



Article

Semantic Segmentation of Mesoscale Eddies in the Arabian Sea: A Deep Learning Approach

Mohamad Abed El Rahman Hammoud ¹, Peng Zhan ^{1,2,3}, Omar Hakla ⁴, Omar Knio ^{1,2} and Ibrahim Hoteit ^{1,2*}

¹ Physical Science and Engineering, King Abdullah University of Science and Technology, Thuwal 23955, Saudi Arabia

² Department of Ocean Science and Engineering, Southern University of Science and Technology, Shenzhen 518055, China

³ Southern Marine Science and Engineering Guangdong Laboratory, Guangzhou 510000, China

⁴ Maroun Semaan Faculty of Engineering and Architecture, American University of Beirut, Beirut 1107, Lebanon

* Correspondence: ibrahim.hoteit@kaust.edu.sa

Abstract: Detecting mesoscale ocean eddies provides a better understanding of the oceanic processes that govern the transport of salt, heat, and carbon. Established eddy detection techniques rely on physical or geometric criteria, and they notoriously fail to predict eddies that are neither circular nor elliptical in shape. Recently, deep learning techniques have been applied for semantic segmentation of mesoscale eddies, relying on the outputs of traditional eddy detection algorithms to supervise the training of the neural network. However, this approach limits the network's predictions because the available annotations are either circular or elliptical. Moreover, current approaches depend on the sea-surface height, temperature, or currents as inputs to the network, and these data may not provide all the information necessary to accurately segment eddies. In the present work, we have trained a neural network for the semantic segmentation of eddies using human-based—and expert-validated—annotations of eddies in the Arabian Sea. Training with human-annotated datasets enables the network predictions to include more complex geometries, which occur commonly in the real ocean. We then examine the impact of different combinations of input surface variables on the segmentation performance of the network. The results indicate that providing additional surface variables as inputs to the network improves the accuracy of the predictions by approximately 5%. We have further fine-tuned another pre-trained neural network to segment eddies and achieved a reduced overall training time and higher accuracy compared to the results from a network trained from scratch.

Keywords: eddy detection; deep neural network; semantic segmentation; Arabian Sea



Citation: Hammoud, M.A.E.R.; Zhan, P.; Hakla, O.; Knio, O.; Hoteit, I. Semantic Segmentation of Mesoscale Eddies in the Arabian Sea: A Deep Learning Approach. *Remote Sens.* **2023**, *15*, 1525. <https://doi.org/10.3390/rs15061525>

Academic Editors: Xiaochen Yang, Qian Du and Yanni Dong

Received: 26 January 2023

Revised: 19 February 2023

Accepted: 22 February 2023

Published: 10 March 2023



Copyright: © 2023 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

Energetic swirls in the ocean, known as eddies, are essential for the transport of heat and material [1] and for energy conversion among dynamics at different scales [2]. In oceanography, eddies with diameters ranging between 10 km and 500 km are known as mesoscale eddies [3]. Mesoscale eddies regulate oceanic mixing and climate by promoting the transport of salt, heat, and carbon on both local and global scales [4,5], and by inducing a strong shear that generates sub-mesoscale processes to produce prominent vertical fluxes in the upper ocean [6,7]. Detecting and tracking mesoscale eddies are therefore important for investigating their properties, transport mechanisms, and their contributions to air–sea coupling processes [8].

Established eddy detection methods depend either on physical or on geometric criteria to classify these mesoscale features from the mainstream flow. Physical criteria involve the determination of dynamical properties such as the magnitude of the sea-level anomaly [9], the velocity gradient [10,11], or the Okubo–Weiss (OW) parameter [12]. Generally, these methods define an eddy based on closed contours of these quantities above a certain threshold

value; however, they are limited by the fact that no single threshold is optimal for the varying oceanic conditions, and the OW parameter is extremely susceptible to noise in sea-surface height (SSH) data [13,14]. To overcome these weaknesses, Chelton et al. 2011 [13] proposed a “threshold-free” method, partitioning the SSH field using a range of monotonically changing thresholds. They further proposed smoothing the SSH field spatially before processing these data to reduce the noise, which may affect the OW parameter.

Unlike the physical criteria, geometric criteria are based on selecting circular or closed streamlines from the quasi-circular flow patterns. One of the most straightforward methods for identifying eddies based on geometric criteria is to identify instantaneous streamlines mapped onto a plane normal to the vortex core [15]. A more sophisticated implementation is to use the winding-angle (WA) method [16], which has been demonstrated to be more efficient and accurate than the Okubo–Weiss method for identifying mesoscale eddies [17]. Nevertheless, a major drawback of established eddy detection techniques is the geometric restriction on the predicted eddies, where these eddies have to be circular or elliptical by construction of the algorithm.

Recently, deep learning techniques have been advancing rapidly, capitalizing on the ability of deep neural networks (DNNs) to learn sophisticated relationships between input and output fields. These advances have been utilized to solve pressing issues in computer vision, pattern recognition, and physics, to name a few examples. For instance, deep learning has been employed to produce high-quality realistic images [18–20], to help with medical assessments of cancer patients [21,22], and to emulate earth systems [23–25]. Deep learning has also been used to detect objects and segment images [26,27] and videos [28]. In particular, semantic segmentation (e.g., Ref. [29]) and instance segmentation (e.g., Refs. [30,31]) aim to delineate all features or individual features, respectively. In segmentation approaches, a DNN is used to output a classification map, of the same size as the input image, which describes the object or class to which a given pixel belongs [32].

Machine learning techniques for detecting eddies were first introduced by Castelani, 2006 [33], who trained a densely connected neural network to detect eddies from sea surface temperature (SST) data. Deep learning techniques were then employed to help improve the detection accuracy. For instance, both Franz et al., 2018 [34], and Lguensat et al., 2018 [35], depended on SSH data to train a convolutional neural network (CNN) to detect eddies. Both groups of investigators employed semantic segmentation approaches to generate classification maps describing the locations and extents of cyclonic and anticyclonic eddies. Moschos et al., 2020 [36], then proposed using SST data rather than SSH data to train a CNN to produce classification maps that identify the locations of cyclonic and anticyclonic eddies. Similarly, Duo et al., 2019 [37], Xu et al., 2019 [38], and Nian et al., 2019 [39], utilized sea-surface anomalies as inputs for the detection and semantic segmentation of mesoscale eddies. More recently, a combination of SST and SSH data and surface speeds were used as inputs to a deep CNN, which predicted the locations and shapes of individual eddies in an instance-segmentation approach [40]. These previously mentioned methods relied on model-generated eddy labels, which limited the predicted shapes of the mesoscale eddies to be either circular or elliptical. In reality, however, the shapes of these eddies may be distorted, necessitating the requirement for more sophisticated labels [41].

In this study, we have trained a DNN to perform the semantic segmentation of cyclonic and anticyclonic eddies in the Arabian Sea by relying on human-labeled and expert-validated annotated data for mesoscale eddies in that sea. Human-based annotations allow for predictions of complex eddy geometries allowing to identify eddies that are neither circular nor elliptical, which is one of the major limitations of non-learning-based algorithms. In particular, our training dataset includes annotations of different shapes that are not necessarily elliptical, which enables the DNN to predict such complex geometries. The annotations are made openly accessible to promote additional developments in learning-based eddy detection. The employed deep neural network comprises a U-Net [29] backbone for feature extraction, with a Pyramid Pooling Module (PPM) [42] as the *decode head* or task-performing network that predicts the output classification maps. We trained this network using as input the SSH, SST, sea-surface salinity (SSS), and zonal and merid-

ional velocity components from the Copernicus Marine Environment Monitoring Service (CMEMS), a global ocean-eddy-resolving reanalysis product [43]. The usage of different surface variables as inputs to the network provides more information about the surface state of the sea and is expected to improve the prediction accuracy compared to the use of a single ocean-surface field. We compared the results obtained by using this trained model to those obtained from another model with the same architecture that we trained using only SSH, SST, or SSS data as the single input, as is typically performed in the literature. We then analyzed the results by examining the sensitivity of the network to different combinations of the input ocean-surface fields. Moreover, to investigate the computational benefits of using “off-the-shelf” networks [44], we compared the results obtained using this DNN to those obtained using a pre-trained ResNet50–FCN [45], which was fine-tuned to predict eddies using a transfer-learning methodology. We also evaluate the performance of a model trained on the Arabian Sea dataset for predicting eddies in the Red Sea to investigate its generalizability for the semantic segmentation of eddies in different basins. This study is the first to analyze the impact of different combinations of ocean-surface fields on the semantic segmentation of eddies, and to examine the utilization of transfer learning for the semantic segmentation of mesoscale eddies.

The remainder of the paper is structured as follows. Section 2 describes the circulation in the Arabian Sea and the data used for training and validating the DNN. The U-Net–PPM, trained from scratch, and the pre-trained ResNet50–FCN are described in Section 3, where the training setup and methodology are presented. The results are presented in Section 4. Finally, Section 5 summarizes the main results and presents the concluding remarks and a discussion of future research directions.

