Mesoscale Eddy Detection and Classification From Sea Surface Temperature Maps With Deep Neural Networks

Mohammad Mahdi Safari, Alireza Sharifi, Javad Mahmoodi, and Dariush Abbasi-Moghadam

***Abstract*—Oceanic eddies are a widespread and important occurrence that plays a vital role in the movement of chemicals and energy within the marine ecosystem. Hence, the astute and precise recognition of these swirling currents may greatly contribute to the progress of our comprehension of oceanography. Due to the continuous breakthroughs in state-of-the-art deep learning technology, the population is witnessing a progressive improvement in the methods used to identify and understand these aquatic characteristics. This study employs sea surface temperature data acquired from the Copernicus Marine and Environment Monitoring Service (CMEMS) in the Atlantic Ocean. The objective is to present EddyNet, a cutting-edge deep-learning framework specifically developed for the automatic identification and categorization of ocean eddies. EddyNet incorporates a pixel-wise classification layer into its neural encoder-decoder architecture. The resulting output is a map that maintains the same dimensions as the input, but each individual pixel is assigned a label indicating its classification as either "0" for noneddy regions, "1" for anticyclonic eddies, or "2" for cyclonic eddies. We propose a new image segmentation method based on the U-net architecture with different convolutional neural network backbones such as VGG16, VGG19, DenseNet121, and MobileNetV2. Our models are built and trained using Python and the Keras library with the Adam optimizer for improved convergence. Our approach uses sparse categorical cross-entropy as the loss function, simplifying the label encoding process for multiclass classification with sparse labels. Initial results show that this method achieves a good balance between computational efficiency and segmentation accuracy, making it suitable for real-time applications.**

***Index Terms*—DenseNet121, MobileNetV2, remote sensing (RS), semantic segmentation U-net, VGG-Net.**

Manuscript received 8 April 2024; accepted 15 May 2024. Date of publication 20 May 2024; date of current version 30 May 2024. This work was supported by Shahid Rajaee Teacher Training University under Grant 4941. *(Corresponding authors: Alireza Sharifi; Dariush Abbasi-Moghadam.)*

Mohammad Mahdi Safari is with the Geoinformatics Engineering Department, Politecnico di Milano, 20133 Milano, Italy (e-mail: mohammadmahdi.safari@mail.polimi.it).

Alireza Sharifi is with the Department of Surveying Engineering, Faculty of Civil Engineering, Shahid Rajaee Teacher Training University, Tehran 16788-15811, Iran (e-mail: a_sharifi@sru.ac.ir).

Javad Mahmoodi is with the Department of Electrical Engineering, Kerman Branch, Islamic Azad University, Kerman 7635168111, Iran (e-mail: javad.mahmoodi@iauk.ac.ir).

Dariush Abbasi-Moghadam is with the Electrical Engineering Department, Shahid Bahonar University of Kerman, Kerman 76169-14111, Iran (e-mail: abbasimoghadam@uk.ac.ir).

Digital Object Identifier 10.1109/JSTARS.2024.3402823

I. Introduction

RECENTLY, crucial occupations involving the understanding of climate change have focused on the detection and comprehension of the escalating phenomenon of global warming, with a specific emphasis on the prediction and surveillance of sea level temperature [1], [2]. Several deep learning (DL) and machine learning (ML) approaches have been developed using graphical and spatial analytic methodologies [3], [4]. These inquiries may provide vital proof demonstrating that the alterations are inducing the elevation of oceans and sea levels. Oceanographic simulations include the study and prediction of ocean currents and streams [1]. Accurate modeling of ocean dynamics and turbulence is crucial in order to enhance our understanding and prediction of climate change and sea level fluctuations [5]. The physical processes responsible for ocean currents have a significant influence on ocean dynamics and the general movement of the ocean [6].

Current patterns of ocean motion have been significantly influenced by the effects of changes in the climate and worldwide warming in recent times. Various oceanic locations often influence the impact of sea surface temperature (SST) fluctuations on anomalies in atmospheric circulation [7]. Mesoscale ocean eddies have the ability to affect the dynamics of the atmosphere at mesoscales mainly via kinetic energy. They also have a regional influence on near-surface wind, cloud features, and rainfall [8]. The ocean's stratification is influenced by substantial currents and the resulting instability on the oceanic mesoscale, which spans around 100 km [9]. These currents also contribute to the transport of heat toward the poles in the climate system [10], [11].

Researchers use to analyze and study eddies to get insights into their impact on ocean climate models [12]. The introduction of altimeter operations and the availability of many altimeters at the same time allowed for the detection of mesoscale eddies in the SST combined objects, leading to sufficient decisions [13], [14]. 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, but the main issue is to establish suitable criteria and thresholds for monitoring mesoscale eddies among physical and mathematical attributes.

The ocean exhibits dynamic, circular currents that have a diameter ranging from 10 to 500 km [13], [15]. These characteristics are crucial in the movement of ocean currents, the distribution of biogeochemical substances, and the management of climate. In addition, we have strengthened the context by using specific instances, recent research discoveries, and a formal style to present a thorough explanation of the importance of mesoscale eddies in the marine system, namely in the North Atlantic area. In this study, SST is considered as an attribute to assist in setting certain maps for analyzing remote sensing (RS) satellite images, which are called SST maps. Anticyclonic and cyclonic eddies are commonly believed to be linked to atypical warm and cold surface waters, respectively [16], [17]. Through the utilization of satellite images, we present evidence that the global ocean contains a significant number of anticyclonic and cyclonic eddies based on warm and cold core water streams in the Gulf Stream. They are also characterized by potential vorticity that is more pronounced near the surface. This study examines a specific type of anticyclonic eddies and cyclonic eddies that are characterized by subsurface-intensified potential vorticity and SST anomalies. Specifically, cyclonic eddies are typically associated with cooler temperatures due to the upwelling of deeper, nutrient-rich waters, while anticyclonic eddies are linked to warmer temperatures resulting from the downwelling of surface waters. These contrasting effects on SST, visible through satellite imagery, are crucial in understanding ocean circulation and marine ecology. Cyclonic eddies are marked by their counterclockwise rotation in the Northern Hemisphere, promoting biological productivity by bringing cold, nutrient-laden water to the surface, while anticyclonic eddies, rotating clockwise, often encapsulate and transport warmer surface waters, impacting local marine life and oceanic carbon uptake differently [18], [19]. Due to their high prevalence, it is necessary to accurately consider and quantify the contribution of cyclonic and anticyclonic when evaluating and defining the impact of oceanic eddies on the global climate.

Potential vorticity is a quantifiable characteristic that combines the rotating motion of a fluid with the consequences of stratification. Understanding the mechanics of fluid motion, especially in geophysical flows like ocean currents, is greatly facilitated by this helpful trait. This phenomenon is particularly evident in the upper layers of the ocean within these small-scale rotating currents, and it can offer vital insights into the intensity and origin of these features [1], [20]. Acquiring an understanding of the attributes and actions of these intermediate-sized swirling currents is essential for enhancing our climate models and precisely predicting changes in the environment. This study employs state-of-the-art neural networks to detect and classify these swirling currents from SST maps, providing a crucial instrument for oceanographers and climatologists.

Consequently, it is essential to comprehend the transmission and characteristics of mesoscale ocean eddies and use this knowledge to detect and examine them, as it has significant importance for worldwide climate change as well as geological and oceanographic studies. This study utilizes four conventional neural networks (CNNs) built using the U-net architecture contains: VGG16, VGG19, DenseNet121, and MobileNetV2 to assess preprocessed satellite images taken in the Atlantic Ocean, specifically focusing on Gulf Stream data in 2019.

This article is divided into six segments. Setion II presents the related works. In Section III, an outline of the suggested approaches and designed models is presented. Also, Section IV emphasizes implementation, findings, and derived outcomes. In Section V, innovations and limitations in this article and outlines of prospective research are discussed. And Section VI offers final remarks and applications of the proposed methodology.

II. RelatedWorks - State-of-the-Art

In the field of computer science, specifically in the area of DL, the task of dealing with a limited amount of labeled data for neural networks requires the use of creative ways to generate or enhance data. This has a substantial effect on traditional methods of identifying mesoscale eddies and their applications in spatial planning and RS. One of the most recent novel research methods focuses on identifying ocean mesoscale eddies for deep transfer learning target recognition. The method is based on YOLOF (You Only Look One Level Feature), which is a widely used framework for goal detection in DL. The study highlights the significant advantages of YOLO series goal detection frameworks in the field of DL. The suggested method surpasses current classification approaches in terms of detection performance, eliminates the impact of threshold change on mesoscale eddy identification, and somewhat enhances identification speed [21], [22].

