UNIONE EUROPEA
Fondo Sociale Europeo

Ministero dell'Università e della Ricerca

PON RICERCA E INNOVAZIONE

REACT EU

UNIVERSITÀ DEGLI STUDI DI TRIESTE

UNIVERSITÀ DEGLI STUDI DI TRIESTE

XXXVII CICLO DEL DOTTORATO DI RICERCA IN

Earth Sciences, Fluid-dynamics and Mathematics.
Interaction and Methods

Novel Geothermal Energy Exploration Methods Through AI-assisted Fracture Analysis and Closed-Loop Heat Exchangers

Settore scientifico-disciplinare: **GEO11**

DOTTORANDO / A
Attilio Molossi

COORDINATORE
PROF. Stefano Maset

SUPERVISORE DI TESI
PROF. Michele Pipan

ANNO ACCADEMICO 2023/2024

Contents

List of Figures

List of Tables

Chapter 1

Overview

1.1 Motivation

The increase in energy demand and the zero emissions goals, leads our society to the need of tackling some of the main challenges for the development of low carbon technologies. Among these, there is geothermal energy, as a very promising clean and renewable source of energy (Lukawski et al., 2014). Despite this potential, geothermal energy showed a slow growth in the past years, given the shortage of favorable policies and the high amount of capital to be invested among other reasons. The recent (predicted) oil crisis, however, setted the conditions to elicit academic and industrial projects to study new solutions, that overcome the major limitations of the geothermal application, in particular for electricity production (Beckers et al., 2014; Reber et al., 2014). Those limitations include uncertainties of all kind and environmental impact (microseismic events, land use), which contributes to the hostility of communities and undermines the interest of policy makers and stakeholders (Manzella et al., 2019). Among the major sources of uncertainty for the exploration and development of geothermal projects there is the inherent uncertainty in subsurface characterization (Hoteit et al., 2023; Gola et al., 2022; Lukawski et al., 2016). This task is largely in the hand of the geoscientists. This underpins the important role of geoscience in this societal transformation challenges we are facing in our time.
To characterize the subsurface is not an easy task and the geoscientists' practice is inherently affected by aleatoric an epistemic uncertainties, like any other scientific field. However, the epistemic uncertainty in geoscience is more remarkable than in other fields, and the role of subjective biases during information extraction from geodata appears evident at various degrees across pretty much all of the geoscientific disciplines (Lacchia et al., 2020; Andrews et al., 2019; Terzaghi, 1965; Gibson et al., 2016).The practices of geoscientists when working with data typically involve performing visual cognitive tasks. These tasks are closely tied to human perception and are enriched by the geoscientist's expertise and specialized knowledge, such as understanding geological history, structural and sedimentary models, domain-specific principles, and leveraging associative memory (Dell'Aversana, 2013). Perception, memory associations, and brain representations of data are fundamental components of human intelligence (Friston

et al., 2012; Euler, 2018; Todd et al., 2012), and they are central to geoscience practice (Shipley et al., 2013; Andrews et al., 2019; Wilson et al., 2020; Terzaghi, 1965; Boersma et al., 2021). However, individual human characteristics—such as mental styles, personal experiences, and cognitive biases—introduce significant variability in outcomes, even for identical tasks on identical data. The geoscientist community thus has a moral imperative to minimize this uncertainty in order to enhance the rigor of scientific conclusions, which have an often underappreciated impact on society (Daniilidis et al., 2017; Bond et al., 2007; Wilson et al., 2019; Terzaghi, 1965; Tannert et al., 2007; Pollack, 2007). For example, in geophysics the seek of objectiveness can be tracked back to the very first steps of advancements in the field. Initially, geophysicist gathered around the same seismic data and confronted their opinion to find the most reliable and robust conceptualization of the subsurface, based on horizon picking. One or two decades later, the rise of computational power gave birth to revolutionizing solutions like seismic attributes (Chopra and Marfurt, 2005) to increase the objectiveness of interpretations with the mathematical calculation of interesting parameters (e.g. pore pressure) from seismic features. With the evolution of our computational capability, nowadays geoscience disciplines are facing an explosion of artificial intelligence (AI) solutions, such as Machine Learning (ML) and Deep Learning (DL), to test new, innovative methods that can help in this seek of objectivity. To go back to the seismic section example, today in literature we can find several papers addressing the use of DL to automate the horizon position from both 2D and 3D seismic data (Roncoroni et al., 2022; Zhang and Zhu, 2021; Geng et al., 2020; Bi et al., 2021; Shi et al., 2020). While the potential of AI-based solutions to mitigate human cognitive biases is increasingly recognized, research on Intelligence Augmentation (IA)—the enhancement of human abilities through information technologies like AI—remains relatively limited (Zhou et al., 2021; Mohanty and Vyas, 2018). The IA paradigm is grounded in the notion that humans and machines can collaborate to leverage their respective strengths, achieving superior results compared to either operating alone. The synergy between AI and IA is intuitive; advances in AI systems naturally support IA's effectiveness. However, there is a fundamental distinction between the two concepts: traditional AI views machines as autonomous systems capable of fully automating tasks, emphasizing the goal of "making machines smarter." In contrast, IA focuses on "making people smarter" by leveraging AI as a technological extension to enhance human capabilities (Zhou et al., 2021; Kline, 1995). Fully autonomous AI systems, carry inherent risks, including the amplification of human biases and the potential for harmful outcomes (Manyika et al., 2019). This risk may be particularly pronounced in geoscientific applications, which generally demand a combination of creativity and objective thinking to uphold scientific rigor. IA seeks to harmonize biological and digital cognitive systems to enhance operational efficiency and generate value. Machines excel in logical, mathematical, and analytical tasks, while humans outperform in abstract reasoning and creative problem-solving (Cavas and Cavas, 2020). Following IA principles, machines are deployed to handle repetitive or routine tasks, enabling humans to focus on non-routine, creative, and abstract reasoning.

In this thesis, we will apply IA principles to develop AI-based solutions aimed at reducing human uncertainty during the analysis and interpretation of critical data for geothermal exploration. One of the most significant sources of uncertainty in subsurface characterization for geothermal exploration and development is the fracture network (Hoteit et al., 2023; Lukawski et al., 2014, 2016; Patterson et al., 2020; Peacock, 2025; Eriçok and Gümrah, 2010; Yang et al., 2020). This property, while challenging to assess, is crucial for geothermal applications, as fractures often serve as the primary natural conduits for fluid flow in the subsurface. These fluids are extracted (and re-injected) to harness the heat transferred from the hot rock to the fluid within the fracture spaces. As in all the other domains of geoscience, fracture characterization is based on the integration of any available data, and hydraulic tests at the sites of the drilled boreholes. Borehole data normally have higher resolutions than other geophysical surface data used in geothermal exploration (e.g. seismic, magnetotellurics-MT) and allow the recognition of finer structures and details of the fracture network. Borehole image logs, for example, enable operators with the visualization of any geological feature in the borehole at variable scales (from mm to cm) (Pöppelreiter et al., 2010b; Sun et al., 2021; Chatterjee et al., 2017; Bouchaala et al., 2019; Dasgupta et al., 2019). In some cases, these images can also support directional drilling, which is commonly employed during the drilling of horizontal wells. This approach has proven to be an effective solution for optimizing geothermal energy extraction (Wielemaker et al., 2020). However, subjectivity remains a significant factor even when analyzing high-resolution data, hindering the ability to confidently and robustly model fractures and their apertures. Consequently, this affects the reliability of long-term thermodynamic predictions and forecasts in specific geothermal scenarios, even when advanced numerical simulations are employed (Boersma et al., 2021; Maerten et al., 2019; Laubach et al., 2019). To overcome the limitations associated with uncertainty, industry leaders have, since 2019, demonstrated the feasibility of cutting-edge solutions with the potential to expand the applicability of geothermal electricity production to a wider range of geological conditions, such as deep borehole heat exchangers (DBHE) or deep closed-loop heat exchangers (DCHE) (Hu et al., 2015) or the (Yuan et al., 2021; Kelly and McDermott, 2022). By leveraging modern and advanced drilling technologies, a DCHE essentially functions as a subsurface radiator created through the drilling of multilateral wells. This represents a groundbreaking advancement in geothermal energy development, as this solution is inherently more scalable and carries fewer risks compared to conventional geothermal systems, which depend on reservoir fluid flow parameters and hydrogeological conditions. However, the extensive use of these advanced technologies is still not financially cost effective, so their feasibility at a global scale remains somehow confined and they are largely untested (Gola et al., 2022; Kelly and McDermott, 2022). The retrofitting of abandoned wells has emerged in recent years as a potential sustainable solution to mitigate the cost limitations of these promising systems (Gharibi et al., 2018; Kaplanoğlu et al., 2019; Sui et al., 2019).

1.2 Related research activity

The results presented in this thesis have been disseminated through scientific papers published in international journals and shared at conferences via oral presentations and posters. For the two main chapters of the thesis, the trackable works are listed below:

- Chapter 2
  - Attilio Molossi and Michele Pipan. Exploiting image logs to reduce drilling hazards: an innovative Artificial Intelligence methodology applied in East Africa. Geophysical Journal International, 235(1):942–950, May 2023. ISSN 0956- 540X, 1365-246X. doi: 10.1093/gji/ggad286. URL https://academic.oup. com/gji/article/235/1/942/7223463. (Molossi and Pipan, 2023)
  - A. Molossi. LWD Image Logs Interpretation (LILI) : An Automated Approach Using Deep Learning. volume 2023, pages 1–5. European Association of Geoscientists & Engineers, September 2023. ISBN 978-2-02-362902-4. doi: 10.3997/2214-4609.2023629020. (Molossi and Pipan, 2023)
  - Attilio Molossi, Giacomo Roncoroni, and Michele Pipan. Efficient Logging- While-Drilling Image Logs Interpretation Using Deep Learning. Petrophysics- The SPWLA Journal of Formation Evaluation and Reservoir Description, 65 (03):365–387, 2024. doi: 10.30632/PJV65N3-2024a5. (Molossi et al., 2024)
  - A. Molossi. Enhancing Geothermal Reservoirs Fracture Interpretation Using Deep Learning: a Case Study from the Western Netherland Basin. In Fifth EAGE Global Energy Transition Conference & Exhibition (GET 2024), pages 1–5, Rotterdam, Netherlands, 2024. European Association of Geoscientists & Engineers. doi: 10.3997/2214-4609.202421218. URL https://www.earthdoc.org/ content/papers/10.3997/2214-4609.202421218. (Molossi, 2024)
- Chapter 3
  - SpiderTherm: Optimizing Geothermal Extraction for Sustainable Energy Transition, Gruppo Nazzionale di Geofisica della Terra Solida (GNGTS 2024), Ferrara, Italy, February 2024
  - SpiderTherm: Optimizing Geothermal Extraction for Sustainable Energy Transition, European Geothermal Ph.D Days (EGPD 2024), Delft, Netherlands, April 2024.

Chapter 2

AI-based methods for borehole image logs analysis automation

2.1 Introduction

The seventh Sustainable Development Goal focuses on green energy production to fulfill the growing energy request. As a potentially ubiquitous, time-continuous, clean form of energy, geothermal energy represents a promising solution for our energetic transition (Manzella et al., 2019; Gola et al., 2022; Lukawski et al., 2014, 2016). Depending on many factors, including both static and dynamic properties of the geothermal reservoir, we can exploit the heat in the underground for power production (Birdsell and Saar, 2019; Juliusson and Horne, 2013; Zarrouk and Moon, 2014). However, the subsurface properties that have to be defined to assess the geothermal potential of any given geological system fail to be defined with a satisfiable degree of confidence and accuracy (Ansari et al., 2017; Yoon et al., 2014; Pandey and Vishal, 2017). The fracture network and their connectivity often form the primary fluid flow paths in the subsurface for the fluids that we are interested in to extract the heat, but the description of such fracture systems is often complex, hindered by uncertainty in measurements, and by human subjectivity (Lukawski et al., 2016; Andrews et al., 2019; Ericok and Gumrah, 2005).

Fracture analysis is, therefore, one of the most important tasks from the geothermal exploration perspective, as it allows to estimate those static properties of the reservoir, such as permeability, based on fractures parameters (Li et al., 2016; Salone, 2024). In geothermal exploration, fracture analysis is normally done using a multi-scale approach, integrating several data with different resolutions (e.g. seismic, magnetotellurics, borehole data), to get the most accurate models of the subsurface (Faiq Adenan et al., 2023; Hosseinzadeh et al., 2023). Despite the extremely high-resolution of some of those data however, the fracture detection from different operators still show high levels of uncertainty (Shipley et al., 2013; Bond et al., 2007; Shipton et al.). Scheiber et al. (2015) observed the low reproducibility of geoscienctists' interpretations, Andrews et al. (2019) showed that fracture analysis tasks using several data and methods for fracture collection are particularly affected by subjective biases, and this correlates with the hypothesis

of a relationship between the mental model and psychology of the interpreters and their results (Shipley et al., 2013; Shipton et al.; Wilson et al., 2020; Euler, 2018). To overcome at least some of the subjective factors related to perception or the attitude of the interpreter (e.g. detail-oriented or less detail-oriented), modern and powerful machine learning (ML) and deep learning (DL) methods are used in all fields of geology and geophysics to support the decisional steps of interpreters in both industry and academia (Roncoroni et al., 2022; Roncoroni, 2024; Alzubaidi et al., 2021; Shi et al., 2020; Gupta et al., 2019; Jin et al., 2019; Roncoroni et al., 2023).

In geothermal exploration, as well as in oil and gas, borehole images (BHI) are among the most commonly used data for performing fracture analysis (Ozkaya and Mattner, 2003; Ernando and Fathoni, 2011; Yang et al., 2020; Boersma et al., 2021; Kato and Sakagawa, 1995; Sulistyowati and Haris, 2021; Mazdarani et al., 2023; Luthi, 2001). These data allow for the derivation of various parameters related to the petrophysical properties of the drilled formations, which are valuable for reservoir characterization (Maerten et al., 2019; Basu et al., 2019). BHIs provide 2D representations of the borehole wall (Pöppelreiter et al., 2010a; Dasgupta et al., 2019; Sun et al., 2021; Chatterjee et al., 2017), capturing geological features with resolutions as fine as millimeters. Despite their high resolution, BHIs are not immune to uncertainty, particularly in fracture analysis tasks, as highlighted in various studies (Ozol et al., 2015; Boersma et al., 2021; Mondol, 2015; Yang et al., 2020). The uncertainty in these analyses significantly impacts the confidence in subsequent scientific and technical assessments related to geothermal reservoir characterization (Lukawski et al., 2016; Witter et al., 2019; Maeso et al., 2015; Pollack, 2007). It is important to note that these high-resolution BHIs are obtained after the drilling of a well, which enables the millimetric precision of the data (Lai et al., 2018).

There is a class of BHI data that are acquired during the drilling of boreholes that belong to the *logging-while-drilling* (LWD), which are characterized by a lower resolution, but can provide with useful insights for decision-making during geosteering of wells, and reduce the risk related, for example, to borehole stability issues (Wang et al., 2019; Dasgupta et al., 2019). These LWD BHI have been proved to be effective in geothermal projects, where they can assist the geosteering of subhorizontal wells (80-85°) in geothermal reservoir, as an intermediate solution for multilateral wells (often a non cost-efficient solution) and deviated or vertical (<35°) wells (Ungemach et al., 2018; Wielemaker et al., 2020). This solution demonstrated the applicability of LWD BHI in geothermal development, since the images (azimuthal density images) were used- in integration with other data- to derive layers of homogeneous poro-permeability properties during the drilling of a well and support decisions (Wielemaker et al., 2020). Unlike the extensive research available on high-resolution BHI, there is relatively little literature exploring the application of machine learning (ML) or deep learning (DL) to logging-while-drilling (LWD) BHI. Nevertheless, LWD BHI data have demonstrated potential for use in geosteering geothermal wells, offering promising opportunities for improving well placement and efficiency. In our work, we proposed two innovative AI-based methods for the real-time automated recognition of fea-

tures in low-resolution borehole images (BHI). The first is a method based on computer vision (CV) combined with a dynamic time warping (DTW) algorithm, referred to as CVML throughout the paper. The second is a deep learning (DL) approach. Both methods are integrated into a framework that supports real-time human interaction with the system during operation, enabling the validation and discarding of results as necessary. This framework employs a human-centric approach inspired by the Intelligence Augmentation (IA) paradigm, which focuses on enhancing human capabilities by harnessing the strengths of AI (Zhou et al., 2021; Skagestad, 1993).

The interaction model we propose seeks to blend machine and human cognitive styles, aiming to increase confidence in the final outcomes. Throughout this work, we delve into the details of machine-level performance and present hypotheses on the potential qualitative changes in human operators' thinking as a result of this collaborative interaction.

In the following sections, we will provide a brief introduction to some of the key features of BHIs, describe the proposed methods in detail, compare them to existing approaches in the literature, and highlight the unique contributions and novelties of our solutions.

2.1.1 Borehole images (BHI)

High-resolution BHIs provide directional structural, geological, and sedimentary information at the millimeter scale, making them critical for reservoir characterization (Ameen and Hailwood, 2008; Keeton et al., 2015; Khoshbakht et al., 2012). As noted earlier, they are 2D representations of petrophysical properties (electrical, acoustic, density) that allow detailed sedimentary and fracture analysis at this fine scale (Prioul et al., 2007; Kosari et al., 2015). Geological features such as bedding, fractures, faults, stylolites, vugs, and breakouts typically appear as sinusoids in BHIs. The classification of these sinusoids (Pattnaik et al., 2020; Prensky, 1999; Folkestad et al., 2012; Lai et al., 2018) represents a 2D approximation of the geological features and is generally done manually by a human interpreter.

However, this interpretation task is often considered a *tedious task* because it is unnecessarily repetitive, monotonous, and time-consuming (Lopes et al., 2018). This has spurred a growing interest in developing AI methods for a wide range of geophysical applications, particularly for BHI interpretation, as discussed further in section 2.1.2. Despite this interest, AI methods are commonly employed under the assumption that they can effectively replicate human behavior and perception. As a result, little attention has been given to evaluating these methods beyond benchmark scores or understanding the implications of human-machine interactions from a human-centric perspective (Liu et al., 2023).

In this thesis, we propose several solutions to improve DL-based methods by employing diverse training strategies aimed at enhancing their alignment with the cognitive abilities humans use for this task. Additionally, we explore the potential implications of these methods, envisioning scenarios of varying complexity where human-machine interaction could play a pivotal role.

One example of high-resolution BHIs is obtained using the Formation MicroImager (FMI), an electric conductivity borehole imaging tool. FMI provides highly detailed conductivity images of borehole walls (Fig. 2.1) and is extensively used for characterizing fractures and collecting essential data for subsequent simulations (Lepillier et al., 2020; Boersma et al., 2021; Rajabi et al., 2010; Jin et al., 2022). High-resolution images like these are acquired using the *wireline* method, where the tool is lowered into the open hole to make measurements after drilling has ceased.
In contrast, LWD BHIs have lower resolution because they are acquired in the dynamic and often disturbed environment of ongoing drilling (Poppelreiter, 2009; Wang et al., 2019). The tool is part of the rotating drilling assembly, and the signals required to generate 2D borehole images are transmitted to the surface via mud pulse telemetry, leveraging the drilling mud (Klotz et al., 2008; Berro and Reich, 2019; Mwachaka et al., 2019). The quality of these images is directly affected by conditions in the borehole, further complicating the interpretation process (Gong et al., 2021).

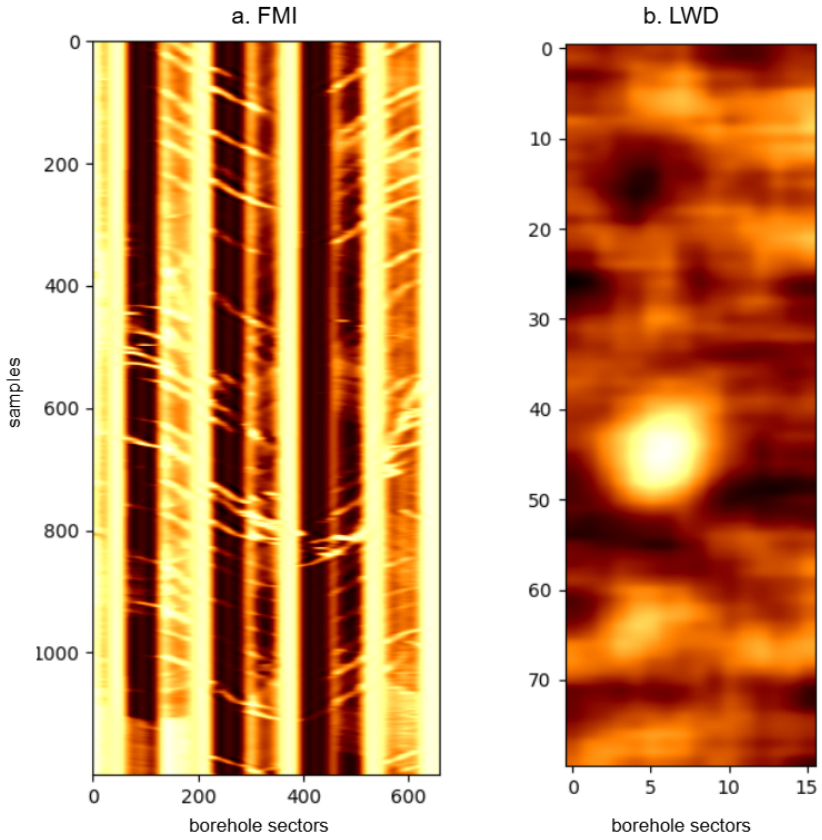

Figure 2.1: Example of a 3-meters interval of an a) FMI and b) LWD azimuthal density borehole image (right).

As mentioned earlier, interpretation results for both wireline and LWD BHIs are subject to uncertainty. In recent years, numerous ML and DL solutions have been proposed in the literature to assist interpreters with this challenging task. In the

next section, we will review the state-of-the-art methods for these problems and highlight how they differ from the approaches proposed in this work.

2.1.2 Related work

In recent decades, significant progress has been made in the detection and characterization of borehole image features. Early methods focused on sinusoidal shape detection using Hough (Hough, 1962) and Radon transforms (van Ginkel et al., 2003). Hall et al. (1996) proposed an approach that combined edge detection, the Hough transform, and an unsupervised neural network called Competitive and Selective Learning for automatic extraction and characterization of features. Thapa et al. (1997) employed a simpler methodology using the Hough transform, where the darkest 10% of pixels were selected, and a 3D search determined the amplitudes, phases, and offsets of potential sinusoids. However, computational limitations at the time made these approaches impractical for production. Similar methodologies were proposed by Glossop et al. (1999), who employed a Laplacian of Gaussian (LoG) filter, and Zhang and Xiao (2009), who combined adaptive histogram equalization with directional filtering before applying the Hough transform. van Ginkel et al. (2003) extended curve detection using Radon transform in a 3D orientation space, but the authors reported poor performance. Assous et al. (2014) introduced a method combining gradient and phase-based approaches to validate congruence and amplitude, utilizing Log-Gabor wavelets for edge validation before sinusoidal detection and estimation, achieving a false positive rate between 2% and 5%. Wang et al. (2007) proposed an edge-detection-based method for rock fracture identification using Support Vector Machines (SVM) with Gaussian kernels, extracting 11 image parameters. While promising, the results did not meet the expected performance.

Al-Sit et al. (2015) improved planar discontinuity detection by combining Gabor filters for feature extraction with the Hough transform for sinusoidal detection. This method achieved high detection rates with a false positive rate of only 1%. Despite these advancements, most of the aforementioned methods were tested on modeled data or low-noise acoustic images. When applied to real acoustic and electrical borehole images with high noise levels, these approaches often produced a significant number of false positives (Anatoli Quintanilla Cruz et al., 2017).

In recent years, the rapid development of deep-learning techniques has significantly improved image processing tasks, including fracture segmentation and extraction. ChunYu et al. (2022) introduced an attention-constrained depth generation network to effectively fill blank regions in electrical imaging logs. Jiang et al. (2021) demonstrated the effectiveness of Convolutional Neural Networks (CNNs) for classifying facies in microresistivity image logs. (Wei et al., 2022) applied Conditional Generative Adversarial Networks (CGANs) (Mirza and Osindero, 2014) to identify horizontal and low-angle fractures in logging images with an accuracy of 90%.Alzubaidi et al. (2022) utilized Mask R-CNN for fracture detection from drilling core images , focusing on sinusoidal fractures, and provided a framework for extracting fractures of similar shapes. Petrov et al. (2023) proposed the first multi-branch deep model for borehole images segmentation.

Compared to traditional approaches, these deep-learning methods have demonstrated remarkable improvements in fracture detection accuracy and robustness. However, all the aforementioned works, whether they are based on ML or DL, focus on high-resolution BHI. Notably, FMI is one of the preferred data for the validation of methods. Furthermore, most of these methods, particularly the deep learning-based approaches, are often described as performing feature extraction (e.g., identifying edges and their associated sinusoids), whereas, in practice, they are primarily engaged in edge detection, which involves localizing edges as regions with high intensity contrast, typically achieved using state-of-the-art neural networks for segmentation. Finally, these methods use field data to train their DL systems and compare the results of the predictions with the results of manual interpretation. Given the high resolution of those data and the extremely high variability of geology as we know it, the use of high-resolution BHI field data for training arguably is a source of bias for the DL systems.

Despite the more extensive use of FMI data in geothermal exploration practices, LWD images can play a crucial role in geosteering navigation for subhorizontal wells (SHW). Their potential to enhance well performance and improve reservoir evaluation has been demonstrated in suburban geothermal projects (Ungemach et al., 2018, 2021). With regard to the DL-based approaches for LWD images for sinusoidal features extraction, Sun et al. (2021) used a 1D approach based on Wavelet Trasform Modulo Maxima (WTMM) and Long-Short Term Memory neural network to pick dipping strata from azimuthal gamma LWD BHI, but the fracture analysis results have not been discussed,

In this work we propose a ML- and a DL-based approach for the automatic sinusoidal feature extraction from LWD borehole image logs. The ML approach (CVML in section 2.2.2) is based on a combination of computer vision (CV) and dynamic time warping (DTW), to dynamically analyze the LWD image and extract the sinusoids. Moreover, we propose a supervised DL method, consisting of two NNs, one for the segmentation of the LWD borehole images (Picknet) and the other to estimate the position of the sinusoids from those segmented images (Fitnet). Two version of Picknet were trained according to the standard (SL) and the curriculum learning (CL) strategies. This differentiation was done to investigate the differences in Fitnet results, given the SL- and CL- Picknet variations, and possible consequences in the human-machine interactions.

Another key aspect of our approach, is that we trained our NNs proposing a synthetic LWD image generator to collect the training dataset. In our experiment, we compare several variants of the same approach, heuristically and selectively changing training strategies and optimization variables (e.g. loss function, weight initialization) and analyze the variations in the results (section 2.3) on synthetic data.

After validating the latter approach, we compare the results of the two methods on field data in terms of prediction time (which we believe is important for real-time applications like geosteering navigation) and benefits and limitations. We use LWD datasets from very far apart areas of the world to validate our methods, to test their robustness across the highly variable borehole data. Within the dataset we include two FMI datasets, introduced in section 2.2.

The use we propose for the implemented methods, follows the semi-automation paradigm, which involves the real-time interaction of the human and the machines, to complete a specific task. Semi-automation has been proposed as a solution in several fields (biology, data labeling, mathematics, geoscience) to limit variability and slowness of humans in tasks that can be easily and quickly handled by machine through their logical processing since the cathode-ray tubes (Desmond et al., 2021; Bennett et al., 1963; Cadena-Herrera et al., 2015). In our method, semi-automation is employed to enable the operator/machine interaction to validate and discard the predictions in real-time. With the rapid growth of Machine Learning (ML) and Deep Learning (DL) in our field, we are arguably capable of addressing nearly all problems traditionally associated with human subjectivity. However, the evaluation of how such methods influence task-specific epistemic uncertainty remains largely unexplored. We argue that semi-automation and the exploration of multiple alternatives using the same machine for identical tasks are essential for the robust validation of DL methods in geoscience. Furthermore, the impact of these methods on human cognitive processes and associated uncertainties should be thoroughly assessed. This aligns with the findings of Andrews et al. (2019); Shipley et al. (2013); Bond et al. (2007); Wilson et al. (2019), which highlighted that individual psychological characteristics play a more critical role in determining outcomes than prior experience. The understanding of how machine can influence our perception, as well as other humans do, is an open topic in computer and cognition sciences for the most common visual tasks of our daily life (Funke et al., 2021; Quintas, 2023; Nakada et al., 2018; Liu et al., 2023). Hopefully, this defines the path for the future of this research.
Finally, some boreholes in the LWD dataset from the Coral Field (Mozambique) were associated with highly risky conditions, that eventually lead to borehole instability and environmental, techno-economical damage to the exploration/production activity. A Mass Transport Deposit (MTC) was interpreted in two boreholes(LWD3 and LWD1) of this dataset by experts (Fonnesu et al., 2020), and this eventually caused instability and loss of instrumentation in one of them (LWD1). In our AI solutions, we also extracted implicit information from the intermediate steps of the automated analysis, to observe if we can extract in real-time some additional supportive information for the decisional processes related to geosteering. We did that for both the CVML and the DL method, conceptualizing different approaches to gather and use the methods results and investigate the possible related geological information (sect. 2.3.6).

2.2 Data and Methods

In this section, we describe the CVML and DL methods proposed in this thesis, along with the field datasets used for their evaluation. To assess the generalization and scalability potential—particularly for the DL method—we utilized real azimuthal density LWD and FMI image logs from diverse drilling projects with varying objectives (Fig. 2.2).
The density LWD data were sourced from the Cascadia subduction zone and Mozambique, acquired for research purposes and gas field exploration, respec-

tively. The FMI logs from the Naaldwijk geothermal field in the West Netherlands Basin (WNB), as interpreted in Boersma et al. (2021), were used to constrain and inform fluid flow simulations aimed at evaluating long-term net energy production. This dataset is publicly available at this link. To evaluate whether our method could replicate human intent, we tested its scalability on these high-resolution BHI datasets. However, since the interpretations in Boersma et al. (2021) were not accessible in sinusoid-related forms, we first validated the upscaling approach using another open FMI dataset. This dataset comprises 115 examples of 2-meter FMI segments, provided by Conoco Phillips, and is available at this link. Unfortunately, the precise locations of these datasets remain undisclosed. Regarding the Mozambique dataset, two boreholes were associated with a massive mass transport complex (MTC) (Fonnesu et al., 2020). MTCs are particularly significant in hydrocarbon exploration but present considerable risks due to their high heterogeneity. In one of these boreholes, drilling operations had to be halted prematurely due to borehole instability issues within the MTC, resulting in material loss. The generation of synthetic training data for the neural network (NN) is detailed in Section 2.2.1. This synthetic data was exclusively used in the DL-based method, as the CVML approach does not rely on NNs and did not require any training.

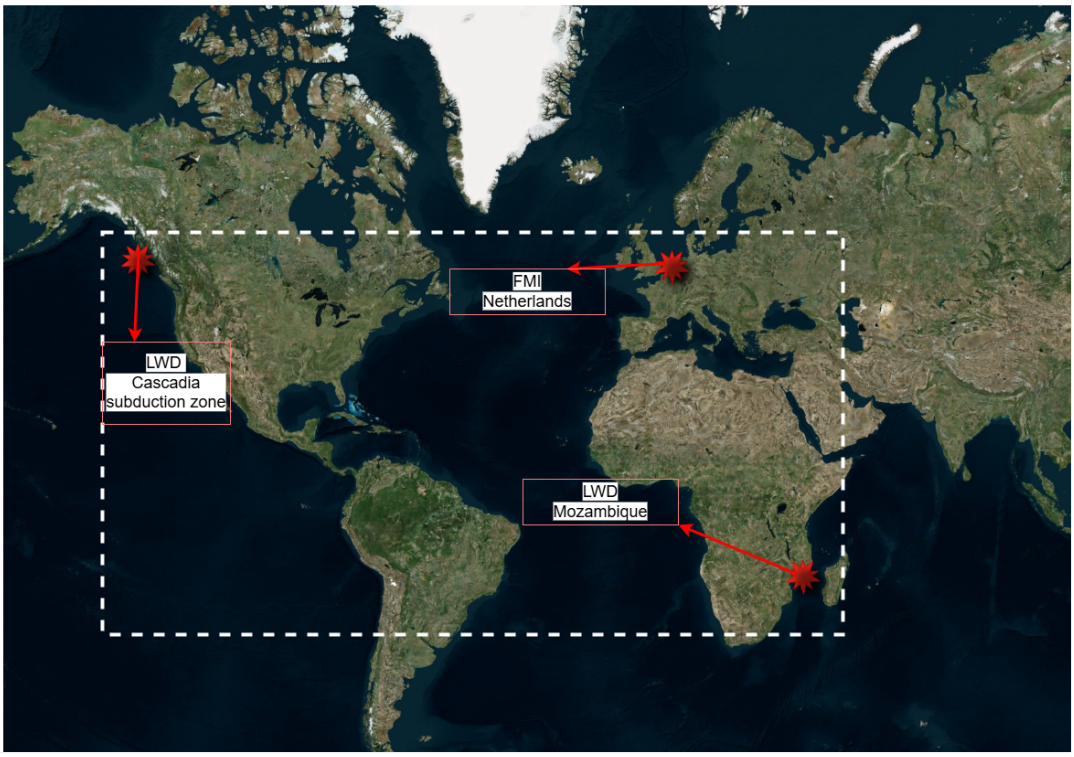

Figure 2.2: Locations of the datasets used to validate the AI-based proposed methods.

2.2.1 Synthetic data

One of the most interesting aspects related to our supervised learning approach is the generation of a synthetic dataset, implementing an algorithm, trying to solve the main challenges related to the representativeness of the dataset and the sparsity of the geological features (Gupta et al., 2019). The representativeness is critical for network training, as the distribution of the dataset defines the learning domain and shapes the model's ability to generalize within that space. The sparsity undermines the applicability of supervised learning to this problem. This is because, especially in high-resolution BHI, several geological features of multiple orders can be observed in non predictable geometrical and structural configurations. The classical NN implementations for a supervised learning problem, require the desired output to be a (pre-defined) fixed size array of desired values, enabling the calculations of all those parameters, e.g. loss function, that make the NN converge towards the optimum in the parameter (weights) space. The strategies we used to address these challenges are discussed in this section.
First, to generate the training dataset, we must define the size of the synthetic images used for network training. This step is critical as it determines the scale at which borehole images are automatically analyzed. As stated above, NN are unable to handle variable-sized training datasets, making consistent input dimensions a key design consideration. In our case, we base the input size on the sampling characteristics of industry-standard Logging While Drilling (LWD) tools, using a depth sampling rate of 0.05 meters and dividing the 360° borehole into 16 sectors. To generate synthetic borehole image windows of 1 meter, we can initialize the data as a zero matrix $\in \mathbb{R}^{20x16}$. Within this zero matrix, we can then define the positions and number of discontinuities arbitrarily. The depth of each interface (edge) is defined by randomly selecting its starting point in the 20×16 initial matrix. From this starting point, a sinusoidal trajectory is defined as:

$$y(x) = A \sin\left(\frac{2\pi}{N}x + \phi\right) \tag{2.1}$$

where:

- The amplitude A is defined as:

$$A = \text{rand}(1, S_{\max}) = \text{rand}(1, \text{rand}(1, 20)) \tag{2.2}$$

- The phase ϕ is defined as:

$$\phi = \text{U}(0, 2\pi) \tag{2.3}$$

- N represents the number of sectors in the borehole image, which is the number of bins into which the tool divides the 360° of the borehole.

- x is a continuous set of values covering the full 360° of the sinusoid in N discrete steps.

In this context:

- A is the amplitude of the sinusoid, randomly sampled between 1 and $S_{\max}$.

- ϕ is the phase of the sinusoid, randomly sampled from a uniform distribution U between 0 and 2π.

- N is the total number of sectors or bins in the borehole image.

- x represents the discrete positions along the sinusoidal wave from 1 to N.

.

After determining the interface location, the pixels corresponding to the sinusoidal edge positions in the original matrix are set to 1, creating a binary matrix where ones mark the edges and zeros fill the rest. This binary matrix is crucial for the workflow design. However, to construct a synthetic raw borehole image panel, we also need to simulate noise before and after convolution, as well as variations in petrophysical properties with different contrasts. For this, we assign a random contrast value within a specified range (2 to 10 in our experiment). To model the transition between petrophysical intervals, we convolve each segment with a kernel of ones, where the kernel's length equals the number of samples in the panels. This convolution smooths the synthetic data by averaging over a window, thereby reducing noise and creating smoother trends. This process helps generate more realistic training data, which is beneficial for training machine learning models or for any application requiring smooth data. Our approach also involves applying least-squares fitting to each detected edge in the synthetic data, allowing us to accurately retrieve the associated best-fitting plane. This technique ensures that the edges are modeled with minimal MSE error, providing a precise estimation of the dip and orientation of geological features. However, it's important to note that training with evenly spaced discontinuities can be problematic. If we were to place the same number of discontinuities in each synthetic panel, the network would be trained to always detect a fixed number of edges, leading significantly limiting the model's generalization to real-world data. We need to find a solution that balances the creation of a representative dataset with the requirements of designing a workflow compatible with the architecture of the NN. This involves ensuring the synthetic data captures realistic variations while aligning with the constraints of the network, such as fixed input sizes. By doing so, we can optimize the training process to produce robust models that perform well on both synthetic and real-world data. In our experiment, we generated panels with a random number of interfaces $D \in [0,2]$.

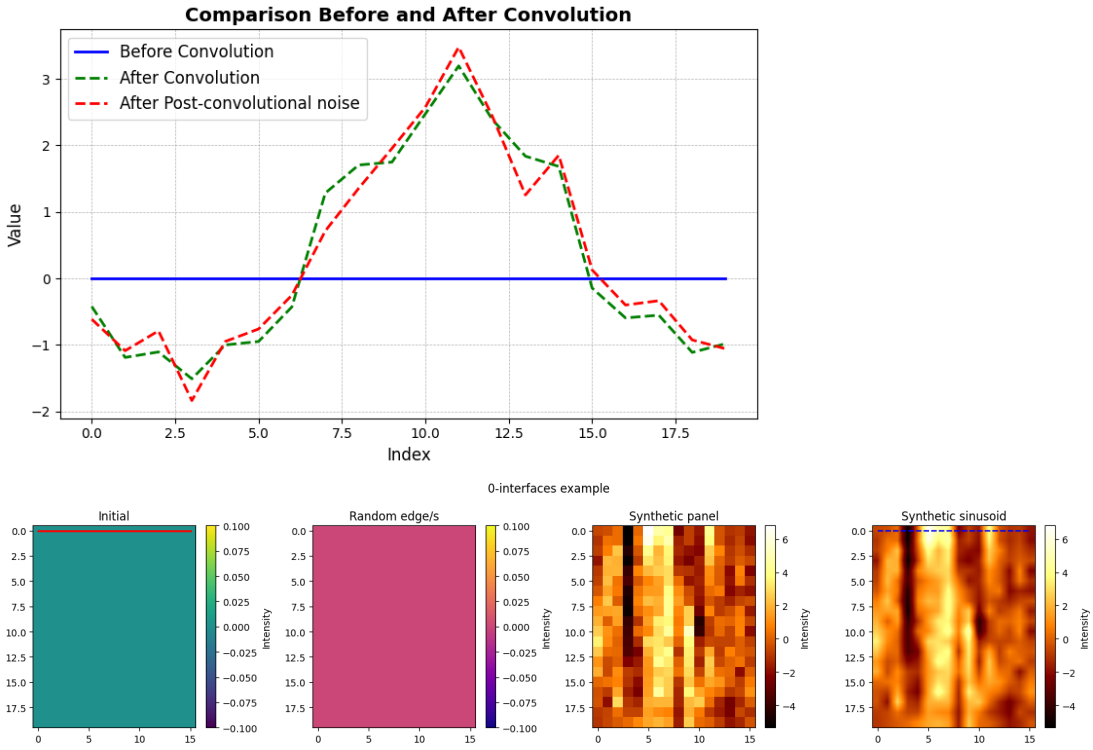

Figure 2.3: Synthetic example with 0 edges in the image.

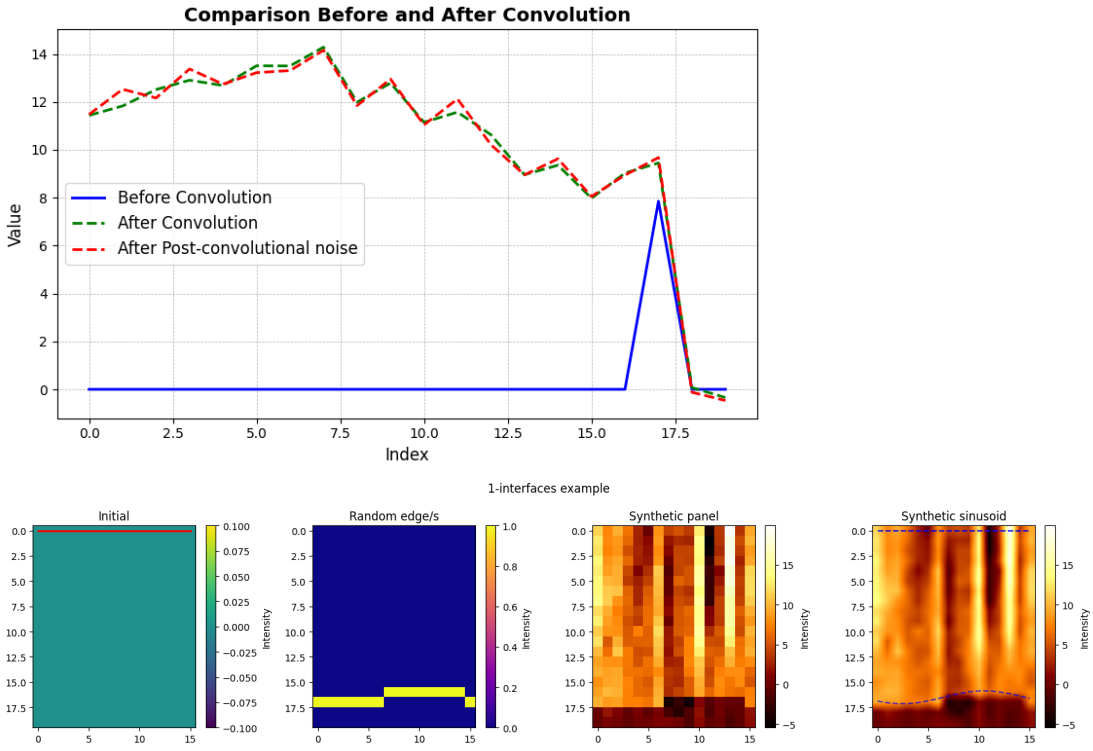

Figure 2.4: Synthetic example with 1 edge in the image.

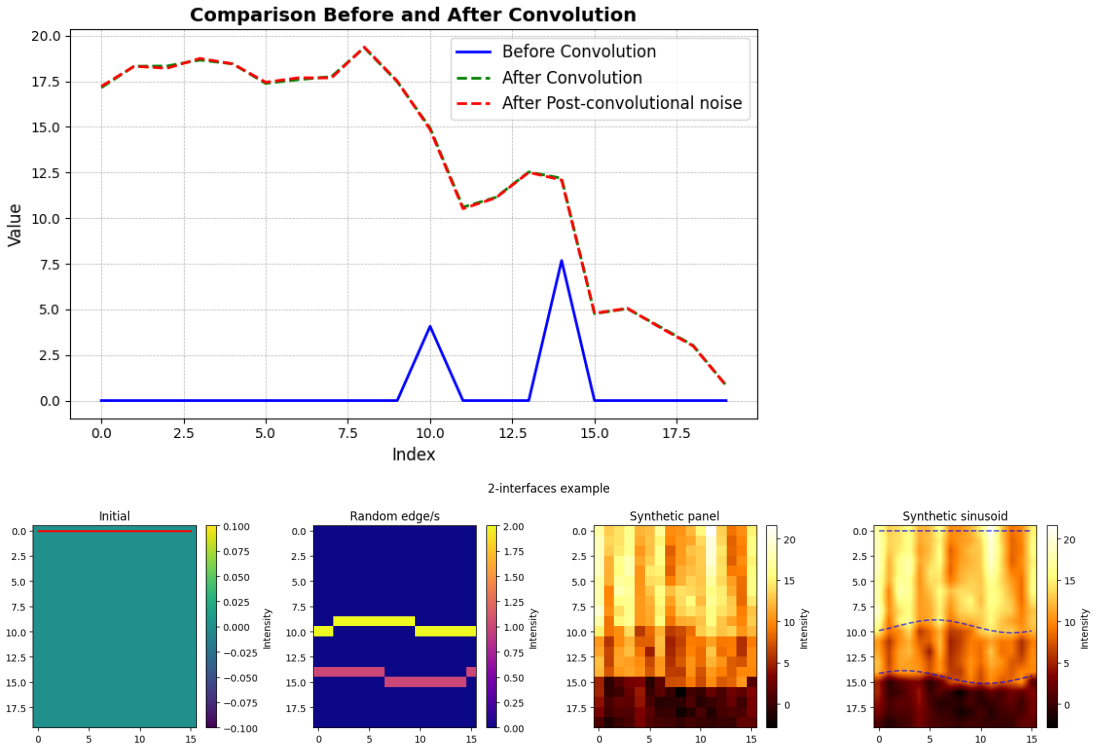

Figure 2.5: Synthetic example with 2 edges in the image.

Our strategy involves zero-initializing a fixed number of vectors for each panel, corresponding to the maximum number of discontinuities present in the image. If one or more edges are added to the panel, the associated zero-initialized vectors are updated to represent the y-coordinates of the best-fitting plane for each edge. This approach ensures that even if the number of edges varies between panels, the network maintains a consistent structure by associating the detected discontinuities with the corresponding best-fitting planes keeping all the others vectors as zeros. An example of that can be seen in Fig. 2.3, 2.4, 2.5.

In section 2.2.4, we will detail how the synthetic dataset was used to evaluate the complexity of the task from the perspective of the DL system under varying conditions in the data generation algorithm. This analysis served as inspiration to test different training strategies aimed at creating alternative digital cognitive styles for the system. We will also explore the differences among these approaches and their potential impact on human intelligence augmentation.

2.2.2 Computer vision & Machine learning - CVML

In this section we will introduce the first method (CVML) we developed for the LWD borehole images automated analysis. We will describe the basic functioning principles of the CV and ML methods that have been combined for this purpose. Finally, we will describe the whole method with details on how we make use of both CV and ML to carry out this task effectively.

Computer vision: the Sobel operator

The computer vision (CV) part of this method is based on the Sobel method, which is an edge detection computer vision methodology. The method relies on the so-called Sobel operator: an efficiently computable isotropic gradient estimate of digitized pictures. This estimate is provided by the vector summation of the 4 possible simple central gradient estimates obtainable in a 3x3 neighborhood, which provides an averaging over the directions of gradient measurements (Sobel, 2014; Thompson et al., 1974). An image is essentially a density or an intensity function, depending on the physical nature of the image source, and a 2D function that maps pixel values P onto an NxN domain of coordinates, such that

$$I(u, v) \in P, \quad \text{where} \quad u, v \in \mathbb{N} \tag{2.4}$$

Several operations can be applied on such images, like point operations and filtering (Burger and Burge, 2016). Point operations are the transformation of an original pixel value a to a value b, such that

$$a = f(I(u, v)), \quad \text{and} \quad b = f(a) \tag{2.5}$$

whereas filters, normally obtain a new pixel value in the output image as a function of a region of pixels $R(u, v)$

$$a = f(I(u, v)), \quad \text{and} \quad b = f(R(u, v)) \tag{2.6}$$

In filtering operations, the filter operator is essentially a matrix $H(i, j)$ with the same size as $R(u, v)$, where each H element specifies the weight to be associated with the corresponding pixel in R. This are definied as *linear* filters, since their outcome is a weighted sum of $R(u, v)$.
Filters are what we need for *edge detection* purposes (Shrivakshan and Chandrasekar, 2012; McIlhagga, 2018; Mehrotra et al., 1992; Sarkar and Boyer, 1991; Torre and Poggio, 1986). Edges are defined as sharp pixel intensity contrasts and their detection by means of computer vision is a very useful operation that allows to reduce the amount of data related to the image, preserving the interesting information, i.e., the edge itself.
Mathematically, the gradient of a function $f(u)$ with respect to an independent variable x corresponds to its first derivative

$$\frac{df(u)}{dx} = \frac{f(u+1) - f(u-1)}{(u-+1) - (u-1)} = \frac{f(u+1) - f(u-1)}{2} \tag{2.7}$$

To estimate the gradients in each direction of the image we express the 3x3 sobel kernels using the following weighting functions

$$H_x = \begin{bmatrix} -1 & 0 & 1 \\ 2 & 0 & 2 \\ -1 & 0 & 1 \end{bmatrix} \quad ; \quad H_y = \begin{bmatrix} 1 & 2 & 1 \\ 0 & 0 & 0 \\ -1 & 2 & -1 \end{bmatrix} \tag{2.8}$$

These filters calculate gradients within the image by means of convolution, so for each directional gradient estimation filter like in equation 2.8, the output image is computed as follows for the x direction

$$G_x = \sum_{i=-\infty}^{+\infty} \sum_{j=-\infty}^{+\infty} I(u-i, v-j) \cdot H_x(i,j) = I(u,v) * H_x(i,j) \tag{2.9}$$

and

$$G_y = \sum_{i=-\infty}^{+\infty} \sum_{j=-\infty}^{+\infty} I(u-i, v-j) \cdot H_y(i,j) = I(u,v) * H_y(i,j) \tag{2.10}$$

The final gradient G is given by the norm of G_x and G_y

$$G = \sqrt{(G_x^2 + G_y^2)} \tag{2.11}$$

Dynamic Time Warping

Dynamic Time Warping (DTW) is a non-linear optimal alignment method between two given numerical sequences (Müller, 2007). The speciality of DTW is that it can be applied to any pair of timeseries, having different sampling and lengths. The alignment of the numerical series is achieved by warping along the time axis to produce the optimal alignment, and this is particularly useful when we need to compare data with time deformations and different *speed*. As a dynamic programming (DP)-based technique, DTW solves the optimal alignment problem by breaking it down to simpler overlapping subproblems, which makes it particularly suitable for problems with optimal substructure. A problem exhibits optimal substructure if an optimal solution to the problem can be constructed from optimal solutions to its subproblems, which are solved multiple times (overlapping subproblems). This overlapping requires some *memoization*, to avoid recalculation of the same subproblem solution during recursion. In DTW each time points alignment between the two timeseries is the subproblem, whose solution is given by the minimum cost, i.e., dissimilarity between the aligned points. DTW compares two time-dependent sequences $X := (x_1, x_2, ..., x_N)$ and $Y := (y_1, y_2, ..., y_N)$ with lengths $N \in \mathbb{N}$ and $M \in \mathbb{N}$, respectively. The sequences may be discrete or with equidistant samples in time, regardless if they are discrete signals or feature sequences. Let's fix a feature space $\mathcal{F}$, such that $x_n, y_m \in \mathcal{F}$ for $n \in N$ and $m \in M$. The comparison between this two features is made by means of a local distance measure defined by the function

$$c : \mathcal{F} \times \mathcal{F} \to \mathbb{R}_{\geq 0} \tag{2.12}$$

where c is the cost function, which is small for similar x and y, otherwise is large. The function c is in the domain of $\mathcal{F}$ where it takes the set of all possible ordered pairs of elements from $\mathcal{F}$ and maps each element from $\mathcal{F} \times \mathcal{F}$ to the set of non-negative real numbers $\mathbb{R}_{\geq 0}$, since c is a distance or similarity function.
The computation of a local similarity measure for each pair of elements x_n and y_m in the end results in a *cost matrix* $C \in \mathbb{R}^{NxM}$, such that $C(n,m) := c(x_n, y_m)$. The optimal alignment can intuitively be interpreted as the set of entries in C that form the minimal overall cost. This set defines the so-called *warping path*, defined as a sequence $p = (p_1, ..., p_L)$, which defines the alignment and satisfies

1. boundary condition: $p1 = (1,1), pL = (M,N)$.

2. monotonicity condition: $n_1 \leq n_2 \leq ... \leq n_L$ and $m_1 \leq m_2 \leq ... \leq m_L$

3. step size condition: $p_{l+1} - pl \in (1,0),(0,1),(1,1)$for $\mathrm{l} \in [1 : L-1]$

A generic warping path is thus associated to a total cost c_p defined as

$$c_p(X,Y) := \sum_{l=1}^{L} c(x_{nl}, y_{ml}) \tag{2.13}$$

but since we are looking for the optimal alignment, the optimal warping path is the final desired results of the DTW method and it's equal to the optimal total cost $c_{p*}(X,Y)$, such that

$$DTW(X,Y) := c_{p*}(X,Y) = \min\{c_p(X,Y)\} \tag{2.14}$$

The CVML algorithm

As a combination of the computer vision Sobel method and DTW, CVML algorithm automatically analyzes borehole images and counts on two fundamental steps. First, the Sobel filters are applied on borehole images to extract the value of the gradients at each pixel position as explained in Section 2.2.2. Starting from an input borehole image I, we obtain the associated gradient image G as in equation 2.11. Let $G_{max} = max(G)$, as the maximum value in the gradient image, we defined a threshold T as a percentage α of G_{max}, as

$$T = \alpha G_{max}, \quad \text{where} \quad 0 < \alpha \leq 1 \tag{2.15}$$

We then define a set P of pixel coordinates (i,j), where the gradient value is greater than T as

$$\mathbf{P} = \{(i,j) | G_{i,j} \geq T\} = \{(i,j) | G_{i,j} \geq \alpha G_{max}\} \tag{2.16}$$

This set of coordinates represents the set of points which are likely to correspond to detectable edges, as they correspond to the higher gradient values. For this, we use P as the intial step for the automated feature correlation in borehole images. For instance, let $\mathbf{P}(i_p, j_p)$ be the coordinate of a pixel $\in \mathbf{P}$ with $j_p = 0$. We define the subsequence S_p as

$$S_p = (I(i_{p-h}, j_p), ..., I(i_{p+h}, j_p)) \tag{2.17}$$

where h is an arbitrary value to be decided according to the desired automated analysis scale and the image resoluton. At this point we can select a set of subsequences $S(j_{p+1})$ in the adjacent image column, such that

$$S(j_{p+1}) = \{S(i, j_{p+1}) \quad | i \in [i_p - H, i_p + H]\} \tag{2.18}$$

where H is the search length in both search directions with respect to i_p. We compute the DTW of S_p and the i^{th}- $S(i, j_{p+1}) \in S(j_{p+1})$ and we select the best adjcent subsequence $S(i*, j_{p+1})$ as

$$S(i*, j_{p+1}) = \arg \min_{i \in [i_p - H, i_p + H]} (DTW(S_p, S(j_{p+1})) \tag{2.19}$$

Finally, we use $S(i*, j_{p+1})$ as the new subsequence and repeat this procedure until 360° coverage of borehole image. In other words, we build a new set of points $\mathbf{P'}$, defined as the union of the original set P and the set of optimal coordinates such that

$$\mathbf{P'} = P \cup \{(i*, j) | j = j_{p+1}, ..., j_{max}\} \tag{2.20}$$

Ultimately, we want to fit a sinusoid to each vector of optimal i coordinate $i_{best} = [i_*^{j_p}]$, which represents the best 2D (planar) approximation of the edge in the original image. Let's the sinusoid be

$$f(j) = Asin(\omega j + \phi) + c \tag{2.21}$$

where A is the amplitude, ω the angular frequency, ϕ the phase, and c the offset of the sinusoid. The fit is done by minimizing the squares residuals between $f(j)$ and the vector of optimal i coordinates as follows

$$(A*, \omega*, \phi*, c*) = \arg \min_{A, \omega, \phi, c} ||i_{best} - f(j)||_2^2 \tag{2.22}$$

where $A*, \omega*, \phi*, c*$ are the optimal sinusoid parameters. In the end, we get the set of detected features $F \subseteq \mathbb{R}^{j_{max}}$ and we say that CVML maps the input borehole image to a multi-dimensional real vector space corresponding to the set of automatically detected geological features in the image:

$$\mathcal{CVML} : I \rightarrow \mathbf{F_{CVML}} \tag{2.23}$$

One keynote about the described CVML methodology is the need to manually set parameters, such as the α coefficient (related to the thresholds) and the search length H that defines the scale of the analysis.

2.2.3 Deep learning

Deep learning is a field of AI that uses parametric models as deep neural networks (DNN) to perform sequential operations on data and eventually find a useful representation of them for our purposes (LeCun et al., 2015). The process of *learning* refers to the optimization problem these networks are carrying out to approximate the function $\varepsilon : \mathbb{R}^d \rightarrow \mathbb{R}$ given ε

$$\bar{y_m} = \varepsilon(x_m) \tag{2.24}$$

where $\bar{y_m}$ and x_m are the m^{th} output and input data, respectively. In *supervised* learning, we then use deep networks to compute the output $\varepsilon_{(m+1)}$ related to the $m + 1^{th}$ input x_{m+1} not by explicit knowledge of the function ε, but using the

$m = 1, 2, ...$ input-output data pairs. This is done through the computation of minimizers of the function, known as loss functions, that compares the output $\bar{y_m}$ with the real output y_m (also known as ground truth). Specifically, a loss function $\mathcal{L} : \mathbb{R} \times \mathbb{R} \to \mathbb{R}_{\geq \nvDash}$ can quantify how good or bad a NN is based on its prediction and the target output or 'ground truth'(these inputs need to be of the same dimension). A loss function compares these two inputs and maps them to a non-negative real number using some kind of operation.
In order to do this, the NN, as hierarchical parametric models, apply sequential linear transformations to the input data, followed by a so-called activation function which introduces non-linearity in the process. The linear transformations are performed at each k^{th} layer of a network with depth d using the matrix $W^k \in \mathbb{R}^{d_{k-1} \times d_k}$. The transformation is applied on the output of the $k-1$ layer $X_{k-1} \in \mathbb{R}^{N \times d_{k-1}}$ to get the output of the k^{th} layer $X_k \in \mathbb{R}^{N \times d_k}$, given by $X_{k-1}W^k$. N is the number of training examples stored in the input dataset $X \in \mathbb{R}^{N \times D}$ (e.g., N digital images, or a dataset of N sampling of D random variables).
For instance, in our problem the input data for our DNN are low-resolution LWD borehole images I, so we say that

$$\mathbf{F_{DL}} = \varepsilon(I) \tag{2.25}$$

Let $\mathbf{F_{DL}}$ denote the set of sinusoids corresponding to approximately planar geological features observed in the borehole image as in 2.23. This prompts several considerations. First, we address the problem as a *regression* task. Specifically, we do not apply DL to classify the detected geological features by their geological meaning in the subsurface. Instead, we aim to automatically retrieve planar approximations of these features by estimating the x- and y-coordinates of their associated sinusoids. Second, in real-world scenarios, the number of geological features we expect to retrieve within a specific region is random and sparse (Gupta et al., 2019), which is one of the main limitations of the supervised learning approach for this problem. Those two points pose challenges for our supervised approach, regarding dataset construction and training.
For dataset construction, we need to address the question: How can we build a dataset that accounts for variability in the number of geological features within borehole images?
For the training phase, we must determine how to train the network to predict varying numbers of sinusoids across different examples of borehole images. We construct the training dataset using a synthetic data generator (section 2.2.1). Within the training dataset, the number of edges in the input images ranges from 0 to 3, prompting corresponding adjustments in target predictions, represented as a set of three zero-initialized vectors. For example, in instances where a single edge is present ($n_{int} = 1$), the target comprises two zero-initialized vectors and one vector containing the sinusoidal coordinates of the edge. When two edges are present ($n_{int} = 2$), the target includes one zero-initialized vector and two vectors with sinusoidal coordinates corresponding to each edge, and so on.
This design choice effectively navigates the challenge of fluctuating edge counts. By structuring the output to accommodate variable edge numbers, the model overcomes the constraints of fixed-size output arrays, allowing it to adapt flexibly

and predict coordinates for different numbers of edges. This approach enhances the model's versatility.
Let the set of sinusoid $\mathbf{F_{DL_m}}$ for the m^{th} borehole image example I_m be defined as

$$\mathbf{F_{DL_m}} = \{f_i \mid i \leq n_{int}\}, \quad \text{with} \quad f_i \in \mathbb{R}_{1,D}, \forall i \leq n_{int}, \tag{2.26}$$

For our specific problem an example of objective function to be optimized (minimized) is given by the well known Mean Squared Error (MSE), defined as

$$\mathcal{L}_{MSE} = \frac{1}{M}[\sum_{m=1}^{M} |\varepsilon(I_m) - \mathbf{F_{DL_m}}|^2 \tag{2.27}$$

which is an adaptation of the L2 norm normalized with respect to the dataset size.
Another key aspect of deep learning is *backpropagation*. This algorithm is used during training to minimize the error in eq. 2.27 between the neural network output and the target and to effectively carry out the optimization procedure. In other words, backpropagation is the part of the training procedure where we optimize the loss function by updating the neural network parameters (weights). It is done calculating the gradient of the loss function with respect to any weight in the network, using the chain rule to propagate the error from the output to the input layer. Let's define y_m as the ground truth and $a^{[L+1]}$ as the activation function in the output layer, where L is the total number of hidden layers. To explain the *backpropagation* idea we calculate δ^{L+1} as

$$\delta^{[L+1]} = a^{[L+1]} - y_m \tag{2.28}$$

For each layer l down to the input layer $\delta^{[l]}$ is

$$\delta^{[l]} = (W^{l+1}\delta^{l+1}) \odot \sigma'(z^{[l]}) \tag{2.29}$$

where $\sigma^{'}(z^{[l]})$ is the derivative function of the activation function at layer l and $z^{[l]}$ is the linear combination of inputs for a particular layer l as it is calculated using the activation function of the previous layer, as follows

$$z^{[l]} = W^{[l]}a^{[l-1]} + b^{[l]} \tag{2.30}$$

with b being the bias term of layer l.
We can now calculate the gradient of the weights of each layer l as

$$\Delta W^{[l]} = a^{T}_{[l-1]}\delta^{l} \tag{2.31}$$

which becomes

$$\Delta W^{[1]} = x^{T}\delta^{l} \tag{2.32}$$

for the input layer, with x being the input. Contrarily, the $\delta^{[l]}$ for the output layer is calculate as the dot product between the activation functions in the layer preceding the output layer and difference between the activation functions in the layer and the output y.

The weight gradient is used to update the weights of each hidden layer during the training process as follows

$$W^{[l]} = -\alpha \Delta W^{[l]} + W^{[l]} \tag{2.33}$$

where α is a learning parameter called *step length* which determines the degree of the adjustments during learning. A small α will result in small adjustments, and viceversa. Note in eq. 2.33 that the update of the weights is done in the opposite direction with respect to the weight gradient and this is the so-called *gradient descent*. This essentially determines the adjustments of weight matrices to follow the same directions of loss reduction.
As explained above, in order to avoid extremely complicated complicated architectures to deal with the possible fluctuating number of edge counts in borehole images, we designed the training dataset to have up to 3 zero-initialized vectors per image, representing the 3 possible edges at the observation scale. The eq. 2.27 thus becomes

$$\mathcal{L} = \frac{1}{M}[\sum_{m=1}^{M} \sum_{j=1}^{3} |i_{jm} - f_{jm}|^2 \tag{2.34}$$

which is backpropagated troughout all the layers of the network in combination with gradient descent for our optimization problem. We can notice straight away that the choice of the activation function and the loss function to use have a fundamental impact on the optimization process, and this determines the importance of backpropagation for deep learning models. However, this is not the only important component. Weight initialization is also very important in neural networks design. The initial points can determine whether the optimization algorithm is able to converge and shows a stable behavior. Normally, weight matrix initialization is done by drawing randomly from Gaussian or uniform distributions, but the initialization choice can be customized according to the particular problem we are facing. For instance, let's build a simple neural network to demonstrate the feasibility of our experiment and the impact of various activation functions and weight initialization on the optimization of our specific problem. The architecture of the network consisted of four layers: an input layer, two hidden layers, and an output layer. The input to the network was a flattened representation of the generated matrix, while the output was expected to replicate the sinusoidal pattern present in the original matrix. The weights for each layer were initialized using both a random sampling from Gaussian distribution and a customized sinusoidal initialization, and we combined these initial values with a ReLu activation function and a sine-based activation functions to compare the results. Training was conducted using backpropagation with gradient descent, with a learning rate of 0.01 to balance between stability and learning speed. Mean Squared Error (MSE) (see eq. 2.27) was employed as the loss function to quantify the deviation between predicted and true sinusoidal patterns. Fig. 2.6-2.7 show the results for a random weight initialization. Fig. 2.6 refers to the NN trained with the ReLu activation function, whereas Fig. 2.7 to the NN trained with a customized sinusoidal activation function. We can observe how this design strategy can benefit the training process and the final results of this simple demonstrative NN.

