POLITECNICO DI TORINO

Master's Degree in Computer Engineering

Master's Degree Thesis

Epicenter Detection integrating Sentinel-1 InSAR data and Deep Learning

Supervisors

Prof. Paolo GARZA

Dott. Daniele REGE CAMBRIN

Candidate

Elena BUCCOLIERO

December 2024

Abstract

This master thesis explores a new approach to improving seismic monitoring, which is often characterized by poor efficiency and accuracy based on current methods. Traditional approaches to generate interferograms from satellite radar data can take a long time and human analysis. In addition, the complexity of SAR image processing makes it difficult to obtain timely information as a disaster response.

The automation of this process is due to the need for faster and more accurate seismic assessments, and the desire to eliminate delays caused by manual data interpretation. This could reduce the possibility of human error and improve the overall reliability of seismic risk assessment. The aim is to incorporate InSAR data and machine learning technologies into a system that functions efficiently and minimizes the human intervention.

The study highlights the potential of integrating remote sensing and deep learning models to optimise earthquake monitoring, in order to achieve faster and more accurate information for disaster preparedness and response. This research aims to lay the foundations for future improvements that will allow real-time seismic monitoring.

A chi c'è stato, a chi c'è e a chi ci sarà.

Table of Contents

List of Tables

List of Figures

Acronyms

SAR
Synthetic Aperture Radar

SLC
Single Look Complex

ESA
European Space Agency

CNN
Convolutional Neural Network

ASF
Alaska Satellite Facility

Chapter 1

Introduction

Earthquake shocks, caused by sudden changes in the earth's crust due to tectonic shifts, have been a persistent natural hazard throughout human history. Their unpredictable nature has often led to catastrophic consequences, causing severe damage to infrastructure, significant economic losses and loss of life. As urban populations continue to grow and infrastructure becomes increasingly complex, seismic monitoring and risk mitigation strategies need to be introduced. To meet this need, advances have been made in technology and research to improve our quality of life in relation to catastrophic events, improving early warning systems and developing more resilient infrastructures.

In the field of seismic monitoring, one of the most promising developments is the use of remote sensing technologies. These technologies, in particular Synthetic Aperture Radar (SAR) and its interferometric applications (InSAR), have revolutionised the way scientists observe and analyse soil deformation associated with seismic activity or even external changes. These systems emit radar signals, measure the waves reflected from the earth's surface and are capable of operating in a variety of weather conditions, both day and night. SAR satellite imagery is a valuable and innovative tool for near real-time field analysis and continuous and reliable monitoring of soil movements.

The Sentinel-1 mission, part of the European Space Agency's (ESA) Copernicus programme, stands out as a key resource in the field of SAR-based earth observation. The mission consists of a constellation of two satellites, Sentinel-1A and Sentinel-1B, which provide high resolution radar and broadband images of the earth's surface. These satellites are equipped with SAR sensors, which can capture different types of images at various levels. SAR products retain both the amplitude and phase information of the radar signal, making them particularly suitable for interferometric analysis. Comparing the phase differences between two images taken at different times, before and after the earthquake, InSAR techniques detect fine ground deformations with high precision. They are preserved and highlighted

by interferograms, which are crucial for assessing the impact of seismic events and essential for training neural networks to perform regression tasks.

Machine learning, especially deep learning techniques such as convolutional neural networks (CNN) and visual transformers (VIT), has proven to be a promising solution for addressing the problem of locating the epicenter through an automated system.

CNN is a class of deep neural networks that have demonstrated remarkable performance in the recognition, classification and extraction of image characteristics. Their architecture, inspired by the visual cortex of the human brain, enables efficient machine learning of spatial hierarchies of features from incoming images. This capability makes CNNs ideal for analyzing complex data and adaptable to identify terrain deformations with precision.

On the other hand, Vision Transformers (ViTs) represent a new approach to deep learning, exploiting transformer-based architectures originally developed for natural language processing. Unlike CNNs, which rely on convolutions to extract local features, ViTs were particularly effective in this study, splitting images into chunks and processing them as sequences, learning global relationships between characteristics using self-care mechanisms.

This capability makes them a powerful tool for tasks requiring long-range dependency capture, such as identifying phase variations in interferograms. The complementary strengths of CNNs and VITs make them robust and concrete tools for advancing automated seismic monitoring.

The main objective of this thesis is to develop a new framework that leverages Sentinel-1A SAR products and deep learning models to automate the process of monitoring earthquakes and localizing epicenters. The research involves, as a first step, collecting data on earthquakes which have already occurred and of which the epicenter is known, and then generating high quality interferograms from the Sentinel-1 data and, finally, the training models to identify the location and predict the epicenter of seismic events.

The analysis begins with a review of known earthquakes, of magnitude not less than *4 mb*, collecting phenomena with their related spatial and temporal information.

Through a search for SAR products covering the area of interest by interrogating the Alaska Satellite Facility (ASF), seismic events were selected for which the images found cover the area near the epicenter, in a useful time frame and not too far from the event detection, collecting data before and after the earthquake occurred.

After obtaining the necessary data, interferograms were generated, from which three interferograms per earthquake, one for each subswath: IW1, IW2, IW3, were obtained. This step was computationally intensive because of the size and complexity of the SAR products. The process was initially tested on SNAP software

and then the workflow was refunded and executed as a graph using the snappy and snapista libraries of Python.

Although subdivided into subswath, the sizes of interferograms obtained were not entirely indifferent so it was strongly necessary to apply a resize to each of them. The applied scaling has collected 256x256 pixels patches until reaching about thirty thousand final samples.

The final dataset containing information on the bands and metadata of the cut-out patches has been split into train, test, val data and then the deep learning models ResNet, ConvNeXt and ViT were introduced.

As a final step, training testing and validation process were run to tune the best hyperparameters and minimize errors. The evaluation metrics allowed to estimate the accuracy of each model and make appropriate comparisons to indicate which deep learning model best suited to predict information from the epicenter and achieved the best performance.

The code developed for this study is available in the GitHub repository at : `https://github.com/bruco27/Epicenter_Detection.git`.
Due to space limitations, only the essential files required to perform the analysis are included in the repository, additional files are stored on the virtual machines *mp1* and *hpc* of Politecnico di Torino and can be accessed upon request.

Chapter 2

Related Works

This research is inspired by a series of studies that have used satellite remote sensing and machine learning techniques to improve earthquake monitoring and analysis. The integration of Interferometric Synthetic Aperture Radar (InSAR) technology with deep learning frameworks has opened new avenues for seismic event detection, analysis, and prediction. This chapter shows all the related works, like Quakeset [1], PhaseNet [2] or ConvNetQuake [3], which have helped to show useful information in the field of geophysics and other fields.

2.1 SAR and InSAR

Synthetic Aperture Radar (SAR) technology, as utilized in the Sentinel-1 mission, has become a cornerstone for monitoring ground deformation in seismic research. Unlike traditional optical imagery, SAR operates in the C band, allowing for day-and-night imaging through clouds and atmospheric disturbances, making it particularly suited for capturing geophysical changes in inaccessible or remote areas [1]. Sentinel-1's Interferometric Wide (IW) mode strikes a critical balance between spatial resolution and area coverage, ensuring that deformation patterns are captured over broad regions with high fidelity. This capability is further enhanced by SAR interferometry (InSAR), which enables precise mapping of ground motion by analyzing the phase differences between SAR images taken at different times.

InSAR principles, as outlined by Ferretti et al. [4], involve the processing of SAR images to extract phase information, which represents the distance between the radar sensor and the ground surface. By comparing the phase values of two or more SAR acquisitions, InSAR measures subtle changes in surface displacement with millimeter-level precision. This technique relies on the coherence of the SAR signal, which can be influenced by factors such as vegetation cover, atmospheric conditions, and temporal spacing between acquisitions. To mitigate these variations, advanced

processing techniques such as atmospheric correction, phase unwrapping, and multi-temporal analysis are employed to ensure accurate displacement measurements.

The QuakeSet project leverages Sentinel-1 SAR data to create a detailed framework for seismic monitoring. The study focuses on Level-1 Ground Range Detected (GRD) products, processed in both VV and VH polarizations. High-resolution samples are prioritized, with SAR backscatter calibrated using the sigma ellipsoid coefficient to ensure consistent and reliable measurements. These preprocessing steps are crucial for maintaining the quality and usability of the data across various geophysical tasks.

A defining feature of QuakeSet is its tri-temporal dataset design, incorporating pre-event, post-event, and neutral images for each of the analyzed earthquakes. The temporal window spans 13 days before and after the seismic event, with careful exclusion of the day of the incident to avoid ambiguities. This tri-temporal approach allows for robust temporal analysis, enabling the detection of subtle ground deformation patterns that might otherwise go unnoticed.

InSAR's ability to measure ground displacement provides a significant advantage over traditional seismic stations, particularly in regions with sparse instrumentation. By capturing the relative phase shifts between SAR images, InSAR facilitates the detection and quantification of surface changes caused by seismic activity. This capability is indispensable for understanding the spatial extent and magnitude of earthquakes, especially in areas where ground-based measurements are not feasible. The inclusion of deformation modeling in InSAR analysis, as described by Ferretti et al., enables the interpretation of ground displacement in terms of underlying geophysical processes such as fault slip and volcanic deformation [4].

QuakeSet also explores the potential for deploying low-resource machine learning models alongside SAR data to enhance its analytical capabilities. Onboard satellite processing using these models can enable real-time detection of affected areas, minimizing data transmission requirements and accelerating disaster response efforts [1]. Tasks such as change detection, magnitude regression, and epicenter localization are supported by this integration of SAR with advanced computational techniques.

The project underscores the transformative role of SAR and InSAR in seismic monitoring, demonstrating their ability to complement or replace traditional seismic networks in certain scenarios. The QuakeSet dataset, paired with detailed preprocessing guidelines and open-source code, provides an invaluable resource for the scientific community. By advancing the use of SAR in geophysics, QuakeSet sets a foundation for scalable and automated solutions in disaster management and seismic analysis.

2.2 Deep Learning and Computer Vision

Deep learning has revolutionized the field of data analysis by enabling the extraction of complex and robust features from diverse datasets. In seismic applications, models like PhaseNet demonstrate the transformative potential of deep learning. Based on a U-Net architecture, PhaseNet automates the picking of seismic arrival times for P and S waves by learning directly from three-component waveform data [2]. This data-driven approach significantly improves upon traditional methods, which rely on manually defined features and are susceptible to variability and noise. By processing over seven million waveform samples spanning 30 years, PhaseNet achieves high F1 scores for P (0.896) and S (0.801) waves, demonstrating remarkable robustness, even in low signal-to-noise ratio (SNR) environments.

The architecture of PhaseNet introduces probabilistic output representations for arrival times, enabling the model to handle noise and uncertainty effectively. This capability is able to cope with the long-standing phenomena in seismic phase picking, such as the accurate detection of S waves, which are often masked by P wave coda interference. The success of PhaseNet highlights the potential of deep learning models to enhance both the precision and scalability of seismic data processing, making it a pivotal tool for modern seismic networks where manual analysis is no longer feasible due to the sheer volume of data [2].

In satellite imagery analysis, deep learning has similarly driven significant advancements. Time-series satellite imagery (SITS) has become a key resource for Earth observation, supporting tasks like land cover classification and vegetation analysis. Models combining convolutional neural networks (CNNs) with recurrent neural networks (RNNs) are particularly effective for processing SITS, as they integrate spatial and temporal data seamlessly. Hybrid architectures demonstrate further adaptability of deep learning to complex environmental data sets even in the absence of data or with a small number of samples tagged [5].

In the domain of earthquake damage detection, deep learning models offer transformative potential. The QuickQuakeBuildings project introduces a dataset that integrates post-event SAR and optical imagery to detect damaged buildings [6]. Deep learning models trained on this dataset leverage SAR's resilience to adverse weather conditions and optical imagery's high interpretability to achieve strong performance, even in urban areas where there are geometric distortions and occlusions. Late fusion architectures, which combine SAR and optical modalities, enhance detection accuracy by integrating complementary information from both data sources. These advancements underscore the versatility of deep learning in handling diverse data modalities for disaster response.

Deep learning and computer vision thus emerge as powerful tools for tackling complex problems in both seismology and satellite remote sensing. By combining

the strengths of advanced architectures with vast datasets, these approaches enable scalable, automated solutions that pave the way for real-time environmental monitoring and disaster management.

2.3 Computer Vision for SAR

Computer vision has become a pivotal tool for processing and analyzing Synthetic Aperture Radar (SAR) imagery, enabling robust solutions to complex issues such as change detection, damage assessment, and environmental monitoring. Unlike traditional image modalities, SAR imagery provides unique benefits, including all-weather, day-and-night imaging capabilities. These features, however, also introduce complexities such as speckle noise and geometric distortions, requiring tailored computer vision approaches for effective analysis.

Among the most widely adopted architectures in SAR-based computer vision tasks are convolutional neural networks (CNNs) and their variants. The U-Net architecture, originally designed for biomedical image segmentation, has been adapted extensively for SAR change detection. Corley et al. revisited the performance of U-Net and its enhanced variants, emphasizing that many reported improvements stem from differences in training practices rather than intrinsic architectural advancements [7]. Enhanced U-Net models with shared encoders have shown particular promise in bi-temporal change detection, highlighting the importance of reproducible benchmarks and standardized evaluations in the field.

In multimodal analysis, SAR imagery is often combined with optical data to overcome individual limitations. For example, the QuickQuakeBuildings dataset leverages both SAR and optical imagery to detect earthquake-induced structural damage [6]. While SAR provides resilience to adverse weather conditions, optical data offers high interpretability, and their late fusion allows models to capitalize on the strengths of both modalities. This multimodal approach has demonstrated improved accuracy in damage detection, even in complex urban environments.

Other applications of computer vision in SAR include seismic activity classification. ConvNetQuake exemplifies the use of CNNs to process SAR data for identifying seismic events [3]. By handling high-dimensional inputs and employing regularized loss functions, ConvNetQuake achieves robust performance in noisy conditions, showcasing the adaptability of CNNs for SAR use.

As SAR imaging systems continue to advance, the role of computer vision is expanding to address diverse geospatial and environmental tasks. Techniques such as data augmentation, domain adaptation, and hybrid model architectures are increasingly employed to enhance the scalability and efficiency of SAR data analysis. These innovations not only improve the accuracy of SAR-based applications but also enable real-time processing for disaster response and monitoring, underscoring

the integral role of computer vision in unlocking the full potential of SAR imagery.

2.4 Earthquakes

Earthquakes are among the most destructive natural phenomena, caused by the sudden release of energy along faults in the Earth's crust. This energy propagates as seismic waves, which are classified as body waves (P-waves and S-waves) and surface waves. Understanding and monitoring earthquakes is crucial for mitigating their impact on infrastructure, populations, and economies. Over the years, advancements in technology and computational methods have significantly improved our ability to detect, locate, and model seismic events.

Seismic monitoring relies on networks of ground-based sensors, known as seismic stations, to record vibrations in the Earth's crust. However, these traditional networks face limitations in remote or sparsely instrumented regions. Innovative approaches combining advanced computational and algorithmic methods have been developed to address these issues. The International Seismological Centre (ISC) has contributed significantly to this field by improving earthquake localization methods. The ISC's enhanced localization algorithm addresses issues such as signal overlap and uneven data quality [8].

Key features of this algorithm include the incorporation of fault structures and a neighborhood algorithm (NA) for refining hypocenter determination. The ISC also integrates later seismic phases and depth-phase stacking to achieve more accurate depth resolutions. These improvements have led to significant advancements in the clustering of seismic event locations, providing a clearer and more reliable representation of global seismicity. Validation tests have shown reduced uncertainties in hypocentral positions, enabling more precise identification of earthquake source mechanisms.

Beyond traditional approaches, computational innovations such as Physics-Informed Neural Networks (PINNs) have revolutionized seismic wave modeling. SeismicNet is an example of leveraging PINNs for efficient simulation of seismic wave propagation in semi-infinite domains [9]. Unlike conventional numerical methods that require extensive labeled data and computational resources, SeismicNet integrates governing physical equations directly into the neural network architecture. This approach allows it to operate effectively even with sparse or noisy observational data.

SeismicNet employs absorbing boundary conditions (ABCs) to handle truncated computational domains, ensuring that wave reflections at the boundaries do not interfere with the simulation. Additionally, it uses domain decomposition strategies, dividing the problem into smaller subdomains for improved scalability and computational efficiency. Numerical experiments demonstrate the model's versatility

across various material distributions and loading conditions, making it a promising tool for forward modeling of seismic wave behavior. The ability to model seismic waves with high accuracy in real-world scenarios highlights the potential of PINNs in advancing earthquake research and engineering applications.

The integration of advanced algorithms and machine learning techniques with traditional seismology represents a paradigm shift in earthquake monitoring and analysis. These tools enable researchers to address longstanding detecting small-magnitude events, improving localization in challenging environments and modeling complex wave interactions. As computational capabilities continue to grow, future advancements are expected to further refine our understanding of earthquakes, enhance early warning systems, and support disaster preparedness efforts.

From improving earthquake catalogs to simulating wave propagation in complex geological settings, the synergy between computational methodologies and seismology exemplifies the transformative impact of interdisciplinary approaches. These innovations pave the way for more robust and efficient seismic monitoring frameworks, offering a clearer understanding of Earth's dynamic processes and helping to mitigate the risks associated with seismic hazards.

