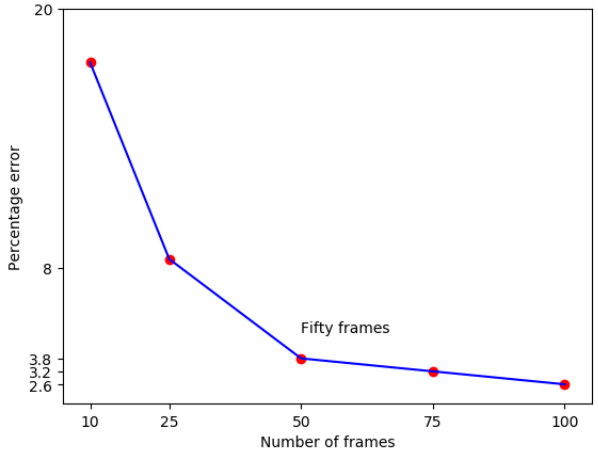


Figure 13: Percentage error vs the number of input frames

Since the temperature evolution happens so fast, the goal is to predict the system's behavior in a short period to correct the inappropriate behaviors. Therefore, 50 frames are the optimum interval chosen because the less than 50 frame interval reduces the model performance, and more than 50 frame reduces the prediction speed.

This choice of 50 frames (5 sec) also seems to be physically founded because in view of Figure 14a, it would seem that there is a characteristic length of about 5 seconds. In addition, the spectrogram corresponding to this test, Figure 14b, also shows that the frequencies received from the system are disturbed in a 5 second periods.

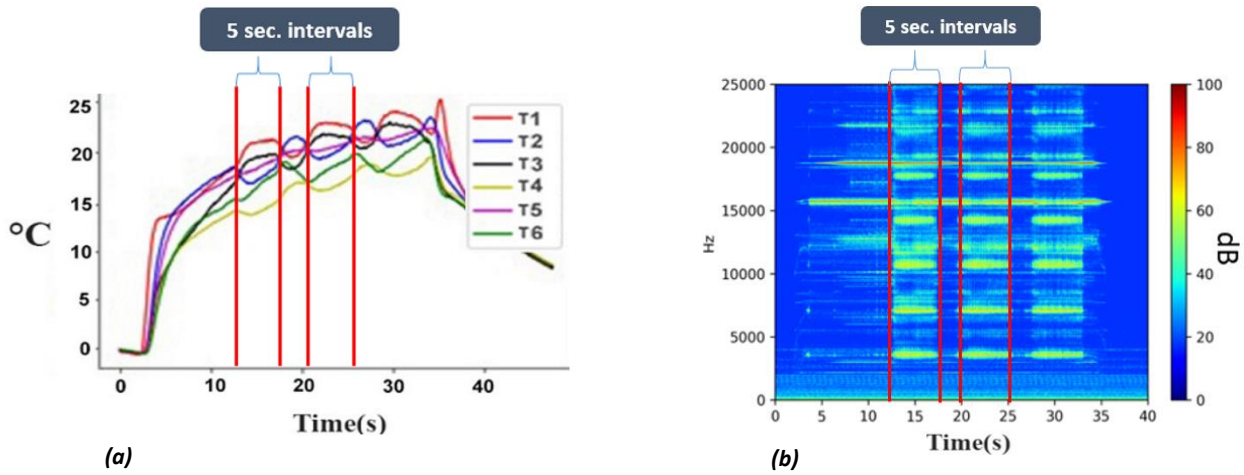


Figure 14a-b: Responses received from the system which change in 5 second intervals.

a) 5 second intervals indicated by red vertical lines (Temperature evolution for test no. 96).
b) 5 second intervals indicated by red vertical lines (Spectrogram corresponding to test no. 96)

According to the error percentage of model for 50 frames as input and the system behavior such as surface temperature evolution and the recorded spectrogram, it can be concluded that the period of 5 seconds is a decisive period in the system behavior. As far as this phenomenon happens in 50 frame or more intervals, it could be generalized to all the following tests.

After finding the best hyperparameters and the best interval, now is the time to train the model and analyze its performance in the training phase.

The training phase consists of sampling the data at 5-second (50 frames) intervals taken randomly from the 95% database which is dedicated to training and validation phase. The goal of this package of 50 frame is to determine the temperature evolution of the six thermocouples at the 51st frame. Then in the test phase (5% of database), by successive iterations and shifting the input data in the DL model, it would be possible to predict the temperature evolutions at all subsequent times. In fact, in the test process, the force data is considered as a parameter which can be controlled by the user when the contact occurs, the rotation speed and the surface temperature data in only the initial 50 frames (5 sec) is considered as known data, and after each prediction, the predicted data enters into the model as an input to predict the temperature evolution in the next step.

III.3.3. Training phase

In order to increase the model ability, a separate model is considered for each thermocouple.

The loss function/epoch (when all the data of the training phase is entered into the model and the weight of the layers is updated in the forward and backward propagation phases) curves in the training and validation phases for each of the thermocouples are shown in Figure 15.

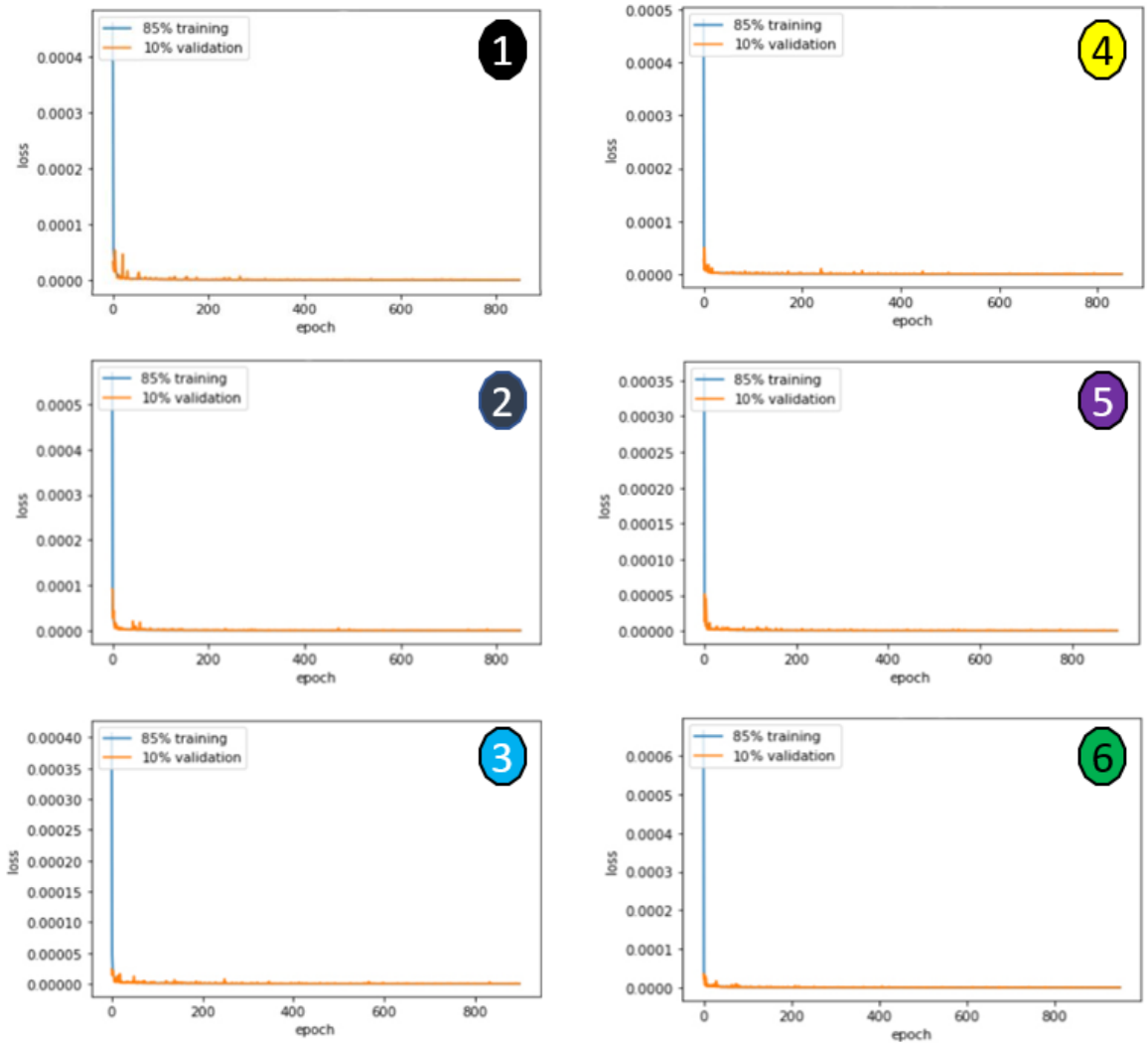


Figure 15: Loss function curves for six DL model associated to each thermocouple.

The training and validation curves for all six thermocouples show that the training phase of the model is well done after the model is trained by this strategy.

The computation time for setting hyper parameters and complete training phase of the model by a server computer (Configuration: Memory: 128 GB, CPU @ 2.50GHz *40, Graphics: 32GB and an SSD 2 TB M.2 NVMe) is less than 24 hours.

III.4. RESULTS AND DISCUSSION

At the end of the training phase, the model is tested on a separate test subset to assess its performance and generalizability. 5% of the dataset is used as a test subset.

As mentioned above, the average error of our 2D-RCNN model on the entire test set is 3.8% per 50 frames of input data, and the calculation time required to predict the entire test is 153ms.

To demonstrate the effectiveness of the model, two specific tests (tests no.18 and no.96) were selected. The temperature evolution in the database can generally be categorized into two types: steady and disturbed evolution. Hence, experiment 18 was chosen as representative of steady temperature evolution, while experiment 96 was chosen as representative of transient temperature evolution, allowing for a comparative analysis of their respective behaviors. For this reason, by examining the input data of these two typical tests, the different behavior of the model is analyzed.

First, test no.18 is analyzed. In order to make predictions, the first 50 frames marked with a red vertical line in Figure 16 have been provided for the model. After the red bar, the AI uses the force data, rotation speed and the results obtained from each step, predicts temperature changes in the next frame.

In Figure 16, the curves shown in red are the results predicted by the DL model, and the curves shown in different colors are the results recorded during the experimental tests. The general trend between the experimental and predicted results by DL model seems to match. In detail, the braking process lasted for 35 seconds. After 35 seconds, the pad-disc pair is no longer in contact, and it is in the free convection cooling phase, but the test presented continued to the 75th sec.

Either in the phase of rising or falling temperature, the DL model seems to be in a good agreement since the errors do not exceed 2.4% comparing the experimental data.

Indeed, this trial seems quite classical concerning the evolutions where the first derivative is always from the same sign on the contact (<35s with a positive sign) and non-contact (>35s with a negative sign) phases. In other words, this test shows a quasi-uniformity of the contact on the whole surface of the pad during the whole braking phase, as shown by the evolution of the temperatures on all the thermocouples and the level of temperatures reached.

According to the evolution of the surface temperature in the experimental test no. 18, although the evolution of the surface temperature in all thermocouples did not take place in the same way, the model is able to detect even the existing peak just before the temperature decrease in thermocouple no.1

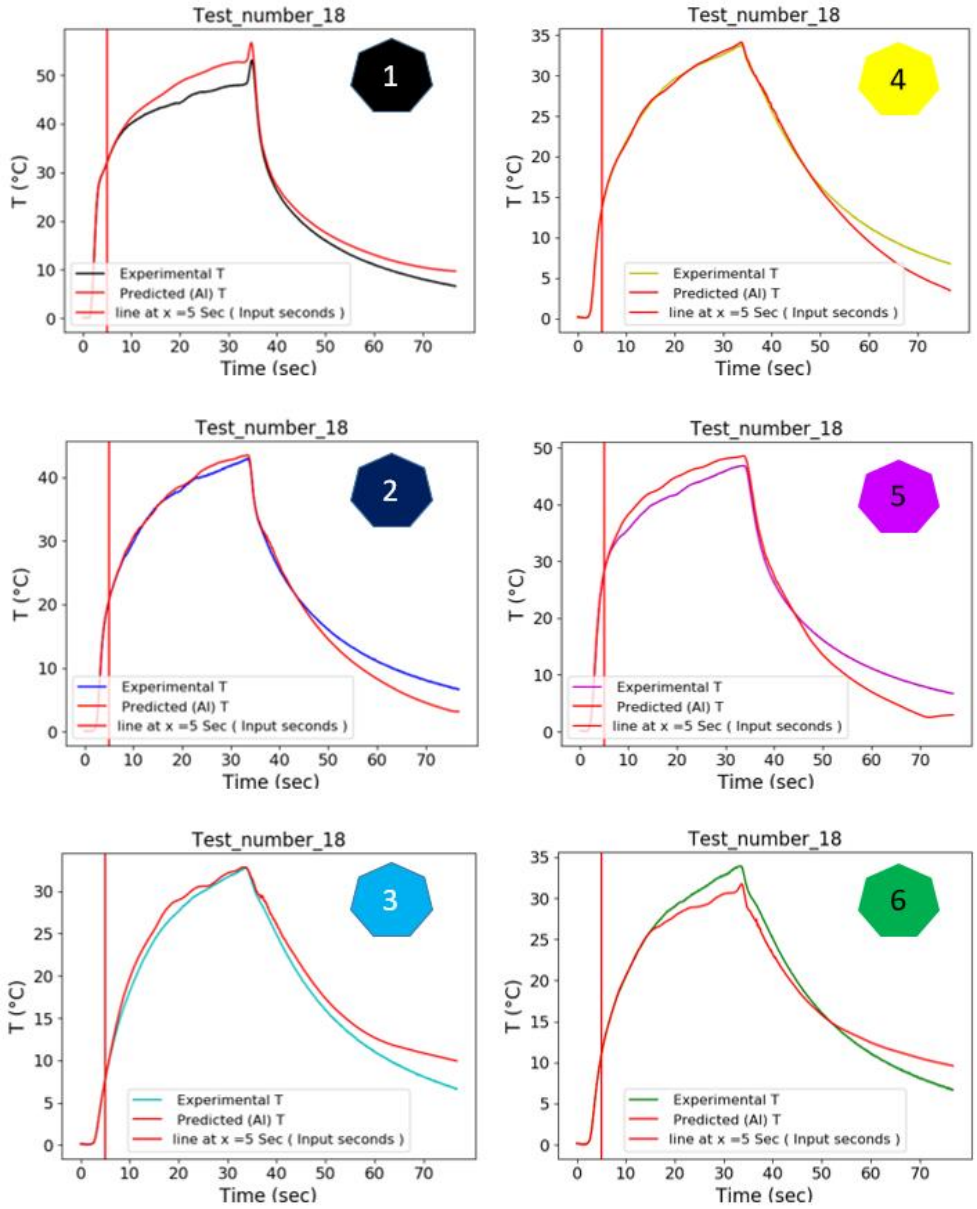


Figure 16: Predicting the temperature evolution of experiment no.18 by the (DL) model

Now, the other trial (no. 96) which is not used in the training phase is analyzed. This trial shows more variations in the braking phase, as shown in Figure 17. Again, the inputs are the first 50 frames, and after the vertical red bar, the DL model is left in autonomy. At the first sight, the global trends are pretty good between the temperature evolutions predicted by the model and the experimental data so that the error percentage for the trial no. 96 is 6.8% (but compared to the total error percentage of the model (3.8 percent), it is large). More precisely, these tests

show more variation (with positive and negative slopes) during the braking phase. These variations are often "captured" by the model, where there may be some errors in predicting the amplitude.

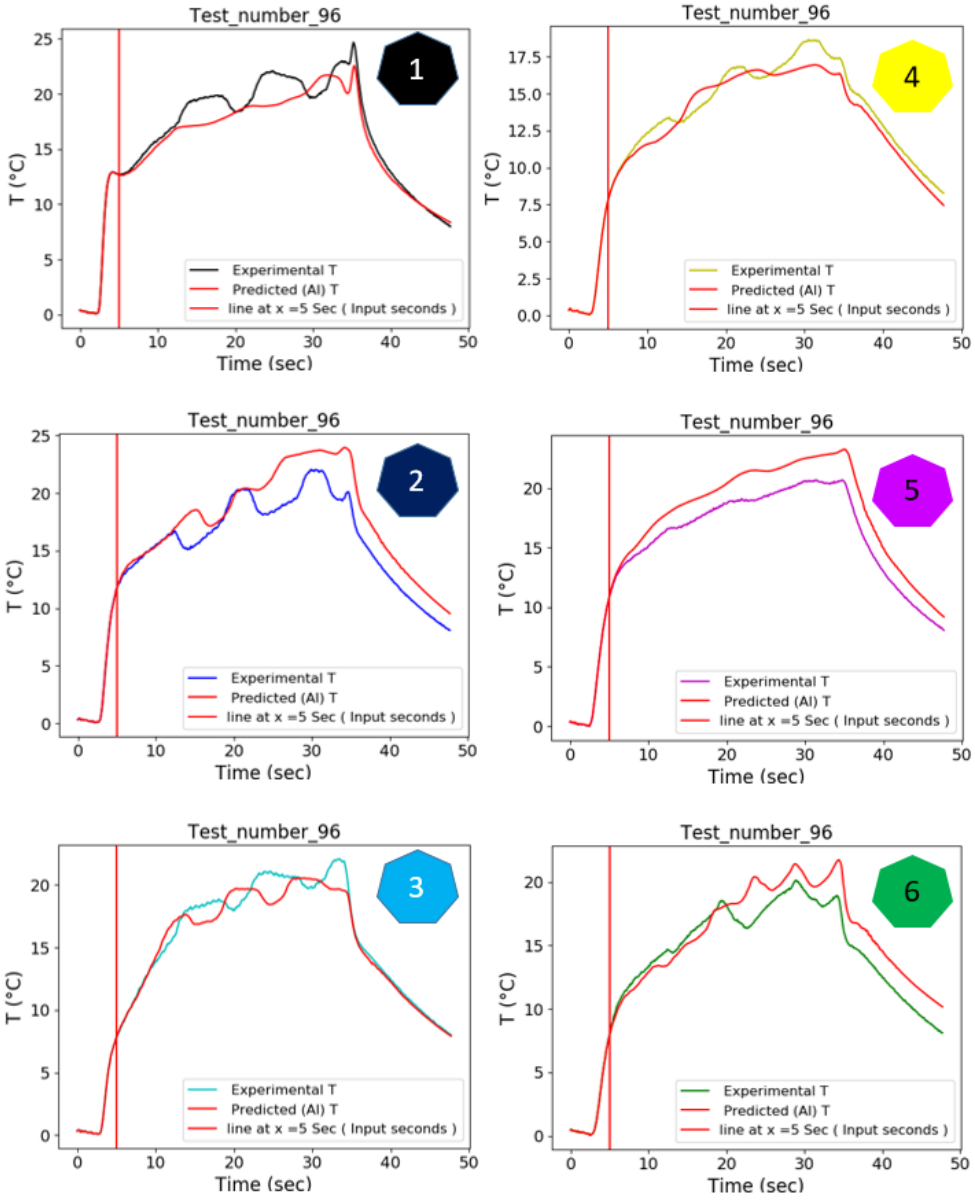


Figure 17: Predicting the temperature evolution of experiment no.96 by the DL model

The results shown before are pretty encouraging but need to be improved in some respects where there are more heterogeneities in the thermal response of the system. This kind of situation occurred in 15% of all cases treated. This is still reasonable, but it is necessary to better understand what is going on to improve the prediction. As mentioned in the instrumentation section, the test bench is equipped with a microphone. And to better

understand the different responses of the system, spectrograms are very useful. The spectrograms of the two tests illustrated above are presented in Figures 18 and 19.

