Highlights

Mesoscale eddy detection and classification from sea surface temperature maps with deep neural networks

- One strategy is designed using the U-Net method based on sea surface temperature maps.
- The results are evaluated due to accuracy, precision, F1 score, Dice score, and Recall regarding the hyperparameters.
Mesoscale eddy detection and classification from sea surface temperature maps with deep neural networks

Tesi di Laurea Magistrale in Geoinformatics Engineering - Ingegneria geoinformatica

Author: Mohammad Mahdi Safari

Student ID: 970076
Advisor: Prof. Dr. Giovanna Venuti
Co-advisors:
Academic Year: 2022-23
Abstract

Eddies in the ocean are a common and important occurrence that is essential to the movement of materials and energy in the marine environment. Therefore, the intelligent and exact identification of these eddies can substantially aid the advancement of our understanding of oceanography. The crowd is seeing a progressive improvement in the techniques used to identify these marine characteristics as a result of the continuous developments in cutting-edge deep learning technology. The Sea Surface Temperature (SST) data from the Copernicus Marine and Environment Monitoring Service (CMEMS) in the Atlantic Ocean are used in this study to present EddyNet, a unique deep-learning architecture developed for the automatic identification and categorization of ocean eddies. A pixel-wise classification layer is added to the convolutional encoder-decoder structure that serves as the foundation of EddyNet. The final product is a map with the same dimensions as the input, but each pixel is labeled as either "0" for non-eddy areas, "1" for anticyclonic eddies, or "2" for cyclonic eddies.

Keywords: Mesoscale Eddy identification, U-Net, Remote Sensing, Deep Learning, pixel-wise classification
Abstract in lingua italiana

Le vortici nell’oceano sono un fenomeno comune ed importante che riveste un ruolo essenziale nel movimento di materiali ed energia nell’ambiente marino. Pertanto, l’identificazione intelligente ed esatta di questi vortici può notevolmente contribuire all’avanzamento della nostra comprensione dell’oceanografia. Il pubblico sta assistendo a un progressivo miglioramento delle tecniche utilizzate per identificare queste caratteristiche marine grazie ai continui sviluppi nella tecnologia avanzata di apprendimento profondo. In questo studio, vengono utilizzati i dati sulla temperatura superficiale del mare (SST) forniti dal Servizio di Monitoraggio Marino e Ambientale Copernicus (CMEMS) nell’Oceano Atlantico per presentare EddyNet, un’architettura di apprendimento profondo unica sviluppata per l’identificazione automatica e la categorizzazione dei vortici oceanici. Uno strato di classificazione pixel-wise viene aggiunto alla struttura encoder-decoder convoluzionale che funge da base di EddyNet. Il prodotto finale è una mappa delle stesse dimensioni dell’input, ma ogni pixel è etichettato come "0" per le aree non vorticose, "1" per i vortici anticiclonici o "2" per i vortici ciclonici.

Parole chiave: Identificazione dei Vortici Mesoscala, U-Net, Teledetectione, Apprendimento Profondo, Classificazione Pixel-wise
# Contents

Abstract \( i \)

Abstract in lingua italiana \( iii \)

Contents \( v \)

Graphical abstract \( 1 \)

1 Introduction \( 3 \)
   1.1 Motivation \( \) \( 3 \)
   1.2 Scope and objectives \( \) \( 4 \)
   1.3 Study contributions \( \) \( 5 \)
   1.4 Organization of the thesis \( \) \( 5 \)

2 State of the Art \( 7 \)
   2.1 Eddy identification \( \) \( 7 \)
   2.2 Deep learning \( \) \( 10 \)

3 Fundamentals \( 15 \)
   3.1 Ocean dynamics and turbulence \( \) \( 15 \)
      3.1.1 Sea surface temperature and sea level anomaly \( \) \( 16 \)
      3.1.2 Mesoscale ocean eddies \( \) \( 17 \)
      3.1.3 The sensitivity of eddies to rising sea levels and climate change \( \) \( 20 \)
      3.1.4 Gulf Stream \( \) \( 22 \)
      3.1.5 Identification of mesoscale ocean eddies \( \) \( 23 \)
   3.2 Artificial neural network \( \) \( 24 \)
      3.2.1 Image segmentation \( \) \( 24 \)
      3.2.2 Introduction of deep learning \( \) \( 25 \)
      3.2.3 Hyperparameter and metric values setup \( \) \( 26 \)
      3.2.4 Unet architecture \( \) \( 28 \)
Graphical abstract

Figure 1: Graphical abstract as outlined in this study.
1 | Introduction

Ocean circulation has been impacted by climate change and global warming nowadays. Several ocean regions often make a difference in how sea surface temperature (SST) changes affect anomalies in atmospheric circulation [1]. Mesoscale ocean eddies are capable of influencing the behavior of the atmosphere at mesoscales primarily through kinetic energy and have a regional impact on near-surface wind, cloud characteristics, and rainfall [2]. As a result, understanding the transmission and features of mesoscale ocean eddies and using this information to identify and analyze them is crucial for both global climate change and geological and oceanography purposes. In this research, a conventional neural network (CNN) constructed using the Unet architecture has been employed to analyze preprocessed satellite images captured in the Atlantic Ocean, encompassing Gulf Stream data from the period between 2017 and 2019. The data has been provided by Bonn University.

1.1. Motivation

In recent years, essential jobs that provide insight into climate change include detecting and comprehending the growing global warming and its repercussions, particularly in the anticipating and monitoring of sea level increases. Numerous deep learning (DL) and machine learning (ML) techniques have so far been based on graphical and spatial analysis methods. These investigations may yield crucial evidence showing that the changes are causing oceans and sea levels to rise. The investigation and forecasting of ocean currents and streams are part of oceanographic simulations. To improve our comprehension and forecast of climate change and sea level changes, reliable modeling of ocean dynamics and turbulence is essential. Physical mechanisms that cause ocean currents have a profound impact on ocean dynamics and the behavior of the ocean as a whole. A crucial concern about the temporal and spatial distribution of aberrant eddies in the global ocean is raised by data from SST that demonstrate that warm anticyclonic eddies and cold cyclonic eddies occasionally occur in some places [3]. The revolving water vortices known as oceanic mesoscale eddies are found across the ocean and have spatial and temporal ranges [4].
Understanding eddy dynamics is essential to comprehending the environmental conditions on Earth. As a result, finding and monitoring mesoscale ocean eddies is essential for producing accurate numerical atmospheric models that anticipate changes in the climate.

The burgeoning field of computer vision research, particularly neural networks, provides interesting applications to achieve more accurate eddy detection in addition to the well-established state-of-the-art methods for identifying eddies. The findings of existing approaches are relatively trustworthy but lack complete automation since they rely on user-defined thresholds, which makes them dependent on user inputs. As a result, the incorporation of AI algorithms has become fascinating. The computer may be trained to recognize eddy patterns on its own by using neural networks. This strategy has the potential to improve eddy detection’s accuracy and effectiveness. Among various concepts in ML and DL methods, this study has utilized the Unet architecture as a DL method for eddy detection. ST-U-shaped network (UNet), a unique semantic segmentation framework for RS satellite images that incorporates the Swin transformer into the traditional CNN-based UNet, is proposed [5].

1.2. Scope and objectives

The purpose of this document is to describe in detail the development process and goals of a patient Mesoscale eddy detection and classification from SST maps with deep neural networks. The primary content care for the mesoscale eddy detection and classification from SST maps. This work evaluates a deep learning-based architecture for automated eddy detection and classification from SST maps. The reference data provided by Bonn University has been defined in Section 4.2. The Unet consists of a convolutional encoder-decoder followed by a pixel-wise classification layer.

There are three goals accomplished in this project. The first was to plot maps with the same size as the input where pixels have the following labels ('0': Non-eddy, '1': anticyclonic eddy, '2': cyclonic eddy). Secondly, provide Python code, the training datasets, and Eddy network weights files. Lastly, our outcome joins the emerging relationship between the remote sensing and machine learning communities that are leading to significant contributions in addressing the segmentation of RS images. To the best of my knowledge, the present work is the first to propose a deep learning-based architecture for pixel-wise classification of eddies, dealing with the challenges of this particular type of data. The project is implemented as a Python program in Anaconda software with a code script developed by Pytorch and TensorFlow libraries in the local host. The database is organized with SST and Gulf Stream information from 2017 to 2019 in the Atlantic Ocean. Gulf
stream data has eddies information containing cyclonic and anticyclonic for all days in each year.

1.3. Study contributions

The presented research uses pre-processed satellite images to resolve the eddy identification challenge utilizing an Unet method based on CNN architectures and standard image processing tools. Common eddy detection techniques divide a learned dataset into test and train data before applying a DL strategy to it. As a result, it is possible to produce a more automated and improved method for eddy detection. The outcomes have also been presented visually on the local host. This visualization encompasses both the hyperparameters utilized during the process and the images that have been predicted by the custom-designed neural network. The given data has been reviewed using the aforementioned approach to achieve appropriate accuracy and loss function.

On the other hand, it provides a comprehensive technique for eddy identification in order to support the understanding and monitoring of oceanography and climate changes. In this regard, employed specific outlines to detect mesoscale ocean eddies:

- preprocessed images in the Atlantic Ocean from 2017 to 2019, plot and visualizing extracting needed information,
- a thoroughly evaluated Unet for semantic segmentation to detect eddies by expanding the spatial analysis to automated monitoring of the current ocean,
- visualizing and showing graphs, hyper-parameters, results, and prediction images from reference images

These complications are carried out using Python components in a system that is almost entirely automated. The most suitable Python backends for building the neural network method are TensorFlow and TensorBoard.

1.4. Organization of the thesis

This study includes six chapters. In the upcoming chapter, this study delves into a comprehensive exploration of ocean eddy identification methods and criteria, elaborated in Section 2.1. Additionally, we provide an extensive review of previous studies and an overview of various DL and ML approaches, which we conceptualize in Section 2.2. The third chapter is dedicated to shedding light on ocean dynamics and turbulences, subdivided into five insightful subsections. Furthermore, we offer a brief introduction
centered around the application of DL and artificial neural networks. In Section 3.2, we
provide a meticulous and detailed explanation of the Unet architecture concept. The
methodology chapter is a repository of essential information, encompassing details about
the study area, dataset, algorithm design, and theoretical underpinnings of the metrics we
employ. Moving forward, the crowd presents the outcomes of our research, complete with
a thorough metric analysis and a comparative examination of our findings against related
studies, all found within the chapter 5. Finally, the conclusion chapter encapsulates a
comprehensive summary of the entire body of work, along with offering valuable insights
and recommendations for future research endeavors.
2 | State of the Art

There is now more need for monitoring and studying the oceans and seas as a result of the rise in sea levels brought on by global warming and environmental concerns. Numerous techniques and algorithms have been created to locate the best solutions and optimize information in order to meet this need. In recent years, optimizing data analysis has become more dependent on the assessment of DL and ML techniques with the use of AI algorithms, and various studies have employed physical and chemical parameters to structure the method to identify the eddies network. In terms of computer vision problems, object identification, and image segmentation sequences, the implementation of the deep neural network, Unet, scores highly. Due to the fact that this study covers both research areas in eddy identification and DL methods, they are feasible to connect the vital and cutting-edge areas of this study. In the following, represented researches include the fields of oceanography, climate change, and environmental matters as well as RS, computer vision, DL, and ML methods during these years.

2.1. Eddy identification

Significant currents and associated instability on the oceanic mesoscale (≈ 100 km) play a key role in determining the ocean’s stratification [6] and aid in the climate system’s poleward heat transfer [7–10]. It is recognized that these circulations’ quasi-stationary component, or what is left over after multi-year averaging, exhibits significant thermal surface manifestations in the form of narrow fronts in SST. The atmospheric’s border layer is seen to be highly connected to them [11–14], and it’s possible that they have an effect on the environment [13, 15, 16]. Mesoscale eddies are widely dispersed over the ocean’s surface and are important for moving mass, energy, heat, and nutrients between different ocean basins [17].

Eddies may be found and studied to assist researchers in understanding their influence on ocean climate models [18]. The SST combined products attained adequate decisions to permit the recognition of mesoscale eddies with the introduction of altimeter operations and the availability of two or more altimeters at the same time [19, 20]. The crowd may
differentiate between three types of eddies using SST maps: cyclonic eddies, anticyclonic eddies, and no eddies. Cyclic eddies are distinguished by their negative SLA, whereas anticyclonic eddies are identified by their positive SLA (Sea Level Anomaly is meaning SST anomaly with respect to a particular mean) [21].