2.5 Earthquake Analysis with Machine Learning

The integration of machine learning into earthquake analysis has revolutionized the field, providing innovative solutions to address the inherent complexities of seismic data processing. Machine learning algorithms excel in handling vast volumes of data, uncovering subtle patterns, and extracting actionable insights from noisy and high-dimensional datasets. These capabilities make them indispensable for tasks such as earthquake detection, phase picking, magnitude estimation, and even forecasting seismic hazards.

One of the main themes of seismology is the detection of low magnitude seismic events, which are often difficult to detect because they are buried in background noise. Traditional approaches, relying on manually engineered features and heuristic thresholds, struggle to achieve robustness and scalability in such scenarios. Deep learning techniques, such as convolutional neural networks (CNNs), have emerged as transformative solutions by automatically learning hierarchical representations directly from raw seismic data. For instance, models like ConvNetQuake [3] demonstrate superior sensitivity and precision, significantly enhancing the detection of minor seismic events. This capability is crucial for improving our understanding of earthquake distributions and identifying precursors to larger seismic events.

In addition to event detection, deep learning models have made groundbreaking advancements in phase picking, the process of identifying seismic wave arrival times. Hybrid architectures, such as EQTransformer, combine convolutional layers,

long short-term memory (LSTM) networks, and attention mechanisms to capture both spatial and temporal features of seismic waveforms [10]. These models excel in noisy environments, achieving unprecedented accuracy in identifying P and S wave phases. The versatility of EQTransformer allows it to be used in real-time applications and retrospective studies, making it a cornerstone in modern seismic analysis.

Expanding on these advancements, transfer learning techniques have enabled the adaptation of models to specific seismic environments, overcoming the scarcity of labeled data. For example, OBSTransformer extends EQTransformer to Ocean Bottom Seismometer (OBS) data by incorporating domain-specific noise reduction strategies and adapting to the unique characteristics of offshore seismic monitoring [10]. This approach enhances phase-picking accuracy in challenging scenarios, such as low signal-to-noise ratios or distant seismic events.

Deep learning has also played a pivotal role in earthquake characterization and damage assessment. By integrating multimodal data—including seismic, optical, and SAR imagery—models trained on deep architectures provide a holistic view of seismic impacts. For example, change detection techniques using U-Nets and Siamese networks accurately identify structural damage and landscape changes following earthquakes [7]. This capability not only accelerates disaster response but also provides critical data for post-event reconstruction and planning.

Beyond detection and damage assessment, machine learning is increasingly applied to seismic hazard prediction and risk assessment. Predictive models, leveraging geological, geophysical, and historical data, estimate earthquake probabilities and potential impacts with high precision. Advanced architectures like graph neural networks (GNNs) and transformer models have shown promise in integrating spatial and temporal data to improve predictions. These efforts are vital for strengthening early warning systems and informing disaster preparedness strategies.

Deep learning also facilitates the analysis of complex geophysical phenomena, such as aftershock sequences, fault slip behavior, and ground shaking patterns. Recurrent neural networks (RNNs) and transformers, capable of processing sequential data, are particularly effective in modeling temporal dependencies in seismic time series. Furthermore, generative models like variational autoencoders (VAEs) and generative adversarial networks (GANs) are being explored for simulating synthetic seismic waveforms, augmenting limited datasets, and testing new hypotheses in a controlled manner.

Despite these advancements, some problems continued to persist. The scarcity of labeled data, particularly for low-magnitude events and rare seismic phenomena, remains a major bottleneck for training and validation. Techniques such as semi-supervised learning, unsupervised pretraining, and data augmentation are actively being developed to address these limitations. Additionally, the interpretability of deep learning models continues to be a critical area of focus, as understanding their

decision-making processes is essential for their adoption in high-stakes applications.

As computational resources and machine learning techniques evolve, the integration of deep learning into seismology is expected to further accelerate. From enhanced detection and phase picking to advanced predictive models and automated damage assessments, these methods are reshaping the field, making seismic analysis more efficient, accurate, and scalable. The growing synergy between machine learning and seismology holds immense potential to improve our understanding of earthquakes and mitigate their risks, paving the way for more resilient societies.

Chapter 3

Dataset

The data set used in this study consists of hundreds of samples that were selected from the report [1], evaluating and selecting earthquakes with magnitudes greater than 4 mb. A further skimming was done by going to find the satellite images covering that particular area, recorded around the time of the earthquake. To develop this dataset, extensive research was conducted using data from both the Copernicus Open Access Hub and the Alaska Satellite Facility (ASF). These platforms are essential resources for acquiring Synthetic Aperture Radar images from Sentinel-1 and other satellite missions, which are critical for monitoring seismic events and understanding their impacts.

3.1 SAR Images

Synthetic aperture radar (SAR) technology generates high resolution images by emitting radar pulses and capturing back-diffused signals from terrestrial targets, allowing day and night observation in all weather conditions. Unlike optical images, SAR preserves the electromagnetic properties and surface characteristics of the area, which are essential for environmental monitoring, terrain analysis, disaster management, and other applications where traditional images may be limited by external conditions. These images also play a crucial role in the high-resolution remote sensing and geo-location applications they include.

3.1.1 Sentinel-1 Mission and Products

The Sentinel-1 mission, an important component of the European Space Agency's (ESA) Copernicus programme, provides coherent SAR observations of the earth's surface through a constellation of twin satellites : Sentinel-1A (S1A) and Sentinel-1B (S1B). Launched a few years ago, these satellites orbit in a solar synchronous

configuration, allowing frequent and consistent images of the same areas. Each satellite is equipped with a SAR sensor, ideal for monitoring both the land surfaces and water and vegetation. The periodic review times of the mission and allow to have high quality SAR data and the same area, at intervals most recent so as to record any environmental changes in a few days.

Sentinel-1 Acquisition Modes

The Sentinel-1 mission includes four acquisition modes, each designed to capture data at different resolutions and swath widths based on specific observational needs. These modes support various polarization configurations—Vertical (V) and Horizontal (H)—to enhance data versatility. The Acquisition Mode are:

- **Interferometric Wide Swath (IW) Mode**: The primary mode for land monitoring, IW mode operates using Terrain Observation with Progressive Scans (TOPS) technology, capturing high-resolution data suitable for interferometric SAR (InSAR) applications. This mode typically supports dual polarizations, such as VV (vertical transmit and receive) and VH (vertical transmit, horizontal receive), which improve ground surface and vegetation detection.
- **Extra Wide Swath (EW) Mode**: The EW mode is optimized for sea ice monitoring and large-area environmental studies. it balances spatial resolution and coverage and, also, supports dual-polarization configurations, VV and VH, to differentiate between ice and open water more effectively.
- **Stripmap (SM) Mode**: The SM mode is the ideal for detailed studies of smaller areas, such as urban monitoring and agricultural applications, it can capture data in single or dual polarizations, offering flexibility for high-resolution observations and enhanced feature discrimination.
- **Wave (WV) Mode**: Designed for open-ocean observations, the WV mode captures small scenes at intervals along the satellite's orbit, supporting single polarization (VV) to provide consistent data for ocean wave monitoring and maritime navigation.

Sentinel-1 SAR Products

Sentinel-1 offers different SAR image layers, level 0, level 1, level 2, which contain 3 types of images based on data processing and the type of information provided:

- **Single Look Complex (SLC)**: SLC images retain both amplitude and phase information for each pixel, making them the most detailed product level. This

level of detail is essential for interferometric applications that rely on precise phase data to detect surface changes, such as land deformation and seismic shifts.

- **Ground Range Detected (GRD)**: GRD images are processed to remove topographic and geometric distortions, projecting data onto a ground-range plane. These images contain only amplitude information, making them suitable for mapping and classification tasks where phase data is not necessary.
- **Ocean (OCN)**: Tailored specifically for ocean monitoring, OCN products provide data on surface wind speed and wave patterns. These products support coastal surveillance, climate research, and maritime navigation.

3.1.2 Single Look Complex (SLC) Images in IW Mode

Since the study is based on seismic monitoring, the most suitable SAR products are, as mentioned above **Single Look Complex (SLC)** images in **Interferometric Wide Swath (IW)** acquisition mode, with dual polarization VV and VH.
SLC images retain complex values for each pixel, preserving both amplitude and phase information.
Each SLC acquisition in IW mode is divided into three sub-zones (IW1, IW2, IW3), each of which is further subdivided into bursts, representing continuous radar acquisitions on specific areas. This detailed structure allows SLC images to capture fine spatial details and preserve phase coherence between scans, making them ideal for surface motion observation, deformation analysis and seismic monitoring.
These images use the Progressive Scan Terrain Observation (TOPS) technique to minimize noise and maintain a constant resolution. This technique improves the accuracy of phase measurements on all lightning and underlays and is essential for studies requiring precise data and high spatial stability.
Each pixel of an SLC image contains a complex value comprising real and imaginary components, representing amplitude and phase. Amplitude indicates the strength of radar backscatter from the ground surface, influenced by factors such as surface roughness and material properties.
The phase, on the other hand, encodes the distance from the radar sensor to the target.
These data were used for the subsequent analyses which, thanks to the phase differences between the successive SAR acquisitions, detect the subtle movements of the soil.
Phase data in the Sentinel-1 SLC images facilitate the generation of interferograms. Interferometric SAR (InSAR) techniques exploit the high phase consistency and stability of SLC images, allowing measurements of soil movement at the millimeter scale.

InSAR is widely used in seismic monitoring, infrastructure stability assessments and geophysical research. The regular acquisition schedule, constant calibration and reliable processing standards of Sentinel-1 make it an ideal choice for InSAR analysis.

3.2 Dataset construction and Image Selection

The dataset construction process began by identifying a list of earthquakes labeled as "event-known" events. These events were pre-categorized as those with reliable and confirmed metadata, including specific details like longitude, latitude, time, and magnitude. This preliminary selection was made to garantee that each sample in the dataset was based on well-documented seismic activity, enabling a robust analysis of the earthquake's impact.

For each earthquake event, the epicenter's location had already been identified in *earthquake.parquet*. Then, using this information, a bounding box was delineated around the epicenter, encompassing the area most likely to experience significant ground deformation. This bounding box was designed to include not only the epicenter but also surrounding regions that may have been affected by seismic waves and secondary ground movements. Focusing on this immediate area, the study ensured that the analysis gathered all relevant data to assess the impact of the earthquake on surface characteristics.

Once the area around the earthquake epicenter was established, an initial search was made to find SLC images recorded within 15 days and covering the specified area. This type of image was carefully selected from SAR products detected by the Sentinel-1 mission.

The research has often produced more than one image, sometimes including several acquisitions during the event period. However, to ensure the quality and consistency of the data set, it was necessary to refine the selection criteria.
First, only those seismic events for which both pre- and post-event images were available were considered, since the presence of both pre- and post-event images is of fundamental importance in the generation of interferograms, These are based on the comparison of phase differences to detect ground displacement. The pre-event image serves as a reference base, while the post-event image highlights changes caused by seismic activity.
Secondly, events for which only pre-event images were available have been excluded from the data set as they do not provide the post-event data needed to detect seismic changes. For events where only post-event images were taken, the image set was supplemented with a manual search in the Data Search of the Alaska Satellite Facility (ASF) to locate the corresponding pre-event images.

This process involved searching for SAR data from the same satellite source

(Sentinel-1), ensuring that the time of acquisition was close to the time of post-event images. This check was essential to produce a valid interferogram, because a large time interval between the image pairs could generate a phase decorrelation and detect an incorrect shift. This manual search process involved matching post-event images with historical data from the same SAR source and the same Sentinel-1 type, to be sure that the acquisition time was close enough to form an effective interferogram.

In addition, it was necessary to select compatible SAR images not only temporally but also spatially between them, and to check that the two images overlapped covering the same area. InSAR techniques are highly sensitive to differences in ground coverage, band frequencies and orbital trajectories. Any discrepancy in these parameters could lead to inaccurate or distorted strain signals, reducing the reliability of interferometric analysis. By further re-checking that the pre-event and post-event images shared similar orbital characteristics, these risks have been minimised, making the process of generating interferograms more robust and accurate.

To obtain correct and consistent results, a manual search was used to find pre-event images that closely matched the post-event images in terms of acquisition parameters. This phase involved the re-examination of the Alaska Satellite Facility (ASF) Data Search to retrieve historical SAR data that cover the same area and have corresponding bands. Images taken from different orbits or with incompatible polarization bands were not suitable for the generation of interferograms, as these differences prevent accurate comparisons of phases.

To produce reliable interferograms, it was essential that the selected image pairs shared a consistent viewing angle and were acquired under similar geometric conditions. Any geometric discrepancy between the images can cause discrepancies that obscure the actual ground movement, making it difficult to distinguish between actual seismic impacts and observed variations. Therefore, a significant part of the data set preparation was to ensure that each image pair met these criteria. In cases where compatible image pairs were not available, certain seismic events had to be excluded from the data set, as only events with well-matched SAR image pairs could be used for an InSAR analysis.

Having performed automated searches and further manual searches to identify pre- and post-event images which might be compatible and suitable for progress with certain analyses, and having made some unavoidable compromises, The final data set comprises 477 samples. They have been selected based on the best available matches, and include a pre-event image and a post-event image for each detected earthquake. Some of these pairs may include images that have been repeated on multiple samples or recorded in the weeks prior to the seismic events. These represent boundary cases of the available images in the dataset that are not to be considered as errors or discrepancies and have been used as samples, correlating

with the analysis, since they include changes in the territory both before and after the event.

Although this approach may introduce some variability, it represents the best possible data set given the constraints of SAR data availability. This carefully selected approach seeks to maintain a high level of accuracy and compliance for interferometric analysis, while offering comprehensive coverage in terms of data quality for understanding earthquake-induced soil deformation.

3.2.1 Dataset Refinement

The dataset used in this study consists of several columns, each of which provides essential information on seismic events and corresponding SAR images used for interferometric analysis. Each column is described in detail below, explaining its relevance for the analysis made later on based on the image before and after the earthquake.

- **index**: This column contains a unique numerical identifier for each earthquake event. Acting as a direct reference, the index allows quick access to individual records and facilitates the organization of data within the dataset.
- **description**: This column provides a descriptive label for each earthquake, detailing the general position and tectonic classification of the affected region. Includes both a geographical reference, such as *border region between Iran and Iraq*, and a tectonic classification based on the type of *Flinn-Engdahl* region. This information contextualizes each event and links it to specific seismic regions, which can influence the characteristics of the earthquake and its potential impact.
- **time**: The timestamp in this column records the exact date and time of the earthquake occurrence and has been formatted as YYYY-MM-DD HH:MM:SS. This precise time information is crucial for chronological analysis, allowing differentiation between pre- and post-event images. It also correlates the spatial distance between the SAR images found and the actual earthquake detection, improving the time context of the study.
- **latitude**: This column specifies the latitude of the earthquake's epicenter, presented in decimal degrees. Latitude, which measures the north-south position, is fundamental for accurately mapping the earthquake's location on the Earth's surface. Along with longitude, this geographic data facilitates spatial analysis and visual representation of the earthquake distribution.
- **longitude**: This column specifies the latitude of the earthquake's epicenter, also in decimal degrees. Longitude measures the east-west position on the

earth's surface and, combined with latitude, accurately locates the epicenter of each earthquake. This data is useful for spatial analysis, allowing seismic events to be superimposed on maps.

- **depth**: The depth of each earthquake epicenter is recorded here, measured in kilometres. Depth plays a significant role in determining the impact of an earthquake, as surface events often produce more pronounced surface effects than deep ones. Understanding depth helps to assess potential damage and understand the nature of seismic activity.
- **magnitude**: This column shows the magnitude of each earthquake, which represents the size or energy released during the event. Magnitude values are essential for classifying the severity of each earthquake and are often used in risk assessment and disaster response planning. Including magnitude, the data set provides a quantifiable measure of earthquake strength.
- **magnitude_type**: This column indicates the specific type of measurement of the magnitude applied to each event.
- **image_ids_before**: This column lists the identifiers for the pre-event SLC images captured by the Sentinel-1 satellite (S1A). These images were acquired before the earthquake, establishing a baseline against which post-event imagery can be compared.
- **image_ids_after**: This column contains the identifiers for the post-event SLC images acquired by Sentinel-1 (S1A) that capture the state of the ground after the earthquake. The presence of post-event imagery is essential for *Interferometric Synthetic Aperture Radar (InSAR)* analysis, as it allows for the detection of changes in ground conditions caused by the earthquake.

In constructing the dataset, earthquake events with a magnitude greater than 4 mb (body-wave magnitude) were selected. This threshold was chosen to focus on significant seismic events that are more likely to result in observable ground deformations, which are suitable for analysis using Interferometric Synthetic Aperture Radar (InSAR) techniques. By applying this magnitude criterion, the dataset emphasizes events with a substantial impact, providing a robust foundation for studying surface displacement patterns associated with moderate to major earthquakes.
Essential parameters and image data for multiple seismic events, including location coordinates, magnitude, depth, and SAR image identifiers for *InSAR* analysis.
The set of these data forms an initial dataset for studying earthquakes. The combination of these allows interferograms and perform *InSAR* analysis to be created, balancing all data relating to time, space and magnitude through SAR products which take pictures before and after the seismic event allowing in-depth examination of impacts on the ground.

3.2.2 Map of Earthquake Epicenters in the Dataset

The selected seismic elements were displayed on a global map to give a concrete representation and geographical location.

The map 3.1 shows a clear spatial projection of events collected in the data set, that is all earthquakes with significant magnitude, contextualizing them in a wider geographical picture.

It highlights the recurrent nature of earthquakes, with many events occurring in the same regions over several years.