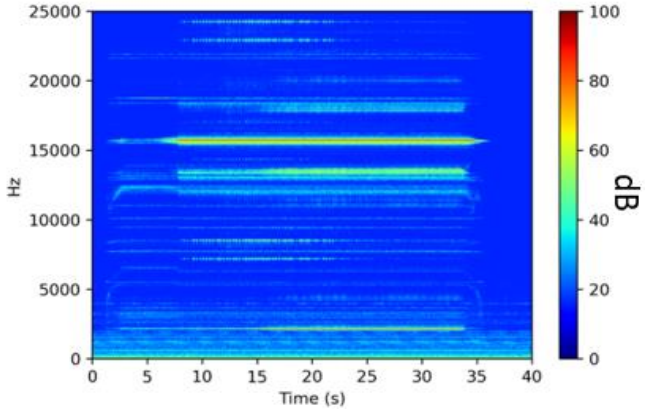


Figure 18: The spectrogram corresponding to experiment no.18

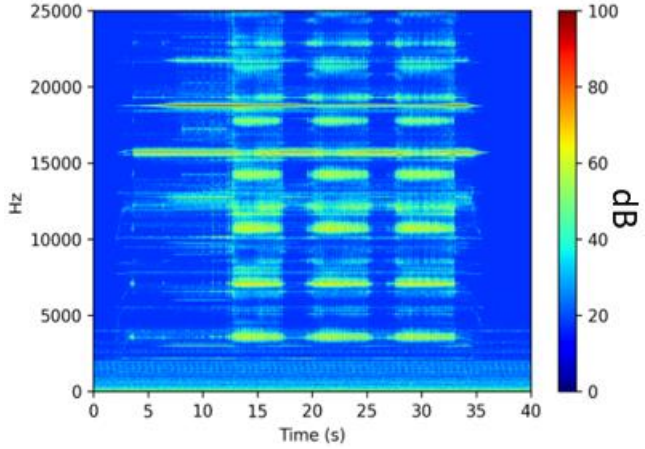


Figure 19: The spectrogram corresponding to experiment no.96

It should be noted that the acoustic characteristics are not the same across different tests. The first spectrogram Figure 18, related to the test no.18, has some clear continuous lines over time without many fluctuations. While on the spectrogram of the test no.96 Figure 19, the response seems more disorganized over time. The difference between these two tests is in three aspects, disc rotation speed, the contact surface evolution and the other is the normal force, that may cause different responses such as acoustic signals from the system or different surface temperature evolution. In fact, looking at the recorded spectrograms, different responses can be divided into two general categories. Spectrograms that have continuous lines and have no undergone sudden frequency changes and have a stable state (class 1). Spectrograms that have sudden frequency changes, the frequencies produced by the system are not continuous over time and are unstable (class 2). On the first class of tests (such as test no.18), the tribological aspects appear to be in a steady state, while on the other class of tests (such as test no.96), they

appear to be transient. In other words, the surface phenomena related to tribology are more chaotic on the second class and must depend on a shorter wavelength. However, by increasing the number of thermocouples, more accurate information on the evolution of the surface temperature may be obtained in order to make a better prediction for the tests in class 2.

Although test no.96 was not able to do the predictions as precisely as no.18 due to disturbed evolution, but test number 93 could predict accurately despite the similarities to test no.96. The results are shown in Figure 20.

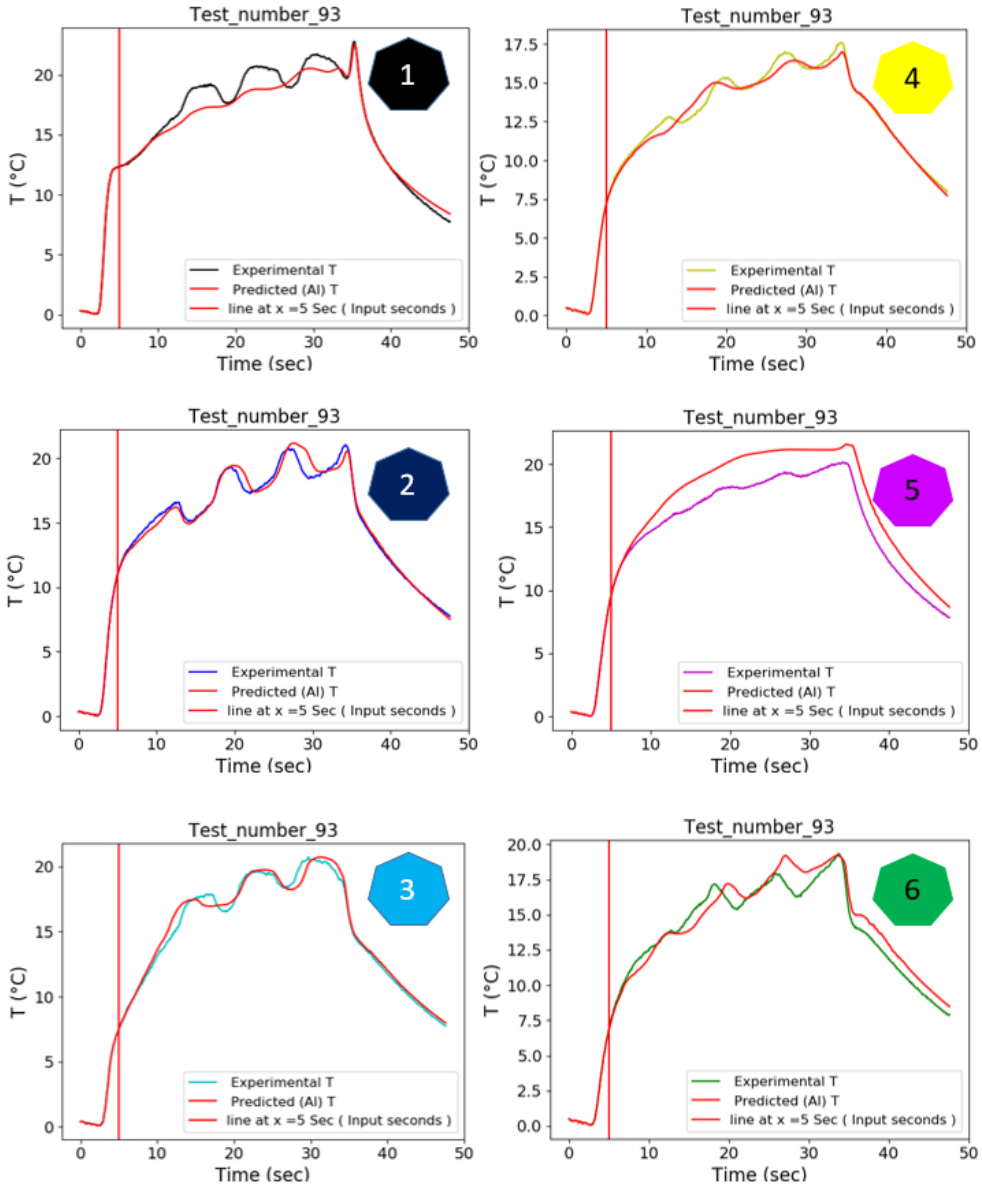


Figure 20: Predicting the temperature evolution of experiment no.93 by the DL model

Here a question is proposed which is why the model had 2 different conclusions in 2 similar tests (93,96), it might be due to the data which has been stored in the database that matches with test 93 requirements.

III.5. A FIRST STEP TO SMART BRAKING

One of the advantages of the proposed model is having the ability to adapt and face dynamic and continuous problems.

As seen in Figures 21a and 21b, the force applied in experiments no. 18 and no. 96 respectively, has a significant impact on the output frequencies, as seen in the corresponding temperature evolution (Figures 16 and 17) and spectrogram (Figures 18 and 19) of the experiments.

In fact, when the pin comes in contact with the disc, due to friction, the kinetic energy is converted into heat. In addition, when a greater force is applied from the pin to the disc, in less time the kinetic energy is converted into thermal energy. This force can also limit the movement of the disc and affect system vibrations.

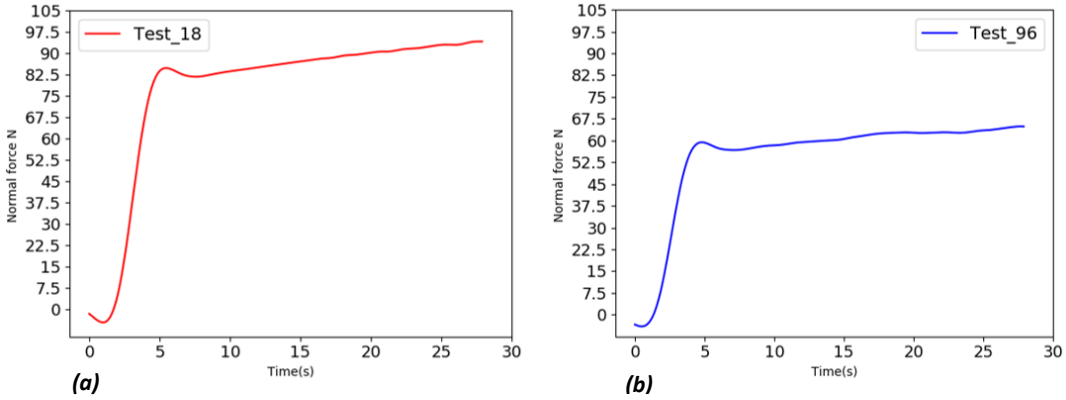


Figure 21a-b: Different normal forces applied in tests no. 18 and no. 96.

a) **The normal force corresponding to experiment no.18.**
b) **The normal force corresponding to experiment no.96.**

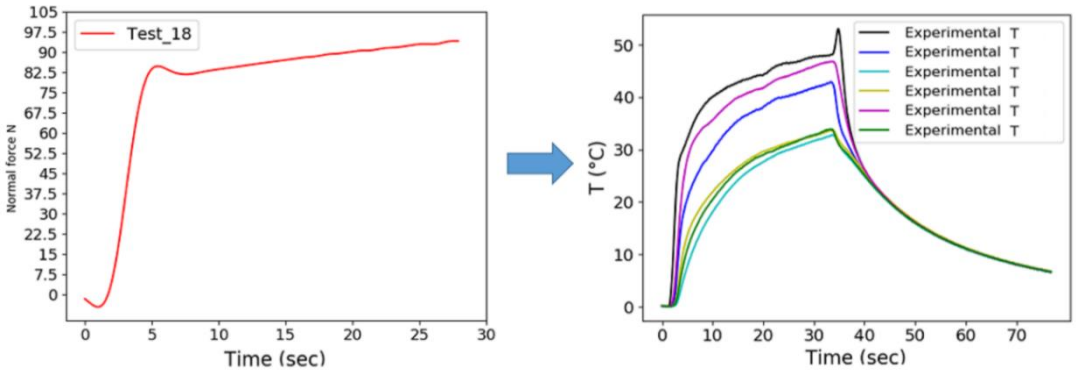


Figure 22: Applied force and experimental temperature evolution (Exp. no. 18).

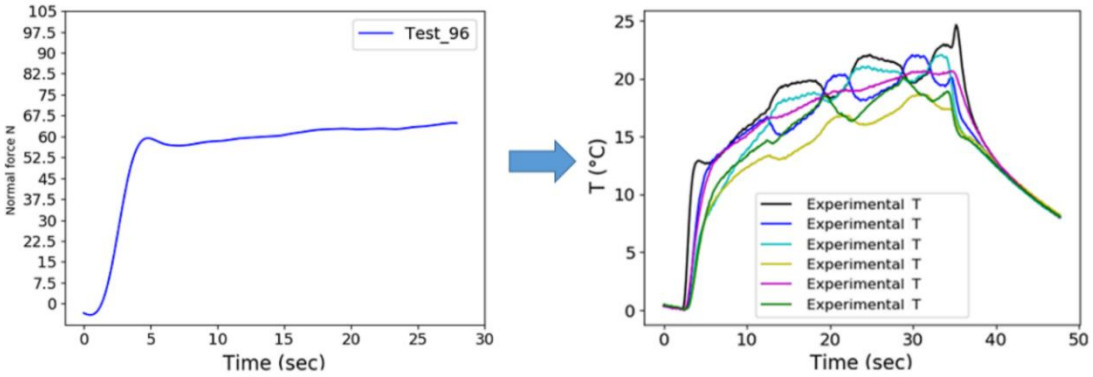


Figure 23: Applied force and experimental temperature evolution (Exp. no. 96).

Force amount in the mentioned experiments shows that when a larger amount of force is applied, the temperature evolution is steady (Figure 22), and when a smaller force is applied, the temperature evolution is not steady (Figure 23), as the associated spectrogram is very noisy. Based on this ability, and the surface temperature evolution sensitivity to the applied force (Figure 24), the idea of predicting the system's responses by changing the influential parameters in the system's behavior is proposed which is the smart braking system.

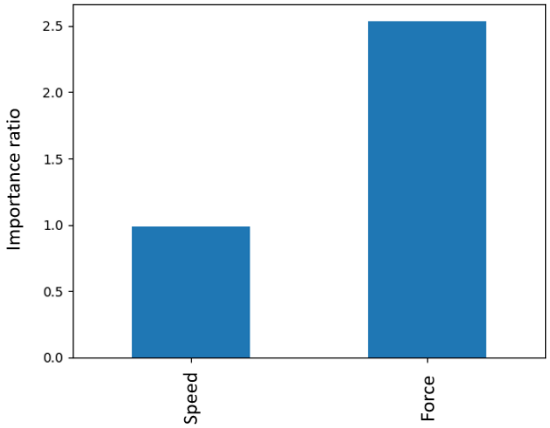


Figure 24: High sensitivity of the surface evolution to the applied load during the contact process

Now, the idea is raised to explore the effects of controlling the braking force as a parameter that can be easily adjusted by the user on the behavior of the system.

To achieve this, the force applied in experiment no. 18 was replicated in experiment no. 96, and the model was tasked with predicting the temperature evolution. Surprisingly it provided significant results. As observed in Figure 25, the model has a completely different prediction from this experiment and the sudden temperature evolution turned into uniform evolution.

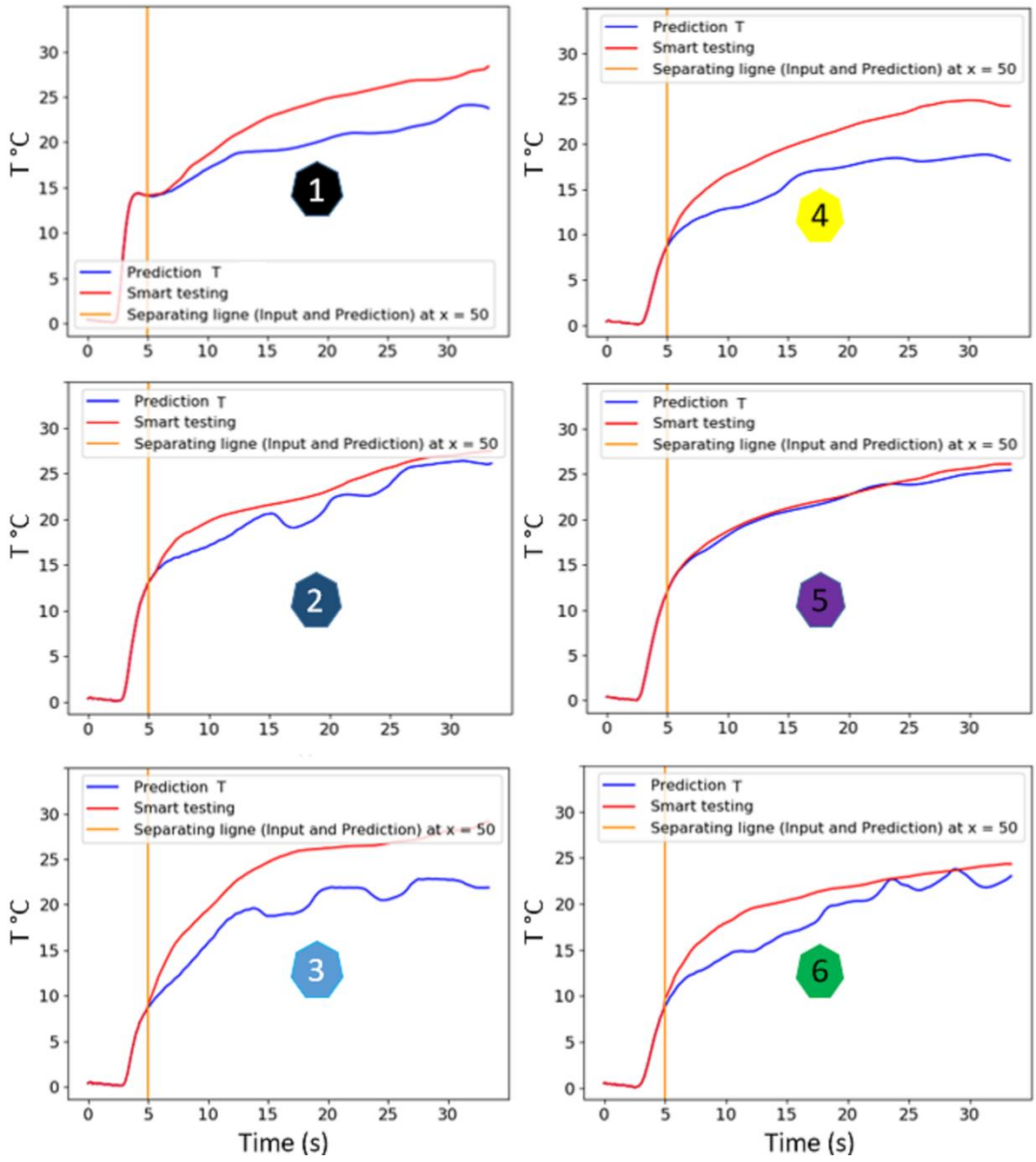


Figure 25: Results of force change to predict system behavior.

In fact, it can be interpreted that controlling the force can completely change the responses of the system and the temperature evolution of the system is highly sensitive to the force applied from the pad to the disc. Because of the applied force, the surface of the pad is uniformly in contact with the surface of the disc and temperature changes are uniform. As the model boasts a short calculation time, it proves to be highly advantageous in generating results quickly. This attribute further strengthens the concept of smart braking. In addition, it was shown that the frequencies produced by the system are unstable and undergo sudden changes when the evolution of temperature undergoes sudden changes. By putting these points together, it can be said that force during the contact process as a key parameter can play an effective role in system responses.

III.6. CONCLUSION

In this chapter, the contact process is investigated as a tribological problem. Since it is not possible to get much information about the state of the interface when the contact surface is closed, eight thermocouples (2 thermocouples were detached during the process and the tests continued with the remaining six) are placed in a pin made of conductive and friction-resistant materials to analyze the state of the contact surface by examining the temperature evolution. Due to the mechanical complexity and the physics of the contact problem which should be considered the history of temperature evolution, a 2D-RCNN has been used for its analysis and prediction.

There are two main motivations for using this architecture, since it has been applied on an experimental dataset including surface temperature changes during the contact process. Firstly, the location of the points involved during this process should be considered, that is why the convolutional neural network is effective. Secondly, the history of these changes should be considered, therefore, recurrent neural network is important.

Due to the heterogeneity of the experimental data, the training process was done discretely, but in the test phase, the prediction was done continuously. The used data for this prediction are the normal forces, disc rotation speed, and surface temperature changes. This data feed the model for an interval including 5 seconds at the beginning of the test, and after the fifth second, normal force and disc rotation speed are considered as known data, but temperature changes are considered unknown.

The error percentage of the model is 3.8%, which is acceptable and by increasing the number of thermocouples and obtaining more information from the interface, its performance could be improved.

The presented method for predicting the surface temperature evolution is interesting because it can predict the evolution of surface temperature in a fraction of a second (153ms) by only having information of less than 5s of the contact process.

The analysis shows that one of the main parameters that has a direct effect on the system behavior is the applied normal force to the contact surface, and in a pin-on-disk system, the force can limit the disc movements, as a result, the system has a more stable behavior in the interface. So, it can be concluded that the behavior of the system can be controlled.

Transition

Contact surfaces are constantly evolving due to the interaction between different mechanisms caused by the force and the contact of two objects in the sliding period. The mechanisms can be mentioned as stress concentration and heterogeneous surface temperature evolution during contact, which can lead to the generation of different topographies and different behavior of tribological systems. For example, the detachment of fine particles from surfaces can cause changes in surface topography. Particles leaving the system can result in environmental pollution or respiratory problems. Particles sticking to contact surfaces due to force or temperature can cause surface roughness, leading to system malfunction and risk of noise pollution or performance issues.

Therefore, knowing and predicting these changes can be useful to prevent such disorders.

But the prediction of these changes is very complicated and hard to make due to the existence of different mechanisms, and heterogeneous materials that lead to different behavior or different roughness.

As it is concluded, roughness is playing a major role in the system function. In Chapter II, it was shown that different roughness can cause different vibration frequencies and in Chapter III, it was shown that different roughness can cause temperature evolution which directly affects the system behavior. In order to prevent system malfunctions, it is necessary to predict the surface evolution. Here, under constant conditions, these predictions are made. To perform this action a model is presented which uses deep learning (DL) using an experimentally generated dataset. The analyzed part is marked in a blue box in the main diagram (Figure IV).