2. Ocean Data

We used the global ocean eddy-resolving reanalysis product from the Copernicus Marine Environment Monitoring Service (CMEMS), which provides reliable and up-to-date information of the global ocean with a comprehensive collection of oceanographic and environmental data and includes information on a wide range of parameters such as temperature, salinity, sea level, ocean currents, and biogeochemical properties. The reanalysis dataset used in this study is available from 1993 to 2020 with a horizontal resolution of $1/12^\circ$, or approximately 8 km [43]. This ocean reanalysis was generated using version 3.1 of the Nucleus for European Models of the Ocean (NEMO) model [46], driven at the surface by the ERA-Interim climate reanalysis dataset [47] of the European Centre for Medium-Range Weather Forecasts (ECMWF). This model assimilates *in situ* data from various sources such as satellites and *in situ* measurements using a reduced-order Kalman filter, including along-track altimeter data, remotely sensed SST data, and *in situ* T/S vertical profiles, and the reanalysis data set was subject to rigorous quality control and validation procedures to ensure its accuracy and reliability, and is continually updated as new observations become available. In this study, daily averaged SSH, SST, SSS, and surface current fields are extracted. The spatial and temporal resolution of the CMEMS product are fine enough to resolve the mesoscale eddies in the Arabian Sea [8]. Figure 1 illustrates a snapshot of the surface currents superimposed on the SSH for a day during the summer season.

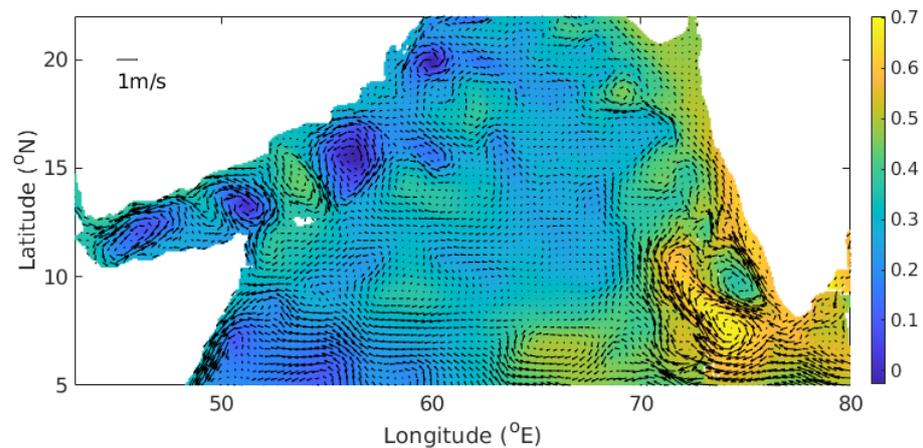


Figure 1. Plot illustrating the SSH and surface currents in the Arabian Sea at a snapshot representative of a typical day during the summer season.

The Somali current along the Somalia coast off northwest Africa makes the Arabian Sea unique compared to other basins with strong western boundary-current systems, since it is not continuous in time but reverses seasonally [48]. This pronounced seasonality is linked to the annually reversing Indian Ocean monsoon. During the southwest monsoon season in summer, anticyclonic oceanic circulation in the Arabian Sea and the northward western-boundary Somali current follow the appearance of a negative curl of the wind stress over the central basin, while the opposite scenario applies during the northeast monsoon season in winter [49]. In addition, the Arabian Sea exhibits pronounced intraseasonal and spatial variability in eddy activity, with enhanced eddy kinetic energy (EKE) during summer, primarily in the western basin. Mesoscale eddies play a key role in modulating the circulation [50] and biological processes [51] in the Arabian Sea—east of the Somalia coast in particular—and they exhibit EKE and eddy-induced transport, which are comparable in magnitude in both the horizontal and vertical directions to other eddy-rich basins in the global oceans [8,52,53].

3. Neural Network

3.1. Neural Network Architecture

In this section, we examine the use of a U-Net–PPM neural-network architecture to segment cyclonic and anticyclonic eddies, and we compare its performance, in terms of prediction accuracy and training time, to that of a pre-trained and fine-tuned Resnet50–FCN. Both network architectures are composed of a CNN backbone for extracting features, followed by a decode head for predicting the classification maps. The first network comprises a U-Net [29] backbone and a PPM [42] decode head, and was trained from randomly initialized model parameters. The second network is a pre-trained CNN that comprises a 50-layer ResNet [54] backbone equipped with a fully convolutional network (FCN) [45] decode head, and we fine-tuned the initial model parameters by training it on our dataset. Both neural network architectures and pipelines are presented in Figure 2, which illustrates the architecture of each DNN along with the convolution blocks used to develop them. This figure portrays the pipeline, showing the input fields, the individual layers/blocks, and the output classification maps.

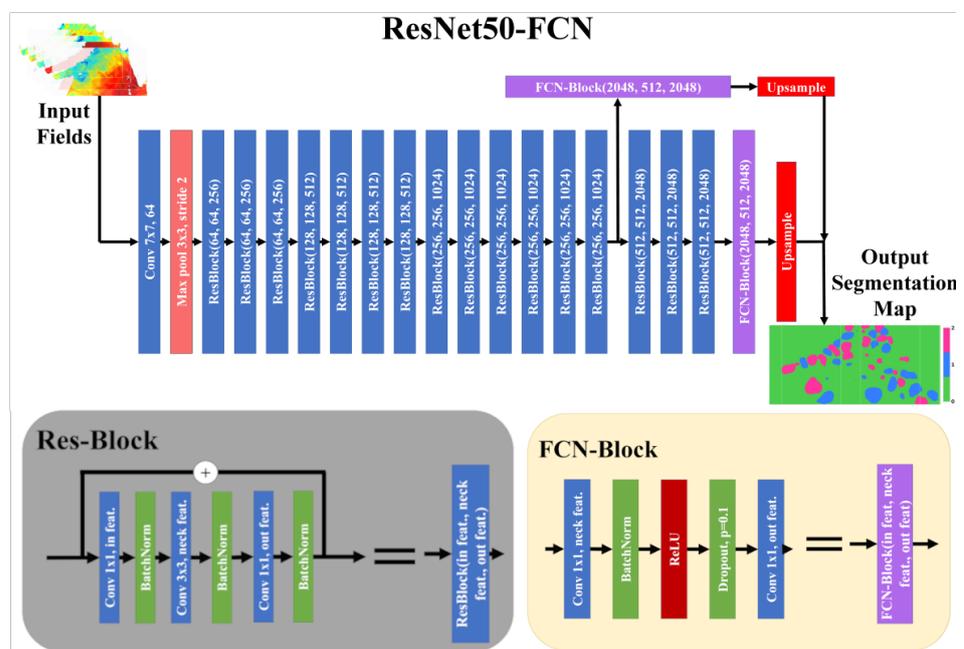
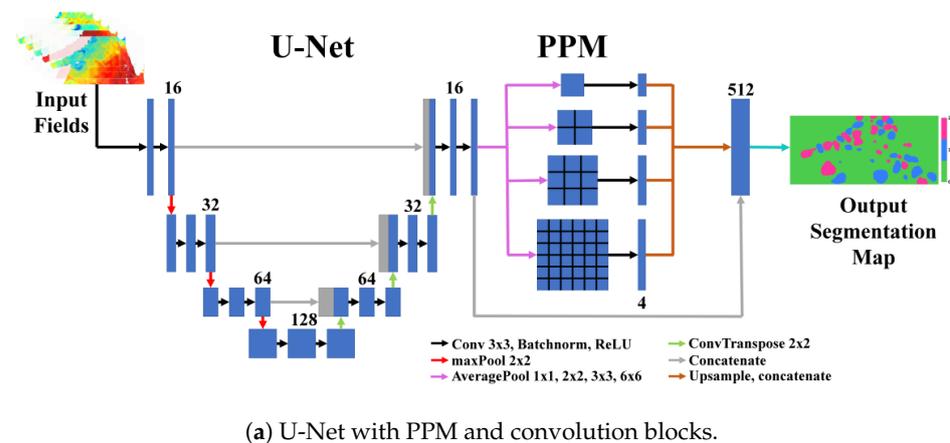


Figure 2. Neural–Network Architectures: Schematics of (a) the U-Net backbone with the PPM decode head, along with the Res- and FCN-blocks, and (b) the ResNet50–FCN that was pre-trained on a subset of the COCO dataset and is available in PyTorch’s torchvision. The deep neural network in (a) is trained by starting from randomized initial parameters, whereas that of (b) was trained from randomly initialized and pre-trained model parameters that are fine-tuned by training on the Arabian Sea dataset.