The figure below displays the global distribution of selected earthquake events, demonstrating areas with concentrated seismic activity:

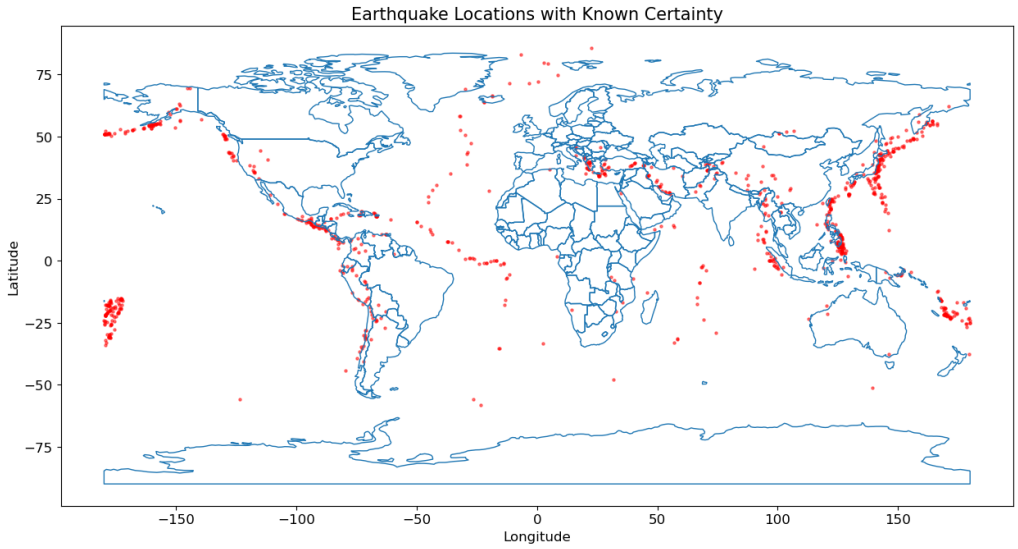

Figure 3.1: Global distribution of earthquakes with red markers for epicenters.

The red marks on the map correspond to the epicenters of earthquakes, traced in various regions of the world, whose concentration is quite evident along the limits of tectonic plates, especially in areas such as the Pacific Belt, where tectonic movements are frequent and intense.

This event highlights a very present seismic movement in these areas and provides a basis for studying the patterns of soil deformation over prolonged periods.

3.3 Analysis of an Earthquake case study

To analyse the impact of a specific seismic event, was focused on a single seismic event and use a step-by-step approach from capturing the pre- and post-event SAR

images, then analyzing its geographic coordinates, to visualizing its epicenter on a map. This section describes the detailed process involved in studying a single seismic event

3.3.1 Downloading SAR Images

The first phase of an earthquake investigation consists of two SAR images: one before the earthquake (pre event) and another shortly after (post-event). These images can be obtained from resources such as the Copernicus Open Access Hub or the Alaska Satellite Facility (ASF) Data Search. The image before the event establishes a baseline, while the image after the event reveals any surface changes caused by seismic activity.

The method used to find and download the SAR products:

The seismic event was identified by its unique timestamp and geographic coordinates (latitude and longitude). Using this information, a search was carried out on the ASF platform to obtain satellite data covering the area around the earthquake epicenter. It was crucial to ensure that both images covered the same area entirely and were superimposed in terms of space, so you have the same Mode, Track, Pass and Polarization. These parameters were essential for reliable interferograms and to minimize geometric distortions.

3.3.2 Verifying the Coordinates

After the images have been taken, the next step is to verify the precise geographical location of the earthquake epicenter. The coordinates (latitude and longitude) were checked to confirm that they corresponded to the target area for analysis. These coordinates are the exact point of origin for seismic activity and serve as a focal point for the visualization and interpretation of soil deformation models.

In this analysis, the coordinates of the epicenter are plotted on a global map. By examining them in the context of geological features, fault lines or tectonic boundaries, further information can be obtained on the factors contributing to the seismic event. This information also helps to define the area of interest for subsequent interferometric analyses, focusing on regions most likely to be affected by the earthquake.

3.3.3 Visualizing on a Map

Once the coordinates of the epicenter are verified, to better understand the location of the earthquake and its geographical context, the epicenter is plotted on a map to also see the area around it.

This procedure involves mapping the latitude and longitude of the epicenter on a map, at an appropriate scale, and the epicenter has been highlighted with a marker, a red point.

If necessary, additional layers such as fault lines, plate boundaries or nearby urban areas can be added to the map to provide contextual information on the earthquake environment.

From the acquisition of data to the geographical display, it provides the basis for a thorough and systematic analysis of an earthquake event.

The display of SAR images on a map not only helps to understand the characteristics of the earthquake, but also provides a solid framework for interpreting data generated in subsequent interferometric analyses which may vary in the presence of mountains, seas or rivers.

The following figure 3.2 shows the epicenter of the first earthquake, marked with a red dot and labeled "Epicenter", providing a clear view of the dimension of the affected area.

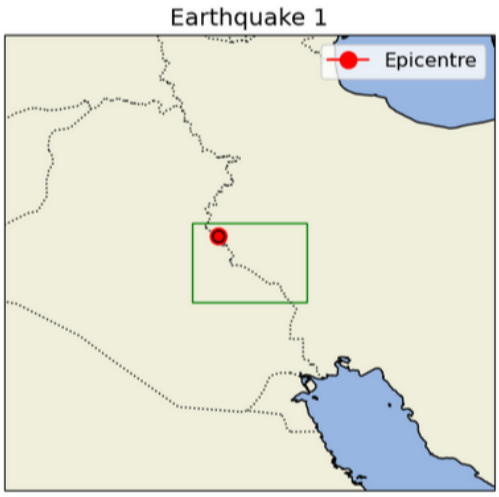

Figure 3.2: Bounding box for SAR pre-event image acquisition.

The green box represents the area of pre-event SAR image acquisition, encompassing the region of interest for interferometric analysis.

The case study earthquake occurred in the Iran-Iraq border region with a magnitude equal to 5.58 mb, on 11 January at around 6am and its related information are:

- **Epicenter's Latitude:** 33.7388
- **Epicenter's Longitude:** 45.7501
- **Magnitude:** 5.58 mb

- **Pre-event SAR Product:** S1A_IW_SLC___1SDV_20180106T025411_20180106T025439_020028_0221F4_0110
- **Post-event SAR Product:** S1A_IW_SLC___1SDV_20180118T025410_20180118T025438_020203_022783_6334

Displaying now the post-event SAR image, was able to confirm that they are perfectly superimposed 3.3

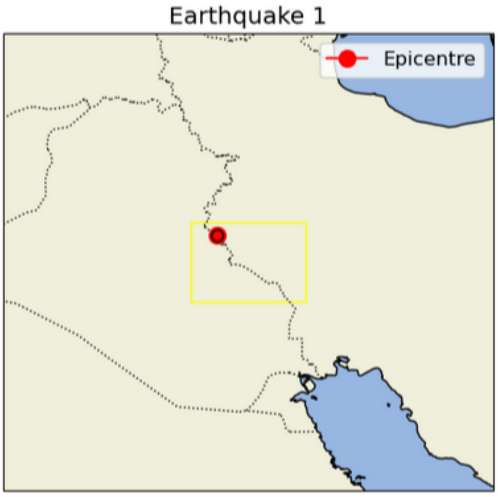

Figure 3.3: Bounding box for SAR post-event image acquisition.

The yellow box represents the area of post-event SAR image acquisition, encompassing the region of interest for interferometric analysis.

Plots provide a clear view of the earthquake epicenter in relation to the surrounding region, allowing precise observation of its location. The red mark identifies the epicenter, while the boxes indicate the areas covered by the pre- and post-event SAR images. By visually locating the epicenter in the SAR image coverage field, this map ensures that the selected images cover the key area of interest for subsequent analysis.

This check is important to confirm that both images capture the same earthquake region and to obtain an accurate analysis of the deformation of the ground, which is crucial for the generation of the interferogram. The aim will be to be able to visualize a concentric phase pattern around the epicenter, a phase difference that indicates a distinct movement of the earth's surface from its surrounding area.

By covering the SAR image at the epicenter, it becomes possible to observe the patterns of soil deformation more clearly, Allow a more accurate interpretation of the seismic impact and the detection of characteristic deformation rings associated with the focal point of the earthquake.

3.4 Extensive Evaluation of Seismic Events

The analysis of individual seismic events has been extended to all earthquakes present in the data set. For each earthquake, SLC images were downloaded both before and after the event to capture surface changes and phase changes.

The process of acquiring SAR products required flexibility, given the different locations of earthquakes. Each seismic event in the data set represents a distinct geographical region, ranging from urban areas and remote landscapes to ocean locations. For example, earthquakes in densely populated areas may present more complex deformation patterns due to infrastructure and different land cover, Whereas those in remote areas often provide a clearer and less obstructed view of natural soil changes.

In many cases, more than one earthquake has been recorded at close range, both within the same region and along tectonically active fault lines. These recurring seismic events have led to overlapping epicentres, sometimes separated by only a short distance or occurring in a short time span. In these cases, the same SAR images could occasionally be reused, provided that the timing of each earthquake was consistent with the dates on which the images were taken.

Several earthquakes in the data set have been located in or near the sea, so additional considerations were necessary because of differences in interferometric models observed on water compared to land. SAR signals above water can be influenced by factors such as sea surface conditions, wind speed and ocean currents, which may introduce noise or variability into the data.

In some cases, however, signals may spread to coastal regions, with potential impacts on infrastructure and ecosystems. All earthquakes were considered in the study to obtain a complete picture of seismic impacts in different environments.

As for example, the figure below 3.4 shows the earthquake near the Gulf of California on 19 January which has its epicenter approximately in the sea.

- **Epicenter's Latitude:** 26.6196
- **Epicenter's Longitude:** -110.9457
- **Magnitude:** 5.65 mb
- **Pre-event SAR Product:** S1A_IW_SLC___1SDV_20180114T132118_20180114T132145_020151_0225F3_53F5
- **Post-event SAR Product:** S1A_IW_SLC___1SDV_20180126T132118_20180126T132145_020326_022B7D_D29E

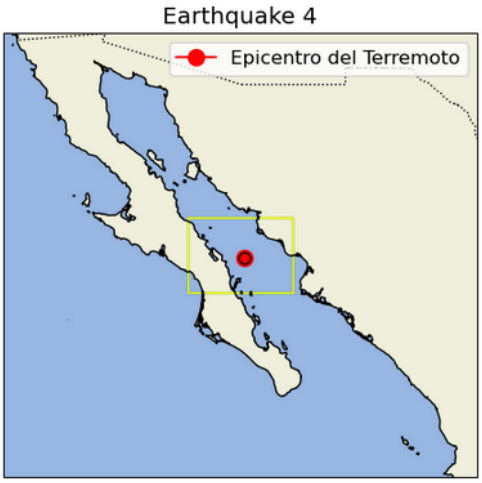

Figure 3.4: Bounding box for SAR pre and post-event image.

Chapter 4

Interferograms

Interferometric synthetic aperture radar (InSAR) is a powerful technique used to detect ground displacement caused by various geophysical processes, including seismic activity.
By comparing phase information between two or more SAR (Synthetic Aperture Radar) images acquired before and after an event, it is possible to generate an interferogram.
This measurement is visually represented to observe the surface deformation as a series of phase fringes of varying intensity.
The generation of interferograms starts with the use of SNAP Desktop, a software tool developed by ESA to process and modify SAR data. Next, the options for changes are identified and then a workflow is generated to be replicated and automated in Python, allowing efficient batch processing of SAR images for multiple seismic events.

4.1 Interferogram Generation

The first step in producing an interferogram is to use the SNAP desktop to process pre- and post-event SAR images. SNAP (Sentinel Application Platform), developed by the ESA, is a powerful tool specifically designed to process Sentinel-1 data, supporting advanced workflows for SAR interferometry. According to SNAP documentation, this platform offers a range of modules and tools ideal for SAR data management, including co-recording, interferogram formation and phase filtering, which are essential for reliable interferometric analysis.

According to the SNAP documentation, the process of generating interferograms involves several detailed steps.

4.1.1 Opening the SAR Images

The pre-event and post-event SAR images were first imported into the SNAP Desktop interface.
These images were acquired from the Alaska Satellite Facility (ASF) and included in the dataset as described earlier. Both images are calibrated and aligned to ensure geometric accuracy, an essential step for producing reliable interferograms.

4.1.2 Coregistration

Coregistration is the process of aligning the two SAR images to ensure that corresponding pixels in both images represent the same ground point. SNAP's Coregistration was used for this step, ensuring precise alignment between the pre-event and post-event images.
This step is crucial because any misalignment can introduce artifacts into the interferogram, making the interpretation of ground displacement more difficult.

4.1.3 Interferogram Formation

Once the images were coregistered, the Interferogram Formation was applied. This generates a phase difference image by comparing the radar signals from the two images.
The result is an interferogram with significant noise, making it difficult to clearly distinguish phase differences. As a result, the phase changes are not easily discernible, and the interferogram does not provide a clear visualization of ground displacement.

4.1.4 Filtering

Deburst, Goldstein Phase Filtering, and Topographic Phase Removal were applied to reduce noise and remove topographic effects, improving the clarity of the phase patterns.
These operations made the interferogram easier to interpret by eliminating burst discontinuities and isolating the ground displacement signals.

4.1.5 Exporting the Interferogram

After generating the interferogram in SNAP Desktop, it can be exported in different formats for further analysis and visualization. The most common formats are hdf5, tiff and BEAM-DIMAP, which is the default format and, then, plot the interferogram to check if the phase difference is clearly visible, emphasizing the location of the epicenter.

While SNAP Desktop provides a user-friendly environment for generating interferograms, this process must be repeated manually for each earthquake event. To streamline the analysis, the workflow was replicated and automated using Python script with the snappy and snapista libraries, allowing for efficient processing of multiple earthquake events.

4.2 Pipeline phases

The workflow for generating an interferogram from two SAR SLC images consists of multiple steps.

4.2.1 Opening the SAR Images

- **Read Operators**: The Read operators serve as the entry point for the SAR images. Both pre-event and post-event SAR data, acquired from the Alaska Satellite Facility (ASF), are loaded into the processing environment. This step is critical for ensuring that the raw SAR data is imported correctly for further processing, maintaining the original format for calibration and interferogram formation.

4.2.2 Coregistration

- **Apply Orbit File**: This operator applies precise orbit information to correct the satellite positioning errors in SAR data. Precise orbits provide better geolocation accuracy compared to restituted orbits, which is essential for accurate image alignment during the coregistration process.

- **TOPSAR-Split**: Sentinel-1 data is acquired in multiple swaths and polarizations. The TOPSAR-Split operator is used to extract only the relevant swath and polarization (e.g., VV or VH) for the region of interest, reducing unnecessary data and making processing more efficient.

- **Back Geocoding** : This step aligns the pre-event and post-event images, ensuring that corresponding pixels in both images represent the same ground location. Coregistration accuracy is essential for the quality of the interferogram, as misalignments lead to errors in phase calculation and displacement measurements.

- **Enhanced Spectral Diversity (ESD)**: ESD is an optional step applied when handling multi-burst SAR data. It corrects range and azimuth shifts between bursts, improving the coregistration accuracy. However, it significantly

increases the processing time and is only recommended for acquisitions where Doppler centroid variations are high.

4.2.3 Interferogram Formation

- **Interferogram Formation**: The phase difference between the coregistered images is calculated to generate an interferogram. This phase contains contributions from surface displacement, topography, and atmospheric disturbances. To focus on ground deformation, other phase components (such as flat-earth and topography) are removed in later steps.

4.2.4 Filtering

- **TOPSAR-Deburst**: It is an optional step, Sentinel-1 acquires data in bursts, which can create discontinuities in the interferogram. The Deburst operator merges these bursts into a continuous strip, eliminating the seamlines, which is essential for producing interpretable and accurate interferograms.
- **Topographic Phase Removal**: This step is optional and removes the phase contribution from the Earth's topography using a digital elevation model (DEM). By subtracting the topographic phase, the interferogram focuses solely on the surface displacement caused by seismic activity, improving the accuracy of the ground deformation measurements.
- **Goldstein Phase Filtering**: The Goldstein Phase Filtering is an optional step,technique is applied to reduce noise in the interferogram. By smoothing out the noisy phase signals while preserving the main deformation patterns, it makes the interferogram clearer and easier to interpret for further analysis.

4.2.5 Export Interferograms

The generation and display of interferograms required experimentation with different file formats to balance data compatibility, usability and integrity. The most popular formats are *BEAM-DIMAP*, *Tiff* and *Hdf5*, each with its own advantages and limitations.

- **Write Operator**: Once the interferogram is processed, the Write operator saves it in the desired output format, such as HDF5 or GeoTIFF, which are commonly used for geospatial analysis. The format selection depends on the tools intended for visualization and further analysis of the displacement data.
- **Visualizing Interferograms** Interferograms were initially saved in *BEAM-DIMAP* format, the default format for many SNAP operations. This format

provides a highly detailed data structure, but was not very practical for visualization:

- **BEAM-DIMAP Format** While BEAM-DIMAP provides a highly detailed data structure, the complexity of the format limited its usability when trying to integrate it into automated processing pipelines, particularly when the goal was to export the interferograms for further analysis outside SNAP.
- **TIFF Format** It was also considered as an alternative format for saving the interferograms, given its wide compatibility with various image processing tools. However, the multi-band structure required to store complex interferometric data (such as phase, coherence, and intensity) was difficult to manage. While TIFF allowed for easy visualization of amplitude and coherence layers, handling the phase information remained problematic, as the interferometric phase data needed further post-processing to extract useful deformation patterns.
- **HDF5 Format** It was adopted for its balance of size efficiency and compatibility across various tools. However, some limitations persisted: while HDF5 solved many of the issues with size and compatibility, the saved files in this format doesn't contain the full phase and intensity files.