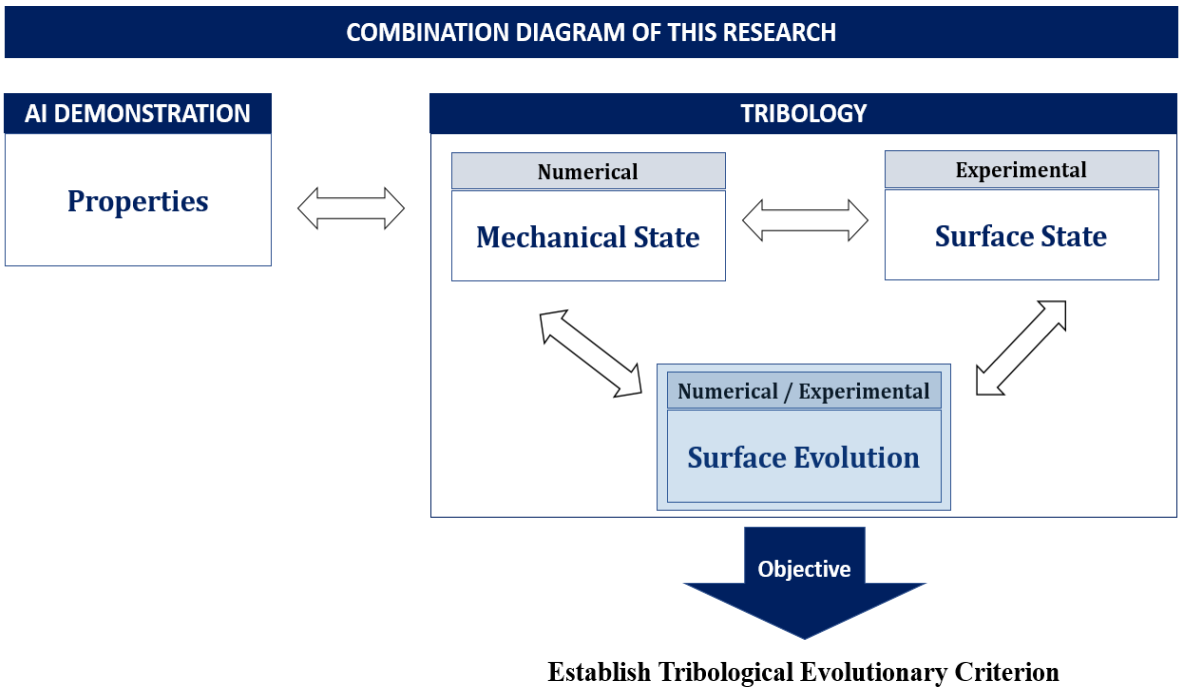


Figure IV: The general scenario of the thesis, which in this chapter the surface state, during the contact process is analyzed.

IV. Prediction of Surface Evolution

Table of contents

Abstract

According to the concept of the tribological circuit, the surface conditions of the first bodies are constantly changing. This is due to a combination of geometrical surface defects, different flow rates and local loadings. Moreover, everything is established in a multi-physical and multi-scale context. The system performances (friction, noise, wear, etc.) are very dependent on these evolutions which have been difficult to model. Indeed, the difficulties in obtaining a relevant model are both on the experimental and numerical sides. From experimental point of view, despite fine measurements, the understanding and the qualification of the phenomena is still a difficult problem. From a numerical point of view, the models (the equation of the phenomena) and their approximation still raise a lively debate in the community. In this chapter, we propose an approach which aims to predict the surface evolution on the pin side. This approach is based on generative adversarial networks (GANs) using an experimental dataset. More precisely, profilometry measurements obtained at the end of a test cycle will be introduced into an artificial intelligence architecture. The results show that the global phenomena are rather well predicted.

Keywords: Wear, Deep Learning, Pin-on-disc configuration.

IV.1. INTRODUCTION

When two surfaces come into contact, the existence of various factors such as surface roughness, sliding speed [179][180], normal force [180], different mechanical properties[181], surface contact temperature [182], etc., cause the evolution of the contact surface. According to the mentioned factors, surface evolution can occur locally or on the entire surface, which is commonly referred to as "wear". There are multiple mechanisms of wear, and the contact surface can be influenced by one or more of these mechanisms, depending on factors such as environmental conditions, type of loading, and others. According to Archard's research [183], several mechanisms can be mentioned:

- **Adhesive wear** is one of the most prevalent mechanisms that causes the modification of the contact surface. When the small particles from the contact surfaces remain in the interface and are stuck to the contact surfaces under the effect of normal force and temperature due to friction, they do not leave the system [184][185][186]. This type of wear has a significant effect on the components of a tribological system [187]. (Figure 1)

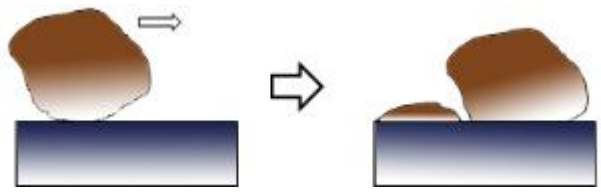


Figure 1: Adhesive wear schematic.

- **Abrasive wear** occurs when two rough surfaces come into contact, and during this process, minute particles separated from the surfaces either leave the system or remain in the interface and slide between the two surfaces [188][189][190]. (Figure 2)

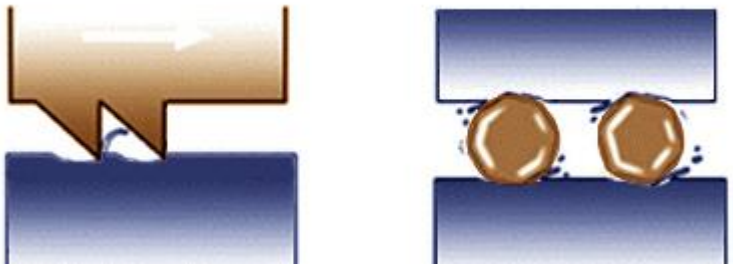


Figure 2: Abrasive wear schematic

- **Corrosive wear** occurs when the contact surfaces are placed in an environment with a risk of corrosion, and during the contact process and sliding, the corroded or oxidized layer is removed

from the contact surface, which causes the surfaces to be continuously exposed to corrosive environmental conditions and accelerate the corrosion process [191][192][193]. (Figure 3)

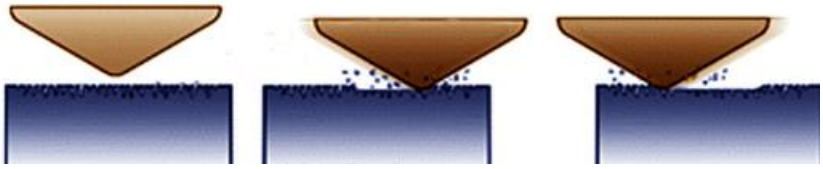


Figure 3: Corrosive wear schematic

- Fatigue wear occurs when the contact surfaces are in contact in a repeating cycle, micro-cracks are formed on the surface or under the contact surface due to the force that the surfaces exert on each other. These cracks grow over time and loads, until they cause the separation of parts from the contact surface [194] [195]. (Figure 4)

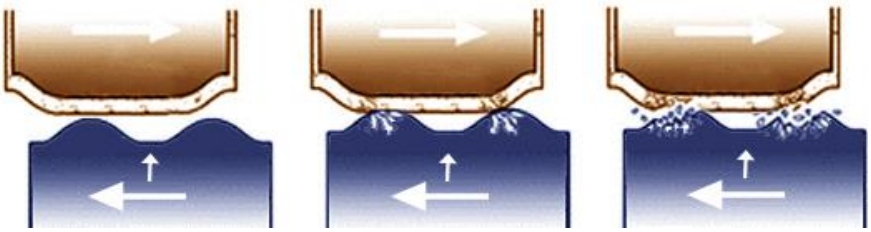


Figure 4: Fatigue wear schematic

These mechanisms vary the contact surface in terms of surface roughness [196], which subsequently change the mechanical responses of the system. Finally, it leads to a disruption or alteration in the behavior of the system. These disturbances can be shown as system squealing [163], friction-induced vibration [197][196] particles emission [198][199], which easily enter the respiratory system due to their small size. These factors affect the performance of the system. In addition, these disorders cause environmental or noise pollution [200][201], which imposes considerable expenses to the industry and treatment sector.

Therefore, knowing the behavior of surface evolution can help to predict and prevent their occurrence. Many different methods were presented for analyzing the surface evolution. For example, [187] studied the complete process of wear particle formation (i.e., nucleation, evolution, and detachment) during adhesive sliding contact using a coarse-grained numerical technique. They realized that by reducing adhesion, the probability of wear particle formation at the asperity contact decreases. [202] Analyzed the influence of the wear partition factor and modeled the wear contact surface based on Archard's wear law. [183] Used the finite element method, and [203] found, through interrupted rotating-bending fatigue tests, that despite the wear mechanism remaining unchanged, the distance from each fretting damage zone boundary

to the contact edge changes with the number of fatigue cycles. [204] Also studied the material response of stainless steel to incremental particle impact and the evolution of surface and subsurface wear with time during erosion-corrosion. [205] presented that the relative impact of pressure sliding amplitude and the number of fretting cycles can be assumed as a single parameter by introducing an energy wear concept.

However, due to the complexity of the contact mechanism, they usually model the evolution of the contact surface using simplified or idealized models, or they might include some parameters such as force, type of material or surface roughness in the model and which are not considered at the same time Despite the complexity of these models, they are known to be time-consuming and require a high level of proficiency [153]. But the evolution of the surface in tribological systems such as friction brakes are subjected to several different mechanisms at the same time, and the factors such as sliding speed, normal force, type of material, contact duration, temperature of the contact surface, roughness, etc. are involved in the surface contact evolution.

Recently, methods based on deep learning have been introduced to study it. For example, [206] tried to classify the wear on the contact surface of a cutting machine by using deep learning, and the model performance was acceptable in some categories and failed to meet expectations in some categories. In addition, this approach has weaknesses in order to classification complex systems such as friction brakes that are performed at high speed and under different environmental conditions and under the effect of forces, temperature and different materials. On the other hand, this model will not be able to predict the surface state, which is one of the main needs of tribology systems. Alternatively, [207] estimates the tool wear by using one-dimensional convolutional networks, but the actual results are far from the predicted results and these models do not predict the evolution of two-dimensional surfaces.

According to the mentioned points, it is necessary to provide a way which can model the evolution of the contact surface in real conditions. For this reason, in this section, a model is presented using Generative Adversarial Networks (GANs) [208] to analyze and predict the evolution of the contact surface. In order to solve this issue, a deep learning model is applied with an experimental database which is described in the following.

IV.2. GENERATING THE EXPERIMENTAL DATABASE

IV.2.1. Presentation

According to the purpose of this section, which is to analyze and predict the evolution of the surface contact state, a pin-on-disk system is used [175][48]. The disc material is 15CrMoV6 steel and the pin is made of composite material. Figure 5 shows the experimental setup.

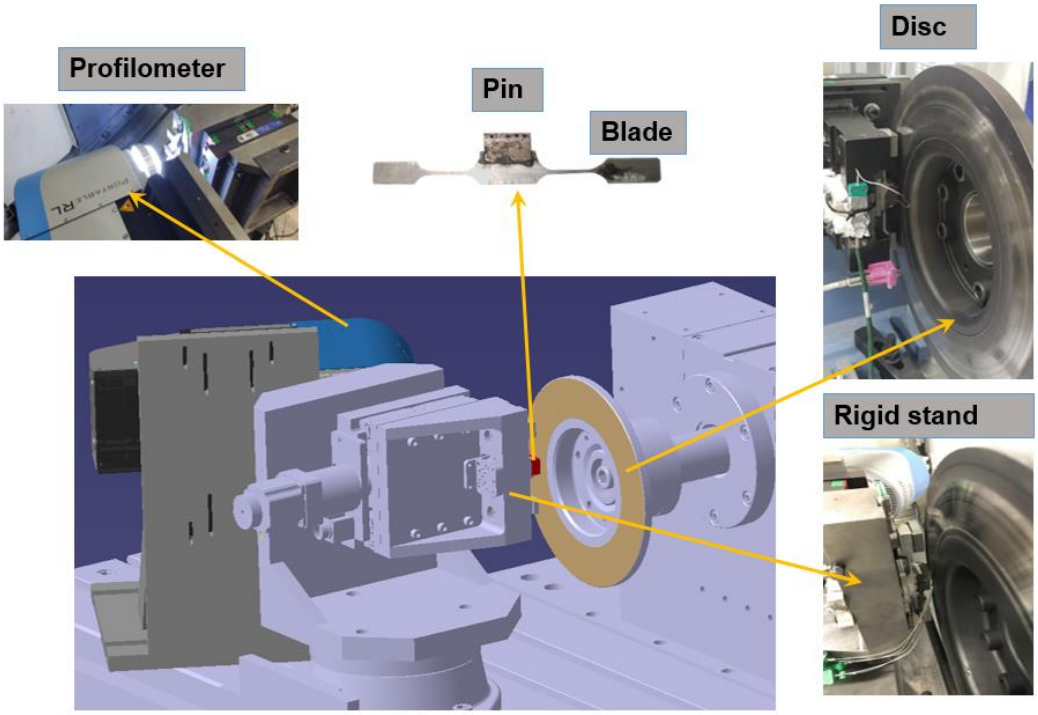


Figure 5: Pin-on-disk system and used equipment to perform experimental tests.

For this purpose, a rigid stand is used, which is connected to the blade at both ends, and a pin with a dimensions of 20 mm x 20 mm is connected in the center of the blade. In order to perform the contact process, a displacement is applied to the stand so the pin is in contact with the disc and the bending created in the blade applies the normal force from the pad side to the disc.

A profilometer is located at a distance of half a meter from the disk and the pin is placed at the average distance of 30 mm from the profilometer after rotation. The pin rotates 90 degrees around the base axis and is placed in front of the profilometer, to perform profilometry, no disassembly is done on the system so that the pin does not suffer shock or vibration and profilometry can be performed with minimal damage. In order to make a fixed reference for performing profilometry, a distance of the profilometer is set to the blade of the pin holder so that the profilometer records the surface changes compared to a fixed reference.

After each 10-contact process, the pin rotates 90 degrees to the profilometer to perform profilometry. Given that the process is time-consuming (approximately 30 minutes, including 15 minutes for machine shutdown, cable connection and disconnection for profilometry, and an additional 15 minutes for restarting). A profilometry was performed after every 10 contact tests. Although this process may cause a smaller database to be created, it can be positive in the sense that the changes made in the contact surface are more for every ten contacts and the model can receive more changes from the evolution of the surface, which may be useful for the model. Figure 6 shows two examples of profilometry (no.1, after first 10 brakes and no.2 after the second 10 brakes).

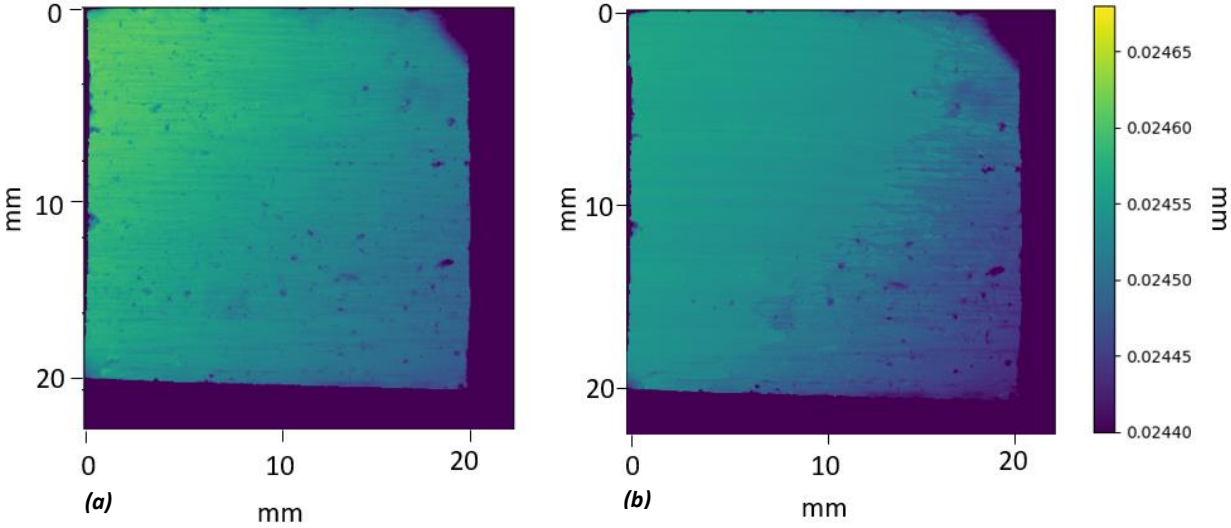


Figure 6a-b: Figure a, shows the profilometry for the complete surface (20*20 mm²) after 10 contact, and Figure b after 20 contact.

a) Profilometry number 1.
b) Profilometry number 2.

IV.2.2. Protocols and illustrations of the experimental results

The experimental test sequence is decomposed into series of tests (330 tests which made 33 profilometry records). The rotation speed is remained constant (drag conditions) for each trial. The distribution of the normal force applied to the surface is shown in Figure 7. In this figure, the blue points represent the tests conducted, the red points represent the profilometry data used during the model training phase, and the yellow points represent the data used during the testing phase.

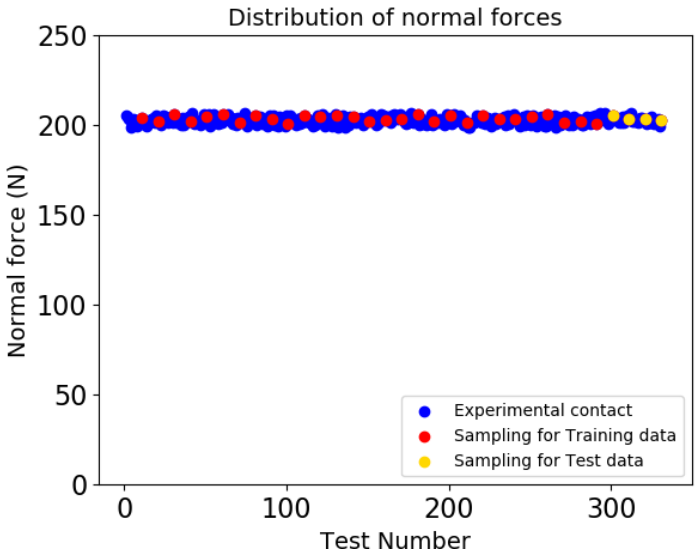


Figure 7: Distribution of normal force in the test series.

Time is another parameter which affects the test results. In fact, in each test, the pin is in contact with the disc for 30 seconds in the form of a drag braking, then the pin is disconnected from the disc for 30 seconds, and the cooling phase takes place, and the pin contacts to the disc again. In Figure 8, the duration of the tests is shown with blue points. The red points indicate the tests where profilometry is performed and the used data in the training phase. The yellow points indicate the tests where profilometry is performed and the data used is in the model testing phase.

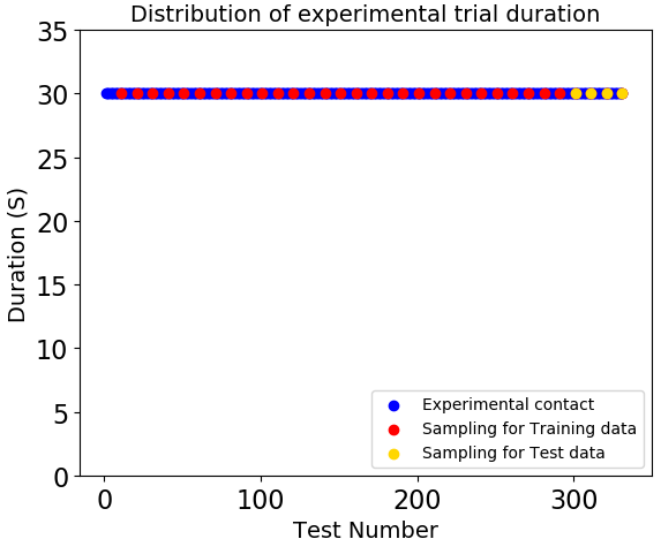


Figure 8: Distribution of tests duration in the test series.