Figure 2a shows the backbone of the first network, which is a U-Net that comprises the following sequence: a convolutional encoder, bottleneck layers, and an expanding convolutional decoder that mirrors the encoder. Specifically, the U-Net’s encoder includes the sequence of a convolution layer with a 3×3 kernel size and a batchnorm followed by ReLU activation function. This sequence is present twice at each level of the encoder, and a maxpooling layer is employed to transition to the next level. The decoder is a reflection of the U-Net encoder, with transposed convolutions and nearest-neighbor upsampling substituted for the convolutions and max-pooling layers, respectively. Finally, a skip connection concatenates the feature maps from the outputs of each encoder layer to the input of the decoder layer at the same level; this has been shown to promote strong gradient propagation and to preserve information from the encoder layers to subsequent layers in the network [29,54].

The second network consists of a 50-layer ResNet CNN backbone formed by a sequence of Res-Blocks, as shown in Figure 2b. The Res-Blocks [54] consist of three consecutive convolution layers, each followed by a BatchNorm layer, with a skip connection between the input and the output of the Res-Block. The convolution layers have convolution kernels of sizes 1×1 , 3×3 and 1×1 , respectively. To achieve 50 layers, this network employs 16 consecutive Res-Blocks along with an initial convolution layer that is followed by a max pooling layer with a stride of 2. The FCN is merged with the 50-layer ResNet by connecting one FCN-block to the 15th Res-Block and another FCN-block to the last Res-Block. The outputs of these FCN-blocks are upsampled and then joined to predict the classification maps. We used a pre-trained model that is readily available in PyTorch's *torchvision* [55] in a transfer-learning [44] setting, with layers that fit the task-dependent dataset substituted for the first and last layers before training the network on the dataset of interest. This approach is commonly adopted in deep learning applications to reduce the training time, where these pre-trained models were extensively trained on large datasets to provide a large set of features that could be utilized in future tasks [56].

In the present work, we have analyzed the performance of both DNNs in segmenting eddies in the Arabian Sea using different combinations of ocean-surface variables as inputs to the DNN. We trained the first DNN using human-annotated labels, which allows for complex geometrical predictions that differ from the outputs of established eddy-detection algorithms, which are limited to circles and ellipses. Importantly, this study differs from that of Fan et al., 2020 [40], who relied on combining the SSH and SST data into one channel, producing a physically inconsistent quantity. Moreover, their analysis does not report the spatial resolution of the input fields nor the geographical location of the region of interest. Finally, the human-based labels used by Fan et al., 2020 [40], are based solely on SSH data, making them susceptible to larger annotation errors than the ones reported here because they miss the circulation information gained by using the velocities.

3.2. Training and Validation Datasets

We inputted the ocean-surface fields to the DNN to predict the probability that a pixel belongs to a cyclonic eddy, an anticyclonic eddy, or neither (i.e., to the background). These three categories comprise the output classes for which the network predicts a softmax probability that a pixel belongs to a given class. Each ocean-surface variable is represented by an input channel to the network, where the total number of input channels is equal to the number of input variables propagated through the DNN. The input fields are propagated through the DNN, which classifies the image point-by-point according to the class with the highest softmax probability, as further discussed in the following section. Finally, the output classification map is a 2D spatial map with the same resolution as the input, and it indicates whether a grid cell is classified to belong to the background, a cyclonic eddy, or an anticyclonic eddy.

The training and validation datasets rely on the CMEMS ocean reanalysis datasets described in Section 2. In particular, we extracted the ocean-surface fields from the years 2001 to 2010 to generate the training dataset and used those from 2011 as the validation dataset. All surface fields—including the velocity-vector fields—were human-labeled and later validated by an expert to confirm the quality of the labels. This is important since a DNN trained using eddy-detection-based data—cf. [35]—can only represent a surrogate of the eddy-detection model. On the other hand, training a DNN using hand-labeled data that contains labels of different shapes and sizes and which are not constrained to be ellipses or circles, is a surrogate for an arbitrary model described by the human-based annotations. Finally, we note that we have removed the last 10 days of 2010 and the first 10 days of 2011 from the training and validation datasets to avoid any leakage of information between the training and validation datasets that may arise from high-frequency correlations that could bias the performance-metric estimates. This results in training and validation datasets comprising 3642 and 354 samples, respectively.

3.3. Loss Function

We trained both networks by minimizing the cross entropy between the predicted and true masks. The cross entropy is commonly used for multi-class semantic-segmentation problems, cf. [57,58]. In multi-class classification, the cross entropy computes a loss for each class label, and the total loss is reduced by taking the mean or sum of all the loss terms. The cross entropy for C unique classes can be expressed as follows:

$$\mathcal{L}(x; \theta) = -\frac{1}{n_x n_y} \sum_{i=1}^{n_x} \sum_{j=1}^{n_y} \sum_{k=1}^{n_c} w_k y_t(x) \log[\hat{y}_\theta(x)], \quad (1)$$

where n_x , n_y , and n_c are the numbers of horizontal grid cells, vertical grid cells, and classes, respectively. The quantities y_t are the reference annotated labels and \hat{y}_θ are the θ -parameterized labels predicted by the network at a given grid cell for an input image $x \in \mathbb{R}^{n_x n_y n_f}$, where n_f is the number of input ocean fields. The loss employs a class-specific weighting w_k , with $\sum_k w_k = 1$, to mitigate issues that may arise from class imbalance [58], since at any given instance the areas covered by cyclonic and anticyclonic eddies are not equal. The classes are weighted by the inverse of percentage of occurrence of a given class throughout the entire training dataset. During training, we omitted the loss for the background class to avoid overestimating the accuracy of the network because it is considerably easier for the network to predict the background class label, leading to an artificial boost in the overall loss without improving the segmentation accuracy.

3.4. Accuracy Metrics

In addition to the loss, we assessed the performance of the trained model by computing a number of accuracy metrics that are commonly used in semantic segmentation. These metrics include the pixel-wise accuracy (\mathcal{A}), the Jaccard index (i.e., the intersection over the union: IoU), the precision (\mathcal{P}), and the recall (\mathcal{R}) [28,45]. If we define a true positive (TP) as an outcome in which the model correctly predicts a positive class, then a true negative (TN) is an outcome in which the model correctly predicts a negative class. Similarly, a false positive (FP) is an outcome in which the model incorrectly predicts a positive class, and a false negative (FN) is an outcome in which the model incorrectly predicts a negative class.

Using this notation, we can define \mathcal{A} as follows:

$$\mathcal{A} = \frac{TP + TN}{TP + TN + FP + FN} \quad (2)$$

a measure of the fraction of pixels in the network labeled correctly. The IoU measures the overlap of the target mask with the predicted output; this is given by

$$\text{IoU} = \frac{|A \cap B|}{|A \cup B|}, \quad (3)$$

where A is the set of positive-class pixels from the prediction and B is the set of positive-class pixels from the true labels. The precision \mathcal{P} is the ratio of TP to all positive examples, which is a measure of the exactness of the model's classification of a positive class; it is given by

$$\mathcal{P} = \frac{TP}{TP + FP}. \quad (4)$$

Finally, \mathcal{R} is a measure of the percentage of detected ground-truth labels, expressed as follows:

$$\mathcal{R} = \frac{TP}{TP + FN}. \quad (5)$$

3.5. Training Setup

We trained the DNN eddy detection models for a total of 20 epochs, each of which randomly passes over 2500 sampled instances of the input fields and the corresponding

labeled classification maps, which totals to 50,000 random samples during training. Training depends on a stochastic-gradient-descent (SGD) optimization algorithm to minimize the loss function, with momentum 0.9 and weight decay 10^{-4} . We performed an initial parameter search to identify the best learning rate (LR) to maximize the model accuracy. In particular, we trained the DNN for a batch size of 10 and LRs of 0.0001, 0.001, 0.01, 0.02, 0.03, 0.04, and 0.1. We selected the model that maximizes the average accuracy at segmenting cyclonic and anticyclonic eddies as the best model and used it to report our results in Section 4. We performed the same parameter search for each case corresponding to different input surface fields, and we present the eddy-segmentation results from the model with the highest accuracy. We utilized a learning-rate scheduler to reduce the LR by a factor of 10 if the accuracy did not increase for at least three epochs; this enables us to take finer optimization steps in efforts to build an improved model.

4. Experiments and Results

4.1. Semantic Segmentation Using Five Ocean-Surface Fields

In this experiment, the outputs of two U-Net-PPM models—one trained using all five surface ocean fields as an input to the DNN while the other is trained using SSH data only—are contrasted qualitatively against the reference annotations. The parameters of the U-Net-PPM model were randomly initialized, and training was performed using the dataset described in Section 3.2. Figure 3 shows the ground-truth labels and the predicted segmentation maps for randomly selected samples for the case that used all five ocean-surface variables to train the neural network. The results suggest that the trained model is able to predict the location of mesoscale eddies quite comparably to the ground-truth labels. The network also predicts additional eddies that were either missed during annotation or are artifacts or “ghost” eddies, as named by Lguensat et al., 2018 [35]. Moreover, these results suggest that the network smooths out the boundaries of the mesoscale predictions by following the natural variations in the surface fields. This is an interesting takeaway because the labels are based on polygons [59], which are not as smooth as the input fields.