During the study, various experiments were carried out to obtain a solution that would balance running times with noise, removing unnecessary operators where possible for better performance. Until a final compromise was reached that made the epicenter and its location visible for phase difference and noise manageable through the use of deep learning models.

4.3 Optimizing the Interferogram Pipeline

The process of generating interferograms, while it can analyse soil deformation, can be time-consuming, especially when large data sets or multiple seismic events are involved. During the initial stages of this study, it became apparent that the generation of a single interferogram could take several hours, depending on the complexity of the pipeline and the available computational resources. This delay has been particularly complicated when trying to extend the process to hundreds of seismic events.

To solve this problem, several pipelines were tested to find the optimal workflow that balances precision and processing speed. Trying to test different approaches, eliminating some operators where possible and looking at the type of result obtained in terms of visualization.

In the various attempts it was tested with and without the various filters, but the initial interferogram contained too much noise and to improve the timing also Enhanced Spectral Diversity (ESD) which is recommended if you use more subswath, as in all the earthquakes of the dataset, but not mandatory.

4.3.1 Initial Pipeline Experiments

The process was initially tested using multiple configurations in SNAP Desktop, with the aim of reducing processing times without compromising interferogram quality. The phases generally followed a common pattern: reading of pre- and post-event SAR images, application of corrections to the orbital file, co-recording, generation of the interferogram, deburring and phase filter application. However, each of these steps could be modified to improve efficiency.

Without any filtering In the first iteration, the pipeline was kept as minimal as possible, without Goldstein Phase Filtering or Enhanced Spectral Diversity (ESD). This reduced the processing steps but resulted in lower-quality interferograms with visible discontinuities in the burst boundaries.

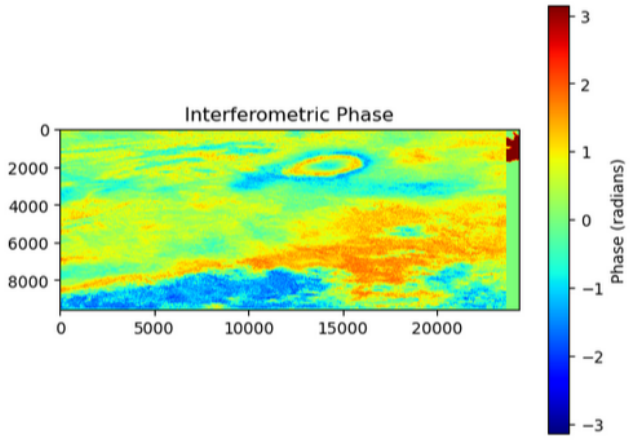

Figure 4.1: Interferogram without any filtering

ESD The second pipeline incorporated ESD, which helps improve the coregistration of SAR images in areas with high Doppler centroid variations, particularly for wide-swath acquisitions. While this configuration improved the coherence of the interferograms, it significantly increased the processing time. In some cases, the generation time stretched beyond several hours for a single interferogram. Despite the improvement in quality, the time required for ESD made it impractical for batch processing across multiple events.

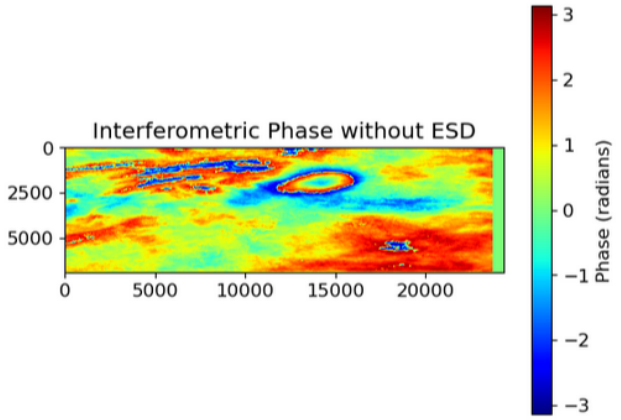

Figure 4.2: Interferogram without ESD

Goldstein Phase Filtering Another experiment was focused on the use of the Goldstein Phase Filtering operation. Operation Goldstein has proved necessary to ensure high quality interferograms, the improvement of image quality has made it a necessary compromise. As can be seen from the final output shown in the image below, which is about the same as the initial unfiltered interferogram.

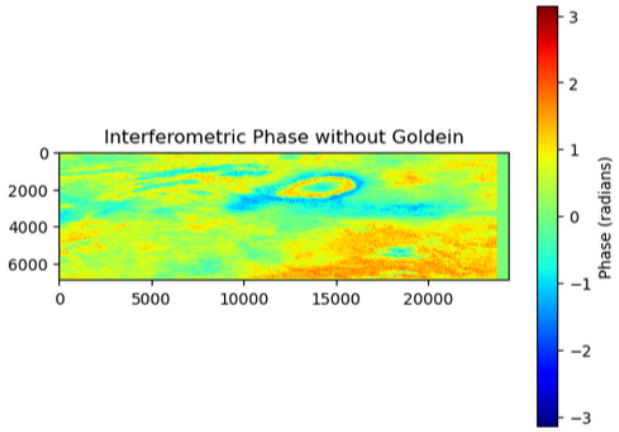

Figure 4.3: Interferogram without Goldstein Phase Filtering

4.4 Final Interferogram Generation Pipeline

The final pipeline adopted for this study was developed after experimenting with various configurations to balance time efficiency and output quality. Each step of the pipeline follows best practices from the SNAP documentation, ensuring the

generation of high-quality interferograms suitable for analyzing ground deformation caused by seismic events. Below is a detailed description of each step in the pipeline:

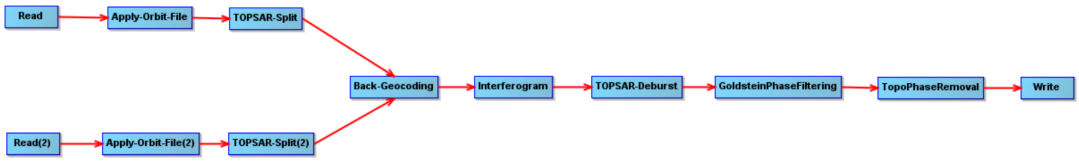

Figure 4.4: Final processing pipeline.

4.4.1 Opening the SAR Images

- **Read Operators 1**: The Read Operator loads the pre-event SAR SLC image, which becomes the source product.
- **Read Operators 2**: The Read Operator loads the post-event SAR SLC image, which becomes the target product.

4.4.2 Coregistration

- **Apply Orbit File 1**: This adds precise orbit information that is automatically downloaded from Sentinel Precise (Auto Download) for pre-event SAR product.
- **Apply Orbit File 2**: This also adds precise orbit information that is automatically downloaded from Sentinel Precise (Auto Download) for post-event SAR product.
- **TOPSAR-Split 1**: For pre-event image, were selected the Subswaths (IW1, IW2, IW3), the polarization (VV) and the number of burst, from 1 to 9.
- **TOPSAR-Split 2**: For post-event image, were selected the Subswaths (IW1, IW2, IW3), the polarization (VV) and the number of burst, from 1 to 9. It is very important to select the same parameters for source and target product.
- **Back Geocoding**: Where the two pre-event and post-event images are aligned, making sure they match and that they are products on which TOPSAR Split operator has been applied. The SRTM 1Sec HGT, Global Digital Elevation Model, was used to correct topographic effects in interferograms, providing high-resolution elevation data [11].

4.4.3 Interferogram Formation

- **Interferogram Formation**: The interferogram generated by the two starting products shows the phase difference of the soil with so much noise that no change or event is recognizable. It was selected again SRTM 1Sec HGT, while the other parameters were used the default ones.

4.4.4 Filtering

- **TOPSAR-Deburst**: This is an optional step to reduce the discontinuity in the interferogram.
- **Topographic Phase Removal**: Useful for removing phase contribution from the Earth's topography using SRTM 1Sec HGT as a digital elevation model (DEM).
- **Goldstein Phase Filtering**: This technique is applied to reduce noise in the interferogram and is an optional step that is important and necessary to achieve a clear output as possible.

4.4.5 Export Interferograms

- **Write Operator**: The operator saves the interferogram in the desired output format and in the desired output folder.
- **Visualizing Interferograms**: The type that was most easily handled was *Hdf5*, so files were saved in this format. They contain various information like band and metadata, but in band there are only the components *coherence (coh)*, *real (i)* and *imaginary (q)*, on which you went to calculate the phase file as:

  $$\arctan(q, i)$$

  Using the real and imaginary parts, the final interferogram plot is obtained showing the phase difference in radiants for that particular earthquake in that specific area with a specific color range.

4.5 Interferogram of an Earthquake case study

The first earthquake was selected for the data set analysis and occurred in a region with known seismic activity. Starting from the SAR products, pre and post event, an interferogram will be generated for each sublayer by selecting all bursts. This case serves as a demonstration of the process of generating interferograms, from

SAR imaging to visualizing soil deformation. The phases show the output obtained with the optimized pipeline described in the previous sections.

4.5.1 Earthquake Details

- **Location**: Iran-Iraq border region
- **Date and Time**: 2018-01-11 at 06:59:31.600
- **Magnitude**: 5.58 mb

4.5.2 SAR SLC Images and Interferogram

Pre- and post-event SAR images were obtained from the Alaska Satellite Facility (ASF) and processed using the workflow, starting with co-recording and ending with phase filtering. Images were saved in HDF5 format, which required manual calculation of the phase component.
As the SAR data of Sentinel-1 are acquired in multiple subswaths : IW1, figure 4.5, IW2, figure 4.6, IW3, figure 4.7, three separate interferograms were generated, one for each subswath.
Interferograms generated for this event show characteristic phase fringes, which represent the movement of the ground caused by seismic activity.

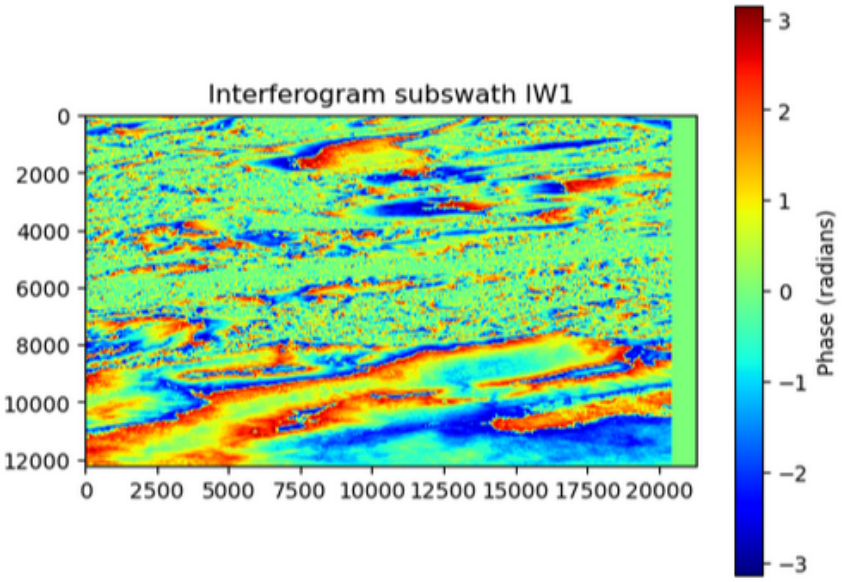

Figure 4.5: IW1

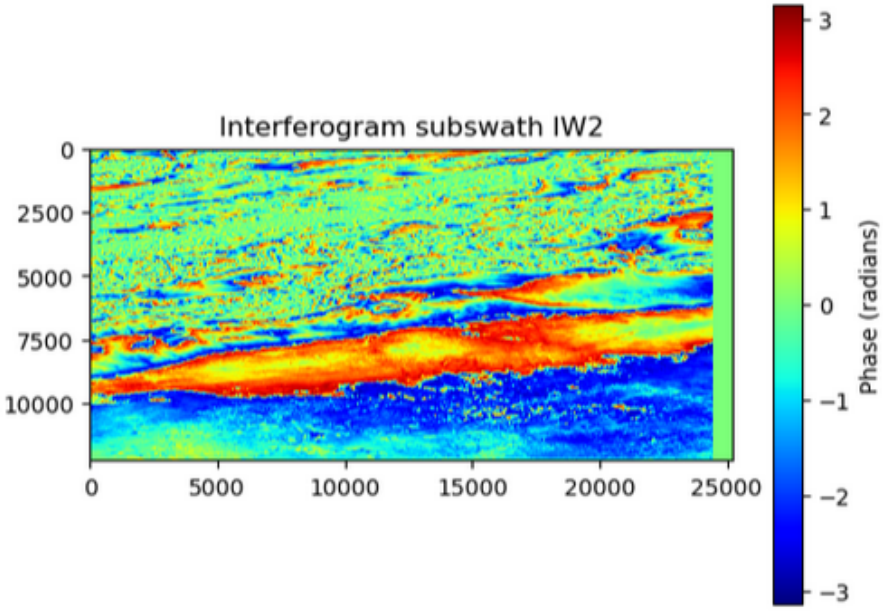

Figure 4.6: IW2

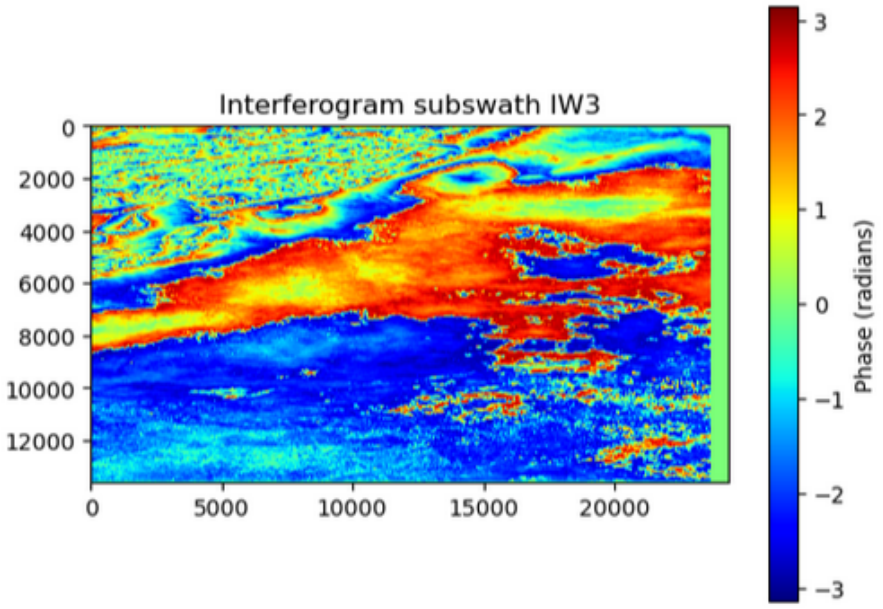

Figure 4.7: IW3

4.5.3 Interpretation of the Interferograms

In the figures above, it is possible to observe the different concentric phase fringes clearly visible in some areas, compared to others. For each subzone IW1, IW2, IW3 there is a corresponding interferogram, as shown by the map display the subwath IW1 contains the epicenter, in fact you can observe a phase variation that corresponds to the cumulative displacement of the terrain, with each complete color cycle representing a certain degree of surface movement. By analysing all three interferograms and following the [4] a wider picture of the ground displacement for a specific earthquake can be obtained:

- **Concentric deformation**: Interferograms show patterns of concentric fringes around the epicenter, indicating the radial movement of the soil caused by the earthquake.
- **Multi-Subswath Coverage**: By generating interferograms for each underlay (IW1, IW2, IW3), a more complete analysis of the earthquake impact is obtained.
- **Area of potential impact**: by analysing the extent of the fringes on all sub-tracts, it is possible to estimate the region most affected by the earthquake and its effects.

The study continues by generating interferograms for each earthquake detected, automating the process through graphs running the SNAP pipeline in python and trying to get as many interferograms as possible.

Chapter 5

Seismic Dataset Development

This chapter describes in detail the process of creating the new dataset for training and testing machine learning models for seismic analysis. Given the time required to generate interferograms, the final dataset is composed of only a few of the selected earthquakes, of which one interferogram is respectively obtained for each subswath. The data preparation started with enlarging the images to obtain a sufficient number of samples, and then calculating distances from the spots to their epicenter to provide a complete spatial context.

5.1 Data Preparation

The first 40 earthquake events from the initial dataset were selected for this study. These events were chosen based on the availability of high-quality pre-event and post-event SAR images, which were processed using Interferometric Synthetic Aperture Radar (InSAR) techniques to capture ground deformation.

SAR products include several sub-waths because of their large size, specifically they correspond to IW1, IW2, IW3, which is why three interferograms were generated for each earthquake. They were created by comparing phase differences between pre-event and post-event SAR images and by application of noise reduction and phase unwrapping techniques.

The results are generally clear and show a phase change, more evident around the epicenter. Sometimes they are not so clear or are in particular areas where the images have not been recorded correctly or have some deviations. In the data set collected, only for earthquake number 5 alone was it not possible to generate the IW2 subswath interferogram for land-related problems, such as possibly the presence of the sea.

5.2 Data Preprocessing

To counter the size of the areas covered by SAR products and the scarcity of data, due to excessive computational time, a resize of each interferogram was made to have a high number of samples so that it is suitable for future analysis.