Moreover, the 3D neural network with a focus on the vertical structure of eddies is used to classify ocean eddies. This study is one of the early attempts to evaluate the effectiveness of DL in identifying vertically arranged eddies. The improved eddy profiles database is generated by using a vertical profile framework that is tightly linked to altimetry sea surface topography. The following is a description of a 3D neural network that focuses on the vertical structure of eddies. This network is based on the residual network (ResNet) and is used to classify eddies as either anticyclonic, cyclonic, or nonexistent. The proposed network might include geographical and dynamic properties as external variables. The proposed network has the capability to represent 3D eddy data and may be extended to a more complex network structure by including 3D convolutions and pooling. Ultimately, classification testing validates the proposed methodology. An especially notable result from the experiments is that the proposed method may improve the capacity to identify eddies by using altimetry-calibrated vertical characteristics with similar classification effectiveness [23].

Another comparable research has used a strategy that combines many approaches. A new combination strategy that

leverages CNN and extreme gradient boosting (XGBoost) advantages may be used to extract the usual vertical feature and eddy from an image. The CNN utilizes the input data analyzed at the lower layers of the neural network to extract vertical features. Furthermore, the model called XGBoost utilizes the complex vectors of variables and account features to classify profiles that are not associated with altimeter-identified eddies (Alt eddy) [24].

The subsequent correspondence introduces the pyramid split attention (PSA) eddy detection U-net architecture (PSA-EDUNet) designed specifically for the purpose of identifying maritime imaging eddies. The PSA-EDUNet has been developed based on the U-net architecture. It has encoder and decoder components to effectively combine low-level and high-level characteristics and avoid the loss of characteristic data via nonlinear connections. Principal component analysis enhances the process of extracting features. The primary criteria for identifying eddies in fusion data are SST and sea level anomaly (SLA). The Kuroshio Extension (KE) and South Atlantic investigations demonstrate that the proposed technology surpasses existing methodologies, particularly in terms of eddy edges and small-scale eddies [25].

The accompanying letter presents developed DL approaches for detecting ocean eddies using semantic segmentation. Contextually efficient pixel-level detection is necessary for semantic segmentation. Two attention modules are designed to tackle this problem. The VGG16-based U-net design, equipped with two attention modules, effectively demonstrates the contextual relationship between the distribution and geographic dimensions. Each pixel or bandwidth incorporates context from the others, as per the links. A residual route is used to substitute the skip link that connects the encoder and decoder units. The results demonstrate that the use of an attention-based deep framework and a unique residual methodology enhances the effectiveness of models compared to the most advanced methods currently available [26]. Duo et al. [27] have developed OEDNet, an oceanic mesoscale eddy automated recognition and localization network. OEDNet leverages DL techniques to enhance previous detection methods. This study utilizes research findings on eddies to address the detection of mesoscale eddies. The approach involves using a training database that contains SLA contour maps with labeled eddy sites. A multilayer deep neural network is constructed and taught to accurately identify the center of mesoscale eddies and generate the corresponding contour. In addition, Aguedjou et al. [16] study the influence of mesoscale eddies on the interaction between the atmosphere and the ocean, focusing on their effects on the exchange of heat and freshwater in the tropical Atlantic Ocean. Their study utilizes a substantial amount of satellite altimetry data spanning multiple years, in conjunction with SST data, to gain a comprehensive understanding of the influence of eddies in these phenomena.

Recent explorations and advancements in DL methodologies have highlighted a research gap in the automation method and precision of mesoscale eddy identification from SST data, specifically utilizing U-net frameworks with varying backbones for enhanced detection and classification in the Gulf Stream of the Atlantic Ocean-based warm and cold core water. Regarding the gap in the previous for detecting and classification and localization of mesoscale in the North Atlantic Ocean based on SST in the Gulf Stream. So, this study has conducted an evaluation of U-net architectures as the principal approach, consisting of a convolutional encoder-decoder followed by a pixel-wise classification layer based on VGG16, VGG19, DenseNet121, and MobileNetV2. The initial visualization involved plot maps with the same size as the input, where pixels have been labeled as follows: "0" for noneddy, "1" for anticyclonic eddy, and "2" for cyclonic eddy. Subsequently, training datasets and EddyNet weights files have been generated. Moreover, our findings signify the growing collaboration between the RS and ML communities, which has resulted in substantial contributions toward addressing the segmentation of SST images. The U-net architectures and implementation models are specifically discussed in the dedicated in the proposed method of the study.

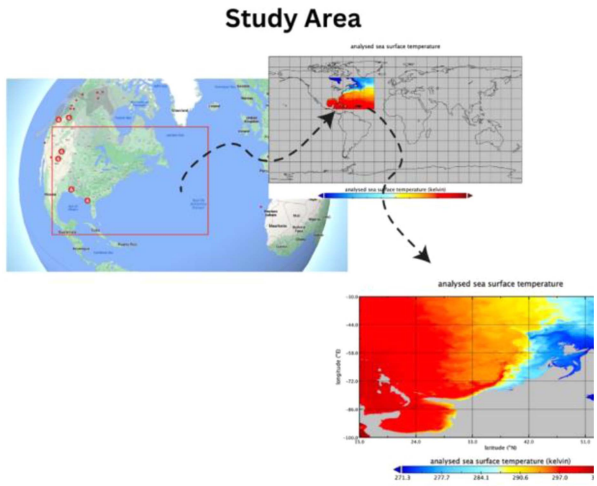


Fig. 1. Study area: North Atlantic Ocean.

III. Material and Methods

A. Study Area

The study area region covers a large portion of the North Atlantic Ocean, spanning from latitude $17.32^{\circ}N$ to $55.50^{\circ}N$ and longitude $41.13^{\circ}W$ to $95.63^{\circ}W$, as shown in Fig. 1 [28]. This vast expanse includes notable geographical features such as the Hudson Bay, the Gulf of Mexico, and the Labrador Sea. Moreover, it spans from the western African shoreline to the eastern coast of the Americas. The North Atlantic Ocean, while often seen as a mostly flat body of water, does exhibit several discernible oceanic traits [29]. An exemplary instance is the Gulf Stream, a vigorous warm oceanic circulation that moves in a northeasterly trajectory over the eastern North American shoreline. The present oceanic conditions have a substantial impact on the climate of adjacent coastal regions. A diverse array of aquatic organisms flourishes in this expansive region, including several fish species, marine invertebrates, and seabirds [30].

The complex marine ecosystems seen in this area result from the delicate interaction between different water currents and

TABLE I
SATELLITE DATASET INFORMATION CREDITED BY COPERNICUS [31]

Dataset classified information	
Full name	ESA SST CCI and C3S reprocessed SST analyses
Product ID	SST GLO SST L4 REP OBSERVATIONS 010 024
Source	Satellite observations
Spatial extent	Global ocean Lat −90° to 90°Lon −180° to 180°
Spatial resolution	0.05° × 0.05°
Temporal extent	1 September 1981 to 31 October 2022
Temporal resolution	Daily
Processing level	Level 4
Variables	Sea Ice Area Fraction Sea Water Temperature (T)
Feature type	Grid
Projection	WGS 84 (EPSG:4326)
Format	NetCDF-4

masses. The North Atlantic has historically been crucial in enabling trade and transit between Europe and North America. The region's strategic significance has resulted in significant marine involvement, emphasizing its essential role as a critical commerce route. This vast research area encompasses several islands and beaches. These include notable geographical areas such as the Azores, a cluster of Portuguese islands, portions of the Caribbean archipelago, and the eastern peripheries of the United States and Canada [30].

The North Atlantic has a negligible impact on the climate of adjacent land regions, particularly in Europe. Nevertheless, the North Atlantic Drift, which is an integral component of the broader North Atlantic Current system, does indeed influence temperatures and contribute to the formation of the European climate. The North Atlantic Ocean receives significant focus for vast research endeavors due to its crucial influence on climate dynamics, marine biology, and oceanography. The many attributes and extensive impacts of this issue make it a captivating topic for scientific investigation and analysis.