In the experiments, the disc rotation speed is 200 rpm, and its distribution is shown in Figure 9. The blue points show the disk rotation speed in the experimental tests. The red dots indicate

the experiments in which profilometry was performed and were used for model training, and the yellow dots indicate the experiments in which profilometry was performed and were used in the model testing phase.

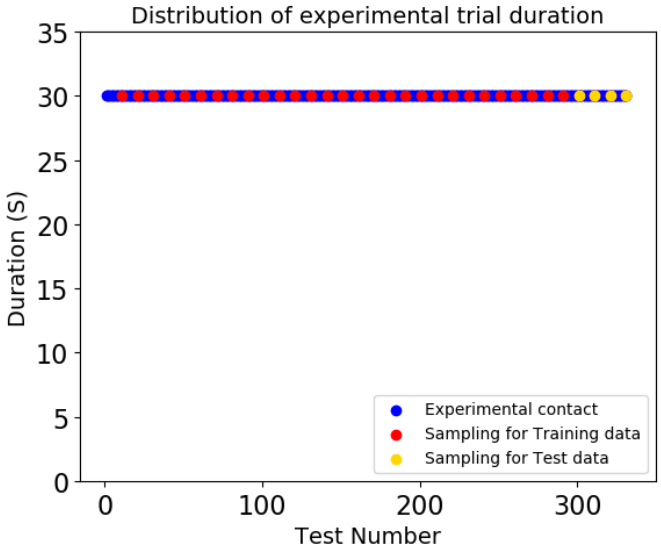


Figure 9: Distribution of normal force in the test series.

IV.3. AI MODEL FOR PREDICTION THE CONTACT SURFACE EVOLUTION

IV.3.1. General presentation of the model

According to the problem that is the contact surface evolution using Generative Adversarial Networks (GANs) seems relevant, because GANs models have both the ability to predict and generate 2D images that fit the needs of this issue. GANs is a deep learning architectures for training a generative model. It was first described in [208], however the initial models worked but were unstable and difficult to train. Therefore a standardized approach called Deep Convolutional Generative Adversarial Networks, or DCGAN, that led to more stable models was later formed by [209][210].

The GAN model architecture includes two sub-models: a generator model for generating new examples and a discriminator model for classifying whether generated examples are real (from the domain) or fake (generated by the generator model).

- **Generator model** is used to generate new plausible examples from the problem domain.
- **Discriminator model** is used to classify examples as real (from the domain) or fake (generated).

Generative adversarial networks are based on a theoretic game scenario in which the generator network must compete against an adversary. The generator network directly produces samples but its adversary, the discriminator network, attempts to distinguish between samples drawn from the training data and samples drawn from the generator [211].

Figure 10 shows the schematic and the learning process of a GANs model.

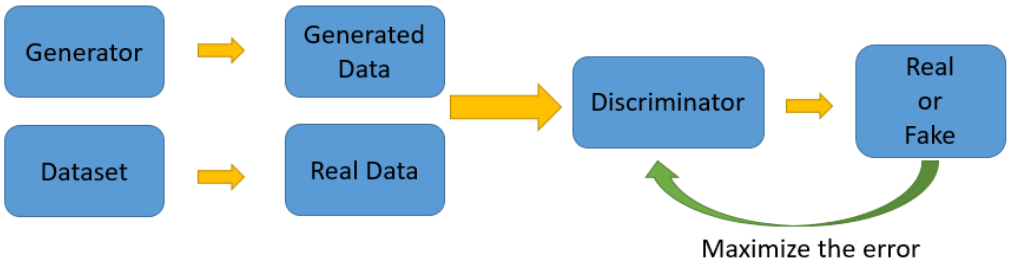


Figure 10: Schematic of a GANs model.

IV.3.2. Global view of the problem resolution

After creating the dataset and choosing the algorithm that suits the needs, it is time to define a scenario to solve the problem. The adopted scenario in this chapter, in order to provide an efficient model consists of five steps. First, experimental tests are performed in such a way that the pin is brought into contact with the surface of the disc with a certain force. Second, after every 10 tests, the state of the contact surface is recorded by profilometry. Third, the data is stored on the dataset. Fourth, the data is entered into the model and the training phase is carried out so that the model finds a correct connection between the inputs and outputs. Finally, after training the model, its ability is checked by data outside the database. The used scenario is shown in Figure 11.

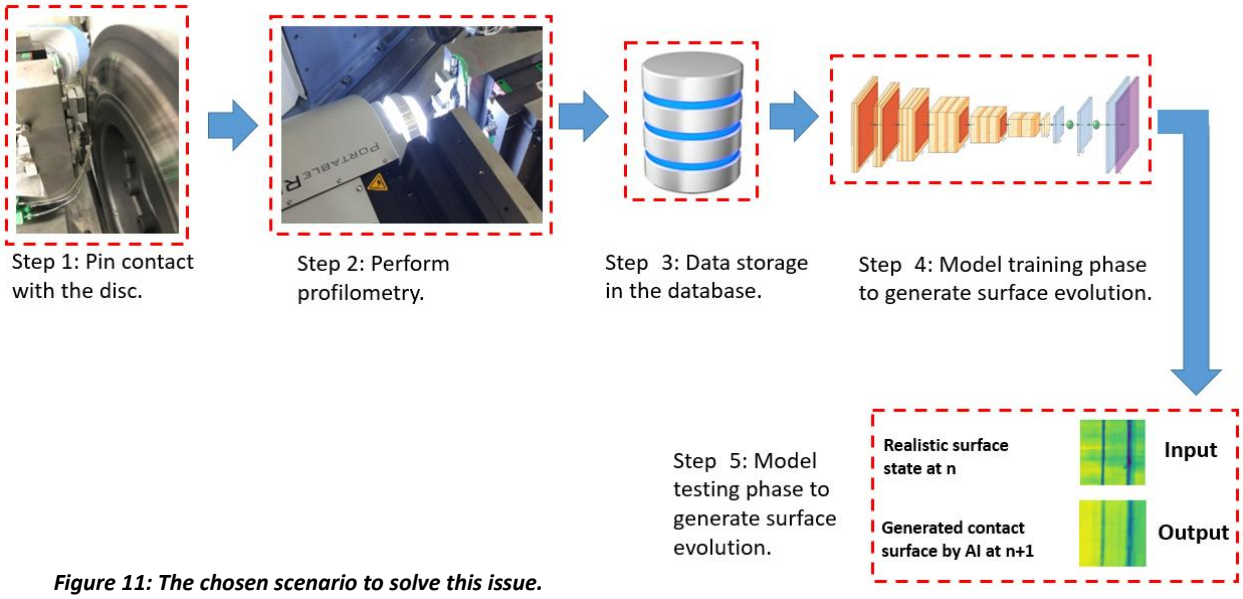


Figure 11: The chosen scenario to solve this issue.

IV.3.3. AI model for surface evolution generating and prediction

The surface contact state is one of the main factors involved in the behavior of tribological systems. It was shown in Chapter II that the occurrence of instability of a pin-on-disk system, is directly related to the roughness of the contact surface. Since the contact surface in a tribological system is constantly evolving under various factors over time and these changes cause different behavior of the system, predicting these changes can be useful to prevent system malfunctions. However, the changes in the contact surface are the result of the action of various mechanisms, and the complexity of this issue has made the presented models not having accurate predictions. For example, by looking at Figures 12a (test no. 90, profilometry no.9) and 12b (test no. 100, profilometry no.10), it can be seen that there are grooves on the contact surface which can be caused by the separation of hard materials such as ceramics at the entrance of the contact surface (shown with a red circle in Figure 12a) and has extracted through it, which has completely changed the contact surface roughness (shown by the red rectangle in photo 12b).

For this reason, in this chapter, prediction of surface roughness evolution is discussed using a DCGANs architecture. In fact, the goal is to give the state of the contact surface at time **n** to the algorithm and predict the state of the surface at time **n+1.** Figure 13 shows the used architecture for this purpose.

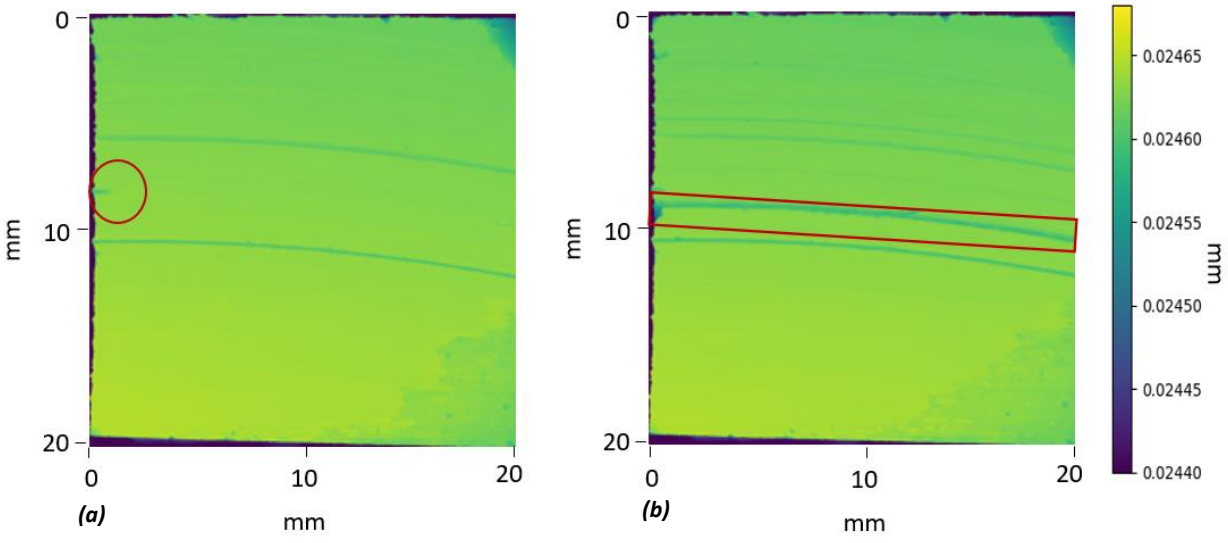


Figure 12: The process of grooving creation

a) Strat point of grooving (test no. 90, profilometry no.9)
b) Grooving creation (test no. 100, profilometry no.10)

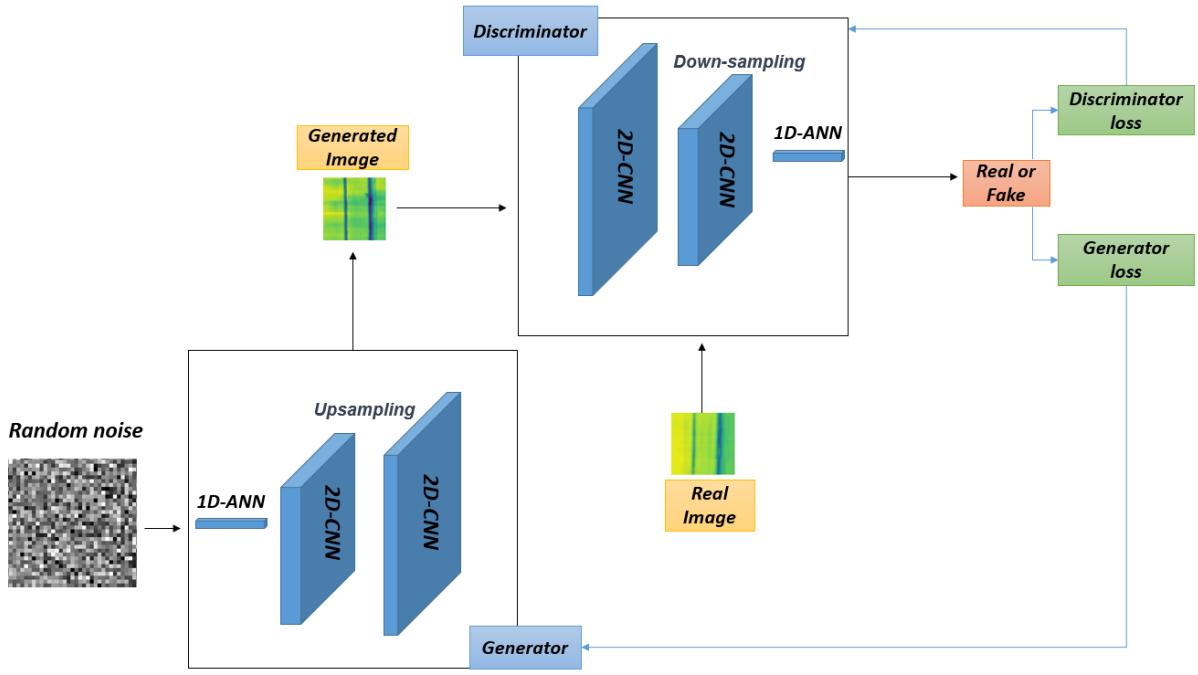


Figure 13: The used mechanism of the model to predict the contact surface evolution.

IV.3.4. Different scenarios to solve the problem

- First strategy: To train the model, the initial 29 profilometry data points were used. The model then aims to learn the underlying relationship between the input and output data, where the input is the surface state at time **n** and the desired target is the surface state at time **n+1**. Since a GANs model consists of two parts, the generator and the discriminator, the hyperparameter adjustment with the iterative strategy is very time-consuming. For this reason, according to the task of the discriminator, which is to classify the surfaces produced by the generating network, the efficient hyperparameter found in Chapter II to classify the roughness of the surface is used. However, the hyperparameters of the generative network were adjusted using an iterative strategy. In fact, the generating network should predict and generate the evolution of the surface so that it makes the minimum difference compared to the actual surfaces and the discriminator model cannot find a difference between the produced results and the actual results. For this reason, the value of the loss function of the GANs model should be reduced to lead the error percentage of the discriminator model increase, which means that the discriminator model cannot find a difference between the generated data and the real data, and the real label is assigned to the generated surfaces. In summary, during the training phase of the GAN model, the hyperparameters of the discriminator are fixed while those of the generator are adjusted. The two components, the generator and the discriminator, work together in an

end-to-end manner to improve the overall performance of the model. The end-to-end training process involves the generator and discriminator models constantly interacting with each other, with the output of one serving as input for the other, until the training is complete and the GAN model is optimized for the task at hand.

The generator's role is to predict and generate the evolution of the surface, minimizing the difference between the actual surface state and the generated one. The discriminator, on the other hand, classifies the generated surfaces and determines their similarity to the real surfaces. By adjusting the hyperparameters of the generator, the goal is to make it harder for the discriminator to distinguish between the real and generated data, thus improving the overall performance of the GAN model. Here, the GAN architecture is a 2D Convolutional Neural Network (2D-CNN) GAN, where both the generator and discriminator components have 2D-CNN structures. The generator's function is to create new data that resembles the real data it has been trained on. For a 2D-CNN GAN, the generator takes a random noise vector as input and produces a 2D output, such as an image. The generator's ability to generate realistic images is honed through its interaction with the discriminator in the adversarial training process. Table 1 shows the used hyperparameters in the discriminator model.

Table 1: Model specifications for classification

Model specifications	
Learning method:	Supervised learning
Algorithm:	Classification via "2D-CNN"
Aim:	Classification of contact surfaces
Number and type of layers:	6 (2D-CNN), 3 Max Pooling and 6 Dense layers
Loss function:	Binary- Crossentropy
Activation function:	sigmoid
Optimization function:	adam
Metric:	Categorical Accuracy

But in order to determine the hyperparameters of the model that generate the contact surface evolution, an iterative strategy is used, the goal is minimizing the value of the loss function, then the data is given to the discriminator model to distinguish the surfaces produced by the algorithm from the real ones. In fact, in order to reduce the training time of the model, four parallel models are used in such a way that in each of the models, one of the loss functions is used as a fixed hyperparameter. Then one of the optimizers is selected and the model starts training by changing the batch sizes and the number of epochs. At the end, the average loss

function is recorded for each configuration. These values for each of the hyperparameters are shown in the green box in Table 2.

Table 2: Model results for tuning hyperparameters (first attempt)

Loss_function :		mean_squared_error	mean_absolute_error	huber_loss	log_cosh
Metrics :	Training MSE	4.378	1.741	7.8821	8.4507
	Validation MSE	2.127	2.957	5.9576	4.4557

Optimizer :		Adagrad	RMSprop	adam	nadam	adamax	adadelta
Metrics :	Training MSE	7.6279	11.1	1.4178	6.387	3.899	5.6397
	Validation MSE	3.0455	6.47208	0.982	2.887	2.36	3.158

Batch Size :		30	60	120
Metrics :	Training MSE	1.412	4.817	4.6437
	Validation MSE	2.75	2.1107	1.5397

Epochs :		200	400	450
Metrics :	Training MSE	1.887	0.944	2.934
	Validation MSE	2.375	1.7922	3.367

The results show that the model is not converged and could not predict the contact surface evolution. In fact, the training and validation results (Table 2) show that the lack of data in the model training phase has caused the model to be unable to establish a relationship between the input and output data, and the model has suffered from poor fit. Also the discriminator has been able to identify with 100% accuracy all the surfaces produced by the model, which shows the weakness of the generator in producing data close to reality. Figure 14 shows the surfaces produced by the generator model. It is clearly shown the model has not been able to update the weight of the layers well due to the lack of data and has produced very poor results.

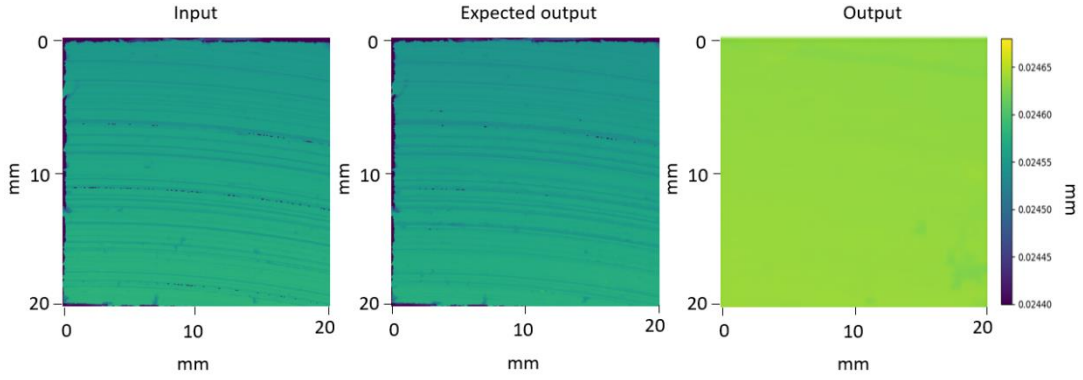


Figure 14: The results obtained from the trained model by 29 profilometry data and tested by 4 profilometry data

- Second strategy: Next, in order to solve the lack of data problem and improve the performance of the generating model, a technique was used to increase the data so that the model has enough data to complete the training phase.

According to the original size of the images, which is 2512x2512 pixels, a window with a size of 512x512 was created and it was moved on the contact surface so maximum of 120 sub-surfaces were generated and it lead the number of data to be increased.

For this reason, each profilometry was divided into 30, 60 and 120 parts, then the training phase of the model was done with different amounts of data.

Since DL models require a large amount of data (even when the data is of lower quality) to obtain satisfactory results to be more effective, the use of data augmentation is common to reduce overfitting and improve the performance[212][213][214].

The minimum loss function in the generative model is related to training the model by dividing each profilometry into 120 parts. In addition, the discriminator model has the maximum error percentage for identifying the data generated by the trained model. Figure 15 shows the error percentage of the discriminator model in the test phase.

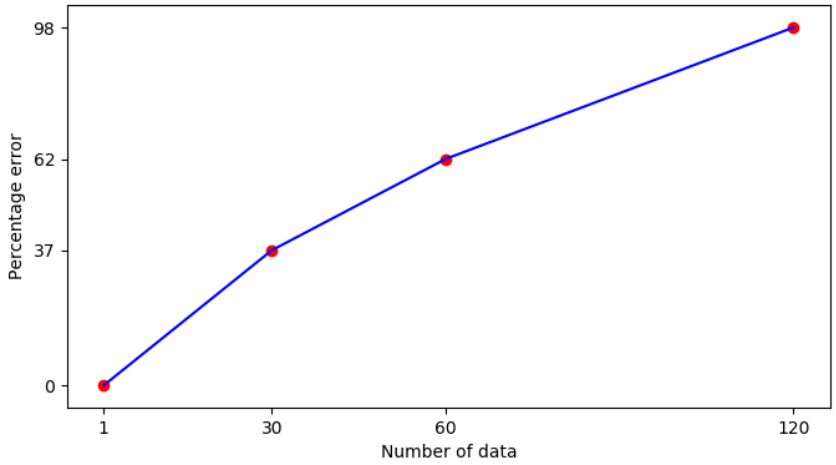


Figure 15: The model percentage error vs. increasing training phase data

Results show that as the number of data increases, the model is able to have better outputs. In fact, it can be concluded that when the profilometry is divided into 120 parts, the surface roughness in a suitable scale is provided to the model and the model can receive the information with good resolution.