The large interferogram generated for each earthquake has been divided into smaller areas for more efficient model training:

- **Patching into 2000x2000 Pixels**: Each interferogram was segmented into patches of size 2000x2000 pixels. These patches were created for each subswath separately, preserving the unique acquisition geometry of the Sentinel-1 SAR images. For each patch was assigned a unique index, to track and organize the data effectively.

Then, to make the data more manageable and suitable for machine learning models, each 2000x2000 pixel patch was resized:

- **Image Resizing to 256x256 Pixels**: The patches were resized to 256x256 pixels. This reduction in size helped decrease computational complexity while preserving essential ground deformation features.

After resizing the data, the new patches obtained were visualized to check the correspond to the relative subswath e to observe their localization on the map.

The following image displays the relative patches about a specific subswath of a SAR product for a specific earthquake 5.1.

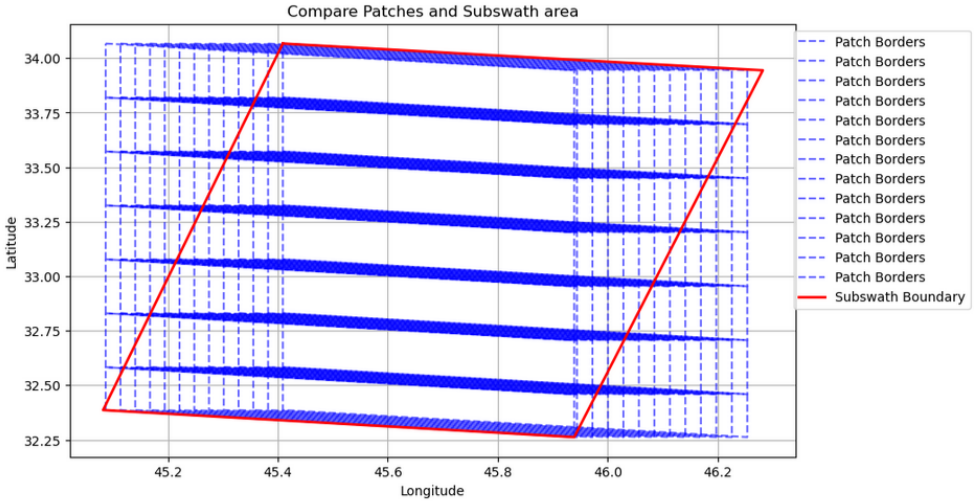

Figure 5.1: Patches resized of the Interferograms about one Subswath.

5.3 Final Dataset

The final data set comprises 40 interferograms, each representing seismic events captured and processed using SAR products.
The interferograms were sized to create uniform and small areas, in order to facilitate analysis, forming a dataset composed of 29973 patches, about thirty thousand samples.
Each data is characterized by three channels, corresponding to the bands from the information contained in the SLC images:

- **Coherence (coh)**: Captures the consistency between pre- and post-event images, indicative of the stability of the terrain.
- **Real (i)**: represents the actual component of the image.
- **Imaginary (q)**: corresponds to the imaginary component of the image.

These channels are the essential and fundamental features for machine learning models to capture ground deformation caused by seismic events.

After scaling of data, each area found is saved in an HDF5 file format, containing one file for each band, subswath and patch, while metadata and other relevant attributes are not saved in the single file.

In this regard, the details containing the geospatial information are extracted and stored in a CSV file so that accurate estimates can be made based on the territorial location of each patch.

This dual structure ensures proper data management and ensures that all the information needed to effectively use machine learning models is stored.

5.3.1 HDF5 File Structure

The HDF5 file serves as the primary repository for the SAR data patches, where each row corresponds to a unique patch derived from the 40 earthquake events. The structure of the final file is organized as follows:

- **Interferogram**: Identifies the interferogram, specifying the earthquake event, the subswath, and the temporal range of SAR image acquisitions.
- **IW Band**: Indicates the interferometric band (IW) used and corresponds to one of the three channels that make up the image.
- **Patch Index**: A unique identifier for each patch, ensuring easy organization and access.

The HDF5 format was chosen for its high performance and advantages in handling large datasets. It allowed efficient reading and writing operations and adaptable to various machine learning models during training.

5.3.2 CSV File Overview

The related CSV file contains all the metadata needed for each patch, providing a complete description of the spatial context of the data. The columns in the CSV file are described below:

- **index**: A unique identifier for each Earthquake for tracking.
- **interferogram**: Identifies the interferogram, linking each patch to its corresponding earthquake event and time range of SAR image acquisitions.
- **band**: Coh, i, q
- **iw_band**: Specifies the Interferometric Wide (IW) band used like coherence, intensity.
- **patch_index**: A unique identifier that corresponds to the one contained in the hdf5 file.
- **patches**: A unique identifier for each patch for analysis.
- **lat_top_left**, **lon_top_left**: Top-left patch coordinates.
- **lat_top_right**, **lon_top_right**: Top-right patch coordinates.
- **lat_bottom_left**, **lon_bottom_left**: Bottom-left patch coordinates.
- **lat_bottom_right**, **lon_bottom_right**: Bottom-right patch coordinates.
- **center_lat**, **center_lon**: The center coordinates of each patch, calculated as the average of the previous coordinates.

This CSV file connects the raw data in the HDF5 file and the associated metadata, to have seamless extraction of the zones and their geographical location which must necessarily be known to help for the predictive model.

5.4 Patches of an Earthquake case study

The seismic events containing their respective interferograms were divided into approximately 800 patches for each. Earthquake 1, taken as a case study, includes 3 interferograms, one for each subswaths IW1, IW2, IW3. Subsequently, the resize was applied and were obtained precisely:

231 patches for IW1 subswath, 77 for each band,

273 for subswath IW2, 91 for each band,

273 for subwath IW3, of which 91 for each band.

The coordinates of these have been extracted from the hdf5 files and projected into the globe to display the correct position by comparing it with the SLC images relating to that particular earthquake.

First, the real areas were visualized at each subswath, going to locate them on the map it was possible to observe that they corresponded with the entire area of the SAR product.

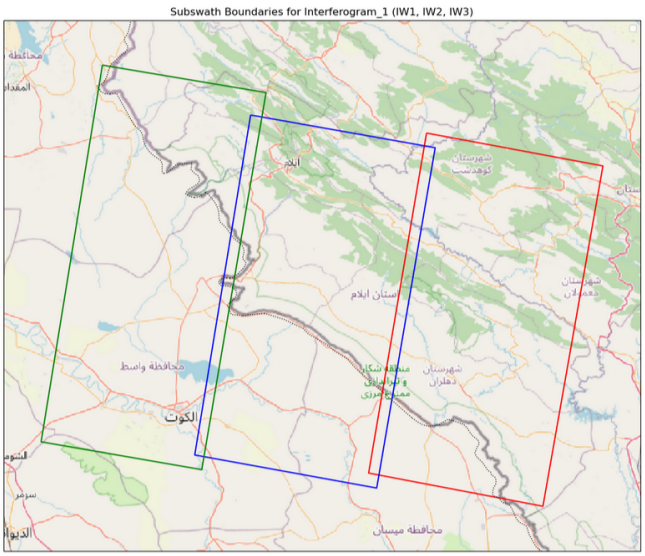

Figure 5.2: Subswaths IW1, IW2 and IW3 about the Earthquake 1

Second, all patches obtained after the resize of each interferogram were mapped to check the position and that the coordinates extrapolated from the various information are actually those corresponding to that area.

Third, in addition to comparing the position of each patch with their relative subswath, latitude and longitude is plotted for further control.

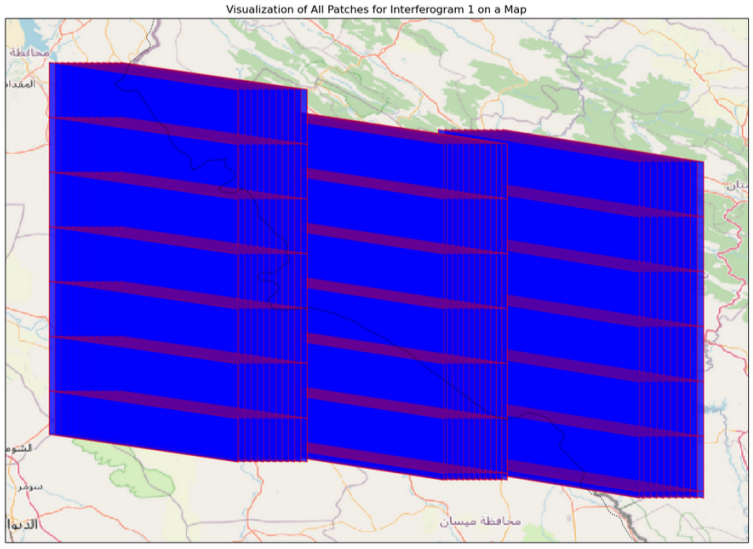

Figure 5.3: Patches resized for each Interferometric Wide (IW)

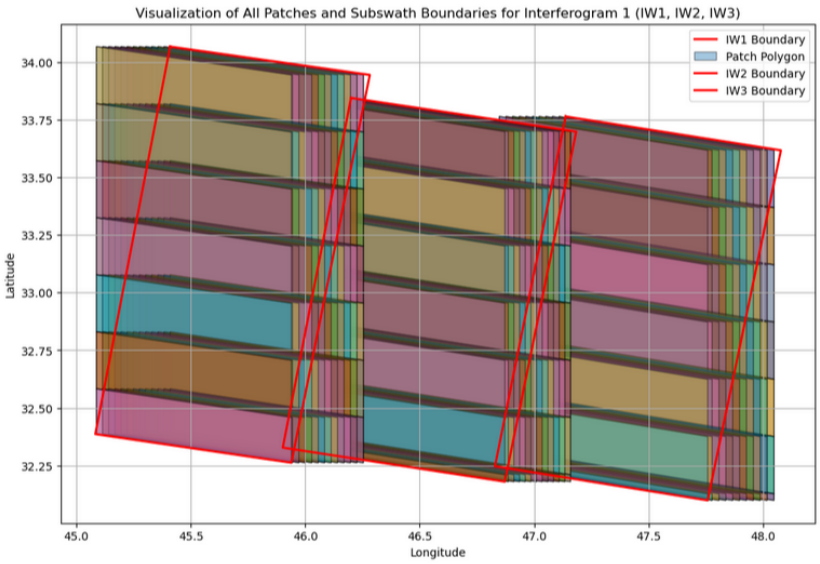

Figure 5.4: Patches resized about the Earthquake 1 with latitude and longitude.

5.5 Data organization and use

The patches extracted from SAR interferograms are stored in an HDF5 file, with metadata stored in a corresponding CSV file, which is why a specific class,

`PatchDataset` has been developed to handle data loading, Preparation and transformation to make them adaptable to the machine learning model.

The data set is structured to achieve efficient analysis and spatial mapping. Each patch is loaded, transformed and stacked into a three-channel tensor, one for each band, making it an input suitable for the various models that will be tested later.

5.5.1 Data Splitting

To evaluate the model's performance effectively, the data set was divided into three sub-sets to ensure a balanced distribution between training, validation and testing. This approach aims to provide a solid framework for model development and evaluation, while reducing as much as possible the risk of overfitting and trying to ensure randomness.

- **Training Set (70 %)**: The training set comprises most of the data set and is used to form the machine learning model. This allows the model to learn about the patterns, relationships and characteristics present in the data, trying to give a good basis for predicting target variables. Using 70% of the data, sufficient diversity and a large amount of information is maintained to generalise results for different seismic events.

- **Validation Set (15 %)**: The validation set is used during the training phase to refine model parameters such as learning speed, batch size and architecture-specific hyperparameters. It also serves as a control point to monitor the performance of the model and identify problems such as overfitting or underfitting.

- **Test Set (15 %)**: The test set is retained until the final evaluation stage. It serves as a non-distorted data set to assess predictive accuracy. This set is critical to determine whether the model can make reliable predictions on completely new data, simulating real-world scenarios. The use of a separate test set prevents information from leaking and ensures the integrity of the assessment process.

The choice of this data subdivision was made in order to preserve the geographical and seismic diversity of the various earthquakes. Selecting 70 % of the file rows corresponding to patches on different interferograms for training, 15 % randomly sampled for validation and 15 % of data reserved exclusively for final evaluation, No overlap between training or validation sub-sets. This distribution helps to maximise the usefulness of available data while maintaining reliable evaluation metrics and ensures a balance between learning and independent evaluation. It

ensures that the data set is used effectively, preserving the integrity of the test data for a non-distorted performance evaluation.

Chapter 6

Machine Learning Models

Deep learning has revolutionized the analysis of complex data sets, including geospatial and information-rich data. The most efficient architectures for solving tasks include convolutional neural networks (CNN) and vision transformers (VIT). The former, with their localized convolutional operations, excel at capturing spatial hierarchies in data, while the latter exploit self-care mechanisms to model global dependencies.
This part of the study provides an overview of these networks such as ResNet, ConvNeXt and ViT, focusing on specific models used for subsequent analyses.

6.1 ResNet

Residual networks (ResNet) represent a breakthrough in convolutional neural networks (CNN), associated with deep learning architectures. Before ResNet, increasing the deep neural networks often led to a gradient-of-loss problem, where gradients of the loss function decreased as they propagated through the network.

This has obstructed effective training and reported a reduction of performance. ResNet introduced residual learning, using fast links to bypass layers, improving deep networks and allowing them to reach new depths.

The ResNet architecture is made up of several features that make it versatile and powerful for various tasks such as:

- **Residual connections**: The central idea of ResNet lies in its residual connections, which allow the network to learn residual mappings instead of attempting to learn direct feature mapping. These connections add the input of a level directly to its output, effectively bypassing one or more levels.
- **Scalability**: ResNet architectures are available in various depths, such as ResNet-18, ResNet-34, ResNet-50, ResNet-101 and even deeper versions.

Scalability allows it to be adapted to various tasks of varying complexity, from lightweight models with limited resources to deeper networks for high performance tasks.

- **Bottleneck Residual Blocks**: to improve the calculation efficiency in deeper versions introduces bottleneck blocks. These blocks use 1x1 convolutions to reduce and then expand the number of channels, reducing the computational burden while maintaining the model's representational capacity.
- **Modularity**: ResNet's modular structure makes it easy to stack layers and build deeper architectures without instability. The use of standard convolutional operations and well defined residual blocks simplifies optimization.

ResNet Applications Due to its robust structure and adaptability, ResNet has been used in various fields, including:

- **Image classification**: it has established new benchmarks for datasets such as ImageNet, achieving cutting-edge results at the time of its introduction.
- **Object Detection and Segmentation**: ResNet provides a solid foundation for popular object detection frameworks, such as Faster R-CNN and Mask R-CNN.

ResNet is also used for tasks such as medical imaging, by detecting diseases in radiographs and segmenting medical images. As in Geospatial Analysis, as shown in this study, it is adapted for regression tasks to predict earthquake epicentres, demonstrating its versatility beyond traditional classification tasks.

The introduction of ResNet marked a change which led to the development of later architectures such as DenseNet and ConvNeXt. Its influence extends beyond computer vision in areas such as natural language processing or speech recognition.

Models Used in this Study There are several variants of ResNet, each with its own characteristics and different depths, for the analysis of seismic events, two models were used, specifically:

- **ResNet-18**: A lightweight variant with 18 layers, making it computationally efficient and suitable for tasks with limited resources. It is a basic, reliable model for tasks with limited computational resources, it has been adapted for the regression activities performed during work to predict the epicentres of earthquakes.
- **ResNet-50**: A deeper variant with 50 layers organized in residual bottleneck blocks, has been used for tasks requiring greater representational capacity.

Reducing parameters through bottleneck blocks, the model is well suited for tasks requiring detailed feature extraction from high-dimensional data.

6.2 ConvNeXt

ConvNeXt is an advanced convolutional neural network (CNN) model, which aims to bridge the performance gap between traditional convolutional models and transformer-based architectures in computer vision. It adapts the principles of visual transformers (ViTs) to improve the efficiency and scalability of CNNs without compromising their simplicity or interpretability.

The ConvNeXt architecture includes several features that enhance its adaptability and performance:

- **Depthwise Convolutions**: ConvNeXt replaces standard convolutions with depthwise separable convolutions, reducing computational complexity but maintaining representativeness.
- **Layer Normalization**: Traditional batch normalization levels are replaced with the level normalization, which is more compatible with large-scale data processing.
- **Inverted bottleneck blocks**: As in the MobileNet architecture, inverted bottlenecks are used to improve feature extraction efficiency.
- **Improved downsampling**: ConvNeXt introduces new levels of downsampling that are simpler and more computationally effective than traditional pooling or creaky convolutions.
- **Expanded Kernel Sizes**: Larger convolutional nuclei are used to increase the receptive field, allowing the model to better capture input data.

The benchmark results indicate that ConvNeXt achieves cutting-edge performance in multiple image classification and object detection tasks, often competing with ViTs which require less computational resources.

ConvNeXt Applications The flexibility and enhanced performance of ConvNeXt make it suitable for various applications:

- **Image Classification**: ConvNeXt has demonstrated competitive results on benchmark datasets, such as ImageNet, establishing itself as a strong alternative to both traditional CNNs and ViTs.

- **Object Detection and Segmentation**: ConvNeXt serves as a robust backbone for object detection frameworks like Faster R-CNN and Mask R-CNN, achieving high accuracy with reduced computational overhead.
- **Geospatial Analysis**: ConvNeXt excels in tasks involving high-resolution spatial data, such as SAR imagery. In this study, it was adapted to predict the distance of SAR image patches from earthquake epicenters, leveraging its ability to extract detailed spatial features.

ConvNeXt bridges the gap between traditional CNNs and Vision Transformers, offering a versatile architecture that benefits from the strengths of both paradigms.