Various studies have been done recently with the goal of automatically recognizing and categorizing eddies. The Three-dimensional mesoscale eddy identification technique centered around pressure anomalies is one of the most common types of approaches that predominate in the literature on the subject. In this work, a three-dimensional (3D) eddy identification and tracking system based on pressure anomalies—which are analogous to SLAs—has been suggested. Despite this, the majority of mesoscale eddy experiments in the area concentrate on surface eddies, necessitating more investigation into the framework and characteristics of 3D eddies. To create a 3D eddy collection in the Kuroshio Extension (KE), the study has used this approach on a 5-year (2008–2012) superior resolution computation output. The trustworthiness of the mathematical result has been demonstrated by the five-year temperature/salinity hydrological parameters and surface eddy distribution. The 3D eddy tracking dataset shows that as eddy existence duration rose, the total quantity of eddies reduced significantly, and more anticyclonic eddies than cyclonic eddies have been presented at times greater than one week. Two 3D eddy-tracking trajectories with a distinct leap in depth and a shift towards the west and equator are displayed with daily fluctuations in the 3D structure. Inevitably is observed as a cylindrical eddy, and its eddy radius is essentially constant throughout all strata, in addition to the bowl, lens, and cone eddies that earlier studies have identified. Significant "negative-positive" saline anomalies, declining current sectors, and negative temperature anomalies are all brought on by cylindrical eddies in the KE area. On the other hand, anticyclonic eddies result in raised current regions, positive temperature anomalies, and "positive-negative" saltiness anomalies. The overall pattern of the current field and temperature/salinity anomalies, which are connected to their organization, are affected differently by the several main types of eddies [21].

Another study addresses the long-term imagery collected by satellite from the South Eastern Arabian Sea (SEAS) coastal waters has examined how mesoscale dynamics affect phytoplankton biomass fluctuation in chlorophyll-a (chl-a) intensity. Satellite-derived chl-a, SLA, SST, and sea surface wind measurements from 1998–2016 have been gathered and analyzed to explore coastal upwelling and mesoscale eddies. The Practical Orthogonal Function and Morlet wavelet analysis estimated statistical fluctuation and indicated considerable seasonal and interannual modification in chl-a concentrations in connection with ecological variables. The Okubo–Weiss criteria identified mesoscale eddies. The summer
period of monsoons (June–September) has recognized cyclonic (cold core) eddies. The summertime thunderstorms had the highest wind-induced upwelling and cyclonic eddies, raising chl-a levels. Coastal upwelling or cyclonic eddies induce chl-a variability, which may be due to seasonal and interannual changes in surface and subsurface materials. El Niño, La Niña, and the Indian Ocean Dipole (IOD) occurred throughout the research period and have been taken into consideration to explain the interannual changes in chl-a and related factors affecting the environment. Strong El Niño, La Niña, and IOD increased chl-a fluctuation. Researchers investigated the link involving chl-a, coastal upwelling, and mesoscale eddies, their relevance, and their impact on SEAS bio-production [22].

Based on research that has been done recently, oceanic mesoscale eddies have enormous thermodynamic anomalies; however, they are mostly investigated in terms of surface temperature and aggregate mean. The crowd uses a theoretical eddy identification and tracking algorithm and a novel corresponding and carrying process to study North Pacific eddy-induced thermodynamic anomalies, their relationship with eddy amplitude, sea surface height (SSH), and the proportion of variation they clarify on sub monthly to interannual time frames. SST, isothermal layer depth (ITD), and upper ocean heat content (HCT) are used to study thermodynamic anomalies. Most eddies have modest magnitudes and thermodynamic abnormalities. Anticyclonic eddies are warm and have a deeper isothermal layer and more heat than cyclonic ones. Some eddies, likely underground ones, have opposing polarity. Thermodynamic parameters have linear associations with eddy amplitude, although scatter and seasonality vary. HCT-amplitude relation scatters the least and has the smallest seasonal change, ITD has the biggest dispersion and changing seasons, and SST is in between. The most eddy-rich area in the North Pacific, the Kuroshio and Oyashio Extension, accounts for about 50% of SSH fluctuation up to season-to-season and ITD and HCT variability up to interannual. Submonthly eddy-induced SST variability is greatest near the Oyashio Extension Front at 40–60%. These findings show how mesoscale eddies affect ocean thermodynamic variability and air-sea connection [23].

An alternative representation is shown in the iso-SST structure; the content of this study offers a unique method for automatically detecting and characterizing mesoscale eddy identification features in RS imagery of SST off Portugal. The dynamics of the investigated region—upwelling currents and oceanography impact generate innumerable and extremely diverse SST structures, characteristics that are fascinating and might possess uniform limitations, and edges related to significant temperature variations that might not belong with a particular eddy—make this assignment difficult. These hinder edge-feature-based image processing, which can automatically detect eddies in other oceanic zones like the Gulf Stream. The suggested system uses iso-SST characteristics involved
with the eddy-related structure to code a governed-by-rules visual recognition procedure. This reveals the position, size, rotation, and symmetry of the eddy-related organization and facilitates the use of SST information to annotate the RS image and assess the observable survey subjectively [24, 25]. A similar study quantifies how mesoscale eddies affect air-sea heat fluxes and associated factors in the South China Sea—inevitably examined combined structures of air-sea fluxes and factors connected to anticyclonic eddies and cyclonic eddies from 2000 to 2015 using imagery from satellites of SST and SSH anomaly and a high-resolution air-sea heat flow item. SST-SSH correlations over eddies are sometimes negative. 56% of anticyclonic eddies are SST$^+$ while 58% of cyclonic eddies are SST$^-$. Eddy amplitude increases the proportion of these eddies, which vary seasonally. All eddies—SST$^+$ and SST$^-$ are composited with SST anomaly, air-sea parameters, and fluxes. All combinations display asymmetric structure, indicating that the fluxes and parameters’ centers (extrema) shift westward and poleward (equatorward) relative to the anticyclonic eddies (cyclonic eddies) cores. Compounds of latent heat flux (LHF), sensible heat flux (SHF), wind speed, air temperature, and particular moisture reveal monopole and dipole trends, respectively [26].

In this study, various parameters employed in different types of research are considered. Among them, satellite images of SST from 2017 to 2019 are daily obtained in the Atlantic Ocean. The SST data have then been categorized into two stages based on the eddy structure, distinguishing cyclonic and anticyclonic streams. This approach allowed for a comprehensive analysis of the ocean dynamics during the specified period, detail of ocean dynamics and turbulence are defined in Section 3.1.

### 2.2. Deep learning

The development of DL has produced impressive advancements and is a game-changing breakthrough in current computer science research. However, obtaining the large amounts of labeled data that neural networks frequently require may be complicated and expensive. Supervised algorithms are frequently enhanced by data creation or data augmentation techniques to overcome this problem. Researchers may resolve shortages of information and improve the overall performance of the models by implementing these tactics. Classical mesoscale eddy identification relies on expert discrimination or threshold setting and is subjective. One of the most recent novel research methods, the study proposes an ocean mesoscale eddy identification method for deep transfer learning target recognition based on YOLOF (You Only Look One Level Feature) in this work according to the major benefits of YOLO series goal detection frameworks in the context of DL. The proposed
approach outperforms more conventional classification techniques in terms of detection operation, avoids the effects of threshold adjustment on mesoscale eddy identification, and somewhat increases identification speed [27, 28].

Another method mentioned in the following is the designed network to evaluate based on DL approaches. Oceanic mesoscale eddies considerably affect energy, issues, and auditory transmission. Interestingly, oceanic mesoscale eddies are detected using a threshold ratio that is too subjective. Since the training established lacks influence, ML approaches are neither sophisticated nor meaningful. This research builds an object recognition network-based mesoscale eddy automated identification and location network, OEDNet, to address the aforementioned issues. To construct the train set, 2D processing of image technology enhances a limited number of accurate eddy samples described by ocean specialists. Then, a deep neural network and feature pyramid network object detection model is optimized for small samples and complicated areas in ocean mesoscale eddies. Experimental findings reveal that the model outperforms the old identification approach and generalizes well across water regions [29]. As a similar concept method, the pyramid scene parsing network (PSPNet), which can fulfill the combination of semantics and specifics, is utilized as the fundamental method in eddy detection procedures with the aim of adjusting to segmentation challenges for multi-scale oceanic eddies. When eddies are found using this AI technique, the findings are comparable to those found using a conventional vector geometry-based (VG) methodology. The AI approach detects higher oceanic eddies than the VG approach, particularly for small-scale eddies. As a result, the current study shows that the AI system may be used to detect oceanic eddies. One of the earliest attempts to link oceanographic investigation with AI technology [30]. Based on the review research has been done by Xu et al. [31], the previous method compared to two other methods. In this review, the quantity, dimensions, and lives of eddies found using the three approaches are compared. Of the three AI-based techniques, the PSPNet algorithm finds the most ocean eddies with regard to eddy counts. Because the Spatial Path is incorporated into the algorithm to prevent the destruction of the eddy edge information, the BiSeNet can discover more large-scale eddies than the other two techniques. Concerning eddy lifetimes, it is noteworthy that the DeepLabV3+ model lacks the capability to track ocean eddies with extended durations. The current limitations of DeepLabV3+ prevent it from effectively capturing the longer lifetimes exhibited by certain oceanic eddies.

Furthermore, the eddy subsurface vertical structure-oriented 3D neural network classifies ocean eddies. This work is one of the initial attempts to test DL for vertically organized eddy recognition. First, the refined eddy profiles database is created from a vertical profile framework, which is closely associated with altimetry sea surface topography. Next,
the eddy vertical structure-oriented 3D neural network based on the residual network (ResNet) classifies eddies as anticyclonic, cyclonic, and no eddies. As external variables, the suggested network can incorporate geographical and dynamic characteristics. The suggested network can describe 3D eddy data and expand to a deeper network architecture using 3D convolutions and pooling. Finally, classification trials verify the suggested technique. A particularly remarkable outcome from trials is that the suggested technique may enhance eddy detection capability by applying altimetry-calibrated vertical characteristics with comparable classification efficiency [32].

Moreover, the other crowd provides a unique technique for detecting eddy signatures on such data that makes use of DL. With the help of an altimetric-based area suggestion, the developers offer the initial collection currently accessible for this job while keeping SST pictures. Inevitably develop a CNN-based classification that successfully recognizes eddy signals in specific cases. This study demonstrates that training on a lesser amount of manually marked information can help overcome the challenge of categorizing an enormous number of automatically maintained pictures. The erratic automated labeling and inherent complexity of the SST data account for the differences in performance between each of the subsets. Oceanographers may be able to validate altimetric eddy identification using SST using this method [33]. Additionally, Kemker and Kanan [34] developed a CNN-based architecture. CNNs excel at processing grid-based topology with spatially dependent elements like photos, making them potential visual identification tools [35]. Postadjian et al. [36] also used CNNs to classify satellite images. Both systems used fully convolutional networks with upsampling following extracting features for pixel-wise semantic segmentation. Encoder-decoder semantic segmentations are common [37]. This design first does feature extraction resembling a CNN classification. CNN predictions are upsampled layer-wise. Down-sampling layering for extracting features can reduce segmentation determination, particularly when using only one upsampling procedure following deep extraction of features.

Furthermore, a similar study has been employed a combination method. The typical vertical characteristic and eddy from an image may be extracted using an innovative combination technique that utilizes CNN and extreme gradient boosting (XGBoost) benefits. CNN extracts vertical characteristics from the input data examined at the bottom of the neural network. secondly, the highly dimensional vectors of variables and other account characteristics are fed into the model known as XGBoost to categorize profiles outside altimeter-identified eddies (Alt eddy). At last, comprehensive trials prove the technique’s efficacy. The CNN-XGBoost simulation classifies 98% of eddies and recaptures 36%. CNN-XGB eddies are compared to Alt eddies for vertical structure and geographical dis-
tribution, showing a piece of high vertical evidence and a tropical ocean eddy zone. The suggested theoretical structure can improve eddy detection and in situ float measurements [38].