In the generator model, the “relu” activation function was used in the encoder layers and the “tanh” activation function was used in the last decoder layer. Therefore, the “relu” activation function was selected in the encoder layers because the purpose of the problem is to predict the evolution of the contact surface and the reason for using the “tanh” activation function in the

last layer is that the input data of the model is normalized between -1 and 1 and since the range of the “tanh” function is between -1 and 1, this function was used.

The results show that the model using “mse” loss function had the best performance. In this way, it can be concluded that the distribution of surface changes during contact is in a certain range and the probability of surface evolution is higher in some places, and the distribution of these changes may be Gaussian distribution, that's why “mse” works better to find the minimum value.

In order to optimize the training process, adam optimization algorithm was chosen since it is an extension of stochastic gradient descent, it can be explained with the same argument and this optimizer randomly finds the minimum value for the model.

Batch size is the next hyperparameter to be set. The results show that batch size 120 has stable and acceptable results. In fact, since each profilometer is divided into 120 parts in each epoch, the 120 data that enter the model are the data of a complete level, that's why the model has presented a better trend by using 120 data.

In the end, the last hyperparameter to consider is the number of epochs. By using 400 epochs, the model was able to minimize the "mse" value. Table 3 shows the results of the iterative strategy to find the best hyperparameters of the generative model, and Table 4 shows the final specifications of the model.

Table 3: Model results for tuning hyperparameters (second attempt)

Loss_function :		mean_squared_error	mean_absolute_error	huber_loss	log_cosh
Metrics :	Training MSE	0.231	0.7537	0.6835	0.291
	Validation MSE	0.1974	0.492	1.489	0.5173

Optimizer :		Adagrad	RMSprop	adam	nadam	adamax	adadelta
Metrics :	Training MSE	0.614	0.535	0.143	0.5617	1.241	0.7736
	Validation MSE	0.8437	0.6482	0.1275	0.3207	0.71	0.5371

Batch Size :		30	60	120
Metrics :	Training MSE	0.095	0.04024	0.018
	Validation MSE	0.1143	0.09346	0.0231

Epochs :		200	400	450
Metrics :	Training MSE	0.00321	0.00014	0.000167
	Validation MSE	0.0071	0.000102	0.00022

Table 4: The final specifications of the generating model.

Model specifications	
Learning method:	Supervised learning
Algorithm:	Regression via "2D-CNN"
Aim:	Prediction of contact surface evolution
Number and type of layers:	7 (2D-CNN) and 7 Conv2DTranspose
Loss function:	mean_squared_error
Activation function:	Relu, Tanh
Optimization function:	adam
Metric:	mean_squared_error

The computation time for setting hyper parameters and complete training phase of the model by a server computer (Configuration: Memory: 128 GB, CPU @ 2.50GHz *40, Graphics: 32GB and a SSD 2 TB M.2 NVMe) is about 12 days.

IV.4. RESULTS AND DISCUSSION

After setting the hyperparameters and training the model in the testing phase, the profilometry data was divided into 120 samples and each of these subparts is entered into the model separately and the model predicts the state of the contact surface at the next moment. After making these predictions for the smaller parts, the subparts are connected together to reconstruct the full contact surface. In Figure 16, the results are presented, where the left image represents the input of the model, depicting the state of the contact surface at time **n** (test no. 310, profilometry no. 31). The middle image displays the state of the contact surface at time **n+1** (test no. 320, profilometry no. 32), and the right image illustrates the predicted state of the contact surface at time **n+1** generated by the model (prediction for profilometry no. 32).

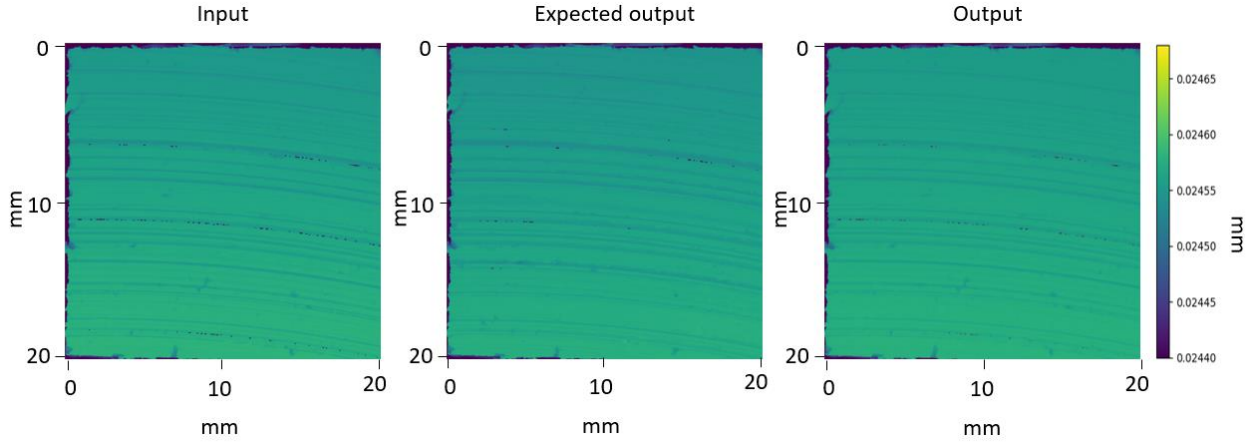


Figure 16: The results obtained by the deep learning model compared to the expected results

The time for making predictions and generating the pin surface evolution is approximately 0.1 seconds. By zooming in, and looking at the pin scale results, it can be seen that the model is able to make good predictions of the contact surface evolution, but there are some differences between the predicted results and the expected data (Figure 17). In fact, these differences can be caused by various factors, such as heterogeneous stress distribution on the contact surface, which causes the temperature to rise in some zones, and leads to the expansion of the pin and disc, separation of the pin material that causes grooves on the surface of the pin and small particles that enter the system during the contact process and stick to the contact surface.

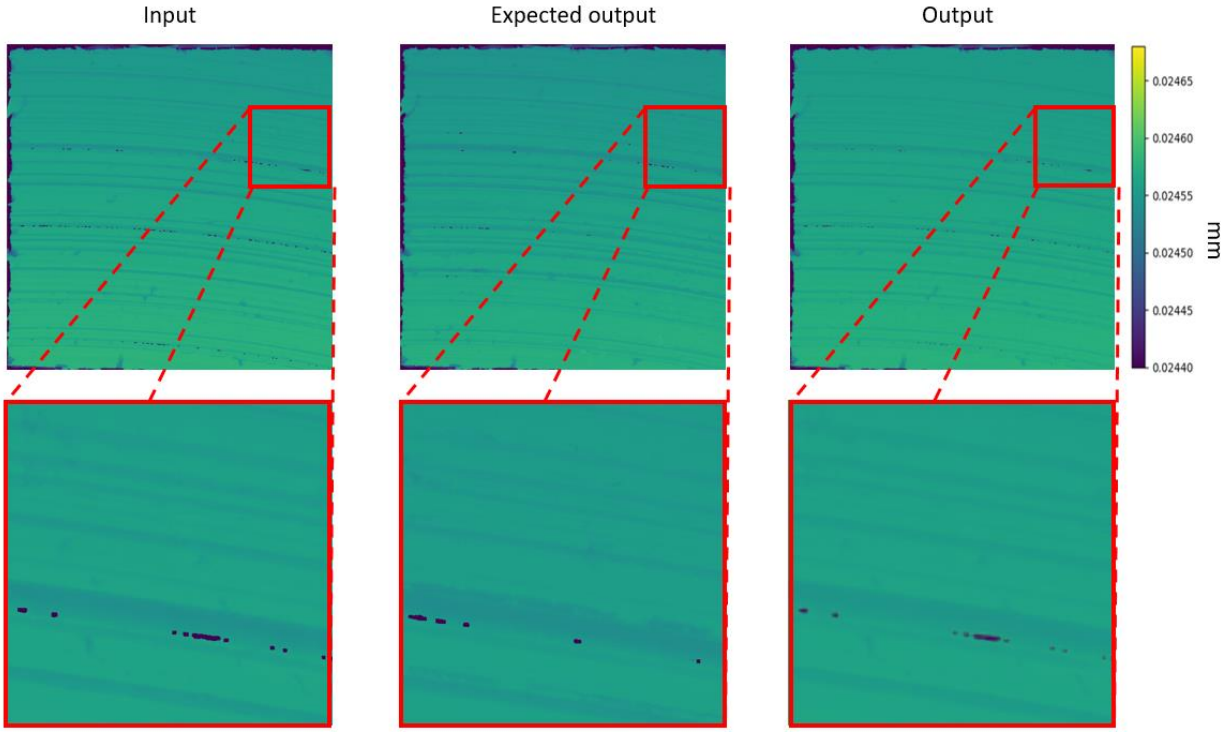


Figure 17: Minor differences in the results obtained compared to the expected results.

In Figure 18, the figure on the right displays the contrast between profilometry no. 31 and no. 32. The middle figure represents the contrast between test no. 31 and no. 32 (predicted), showing the difference between the AI results and the real profilometry to provide a clear visual comparison. The left figure illustrates the contrast between profilometry no. 32 and no. 32 (predicted). The results indicate slight differences between the predictions and the actual state, allowing for a comprehensive evaluation of the accuracy of the AI predictions.

In fact, it can be concluded that in smaller scales, the process of surface evolution is different compared to the process of surface evolution at the larger scales, because different mechanisms are involved in these scales, the model has less accuracy in production and predictions,

however, the discriminator model could not find a significant difference between the generated and the expected data.

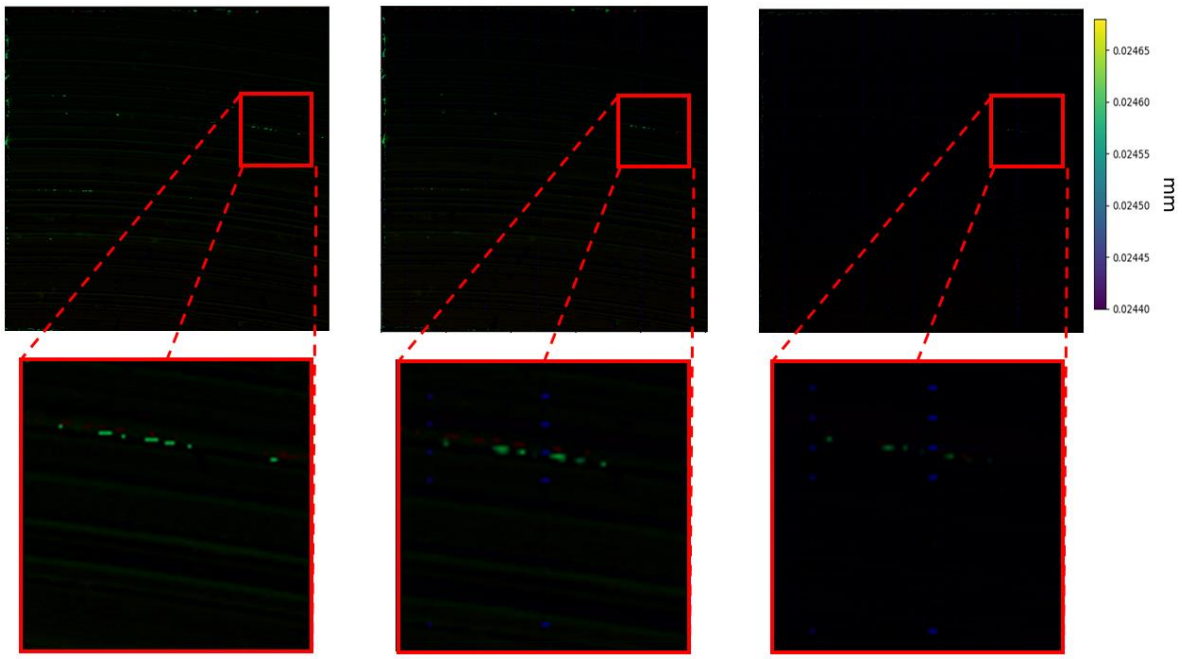


Figure 18: The blue spots indicate the border lines of the sub surfaces which have been overlapped. In the final regenerating they will appear as the blue spots.

In fact, the results can be interpreted from several aspects. In the original scale, with a general look at the results, it can be seen that the model has been able to obtain acceptable results using the analysis of the surface condition. In the smaller scale, it is possible to see differences in some points of the contact surface in experimental tests and predicted results, which can be caused by the presence of various materials in some areas of the surface or the concentration of stress in some areas due to the height differences in the contact surface. It may lead to the local increase of the contact surface in limited areas and the occurrence of heterogeneous surface temperature evolution in the contact surface and as a result different behavior of the contact surface. In the smallest scale, the presence of microcracks on the contact surface can be considered, which can be caused by normal force, tangential force, system vibration or sliding speed on the contact surface and lead to the separation of the contact surface materials and as a result changes random and complex contact surface.

IV.5. CONCLUSION

In this chapter, the evolution of the contact surface was investigated, and a model was presented to predict the surface evolution. In order to make these predictions, a GANs model was used to predict surface state at time **n+1** by examining and analyzing the surface state at time **n**. In order to train the model, an experimental database was used, the used data to feed deep learning model is the surface profilometry of the pin surface in a pin-on-disk system. It was demonstrated that the data recorded in the experimental phase is insufficient for training the model. Therefore, a technique was employed to augment the data, enabling the model to establish the relationship between the inputs and outputs more effectively. As a result, the model can better capture the patterns and dependencies within the data. It was shown that GANs models have the ability to predict surface evolution, and acceptable results were obtained. Although slight differences can be observed in certain areas, these variations can be attributed to various factors, including particle separation from the contact surface, elevated contact surface temperature, disc material and thickness, and more. By incorporating additional data, such as contact surface temperature throughout the contact process, stress distribution on the pin surface, state of the contact surface on the disc side, mechanical properties of the pin materials, and more profilometry measurements, the model's understanding of the contact process will be enhanced.

In fact, the presented model is able to generate the surface evolution of the contact surface in 0.1 sec. and before the contact occurs, and make it possible to predict the system responses. By using these predictions can prevent system disturbances such as system instability or the release of fine particles by the system can be identified before they occur, and preparations can be taken to prevent them in order to avoid huge costs in health, environment and after-sales services.

Conclusions and Perspectives

Conclusion:

In this thesis, mechanical issues have been addressed through artificial intelligence. Classification and regression problems were handled using data from other models or directly from experimental tests. More precisely, apart from the first chapter, which is a demonstrator of artificial intelligence for the mechanical community, the main problem is dedicated to tribology. This science contains many difficulties because it is multi-scale, multi-physical and evolutionary which have crucial effects on the prediction process. To simplify understanding the process, a 3-step categorization of tribology are proposed in the following.

- The first consisted in establishing the mechanical state according to a system considering rough surfaces.
- The second is aimed to physically identify the contact states according to the history effect of the contact.
- Finally, the third part is devoted to the evolution of contact states by trying to link them with the surface state.

Although this breakdown could be improved, a number of strong results have been proposed within this 3. step breakdown. At first, it appears possible to establish a link between the frequencies of a system and the contact surface. This link highlights a strong multi-scale relationship. Still, on this aspect of "mechanical state" and relying on AI, stress fields were always determined in the presence of irregular surfaces. This point of view will certainly in the long term make it possible to identify relevant criteria. In the second step, the prediction of contact localization was established based on temperature rises. Although it was necessary to rely on a few thermocouples embedded within the pad, the results deduced from the AI seem to indicate good relevance and good robustness despite various phenomena related to the parameterization of the input data (force, speed, test duration). This approach could highlight that the duration of the story effect (indicated as frame number in chapter 3) is important to consider. Finally, in order to better understand the evolution of surfaces, an approach based on a GAN scheme showed that the global point of view has been well characterized. More local phenomena (local grooves, heterogeneities, etc.) still need to be improved by relying on more experimental data (microstructure, history effect, frequency spectrum, etc.).

In general, mechanical problems, especially tribological problems, are intrinsically complex due to their nature, and the behavior of multi-physical, multi-scale, and evolutionary systems is so complex that there are many nonlinear behaviors in these fields, hence machine learning algorithms or even classical neural networks and they cannot be the right and effective choice to solve these problems. For example, in the process of contact, this mechanism occurs at the interface of two objects, and the spatial location of the contact points should be considered so that the model can have a proper picture of the details of the contact, hence the use of two-dimensional convolutional neural networks is a good choice because not only it considers the two-dimensional form of the problem, but also takes the interaction between different points of the contact surface into account.

Another effective factor in the behavior of tribological systems is the history of the contact, in which the model would be able to provide an appropriate understanding of the process and subsequently a precise prediction of the system's responses by considering the contact history, as far as the algorithms existing in the field of machine learning do not have such ability and the history of the problem cannot be taken into consideration using these algorithms, the use of deep learning algorithms and, more precisely, recurrent neural network algorithms can be a good choice in dealing with tribological problems, because contact history plays a key role in these problems and cannot be ignored.

It must be noted that, most of the existing hyperparameters and tools related to artificial intelligence and more precisely deep learning have been developed to use artificial intelligence algorithms for solving various problems rather than mechanical problems, which can be a key point because by developing a tool that fulfills the needs of mechanical issue the accuracy of the models would be increased and also the training time for the training phase could be decreased which is essential in calculation process.

Another point which plays an important role in mechanical issues is the geometry of the studied objects, because, by changing the geometry of the system, the received responses will be different. Therefore, the development of a tool that can classify and predict the appropriate responses of the system according to different geometries is very important and can be greatly helpful for the generalization ability of the models.

In summary in this thesis, by combining the chapters related to tribology, a comprehensive model can be reached in such a way that by analyzing contact location and contact surface changes, a model with high accuracy could be presented to predict roughness changes. Then the results of this model are entered into the model presented in chapter two, and according to the roughness of the contact surface, the risk of system instability is analyzed, and this cycle

can be completed for the entire life of the contact surface in less than a few seconds without performing complex calculations and time-consuming to predict the risks facing the system, so it is less expensive and time-consuming than many experimental tests.

In short, we look forward to have demonstrated within this manuscript that artificial intelligence is the most complementary and effective tool for the classical modeling due to its capabilities and capacities which can make it outstanding from the other methods and due to its characteristics, it has a strong potential to be developed in the numerical-experience coupling which can continuously result to time and energy reduction.

Perspectives:

In this part, the perspectives related to each chapter are presented separately, and in fact, it shows the potential for the continuation of the path of each chapter. This analysis provides a detailed overview of the key aspects of each chapter and highlights the potential opportunities for further exploration and development. The information presented in this analysis serves as a valuable resource for researchers and practitioners who are interested in the continued growth and evolution of the field. In fact, the analysis and insights presented in this chapter demonstrate the potential for further exploration in each area, and provide valuable guidance for future research and development endeavors.

In Chapter I, artificial neural networks were used to classify the behavior law and 2D-CNN were used to determine mechanical properties, which results showed the artificial intelligence capabilities in the mechanical field. In this research, the ability of artificial intelligence to determine the constitutive law parameters for relatively complex materials was shown and good results were obtained.

Therefore, this idea can be expanded, and this method can be used to determine the mechanical properties of heterogeneous materials. (Figure 1)

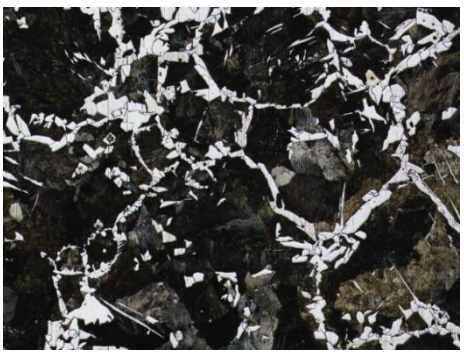


Figure 1: A microscopic heterogonous material example (Nagesh et al. 2021-2024)

For this purpose, the method introduced in this chapter might be used to determine the mechanical properties of heterogeneous materials, which are very complicated due to their behavior. But by using the proposed method, it is possible to record the real responses of the system with heterogeneous materials, and according to these responses, the material constitutive law and its parameters would be calculated.