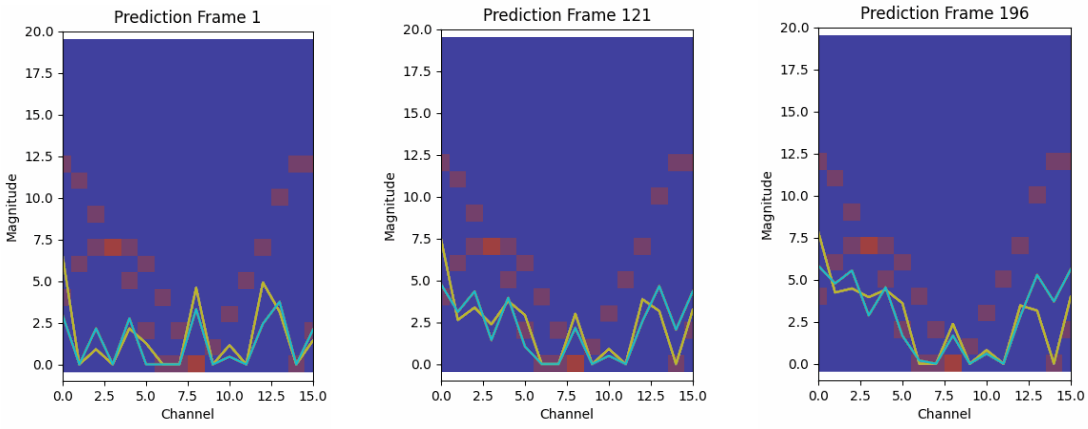

Figure 2.6: Random weight initialization and ReLu activation function

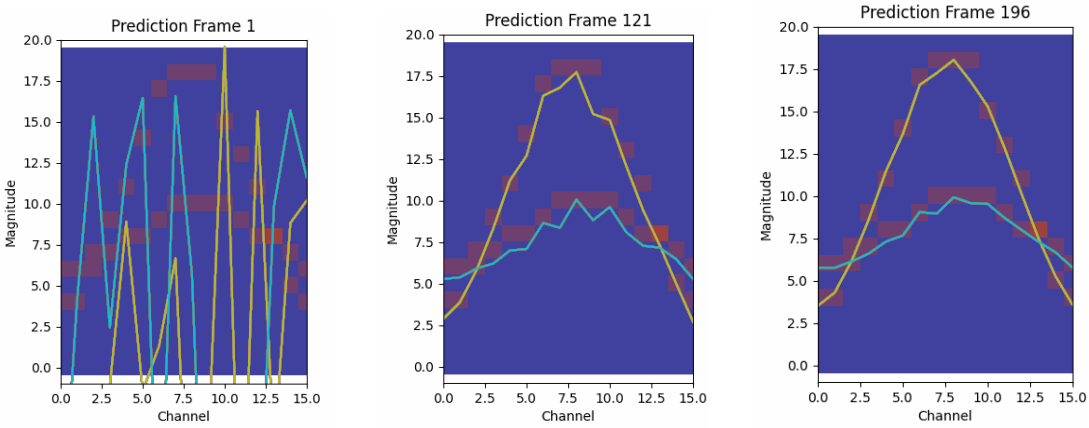

Figure 2.7: Random weight initialization and sine activation function

The same activation functions comparison was done for the sinusoidal weights initialization and the results are shown in Fig. 2.8-2.9, for the ReLu and the sinusoidal activation functions respectively. We can observe from the figure that the combination of the sinusoidal weights initialization we could achieve the best results, for the same number of iterations (epochs). However, the impact of the initial weights distribution appears to affect the final results to a less extent, with respect to the activation function. Fig. 2.7 and 2.8 refer to the sinusoidal activation function combined with a random and a sinusoidal weight initialization function, and both show a faster convergence with respect to the ReLu.

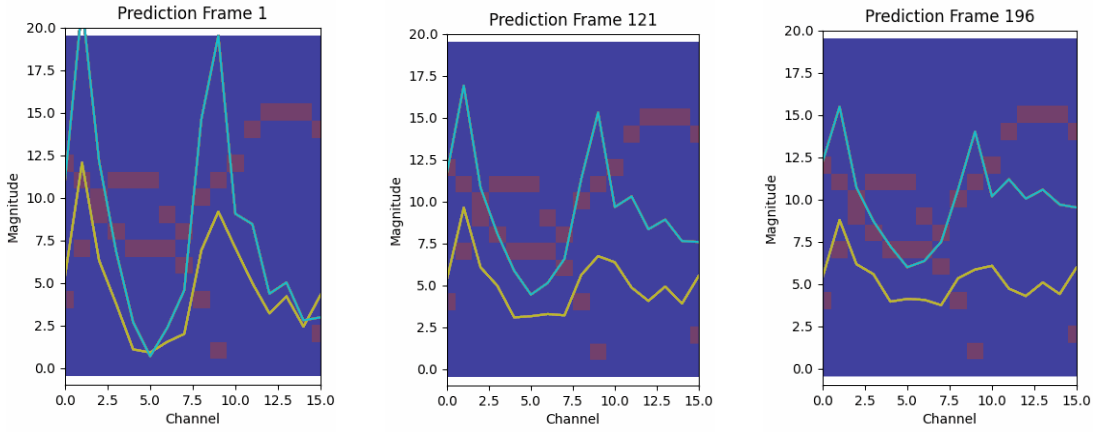

Figure 2.8: Sine weight initialization and ReLu activation function

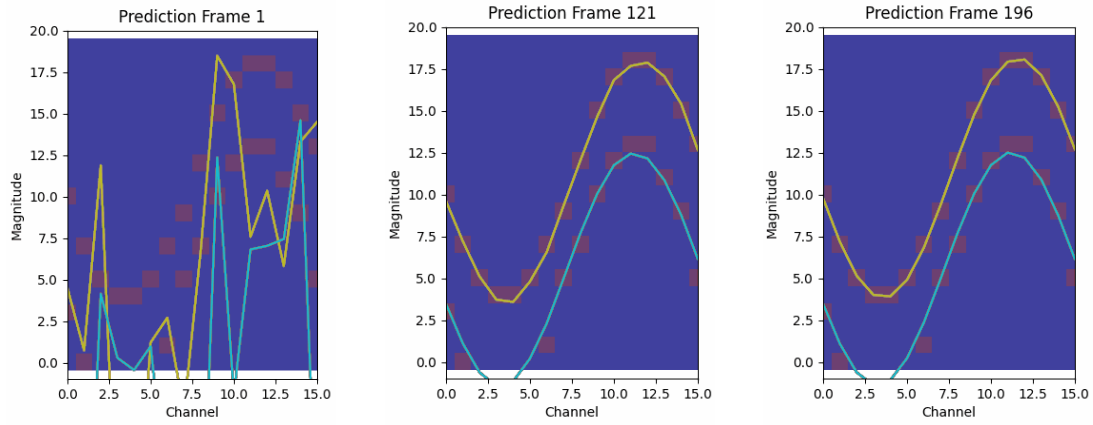

Figure 2.9: Sine weight initialization and sine activation function

This demonstrative supervised-learning example using a simple neural network (NN) was designed to introduce the fundamental concepts behind the learning process and showcase the applicability of a regression framework for automated feature analysis in borehole images. However, our DL approach is more complex than the NN used in this example, which is a Fully Connected Neural Network (FCNN). An FCNN consists of a series of layers where all weights are mutually connected. We chose the FCNN for demonstration purposes because it is a broadly applicable type of NN, as it makes no specific assumptions about the input. However, this generality also makes FCNNs less effective than specialized NNs that are tailored to specific input structures. For instance, Convolutional Neural Networks (CNNs) are designed specifically for image data. They leverage the assumption that input data are images by applying convolutional operations using sliding filters (kernels) of arbitrary sizes across the image.
In our actual DL system, we rely more on CNNs, Picknet , than FCNNs, Fitnet. Additionally, the task-specific design choices for the DL system differ significantly due to the use of specialized architectures. The DL method described in the following sections involves two distinct deep neural networks (DNNs) designed

to emulate the human cognitive process during the interpretation of BHI. These architectures are tailored to enhance performance on the specific challenges posed by BHI analysis.

Picknet: a U-Net architecture

Binary segmentation can be very effectively performed by a U-Net (Ronneberger et al., 2015). This is a symmetric architecture (Fig. 2.10) corresponding to a down-sampling and up-sampling path.

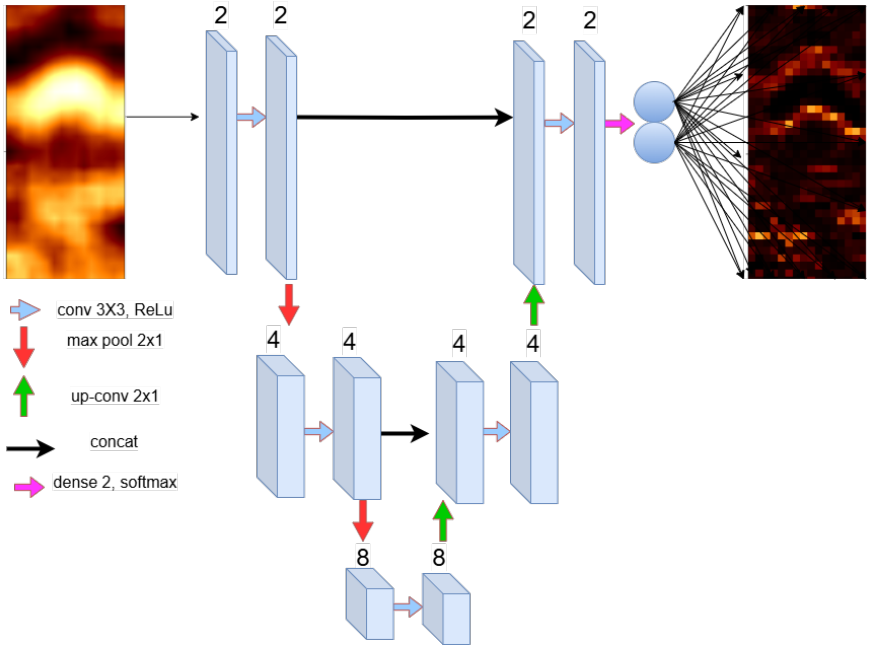

Figure 2.10: Picknet is a U-Net architecture for 2D image processing. The input image passes through convolutional and max-pooling layers for feature extraction, followed by up-convolutional layers to reconstruct spatial details. The concatenations (also known as skip-connections) combine high-resolution features from the encoder with decoder layers, and a final dense softmax layer generates the output classification map.

The down-sampling path consists of three layers of 3x3 convolutions followed by a rectified linear unit activation function (ReLu) and a 2x1 max pooling with a stride of 2, to keep the information related to each channel in the datum. At each down-sampling step, we double the number of feature maps, and because of this, the system can extract the local information from a specific input to proceed with binary segmentation. In the expansion path, each step consists of up-sampling the feature maps, followed by a convolution that halves the number of feature maps, and finally a concatenation between the deconvolution and the feature map of the corresponding down-sampling stage. Over the final feature map, we compute a pixel-wise softmax combined with the categorical cross-entropy loss function (Mannor et al., 2005)

$$loss = \sum_{i=0}^{C} t_i log(s_i) \tag{2.35}$$

where C = 2 is the number of classes, t is the expected class and s is the network score for each class among the C.
Finally, a 1x1 convolution, followed by a linear activation function, is used to map the features to the corresponding classes. We used the U-Net with Tensorflow, a CUDA-compatible implementation (Abadi et al., 2016). As an optimizer, we used Adam, i.e., a method for stochastic optimization of a parametrized objective function, with little memory requirements thanks to the fact that it only requires first-order gradients (Kingma and Ba, 2017). The curriculum learning (CL) (Bengio et al., 2009) strategy consists of two training steps of the net: a pre-training and a fine-tuning step. Pre-training is done by assigning a simplified training set, composed of synthetic LWD image logs. Then, the fine-tuning stage was done on more complex training examples. The standard learning strategy, on the contrary, does not require this sorting of training instances according to complexity. The distinction between the training instances is discussed in section 2.2.4 and 2.2.5. We distinguish the network trained with the CL strategy from the same architecture with standard training, hereinafter referred to as CL-PickNet and SL-PickNet, respectively. SL-PickNet was trained for 200 epochs, with a learning rate of 0.01, on the 80% of the 1 million synthetic images original training set, whereas the remaining 20% was used as a test set. CL-PickNet was pre-trained for 200 epochs on 80% of the original training set, followed by 10 epochs of fine-tuning on 50000 complex synthetic images. In our method, we propose to use the Picknet segmentation maps as input to automatically fit the sinusoids corresponding to the segmented edges in the panel. This is the main reason behind the introduction of the curricolar approach. As this whole method is thought to hopefully become a semi-automated BHI analysis tool, we wanted to observe possible impactful differences in the final results for different training strategies.

Fitnet:CNN, Fully-connected layers and dropout

The FCNN architecture, hereinafter referred to as FitNet, was designed to perform the sinusoidal fitting of the segmented edges. It was trained using the binary masks of our synthetic dataset and the set of sinusoids corresponding to the sinusoidal regression of the edges in the images as input and output, respectively (see section 2.2.1). Precisely, the output of FitNet corresponds to the depth coordinates of the points in each sinusoid. Within the training dataset, the range of edges in the input images varies from 0 to 3, prompting corresponding adjustments in target predictions, which correspond to a set of three zero-initialized vectors. As an example, in instances where a singular edge is present, the corresponding target comprises two zero-initialized vectors and one vector housing the sinusoidal coordinates linked to the solitary edge. When two edges are present, the corresponding target comprises one zero-initialized vector and two vectors housing the sinusoidal coordinates linked to the two edges.

This intentional design choice was made to effectively navigate the challenge of accommodating fluctuating edge counts. By doing so, the model circumvents the intricacies associated with training a network constrained by a fixed size for output arrays. The inherent adaptability in predicting coordinates for a dynamic number of edges underscores the model's versatility and its ability to perform optimally across diverse scenarios.
The network architecture includes two 3x3 convolutional layers and a final dense layer (Fig. 2.11, all followed by a ReLu activation function (Agarap, 2019)) and initially combined with a Mean Squared Error loss function defined as

$$MSE_{loss} = \frac{1}{N}\sum_{i=0}^{N}(y_{true} - y_{pred})^2 \tag{2.36}$$

where N is the number of predictions, y_{true} and y_{pred} are the expected and predicted value of the output, respectively.

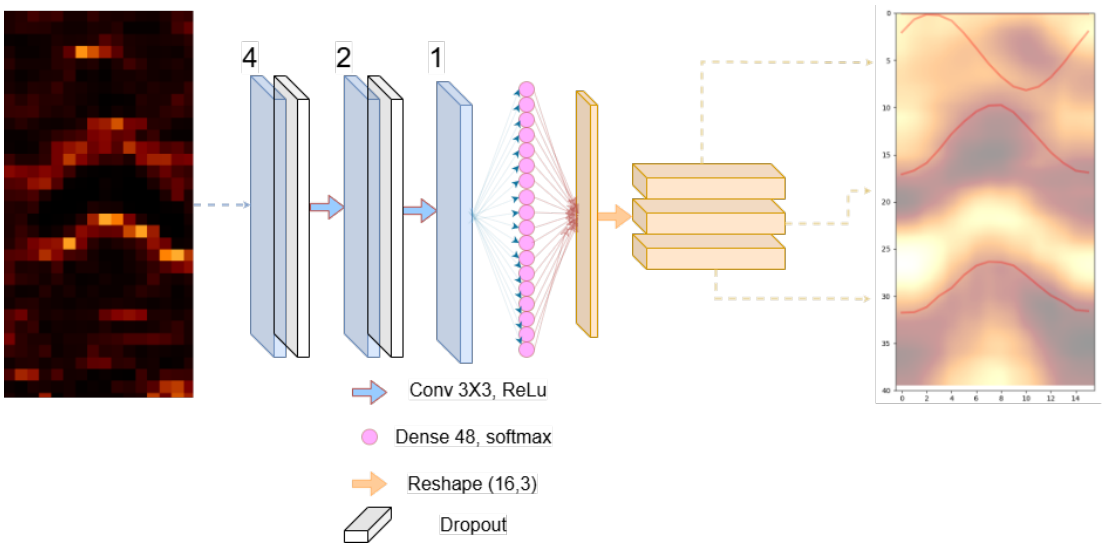

Figure 2.11: Fitnet is a neural network architecture designed for sequence-to-sequence modeling of 2D spatial data. The input is a 2D segmentation map passed through multiple convolutional layers with 3x3 kernels and ReLU activations. The final feature map is reshaped and processed by a fully connected layer with 48 softmax outputs. The final reshaped output is interpreted as a 2D representation of predicted features.

Each convolutional layer is followed by a dropout layer, int he spirit of the Monte-Carlo Dropout (MCD) methodology (Gal and Ghahramani, 2016). Monte Carlo (MC) dropout is a technique used in deep learning, to estimate model uncertainty. It works by applying dropout (i.e., a regularization method that randomly drops out neurons during training) at both training and inference. Normally, dropout is disabled during inference, but in MC dropout, multiple stochastic forward passes are performed with dropout enabled. We considered MCD as an efficient way to estimate the uncertainty of the predicted features, in the hope of extracting valuable geological information at the feature level.

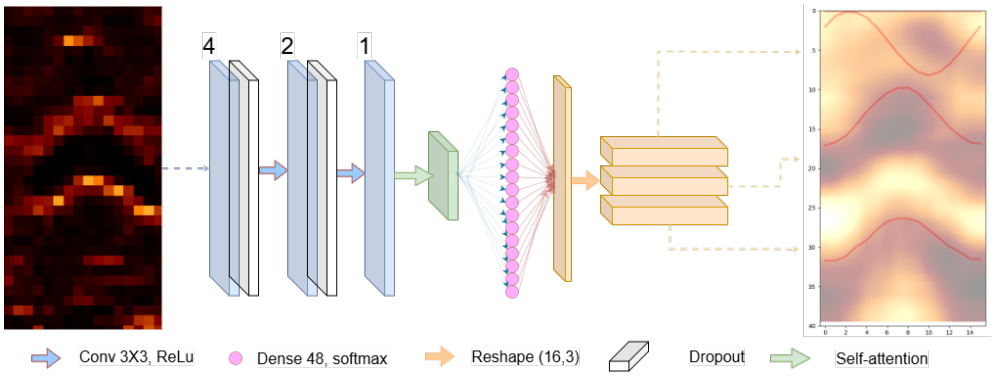

Figure 2.12: Fitnet integrating a self-attention block (green square).

As previously mentioned, the supervised learning paradigm for the problem of automated detection can be problematic, especially for problems related to the sparcity of geological features in BHI (Gupta et al., 2019). It is well known, that machine and deep learning system benefit task-oriented optimization designs (Sun, 2019), even allowing for the formulation of customized solutions (Rani et al., 2022; Zeng et al., 2017; Huang et al., 2020). During training, for reasons that will be later explained along with the results description in section 2.3, we noticed some task-specific design issues, which we heuristically addressed step by step during the analysis of results of both synthetic and field data. One of the key design choices we made, was the introduction of a self-attention block after the last convolutional layer (Fig. 2.12 and 2.13). This mechanism is also inspired by the focued attention observed in humans during their most common cognitive activities (Corbetta and Shulman, 2002; Rensink, 2000), where they conscously steer their attention towards specific objects during a task. The self-attention block receives the input from the last convolutional layer in the network, right before the fully connected part (see Fig. 2.11). This input source data of the attention block is called *key* and compute also a *query*, which is a representation vector of each element in the input at a time. These two are then combined in a multiplicative way (*multiplicative attention*), using the dot-product (Luong, 2015) and a sigmoid activation function that reflects the queries relevance with respect to the keys. Simply, the self-attention block (Brauwers and Frasincar, 2021; Yang et al., 2022; Niu et al., 2021; Hafiz et al., 2021) is added to make Fitnet focus on the relevant parts of its input (i.e., the Picknet segmentation maps) to fit the features.
FitNet was implemented using Tensorflow and trained for 200 epochs, with Adam as an optimizer and a learning rate of 0.01, on 80% of the original training set.

2.2.4 NN Sensitivity

We investigate the sensitivity of both the segmentation and curve fitting networks. By adjusting the tunable parameters that influence the quality and representativeness of synthetic data, we observe how these changes affect network errors. This allows us to make informed considerations about the design of the training process and the data used. In fact, whenever the networks exhibit an increasing

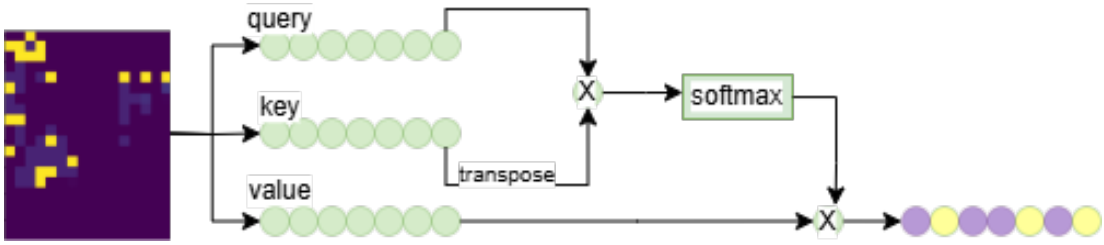

Figure 2.13: Illustration of the self-attention mechanism. The input consists of query, key, and value vectors derived from the input data. The query and key vectors are combined using a dot product, scaled, and passed through a softmax function to calculate attention weights. These weights are then used to weight the value vectors, producing a context-aware output that captures relationships between the input elements.

error due to specific variations in the parameters, it indicates that the complexity of the training examples has been elevated by these changes. For instance, we observe a rise in error as the number of interfaces or the amount of shift increases, particularly in the case of the segmentation network (see Fig. 2.14). This suggests that by having control over the synthetic data, we can begin to consider leveraging the complexity of training examples as part of the training strategy. This opens up the possibility of tailoring the difficulty of the data to optimize network performance. The sinusoidal shift magnitude defines the amplitude of the sinusoid within the panel, which affects the networks' ability to detect high dip magnitude features in real data. As shown in Fig. 5, the Fitnet error actually decreases as the maximum shift of the edges increases. This is likely because, due to weight initialization, the gradient descent function finds it easier to converge with higher maximum shifts. This suggests that, unlike the deterministic approach described in Section 2.2.2, the automated detection of features with high apparent dip magnitudes could potentially yield better results.

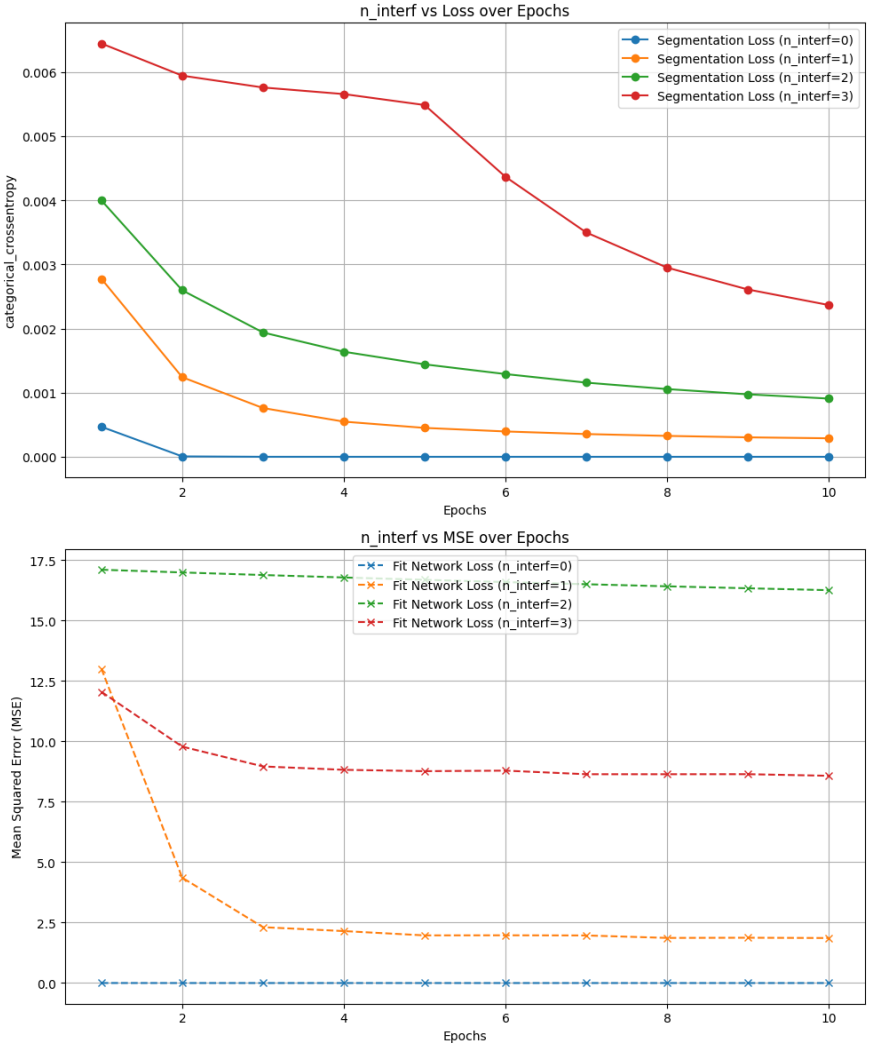

Figure 2.14: PickNet categorical cross-entropy and Fitnet MSE for the different number of interfaces of the synthetic examples.

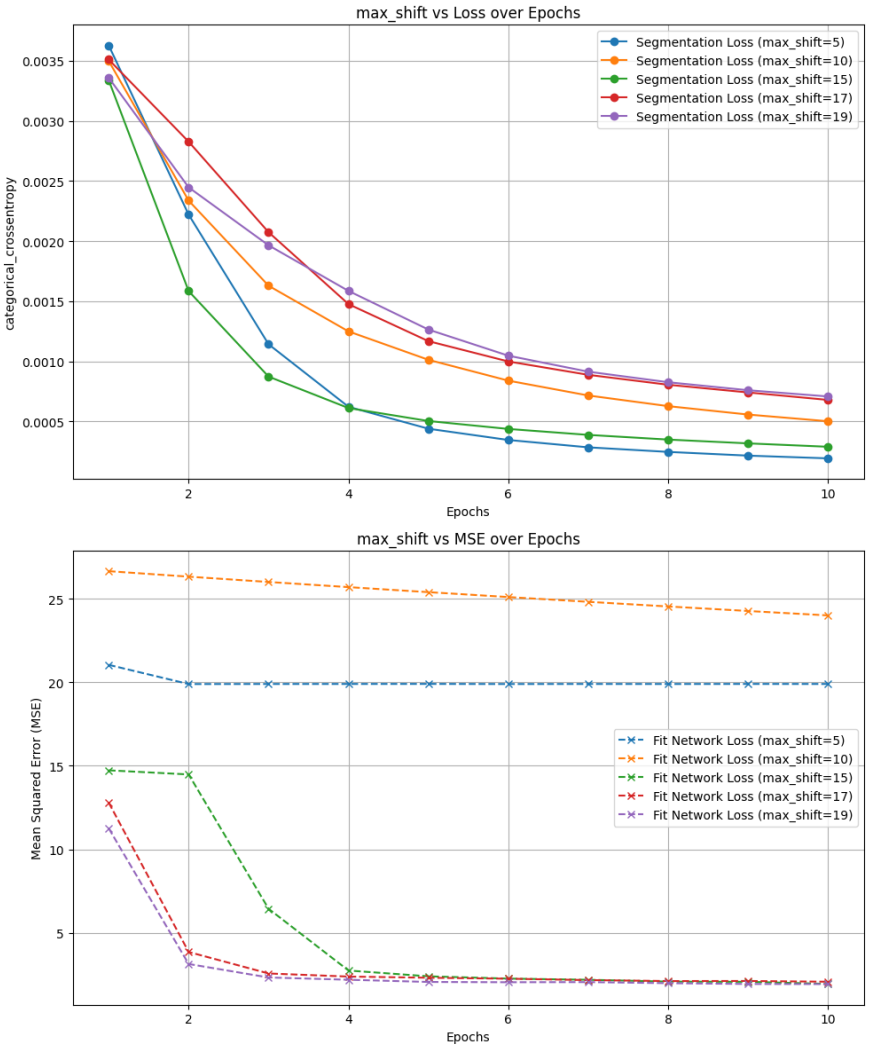

Figure 2.15: PickNet categorical cross-entropy and Fitnet MSE for the different sinuosidal shift in the synthetic examples.

On the contrary, Picknet exhibits the opposite behavior. The categorical cross-entropy loss increases as the maximum shift grows. Similarly to the increasing number of interfaces, in Fig. 2.15 we observe that training examples featuring high apparent dip magnitudes are particularly challenging for the segmentation network. This suggests that such examples introduce greater complexity.

The suitability of the U-Net architecture for automating the analysis of low-resolution images, such as LWD borehole images, is demonstrated by the behavior shown in Fig. 6. In this figure, PickNet exhibits consistent loss function trends, with the exception of the lowest noise level, where the function appears smoother. This highlights the architecture's effectiveness even with noisy, low-resolution data.

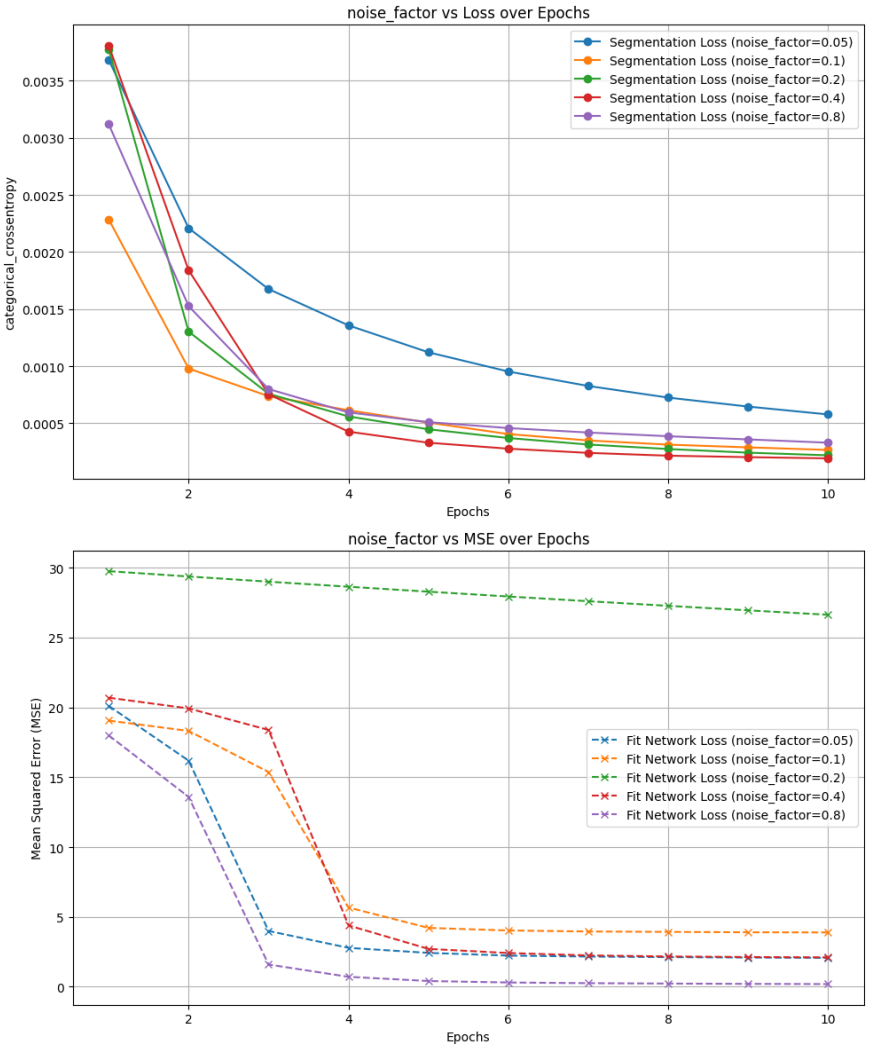

Figure 2.16: PickNet categorical cross-entropy and Fitnet MSE for the different noise levels in the synthetic examples.

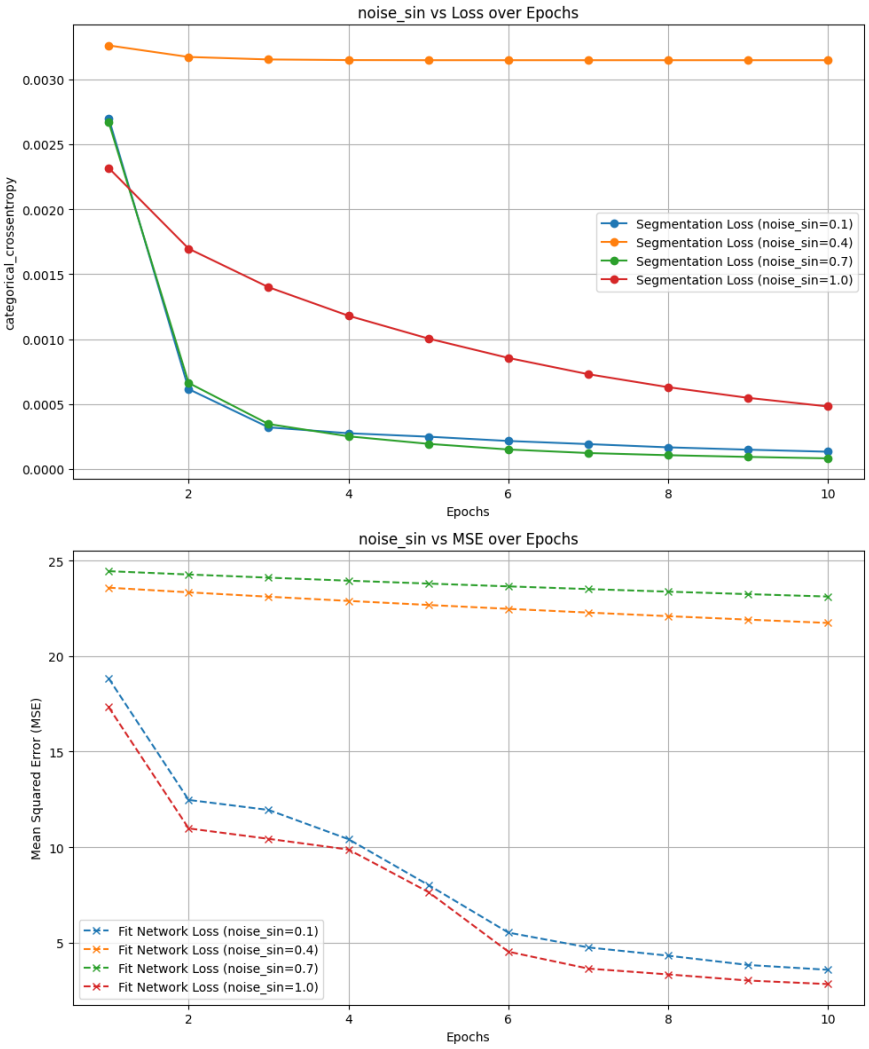

Figure 2.17: PickNet categorical cross-entropy and Fitnet MSE for the different edge irregularities in the synthetic examples.

An additional control parameter for data complexity in neural network architectures is the noise added to the sinusoidal pattern described by the edges. To ensure the training dataset is representative, the features in the image must appear realistic. Since the sinusoidal pattern approximates a 2D representation of a real geological contrast, it's highly likely that in practice, the sinusoidal patterns observed in borehole images do not manifest as perfect sinusoids, but rather exhibit some degree of irregularity. We control this irregularity by introducing noise to the edge sinusoidal pattern, where the noise level is defined by a parameter we set. The roughness is simulated by randomly sampling a float between 0 and this predefined noise level, ensuring that the perturbations to the pattern reflect the natural variability of the geological structures.As shown in Fig. 2.17, the introduced irregularity impacts the neural network optimization performance, highlighting the suitability of this parameter as a control measure for data complexity. However, it is evident that the noise parameter influences Picknet and Fitnet in distinct ways, even though both networks exhibit an increasing trend

in the final loss function as the noise level increases. This suggests that while the noise parameter effectively controls complexity, its effect on different architectures may vary, necessitating careful tuning for optimal performance in each case.

2.2.5 Training strategies

The concept of deep learning (DL) is inherently inspired by the nature of the human brain, and certain training strategies for neural networks (NNs) mirror how humans learn. One such strategy is curriculum learning (CL), which involves a pre-defined order of learning steps, each associated with a different level of complexity (Bengio et al., 2009; Soviany et al., 2022). This approach has proven highly effective in enhancing generalization and improving the rate of convergence across various models, including those in computer vision (Wang et al., 2022b). As discussed earlier, our DL-based method employs a supervised learning approach using meter-scale examples of synthetic LWD borehole images. Sections 2.2.1 and 2.2.4 introduced the synthetic data generation algorithm and examined the sensitivity of the two NN architectures that form the basis of our method. In the latter, we analyzed how NN errors varied with changes in the control parameters of the synthetic data generation, yielding valuable insights into data complexity. This control over complexity serves as the foundation for implementing a curriculum learning strategy. We compare the performance of Picknet and Fitnet under standard learning (SL) practices—without complexity selection—and curriculum learning (CL). Since Picknet's segmentation map is used as input for Fitnet to extract sinusoidal features, this comparison allows us to visually and quantitatively evaluate how segmentation maps produced through SL and CL impact Fitnet's results. Given that the overarching aim of this approach is to develop a tool to assist human operators with this challenging task effectively, we believe offering alternative configurations of the same DL solutions is essential for a thorough understanding of their benefits and limitations. For example, Andrews et al. (2019) found that, in tasks involving fracture analysis in high-resolution data, working in groups was generally beneficial for reducing task-related uncertainty. Specifically, more detail-oriented individuals often influenced less detail-oriented colleagues, improving overall group performance. These findings underscore the importance of diverse perspectives in uncertain tasks. Our tool, however, focuses on the interaction between a human operator and the machine to complete the task efficiently. Drawing inspiration from studies such as Andrews et al. (2019); Shipley et al. (2013); Shipton et al. (2020); Bond et al. (2007), we pose key questions: Can the same NN architecture, trained using different strategies, effectively assist the human operator? To what extent? What level of detail do these strategies provide? The dynamics of human-machine interaction, including their respective errors, are vital concepts to explore, even in more ordinary cognitive tasks such as object recognition in everyday images (Liu et al., 2023). We propose that, to some extent, diversifying data and assistance methods for the same task could be analogous to placing an individual in a heterogeneous group of interpreters with varying expertise and cognitive styles. Ultimately, future studies could investigate

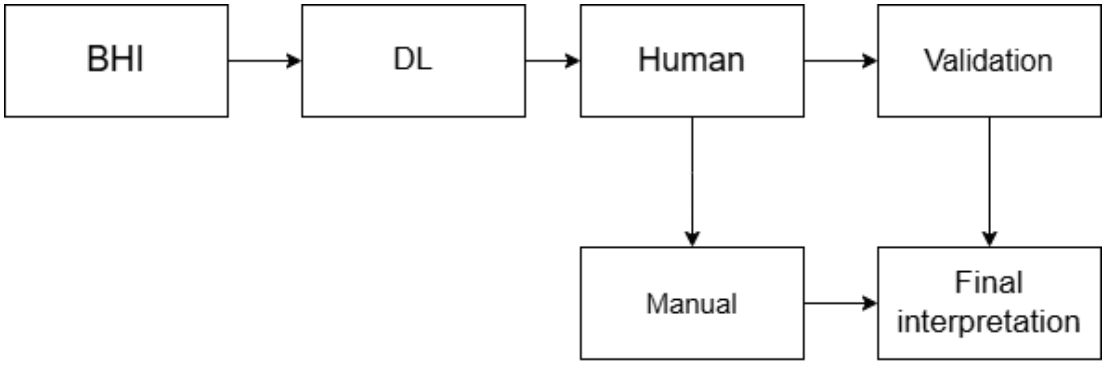

Figure 2.18: Scheme of the proposed semiautomation workflow.

operator uncertainty with and without machine support across various configurations of the same machine. Such studies could provide deeper insights into how humans conduct fracture analysis on low-resolution data and how machine-aided approaches influence their performance and confidence.

2.2.6 Human-Machine collaboration

We conceptualize the problem as a semi-automation one. In this framework, machine predictions are not considered definitive but are subject to human validation (Fig. 2.18), with the aim of maximizing the accuracy of final results. If we assume an individual working alone on an interpretation task, providing assistance via predictions from different training strategies can be seen as analogous to integrating the individual into a team comprising two entities with distinct cognitive styles: SL-PickNet and CL-PickNet.
To address this, we developed a framework for post-hoc validation of FitNet predictions, leveraging segmentation maps from SL-PickNet and CL-PickNet on real-world data. An example of this approach is discussed in the results section. However, we believe that a more comprehensive study requires access to additional data and the organization of a workshop-style event to collect insights from multiple individuals. In this regard, we believe that the DL-methods have already demonstrated a high potential in assisting humans with their inherently uncertain interpretive task. As previously discussed, for a great part of the subjectivity-related problems several options and approaches have already been proposed in literature for the same task (e.g., all the fracture analysis methods described in section 2.1.2), but little has been done in evaluating how the various machine alternatives impact the work of academic and industry expert both individually and in groups. In section 2.3.5, we report some preliminary results of human and machine analysis for comparison purposes, based on the results of manual analysis and of the post-hoc validation. To do so, we created a window-wise manual analysis framework that enabled us to have equivalent settings for the manual and semi-automated BHI analysis. As indicated by the results, this could represent a possible way to identify mechanisms involved in the two approaches to get the final results and gain insights on the human and machine perceptions of the same input. For this preliminary results (Fig. 2.66 to 2.68, we followed

the following steps:

1. manual analysis on sample-wise shifted windows of the original BHI, and for each picked feature we store the window index in which it was correlated, its amplitude and average depth position.

2. post-hoc validation keeping track of the window index, amplitude and mean depth position of each validated feature.

3. calculate center of mass and focal entropy for each input window

4. identify nodes and branches of the resulting fracture set

5. visualize and compare feature depth positions, window index, input center of mass and entropy, spatial distribution of topological elements

To preliminary evaluate the potential of fine-tuning the original NN using the validated features, we kept the NN training active for 1 epoch per each validated feature after step 2 in the workflow described above. Then, we repeat the validation using the fine-tuned NN and observe the differences (Fig. 2.66 and 2.68) in the experimental results.

2.2.7 Scalability test

As introduced earlier in this section, adapting our method for FMI data required designing an approach to fit the data into the NN's required input size. To achieve this, we implemented a straightforward method based on downsampling FMI windows of arbitrary size, feeding the resampled images to the NNs, and then restoring the images to their original dimensions. This process, however, comes with a significant trade-off—an extreme loss of resolution (approximately 98%), degrading the image from a resolution of about $\approx 5\,\text{mm}$ to $\approx 5\,\text{cm}$. For the Conoco Phillips open dataset, we used interpretations provided by company experts as comparison terms to evaluate how well the predictions aligned with human intent. However, previous studies on human uncertainty during fracture analysis (e.g., (Andrews et al., 2019; Scheiber et al., 2015)) suggest that experience is not necessarily a reliable indicator of the quality or detail of interpretation results. While quantitative measures of mismatch between predictions and human interpretations are valuable for assessing the accuracy of automated analysis, they are insufficient to assert that this tool effectively reduces operator uncertainty in this task. Regarding the scalability of the CVML method, we conducted tests on an interval of FMI data from the Netherlands. Despite predictable methodological limitations—such as the blank stripes commonly present in high-resolution FMI data due to the tool's configuration—CVML faced challenges. These blank strips are detected as significant edges by the CVML algorithm, leading to inaccuracies. Furthermore, CVML exhibited an increase in computational time when applied to FMI data, further undermining its practicality and scalability for such high-resolution datasets.

2.3 Results

In this section, we will analyze and discuss the outcomes of our methodologies. First, we will examine the results obtained using the CVML methodology presented in Section 2.3.1, highlighting its key findings and performance metrics. Next, we will focus on the results of the DL-based methodology described in Section 2.2.3, evaluating its effectiveness and potential.

2.3.1 CVML results

As explained in section , the CVML method maps any borehole image to a set of sinusoids automatically analyzed as geological surfaces, by means of a combination of the Sobel and the DTW methods. The sinusoids are found as the sine-curve least-squares fitting of the sets of points related to each edge. This relation is based on maximum similarity calculated with DTW. In order to do this, we then need to define an initial set of points locations for the dynamic calculation of similarity. This set of points includes all the points in the image where the sobel gradient value ecceeds an arbitrary user-defined threshold, as explained in eq 2.16. This implies that the final set of features $\mathbf{F_{CVML}}$ (eq. 2.23) varies, according to the threshold we define. It is intuitive to anticipate a reduction in the number of initial points used for similarity calculations as the threshold increases. Higher thresholds impose stricter criteria for similarity, leading to the exclusion of points that do not meet these more stringent conditions. This results in a smaller set of initial points qualifying for further analysis. We also calculate a confidence measure for each feature in $\mathbf{F_{CVML}}$ as the percentage of Sobel values having the same sign along the trajectory described by the feature coordinates. The main goal of this confidence estimation is to capture some information related to the reliability of each automatically detected feature. Since the Sobel method essentially computes a 2D derivative of the image, the consistency of the Sobel values along the feature trajectory can describe the coherency of the feature in the image, which may reflect the sharpness of the edge approximated by the feature. If this is true, for higher values of the α value defining the threshold T (eq. 2.15) the mean confidence in $\mathbf{F_{CVML}}$ should also increase. We test these concepts on the LWD field dataset introduced at the beginning of this section. Figure 2.19a shows how confidence changes over threshold for all the borehole images in the dataset. We can observe a general increasing of the mean confidence values within the CVML set of features $\mathbf{F_{CVML}}$ for α going from 0.3 to 0.7. All the data show a higher mean confidence for $\alpha = 0.7$. However, the confidence doesn't grow at the same rate for all the LWD BHI and in one case (LWD3), it even shows a slight reduction in the range of α going from 0.4 to 0.6.
The size of the $\mathbf{F_{CVML}}$ set exhibits a significant reduction for higher thresholds, as expected (Fig. 2.19b). The curve number exhibits a power-law relationship with the threshold T across all datasets and demonstrates varying sensitivity to changes in the threshold. This sensitivity is reflected in the estimated slope from the bilogarithmic plot of the number of curves and the thresholds (Fig 2.19c-d). The power-law relationship reveals an underlying self-similar and hierarchical

structure in the CVML feature set with respect to the threshold T, which acts as a refinement parameter. This implies that the feature set associated with lower thresholds includes those from higher thresholds, highlighting the capability of the CVML algorithm to effectively capture the underlying geometrical processes governing the BHI. (Fig. 2.20). The threshold, therefore, appears as a very important parameter in the CVML algorithm, defining the level of structural detail of the outcome. To better understand its effects on the final CVML results, we computed a measure of spatial overlap across the thresholds for each dataset, as a measure of spatial sensitivity. We found a very high overlap for all the datasets, showing values that are approximately equal to 1 (Fig.2.22). A high spatial overlap may reinforce the self-similar feature set hypothesis, confirming that the spatial positions of the feature does remain approximately constant as T changes. This also implies that the refinement process, driven by T, filters the features in F based on certain implicit properties of the features, rather than their spatial arrangement.

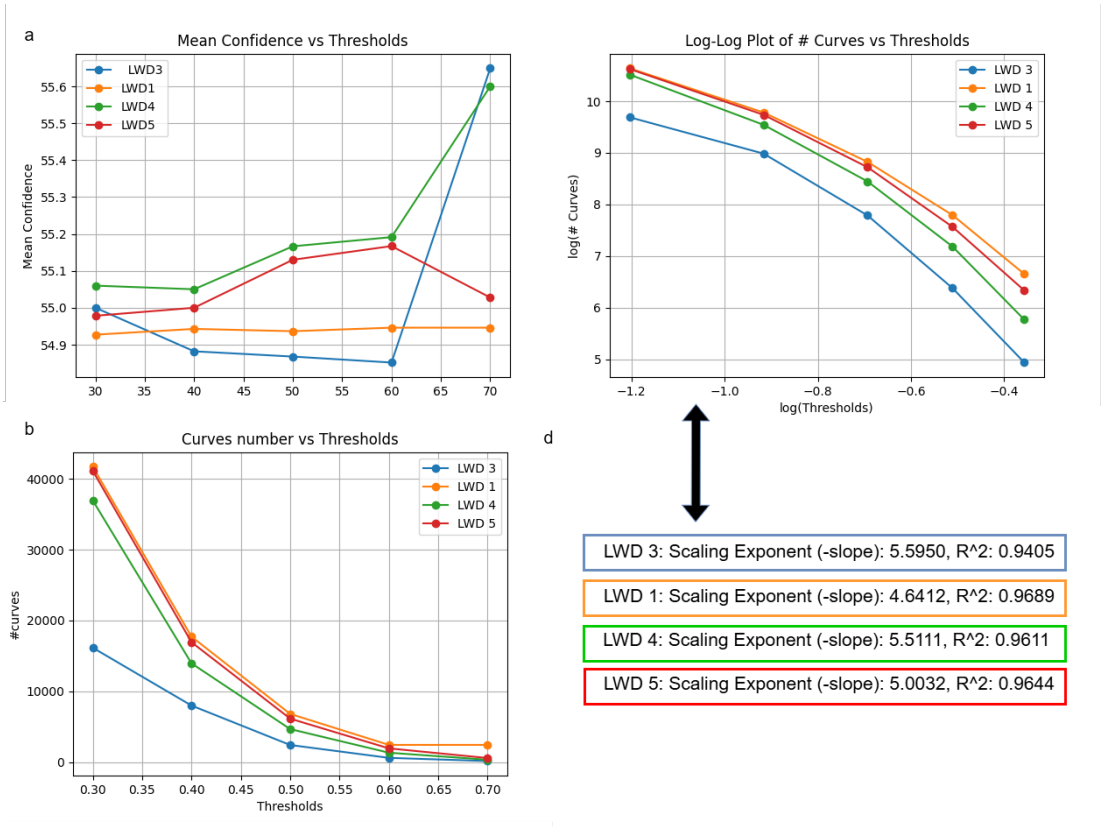

Figure 2.19: The figure shows a) the confidence calculated using the sobel map increasing over thresholds and b) the number of curves in the set F dramatically changing over thresholds, as the results of a smaller number of points in the set P' (eq. 2.20).

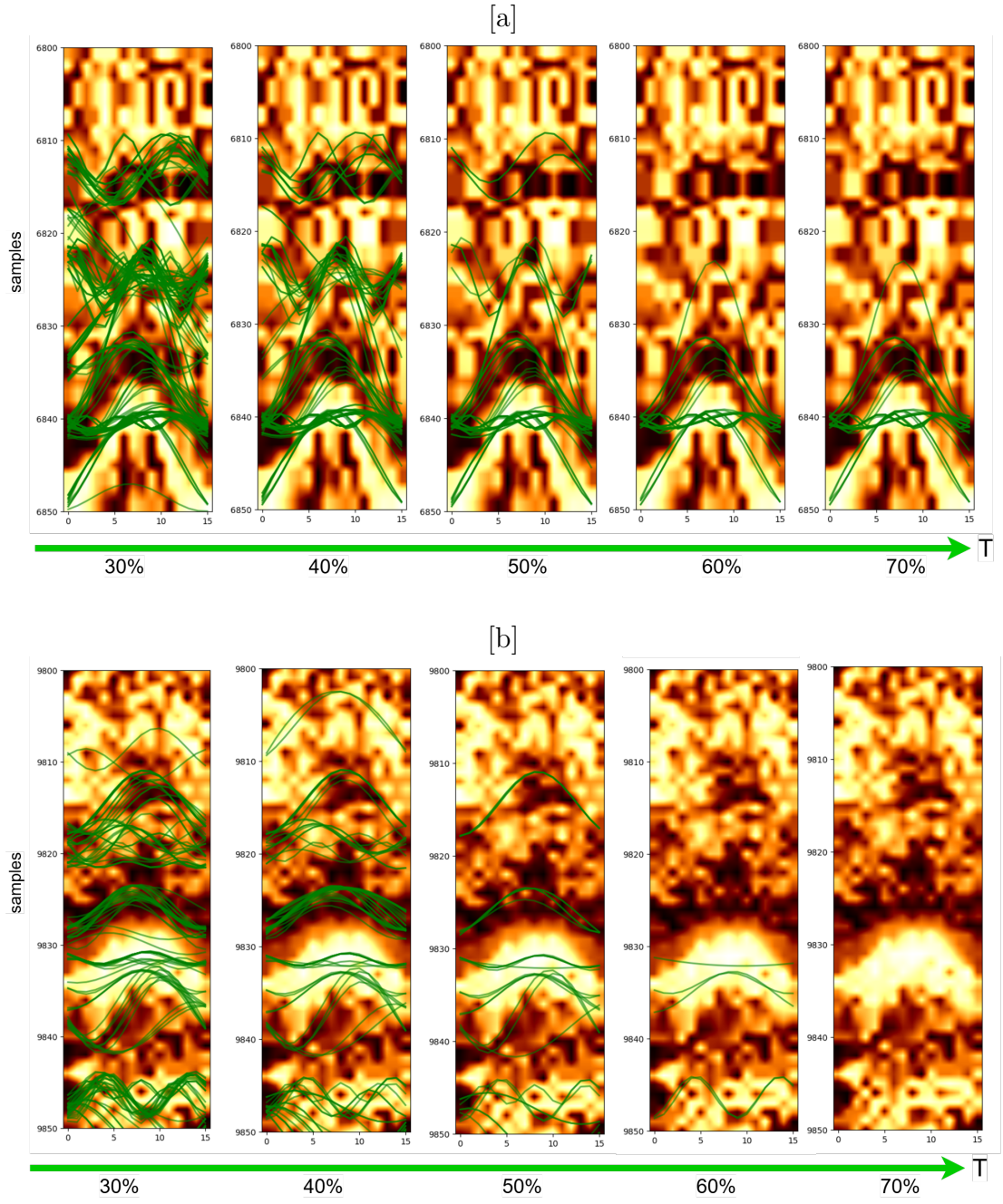

Figure 2.20: The dynamics of the **F** set as the T value changes for different values of α. (a) LWD1. (b) LWD3.

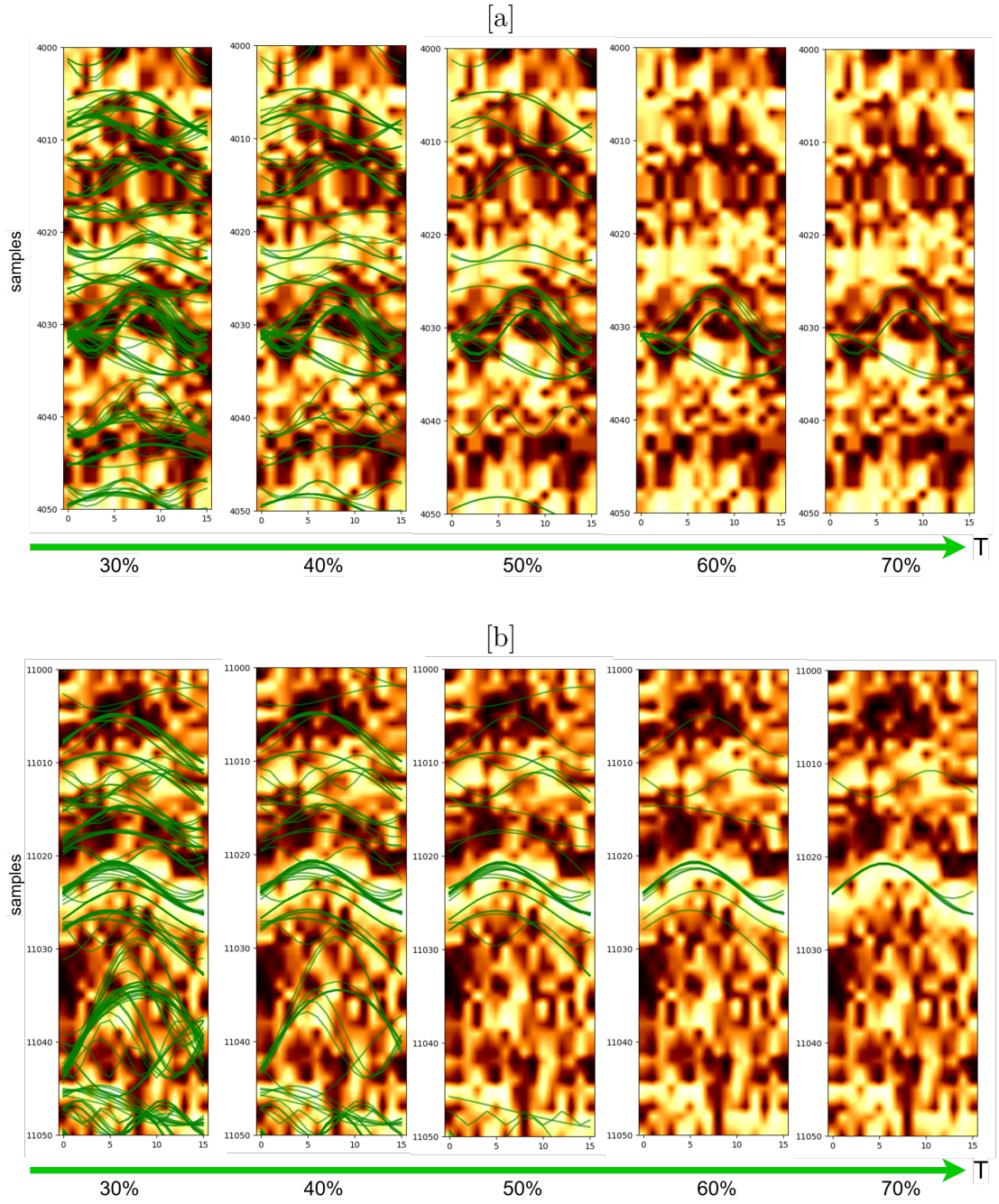

Figure 2.21: The dynamics of the **F** set as the T value changes for different values of α.(a) LWD4, (b) LWD5.

In this regards, we analyze the distribution of BHI values above and below each CVML predicted feature as an initial approach to understanding the implicit feature properties governing the threshold-based refinement process. By calculating the difference ($|\Delta\rho|$) in the mean borehole image values across these features and averaging the results over the entire curve, we can observe how this parameter changes with varying thresholds. This parameter is intended to quantify the difference in petrophysical properties within the sub-intervals divided by

the CVML features. We hypothesize that this difference provides a quantitative representation of the geological class associated with the feature, with features corresponding to bedding planes acting as boundaries between intervals with distinct petrophysical characteristics. In our dataset, we observe that $|\Delta\rho|$ generally increases with the threshold, except in the case of the LWD3 dataset (Fig. 2.23). This confirms that the T-base refinement procedure in the CVML algorithm can provide a mapping function of the original LWD borehole image I (eq. 2.23) with several levels of structural analysis detail and that the importance of the contrasts associated with the features set F progressively grows as α, and consequently T, grows.

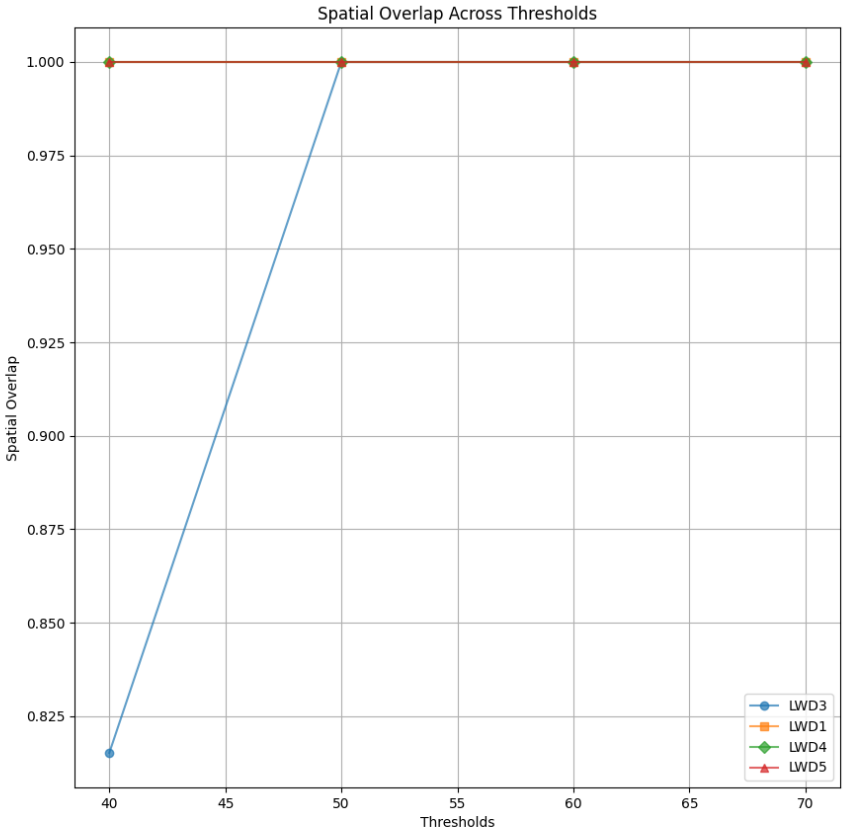

Figure 2.22: Spatial overlap of the curves in the set F across thresholds. The high spatial overlap denotes a refinement process capturing some intrinsic features properties, rather than the spatial distribution.

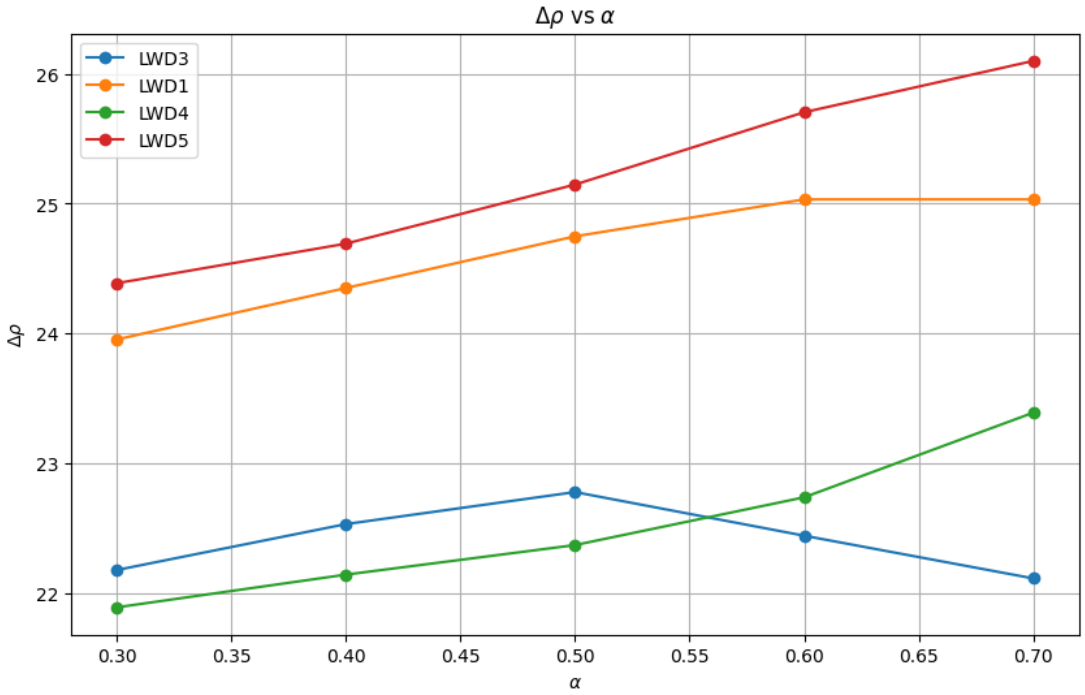

Figure 2.23: $|\Delta\rho|$ as a function of α across the dataset. The figure demonstrates that for higher thresholds the CVML algorithm is progressively focusing on the clearly represented features in the borehole image, normally associated with strong petrophisical changes in the interval.

In general, we can state that the performance of the CVML algorithm varies sensibly with the threshold, as does the computational time (Fig. 2.24). Across datasets, computational time decreases non-linearly as the threshold increases, with diminishing returns at higher values. LWD1 demonstrates a sharp decline, stabilizing at lower computational times as the threshold rises, while LWD3 consistently achieves the lowest times overall. Conversely, LWD4 exhibits the highest computational times, even at higher thresholds, reflecting some dependence on the borehole image log length. These observations highlights the algorithm's sensitivity to dataset characteristics and the need of a trade-off between computational cost and threshold values for real-time application purposes. Table 2.1 summarizes the computational time on a Lenovo Ideapad Gaming 3 machine, along with the threshold value for different α values (eq. 2.15) and the size, i.e., the number of points in $\mathbf{P}'$ (eq 2.20).