Model Used For the analysis of seismic events discussed in this project, the ConvNeXt model used is:

- **ConvNeXt-Tiny**: the smallest variant in the ConvNeXt family, is designed to balance performance and computational efficiency. It features depthwise separable convolutions and patch-based processing, making it well-suited for large-scale datasets. In this study, ConvNeXt-Tiny was fine-tuned for regression tasks to predict earthquake epicenters, demonstrating superior accuracy and efficiency compared to traditional CNNs.

ConvNeXt's modernized design and adaptability ensure its relevance across diverse domains, making it a valuable model in seismic data analysis and other similar applications.

6.3 Vision Transformer (ViT)

Vision Transformer (ViT) represents a change in deep learning for computer vision tasks. Unlike convolutional neural networks (CNN), which are based on localized convolutional operations, ViT applies transformer-based self-attention mechanisms originally developed for natural language processing to image data. The new approach allows long-range dependencies and global patterns to be captured in images, and is particularly effective for tasks requiring a holistic understanding of space structures.

The ViT architecture has several innovative features that distinguish it from convolutional models:

- **Patch Embedding**: instead of processing images pixel by pixel or through convolutions, ViT splits an image into patches of fixed size. Each patch is flattened into a vector, which acts as an input token for the transformer.

- **Positional Encoding**: to preserve spatial information within the transformer, position encoding is added to patch embeddings. This ensures that the model understands spatial relationships between patches, a crucial aspect of image analysis.
- **Self-preservation mechanism**: the core of ViT lies in its multi-head self-attention mechanism, which models the interactions between all patches in the image. This allows the model to capture local and global dependencies in the input data.
- **Layer Normalization**: Level normalization is applied instead of batch normalization, ensuring stability during training and improving performance for large-scale data sets.

ViT Applications The innovative design of ViT makes it a versatile tool for various computer vision tasks, including:

- **Image Classification**: ViT has set new benchmarks on large-scale image datasets like ImageNet, demonstrating its ability to outperform CNNs in recognizing diverse image categories.
- **Object Detection and Segmentation**: With its global self-attention mechanism, ViT is well-suited for detecting objects and segmenting regions within images, particularly in applications requiring high precision.
- **Geospatial Analysis**: ViT's capability to model global dependencies makes it ideal for geospatial tasks, such as analyzing satellite imagery and detecting patterns in seismic data. In this study, it was adapted to predict the distance of SAR image patches from earthquake epicenters, leveraging its advanced spatial representation capabilities.

Model Used The specific Vision Transformer model employed in this study is:

- **ViT-Patch16-224**: This variant of ViT divides images into 16x16 patches and processes them using a 224x224 resolution input. It features 12 transformer encoder layers and 768 hidden units, allowing it to effectively capture complex spatial patterns in high-resolution images. ViT-Base-Patch16-224 was fine-tuned for regression tasks in this study, demonstrating its ability to model global deformation patterns and predict earthquake epicenters with high accuracy.

Vision Transformer (ViT) represents a significant advancement in computer vision, combining the strengths of transformer-based architectures with the needs

of image analysis. Its adaptability and ability to handle large-scale, high-resolution data make it an invaluable tool for seismic data analysis and other complex applications.

6.4 Evaluation Metrics

To evaluate the performance of the models used in this study, three key regression metrics were employed. These metrics offer a comprehensive understanding of the model's accuracy and reliability in predicting distances or epicenter coordinates from patches to earthquake epicenters.

- **Mean Absolute Error (MAE)**: MAE measures the average absolute difference between the predicted and actual values, providing an interpretable metric for evaluating prediction accuracy. It directly reflects the average error magnitude, without considering the direction of the error.

$$MAE = \frac{1}{N}\sum_{i=1}^{N}|y_i - \hat{y}_i|$$

  Lower MAE values indicate better model performance, as they signify smaller average deviations from the true values.

- **Mean Squared Error (MSE)**: MSE quantifies the average squared difference between predicted and actual values. This metric is sensitive to large errors, as it penalizes significant deviations more heavily than smaller ones, making it a robust measure for detecting substantial inaccuracies.

$$MSE = \frac{1}{N}\sum_{i=1}^{N}(y_i - \hat{y}_i)^2$$

  MSE is particularly useful for assessing the overall accuracy of the model, as it emphasizes the need to minimize larger errors.

- **Root Mean Squared Error (RMSE)**: RMSE provides the square root of the MSE, translating the metric back into the same unit as the target variable. This makes RMSE an intuitive measure of prediction error, aligning with the scale of the actual distances.

$$RMSE = \sqrt{MSE}$$

  Lower RMSE values indicate that the model predictions are closer to the actual values, with fewer significant outliers.

Interpretation of Metrics Each metric provides unique insights into model performance:

- **MAE** highlights the average magnitude of errors, offering an easily interpretable measure of typical performance.
- **MSE** emphasizes larger errors, making it particularly useful for identifying models that consistently produce extreme deviations.
- **RMSE** combines the interpretability of MAE with the sensitivity of MSE to large errors, offering a balanced view of model accuracy.

By analyzing these metrics together, a comprehensive evaluation of the models' effectiveness in seismic data regression tasks was achieved, providing a differentiated understanding of their strengths and areas for improvement.

Chapter 7

Epicenter Detection

Seismic analysis is carried out by combining the data collected with selected machine learning models and then presenting them with regression tasks.

Distance : The first task to be solved and predicted is the distance from the patches to the seismic epicentres, which indicates an estimate of how far away or not the earthquake is.

Point : The second task tries to predict exactly the point of the epicenter, in terms of coordinates: latitude and longitude.

Attempting to make a prediction using machine learning models of the deformation of the soil relative to the epicenter of an earthquake, several critical aspects have been found. During the survey, an attempt is made to capture soil changes using interferograms and analyse them to obtain relevant information on how seismic activity has affected surrounding areas. The ultimate aim is to provide an estimate of the epicenter, in terms of coordinates or distances, to speed up the assessment of the magnitude of seismic effects.

7.1 Distance from Epicenter to Patch

The first task to be performed and tested was to predict the distance between a reference point and the epicenter, trying to minimize the error. To make an accurate estimate, it was necessary to understand the spatial relationship between the earthquake epicenter and the patch area, so the distance between the epicenter and the center of each patch was calculated using wgs84 coordinates.

To calculate the distance between the earthquake epicenter and the center of each patch was applied the *Haversine formula*. This formula computes the great-circle

distance between two points on the Earth's surface based on their latitude and longitude, accounting for the Earth's curvature.

The data obtained has been stored and saved in a new column of the csv file containing all the metadata related to each patch:

- **distance**: The computed distance between the epicenter and the patch center, serving as a key target variable for regression models.

The distances calculated in metres were crucial to analyse the spatial relationship between the earthquake epicenter and the patches, allowing the study of the affected areas with precision.

As for example, a random patch of Earthquake case study, about the subswath IW1 has the following relevant informations:

- **Patch Center**: Latitude: 33.5663, Longitude: 47.5934.
- **Epicenter**: Latitude: 33.7388, Longitude: 45.7501.
- **Calculated Distance**: 202279 m.

Plotting the distance and location of the patch as in the figure 7.1 gives an idea about the spatial location of the data.

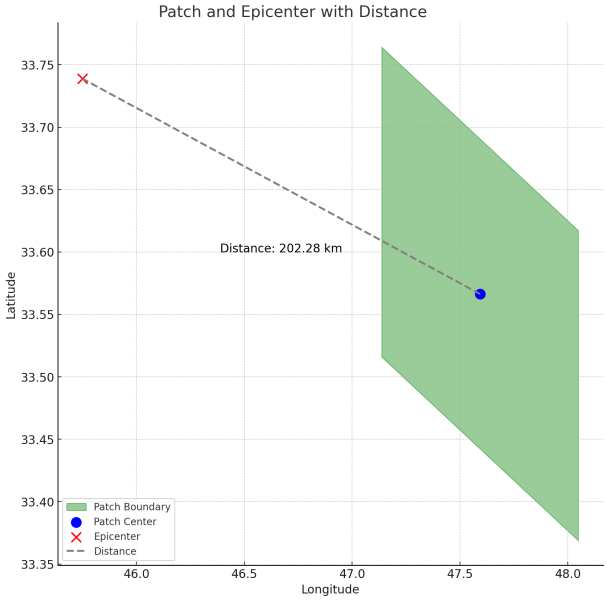

Figure 7.1: The spatial relationship between the patch and the earthquake epicenter, providing insight into their relative positioning.

The first trainings have obtained poor results obtained, as shown in the table **??**, so, it was considered appropriate to introduce the coordinates related to the center of the patch. Using coordinates as input aids for the models, attempts have been made to make further attempts with the aim of minimising error.

The initial idea was to carry out an analysis by trying to reduce the calculated distance, so in addition to exploiting a spatial knowledge given by latitude and longitude, it was normalized and calculated the distance based on the normalized coordinates remaining in the range [0,1], so that you have more manageable values. The new distances was calulated with the formula of the *distance euclidean* and inserted in a new column called:

- **normalized_distance**: The new distance between the epicenter and the patch center acts as a new target variable.

Given the results obtained, not yet satisfactory, placed in the table 7.2 a further attempt was made, in which the coordinates of the center of the patch are inserted as input, acting as additional channels.

In this case the coordinates **center_lat** and **center_lon** relative to that patch and distance, calculated previously, were used and the best results obtained.

For each machine learning model some modifications and optimizations have been made, changing hyperparameters and loss function.

7.1.1 Models

The study used each of the machine learning models selected above to predict the distance from the epicenter, using SAR patch data stored in the HDF5 file. The training process used textttPyTorch and the textttimm library to instantiate and configure pre-trained models. They have been modified to test various parameters such as batch size, learning rate, optimizer and lossing functions in order to improve regression tasks.

The input data, for each model, consisted of SAR image patches with 3 channels:

- **Coherence** (`coh`)
- **Real** (`i`)
- **Imaginary** (`q`)

For subsequent attempts, the following data have also been added like extra channels:

- **Normalized Center Latitude** (`center_lat`)
- **Normalized Center Longitude** (`center_lon`)

Therefore latitude and longitude values were normalized, if necessary, to the range $[0, 1]$ using:

$$\text{Normalized Latitude} = \frac{\text{center\_lat} + 90}{180}$$

$$\text{Normalized Longitude} = \frac{\text{center\_lon} + 180}{360}$$

This normalization ensured consistent scaling across geographic coordinates and predicted distance.

ResNet

Residual networks (ResNets) emerged as a revolutionary architecture in deep learning and have been used to analyze complex data such as spatial and spectral information embedded in synthetic aperture radar (SAR) data, making it a suitable choice for performing regression tasks such as predicting the distance from the epicenter of an earthquake.

ResNet-18 ResNet-18 has been made adaptable for regression tasks, to obtain the forecast of the distance from the center of the patch to the epicentres of relative earthquakes. By incorporating residual connections, the model addresses the problem of the gradient of disappearance, allowing deeper networks to learn complex representations of characteristics effectively.

The input data for ResNet-18 consisted of patches of SAR images represented as three-channel tensors. These channels correspond to the coh, real and imaginary coherence components of the SAR data, providing a complete view of the phase and amplitude variations in the images.

Initially the results obtained were very high, although approximating distances in the order of kilometers, the model could not predict with an acceptable error, then the parameters were changed and the target variables were approximated to 100 km.

In order to improve the performance of the model reported on the table 7.4, it was decided to add spatial information of latitude and longitude relative to the center of the patch.

In fact, each patch was coupled to its geographic coordinates of the center of the patch (latitude and longitude) and were inserted as additional input data to give a spatial structure to the model, so as not to predict a point without any reference.

Variable targets are the coordinates representing the epicenter of the earthquake and serve as output for the regression task. The inclusion of spatial data ensures that the model learns the relationships between input characteristics and the geographical distribution of seismic activity.

Subsequently, normalization was applied to the input data to obtain lower results : `center_lat`, `center_lon` were scaled to the range [**0, 1**].

The training process works with 20 initial periods and then goes on to define 30 as the best number in terms of losses. This model is best suited with a low learning rate of 0.001, while the loss functions used are: one for training `CombinedLoss` and one for testing `MSEloss`. The model has been optimized using `AdamW` with weight decay to avoid overweight, and with a scheduler `ExponentialLR`.

The Evaluation Metrics used to test the performance of ResNet-18 to evaluated regression task :

- **Mean Absolute Error (MAE)**: Measures the average absolute difference between predicted and actual values, providing an interpretable metric for real-world distances.
- **Mean Squared Error (MSE)**: Quantifies the average squared differences between predictions and ground truth, penalizing larger errors more heavily.
- **Root Mean Squared Error (RMSE)**: Translates the MSE back into the same units as the target variable, offering an intuitive measure of typical errors.

The ResNet18 model has achieved high computational accuracy and efficiency, demonstrating its suitability for large-scale geospatial regression tasks such as predict the ditance from earthquake epicentres.

This is why, for this model, the R^2 *Score* was also evaluated to evaluate the results obtained and therefore which model provides a better compliance with those subsequently obtained with ViT and ConvNeXt.

ResNet-50 ResNet50 was used because it is adaptable for tasks with several features and represents a more advanced variant of the ResNet family. The input data used are SAR images represented as three-channel tensors, corresponding to the real and imaginary coherence components of the SAR data. These channels provided a detailed representation of phase and amplitude variations.

To improve the unsatisfactory initial results, several modifications were made, following also those used for ResNet18, such as the number of eras and the loss function, in order to have a discrete error rounded to hundreds of km.

Later, additional spatial data were introduced, the geographic coordinates of the center of the area (latitude and longitude) as additional inputs, incorporating the spatial context directly into the model and applying normalization.

ResNet-50 was trained over 30 epochs using a batch of 16, with checkpointing implemented every 5 epochs to preserve progress in the event of interruptions. The training process leveraged the `CombinedLoss` function, which integrates mean

square error (MSE) with smooth L1 loss to balance sensitivity to large errors and robustness to outliers. It was optimized using `AdamW` and learning speed adjustments were handled through a `ExponentialLR` scheduler.

To evaluate performance, the regression metrics employed:

- **Mean Absolute Error (MAE)**: Captures the average absolute deviation between predicted and actual values, offering a straightforward interpretation of real-world accuracy.
- **Mean Squared Error (MSE)**: Penalizes larger deviations more heavily, serving as the primary metric for optimization.
- **Root Mean Squared Error (RMSE)**: Converts MSE back to the original units, providing an intuitive measure of typical prediction errors.

This model report the worst results despite of the good ones about ResNet18, but it was not entirely unexpected, as it is a significantly deeper model compared to the previous model.

ConvNeXt

In this study, ConvNeXt was adapted to the regression task, leveraging its efficient processing capabilities and modernized convolutional operations to handle the SAR image patches.

Input data for ConvNeXt, as for ResNet, consisted of patches of SAR images represented as three-channel tensors, corresponding to the coherence, real and imaginary components of the SAR data. These channels have provided a detailed representation of the phase and amplitude variations, which are essential for accurately predicting the epicentres of earthquakes.

The model used is ConvNeXt-Tiny, which is the smallest of the ConvNeXt family that combines CNN efficiency with Vision transformers to improve performance on geospatial data.

Convnext-tiny The ConvNeXt-Tiny architecture was used to predict distance from epicentres by analyzing SAR data patches.

The geographic coordinates of the center of the zone (latitude and longitude) were added as additional inputs and iniseme to the target variables, latitude and longitude of the epicenter, were normalized to the range **[-1, 1]** to ensure consistency with normalised input characteristics and facilitate the optimisation of the stable model. Similarly, SAR channels have been normalized, contributing to numerical stability during training.

ConvNext has been trained on 30 epochs and a batch size equal to 16. The training process used the `MSELoss` to balance sensitivity to large errors and robustness to outliers. The model was optimized using `AdamW` with weight decay and learning speed adjustments were handled through a `ExponentialLR` scheduler and a learning rate equal to 0.00001.

The following regression metrics were used to measure the model's performances:

- **Mean Absolute Error (MAE)**: Represents the average magnitude of prediction errors without considering their direction.
- **Mean Squared Error (MSE)**: Highlights larger errors by squaring the deviations from actual values.
- **Root Mean Squared Error (RMSE)**: Offers an error metric that aligns with the original scale of the target variable.

The ConvNeXt-Tiny model has achieved high computational accuracy and efficiency, demonstrating its suitability for large-scale geospatial regression tasks such as localizing earthquake epicentres in terms of distance.

The results obtained represents the error rounded about 100 kilometers of difference from real data,demonstrating its suitability for large-scale geospatial regression tasks such as localizing earthquake epicentres.

Also for this model, the R^2 *Score* was evaluated to compare the results obtained and to decide which model predict better between the others ViT and ResNet18.

ViT

Vision Transformer is not based on gyres, it divides images into pieces and processes them as tokens. This allows ViT to capture long-range dependencies in images, making it particularly suitable for complex tasks.
The ViT-Base-Patch16-224 model was used in this analysis to perform regression tasks.

Vit-patch16-224 The ViT-Base-Patch16-224 has a transformer-based architecture for computer vision tasks and is equipped with 3 main input channels: coherence, real, imaginary. The target variables correspond to the distances calculated above. The results obtained, as can be seen from the ref table, are not optimal if the parameters of the model have been changed accordingly.

In subsequent training, the normalized latitude and normalized longitude were also added. in the textbf[0, 1] range, corresponding to the center of each patch and the normalized distance is used as target variable. The SAR channels were also normalized, contributing to numerical stability during training, but the final output was not completely exhaustive.