B. Satellite Data Observation

The fundamental dataset supporting this study is the ESA SST CCI and C3S global SST Reprocessed product. Table I shows the satellite dataset information. This dataset is crucial for comprehending and examining global-scale SST dynamics. The dataset provides a detailed representation of the daily average SST at a depth of 20 cm, presented in a spatial grid resolution of 0.05° by 0.05°. The dataset is carefully compiled using satellite data obtained from three different sources: the Advanced Along-Track Scanning Radiometer (AATSRs), the Sea and Land Surface Temperature Radiometer (SLSTR), and the Advanced Very High-Resolution Radiometer (AVHRR) series of sensors, as thoroughly explained by the work in [31] and [32]. The combination of several satellite sensors yields a comprehensive and varied dataset, effectively capturing the intricacies of SST fluctuations across various geographical areas.

TABLE II
PERFORMANCE METRICS COMPARISON OF CONVOLUTIONAL NEURAL NETWORK MODELS

	Unet-VGG16	Unet-VGG19	Unet-Mobile NetV2	Unet-DenseNet121
Precision	88.97	84.79	91.37	**96.05**
Recall	83.31	78.67	90.01	**95.56**
F1 score	86.03	81.58	90.67	**95.79**
Accuracy	96.77	95.84	97.98	**99.37**
Dic class0	98.11	97.52	98.73	**99.43**
Dic class1	86.02	81.55	90.67	**95.79**
Dic class2	86.14	81.99	91.11	**95.98**
IOU class0	96.21	95.07	97.44	**98.84**
IOU class1	71.74	64.038	80.45	**91.47**
IOU class2	71.83	64.72	81.15	**91.6**

This table presents a comparative analysis of Unet-VGG16, Unet-VGG19, Unet-DenseNet121, and Unet-MobileNetV2 in terms of precision, recall, F1 score, accuracy, Dice coefficient, and IoU for classifying different eddy conditions.
Bold values represent the best significant results.

The ESA SST dataset has been selected for its superior resolution and timeliness throughout the study period, which are essential for effectively capturing mesoscale eddy features.

The ESA SST CCI and C3S level 4 analyses are produced using a complex processing pipeline called the operational SST and sea ice analysis (OSTIA) system [33]. This system accurately and carefully analyzes the satellite data it receives, resulting in a daily study of SST with a high grid resolution of around 5 km. The increased resolution greatly improves the dataset's capacity to record small-scale variations in oceanic temperatures. In this section, the study designates the aforementioned dataset as the foundational training and testing repository for our deep neural network-based U-net algorithm. This investigation harnesses a year in 2019, during which we assimilate daily sensed and detected SST satellite images. The decision to prioritize the year 2019 for eddy identification has been made in order to obtain a comprehensive and precise understanding of the most recent characteristics of eddies. This choice also ensures that the study is based on the most up-to-date and complete information, enabling a consistent examination of the current effects and patterns of eddies. The corresponding SST maps, essential for our research, are thoughtfully furnished by the Copernicus Marine Environment. The SST data underwent spatial and temporal filtering procedures to accurately separate mesoscale signals, guaranteeing that the DL model had been trained on features that are distinctively associated with mesoscale eddies.

C. Proposed Method

In our research, we employed a U-net framework augmented with various backbone networks to identify and categorize oceanic eddies within our dataset. This approach allowed us to classify each pixel into one of three distinct classes: noneddy regions, anticyclonic eddies, or cyclonic eddies. The U-net's

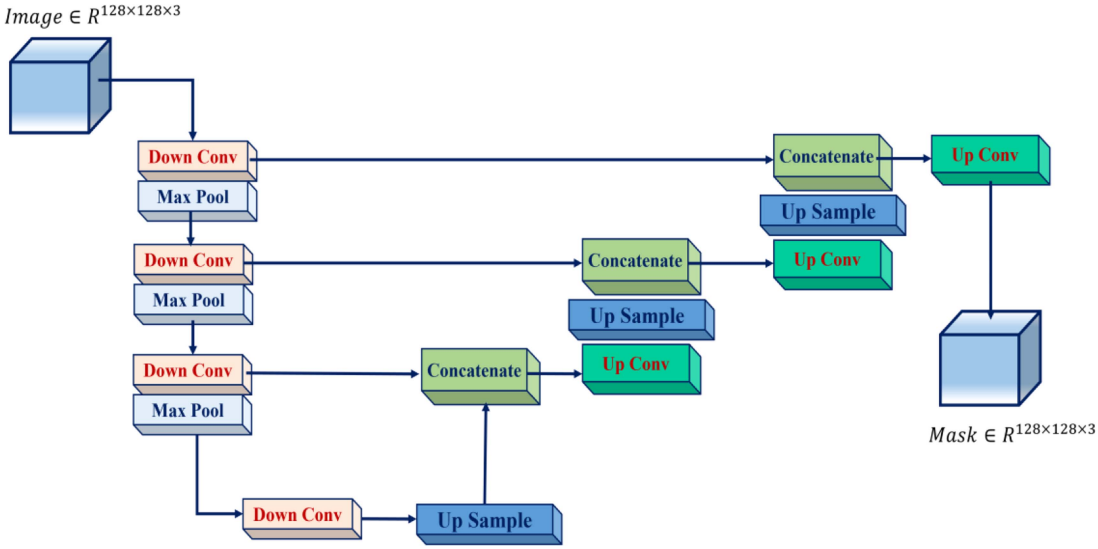


Fig. 2. U-net architecture.

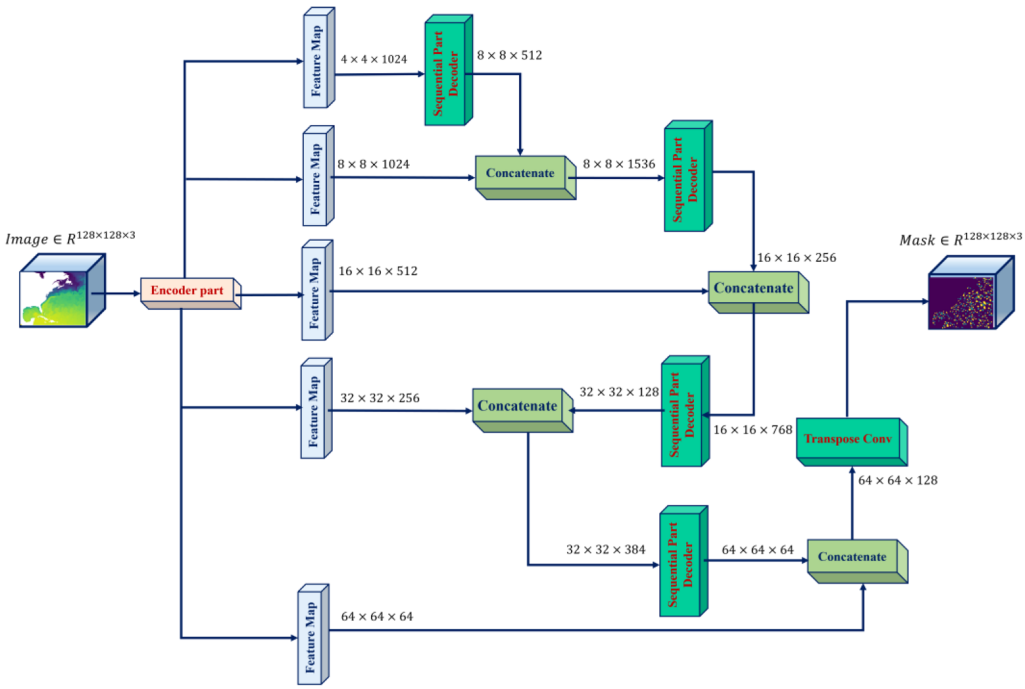


Fig. 3. Architecture of the U-net with DenseNet 121 as a backbone.

architecture, characterized by its distinctive U-shaped design as depicted in Fig. 2, is the inspiration behind its nomenclature. This structure is pivotal for the network's ability to effectively capture both contextual and localized information, thereby enhancing its segmentation capabilities.