In the following mentioned some similar academic studies have been used the Unet method as the main eddy identification. such as a deep neural network-based framework for eddy detection from several RS sources is proposed in the study as the multi-modal U-Net model. By combining SSH and SST data, the novel suggested approach for eddy identification enhances its precision and efficacy in comparison to earlier methods. RS images, such as those showing SSH, SST, etc., may be employed to identify mesoscale eddies. The majority of autonomous eddy detection algorithms in use today have been developed using a specific type of RS data. To assure the precision and effectiveness of eddy identification, there is a need for an autonomous eddy detection method that can fully use data from many RS sources [39]. The following letter presents the pyramid split attention (PSA) eddy detection U-Net architecture (PSA-EDUNet) for marine imaging eddy identification. U-Net inspired the PSA-EDUNet, which has encoder and decoder elements to efficiently integrate inferior and senior features and prevent characteristic data from being lost through nonlinear connection mode. PAS improves the extraction of characteristics. SST and SLA are the key eddy identification criteria for fusion data. The Kuroshio Extension (KE) and South Atlantic studies show that the suggested technique outperforms previous methods, notably for eddy margins and small-scale eddies [40]. DL methods for eddy detection are quite young. The enclosed letter proposes another semantic segmentation-based DL for ocean eddy detection. Semantic segmentation requires context-efficient pixel-level detection. Two modules for attention address this issue. VGG16-based U-Net architecture with two attention modules shows the contextual connection in the distribution and geographic dimensions. In accordance with connections, each pixel or bandwidth includes context from the others. A residual route replaces the skip link between encoder and decoder units. The studies reveal that an attention-based deep framework and novel residual approach increase the efficiency of models over state-of-the-art approaches [41]. In addition, another similar research introduces EddyNet, a DL-based architecture for automatic eddy detection and classification from Copernicus Marine and Environment Monitoring Service (CMEMS) SSH maps. A convolutional encoder-decoder with a pixel-by-pixel classifying structure makes up EddyNet. The result is a map of comparable size as the input with the three labels shown under each of the pixels [42].

After a thorough examination of diverse research in recent years and the utilization of DL methods, this study has conducted an evaluation of Unet as the principal approach, con-
sisting of a convolutional encoder-decoder followed by a pixel-wise classification layer. The initial visualization involved plot maps with the same size as the input, where pixels were labeled as follows: '0' for Non-eddy, '1' for anticyclonic eddy, and '2' for cyclonic eddy. Subsequently, training datasets and Eddy Net weights files have been generated. Moreover, our findings signify the growing collaboration between the RS and ML communities, which has resulted in substantial contributions towards addressing the segmentation of SST images. The Unet architecture and implementation models are specifically discussed in the dedicated Section 3.2.4 of the study.
3 | Fundamentals

The pair of primary facets of this research’s focus are stand-alone image segmentation and DL issues, each with a different emphasis. The detection of mesoscale ocean eddies over a three-year period, driven by a physiologically motivated purpose, is the key application paradigm and goal of this work. As described in Section 2.1, some existing methods that use SST are based on temperature comparisons.

The study uses the most recent developments in deep neural networks to accomplish eddy identification in order to overcome this difficulty. Modern representations for a variety of tasks, including object detection, recognition, and recognizing individuals, have quickly evolved using DL techniques. As a result, deep architectures offer a strong framework for efficiently detecting coherent patterns in SST, that can extract from SLA concept meaning.

By creating and training a DL-based technique for eddy analysis in the future, this study aims to close the gap between RS methodologies and DL methods. With this strategy, new and exciting opportunities to understand hybrid dynamic processes in the Atlantic Ocean are promised. The crowd would be given through the creation, which is the contribution and geophysical interdependencies of mesoscale ocean eddies in the sections that follow. The most well-known eddy identification methods will also be covered. The broad study field of deep feature representations and neural networks is set to be covered, including the fundamental concepts, mathematical formalisms, and significant developments in this domain.

3.1. Ocean dynamics and turbulence

The extremely turbulent character of the oceanic flow, which spans a wide range of sizes from 1,000 km down to 1 m, and even smaller size scales, is well known [43]. Nonlinear scale connections that cause this turbulence make it possible to transmit energy both upmarket and downscale. Mesoscale eddies, which may range in size from 100 to 300 km, are crucial to many marine events. They not only participate considerably in the
turbulent distribution and transfer of trace elements like temperature and carbon dioxide throughout marine basins [43–46] but also contribute to a sizeable fraction of the total kinetic energy in the oceans [43].

Satellite altimeters successfully monitor the kinetic energy associated with mesoscale eddies from space, offering insightful data on their dynamics. Additionally, these eddies are explicitly resolved in contemporary, realistic ocean simulations, furthering our understanding of their behavior and impact on ocean dynamics [47]. The upper-ocean ecosystem and air-sea dynamics, including movement, energy, heat, and chemical fluxes, are all significantly influenced by ocean dynamics and turbulences. Additionally, these turbulences can have an impact on economic operations, including offshore exploration, shipping, and fishing. In practical terms, precise modeling of surface amplitudes is crucial for many applications, such as particle identification and spilled oil modeling [48], atmospheric and environmental modeling [49], and the determination of the risk of flooding along coastlines [50]. The 1930s revealed the Gulf Stream and Kuroshio jet currents’ eddies and meandering [51–53]. At this point in time, both big eddy forms (up to 200 km in diameter) and minor eddies (up to 20 km) have been recognized by water temperature. In the 1950s and 1960s, synoptic hydrological surveys have been carried out using multiple ships (Operation Cabot [54]), radio navigational instruments, bathythermographs, electromagnetic current meters, and neutral buoyancy floats have been incorporated into oceanography, which helped detect eddies. Eddy formations were identified in the Gulf Stream and Kuroshio areas, the Gulf of Mexico [55], the East Australian Current zone [56, 57], and other portions of the World Ocean [54, 58]. However, eddies were not yet seen as a global phenomenon or a unique category of marine processes. Oceanology lacks mesoscale dynamic process theory, unlike atmospheric sciences [59].

3.1.1. Sea surface temperature and sea level anomaly

Global SST fields serve as a diagnostic tool for comparing to SST generated by ocean simulations, an oceanic boundary requirement for atmospheric theories, and a means of tracking climate change. The SST field could be the most well-known ocean characteristic on a global scale since the SST can be measured via satellites [60]. The crowd may differentiate between two types of eddies using SST maps: cyclonic eddies and anticyclonic eddies. Cyclic eddies are distinguished by their negative SLA, whereas anticyclonic eddies are identified by their positive SLA [42]. The geostrophic movement caused by mesoscale eddies may result in the sea surface rising or falling, which satellite altimeter measurements are capable of detecting. Eddies are often recognizable in the SST anomaly area as confined spaces. It is clear that the SST anomaly map has a significant degree of resem-
blance; frameworks with high/low SSH anomalies match those with high(warm)/low(cold) SST anomalies. In order to distinguish mesoscale eddies from the SST anomaly subject matter, a contained zone of SST anomaly has been used [61]. High SST anomaly eddies constitute warm eddies (anticyclones), whereas low SST anomaly eddies represent cold eddies (cyclones) [62].

In geophysical fluids, stratified and rotating severely impede vertical velocities throughout an extensive set of horizontal scales $L_h$, as well as above the deformation radius $(R_d)$, generally 30 km in the ocean. Vertical velocities usually range from a few meters to many tens of meters per day for submesoscale motions, $L_h; O(1)$ km, although their strength rises at shorter scales [63]. Vertical transfer of thermal energy and biogeochemical tracers, however lower than horizontal advection, is vital to the World Ocean’s functioning [64]. Coastal upwelling favorable winds or positive wind stress curl can cause vertical components in enormous scale circulation (100 km and bigger). At smaller scales (1–100km), intermittent vertical velocities have been linked to mesoscale and submesoscale turbulence and can be caused by forced and unforced movements like frontogenesis, baroclinic instability, or air–sea interactions, which may couple [65]. They are responsible for vertical flows that have been challenging to measure but are generally accepted that they serve an important part in heat [66–68] and salt budgets [69], carbon and nutrient cycles [70, 71], and oceanic biodiversity [64]. Vertical fluxes related to 3D turbulence also fluctuate at sizes smaller than the submesoscale [72].

3.1.2. Mesoscale ocean eddies

Mesoscale eddy-wind interactions, or the actual effect carried out by winds on ocean eddies, is one of the physical mechanisms that may have a large influence on ocean eddies and the energy they transport. It has been recognized [73–75] that atmospheric wind monitoring is methodically SST currents in the ocean if the corresponding movement across the atmosphere and underlying surface ocean is included into consideration in the surface stress calculation (so-called "relative wind stress effect"). Simple scalability analysis indicates that damping by the proportion of wind stress should largely work on mesoscale eddies since eddies predominate kinetic energy at the ocean surface [74, 76], world’s SST anomaly has been shown in Figure 3.1.
Consequently, the negative wind work on ocean eddies usually corresponds to the reduction in wind power input to the ocean when the relative wind stress is employed in the energy estimate [76]. To current understanding, no direct observational studies have yet been done that distinctly show and quantify how much the relative wind stress dampens the influence of ocean eddies, particularly on a global scale. The mathematical modeling of eddy-wind interaction frequently accepts that the underlying wind field is evenly distributed over the lateral extent of the eddies because atmospheric winds have a tendency to change on a great deal wider spatial scales than mesoscale ocean eddies. As a result, it frequently emphasizes the vortex framework of ocean eddies and its function in producing the anomalous compared wind stress curl on the scale of the eddies [74, 75]. On the opposite, it has been demonstrated that the surface wind stress displays significant geographical fluctuations that are characterized by cyclonic and anticyclonic wind stress curl in the subpolar and subtropical oceans, respectively. The effect of the enormous scale wind stress curl on the power input to mesoscale ocean eddies is poorly understood. This study employs satellite measurements to quantify the impact of air winds on mesoscale eddies in the global ocean for anticyclonic and cyclonic (Figure 3.2) [78].
The conventional wisdom holds that mesoscale cyclonic eddies and anticyclonic eddies are connected to anomalously warm and cold surfaces and subsurface cores, accordingly, and are likewise characterized by surface-intensified potential vorticity. This study is focused on a particular category of anticyclonic eddies and cyclonic eddies that have been characterized by subsurface-intensified potential vorticity and SST anomalies of a reverse orientation to the conventional eddies [71, 80, 81], i.e., anticyclonic eddies with SST warmer (WAEs), and cyclonic eddies with SST colder (CCEs), than the water surrounding them beyond the eddies. Although WAEs and CCEs have been seen by satellite a few times in boundary currents and marginal sea regions [82–84], it is unclear how common these unusual eddies occur on a global scale. Additionally, investigations depending just on satellite data don’t reveal anything regarding WAEs and CCEs’ underlying features. As an example, it is unclear if the cold cores of CCEs and warm cores of WAEs discovered in satellite observations reach tens, hundreds, or even thousands of meters through the depth of the ocean or if they are restricted to relatively near the sea surface [85]. Determining air-sea interactions may need a special focus on the distribution and upper ocean structure of WAEs and CCEs throughout the global ocean. The atmosphere and ocean interior are connected and communicate with one another through the air-sea boundary layer. According to Frenger et al. [2], mesoscale eddies can drastically alter fluxes at the air-sea interface. As an illustration, warm (cold) surface water commonly associated with AEs (CEs) causes anomalous upward (downward) air-sea heat fluxes, which in turn strengthen (weaken) surface wind stress by increasing (decreasing) vertical turbulent mixing and downward momentum transport in the atmosphere boundary layer [17, 85]. Additionally,
it has been observed that anticyclonic eddies and cyclonic eddies, which are caused by anomalous air-sea thermal fluxes linked to eddy-induced SST anomalies [86], have been observed to deepen and shoal the surface mixed layer, respectively [87, 88]. Due to varied causes in various places, mesoscale eddies are also known to cause unique top ocean biologic reactions.

3.1.3. The sensitivity of eddies to rising sea levels and climate change

Eddies affect ocean-atmosphere interactions. Eddy anomalies in SST, salt content, and ocean circulation can impact turbulence flows at the air–sea interaction [89]. Their influence on air–sea interactions have been examined for over twenty years. At modest scales, SST fluctuation may cause air–sea thermal fluxes. Another effect [12] explains how air temperature affects near-surface atmospheric stability, vertical motion transfer, and surface pressure. Eddies alter air–sea motion exchanges through the ocean’s surface. Ocean eddies friction on the atmosphere creates a wind stress curl [90], lowering eddy kinetic energy and lifespan and altering enormous scale currents. Eddy-induced stress curl causes small-scale Ekman pumping [91]. Eddies modulate air–sea flows, although their effects vary, due to these investigations [12] have found that regional determinants affect air–sea fluxes’ small-scale variation.