A further attempt was made with the input coordinates and distances used as target variables without normalization, obtaining reasonable results, rounding up the distance in the order of hundreds of kilometers.

The training process used 30 epochs and a learining rate equals 0.0001 and a function loss `MSELoss`.

Vit model performance was evaluated using:

- **Mean Absolute Error (MAE)**: Calculates the average size of the differences between predicted and actual values.
- **Mean Squared Error (MSE)**: Emphasizes larger errors by squaring the differences to amplify their impact.
- **Root Mean Squared Error (RMSE)**: Measures error in the same units as the target variable for easier interpretation.

The model has reported computational accuracy and efficiency, obtaining lower errors than models tested in previous periods, proving to be remarkably adaptable to predict the distance from the epicentres of earthquakes.

7.1.2 Results

Each model's performance was evaluated based on its accuracy in predicting distances to earthquake epicenters, as well as computational efficiency. Results were compared to identify the best-performing architecture for the regression task. Key metrics included Mean Squared Error (MSE), Mean Absolute Error (MAE), and Root Mean Squared Error (RMSE), which provided insight into each model's precision.

Model	**MAE**	**MSE**	**RMSE**
ResNet18	16.1775	1756.4472	41.9631
ResNet50	13.6538	1860.0637	43.1951
ConvNeXt	17.1714	1854.4116	43.1309
ViT	17.1115	1854.4229	43.1311

Table 7.1: Performance comparison of models with distance predictions

Further optimizations were implemented, including normalization of distances and hyperparameter tuning. Models were trained with normalized distances in the range [0,1]. The optimizations included combining loss functions and applying data

augmentation (excluding spatial transformations). Results after 30 epochs, with learning rates tuned for each model, are summarized below.

Model	**MAE**	**MSE**	**RMSE**
ResNet18	0.0627	0.0160	0.1263
ResNet50	0.0339	0.0111	0.1052
ConvNeXt	0.0333	0.0108	0.1037
ViT	0.0338	0.0089	0.0942

Table 7.2: Models performance with normalized coordinates and distance

The results obtained can not be compared with the others beacuse was used a different metrics and approach to minimize the error, but in terms of real distance it was necessary do another training without normalization.

Finally, the models were re-evaluated using the center coordinates of the patches as inputs, with distances rounded to approximately 100 km. These modifications aimed to further refine the performance by addressing spatial variability in the data.

Model	**MAE**	**MSE**	**RMSE**
ResNet18	7.7456	824.4542	28.7063
ResNet50	10.3901	1211.0259	34.8051
ConvNeXt	7.0103	1114.1005	33.3781
ViT	6.2251	990.5480	31.4730

Table 7.3: Models performance with center patch coordinates

Best Model

The best performing model was selected based on its ability to reduce prediction error consistently across the test set. Considerations for this task took into account various factors and R^2 score was also introduced. The selected model demonstrated a strong balance between accuracy and computational feasibility, providing a reliable basis for further applications in seismic analysis.

Table 7.4 presents the performance of models on the distance prediction task, with scaled distances of 100 km. The results provide information on the predictive

Model	**MAE**	**MSE**	**RMSE**	R^2
ResNet18	7.7456	824.4542	28.7063	0.5570
ConvNeXt	7.0103	1114.1005	33.3781	0.4011
ViT	6.2251	990.5480	31.4730	0.4675

Table 7.4: Performance comparison of models with center patch coordinates

accuracy of each model and the ability to explain variance in data.

The ViT model demonstrated the best overall accuracy, with the lowest MAE and RMSE among all models, making it the most reliable for distance prediction tasks. Its slightly lower R^2 compared to ResNet18 indicates room for improvement in explaining the variance in the data, but, as the image shows in 7.2.

While ResNet18 achieved the highest R^2 score, indicating strong ability to capture variance, its higher MAE and RMSE values suggest a large discrepancy in the forecasts compared to ViT. The figure 7.2 displays significant fluctuations in validation MAE that indicate instability and possible overfitting at certain epochs.

Moreover, ConvNeXt served as a balanced performer, with competitive MAE but a lower R^2 score compared to ResNet18, showing potential for further refinement.

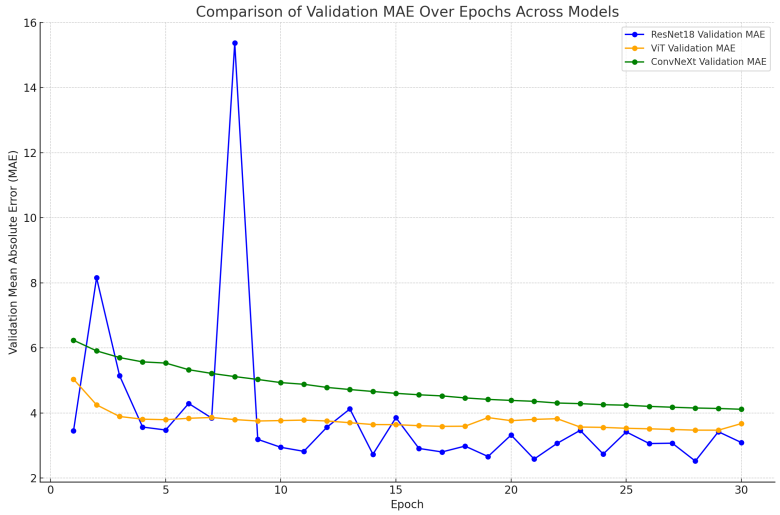

Figure 7.2: MAE of different model to choose the best

Overall, the predicted distances were scaled by 100 km, making the errors relatively large compared to the expected distances. This highlights the models' challenges in achieving fine-grained accuracy, particularly for smaller distances and the final results demonstrate that ViT is the most accurate model for distance

prediction, offering the best balance of low prediction errors and moderate variance explanation.

7.2 Point of the Epicenter

The second problem to be solved by exploiting neural networks is to focus on the precise location of the epicenter of a given earthquake.

The final goal is to have a forecast of its coordinates, hence latitude and longitude of the point, which is why the coordinates previously known as *"latitude"* and *"longitude"* are added to the csv file so that they have a match between all the patches relaitive to that earthquake and its epicenter. These coordinates are fundamental because they form the ground truth on which to rely and serve as a reference for monitoring whether the model is adaptable or not.

The coordinates have been added in two new columns, each row containing the epicenter of that specific earthquake:

- **epicenter_lat**: The latitude of the earthquake epicenter.
- **epicenter_lon**: The longitude of the earthquake epicenter.

To observe all the data obtained and check their position in relation to each patch and each subswath, they were plotted together, also going to locate the epicenter.

The geographical distribution of the zones of a selected earthquake is shown in figure 7.3, it provides a visual context for the data set.

7.2.1 Models

Starting with the machine learning models selected above, each model has been optimized to predict the coordinates of the epicenter, using SAR patch data stored in the HDF5 file. The training process used `PyTorch` and the `timm` library to instantiate and configure pre-trained models. They have been modified to test the various parameters, such as batch size, learning rate, and optimizer in order to improve regression tasks.

The final input data, for each model, consisted of SAR image patches with five channels:

- **Coherence** (`coh`)
- **Real** (`i`)
- **Imaginary** (`q`)

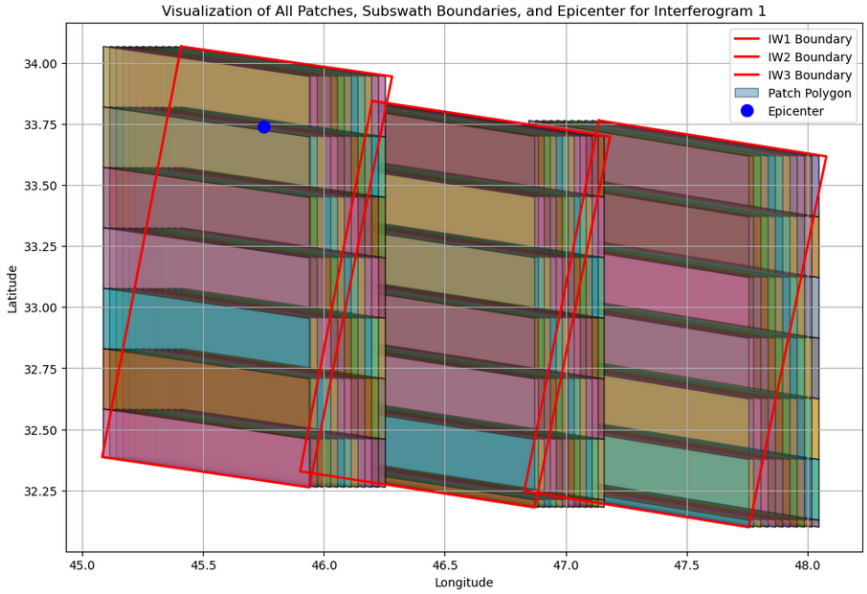

Figure 7.3: Patches resized related to the case study Earthquake.

- **Normalized Center Latitude** (`center_lat`)
- **Normalized Center Longitude** (`center_lon`)

This time the latitude and longitude values were normalized to the range $[-1, 1]$ using the formula:

$$\text{Normalized Latitude} = \frac{2(\text{center\_lat} + 90)}{180} - 1$$

$$\text{Normalized Longitude} = \frac{2(\text{center\_lon} + 180)}{360} - 1$$

This normalization ensured consistent scaling across geographic coordinates.

ResNet

Residual networks (ResNets) emerged as a revolutionary architecture in deep learning and have been used to analyze complex data such as spatial and spectral information embedded in synthetic aperture radar (SAR) data, making it a suitable choice for performing regression tasks such as predicting the epicenter of an earthquake.

ResNet-18 ResNet-18 has been made adaptable for regression tasks, to obtain the forecast of the coordinates of the epicentres of earthquakes. By incorporating residual connections, the model addresses the problem of the gradient of disappearance, allowing deeper networks to learn complex representations of characteristics effectively.

The input data for ResNet-18 consisted of patches of SAR images represented as three-channel tensors. These channels correspond to the coh, real and imaginary coherence components of the SAR data, providing a complete view of the phase and amplitude variations in the images.

Initially the results obtained, although modifications were made to the hyperparameters, were very poor and not satisfactory.

In fact, each patch was coupled to its geographic coordinates of the center of the patch (latitude and longitude) and were inserted as additional input data to give a spatial structure to the model, so as not to predict a point without any reference.

Variable targets are the coordinates representing the epicenter of the earthquake and serve as output for the regression task. The inclusion of spatial data ensures that the model learns the relationships between input characteristics and the geographical distribution of seismic activity.

Subsequently, normalization was applied to the input data to ensure numerical stability and accelerate convergence during training and SAR channels were normalized. At the same time, normalization was applied to target variables, the latitude and longitude values of the epicenter were resized to maintain consistency with the input characteristics. Normalized coordinates for the center and epicenter (`center_lat`, `center_lon`, `epicenter_lat`, `epicenter_lon`) were scaled to the range **[-1, 1]**.

The training process was conducted over several epochs, typically 20 to 30, with batch size of 16 samples and with control points saved every 5 epochs to safeguard against potential interruptions. The loss functions used, one for training `CombinedLoss` while the mean square error (MSE) was used as primary metric to guide model optimization. The model was optimized using `AdamW` with weight decay to avoid overweight, and the learning speed was adjusted dynamically using a `ExponentialLR` scheduler.

The Evaluation Metrics used to test the performance of ResNet-18 to evaluated regression task :

- **Mean Absolute Error (MAE)**: Measures the average absolute difference between predicted and actual values, providing an interpretable metric for real-world distances.

- **Mean Squared Error (MSE)**: Quantifies the average squared differences between predictions and ground truth, penalizing larger errors more heavily.

- **Root Mean Squared Error (RMSE)**: Translates the MSE back into the same units as the target variable, offering an intuitive measure of typical errors.

ResNet-18 demonstrated its effectiveness in leveraging the coherence, real, and imaginary components of SAR data to predict the latitude and longitude of earthquake epicenters. Its lightweight architecture made it computationally efficient while obtaining an high error. This adaptability highlights a baseline for further exploration and refinement.

ResNet-50 ResNet-50 was used as a deeper variant of the ResNet family. The input data used are SAR image patches represented as three-channel tensors, corresponding to the coherence, real and imaginary components of the SAR data. These channels provided a detailed representation of phase and amplitude variations.

To improve on the unsatisfactory initial results, additional spatial data has been incorporated. Geographic coordinates of the centre of the area (latitude and longitude) were added as additional inputs, incorporating spatial context directly into the model and ensuring that forecasts were based on geographical relevance.

Then, normalization was applied to the input data and to target variables, like for ResNet18, to maintain consistency with the input characteristics.

ResNet-50 was trained over 20-30 epochs using a batch of 16, with checkpointing implemented every 5 epoch to preserve progress in the event of interruptions. The training process leveraged the `CombinedLoss` function, which integrates mean square error (MSE) with smooth L1 loss to balance sensitivity to large errors and robustness to outliers. The model was optimized using `AdamW` with weight decay and learning speed adjustments were handled through a `ExponentialLR` scheduler.

To evaluate performance, the following regression metrics were employed:

- **Mean Absolute Error (MAE)**: Captures the average absolute deviation between predicted and actual values, offering a straightforward interpretation of real-world accuracy.

- **Mean Squared Error (MSE)**: Penalizes larger deviations more heavily, serving as the primary metric for optimization.

- **Root Mean Squared Error (RMSE)**: Converts MSE back to the original units, providing an intuitive measure of typical prediction errors.

Its use of choke blocks and improved depth provided a balance between computational efficiency and performance, but did not lead to very reliable or noteworthy results for high-dimensional regression tasks, in particular to predict the coordinates of the epicenter.

ConvNeXt

In this study, ConvNeXt was adapted to the regression task, leveraging its efficient processing capabilities and modernized convolutional operations to handle the SAR image patches.

ConvNeXt's architecture allowed for effective feature extraction across spatial scales, providing advantages in handling high-resolution SAR data.

The model used is ConvNeXt-Tiny, which is the smallest of the ConvNeXt family that was inspired by Vision transformers, combining CNN efficiency with advances in transformer style to improve performance on geospatial data.

Convnext-tiny The ConvNeXt-Tiny architecture was used to predict earthquake epicentres by analyzing SAR data patches with five input channels: coherence, real, imaginary, normalized latitude and normalized longitude. The model focuses on scalability, computational efficiency and advanced feature extraction.

The geographic coordinates of the center of the zone (latitude and longitude) were added as additional inputs and iniseme to the target variables, latitude and longitude of the epicenter, were normalized to the range **[-1, 1]** to ensure consistency with normalised input characteristics and facilitate the optimisation of the stable model. Similarly, SAR channels have been normalized, contributing to numerical stability during training.

ConvNext has been trained on 20-30 epochs using a batch of 16. The training process used the `CombinedLoss` like loss function, which integrates mean square error (MSE) to balance sensitivity to large errors and robustness to outliers. The model was optimized using `AdamW` with weight decay and learning speed adjustments were handled through a `ExponentialLR` scheduler and a learning rate equal to 0.00001.

The following regression metrics were used to assess performance:

- **Mean Absolute Error (MAE)**: Represents the average magnitude of prediction errors without considering their direction.
- **Mean Squared Error (MSE)**: Highlights larger errors by squaring the deviations from actual values.
- **Root Mean Squared Error (RMSE)**: Offers an error metric that aligns with the original scale of the target variable.

Ground-based forecasts and truth values are de-normalized to real world coordinates for evaluation. The metrics are calculated for the test data set, while the predictions are saved in a CSV file along with true values for further analysis.

The ConvNeXt-Tiny model has achieved high computational accuracy and efficiency, demonstrating its suitability for large-scale geospatial regression tasks such as localizing earthquake epicentres.

This is why, for this model, the R^2 *Score* was also evaluated because reflects how well the model's predictions explain the variance in the actual data and is essential to evaluate the results obtained and therefore which model provides a better compliance with those subsequently obtained with ViT.

ViT

Vision Transformer (ViT) introduces a novel approach by applying transformer-based attention mechanisms to image data. Rather than relying on convolutions, ViT divides images into patches and processes them as tokens, similar to natural language processing models. This enables ViT to capture long-range dependencies within the images, making it particularly suited for complex spatial tasks.

The ViT-Base-Patch16-224 model was used in this analysis and has been initialized using the `timm` library with pre-trained weights on ImageNet.

Vit-patch16-224 The ViT-Base-Patch16-224 has a transformer-based architecture for computer vision tasks and is adapted to 3 main input channels: coherence, real, imaginary, To which are also added the normalized latitude and normalized longitude relative to the center of each patch to which the refill is made.

The geographic coordinates of the epicenter (latitude and longitude) were taken as target variables and all coordinates were normalized in the textbf range[-1, 1] to ensure consistency with the input characteristics and facilitate the optimization of the stable model. Similarly, SAR channels have been normalized, contributing to numerical stability during training.

The ViT model was trained over 20-30 ages using a net value of 0.0001. The training process used the `CombinedLoss` function, which integrates the mean square error (MSE) to balance sensitivity to large errors and robustness to outliers. The model was optimized using `AdamW` with weight decay and learning speed adjustments were handled through a `ExponentialLR` scheduler.

Model performance was evaluated using:

- **Mean Absolute Error (MAE)**: Calculates the average size of the differences between predicted and actual values.
- **Mean Squared Error (MSE)**: Emphasizes larger errors by squaring the differences to amplify their impact.
- **Root Mean Squared Error (RMSE)**: Measures error in the same units as the target variable for easier interpretation.