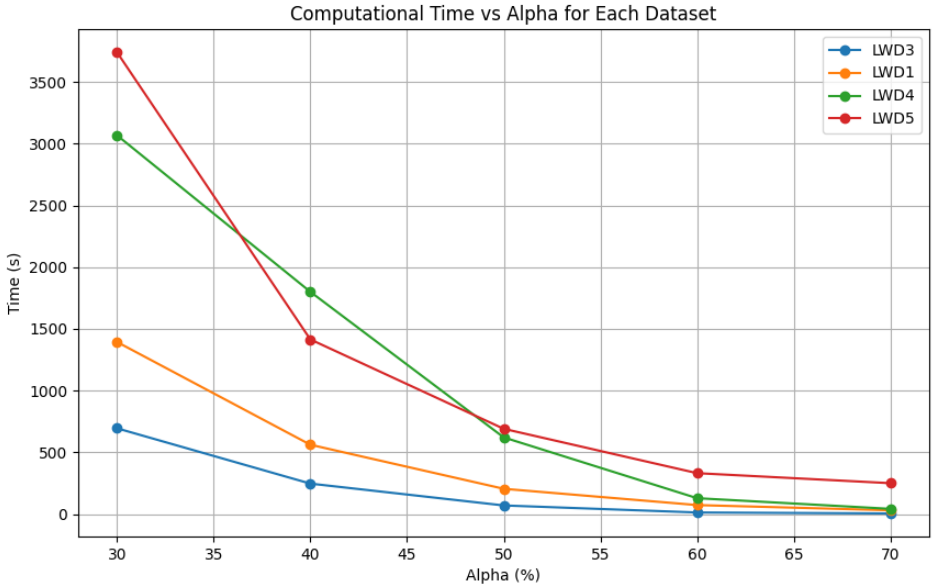

Figure 2.24: Computational time as a function of threshold values (alpha) for four datasets (LWD1, LWD3, LWD4, LWD5). The plot highlights the non-linear decrease in computational time as the threshold increases, with distinct behaviors for each dataset.

Well time [s]	α **x 100**					depth [m]
T/size	**30%**	**40%**	**50%**	**60%**	**70%**	
LWD1	1394,39s	561,61s	204,60s	72,64s	30,59s	728m
T value	109,71	148,28	182,86	219,43	256,0	
P' size	41725	17711	6796	2422	773	
LWD3	694,86s	246,62s	69,61s	12,60s	4,83s	634m
T value	109,91	145,35	181,70	218,04	254,37	
P' size	37116	7981	2406	588	139	
LWD4	3069,00 s	1800,60 s	621,00s	128,42s	42,00s	783m
T value	99,63	132,84	166,06	199,27	232,48	
P' size	36955	13945	4653	1317	319	
LWD5	3742,24 s	1414,82 s	690,63s	90,64s	331,21s	765m
T value	107.14	142.86	178.57	214.29	250.01	
P' size	41154	16933	6130	1929	562	

Table 2.1: Summary of CVML performances for LWD field dataset, considering several alpha values. Each row in the table refers to a borehole in the LWD dataset. For each borehole the table shows the computational time in seconds, the threshold value and the number of points in P' that will be used to retrieve the set of curves F.

2.3.2 DL results

U-Net segmentation results

We analyze the results of the LWD borehole images U-Net segmentation on a batch of synthetic test examples (Tab. 2.2) and, successively, on the field dataset introduced at the beginning of the Results section. Specifically, our segmentation problem consists of a pixel-wise binary classification according to the probability of the pixel belonging to an edge. Considering the borehole images measurement and analysis principles, the edges are associated with approximately sinusoidal patterns representing the dipping geological surfaces. Additionally, this is an unbalanced binary segmentation problem, since the maximum number of non-zero pixels in a 320 pixels training example is 48, given how we structured our synthetic data (section 2.2.1). To overcome the class imbalance we applied a sample weighting function to the synthetic test dataset used for evaluation as in eq. 2.37, so that the weight of each sample corresponds to the reciprocal of its class frequency in the training example.

$$w_{n,c} = \frac{1}{C} \tag{2.37}$$

where C is the class frequency. For both the trained versions of the U-Net, the training time is largely unaffected at approximately 109s for a test synthetic dataset with 210000 instances (Tab.2.2). Note that 210000 examples would virtually correspond to a total of 210 km of borehole images, which makes the execution time of this evaluation procedure extremely low, compared to the CVML execution times (see Tab. 2.1).

	Model	Metrics		
		Cross Entropy	Accuracy	Time
No Pixel Weight	SL	0.2209	0.9592	107s
	CL	0.2046	0.9474	108s
Pixel Weight	SL	0.0067	0.8778	109s
	CL	0.0074	0.7689	109s

Table 2.2: Comparison of metrics with and without sample weight for SL and CL models on 210k synthetic test instances.

The accuracy for both the standard (SL) and the curriculum (CL) learning U-Nets decreases with the pixel weighting function active during evaluation. The accuracy for the SL and CL U-Nets without pixel weighting is very similar at 0,9592 and 0,9474, respectively, whereas the scenario with pixel weighting is associated with larger accuracy differences of the values but sharply lower cross-entropy (see Tab. 2.2 and Fig. 2.25).

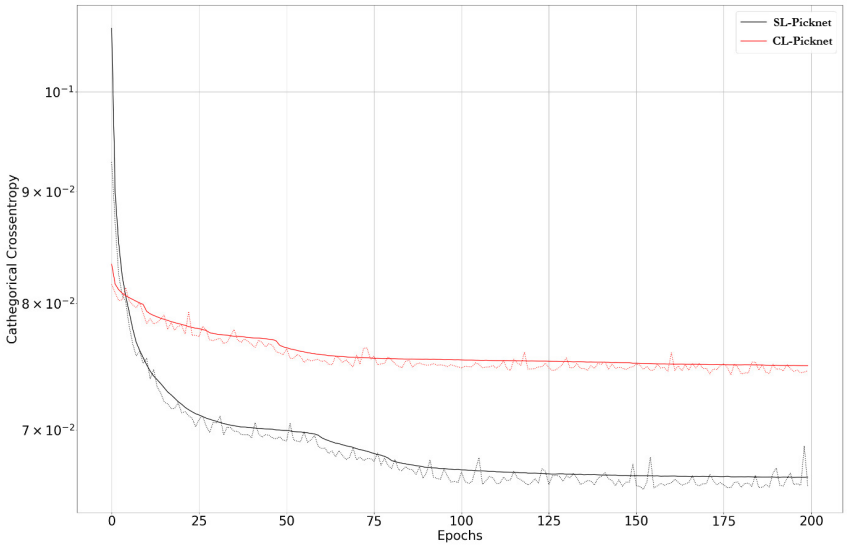

Figure 2.25: Loss functions of SL-Picknet (black) and CL-Picknet (red).

Figures 2.26 and 2.27 show the segmentation results for three different levels of complexity in the test set. Visually, it appears clear that the CL-Picknet exhibits a different behavior compared to its SL counterpart. In particular, it's clear from the predicted segmentation maps that SL-Picknet is more focused on the location of the actual edge compared to the CL counterpart. Moreover, in the synthetic examples where no edges were present, SL-Picknet predicts a weak background noise compared to the higher background noise and curved shapes in the CL-Picknet maps. This explains the lower overall accuracy of the latter. However, considering the most complex synthetic example with two edges, one may conclude that the CL-Picknet segmentation map represent a better result, and the the lower performance level may be due to a greater false positive rate, i.e., the misclassification of zero-valued pixels as positive classes. This demonstrates a higher sensitivity towards the edge class, which, may be a preferable solution in our real-time application problem. The observable differences in the segmentation maps are expected to impact the Fitnet network performance, during the fitting of sinusoids for each segmented edge. For example, SL-Picknet focuses primarily on the most prominent edges in the segmentation map. However, it tends to under-segment more intricate edge patterns, resulting in a loss of finer structural details. On the other hand, CL-Picknet, trained with a more complex strategy, demonstrates superior sensitivity in capturing the intricacies of edge patterns, effectively representing the complexity of the image. However, this comes at the cost of increased higher levels of background noise in the segmentation map.

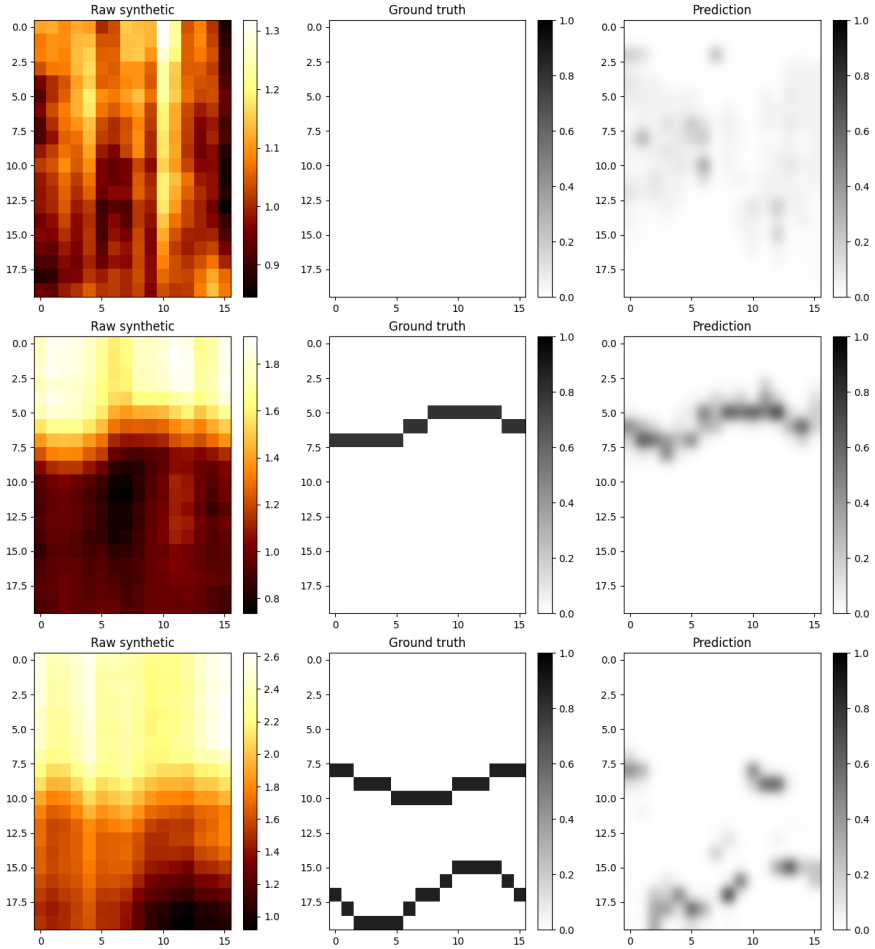

Figure 2.26: SL-Picknet segmentation result on three synthetic examples with increasing complexity.

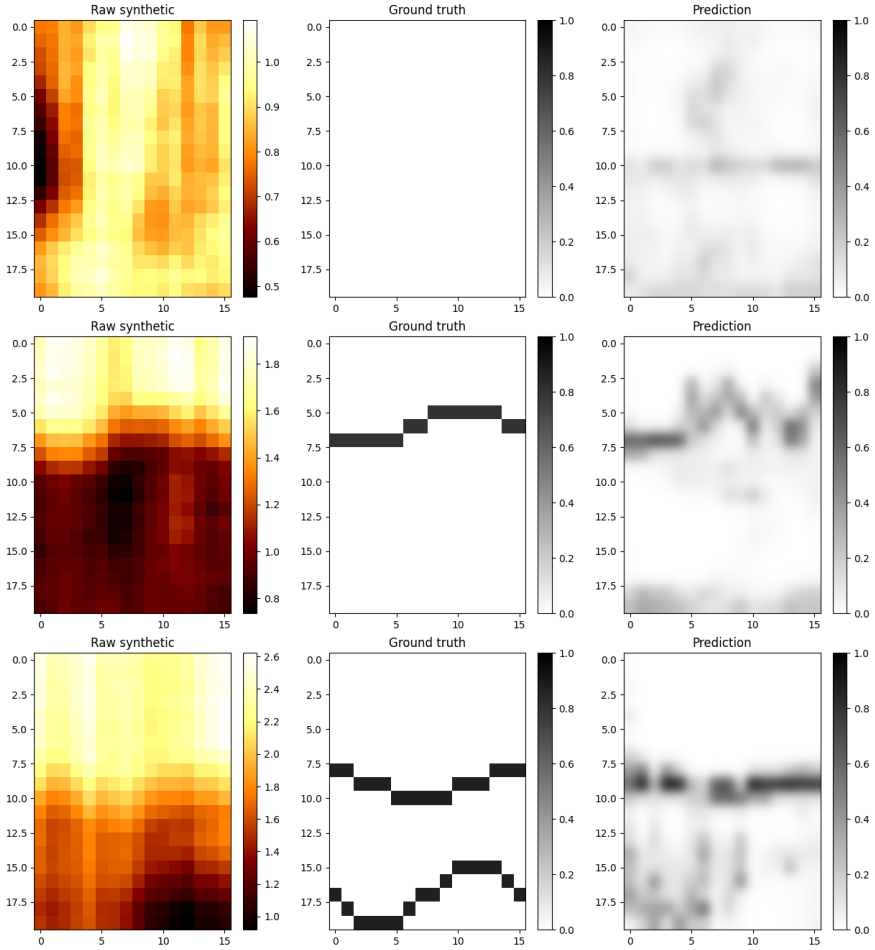

Figure 2.27: CL-Picknet segmentation result on three synthetic examples with increasing complexity. CL-Picknet demonstrates a distinct behavior compared to its SL counterpart. The predicted segmentation maps reveal that SL-Picknet places more emphasis on the exact location of the geological edge, while CL-Picknet shows a broader focus. Where no edges are present, SL-Picknet predicts a faint background noise, whereas CL-Picknet displays more pronounced background noise.

We visually compare the results of both versions on the field data (Fig. 2.28 and 2.29). We see the predictions of the networks in some windows within a 2.5 meters (50 samples) of LWD BHI with strong contrasts, to better appreciate the differences in results. Moreover, we also plot a restored segmentation map of the same interval, that corresponds to the averages of all the prediction windows in the considered interval. From these figures one can claim that both the Picknet versions appear to capture the broad structure of the associated intervals and attempt to identify various orders of detailed structures within them, since in both cases the segmented regions in the segmentation maps tend to extend beyond the stronger contrasts. However, CL-Picknet exhibits a much higher contrasts and sharp transitions in the segmentation maps, that correspond to sharp variations in the input image.
The prediction time for SL-Picknet and CL-Picknet in the LWD dataset is listed

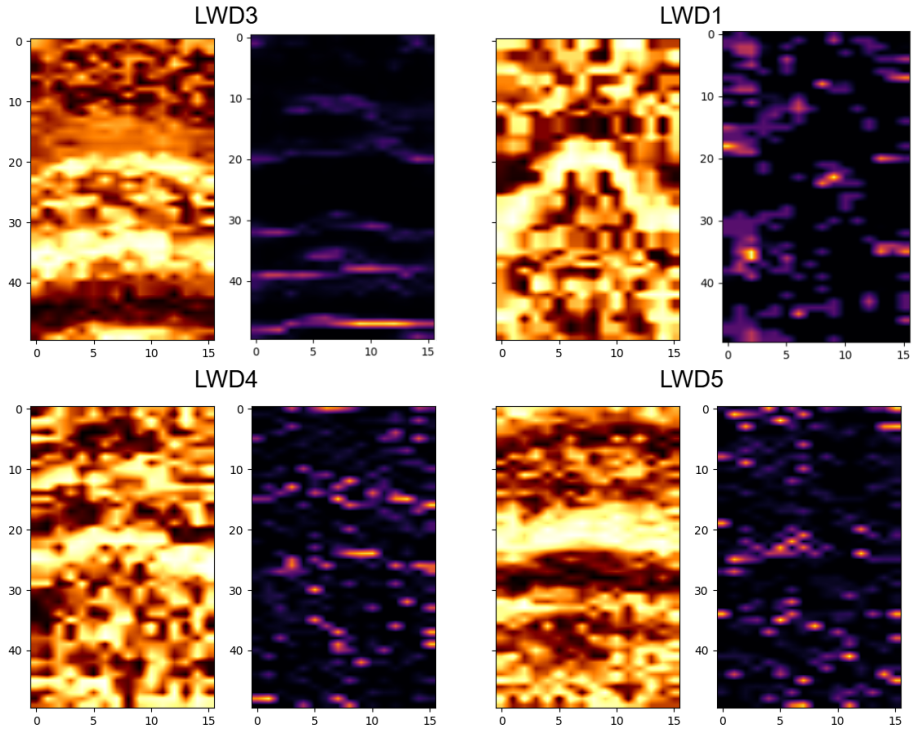

Figure 2.28: SL restored segmentation maps for the LWD borehole image logs in the field dataset

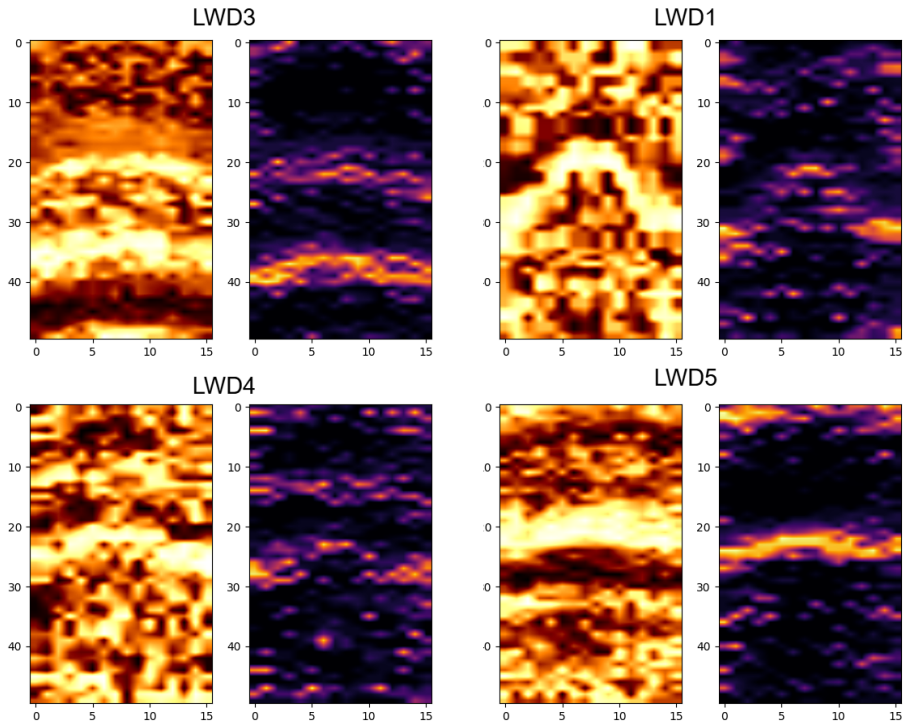

Figure 2.29: CL restored segmentation maps for the LWD borehole image logs in the field dataset

in Table 2.3 and is the same for both of them.

	LWD dataset			
	LWD1	LWD3	LWD4	LWD5
Time [s]	13s	15s	16s	15s
n windows	12684	14721	15678	15319

Table 2.3: Picknet prediction times across the LWD borehole image logs in the field dataset.

These differences in the segmentation results are expected to affect the Fitnet predictions in a way that may, consequently, impact the semiautomated procedure. In the next section, we will describe the differences in Fitnet outcomes using SL- and CL-Picknet results and how it could impact human decisions.

Fitnet results

In our proposed method, the 1-meter interval segmentation maps are used as input in the Fitnet network to automatically retrieve the features planar approximation associated with each segmented edge. The different behaviors of the Picknet versions are likely to have an impact on Fitnet results, which are the final objective of the method. Initially, Fitnet was trained on the whole batch of training data, so that the images that were used as target in Picknet training are now used as input, and the output is a 48-elements vectors, i.e. the reshaped set of 3 vectors with 16 elements, consisting of depth coordinates of the sinusoidal segmented edges in the image. This initial Fitnet version (Fitnet-S) trained using a simple MSE loss function on the whole dataset, reveals important aspects of this standard training strategy with respect to our optimization problem. In particular, the results lead to two considerations. First, the MSE as an optimization metric alone, with no kind of regularization or penalty factor added to it, could represent an ambigous parameter and be somehow misleading the optimization problem towards a non-optimal solution, despite a sharp decreasing of the MSE loss function over epochs during training. This may come from the MSE optimization being carried out comparing all the 48 elements in the dense output vector and the reshaped 48-shaped ground truth vectors and no information or constrain is introduced, related to equivalence of these 48 coordinates values with the set of 3x16 independent vectors of coordinates. Additionally, the results reveal that this DL-based algorithm may benefit a more sophisticated architecture that involves the learning of the non-zero pixel positions and the output values direct relation. With respect to this latter point, we introduced in the network the self-attention mechanism introduced in section 2.2.3. Compared to the more trivial counterpart, this new Fitnet version (Fitnet-S-SA) shows a more stable convergence, but ends with a slightly greater MSE error in the final training epoch (Fig. 2.30).

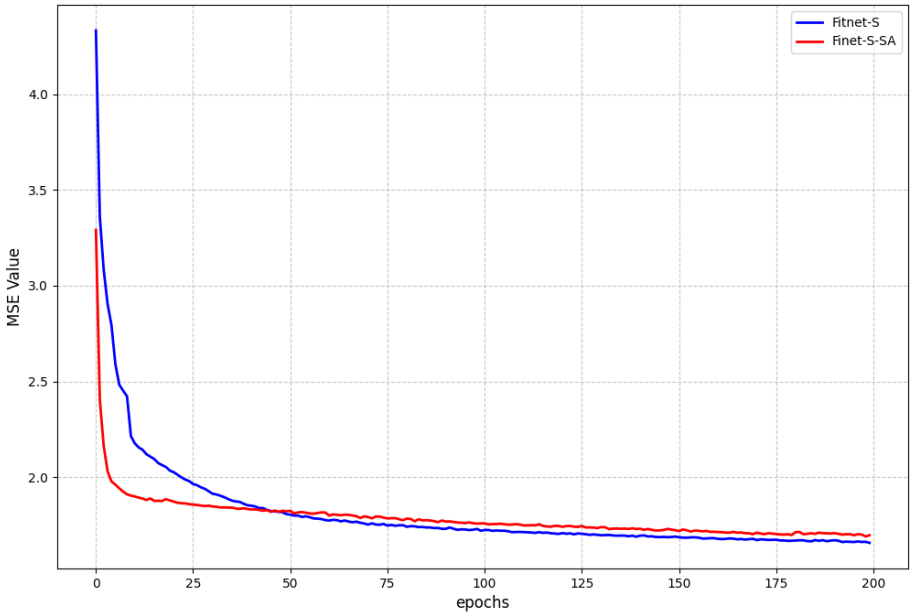

Figure 2.30: Mean Squared Error (MSE) comparison during training for the Standard model and the Standard model with Self-Attention (SA). The plot illustrates the convergence of MSE over 200 epochs.

We also compared the results of this two Fitnet versions, with and without the self-attention mechanism, on test examples with varying number of edges, to observe the results on varying input complexity. For each test example we collect the synthetic BHI and we perform a prediction using both SL-Picknet and CL-Picknet. At this point we use the associated synthetic binary image to collect a baseline prediction associated with an ideal binary output, then the actual SL-Picknet and the CL-Picknet segmentation maps are passed through both Fitnet-S and Fitnet-S-SA with and without the self-attention mechanism. We then compute the Mean Absolute Error (MAE) for the predictions associated with each input panel, and compare the results. Generally, the results are consistent with what the loss functions (of all the Picknet and Fitnet versions described so far) tell us. Figures 2.31a-2.33a and Figures 2.31b-2.33b show the difference between the two versions of Fitnet-S-SA and Fitnet-S on test examples of varying complexity. Fitnet-S is associated with a lower final loss function value, which may appear as an indication of better performance coming from this network, but when we visualize the results on the synthetic dataset, we realize that that is not the case.
On the contrary, Fitnet-S appears quite naive and performs poorly even during the baseline prediction on the ideal binary input, whereas Fitnet-S-SA shows a drastic improvement of the baseline MAE metric. Nevertheless, the example shown in Fig. 2.31a suggests that this enhanced Fitnet version could still be producing sub-optimal or poor results, especially for complex synthetic images. Morever, since this method was initially implemented to be applied for real-time

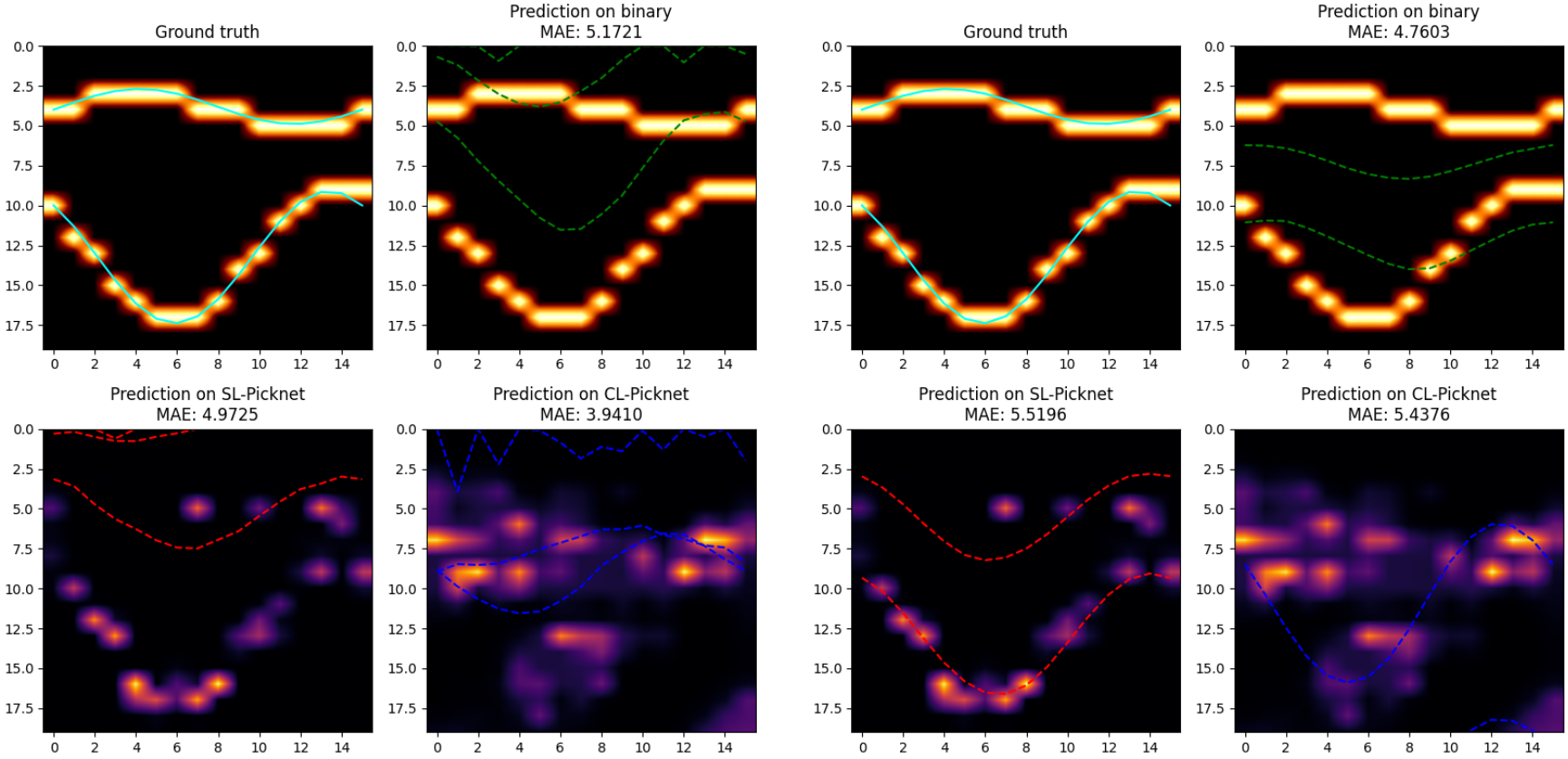

Figure 2.31: Comparison of MSE results with and without Self-Attention (SA).

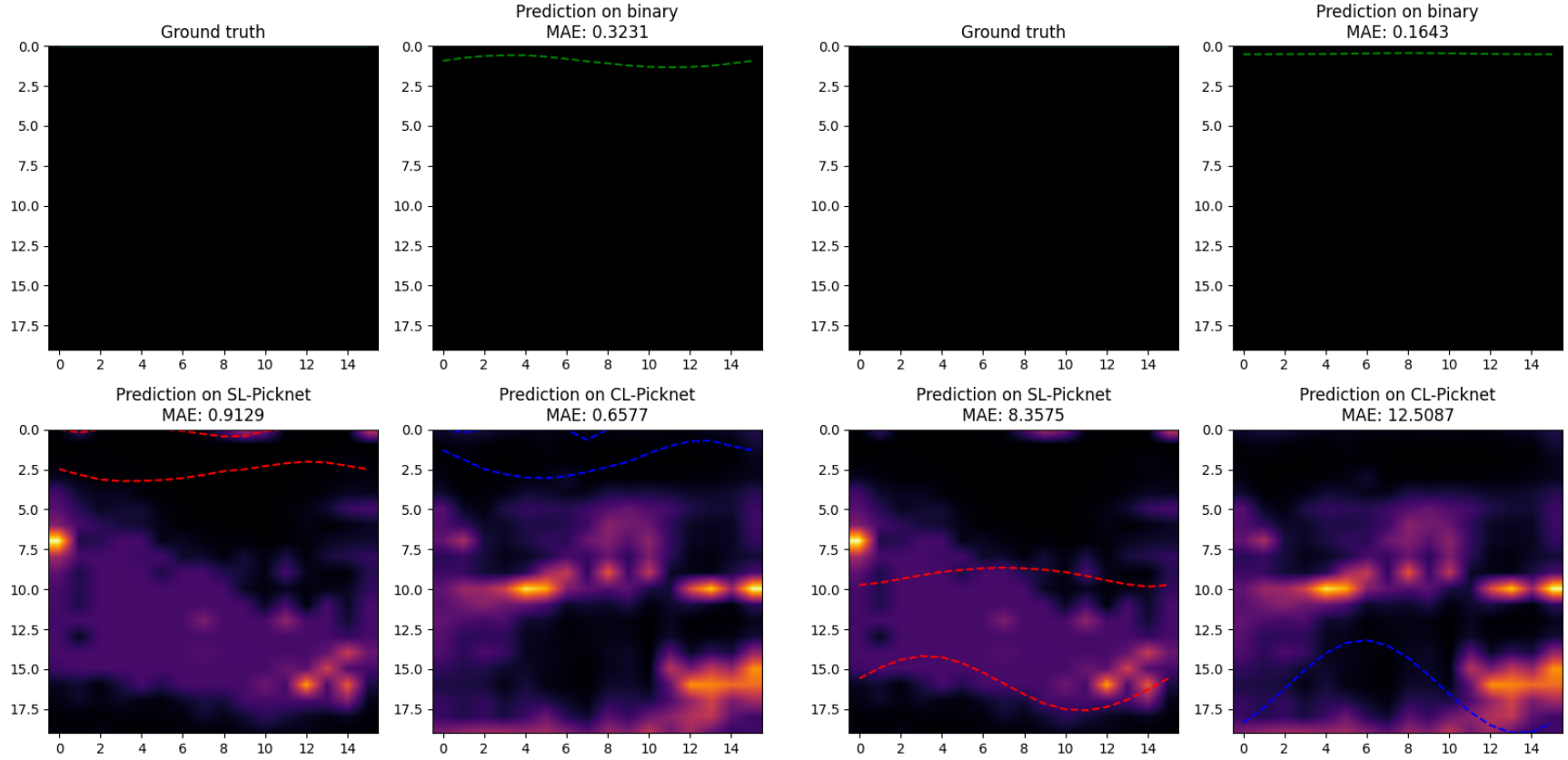

Figure 2.32: Comparison of MSE results with and without Self-Attention (SA).

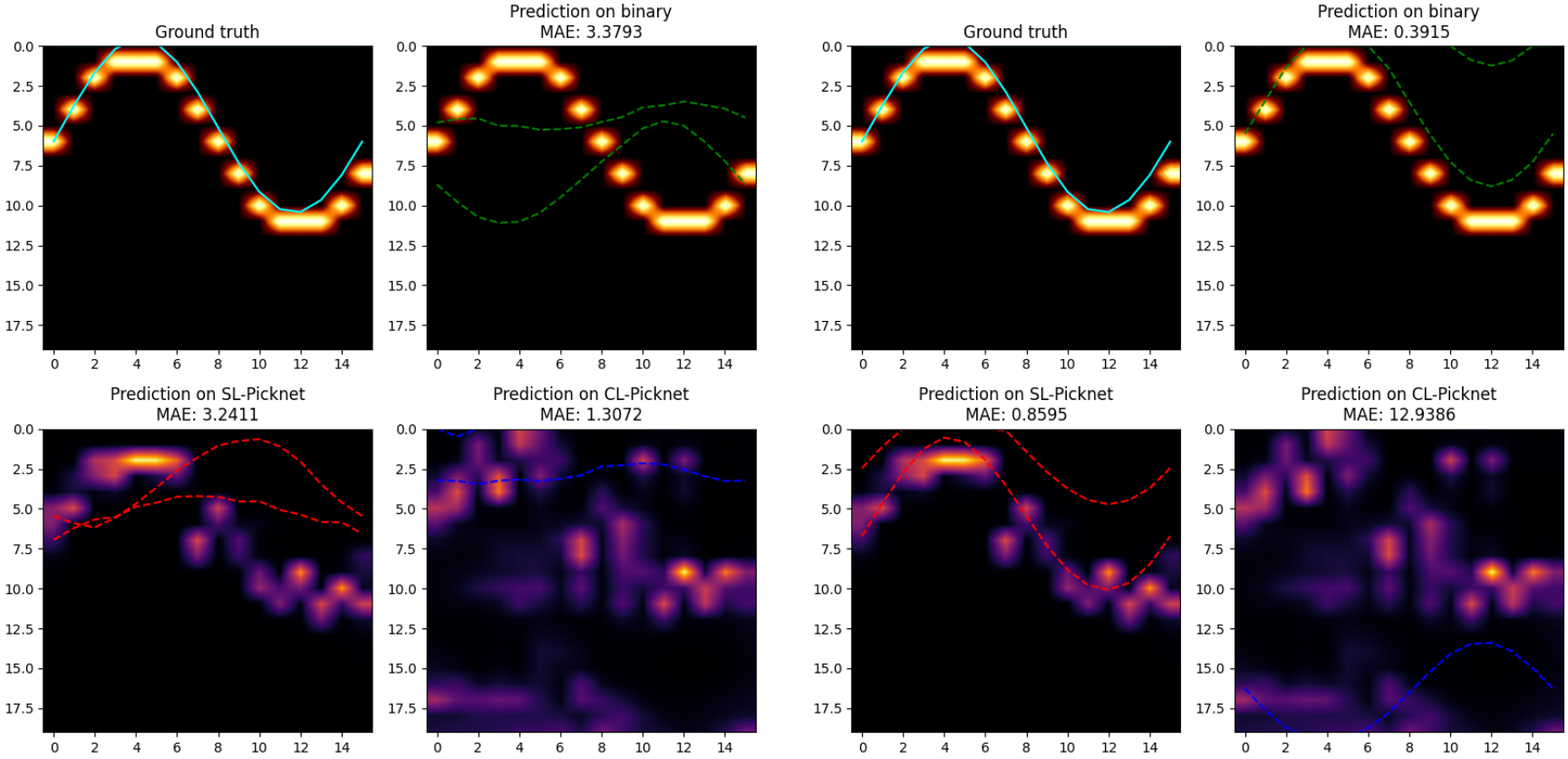

Figure 2.33: Comparison of MAE results: (left) with and (right) without Self-Attention (SA).

automated analysis of low-resolution LWD BHI, the binary input might not be a realistic input to analyze how the whole method is performing. We make Fitnet-S and Fitnet-S-SA predict the set of features associated with the SL-Picknet and CL-Picknet representations of the same binary input used for the baseline prediction. When we compared the SL- and CL-Picknet results on synthetic images (Fig. 2.26-2.27) we noticed that CL-Picknet was associated with an overall lower performance level, but we referred this to a higher sensitivity towards the presence of edges in the input image. We supposed that this would have an effect on Fitnet performance. This effect is particularly interesting in Fig. 2.31 and 2.33. Fig. 2.31a and Fig. 2.33a suggest that for Fitnet-S the CL-Picknet input could represent the best input option to maximize performances. However, the results displayed in Fig. 2.31b and 2.33b, indicate that for an improved Fitnet version like Fitnet-S-SA the SL-Picknet segmentation results may represent the option that produces the closer results to the baseline prediction. This may be caused by the fat that in SL-Picknet segmentation maps the majority of the energy tends to concentrate on one edge at a time, whereas the CL-Picknet segmentation map highlight noise presence, e.g. the observable scattering of activated pixels around the main detected edge. Despite this improvement however, the MSE still seems to be a trivial optimization choice for the challenging problem at hand. We formulated an alternative customized loss function. In this alternative we calculate a $\mathbb{R}^{3x3}$ cost matrix of the reshaped Fitnet output and ground truth (from one single vector $\mathbb{R}^{48}$ to matrix $\mathbb{R}^{3x16}$). We perform an optimal assignment algorithm that determines the rows and columns indices pairs that minimize the total cost of the matrix. The custom loss function is defined as follows. Let $\mathbf{y}_{\text{pred}} \in \mathbb{R}^{B\times 3\times 16}$ represent the predicted values reshaped into a batch of 3×16 matrices, and let $\mathbf{y}_{\text{true}} \in \mathbb{R}^{B\times 3\times 16}$ represent the ground truth values similarly reshaped. For each batch $i \in \{1, \ldots, B\}$, we compute the cost matrix $\mathbf{C}_i \in \mathbb{R}^{3\times 3}$ such that

$$\mathbf{C}_i[j,k] = \|\mathbf{y}_{\text{pred},i}[j,:] - \mathbf{y}_{\text{true},i}[k,:]\|_2,$$

where $j, k \in \{1, 2, 3\}$. Using this cost matrix, the Hungarian algorithm is applied to find the optimal matching indices, denoted $\mathcal{M}_i = (r, c)$, where r and c are row and column indices, respectively. The matching minimizes the total cost:

$$\mathcal{M}_i = \arg\min_{r,c} \sum_{j=1}^{3} \mathbf{C}_i[r[j], c[j]].$$

To penalize deviations in the matching, we compute the squared L_2-norm of the differences between the matched row and column indices:

$$\text{Penalty}_i = \sum_{j=1}^{3} (r[j] - c[j])^2.$$

The loss for the i-th batch is then the sum of the minimal matching cost and the penalty:

$$\text{Loss}_i = \sum_{j=1}^{3} \mathbf{C}_i[r[j], c[j]] + \text{Penalty}_i.$$

Finally, the total loss is obtained by averaging over all batches:

$$\text{Total Loss} = \frac{1}{B} \sum_{i=1}^{B} \text{Loss}_i.$$

We will refer to this new Fitnet version to as Fitnet-S-SA-H. The Fitnet-S-SA-H loss function in this case appears shifted towards higher error levels (Fig. 2.34), but reflects the same stable behavior over epochs as Fitnet-S-SA, thanks to the self-attention mechanism. However, the loss function may again be a misleading parameter to derive conclusions on the performances of the three Fitnet versions implemented so far, since the higher Fitnet-S-SA-H error level might be motivated by the additional penalty factor, through which we try to compensate for the MSE limitations. It is, therefore, necessary to proceed with a visual analysis of the results, as shown in Fig. 2.31- 2.33. In Fig.2.33 we can now see a clear example of how the MSE can be an ambiguous training parameter for our problem formulation, and despite a higher prediction error, the Fitnet-S-SA-H shows better results, especially in this more complex case. Similarly to what observed during the Fitnet-S-SA and Fitnet-S observation (Fig. 2.31-2.33), Fitnet-S-SA-H shows that the better results can be obtained from the SL-Picknet segmentation maps on test synthetic data, which is more similar to the binary baseline with respect to its CL counterpart.

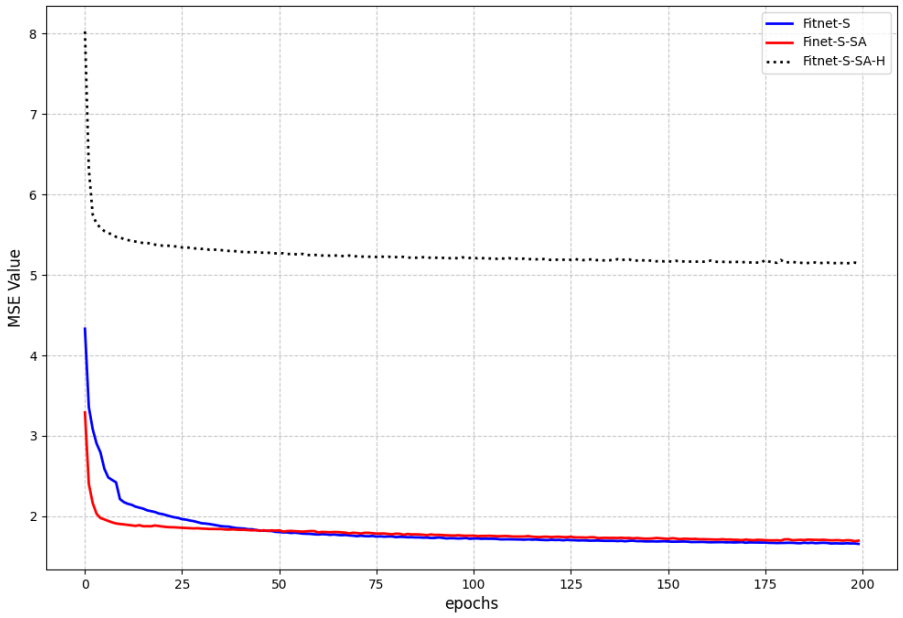

Figure 2.34: Loss function comparison between Fitnet-S, Fitnet-S-SA, Fitnet-S-SA-H

The Fitnet-S-SA and Fitnet-S-SA-H comparison highlighted two main aspects of our proposed method performance on synthetic data. First, our approach benefits from penalizing the MSE loss function with some parameter encoding the information that the $\mathbb{R}^{48}$ output vector is actually to be interpreted as a $\mathbb{R}^{3x16}$ matrix consisting of three sinusoidal features in the image. Second, despite the overall improvement, there is still a clear dependence of the prediction MAE test error on the number of edges to be identified in the input panel. This suggested that the whole approach can potentially be improved by considering a curriculum learning strategy for Fitnet training as well, which involved a pre-training for 200 epochs on a larger dataset 0-edges and 1-edge examples, and a subsequent fine-tuning for 100 epoch on a smaller dataset with 2-edges examples only. We will refer to this new curricular Fitnet version to as Fitnet-C-SA-H and compare the final loss function of both the pretraining and the finetuning step (Fig. 2.38). We can clearly see in the figure that the Fitnet-C-SA-H reaches the lowest loss function values after 200 epochs. Table 2.4 summarizes the final loss value, and training and prediction times for all the compared Fitnet versions.

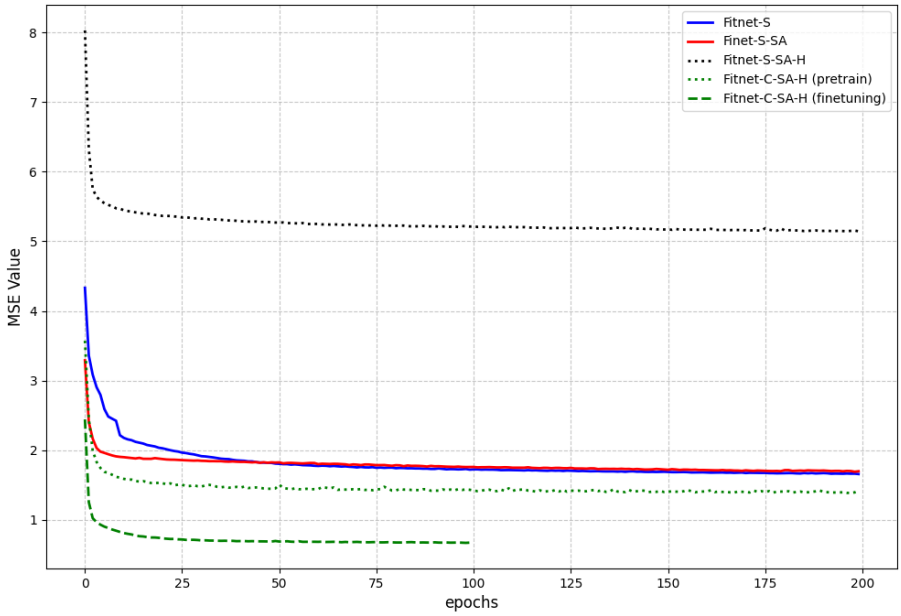

Figure 2.38: Loss functions of all the tested Fitnet options. Note that greater error levels are not necessarily associated with worse performance on the test examples.

Fitnet	final loss	Training Time	Window Prediction Time
Fitnet-S	1.657	777 s	0.0018 s
Fitnet-S-SA	1.697	939 s	0.0019 s
Fitnet-S-SA-H	5.144	11844 s	0.0021 s
Fitnet-C-SA-H	0.672	17874 s	0.0023 s

Table 2.4: Final loss value, training and window prediction time for all the tested Fitnet options. For all the Fitnet tested alternatives, the window prediction time is in the order of milliseconds, outperforming the CVML algorithm by several orders of magnitude.

Fig. 2.39-2.41 show the Fitnet-C-SA-H results on the reference synthetic test examples. The figures confirms the improvements of the Fitnet-C-SA-H especially in the 2-edges examples in Fig. 2.39, which can be appreciated both in visual terms and in terms of MAE metric. It appears clear that combining the curricular strategy with the customized MSE loss function and self-attention represents the best Fitnet results on such complex synthetic test data. The penalized MSE loss function oercomes the ambiguity of having more than one edge in the panel but only one output vector. However, this apparently negatively affects the results of the more simple cases in Fig. 2.40 and 2.41 as indicated by the MAE greater values compared to the Fitnet-S-SA-H ones, which is likely to be related to the greater number of non-zero predicted features by Fitnet-C-SA-H. Another important aspect here, is that this is the test example associated with the most similar predictions on the SL- and CL- Picknet maps. This can be appreciated in Fig. 2.42, where the cosine similarity of the predictions on SL-Picknet and CL-Picknet inputs are plotted over the number of edges in the examples. The figure shows cosine similarity values included in a range between 0.8 and 1 regardless of the number of edges in the panels for Fitnet-C-SA-H and Fitnet-S-SA-H , whereas appear more scattered for the other two, suggesting that this improvement may be a direct effect of the reformulated loss function. In Fig. 2.41, for example, we can see that the single edge is accurately fitted by both Fitnet-S-SA-H and Fitnet-C-SA-H, respectively, using the binary segmentation maps. Nevertheless, the predictions on CL-Picknet segmentation map of the test instances indicate that this version of Picknet can be confusing for any Fitnet alternative, particularly in the less complex instances.

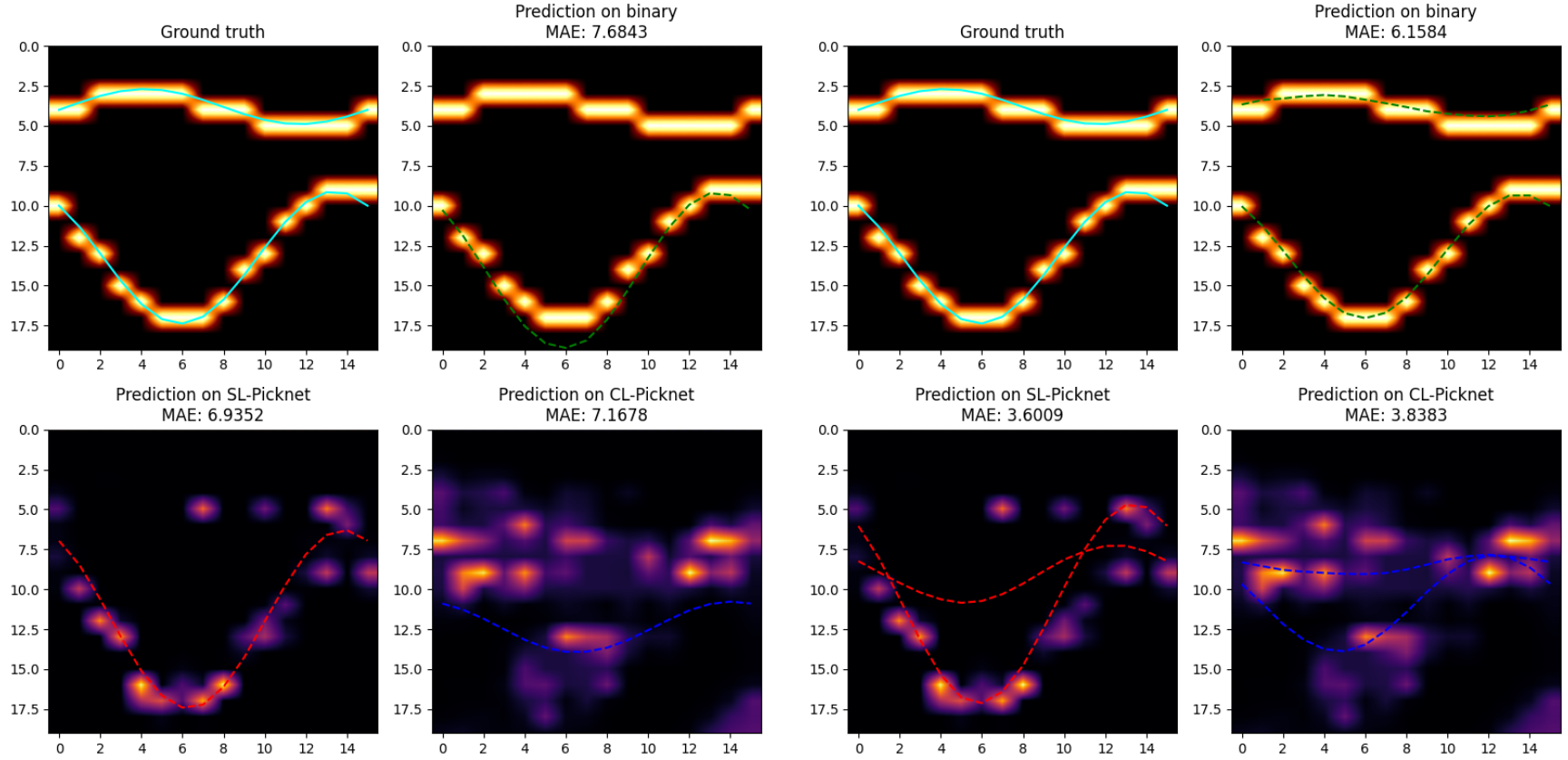

Figure 2.39: Comparison of MAE results: (left) with Self-Attention (SA) combined with custom loss function and (right) with SA combined with the custom loss function and curriculum learning (CL).

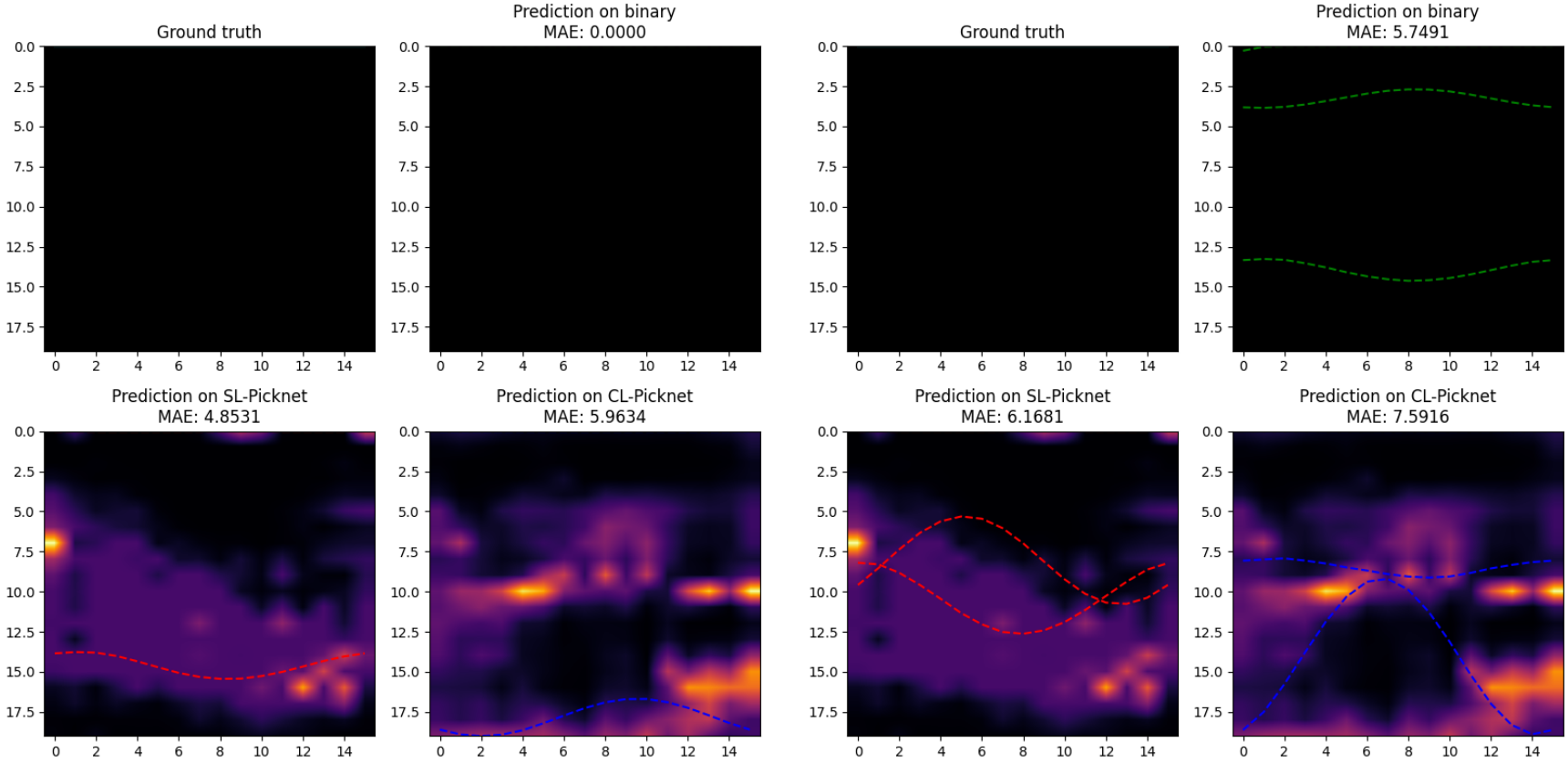

Figure 2.40: Comparison of MAE results: (left) with Self-Attention (SA) combined with custom loss function and (right) with SA combined with the custom loss function and curriculum learning (CL)

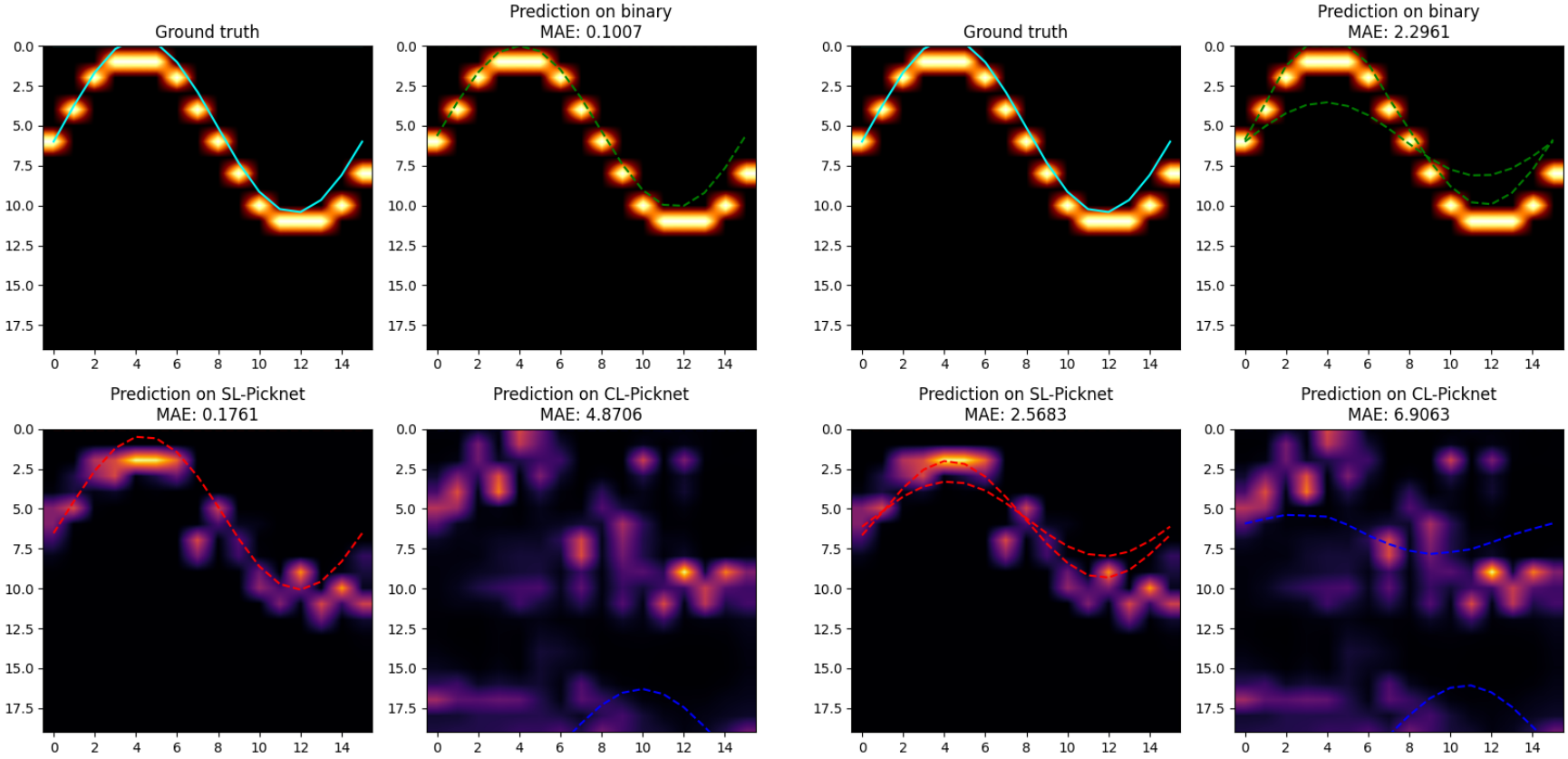

Figure 2.41: Comparison of MAE results: (left) with Self-Attention (SA) combined with custom loss function and (right) with SA combined with the custom loss function and curriculum learning (CL)

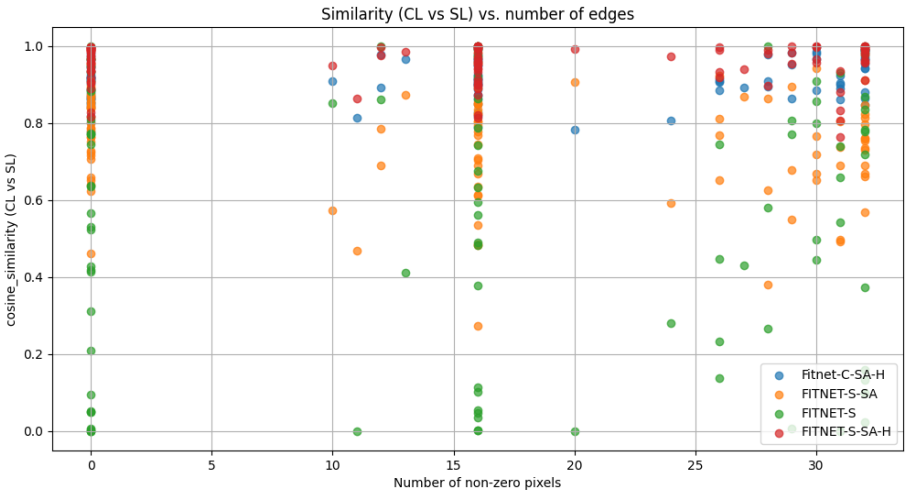

Figure 2.42: Cosine similarity vs number of non-zero pixels in panel for the four Fitnet versions. The Fitnet version with SA, custom loss function and curriculum learning (CL) (Fitnet-C-SA-H) shows maximum similarity, regardless of the number of edges in the input image.

It is interesting to note that a common observation amongst all the compared Fitnet versions and their performance on all the possible input representations is that, excluding the baseline binary panel, the SL-Picknet segmentation map is apparently the preferable segmentation map to be later fed in Fitnet to quickly and accurately find the trace of sinusoidal depth coordinates that correspond with the edges segmented in the panel. However, every DL method that was trained using synthetic data, in order to be fully validated needs to be evaluated on field data. Normally, field data are more complex and challenging than synthetics, which might cause the failure of such DL systems in real case scenarios. Especially in the problem we try to address with the proposed DL-based method, field data are characterized by high variability determined by site-dependent geological configurations in the subsurface and instrumental artifacts that can easily arise and corrupt the data in the highly dynamic and disturbed environment of a borehole, in particular in LWD. Moreover, we conceptualize our method in such a way that it could detect a maximum number of three edges per input image and in a real-time fashion. To simulate such an application, we applied the method on field data differently than we did for the synthetic test examples discussed in this section. We propose a sample-wise shifting window technique to simulate the the method predictions collection at each new sample in the borehole. Intuitively, we expect this to generate a redundancy in the predictions for each meter of LWD BHI, which is why the validation process is such an important component of the proposed approach. In the next section, we will describe and discuss the results of the Fitnet alternatives presented in this section on the LWD BHI dataset, using both the SL- and CL-Picknet segmentation maps.

2.3.3 Field data

Differently from synthetic data, field data may present features that impact on both the SL- and CL-Picknet performances and, consequently, on Fitnet. For instance, the real geological surfaces may have a wider range of variability in terms of scattering around the main sinusoidal trace of interest, or be associated with a wider range of contrasts. This is likely to produce different results from the synthetic data scenario, which can eventually lead to different conclusions, especially with respect to the real-time or near real-time applicability of the proposed method, which is the main objective for applications in the LWD domain. For consistency in the presentation of the results, we will now consider the same LWD BHI intervals as in Fig. 2.26-2.27. We visualize and compare the Fitnet-S and Fitnet-S-SA results using the SL- and CL-Picknet segmentation maps. Note that the manual interpretation, that is often used in these cases as the ground truth, of the field data used here was unavailable. In this paragraph we will consider LWD BHI intervals of 50 samples (2.5 m). Each interval will be binned in 50 prediction windows with 20 samples, using a sample-wise sliding approach to perform predictions as anticipated in section 2.3.2. This marks a very important difference with respect to what we observed in the section where we observed the results on synthetic data. Every meter (20 samples) in the field borehole image log is then binned into 20, and each sample row is "seen" 20 times by Fitnet. This is also expected to produce redundancy in the predictions, since each edge, or at least a part of it, in the prediction windows is likely to be seen multiple times by Fitnet. In the following paragraphs, we will visually appreciate the different results of the discussed Fitnet options on the same LWD BHI dataset introduced at the beginning of section 2.3. The results are summarize in Tab. 2.5 and presented as follows: first, we summarize the results of all the discussed Fitnet versions by calculating the cosine similarity of predictions on the CL-Picknet and SL-Picknet segmentation maps, as we did in Fig. 2.42. This parameter reflects the robustness of the Fitnet versions on field data, capturing the amount of variation of the predicted features on the different segmentation results. Second, to overcome the lack of ground truth, we compared the predictions of every Fitnet version with the CVML results in the same reference intervals. Specifically, we start by randomly chosing a CVML feature as a comparison term, which we will refer to as the CVML *marker*. Note that we refer to the set of CVML features corresponding to the maximum threshold value of 70% (section 2.3.1). We calculate the prediction window at which each marker starts to be included in the window itself and therefore is visible from Picknet and Fitnet perspective. This value is found under the marker column in Tab. 2.5. Finally, we compare the prediction window number at which we find the closest Fitnet prediction to the marker for the SL-Picknet and CL-Picknet segmentation maps. We believe that this can lead to important considerations on the various discussed Fitnet versions and their interaction with Picknet, especially with respect to the real-time applicability of the proposed method on field data, its potential and limitations. Each reported result in Tab. 2.5 refers to the Fig. 2.43-2.50. For the sake of readability, each one of those figures is discussed under the subsection entitled with the name we gave to the associate Fitnet version.