In Chapter II, classifications and predictions were made to predict the risk of system instability, and acceptable results were obtained. To apply deep learning, loss functions, activation functions, optimizer, and existing functions of deep learning were used. In addition, by increasing the number of data and an iterative strategy to adjust hyperparameters, an optimal model for classification and predictions was presented. We believe that by developing loss functions, optimizer and activation functions tailored to the needs of the contact problem, results can be obtained with a smaller amount of data in the training phase.

For example, the activation function used in this section to predict instability frequencies is relu, (Figure 2) which is a linear function and if it can be developed to a nonlinear one such as the one shown in Figure 3, it would facilitate the model's training and generalizability process. As it can be seen from Figure 3, which is based on Hertz contact theory, the function goes through a different process. So, if an activation function would be developed following Hertz contact theory, it might cause the results of the model to be improved.

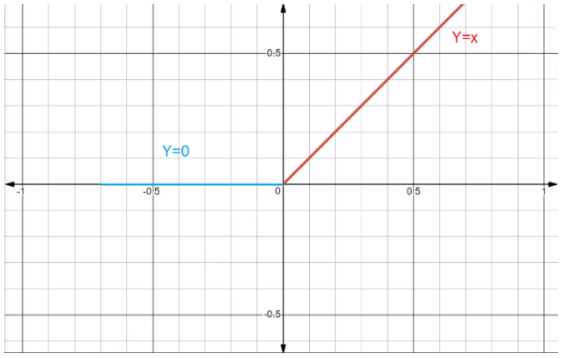


Figure 2: The used activation function curve for the prediction model

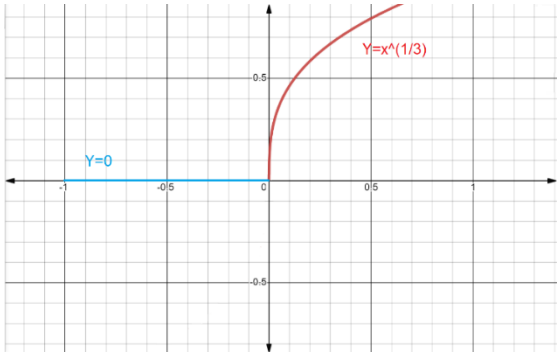


Figure 3: Proposed activation function based on Hertz theory

In Chapter III, the obtained results indicate that more studies can be done to improve the prediction of dynamic responses such as generated frequencies by the system. Since it was shown that there is a direct relationship between the temperature evolution and the frequencies generated by the system, this idea can be proposed to predict the generated frequencies by using surface temperature evolution and the force applied to the surface. For this purpose, the temperature and force data recorded from the experimental tests and the first few seconds of the spectrogram can be considered as input data and then the evolution of the received frequencies from the system could be predicted. Figure 4 shows the results that are not satisfactory.

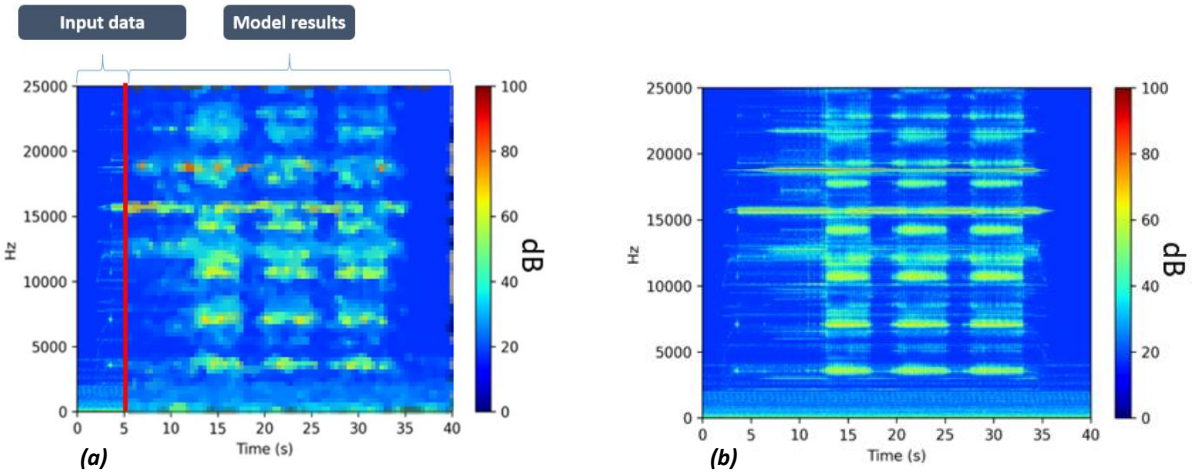


Figure 4: The model performance using the first 5 seconds data.

a) Model training phase and test results (spectrogram corresponding to experiment no.96).
b) The spectrogram corresponding to experiment no.96.

Therefore, better predictions could be made by providing more features such as surface roughness, increasing the number of thermocouples and system geometry to the model.

In Chapter IV, despite some weaknesses in certain areas, the model successfully obtained satisfactory results. It can be improved by incorporating additional physical data, such as the stress field.

Due to the roughness of the contact surfaces, the stress distribution is not uniform, leading to stress concentration in certain areas during the contact process. As a result, the evolution of the surface and surface temperature in some areas can be significantly different from other areas. Knowing the stress distribution during the contact process can thus be a critical factor in obtaining more accurate results. Therefore, it would be useful to develop a model that can

predict the stress field distribution based on the surface roughness. Figure 5 shows the results obtained from the deep learning model trained using numerical data, where the stress field is predicted using the surface roughness. Figure 5 displays the results, with the left image being the input of the model (the roughness), the middle image being the stress field obtained from a finite element model, and the right image being the stress field produced by the deep learning model.

In fact, presenting a model having several different effective inputs in the contact process can lead to more coherent models, so that the models will be able to predict the contact surface evolution more efficiently.

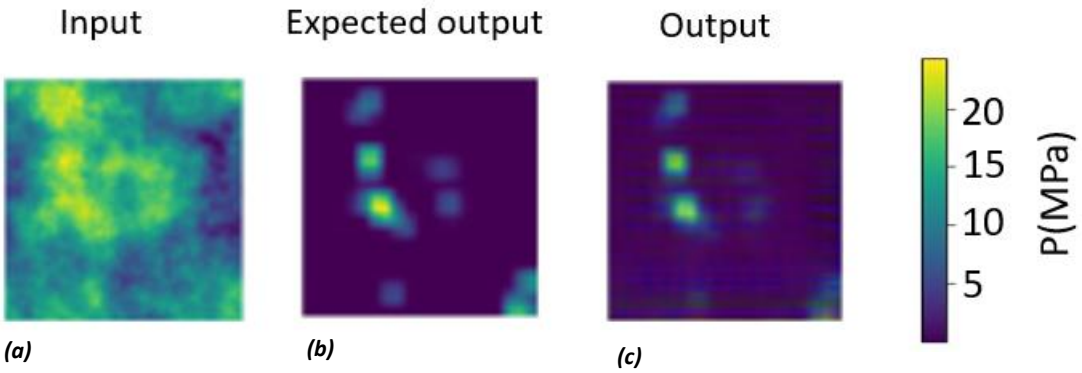


Figure 5: Generating stress field using deep learning model

- ***a) Surface roughness.***
- ***b) Stress field calculated by finite element model***
- ***c) Stress field generated AI model***

The content of this research presented at various conferences [1][2][3][4]. Also two research papers [5][6] were written and submitted for publication. The papers are currently under review for acceptance.

The author's work has been evaluated and discussed through these presentations and potential publications.

[1] N. Motamedi, V. Magnier, and H. Wannous, "Link between roughness and squeal: prediction by deep learning approach," 7th World Tribology Congress (WTC), July 7-12, 2022

[2] N. Motamedi, V. Magnier, and H. Wannous, "Identification of material properties from kinematic fields via deep learning," 15th National Conference on Structural Calculations (CSMA 2022), May 2022

[3] N. Motamedi, Vincent Magnier, Hazem Wannous, V-V Lai, J-F Brunel, and P. Dufrénoy. "Using Deep Learning to Identify the Evolution of Contact Localization and the Consequence

on the Friction-Induced Vibration." In Proceedings of Euro-Brake 2022 International Braking Technology Community & Event, May 2022.

[4] N. Motamedi, Vincent Magnier, and Hazem Wannous. "Utilisation du deep learning dans la prédiction des instabilités vibratoires: Application sur un système pion-disque avec défauts de surfaces." In Journée Internationale Francophone de Tribologie (JIFT), June 2021.

[5] N. Motamedi, V. Magnier, and H. Wannous, "Identification of Elastic Properties for Transversely Isotropic Materials via Deep Learning Exploiting the Kinematic Fields," European Journal of Mechanics / A Solids (Under Review).

[6] N. Motamedi, V. Magnier, and H. Wannous, "Towards the identification of the link between the contact roughness and the friction-induced vibration: use of deep learning," European Journal of Mechanics / A Solids (Under Review).

REFERENCES

[1] E. K. McLean and A. Tarnopolsky, “Noise, discomfort and mental health a review of the socio-medical implications of disturbance by noise,” *Psychological Medicine*, vol. 7, no. 1. pp. 19–62, 1977, doi: 10.1017/S0033291700023138.

[2] E. B. Blanchard, E. J. Hickling, T. C. Buckley, A. E. Taylor, A. Vollmer, and W. R. Loos, “Psychophysiology of Posttraumatic Stress Disorder Related to Motor Vehicle Accidents: Replication and Extension,” *J. Consult. Clin. Psychol.*, vol. 64, no. 4, pp. 742–751, 1996, doi: 10.1037/0022-006X.64.4.742.

[3] J. Dobson, “Toxicological aspects and applications of nanoparticles in paediatric respiratory disease,” *Paediatr. Respir. Rev.*, vol. 8, no. 1, pp. 62–66, 2007, doi: https://doi.org/10.1016/j.prrv.2007.02.005.

[4] J. Krutmann *et al.*, “Pollution and skin: From epidemiological and mechanistic studies to clinical implications,” *J. Dermatol. Sci.*, vol. 76, no. 3, pp. 163–168, 2014, doi: https://doi.org/10.1016/j.jdermsci.2014.08.008.

[5] M. Ammasi Krishnan *et al.*, “Effects of ambient air pollution on respiratory and eye illness in population living in Kodungaiyur, Chennai,” *Atmos. Environ.*, vol. 203, pp. 166–171, 2019, doi: https://doi.org/10.1016/j.atmosenv.2019.02.013.

[6] G. Straffelini, R. Ciudin, A. Ciotti, and S. Gialanella, “Present knowledge and perspectives on the role of copper in brake materials and related environmental issues: A critical assessment,” *Environ. Pollut.*, vol. 207, pp. 211–219, 2015, doi: https://doi.org/10.1016/j.envpol.2015.09.024.

[7] M. Vojtíšek-Lom *et al.*, “Effects of braking conditions on nanoparticle emissions from passenger car friction brakes,” *Sci. Total Environ.*, vol. 788, p. 147779, 2021, doi: https://doi.org/10.1016/j.scitotenv.2021.147779.

[8] A. I. Vakis *et al.*, “Modeling and simulation in tribology across scales: An overview,” *Tribol. Int.*, vol. 125, pp. 169–199, 2018, doi: https://doi.org/10.1016/j.triboint.2018.02.005.

[9] H. Hertz, “Ueber die Berührung fester elastischer Körper.,” 1882.

[10] K. L. Johnson, K. Kendall, and aAD Roberts, “Surface energy and the contact of elastic solids,” *Proc. R. Soc. London. A. Math. Phys. Sci.*, vol. 324, no. 1558, pp. 301–313, 1971.

[11] B. V Derjaguin, V. M. Muller, and Y. . Toporov, “Effect of contact deformations on

the adhesion of particles," *J. Colloid Interface Sci.*, vol. 53, no. 2, pp. 314–326, 1975, doi: https://doi.org/10.1016/0021-9797(75)90018-1.

[12] A. Constantinescu, A. M. Korsunsky, O. Pison, and A. Oueslati, "Symbolic and numerical solution of the axisymmetric indentation problem for a multilayered elastic coating," *Int. J. Solids Struct.*, vol. 50, no. 18, pp. 2798–2807, 2013, doi: https://doi.org/10.1016/j.ijsolstr.2013.04.017.

[13] C. Bagault, D. Nélias, M. C. Baietto, and T. C. Ovaert, "Contact analyses for anisotropic half-space coated with an anisotropic layer: Effect of the anisotropy on the pressure distribution and contact area," *Int. J. Solids Struct.*, vol. 50, no. 5, pp. 743–754, 2013, doi: https://doi.org/10.1016/j.ijsolstr.2012.11.002.

[14] S. J. Chidlow and M. Teodorescu, "Sliding contact problems involving inhomogeneous materials comprising a coating-transition layer-substrate and a rigid punch," *Int. J. Solids Struct.*, vol. 51, no. 10, pp. 1931–1945, 2014, doi: https://doi.org/10.1016/j.ijsolstr.2014.02.003.

[15] G. Stan and G. G. Adams, "Adhesive contact between a rigid spherical indenter and an elastic multi-layer coated substrate," *Int. J. Solids Struct.*, vol. 87, pp. 1–10, 2016, doi: https://doi.org/10.1016/j.ijsolstr.2016.02.043.

[16] C. Putignano, G. Carbone, and D. Dini, "Mechanics of rough contacts in elastic and viscoelastic thin layers," *Int. J. Solids Struct.*, vol. 69–70, pp. 507–517, 2015, doi: https://doi.org/10.1016/j.ijsolstr.2015.04.034.

[17] F. P. Bowden and D. Tabor, "The area of contact between stationary and moving surfaces," *Proc. R. Soc. London. Ser. A. Math. Phys. Sci.*, vol. 169, no. 938, pp. 391–413, 1939.

[18] B. N. J. Persson, O. Albohr, U. Tartaglino, A. I. Volokitin, and E. Tosatti, "On the nature of surface roughness with application to contact mechanics, sealing, rubber friction and adhesion," *J. Phys. Condens. matter*, vol. 17, no. 1, p. R1, 2004.

[19] F. Renaud, G. Chevallier, J.-L. Dion, and G. Taudière, "Motion capture of a pad measured with accelerometers during squeal noise in a real brake system," *Mech. Syst. Signal Process.*, vol. 33, pp. 155–166, 2012.

[20] "Frequency and time simulation of squeal instabilities." .

[21] T. K. Kao, J. W. Richmond, and A. Douarre, "Brake disc hot spotting and thermal judder: an experimental and finite element study," *Int. J. Veh. Des.*, vol. 23, no. 3–4, pp. 276–296, 2000.

[22] P. Dufrénoy and D. Weichert, "A thermomechanical model for the analysis of disc

brake fracture mechanisms," *J. Therm. Stress.*, vol. 26, no. 8, pp. 815–828, 2003.

[23] D. Naidoo Ramasami, "Influence of friction material & test sequence on disc brake squeal." Lille 1, 2014.

[24] R. Mann, "Experiments and thermomechanical modelling of braking application & friction material characterization with loading history effect." Lille 1, 2017.

[25] N. Hoffmann, M. Fischer, R. Allgaier, and L. Gaul, "A minimal model for studying properties of the mode-coupling type instability in friction induced oscillations," *Mech. Res. Commun.*, vol. 29, no. 4, pp. 197–205, 2002.

[26] V. Magnier, J. F. Brunel, and P. Dufrénoy, "Impact of contact stiffness heterogeneities on friction-induced vibration," *Int. J. Solids Struct.*, vol. 51, no. 9, pp. 1662–1669, 2014, doi: 10.1016/j.ijsolstr.2014.01.005.

[27] Y. Waddad, "Multiscale thermomechanical strategies for rough contact modeling : application to braking systems." 2017.

[28] Y. Waddad, V. Magnier, P. Dufrénoy, and G. De Saxcé, "A multiscale method for frictionless contact mechanics of rough surfaces," *Tribol. Int.*, vol. 96, pp. 109–121, 2016, doi: 10.1016/j.triboint.2015.12.023.

[29] Y. Waddad, V. Magnier, P. Dufrénoy, and G. De Saxcé, "Multiscale thermomechanical modeling of frictional contact problems considering wear – Application to a pin-on-disc system," *Wear*, vol. 426–427, pp. 1399–1409, 2019, doi: 10.1016/j.wear.2018.12.063.

[30] W. Yassine, "Multiscale thermomechanical strategies for rough contact modeling : application to braking systems." 2017.

[31] M. Renouf, F. Dubois, and P. Alart, "A parallel version of the non smooth contact dynamics algorithm applied to the simulation of granular media," *J. Comput. Appl. Math.*, vol. 168, no. 1–2, pp. 375–382, 2004.

[32] M. Renouf, H.-P. Cao, and V.-H. Nhu, "Multiphysical modeling of third-body rheology," *Tribol. Int.*, vol. 44, no. 4, pp. 417–425, 2011.

[33] J. J. Moreau, "Une formulation du contact a frottement sec; application au calcul numerique," *Comptes Rendus l'Academie des Sci. Ser. II*, vol. 302, no. 13, pp. 799–801, 1986.

[34] V.-H. Nhu, "Dialogues numériques entre échelles tribologiques." INSA de Lyon, 2013.

[35] M. Müller and G. P. Ostermeyer, "A Cellular Automaton model to describe the three-dimensional friction and wear mechanism of brake systems," *Wear*, vol. 263, no. 7–12,

pp. 1175–1188, 2007.

[36] S. G. Psakhie *et al.*, “Movable cellular automata method for simulating materials with mesostructure,” *Theor. Appl. Fract. Mech.*, vol. 37, no. 1–3, pp. 311–334, 2001.

[37] A. I. Dmitriev, M. Schargott, and V. L. Popov, “Direct modelling of surface topography development in a micro-contact with the movable cellular automata method,” *Wear*, vol. 268, no. 7–8, pp. 877–885, 2010.

[38] A. Essosnam, “Wear and contact modelling of heterogeneous material using multi-scale method .” 2020.

[39] J. W. Ringsberg, M. Loo-Morrey, B. L. Josefson, A. Kapoor, and J. H. Beynon, “Prediction of fatigue crack initiation for rolling contact fatigue,” *Int. J. Fatigue*, vol. 22, no. 3, pp. 205–215, 2000, doi: https://doi.org/10.1016/S0142-1123(99)00125-5.

[40] U. D. Schwarz, “A generalized analytical model for the elastic deformation of an adhesive contact between a sphere and a flat surface,” *J. Colloid Interface Sci.*, vol. 261, no. 1, pp. 99–106, 2003, doi: https://doi.org/10.1016/S0021-9797(03)00049-3.

[41] A. F. Bower and K. L. Johnson, “The influence of strain hardening on cumulative plastic deformation in rolling and sliding line contact,” *J. Mech. Phys. Solids*, vol. 37, no. 4, pp. 471–493, 1989, doi: https://doi.org/10.1016/0022-5096(89)90025-2.

[42] D. Majcherczak, P. Dufrenoy, and Y. Berthier, “Tribological, thermal and mechanical coupling aspects of the dry sliding contact,” *Tribol. Int.*, vol. 40, no. 5, pp. 834–843, 2007.

[43] “Surface temperatures in sliding contact : M. J. Furey, ASLE Trans., 7 (2) (1964) 133; 11 figs., 7 tables, 16 refs,” *Wear*, vol. 8, no. 2, p. 143, 1965, doi: 10.1016/0043-1648(65)90262-0.

[44] J. Bos and H. Moes, “Frictional Heating of Tribological Contacts,” *J. Tribol.*, vol. 117, no. 1, pp. 171–177, Jan. 1995, doi: 10.1115/1.2830596.

[45] J. Wu, B. Ma, H. Li, and L. Wang, “The Temperature Field of Friction Disc in Wet Clutch Involving the Unconventional Heat Dissipation on the Contact Surface,” *Tribol. Trans.*, vol. 64, no. 1, pp. 1–9, Jan. 2021, doi: 10.1080/10402004.2020.1751366.

[46] L. Karthikeyan, V. S. Senthilkumar, V. Balasubramanian, and S. Natarajan, “Mechanical property and microstructural changes during friction stir processing of cast aluminum 2285 alloy,” *Mater. Des.*, vol. 30, no. 6, pp. 2237–2242, 2009, doi: https://doi.org/10.1016/j.matdes.2008.09.006.