Eddies are in charge of regulating the temperature, volume, and mixing of the ocean’s waters. This mixing is caused by the spinning movement of the ocean’s depths. As a result, it brings up some intriguing points regarding the effects of environmental degradation and rising sea levels.

\[ \rho = \frac{m}{v} \]  

(3.1)

Sea level fluctuations (Equation (3.1)) are essentially related to the current water density \( \rho \), where \( m \) represents mass, and \( V \) represents volume. The density is significantly influenced by pressure, temperature, and nutritional value, particularly salt. The volume is inversely proportional to mass and density, therefore, when the temperature rises, materials or volumes increase, and their densities decrease. As a result, the sea level increases locally due to the eddies’ ability to transmit heat. Additionally, compared to pure water at an identical temperature, seawater has a larger density due to its higher saline content. The distribution of heat in the seas is thus very susceptible to eddy dynamics. The ocean circulation is further distinguished by its forward motion, and energy exchanged
Eddies affected wind speeds, cloud cover, and wetness in the Southern Ocean [2, 17]. Eddies change mixed-layer depth and air–sea interactions [88]. In addition, eddies impact [93] has demonstrated that ocean submesoscale activity can be dampened by winds. Eddies affect air–sea motion exchanges and the antarctic circumpolar current (ACC) energy balance and stratification [94]. Eddies may cause Southern Ocean multi-decadal variation [95]. Eddies affect air–sea interactions, although whether they do in the ice-covered Southern Ocean is unknown. High-latitude eddies are lower than other oceanic eddies [17, 95], making it harder for the atmosphere to react to them. The Southern Ocean has high winds as well as regular thunderstorms, which may make ocean currents less important for wind velocity than in other eddy-prone locations. Reduced sea ice temperatures may cause temperature-inverted shapes, stabilizing the low-level atmosphere and reducing its susceptibility to surface conditions. Thermal variations among sea ice and leads or polynyas may cause mesoscale convection [96]. Furthermore, heavy sea ice can block mesoscale air–sea interactions throughout winter, and Eddies may leave their mark on sea ice, which in turn can leave their mark on the environment as a whole, below represents the distribution of mesoscale eddies over ocean areas in Figure 3.3.

![Figure 3.3: Ocean Models: Representation The issue of Mesoscale Eddies [97].](image)

Determined, localized SST anomalies may indicate that an ocean circulation, like the Gulf Stream circulation off the Atlantic Ocean coast, has departed from its regular route or is more or less intense than normal, discussed in Section 3.1.4. Long-term SST anomalies may indicate global warming. The Northern Hemisphere summer warm anomaly looks to worsen. Sea ice is receding to a smaller region in the summer, leaving places that
are utilized to be encased by ice accessible to the sea. SST anomalies are technical and functional. In coastal locations, abnormal temperatures might favor one organism over another, leading one type of bacteria, algae, or fish to thrive or decrease. Warm sea surface temperature anomalies can alert administrators of natural resources to bleaching coral reefs [77]. The Figure 3.4 depicts the SST that has been observed on the initial day of data collection within the study area during the year 2017.

3.1.4. Gulf Stream

Massive ocean currents like the Gulf Stream transport warm water from the Gulf of Mexico to the Atlantic Ocean. It stretches across the whole eastern coastline of both Canada and the United States (Figure 3.5)[98]. The Gulf Stream has been compared to a river in the ocean, and for a large portion of its length, its surface of 100 m is where its greatest signatures of saltiness, temperature, and movement may be found [99]. From the Florida Current’s first concentration of circulation, resulting in flows from the Dry Tortugas to Nova Scotia, transfer in the Gulf Stream rises by about five times [100]. The increase is primarily caused by the entrainment of Sargasso Sea waters and, to a lesser degree, by the shelves and Slope Water. Seasonal variations in the Florida Current’s velocities and transport volumes lead to varied embedded zooplankton quantities and source locations [101].
From Cape Hatteras to the Grand Banks and beyond, the Gulf Stream Meander Region (GSMR) flows eastward. It is distinguished by a complicated hydrography and convoluted course, which has a considerable influence on biological mechanisms and production [102]. Meanders form and grow in intensity in this dynamic locale, cold- and warm-core rings diverge to the south and north, and water from various domains is mixed with the flow [103, 104].

3.1.5. Identification of mesoscale ocean eddies

Several physical or geometrical criteria-based autonomous eddy detection techniques have been established, and they may be categorized into three groups:

- the physical attribute approach, such as the Okubo-Weiss parameter method [105]
- the flow geometry method, which involves the winding-angle method [106, 107] and the vector geometry method [108]
- the SST-based method [109].

But for three explanations, not all methods can be used to locate mesoscale eddies in the Atlantic Ocean, primarily because it can eliminate noise and excessive eddy detections, the SST-based technique outperforms the Okubo-Weiss parameter method [2]. Secondly, the current observable data in the Atlantic Ocean cannot provide the higher resolution data needed by the flow geometry algorithms to get an accurate flow field to locate eddies [35]. As a result, the SST-based approach created by U-net [17] has been employed in this
work as the DL method (Subsection 3.2.4). By building on the latest advances in DL for image segmentation, "EddyNet" (Section 4.3)—a deep neural network for automatic eddy recognition and categorization using SST maps given by the Copernicus Marine and Environment Monitoring Service (Section 4.2)—presents itself. Concepts from popular image segmentation structures, particularly U-shaped systems like U-Net, served as inspiration for EddyNet.

3.2. Artificial neural network

The primary premise of artificial neural networks is to develop a deep description of past information, taking the scientific enthusiasm from Section 2.2 into consideration. Since they are now used in the majority of computing platforms, neural networks have great potential to advance computer vision technology. The concept and theoretical framework of neural networks are described in more detail in the following. In order to create a deep forwarding network, we must first cover the fundamentals of perceptrons and neurons. We concentrated on convolutional neural networks in depth and their distinctions from conventional nets, the benefits, and drawbacks, in order to develop an eddy detection on SST maps.

3.2.1. Image segmentation

Begin to define the segmentation of images before discussing why U-Net is so often used for image segmentation assignments. Among the several fascinating uses for AIs is computer vision. The most frequent computer vision applications are the recognition of objects and image classification. Predicting whether an image matches class A or class B or class C is the process of image classification for three classes. The entire image is given the expected label. subsequently want to determine what class the image belongs to, classification is useful [110].

Contrarily, object detection goes a step further by foretelling where an object will appear in an input image. By creating a bounding box surrounding an item in the image, then may localize it there. The contents of the image may be found and tracked with a method of detection. Classification and localization are two concepts that may be used to form image segmentation. Segmenting an image entails dividing it into smaller, referred to as segments. To comprehend what is provided in an image at the pixel level, segmentation is employed. It offers detailed knowledge of the image as well as the limits and forms of the items. The result of image segmentation is a mask, every component of which denotes the class to which a given pixel belongs. Let’s use an example to better grasp this Figure 3.6.
The submitted image of an SST map in the study area is seen above on the left. It’s the assignment to make the eddies stand out from the surroundings. As a result, the result is three output classes: no eddy [0], anticyclonic eddy [1], and cyclonic eddy [2]. However, it’s necessary to know the eddies’ precise placement in the image in order to distinguish them from their background and detect each other. Pixel by pixel, image segmentation addresses the aforementioned issue. Inevitably want to distinguish different pixels and group comparable pixels. The crowd is going to categorize each pixel to determine whether it belongs to the eddies classes or the backdrop at that particular location. The other pixels might all have the label 0, whereas all the pixels that our model predicts as belonging to the eddies ought to have the labels 1 and 2. The above-described approach would have resulted in the creation of a mask of our input image, and at the conclusion of this pixel-by-pixel categorization, this would also have identified the eddies’ precise location in SST map images in the study area. Attempt to discover the U-Net model now that there is a better understanding of the process (Section 3.2.4).

3.2.2. Introduction of deep learning

Any classification task’s purpose is to give each pixel a class label. The crowd can make distinctions among supervised and unsupervised classifications based on the information that is accessible. Approaches for supervised instruction are built on prior information, such as ground truth data. Employing labels, the ground truth reflects the actual class identification. The training step involves learning a characteristic illustration through training data using labeled data and the associated dataset. Unsupervised classifications, in opposition, are not dependent on prior information; instead, they combine related observations to group them into the most comparable class. The k-means approach is the most often used method.
Unsupervised learning, meanwhile, is what people actually learn. As a result, this is going to be among the fastest-growing fields for DL and ML studies in the future. DL classification algorithms are typically supervised, however early attempts use neural nets like U-Net to obtain favorable outcomes [111]. Common ML methods typically produce worthwhile outcomes for simple situations. Currently, the quantity and amount are always growing. DL models may use enormous volumes of data more effectively than traditional ML techniques, though. The visualization of data using deep layer structures is the objective of DL. The crowd can recreate individual impressions of an item or component by disassembling these structures. Comparable to multipolar neurons in the human brain, this function. In further detail, the neurons are triggered by signals that are sent throughout the axons and dendrites, presuming the human brain is a sophisticated, highly linked network. The aforementioned are the neurons’ subsidiaries and interactions. The cell that gets the stimulation is known as a neuron. Now, an artificial neural network can be realized by DL. Then, it can enable robots to make complicated choices similar to sophisticated neurological systems [112, 113].

### 3.2.3. Hyperparameter and metric values setup

Essential variables are necessary for controlling the learning trajectory and model capacity in every ML method. The main objective is to make a respectable generalization mistake. These crucial parameters, sometimes referred to as hyperparameters, need to be preestablished before being used. Convolutional neural network outputs are substantially dependent on these factors. Their value, in particular, has a substantial impact on the provision of memory and time spent on computation, which has a major impact on model quality [113].

The learning method is applied to the set of hyperparameters, which should ideally avoid overfitting and underfitting. The selection of hyperparameters can be either manually or automatically for each model or layer. Several automated algorithms have been suggested to do this. Nevertheless, the increased computational costs caused by automated optimization are particularly noticeable. The grid search and the random search are the two most prevalent applicable methods. The model is trained for several sets of hyperparameters in the grid search scenario, such as a range of learning rate values (0.1, 0.01, $10^{-3}$, $10^{-4}$, and $10^{-5}$). The best performance value is then chosen. The random search, in contrast, assumes that there is a marginal distribution of hyperparameters from which to sample. The ability to examine a larger range of hyperparameter settings is the main benefit here. Even though the random search frequently produces better results, it does not converge since the parameters are not continually optimized [111, 114].
Both automated methods have the benefit of eliminating the requirement for an in-depth grasp of how different hyperparameter values affect model capacity. Validation sets are frequently used in the hyperparameter optimization process. The validation set is a separate collection from the training set, just like the test set is. In ordinary circumstances, the validation set is a smaller subset than the training samples. This indicates a split where one set is used to train the classifier and the second dataset refines the learning process with a smaller sample size, focusing on hyperparameter optimization. Cross-validation is frequently used to carry out this segmentation in tandem with the ongoing training procedure [113]. Convolutional neural networks require careful consideration of a number of hyperparameters, including [113, 114]:

- **Learning rate**: The weight modifications during training are controlled by the learning rate, which plays a crucial role. Each derivative is multiplied by a fractional value between 0 and 1, according to this assumption. This parameter has a big impact on how well the model can learn. It is a key element of the optimisation process, together with learning rate degradation and momentum. Furthermore, it is crucial to have the best calibration for this parameter,

- **Epochs**: The number of training iterations carried out until convergence is determined by the number of epochs. Implementing an early halting mechanism is a simple method for figuring out the ideal value. The training process is stopped by this function after the model reaches its greatest validation accuracy. It is also possible to integrate learning rate degradation for each iteration,

- **Minibatches**: Data subsets are represented by minibatches. The sizes of these subsets that are chosen frequently have a substantial influence on the results of neural networks. These sizes typically vary from one to a few hundred occurrences,

- **Activation function**: Due to the common problem of the vanishing gradient, the ReLU activation function is typically used in neural networks instead of the sigmoid activation function. The choice of an activation function often depends on the specifics of the situation at hand, however, the ReLU function regularly yields positive results and powerful performances,

- **Loss function**: A loss function or objective function, is an essential element in the training of neural networks and ML algorithms. It measures the difference between the actual ground truth values found in the training data and the projected values produced by a model. A loss metric’s main goal is to give a gauge of how well the model’s predictions match the actual values, effectively rating the model’s effectiveness. During this study has been used Cross-Entropy Loss, measures the
dissimilarity between predicted class probabilities and true class labels and Dice coefficient. Evaluates the overlap between the predicted segmentation mask and the ground truth mask.