Fitnet	LWD	Metrics			
		sim(CL,SL)	marker	SL window	CL window
-S	1	0.1398	1	11	13
	3	0.1601	22	31	33
	4	0.2634	12	26	27
	5	0.2629	7	14	22
-S-SA	1	0.2661	1	11	6
	3	0.3599	22	34	22
	4	0.4321	12	26	9
	5	0.3708	7	20	11
-S-SA-H	1	0.7146	1	11	6
	3	0.7118	22	34	27
	4	0.7829	12	21	17
	5	0.7163	7	18	9
-C-SA-H	1	0.8306	1	5	2
	3	0.8099	22	31	24
	4	0.8324	12	21	18
	5	0.8584	7	18	10

Table 2.5: Comparison of cosine similarity between the Fitnet predictions on CL-Picknet and SL-Picknet segmentation maps for each Fitnet version.

Fitnet-S As the simplest of the Fitnet versions we trained, Fitnet-S is the one who appears to be less robust, according to the cosine similarity of the predictions on SL- and CL- Picknet segmented panels. The impact of the segmentation results is particularly visible in Fig. 2.43b and 2.44b. Both these figures show how the higher CL-Picknet sensitivity is highly noisy and afects Fitnet-S ability to fit the edges with the same smoothed and clean sinusoids we see in Fig. 2.43a and 2.44a. However, even in the latter figures the results cannot be considered satisfactory, since, at least for Fig. 2.44a, where the marker can be considered as a less questionable comparison term, there are no Fitnet-S predictions that approximates the marker itself.

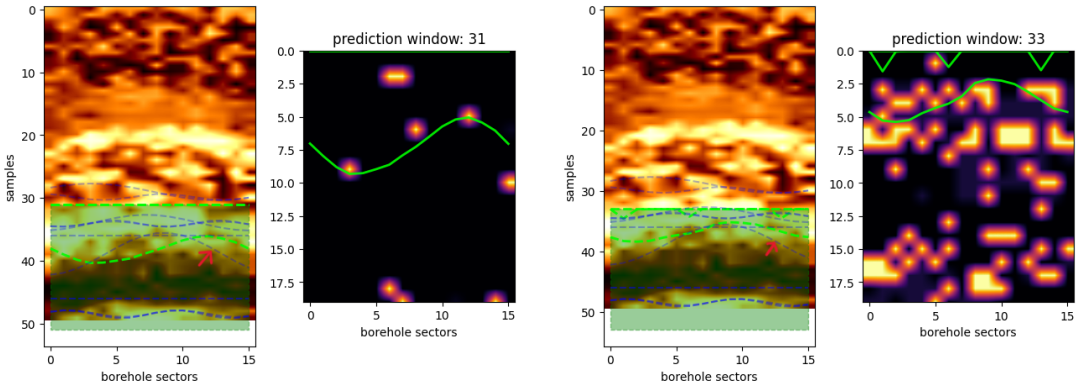

Figure 2.43: Fitnet-S on left: SL-Picknet segmentation map; right: CL-Picknet segmentation map. The red arrow indicates the randomly selected marker (blue line) (blue line) among the CVML predictions in the LWD3 interval.

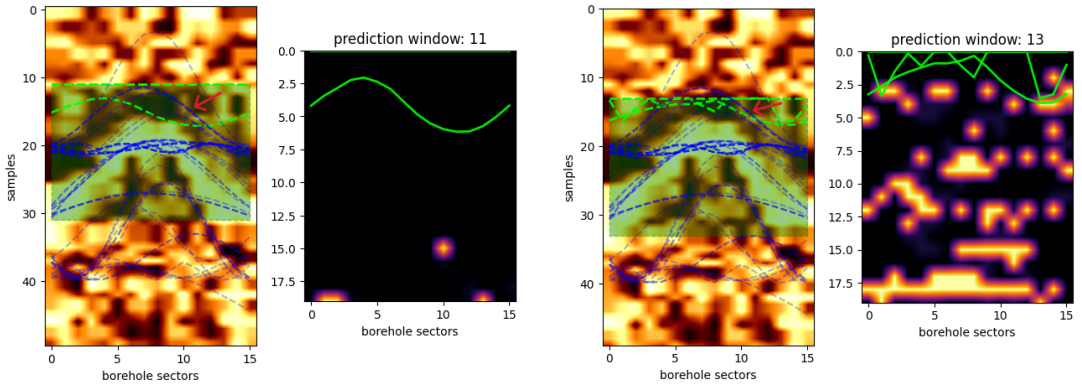

Figure 2.44: Fitnet-S on left: SL-Picknet segmentation map; right: CL-Picknet segmentation map. The red arrow indicates the randomly selected marker (blue line) among the CVML predictions in the LWD1 interval.

Fitnet-S-SA Fitnet-S-SA is te version who was trained in combination with a self-attention mechanism in the attempt to improve the results. This seems to be the case, if we observe Fig. 2.45-2.46. The predictions, in fact, now appear to be more similar t the CVML marker in both the SL- and CL- Picknet results (Fug. 2.45a-b). However, there are some clues in this predictions examples that are worth noting. Differently from the Fitnet-S case (section 2.3.3), this time we reach the closest prediction in shallower prediction windows. Moreover, in Fig. 2.45b, one may notice that the Fitnet-S-SA prediction we chose is partially exceeding the prediction window. This might represent an advantage of the CL-Picknet segmentation map, in terms of real-time or near real-time applicability, as it certainly suggests that this segmentation map somehow provides more information than its SL counterpart, and this information can be focused closer to the lower boundary. Despite the improvement, Fig. 2.46b still indicates some undesired errors, such as the sawtooth-like features predicted at the top boundary of the segmented panel.

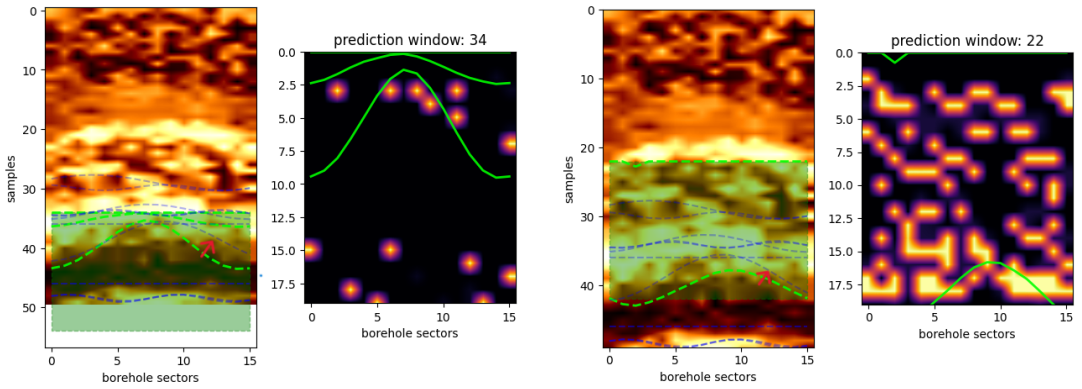

Figure 2.45: Fitnet-S-SA on left: SL-Picknet segmentation map; right: CL-Picknet segmentation map. The red arrow indicates the randomly selected marker (blue line) among the CVML predictions in the LWD3 interval.

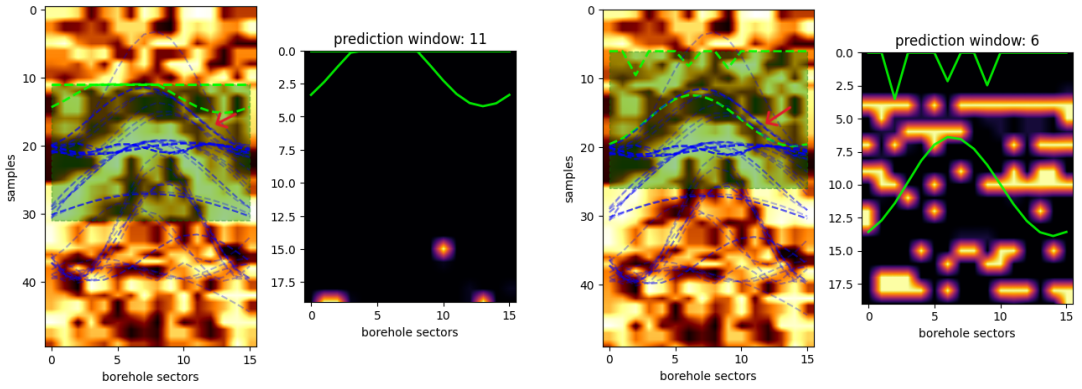

Figure 2.46: Fitnet-S-SA on left: SL-Picknet segmentation map; right: CL-Picknet segmentation map. The red arrow indicates the randomly selected marker (blue line) among the CVML predictions in the LWD1 interval.

Fitnet-S-SA-H Observing the Fitnet-S-SA results on synthetic data (section 2.3.2), we found that Fitnet would benefit from a penalty factor in the loss function in order to overcome the MSE loss limitations for our specific problem. We saw on synthetic data that this would produce a little improvement to the all process, but still with some error dependence on the example complexity. Field data are way more complex than synthetic cases, normally, which makes Fitnet-S-SA-H more likely to behave like in the complex case in Fig. 2.39a. It turns out that, particularly in this case, the results on synthetic data were somehow indicative of the performance on field data. The SL-Picknet is the segmentation map driving the best Fitnet-S-SA-H results (Fig. 2.47a-2.48a), compared with the CL-Picknet input, whereas the more confusing CL-Picknet panel determines failure in approximating the marker, but given the generally bad results for these examples, there is no interest in discussing the prediction timing for these figures.

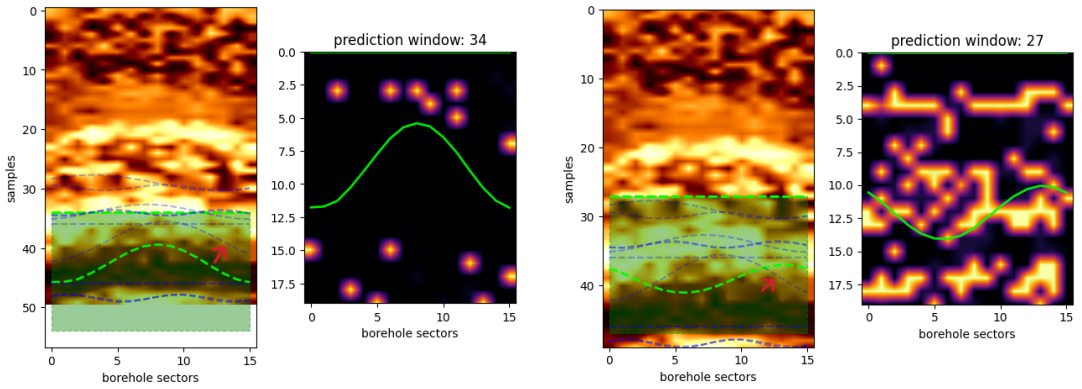

Figure 2.47: Fitnet-S-SA-H on left: SL-Picknet segmentation map; right: CL-Picknet segmentation map. The red arrow indicates the randomly selected marker (blue line) among the CVML predictions in the LWD3 interval.

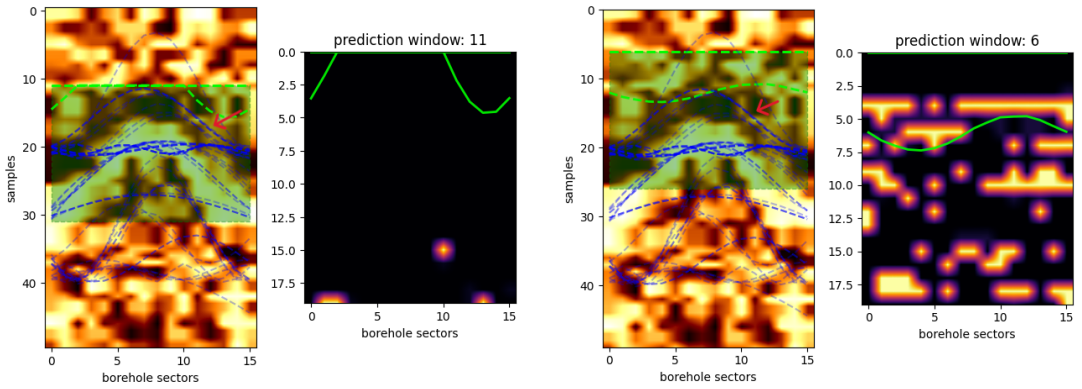

Figure 2.48: Fitnet-S-SA-H on left: SL-Picknet segmentation map; right: CL-Picknet segmentation map. The red arrow indicates the randomly selected marker (blue line) among the CVML predictions in the LWD1 interval.

Fitnet-C-SA-H Finally, we trained Fitnet-C-SA-H to try solve the clear error dependence on synthetic data complexity for Fitnet-S-SA-H, and, arguably, obtained the best Fitnet versions, as was already indicated in the results on synthetic data (Fig. 2.39-2.41). The field data confirm these results, highlighting a good similarity of the selected predictions and the marker, and at the same time what we observed for the CL map in Fig. 2.45b,i.e., that the CL-Picknet segmentation map drives an early detection of the possible features in the panel, compared to the SL one.

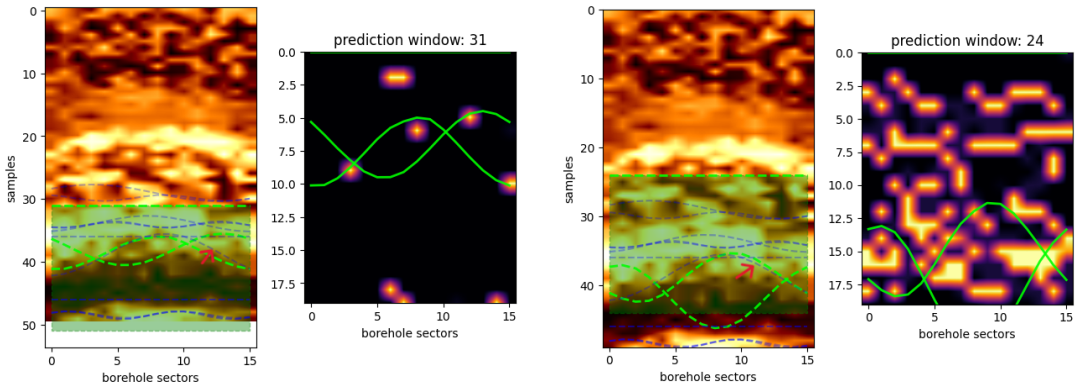

Figure 2.49: Fitnet-C-SA-H on left: SL-Picknet segmentation map; right: CL-Picknet segmentation map. The red arrow indicates the randomly selected marker (blue line) among the CVML predictions in the LWD3 interval.

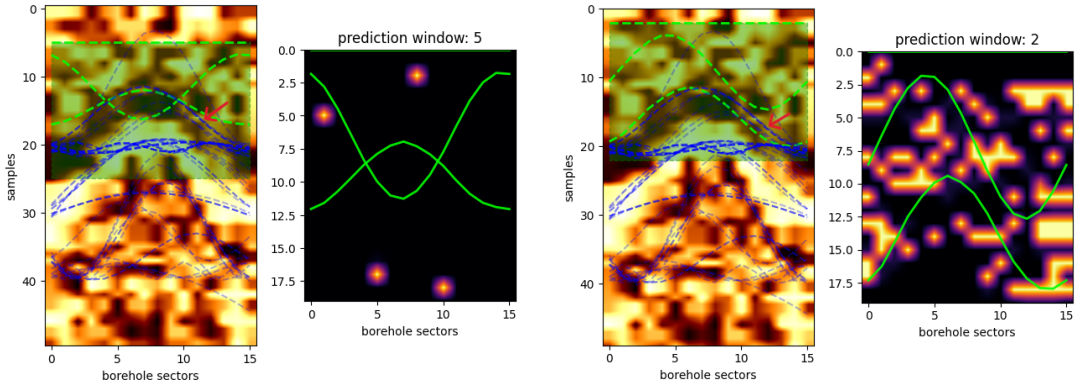

Figure 2.50: Fitnet-C-SA-H on left: SL-Picknet segmentation map; right: CL-Picknet segmentation map. The red arrow indicates the randomly selected marker (blue line) among the CVML predictions in the LWD1 interval.

In summary, the Fitnet results on field data aligned with the synthetic data results. They both confirmed Fitnet-C-SA-H as the best Fitnet version, that overcomes some of the key limitations identified during the implementation of the proposed method, but at the cost of a greater computational complexity that leads to the maximum training time compared to all the other Fitnet versions. Prediction time is also slightly larger. We analyzed the results comparing them with randomly selected CVML predictions, called as *markers*, to increase the objectivity in the analysis. However, it's useful to keep in mind two fundamental questions: (1) that the CVML method has its own limitations, which makes the marker itself somehow questionable, and (2) the marker was randomly selected, which implies that some of the bad results shown in the figures above could have been a good result if we randomly selected a different marker. We discussed the inherent limitations of the DL approach and tackled them proposing some motivated solutions to improve the performances. Since we obtained the expected improvements during the evaluation on both synthetic and field data we, therefore, might consider the DL-based method results on field data as good, and to

have demonstrated the applicability of a synthetic data based supervised learning approach for the automated BHI analysis based on DL, avoiding the use of field data and the manual interpretations as ground truths. Further analysis on field data results will be done in the next sections.

Confidence

Both the CVML and DL methods rely on some transformation of the raw BHI, in order to proceed with the automated features detection. We described in section 2.3.1 that we made use of the sobel results to compute the confidence of each detected feature. Differently, the DL method uses Picknet (section 2.2.3) to transform the raw image in a segmented binary image, with nonzero values associated to the sinusoidal edges it includes, if any. This segmentation map is then used to fit the best matching sinusoid of each edge. This suggests that this intermediate step that is required in both approaches, should contain some highly relevant information about both the raw input data, and the corresponding features. This gives rise to some questions: (i) *can we leverage the information included in this intermediate step to further push forward our understanding of the results?* and (ii) *how reliable is this information?*. Our segmentation task involves pixel-wise binary classification, where each pixel in the binary segmentation map represents the probability of that pixel being associated with an edge in the raw LWD BHI. We apply a sample-wise windowing technique to the raw borehole images, and for each window, we generate the corresponding segmentation map. Once all segmentation panels are collected, we restore the original raw data shape by averaging the segmentation maps for each meter (see Fig. 2.28-2.29). We propose to leverage those restored maps generated by both SL-Picknet and CL-Picknet to formulate a confidence measure of the predicted features. We start by analyzing the distribution of the standard deviation of the pixel values over all the segmented panels for the LWD datasets (Fig. 2.51). The panels standard deviation is higher if we use the SL-Picknet version for segmentation (see section 2.3.2). The analysis of panel standard deviations for the LWD3 and LWD1 datasets reveals significant differences in variability between the two Picknet versions. In both datasets, CL-Picknet panels exhibit narrowly distributed standard deviations around low values, indicating minimal variability and greater consistency. Specifically, in the LWD3 dataset, the CL distribution peaks near $\sigma_{\text{panel}} \approx 0.005$, while in the LWD1 dataset, the peak occurs around $\sigma_{\text{panel}} \approx 0.02$. In contrast, SL-Picknet panels show broader distributions with higher central values, reflecting greater variability. For the LWD3 dataset, the "SL" standard deviations are centered around $\sigma_{\text{panel}} = 0.15$, whereas in the LWD1 dataset, the distribution centers around $\sigma_{\text{panel}} = 0.25$. These results consistently highlight that "CL" panels are more stable and exhibit lower variability than "SL" panels, suggesting intrinsic differences in their characteristics, measurement precision, or stability across both datasets.

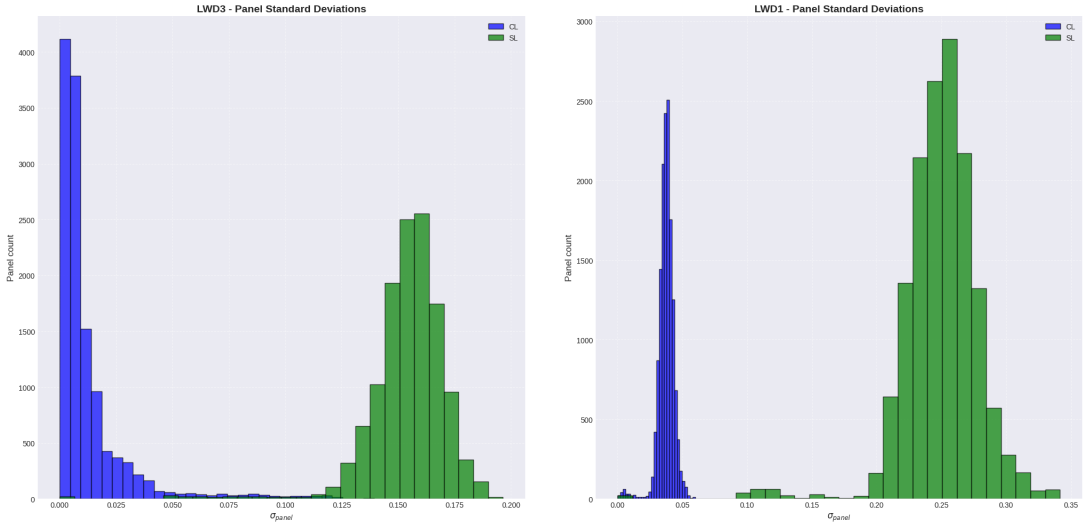

Figure 2.51: Left: histogram of LWD3 panel standard deviations comparing two categories: CL (blue) and SL (green). CL shows minimal variability (peaking near 0.005), whereas SL exhibits a higher variability (centered around 0.15). Right: histogram of LWD1 panel standard deviations comparing two categories: CL (blue) and SL (green). CL displays low variability (peaking near 0.02), whereas SL shows a higher variability (centered around 0.25).

However, interpreting these results can be challenging, as one might initially assume that lower standard deviation indicates better performance. Conversely, a higher pixel-wise edge probability from one network could suggest more robust behavior, as it may consistently assign high probabilities to real edges and low probabilities to non-edges. We define the segmentation-based confidence C_{sin} for each sinusoid sin as:

$$C_{sin} = \sum_{i=1}^{N} w_1 \cdot SL_{sin,i} + w_2 \cdot CL_{sin,i} \tag{2.38}$$

where N represents the number of elements in the sinusoid (corresponding to the number of sectors in the BHI), w_1 is a weight applied to the value of the i-th pixel along the sinusoid on the SL segmentation map $SL_{sin,i}$, and w_2 is a weight applied to the i-th pixel value from the CL segmentation map $CL_{sin,i}$. To validate the confidence measure, we adjust w_1 and w_2 such that

$$w_2 = 1 - w_1 \tag{2.39}$$

We then observe how the set of detected features changes after filtering out all features with $C_{sin} \geq 0.5$, allowing us to assess the impact of favoring either of the two segmentation maps (SL or CL) in the confidence calculation, assigning a greater weight to it. As we can observe in 2.52 and 2.53 for w1 values from 0.1 to 0.9 the number of sinusoids with $C_{sin} \geq 0.5$ decreases, which means that relying more on the SL segmentation map for the confidence measure, the whole confidence set would shift towards lower values, and less of them will exceed the 0.5 threshold. At the same time, relatively high confidences are more likely to be

related to evident petrophysical contrasts in the raw BHI, which suggests that SL-Picknet and CL-Picknet can represent two interchangeable solutions for an hypotetical interpreter who is serving of the proposed method for his specific interpretation goals.

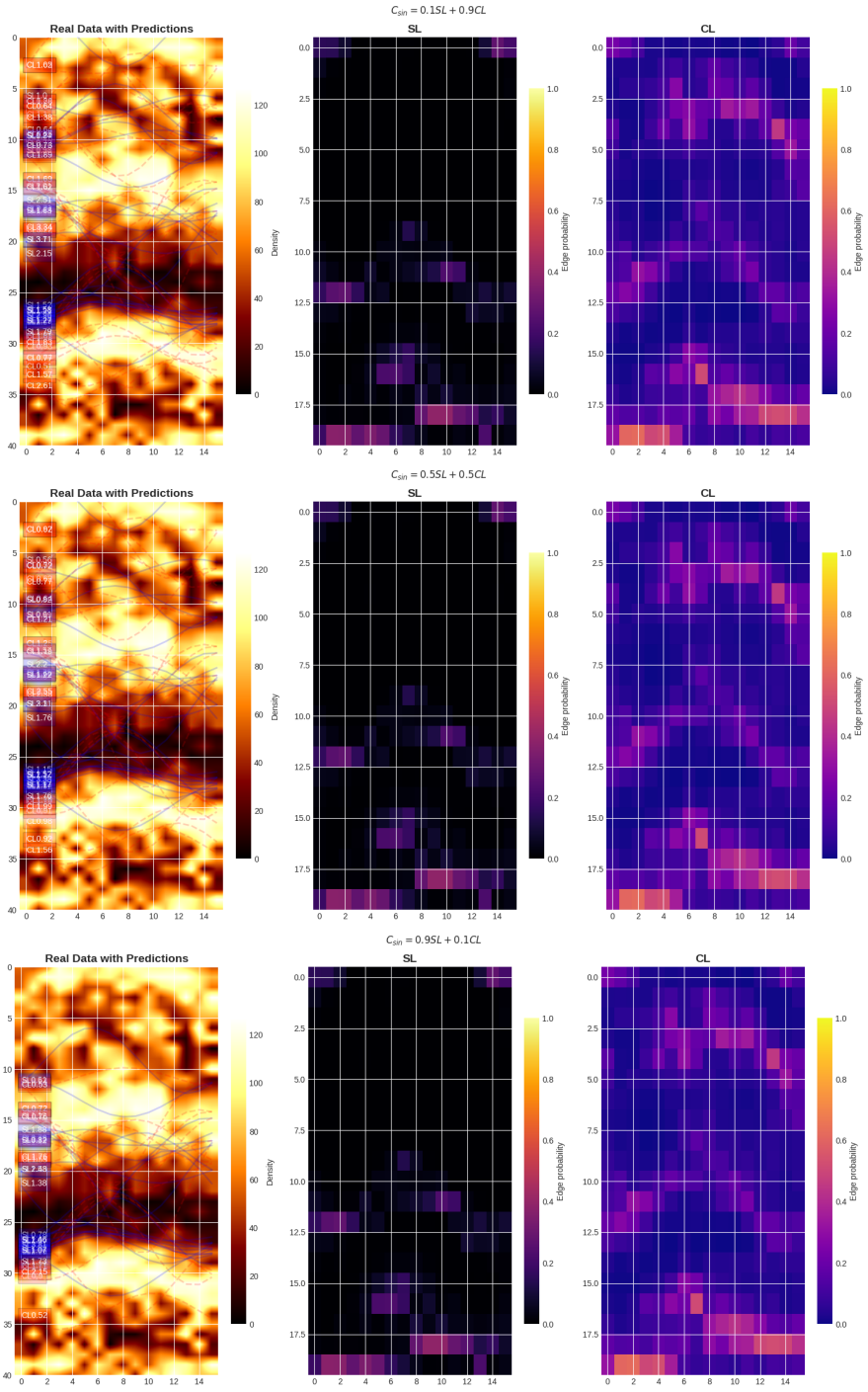

Figure 2.52: LWD3: The set of predicted features filtered of features with confidence lower than 0.5, for different values of w1 and w2 (see eq. 2.39).

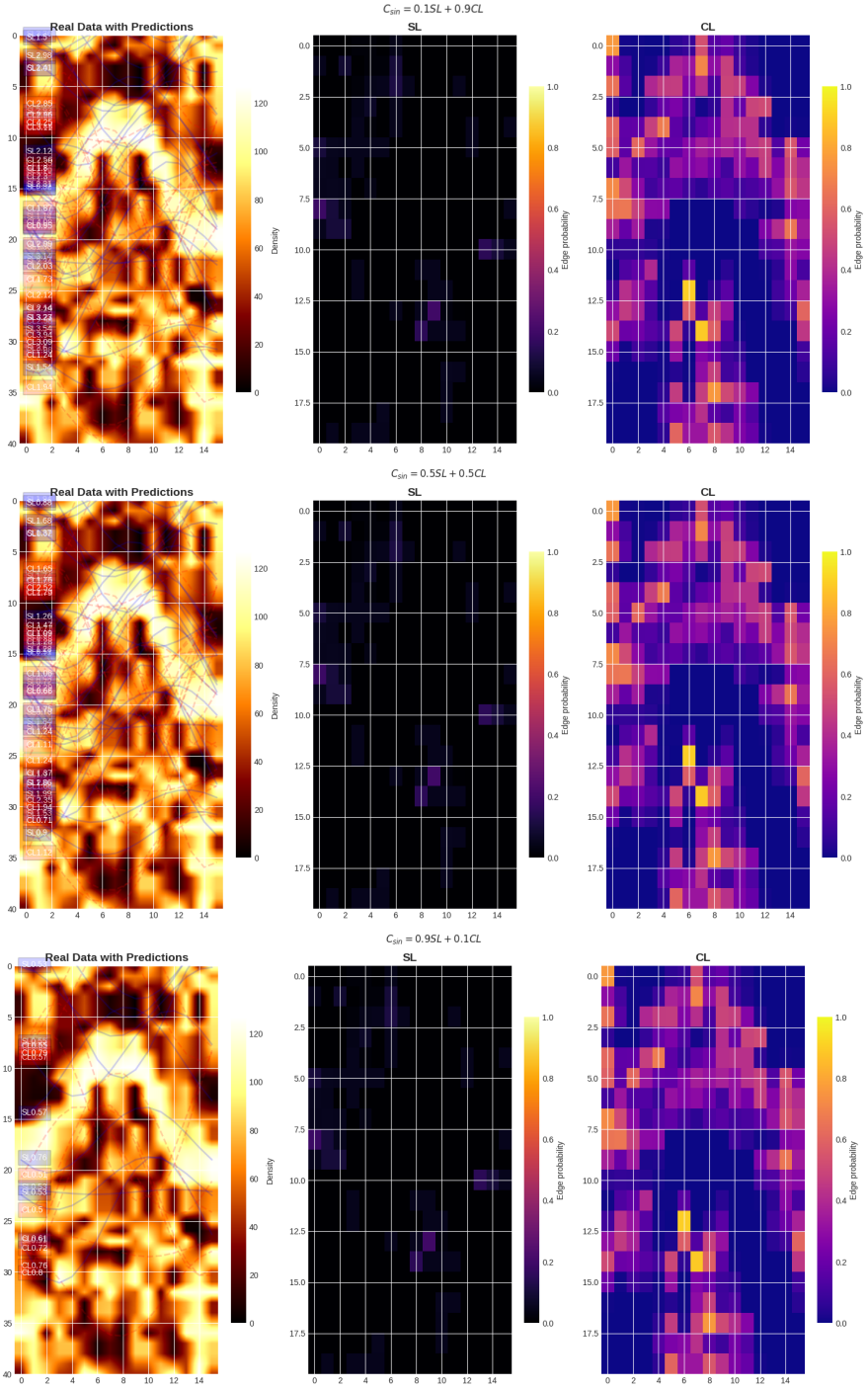

Figure 2.53: LWD1: The set of predicted features filtered of features with confidence lower than 0.5, for different values of w1 and w2 (see eq. 2.39).

During manual interpretation, the features correlated by the human interpreter are then classified according to its associated geological discontinuity. The main classes of interest are beddings, faults and fractures. Beddings generally separate intervals of various thicknesses with coherent petrophysical properties, and they normally appear as lower amplitude and more regular sinusoids in a vertical or near-vertical well, compared to fractures and faults. Fractures can be associated

with irregular features with different diagnostic attributes depending on the type of fracture (empty, filled, cemented), the type of borehole image (density, acoustic, resisitivity), etc. This classification task is normally combined with some prior domain knowledge of the interpreter. However, we can try to harness the DL results to quantify some diagnostic attributes for each predicted feature and see if they can be correlated with the type of geological discontinuity. For each feature, we associate two sets of values $v1$ and $v2$ collected from n-samples above and below the feature, respectively, similarly to what we did for the CVML method (see section 2.3.1). We estimate v1 and v2 distributions and use the first two statistical moments ($\bar{v}$, σ_v) to build a filtering criterion.

$$\bar{v2}_{sin} \geq \bar{v1}_{sin} + \Gamma \cdot \sigma_{v1,sin} \tag{2.40}$$

where $\bar{v1}$ and $\bar{v2}$ are the mean values in the two sets, σ is the standard deviation, and Γ is a factor $\in [0, 1]$ that we apply to σ to soften the constraint regarding the values differences above and below the prediction. Figures 2.54, 2.55, 2.56 show the different results for $\Gamma = 1, \frac{1}{2}, \frac{1}{3}$.

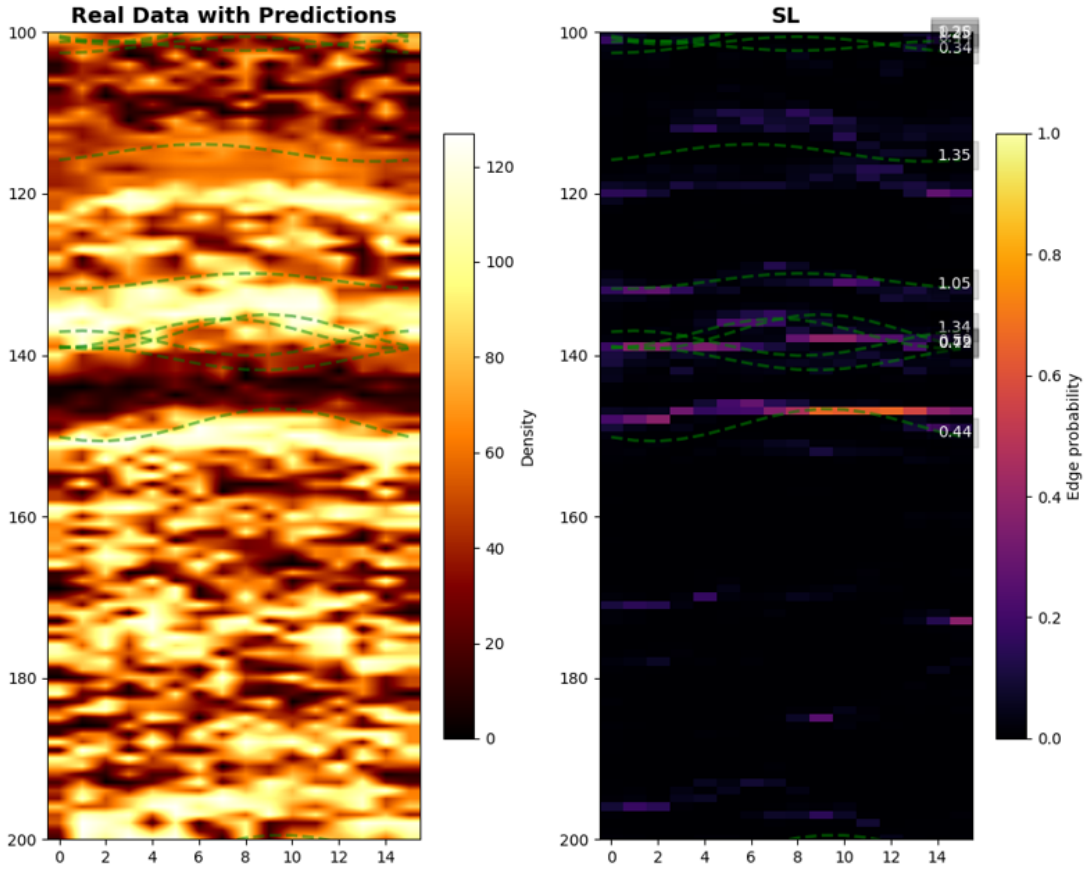

Figure 2.54: **CLASS** α for the sigma factor $\Gamma = 1$

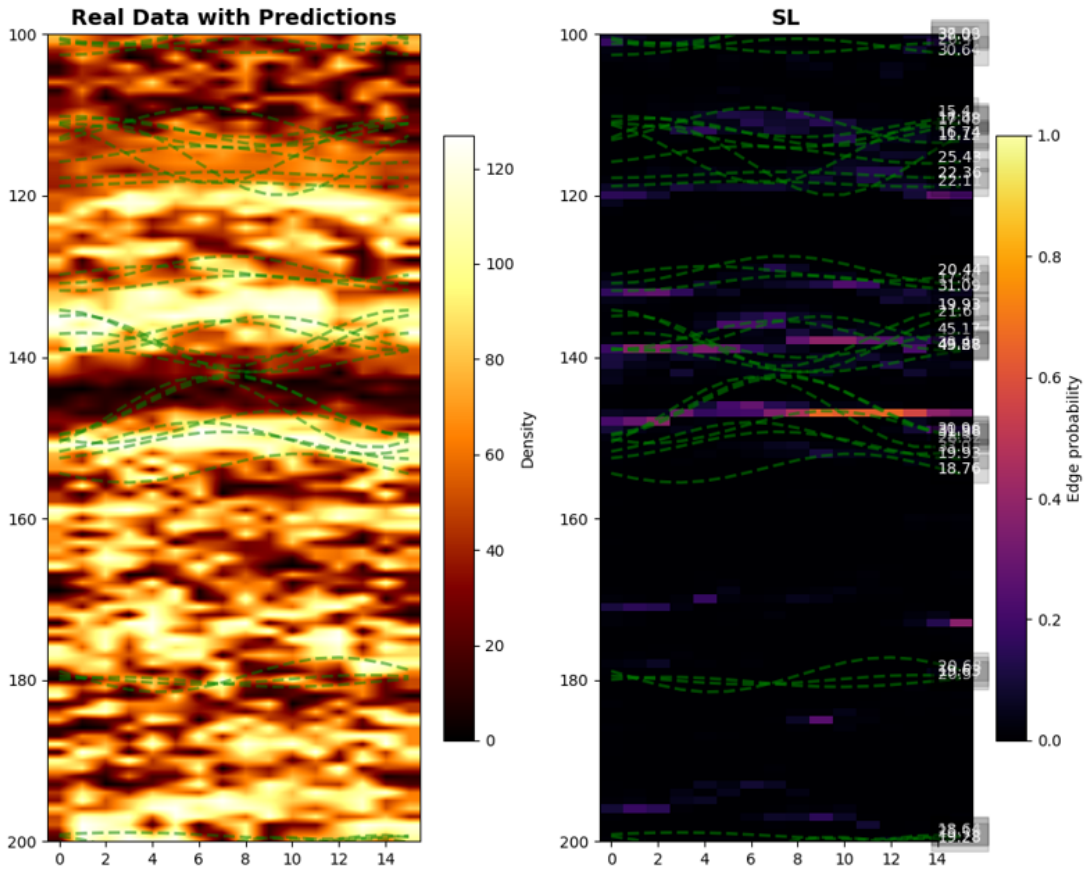

Figure 2.55: **CLASS** β for the sigma factor $\Gamma = 1/2$

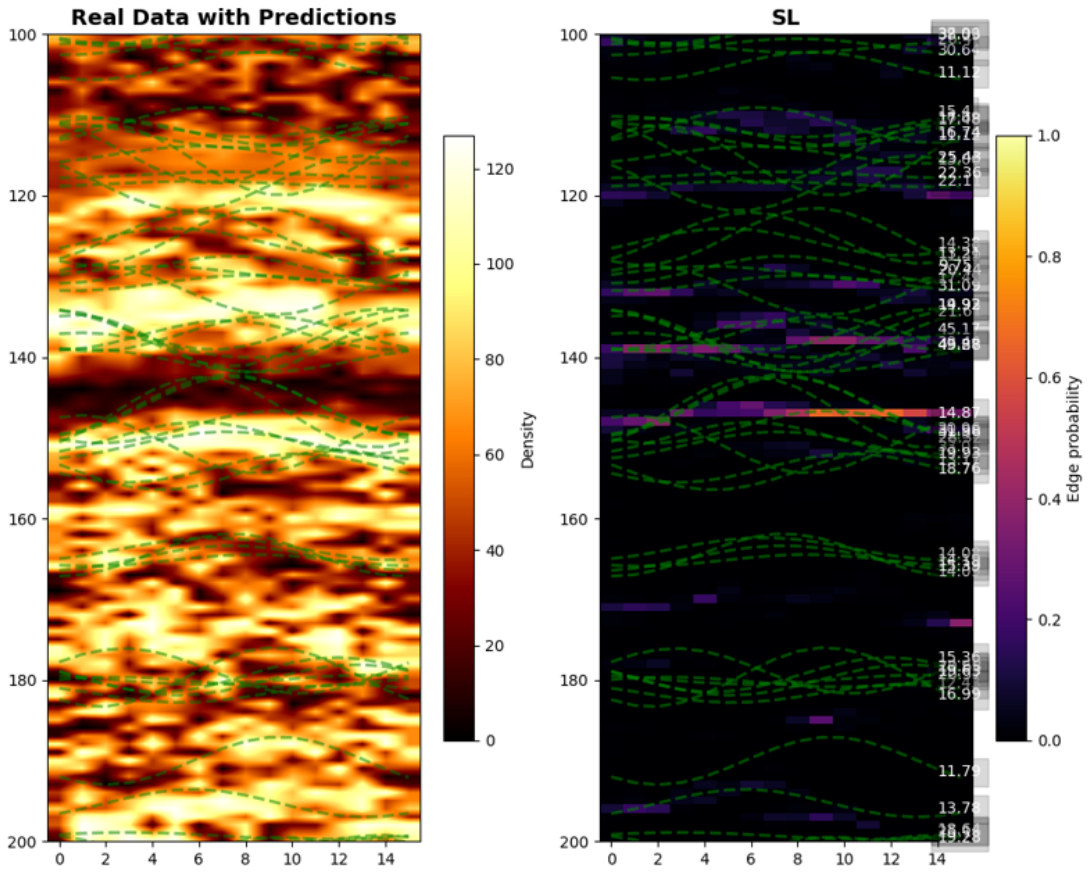

Figure 2.56: **CLASS** γ for the sigma factor $\Gamma = 1/3$

We see that in this case as well, by decreasing Γ we increase the number of features to keep after the filtering and that the features gradually increase their dispersion around the most prominent petrophysical contrasts. Moreover, we can observe in 2.57 that the Δ_ρ is decreasing for higher sigma factor values , where

$$\Delta_\rho = \frac{1}{M} \sum_{sin=1}^{M} |v1_{sin} - v2_{sin}| \tag{2.41}$$

and

$$sigmafactor = 1/\Gamma \tag{2.42}$$

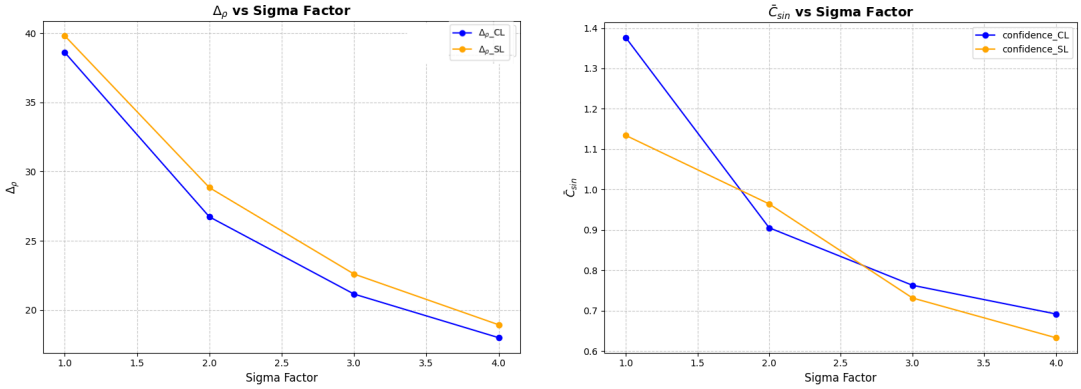

Figure 2.57: LWD3. Left: $\Delta\rho$ vs sigma factor. Right: mean confidence $\bar{C}_{sin}$ vs sigma factor.

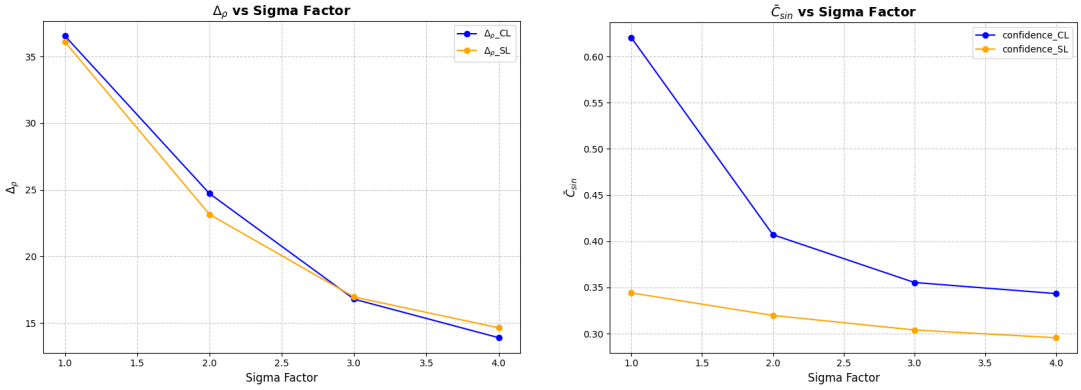

Figure 2.58: LWD1. Left: $\Delta\rho$ vs sigma factor. Right: mean confidence $\bar{C}_{sin}$ vs sigma factor.

Fig. 2.57 shows that confidence tends to decrease with increasing *sigma factor*. This implies that deep-learning models are more certain in their predictions when the *sigma factor* is lower, which reinforces the hypothesis that lower *sigma factor* values are more likely to be related to bedding features, associated with relatively sharp formation changes. Higher contrast regions generally provide clearer, more distinct features that facilitate more confident predictions.

Furthermore, this relationship can also be explained in terms of the confidence weighting mechanism used by the model (refer to eq. 2.39). When the contrasts are high, the features extracted are more pronounced, allowing the NN to identify characteristic patterns more effectively, thereby leading to higher confidence values. This aligns with the idea that geological categories with distinct, high-contrast features are easier for the model to categorize with certainty.

It might also be beneficial to investigate whether an optimal range of *sigma factor* exists that provides a balance between minimizing noise and retaining enough feature contrast for confident predictions. Understanding this balance could improve prediction accuracy and confidence in geological interpretations.

2.3.4 Results integration

In section 2.3.3, we visually observed and analyzed the results of the proposed DL-based approach. In the lack of a manual interpretation, we used the CVML results described in section 2.3.1 as a baseline to compare the performances of several configurations of such DL-system (Fig. 2.43-2.50). Despite the differences in the results, which may also given by both the random selection of the CVML marker and the inherent differences generated by the sample-wise sliding window technique, the comparison indicated that the results of the two method can eventually converge to similar automated analysis of the borehole image, after some efficient integration. To do that, we started by solving the rectangular *linear sum assignment problem (LSAP)* to find the Fitnet predictions $\mathbf{F}_{DL}$ that best matches the CVML set of features $\mathbf{F}_{CVML}$ (see section 2.2.2). Let's denote the set of Fitnet features as $\mathbf{F}_{DL}$. We speak of our LSAP as a rectangular type LSAP, because the size of the two sets is likely to be different, due to different implementations. Let's denote the sizes of the $\mathbf{F}_{CVML}$ and $\mathbf{F}_{DL}$ as $\mathbf{N}_{CVML}$ and $\mathbf{N}_{DL}$, respectively. We expect $\mathbf{N}_{DL} \geq \mathrm{N}_{CVML}$, especially for higher threshold values in the CVML algorithm (see section 2.3.1). We can now build the $\mathbf{N}_{DL} \times \mathrm{N}_F$ cost matrix $\mathbf{C}$, where each entry is equal to the MSE measure of the correspond pair of $\mathbf{F}_{DL}$ and $\mathbf{F}_{CVML}$ features (Fig. 2.59).

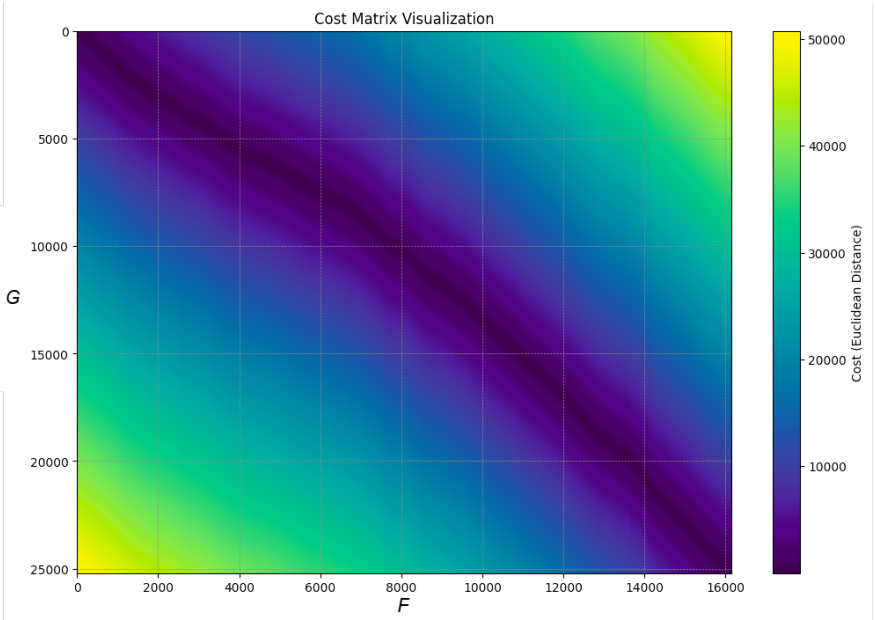

Figure 2.59: An example of cost matrix built using $\mathbf{F}_{DL}$ and $\mathbf{F}$ from the LWD3 borehole image in the field dataset. Note that G (columns) is greater that F(rows), which makes our LSAP problem a rectangular problem.

The 2D rectangular assignment problem involves selecting one element from each column and at most one element from each row, with the goal of minimizing the total sum of the selected elements in $\mathbf{C}$. Mathematically, this expressed as in equation 2.43.

$$\mathbf{X}^* = \arg\min_{\mathbf{X}} \sum_{i=1}^{N_R} \sum_{j=1}^{N_C} c_{i,j} x_{i,j} \tag{2.43}$$

subject to

$$\sum_{j=1}^{N_C} x_{i,j} = 1 \quad \forall i \tag{2.43.2}$$

$$\sum_{i=1}^{N_R} x_{i,j} \leq 1 \quad \forall j \tag{2.43.3}$$

$$x_{i,j} \in \{0, 1\} \quad \forall i, j \tag{2.43.4}$$

Equivalent to

$$x_{i,j} \geq 0 \quad \forall x_{i,j} \tag{2.43.5}$$

We repeated this procedure for every threshold in the CVML method and calculated the total error in each set of matched $\mathbf{F}_{DL}$ and $\mathbf{F}_{CVML}$ features as the average of all the matched pairs MSE. In this way we want to observe if there is a correlation between the degree of similarity between $\mathbf{F}_{DL}$ and $\mathbf{F}_{CVML}$ and the threshold used in the CVML algorithm. If so, it would imply that the threshold needed in the CVML method not only impacts the CVML results, but could also enhance the alignment between the $\mathbf{F}_{DL}$ and $\mathbf{F}_{CVML}$ features, potentially leading to improvements in the semiautomated analysis, by means of this simple integration. Referring to the same dataset associated with the cost matrix in Fig. 2.59, we can see in Fig. 2.60 the example of LWD3 and LWD1, respectively, which confirm the expected threshold impact on the similarity of the two methods results. The results indicate an efficient integration of the two features sets through this simple LSAP approach, would be a useful tool to filter the $\mathbf{F}_{DL}$ features set of some redundant predictions, taking advantage of the information provided by the CVML algorithm so that $\mathbf{F}_{DL} \approx \mathbf{F}_{CVML}$. By means of LSAP, we could therefore achieve an approximately equal spatial distribution of the outcomes of the two methods (Fig. 2.61).

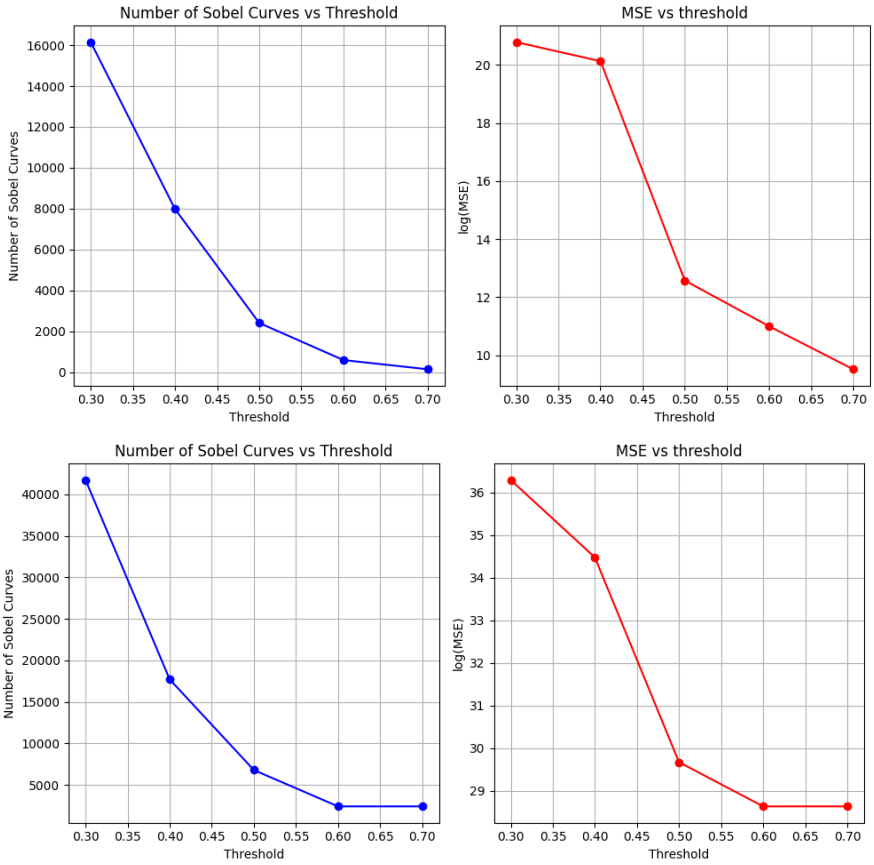

Figure 2.60: Left: number of CVML curves vs thresholds. Right: the logarithm of the average MSE in the set of matched features decreases for higher thresholds, suggesting the impact of threshold in the quality of the integrated results. The upper figure refers to the LWD3 dataset, whereas the lower refers to LWD1.

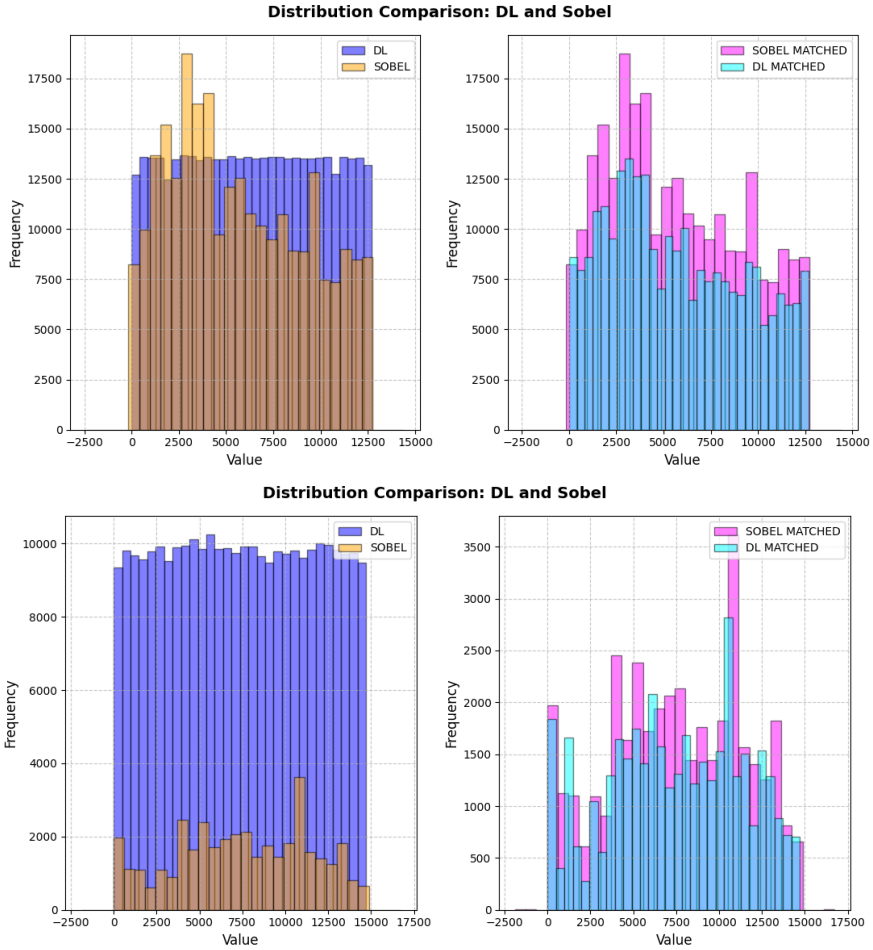

Figure 2.61: Top left: Histogram comparing the distributions of DL (blue) and Sobel (orange) values. The distributions are overlaid to highlight differences and overlap, with DL exhibiting slightly more uniform variability compared to Sobel. Top right: Histogram showing matched distributions for Sobel (pink) and DL (cyan) values. Bottom left: Histogram comparing the distributions of DL (blue) and Sobel (orange) values with a broader range. The plot illustrates differences in spread and variability, with DL values appearing more evenly distributed. Bottom right: Histogram showing matched distributions for Sobel (pink) and DL (cyan) values over an extended range. This plot focuses on the aligned sections of the distributions after matching.

However, with respect to the real-time or near real-time application of these methods on LWD borehole images, some considerations must be done. Such a near real-time application requires a reasonably high speed of the system, in order to fit the original purpose. In practice, this means that the system must be able to quickly provide with high-quality results and information that effectively supports operations and decision-making during drilling. We saw in this section that the CVML and the DL results can be integrated to find a two final approximately equal sets of features of the same LWD BHI, with similar spatial distributions. To achieve these results in near-real time would require, nevertheless, the two methods to be interoperable. This is not the case, unfortunately, since the pre-

diction time for 1-meter of borehole image is in the order of the seconds for the DL method to run both Picknet and Fitnet, whereas takes several minutes to the CVML algorithm (see Tab 2.4 and 2.1).

2.3.5 Semiautomation

The adoption of the proposed method as an interactive software, whether operating on a server or a user's local machine, necessitates addressing the crucial aspects of explainability. Explainability not only fosters user trust and enhances the overall experience with the tool but also provides valuable scientific insights into the behavior of deep neural networks (DNNs), which are often perceived as opaque "black boxes."
This section explores field data results to highlight key considerations in the context of human-machine collaboration (introduced in sect. 2.2.6). In Section 2.3.3, we provided an overview of the two versions of Picknet and their influence on the subsequent Fitnet prediction step. Specifically, it was observed that segmentation maps produced by CL-Picknet exhibit a higher density of non-zero pixels, suggesting a greater sensitivity to edge features in the input image. This distinction is further illustrated in Fig. 2.51, which shows the standard deviation of pixel values across all 1-meter segmentation maps in LWD3. The higher panel standard deviation associated with SL-Picknet aligns with tghe evidence that this Picknet alternative focuses on segmenting one edge at a time. Conversely, CL-Picknet captures additional, and more subtle details in the raw BHI. This characteristic allows CL-Picknet to detect features earlier than SL-Picknet (see Table 2.5), making it particularly well-suited for real-time or near real-time applications. In Section 2.2.6, we introduced the fundamental concept underlying the practical use of these tools: the semi-automation of the process, wherein humans and machines collaborate to achieve higher performance. This semi-automation can be envisioned as a human (user) validating and discarding feature predictions. During this validation process, understanding the underlying prediction mechanisms becomes critical to understand how the human can be influenced by different types of machines trained for the same tasks. Here, we make some examples of comparison of the guidance provided to the user by the two Picknet versions and analyze the semi-automation of this task from a human-machine collaboration perspective. As discussed in Section 2.2.6, the accuracy of fracture detection by humans working on diverse datasets depends significantly on their cognitive style. Individuals working alone often exhibit cognitive tendencies—such as sensitivity (detail-oriented) or intuition (less detail-oriented)—that influence their performance. When working collaboratively, the sensitive style can often influence and complement the intuitive style. Building on this foundation, we attempt to extend these ideas to the two versions of Picknet and explore how they align with these cognitive styles. Specifically, we assess whether each version acts as a "detail-oriented" or "intuitive" collaborative entity. For this purpose, we analyze a 100 samples (5 meters) interval, which includes sub-intervals with different degree of complexity. We tried to quantify this complexity by measuring the variance of the pixel values in the input panel, similarly to what we did for the

Picknet panels (Fig. 2.51). This is helpful to emphasize the analogies with the aforementioned human cognitive styles impacting the fracture detection outcome, and to enhance the interpretability of our method, including how the implicit information like the confidence score can be used in our proposed semiautomation framework. The version of Fitnet we used for this is Fitnet-C-SA-H, as it corresponds to the best Fitnet version and the one with maximum cosine similarity of the features predicted on the two Picknet versions (section 2.3.2, Tab. 3.3). First of all, we observe how this same similarity changes and if it correlates the input variance sequence, across the reference interval of field data (Fig. 2.62). We see in the figure that these two quantities are related and the lower variance areas in the interval coincide with the higher similarity values.

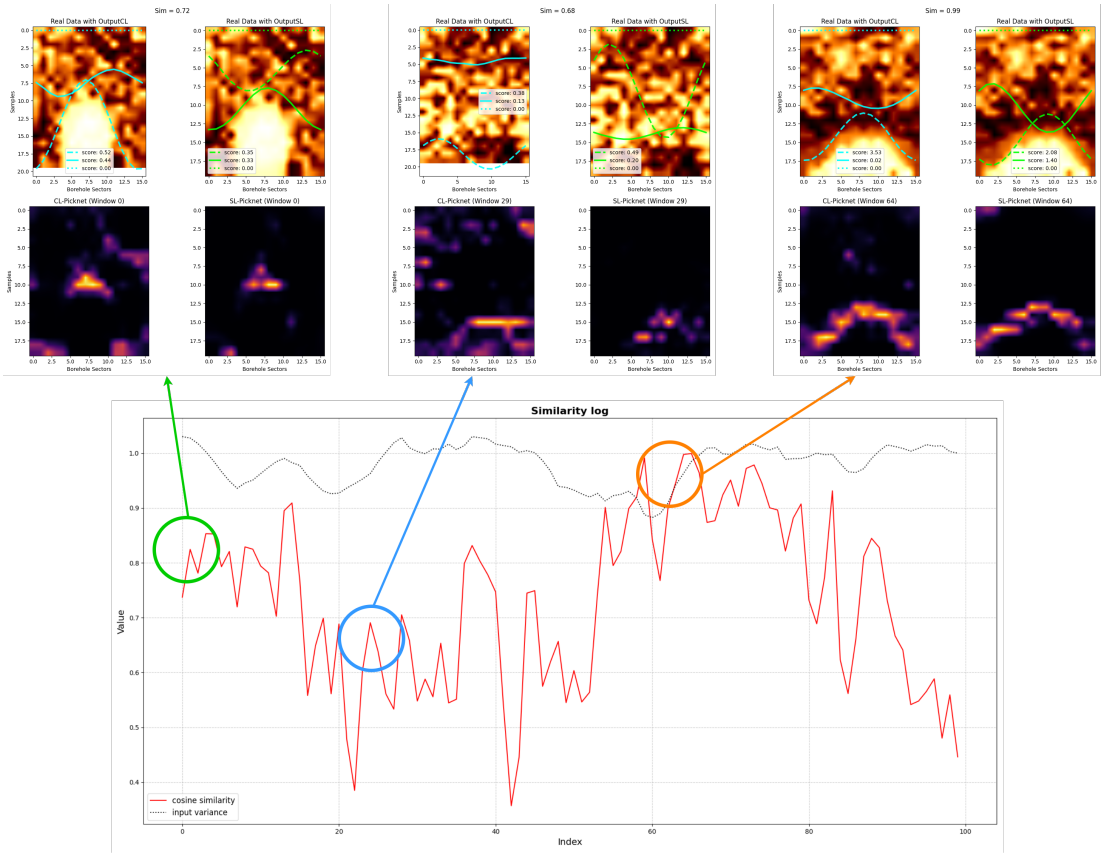

Figure 2.62: The figure illustrate the similarity of the features predicted using CL and SL Picknet as input and its relation with the variance in the input BHI.