The model has achieved high computational accuracy and efficiency, obtaining lower predictions and metrics than models tested in previous periods, showing to

be remarkably adaptable to the location of earthquake epicentres. The forecasts have been saved next to the soil truth values for further analysis.

The ViT's ability to model global dependencies has proved beneficial in accurately predicting earthquake epicentres, making it a valuable tool for seismic data analysis. The model's attention mechanism allowed for focusing on the relevant areas within image patches, potentially improving the accuracy of point coordinate predictions.

7.2.2 Results

Each model's performance was evaluated based on accuracy in predicting distances to earthquake epicenters, as well as computational efficiency. Results were compared to identify the best-performing architecture for the regression task at hand. Key metrics included Mean Squared Error (MSE) and Mean Absolute Error (MAE), which provided insight into each model's precision in the regression task.

The best-performing model was identified based on its ability to minimize error consistently across the test set, along with considerations for computational efficiency.

The initial experiments, which excluded the coordinates of the center patch, resulted in poor performance across all models. Subsequently, normalized patch coordinates were incorporated as input features, significantly improving the models' ability to predict epicenter distances.

The models were evaluated on the test set using denormalized epicenter coordinates. Table 7.5 presents the performance of all tested models:

Model	**MAE**	**MSE**	**RMSE**
ResNet-18	55.5475	6524.7121	80.7757
ResNet-50	56.4612	6200.3866	78.7425
ConvNeXt-Tiny	28.1063	3718.7980	60.9819
ViT-Patch16-224	16.5482	997.5620	31.5842

Table 7.5: Performance comparison of models on epicenter localization.

From these results, it is evident that ConvNeXt-Tiny and ViT-Patch16-224 outperformed ResNet-18 and ResNet-50, with ViT-Patch16-224 achieving the lowest errors across all metrics.

Best Model Analysis

To further refine the evaluation, only the two best-performing models, ConvNeXt-Tiny and ViT-Patch16-224, were analyzed using the R^2 score as an additional metric. The updated comparison is shown in Table 7.6:

Model	**MAE**	**MSE**	**RMSE**	**R^2**
ConvNeXt-Tiny	28.1063	3718.7980	60.9819	0.3701
ViT-Patch16-224	16.5482	997.5620	31.5842	0.6482

Table 7.6: Performance comparison of best models.

ViT-Patch16-224 demonstrated the best overall performance, achieving the lowest MSE, RMSE, and MAE, along with the highest R^2 score. These results establish it as the most effective model for predicting distances to earthquake epicenters.

ConvNeXt-Tiny also performed well, particularly in terms of MAE, but in terms of error values it is less accurate than ViT-Patch16-224.

Generally, coordinate integration has proved essential to improve the performance of each model, highlighting the importance of spatial information in this regression task.

7.3 Overall Results and Model Evaluation

The evaluation of multiple deep learning models focused on their predictive performance in seismic analysis activities, specifically to locate the epicenter. The performance of all models was quantified using the mean absolute error (MAE), mean square error (MSE), mean square root error (RMSE) and R^2 score to have a final comparison on each task produced.

7.3.1 Task 1: Distance Prediction

The results highlight significant differences among the models in their ability to predict distances to earthquake epicenters, as the table 7.4 reports:

- **ViT** consistently achieved the best overall performance, with the lowest MAE, MSE, and RMSE across all configurations. It demonstrated strong generalization capabilities and was particularly effective in scenarios where spatial variability was addressed.

- **ResNet18** showed the highest R^2 score, indicating superior ability to explain variance in the data. However, its MAE and RMSE were higher than those of ViT, suggesting less precise predictions overall.

- **ConvNeXt** performed moderately, striking a balance between prediction accuracy and variance explanation. While its errors were higher than ViT's, it provided consistent results across the test set.

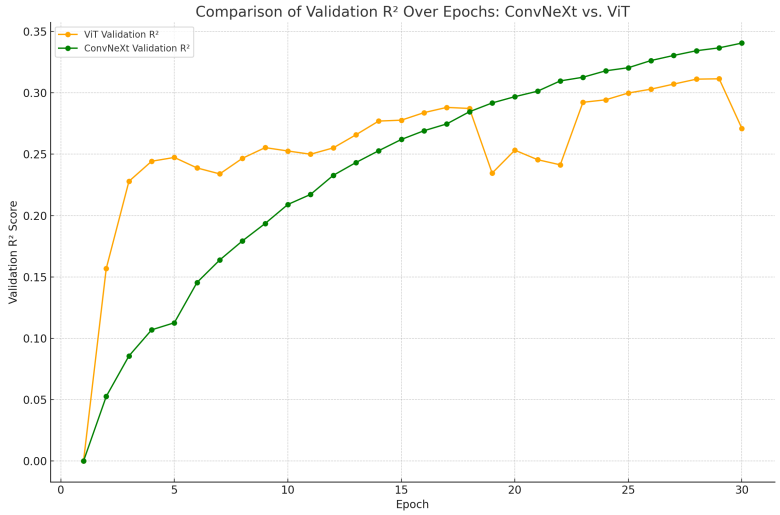

Figure 7.4: Comparison of Validation R^2 Scores Between ConvNeXt and ViT

The figure 7.4 illustrates the progression of R^2 scores across epochs for ConvNeXt and ViT. ViT consistently achieves higher R^2 scores, reflecting its superior ability to explain variance in the validation data. ConvNeXt demonstrates stable R^2 performance but remains consistently below ViT, underscoring the latter's advantage in capturing complex data patterns and generalizing effectively to unseen data.

Overall, ViT emerged as the most reliable model for the distance prediction task, achieving a balance between accuracy and computational efficiency. The observed differences between the true and predicted distances ranged from approximately tens of metres to a maximum of thousands of km. These findings underscore the challenges of distance prediction in seismic analysis and the importance of selecting appropriate models for such tasks.

7.3.2 Task 2: Epicenter Localization

For the task of epicenter localization, the models were evaluated based on their ability to predict the precise coordinates of earthquake epicenters. Initial experiments, conducted without spatial context, revealed poor performance across all models. This highlighted the necessity of incorporating spatial information to improve localization accuracy. When normalized patch coordinates were introduced as input features, all models demonstrated substantial improvements, leveraging the additional spatial context to better capture positional relationships within the data.

Among the models, **ViT** delivered the most accurate predictions, achieving the lowest MAE, MSE, and RMSE. It consistently outperformed other models in aligning predictions with true epicenter locations. **ConvNeXt** also demonstrated competitive performance, particularly in reducing MAE; however, its higher overall errors and lower R^2 score indicated slightly lower precision compared to ViT.

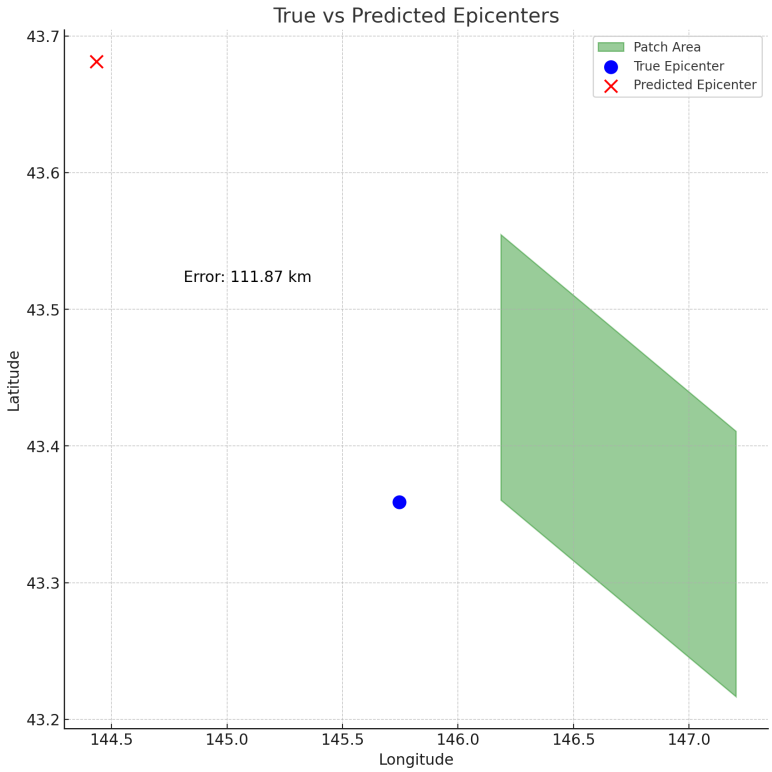

Figure 7.5: Comparison of true and predicted epicenters for the patch of the case-study earthquake.

As illustrated in Figure 7.5, the ViT model predicted an epicenter location (red cross) that, despite being offset from the true epicenter (blue dot), remains within

the general vicinity of the patch boundary (green polygon). The prediction error for this case was approximately 111.87 km, demonstrating the challenges of precise localization.

The inclusion of normalized patch coordinates proved crucial in improving model performance. Normalization ensured consistent scaling of coordinates, reducing biases caused by variations in magnitude and enabling the models to better understand spatial relationships within the data. These findings highlight the importance of spatial preprocessing in enhancing the accuracy of epicenter localization tasks.

On the whole, the ViT model was considered to be the most effective for localizing epicentres, demonstrating robustness and superior generalization across various cases. ConvNeXt provided reasonable results but fell short in terms of overall precision compared to ViT, as shown as in the plot below 7.6

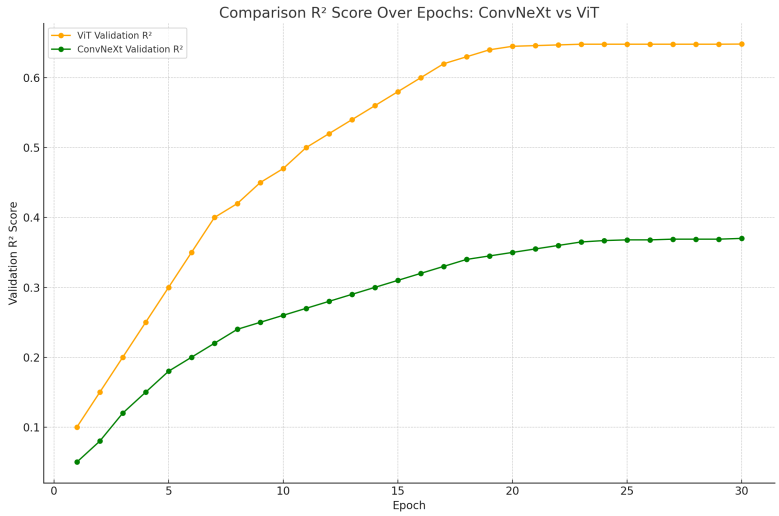

Figure 7.6: Comparison of R2 score.

Chapter 8

Conclusion

This thesis addressed the problem of using machine learning models to predict earthquake epicentres and distances using data from the Sentinel-1 synthetic aperture radar (InSAR). Through the application of advanced deep learning architectures and accurate data preprocessing, we have developed a new approach to seismic analysis, demonstrating the potential of these methods to integrate traditional seismological studies.

The research work carried out is structured in such a way as to observe all the various stages of the analysis. From outlining the motivations for using InSAR data and machine learning in seismic analysis, to highlighting the need for accurate and efficient methods to predict earthquake epicentres.

Subsequently, the Sentinel-1 data were prepared in a timely manner to generate the final interferograms from the SLC images and finally resize and extract the patches and metadata as samples of the final dataset.

Special attention goes to the experimental part, where after selecting the various models of deep learning, ResNet, ConvNeXt and Vision Transformer (ViT), were evaluated for regression tasks aimed at predicting distances and coordinates relative to the earthquake epicenter.

Among these, the ViT model has proved to be the most efficient, demonstrating superior performance in capturing complex space models within SAR data. Its transformer-based architecture, which processes images as patch sequences, has made it possible to model long range dependencies effectively, The results are more accurate and generalizable than traditional CNN-based architectures.

ConvNeXt, on the other hand, provided a solid performance base. Although not achieving the same level of precision as ViT, its design, inspired by modern transformer architectures and preserving the simplicity of convolving networks, offered a robust and computationally efficient alternative. Both models have significantly outperformed ResNet, particularly in handling the complex and high-size characteristics of SAR data, Although each has shown unique strengths and

weaknesses in terms of noise sensitivity and generalisation through different seismic events.

Despite these promising results, some gaps and problems remain due to the complexity of the data.

The main difficulty of seismic analysis was the generation of interferograms which, being computationally expensive, often took several hours even for a single image. This considerable time-consuming process has limited the speed of testing and iterating models, and also limited the data obtained and used in the final dataset.

Data magnification techniques have been carefully applied, but the lack of sufficient diversity in data remains a limiting factor. In addition, SAR data, due to their high dimensionality, contain complex spatial information, making it difficult to capture significant features.

Based on the results of this thesis, various avenues for future research and improvement are proposed.

Firstly, the expansion of the data set with additional images from Sentinel-1, with more interferograms or with noise minimisation of these, could improve the robustness and generalizability of the model. In addition, adequate resources should be available to speed up the workflow and make it more accessible by improving patch extraction speeds and enabling faster testing.

This thesis represents a significant step toward the application of machine learning in seismic analysis, demonstrating the potential of combining deep learning with traditional methodologies to predict earthquake epicenters. Despite challenges related to time constraints, data limitations, and the complexity of SAR images, the models developed in this study achieved encouraging results, with ViT standing out as the most effective approach.

The work highlights the importance of interdisciplinary approaches in addressing seismic challenges, laying a foundation for future research to build upon. By refining these methods and addressing the limitations outlined, future studies can further enhance seismic analysis and contribute to mitigating the risks associated with earthquakes.

Acknowledgements

I would like to thank all the people who have accompanied and supported me during this course of studies, and without whom this goal would not have been possible. First of all, I thank my rapporteur, Prof. Paolo Garza and my co-rapporteur Daniele Rege Cambrin, for their availability, patience and valuable guidance. Thanks to their advice and experience, I have been able to overcome many difficulties and carry out this research with determination and a critical spirit. I have to thank HPC@POLITO and SMARTDATA@POLITO for the provision of computational resources on which to work this project.

I thank my family, who even from afar has always encouraged me and pushed me to go forward, despite the difficulties and the unforeseen every day.
I thank my uncles, my second family, on whom I could count in every single moment, from the most banal to the most critical.
Finally, I thank the one who always believed in me, since childhood, my grandfather.

Thanks to those who are no longer here. To those who in a short time has become a fundamental part of my life, to all those who have always been by my side, especially at this strange moment of my life and to those who have managed to make my days less heavy or to those who, without knowing it, made this path less steep. To those who reminded me that it takes very little to be happy.

Bibliography

[1] Daniele Rege Cambrin and Paolo Garza. «QuakeSet: A Dataset and Low-Resource Models to Monitor Earthquakes through Sentinel-1». In: *Proceedings of the 21st ISCRAM Conference* (2024) (cit. on pp. 4, 5, 12).

[2] Weiqiang Zhu and Gregory C. Beroza. «PhaseNet: A Deep-Neural-Network-Based Seismic Arrival Time Picking Method». In: *Geophysical Journal International* XX (2018), pp. XX–XX (cit. on pp. 4, 6).

[3] José Augusto Proença Maia Devienne. «Convolutional neural network for earthquake detection». In: *Big Data to Earth Scientists* (2020) (cit. on pp. 4, 7, 9).

[4] Alessandro Ferretti, Andrea Monti-Guarnieri, Claudio Prati, Fabio Rocca, and Didier Massonnet. *InSAR Principles: Guidelines for SAR Interferometry Processing and Interpretation*. European Space Agency. 2007. URL: `https://www.esa.int/esapub/tm/tm19/TM-19_ptA.pdf` (cit. on pp. 4, 5, 36).

[5] Lynn Miller, Charlotte Pelletier, and Geoffrey I. Webb. «Deep Learning for Satellite Image Time Series Analysis: A Review». In: *IEEE Transactions on Geoscience and Remote Sensing* XX.XX (2023), pp. XX–XX (cit. on p. 6).

[6] Yao Sun, Yi Wang, and Michael Eineder. «QuickQuakeBuildings: Post-earthquake SAR-Optical Dataset for Quick Damaged-building Detection». In: *IEEE Transactions on Geoscience and Remote Sensing* (2024) (cit. on pp. 6, 7).

[7] Isaac Corley, Caleb Robinson, and Anthony Ortiz. «A CHANGE DETECTION REALITY CHECK». In: *arXiv Preprint* (2024). URL: `https://arxiv.org/abs/2402.06994` (cit. on pp. 7, 10).

[8] István Bondár and Dmitry Storchak. «Improved location procedures at the International Seismological Centre». In: *Geophysical Journal International* 186 (2011), pp. 1220–1244. DOI: `10.1111/j.1365-246X.2011.05107.x` (cit. on p. 8).

[9] Pu Rena, Chengping Rao, Su Chen, Jian-Xun Wang, Hao Sun, and Yang Liu. «SeismicNet: Physics-informed neural networks for seismic wave modeling in semi-infinite domain». In: *Physics of the Earth and Planetary Interiors* XX (2022), pp. XX–XX (cit. on p. 8).

[10] Alireza Niksejel and Miao Zhang. «OBSTransformer: A Deep-Learning Seismic Phase Picker for OBS Data Using Automated Labelling and Transfer Learning». In: *arXiv Preprint* (2023). URL: `https://arxiv.org/abs/2312.06587` (cit. on p. 10).

[11] U.S. Geological Survey. *SRTM 1 Arc-Second Global Digital Elevation Model (HGT)*. Accessed November 2024. NASA Jet Propulsion Laboratory. 2000. URL: `https://earthexplorer.usgs.gov/` (cit. on p. 32).