Similarly, Figure 4 shows that the eddy locations predicted using only SSH data as input are also visually comparable to the ground truth labels. The shapes of the mesoscale features, however, are not as accurate as those predicted by the DNN trained on all the ocean-surface variables. In particular, this model enlarges and over-smooths the predicted regions that correspond to mesoscale eddies. The trained model’s predictions also merge multiple labels together and exhibit some artifacts in the predicted segmentation masks. Compared to the case in which all the surface variables were used as input, these predictions are less accurate, as indicated by the prediction masks, inaccurately predicting extremely small eddies or overestimating the corresponding prediction areas. While this section presents a qualitative analysis of the network’s performance at segmenting eddies using different input variables, the following section quantifies the performance boost achieved from using multiple input fields as opposed to using only SSH data.

4.2. Sensitivity of Segmentation Accuracy to the Input Surface Fields

The results of the previous section indicates that the predicted eddies are better captured qualitatively when all five surface variables are used as input to the DNN in comparison to only using the SSH fields. In this section, the performance of the U-Net-PPM DNN model, with randomly initialized trainable parameters, is quantified for different combinations of input ocean-surface fields. The input layer of the network is the only component of the DNN that is modified in this numerical experiment, varying the number of input channels to correspond to the number of input fields. Note that the training parameters, including the LR and batch size, are held fixed throughout these experiments to allow a fair comparison between different models.

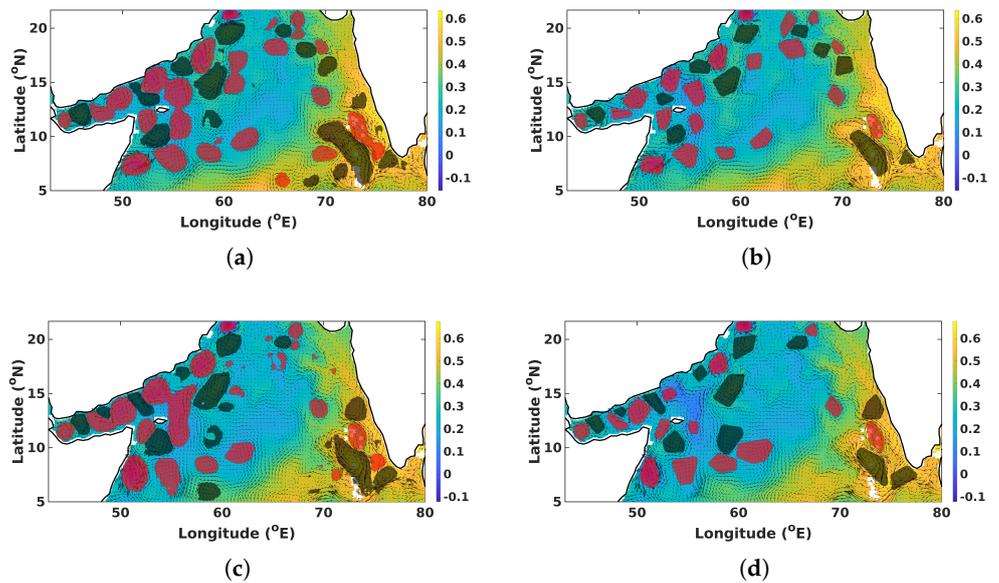


Figure 3. Ground–truth and predicted segmentation masks for the U-Net–PSPnet deep neural network using all five ocean-surface fields. The segmentation maps superpose the SSH data and the surface-velocity vectors. The plots demonstrate that the network captures all the labeled eddies with smoother boundaries. Here, the cyclonic eddies are colored red and the anticyclonic eddies are colored black. The color bars at the sides of each panel indicate the SSH values in meters; (a) Prediction—Sample 1; (b) Ground Truth—Sample 1; (c) Prediction—Sample 2; (d) Ground Truth—Sample 2.

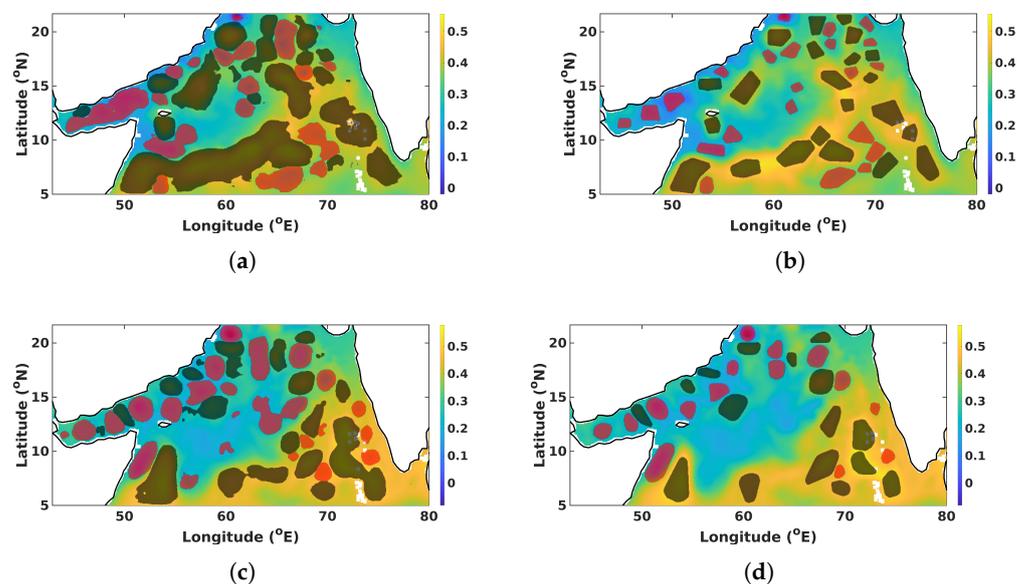


Figure 4. Ground-truth and predicted segmentation masks for the U-Net–PSPnet deep neural network using only the SSH data as input. The segmentation maps are superposed on the SSH data only to show the input to the network alongside the output masks. The plots demonstrate that the network captures the labeled eddies with smoother boundaries than those in the ground-truth plots but lumps some labels together. Here, the cyclonic eddies are colored red and the anticyclonic eddies are colored black. The color bars at the sides of each panel indicate the SSH values in meters; (a) Prediction—Sample 1; (b) Ground Truth—Sample 1; (c) Prediction—Sample 2; (d) Ground Truth—Sample 2.

Table 1 reports the accuracy metrics for the different combinations of the input variables. The results suggest that using all five ocean-surface variables yields a 2% boost in \mathcal{A} , in comparison to using only SSH as input, and a 5% increase in \mathcal{A} in comparison to using SST or SSS as input. In addition, using SSH or the velocities alone yields a value of \mathcal{A} that is roughly 2% greater than using either—or even both—SSS or SST as input. This indicates that both SSH and the velocities are sufficient to achieve the semantic segmentation of eddies with high \mathcal{A} values; however, using all the ocean variables as input provides the neural network with a rich flow of information, which achieves the highest \mathcal{A} values. The mean IoU (mIoU) varies in a way similar to \mathcal{A} , with the largest mIoU obtained when all the ocean-surface fields are used as inputs. Note that the values of the mIoU are expected to be smaller than the \mathcal{A} values because the mIoU penalizes wrong predictions, whereas \mathcal{A} does not.

Table 1. Accuracy Metrics for different inputs to the neural network: The inputs to the U-Net–PSPnet model are varied systematically, and the corresponding accuracy metrics are evaluated and reported in the table. The table shows that the accuracy metrics are improved by approximately 5% when all five ocean-surface variables are used, as opposed to using only the SSH data.