Regarding metric values have also indicated below:

- **F1 score**: The F1 score is a widely used evaluation metric in ML, particularly in tasks involving classification. It provides a balanced measure of a model’s precision and recall, offering insight into the model’s ability to correctly identify positive instances while minimizing false positives and false negatives. The F1 score ranges between 0 and 1, with higher values indicating better model performance. A perfect F1 score of 1 means that the model achieves both perfect precision and recall. Precision is the ratio of correctly predicted positive instances to the total instances predicted as positive. It measures the accuracy of positive predictions. Recall, also known as sensitivity or true positive rate, is the ratio of correctly predicted positive instances to the total actual positive instances. It measures the model’s ability to identify all positive instances.

- **Training metric**: It measures the proportion of correctly predicted instances in the training dataset relative to the total number of instances. Training accuracy is an essential indicator of how well a model has learned from the training data. Training accuracy, and training loss have been visualized in (section 5).

### 3.2.4. Unet architecture

Convolutional networks are frequently employed for the classification of images in which the result applied to an image is a single class label. The intended results, or the assignment of a class label to each pixel, should incorporate localization in many visual assignments, particularly in image segmentation processing. In order to determine the class labeling on every pixel, Ciresan et al. [115] trained a network in a sliding-window configuration using the local region (patch) around the pixel as input. The designed network is able to initially localize. Second, there are far more patches in the training data than there are training images. Clearly, the Ciresan et al. [115] method has two shortcomings. First of all, there are a lot of redundancies as a result of overlapped areas, and it is fairly sluggish because the network must be performed individually for each patch. Furthermore, there is a compromise between contextual use and the accuracy of localization. While smaller patches only enable the network to observe a limited amount of information, larger patches necessitate additional max-pooling layers, which decreases the localization accuracy. A classification algorithm with output that incorporates into
consideration the characteristics from several layers has been suggested by more recent techniques [116, 117]. Contextual usage and effective localization are both feasible simultaneously.

Regarding the drawback noted in the earlier way, a more tasteful and well-regarded architecture—the so-called "fully convolutional network"—emerged [118]. Based on this network, Unet architecture has been created, which requires fewer training images and produces more accurate segmentation predictions. The primary idea of [118] is to add additional layers to a typical contractual network, replacing the pooling operators with upsampling operators. As a result, the output’s resolution is increased by these layers. High-resolution characteristics from the contracting path are mixed with the output that has been upsampled in an attempt to localization. On the basis of this knowledge, a subsequent convolution layer may subsequently learn to put together a further exact result. One significant change to the Unet design is the addition of several channel features to the upsampling section, which enables the network to convey contextual details to higher-resolution layers. As a result, the expanding path produces a U-shaped architecture that is roughly symmetric to the contraction path (see Figure 3.7). The segmentation map only comprises the pixels for which the whole context is present in the input picture. The network does not include any entirely connected layers and just employs the valid fraction of every convolution. By using an overlap-tile technique, this solution enables the smooth segmentation of any huge pictures. The omitted context is obtained by mirroring the given image to attempt to predict the pixels in the edges area of the image. To make use of the network for enormous images, this tiling technique is crucial since otherwise, the GPU RAM would impose a resolution restriction [119].
Figure 3.7: U-net simple sample architecture, a multiple channels characteristic map correlates to each blue box. On the upper part of the element, there is a channel count indicator. At the lower left corner of the sample, the i-j dimension is displayed. Replicated feature maps are represented by white boxes. The various processes are shown by the indications. [119].

Since there is not much training data accessible for the duties we perform, this study utilizes excessive data augmentation by adding elasticity deformations to the training images that are available. As a result, the network may learn to be invariant to these deformations without having to observe them in the labeled image database. This can be crucial for RS image segmentation because realistic deformations can be effectively mimicked, and deformation evolved to be the most frequently occurring variable in the material.

Figure 3.7 shows the network architecture in detail. It comprises an expanded pathway on the right, which is The decoder network, and a decreasing path on the left, which is called The encoder network. The contraction route adheres to the standard convolutional network structure. Two 3x3 convolutions (unpadded convolutions) are applied repeatedly, and after each one, a rectified linear unit (ReLU) and a 2x2 max pooling operation with stride 2 are applied for downsampling. The crowd quadruples the number of characteristic channels with each downsampling step. An upsampling of the feature map continues with a 2x2 convolution ("up-convolution") that cuts the number of feature channels in half, a combination with the accordingly cropped feature map from the path of reduction, and two 3x3 convolutions, each complied with by a ReLU, at each stage of the expanding path.
In order to compensate for the loss of boundary pixels in each convolution, cropping is required. Each of the 64-component feature vectors is mapped to the required number of classes in the final layer using a 1x1 convolution. The network includes 23 convolutional layers in total. It is crucial to choose the input tile size to ensure that all 2x2 max-pooling processes are performed to a layer with an even x- and y-dimension. This will enable a smooth tiling of the output segmentation map [119]. The following outlines the process described in this study [112, 113, 119]:

**Encoder Path:**

- **Input Image:** The U-Net architecture commences with the input image intended for segmentation. This image could be grayscale or multi-channel, in this study has been used SST map data in the Atlantic Ocean (Section 4.2).

- **Convolutional Layers:** The input image is routed across a number of convolutional layers in the encoder path. The input image is processed by a series of learnable filters applied by each convolutional layer, which captures diverse characteristics at varying scales. These filters aid in the image's structure and feature detection. As a result, process has been obtained three eddy classes.

- **ReLU Activation:** Following every convolutional layer, an element-wise application of the ReLU activation function takes place. The ReLU activation introduces non-linearity into the network by preserving positive values and setting negative values to zero. This fosters the network’s ability to capture intricate relationships within the data.

- **Max Pooling:** At periodic intervals, subsequent to several convolutional layers, max pooling is executed. Max pooling serves to diminish the spatial dimensions of the feature maps while retaining vital information. This process entails selecting the maximum value from a local region of the input, resulting in an effective downsampling of the feature maps.

- **Feature Map Dimension Reduction:** As the encoder progresses, the count of feature channels generally increases due to the application of more filters, while the spatial dimensions decrease due to max pooling. This sequence of operations facilitates the extraction of hierarchical features, encompassing both finer details and broader contextual insights.

**Bottleneck layer:**

- **Bottleneck layer:** The bottleneck layer connects the encoder and decoder networks. According to the illustration above, this corresponds to the layer that is at
the bottom. It has two convolutional layers and then Relu. The ultimate illustration of the feature map is the output of the bottleneck.

Decoder Path:

- **Up-Convolution (Transpose Convolution)**: The decoder path aims to upsample the feature maps to match the size of the original input image. Up-convolutions (also called transpose convolutions or deconvolutions) are used to achieve this. These operations "undo" the downsampling performed by max pooling.

- **Skip Connections**: One of the distinctive features of the U-Net architecture is the use of skip connections. These connections link the corresponding layers from the encoder to the decoder path. Skip connections allow the decoder to access both high-level context and fine-grained details from the encoder, aiding in accurate localization.

- **Concatenation**: At each decoder step, the upsampled feature maps are concatenated with the corresponding feature maps from the encoder path. This concatenated feature map contains information from both the upsampled features and the skip connections, effectively combining global and local contexts.

- **Convolutional Layers**: After concatenation, the combined feature map goes through a series of convolutional layers. These layers help refine the segmented output and adapt it to the task’s requirements.

- **ReLU Activation**: Similar to the encoder, ReLU activation functions are applied after each convolutional layer to introduce non-linearity.

- **Final Layer**: The final layer of the decoder employs a convolutional layer with a softmax activation function. This produces a probability distribution over different classes for each pixel in the output segmentation mask. The class with the highest probability is assigned to each pixel.
Methodology

A novel approach in geoscience research is the use of a neural network to detect coherent eddy patterns. The study region, data, and resulting notion of a framework for identifying mesoscale ocean eddies are presented in this chapter. Given the impressive results that the DL algorithm produces in terms of pattern recognition, it is obvious to complete this job using deep neural networks, in particular, Unet. Additionally, this study advises combining SST image processing with neural network detections. Here, we outline our study area, data from satellite altimetry observations that have been preprocessed, as well as the framework that it has suggested. Also provide a general overview of the major components of our structure, their relationships, and interactions. In the following define the headers, processes, input, and output parameters of these routines for this purpose (Figure 4.1).

Graphical Abstract

![Graphical Abstract](image_url)

Figure 4.1: Graphical abstract as outlines in this study.
4.1. Study area

The study area chosen for examination encompasses a significant expanse within the North Atlantic Ocean, ranging between latitudes 17.32°N to 55.50°N and longitudes 41.13°W to 95.63°W, as depicted in the accompanying Figure 4.2 [120]. This extensive scope includes prominent geographical features such as the Hudson Bay, the Gulf of Mexico, and the Labrador Sea. Furthermore, it extends from the western African coastline to the eastern seaboard of the Americas. The North Atlantic Ocean, while generally considered a relatively level expanse of water, does possess certain distinctive oceanic characteristics. One notable example is the Gulf Stream, a robust warm ocean current that flows in a northeasterly direction along the eastern North American coastline. This current plays a significant role in influencing the climate of nearby coastal areas.

![Study Area Diagram]

Within this vast area, a variety of marine life thrives, encompassing different fish species, marine creatures, and seabirds [121]. The intricate marine ecosystems characteristic of this zone emerge from the intricate interplay among various water currents and masses.
Throughout its history, the North Atlantic has played a pivotal role in facilitating commerce and transportation between Europe and North America. The region’s strategic importance has led to substantial maritime engagement, highlighting its vital function as a crucial trade route. Several islands and coastlines are distributed across this extensive research area. These include prominent locations like the Azores, a group of Portuguese islands, segments of the Caribbean archipelago, and the eastern edges of the United States and Canada.

The North Atlantic doesn’t exert a substantial influence on the climate of nearby land areas, especially in Europe. However, the North Atlantic Drift, a key part of the larger North Atlantic Current system, does play a role in adjusting temperatures and shaping the European climate. The North Atlantic Ocean captures considerable attention for extensive research efforts, owing to its vital role in shaping climate dynamics, marine biology, and oceanography. Its varied characteristics and wide-ranging effects make it an intriguing subject for scientific exploration and study.

### 4.2. Satellite data observations

The foundational dataset underpinning this research is the ESA SST CCI and C3S global SST Reprocessed product. This dataset constitutes the pivotal component in the pursuit of understanding and analyzing SST dynamics at a global scale. The dataset offers a comprehensive depiction of the daily average SST at a depth of 20 cm, manifested in a spatial grid resolution of 0.05° x 0.05°.

The dataset is meticulously constructed from satellite data derived from three distinct sources: the Advanced Along-Track Scanning Radiometer (A)ATSRs, the Sea and Land Surface Temperature Radiometer (SLSTR), and the Advanced Very High-Resolution Radiometer (AVHRR) series of sensors, as meticulously elucidated by Merchant et al. (2019) [122]. These satellite sensors collectively provide an expansive and diverse dataset, capturing the complexities of SST variations across different geographical regions.

To yield the ESA SST CCI and C3S level 4 analyses, an intricate processing pipeline known as the Operational Sea Surface Temperature and Sea Ice Analysis (OSTIA) system is harnessed [123]. This system meticulously processes the input satellite data, culminating in a daily analysis of SST that boasts an impressive grid resolution of approximately 5 km or 1/20°. This heightened resolution significantly enhances the dataset’s ability to capture fine-scale temperature fluctuations in oceanic regions.

One distinguishing characteristic of this dataset is its reliance solely on satellite data
Table 4.1: Data information abstraction [125].

processed by the ESA SST CCI and C3S projects. This deliberate approach ensures data consistency and stability, mitigating potential inconsistencies that might arise from heterogeneous data sources. By utilizing exclusively ESA-processed data from (A)ATSR, SLSTR, and AVHRR sensors, the dataset attains a level of reliability and coherence essential for scientific research and analysis.

Augmenting the dataset’s comprehensiveness is the integration of reprocessed sea-ice concentration data from the European Organisation for the Exploitation of Meteorological Satellites (EUMETSAT) Ocean and Sea Ice Satellite Application Facility (OSI-SAF) projects (OSI-450 and OSI-430) as expounded by Lavergne et al. (2019) [124]. This infusion of sea-ice concentration data enriches the dataset by providing valuable contextual information on sea-ice dynamics in conjunction with SST variations.