This emphasizes the impact of the Picknet training strategy on Fitnet and, by extension, the entire semiautomation process. We examine examples of Fitnet features from both Picknet alternatives within the same sub-interval. Since the outcomes are more similar in low-variance sub-intervals, we begin our analysis with high-variance sub-intervals to better observe significant deviations (Fig. 2.63).
The figure illustrates that in this interval, where strong and distinct contrasts are absent, the SL-Picknet segmentation maps produce a more robust Fitnet prediction. This robustness is evident as the features remain consistently aligned

with the same edge across the series of subsequent windows. In contrast, the CL-Picknet segmentation maps exhibit a more dynamic prediction behavior, where the detected features shift with the sample-wise sliding of the prediction window. Specifically, the SL-Picknet segmentation map directs Fitnet to concentrate on more pronounced features in the image throughout the sub-interval, ensuring consistent identification of structures with high stability. Conversely, the CL-Picknet segmentation map encourages Fitnet to more variable predictions and, consequently, consider subsidiary elements as well. This is also evident in Fig 2.64, where we observe the manual interpretations from the same interpreter in two intervals with different complexities (note that the manual picking were done *ad-hoc* for this example of semiautomation).

In Fig 2.64-A, for example, if we look between sample 20 and 30, we see a low-angle low-density feature that may be a low-angle fracture. This fracture was picked by the human interpreter, but Fitnet suggested a similar feature only when predicting on the CL-Picknet segmentation map. However, on the SL-Picknet segmentation maps the validations show a better alignment to the manual pickings in correspondence of the most prominent contrasts. This somehow confirms what we discussed during the description of Fig. 2.63, which is that CL-Picknet may be a more detailed-oriented tool, compared to the more robust SL counterpart. This also underpins their complementarity, which, as previously mentioned for the human working in groups, can be beneficial to the collaboration experience and the final results quality. In Fig 2.64-B an example with much less evident contrasts is shown. As expected (see Fig. 2.62), we can observe a more scattered pattern of the manual and the validation results. The manual interpretation, however, seem to be cautious and conservative, with the quasi-fully represented sinusoids picked, wheras other, more subtle features have been ignored. This scenario represents an example of how a conservative interpreter might be influenced to reconsider his perspective on a selected feature (e.g., blue arrows), validate their interpretation (green arrows), or incorporate additional elements to complete the overall puzzle (red arrows). Furthermore, evidence of the differing input perceptions and their impact on Fitnet's focus on detailed structures is demonstrated by the fact that all validations, which either confirmed or partially confirmed manual interpretations, were derived from the CL-Picknet segmentation maps for this challenging interval.

From a semi-automation perspective, this distinction suggests that a human collaborator working with these two alternatives of the same NN architecture could receive different suggestions: one focusing on prominent features within the borehole image and the other highlighting subtle details. Moreover, these features are informed by implicit information from the models, such as prediction confidence (section 2.3.3) and interval difficulty (Fig. 2.62), further enhancing the collaborative process. We believe approach can effectively reduce subjective bias during analysis and improve the reliability of models trained with the resulting data. For instance, subtle details in BHI log interpretation, such as small fractures near or crossing a larger one, could significantly influence fluid flow parameters in thermodynamic subsurface models. Detecting these details is therefore critical for accurate modeling and decision-making. As discussed in the Introduction

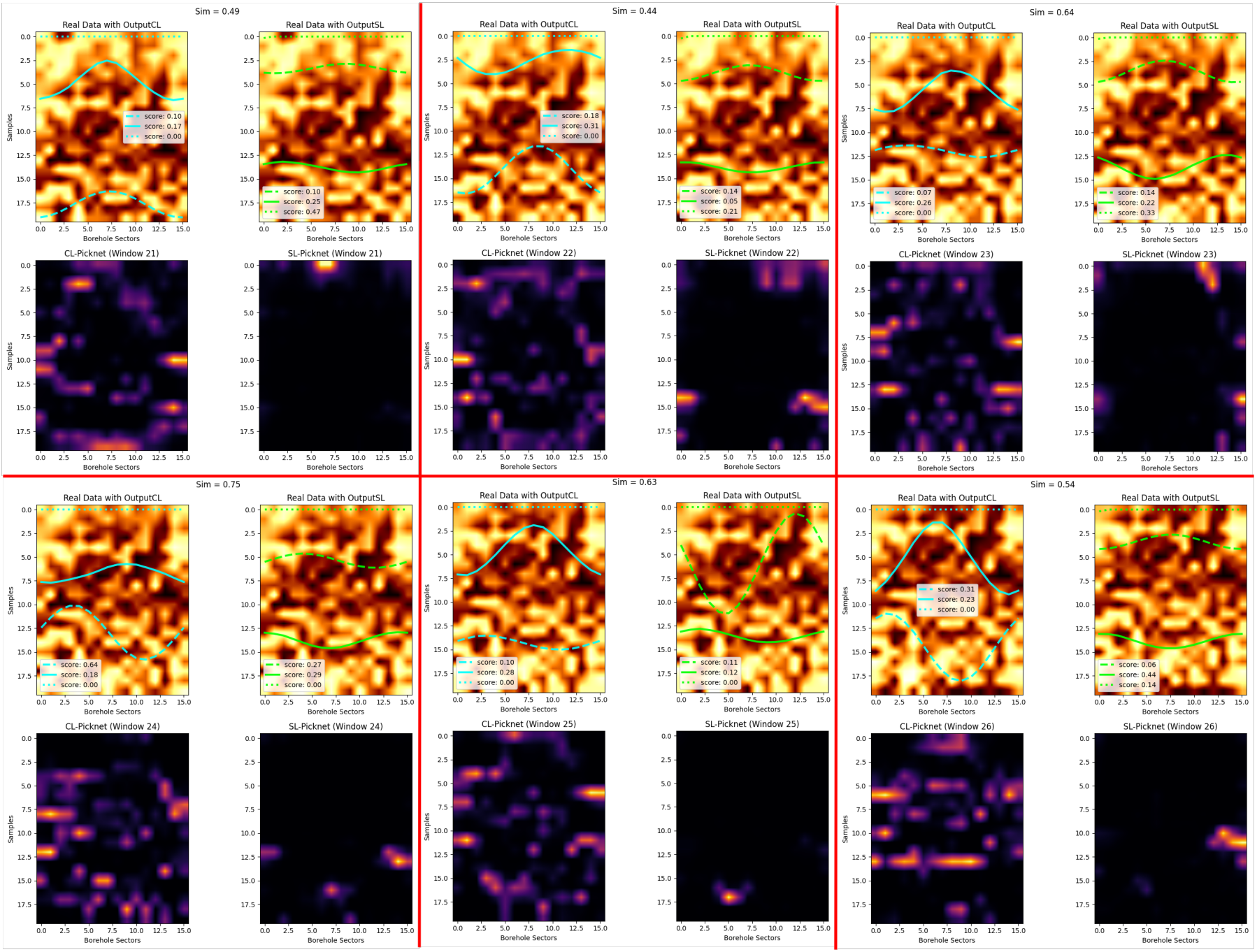

Figure 2.63: Fitnet-C-SA-H on CL-Picknet and SL-Picknet segmentation maps for these complex subsequent prediction windows.

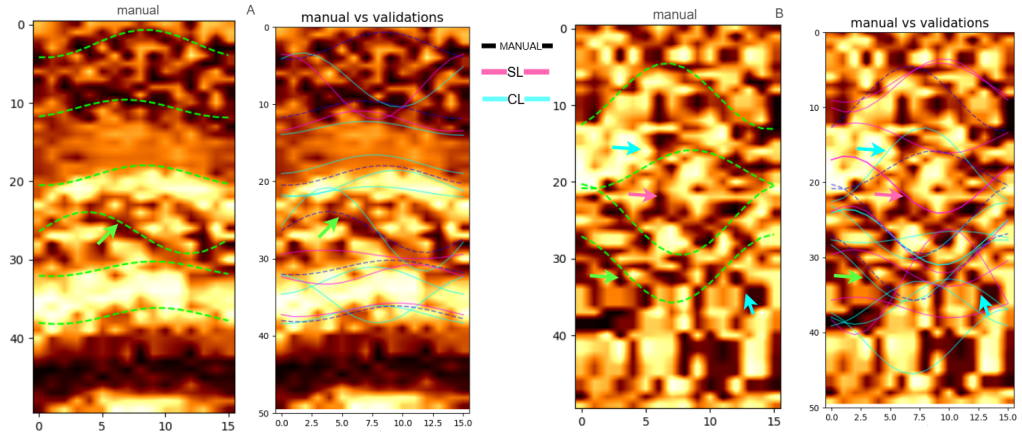

Figure 2.64: Manual interpretations and validations on both the CL-Picknet and SL-Picknet for two intervals with different complexity. Green arrows indicate the manual features that were confimed by the validations. Blue arrows indicate the manual features that could be questioned by the validation and modified accordingly, depending on the interpreter. Red arrows indicate the validations that don't correspond to any manual picking, but can represent a valid alternative.

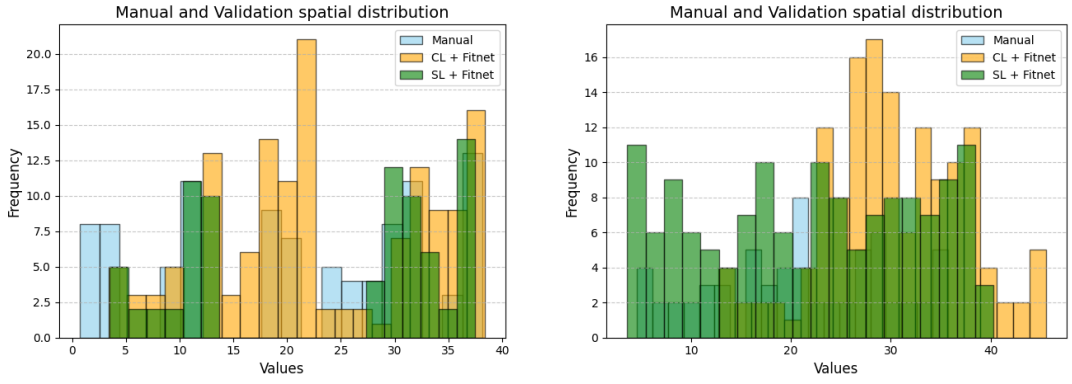

Figure 2.65: Distributions of coordinates in the intervals associated with (left) fig. 2.64-A and (right) 2.64-B.

of this chapter, previous studies have demonstrated that final fracture analyses by different interpreters often yield uncertain results, even when working with high-resolution data surpassing that of BHIs to varying degrees, depending on the resolution method used (e.g., linear or circular scanlines, window sampling, topology sampling). Consequently, achieving reproducible and meaningful fracture network interpretations remains a challenge. Addressing these issues—such as subjective bias among interpreters—could enhance confidence in modeling and engineering decisions related to geothermal resource estimation and management. While a statistically significant study on subjective bias was not feasible due to limited human resources, we present here the results of manual and semi-automated analyses conducted by a single interpreter (the thesis author) on LWD BHI data. Efforts were made to maintain consistent experimental conditions for both manual analysis and validation. This represents an initial attempt to collect variables from both processes, allowing for comparability of results and potential insights into human versus machine perception and performance. For instance, we create a framework for a sample-wise shifted window-based manual analysis of several intervals of BHI, in order to keep track of the window where humans would locate a sinusoid and its mean depth position and amplitude. Seemingly, on the human-machine collaboration side during validation we can keep track of the prediction window where the human decided to validate the suggestions coming from the machine, and their mean depth position and amplitude. We can also try to relate those observations with some features of the input data in each window to gain insights about possible correlations, i.e., the depth coordinate of the center of mass in the input window and the focal entropy. Additionally, the final results of manual and DL-assisted BHI interpretation can be compared in topological terms, analyzing the spatial distribution of nodes and branches defined by the intersections of the correlated features, on which the evaluation of the fracture network connectivity- closely related to formation permeability- depends (Fig. 2.66-c and 2.67-c). For instance, we did this for the same two intervals as in Fig. 2.64. For the interval in Fig. 2.64-A, wee see in Fig. 2.66 an example of the human-machine comparison studies as described above, where we observe an positive trend of the window focal-entropy value (Fig. 2.66-b) from window 10 (i.e., extending from depth samples 10 to 30) to 30 (from depth samples 30 to 50). In this sub-interval, can see that the features that have been validated were at located in deeper portions, with respect to center of mass of the associated windows. The green arrows in Fig. 2.66-a and Fig. 2.66-c indicate two almost identical correlated features- one from the manual interpretations, and the other from the validation procedure. Fig. 2.66-a indicates that, for this example using the curriculum learning segmentation map during the validation, the position of the features suggested by the machine could be influenced by the focal entropy in the raw input in a way that differs from the human. As we can observe, this feature have been detected in different input windows in the two scenarios and, specifically, the machine suggested its presence at an earlier stage (window 20, from depth sample 20 to 40) compared to when the human manually picked it (window 27, from 27 to 47). For the other interval example-more heterogeneous, with no evident layers in the corresponding BHI- we can see that more evidently

than in the previous case, the manually correlated features (Fig. 2.67-a) are closer to the position of the center of mass of the input window, whereas the validated features exceed its coordinates pretty much in all the prediction windows in the sub-interval.This suggests that the focal entropy within the input window—representing the randomness of pixel values in an image—and the position of the center of mass within the window may influence human and machine perception differently. These two input-related quantities could be valuable in future studies for a more detailed analysis of their differences, limitations, and potential for integration.

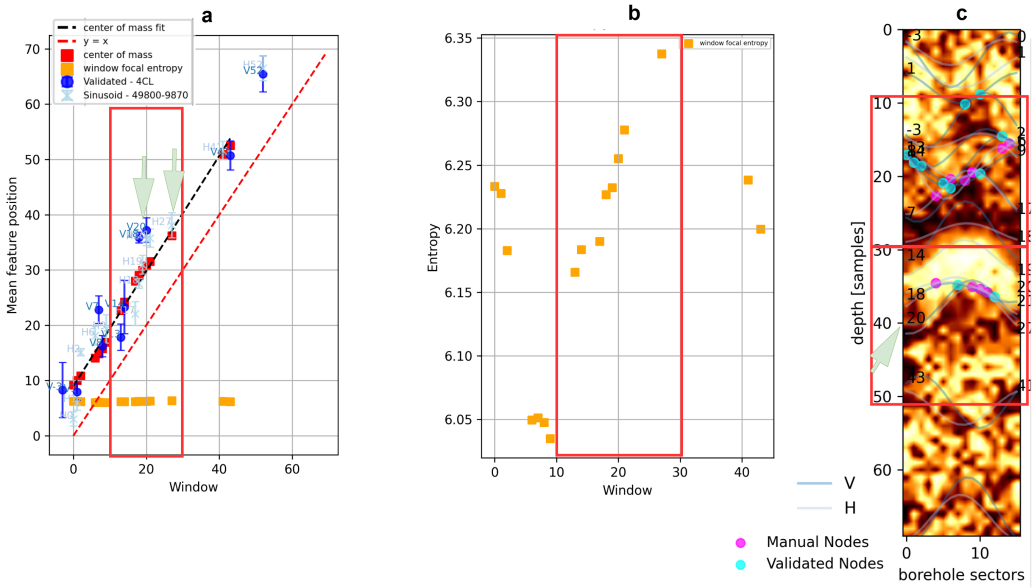

Figure 2.66: A comprehensive plot comparing the manual (H) and semi-automated (V) analyses for a high-contrast interval. (a) The validated (V-blue dot) and human (H-light blue cross) detected features are plotted, with the x-axis representing the window index in which the feature was detected and the y-axis showing the mean depth position of the feature in samples. The length of each bar is proportional to the feature's amplitude. Red squares represent the y-coordinates of window centers of mass, and the black dotted line indicates the linear regression of the set of centers of mass across all considered windows. A red dotted line is added for reference, representing $y = x$. Green arrows highlight two features that can be identified in (c). (b) The focal entropy-window plot, where the red rectangle highlights a positive entropy trend from window 10 to 30. (c) Validated and manually picked sinusoids, with blue dots and magenta dots representing nodes of the validated and manual feature sets, respectively.

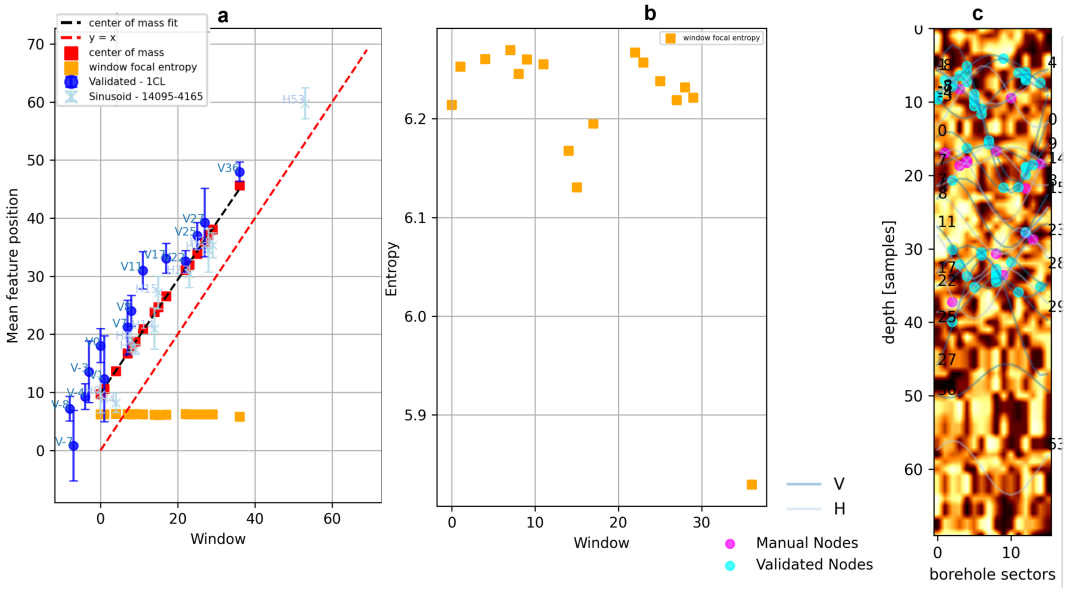

Figure 2.67: A comprehensive plot comparing the manual (H) and semi-automated (V) analyses for a low-contrast interval. (a) The validated (V-blue dot) and human (H-light blue cross) detected features are plotted, with the x-axis representing the window index in which the feature was detected and the y-axis showing the mean depth position of the feature in samples. The length of each bar is proportional to the feature's amplitude. Red squares represent the y-coordinates of window centers of mass, and the black dotted line indicates the linear regression of the set of centers of mass across all considered windows. A red dotted line is added for reference, representing $y = x$. Green arrows highlight two features that can be identified in (c). (b) The focal entropy-window plot, where the red rectangle highlights a positive entropy trend from window 10 to 30. (c) Validated and manually picked sinusoids, with blue dots and magenta dots representing nodes of the validated and manual feature sets, respectively.

As anticipated in the introduction, in the semiautomation framework also enables the use of the use of the validations to fine-tune the NN parameters (weights)-trained on synthetic data- based on real world examples. After this operation, the validation is repeated using the fine-tuned weights and we observe the differences in the window analysis results and distributions of topological elements. For instance, the validations in Fig. 2.66 for this less chaotic BHI interval, were used to fine-tune the NN weights for 1 epoch per validation, and the results of the repeated validation process are shown in Fig. 2.68 and 2.69-b. We can observe in Fig. 2.68-a that, especially in the sub-interval marked by the red rectangle, the validations results have changed compared to the original NN. In particular, considering the same features indicated by the green arrows (Fig. 2.68-a and -c) we see that now the same feature is validated in a deeper window (33, from depth samples 33 to 53) compared to the original NN (window 20, from depth samples 30 to 50) and the topological elements have also slightly changed from the original validation (Fig.2.69-a) to the validation on the fine-tuned NN (Fig. 2.69-b).

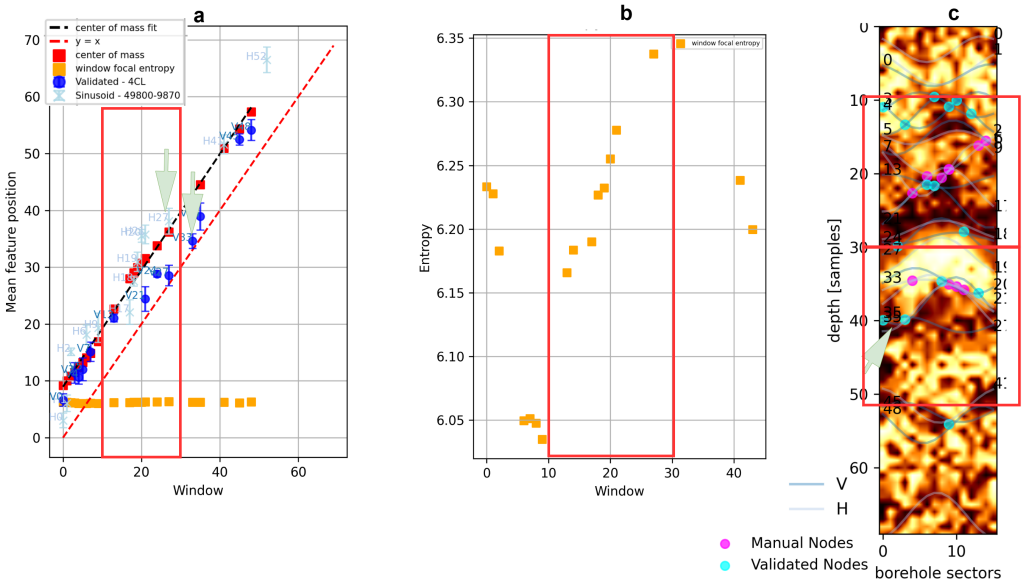

Figure 2.68: A comprehensive plot comparing the manual (H) and semi-automated (V) analyses for a high-contrast interval after NN fine-tuning based on validation. (a) The validated (V-blue dot) and human (H-light blue cross) detected features are plotted, with the x-axis representing the window index in which the feature was detected and the y-axis showing the mean depth position of the feature in samples. The length of each bar is proportional to the feature's amplitude. Red squares represent the y-coordinates of window centers of mass, and the black dotted line indicates the linear regression of the set of centers of mass across all considered windows. A red dotted line is added for reference, representing $y = x$. Green arrows highlight two features that can be identified in (c). (b) The focal entropy-window plot, where the red rectangle highlights a positive entropy trend from window 10 to 30. (c) Validated and manually picked sinusoids, with blue dots and magenta dots representing nodes of the validated and manual feature sets, respectively.

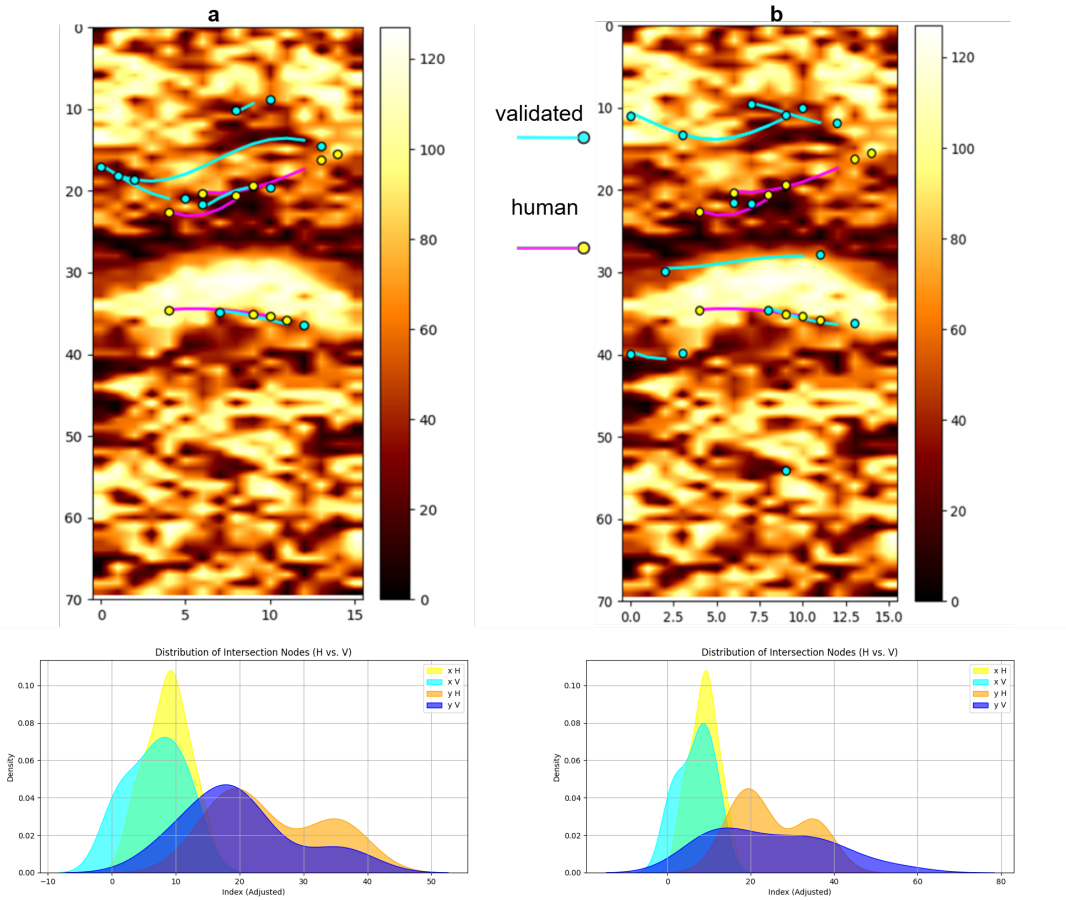

Figure 2.69: Nodes and branches distributions from the (a) manual and validation on orignal NNs and (b) manual and validation on fine-tuned NNs using the validations. Below each BHI, the spatial (x,y) distribution of manual (xH, yH) and machine (xV, yV) x nodes.

Collecting those results for several interpreters and comparing them for subintervals with varying levels of complexity might be useful to highlight some important components of the human and the machine level performance and gain relevant insights on human and machine perception that would help quantifying the subjective bias and assess the impact of the machine assistance for such task.

2.3.6 Geological interpretation

Two wells in the LWD dataset (LWD1, LWD3) introduced in section 2.3 were associated with the presence of Mass Transport Complex (MTC) crossed by the borehole, which caused borehole stability issues in one of them, leading to material and economic loss. To further validate the proposed methods for automating borehole image log analysis and explore their potential in providing interpreters with meaningful geological insights beyond the mere feature predictions, we utilized their results to derive quantities at the feature and interval levels. This allowed us to examine whether these derived quantities could help distinguish intervals within the MTC from those outside it. In this section we will see how

we addressed this challenge using both the CVML and the DL-based methods, and will describe the results. Note that for the DL method, the described results refer to a standalone application of the method, considering all the predictions by the method, not just the validated ones. (Fig. 2.18).

CVML method

We collect the main implicit information from the CVML method as a confidence score calculated using the Sobel map information. For each curve predicted by the algorithm in the LWD data, we measure a score as the percentage of points along the curve that share the same sign of the sobel value in the map. At this point, we simply analyze the distributions of confidence and the dip magnitude (derived from the sinusoidal parameters) inside and outside the MTC intervals interpreted by geologists (see Fonnesu et al. (2020) for a more detailed description of the MTC (Fig. 2.70, 2.7 and Tab.2.6, 2.7).

Table 2.6: DIP and Confidence statistics for LWD3

Metric	**Interval**	**1st Quartile**	**Average**	**3rd Quartile**
DIP	Inside MTC_370	33%	47%	61%
	Outside MTC_370	37%	49%	60%
Confidence	Inside MTC_370	33%	47%	62%
	Outside MTC_370	42%	47%	55%

Metric	**Interval**	**1st Quartile**	**Average**	**3rd Quartile**
DIP	Inside MTC_370 (No Instability)	51%	56%	63%
	Instability	65%	63%	68%
	Outside MTC_370	42%	50%	59%
Confidence	Inside MTC_370 (No Instability)	47%	55%	63%
	Instability	59%	65%	70%
	Outside MTC_370	37%	48%	60%

Table 2.7: DIP and Confidence statistics for LWD1

One of the two boreholes that are associated with this geological interpretations, also incurred in borehole stability issues within the interpreted MTC interval (LWD1). For this well, we further observe the distribution in the instability zone (Fig. 2.71, Tab 2.7). The results inspired the use of a very simple rule to use the information at the event (feature) level (sobel confidence score and dip magnitude) to extract information at the interval level. The rule consists in classifying a 10-meter interval into the MTC/NON-MTC categories if boths the percentage of events with a score higher than 75% and a dip magnitude higher than 45° exceed 50%. We evaluate this as an additional feature of the CVML tool that can assist interpreters in real-time, with some geologically meaningful information of the drilled intervals. This additional part of the method would consist of an alert system with two alert levels, whether the rule condition is met for only one or

both the derived parameters. Finally, we compare the alerting intervals and the level of alert with the interpreted MTC depths in the two wells (Fig. 2.72- 2.73) and the results are shown in Tab. 2.8. We see that the level 2 alert is the weakest in both cases. However we notice that the level 2 alert succeeds in identifying the instability zone with an 80% accuracy in the problematic interval (Fig.2.73).

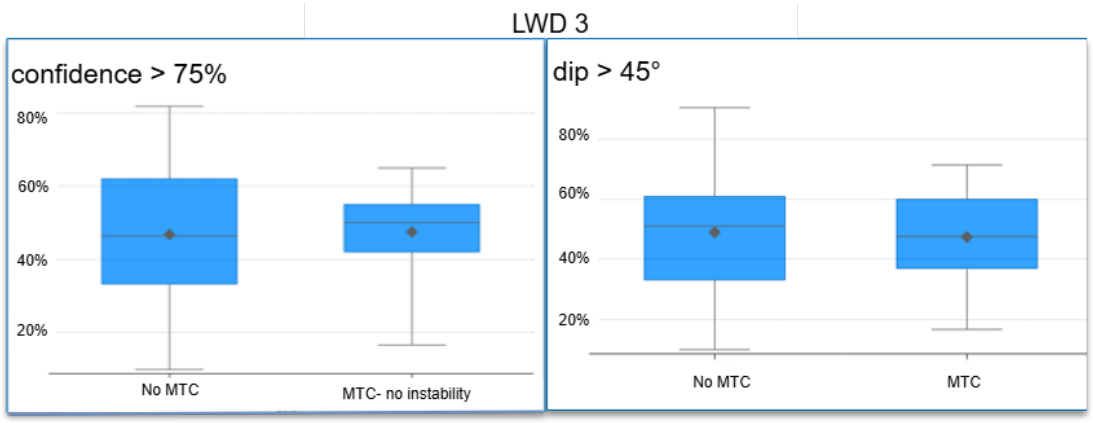

Figure 2.70: Confidence and dip magnitude distribution inside and outside the MTC for LWD3. In this borehole, there is a slightly different distribution of confidence and dip magnitude of the CVML features.

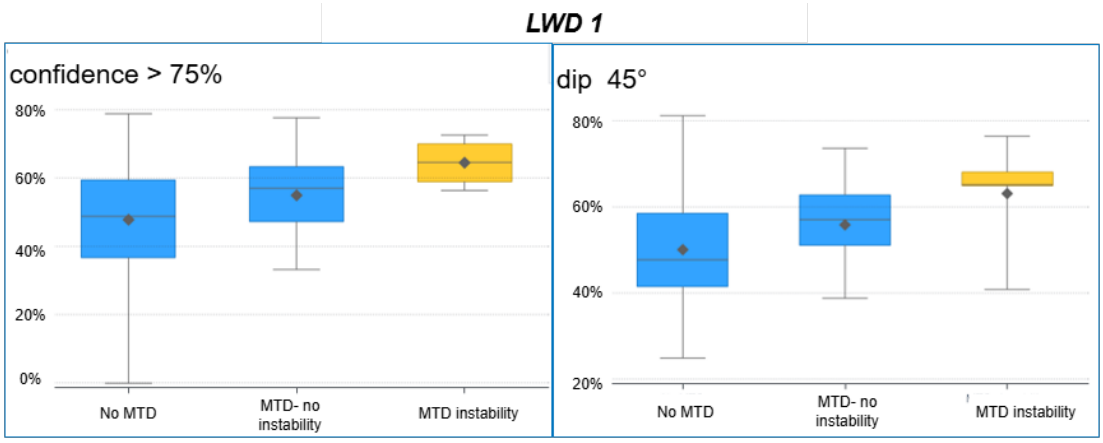

Figure 2.71: Confidence and dip magnitude distribution inside and outside the MTC (and the instability zone) for LWD1. DIfferently from the LWD3 case, here we see clearly different distribution of confidence and dip magnitude values of the CVML features in and out the MTC. In particular, the interval where instability issues occurred (yellow box).

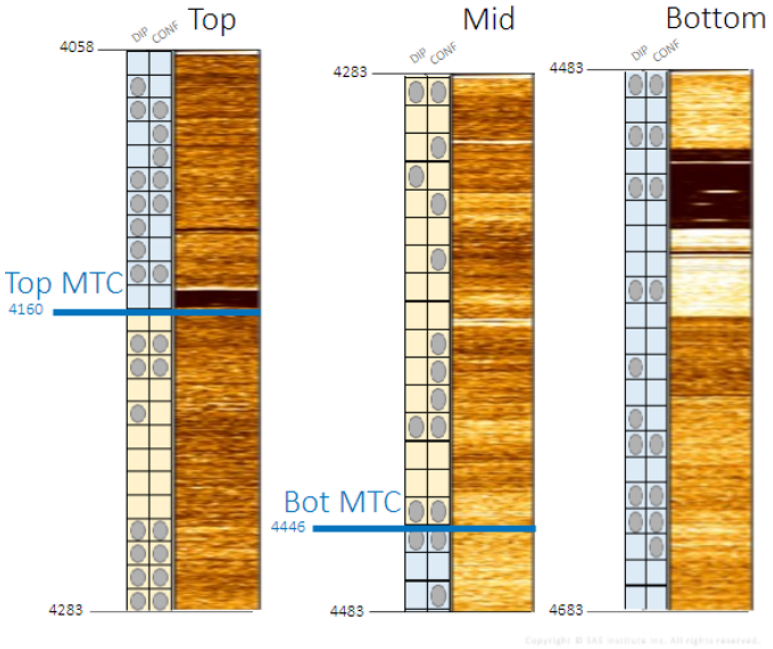

Figure 2.72: The distribution of the alerting intervals depths compared to the interpreted MTC depth for LWD3

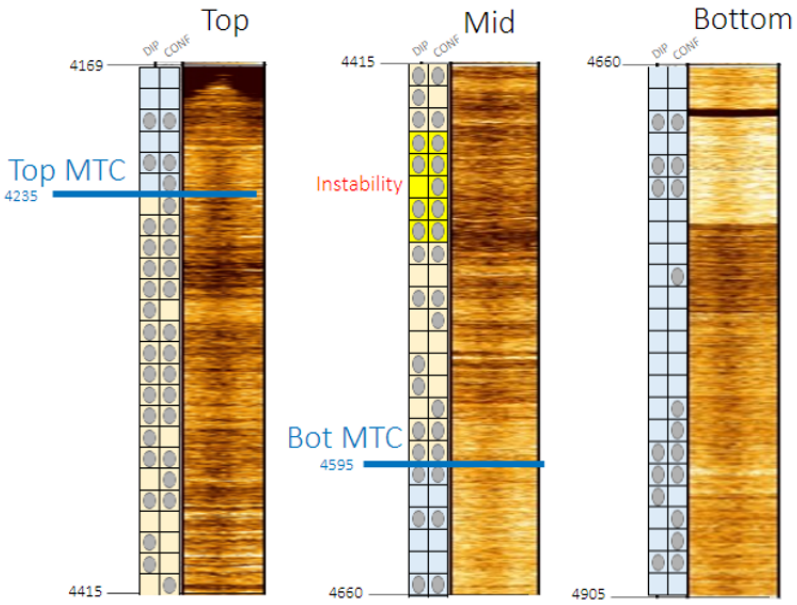

Figure 2.73: The distribution of the alerting intervals depths compared to the interpreted MTC depth for LWD1

Test	Alert level	Precision	Recall
LWD3	1-DIP	0.3929	0.3929
	1-CONF	0.4839	0.5172
	2	0.4286	0.3103
LWD1	1-DIP	0.7000	0.7778
	1-CONF	0.6136	0.7297
	2	0.6563	0.5676

Table 2.8: Precision and Recall for the three levels of alert in LWD3 and LWD1. As expected from the distributions of confidence and dip values, the performance of the classification rule based on confidence and dip magnitude of the CVML features are better in the LWD1 borehole.

DL method

In section 2.2.3 we described the baseline Fitnet architecture as a combination of convolutional layers, dropout layers and a fully connected layer for the sinusoidal fitting of edges in the binary input image. We also described the role of dropout layers during training and prediction steps. In addition to this uncertainty estimates, we calculated the apparent dip magnitude of the predicted sinusoidal features in $\mathbf{F}_{DL}$ and integrated these other information in the same dataset with segmentation maps, and the density values in the BHI at the feature location, for each predicted feature (Fig.2.74). After integration, we conducted a Principal Component Analysis (PCA) to reduce the dimensionality of the dataset. The dataset was reduced to two axes corresponding to the first two principal components, which capture the maximum variance within the data. By doing so, we are essentially reducing the dataset generated by a standalone application of the method to a binary classification problem (MTC/Non-MTC) with two features capturing the maximum variance in the integrated dataset, to overcome the redundancy in the original dataset, as is determined by Fitnet implementation. We then proceed with the binary classification using an unsupervised learning method as the Agglomerative Clustering (AC) technique, imposing a 2-cluster separation of the data points in PCA space. We used a cosine-similarity metric to achieve the cluster division, since it maximized the silhouette scores of the clusters (Fig. 2.75).

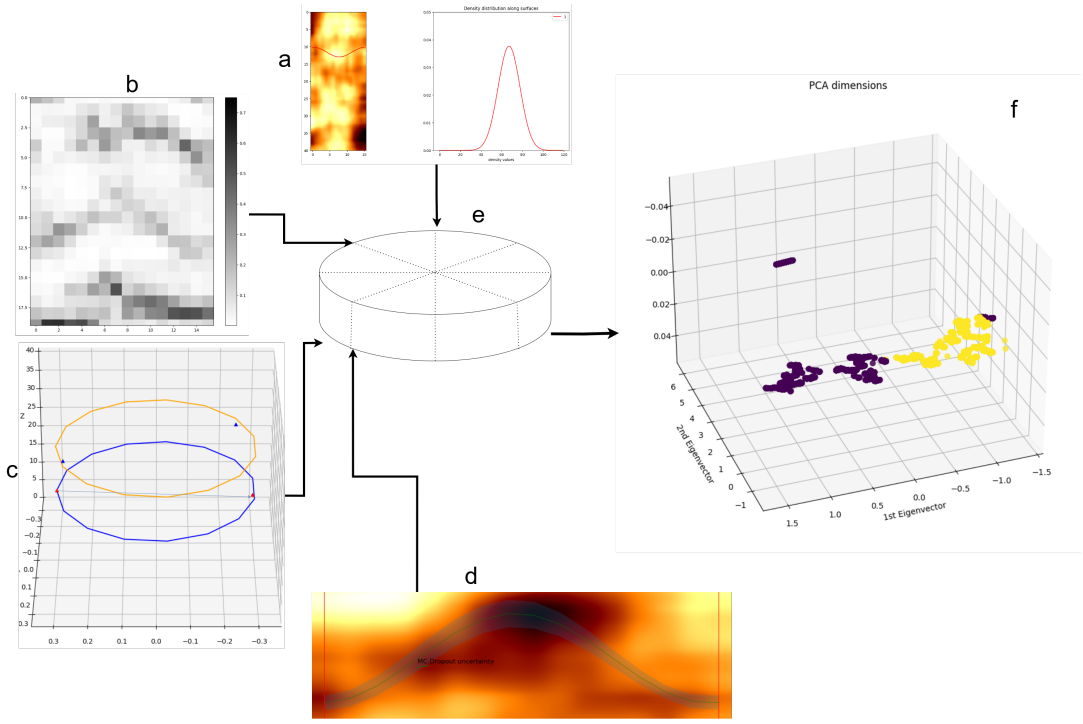

Figure 2.74: a: distribution of petrophysical parameters along the features. b: segmentation maps. c: calculated dip magnitude for each feature. d: dropout uncertainty estimate for each feature. f: PCA for dimensionality reduction of the integrated dataset.

Finally we compared the AC resulting cluster separation with the manual MTC detection for both LWD1 and LWD3 (Fig. 2.76), and compared the results in therm of Precision and Recall, to also evaluate the sensitivity towards the MTC class (Tab. 2.9). The success of the clustering approach is heavily dependent on the presence of underlying patterns in this reduced deep learning-derived dataset. For both boreholes, the clusters align well with the manual MTC and NON-MTC classifications. Furthermore, the binary agglomerative classifier exhibits high sensitivity toward the "alert" (positive) class in both boreholes. This means it could effectively identify critical intervals for those two boreholes, which is particularly crucial in a safety-critical application like this. Interestingly, less precision is observed for the LWD3 borehole image, which correlates with the CVML results (see Tab. 2.8). In this case however, the different results may be related to the machine perception of the field data.

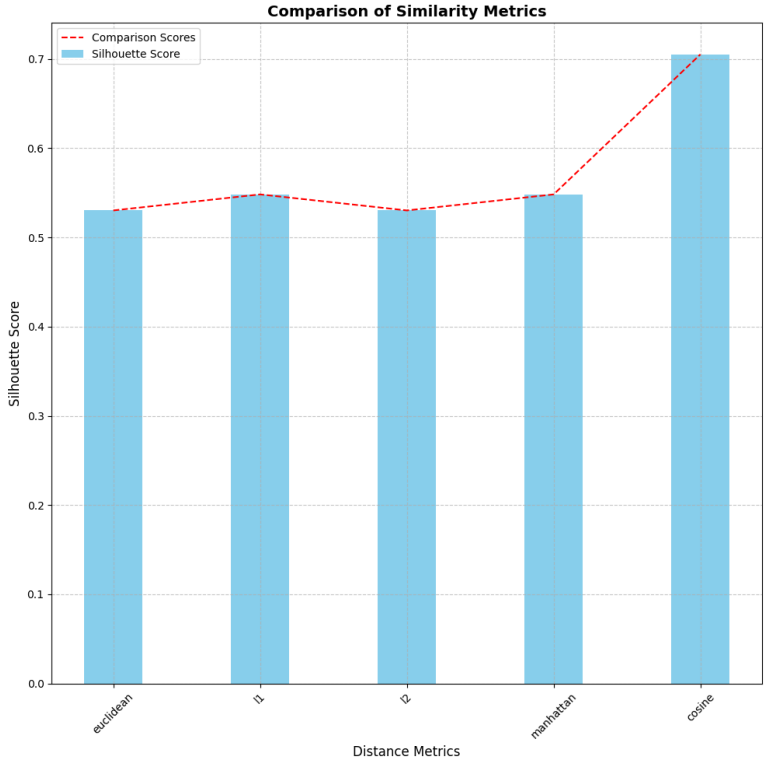

Figure 2.75: The figure compares silhouette scores across different distance metrics: Euclidean, L1, L2, Manhattan, and Cosine. The bar heights represent silhouette scores (blue), while the red dashed line shows the comparison scores for each metric. Cosine achieves the highest silhouette score, indicating superior performance in clustering similarity compared to other metrics.

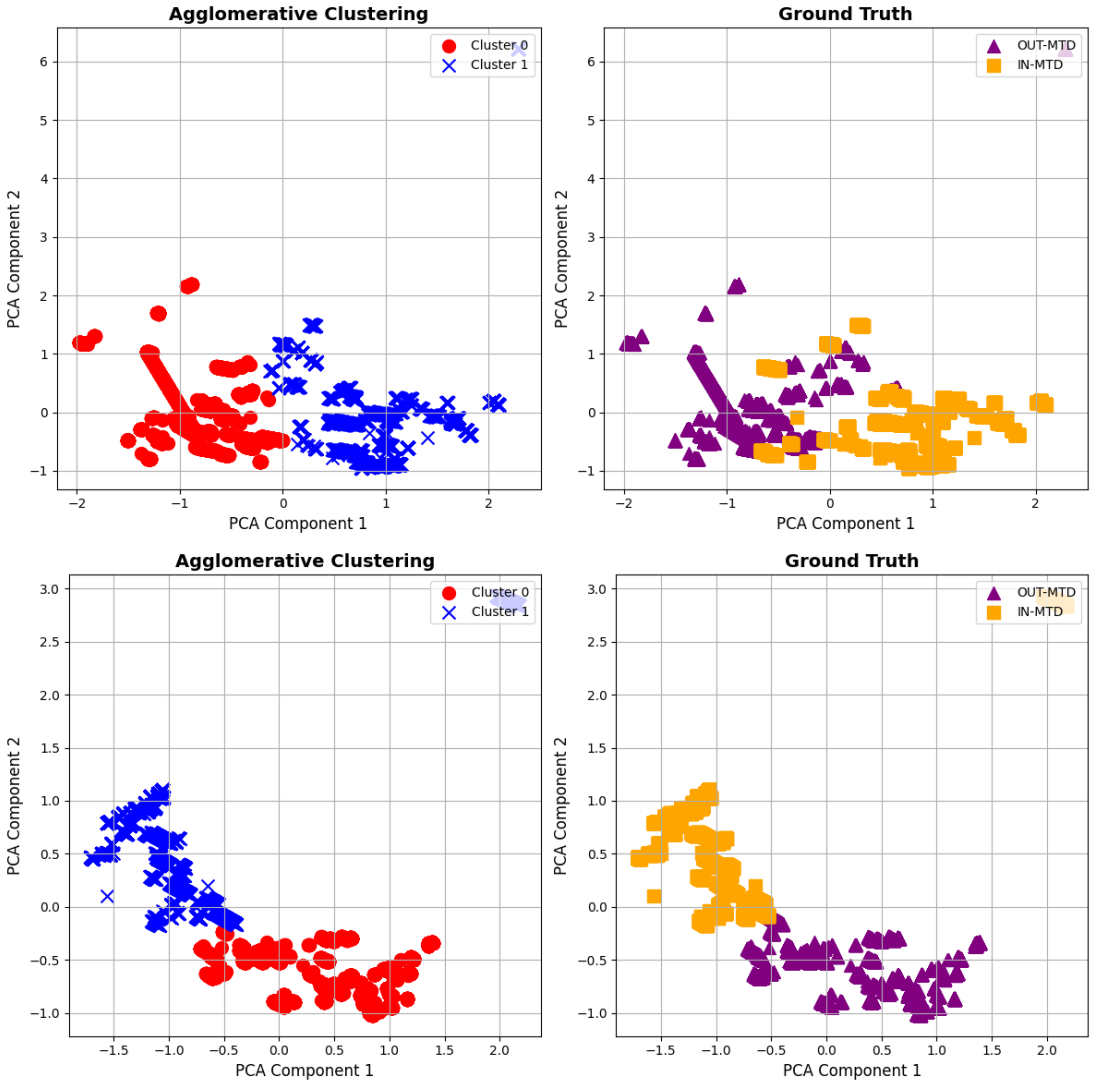

Figure 2.76: Up: agglomerative clustering results for the LWD3 BHI on the PCA space. Down: agglomerative clustering results for the LWD1 (instability) BHI on the PCA space.

	Precision	**Recall**
LWD3	0.77	0.91
LWD1	0.96	0.99

Table 2.9: Precision and Recall for the two boreholes illustrated in Fig. 2.76.

Fig. 2.77 shows the panel standard deviation (σ_{segm}) (as in Fig 2.51) in a scatter plot with the variance of the associated input interval (σ_{raw}) (see Fig. 2.62). Both CL-Picknet and SL-Picknet respond differently to the two field datasets. This is also true for the CVML method (sect. 2.3.1), in the sense that the distribution of the CVML implicit information were sharply different between LWD3 and

LWD1, may be related to the fact that the borehole stability issues only occurred in LWD1.

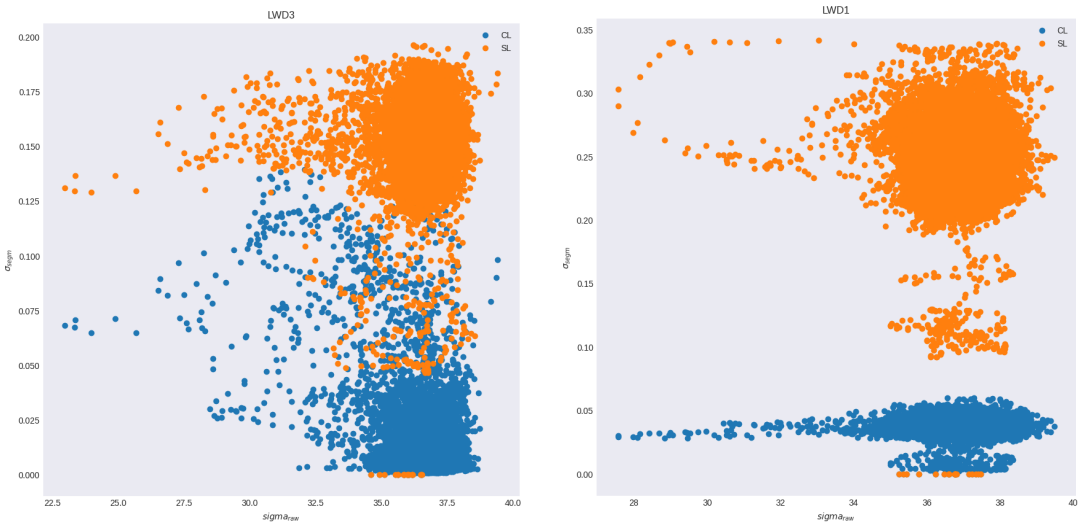

Figure 2.77: Panel standard deviation (σ_{segm}) as a function of the input variance (σ_{raw}) for (right) LWD3 and (left) LWD1.

2.3.7 Scalability

A scalable method is a method that maintains its level of performance for a varying- usually growing- size of the problem. In that sense, the system may be considered scalable if it can upgrade to adapt to larger problems, or reducing the problem size minimizing the performance quality reduction. We apply this scalability concept to our problem. As anticipated, we formulated the synthetic data-based training of the network and their architectures according to the traditional sampling parameters of LWD borehole imaging tools. The resulting images, however, are low-resolution images (vertical sampling $\approx$ 0.05 m) compared to the wireline images, such as the FMI (vertical sampling $\approx$ 0.005 m) that are more used in the geothermal exploration industry. Referring to LWD borehole images to drive the choice of out synthetic images parameters, the implementation of a meter-scale automated analysis approach was eased by the resulting image size of 20 depth samples and 16 borehole sectors. One meter of FMI-like synthetic image would therefore take 200 samples in depth, and the borehole sectors are also in the order of hundreds, so that the 1-meter synthetic borehole image size could be 200x660. Moreover, a greater resolution in the real world images also means a higher number of hierarchical features to be encompassed by the synthetic images generator and a greater strive in achieving the synthetic dataset required representativeness. For instance, the FMI borehole images usually come with vertical blank strips generated by the seams between the arms of the tool, that are not present in LWD data. To test the scalability of the proposed DL method, we applied it to a set of 115 FMI images representing 2-meter intervals, where each FMI interval is associated with one fracture, resulting from the manual interpretation of experienced petrophysicists. Some fractures are full,

others are only partially visible in the image. Each image in the FMI dataset is a 1000x240 image. To resize the full images to our method input size and wanting to perform the automated analysis at the meter scale, we considered sample-wise sliding windows with size 500x240. Resizing those images to 20x16 means to work with a 99% of resolution loss and a distorted image, as the aspect ratio varies from 2,08 to 1,25. We then applied the optimal matching algorithm introduced in section 2.3.4 to find the prediction across all the windows that best matches the manually picked fracture. In this way, we aim at speeding up the semiautomation process, using the human picking as a validation term. Despite the image degradation related to resolution loss and distorsion, this test showed that method could effectively approximate the human results, in particular for those examples with fully represented sinusoids associated with higher contrasts (see Fig. 2.78). However, in cases the sinusoids are fully represented but very subtle or with other more prominent features crossing, the negative impact of image degradation is evident (see Fig. 2.79). Another case where the human outcome was unmatched, is the vertical fractures case, where essentially the fracture coordinates exceed the image bounds (see Fig. 2.80). These observed inaccurate results prompt the question of whether adjusting the window length for resizing could lead to improved outcomes. We modified the window length to the full image vertical size (1000 samples), and compared the results (Fig. 2.81. What we observe is a decrease in the MSE from the two fractures, which goes from 560,79 to 63,35.

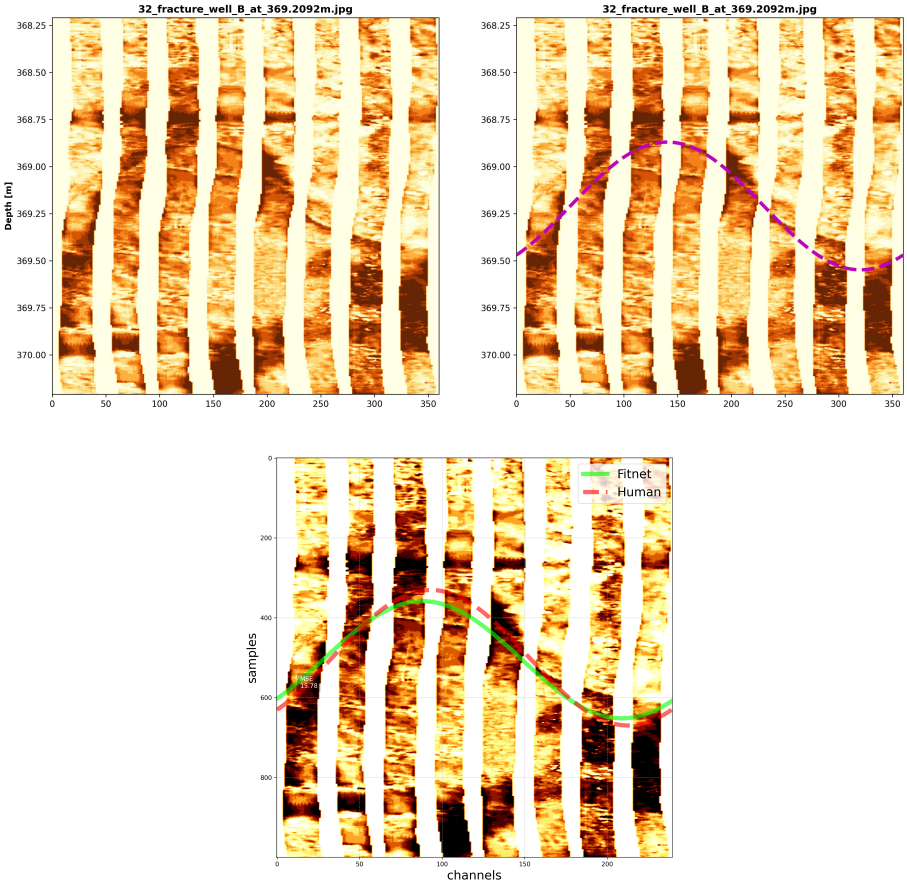

Figure 2.78: Instance n.32 in the Conoco Phillips FMI dataset used in the scalability test.

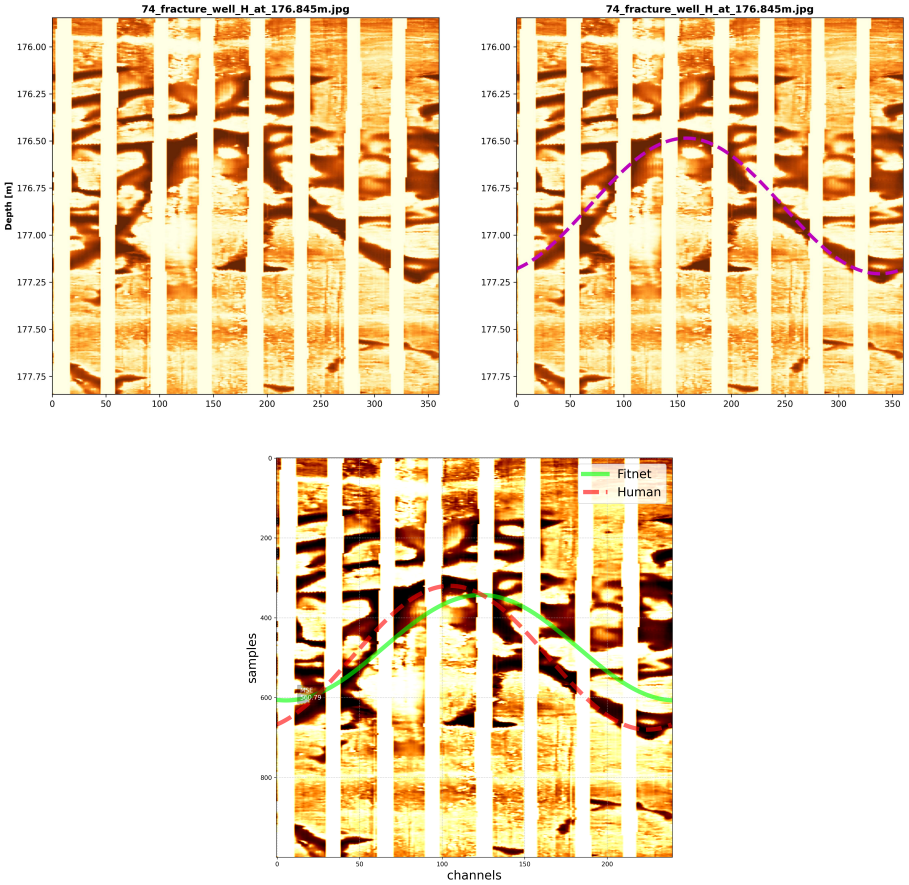

Figure 2.79: Instance n.74 in the Conoco Phillips FMI dataset used in the scalability test.

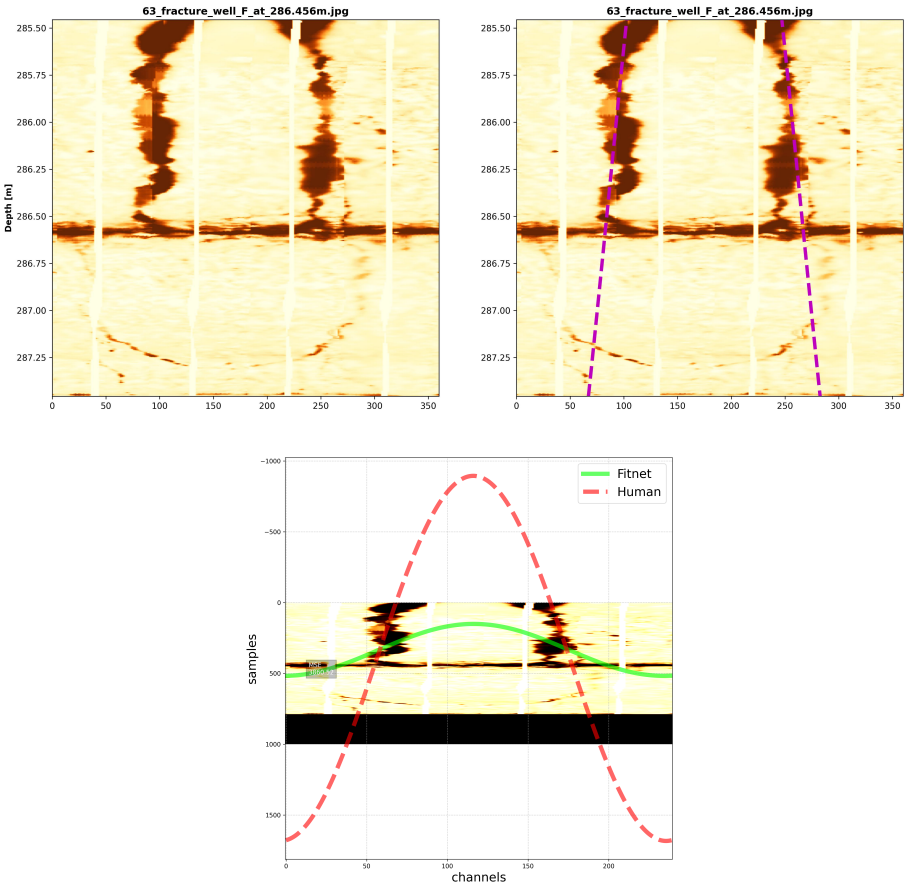

Figure 2.80: Instance n.63 in the Conoco Phillips FMI dataset used in the scalability test.

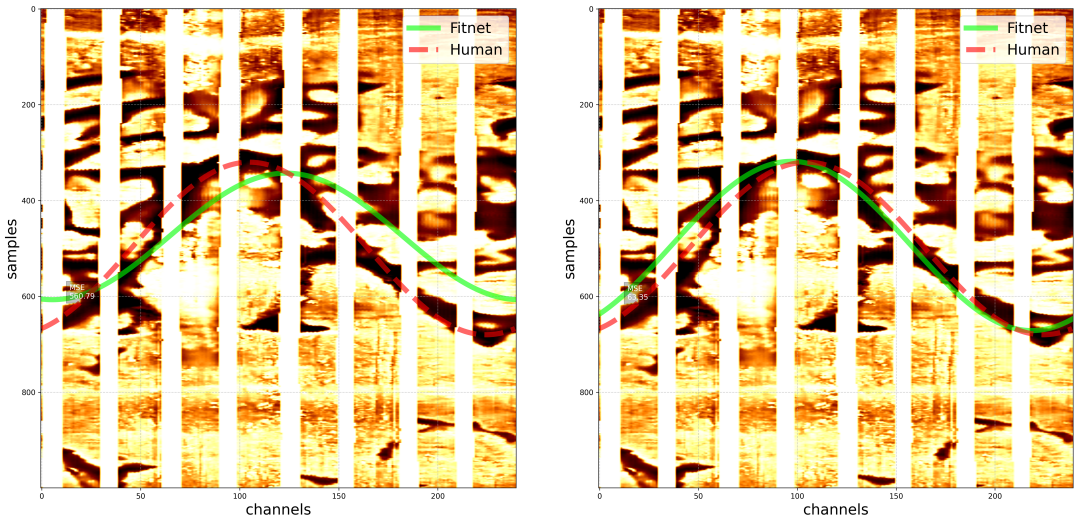

Figure 2.81: Instance n.74 in the Conoco Phillips FMI dataset used in the scalability test. Left: window size of 500 samples. Right: window size of 800 samples.

This observation underscores the potential of a multi-scale approach facilitated

by the window-sliding technique, which could impact results and better align the predictions with human intent. For instance, we could reverse the process and analyze how the results vary within the same interval as depicted in Fig. 2.81, focusing on another distinct but less-represented fracture in the image (Fig. 2.82). This fracture intersects with the one previously enhanced by extending the window length, providing additional insights into the influence of window size. In this case, the best result is obtained for the 500 samples window (MSE = 187.95), compared to the 800 samples window (MSE = 230.14).

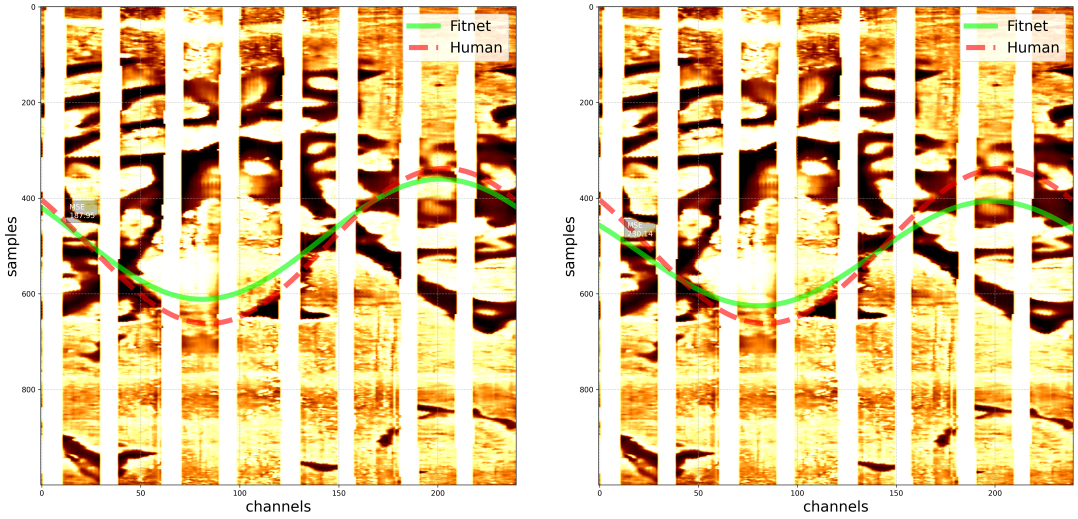

Figure 2.82: Instance n.75 in the Conoco Phillips FMI dataset used in the scalability test. Left: window size of 500 samples. Right: window size of 800 samples.