Input Variables	Accuracy		Mean IoU		\mathcal{P}		\mathcal{R}	
	Cyclonic	Anticyclonic	Cyclonic	Anticyclonic	Cyclonic	Anticyclonic	Cyclonic	Anticyclonic
SSH	0.92871	0.93942	0.45984	0.42608	0.47306	0.44126	0.94267	0.92531
SST	0.89841	0.90285	0.35974	0.30395	0.37709	0.31804	0.88659	0.87279
SSS	0.89623	0.90300	0.34649	0.28808	0.36821	0.30932	0.85457	0.80754
SSH and SST	0.92809	0.93197	0.45975	0.40122	0.47049	0.41150	0.95272	0.94140
SSH and SSS	0.93942	0.94252	0.49387	0.43301	0.51596	0.45322	0.92022	0.90661
SSS and SST	0.90210	0.90361	0.36783	0.30508	0.38596	0.31913	0.88679	0.87388
Velocities	0.92511	0.92935	0.44640	0.38892	0.45945	0.40091	0.94016	0.92856
SSH, SSS, and SST	0.92341	0.92849	0.44366	0.38731	0.45447	0.39805	0.94913	0.93487
SST, SSS, and velocities	0.92524	0.93283	0.44927	0.40000	0.46097	0.41388	0.94653	0.92264
SSH, SSS, and velocities	0.92783	0.92855	0.45525	0.38625	0.46985	0.39846	0.93610	0.92648
SST, SSH, and velocities	0.91893	0.92540	0.43280	0.37998	0.44072	0.38907	0.96017	0.94211
SSH, SSS, SST, and velocities	0.94340	0.94855	0.49602	0.45173	0.49095	0.44603	0.93994	0.92783

Finally, this table shows that the network is able to not only achieve large values of \mathcal{A} but also to achieve \mathcal{R} values that are much larger than the \mathcal{P} values. Combining these results with those of Figures 3 and 4 demonstrates that the network predicts the majority of the positive classes; i.e., it predicts the eddies correctly. However, the areal coverage of the eddies is overestimated, which adversely affects the \mathcal{P} values, since the eddies are not perfectly covered by the prediction. These results are also in agreement with the predicted semantic-segmentation maps, in which the regions of the predicted labels are almost always larger than those of the true labels. The \mathcal{P} values may also be lower than the \mathcal{R} values because of class imbalance, where the background covers a larger area than the eddies. Moreover, a high \mathcal{R} value with a lower \mathcal{P} value may also be attributable to the network’s predictions of eddies that were not labeled; i.e., either missed during annotation or non-existent. Finally, the results also indicate that the effects of the surface variables on \mathcal{P} and \mathcal{R} are similar to their effects on \mathcal{A} and mIoU.

Figure 5 presents the neural-network predictions for each combination of the input variables considered in this sensitivity study. This figure also includes the reference human labels to enable straightforward comparisons of the predictions with the reference for each case. When only a single ocean variable is used as input to the DNN, several fictitious eddies are predicted, as indicated by the artifacts, and the areas of the predicted eddies are overestimated. This is similar to the results of Lguensat et al., 2018 [35]. As the number of input variables is increased, the network predictions improve, as indicated by the predicted regions better covering the actual eddies, as given by the labels. In all but the last case, the DNN prediction does not split the eddy boundaries correctly, as is apparent from the continuous predictions that lump several of the labeled eddies into a single predicted eddy. This problem was largely mitigated when the SSH, SST, SSS, and the velocities were used

as input to the network. In that case, the predictions better delineate the eddy boundaries, with a smaller numbers of artifacts than in the other cases.

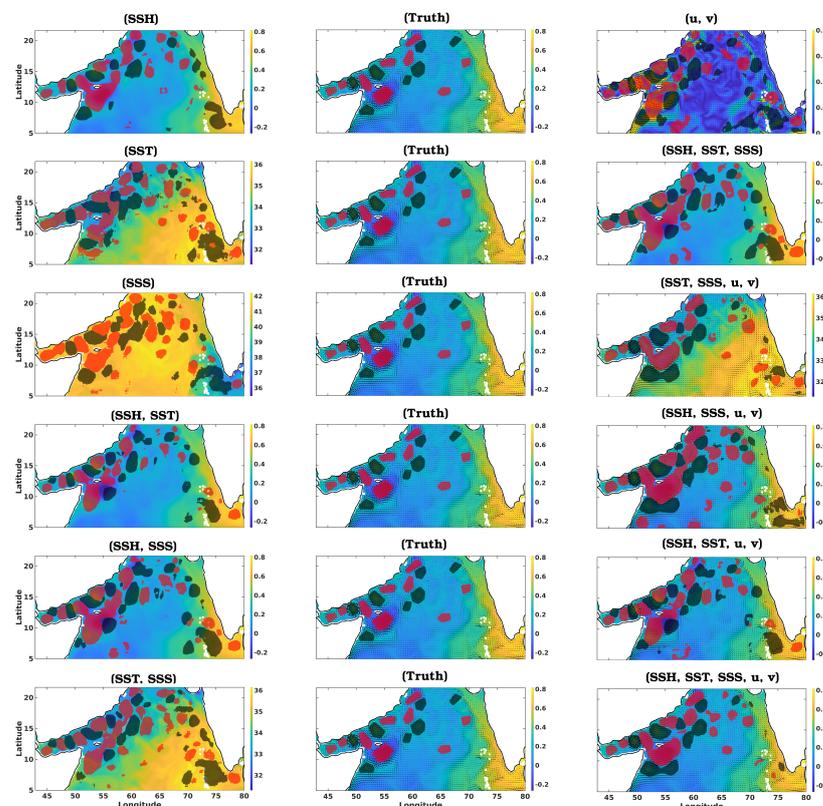


Figure 5. Illustration of the neural network predictions (**left** and **right** columns) for DNNs trained using the input variables indicated by the titles of each subplot, along with the true human labels (**center** column). We used the same input field for all the predictions, and the same truth is repeated in each row to make the comparison easier for each case. Here, cyclonic eddies are colored red and the anticyclonic eddies are colored black. Note that the background and the color bars correspond to the first ocean–surface variable indicated by the subfigure title. The velocity vector field is overlaid on top of the scalar field whenever applicable.

4.3. Transfer Learning Using a Pre-Trained Model

Various pre-trained models are readily available for deployment for a variety of tasks, including object detection [60], human-pose determination [61], and semantic segmentation [28]. These models are generally trained extensively on large datasets—for example, on ImageNet [62]—which helps the network to learn a rich collection of features. These pre-trained models can be further tuned for a particular task using a specific dataset by—for example, in classification problems—reinitializing the input and output layers to accommodate the input dimensions and the output classes required by the new task of interest. Here, we explored the use of a pre-trained ResNet50–FCN [45,63] for the semantic segmentation of mesoscale features in the Arabian Sea using the five ocean–surface fields employed above. The model was pre-trained on a subset of *COCO train2017* using 20 categories that also occur in the *Pascal VOC* dataset. This model achieves an \mathcal{A} of 91.4% and mIoU of 60.5% on *COCO val2017*. We reinitialized the input layer to accommodate five channels, corresponding to the five ocean–surface fields. The classifier was reinitialized to predict three channels, which correspond to the cyclonic-eddy, anticyclonic-eddy, and background classes. The model parameters were fine-tuned by training the model using our human-annotated Arabian Sea ocean-eddies dataset. Furthermore, the same ResNet50–FCN model was trained starting from randomly initialized trainable parameters using the same training dataset, and the resulting

predictions were contrasted to the reference annotations and the predictions of the U-net–PPM and fine-tuned ResNet50–FCN.

Figure 6 shows the reference and predicted segmentation masks for the U-net–PPM, ResNet50–FCN, and fine-tuned ResNet50–FCN. The figure shows that all the predictions perform comparably at detecting the mesoscale eddies. The U-net–PPM predicts several “ghost” eddies that are not present in the ResNet50–FCN predictions. In contrast, both ResNet50–FCN predictions are smoother and have less discontinuous predictions. This could be attributed to the ResNet50–FCN model having ~ 35 million trainable parameters, far more than the ~ 1 million of the U-net–PPM. The figure indicates that the ResNet50–FCN predictions have less “ghost” eddies for the case of the pre-trained model as opposed to the one trained starting from randomly initialized trainable parameters. The predictions are smoother than the labels; however, small artifacts that do not resemble eddies were noticeable in a few samples. Nevertheless, the accuracy of the model is greater than the accuracy of the U-Net–PPM trained with all five ocean-surface variables. This indicates that the pre-trained model possesses a number of features that can be fine-tuned to improve the predictability of eddies.

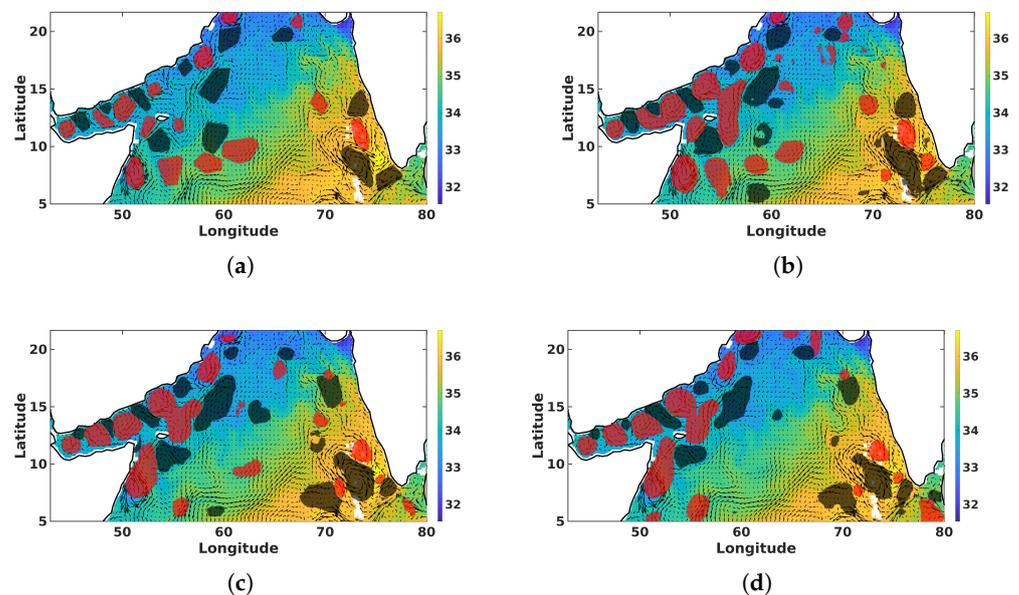


Figure 6. Transfer Learning: Predicted and ground-truth segmentation maps for some randomly selected samples. We obtained these predictions using the fine-tuned ResNet50–FCN model, with all five ocean-surface variables as inputs to the trained model. Cyclonic eddies are colored red and anticyclonic eddies are colored black. The color bars at the sides of each panel indicate the SSH values in meters; (a) Reference; (b) U-Net–PPM; (c) ResNet50–FCN; (d) Pretrained ResNet50–FCN.