In this section, the study designates the aforementioned dataset as the foundational training and testing repository for our deep neural network-based Unet algorithm. This investigation harnesses a span of three years, spanning from 2017 to 2019, during which we assimilate daily sensed and detected SST satellite images. The corresponding SST maps, essential for our research, are thoughtfully furnished by the Copernicus Marine Environment.
To accommodate memory constraints, our model’s input images are configured to 512 x 512 pixels. Within this dataset, a total of 2190 cyclonic and anticyclonic satellite images extracted from January 1st, 2017 to December 31st, 2019, play a pivotal role. Additionally, a set of 12 netCDF files, housing SST information, assumes the role of validation data in our study. Delving further, our training dataset comprises 90% of the images, effectively fueling the learning process. In a stratified approach, the images spanning from January 1st, 2017 to the conclusion of May 2017 are set aside exclusively for testing the efficacy of our architecture.

Our geographical scope encapsulates the North Atlantic Ocean region, a canvas vividly illustrated in Figure 4.2. A noteworthy feature of this dataset is the languid nature of its dynamics; individual eddies can endure for several days or even extend beyond a year. Crucially, the dataset undergoes preprocessing to exclude the uppermost region devoid of detected eddies.

A pivotal phase involves the meticulous crafting of segmentation masks for training patches. This intricate procedure entails the generation of polygonal shapes by projecting the SST data onto the nearest lattices within the 0.05° grid. Pixels encompassed within each polygon are subsequently assigned labels corresponding to the class of the polygon, signifying the nature of the eddy: ‘0’ for non-eddy/land/no data, ‘1’ for anticyclonic eddies, and ‘2’ for cyclonic eddies. A visual exemplar of the paired SST map, segmentation map drawn from the training dataset is vividly presented in Figure 3.6. This meticulous amalgamation of data sets, aligned with rigorous preprocessing and labeling, forms the bedrock of our exploration into the U-Net architecture’s capabilities and has been discussed in the following.

4.3. Eddy identification by Unet

The EddyNet architecture draws inspiration from the renowned U-Net framework [119]. The inception of EddyNet involves a dual-track progression, commencing with an encoding (downsampling) trajectory spanning three distinctive stages. Each stage features a tandem of 3 x 3 convolutional layers, sequenced by either the Scaled Exponential Linear Unit (SELU) activation function denoted as EddyNet, or by the established duo of classical ReLU activation and Batch Normalization, referred to as EddyNet. This is followed by a 2 x 2 max pooling layer, orchestrating a halving of input resolution. Meanwhile, the decoding (upsampling) trajectory employs transposed convolutions, also known as deconvolutions, to reinstate the initial resolution. EddyNet, akin to UNet, seamlessly integrates skip connections, facilitating the transmission of information from the contracting path
to the expansive counterpart to account for early-stage insights.

Initial attempts utilizing the original U-Net architecture revealed overfitting tendencies, as the architecture’s capacity exceeded the available training samples. A meticulous iterative process, coupled with hyperparameter fine-tuning, culminated in a definitive selection: a streamlined 3-stage design, each stage equipped with [64, 128, 256, 512] filters, as depicted in Figure 3.7. This design imbues EddyNet with a parsimonious parameter count compared to prevalent architectures, leading to economical memory utilization. EddyNet’s proficiency in mastering data is underscored by its potential for overfitting, affirming its aptitude for unraveling the intricate nonlinearities embedded within eddy detection and classification, thereby addressing the inverse problem.

Amid the U-Net’s stratagem lies the core tenet of capturing multifaceted features across diverse granularities. This paradigm materializes through a sequence of down-sampling and up-sampling phases. The "features" parameter exercises dominion over the quantity of feature channels, synonymous with filters, at each tier of these stages. Augmented with dropout layers, EddyNet undergoes regularization, fortifying its performance in validation loss. The assimilation of SELU theory into CNNs demands meticulous adaptation. Based on previous experiments, the employment of SELU activations in a Unet-like structure induced volatile losses, occasionally spiraling out of control. The study hypothesizes that this may arise from skip connections that potentially contravene SELU’s coveted self-normalization trait. Consequently elected to retain Batch Normalization and ReLU in EddyNet, after every occurrence of maxpooling, transposed convolution, and concatenation layers (Figure 4.3).

![Figure 4.3: The outline of the designed Unet has been used in the study.](image)
Methodology

A framework for eddy detection and classification in a computationally judicious manner, shown in below pseudo code 4.1 has been designed in this study.

Algorithm 4.1 U-Net Architecture for Eddy Detection

1: Initialize U-Net architecture components
2: for each down-sampling stage do
3: Create DoubleConv block for down-sampling
4: Adjust input channels for next stage
5: end for
6: for each up-sampling stage do
7: Create transposed convolution layer
8: Create DoubleConv block for up-sampling
9: end for
10: Create DoubleConv bottleneck layer
11: Create final convolution layer
12: Initialize skip connections list
13: for each down-sampling block do
14: Apply DoubleConv block
15: Store intermediate output in skip connections
16: Apply max pooling
17: end for
18: Apply bottleneck DoubleConv block
19: Reverse skip connections list
20: for each up-sampling block do
21: Apply transposed convolution
22: Retrieve corresponding skip connection
23: if output shape differs from skip connection shape then
24: Resize output to match skip connection shape
25: end if
26: Concatenate skip connection and output
27: Apply DoubleConv block
28: end for
29: Apply final convolution layer
30: return Final output

The U-Net architecture has been developed using Python 3.9.12 using the TensorFlow and Torch library as backend. The crowd has converted all the images into the same dimen-
sions of 512*512 pixels before feeding the images into the model for training and testing. We also transformed images into grayscale images and another scenario using RGB band to better visualization by TensorFlow. By varying all these above hyperparameters, all three possible variations as mentioned in the base Unet model have been generated and trained on two types of datasets for different ranges epochs, and the performance of the models has been measured and compared based on dice coefficient, accuracy, binary cross-entropy and also other hyperparameters in section 4.4. The segmented outputs corresponding to each trial have been generated. As mentioned in section 3.1.2, cyclonic has cooler than average temperatures compared to the surrounding ocean water or higher intensity or distinctive color variations, in opposite, moderate to high intensity, capturing warm-core temperature anomalies indicative of anticyclonic eddies.

Additionally, to create segmentation masks that correspond to the labeled classes. These masks will have the same dimensions as the original images, with each pixel labeled according to its class. For each input SST image, there is a corresponding target segmentation mask. This mask is created based on manually labeled data, where each pixel is assigned a class label (0, 1, or 2) representing "no eddy," "cyclonic eddy," or "anticyclonic eddy." In the following divide dataset into training, validation, and testing subsets. It’s important to ensure that each subset has a representative distribution of the three classes to avoid bias in the designed model’s performance evaluation. For the U-Net architecture, the input will be the original SST map satellite images, and the desired output will be the corresponding segmentation masks that have been created. Train the U-Net model using the prepared dataset, where the input is SST images and the target is the corresponding segmentation mask. During training, the U-Net learns to predict pixel-wise class labels. Once the model is trained, model can be used it to predict class labels for unseen SST images. Apply the trained model to new SST images to generate segmentation maps.

4.4. Hyperparameters and metric values

4.4.1. Dice score

On the training and test sets for the supervised learning prediction, assessment methods have been used to gauge the effectiveness of the learning model. The eddy net is visible in the output of the Unet DL-based detection framework, displayed as a satellite RS image. To distinguish across the original and predicted eddy-detected results, a comparison involving the output image and the original growth truth image has been generated. The dice coefficient has been considered to be a more logical way to gauge how well the image segmentation problem produced results [126]. It has been utilized to calculate the 0 to
1 range of overlap between the two images. The perfect and total overlap is represented by a dice coefficient value of 1. The dice coefficient has often been utilized to assess the model in current RS imaging issues. In order to measure the effectiveness of image segmentation, the dice score/coefficient has been used. The dice score is a measurement used to determine how similar two items are. The score for a certain class \( C \) is the amount of overlap between the predicted region and its actual counterpart. It is defined as the size of the intersection of the two segmentations divided by the total size of the two objects. The degree of spatial overlap between the predicted segmentation \( Y \) and the ground truth segmentation \( X \) is determined by the dice coefficient. Equation 4.1 is the definition of the dice score/coefficient calculation formula [127].

\[
\text{Dice Coefficient} = \frac{2 \cdot |X \cap Y|}{|X| + |Y|}
\]  

\[ 4.1 \]

4.4.2. Binary cross-entropy loss

Log loss is another name for the binary cross-entropy loss. It gauges how well a categorization model is doing. Its value ranges from 0 to 1. When a data set’s anticipated probability begins to deviate from its actual label, cross-entropy loss grows. A model should have a loss of 0 if we want it to be flawless. When \( M = 2 \) classes are involved in a binary cross-entropy loss, cross-entropy may be calculated as:

\[
- \sum_{c=1}^{M} y_{o,c} \log p_{o,c}
\]

\[ 4.2 \]

where \( M \) is equal to number of classes, \( \log \) indicate to natural log. \( y \) is binary indicator (values vary from 0 to 1) if class label \( c \) is the correct classification for observation \( o \), \( p \) shows predicted probability observation \( o \) is of class \( c \) [128].

4.4.3. F1 score

The F1 score is a well-known metric in binary classification that balances both precision and recall. Precision measures the accuracy of positive predictions, while recall assesses the ability to capture positive instances. In the context of geographical image segmentation, these notions translate to accurately identifying pixels belonging to the target class (high precision) and correctly capturing all instances of the target class (high recall). The F1 score provides a harmonious blend of these two metrics, making it suitable for segmentation tasks where the trade-off between precision and recall is crucial. The primary
focus of this research is to integrate the F1 score as a loss function into the Unet architecture. This involves adapting the Unet’s loss calculation to incorporate the F1 score while optimizing the network’s weights (Eq. 4.3). The F1-based loss can be tailored to consider class imbalances and can be extended for multi-class segmentation tasks. The performance of the Unet architecture with the F1-based loss will be compared against traditional Unet variants utilizing standard loss functions. This research bridges the gap between the Unet architecture and the F1 score, providing a more comprehensive assessment of segmentation accuracy.

\[ F1 = \frac{2 \times \text{True Positives}}{2 \times \text{True Positives} + \text{False Positives} + \text{False Negatives}} \]  \hspace{1cm} (4.3)

where True Positives (TP) is the number of instances that are correctly predicted as positive, The number of instances that are incorrectly predicted as positive has been defined as False Positives (FP), False Negatives (FN) is equal to The number of instances that are incorrectly predicted as negative. As mentioned in first, the F1 score can be computed with recall and precision [129].

\[ F1 = \frac{2 \times \text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}} \]  \hspace{1cm} (4.4)

This formula takes into account both precision and recall, providing a balanced measure of a model’s performance for a specific class. both are calculated mathematically below [129].

\[ \text{Precision} = \frac{\text{True Positives}}{\text{True Positives} + \text{False Positives}} \]  \hspace{1cm} (4.5)

Precision is the ratio of true positive predictions to the total predicted positives. It measures how many of the positive predictions were actually correct [129].

\[ \text{Recall} = \frac{\text{True Positives}}{\text{True Positives} + \text{False Negatives}} \]  \hspace{1cm} (4.6)

Recall is the ratio of true positive predictions to the total actual positives. It measures the ability of the model to correctly capture all instances of the positive class [129].
5 | Result and discussion

The result section of this master thesis presents the outcomes of the U-Net algorithm implementation for the detection and segmentation of mesoscale eddies in the Atlantic Ocean from SST images, they are mentioned below. The U-Net algorithm has been trained on a dataset of referenced images—1888 images as train data—to predict the presence of eddies in test images. The results are evaluated based on the comparison between the predicted and referenced images for each epoch, as well as the metrics of training accuracy, loss function, recall, precision, F1 score, and other hyperparameters mentioned in Section 4.4. The U-Net algorithm is implemented using different parameters and scenarios, such as learning rate, kernel size, width, and height images, to optimize its performance.

Overall, the results presented in this section provide insights into the effectiveness of the U-Net algorithm for the detection and segmentation of mesoscale eddies in ocean RS images. The findings can contribute to the development of more accurate and efficient methods for studying the dynamics of the ocean and its impact on climate and marine ecosystems.