[47] M. Bashir, A. Qayoum, and S. Saleem, “Analysis of frictional heating and thermal expansion in a disc brake using COMSOL,” in *Journal of Physics: Conference Series*,

2019, vol. 1240, no. 1, p. 12094.

[48] V.-V. Lai, I. Paszkiewicz, J.-F. Brunel, and P. Dufrénoy, “Squeal occurrence related to the tracking of the bearing surfaces on a pin-on-disc system,” *Mech. Syst. Signal Process.*, vol. 165, p. 108364, 2022, doi: https://doi.org/10.1016/j.ymssp.2021.108364.

[49] S. A. Albatlan, “Study effect of pads shapes on temperature distribution for disc brake contact surface,” *Int. J. Eng. Res. Dev.*, vol. 8, no. 9, pp. 62–67, 2013.

[50] H. S. Qi and A. J. Day, “Investigation of disc/pad interface temperatures in friction braking,” *Wear*, vol. 262, no. 5–6, pp. 505–513, 2007.

[51] C. Cravero and D. Marsano, “Flow and thermal analysis of a racing car braking system,” *Energies*, vol. 15, no. 8, p. 2934, 2022.

[52] P. Wang, H. Shen, R. Dou, Q. Ye, M. Yu, and Q. Xu, “Research on Brake Thermal Simulation Technology Based on Heat Flux Theory,” in *2022 8th International Conference on Mechatronics and Robotics Engineering (ICMRE)*, 2022, pp. 6–11.

[53] A. Akay, “Acoustics of friction,” *J. Acoust. Soc. Am.*, vol. 111, no. 4, pp. 1525–1548, 2002.

[54] N. M. Kinkaid, O. M. O’Reilly, and P. Papadopoulos, “Automotive disc brake squeal,” *J. Sound Vib.*, vol. 267, no. 1, pp. 105–166, 2003.

[55] M. Eriksson, F. Bergman, and S. Jacobson, “Surface characterisation of brake pads after running under silent and squealing conditions,” *Wear*, vol. 232, no. 2, pp. 163–167, 1999.

[56] G. P. Ostermeyer, “Friction and wear of brake systems,” *Forsch. im Ingenieurwes.*, vol. 66, no. 6, pp. 267–272, 2001, doi: 10.1007/s100100100063.

[57] T. Grigoratos and G. Martini, “Brake wear particle emissions: a review,” *Environ. Sci. Pollut. Res.*, vol. 22, no. 4, pp. 2491–2504, 2015, doi: 10.1007/s11356-014-3696-8.

[58] D. H. Buckley, *Surface effects in adhesion, friction, wear, and lubrication*, vol. 5. Elsevier, 1981.

[59] A. Devaraju, “A critical review on different types of wear of materials,” *Int. J. Mech. Eng. Technol*, vol. 6, no. 11, pp. 77–83, 2015.

[60] M. M. Khruschov, “Principles of abrasive wear,” *Wear*, vol. 28, no. 1, pp. 69–88, 1974.

[61] K. V Pool, C. K. H. Dharan, and I. Finnie, “Erosive wear of composite materials,” *Wear*, vol. 107, no. 1, pp. 1–12, 1986.

[62] Y. Berthier, L. Vincent, and M. Godet, “Fretting fatigue and fretting wear,” *Tribol. Int.*, vol. 22, no. 4, pp. 235–242, 1989.

[63] S. Maya-Johnson, J. F. Santa, and A. Toro, “Dry and lubricated wear of rail steel under rolling contact fatigue-Wear mechanisms and crack growth,” *Wear*, vol. 380, pp. 240–250, 2017.

[64] W. J. Schumacher, “Corrosive wear principles,” *Mater. Performance;(United States)*, vol. 32, no. 7, 1993.

[65] H. Zhang, R. Goltsberg, and I. Etsion, “Modeling Adhesive Wear in Asperity and Rough Surface Contacts: A Review,” *Materials (Basel).*, vol. 15, no. 19, p. 6855, 2022.

[66] H. Zhang and I. Etsion, “Evolution of adhesive wear and friction in elastic-plastic spherical contact,” *Wear*, vol. 478, p. 203915, 2021.

[67] J. Zuo, X. Wang, S. Zhou, and F. Yang, “Simulation and experimental study on abrasive wear of brake discs,” *Tribol. Trans.*, pp. 1–11, 2022.

[68] N. Axen and S. Jacobson, “A model for the abrasive wear resistance of multiphase materials,” *Wear*, vol. 174, no. 1–2, pp. 187–199, 1994.

[69] C. Li, G. Bin, J. Li, and Z. Liu, “Study on the erosive wear of the gas-solid flow of compressor blade in an aero-turboshaft engine based on the Finnie model,” *Tribol. Int.*, vol. 163, p. 107197, 2021.

[70] M. S. ElTobgy, E. Ng, and M. A. Elbestawi, “Finite element modeling of erosive wear,” *Int. J. Mach. Tools Manuf.*, vol. 45, no. 11, pp. 1337–1346, 2005.

[71] K. Pereira, T. Yue, and M. A. Wahab, “Multiscale analysis of the effect of roughness on fretting wear,” *Tribol. Int.*, vol. 110, pp. 222–231, 2017.

[72] T. Yue and M. A. Wahab, “Finite element analysis of fretting wear under variable coefficient of friction and different contact regimes,” *Tribol. Int.*, vol. 107, pp. 274–282, 2017.

[73] J. Tunna, J. Sinclair, and J. Perez, “A review of wheel wear and rolling contact fatigue,” *Proc. Inst. Mech. Eng. Part F J. Rail Rapid Transit*, vol. 221, no. 2, pp. 271–289, 2007.

[74] I. G. Goryacheva and E. V Torskaya, “Modeling of fatigue wear of a two-layered elastic half-space in contact with periodic system of indenters,” *Wear*, vol. 268, no. 11–12, pp. 1417–1422, 2010.

[75] I. Garcıa, D. Drees, and J.-P. Celis, “Corrosion-wear of passivating materials in sliding contacts based on a concept of active wear track area,” *Wear*, vol. 249, no. 5–6, pp. 452–460, 2001.

[76] M. Stender, M. Tiedemann, D. Spieler, D. Schoepflin, N. Hoffmann, and S. Oberst,

"Deep learning for brake squeal: Brake noise detection, characterization and prediction," *Mech. Syst. Signal Process.*, vol. 149, p. 107181, 2021, doi: 10.1016/j.ymssp.2020.107181.

[77] J. Brownlee, "Deep Learning for Computer Vision Image Classification , Object Detection , and Face Recognition in Python UNLOCK Computer Vision With Deep Learning," *Deep Learn. Comput. Vis.*, pp. 1–21, 2019, [Online]. Available: https://machinelearningmastery.com/deep-learning-for-computer-vision/.

[78] N. Grochevaia, "Prediction of Brake Squeal: A Deep Learning Approach Analysis byMeans of Recurrent Neural Networks," 2020.

[79] M. Gorzinmataee, "Implementation of Machine Learning Techniques in Brake NVH Using Computational and Experimental Data." 2020.

[80] A. Bouchot, A. Ferrieux, J. Debayle, G. Mollon, and S. Descartes, "Image processing applied to tribological dry contact analysis," *Wear*, vol. 476, p. 203748, 2021, doi: https://doi.org/10.1016/j.wear.2021.203748.

[81] T. M. Atanackovic and A. Guran, *Theory of elasticity for scientists and engineers*. Springer Science & Business Media, 2000.

[82] M. Bonnet and A. Constantinescu, "Inverse problems in elasticity," *Inverse Probl.*, vol. 21, no. 2, 2005, doi: 10.1088/0266-5611/21/2/R01.

[83] A. Constantinescu and N. Tardieu, "On the identification of elastoviscoplastic constitutive laws from indentation tests," *Inverse Probl. Eng.*, vol. 9, no. 1, pp. 19–44, 2001, doi: 10.1080/174159701088027751.

[84] R. Forestier, Y. Chastel, and E. Massoni, "3D inverse analysis model using semi-analytical differentiation for mechanical parameter estimation," *Inverse Probl. Eng.*, vol. 11, no. 3, pp. 255–271, 2003, doi: 10.1080/1068276031000135908.

[85] A. Gavrus, E. Massoni, and J. L. Chenot, "An inverse analysis using a finite element model for identification of rheological parameters," *J. Mater. Process. Technol.*, vol. 60, no. 1–4, pp. 447–454, 1996, doi: 10.1016/0924-0136(96)02369-2.

[86] R. Mahnken and E. Stein, "Parameter identification for viscoplastic models based on analytical derivatives of a least-squares functional and stability investigations," *Int. J. Plast.*, vol. 12, no. 4, pp. 451–479, 1996, doi: 10.1016/S0749-6419(95)00016-X.

[87] N. Tardieu and A. Constantinescu, "On the determination of elastic coefficients from indentation experiments," *Inverse Probl.*, vol. 16, no. 3, pp. 577–588, 2000, doi: 10.1088/0266-5611/16/3/303.

[88] N. D. BERGER A., WITZ J.F., EL BARTALI A., LIMODIN N., DUBAR M., "Pixel-

wise full-field strain measurements for analysis of strain heterogeneities with regards to the material microstructure," *Photomech. IDICS*, 2020.

[89] S. M. A., *Image correlation for shape motion and deformation measurements : basic concepts, theory and applications / Michael A. Sutton, Jean-José Orteu, Hubert Schreier*. Boston: Springer, 2009.

[90] A. Guery, F. Hild, F. Latourte, and S. Roux, "Identification of crystal plasticity parameters using DIC measurements and weighted FEMU," *Mech. Mater.*, vol. 100, pp. 55–71, 2016, doi: 10.1016/j.mechmat.2016.06.007.

[91] K. T. Kavanagh and R. W. Clough, "Finite element applications in the characterization of elastic solids," *Int. J. Solids Struct.*, vol. 7, no. 1, pp. 11–23, 1971.

[92] J. L. Ericksen, "Large elastic deformations and non-linear continuum mechanics: A. E. Green and J. E. Adkins: Clarendon Press, Oxford, 1960. xiii + 348 pp., 55s," *J. Mech. Phys. Solids*, vol. 9, no. 3, pp. 213–214, 1961.

[93] L. R. G. Treloar, "Strains in an inflated rubber sheet, and the mechanism of bursting," *Rubber Chem. Technol.*, vol. 17, no. 4, pp. 957–967, 1944.

[94] R. S. Rivlin and D. W. Saunders, "Large Elastic Deformations of Isotropic Materials. VII. Experiments on the Deformation of Rubber," *Philos. Trans. R. Soc. Lond. A.*, vol. 243, no. 865, pp. 251–288, 1951.

[95] R. W. Ogden, "Large Deformation Isotropic Elasticity: On the Correlation of Theory and Experiment for Compressible Rubberlike Solids," *Proc. R. Soc. Lond. A. Math. Phys. Sci.*, vol. 328, no. 1575, pp. 567–583, 1972.

[96] G. A. Maugin, "Mechanics of solid materials: J. LEMAÎTRE and J.-L. CHABOCHE, Cambridge University Press, U.K. (Originally published as Mécanique des Matériaux Solides. Dunod, Paris, 1985), 1990; ISBN 0-521-32853-5. English translation from the French by B. SHRIVASTAVA," *Int. J. Eng. Sci.*, vol. 29, no. 8, pp. 1017–1018, 1991.

[97] E. Verron, "Modèles hyperélastiques pour le comportement mécanique des élastomères," *Tech. l'ingénieur*, vol. AM8210 V1, no. 0, p. 28, 2018.

[98] L. Meunier, G. Chagnon, D. Favier, L. Orgéas, and P. Vacher, "Mechanical experimental characterisation and numerical modelling of an unfilled silicone rubber," *Polym. Test.*, vol. 27, no. 6, pp. 765–777, 2008, doi: 10.1016/j.polymertesting.2008.05.011.

[99] T. Kirchdoerfer and M. Ortiz, "Data-driven computational mechanics," *Comput. Methods Appl. Mech. Eng.*, vol. 304, pp. 81–101, 2016, doi: 10.1016/j.cma.2016.02.001.

[100] T. Kirchdoerfer and M. Ortiz, “Data Driven Computing with noisy material data sets,” *Comput. Methods Appl. Mech. Eng.*, vol. 326, pp. 622–641, 2017, doi: 10.1016/j.cma.2017.07.039.

[101] T. Kirchdoerfer and M. Ortiz, “Data-driven computing in dynamics,” *Int. J. Numer. Methods Eng.*, vol. 113, no. 11, pp. 1697–1710, 2018, doi: 10.1002/nme.5716.

[102] R. Ibáñez, E. Abisset-Chavanne, E. Cueto, and F. Chinesta, “Data-driven in computational plasticity,” in *AIP conference proceedings*, 2018, vol. 1960, no. 1.

[103] M. Dalémat, “Une expérimentation réussie pour l’identification de la réponse mécanique sans loi de comportement : Approche data-driven appliquée aux membranes élastomères,” p. 212, 2019, [Online]. Available: https://tel.archives-ouvertes.fr/tel-02506891.

[104] A. Leygue, M. Coret, J. Réthoré, L. Stainier, and E. Verron, “Data-based derivation of material response,” *Comput. Methods Appl. Mech. Eng.*, vol. 331, pp. 184–196, 2018, doi: 10.1016/j.cma.2017.11.013.

[105] H. Yang, X. Guo, S. Tang, and W. K. Liu, “Derivation of heterogeneous material laws via data-driven principal component expansions,” *Comput. Mech.*, vol. 64, no. 2, pp. 365–379, 2019, doi: 10.1007/s00466-019-01728-w.

[106] N. Massara, E. Boccaleri, M. Milanesio, and M. Lopresti, “IETeasy: An open source and low-cost instrument for impulse excitation technique, applied to materials classification by acoustical and mechanical properties assessment,” *HardwareX*, vol. 10, pp. e00231–e00231, 2021.

[107] J. McCarthy, M. L. Minsky, N. Rochester, and C. E. Shannon, “A proposal for the Dartmouth summer research project on artificial intelligence,” *AI Mag.*, vol. 27, no. 4, pp. 12–14, 2006.

[108] J. Brownlee, *Machine learning mastery with Python: understand your data, create accurate models, and work projects end-to-end. Machine Learning Mastery,*. 2016.

[109] J. Brownlee, *“Master Machine Learning Algorithms: Discover How They Work and Implement Them From Scratch, 2016.” URL https://books. google. ca/books.* .

[110] J. Brownlee, *Deep learning with Python: develop deep learning models on Theano and TensorFlow using Keras. Machine Learning Mastery, 2016.* .

[111] J. Brownlee, *Deep learning with Python: develop deep learning models on Theano and TensorFlow using Keras*. Machine Learning Mastery, 2016.

[112] S. Khanse, P. Bhandari, R. Singru, N. Runwal, and A. Dharane, “Comparative Study of Genetic Algorithm and Artificial Neural Network for Multi-class Classification

based on Type-2 Diabetes Treatment Recommendation model," in *2020 Sixth International Conference on Parallel, Distributed and Grid Computing (PDGC)*, 2020, pp. 538–543.

[113] S.-C. Wang, "Artificial neural network," in *Interdisciplinary computing in java programming*, Springer, 2003, pp. 81–100.

[114] W. S. Sarle, "Neural Networks and Statistical Models," *SAS USers Group International Conference*. p. 13, 1994.

[115] J.-H. Wang, J.-H. Jiang, and R.-Q. Yu, "Robust back propagation algorithm as a chemometric tool to prevent the overfitting to outliers," *Chemom. Intell. Lab. Syst.*, vol. 34, no. 1, pp. 109–115, 1996.

[116] H. Jabbar and R. Z. Khan, "Methods to avoid over-fitting and under-fitting in supervised machine learning (comparative study)," *Comput. Sci. Commun. Instrum. Devices*, vol. 70, 2015.

[117] G. Panchal, A. Ganatra, P. Shah, and D. Panchal, "Determination of over-learning and over-fitting problem in back propagation neural network," *Int. J. Soft Comput.*, vol. 2, no. 2, pp. 40–51, 2011.

[118] A. P. Piotrowski and J. J. Napiorkowski, "A comparison of methods to avoid overfitting in neural networks training in the case of catchment runoff modelling," *J. Hydrol.*, vol. 476, pp. 97–111, 2013.

[119] K. Y. Chan, C. K. Kwong, T. S. Dillon, and Y. C. Tsim, "Reducing overfitting in manufacturing process modeling using a backward elimination based genetic programming," *Appl. Soft Comput.*, vol. 11, no. 2, pp. 1648–1656, 2011.

[120] W. M. P. Van der Aalst, V. Rubin, H. M. W. Verbeek, B. F. van Dongen, E. Kindler, and C. W. Günther, "Process mining: a two-step approach to balance between underfitting and overfitting," *Softw. Syst. Model.*, vol. 9, no. 1, pp. 87–111, 2010.

[121] S. Albawi, T. A. Mohammed, and S. Al-Zawi, "Understanding of a convolutional neural network," in *2017 international conference on engineering and technology (ICET)*, 2017, pp. 1–6.

[122] J. Gu *et al.*, "Recent advances in convolutional neural networks," *Pattern Recognit.*, vol. 77, pp. 354–377, 2018.

[123] M. Raissi, P. Perdikaris, and G. E. Karniadakis, "Physics Informed Deep Learning (Part I): Data-driven Solutions of Nonlinear Partial Differential Equations," 2017.

[124] M. Raissi, P. Perdikaris, and G. E. Karniadakis, "Physics Informed Deep Learning (Part II): Data-driven Discovery of Nonlinear Partial Differential Equations," 2017.

[125] T. Chen and C. Guestrin, “Xgboost: A scalable tree boosting system,” in *Proceedings of the 22nd acm sigkdd international conference on knowledge discovery and data mining*, 2016, pp. 785–794.

[126] F. Massi, Y. Berthier, and L. Baillet, “Contact surface topography and system dynamics of brake squeal,” *Wear*, vol. 265, no. 11–12, pp. 1784–1792, 2008, doi: 10.1016/j.wear.2008.04.049.

[127] N. M. Kinkaid, O. M. O’Reilly, and P. Papadopoulos, “Automotive disc brake squeal,” *J. Sound Vib.*, vol. 267, no. 1, pp. 105–166, 2003, doi: 10.1016/S0022-460X(02)01573-0.

[128] F. Massi, L. Baillet, O. Giannini, and A. Sestieri, “Brake squeal: Linear and nonlinear numerical approaches,” *Mech. Syst. Signal Process.*, vol. 21, no. 6, pp. 2374–2393, 2007, doi: 10.1016/j.ymssp.2006.12.008.

[129] F. Massi and O. Giannini, “Effect of damping on the propensity of squeal instability: an experimental investigation,” *J. Acoust. Soc. Am.*, pp. 2017–2023, 2008.

[130] M. R. North, *Disc brake squeal: a theoretical model*. MIRA Noneaton, 1972.

[131] A. Akay, J. Wickert, and Z. Xu, “Investigation of mode lock-in and friction interface,” *Final Report, Dep. Mech. Eng. Carnegie Mellon Univ.*, 2000.

[132] J. Flint and J. Hulten, “Lining-deformation-induced modal coupling as squeal generator in a distributed parameter disc brake model,” *J. Sound Vib.*, vol. 254, no. 1, pp. 1–21, 2002.

[133] O. Giannini, A. Akay, and F. Massi, “Experimental analysis of brake squeal noise on a laboratory brake setup,” *J. Sound Vib.*, vol. 292, no. 1–2, pp. 1–20, 2006, doi: 10.1016/j.jsv.2005.05.032.

[134] O. Giannini and A. Sestieri, “Predictive model of squeal noise occurring on a laboratory brake,” *J. Sound Vib.*, vol. 296, no. 3, pp. 583–601, 2006, doi: 10.1016/j.jsv.2006.02.022.

[135] H. Ouyang, W. Nack, Y. Yuan, and F. Chen, “Numerical analysis of automotive disc brake squeal: a review,” *Int. J. Veh. Noise Vib.*, vol. 1, no. 3–4, pp. 207–231, 2005.

[136] J. D. Fieldhouse, D. Bryant, and C. J. Talbot, “The Influence of Pad Abutment on the Generation of Brake Noise.,” *Int. J. Veh. Struct. Syst.*, vol. 3, no. 1, 2011.