U-net is a DL architecture for image segmentation. U-net is designed as a convolutional network with two main components: an encoder for feature extraction and a decoder for feature localization and resolution restoration. The encoder part of U-net captures the essential features of the image while reducing its dimensionality through max-pooling. Subsequently, the decoder portion employs up-sampling and deconvolution techniques to incrementally recover the image's resolution and precisely localize the features. U-net also uses skip connections to concatenate the features from the encoder and the decoder, which helps to preserve the spatial information and improve the segmentation accuracy. The backbone of U-net is the network that is used as the encoder. Different backbones can have different effects on the performance and efficiency of U-net. Specifically, changing the backbone of U-net can be a tradeoff between accuracy and speed, depending on the application and the available resources. In this article, we exploit DenseNet121 as the backbone of the U-net. As depicted in Fig. 3, our specific U-net implementation integrates DenseNet121 as its encoder backbone. The network

processes input images of size 128 × 128, propelling them through the layered architecture. The encoder generates feature maps of descending spatial dimensions but increasing depth, specifically sizes of 4 × 4 × 1024, 8 × 8 × 1024, 16 × 16 × 512, 32 × 32 × 256, and 64 × 64 × 64. These progressively reduced representations are then channeled into the decoder part for detailed reconstruction. Within the U-net architecture, the decoder undergoes a systematic sequence of stages to refine features and expand dimensionality. For example, a feature map of 4 × 4 × 1024 is subjected to up-sampling and deconvolution to get a new size of 8 × 8 × 512. The extended feature map is joined with an 8 × 8 × 1024 feature map from the encoder, resulting in a merged feature map of 8 × 8 × 1536.

Afterward, the feature map undergoes further processing by a subsequent layer in the decoder, resulting in a feature map with dimensions of 16 × 16 × 256. The map is combined with a 16 × 16 × 512 feature map obtained from the encoder, resulting in a composite feature map of 16 × 16 × 768. Subsequently, the procedure proceeds by using an additional decoder layer to convert the feature map into a size of 32 × 32 × 128. The 32 × 32 × 256 feature map from the encoder is combined with this, resulting in a 32 × 32 × 384 feature map. Afterward, an additional level of the decoder is used to generate a feature map with dimensions of 64 × 64 × 64. This feature map is then combined with a feature map obtained from the encoder, which also has dimensions of 64 × 64 × 64. The result is a feature map with dimensions of 64 × 64 × 128. Ultimately, a transpose convolutional layer is used to manipulate this feature map and generate the ultimate output mask, which has dimensions of 128 × 128 × 3. The precise sequence of up-sampling, merging, and processing is crucial for the U-net to precisely locate and define the target features in the output mask.

The selection of the 128 × 128 pixel size for input images in our U-net-based segmentation approach has been driven by a balance between computational efficiency and the level of detail necessary for accurate eddy detection. This size is sufficient to capture the essential features of mesoscale eddies while maintaining manageable computational demands for processing large datasets. It allows for the inclusion of enough spatial context to effectively identify and classify eddy and noneddy regions, leveraging the U-net architecture's capability to handle both contextual and localized information. This dimensionality also aligns well with the U-net's encoder-decoder mechanism, where the encoder reduces the dimensionality through max-pooling, and the decoder employs up-sampling and deconvolution to recover resolution and localize features accurately. The chosen size ensures that the network can efficiently process feature maps through its layers, as demonstrated by the systematic reduction and expansion of spatial dimensions within the network architecture. This careful consideration of input image size thus supports the model's overall performance in segmenting satellite imagery for oceanic eddy detection.

In our U-net architecture, we employ four distinct backbone networks: VGG16 [34], VGG19 [34], DenseNet 121 [35], and MobileNet V2 [36]. Of these, DenseNet 121 emerges as the most effective in terms of performance. The backbone weights are initialized using pretraining on the ImageNet dataset, which confers several advantages. This initialization strategy shortens the training duration for the terminal layers, expedites the convergence process, and enhances the overall efficacy of the network. MobileNetV2 and DenseNet121 are CNNs used for image processing applications. However, they vary greatly in terms of the complexity and number of parameters in their structures. MobileNetV2 is renowned for its high efficiency, achieved by the use of depth-wise separable convolutions in combination with inverted residual structures, resulting in a reduction in parameter count and computing burden. The sleek appearance of this product makes it well-suited for settings with minimal resources.

Conversely, DenseNet121, a modified version of the DenseNet architecture, employs a network topology that is densely linked. This model is composed of four dense blocks, each consisting of 6, 12, 24, and 16 convolutional layers accordingly. It is distinguished by its distinctive connection structure, where every layer gets input from all the layers that come before it. The network concludes with a fully connected layer, which is then followed by a SoftMax classification layer. Because of its extensive connection and several layers, DenseNet121 is regarded as a network that requires more parameters. This means it demands more computing resources, but it also enhances its ability to handle difficult visual recognition jobs effectively. MobileNetV2 is an enhanced version of the original MobileNet design. It incorporates depth-wise separable convolutions together with inverted residual structures to reduce the number of parameters and computing requirements. MobileNetV2 is designed to be efficient, with 17 inverted residual blocks that have different expansion factors and output channels. The model is finalized with a fully connected layer, which is then followed by a SoftMax layer for the purpose of classification. MobileNetV2, which consists of almost 3.5 million parameters, is specifically designed to prioritize speed and efficiency. This makes it highly ideal for devices with limited processing capabilities, such as mobile and embedded devices.

On the other hand, DenseNet121, which has around 8 million parameters, is highly praised for its precision and resilience. The reason for this is its compact block structure, which improves the transmission and reutilization of characteristics, making it very proficient at managing intricate and diverse picture collections. In summary, MobileNetV2 and DenseNet121 provide a harmonious combination of speed, efficiency, accuracy, and complexity, with each being specifically designed for certain circumstances and applications. MobileNetV2 is the preferred choice for achieving fast and efficient performance in contexts with limited resources, whereas DenseNet121 stands out in situations where precision and robustness are of utmost importance. VGG16 and VGG19 are two configurations of the VGG model, which is characterized by its use of small 3 × 3 convolutional filters stacked in multiple layers before a max-pooling layer. The primary distinction between VGG16 and VGG19 lies in the number of convolutional layers: VGG16 comprises 13, while VGG19 includes 16. Each model concludes with three dense layers and a SoftMax layer for output classification. Both VGG16 and VGG19 are considered to be models with substantial parameter counts, approximately 138 million and

144 million, respectively, reflecting their more complex and expansive architectures.

The selection of neural network architectures in this study has been driven by the need to balance accuracy, computational efficiency, and model complexity. The U-net architecture, known for its efficacy in image segmentation tasks, has been chosen as the foundation for all models due to its robust performance in feature extraction and localization. The VGG16 and VGG19 models were selected for their deep architectures and strong feature extraction capabilities, providing a benchmark for segmentation performance. DenseNet121 has been chosen for its efficient feature propagation and reuse, which is beneficial in identifying complex patterns in SST data. Finally, MobileNetV2 is included for its lightweight architecture and rapid processing capabilities, making it ideal for real-time analysis in computationally constrained environments. Each model thus caters to different requirements of the segmentation task, from the VGG models providing detailed feature extraction to the MobileNetV2 ensuring efficiency and speed. This comprehensive approach allows for a robust evaluation of varying architectural strengths in the context of mesoscale eddy detection from satellite imagery.

D. Evaluation Metrics for Convolutional Neural Network Performance in Mesoscale Eddy Segmentation

Within the theoretical framework of our research, we have used a range of measures to evaluate the effectiveness of neural network models in dividing satellite pictures into segments. The Intersection over Union (IoU) is a widely used statistic in segmentation tasks that quantifies the degree of overlap between the predicted segment (the region recognized by the model) and the ground truth (the actual region). The ratio of the intersection of the projected and actual areas to their union is calculated. A higher IoU value signifies a larger degree of overlap, which in turn indicates superior performance of the model [37].

$$IoU = \frac{\text{Area of Overlap between Prediction and Ground Truth}}{\text{Area of Union between Prediction and Ground Truth}}. \quad (1)$$

The Dice Score, sometimes referred to as the Dice coefficient, is a similarity statistic used to assess the precision of the segmentation. The model's performance is assessed by determining the extent of overlap between the predicted segmentation and the ground truth. This is achieved by multiplying the overlap size by two and dividing it by the total of the pixels in both segmentations. Similar to the IoU metric, a higher Dice Score signifies superior performance [38].

$$Dice = \frac{2\times\text{Area of Overlap between Pred and GT}}{\text{Total Num of Pixels in Pred}+\text{Total Num of Pixels in GT}}. \quad (2)$$

In this case, accuracy refers to the degree of correctness of the model. It is determined by summing the number of accurate predictions (true positives and true negatives) and dividing it by the total number of predictions produced. True positives (TP) refer to the accurately detected eddy pixels, while true negatives (TN) represent the properly recognized noneddy pixels. On the other hand, false negatives (FN) indicate the eddy pixels that the model mistakenly classified as noneddy [39].