Figure 7 outlines the evolution of \mathcal{A} for both cyclonic (C) and anticyclonic (AC) eddies over time, where each marker represents the value of \mathcal{A} at the end of an epoch. The results indicate that the pre-trained ResNet50–FCN model is able to achieve an average \mathcal{A} higher than that of U-Net–PPM for all time steps. Specifically, the pre-trained ResNet50–FCN reaches an average \mathcal{A} of approximately 96% in less than 30 min of training, whereas the U-Net–PPM model achieves an average \mathcal{A} of about 95% in approximately 200 min. This demonstrates that fine-tuning a pre-trained model can help to reduce the consumption of computational resources for such tasks, as the computational time was reduced more than six-fold by fine-tuning the pre-trained network. Furthermore, the randomly initialized ResNet50–FCN was also able to achieve accuracy values larger than those of the U-net–PPM, however, such accuracy values were achieved after about 235 min of training, which is expected from using a large model.

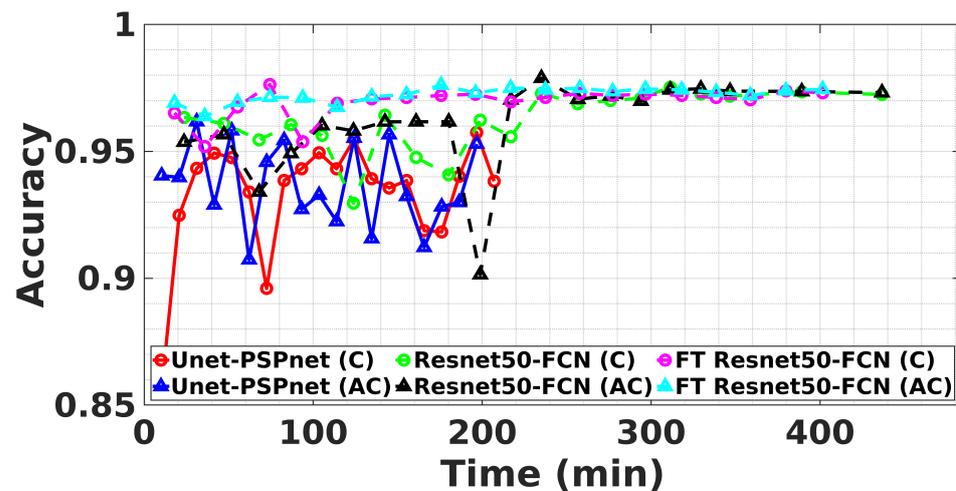


Figure 7. Comparison Learning Curves: Evolution of the pixel accuracy of the U–Net–PSPnet model trained from a random initialization and of a pre–trained ResNet50–FCN model that has been fine-tuned. The curves show that the pre–trained ResNet50–FCN models achieve an accuracy above 95% much faster than the U–Net–PSPnet model. **Note:** FT refers to fine-tuned.

4.4. Generalization to the Red Sea

We finally investigated the performance of the pre-trained model in the semantic segmentation of eddies in a different eddy-rich sea, the Red Sea, using the DNN trained with data labeled for the Arabian Sea oceanic variables. In general, the model’s performance is expected to degrade when the inputs are changed to a different sea because of the change in the input distributions. To detect eddies in the Red Sea, we used the fine-tuned ResNet50–FCN using all five ocean-surface variables (SSH, SST, SSS, and the velocities) as input, as they include the richest set of features and achieve the highest accuracy, as discussed in previous sections.

Figure 8 shows the detected-eddy segmentation masks superimposed on the SSH data and the velocity-vector field. Since no ground-truth labels are available for the Red Sea, we performed a qualitative visual assessment of the performance of the model. Specifically, this figure shows that the model is able to identify most, if not all, large-scale mesoscale eddies in the Red Sea. In comparison to those in the Arabian Sea, eddies in the Red Sea are generally smaller and are not greatly deformed from circular shapes. The detected eddies flank each other in the central and northern Red Sea, and their sizes can become as large as the whole width of the basin. This is due to the effect of the large meridional density gradient that energizes the mesoscale dynamics under the topographic constraints [64].

While some eddies are not detected by this neural network, we suspect that this may have occurred because the resolution of the data may be too coarse to enable eddy segmentation in the narrow Red Sea basin [2]. Nevertheless, the network is able to identify several large and semi-persistent eddies that are known to exist in the central Red Sea [2,65–67]. We expect that better performance could be achieved by retraining the DNN using additional data labeled for the Red Sea. Another approach to mitigate the effects of this domain shift is to perform domain adaptation to push the input distribution closer to the one on which the DNN was trained [68,69].

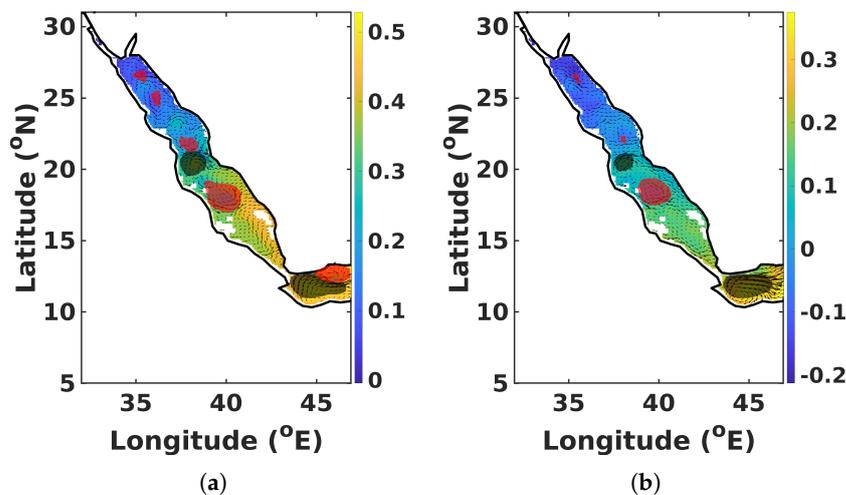


Figure 8. Red Sea Predictions: Predicted semantic-segmentation maps obtained using the fine-tuned ResNet50–FCN model. Since no ground-truth labels are available, we assessed the performance of the model without using any domain generalization to improve the predictions. Despite the facts that the network was not trained for the Red Sea and that the data for this narrow basin are coarse, the network is able to predict the rough locations of the mesoscale eddies. The plot shows the cyclonic eddies colored red and the anticyclonic eddies colored black. The scalar field shown in the plots corresponds to the SSH values in meters; (a) Sample 1; (b) Sample 2.

5. Discussion and Conclusions

We have utilized human-annotated labels for cyclonic and anticyclonic eddies in the Arabian Sea to train a deep neural network for the semantic segmentation of these mesoscale features. We first trained a deep neural network (DNN) with a U-Net feature extractor and a PPM decode head starting from random initial parameters. We explored using only the SSH data as input to the DNN, as is commonly done in the literature, in contrast to using multiple surface-state variables (SSH, SST, SSS, and the velocities). The results demonstrate that using all five ocean-surface variables yields a higher accuracy for eddy detection than using only the SSH data as input.

Accordingly, we varied the input variables systematically to assess the sensitivity of the DNN to these inputs. Our numerical experiments showed that the accuracy of the DNN predictions is most sensitive to the SSH and velocities, and we obtained a boost of 2–3% in accuracy when we included these variables as inputs, as opposed to using only the SST or SSS data. We also explored the use of a transfer-learning technique for the semantic segmentation of eddies. In particular, a ResNet50 backbone with an FCN decode head that were pre-trained on a subset of the COCO dataset were fine-tuned by training them on the Arabian Sea eddies dataset. The results of this experiment demonstrated that using a pre-trained model improves the prediction accuracy by approximately 2%. In addition, a pre-trained model reduced the training time by more than half, by simply fine-tuning the learned convolution kernels to the dataset at hand.