5.1. Testing and training data scenarios

In the pursuit of algorithmic perfection, the imperative of initially subjecting one’s computational methodologies or models to a subset of the data bears notable significance. This practice is underpinned by multifarious reasons that collectively accentuate its strategic value within the broader landscape of data processing and analysis. Foremost, this preliminary endeavor serves as a preliminary crucible to assess the integrity of one’s code and identify potential aberrations before extending its application to the entirety of the dataset. By scrutinizing the algorithm’s performance on a reduced scale, potential discrepancies and glitches can be expeditiously unearthed and rectified. This proactive approach not only conserves valuable time but also mitigates resource expenditure, precluding the necessity to re-run the algorithm comprehensively in the event of unanticipated issues. However, this subset should inherently encompass both training and test data for a holistic evaluation.
In parallel, this preliminary testing regimen cultivates an environment conducive to experimentation and parameter optimization. The scope to maneuver and explore diverse hyperparameters, model attributes such as kernel size and padding, and preprocessing techniques is accentuated within this controlled setting. It provides a sandbox wherein the algorithm’s behavior under distinct configurations can be expeditiously gauged without necessitating the full-scale engagement of the dataset. This iterative process empowers the algorithm’s refinement without the encumbrance of processing the complete dataset at each iteration.

Yet, the crux of this practice rests upon the foundation of representativeness. The subset chosen for testing must faithfully mirror the broader dataset’s characteristics, ensuring that conclusions drawn are translatable to the dataset as a whole. This involves meticulous consideration in the selection process, eschewing randomly continuous choices for more informed decisions. Moreover, the partitioning of distinct subsets for testing and validation is pivotal. This bifurcation circumvents the peril of overfitting and endorses the algorithm’s capacity to generalize adeptly when exposed to novel data.

In synthesis, the prelude of scrutinizing algorithms on a data subset prior to engaging the complete dataset emerges as a pragmatic and astute protocol. It epitomizes the synergy of computational efficacy and methodological soundness. While economizing resources and fostering experimentation in the Unet method, this practice’s integrity is safeguarded by the selection of a representative subset and the partitioning of discrete subsets for validation. In the quest for algorithmic mastery, this practice emerges as a cornerstone, aligning theoretical considerations with operational finesse in the following scenarios.

5.1.1. Eddy detection examination and prediction in assumption one

The hyper-parameters, represented in Table 5.1, have been used in this case are carefully selected to optimize the performance of the model. The learning rate is set to 0.001 to ensure that the model is able to make small adjustments to the weights during training, and the optimization algorithm adjusts the model’s parameters during the training process. The learning rate determines the step size that the optimizer takes during the parameter updates, and a high learning rate may cause the optimizer to overshoot the minimum and diverge, while a low learning rate may result in slow convergence. Therefore, a balance between rapid convergence and stable optimization is important to achieve optimal performance of the neural network. The image size was set to 128x128 to balance computational efficiency and the ability to capture sufficient spatial detail. The model has
5 | Result and discussion

<table>
<thead>
<tr>
<th>First Hyperparameters</th>
</tr>
</thead>
<tbody>
<tr>
<td>Learning rate</td>
</tr>
<tr>
<td>Epochs</td>
</tr>
<tr>
<td>Image width and height</td>
</tr>
<tr>
<td>Batch size</td>
</tr>
<tr>
<td>Kernel size</td>
</tr>
<tr>
<td>Number of workers</td>
</tr>
<tr>
<td>Padding</td>
</tr>
<tr>
<td>Stride</td>
</tr>
</tbody>
</table>

Table 5.1: Hyper-parameters of first assumption

<table>
<thead>
<tr>
<th>First Assumption</th>
</tr>
</thead>
<tbody>
<tr>
<td>Training Accuracy</td>
</tr>
<tr>
<td>Training Loss</td>
</tr>
<tr>
<td>Dice Score</td>
</tr>
<tr>
<td>Precision</td>
</tr>
<tr>
<td>F1 Score</td>
</tr>
<tr>
<td>Recall Class 0</td>
</tr>
<tr>
<td>Recall Class 1</td>
</tr>
<tr>
<td>Recall Class 2</td>
</tr>
</tbody>
</table>

Table 5.2: Results of first assumption

been trained for 80,000 epochs to ensure that it has enough time to converge to a stable solution. A batch size of 3 and 2 workers is used to balance memory usage and training speed. A kernel size of 3 with padding the receptive field of the convolution operation, which is essentially the region of the input data that is considered when computing the output value, and a stride of 1 is used for the convolutional layers to capture both local and global features.

The model achieved promising results with a Dice score of nearly 0.8 and a loss of about 0.2 on the test dataset, demonstrating the effectiveness of the chosen hyper-parameters. Upon analysis of this case, it is found that the dice score initially reached a value of approximately 36.7 percent after 10k epochs but later dropped to 31.99 percent. Additionally, the training loss reached a value of about 0.07 which indicates that the model is effectively learning the features from the input images.

However, further exploration is required to identify the reasons for the drop in Maybe. The study can say that the Dice score is a measure of the similarity between the predicted segmentation map and the reference segmentation map. A Dice score of 31.99 percent means that the predicted segmentation map is 31.99 percent similar to the reference segmentation map. However, it is concerning that the Dice score fell down after 10k
epochs when it reached 36.7 percent. This might indicate that the model started to over-fit the training data after 10k epochs, leading to a decrease in the accuracy of the predictions on the test data.

Furthermore, the training loss of 0.07 indicates that the model has achieved a good level of convergence during training. A lower training loss means that the model is better at predicting the correct segmentation labels. However, it is important to note that a low training loss does not always guarantee good performance on the test data. The model’s performance on the test data should always be evaluated using metrics such as the Dice score.
5.1.2. Eddy detection examination and prediction in assumption two

In this case, the recommended learning rate of 0.001 is for the optimization of the neural network during the training process. Specifically, it aims to balance between achieving a fast convergence of the model’s parameters to a minimum of the loss function and maintaining stable optimization without overshooting the minimum. The image width and height of 512 is a good choice as it provides enough detail for the model to learn while still being computationally feasible. An epoch of 30k is a long training duration but can be appropriate depending on the complexity of the model and the size of the dataset. A batch size of 3 can help to avoid overfitting and ensure that each batch is representative of the overall dataset. With two workers, the data can be efficiently loaded into the model during training, leading to faster training times. In addition, a kernel size of 3 with padding of 1, because a padding value of 1 has been used along with a kernel size of 3 and a stride of 1, so Padding is used to add extra pixels around the edges of an image, which helps to preserve the spatial dimensions of the output feature maps. In this case, the padding value of 1 helps to ensure that the output feature maps have the same spatial dimensions as the input images. This is important because it enables the model to learn features that are relevant to the task at hand without losing any information due to the reduction in spatial dimensions that can occur when using larger kernel sizes or strides, and the stride of 1 would be appropriate for the task. Overall, these hyperparameters can provide an appropriate balance between model performance and training efficiency when using the entire test dataset, as shown in all mentioned parameters in Table 5.3.

After training this option, a dice score of nearly 40 percent and a loss of about 0.6 indicate that the performance of the model is reasonable. A performance level of approximately 40 percent implies that the model has successfully grasped certain essential characteristics of the objects being targeted in the images. However, there remains ample opportunity for enhancement. The recorded loss value of 0.6 suggests that the model isn’t excessively focused on the training data and is continuing to learn from it. Enhancements to the model’s effectiveness can be achieved through adjustments to hyperparameters, refining data preprocessing procedures, enhancing augmentations, and augmenting the volume of training data. Based on mentioned parameter in Table 5.4, there could be several reasons why the model does not reach a better dice score and loss with the given hyperparameters. One possibility is that the chosen learning rate is not optimal for this specific dataset and architecture, which could have led to slower convergence and suboptimal performance. An alternative explanation could be that the current number of training epochs might not
### Second Hyperparameters

<table>
<thead>
<tr>
<th>Hyperparameter</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Learning rate</td>
<td>0.001</td>
</tr>
<tr>
<td>Epochs</td>
<td>30000</td>
</tr>
<tr>
<td>Image width and height</td>
<td>512</td>
</tr>
<tr>
<td>Batch size</td>
<td>3</td>
</tr>
<tr>
<td>Kernel size</td>
<td>3</td>
</tr>
<tr>
<td>Number of workers</td>
<td>2</td>
</tr>
<tr>
<td>Padding</td>
<td>1</td>
</tr>
<tr>
<td>Stride</td>
<td>1</td>
</tr>
</tbody>
</table>

Table 5.3: Hyper-parameters of second assumption

<table>
<thead>
<tr>
<th>Second Assumption</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Training Accuracy</td>
<td>70.19</td>
</tr>
<tr>
<td>Training Loss</td>
<td>0.586</td>
</tr>
<tr>
<td>Dice Score</td>
<td>39.22</td>
</tr>
<tr>
<td>Precision</td>
<td>0.8783</td>
</tr>
<tr>
<td>F1 Score</td>
<td>0.3922</td>
</tr>
<tr>
<td>Recall Class 0</td>
<td>0.8406</td>
</tr>
<tr>
<td>Recall Class 1</td>
<td>0.1954</td>
</tr>
<tr>
<td>Recall Class 2</td>
<td>0.1618</td>
</tr>
</tbody>
</table>

Table 5.4: Results of second assumption

be adequate for achieving convergence. Extending the training duration could potentially lead to improved performance outcomes.
Regarding the variation observed in the upper right corner of the predicted images (Figure 5.2), this dissimilarity can arise from several contributing factors. It’s plausible that the training dataset lacked a sufficient number of instances with analogous attributes in that specific region, potentially resulting in an inadequate depiction of that area within the model. Another possibility is that the model architecture itself is not capable of
capturing the specific features present in that region, and a more complex or specialized architecture might be needed. It could also be a result of the chosen hyperparameters, such as the kernel size, padding, and stride, which could affect the model’s ability to capture fine details in the image. Further investigation and experimentation with different hyperparameters and architectures might help to improve the model’s performance and address these issues.

5.2. **Contrasting Outcomes with Comparable Research Endeavors**

There are various studies proceeding on the same study, in this thesis, based on limitations in time and master acknowledgment, has been tried to reach logical and sufficient results. If compare two assumptions result to other related studies mentioned in Section 2.2.

The first study, done by Lguensat et al. 2017 [42], introduces EddyNet, an innovative DL architecture designed for the automated identification and categorization of eddies using SSH maps sourced from the Copernicus Marine and Environment Monitoring Service (CMEMS). EddyNet comprises a convolutional encoder-decoder framework, which is subsequently followed by a pixel-wise classification layer. The resulting output is a map of the same dimensions as the input, wherein individual pixels are assigned labels denoting the following categories: '0' for Non eddies, '1' for anticyclonic eddies, and '2' for cyclonic eddies.

In the mentioned study has been utilized the Keras framework with a Tensorflow backend as the foundation for this work. EddyNet has trained on an Nvidia K80 GPU card employing the ADAM optimizer and mini-batches consisting of 16 maps. For weight initialization, The crowed has employed truncated Gaussian distributed weights with a mean of zero and a variance of 2/number of input units for EddyNet, while EddyNet S used weights drawn from a truncated Gaussian distribution with a mean of zero and a variance of 1/number of input units. The study’s training dataset is divided into 4080 images for training and 1020 for validation. Additionally, it has incorporated an early-stopping strategy, which terminated the learning process when the validation dataset loss failed to improve over five consecutive epochs. The final weights for EddyNet are those associated with the lowest validation loss.

Furthermore, they have conducted a comparative analysis between EddyNet and EddyNet S, focusing on the utilization of classical ReLU+BN and SELU activation functions and have also compared the use of the Dice Loss, an overlap-based metric, with the con-
Table 5.5: Comparison of Training Results for EddyNet and EddyNet S by Lguensat et al. 2017 [42].

<table>
<thead>
<tr>
<th>Method</th>
<th>Train Loss</th>
<th>Mean Dice Coef</th>
<th>Global Accuracy</th>
</tr>
</thead>
<tbody>
<tr>
<td>EddyNet</td>
<td>Dice Loss</td>
<td>0.772</td>
<td>88.60%</td>
</tr>
<tr>
<td></td>
<td>CCE</td>
<td>0.762</td>
<td>89.92%</td>
</tr>
<tr>
<td>EddyNet S</td>
<td>Dice Loss</td>
<td>0.764</td>
<td>88.98%</td>
</tr>
<tr>
<td></td>
<td>CCE</td>
<td>0.758</td>
<td>89.83%</td>
</tr>
</tbody>
</table>

conventional Categorical Cross-Entropy (CCE) loss function, but in thesis study has been employed ReLU as activation function because of mentioned reason in Section 4.3. Table 5.5 summarizes the results of these four combinations in terms of global accuracy and mean Dice coefficient, which is computed on 50 random sets of 360 SSH maps measuring $120 \times 120$ from the year 2012. Overall, the Dice Loss resulted in a superior mean Dice coefficient compared to training with the CCE loss. Regarding the impact of the activation function, our research showed that EddyNet achieved better metrics, albeit at the cost of a longer training duration. Regarding the comparison of the model designed in this study and thesis, both studies have the same accuracy rate, but the designed model in this study has performed better in the Dice score coefficient.