$$\text{Accuracy} = \frac{\text{True Positives} + \text{True Negatives}}{\text{Total Population}}. \quad (3)$$

Precision is a quantification of accuracy or excellence. Precision is calculated as the ratio of true positives to the sum of true positives and erroneous positives, where false positives refer to noneddy pixels that are mistakenly recognized as eddies. High accuracy is characterized by a minimal occurrence of false positives [40].

$$\text{Precision} = \frac{\text{True Positives}}{\text{True Positives} + \text{False Positives}}. \quad (4)$$

The F1 Score is calculated as the harmonic mean of accuracy and recall, effectively balancing the two metrics by including both erroneous positives and FN. The F1 Score is essential since it considers both the accuracy and recall of the model, which is the ratio of TP to the sum of TP and FN [41].

$$\text{F1 Score} = 2 \times \frac{\text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}}. \quad (5)$$

The recall, therefore, signifies the model's capacity to correctly detect all relevant instances (all real eddies) [41].

$$\text{Recall} = \frac{\text{True Positives}}{\text{True Positives} + \text{False Negative}}. \quad (6)$$

Collectively, these measures provide a comprehensive perspective on the model's performance, with each statistic reflecting a distinct facet of the segmentation process. Achieving high scores in these criteria indicates a model that is both accurate and precise, and capable of reliably detecting eddies in diverse marine satellite data.

IV. Results

The neural networks have been trained on the satellite imagery dataset from the North Atlantic Ocean in 2019, resulting in segmentation maps that are visually and numerically distinguishable. Our methodology is implemented using Python with the Keras framework, leveraging the Adam optimization algorithm, and also, we trained the model on a Google Colab with 12.7 GB of memory and T4 as GPU. We configure the dataset with a batch size of 16 and undertake training over 100 epochs, allocating 10% of the data for validation purposes. The chosen loss function is sparse categorical cross-entropy, which is apt for multiclass categorization tasks involving sparsely labeled data. This means that instead of using a one-hot encoded vector, each sample's label is denoted by a single integer. Fig. 4 illustrates the satellite image, ground truth, and predicted image for different backbones. To train our DL models, we carefully selected a ground truth dataset consisting of mask images that depict the borders of mesoscale eddies. As mentioned before, we have been assisted by SST data that implement our mask images for labeling eddies based on temperature as cyclonic and anticyclonic eddies regarding cold and warm core water in our study area. The CMEMS set up and confirmed these masks, which indicate the locations of eddies. These masks serve as a dependable foundation for training our model and guarantee the

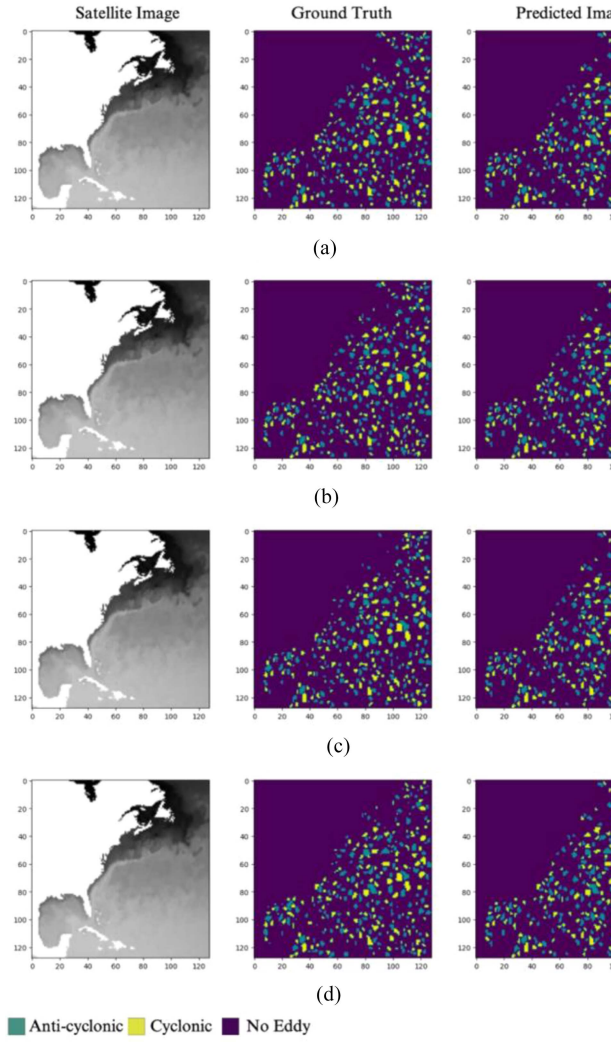


Fig. 4. Comparative analysis of mesoscale eddy detection: the first column depicts the original satellite imagery, while the subsequent columns display ground truth and the segmentation results for anticyclonic, cyclonic, and no eddy conditions using (a) Unet-DenseNet121, (b) Unet-MobileNetV2, (c) Unet-VGG16, and (d) Unet-VGG19 neural network models, respectively.

accuracy of eddy identification and segmentation. By developing code script regarding visualization of the mentioned images by using Mat plot library in Python to extract temperature information from NetCDF (credited by CMEMS) file images and sorting and representing based on time, coordinate, temperature and exporting ground truth classified images via temperature in interested area based on three mentioned classes.

The images shown here illustrate the relative efficacy of the Unet-VGG16, Unet-VGG19, Unet-DenseNet121, and Unet-MobileNetV2 models in accurately detecting and categorizing different attributes in the photographs. The first figure in the series depicts the unaltered grayscale input satellite image. The following images depict the segmentation maps generated by the corresponding neural network models. Every hue in the segmentation maps corresponds to a distinct categorized attribute or category in the examined image.

TABLE III
COMPARATIVE ANALYSIS OF U-NET ARCHITECTURES: PARAMETER DISTRIBUTION ACROSS DIFFERENT CONVOLUTIONAL BACKBONE NETWORKS

	Unet-VGG16	Unet-VGG19	Unet-Mobile NetV2	**Unet-DenseNet 121**
Total parameters	20177859	25487555	6504227	**16408259**
Trainable parameters	5461251	5461251	4660323	**9368835**
Nontrainable parameters	14716608	20026304	1843004	**7039424**

Bold values represent the best significant results.

The outcomes of the training dataset are summarized in Table II. The Unet-DenseNet121 model demonstrated superior performance compared to the other models in almost all parameters, attaining the best precision (96.05%), recall (95.56%), F1 Score (95.79%), and accuracy (99.37%). In addition, this model demonstrated higher Dice coefficient scores for class 0 (99.43%), class 1 (95.79%), and class 2 (95.98%), as well as IoU scores for class 0 (98.84%), class 1 (91.47%), and class 2 (91.6%). The Unet-MobileNetV2 also demonstrated impressive performance, with precision, recall, F1 Score, and accuracy statistics closely behind those of the Unet-DenseNet121. The Dice coefficient and IoU scores exhibited comparable high values, suggesting a resilient model, but somewhat less efficient than the Unet-DenseNet121. In contrast, the performance of Unet-VGG16 and Unet-VGG19 has been worse when compared to the other two models. In terms of various metrics, the VGG16 variation performed somewhat better than the VGG19. However, both models significantly underperformed compared to the Unet-DenseNet121 and Unet-MobileNetV2 models, especially when it came to the IoU scores for classes 0, 1, and 2. This suggests that the VGG models had a lesser capability to accurately define the borders of these classes.

In the analysis of mesoscale eddy detection within North Atlantic Ocean satellite imagery, the Unet-DenseNet121 model notably outperformed its counterparts, delivering top-tier results across precision, recall, F1 score, and accuracy metrics. The model's effectiveness is further underscored by its superior Dice coefficient and IoU scores for the classified eddy features. While the Unet-MobileNetV2 followed closely, presenting a viable option for efficient and accurate segmentation, the VGG-based models, Unet-VGG16 and Unet-VGG19, lagged in performance, particularly in their precision of class boundary delineation. These insights not only demonstrate the advancements in neural network applications for oceanographic studies but also pave the way for future enhancements in satellite imagery analysis.