Finally, we evaluated the performance of the eddy detection DNN in qualitatively predicting the eddies in the Red Sea. We used the ResNet50–FCN model that was fine-tuned on the Arabian Sea dataset to predict the eddies in the Red Sea. The results show that the model predicts the larger mesoscale eddies but fails to predict the smaller eddies. The performance of this DNN could be improved by using data labeled for the Red Sea, using domain adaptation and generalization techniques, or using higher-resolution data to train the DNN.

This work could be extended further to track or even predict the locations and shapes of eddies. In addition, one might be interested in applying domain-adaptation and generalization techniques to construct a model to predict eddies on a global scale. Moreover, one could consider examining the performance of these models when remote-sensing data is

used as input or when higher-resolution input fields are used to capture smaller features in the ocean. Finally, we plan to investigate an instance-segmentation approach to detect and delineate individual eddies in the ocean.

Author Contributions: Conceptualization, M.A.E.R.H., P.Z., O.K., I.H.; methodology, M.A.E.R.H. and P.Z.; software, M.A.E.R.H.; validation, P.Z., O.K., I.H.; formal analysis, M.A.E.R.H., P.Z., O.K., I.H.; investigation, M.A.E.R.H.; resources, O.K. and I.H.; data curation, M.A.E.R.H., P.Z., O.H.; writing—original draft preparation, M.A.E.R.H. and P.Z.; writing—review and editing, M.A.E.R.H., P.Z., O.H., O.K., I.H.; visualization, M.A.E.R.H. and P.Z.; supervision, O.K. and I.H.; project administration, O.K. and I.H.; funding acquisition, O.K. and I.H. All authors have read and agreed to the published version of the manuscript.

Funding: The research reported in this publication was supported by the Virtual Red Sea Initiative Award #REP/1/3268-01-01.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: The datasets used in this study are publicly available in <https://doi.org/10.5061/dryad.ht76hdm3>.

Acknowledgments: We thank Issam Lakkis for his helpful discussions. We also thank Jad Bhamdouni, Louis Youssef, Nour Qaraqira, and Omar AlLahham for their help with acquiring the annotation.

Conflicts of Interest: The authors declare no conflict of interest.

References

- Kundu, P.K.; Cohen, I.M.; Dowling, D.R. Chapter 10—Boundary Layers and Related Topics. In *Fluid Mechanics*, 6th ed.; Kundu, P.K., Cohen, I.M., Dowling, D.R., Eds.; Academic Press: Boston, MA, USA, 2016; pp. 469–532.
- Zhan, P.; Subramanian, A.C.; Yao, F.; Kartadikaria, A.R.; Guo, D.; Hoteit, I. The eddy kinetic energy budget in the Red Sea. *J. Geophys. Res. Ocean.* **2016**, *121*, 4732–4747. [[CrossRef](#)]
- Tansley, C.E.; Marshall, D.P. Flow past a Cylinder on a β Plane, with Application to Gulf Stream Separation and the Antarctic Circumpolar Current. *J. Phys. Oceanogr.* **2001**, *31*, 3274–3283. [[CrossRef](#)]
- McWilliams, J.C. The nature and consequences of oceanic eddies. In *Ocean Modeling in an Eddying Regime*; American Geophysical Union: Washington, DC, USA, 2008; pp. 5–15.
- Sommer, J.L.; d’Ovidio, F.; Madec, G. Parameterization of subgrid stirring in eddy resolving ocean models. Part 1: Theory and diagnostics. *Ocean. Model.* **2011**, *39*, 154–169. [[CrossRef](#)]
- Su, Z.; Wang, J.; Klein, P.; Thompson, A.F.; Menemenlis, D. Ocean submesoscales as a key component of the global heat budget. *Nat. Commun.* **2018**, *9*, 1–8. [[CrossRef](#)]
- Zhan, P.; Guo, D.; Krokos, G.; Dong, J.; Duran, R.; Hoteit, I. Submesoscale Processes in the Upper Red Sea. *J. Geophys. Res. Ocean.* **2022**, *127*, e2021JC018015. [[CrossRef](#)]
- Zhan, P.; Guo, D.; Hoteit, I. Eddy-Induced Transport and Kinetic Energy Budget in the Arabian Sea. *Geophys. Res. Lett.* **2020**, *47*, e2020GL090490. [[CrossRef](#)]
- Chaigneau, A.; Eldin, G.; Dewitte, B. Eddy activity in the four major upwelling systems from satellite altimetry (1992–2007). *Prog. Oceanogr.* **2009**, *83*, 117–123. [[CrossRef](#)]
- Isern-Fontanet, J.; García-Ladona, E.; Font, J. Identification of marine eddies from altimetric maps. *J. Atmos. Ocean. Technol.* **2003**, *20*, 772–778. [[CrossRef](#)]
- Kurian, J.; Colas, F.; Capet, X.; McWilliams, J.C.; Chelton, D.B. Eddy properties in the California current system. *J. Geophys. Res. Ocean.* **2011**, *116*. [[CrossRef](#)]
- Chelton, D.B.; Schlax, M.G.; Samelson, R.M.; de Szoeke, R.A. Global observations of large oceanic eddies. *Geophys. Res. Lett.* **2007**, *34*. [[CrossRef](#)]
- Chelton, D.B.; Schlax, M.G.; Samelson, R.M. Global observations of nonlinear mesoscale eddies. *Prog. Oceanogr.* **2011**, *91*, 167–216. [[CrossRef](#)]
- Zhan, P.; Subramanian, A.C.; Yao, F.; Hoteit, I. Eddies in the Red Sea: A statistical and dynamical study. *J. Geophys. Res. Ocean.* **2014**, *119*, 3909–3925. [[CrossRef](#)]
- Robinson, S.K. Coherent motions in the turbulent boundary layer. *Annu. Rev. Fluid Mech.* **1991**, *23*, 601–639. [[CrossRef](#)]
- Sadarjoen, I.A.; Post, F.H. Detection, quantification, and tracking of vortices using streamline geometry. *Comput. Graph.* **2000**, *24*, 333–341. [[CrossRef](#)]
- Chaigneau, A.; Gizolme, A.; Grados, C. Mesoscale eddies off Peru in altimeter records: Identification algorithms and eddy spatio-temporal patterns. *Prog. Oceanogr.* **2008**, *79*, 106–119. [[CrossRef](#)]