In general, we observed that the size of the window can impact the quality of the best-matching prediction depending on the dipping angle pf the surface. Let's for instance consider window sized of 200, 500, 800 and 1000. For each window size we find the best matching prediction across all windows dividing the original interval, and compute the MSE. Let's for example, consider the instances indicated with indexes 80, 81, 82 (Fig. 2.83) . (Fig. 2.83)-a shows the dipping angle of the three selected fractures, which represent low-angle (80), medium-angle (81) and high-angle (82) fractures, respectively. The minimum MSE values for examples 80 and 81 were observed at a window length of 200, while for example 82, the minimum MSE occurred at a window length of 1000. This highlights the influence of window length, which reflects the scale of automated analysis, on the final results. These findings guided the application choices detailed later for the WNB case study. However, it is important to note that MSE can sometimes be a misleading metric for evaluating these results. For instance, in example 81, the minimum MSE was identified at a window length of 200 samples, yet visually, the best results were evidently achieved with a window length of 500 samples.

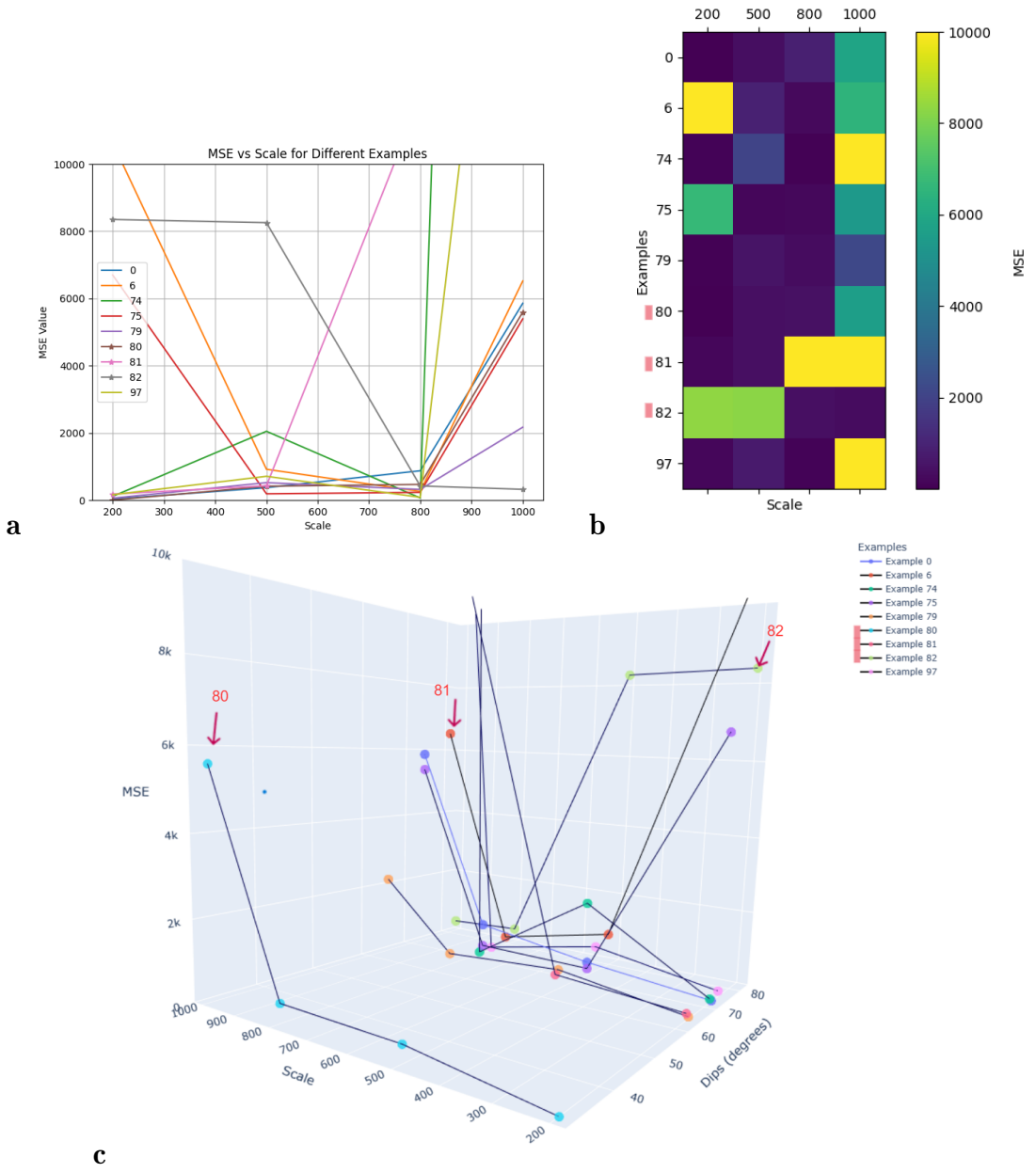

Figure 2.83: a) Plot of MSE vs. scale; b) Heatmap of MSE across different scales. Examples 80, 81, and 82 represent low-angle, medium-angle, and high-angle fractures, respectively. Notably, for the high-angle fracture, the MSE decreases as the prediction depth window length (in samples) increases. c) 3D plot showing the relationship between scale, MSE, and dip angles for different examples, with points connected by lines to highlight MSE trends within each example. Samples 80, 81, and 82, representing low-angle, medium-angle, and high-angle fractures, respectively, are marked by red arrows. Notably, for the medium- and low-angle fractures (examples 81 and 80), the minimum MSE corresponds to a window length of 800, whereas for the high-angle fracture (example 82), the minimum MSE occurs at the full length of the interval.

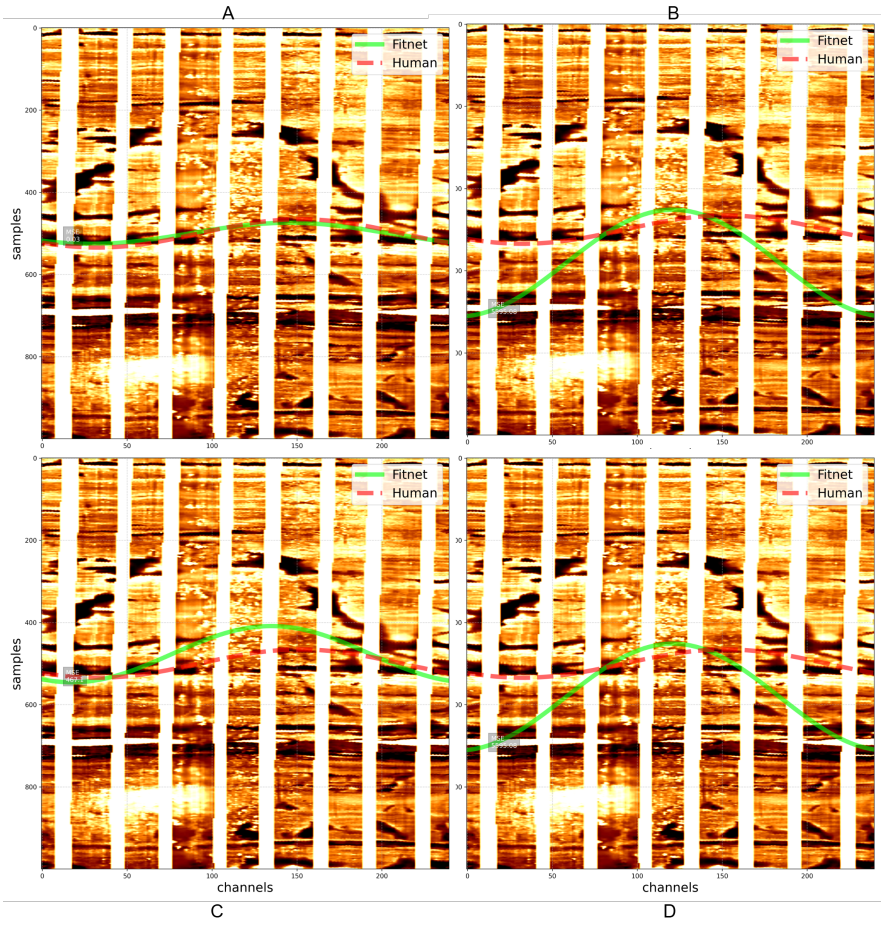

Figure 2.84: **Example 80. A)** Predictions using a window length of 200 samples. **B)** Predictions using a window length of 500 samples. **C)** Predictions using a window length of 800 samples. **D)** Predictions using a window length of 1000 samples.

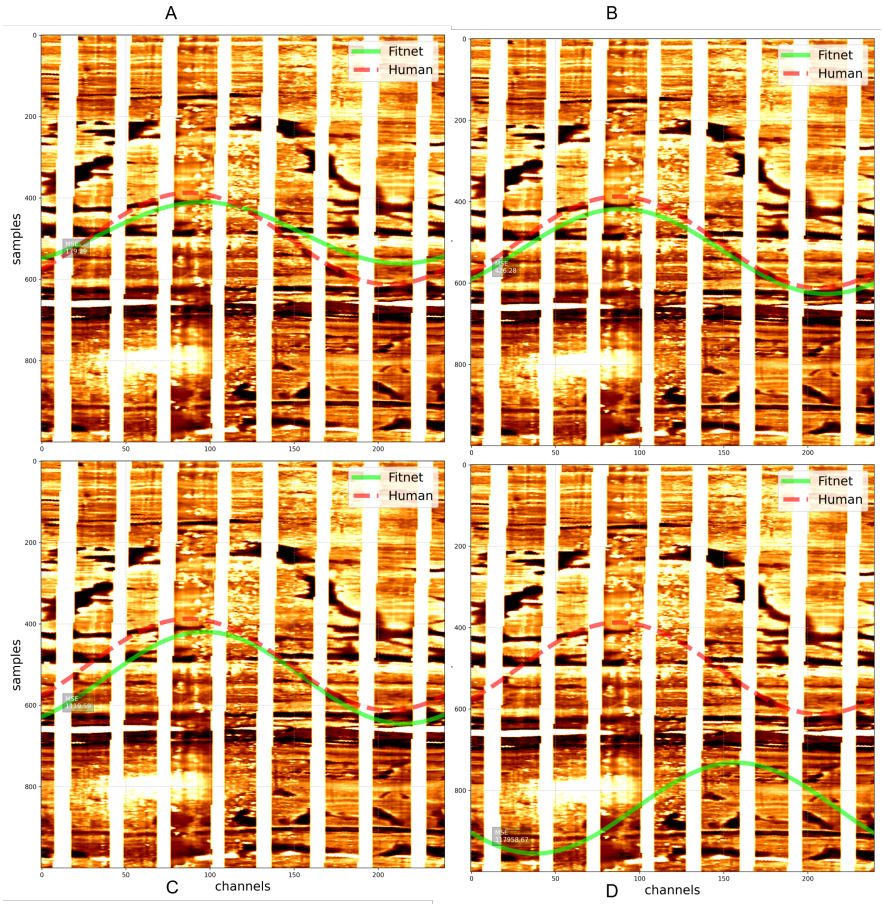

Figure 2.85: **Example 81. A)** Predictions using a window length of 200 samples. **B)** Predictions using a window length of 500 samples. **C)** Predictions using a window length of 800 samples. **D)** Predictions using a window length of 1000 samples.

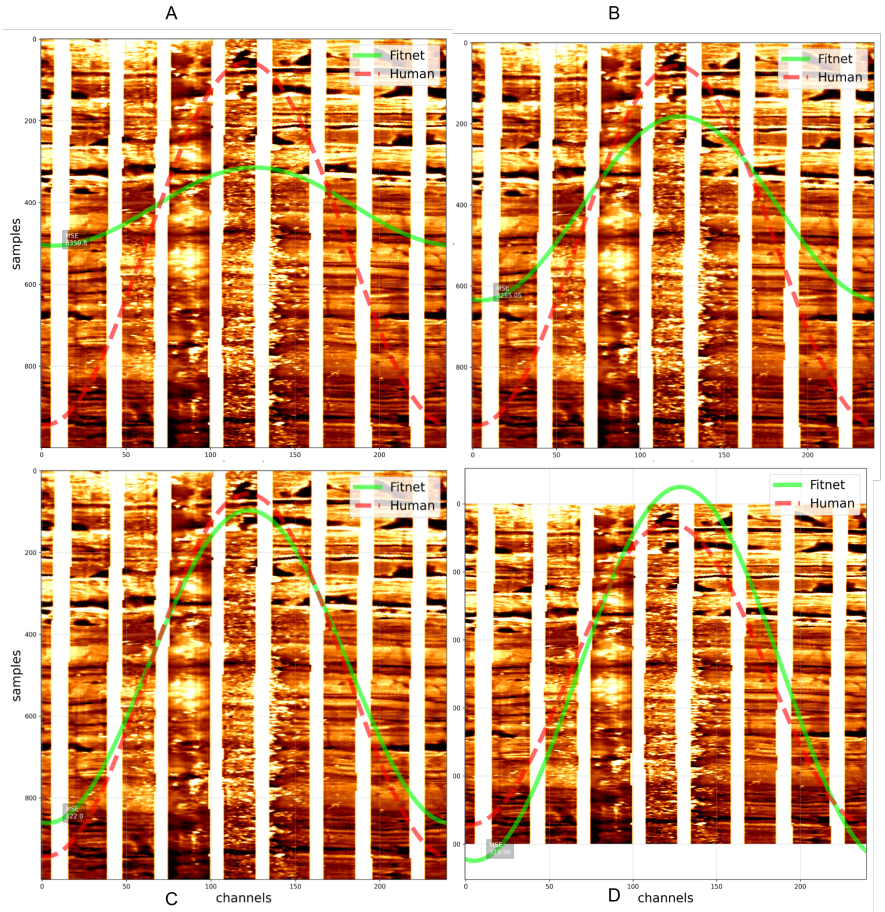

Figure 2.86: **Example 82. A)** Predictions using a window length of 200 samples. **B)** Predictions using a window length of 500 samples. **C)** Predictions using a window length of 800 samples. **D)** Predictions using a window length of 1000 samples.

These results highlight the adaptability of the NNs used in this study, that can be used on any kind of BHI. Nevertheless, this approach for the method upgrading does not prove its full scalability. For example, we expect that the downsampling of the input image and the blank strips negatively affects the reliabilty of the confidence score (section 2.3.3) as we discussed it for LWD datasets. Furtherly, we adapt the method using the same logic on FMI BHI from one well (NLW-GT-01) geothermal exploration project in the West Netherland Basin (WNB). Besides being geothermal exploration data, the particular interest on this data rises as they have been interpreted by multiple interpreters. We used this FMI data in the attempt to apply the proposed semi-automation method and effectively reproduce the results of a manual interpretation calibrated with core-data. We proceed with a visual comparison, as we did not have access to any comparable format of the interpreted fractures in these data to conduct a quantitative evaluation. Also for this WNB high-resolution images, we had to lose more than 90% resolution in the image, and we entrust the NNs to exploit the relevant information in the degraded FMI windows to provide the human with acceptable guesses at multiple scales (Fg. 2.87). The validated fractures could significantly

approximate the manual interpreted ones (Fig. 2.88). These two cases represent a set of fully represented small conductive fractures and a conductive fault with variable ranges of aperture, that can contribute to the permeability enhancement of this interesting formation for geothermal exploration in the WNB. However, this conservative interpretation focuses on the fractures visible in the core, but other discontinuities in the FMI data have been identified as either hydraulically or drilling-induced fractures. These additional features may also be detected by the semi-automation procedure(Fig. 2.89). Fractures are not the only class of geological features we can recognize and detect in a FMI. For instance, the first clear geological contrast that we can observe at approximately 4255 m depth or sample 200 in Fig. 2.88 and Fig.2.89 is a *bedding*. During the validation procedure, we noticed that some predictions were accurately suggesting the depth of those events, especially at the cm- scale (Fig. 2.91).

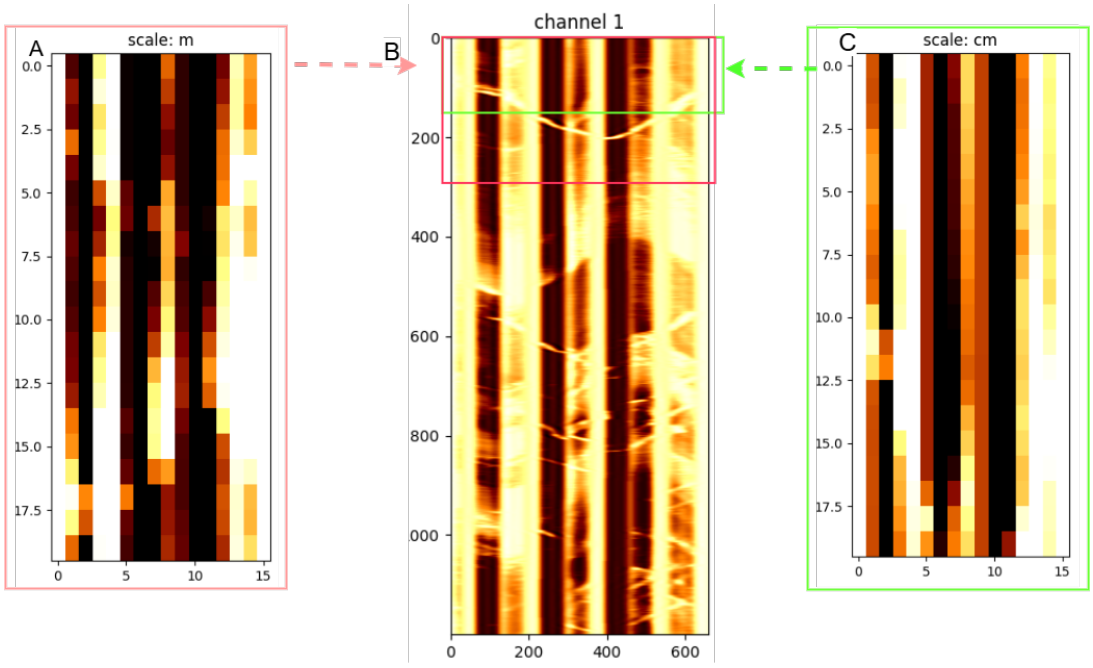

Figure 2.87: A) Resized image corresponding to the red box in B. B) An interval of 1200 samples (≈ 3.40 meters) showing the 3 FMI intervals. C) Resized image corresponding to the green box in B.

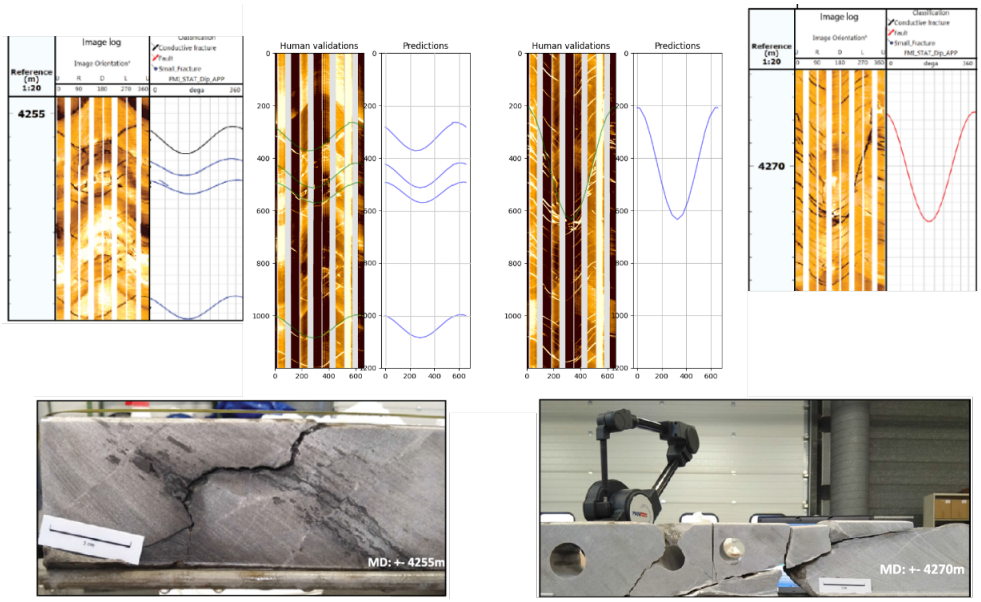

Figure 2.88: Comparison of the interpretations published by (Boersma et al., 2021) with the validated DL predictions. The interpretation in the reference paper is described as a conservative interpretation with fully represented sinusoids that we successfully replicated through our scalability assessment.

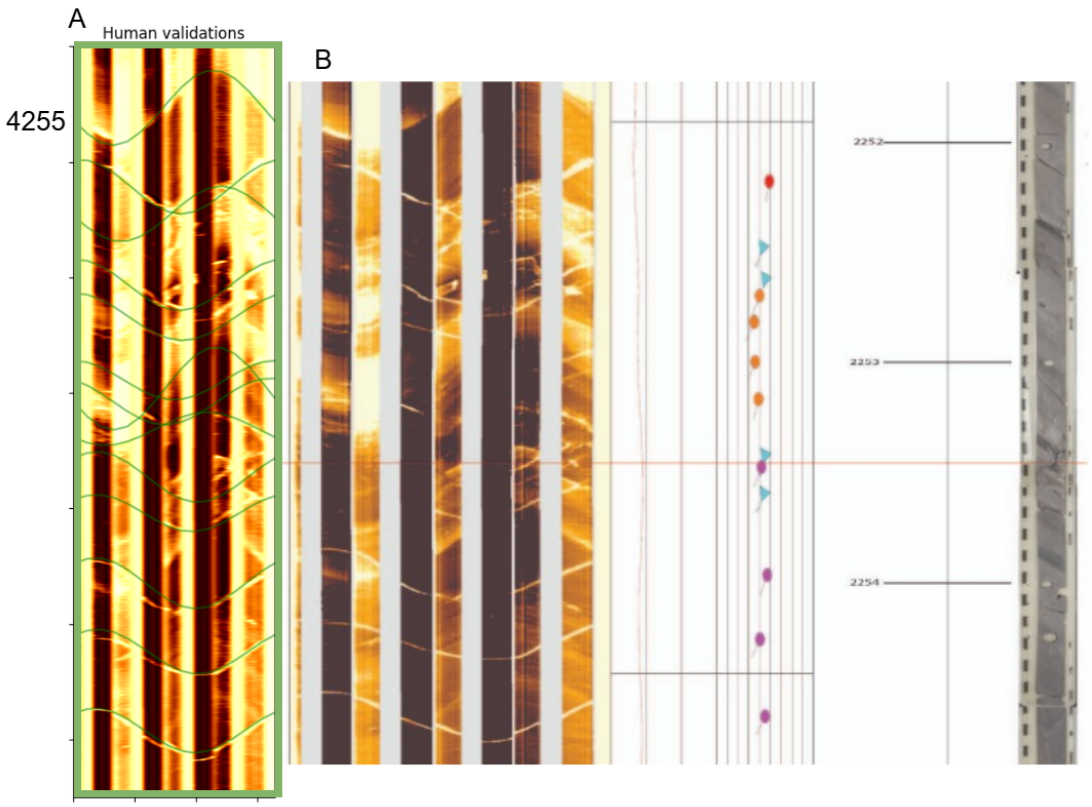

Figure 2.89: A) Validated deep learning (DL) predictions using a less conservative approach reveal discontinuities attributed to drilling-induced fractures that are absent in core data. B) The less conservative interpretation is visualized with tadpoles and core data for the interval starting at a depth of 4255.

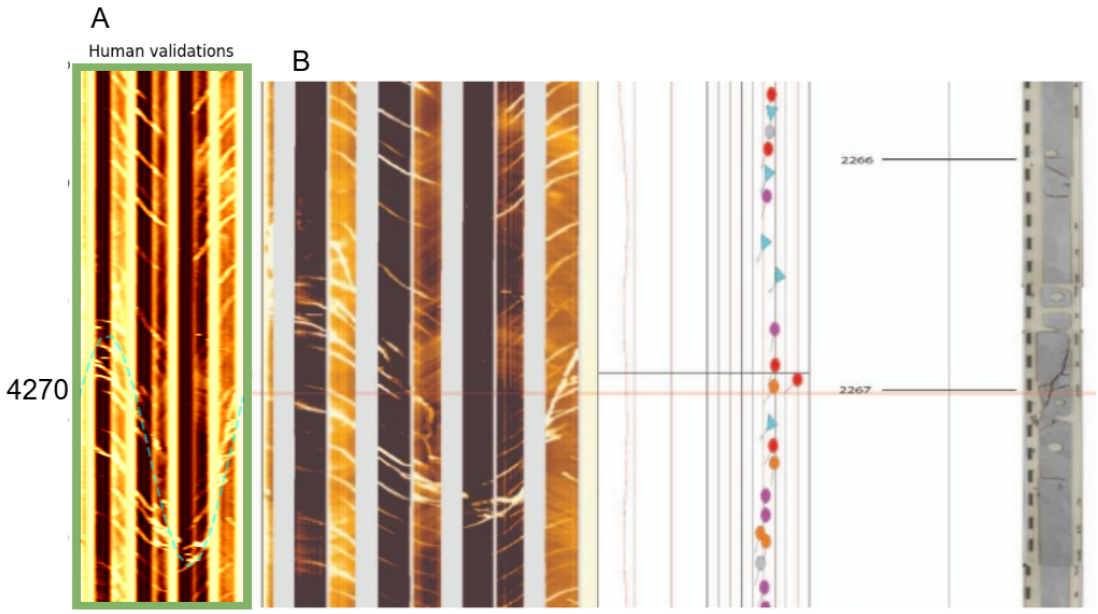

Figure 2.90: A) Validated deep learning (DL) predictions replicating the conservative approach in (Boersma et al., 2021) that revealed a high-dippping fracture visible in core data. B) The less conservative interpretation is visualized with tadpoles and core data for the interval starting at a depth of 4270.

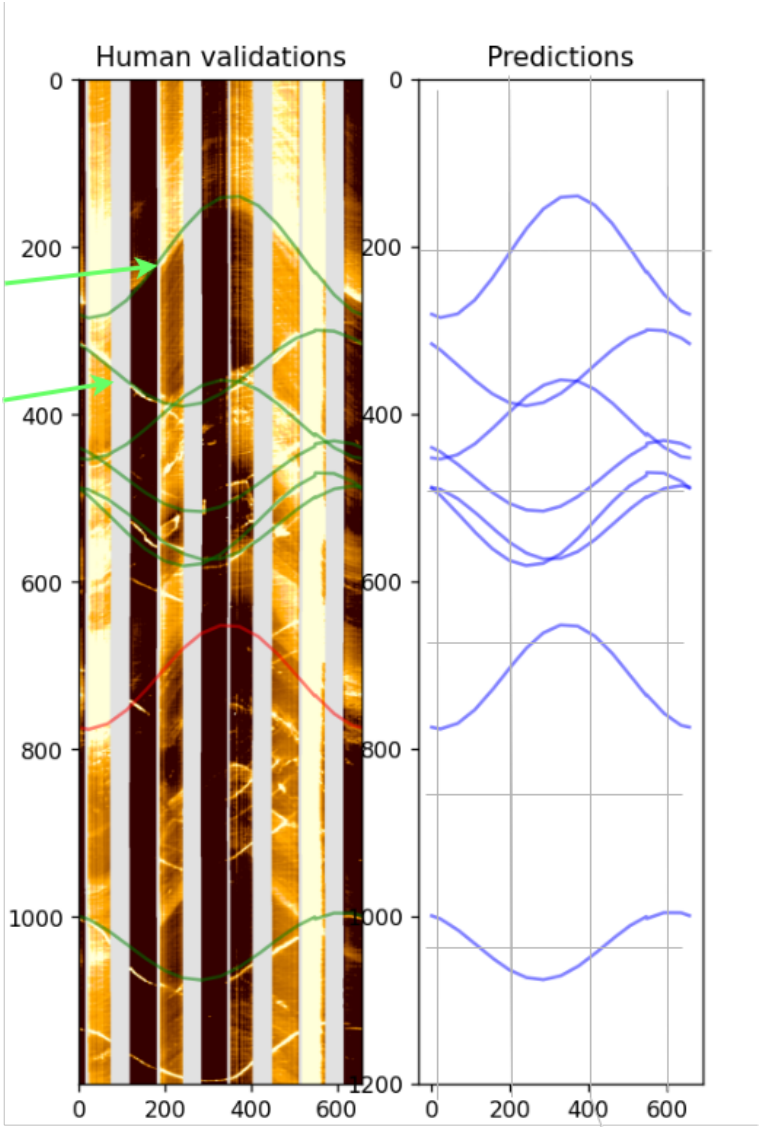

Figure 2.91: Validated DL predictions using a less conservative approach identify bedding and fractures at both the centimeter (green) and meter (red) scale. The sinusoids indicated by the green arrows are the one considered in Tab.2.10.

We computed the $\Delta\rho$ value for the bedding and its closest fracture (marked by green arrows in Fig. 2.90) as in eq. 2.41, for four different accurate suggestions (A,B,C,D) of these same two events detected in different subsequent windows (Tab. 2.10). What we observed, was a general higher $\Delta\rho$ for the bedding plane, which correlates with the results we observed on the LWD density azimuthal dataset (see section 2.3.3). This denotes the validity of such formulated parameter as an additional information of the predicted features also in the higher resolution and the different petrophysical parameter measured in this FMI data, compared to the LWD data (see Fig.2.54-2.56).

With regard to the CVML methodology, the scalability tests were more straightforward, as the only parameters that needed to be adjusted in the algorithm to adapt it to FMI images were the step length H and the number of borehole sec-

$\Delta\rho$	Bedding	Fracture
A	33.09	-3.68
B	35.98	-2.05
C	39.22	-6.06
D	42.19	1.23

Table 2.10: A table showing the relationship between FMI change (Δ) associated with bedding, and fracture.

tors (Section 2.2.2, Eq. 2.18). As expected, there are several reasons why the CVML method fails on the same FMI interval shown in Fig. 2.90. Fig. 2.92 illustrates that the set of correlated features is negatively impacted by the presence of blank strips, which are typically present in FMI images. Furthermore, the computational time for processing this interval exceeded 5 hours, whereas the prediction time for the DL method remained unchanged.

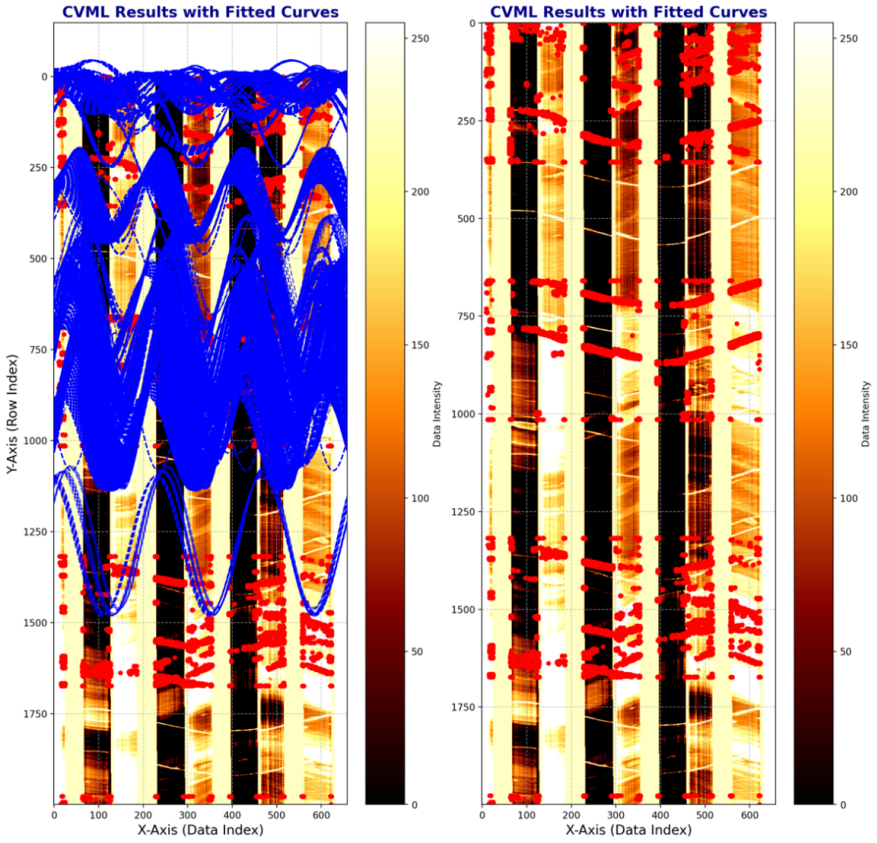

Figure 2.92: Scalability results of the CVML method. The poor results are determined by the blank strips in the image.

2.4 Conclusion

We presented an integrated framework for automating borehole image log analysis, combining computer vision -based (CVML) and deep learning (DL) meth-

ods to assist geological interpretation in real-time or near-real-time drilling operations. The CVML method applies Sobel edge detection and Dynamic Time Warping (DTW) to extract features and approximate geological surfaces, while the DL approach uses segmentation (PickNet) and feature-fitting networks (FitNet) to detect and fit sinusoids that represent geological discontinuities. The results of the two methods are aligned using a linear sum assignment approach, achieving consistent feature detection and enabling integrated analysis.

A key contribution of this work is the demonstration of the validity of synthetic data for training DL models in the context of borehole image analysis, a topic that understandably raises skepticism among interpreters of those type of data. While synthetic data may lack full representativeness of real-world scenarios, the results presented here show that models trained on these datasets deliver at least acceptable and often good performance on field datasets. Both CVML and DL methods effectively identify geological features such as bedding planes and fractures and can provide with information at the interval level, providing meaningful geological insights through derived quantities like apparent dip magnitude, density contrasts, and uncertainty estimates of the predicted features. These results validate the synthetic data and supervised-learning framework as a practical alternative when real-world annotated data is scarce for this problem.

Both methods, however, have limitations. The CVML approach is computationally intensive, making it unsuitable for real-time applications, and its reliance on a threshold parameter introduces subjectivity that can impact results. The DL method, while faster and adaptive, depends heavily on the quality and representativeness of the synthetic data. Its predictions, though generally reliable, can include noise or artifacts, particularly when handling ambiguous or poorly represented geological features. Additionally, the DL system is limited by the lack of an established metric to measure how well synthetic data represents real-world borehole images, especially for high-resolution datasets. The This work demonstrates the potential of combining deterministic and machine learning methods to automate borehole image analysis and provide actionable insights. It also highlights the need for future research to improve the representativeness of synthetic data by developing tools to quantify and enhance its alignment with real-world scenarios. Such efforts would further close the gap between synthetic and field data performance, ensuring more robust and reliable automated solutions for subsurface interpretation.

Chapter 3

Exploring sustainable geothermal power: deep closed-loop heat exchangers and Long-Short Term Memory applications

3.1 Introduction

Geothermal energy offers at least three key advantages: (i) it is a constant and widely available energy source; (ii) it supports thermal and electrical energy needs across residential and industrial sectors; and (iii) it plays a vital role in advancing the renewable energy transition. Geothermal fluids are especially versatile, with applications depending on their temperature range. At 40–60 °C, they are used for greenhouse heating and aquaculture; at 60–70 °C, for building air conditioning and domestic water heating; at 70–80 °C and above, for industrial processes and absorption chillers in refrigeration cycles; and beyond 100 °C, for electricity generation in binary plants (Gola et al., 2022; DiPippo, 2007). However, the development of traditional geothermal projects is often hindered during exploration by uncertainties related to the physical conditions of the reservoir (e.g., fluid pressure) and its petrophysical properties (such as porosity and permeability). Pipe corrosion, scaling, subsidence, and induced seismicity can sometimes pose challenges during geothermal resource exploitation, potentially affecting its commercial success (Chen et al., 2020; Palomo et al., 2022; Witter et al., 2019). Additionally, the high development costs—particularly those related to drilling operations—remain a significant barrier to geothermal projects. A promising solution to address these costs is the reuse of sterile or depleted deep wells originally drilled by the petroleum industry. Retrofitting abandoned wells provides multiple benefits: it reduces drilling expenses, takes advantage of existing knowledge about subsurface geological and physical conditions, lowers environmental impact, and enables the deployment of both conventional and non-conventional geothermal systems (Alimonti and Soldo, 2016; Nadkarni et al., 2022; Alimonti et al., 2016). A promising alternative to overcome the practical limitations of conventional hydrothermal systems is the deep borehole heat exchanger (DBHE). This approach

uses coaxial pipes installed in single wells, which can often be repurposed from abandoned ones. In DBHE systems, a fluid is injected into the central pipe, absorbs heat from the surrounding rocks, and rises through the annulus (Doran et al., 2021; Alimonti et al., 2016; Kujawa et al., 2006; Lund, 2003). However, their efficiency remains relatively low, typically producing only a few hundred kWth, which limits their competitiveness compared to conventional hydrothermal systems. In such cases, the primary benefits include reduced drilling costs, detailed knowledge of well lithology, minimized environmental impact, decreased corrosion and scaling effects, and the ability to harness both conventional and unconventional geothermal systems. The first operational DBHE system was developed by Mining Support Ltd in Kiskunhalas, Hungary, by repurposing an abandoned hydrocarbon well for geothermal use. Recently, a new approach has been introduced to enhance the efficiency of closed-loop geothermal systems by using sub-horizontal connections between two or more wells, leveraging multilateral well technology (Hu et al., 2021; Liao et al., 2021; Yuan et al., 2021). These deep closed-loop heat exchanger (DCHE) systems outperform DBHE systems by increasing the circuit length, heat exchange surface, and the residence time of the fluid within the pipes. The effectiveness of these configurations was demonstrated by *Eavor*^{TM} at the Derek Riddell Demonstration Facility in Alberta, Canada, in 2019. In 2023, this technology was implemented in Europe at the Geretsried drilling site in Germany. The *Eavor-Loop*^{TM} in Geretsried is set to provide district heating and electricity for the entire region within the next four years, showcasing its potential to revolutionize energy supply and become a transformative solution in the sector. Despite the high potential, this solution is far from being cost effective, which undermines its large scale employment (at least temporarily).

The energy performance of DCHE systems depends on several parameters, each influencing thermal performance differently and to varying degrees, depending on the period over which the thermal response of the DCHE is observed. Gola et al. (2022) conducted a sequence of heuristic simulations, altering the values of environmental, design, and operational variables within discrete ranges to analyze their effect on production temperature at the end of the reference period. Sensitivity analysis, using regression techniques (Saltelli et al., 2008), identified key variables such as: (i) the undisturbed geothermal gradient; (ii) thermal diffusivity; (iii) the vertical and horizontal dimensions of the closed-loop system; (iv) the internal well diameter; (v) the thermal conductivity and roughness of the well completion materials; and (vi) operational parameters like flow rate, heat extracted by the primary fluid, and injection temperature. Results indicate that long-term production temperature is primarily influenced by the geothermal gradient among environmental variables, maximum depth of the closed-loop among design variables, and flow rate and heat exchange among operational factors.

This study aims to analyze the long-term performance of a novel DCHE configurations by varying the number of reused exploratory wells (N), the length of their vertical (H) and horizontal (L) sections for their connection, flow rate per well (q), injection temperature (Tin), heat exchange (T), and geothermal gradient (G geo). The primary objective is to identify conditions under which a minimum

long-term production temperature of 100°C—necessary for binary power plants dedicated to electricity generation (DiPippo, 2007)—can be achieved. Performance predictions for geothermal systems under different conditions are typically conducted using numerical simulations. While reliable, these simulations are time-consuming. Several deep learning techniques have already been validated for predicting geothermal system performance (Gangwani et al., 2020; Shi et al., 2021; Wang et al., 2022a). This work explores the applicability of a Long Short-Term Memory (LSTM) neural network to predict long-term production temperature based on the aforementioned variables, to get insights into the long-term production temperature forecasting in the proposed DCHE configurations, with the geothermal gradient as the only geological information int he process, since in our models we assumed no lateral variation for it and considered, therefore, relatively simple geological contexts. In Italy, for example, geothermal reservoirs are abundant, but the high-temperature ones (>150°C) that are of interest for power production, tend to be in tectonically active and fault controlled systems (Manzella et al., 2019; Santilano et al., 2015). An extraordinary case in Italy is Larderello (Tuscany), where the first geothermal power plant in the world came into being. Tuscany, in general, is an italian region with a remarkable green energy contribution of about 30%. The contribution of green energy over the total national demand however, is still limited to 2% (Manzella et al., 2019). Since the 20th century, more than 2,000 exploratory wells have been drilled in Italy, primarily for hydrocarbon exploration and, to a lesser extent, geothermal monitoring. Data on these wells are stored in public databases (Trumpy and Manzella, 2017). The ultimate goal of this study is to evaluate the potential application of these DCHE to two locations in Central Italy- one in Larderello (Tuscany) and the other in Cesano (Lazio), with high densities of exploratory wells. By doing so, we aim to propose an explorable solution for sustainable use of deep resources and to broaden geothermal electricity production to a wider range of geological conditions. We hope to contribute to the improvement of the role of geothermal energy within the European energy mix by 2050, by means of sustainable state-of-the-art approach.

3.2 Methods

In this section, we will describe the methodologies employed in this DCHE innovative configuration study, based on the reused of abandoned wells. We introduce the geological setting of the areas of our DCHE simulations in real-cases scenarios. Then, we describe the numerical models that were used for our multiphysics DCHE simulations and their results. Finally, we introduce the LSTM application to this problem. As anticipated in the introduction, we intend to test this LSTM not with the purpose of demonstrating the predictability of the production temperature, but to somehow assess its forecasting potential, being informed essentially with operational variables, with except to the geothermal gradient (Fig. 3.1).

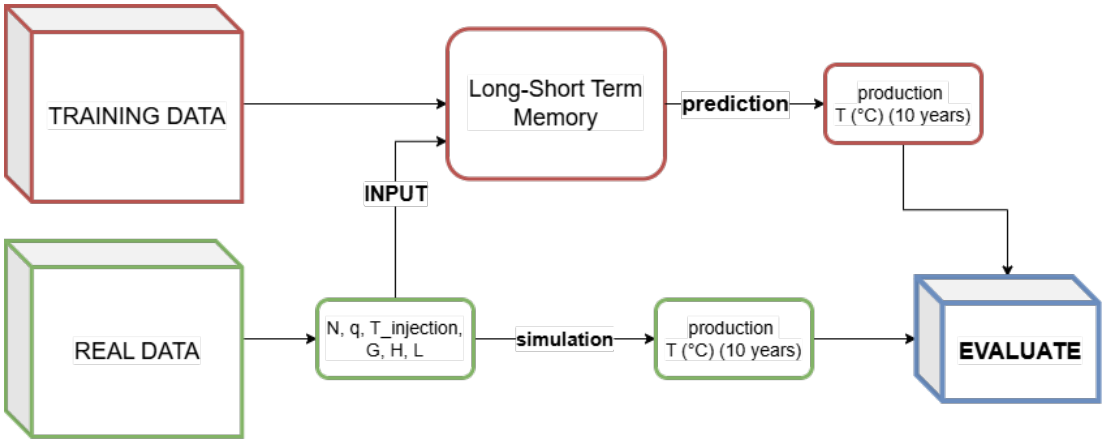

Figure 3.1: Scheme of the LSTM training and real-site prediction workflow.

3.2.1 Geological and geothermal setting of case study areas

Two geothermal regions were chosen to evaluate the thermal performance of a DCHE system using pre-existing deep boreholes for geothermal power production. These regions, Larderello and Cesano-Sabatini, are located within the internal sector of the Northern Apennines and belong to the Tuscan and Roman Magmatic Provinces, respectively (Peccerillo and Donati, 2003) (Fig. 3.2).

The geological structure of this area is characterized by compressional fold-and-thrust belts with nappes predominantly oriented northeastward or eastward. Both sites share common geological traits, such as their location within a convergent margin influenced by post-orogenic extension and magmatic activity during the Pliocene-Pleistocene (intrusive in Larderello) and Pleistocene-Recent (effusive in Cesano-Sabatini). These sites are notable for hosting high-temperature geothermal reservoirs ($T > 150\,°C$) primarily governed by secondary permeability.

Extensive geothermal exploration between the 1970s and 1990s in Southern Tuscany and Northern Latium revealed significant geological and geothermal data, distinguishing the features of the two regions. In the Larderello area, drilling reached the regional metamorphic basement, known as the Gneiss Complex (Precambrian to Lower Paleozoic). The overlying unit is composed of metamorphic Paleozoic formations (micaschists, phyllites) and of quartzitic, evaporitic and carbonate formations (Triassic), which were tectonically placed above the autochthonous metamorphic basement during the Alpine orogeny (Tortonian), constituting the so-called Tectonic Wedges Complex (Bertini et al., 2006). Above these units lies the Tuscan Nappe, composed mainly of evaporitic-to-carbonate deposits (Upper Triassic to Upper Cretaceous). The sedimentary Tuscan Nappe overthrusts the metamorphic ones and it is made up of a mainly evaporitic to carbonate sequence (Upper Trias to Upper Cretaceous). The Tuscan Nappe experienced extensional tectonic processes, which locally determined the direct superposition of the overlying Ligurian Units both on its oldest Upper Triassic formations and on the metamorphic ones. The allochthonous Ligurian Units (Lower Cretaceous to Eocene) are made up of shales, marly limestones and em-

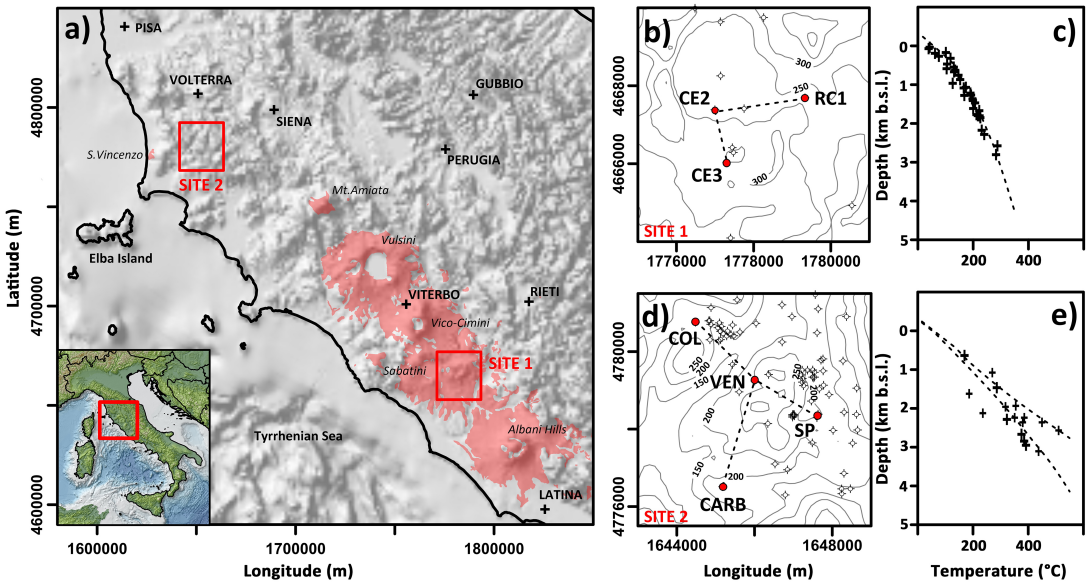

Figure 3.2: The figure provides an overview of the Cesano (Site 1) and Larderello (Site 2) geothermal fields in Italy. Panel (a) shows the geographical location of the two sites, with Cesano and Larderello highlighted in red boxes. The inset map provides a regional context within Italy. Panel (b) displays a contour map of the Cesano site with key locations (CE2, CE3, and RC1), and Panel (c) plots depth (km b.s.l.) against temperature (°C) for Cesano, showing a clear geothermal gradient. Panel (d) illustrates the contour map of the Larderello site with important wells (COL, VEN, SP, and CARB), while Panel (e) presents the depth-temperature relationship for Larderello, also highlighting its geothermal gradient. The maps and plots provide insights into the geothermal characteristics and temperature-depth profiles of the two sites.

bedded ophiolitic bodies, constituting the impervious cap-rock of the deep-seated geothermal reservoir.
In northern Latium the deepest boreholes tap into the dolomitic–anhydrite Triassic formations of the Tuscan and Umbria-Marche Series. The succession continues upward with the Jurassic-Eocene carbonate multilayer consisting of shallow water carbonate platform deposits and of calcareous-marly pelagic basin units. The geothermal reservoir is hosted in the Meso-Cenozoic carbonate sequence and it presents wide-spread self-sealing phenomena related to the recent thermometamorphic and hydrothermal events (Funiciello et al., 1979; Cavarretta and Tecce, 1987).
Data on the region's geothermal potential were obtained through stratigraphic and bottom-hole temperature (BHT) profiles from deep hydrocarbon wells (Gola et al., 2022). These wells typically exhibit conductive temperature profiles, with the Cesano area showing an average geothermal gradient of $90\,^{\circ}\mathrm{C\,km^{-1}}$. The litho-thermal units explored in the Cesano wells CE2, CE3 and RC1 (Pasquale et al., 2011; Pauselli et al., 2019; Dalla Santa et al., 2020) are summarized in Tab.3.1.

Lithology	**Age**	λ **(W/mK)**	c_p **(J/kgK)**	ρ **(kg/m^3)**	**CE2 (m)**	**CE3 (m)**	**RC1 (m)**
Volcanites and clay	Pl-Plst	1.2	1250	2048	230	250	225
Quartzose sandstones and argillites	K-Olig	2.6	1250	2360	-170	-325	-545
Mudstones and marls	u-K-Eoc	2.5	1250	2368	-1115	-1087	-770
Marly/cherty limestones	m-J-l-K	2.9	1100	2700	-1593	-1340	-1055
Massive limestones	l-J	3.5	1100	2700	-1730	-1460	-1475
Dolomitic limestone and anhydrites	u-Tr	4.5	1100	2800	—	—	-2425

Table 3.1: Geological units in Cesano and associated thermal properties. Thermal conductivity (λ), volumetric heat capacity (c_p), density (ρ), and depths (TVD) to the tops of litho-thermal units are displayed. Ages: Q – Quaternary, Pl – Pliocene, Mioc – Miocene, Olig – Oligocene, Eoc – Eocene, K – Cretaceous, J – Jurassic, Tr – Triassic, P – Permian, u – upper, m – middle, l – lower.

With regards to the Larderello area (Larderello-Travale Geothermal Area-LGTA), it has been a key focus for studying geothermal systems, due to its potential for unconventional geothermal resources. Two main geothermal reservoirs exist in this area: a shallow reservoir located in evaporite-carbonate units at depths of about 0.7–1.0 km below ground level (b.g.l.), with temperatures ranging from 150°C to 260°C, and a deep reservoir hosted in metamorphic rocks and Neogene

granitoids at depths of 2.5–4.0 km b.g.l., with temperatures reaching 300°C to 350°C (Gola et al., 2017; Romagnoli et al., 2010). Exploration in the early 1980s in the Lago area, revealed high-temperature (>400°C) and high-pressure (>24 MPa) fluids below the currently exploited hydrothermal systems, encountered at approximately 2900 m depth in the San Pompeo 2 well (one of the injection wells we selected for our simulation), located in the southern portion of the Larderello geothermal field, targeted deep basement layers beneath the exploited reservoir to extract geothermal fluids for electricity generation (Bertini et al., 2006). The San Pompeo 2 well, located in the southern part of the Larderello geothermal field, was aimed to the production of geothermal fluids from the deep basement layers, where an anomalous seismic horizons of bright-spots type was recognized. Below the sedimentary cap rock and the presently exploited geothermal reservoir, the well crossed a thick sequence of metamorphic formations of Paleozoic age (phyllites and micaschists) affected by a widespread hydrothermal metamorphism. High temperature hydrothermal and contact metamorphism mineral assemblage prove the presence of a nearby apophysis of an Alpine granite intrusion. In-hole geophysical logs confirmed the existence of a deep seismic reflecting horizon which probably corresponds to fractured levels at bottom hole in the basement. Here a fluid with $T \geq 400$°C and $P \geq 240$ bars was found(Gola et al., 2017). Two related seismic markers, the H-horizon and K-horizon, have been identified throughout the LTGA (BATINI et al., 1984; Cameli et al., 1993). The H-horizon, associated with the thermo-metamorphic aureole of Neogene granitoids in some areas has yielded superheated steam in several wells (Bertini et al., 2006). The deeper K-horizon, more spatially continuous and marked by bright spot features, has not been fully drilled, with the exception of the San Pompeo 2 well. Thermobaric conditions at the K-horizon (P 30 MPa and T > 400°C) differ significantly from the sub-hydrostatic pressures characteristic of the exploited reservoirs. This raises questions about its origin, which some suggest could be linked to recent magmatic activity (Gola et al., 2017). We used the K-horizon as bottom horizon in our simulation for the Larderello area (see section 3.2.3).

Lithology	**Porosity**	λ **(W/mK)**	ρ **(kg/m^3)**	c_p **(J/kgK)**	**Radiogenic Heat [μW/m^3]**
Cover (Neogene & Ligurian)	0.1	2.5	2400	900	1.0
Tuscan	0.06	3.0	2700	850	0.5
TWC	0.05	3.5	2800	900	0.8
Metamorphic basement	0.01	3.0	2770	850	2.0
Intrusive bodies (melt)	-	2.5	2500	1300	-
Intrusive bodies (solid)	-	3.0	2650	850	2.0

Table 3.2: Geological units in Larderello and associated thermal properties.

3.2.2 Numerical models

In developing the numerical DCHE model, thermal conduction and advection were assumed as the primary heat transfer mechanisms within the borehole and geological formations. The large length-to-diameter ratio of the DCHE ensures fully developed fluid flow, such that all velocity components normal to the borehole axis are zero. The model solves for the one-dimensional velocity, pressure, and temperature of a Newtonian fluid circulating inside the closed loop, exchanging heat with the surrounding rocks, whose temperature increases with depth. Heat transfer at the wall-fluid interface is modeled as pure conduction using Newton's law of cooling. Freshwater is assumed as the working fluid, and its thermodynamic properties are defined using the IAPWS-IF97 formulation.
The governing momentum and mass conservation equations for the fluid are:

$$\rho_f \frac{\partial u}{\partial t} + \rho_f u \frac{\partial u}{\partial z} = -\frac{\partial p}{\partial z} - f_D \frac{\rho_f u^2}{2D} + F, \tag{3.1}$$

$$\frac{\partial \rho_f}{\partial t} + \frac{\partial (\rho_f u)}{\partial z} = 0, \tag{3.2}$$

where ρ_f is the fluid density (kg m^{-3}), u is the tangential fluid velocity (m s^{-1}), p is the pressure (Pa), f_D is the Darcy-Weisbach friction factor (dimensionless), F is the gravitational force per unit volume (N m^{-3}), D is the borehole diameter (m), and z is the vertical coordinate (m). The pressure drop along the wellbore is calculated using the Haaland equation for the Darcy-Weisbach friction factor:

$$f_D = \left[-1.8 \log_{10} \left(\frac{\epsilon}{3.7D} + \frac{6.9}{\mathrm{Re}} \right) \right]^{-2}, \tag{3.3}$$

where $\mathrm{Re} = \frac{\rho_f u D}{\mu_f}$ is the Reynolds number, μ_f is the fluid dynamic viscosity (Pa s), and ϵ is the surface roughness of the borehole wall (m). Additional pressure losses due to changes in borehole diameter are accounted for as:

$$\Delta p = K \frac{\rho_f u^2}{2}, \tag{3.4}$$

where K is the loss coefficient determined by the ratio of cross-sectional areas.
The energy conservation equations for the fluid and surrounding rocks are:

$$\rho_f c_p \frac{\partial T_f}{\partial t} + \rho_f c_p u \frac{\partial T_f}{\partial z} = \lambda_f \frac{\partial^2 T_f}{\partial z^2} + Q_w, \tag{3.5}$$

$$\rho_r c_p \frac{\partial T_r}{\partial t} = \lambda_r \nabla^2 T_r, \tag{3.6}$$

where T_f and T_r are the fluid and rock temperatures (K), c_p is the specific heat (J kg^{-1} K^{-1}), λ is the thermal conductivity (W m^{-1} K^{-1}), and Q_w represents the radial heat exchange between the fluid and rocks. The radial heat transfer is described by:

$$Q_w = h(T_f - T_r), \tag{3.7}$$

where h is the heat transfer coefficient, which depends on the fluid properties and flow conditions. The Nusselt number (Nu) for turbulent flow in circular pipes is given by:

$$\mathrm{Nu} = 0.023\,\mathrm{Re}^{0.8}\,\mathrm{Pr}^{0.3}, \tag{3.8}$$

where $\mathrm{Pr} = \frac{c_p \mu}{\lambda}$ is the Prandtl number.
For cased boreholes with multiple cylindrical shells, the thermal conductivity of the borehole wall is computed as the weighted sum of resistances across the shells, assuming quasi-static heat transfer.
The governing equations (Eqs. 1–6) were implemented and solved using the finite element method in COMSOL Multiphysics®, with the MUMPS direct solver for sparse linear systems. At the injection wellhead, the boundary conditions include a fixed injection temperature (T_{in}) and volumetric flow rate (q_v), while at the production wellhead, heat flow (Q_{out}) and pressure (p_o) constraints are applied. A minimum outlet pressure of $p_o > 1$ bar prevents fluid phase changes at high outflow temperatures.
The thermal regime in the surrounding rocks is assumed conductive, with surface temperature (T_o) set to the mean annual air temperature. Thermally insulating lateral boundaries are defined beyond the DCHE influence radius, and the initial conditions include an isothermal fluid profile and steady-state conductive temperature in the rocks.

Closed-loop configuration

The classical configuration of a DCHE consists of two boreholes—an injection well and a production well—connected at depth by a horizontal section and an insulated surface pipeline (Gola et al., 2022). These components, along with the casing and tubing materials, are critical to maintaining the closed loop's efficiency as a primary fluid circulates to transfer heat to the geothermal plant.
In our simulations, we evaluated systems with varying parameters, including the number of reused exploration wells (N), vertical depth (H), horizontal section length (L), flow rate per well (Q), injection temperature (T_{in}), heat exchange (ΔT), and geothermal gradient (G_{geo}). The wells are drilled vertically to a depth H, where they deviate following a tangent build curve with a medium-radius build rate, achieving a 45° inclination. Drilling then proceeds tangentially for a specified length before transitioning back to a vertical orientation using the same build rate. This deviation results in a horizontal offset (R) between the end of the vertical section and the start of the horizontal section, which has a length L. The horizontal wells are connected to a central production well, and the production temperature (T_{prod}) is determined over a simulation period of up to ten years.
This study evaluates the long-term performance of this innovative DCHE configuration to identify conditions capable of achieving the recommended minimum production temperature of 100°C for binary geothermal power plants. Table 3.3 summarizes the parameter combinations used in the simulations.
The flow rate in each horizontal well was adjusted to achieve a total flow rate of 70–100 l/s at the production well, representing a typical range for geothermal power production. The total flow rate is expressed as:

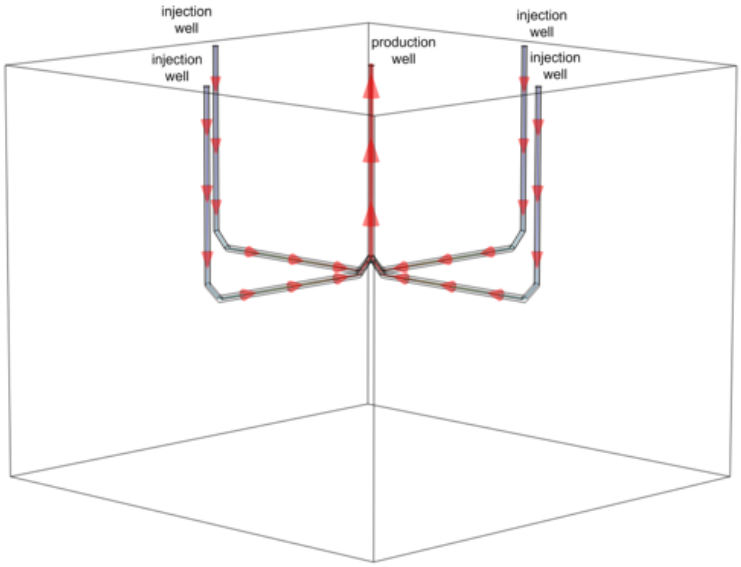

Figure 3.3: Example of the proposed DCHE system configuration. The red arrows in correspondence of the borehole have a size proportional to the flow rate. At the production well the flow rate is given by the sum of the injection wells rates.

$$q_{\text{tot}} = \sum_{n=1}^{N} q_n, \tag{3.9}$$

where q_n is the flow rate of each well.

For each configuration of the variables (see Table 3.3), a multiphysics simulation was performed using COMSOL®. These simulations provided the production temperature profiles used as the target for LSTM training (Fig. 3.4-3.7). All simulations refer to variable numbers of boreholes (N), with variable flow rates (q) and injection temperatures (T_{in}), although they are assumed to be constant over the 10-year period for each simulation. The physical properties of the medium, such as volumetric heat capacity (c_p), density (ρ_r), and thermal conductivity (λ_r), were kept unchanged for all simulations and remained constant within the geological medium where the DCHE was simulated. However, for the case study simulations, the thermal properties of the litho-thermal units, as described in Tables 3.1 and 3.2, were considered, along with the real coordinates of the existing wells. We believe that performing the simulation in such conditions, is more useful to get insights about the potential of these systems in simpler geological conditions, compared to those observed in correspondence of high-temperature geothermal systems. For this homogeneous geological medium, we set $\rho_r c_p = 2,75 \times 10^6 [JK^1m^3]$ and $\lambda_r = 2,5 [WK^1m^1]$.

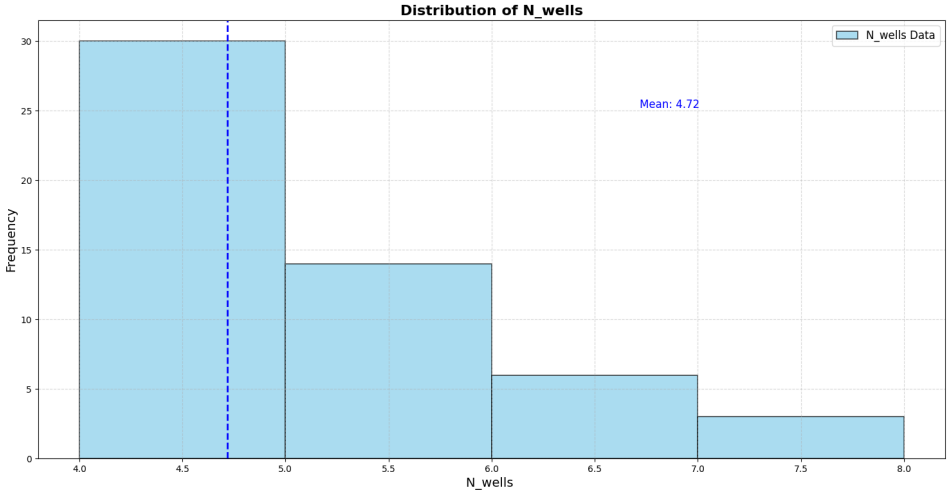

Figure 3.4: Histogram representing the distribution of the number of wells in the dataset.

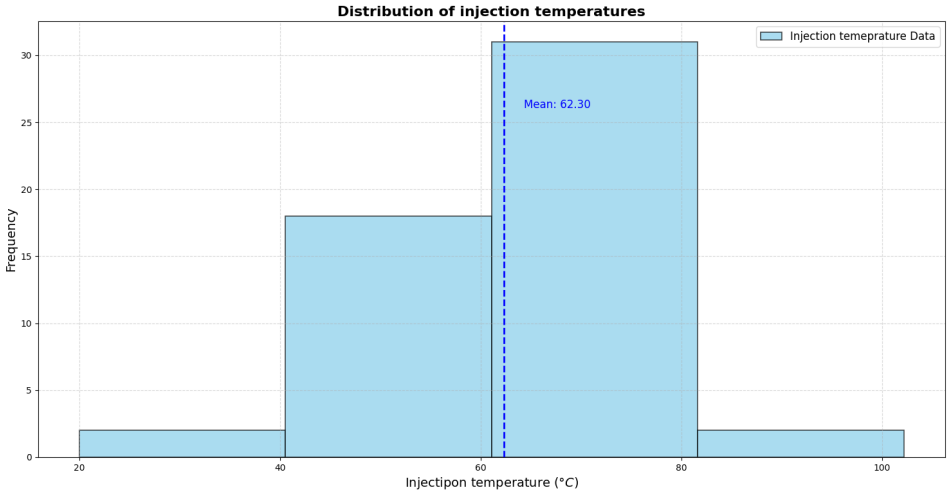

Figure 3.5: Histogram showing the distribution of injection temperatures in the dataset. The mean injection temperature is 62.30°C, highlighted by the blue dashed line.