In the comparative analysis of U-net architectures with different backbones, Table III presents a summary of the total, trainable, and nontrainable parameters for each model variant: U-net with VGG16, U-net with VGG19, U-net with DenseNet121, and U-net with MobileNetV2 backbones. The Unet-VGG16 model comprises a total of 20177859 parameters, of which 5461251

are trainable. The remainder, 14716608 parameters, are nontrainable, indicating that a significant portion of the network's parameters is fixed during training, likely due to the incorporation of pretrained weights. Unet-VGG19 shows an increase in complexity with a total of 25487555 parameters. Trainable parameters remain consistent with the Unet-VGG16 model at 5461251, suggesting a similar strategy for training. However, the nontrainable parameter count rises to 20026304, reflecting the model's deeper architecture. The Unet-DenseNet121 architecture demonstrates a total parameter count of 16408259, a reduction compared to the VGG-based models. It also has a higher number of trainable parameters at 9368835, suggesting a more flexible model that adapts more substantially during training. The nontrainable parameters are significantly lower at 7039424, which may suggest a leaner pretraining component or a more efficient parameter utilization.

Finally, the Unet-MobileNetV2 architecture boasts the fewest parameters with a total of 6504227. Of these, 4660323 are trainable, which is relatively high given the total parameter count and indicative of the network's design for efficient training. The nontrainable parameters are the lowest at 1843004, aligning with the network's goal of being lightweight and suitable for environments with constrained computational resources. From these figures, it can be inferred that each U-net variant offers different advantages. The VGG-based U-nets, with their high number of nontrainable parameters, may benefit from faster convergence times due to pretrained features. In contrast, Unet-DenseNet121, with its balanced distribution of trainable parameters, might offer a good tradeoff between flexibility and efficiency. Unet-MobileNetV2 stands out for its minimal parameter count and high proportion of trainable parameters, which is likely to offer rapid training times and model adaptability, albeit potentially at the cost of reduced feature extraction capability due to fewer parameters. This distribution of parameters impacts the models' training dynamics and their suitability for different deployment scenarios, a critical discussion point for the results and discussion section of an academic paper.

V. Discussion

The segmentation map generated by the Unet-DenseNet121 model exhibits a remarkable amount of detail and accuracy, which is consistent with the reported high quantitative metrics. The model demonstrates a discernible differentiation among several characteristics, and it has a reduced rate of incorrectly categorized pixels in comparison to the other models. The accuracy of the segmentation indicates that the DenseNet121 backbone allows the model to comprehend the geographical context inside the satellite images in a more detailed manner, resulting in improved accuracy when predicting class labels. The segmentation map of Unet-MobileNetV2, albeit less intricate than that of Unet-DenseNet121, exhibits a significant level of precision. The little integration of classes may stem from the architectural compromises made to enhance efficiency in the creation of MobileNetV2. However, its performance is praiseworthy, particularly when considering its low weight and ability to be used in systems with limited computing power. Conversely, the segmentation maps generated by the Unet-VGG16 and Unet-VGG19 models have a higher level of noise and a lower level of separation across classes. The elevated levels of misclassification may be clearly seen by the presence of dispersed and uneven color patches that deviate from the forms and bounds in the original picture. The observed visual results support the lower precision and recall scores obtained by these models. This indicates that although VGG architectures are strong for image classification tasks, they may not be as efficient for the detailed pixel-level predictions needed for semantic segmentation in complex images such as satellite data. The visual outcomes emphasize the significance of choosing a suitable neural network structure for certain tasks. In addition, they emphasize the potential advantages of using more sophisticated models, such as Unet-DenseNet121, in situations when precision is of utmost importance. Nevertheless, it is crucial to take into account the compromises in terms of computing expenditure and inference duration, which might be pivotal limitations in practical scenarios.

The Unet-DenseNet121's superior performance can be attributed to its architectural features. DenseNet121 is known for its efficient feature propagation and reuse, which likely contributed to the model's enhanced capability in recognizing patterns and features in the satellite imagery. Its consistently high scores across precision, recall, and the F1 Score suggest not only that it is good at predicting the correct classes but also at minimizing false positives and negatives. The Unet-MobileNetV2's results were also noteworthy, especially considering that MobileNetV2 is designed to be a lightweight model suited for mobile and edge devices. Its performance underscores the potential for deploying effective segmentation models in resource-constrained environments, although with a slight tradeoff in the precision of class boundary delineation compared to the Unet-DenseNet121. The lower performance of the VGG-based models could be due to the inherent architecture limitations, which may lead to less effective feature extraction in the context of the specific segmentation tasks required by this study. It is also possible that these models are more prone to overfitting, given their large number of parameters and deeper architectures, which might not generalize as well to the varied features present in satellite oceanic imagery. In conclusion, our findings indicate that Unet-DenseNet121 is the most suitable model for semantic segmentation tasks in the domain of satellite oceanic imagery for the North Atlantic Ocean, given its high precision in classifying various objects or phenomena. Future work should focus on validating these results across different oceanic regions and in different conditions to ensure the robustness and generalizability of the Unet-DenseNet121 model. Additionally, the exploration of further model optimizations and the potential for real-time application in edge devices using models like Unet-MobileNetV2 could be valuable avenues for research (shown in Fig. 5 and Table II). While our study identifies the Unet-DenseNet121 as a robust CNN model for segmenting mesoscale eddies, its application is limited by the use of data from only the North Atlantic and for a single year, which may not be representative

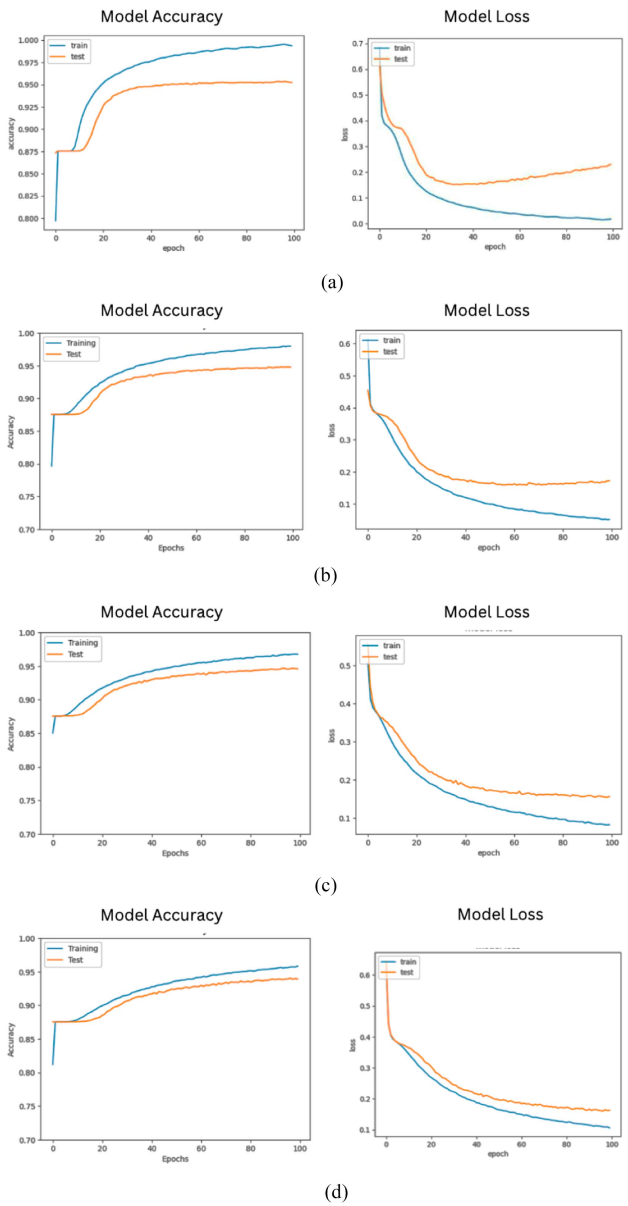


Fig. 5. Training performance of different convolutional neural network architectures: the graphs display the progression of model accuracy and loss over epochs for (a) Unet-DenseNet121, (b) Unet-MobileNetV2, (c) Unet-VGG16, and (d) Unet-VGG19 when applied to satellite imagery segmentation tasks, highlighting their learning efficiency and convergence behavior during the training process.

of other regions or periods. Additionally, the computational intensity of the model could be prohibitive for real-time analysis on platforms with limited resources. The study also did not investigate ensemble models or the interpretability of the CNN decisions, which are critical for operational trustworthiness.