This paper [40] introduces the Pyramid Split Attention Eddy Detection Unet (PSA-EDUNet) architecture, abbreviated as PSA-EDUNet, with a focus on identifying oceanic eddies in remote sensing imagery of the ocean. The inspiration behind PSA-EDUNet draws from the Unet framework, known for its encoder and decoder components, which effectively integrate lower and higher-level features while preserving valuable feature information through non-linear connections. Furthermore, the PSA module is introduced to enhance feature extraction within the network. When it comes to fusing data for eddy identification, the primary criteria revolve around sea surface characteristics, specifically SST and SLA. The experimental evaluation is conducted in the Kuroshio Extension (KE) and South Atlantic (SAO) regions, specifically within the longitude range of 3.875-59.875° W and latitude range of 5.375-61.375° S. The results of this research showcase the superiority of the proposed method when compared to other existing approaches, particularly in the context of detecting eddy boundaries and small-scale eddies.

The RMSprop optimizer is used to optimize the experiment in this study. The loss function is the cross entropy loss function, and the learning rate is le-3. Each small batch is composed of 8 images. Additionally, the training set loss stopped improving for 30 consecutive epochs, the strategy of reducing the learning rate will be employed to improve the learning process. The study’s code is written using the Pytorch DL library. Training
Table 5.6: Performance Metrics for PSA-EDUNet on Kuroshio Extension (KE) and South Atlantic Ocean (SAO) [40]

<table>
<thead>
<tr>
<th>Metric</th>
<th>KE</th>
<th>SAO</th>
</tr>
</thead>
<tbody>
<tr>
<td>Accuracy</td>
<td>94.06%</td>
<td>94.18%</td>
</tr>
<tr>
<td>Precision</td>
<td>88.72%</td>
<td>89.13%</td>
</tr>
<tr>
<td>Recall</td>
<td>88.62%</td>
<td>88.89%</td>
</tr>
</tbody>
</table>

is performed using an Nvidia GeForce RTX 3090 GPU. To confirm the performance of the PSA-EDUNet, the comparative experiment results are presented in Table 5.7. In the KE region, the accuracy of the paper’s model can reach 0.9406, which is 0.1556 higher than the thesis experiment; while in the SAO region, the accuracy of the mentioned model is 0.9418, which is increased by 0.1568 respectively based on the thesis models. The proposed PSA-EDUNet has the most brilliant performance. Our Unet model precision has comparatively better performance in eddy identification.

Table 5.7: Performance Metrics for thesis outcomes in the study area.

<table>
<thead>
<tr>
<th>Metric</th>
<th>Atlantic Ocean</th>
</tr>
</thead>
<tbody>
<tr>
<td>Accuracy</td>
<td>78.5%</td>
</tr>
<tr>
<td>Precision</td>
<td>92.5%</td>
</tr>
<tr>
<td>Recall</td>
<td>69.4%</td>
</tr>
<tr>
<td>Dice score</td>
<td>39.2%</td>
</tr>
</tbody>
</table>
Conclusion

The appealing field of applying recent developments in DL-based image segmentation for a difficult ocean RS profession, especially the identification and classification of eddies from SST maps, has been explored in this paper. Our original contribution, EddyNet, is a deep neural network architecture that is influenced by well-known designs and ideas that are well-liked in the computer vision field. This study has worked severely to apply this understanding to the process of eddy classification, overcoming a variety of challenging obstacles along the way.

This work focuses on the use of state-of-the-art feature representation backbones in conjunction with cutting-edge semantic segmentation technology to tackle the problem of autonomously recognizing ocean eddies in the northern part of the Atlantic Ocean. We incorporated the Unet module into our system in an effort to obtain more accurate boundary information and detect mesoscale eddies more successfully. Our research has shown that the high-resolution representation technique outperforms other models already in use for ocean eddy identification. This cutting-edge representation technique represents a significant advancement in existing technology and promises improved eddy detection and border delineation accuracy and dependability.

Future research projects will focus on utilizing temporal quantities of SST data, as suggested by the encouraging results in the preceding sections. Using the knowledge we have gathered so far, we intend to create a 3D version of our model. Additionally, we are thinking about including other surface data, including SSH, which might improve our eddy identification skills. Additionally, we as a community hope to expand the geographic reach of EddyNet while evaluating its functionality and applicability in many parts of the world. This broadening could show how useful it is outside the boundaries of our present research topic. It is important to note that post-processing activities, such as verifying discovered eddies against additional standards and creating eddy tracking techniques, have not been looked into in this work. These improvements have the potential to improve and expand the capabilities of our strategy.
Bibliography


[87] F. Dufois, N. J. Hardman-Mountford, J. Greenwood, A. J. Richardson, M. Feng, and R. J. Matear, “Anticyclonic eddies are more productive than cyclonic eddies in


[108] F. Nencioli, C. Dong, T. Dickey, L. Washburn, and J. C. McWilliams, “A vector geometry-based eddy detection algorithm and its application to a high-resolution numerical model product and high-frequency radar surface velocities in the southern


[120] D. B. Zwiefelhofer. (No Date) Find latitude and longitude. Website: FindLatitude-


Abbreviations

Sea Surface Temperature ... SST
Remote Sensing ... RS
Machine Learning ... ML
Conventional Neural Network ... CNN
U network ... Unet
Artificial Intelligence ... AI
Sea Level Anomaly ... SLA
Chlorophyll-a ... Chl-a
Indian Ocean Dipole ... IOD
South Eastern Arabian Sea ... SEAS
Isotherms Layer Depth ... ITD
Upper Ocean Heat Content ... HCT
Sea Surface High ... SSH
Latent Heat Flux ... LHF
Sensible Heat Flux ... SHF
You Only Look One Level Feature ... YOLO
Extreme Gradient Boosting ... XGBoost
Altimeter Identification Eddies ... Alt Eddy
Karoshio Extension ... KE
Copernicus Marine and Environmental Monitoring Service ... CMEMS
Residual network ... ResNet
Vector Geometry-Based ... VG
Pyramid Scene Parsing Network ... PSPNet
Bilateral Segmentation Network ... BiSeNet
Deep Laboratory V3+ ... DeepLabV3+
Deformation Radius ... RD
Warm Anticyclonic Eddies ... WAEs
Cold Cyclonic Eddies ... CCEs
Gulf Stream Meander Region ... GSMR
Antarctic Circumpolar Current ... ACC
Rectified Linear Unit ... ReLU
Advance Along-Track Scanning Radiometer ... (A)STSR
Sea and Level Surface Temperature Radiometer ... SLSTR
Advance Very High-Resolution Radiometer ... AVHRR
Operational Sea Surface Temperature and Sea Ice Analysis ... OSTIA
European Organization for the Exploration of Meteorological Satellite ... EUMETSAT
Ocean and Sea Ice Satellite Application Facility ... COSI-SAF
Scaled Exponential Linear Unit ... SELU
## List of Figures

<table>
<thead>
<tr>
<th>Figure</th>
<th>Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Graphical abstract as outlined in this study.</td>
<td>1</td>
</tr>
<tr>
<td>3.1</td>
<td>Sea surface temperature anomalies, credited by NASA [77].</td>
<td>18</td>
</tr>
<tr>
<td>3.2</td>
<td>Diagrams of cyclonic and anticyclonic eddies in three dimensions are shown in (a) and (b), respectively. The red arrow indicates the direction of the eddy current in the northern hemisphere. The various hues of color show the level of water column temperature, with the red area having the greatest temperature and the blue part having the lowest [79].</td>
<td>19</td>
</tr>
<tr>
<td>3.3</td>
<td>Ocean Models: Representation The issue of Mesoscale Eddies [97].</td>
<td>21</td>
</tr>
<tr>
<td>3.4</td>
<td>SST map on 1 January 2017 in the study area.</td>
<td>22</td>
</tr>
<tr>
<td>3.5</td>
<td>Gulf Stream Currents [98].</td>
<td>23</td>
</tr>
<tr>
<td>3.6</td>
<td>Input and output of image segmentation in the study area.</td>
<td>25</td>
</tr>
<tr>
<td>3.7</td>
<td>U-net simple sample architecture, a multiple channels characteristic map correlates to each blue box. On the upper part of the element, there is a channel count indicator. At the lower left corner of the sample, the i-j dimension is displayed. Replicated feature maps are represented by white boxes. The various processes are shown by the indications. [119].</td>
<td>30</td>
</tr>
<tr>
<td>4.1</td>
<td>Graphical abstract as outlines in this study.</td>
<td>33</td>
</tr>
<tr>
<td>4.2</td>
<td>Study area.</td>
<td>34</td>
</tr>
<tr>
<td>4.3</td>
<td>The outline of the designed Unet has been used in the study.</td>
<td>38</td>
</tr>
<tr>
<td>5.1</td>
<td>The visualization of outcomes by TensorFlow in assumption one.</td>
<td>46</td>
</tr>
<tr>
<td>5.2</td>
<td>The visualization of outcomes by TensorFlow in assumption two.</td>
<td>49</td>
</tr>
</tbody>
</table>
## List of Tables

<table>
<thead>
<tr>
<th>Table</th>
<th>Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>4.1</td>
<td>Data information abstraction [125].</td>
<td>36</td>
</tr>
<tr>
<td>5.1</td>
<td>Hyper-parameters of first assumption</td>
<td>45</td>
</tr>
<tr>
<td>5.2</td>
<td>Results of first assumption</td>
<td>45</td>
</tr>
<tr>
<td>5.3</td>
<td>Hyper-parameters of second assumption</td>
<td>48</td>
</tr>
<tr>
<td>5.4</td>
<td>Results of second assumption</td>
<td>48</td>
</tr>
<tr>
<td>5.5</td>
<td>Comparison of Training Results for EddyNet and EddyNet S by Lguensat et al. 2017 [42].</td>
<td>51</td>
</tr>
<tr>
<td>5.6</td>
<td>Performance Metrics for PSA-EDUNet on Kuroshio Extension (KE) and South Atlantic Ocean (SAO) [40]</td>
<td>52</td>
</tr>
<tr>
<td>5.7</td>
<td>Performance Metrics for thesis outcomes in the study area.</td>
<td>52</td>
</tr>
</tbody>
</table>
## List of Symbols

<table>
<thead>
<tr>
<th>Variable</th>
<th>Description</th>
<th>SI unit</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\rho$</td>
<td>Water density</td>
<td>kg/m$^3$</td>
</tr>
<tr>
<td>$V$</td>
<td>Volume</td>
<td>m$^3$</td>
</tr>
<tr>
<td>$m$</td>
<td>Mass</td>
<td>Kg</td>
</tr>
<tr>
<td>$Y$</td>
<td>Predicted Segmentation</td>
<td></td>
</tr>
<tr>
<td>$X$</td>
<td>Ground truth Segmentation</td>
<td></td>
</tr>
<tr>
<td>$M$</td>
<td>Number of Classes</td>
<td></td>
</tr>
<tr>
<td>$y$</td>
<td>Binary Indicator</td>
<td></td>
</tr>
<tr>
<td>$c$</td>
<td>Correct Classification for Observation $o$</td>
<td></td>
</tr>
<tr>
<td>$p$</td>
<td>Prediction Probability Observation is of class $c$</td>
<td></td>
</tr>
</tbody>
</table>
Acknowledgements

I would like to express my deepest gratitude to my thesis supervisor, Prof. Dr. Giovanna Venuti for her guidance and support throughout the duration of my master’s thesis. Her expertise and valuable feedback have been instrumental in shaping this research project on mesoscale eddy detection and classification from sea surface temperature maps with deep neural networks.

I would also like to thank the staff and faculty of Bonn University, Germany, and Politecnico di Milano, Italy, particularly the Erasmus program, for providing me with the necessary resources and facilities to carry out this nice communication between universities. Additionally, I extend my appreciation to my colleagues and friends who have been a constant source of motivation and encouragement throughout this journey.

Furthermore, I would like to express my gratitude to the data providers for making the sea surface temperature data available, without which this research would not have been possible.

Finally, I am deeply grateful to my family for their unwavering support, love, and encouragement, which have been the driving force behind my academic pursuits.