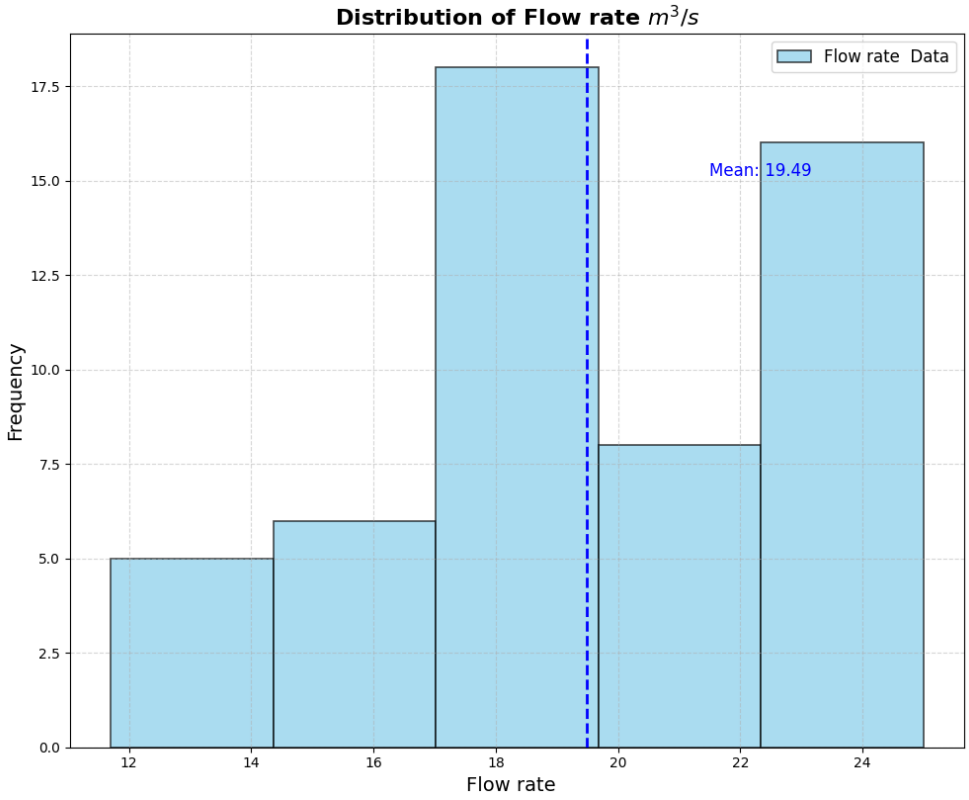

Figure 3.6: Histogram showing the distribution of flow rates in the dataset. The mean flow rate is indicated by the blue dashed line at 19.49 l/s.

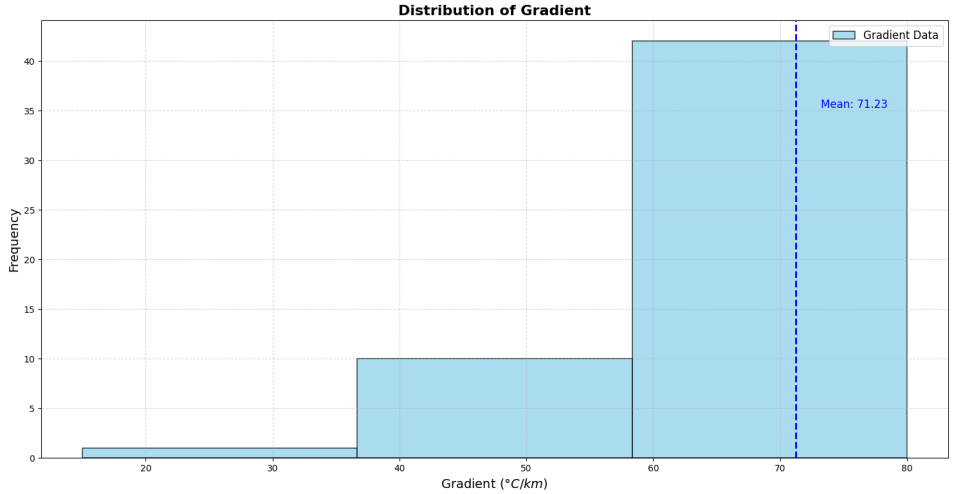

Figure 3.7: Histogram illustrating the distribution of geothermal gradients in the dataset. The mean gradient is 71.23°C/km, indicated by the blue dashed line.

The full dataset consists of 53 training examples, which serves as the target for the LSTM training task. Figure 3.8 and 3.9 provides a comprehensive view of all temperature profiles in the dataset, including the mean profile and a confidence interval of $\pm 1\sigma$. The graph illustrates the production temperature profiles over 3600 days. Initially, the mean production temperature starts above 150°C and

decreases rapidly within the first few hundred days. This decline is followed by a stabilization phase, where the temperature gradually decreases, reaching a final value of 94.26°C at the end of the simulation period.

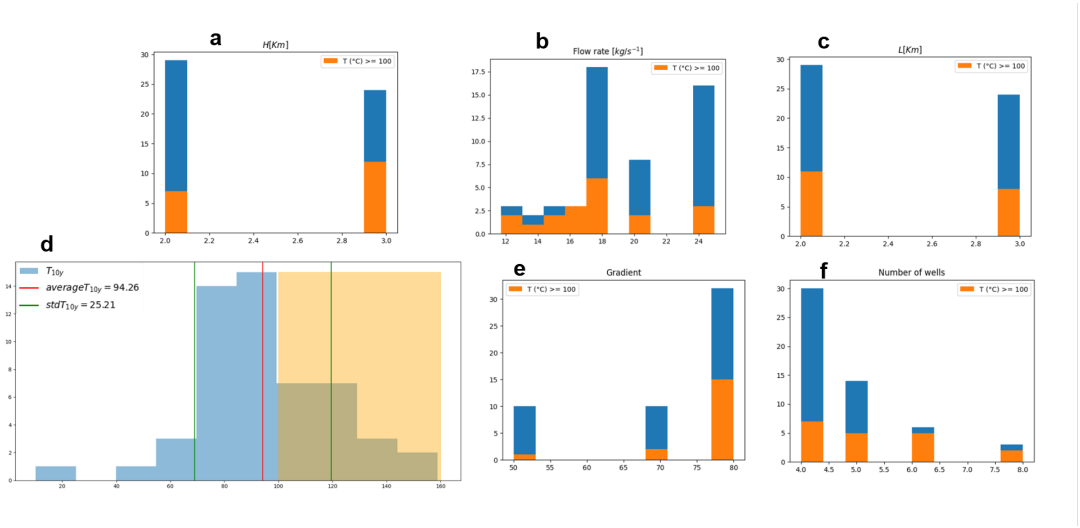

Figure 3.8: Comprehensive view of the production temperature results in the synthetic set. The production temperature (T_{prod}) after 10 years circulation was found greater than 100 °C in the 32% of the cases for different configurations of the considered operational variables. Most of the successful simulations are evidently associated with the highest geothermal gradient, intermediate flow rate values and 3 km vertical depth.

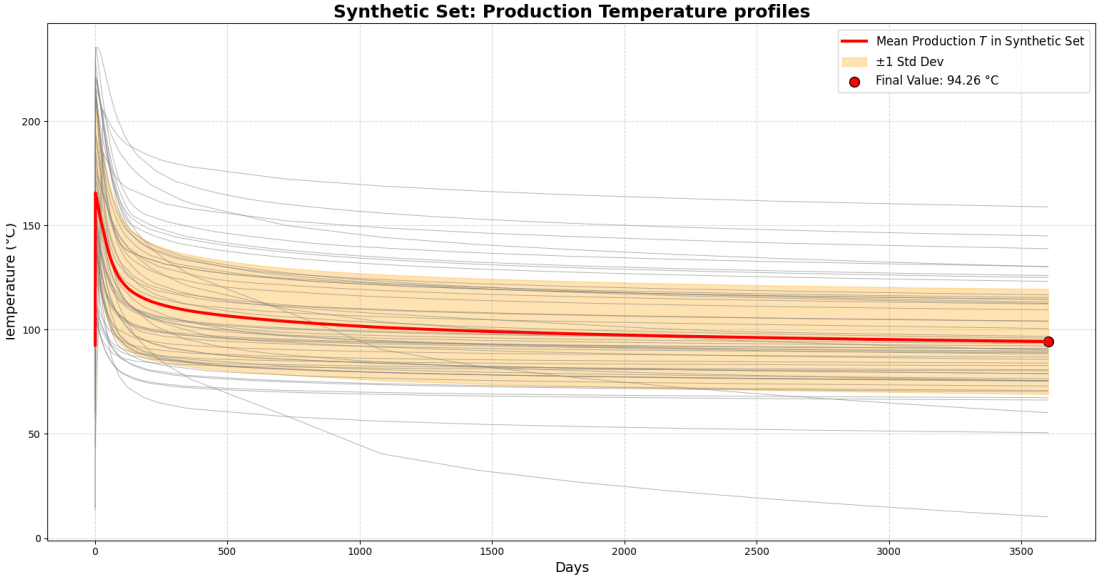

Figure 3.9: Production temperature profiles of the synthetic dataset over time. The solid orange line represents the mean production temperature, with the shaded region indicating ±1 standard deviation. The final temperature is indicated as 94.26°C.

To evaluate the thermal performance of DCHE systems in the Cesano site case study, the paths of three wells- 2 injection wells and one production well- were reconstructed using available drilling survey data. The same was done for Larderello site, but considering three injection wells and one production well as in Fig. 3.2 .The minimum curvature method (Pauselli et al., 2019) was applied to transform measured depth (MD) data into true vertical depth (TVD), using wellbore inclination and azimuth measurements at survey points. Vertical and horizontal displacements were calculated for each depth interval to reconstruct the well trajectories.

N	**H [km]**	**L [km]**	$T_{\textbf{injection}}$ **(°C)**	**G (°C/km)**	**q [l/s]**
3	2, 3	2, 3	50, 70	50, 70, 80	25 , 17.5
4	2, 3	2, 3	50, 70	50, 70, 80	25, 17.5
5	2, 3	2, 3	50, 70	50, 70, 80	20, 14
6	2, 3	2, 3	50, 70	50, 70, 80	11.7, 16.7

Table 3.3: Table showing the configurations of N, H, L, $T_{\text{injection}}$, G, and q.

3.2.3 Case studies simulations

To perform the simulations of the proposed DCHE configurations in the two sites described in section 3.2.1, we select an arbitrary numbers of wells from the available open dataset in Trumpy and Manzella (2017). For Site 1 (Cesano, Lazio) we selected three wells (CE2, CE3, RC1) setting two of them as injectors and one production well (Fig. 3.10). We simulated a 10 years circulation with a fixed injection temperature of 50 °C, a total flow rate at the production well of 70 l/s and a geothermal gradient of 90 °C/km (see section 3.2.1). The horizontal connection of the three wells is at approximately 3 km depth and their distance varies between 2 and 3 km. The Larderello area, on the other side, is related to a more complex geology (see section 3.2.1). We select three injection wells (COL, CARB, SP in Fig. 3.2.1) and one production well (VEN), connected at 2 km depth (Fig. 3.11) and have an horizontal distance of about 3 km. We simulated for a 100 °C/km gradient and a total flow rate of 72 l/s.

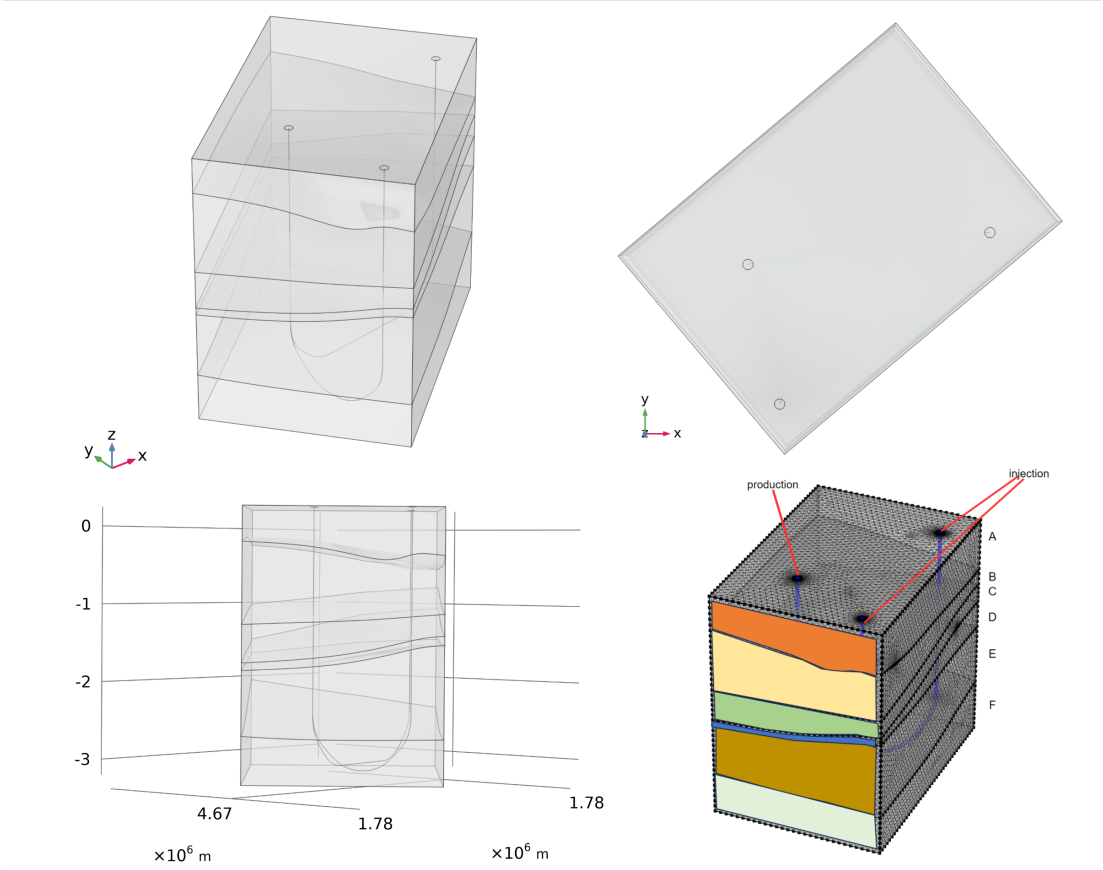

Figure 3.10: The DCHE geological and thermodynamical model of the simulation at Site 1 (Cesano, Lazio). The geological units refer to Tab.3.1

.

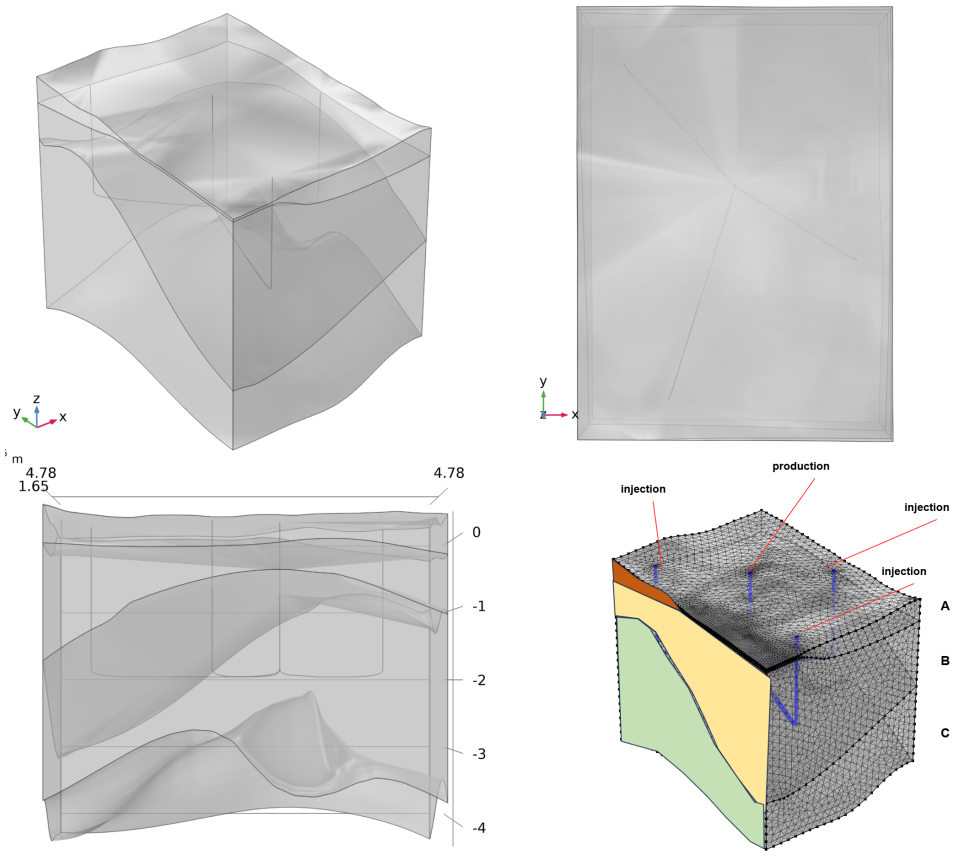

Figure 3.11: The DCHE geological and thermodynamical model of the simulation at Site 1 (Cesano, Lazio). The geological units refer to Tab.3.2

After the simulations, we calculated the thermal (P_{th}) and electric (P_e) power as follows (Zarrouk and Moon, 2014; Gola et al., 2022)

$$P_{\text{th}} = q_w(\rho c_p)_f \cdot (T_{\text{prod}} - T_{\text{in}}) \tag{3.10}$$

$$P_e = P_{\text{th}} \cdot [6.9681 \times \ln(T_{\text{INLET}}) - 29.713] \tag{3.11}$$

3.2.4 Long-Short Term Memory

The Long Short-Term Memory (LSTM) model, introduced by Hochreiter and Schmidhuber (1997), is a type of Recurrent Neural Network (RNN) that addresses one of the critical limitations of traditional RNNs: the inability to learn long-term dependencies effectively. This limitation arises primarily due to the vanishing and exploding gradient problem, which occurs during the training of models over long sequences (Roncoroni et al., 2022; Wang et al., 2022a). LSTMs overcome this issue by introducing memory cells and a gating mechanism, enabling them to retain and regulate information across extended time intervals.

Typically, an LSTM is composed of several core components, each with a function for managing the flow of information (Greff et al., 2017). The cell state,(c_t), a forget gate (f_t), an input gate (i_t), and an output gate (o_t).
The memory cells store information across time steps. The input gate regulates the amount of new information entering the cell, while the forget gate determines how much information from the previous cell state is retained or discarded. Finally, the output gate controls how much of the current cell state is used to produce the output at a given time step. Together, these components work in a recurring cycle across time steps, enabling the LSTM to selectively retain important information and discard irrelevant details (Wei et al., 2021; Al-Selwi et al., 2024).
To describe a forward pass in an LSTM model, let us assume a model with N processing blocks and M inputs. The function of the input block is to update the input block component, given by the combination of the input $x(t)$ and the output of the previous iteration $y(t-1)$:

$$z_t = g(W_z x_t + R_z y_{t-1} + b_z), \tag{3.12}$$

where W_z and R_z are the sets of input and recurrent weights associated with input $x(t)$ and output $y(t-1)$, respectively, and b_z is the bias vector. g is a non-linear activation function, usually being the *hyperbolic tangent* $tanh(x)$.
The input gate is then updated, given the input $x(t)$, the output $y(t-1)$, and the cell value at the previous iteration $c(t-1)$:

$$i_t = \sigma(W_i x_t + R_i y_{t-1} + p_i \cdot c_{t-1} + b_i), \tag{3.13}$$

where W_i, R_i, and p_i are the weights associated with the input $x(t)$, output $y(t-1)$, and cell $c(t-1)$, respectively, and b_i is the bias vector. σ is a logistic sigmoid activation function used as sigmoid usually used for the gates to select and regulate the information directed to the cell state c_t. $\cdot$ is the element-wise Hadamard product.
The forget gate determines the amount of information in $c(t-1)$ to be removed:

$$f_t = \sigma(W_f x_t + R_f y_{t-1} + p_f \cdot c_{t-1} + b_f), \tag{3.14}$$

where W_f, R_f, and p_f are the weights associated with the input $x(t)$, output $y(t-1)$, and cell $c(t-1)$, respectively, and b_f is the bias vector.
The cell value is computed at each iteration to determine $c(t)$ as follows:

$$c_t = z_t \cdot i_t + c_{t-1} \cdot f_t. \tag{3.15}$$

The LSTM output can now be derived from the cell state, through the output gate, which has the same regulatory function as the other gates. The output gate is updated using the input $x(t)$, the output $y(t-1)$, and the cell value $c(t-1)$:

$$o_t = \sigma(W_o x_t + R_o y_{t-1} + p_o \cdot c_{t-1} + b_o), \tag{3.16}$$

where W_o, R_o, and p_o are the weights associated with the input $x(t)$, output $y(t-1)$, and cell $c(t-1)$, respectively, and b_o is the bias vector.

Finally, the block output combines the cell state value with the output gate:

$$y_t = g(c_t \cdot o_t). \tag{3.17}$$

The training of an LSTM involves optimizing the network's parameters (weights) using Backpropagation Through Time (BPTT) (Werbos, 1990). This algorithm allows the computation of gradients needed to adjust the weights connecting the different parts of the network.

In this study, we propose the use of an LSTM network to predict the production temperature profile over a 10-year period of our proposed DCHE configuration. The input to the LSTM includes the environmental and operational variables mentioned in the introduction, as well as the parameters that define the DCHE model configuration (see Table 3.3). By leveraging the LSTM's capability to model long-term temporal dependencies, we want to assess how predictable the mid-long term production temperature of our DCHE systems, based on the operational variables and the gradient. Details on the training data and design are given in the following section.

Training details

We trained an LSTM network to predict the production temperature profile over a period of 10 years. The neural network consists of four hidden layers with 21, 14, 7, and 1 neurons, respectively, followed by a sigmoid activation function, as shown in Fig. 3.12. The layers were trained using the Adam optimizer (Kingma and Ba, 2017) and a Mean Absolute Error (MAE) loss function. The training lasted 400 epochs with a batch size of 32. The implementation was carried out using TensorFlow with CUDA-compatible support (Abadi et al., 2016). The dataset consists of normalized time series corresponding to the variables listed in Table 3.3. The distributions of these variables in the dataset are shown in Figs. 3.4 to 3.7. Given the limited number of training examples in the dataset, we compared the performance of two networks with identical architectures: one trained with a standard strategy and the other using a curriculum learning approach. Curriculum learning involves dividing the training dataset into two subsets based on the complexity of the examples. The first subset, containing simpler examples, is used for pre-training. Fine-tuning is then performed on the pre-trained weights using the subset of more complex instances (Bengio et al., 2009; Soviany et al., 2022). In our case, DCHE configurations with $N = 4$ were used for the pre-training dataset, while the remaining configurations were used for fine-tuning. Note that, this LSTM application is not meant to provide with conclusive insights about the applicability of LSTM to predict the long-term temperature of the proposed DCHE configuration, but it's a test application to assess the generalization potential of such specialized architecture for this task.

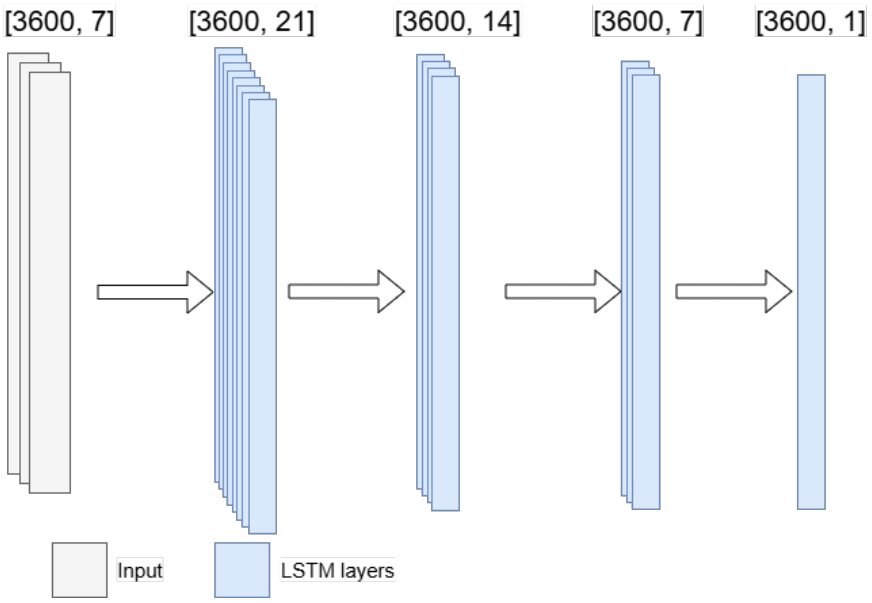

Figure 3.12: Architecture of the neural network with four LSTM layers containing 21, 14, 7, and 1 neuron, respectively.

3.3 Results

The results of the simulations gave production tempertures higher than 100°C in 32% of the cases, which was the first indication of a potential of these types of DCHE for power production. This motivated the simulations in the real case scenarios of Cesano and Larderello using the available data. In Cesano, Site 1 in Fig. 3.2, the simulations results gave a final production temperature of 139,56 °C after 10 years of circulation (Fig. 3.13). The calculated thermal and electrical power and their cumulative curve over time are shown in Fig. 3.14. The average thermal production for Site 1 was 23.7 MW_{th}, corresponding to an average electric power of 1.30 MW_e, resulting in a total average annual power production of 41 TW_e/y in the 10-years period. Similarly in Larderello, Site 2 in Fig. 3.2, the simulations results gave a final production temperature of 138,68 °C after 10 years of circulation (Fig. 3.15). The calculated thermal and electrical power and their cumulative curve over time are shown in Fig. 3.16. The average thermal production for Site 2 was 30.8 MW_{th}, corresponding to an average electric power of 1.66 MW_e, which results in a total average annual power production of 52 TW_e/y in the 10-years period.

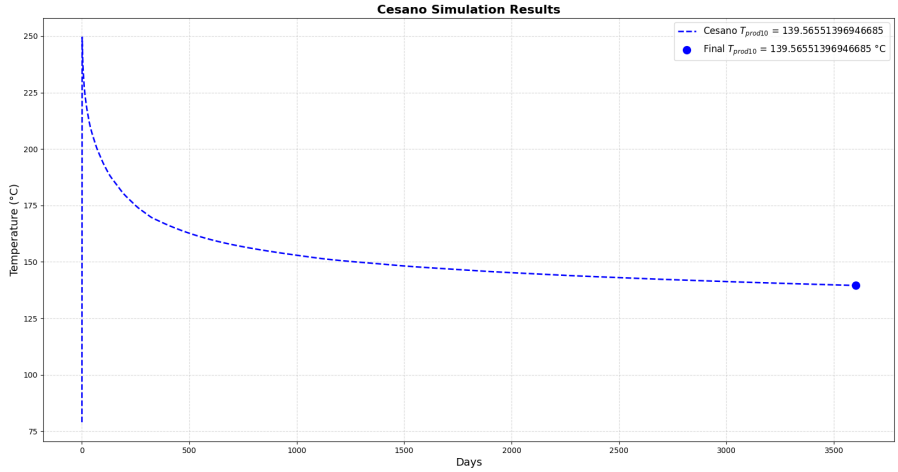

Figure 3.13: Temperature evolution over time in the Cesano numerical simulation, reaching a final temperature of 139.57°C.

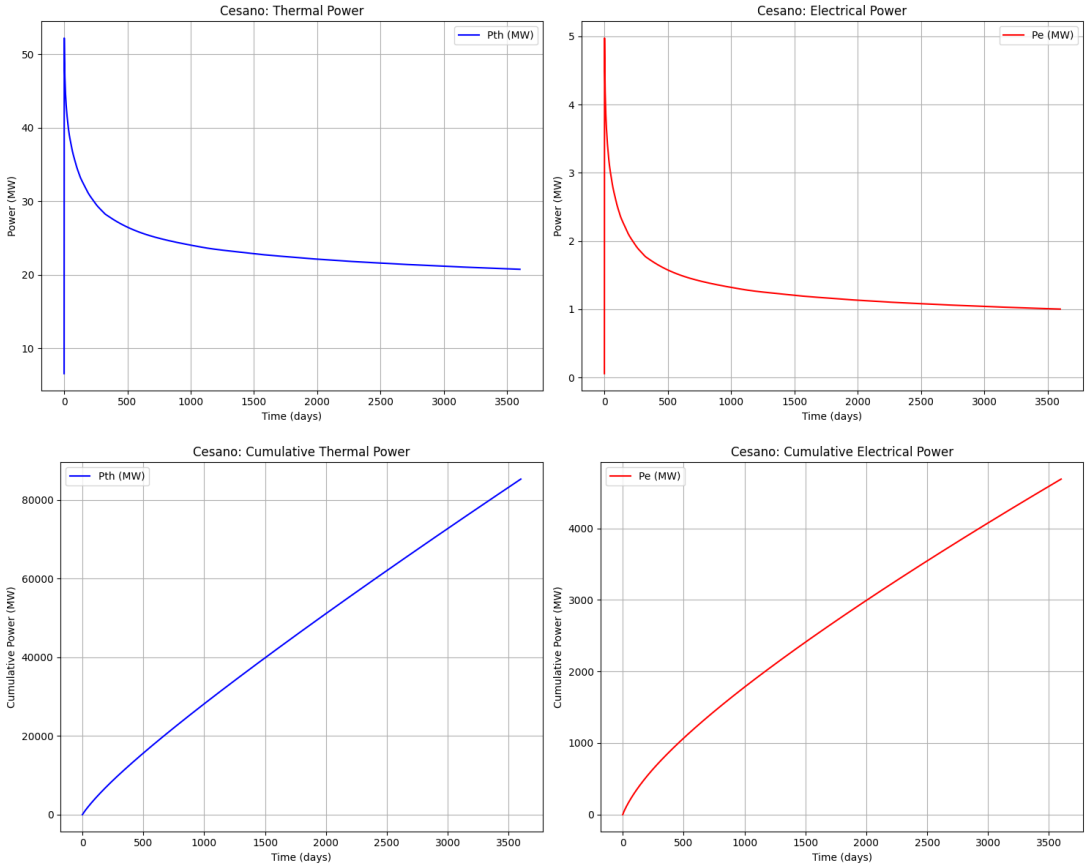

Figure 3.14: CESANO-SABATINI. Up: thermal power and electrical power. Down: cumulative thermal and electrical power.

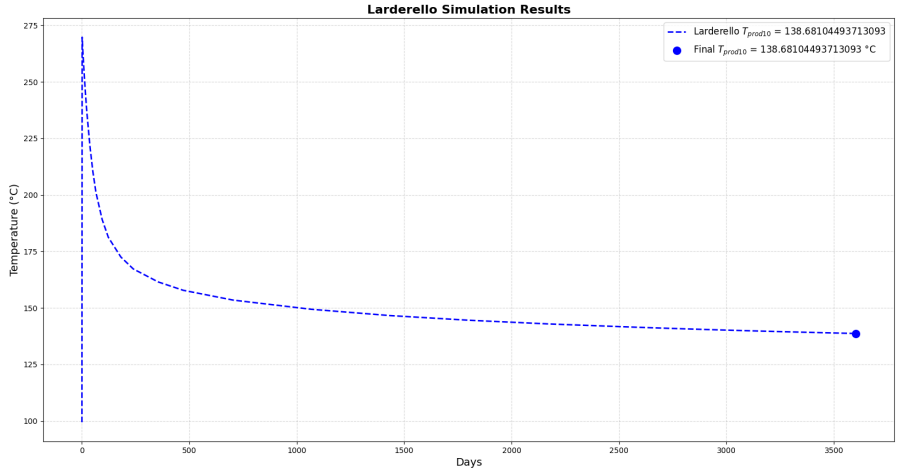

Figure 3.15: Temperature evolution over time in the Larderello numerical simulation, reaching a final temperature of 138.68°C.

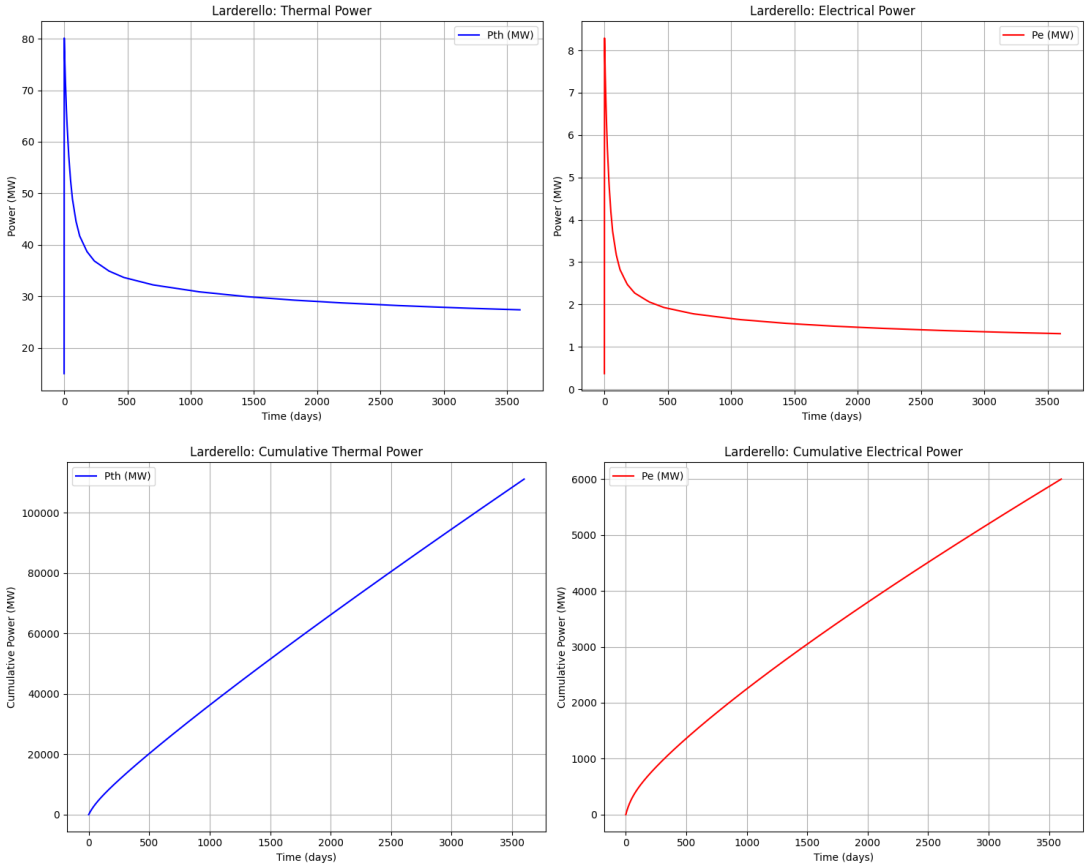

Figure 3.16: LARDERELLO. Up: thermal power and electrical power. Down: cumulative thermal and electrical power.

We tested the applicability of LSTM NN in forecasting the production temperature profile in real-case scenarios, after beinig trained on generic simulation results, with no geological information (sect. 3.2.4). Figure 3.17 compares the mean absolute error (MAE) during training for two neural network training strategies: Curriculum Training (in blue) and Standard Training (in orange). Curriculum Training consistently resulted in lower loss compared to Standard Training. The blue line indicates that Curriculum Training achieved faster convergence and maintained a more stable learning process, ultimately reaching a lower final MAE of 0.04. Conversely, Standard Training started with a much higher loss and exhibited instability during the early epochs, eventually converging to a final MAE of 0.07.
These results highlight that Curriculum Training effectively enhances the model's accuracy, leading to better performance with a lower MAE by the end of training.

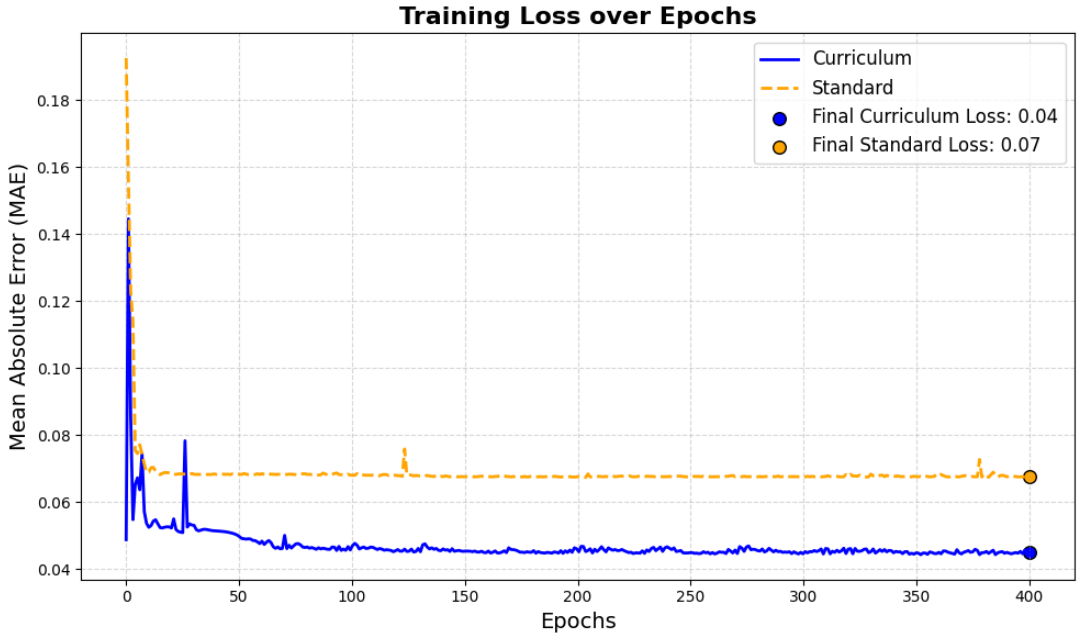

Figure 3.17: Training Loss over Epochs for Curriculum vs Standard Training. The blue line represents the curriculum training approach, while the orange dashed line represents standard training. The final losses are indicated: Curriculum Training reached a final MAE of 0.04, while Standard Training reached 0.07.

The two versions of the LSTM model were evaluated on real case scenarios from Cesano and Larderello. For the Cesano real-case scenario, the final production temperature (after 10 years) predictions errors for the curriculum and standard LSTM models are +2.4 °C and −41.1 °C, respectively. This confirms the superior performance of the curriculum strategy, with the standard LSTM prediction aligning with the long-term production temperature average value in the training dataset. However, the Cesano case is significantly different from the Larderello case, particularly in its geological setting, which is simpler in the thermodynamic simulations. In the Larderello scenario, the model performances are similar, showing errors of +42.1 °C and −47.8 °C for the curriculum and standard LSTM mod-

els, respectively, showing a consistent performance of the standard LSTM and a significant accuracy redcution in the other.

	CL	**SL**
MAE	0.04	0.07
CES	+2.4°C	-41.1°C
LAR	+42.1°C	-47.8°C

Table 3.4: Comparison of Curriculum Learning (CL) and Standard Learning (SL) Performance Metrics. The table shows Mean Absolute Error (MAE) and the temperature deviations for the Cesano (CES) and Larderello (LAR) simulations. Curriculum Learning consistently outperforms Standard Learning, achieving lower MAE and closer temperature predictions, with deviations of +2.4°C for Cesano and +42.1°C for Larderello, compared to larger negative deviations in Standard Learning.

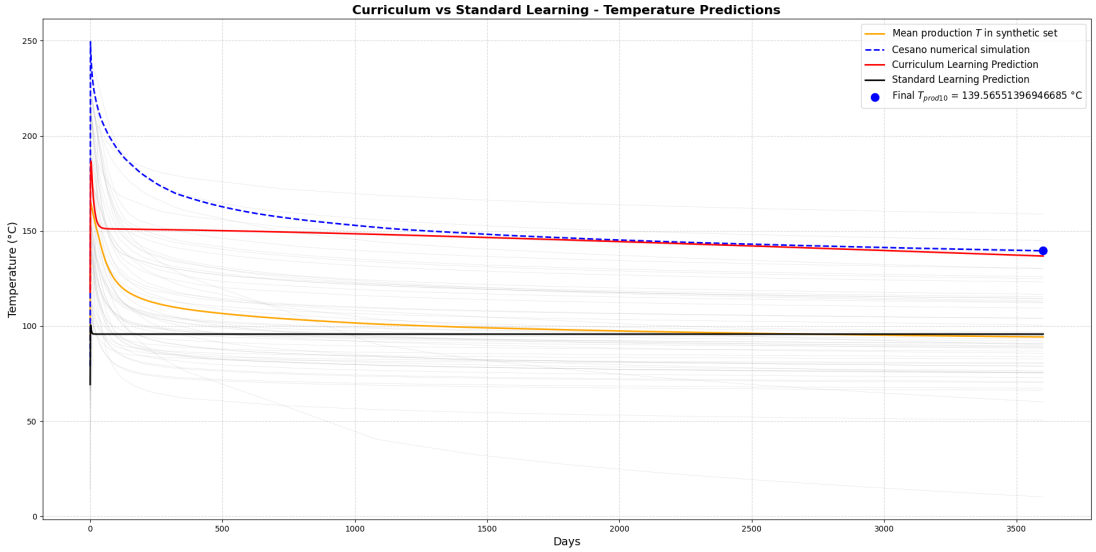

Figure 3.18: Comparison of Curriculum Learning and Standard Learning predictions against the Cesano simulation.

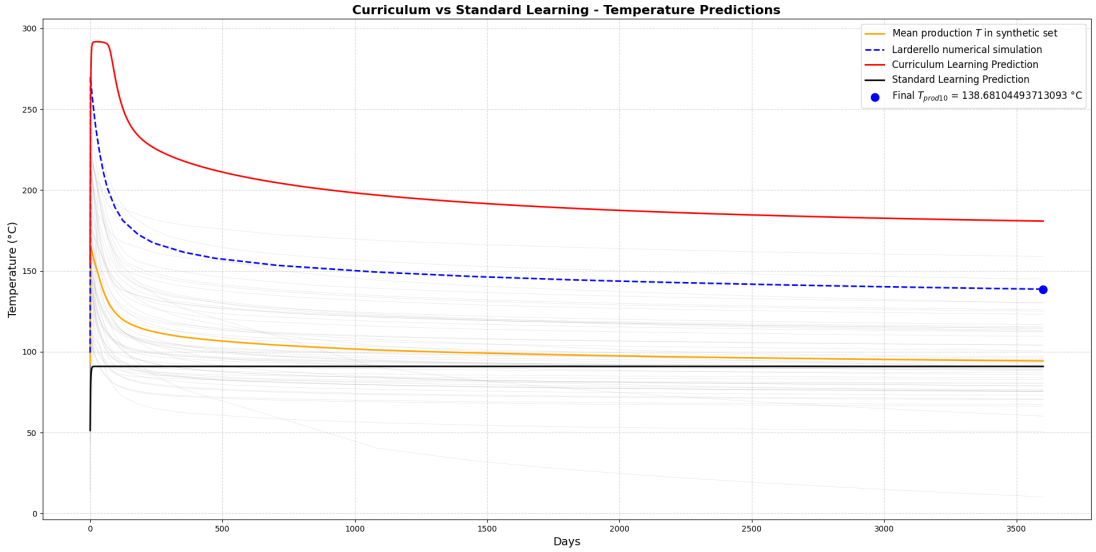

Figure 3.19: Comparison of Temperature Evolution in Larderello Simulation and Curriculum vs Standard Learning Predictions. The top plot shows the temperature decline over time in the Larderello numerical simulation, while the bottom plot contrasts predictions made by Curriculum Learning and Standard Learning models.

3.4 Discussion

Unlike other works that present simulations for 20-30 year periods, we limit our simulation to 10 years, given the rapid exponential drop in the production temperature profile typically observed during the initial phases of radial cooling at the borehole-rock interface. After this phase, a slower, more linear rate of decrease becomes evident.

We preliminarily estimate the potential contribution of the geothermal energy in fulfilling residential electrical energy demands, considering standard 90 m^2 housing units accommodating families of four, with an annual energy demand of 3600 kWh/ (Besagni et al., 2020). At Site 1 (Cesano), the recoverable power of 41 TWe-year can sustain approximately 1300 such housing units, compared to the 25 TWe-year and 793 found in Gola et al. (2022) in the same site. At Site 2 (Larderello), characterized by a very high geothermal gradient, the recoverable power of 52 TWe-year can fulfill the demand of approximately 1649 residential units.

The differing performance levels of the LSTM model observed in the Cesano and Larderello cases may be attributed to the specific nature of the training process and the distinct characteristics of the geothermal systems at each site. The LSTM model was trained to predict the medium- to long-term temperature behavior of geothermal systems configured with DCHE wells, using operational parameters and the geothermal gradient as input features. Notably, the training process did not incorporate detailed geological data, meaning the model's predictions are primarily based on generalized thermal gradients and operational conditions rather

than site-specific geological characteristics. This limitation could contribute to variations in model performance between the two geothermal fields.
In the Cesano case, the geological setting is relatively simpler, characterized by quasi-horizontal stratification and smooth contrasts in thermophysical properties. This simplicity contributes to a resemblance to a thermally homogeneous medium, which aligns closely with the conditions used to generate the synthetic data for training the LSTM model. As a result, the LSTM, particularly when trained with curriculum learning, demonstrated more accurate predictions. This underscores the importance of gradually introducing increasingly complex data during the training process, as it enhances the model's ability to generalize and maintain robustness, especially when the available training data is limited.
In contrast, the Larderello scenario presents significantly greater geological complexity. The geothermal gradient in the Larderello area is influenced by intricate geological structures that diverge from the simplified, homogeneous thermal characteristics used in the synthetic training data. This discrepancy posed a challenge for both the standard and curriculum-trained LSTM models in accurately predicting the temperature evolution. The model's lack of explicit geological information limits its ability to fully capture the effects of complex subsurface conditions, which is particularly evident in Larderello, where geological heterogeneity plays a critical role in thermal behavior.
This analysis may indicate that while the curriculum learning strategy offers advantages in generalizing from synthetic data, the absence of detailed geological data determines the the model's failure in the production temperature forecasting for geologically complex settings.

3.5 Conclusion

In this work, we investigated the power production performance of deep closed-loop heat exchangers (DCHEs) by reutilizing abandoned wells, aiming to propose a sustainable and cost-effective solution to enhance their efficiency. We assessed the power production potential of the proposed DCHE configuration using numerical models to perform medium- to long-term simulations, varying operational parameters while considering the geothermal gradient as the sole environmental variable. The proposed DCHE approach was applied to two real site-specific study areas in Italy, utilizing existing data from open databases to inform the simulations. These sites, located in distinct geothermal regions with varying geological complexity, yielded power plant sizes ranging from 1 to 1.3 MWe. These results confirm that the proposed multilateral well configurations can reliably sustain electrical energy production, demonstrating the feasibility and effectiveness of the DCHE concept. We tested the applicability of an LSTM model for forecasting 10-year production temperatures at the study sites, focusing primarily on operational variables. The model was trained on a dataset derived from simulations in lithothermically homogeneous media, and we compared two training strategies: curriculum learning (CL) and standard learning. Our findings indicate that the LSTM model trained with the more sophisticated CL strategy significantly outperformed the standard approach in the simplest real-case

scenario (Site 1). However, both training strategies failed to accurately predict outcomes in the geologically more complex scenario (Site 2).
In conclusion, DCHEs designed with wells interconnected by multilateral legs offer a promising solution for geothermal power production, particularly by reusing dry or depleted deep geothermal and hydrocarbon wells. This approach has the potential to reduce geothermal project costs while enhancing sustainability.

Chapter 4

Conclusions

Machine learning, especially deep learning, has witnessed an increasing interest in all sorts of geophysical fields, thanks to its ability to handle big data and the computational power available nowadays. In geophysics, this is generally done in the attempt to increase accuracy and avoid manual, tedious, or intensive tasks. However, gaining insights into machine-level performance compared to human-level performance for each specific task is not as common as it should be. This is certainly due to the complexity and inherent uncertainty of geophysical tasks, which often does not enable a straightforward formulation of a human-machine evaluation framework.

As part of this research project, we aimed to deliver promising and innovative methodologies based on AI to optimize the exploration and exploitation of deep geothermal resources, particularly interesting for power production purposes. Geothermal exploration is now facing the need to accelerate its growth rate to enhance the contribution of renewable energy sources to the energy mix. Nevertheless, its growth is undermined by uncertainties in subsurface properties, related to both aleatoric and epistemic uncertainties, with the latter heavily affecting one of the most important geoscientific tasks in geothermal exploration: fracture analysis. To overcome these limitations and the operational problems related to the use of aquifer brines as carrier fluids, several solutions have been proposed that rely on the use of closed-loop heat exchangers. These solutions consist of the multilateral connection of subhorizontal wells to create a closed loop where fluid circulation can be imposed, exploiting heat exchange at the borehole-rock interface. However, DCHE projects are not always cost-effective. Sub-horizontal wells have proven effective in enhancing geothermal project efficiency, as they increase the available surface area for heat exchange compared to vertical or sub-vertical doublets. The navigation of such horizontal drilling is guided by LWD geophysical data, like LWD density borehole images, which assist in formation and fracture identification during well drilling, supporting decision-making during geosteering operations.

In this thesis, we focused on two main aspects, addressing challenges that affect the growth of geothermal energy.

In Chapter 2, we proposed two innovative AI-based solutions—one based on ML and the other on DL—to assist the manual strata and fracture analysis of LWD

borehole images in real-time and support decisions in geosteering wells through semiautomation. For the DL method, we used a supervised learning approach based on synthetic borehole images, representing another key innovative aspect of this methodology. Moreover, we heuristically made design choices for NN training to improve performance to a satisfactory level on both synthetic and field data. Both the ML and DL methodologies showed promising results for automated strata and fracture picking, based on a mathematical representation of the input borehole images, thus using more objective criteria than those employed during manual elaboration. Regarding real-time decision support, the best DL method (Fitnet-C-SA-H) showed more promising results, outperforming the ML counterpart in terms of prediction time. Additionally, by controlling the complexity of the training data, we trained two different alternatives of the same DL method based on curriculum learning and standard learning principles. Our findings indicate that these two training approaches impact final results and human-machine interactions in complementary ways, i.e., with different orientations to detail, to the extent that they could be combined into a more complete and objective decision-support tool. Another significant aspect of DL outperforming ML for this task is scalability. We adapted the method to high-resolution borehole images, such as FMI, extensively used in geothermal exploration and associated with manually picked fractures. The results underline the ability of the method to align with human intentions, despite the higher resolution data compared to the training data. However, recycling this supervised learning approach for FMI or high-resolution data is not straightforward, as the generation of synthetic data would require an extremely high number of features in the images to build a representative dataset. This marks the road for future research.

In Chapter 3, we proposed a deep closed-loop heat exchanger configuration, based on the mutual horizontal connection of dry, sterile, or abandoned wells, for geothermal power production. This was conceived as an intermediate solution for power production, balancing cost-effectiveness and efficiency between highly promising but expensive modern solutions like *Eavor-Loop*^{TM} and the sustainable but less efficient deep borehole heat exchangers (DBHE). We first investigated the potential of such DCHEs in media with homogeneous thermal properties and no geological information using numerical simulations of 10-year circulation periods with varying operational configurations. We then applied the same DCHE concepts to site-specific study areas with medium-to-high and high geothermal gradients, using available geological information for the simulations, and extracted production temperature results to compute the average and total generated electrical power. Our findings demonstrate that this can be a relatively sustainable and cost-effective solution for power production, providing energy for more than 1300 residential units per year in both cases. To preliminarily assess the predictability of production temperature profiles based mainly on operational variables, we also trained an LSTM model on the dataset derived from the simulations in homogeneous media, again comparing the results of different training strategies (curriculum learning and standard learning). The predictions of these networks highlighted that the curriculum learning network, trained with pretraining and fine-tuning on instances with increasing complexity, outperformed

the counterpart in predicting the long-term production temperature for the real site with simpler geology, i.e., more resembling the media exchanging heat with the DCHE in our simulations. However, for the site with the most complex geology, both approaches failed, arguably indicating that this LSTM architecture trained mostly on operational variables (except for the geothermal gradient) to forecast the mid-long term production temperature of the proposed DCHE configuration, have greater forecasting ability in simpler geological real-case scenarios. However, it is important to note that the error levels presented in Table 2.3 correspond to the final production temperature value after 10 years of circulation. This implies that relying on the forecasted 10-year temperature profile to estimate thermal and electric power could lead to erroneous assessments of geothermal potential for both the CL and SL LSTMs. We believe that these two chapters address challenges that can significantly contribute to geothermal energy growth. First, we implemented innovative, reliable, and—in one case—scalable AI-based methodologies to support crucial decision-making during the drilling of horizontal wells, which has demonstrated a positive impact on geothermal projects. The scalability tests of the LWD-specific DL method indicate the potential of the proposed solution for application to high-resolution images. However, a more tailored approach—one that accounts for the higher dimensionality of the problem and incorporates additional features unique to high-resolution BHI—is required to validate this synthetic data-driven supervised learning approach for automated BHI analysis. Second, we proposed a configuration of DCHE based on horizontal drilling to connect abandoned wells, increasing sustainability and reducing operational costs for electrical power production (e.g., submersible pumps in the borehole), and applied the concepts to two test sites, where simulations foresee plant sizes ranging from 1 to 1.3 MW_e for medium-to-high and high geothermal gradients, respectively.

4.1 Future Research Directions

1. **Extending AI-Based Borehole Image Analysis:**
    - Building on the success of the proposed methods for logging-while-drilling (LWD) images, future work will focus on adapting and testing these approaches on high-resolution borehole image (BHI) datasets, such as Formation MicroImager (FMI) data.
    - Particular emphasis will be placed on refining synthetic data generation techniques to ensure they effectively capture the complexities and variability of high-resolution geological features.
2. **Geothermal Potential Mapping:**
    - A key future objective is to create a detailed geothermal energy potential map for Italian regions densely populated with abandoned wells.
    - This initiative will evaluate the feasibility of reusing these wells for deep closed-loop heat exchangers (DCHE) based on the proposed configuration.

- The mapping will consider geological, thermodynamic, and economic factors, providing actionable insights for sustainable energy development and regional energy planning.

Bibliography

Martín Abadi, Paul Barham, Jianmin Chen, Zhifeng Chen, Andy Davis, Jeffrey Dean, Matthieu Devin, Sanjay Ghemawat, Geoffrey Irving, Michael Isard, Manjunath Kudlur, Josh Levenberg, Rajat Monga, Sherry Moore, Derek G. Murray, Benoit Steiner, Paul Tucker, Vijay Vasudevan, Pete Warden, Martin Wicke, Yuan Yu, and Xiaoqiang Zheng. TensorFlow: A system for large-scale machine learning, May 2016. URL http://arxiv.org/abs/1605.08695. arXiv:1605.08695 [cs].

Abien Fred Agarap. Deep Learning using Rectified Linear Units (ReLU), February 2019. URL http://arxiv.org/abs/1803.08375. arXiv:1803.08375 [cs, stat].

Safwan Mahmood Al-Selwi, Mohd Fadzil Hassan, Said Jadid Abdulkadir, Amgad Muneer, Ebrahim Hamid Sumiea, Alawi Alqushaibi, and Mohammed Gamal Ragab. RNN-LSTM: From applications to modeling techniques and beyond—Systematic review. *Journal of King Saud University - Computer and Information Sciences*, 36(5):102068, June 2024. ISSN 1319-1578. doi: 10.1016/j.jksuci.2024.102068. URL https://www.sciencedirect.com/science/article/pii/S1319157824001575.

Waleed Al-Sit, Waleed Al-Nuaimy, Matteo Marelli, and Ali Al-Ataby. Visual texture for automated characterisation of geological features in borehole televiewer imagery. *Journal of Applied Geophysics*, 119:139–146, August 2015. ISSN 0926-9851. doi: 10.1016/j.jappgeo.2015.05.015. URL https://www.sciencedirect.com/science/article/pii/S0926985115001718.

C. Alimonti and E. Soldo. Study of geothermal power generation from a very deep oil well with a wellbore heat exchanger. *Renewable Energy*, 86:292–301, February 2016. ISSN 0960-1481. doi: 10.1016/j.renene.2015.08.031. URL https://www.sciencedirect.com/science/article/pii/S0960148115302317.

C. Alimonti, D. Berardi, D. Bocchetti, and E. Soldo. Coupling of energy conversion systems and wellbore heat exchanger in a depleted oil well. *Geothermal Energy*, 4(1):11, September 2016. ISSN 2195-9706. doi: 10.1186/s40517-016-0053-9. URL https://doi.org/10.1186/s40517-016-0053-9.

Fatimah Alzubaidi, Peyman Mostaghimi, Pawel Swietojanski, Stuart R. Clark, and Ryan T. Armstrong. Automated lithology classification from drill core images using convolutional neural networks. *Journal of Petroleum Science*

and Engineering, 197:107933, February 2021. ISSN 0920-4105. doi: 10.1016/j.petrol.2020.107933. URL https://www.sciencedirect.com/science/article/pii/S0920410520309888.

Fatimah Alzubaidi, Patrick Makuluni, Stuart R. Clark, Jan Erik Lie, Peyman Mostaghimi, and Ryan T. Armstrong. Automatic fracture detection and characterization from unwrapped drill-core images using mask R–CNN. *Journal of Petroleum Science and Engineering*, 208:109471, January 2022. ISSN 0920-4105. doi: 10.1016/j.petrol.2021.109471. URL https://www.sciencedirect.com/science/article/pii/S0920410521011141.

Mohammed S. Ameen and Ernest A. Hailwood. A new technology for the characterization of microfractured reservoirs (test case: Unayzah reservoir, Wudayhi field, Saudi Arabia). *AAPG Bulletin*, 92(1):31–52, 2008. ISSN 0149-1423. doi: 10.1306/08200706090. URL https://archives.datapages.com/data/bulletns/2008/01jan/BLTN06090/BLTN06090.HTM. Publisher: American Association of Petroleum Geologists.

Rommel Anatoli Quintanilla Cruz, Diego Carriço Cacau, Renato Moraes dos Santos, Evandro Jose Ribeiro Pereira, Fabiana Rodrigues Leta, and Esteban Gonzalez Clua. Improving Accuracy of Automatic Fracture Detection in Borehole Images with Deep Learning and GPUs. In *2017 30th SIBGRAPI Conference on Graphics, Patterns and Images (SIBGRAPI)*, pages 345–350, October 2017. doi: 10.1109/SIBGRAPI.2017.52. URL https://ieeexplore.ieee.org/abstract/document/8097332. ISSN: 2377-5416.

Billy J. Andrews, Jennifer J. Roberts, Zoe K. Shipton, Sabina Bigi, M. Chiara Tartarello, and Gareth Johnson. How do we see fractures? Quantifying subjective bias in fracture data collection. *Solid Earth*, 10(2):487–516, April 2019. ISSN 1869-9510. doi: 10.5194/se-10-487-2019. URL https://se.copernicus.org/articles/10/487/2019/. Publisher: Copernicus GmbH.

Esmail Ansari, Richard Hughes, and Christopher D. White. Statistical modeling of geopressured geothermal reservoirs. *Computers & Geosciences*, 103:36–50, 2017. URL https://www.sciencedirect.com/science/article/pii/S0098300417301905. Publisher: Elsevier.

Said Assous, Peter Elkington, Stuart Clark, and James Whetton. Automated detection of planar geologic features in borehole images. *GEOPHYSICS*, 79(1): D11–D19, January 2014. ISSN 0016-8033. doi: 10.1190/geo2013-0189.1. URL https://library.seg.org/doi/full/10.1190/geo2013-0189.1. Publisher: Society of Exploration Geophysicists.

Pramit Basu, Stephanie Morris, Stephen Morris, Nicklas Ritzmann, and Yan Zheng. Categorising Risk for Drilling and Completions Using High-Resolution Acoustic LWD Image Logs. OnePetro, September 2019. doi: 10.2118/195795-MS. URL https://dx.doi.org/10.2118/195795-MS.

F BATINI, G. BERTINI, A. BOTTAI, P. D. BURGASSI, G. CAPPETTI, G. GIANELLI, and M. PUXEDDU. San Pompeo 2 deep well: a high temperature and high pressure geothermal system. *San Pompeo 2 deep well: a high temperature and high pressure geothermal system*, pages 341–353, 1984.

Koenraad F. Beckers, Maciej Z. Lukawski, Brian J. Anderson, Michal C. Moore, and Jefferson W. Tester. Levelized costs of electricity and direct-use heat from Enhanced Geothermal Systems. *Journal of Renewable and Sustainable Energy*, 6(1):013141, February 2014. ISSN 1941-7012. doi: 10.1063/1.4865575. URL https://doi.org/10.1063/1.4865575.

Yoshua Bengio, Jérôme Louradour, Ronan Collobert, and Jason Weston. Curriculum learning. In *Proceedings of the 26th Annual International Conference on Machine Learning - ICML '09*, pages 1–8, Montreal, Quebec, Canada, 2009. ACM Press. ISBN 978-1-60558-516-1. doi: 10.1145/1553374.1553380. URL http://portal.acm.org/citation.cfm?doid=1553374.1553380.

J. H. Bennett, W. B. Easton, J. R. Guard, and T. H. Mott Jr. Introduction to semi-automated mathematics. *Final Rep. No. AFCRL*, pages 63–180, 1963.

Mouhammed Jandal Berro and Matthias Reich. Laboratory investigations of a hybrid mud pulse telemetry (HMPT) – A new approach for speeding up the transmitting of MWD/LWD data in deep boreholes. *Journal of Petroleum Science and Engineering*, 183:106374, December 2019. ISSN 0920-4105. doi: 10.1016/j.petrol.2019.106374. URL https://www.sciencedirect.com/science/article/pii/S0920410519307958.

Giovanni Bertini, M. Casini, G. Gianelli, and Enrico Pandeli. Geological structure of a long-living geothermal system, Larderello, Italy. 2006. URL https://flore.unifi.it/handle/2158/255217. Accepted: 2015-04-28T13:14:49Z.

Giorgio Besagni, Marco Borgarello, Lidia Premoli Vilà, Behzad Najafi, and Fabio Rinaldi. MOIRAE – bottom-up MOdel to compute the energy consumption of the Italian REsidential sector: Model design, validation and evaluation of electrification pathways. *Energy*, 211:118674, November 2020. ISSN 0360-5442. doi: 10.1016/j.energy.2020.118674. URL https://www.sciencedirect.com/science/article/pii/S0360544220317825.

Zhengfa Bi, Xinming Wu, Zhicheng Geng, and Haishan Li. Deep Relative Geologic Time: A Deep Learning Method for Simultaneously Interpreting 3-D Seismic Horizons and Faults. *Journal of Geophysical Research: Solid Earth*, 126(9):e2021JB021882, September 2021. ISSN 2169-9313, 2169-9356. doi: 10.1029/2021JB021882. URL https://agupubs.onlinelibrary.wiley.com/doi/10.1029/2021JB021882.

Daniel T. Birdsell and Martin O. Saar. Use Of A Coupled Thermo-Hydro-Mechanical Model To Constrain The Risk Of Ground Surface Deformation Due To Subsurface Energy Storage And Production. In *AAPG European Region, 3rd Hydrocarbon Geothermal Cross Over Technology Workshop*,

2019. URL https://www.searchanddiscovery.com/abstracts/html/2019/geneva-90346/abstracts/2019.ER.Geneva.11.html.

Quinten D. Boersma, Pierre Olivier Bruna, Stephan de Hoop, Francesco Vinci, Ali Moradi Tehrani, and Giovanni Bertotti. The impact of natural fractures on heat extraction from tight Triassic sandstones in the West Netherlands Basin: a case study combining well, seismic and numerical data. *Netherlands Journal of Geosciences*, 100:e6, January 2021. ISSN 0016-7746, 1573-9708. doi: 10.1017/njg.2020.21.

C.E. Bond, A.D. Gibbs, Z.K. Shipton, and S. Jones. What do you think this is? "Conceptual uncertainty" in geoscience interpretation. *GSA Today*, 17(11): 4, 2007. ISSN 1052-5173. doi: 10.1130/GSAT01711A.1. URL http://www.geosociety.org/gsatoday/archive/17/11/pdf/i1052-5173-17-11-4.pdf.

Fateh Bouchaala, Mohammed Y. Ali, Jun Matsushima, Youcef Bouzidi, Eric M. Takam Takougang, Aala A. I. Mohamed, and Akmal Sultan. Azimuthal investigation of compressional seismic-wave attenuation in a fractured reservoir. *GEOPHYSICS*, 84(6):B437–B446, November 2019. ISSN 0016-8033. doi: 10.1190/geo2019-0079.1. URL https://library.seg.org/doi/full/10.1190/geo2019-0079.1. Publisher: Society of Exploration Geophysicists.

Gianni Brauwers and Flavius Frasincar. A general survey on attention mechanisms in deep learning. *IEEE Transactions on Knowledge and Data Engineering*, 35(4):3279–3298, 2021. URL https://ieeexplore.ieee.org/abstract/document/9609539/?casa_token=TuIfv9ZWxNwAAAAA:xXf4Pe4DILs3BbvK9ZO6PbLwYGWbInF1_RRQlvvtAO_GHMzXY-p7-yHK-25XlzOUyMZ9he4. Publisher: IEEE.