VI. Conclusion

In this comparative analysis of four advanced CNN models applied to satellite imagery of the North Atlantic Ocean, we have determined the Unet-DenseNet121 to be the superior architecture for segmenting mesoscale eddies into cyclonic, anticyclonic, and noneddy categories. The Unet-DenseNet121's performance stands out with the highest Dice coefficient and IoU values, indicating its exceptional ability to delineate the boundaries of eddies accurately. Following closely was the Unet-MobileNetV2 model, notable for its efficiency in situations where computational resources are limited. Its lightweight structure is particularly beneficial for onboard satellite processing or real-time analysis, making it a practical option for immediate application. In contrast, the Unet-VGG16 and Unet-VGG19 models, despite being foundational in DL, lagged behind in segmentation quality, as reflected in their lower IoU scores. This suggests a potential need for improvement or the possibility that they may be outperformed by more contemporary architectures. The quantitative findings are bolstered by visual inspections of segmentation outputs and training performance graphs. The Unet-DenseNet121 consistently produced segmentation maps with high fidelity to the original satellite images and exhibited stable learning curves, indicative of reliable learning and generalization capabilities. The implications of this research are twofold. First, the study underlines the enhanced accuracy and efficiency that advanced CNNs can bring to the identification of oceanic features, which is vital for monitoring marine environments and understanding their temporal dynamics. Second, it paves the way for future research, emphasizing the need for extensive validation of these models across diverse datasets, exploration of ensemble methods to leverage each architecture's strengths, and investigation into the transparency of model decisions to build trust in AI-driven analyses. The integration of these CNN models into operational frameworks used by meteorological and oceanographic institutions could revolutionize real-time analysis and monitoring. By harnessing the computational power and pattern recognition prowess of CNNs, we can significantly improve the timeliness and accuracy of oceanographic assessments. This study serves as a critical evaluation of CNN architectures for mesoscale eddy detection in satellite imagery, contributing to the advancement of oceanographic research and environmental monitoring. Future endeavors should aim to refine these DL models, tailor them to the complexities of oceanographic data, and incorporate them into holistic systems that foster sustainable ocean management and protection.

References

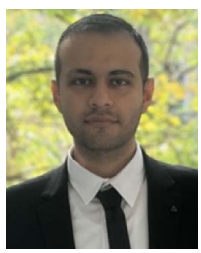
[1] H. Sasaki, P. Klein, B. Qiu, and Y. Sasai, "Impact of oceanic-scale interactions on the seasonal modulation of ocean dynamics by the atmosphere," *Nature Commun.*, vol. 5, no. 1, pp. 1–8, Dec. 2014, doi: 10.1038/ncomms6636.

[2] J. Röhrs et al., "Surface currents in operational oceanography: Key applications, mechanisms, and methods," *J. Oper. Oceanogr.*, vol. 16, no. 1, pp. 60–88, Jan. 2023, doi: 10.1080/1755876X.2021.1903221.

[3] S. M. M. Nejad, D. Abbasi-Moghadam, A. Sharifi, N. Farmonov, K. Amankulova, and M. László, "Multispectral crop yield prediction using 3D-convolutional neural networks and attention convolutional LSTM approaches," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 16, pp. 254–266, Nov. 2023, doi: 10.1109/JSTARS.2022.3223423.

[4] M. Esmaeili, D. Abbasi-Moghadam, A. Sharifi, A. Tariq, and Q. Li, "Hyperspectral image band selection based on CNN embedded GA (CNNeGA)," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 16, pp. 1927–1950, Feb. 2023, doi: 10.1109/JSTARS.2023.3242310.