[137] U. von Wagner, H. Gödecker, and S. Schlagner, “Brake squeal-modeling and experimental investigation using a work criterion,” *Int. J. Veh. Struct. Syst.*, vol. 3, no. 1, pp. 21–27, 2011, doi: 10.4273/ijvss.3.1.03.

[138] T. Butlin and J. Woodhouse, “Friction-induced vibration: Quantifying sensitivity and

uncertainty," *J. Sound Vib.*, vol. 329, no. 5, pp. 509–526, Mar. 2010, doi: 10.1016/j.jsv.2009.09.026.

[139] A. Nobari, H. Ouyang, and P. Bannister, "Uncertainty quantification of squeal instability via surrogate modelling," *Mech. Syst. Signal Process.*, vol. 60, pp. 887–908, 2015, doi: 10.1016/j.ymssp.2015.01.022.

[140] S. Oberst and J. C. S. Lai, "Statistical analysis of brake squeal noise," *J. Sound Vib.*, vol. 330, no. 12, pp. 2978–2994, 2011, doi: 10.1016/j.jsv.2010.12.021.

[141] A. Hanselowski and M. Hanss, "Analysis of epistemic uncertainty for the friction-induced vibration," *ZAMM - J. Appl. Math. Mech. / Zeitschrift für Angew. Math. und Mech.*, vol. 94, no. 11, pp. 933–944, 2014, doi: 10.1002/zamm.201300299.

[142] Z. Zhang, S. Oberst, and J. C. S. Lai, "On the potential of uncertainty analysis for prediction of brake squeal propensity," *J. Sound Vib.*, vol. 377, pp. 123–132, 2016, doi: 10.1016/j.jsv.2016.05.023.

[143] A. Renault, F. Massa, B. Lallemand, and T. Tison, "Experimental investigations for uncertainty quantification in brake squeal analysis," *J. Sound Vib.*, vol. 367, pp. 37–55, 2016, doi: 10.1016/j.jsv.2015.12.049.

[144] M. Stender, M. Tiedemann, D. Spieler, D. Schoepflin, N. Hoffmann, and S. Oberst, "Deep learning for brake squeal: Vibration detection, characterization and prediction," *arXiv*, 2020.

[145] V. Magnier, E. Roubin, J. B. Colliat, and P. Dufrénoy, "Methodology of porosity modeling for friction pad: Consequence on squeal," *Tribol. Int.*, vol. 109, no. December 2016, pp. 78–85, 2017, doi: 10.1016/j.triboint.2016.12.026.

[146] P. Dufrenoy, V. Magnier, Y. Waddad, J. F. Brunel, and G. De Saxce, "A Multiscale Model of a Disc Brake Including Material and Surface Heterogeneities," *SAE Int. J. Passeng. Cars - Mech. Syst.*, vol. 9, no. 3, pp. 1136–1143, 2016, doi: 10.4271/2016-01-1911.

[147] B. N. J. Persson, O. Albohr, U. Tartaglino, A. I. Volokitin, and E. Tosatti, "On the nature of surface roughness with application to contact mechanics, sealing, rubber friction and adhesion," *J. Phys. Condens. Matter*, vol. 17, no. 1, 2005, doi: 10.1088/0953-8984/17/1/R01.

[148] D. N. Ramasami, "Thèse pour l ' obtention du grade de Docteur de l ' Université Lille 1 Mention Mécanique présentée par Influence of friction material & test sequence on disc brake squeal," 2014.

[149] D. Martin, "Etude multi-échelle du crissement : dispositif expérimental et éléments de

compréhension / Martin Duboc ; sous la direction de Philippe Dufrénoy et de Jean-François Brunel et de Vincent Magnier," *Etude multi-échelle du crissement : dispositif expérimental et éléments de compréhension / Martin Duboc ; sous la direction de Philippe Dufrénoy et de Jean-François Brunel et de Vincent Magnier*. 2013.

[150] K. Bonnay, V. Magnier, J.-F. Brunel, P. Dufrenoy, and G. de Saxcé, "Influence of material and surface heterogeneities on mode lock-in in contact problem," Jun. 2013, [Online]. Available: https://hal.archives-ouvertes.fr/hal-01768675.

[151] E. Arfa, V. Magnier, P. Dufrénoy, and G. De Saxcé, "Contact modelling of highly heterogeneous friction material for braking applications," *Mech. Ind.*, vol. 21, no. 5, 2020, doi: 10.1051/meca/2020031.

[152] "http://hpc.univ-lille.fr/cluster-hpc-htc/calculer."

[153] Y. Waddad, V. Magnier, P. Dufrénoy, and G. De Saxcé, "Multiscale thermomechanical modeling of frictional contact problems considering wear – Application to a pin-on-disc system," *Wear*, vol. 426–427, no. September 2018, pp. 1399–1409, 2019, doi: 10.1016/j.wear.2018.12.063.

[154] Q. De Smedt, H. Wannous, and J. P. Vandeborre, "Heterogeneous hand gesture recognition using 3D dynamic skeletal data," *Comput. Vis. Image Underst.*, vol. 181, no. September 2018, pp. 60–72, 2019, doi: 10.1016/j.cviu.2019.01.008.

[155] D. J. Hand and R. J. Till, "A simple generalisation of the area under the ROC curve for multiple class classification problems," *Mach. Learn.*, vol. 45, no. 2, pp. 171–186, 2001.

[156] D. J. Hand, "Measuring classifier performance: a coherent alternative to the area under the ROC curve," *Mach. Learn.*, vol. 77, no. 1, pp. 103–123, 2009.

[157] M. B. Kursa and W. R. Rudnicki, "Feature selection with the Boruta package," *J. Stat. Softw.*, vol. 36, pp. 1–13, 2010.

[158] Y. Waddad, V. Magnier, P. Dufrenoy, and G. de Saxcé, "A new contact model for multi-layered solids with rough surfaces," *Tribol. Lett.*, vol. 65, no. 4, 2017, doi: 10.1007/s11249-017-0941-6.

[159] S. Hyun, L. Pei, J. F. Molinari, and M. O. Robbins, "Finite-element analysis of contact between elastic self-affine surfaces," *Phys. Rev. E - Stat. Physics, Plasmas, Fluids, Relat. Interdiscip. Top.*, vol. 70, no. 2, p. 12, 2004, doi: 10.1103/PhysRevE.70.026117.

[160] H. Hetzler and K. Willner, "On the influence of contact tribology on brake squeal," *Tribol. Int.*, vol. 46, no. 1, pp. 237–246, 2012, doi: https://doi.org/10.1016/j.triboint.2011.05.019.

[161] F. Bergman, M. Eriksson, and S. Jacobson, “Influence of disc topography on generation of brake squeal,” *Wear*, vol. 225, pp. 621–628, 1999, doi: 10.1016/S0043-1648(99)00064-2.

[162] H. A. Sherif, “Investigation on effect of surface topography of pad/disc assembly on squeal generation,” *Wear*, vol. 257, no. 7, pp. 687–695, 2004, doi: 10.1016/j.wear.2004.03.015.

[163] J.-J. Sinou, D. Lenoir, S. Besset, and F. Gillot, “Squeal analysis based on the laboratory experimental bench ‘Friction-Induced Vibration and noisE at École Centrale de Lyon’ (FIVE@ECL),” *Mech. Syst. Signal Process.*, vol. 119, pp. 561–588, 2019, doi: 10.1016/j.ymssp.2018.07.006.

[164] M. Graf and G.-P. Ostermeyer, “Friction-induced vibration and dynamic friction laws: Instability at positive friction–velocity-characteristic,” *Tribol. Int.*, vol. 92, pp. 255–258, 2015, doi: 10.1016/j.triboint.2015.06.019.

[165] M. Kchaou, A. R. Mat Lazim, M. K. Abdul Hamid, and A. R. Abu Bakar, “Experimental studies of friction-induced brake squeal: Influence of environmental sand particles in the interface brake pad-disc,” *Tribol. Int.*, vol. 110, pp. 307–317, 2017, doi: 10.1016/j.triboint.2017.02.032.

[166] Z. Y. Xiang, J. L. Mo, H. Ouyang, F. Massi, B. Tang, and Z. R. Zhou, “Contact behaviour and vibrational response of a high-speed train brake friction block,” *Tribol. Int.*, vol. 152, p. 106540, 2020, doi: 10.1016/j.triboint.2020.106540.

[167] A. Heussaff, L. Dubar, T. Tison, M. Watremez, and R. F. Nunes, “A methodology for the modelling of the variability of brake lining surfaces,” *Wear*, vol. 289, pp. 145–159, 2012, doi: 10.1016/j.wear.2012.04.002.

[168] Y. Xiao *et al.*, “Evolution of contact surface characteristics and tribological properties of a copper-based sintered material during high-energy braking,” *Wear*, vol. 488–489, p. 204163, 2022, doi: 10.1016/j.wear.2021.204163.

[169] V. V. Lai, I. Paszkiewicz, J. F. Brunel, and P. Dufrénoy, “Multi-scale contact localization and dynamic instability related to brake squeal,” *Lubricants*, vol. 8, no. 4, 2020, doi: 10.3390/LUBRICANTS8040043.

[170] D. Majcherczak, P. Dufrenoy, and Y. Berthier, “Tribological, thermal and mechanical coupling aspects of the dry sliding contact,” *Tribol. Int.*, vol. 40, no. 5, pp. 834–843, 2007, doi: 10.1016/j.triboint.2006.08.004.

[171] D. Majcherczak, P. Dufre´noy, and M. Nai¨t-Abdelaziz, “Third Body Influence on Thermal Friction Contact Problems: Application to Braking,” *J. Tribol.*, vol. 127, no.

1, pp. 89–95, 2005, doi: 10.1115/1.1757490.

[172] N. Singla, J.-F. Brunel, A. Mège-Revil, H. Kasem, and Y. Desplanques, “Experiment to Investigate the Relationship Between the Third-Body Layer and the Occurrence of Squeals in Dry Sliding Contact,” *Tribol. Lett.*, vol. 68, no. 1, pp. 1–11, 2019, doi: 10.1007/s11249-019-1244-x.

[173] S. Oberst, J. C. S. Lai, and S. Marburg, “Guidelines for numerical vibration and acoustic analysis of disc brake squeal using simple models of brake systems,” *J. Sound Vib.*, vol. 332, no. 9, pp. 2284–2299, 2013, doi: 10.1016/j.jsv.2012.11.034.

[174] R. Mann, V. Magnier, J. F. Brunel, F. Brunel, P. Dufrénoy, and M. Henrion, “Relation between mechanical behavior and microstructure of a sintered material for braking application,” *Wear*, vol. 386–387, no. May, pp. 1–16, 2017, doi: 10.1016/j.wear.2017.05.013.

[175] M. Duboc, V. Magnier, J.-F. Brunel, and P. Dufrénoy, “Experimental set-up and the associated model for squeal analysis,” *Mech. Ind. an Int. J. Mech. Sci. Eng. Appl.*, vol. 21, no. 2, p. 204, 2020, doi: 10.1051/meca/2019083.

[176] D. E. Rumelhart, G. E. Hinton, and R. J. Williams, “Learning representations by back-propagating errors,” *Nature*, vol. 323, no. 6088, pp. 533–536, 1986.

[177] I. Goodfellow, Y. Bengio, and A. Courville, “Deep Learning: Machine Learning Book,” p. 785, 2016, [Online]. Available: http://www.deeplearningbook.org/.

[178] B. Jason, “Predict the Future with MLPs, CNNs and LSTMs in Python,” *Ml*, vol. 1, no. 1, pp. 1–50, 2018.

[179] L. Mao, M. Cai, and G. Wang, “Effect of rotation speed on the abrasive–erosive–corrosive wear of steel pipes against steel casings used in drilling for petroleum,” *Wear*, vol. 410–411, pp. 1–10, 2018, doi: 10.1016/j.wear.2018.06.002.

[180] S. Toumi, S. Fouvry, and M. Salvia, “Prediction of sliding speed and normal force effects on friction and wear rate evolution in a dry oscillating-fretting PTFE/Ti-6Al-4V contact,” *Wear*, vol. 376–377, no. PB, pp. 1365–1378, 2017, doi: 10.1016/j.wear.2017.02.021.

[181] G. Li, S. G. Qu, Y. X. Pan, and X. Q. Li, “Effects of the different frequencies and loads of ultrasonic surface rolling on surface mechanical properties and fretting wear resistance of HIP Ti–6Al–4V alloy,” *Appl. Surf. Sci.*, vol. 389, pp. 324–334, 2016, doi: 10.1016/j.apsusc.2016.07.120.

[182] Y. Waddad, V. Magnier, P. Dufrénoy, and G. De Saxcé, “Heat partition and surface temperature in sliding contact systems of rough surfaces,” *Int. J. Heat Mass Transf.*,

vol. 137, pp. 1167–1182, 2019, doi: 10.1016/j.ijheatmasstransfer.2019.04.015.

[183] J. F. Archard, “Contact and Rubbing of Flat Surfaces,” *J. Appl. Phys.*, vol. 24, no. 8, pp. 981–988, 1953, doi: 10.1063/1.1721448.

[184] R. K. Paretkar, J. P. Modak, and A. V Ramarao, “An approximate generalized experimental model for dry sliding adhesive wear of some single-phase copper-base alloys,” *Wear*, vol. 197, no. 1–2, pp. 17–37, 1996.

[185] X. Yin and K. Komvopoulos, “An adhesive wear model of fractal surfaces in normal contact,” *International Journal of Solids and Structures*, vol. 47, no. 7–8. pp. 912–921, 2010, doi: 10.1016/j.ijsolstr.2009.12.003.

[186] L. J. Yang, “An integrated transient and steady-state adhesive wear model,” *Tribol. Trans.*, vol. 46, no. 3, pp. 369–375, 2003.

[187] R. Aghababaei, “Effect of adhesion on material removal during adhesive wear,” *Physical Review Materials*, vol. 3, no. 6. 2019, doi: 10.1103/PhysRevMaterials.3.063604.

[188] D. Forsstrom and P. Jonsen, “Calibration and validation of a large scale abrasive wear model by coupling DEM-FEM Local failure prediction from abrasive wear of tipper bodies during unloading of granular material,” *Eng. Fail. Anal.*, vol. 66, pp. 274–283, 2016, doi: 10.1016/j.engfailanal.2016.04.007.

[189] A. ASTM, “Standard terminology relating to wear and erosion,” *G40*, 2013.

[190] G. Bolelli *et al.*, “Sliding and abrasive wear behaviour of HVOF- and HVAF-sprayed Cr3C2–NiCr hardmetal coatings,” *Wear*, vol. 358–359, pp. 32–50, 2016, doi: 10.1016/j.wear.2016.03.034.

[191] J. Cheng, Z. Wang, X. Gan, Q. Lei, Z. Li, and K. Zhou, “Corrosion and corrosive-wear behaviors of a high strength and toughness Cu–15Ni–8Sn alloy in seawater,” *Mater. Corros.*, vol. 71, no. 4, pp. 593–607, 2020, doi: 10.1002/maco.201911323.

[192] C. Ouyang, S. Zhu, and D. Y. Li, “Corrosion and corrosive wear behavior of WC–MgO composites with and without grain-growth inhibitors,” *J. Alloys Compd.*, vol. 615, pp. 146–155, 2014, doi: 10.1016/j.jallcom.2014.06.137.

[193] Y. Gu, X. Zheng, Q. Liu, H. Ma, L. Zhang, and D. Yang, “Investigating Corrosion Performance and Corrosive Wear Behavior of Sol–gel/MAO-Coated Mg Alloy,” *Tribol. Lett.*, vol. 66, no. 3, pp. 1–15, 2018, doi: 10.1007/s11249-018-1052-8.

[194] R. Kumar, M. Antonov, P. Klimczyk, V. Mikli, and D. Gomon, “Effect of cBN content and additives on sliding and surface fatigue wear of spark plasma sintered Al2O3-cBN composites,” *Wear*, vol. 494–495, p. 204250, 2022, doi: 10.1016/j.wear.2022.204250.

[195] S. L. A., *Tribo-Fatigue : Wear-Fatigue Damage and its prediction*. Berlin, Heidelberg: Springer Berlin Heidelberg, 2005.

[196] A. Y. Wang, J. L. Mo, X. C. Wang, M. H. Zhu, and Z. R. Zhou, "Effect of surface roughness on friction-induced noise: Exploring the generation of squeal at sliding friction interface," *Wear*, vol. 402–403, pp. 80–90, 2018, doi: 10.1016/j.wear.2018.02.005.

[197] V.-A. Orang, *Friction-Induced Vibration in Lead Screw Drives* . New York, NY: Springer New York Imprint: Springer, 2011.

[198] J. Park *et al.*, "Analysis of wear induced particle emissions from brake pads during the worldwide harmonized light vehicles test procedure (WLTP)," *Wear*, vol. 466–467, p. 203539, 2021, doi: 10.1016/j.wear.2020.203539.

[199] S. Abbasi, A. Jansson, L. Olander, U. Olofsson, and U. Sellgren, "A pin-on-disc study of the rate of airborne wear particle emissions from railway braking materials," *Wear*, vol. 284–285, pp. 18–29, 2012, doi: 10.1016/j.wear.2012.01.016.

[200] M. Loïc, "Crissement des freins de giration de nacelles d'éoliennes : éléments de compréhension et influence du matériau de friction." s.n., S.l, 2012.

[201] Y. Soeta and R. Shimokura, "The Impact of External Environments and Wheel-Rail Friction on Noise inside a Train Car," *Noise Vib. Worldw.*, vol. 43, no. 8, pp. 9–22, 2012, doi: 10.1260/0957-4565.43.8.9.

[202] L. Mattei and F. Di Puccio, "Influence of the wear partition factor on wear evolution modelling of sliding surfaces," *International Journal of Mechanical Sciences*, vol. 99. pp. 72–88, 2015, doi: 10.1016/j.ijmecsci.2015.03.022.

[203] Y. Zhang, L. Lu, Y. Gong, J. Zhang, and D. Zeng, "Fretting wear-induced evolution of surface damage in press-fitted shaft," *Wear*, vol. 384–385. pp. 131–141, 2017, doi: 10.1016/j.wear.2017.05.014.

[204] S. S. Rajahram, T. J. Harvey, J. C. Walker, S. C. Wang, R. J. K. Wood, and G. Lalev, "A study on the evolution of surface and subsurface wear of UNS S31603 during erosion-corrosion," *Wear*, vol. 271, no. 9–10. pp. 1302–1313, 2011, doi: 10.1016/j.wear.2010.11.018.

[205] C. Paulin, S. Fouvry, and C. Meunier, "Finite element modelling of fretting wear surface evolution: Application to a Ti–6A1–4V contact," *Wear*, vol. 264, no. 1, pp. 26–36, 2008, doi: 10.1016/j.wear.2007.01.037.

[206] G. Mart, G. Terrazas, and S. Ratchev, "Tool wear classification using time series imaging and deep learning," pp. 3647–3662, 2019.

[207] P.-M. Huang and C.-H. Lee, “Estimation of Tool Wear and Surface Roughness Development Using Deep Learning and Sensors Fusion,” *Sensors (Basel).*, vol. 21, no. 16, p. 5338, 2021, doi: 10.3390/s21165338.

[208] I. Goodfellow *et al.*, “Generative adversarial nets,” *Adv. Neural Inf. Process. Syst.*, vol. 27, 2014.

[209] A. Radford, L. Metz, and S. Chintala, “Unsupervised representation learning with deep convolutional generative adversarial networks,” *arXiv Prepr. arXiv1511.06434*, 2015.

[210] Y. Yu, Z. Gong, P. Zhong, and J. Shan, “Unsupervised representation learning with deep convolutional neural network for remote sensing images,” in *International conference on image and graphics*, 2017, pp. 97–108.

[211] J. Brownlee, “Generative Adversarial Networks with Python Deep Learning Generative Models for Image Synthesis and Image Translation.”

[212] L. Perez and J. Wang, “The Effectiveness of Data Augmentation in Image Classification using Deep Learning,” 2017, [Online]. Available: http://arxiv.org/abs/1712.04621.

[213] N. Romero, M. Gutoski, L. Hattori, and H. S. Lopes, “The Effect of Data Augmentation on the Performance of Convolutional Neural Networks,” no. October, pp. 1–12, 2019, doi: 10.21528/cbic2017-51.

[214] K. Chatfield, K. Simonyan, A. Vedaldi, and A. Zisserman, “Return of the devil in the details: Delving deep into convolutional nets,” *arXiv Prepr. arXiv1405.3531*, 2014.