Wilhelm Burger and Mark J. Burge. *Digital Image Processing: An Algorithmic Introduction Using Java*. Texts in Computer Science. Springer London, London, 2016. ISBN 978-1-4471-6683-2 978-1-4471-6684-9. doi: 10.1007/978-1-4471-6684-9. URL http://link.springer.com/10.1007/978-1-4471-6684-9.

Daniela Cadena-Herrera, Joshua E. Esparza-De Lara, Nancy D. Ramírez-Ibañez, Carlos A. López-Morales, Néstor O. Pérez, Luis F. Flores-Ortiz, and Emilio Medina-Rivero. Validation of three viable-cell counting methods: Manual, semi-automated, and automated. *Biotechnology Reports*, 7:9–16, September 2015. ISSN 2215-017X. doi: 10.1016/j.btre.2015.04.004. URL https://www.sciencedirect.com/science/article/pii/S2215017X15000235.

G. M. Cameli, I. Dini, and Domenico Liotta. Upper crustal structure of the Larderello geothermal field as a feature of post-collisional extensional tectonics (Southern Tuscany, Italy). 1993. URL https://ricerca.uniba.it/handle/11586/1386. Accepted: 2015-07-23T10:17:44Z.

G. Cavarretta and F. Tecce. Contact metasomatic and hydrothermal minerals in the SH2 deep well, Sabatini volcanic district, Latium, Italy. *Geothermics*,

16(2):127–145, January 1987. ISSN 0375-6505. doi: 10.1016/0375-6505(87)90061-7. URL https://www.sciencedirect.com/science/article/pii/0375650587900617.

Bulent Cavas and Pinar Cavas. Multiple Intelligences Theory—Howard Gardner. In Ben Akpan and Teresa J. Kennedy, editors, *Science Education in Theory and Practice: An Introductory Guide to Learning Theory*, pages 405–418. Springer International Publishing, Cham, 2020. ISBN 978-3-030-43620-9. doi: 10.1007/978-3-030-43620-9_27. URL https://doi.org/10.1007/978-3-030-43620-9_27.

Rima Chatterjee, Saurabh Datta Gupta, and Partha Pratim Mandal. Fracture and stress orientation from borehole image logs: A case study from Cambay basin, India. *Journal of the Geological Society of India*, 89:573–580, 2017. URL https://idp.springer.com/authorize/casa?redirect_uri=https://link.springer.com/article/10.1007/s12594-017-0646-3&casa_token=slQXLju5wkIAAAAA:sbhNe1StiwuuJTD2KKNem-SouZ0iXp1kQOcHhZhNyv_02Cumt3SHlJ4W9Zbb2XpKncS50kdsyuDtlW0. Publisher: Springer.

Siyuan Chen, Qi Zhang, Philip Andrews-Speed, and Benjamin Mclellan. Quantitative assessment of the environmental risks of geothermal energy: A review. *Journal of Environmental Management*, 276:111287, December 2020. ISSN 0301-4797. doi: 10.1016/j.jenvman.2020.111287. URL https://www.sciencedirect.com/science/article/pii/S0301479720312111.

Satinder Chopra and Kurt J. Marfurt. Seismic attributes — A historical perspective. *GEOPHYSICS*, 70(5):3SO–28SO, September 2005. ISSN 0016-8033. doi: 10.1190/1.2098670. URL https://library.seg.org/doi/abs/10.1190/1.2098670. Publisher: Society of Exploration Geophysicists.

D. U. ChunYu, Xing Qiang, Zhang JinYan, Wang Jun, L. I. U. BaoDi, and Wang YanJiang. Blank strips filling for electrical logging images based on attention-constrained deep generative network. *Progress in Geophysics*, 37(4):1548–1558, August 2022. ISSN 1004-2903. doi: 10.6038/pg2022FF0320. URL http://www.progeophys.cn//en/article/doi/10.6038/pg2022FF0320. Publisher: Progress in Geophysics.

Maurizio Corbetta and Gordon L. Shulman. Control of goal-directed and stimulus-driven attention in the brain. *Nature Reviews Neuroscience*, 3(3):201–215, March 2002. ISSN 1471-0048. doi: 10.1038/nrn755. URL https://www.nature.com/articles/nrn755. Publisher: Nature Publishing Group.

Giorgia Dalla Santa, Antonio Galgaro, Raffaele Sassi, Matteo Cultrera, Paolo Scotton, Johannes Mueller, David Bertermann, Dimitrios Mendrinos, Riccardo Pasquali, Rodolfo Perego, Sebastian Pera, Eloisa Di Sipio, Giorgio Cassiani, Michele De Carli, and Adriana Bernardi. An updated ground thermal properties database for GSHP applications. *Geothermics*, 85:101758, May 2020. ISSN 0375-6505. doi: 10.1016/j.geothermics.2019.101758. URL https://www.sciencedirect.com/science/article/pii/S0375650519301944.

Alexandros Daniilidis, Betül Alpsoy, and Rien Herber. Impact of technical and economic uncertainties on the economic performance of a deep geothermal heat system. *Renewable Energy*, 114:805–816, December 2017. ISSN 0960-1481. doi: 10.1016/j.renene.2017.07.090. URL https://www.sciencedirect.com/science/article/pii/S0960148117307176.

Troyee Dasgupta, Swagato Dasgupta, and Soumyajit Mukherjee. Image Log Interpretation and Geomechanical Issues. In Soumyajit Mukherjee, editor, *Teaching Methodologies in Structural Geology and Tectonics*, pages 237–251. Springer Singapore, Singapore, 2019. ISBN 9789811327803 9789811327810. doi: 10.1007/978-981-13-2781-0_10. URL http://link.springer.com/10.1007/978-981-13-2781-0_10. Series Title: Springer Geology.

Paolo Dell'Aversana. *Cognition in geosciences: the feeding loop between geodisciplines, cognitive sciences and epistemology*. May 2013.

Michael Desmond, Evelyn Duesterwald, Kristina Brimijoin, Michelle Brachman, and Qian Pan. Semi-Automated Data Labeling. In *Proceedings of the NeurIPS 2020 Competition and Demonstration Track*, pages 156–169. PMLR, August 2021. URL https://proceedings.mlr.press/v133/desmond21a.html. ISSN: 2640-3498.

Ronald DiPippo. Ideal thermal efficiency for geothermal binary plants. *Geothermics*, 36(3):276–285, June 2007. ISSN 0375-6505. doi: 10.1016/j.geothermics.2007.03.002. URL https://www.sciencedirect.com/science/article/pii/S0375650507000375.

Hannah R. Doran, Theo Renaud, Gioia Falcone, Lehua Pan, and Patrick G. Verdin. Modelling an unconventional closed-loop deep borehole heat exchanger (DBHE): sensitivity analysis on the Newberry volcanic setting. *Geothermal Energy*, 9(1):4, December 2021. ISSN 2195-9706. doi: 10.1186/s40517-021-00185-0. URL https://geothermal-energy-journal.springeropen.com/articles/10.1186/s40517-021-00185-0.

O. Ericok and F. Gumrah. Uncertainty Assessment in Reserve Estimation of a Naturally Fractured Reservoir. OnePetro, June 2005. doi: 10.2118/2005-150. URL https://dx.doi.org/10.2118/2005-150.

Ö Eriçok and F. Gümrah. Naturally Fractured Carbonate Oil Reservoir: Reserve Estimation Study. *Energy Sources, Part A: Recovery, Utilization, and Environmental Effects*, February 2010. doi: 10.1080/15567030701744898. URL https://www.tandfonline.com/doi/abs/10.1080/15567030701744898. Publisher: Taylor & Francis Group.

Zeindra Ernando and Achmad Fathoni. Volcanic reservoir characterization of Jatibarang formation based on an integrated study of petrography, core, FMI, and well log. 2011. URL https://archives.datapages.com/data/ipa_pdf/2011/IPA11-G-166.pdf. Publisher: Indonesian Petroleum Association.

Matthew J. Euler. Intelligence and uncertainty: Implications of hierarchical predictive processing for the neuroscience of cognitive ability. *Neuroscience & Biobehavioral Reviews*, 94:93–112, November 2018. ISSN 0149-7634. doi: 10.1016/j.neubiorev.2018.08.013. URL https://www.sciencedirect.com/science/article/pii/S0149763418302045.

Mohammad Faiq Adenan, Ebrahim Fathi, Tim Carr, and Brian Panetta. Machine learning-based workflow for identifying fractures and baffles from Formation Micro Imager (FMI) log: A practical application in Illinois Basin Decatur Project (IBDP). In *SEG International Exposition and Annual Meeting*, pages SEG–2023. SEG, 2023. URL https://onepetro.org/SEGAM/proceedings-abstract/IMAGE23/All-IMAGE23/539435.

Atle Folkestad, Zbynek Veselovsky, and Paul Roberts. Utilising borehole image logs to interpret delta to estuarine system: A case study of the subsurface Lower Jurassic Cook Formation in the Norwegian northern North Sea. *Marine and Petroleum Geology*, 29(1):255–275, January 2012. ISSN 0264-8172. doi: 10.1016/j.marpetgeo.2011.07.008. URL https://www.sciencedirect.com/science/article/pii/S0264817211001735.

Marco Fonnesu, Denis Palermo, Mauro Galbiati, Marco Marchesini, Enrico Bonamini, and Daniel Bendias. A new world-class deep-water play-type, deposited by the syndepositional interaction of turbidity flows and bottom currents: The giant Eocene Coral Field in northern Mozambique. *Marine and Petroleum Geology*, 111:179–201, January 2020. ISSN 0264-8172. doi: 10.1016/j.marpetgeo.2019.07.047. URL https://www.sciencedirect.com/science/article/pii/S0264817219303575.

Karl Friston, Michael Breakspear, and Gustavo Deco. Perception and self-organized instability. *Frontiers in computational neuroscience*, 6: 44, 2012. URL https://www.frontiersin.org/articles/10.3389/fncom.2012.00044/full. Publisher: Frontiers Media SA.

R. Funiciello, G. Mariotti, M. Parotto, M. Preite-Martinez, F. Tecce, R. Toneatti, and B. Turi. Geology, mineralogy and stable isotope geochemistry of the cesano geothermal field (Sabatini Mts. volcanic system, Northern Latium, Italy). *Geothermics*, 8(1):55–73, January 1979. ISSN 0375-6505. doi: 10.1016/0375-6505(79)90066-X. URL https://www.sciencedirect.com/science/article/pii/037565057990066X.

Christina M. Funke, Judy Borowski, Karolina Stosio, Wieland Brendel, Thomas SA Wallis, and Matthias Bethge. Five points to check when comparing visual perception in humans and machines. *Journal of Vision*, 21(3):16–16, 2021. URL https://jov.arvojournals.org/article.aspx?articleid=2772393. Publisher: The Association for Research in Vision and Ophthalmology.

Yarin Gal and Zoubin Ghahramani. Dropout as a Bayesian Approximation: Representing Model Uncertainty in Deep Learning, October 2016. URL http://arxiv.org/abs/1506.02142. arXiv:1506.02142 [cs, stat].

Pranav Gangwani, Jayesh Soni, Himanshu Upadhyay, Santosh Joshi, and Master's Student. A Deep Learning Approach for Modeling of Geothermal Energy Prediction. 18(1), 2020.

Zhicheng Geng, Xinming Wu, Yunzhi Shi, and Sergey Fomel. Deep learning for relative geologic time and seismic horizons. *GEOPHYSICS*, 85(4):WA87–WA100, July 2020. ISSN 0016-8033, 1942-2156. doi: 10.1190/geo2019-0252.1. URL https://library.seg.org/doi/10.1190/geo2019-0252.1.

Shabnam Gharibi, Emad Mortezazadeh, Seyed Jalaledin Hashemi Aghcheh Bodi, and Ali Vatani. Feasibility study of geothermal heat extraction from abandoned oil wells using a U-tube heat exchanger. *Energy (Oxford)*, 153:554–567, 2018. URL https://inis.iaea.org/search/search.aspx?orig_q=RN:52120534.

Hazel Gibson, Iain S. Stewart, Sabine Pahl, and Alison Stokes. A "mental models" approach to the communication of subsurface hydrology and hazards. *Hydrology and Earth System Sciences*, 20(5):1737–1749, May 2016. ISSN 1027-5606. doi: 10.5194/hess-20-1737-2016. URL https://hess.copernicus.org/articles/20/1737/2016/. Publisher: Copernicus GmbH.

K. Glossop, P. J. G. Lisboa, P. C. Russell, A. Siddans, and G. R. Jones. An Implementation of the Hough Transformation for the Identification and Labelling of Fixed Period Sinusoidal Curves. *Computer Vision and Image Understanding*, 74(1):96–100, April 1999. ISSN 1077-3142. doi: 10.1006/cviu.1999.0747. URL https://www.sciencedirect.com/science/article/pii/S1077314299907476.

Gianluca Gola, Giovanni Bertini, Marco Bonini, Serena Botteghi, Andrea Brogi, Roberto De Franco, Andrea Dini, Assunta Donato, Giovanni Gianelli, Domenico Liotta, Adele Manzella, Domenico Montanari, Giordano Montegrossi, Lorenzo Petracchini, Giovanni Ruggieri, Alessandro Santilano, Davide Scrocca, and Eugenio Trumpy. Data integration and conceptual modelling of the Larderello geothermal area, Italy. *Energy Procedia*, 125:300–309, September 2017. ISSN 1876-6102. doi: 10.1016/j.egypro.2017.08.201. URL https://www.sciencedirect.com/science/article/pii/S1876610217337049.

Gianluca Gola, Eloisa Di Sipio, Marina Facci, Antonio Galgaro, and Adele Manzella. Geothermal deep closed-loop heat exchangers: A novel technical potential evaluation to answer the power and heat demands. *Renewable Energy*, 198:1193–1209, 2022. URL https://www.sciencedirect.com/science/article/pii/S096014812201237X. Publisher: Elsevier.

Bo Gong, Ela Manuel, Youfang Liu, David Forand, Tom Malizia, Vahid Tohidi, and Alex Saldana. Interpretation of LWD Acoustic Borehole Image Logs: Case Studies from North American Shale Plays. OnePetro, May 2021.

doi: 10.30632/SPWLA-2021-0085. URL https://dx.doi.org/10.30632/SPWLA-2021-0085.

Klaus Greff, Rupesh K. Srivastava, Jan Koutník, Bas R. Steunebrink, and Jürgen Schmidhuber. LSTM: A Search Space Odyssey. *IEEE Transactions on Neural Networks and Learning Systems*, 28(10):2222–2232, October 2017. ISSN 2162-2388. doi: 10.1109/TNNLS.2016.2582924. URL https://ieeexplore.ieee.org/abstract/document/7508408?casa_token=VXb6gF7KjJcAAAAA:3HiQ0ZJf6DafNH0jVZRMWFi4tsGqDbFMLL812niXZtP_2yLpVWUg5tz58hCJqwKHNZtKL7fbwg. Conference Name: IEEE Transactions on Neural Networks and Learning Systems.

Kinjal Dhar Gupta, Valentina Vallega, Hiren Maniar, Philippe Marza, Hui Xie, Koji Ito, and Aria Abubakar. A Deep-Learning Approach for Borehole Image Interpretation. OnePetro, June 2019. doi: 10.30632/T60ALS-2019_BB. URL https://dx.doi.org/10.30632/T60ALS-2019_BB.

Abdul Mueed Hafiz, Shabir Ahmad Parah, and Rouf Ul Alam Bhat. Attention mechanisms and deep learning for machine vision: A survey of the state of the art, June 2021. URL http://arxiv.org/abs/2106.07550. arXiv:2106.07550 [cs].

Jonathan Hall, Marco Ponzi, Mauro Gonfalini, and Giorgio Maletti. Automatic Extraction And Characterisation Of Geological Features And Textures Front Borehole Images And Core Photographs. OnePetro, June 1996. URL https://dx.doi.org/.

Sepp Hochreiter and Jürgen Schmidhuber. Long Short-Term Memory. *Neural Computation*, 9(8):1735–1780, November 1997. ISSN 0899-7667. doi: 10.1162/neco.1997.9.8.1735. URL https://ieeexplore.ieee.org/abstract/document/6795963. Conference Name: Neural Computation.

Sirous Hosseinzadeh, Ali Kadkhodaie, David A. Wood, Reza Rezaee, and Rahim Kadkhodaie. Discrete fracture modeling by integrating image logs, seismic attributes, and production data: a case study from Ilam and Sarvak Formations, Danan Oilfield, southwest of Iran. *Journal of Petroleum Exploration and Production Technology*, 13(4):1053–1083, April 2023. ISSN 2190-0558, 2190-0566. doi: 10.1007/s13202-022-01586-y. URL https://link.springer.com/10.1007/s13202-022-01586-y.

Hussein Hoteit, Xupeng He, Bicheng Yan, and Volker Vahrenkamp. Uncertainty quantification and optimization method applied to time-continuous geothermal energy extraction. *Geothermics*, 110:102675, May 2023. ISSN 0375-6505. doi: 10.1016/j.geothermics.2023.102675. URL https://www.sciencedirect.com/science/article/pii/S0375650523000299.

Paul V. C. Hough. Method and means for recognizing complex patterns, December 1962. URL https://patents.google.com/patent/US3069654/en.

L. M. Hu, T. S. Ge, Y. Jiang, and R. Z. Wang. Performance study on composite desiccant material coated fin-tube heat exchangers. *International journal of heat and mass transfer*, 90:109–120, 2015. URL https://www.sciencedirect.com/science/article/pii/S0017931015006559?casa_token=5BOv72gJ3L8AAAAA:EBdABRjn8ArbnNuxCHSHxIaB_yVsS2S3wm7dN0vdQhJ9zXq2wU6ft3RaBnZjkG39zDNixsQ9. Publisher: Elsevier.

Xincheng Hu, Jonathan Banks, Yunting Guo, and Wei Victor Liu. Retrofitting abandoned petroleum wells as doublet deep borehole heat exchangers for geothermal energy production—a numerical investigation. *Renewable Energy*, 176:115–134, October 2021. ISSN 0960-1481. doi: 10.1016/j.renene.2021.05.061. URL https://www.sciencedirect.com/science/article/pii/S0960148121007370.

Jida Huang, Hongyue Sun, Tsz-Ho Kwok, Chi Zhou, and Wenyao Xu. Geometric deep learning for shape correspondence in mass customization by three-dimensional printing. *Journal of Manufacturing Science and Engineering*, 142(6):061003, 2020. URL https://asmedigitalcollection.asme.org/manufacturingscience/article-abstract/142/6/061003/1081979. Publisher: American Society of Mechanical Engineers.

Jiajun Jiang, Rui Xu, Scott C. James, and Chicheng Xu. Deep-Learning-Based Vuggy Facies Identification from Borehole Images. *SPE Reservoir Evaluation & Engineering*, 24(01):250–261, February 2021. ISSN 1094-6470. doi: 10.2118/204216-PA. URL https://doi.org/10.2118/204216-PA.

Guodong Jin, Huilin Xing, Tianbin Li, Rongxin Zhang, Junbiao Liu, Zhiwei Guo, and Zihan Ma. An Integrated Approach of Numerical Well Test for Well Intersecting Fractures Based on FMI Image. *Lithosphere*, 2021(Special 1):4421135, January 2022. ISSN 1941-8264. doi: 10.2113/2022/4421135. URL https://doi.org/10.2113/2022/4421135.

Yuchen Jin, Xuqing Wu, Jiefu Chen, and Yueqin Huang. Using a Physics-Driven Deep Neural Network to Solve Inverse Problems for LWD Azimuthal Resistivity Measurements. OnePetro, June 2019. doi: 10.30632/T60ALS-2019_IIII. URL https://dx.doi.org/10.30632/T60ALS-2019_IIII.

Egill Juliusson and Roland N. Horne. Optimization of injection scheduling in fractured geothermal reservoirs. *Geothermics*, 48:80–92, 2013. URL https://www.sciencedirect.com/science/article/pii/S0375650513000412. Publisher: Elsevier.

Murat A. Kaplanoğlu, Alper Baba, and Gulden Gokcen Akkurt. Use of abandoned oil wells in geothermal systems in Turkey. *Geomechanics and Geophysics for Geo-Energy and Geo-Resources*, 6(1):2, October 2019. ISSN 2363-8427. doi: 10.1007/s40948-019-00125-0. URL https://doi.org/10.1007/s40948-019-00125-0.

Osamu Kato and Y. Sakagawa. Characteristics of fractures based on FMI logs and cores in well WD-1 in the Kakkonda geothermal field, Japan. Technical report, Geothermal Resources Council, Davis, CA (United States), 1995. URL https://www.osti.gov/biblio/175643.

Gabriela I. Keeton, Matthew J. Pranter, Rex D. Cole, and Edmund R. (Gus) Gustason. Stratigraphic architecture of fluvial deposits from borehole images, spectral-gamma-ray response, and outcrop analogs, Piceance Basin, Colorado. *AAPG Bulletin*, 99(10):1929–1956, October 2015. ISSN 0149-1423. doi: 10.1306/05071514025. URL https://doi.org/10.1306/05071514025.

Joseph Kelly and Christopher McDermott. Numerical modelling of a deep closed-loop geothermal system: Evaluating the Eavor-Loop. *AIMS Geosciences*, 8(2):175–212, February 2022. ISSN 2471-2132. doi: 10.3934/geosci.2022011. URL https://www.research.ed.ac.uk/en/publications/numerical-modelling-of-a-deep-closed-loop-geothermal-system-evalu. Publisher: AIMS Press.

F. Khoshbakht, M. Azizzadeh, H. Memarian, G. H. Nourozi, and S. A. Moallemi. Comparison of electrical image log with core in a fractured carbonate reservoir. *Journal of Petroleum Science and Engineering*, 86-87:289–296, May 2012. ISSN 0920-4105. doi: 10.1016/j.petrol.2012.03.007. URL https://www.sciencedirect.com/science/article/pii/S0920410512000605.

Diederik P. Kingma and Jimmy Ba. Adam: A Method for Stochastic Optimization, January 2017. URL http://arxiv.org/abs/1412.6980. arXiv:1412.6980 [cs].

Stephen Jay Kline. *Conceptual Foundations for Multidisciplinary Thinking*. Stanford University Press, January 1995. ISBN 978-0-8047-6393-6. doi: 10.1515/9780804763936. URL https://www.degruyter.com/document/doi/10.1515/9780804763936/html.

C. Klotz, P. Bond, I. Wasserman, and S. Priegnitz. A New Mud Pulse Telemetry System for Enhanced MWD/LWD Applications. OnePetro, March 2008. doi: 10.2118/112683-MS. URL https://dx.doi.org/10.2118/112683-MS.

Ehsan Kosari, Sajjad Ghareh-Cheloo, Ali Kadkhodaie-Ilkhchi, and Abbas Bahroudi. Fracture characterization by fusion of geophysical and geomechanical data: a case study from the Asmari reservoir, the Central Zagros fold-thrust belt. *Journal of Geophysics and Engineering*, 12(1):130–143, February 2015. ISSN 1742-2132. doi: 10.1088/1742-2132/12/1/130. URL https://doi.org/10.1088/1742-2132/12/1/130.

Tomasz Kujawa, Władysław Nowak, and Aleksander A. Stachel. Utilization of existing deep geological wells for acquisitions of geothermal energy. *Energy*, 31(5):650–664, April 2006. ISSN 0360-5442. doi: 10.1016/j.energy.2005.05.002. URL https://www.sciencedirect.com/science/article/pii/S0360544205001027.

Anthea Lacchia, Geertje Schuitema, and Fergus McAuliffe. The human side of geoscientists: comparing geoscientists' and non-geoscientists' cognitive and affective responses to geology. *Geoscience Communication*, 3(2):291–302, September 2020. doi: 10.5194/gc-3-291-2020. URL https://gc.copernicus.org/articles/3/291/2020/. Publisher: Copernicus GmbH.

Jin Lai, Guiwen Wang, Song Wang, Juntao Cao, Mei Li, Xiaojiao Pang, Chuang Han, Xuqiang Fan, Liu Yang, Zhibo He, and Ziqiang Qin. A review on the applications of image logs in structural analysis and sedimentary characterization. *Marine and Petroleum Geology*, 95:139–166, August 2018. ISSN 0264-8172. doi: 10.1016/j.marpetgeo.2018.04.020. URL https://www.sciencedirect.com/science/article/pii/S0264817218301946.

S. E. Laubach, R. H. Lander, L. J. Criscenti, L. M. Anovitz, J. L. Urai, R. M. Pollyea, J. N. Hooker, W. Narr, M. A. Evans, S. N. Kerisit, J. E. Olson, T. Dewers, D. Fisher, R. Bodnar, B. Evans, P. Dove, L. M. Bonnell, M. P. Marder, and L. Pyrak-Nolte. The Role of Chemistry in Fracture Pattern Development and Opportunities to Advance Interpretations of Geological Materials. *Reviews of Geophysics*, 57(3):1065–1111, 2019. ISSN 1944-9208. doi: 10.1029/2019RG000671. URL https://onlinelibrary.wiley.com/doi/abs/10.1029/2019RG000671. _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1029/2019RG000671.

Yann LeCun, Yoshua Bengio, and Geoffrey Hinton. Deep learning. *Nature*, 521 (7553):436–444, May 2015. ISSN 1476-4687. doi: 10.1038/nature14539. URL https://www.nature.com/articles/nature14539. Number: 7553 Publisher: Nature Publishing Group.

Baptiste Lepillier, Pierre-Olivier Bruna, David Bruhn, Eivind Bastesen, Alexandros Daniilidis, Óscar Garcia, Anita Torabi, and Walter Wheeler. From outcrop scanlines to discrete fracture networks, an integrative workflow. *Journal of Structural Geology*, 133:103992, April 2020. ISSN 0191-8141. doi: 10.1016/j.jsg.2020.103992. URL https://www.sciencedirect.com/science/article/pii/S0191814119303050.

Tianyu Li, Sogo Shiozawa, and Mark W. McClure. Thermal breakthrough calculations to optimize design of a multiple-stage Enhanced Geothermal System. *Geothermics*, 64:455–465, November 2016. ISSN 0375-6505. doi: 10.1016/j.geothermics.2016.06.015. URL https://www.sciencedirect.com/science/article/pii/S0375650516300712.

Youqiang Liao, Xiaohui Sun, Baojiang Sun, Zhiyuan Wang, Jintang Wang, and Xuerui Wang. Geothermal exploitation and electricity generation from multibranch U-shaped well–enhanced geothermal system. *Renewable Energy*, 163:2178–2189, January 2021. ISSN 0960-1481. doi: 10.1016/j.renene.2020.10.090. URL https://www.sciencedirect.com/science/article/pii/S0960148120316566.

Minghao Liu, Jiaheng Wei, Yang Liu, and James Davis. Do humans and machines have the same eyes? Human-machine perceptual differences on image classification, April 2023. URL http://arxiv.org/abs/2304.08733. arXiv:2304.08733 [cs].

Daniel S. Lopes, Rafael Kuffner dos Anjos, and Joaquim A. Jorge. Assessing the usability of tile-based interfaces to visually navigate 3-D parameter domains. *International Journal of Human-Computer Studies*, 118:1–13, October 2018. ISSN 1071-5819. doi: 10.1016/j.ijhcs.2018.05.005. URL http://www.scopus.com/inward/record.url?scp=85047941372&partnerID=8YFLogxK.

Maciej Z. Lukawski, Brian J. Anderson, Chad Augustine, Louis E. Capuano, Koenraad F. Beckers, Bill Livesay, and Jefferson W. Tester. Cost analysis of oil, gas, and geothermal well drilling. *Journal of Petroleum Science and Engineering*, 118:1–14, June 2014. ISSN 0920-4105. doi: 10.1016/j.petrol.2014.03.012. URL https://www.sciencedirect.com/science/article/pii/S0920410514000813.

Maciej Z. Lukawski, Rachel L. Silverman, and Jefferson W. Tester. Uncertainty analysis of geothermal well drilling and completion costs. *Geothermics*, 64:382–391, November 2016. ISSN 0375-6505. doi: 10.1016/j.geothermics.2016.06.017. URL https://www.sciencedirect.com/science/article/pii/S0375650516300736.

John W. Lund. The use of downhole heat exchangers. *Geothermics*, 32 (4):535–543, August 2003. ISSN 0375-6505. doi: 10.1016/j.geothermics.2003.06.002. URL https://www.sciencedirect.com/science/article/pii/S0375650503000853.

Minh-Thang Luong. Effective approaches to attention-based neural machine translation. *arXiv preprint arXiv:1508.04025*, 2015. URL https://courses.grainger.illinois.edu/cs546/sp2018/Slides/Mar15_Luong.pdf.

Stefan M. Luthi. Density Borehole Imaging. In Stefan M. Luthi, editor, *Geological Well Logs: Their Use in Reservoir Modeling*, pages 147–153. Springer, Berlin, Heidelberg, 2001. ISBN 978-3-662-04627-2. doi: 10.1007/978-3-662-04627-2_7. URL https://doi.org/10.1007/978-3-662-04627-2_7.

Laurent Maerten, Xavier Legrand, Claire Castagnac, Marie Lefranc, Jean-Pierre Joonnekindt, and Frantz Maerten. Fault-related fracture modeling in the complex tectonic environment of the Malay Basin, offshore Malaysia: An integrated 4D geomechanical approach. *Marine and Petroleum Geology*, 105:222–237, July 2019. ISSN 0264-8172. doi: 10.1016/j.marpetgeo.2019.04.025. URL https://www.sciencedirect.com/science/article/pii/S0264817219301758.

Carlos Maeso, Isabelle Dubourg, Daniel Quesada, and Walid Abou El-Nour. Uncertainties in fracture apertures calculated from electrical borehole images. In *International Petroleum Technology Conference*, page D031S030R005. IPTC, 2015. URL https://onepetro.org/IPTCONF/proceedings-abstract/15IPTC/3-15IPTC/153647.

Shie Mannor, Dori Peleg, and Reuven Rubinstein. The cross entropy method for classification. In *Proceedings of the 22nd international conference on Machine learning*, ICML '05, pages 561–568, New York, NY, USA, 2005. Association for Computing Machinery. ISBN 978-1-59593-180-1. doi: 10.1145/1102351.1102422. URL https://doi.org/10.1145/1102351.1102422.

James Manyika, Jake Silberg, and Brittany Presten. What do we do about the biases in Al. *Harvard Business Review*, 2019. URL http://www.nwscrs.org/uploads/4/1/5/1/41511209/biases_detected_by_ai.pdf.

Adele Manzella, Davide Serra, Gabriele Cesari, Eleonora Bargiacchi, Maurizio Cei, Paolo Cerutti, Paolo Conti, Geoffrey Giudetti, Mirco Lupi, and Maurizio Vaccaro. *Geothermal Energy Use, Country Update for Italy*. June 2019.

Amir Mazdarani, Ali Kadkhodaie, David A. Wood, and Zohreh Soluki. Natural fractures characterization by integration of FMI logs, well logs and core data: a case study from the Sarvak Formation (Iran). *Journal of Petroleum Exploration and Production Technology*, 13(5):1247–1263, May 2023. ISSN 2190-0558, 2190-0566. doi: 10.1007/s13202-023-01611-8. URL https://link.springer.com/10.1007/s13202-023-01611-8.

William McIlhagga. Estimates of edge detection filters in human vision. *Vision Research*, 153:30–36, 2018. URL https://www.sciencedirect.com/science/article/pii/S0042698918302050. Publisher: Elsevier.

Rajiv Mehrotra, Kameswara Rao Namuduri, and Nagarajan Ranganathan. Gabor filter-based edge detection. *Pattern recognition*, 25(12):1479–1494, 1992. URL https://www.sciencedirect.com/science/article/pii/003132039290121X. Publisher: Elsevier.

Mehdi Mirza and Simon Osindero. Conditional Generative Adversarial Nets, November 2014. URL http://arxiv.org/abs/1411.1784. arXiv:1411.1784 [cs].

Soumendra Mohanty and Sachin Vyas. *How to Compete in the Age of Artificial Intelligence: Implementing a Collaborative Human-Machine Strategy for Your Business*. Apress, September 2018. ISBN 978-1-4842-3808-0. Google-Books-ID: RV5wDwAAQBAJ.

A. Molossi. Enhancing Geothermal Reservoirs Fracture Interpretation Using Deep Learning: a Case Study from the Western Netherland Basin. In *Fifth EAGE Global Energy Transition Conference & Exhibition (GET 2024)*, pages 1–5, Rotterdam, Netherlands,, 2024. European Association of Geoscientists & Engineers. doi: 10.3997/2214-4609.202421218. URL https://www.earthdoc.org/content/papers/10.3997/2214-4609.202421218.

Attilio Molossi and Michele Pipan. Exploiting image logs to reduce drilling hazards: an innovative Artificial Intelligence methodology applied in East Africa.

Geophysical Journal International, 235(1):942–950, May 2023. ISSN 0956-540X, 1365-246X. doi: 10.1093/gji/ggad286. URL https://academic.oup.com/gji/article/235/1/942/7223463.

Attilio Molossi, Giacomo Roncoroni, and Michele Pipan. Efficient Logging-While-Drilling Image Logs Interpretation Using Deep Learning. *Petrophysics-The SPWLA Journal of Formation Evaluation and Reservoir Description*, 65 (03):365–387, 2024. URL https://onepetro.org/petrophysics/article/65/03/365/545754. Publisher: OnePetro.

Nazmul Haque Mondol. Well Logging: Principles, Applications and Uncertainties. In Knut Bjørlykke, editor, *Petroleum Geoscience*, pages 385–425. Springer Berlin Heidelberg, Berlin, Heidelberg, 2015. ISBN 978-3-642-34131-1 978-3-642-34132-8. doi: 10.1007/978-3-642-34132-8_16. URL https://link.springer.com/10.1007/978-3-642-34132-8_16.

Saleh M. Mwachaka, Aiping Wu, and Qingqing Fu. A review of mud pulse telemetry signal impairments modeling and suppression methods. *Journal of Petroleum Exploration and Production Technology*, 9(1):779–792, March 2019. ISSN 2190-0566. doi: 10.1007/s13202-018-0483-y. URL https://doi.org/10.1007/s13202-018-0483-y.

Meinard Müller. *Information Retrieval for Music and Motion*. Springer Berlin Heidelberg, Berlin, Heidelberg, 2007. ISBN 978-3-540-74047-6 978-3-540-74048-3. doi: 10.1007/978-3-540-74048-3. URL http://link.springer.com/10.1007/978-3-540-74048-3.

Kabir Nadkarni, Lianne M. Lefsrud, Daniel Schiffner, and Jonathan Banks. Converting oil wells to geothermal resources: Roadmaps and roadblocks for energy transformation. *Energy Policy*, 161:112705, February 2022. ISSN 0301-4215. doi: 10.1016/j.enpol.2021.112705. URL https://www.sciencedirect.com/science/article/pii/S030142152100570X.

Masaki Nakada, Honglin Chen, and Demetri Terzopoulos. Deep learning of biomimetic visual perception for virtual humans. In *Proceedings of the 15th ACM Symposium on Applied Perception*, pages 1–8, Vancouver British Columbia Canada, August 2018. ACM. ISBN 978-1-4503-5894-1. doi: 10.1145/3225153.3225161. URL https://dl.acm.org/doi/10.1145/3225153.3225161.

Zhaoyang Niu, Guoqiang Zhong, and Hui Yu. A review on the attention mechanism of deep learning. *Neurocomputing*, 452: 48–62, 2021. URL https://www.sciencedirect.com/science/article/pii/S092523122100477X?casa_token=xEUpX_dOhLQAAAAA:bn37iWlddJXPFvb0CEBnTO5Y3Jyq8BAGKiGwsI5PZ5vLHeKb8mLl6rCs89v9-oZlYl7Vqid1. Publisher: Elsevier.

Sait I. Ozkaya and Joerg Mattner. Fracture connectivity from fracture intersections in borehole image logs. *Computers & Geosciences*, 29(2):143–153, 2003. URL https://www.

sciencedirect.com/science/article/pii/S0098300402001139?casa_token=RbX1XytB_6kAAAAA:UZABXd1flBsAWcJQ8z_b17m-QBehdo5iKBWdbvD-9hlmMxE5Mi6BkMvw6MohrxIynKE06s15. Publisher: Elsevier.

C. Ozol, E. Climer, R. Aiello, D. Di Tommaso, M. Borghi, N. Cominesi, and M. Medaglia. The Use of Real-time Microlmager Data to Validate Structural Uncertainties. Madrid, Spain, June 2015. doi: 10.3997/2214-4609.201412647. URL http://www.earthdoc.org/publication/publicationdetails/?publication=80401.

Elisabet Palomo, Antonio Colmenar-Santos, and Enrique Rosales-Asensio. Measures to Remove Geothermal Energy Barriers in the European Union. In Elisabet Palomo, Antonio Colmenar-Santos, and Enrique Rosales-Asensio, editors, *Potential of Low-Medium Enthalpy Geothermal Energy: Hybridization and Application in Industry*, Green Energy and Technology, pages 9–45. Springer International Publishing, Cham, 2022. ISBN 978-3-030-95626-4. doi: 10.1007/978-3-030-95626-4_2. URL https://doi.org/10.1007/978-3-030-95626-4_2.

S. N. Pandey and Vikram Vishal. Sensitivity analysis of coupled processes and parameters on the performance of enhanced geothermal systems. *Scientific reports*, 7(1):17057, 2017. URL https://www.nature.com/articles/s41598-017-14273-4. Publisher: Nature Publishing Group UK London.

V. Pasquale, G. Gola, P. Chiozzi, and M. Verdoya. Thermophysical properties of the Po Basin rocks. *Geophysical Journal International*, 186(1):69–81, July 2011. ISSN 0956-540X. doi: 10.1111/j.1365-246X.2011.05040.x. URL https://doi.org/10.1111/j.1365-246X.2011.05040.x.

Jeremy R. Patterson, Michael Cardiff, and Kurt L. Feigl. Optimizing geothermal production in fractured rock reservoirs under uncertainty. *Geothermics*, 88:101906, November 2020. ISSN 0375-6505. doi: 10.1016/j.geothermics.2020.101906. URL https://www.sciencedirect.com/science/article/pii/S037565052030198X.

Sonali Pattnaik, Songhua Chen, Adly Helba, and Shouxiang Ma. Automatic Carbonate Rock Facies Identification with Deep Learning. OnePetro, October 2020. doi: 10.2118/201673-MS. URL https://dx.doi.org/10.2118/201673-MS.

C. Pauselli, G. Gola, P. Mancinelli, E. Trumpy, M. Saccone, A. Manzella, and G. Ranalli. A new surface heat flow map of the Northern Apennines between latitudes 42.5 and 44.5 N. *Geothermics*, 81:39–52, September 2019. ISSN 0375-6505. doi: 10.1016/j.geothermics.2019.04.002. URL https://www.sciencedirect.com/science/article/pii/S0375650518301421.

D. C. P. Peacock. The *certainty matrix* for fault data and interpretations. *Geothermics*, 125:103197, January 2025. ISSN 0375-6505. doi: 10.1016/j.

geothermics.2024.103197. URL https://www.sciencedirect.com/science/article/pii/S0375650524002839.

A. Peccerillo and C. Donati. The tuscan magmatic province. 72:27–39, January 2003.

A. Petrov, N. Belayouni, M. Belouahchia, L. Xin, Y. Fan, and A. Noufal. Deep Learning Algorithms Based Approach for AI Derived Borehole Images Automatic Interpretation. OnePetro, October 2023. doi: 10.2118/216488-MS. URL https://dx.doi.org/10.2118/216488-MS.

Henry N. Pollack. Scientific uncertainty and public policy: Moving on without all the answers. *GSA Today*, 17(3):28, 2007. ISSN 1052-5173. doi: 10.1130/GSAT01703GW.1. URL http://www.geosociety.org/gsatoday/archive/17/3/pdf/i1052-5173-17-3-28.pdf.

Michael Poppelreiter. Borehole Image-Log Technology: Application Across the Exploration and Production Life Cycle. pages 1–13. October 2009. ISBN 978-0-89181-373-6.

Stephen E. Prensky. Advances in borehole imaging technology and applications. *Geological Society, London, Special Publications*, 159(1):1–43, January 1999. ISSN 0305-8719, 2041-4927. doi: 10.1144/GSL.SP.1999.159.01.01. URL https://www.lyellcollection.org/doi/10.1144/GSL.SP.1999.159.01.01.

R. Prioul, A. Donald, R. Koepsell, T. Bratton, C. Signer, A. Boyd, and A. Etchecopar. Fracture-induced Sonic Anisotropy Characterization Using a Combination of Borehole Image and Sonic Logs. page cp. European Association of Geoscientists & Engineers, September 2007. ISBN 978-90-73781-74-0. doi: 10.3997/2214-4609.20146715. URL https://www.earthdoc.org/content/papers/10.3997/2214-4609.20146715. ISSN: 2214-4609.

M. Pöppelreiter, C. García-Carballido, and M. Kraaijveld. Borehole Image Log Technology: Application Across the Exploration and Production Life Cycle. volume 92, pages 1–13. American Association of Petroleum Geologists, 2010a. ISBN 978-0-89181-373-6 978-1-62981-026-3. doi: 10.1306/13181274M923406. URL http://pubs.geoscienceworld.org/books/book/1263/chapter/107077960/Borehole-Image-Log-TechnologyApplication-Across. Book Title: Dipmeter and Borehole Image Log Technology.

Michael Pöppelreiter, García-Carballido Carmen, and Martin Kraaijveld. Borehole Image Log Technology: Application Across the Exploration and Production Life Cycle. In M. Pöppelreiter, C. García-Carballido, and M. Kraaijveld, editors, *Dipmeter and Borehole Image Log Technology*, volume 92, page 0. American Association of Petroleum Geologists, January 2010b. ISBN 978-1-62981-026-3. doi: 10.1306/13181274M923406. URL https://doi.org/10.1306/13181274M923406.

João Quintas. Improving Visual Perception of Artificial Social Companions Using a Standardized Knowledge Representation in a Human-Machine Interaction Framework. *International Journal of Social Robotics*, 15(3):425–444, 2023. URL https://idp.springer.com/authorize/casa?redirect_uri=https://link.springer.com/article/10.1007/s12369-021-00859-6&casa_token=tkrS_h9j0gQAAAAA:sYqneoHPJPi1fMXu3INCiMYdv2H5ROLfRE9pzS5KZAiUYumFDz_oI45l3EKIh_e_-z7f7szQT7pxxXo. Publisher: Springer.

Mojtaba Rajabi, Shahram Sherkati, Bahman Bohloli, and Mark Tingay. Subsurface fracture analysis and determination of in-situ stress direction using FMI logs: An example from the Santonian carbonates (Ilam Formation) in the Abadan Plain, Iran. *Tectonophysics*, 492(1):192–200, September 2010. ISSN 0040-1951. doi: 10.1016/j.tecto.2010.06.014. URL https://www.sciencedirect.com/science/article/pii/S0040195110002672.

Ritu Rani, Sandhya Pundhir, Amita Dev, and Arun Sharma. An Optimized Flower Categorization Using Customized Deep Learning. *International Journal of Computational Intelligence and Applications*, 21(04):2250029, December 2022. ISSN 1469-0268, 1757-5885. doi: 10.1142/S1469026822500298. URL https://www.worldscientific.com/doi/10.1142/S1469026822500298.

T.J. Reber, K.F. Beckers, and J.W. Tester. The transformative potential of geothermal heating in the U.S. energy market: A regional study of New York and Pennsylvania. *Energy Policy*, 70:30–44, 2014. doi: 10.1016/j.enpol.2014.03.004.

Ronald A. Rensink. The Dynamic Representation of Scenes. *Visual Cognition*, 7(1-3):17–42, January 2000. ISSN 1350-6285. doi: 10.1080/135062800394667. URL https://doi.org/10.1080/135062800394667. Publisher: Routledge _eprint: https://doi.org/10.1080/135062800394667.

Paolo Romagnoli, Alessia Arias, Antonio Barelli, Maurizio Cei, and Michele Casini. An updated numerical model of the Larderello–Travale geothermal system, Italy. *Geothermics*, 39(4):292–313, December 2010. ISSN 0375-6505. doi: 10.1016/j.geothermics.2010.09.010. URL https://www.sciencedirect.com/science/article/pii/S0375650510000507.

Giacomo Roncoroni. Long-Short-Term Memory in Active Wavefield Geophysical Methods. 2024. URL https://arts.units.it/handle/11368/3071821. Publisher: Università degli Studi di Trieste.

Giacomo Roncoroni, Emanuele Forte, Luca Bortolussi, and Michele Pipan. Efficient extraction of seismic reflection with Deep Learning. *Computers & Geosciences*, 166:105190, 2022. URL https://www.sciencedirect.com/science/article/pii/S0098300422001431?casa_token=eDxBMzGjiqYAAAAA:wqo1qSWs2kI6QYq92677zSCIc2u9hmk8NtNlHNR4KuP-Zq5q3iZRZhDJBF8tgg_sWjBuSo5z. Publisher: Elsevier.

Giacomo Roncoroni, Emanuele Forte, and Michele Pipan. Merging gated frequency-modulated continuous-wave Mars2020 RIMFAX ground-penetrating radar data. *GEOPHYSICS*, 88(2):A7–A12, March 2023. ISSN 0016-8033, 1942-2156. doi: 10.1190/geo2022-0466.1. URL https://library.seg.org/doi/10.1190/geo2022-0466.1.

Olaf Ronneberger, Philipp Fischer, and Thomas Brox. U-Net: Convolutional Networks for Biomedical Image Segmentation, May 2015. URL http://arxiv.org/abs/1505.04597. arXiv:1505.04597 [cs].

Rosanna Salone. Integration of Geophysical and Thermo-Fluid Dynamic Numerical Modelling to Study Natural Gas Emissions in Complex Geological Contexts. 2024. URL https://www.iris.unina.it/handle/11588/989223.

Andrea Saltelli, Marco Ratto, Terry Andres, Francesca Campolongo, Jessica Cariboni, Debora Gatelli, Michaela Saisana, and Stefano Tarantola. *Global Sensitivity Analysis: The Primer*. John Wiley & Sons, February 2008. ISBN 978-0-470-72517-7. Google-Books-ID: wAssmt2vumgC.

Alessandro Santilano, Adele Manzella, Giovanni Gianelli, Assunta Donato, Gola Gianluca, Isabella Nardini, Eugenio Trumpy, and Serena Botteghi. Convective, intrusive geothermal plays: what about tectonics? *Geothermal Energy Science*, 3:51–59, September 2015. doi: 10.5194/gtes-3-51-2015.

Sudeep Sarkar and Kim L. Boyer. On optimal infinite impulse response edge detection filters. *IEEE Transactions on Pattern Analysis & Machine Intelligence*, 13(11):1154–1171, 1991. URL https://www.computer.org/csdl/journal/tp/1991/11/i1154/13rRUwI5TYl. Publisher: IEEE Computer Society.

Thomas Scheiber, Ola Fredin, Giulio Viola, Alexandra Jarna, Deta Gasser, and Renata Łapińska Viola. Manual extraction of bedrock lineaments from high-resolution LiDAR data: methodological bias and human perception. *GFF*, 137(4):362–372, October 2015. ISSN 1103-5897. doi: 10.1080/11035897.2015.1085434. URL https://doi.org/10.1080/11035897.2015.1085434. Publisher: Taylor & Francis _eprint: https://doi.org/10.1080/11035897.2015.1085434.

Yu Shi, Xianzhi Song, and Guofeng Song. Productivity prediction of a multilateral-well geothermal system based on a long short-term memory and multi-layer perceptron combinational neural network. *Applied Energy*, 282:116046, January 2021. ISSN 0306-2619. doi: 10.1016/j.apenergy.2020.116046. URL https://www.sciencedirect.com/science/article/pii/S0306261920314811.

Yunzhi Shi, Xinming Wu, and Sergey Fomel. Waveform embedding: Automatic horizon picking with unsupervised deep learning. *GEOPHYSICS*, 85(4):WA67–WA76, July 2020. ISSN 0016-8033, 1942-2156. doi: 10.1190/geo2019-0438.1. URL https://library.seg.org/doi/10.1190/geo2019-0438.1.

Thomas F. Shipley, Basil Tikoff, Carol Ormand, and Cathy Manduca. Structural geology practice and learning, from the perspective of cognitive science. *Journal of Structural Geology*, 54:72–84, September 2013. ISSN 0191-8141. doi: 10.1016/j.jsg.2013.07.005. URL https://www.sciencedirect.com/science/article/pii/S0191814113001223.

Z. K. Shipton, J. J. Roberts, E. L. Comrie, Y. Kremer, R. J. Lunn, and J. S. Caine. Fault fictions : systematic biases in the conceptualization of fault zones. URL https://core.ac.uk/outputs/195294655/?utm_source=pdf&utm_medium=banner&utm_campaign=pdf-decoration-v1.

Z. K. Shipton, J. J. Roberts, E. L. Comrie, Y. Kremer, R. J. Lunn, and J. S. Caine. Fault fictions: systematic biases in the conceptualization of fault-zone architecture. *Geological Society, London, Special Publications*, 496(1):125–143, January 2020. doi: 10.1144/SP496-2018-161. URL https://www.lyellcollection.org/doi/abs/10.1144/sp496-2018-161. Publisher: The Geological Society of London.

G. T. Shrivakshan and Chandramouli Chandrasekar. A comparison of various edge detection techniques used in image processing. *International Journal of Computer Science Issues (IJCSI)*, 9(5):269, 2012. URL https://citeseerx.ist.psu.edu/document?repid=rep1&type=pdf&doi=685e8ecc440343d59f157ca1add377b956b2199d. Publisher: Citeseer.

Peter Skagestad. Thinking with machines: Intelligence augmentation, evolutionary epistemology, and semiotic. *Journal of Social and Evolutionary Systems*, 16(2):157–180, January 1993. ISSN 1061-7361. doi: 10.1016/1061-7361(93)90026-N. URL https://www.sciencedirect.com/science/article/pii/106173619390026N.

Irwin Sobel. An Isotropic 3x3 Image Gradient Operator. *Presentation at Stanford A.I. Project 1968*, February 2014.

Petru Soviany, Radu Tudor Ionescu, Paolo Rota, and Nicu Sebe. Curriculum Learning: A Survey, April 2022. URL http://arxiv.org/abs/2101.10382. arXiv:2101.10382 [cs].

D. Sui, E. Wiktorski, M. Røksland, and T. A. Basmoen. Review and investigations on geothermal energy extraction from abandoned petroleum wells. *Journal of Petroleum Exploration and Production Technology*, 9(2):1135–1148, 2019. URL https://go.gale.com/ps/i.do?id=GALE%7CA583531064&sid=googleScholar&v=2.1&it=r&linkaccess=abs&issn=21900558&p=AONE&sw=w. Publisher: Springer.

E. Sulistyowati and A. Haris. Integration of Borehole Image and Sonic to Evaluate Critically-Stressed Fractures to Optimize Production at FORGE Geothermal Field. In *Journal of Physics: Conference Series*, volume 2019, page 012085. IOP Publishing, 2021. URL https://iopscience.iop.org/article/10.1088/1742-6596/2019/1/012085/meta. Issue: 1.

Qifeng Sun, Na Li, Youxiang Duan, Hongqiang Li, and Haiquan Tang. Logging-while-drilling formation dip interpretation based on long short-term memory. *Petroleum Exploration and Development*, 48(4):978–986, August 2021. ISSN 1876-3804. doi: 10.1016/S1876-3804(21)60082-4. URL https://www.sciencedirect.com/science/article/pii/S1876380421600824.

Ruoyu Sun. Optimization for deep learning: theory and algorithms, December 2019. URL http://arxiv.org/abs/1912.08957. arXiv:1912.08957 [cs].

Christof Tannert, Horst-Dietrich Elvers, and Burkhard Jandrig. The ethics of uncertainty. *EMBO reports*, 8(10):892–896, October 2007. ISSN 1469-221X. doi: 10.1038/sj.embor.7401072. URL https://www.embopress.org/doi/full/10.1038/sj.embor.7401072. Num Pages: 896 Publisher: John Wiley & Sons, Ltd.

Ruth D. Terzaghi. Sources of Error in Joint Surveys. *Géotechnique*, 15(3): 287–304, September 1965. ISSN 0016-8505. doi: 10.1680/geot.1965.15.3.287. URL https://www.icevirtuallibrary.com/doi/10.1680/geot.1965.15.3.287. Publisher: ICE Publishing.

Bhaskar B. Thapa, Paul Hughett, and Kenzi Karasaki. Semi-automatic analysis of rock fracture orientations from borehole wall images. *GEOPHYSICS*, 62(1): 129–137, January 1997. ISSN 0016-8033, 1942-2156. doi: 10.1190/1.1444112. URL https://library.seg.org/doi/10.1190/1.1444112.

Michael Thompson, Richard O. Duda, and Peter E. Hart. Pattern Classification and Scene Analysis. In *Leonardo*, volume 7, page 370, 1974. doi: 10.2307/1573081. URL https://www.jstor.org/stable/1573081?origin=crossref. ISSN: 0024094X Issue: 4 Journal Abbreviation: Leonardo.

Juanita Todd, Erich Schröger, and István Winkler. Introductory notes on "Predictive information processing in the brain: Principles, neural mechanisms, and models". *International Journal of Psychophysiology*, 83(2):119, February 2012. ISSN 0167-8760. doi: 10.1016/j.ijpsycho.2012.01.015. URL https://www.sciencedirect.com/science/article/pii/S0167876012000177.

Vincent Torre and Tomaso A. Poggio. On edge detection. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, (2): 147–163, 1986. URL https://ieeexplore.ieee.org/abstract/document/4767769/?casa_token=g0wUVX7EFHMAAAAA:3xlIo9R7Owpo_w_faMH8p7k-j3cQVV43DDs3tBL6yqIKW1qoIxRRSx10IVaiA4gJgvwvpqQY. Publisher: IEEE.

Eugenio Trumpy and Adele Manzella. Geothopica and the interactive analysis and visualization of the updated Italian National Geothermal Database. *International Journal of Applied Earth Observation and Geoinformation*, 54:28–37, February 2017. ISSN 1569-8432. doi: 10.1016/j.jag.2016.09.004. URL https://www.sciencedirect.com/science/article/pii/S0303243416301556.

P. Ungemach, M. Antics, and M. Davaux. An innovative Geothermal Well Architecture and Reservoir Evaluation Concept - the Cachan Subhorizontal Doublet. page cp. European Association of Geoscientists & Engineers, November 2018. ISBN 978-94-6282-269-6. doi: 10.3997/2214-4609.201802929. URL https://www.earthdoc.org/content/papers/10.3997/2214-4609.201802929. ISSN: 2214-4609.

Pierre Ungemach, Miklos Antics, Davide Di Tommaso, and Filippo Casali. Real Time Geosteering Integrated Services. A Key Issue in Maximizing Geothermal Exposure and Minimizing Drilling and Completion Risks. A Paris Basin Case Study. OnePetro, March 2021. doi: 10.2118/204012-MS. URL https://dx.doi.org/10.2118/204012-MS.

M. van Ginkel, M. A. Kraaijveld, L. J. van Vliet, E. P. Reding, P. W. Verbeek, and H. J. Lammers. Robust Curve Detection Using a Radon Transform in Orientation Space. In Josef Bigun and Tomas Gustavsson, editors, *Image Analysis*, pages 125–132, Berlin, Heidelberg, 2003. Springer. ISBN 978-3-540-45103-7. doi: 10.1007/3-540-45103-X_18.

Haijing Wang, Ed Stockhausen, Dennis F. Wyatt, Jr., and Dean Gulick. Modeling of Azimuthal Gamma-Ray Tools for Use in Geosteering in Unconventional Reservoirs. *Petrophysics - The SPWLA Journal of Formation Evaluation and Reservoir Description*, 60(01):93–112, February 2019. ISSN 1529-9074. doi: 10.30632/PJV60N1-2019a8. URL https://doi.org/10.30632/PJV60N1-2019a8.

Nanzhe Wang, Haibin Chang, Xiangzhao Kong, Martin O. Saar, and Dongxiao Zhang. Deep learning based closed-loop optimization of geothermal reservoir production, April 2022a. URL http://arxiv.org/abs/2204.08987. arXiv:2204.08987 [cs, eess].

Weixing Wang, Haijun Liao, and Ying Huang. Rock fracture tracing based on image processing and SVM. In *Third International Conference on Natural Computation (ICNC 2007)*, volume 1, pages 632–635, August 2007. doi: 10.1109/ICNC.2007.643. URL https://ieeexplore.ieee.org/abstract/document/4344267?casa_token=IwqNd0TDpNMAAAAA:eSZHBQRB2HItGSS_srViLinDXPm9dYrwDqC_NJsnvvL1a_oURyCSfkrEnpdUQ-UIZt1EHGA. ISSN: 2157-9563.

Xin Wang, Yudong Chen, and Wenwu Zhu. A Survey on Curriculum Learning. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 44(9):4555–4576, September 2022b. ISSN 1939-3539. doi: 10.1109/TPAMI.2021.3069908. URL https://ieeexplore.ieee.org/abstract/document/9392296?casa_token=hdqjWKcGMAcAAAAA:2TfHaozQ46TqjaJ5223ojPpzmxpcObYJK6nuIODMX5tcRKxqeGFrSmw9y9V95emhYF_-HC4. Conference Name: IEEE Transactions on Pattern Analysis and Machine Intelligence.

Boyang Wei, Baozhi Pan, Qiuli Yin, Chaoguo Tian, and Xiaoya Guo. Identification of image logging data based on conditional generation adversarial network. *Geophysical Prospecting for Petroleum*, 59(2):295–302, April 2022. ISSN 1000-1441. doi: 10.3969/j.issn.1000-1441.2020.02.016. URL https://www.sciengine.com/GPP/doi/10.3969/j.issn.1000-1441.2020.02.016. Publisher: Beijing Zhongke Journal Publising Co. Ltd.

Xin Wei, Lulu Zhang, Hao-Qing Yang, Limin Zhang, and Yang-Ping Yao. Machine learning for pore-water pressure time-series prediction: Application of recurrent neural networks. *Geoscience Frontiers*, 12(1):453–467, January 2021. ISSN 1674-9871. doi: 10.1016/j.gsf.2020.04.011. URL https://www.sciencedirect.com/science/article/pii/S1674987120301134.

P.J. Werbos. Backpropagation through time: what it does and how to do it. *Proceedings of the IEEE*, 78(10):1550–1560, October 1990. ISSN 1558-2256. doi: 10.1109/5.58337. URL https://ieeexplore.ieee.org/document/58337. Conference Name: Proceedings of the IEEE.

Erik Wielemaker, Chiara Cavalleri, Leon Dahlhaus, Alejandra Reynaldos, Giovanni Sosio, Pierre Ungemach, Miklos Antics, and Melanie Davaux. Delineating the Geothermal Structure and Flow Properties in a Sub-Horizontal Well with the use of Wireline and LWD Data in a Multiphysics Approach. OnePetro, June 2020. doi: 10.30632/SPWLA-5065. URL https://dx.doi.org/10.30632/SPWLA-5065.

Cristina G. Wilson, Clare E. Bond, and Thomas F. Shipley. How can geologic decision-making under uncertainty be improved? *Solid Earth*, 10(5):1469–1488, September 2019. ISSN 1869-9510. doi: 10.5194/se-10-1469-2019. URL https://se.copernicus.org/articles/10/1469/2019/. Publisher: Copernicus GmbH.

Cristina G. Wilson, Thomas F. Shipley, and Alexandra K. Davatzes. Evidence of vulnerability to decision bias in expert field scientists. *Applied Cognitive Psychology*, 34(5):1217–1223, 2020. ISSN 1099-0720. doi: 10.1002/acp.3677. URL https://onlinelibrary.wiley.com/doi/abs/10.1002/acp.3677. _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1002/acp.3677.

Jeffrey B. Witter, Whitney J. Trainor-Guitton, and Drew L. Siler. Uncertainty and risk evaluation during the exploration stage of geothermal development: A review. *Geothermics*, 78:233–242, March 2019. ISSN 0375-6505. doi: 10.1016/j.geothermics.2018.12.011. URL https://www.sciencedirect.com/science/article/pii/S0375650518303183.

Jingye Yang, Guozheng Kang, and Qianhua Kan. A novel deep learning approach of multiaxial fatigue life-prediction with a self-attention mechanism characterizing the effects of loading history and varying temperature. *International Journal of Fatigue*, 162:106851, 2022. URL https://www.sciencedirect.com/science/article/pii/S0142112322001268?casa_token=wyUPL4ACQi0AAAAA:UbOMn3wwVOQ2_

nRihx_W7jviVKsT8fkb845qcKEPsMuBcJgMbFbdOgO61OqRZS1xhbeea7FU. Publisher: Elsevier.

Yong Yang, Ming Zhang, Aifang Bie, Jietang Lv, Wenqi Zhang, Zehong Cui, and Zhaohui Xia. A Case Study of Subsurface Uncertainty Analysis in Modelling Carbonate Reservoir. In Jia'en Lin, editor, *Proceedings of the International Field Exploration and Development Conference 2018*, pages 851–864, Singapore, 2020. Springer. ISBN 9789811371271. doi: 10.1007/978-981-13-7127-1_79.

Jeoung Seok Yoon, Arno Zang, and Ove Stephansson. Numerical investigation on optimized stimulation of intact and naturally fractured deep geothermal reservoirs using hydro-mechanical coupled discrete particles joints model. *Geothermics*, 52:165–184, 2014. URL https://www.sciencedirect.com/science/article/pii/S0375650514000108. Publisher: Elsevier.

Wanju Yuan, Zhuoheng Chen, Stephen E. Grasby, and Edward Little. Closed-loop geothermal energy recovery from deep high enthalpy systems. *Renewable Energy*, 177(C):976–991, 2021. ISSN 0960-1481. URL https://econpapers.repec.org/article/eeerenene/v_3a177_3ay_3a2021_3ai_3ac_3ap_3a976-991.htm. Publisher: Elsevier.

Sadiq J. Zarrouk and Hyungsul Moon. Efficiency of geothermal power plants: A worldwide review. *Geothermics*, 51:142–153, July 2014. ISSN 0375-6505. doi: 10.1016/j.geothermics.2013.11.001. URL https://www.sciencedirect.com/science/article/pii/S0375650513001120.

Hanqing Zeng, Chi Zhang, and Viktor Prasanna. Fast generation of high throughput customized deep learning accelerators on FPGAs. In *2017 International Conference on ReConFigurable Computing and FPGAs (ReConFig)*, pages 1–8. IEEE, 2017. URL https://ieeexplore.ieee.org/abstract/document/8279792/.

Hao Zhang and Peimin Zhu. 3D seismic horizons extraction based on deep learning. In *SEG International Exposition and Annual Meeting*, page D011S068R005. SEG, 2021. URL https://onepetro.org/SEGAM/proceedings-abstract/IMAGE21/1-IMAGE21/471634.

Xiang Zhang and Xiaoling Xiao. Detection of fractures in borehole image. In *MIPPR 2009: Automatic Target Recognition and Image Analysis*, volume 7495, pages 1043–1048. SPIE, October 2009. doi: 10.1117/12.832608. URL https://www.spiedigitallibrary.org/conference-proceedings-of-spie/7495/749541/Detection-of-fractures-in-borehole-image/10.1117/12.832608.full.

Lina Zhou, Souren Paul, Haluk Demirkan, Lingyao Yuan, Jim Spohrer, Michelle Zhou, and Julie Basu. Intelligence Augmentation: Towards Building Human-Machine Symbiotic Relationship. *AIS Transactions on Human-Computer Interaction*, 13(2):243–264, June 2021. ISSN 1944-3900. doi: 10.17705/1thci.00149. URL https://aisel.aisnet.org/thci/vol13/iss2/5.

Chapter 5

Acknowledgments

First and foremost, I would like to express my deepest gratitude to my supervisor, Michele Pipan, for his invaluable guidance throughout my research journey. His encouragement and ability to create opportunities for travel and research have greatly enriched this work and my academic experience.
I am sincerely thankful to Giacomo Roncoroni, whose collaboration and intellectual exchange have been both inspiring and instrumental in shaping this thesis. His perspective as an esteemed colleague has been of immense value.
A special thanks to Davide Baldini from ENI, whose innovative ideas served as a foundation for some of the most critical aspects of this research. His influence on this work cannot be overstated.
To Gianluca Gola, I am profoundly grateful for his unwavering patience and steadfast support. His passion for seeking innovative and practical solutions in the geothermal field has been a constant source of inspiration. His dedication to increasing green energy's contribution to our country is deeply admirable.
I also extend my heartfelt appreciation to Emilio Cecchetti. His early interest in my work was a turning point, sparking a sequence of events that highlighted the potential and innovation of the proposed methods, when I was lost in discouragement. His efforts in acquiring the data for testing the scalability of these methods, particularly for the Netherlands case study, have laid the groundwork for future research endeavors.