[5] L. Siegelman et al., "Enhanced upward heat transport at deep submesoscale ocean fronts," *Nature Geosci.*, vol. 13, no. 1, pp. 50–55, Jan. 2020, doi: 10.1038/s41561-019-0489-1.
[6] B. Sun, G. W. Kattawar, P. Yang, M. S. Twardowski, and J. M. Sullivan, "Simulation of the scattering properties of a chain-forming triangular prism oceanic diatom," *J. Quantitative Spectrosc. Radiative Transfer*, vol. 178, pp. 390–399, Jul. 2016, doi: 10.1016/j.jqsrt.2016.02.035.
[7] C. Chu et al., "Effect of the tropical Pacific and Indian Ocean warming since the late 1970s on wintertime Northern Hemispheric atmospheric circulation and East Asian climate interdecadal changes," *Climate Dyn.*, vol. 50, no. 7–8, pp. 3031–3048, Apr. 2018, doi: 10.1007/s00382-017-3790-y.
[8] I. Frenger, N. Gruber, R. Knutti, and M. Münnich, "Imprint of southern ocean eddies on winds, clouds and rainfall," *Nature Geosci.*, vol. 6, no. 8, pp. 608–612, Jul. 2013, doi: 10.1038/ngeo1863.
[9] P. B. Rhines and W. R. Young, "Homogenization of potential vorticity in planetary gyres," *J. Fluid Mechanics*, vol. 122, no. 1, Sep. 1982, Art. no. 347, doi: 10.1017/S0022112082002250.
[10] J. Marshall, H. Jones, R. Karsten, and R. Wardle, "Can eddies set ocean stratification?," *J. Phys. Oceanogr.*, vol. 32, no. 1, pp. 26–38, 2002, doi: 10.1175/1520-0485(2002)032<0026:CESOS>2.0.CO;2.
[11] A. Voorhis, E. H. Schroeder, and A. Leetmaa, "The influence of deep mesoscale eddies on sea surface temperature in the North Atlantic subtropical convergence," *J. Phys. Oceanogr.*, vol. 6, pp. 953–961, 1976, Accessed: Jul. 23, 2023. [Online]. Available: https://journals.ametsoc.org/view/journals/phoc/6/6/1520-0485_1976_006_0953_tiodme_2_0_co_2.xml
[12] J. Le Sommer, F. d'Ovidio, and G. Madec, "Parameterization of subgrid stirring in eddy resolving ocean models. Part 1: Theory and diagnostics," *Ocean Model*, vol. 39, no. 1–2, pp. 154–169, 2011, doi: 10.1016/j.ocemod.2011.03.007.
[13] J. H. Faghmous, I. Frenger, Y. Yao, R. Warmka, A. Lindell, and V. Kumar, "A daily global mesoscale ocean eddy dataset from satellite altimetry," *Sci. Data*, vol. 2, no. 1, pp. 1–16, Jun. 2015, doi: 10.1038/sdata.2015.28.
[14] A. Pascual, Y. Faugère, G. Larnicol, and P. Y. Le Traon, "Improved description of the ocean mesoscale variability by combining four satellite altimeters," *Geophysical Res. Lett.*, vol. 33, no. 2, pp. 1–4, Jan. 2006, doi: 10.1029/2005GL024633.
[15] D. B. Chelton, M. G. Schlax, and R. M. Samelson, "Global observations of nonlinear mesoscale eddies," *Prog. Oceanogr.*, vol. 91, no. 2, pp. 167–216, Oct. 2011, doi: 10.1016/j.pocean.2011.01.002.
[16] H. M. A. Aguedjou, A. Chaigneau, I. Dadou, Y. Morel, E. Baloïtcha, and C. Y. Da-Allada, "Imprint of mesoscale eddies on air-sea interaction in the Tropical Atlantic Ocean," *Remote Sens.*, vol. 15, no. 12, Jun. 2023, Art. no. 3087, doi: 10.3390/RS15123087.
[17] Q. Ni, X. Zhai, X. Jiang, and D. Chen, "Abundant cold anticyclonic eddies and warm cyclonic eddies in the global ocean," *J. Phys. Oceanogr.*, vol. 51, no. 9, pp. 2793–2806, Sep. 2021, doi: 10.1175/JPO-D-21-0010.1.
[18] C. Dong et al., "The near-global ocean mesoscale eddy atmospheric-oceanic-biological interaction observational dataset," *Sci. Data*, vol. 9, no. 1, pp. 1–13, Jul. 2022, doi: 10.1038/s41597-022-01550-9.
[19] R. Tang, Y. Yu, J. Xi, W. Ma, and Y. Wang, "Mesoscale eddies induce variability in the sea surface temperature gradient in the Kuroshio Extension," *Front. Mar. Sci.*, vol. 9, Jul. 2022, Art. no. 926954, doi: 10.3389/FMARS.2022.926954/BIBTEX.
[20] A. Czaja and N. Blunt, "A new mechanism for ocean-atmosphere coupling in midlatitudes," *Quart. J. Roy. Meteorological Soc.*, vol. 137, no. 657, pp. 1095–1101, Apr. 2011, doi: 10.1002/QJ.814.
[21] L. Cao, D. Zhang, X. Zhang, and Q. Guo, "Detection and identification of mesoscale eddies in the South China Sea based on an artificial neural network model—YOLOF and remotely sensed data," *Remote Sens.*, vol. 14, no. 21, Oct. 2022, Art. no. 5411, doi: 10.3390/rs14215411.
[22] L. Cao, D. Zhang, Q. Guo, and J. Zhan, "Ocean mesoscale eddies identification based on YOLOF," in *Proc. IEEE Int. Geosci. Remote Sens. Symp.*, 2022, pp. 1177–1180, doi: 10.1109/IGARSS46834.2022.9883834.
[23] B. Huang, L. Ge, X. Chen, and G. Chen, "Vertical structure-based classification of oceanic eddy using 3-D convolutional neural network," *IEEE Trans. Geosci. Remote Sens.*, vol. 60, Aug. 2022, Art. no. 4203614, doi: 10.1109/TGRS.2021.3103251.
[24] X. Chen, G. Chen, L. Ge, B. Huang, and C. Cao, "Global oceanic eddy identification: A deep learning method from Argo profiles and altimetry data," *Front. Mar. Sci.*, vol. 8, May 2021, Art. no. 25642, doi: 10.3389/fmars.2021.646926.
[25] N. Zhao, B. Huang, J. Yang, M. Radenkovic, and G. Chen, "Oceanic eddy identification using pyramid split attention U-net with remote sensing imagery," *IEEE Geosci. Remote Sens. Lett.*, vol. 20, Feb. 2023, Art. no. 1500605, doi: 10.1109/LGRS.2023.3243902.
[26] S. John Saida and S. Ari, "MU-Net: Modified U-Net architecture for automatic ocean eddy detection," *IEEE Geosci. Remote Sens. Lett.*, vol. 19, Nov. 2022, Art. no. 1507005, doi: 10.1109/LGRS.2022.3225140.
[27] Z. Duo, W. Wang, and H. Wang, "Oceanic mesoscale eddy detection method based on deep learning," *Remote Sens.*, vol. 11, no. 16, Aug. 2019, Art. no. 1921, doi: 10.3390/rs11161921.
[28] "Atlantic Ocean," Accessed: Jan. 27, 2024. [Online]. Available: https://www.findlatitudeandlongitude.com/l/Atlantic+Ocean/5748743/
[29] Y. Zhong and A. Bracco, "Submesoscale impacts on horizontal and vertical transport in the Gulf of Mexico," *J. Geophysical Res.: Oceans*, vol. 118, no. 10, pp. 5651–5668, 2013, doi: 10.1002/jgrc.20402.
[30] A. McQuatters-Gollop et al., "Assessing the state of marine biodiversity in the Northeast Atlantic," *Ecological Indicators*, vol. 141, Aug. 2022, Art. no. 109148, doi: 10.1016/J.ECOLIND.2022.109148.
[31] "ESA SST CCI and C3S reprocessed sea surface temperature analyses | Copernicus Marine MyOcean viewer," Accessed: Jan. 27, 2024. [Online]. Available: https://data.marine.copernicus.eu/product/SST_GLO_SST_L4_REP_OBSERVATIONS_010_024/description
[32] C. J. Merchant et al., "Satellite-based time-series of sea-surface temperature since 1981 for climate applications," *Sci. Data*, vol. 6, no. 1, Dec. 2019, Art. no. 25495, doi: 10.1038/S41597-019-0236-X.
[33] S. Good et al., "The current configuration of the OSTIA system for operational production of foundation sea surface temperature and ice concentration analyses," *Remote Sens.*, vol. 12, no. 4, Feb. 2020, Art. no. 720, doi: 10.3390/RS12040720.
[34] K. Simonyan and A. Zisserman, "Very deep convolutional networks for large-scale image recognition," in *Proc. 3rd Int. Conf. Learn. Representations*, Sep. 2014, Accessed: Feb. 05, 2024, pp. 1–14. [Online]. Available: https://arxiv.org/abs/1409.1556v6
[35] G. Huang, Z. Liu, L. Van Der Maaten, and K. Q. Weinberger, "Densely connected convolutional networks," in *Proc. IEEE 30th Conf. Comput. Vis. Pattern Recognit.*, 2017, pp. 2261–2269, doi: 10.1109/CVPR.2017.243.
[36] M. Sandler, A. Howard, M. Zhu, A. Zhmoginov, and L. C. Chen, "MobileNetV2: Inverted residuals and linear bottlenecks," in *Proc. IEEE Comput. Soc. Conf. Comput. Vis. Pattern Recognit.*, 2018, pp. 4510–4520, doi: 10.1109/CVPR.2018.00474.
[37] M. A. Rahman and Y. Wang, "Optimizing intersection-over-union in deep neural networks for image segmentation," in *Proc. Adv. Vis. Comput.*, pp. 234–244, vol. 10072, 2016, doi: 10.1007/978-3-319-50835-1_22/FIGURES/3.
[38] C. H. Sudre, W. Li, T. Vercauteren, S. Ourselin, and M. J. Cardoso, "Generalised Dice overlap as a deep learning loss function for highly unbalanced segmentations," in *Proc. Deep Learn. Med. Image Anal. Multimodal Learn. Clin. Decis. Support*, Jul. 2017, vol. 10553, pp. 240–248, doi: 10.1007/978-3-319-67558-9_28.
[39] A. Serwa and S. Elbialy, "Enhancement of classification accuracy of multi-spectral satellites' images using Laplacian pyramids," *Egyptian J. Remote Sens. Space Sci.*, vol. 24, no. 2, pp. 283–291, Aug. 2021, doi: 10.1016/J.EJRS.2020.12.006.
[40] C. Goutte and E. Gaussier, "A probabilistic interpretation of precision, recall and F-score, with implication for evaluation," in *Proc. Adv. Inf. Retrieval*, 2005, vol. 3408, pp. 345–359, doi: 10.1007/978-3-540-31865-1_25/COVER.
[41] R. Yacouby and D. Axman, "Probabilistic extension of precision, recall, and F1 score for more thorough evaluation of classification models," in *Proc. 1st Workshop Eval. Comparison NLP Syst.*, 2020, pp. 79–91, doi: 10.18653/V1/2020.EVAL4NLP-1.9.



Mohammad Mahdi Safari was born in Saveh, Iran, in 1996. He received the B.S. degree in geoinformatics engineering from Shahid Rajaee University, Tehran, Iran, in 2020, and the M.S. degree in geoinformatics engineering from Politecnico di Milano, Milan, Italy, in 2023.

His master's thesis was completed in the Erasmus program with Bonn University, Germany. He has both academic and industry experience, having worked on GIS and remote sensing projects during his master's and bachelor's studies, as well as more than two years of R&D experience as a Geoinformatics Engineer. His research interests include remote sensing, GIS, machine learning, deep learning, and environmental issues.

Alireza Sharifi was born in Tehran, Iran, in 1981. He received the M.Sc. and Ph.D. degrees in remote sensing engineering from the University of Tehran, Tehran, Iran, in 2008 and 2015, respectively.

He is currently an Associate Professor of Remote Sensing with the Faculty of Civil Engineering, Shahid Rajaee Teacher Training University, Tehran, Iran. His current research activities are remote sensing, time series analysis, and satellite image processing. In particular, he is involved in the GEOAI program for food security and environmental monitoring.

Javad Mahmoodi was born in Kerman, Iran, in 1981. He received the M.Sc. degree in electrical engineering from Semnan University, Kerman, Iran, in 2006, and the Ph.D. degree in communication engineering from Shahid Bahonar University, Kerman, Iran, in 2024.

He is currently with the Department of Electrical Engineering, Kerman Branch of Islamic Azad University, Kerman, Iran, serving as an Assistant Professor. His research interests lie in the areas of image processing, object detection, remote sensing, and video processing.

Dariush Abbasi-Moghadam received the B.S. degree in electrical engineering from Shahid Bahonar University, Kerman, Iran, in 1998, and the M.S. and Ph.D. degrees in electrical engineering from Iran University of Science and Technology, Tehran, Iran, in 2001 and 2011, respectively.

From 2001 to 2003, he was primarily with the Advanced Electronic Research Center—Iran and worked on the design and analysis of satellite communication systems. In 2004, he joined the Iranian Telecommunications Company, Tehran, as a Research Engineer. He is currently with the Department of Electrical Engineering, Shahid Bahonar University of Kerman, Kerman, Iran, as an Associate Professor. His research interests are in the areas of wireless communications, satellite communication systems, remote sensing, and signal processing.