18. Abdal, R.; Qin, Y.; Wonka, P. Image2StyleGAN: How to Embed Images Into the StyleGAN Latent Space? In Proceedings of the 2019 IEEE/CVF International Conference on Computer Vision (ICCV), Seoul, Republic of Korea, 27–28 October 2019; pp. 4431–4440. [\[CrossRef\]](#)
19. Karras, T.; Aittala, M.; Laine, S.; Härkönen, E.; Hellsten, J.; Lehtinen, J.; Aila, T. Alias-Free Generative Adversarial Networks. In Proceedings of the 35th Conference on Neural Information Processing Systems (NeurIPS), online, 6–14 December 2021.
20. Rombach, R.; Blattmann, A.; Lorenz, D.; Esser, P.; Ommer, B. High-resolution image synthesis with latent diffusion models. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, New Orleans, LA, USA, 18–24 June 2022; pp. 10684–10695.
21. Wang, G.; Li, W.; Ourselin, S.; Vercauteren, T. Automatic brain tumor segmentation using cascaded anisotropic convolutional neural networks. In Proceedings of the International MICCAI Brainlesion Workshop, Quebec City, QC, Canada, 14 September 2017; pp. 178–190.
22. Wu, N.; Phang, J.; Park, J.; Shen, Y.; Huang, Z.; Zorin, M.; Jastrzebski, S.; Févry, T.; Katsnelson, J.; Kim, E.; et al. Deep Neural Networks Improve Radiologists' Performance in Breast Cancer Screening. *IEEE Trans. Med. Imaging* **2019**. [\[CrossRef\]](#)
23. Rasp, S.; Pritchard, M.S.; Gentine, P. Deep learning to represent subgrid processes in climate models. *Proc. Natl. Acad. Sci. USA* **2018**, *115*, 9684–9689. [\[CrossRef\]](#)
24. Beucler, T.; Pritchard, M.; Rasp, S.; Ott, J.; Baldi, P.; Gentine, P. Enforcing Analytic Constraints in Neural Networks Emulating Physical Systems. *Phys. Rev. Lett.* **2021**, *126*. [\[CrossRef\]](#)
25. Keisler, R. Forecasting Global Weather with Graph Neural Networks. *arXiv* **2022**, arXiv:2202.07575.
26. Noh, H.; Hong, S.; Han, B. Learning deconvolution network for semantic segmentation. In Proceedings of the IEEE International Conference on Computer Vision, Santiago, Chile, 7–13 December 2015; pp. 1520–1528.
27. Yurtkulu, S.C.; Şahin, Y.H.; Unal, G. Semantic segmentation with extended DeepLabv3 architecture. In Proceedings of the 2019 27th Signal Processing and Communications Applications Conference (SIU), Sivas, Turkey, 24–26 April 2019; pp. 1–4.
28. Contributors, M. MMSegmentation: OpenMMLab Semantic Segmentation Toolbox and Benchmark. 2020. Available online: <https://github.com/open-mmlab/mms Segmentation> (accessed on 23 February 2023).
29. Ronneberger, O.; Fischer, P.; Brox, T. U-Net: Convolutional Networks for Biomedical Image Segmentation. In Proceedings of the MICCAI 2015: 18th International Conference, Munich, Germany, 5–9 October 2015. [\[CrossRef\]](#)
30. He, K.; Gkioxari, G.; Dollár, P.; Girshick, R. Mask r-cnn. In Proceedings of the IEEE International Conference on Computer Vision, Venice, Italy, 22–29 October 2017; pp. 2961–2969.
31. Bochkovskiy, A.; Wang, C.Y.; Liao, H.Y.M. Yolov4: Optimal speed and accuracy of object detection. *arXiv* **2020**, arXiv:2004.10934.
32. Peng, C.; Zhang, X.; Yu, G.; Luo, G.; Sun, J. Large Kernel Matters – Improve Semantic Segmentation by Global Convolutional Network. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Honolulu, HI, USA, 21–26 July 2017; pp. 4353–4361.
33. Castellani, M. Identification of eddies from sea surface temperature maps with neural networks. *Int. J. Remote. Sens.* **2006**, *27*, 1601–1618. [\[CrossRef\]](#)
34. Franz, K.; Roscher, R.; Milioto, A.; Wenzel, S.; Kusche, J. Ocean Eddy Identification and Tracking Using Neural Networks. In Proceedings of the IGARSS 2018—2018 IEEE International Geoscience and Remote Sensing Symposium, Valencia, Spain, 23–27 July 2018; pp. 6887–6890. [\[CrossRef\]](#)
35. Lguensat, R.; Sun, M.; Fablet, R.; Tandeo, P.; Mason, E.; Chen, G. EddyNet: A Deep Neural Network For Pixel-Wise Classification of Oceanic Eddies. In Proceedings of the IGARSS 2018—2018 IEEE International Geoscience and Remote Sensing Symposium, Valencia, Spain, 23–27 July 2018; pp. 1764–1767. [\[CrossRef\]](#)
36. Moschos, E.; Schwander, O.; Stegner, A.; Gallinari, P. Deep-SST-Eddies: A Deep Learning Framework to Detect Oceanic Eddies in Sea Surface Temperature Images. In Proceedings of the ICASSP 2020—2020 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), Barcelona, Spain, 4–8 May 2020; pp. 4307–4311. [\[CrossRef\]](#)
37. Duo, Z.; Wang, W.; Wang, H. Oceanic Mesoscale Eddy Detection Method Based on Deep Learning. *Remote. Sens.* **2019**, *11*, 1921. [\[CrossRef\]](#)
38. Xu, G.; Cheng, C.; Yang, W.; Xie, W.; Kong, L.; Hang, R.; Ma, F.; Dong, C.; Yang, J. Oceanic Eddy Identification Using an AI Scheme. *Remote. Sens.* **2019**, *11*, 1349. [\[CrossRef\]](#)
39. Nian, R.; Cai, Y.; Zhang, Z.; He, H.; Wu, J.; Yuan, Q.; Geng, X.; Qian, Y.; Yang, H.; He, B. The Identification and Prediction of Mesoscale Eddy Variation via Memory in Memory With Scheduled Sampling for Sea Level Anomaly. *Front. Mar. Sci.* **2021**, *8*, 1689. [\[CrossRef\]](#)
40. Fan, Z.; Zhong, G.; Li, H. A Feature Fusion Network for Multi-modal Mesoscale Eddy Detection. In Proceedings of the 27th International Conference on Neural Information Processing (ICONIP2020), Bangkok, Thailand, 23–27 November 2020; pp. 51–61.
41. Chen, G.; Han, G.; Yang, X. On the intrinsic shape of oceanic eddies derived from satellite altimetry. *Remote. Sens. Environ.* **2019**, *228*, 75–89. [\[CrossRef\]](#)
42. Zhao, H.; Shi, J.; Qi, X.; Wang, X.; Jia, J. Pyramid Scene Parsing Network. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Honolulu, HI, USA, 21–26 July 2017.
43. Ferry, N.; Parent, L.; Garric, G.; Barnier, B.; Jourdain, N.C. Mercator global Eddy permitting ocean reanalysis GLORYS1V1: Description and results. *Mercat.-Ocean. Q. Newsl.* **2010**, *36*, 15–27.
44. Weiss, K.; Khoshgoftaar, T.M.; Wang, D. A survey of transfer learning. *J. Big Data* **2016**, *3*, 1–40. [\[CrossRef\]](#)

45. Long, J.; Shelhamer, E.; Darrell, T. Fully convolutional networks for semantic segmentation. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Boston, MA, USA, 7–12 June 2015; pp. 3431–3440.
46. Madec, G.; Bourdallé-Badie, R.; Bouttier, P.A.; Bricaud, C.; Bruciaferri, D.; Calvert, D.; Chanut, J.; Clementi, E.; Coward, A.; Delrosso, D.; et al. NEMO Ocean Engine. 2017. Available online: <https://www.nemo-ocean.eu/doc/> (accessed on 23 February 2023).
47. Hersbach, H.; Bell, B.; Berrisford, P.; Hirahara, S.; Horányi, A.; Muñoz-Sabater, J.; Nicolas, J.; Peubey, C.; Radu, R.; Schepers, D.; et al. ERA5 hourly data on single levels from 1979 to present. Copernicus Climate Change Service (C3S) Climate Data Store (CDS). 2018. Available online: <https://cds.climate.copernicus.eu/cdsapp#!/dataset/reanalysis-era5-single-levels?tab=overview> (accessed on 23 February 2023).
48. Schott, F.A.; Xie, S.P.; McCreary, J.P., Jr. Indian Ocean circulation and climate variability. *Rev. Geophys.* **2009**, *47*. [[CrossRef](#)]
49. Wang, H.; McClean, J.L.; Talley, L.D.; Yeager, S. Seasonal cycle and annual reversal of the Somali Current in an eddy-resolving global ocean model. *J. Geophys. Res. Ocean.* **2018**, *123*, 6562–6580. [[CrossRef](#)]
50. Fischer, A.S.; Weller, R.A.; Rudnick, D.L.; Eriksen, C.C.; Lee, C.M.; Brink, K.H.; Fox, C.A.; Leben, R.R. Mesoscale eddies, coastal upwelling, and the upper-ocean heat budget in the Arabian Sea. *Deep. Sea Res. Part II Top. Stud. Oceanogr.* **2002**, *49*, 2231–2264. [[CrossRef](#)]
51. Sergey, P. Mesoscale Eddies of Arabian Sea: Physical-biological Interactions. *Int. J. Mar. Sci.* **2012**, *2*, 51–56. [[CrossRef](#)]
52. Scharffenberg, M.G.; Stammer, D. Seasonal variations of the large-scale geostrophic flow field and eddy kinetic energy inferred from the TOPEX/Poseidon and Jason-1 tandem mission data. *J. Geophys. Res.* **2010**, *115*. [[CrossRef](#)]
53. Rouillet, G.; Capet, X.; Maze, G. Global interior eddy available potential energy diagnosed from Argo floats. *Geophys. Res. Lett.* **2014**, *41*, 1651–1656. [[CrossRef](#)]
54. He, K.; Zhang, X.; Ren, S.; Sun, J. Deep Residual Learning for Image Recognition. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Las Vegas, NV, USA, 27–30 June 2016.
55. Paszke, A.; Gross, S.; Massa, F.; Lerer, A.; Bradbury, J.; Chanan, G.; Killeen, T.; Lin, Z.; Gimelshein, N.; Antiga, L.; et al. PyTorch: An Imperative Style, High-Performance Deep Learning Library. In Proceedings of the Annual Conference on Neural Information Processing Systems 2019 (NeurIPS 2019), Vancouver, BC, Canada, 8–14 December 2019; Volume 32.
56. Wang, K.; Gao, X.; Zhao, Y.; Li, X.; Dou, D.; Xu, C.Z. Pay attention to features, transfer learn faster CNNs. In Proceedings of the International Conference on Learning Representations, New Orleans, LA, USA, 6–9 May 2019.
57. Garcia-Garcia, A.; Orts-Escolano, S.; Oprea, S.; Villena-Martinez, V.; Martinez-Gonzalez, P.; Garcia-Rodriguez, J. A survey on deep learning techniques for image and video semantic segmentation. *Appl. Soft Comput.* **2018**, *70*, 41–65. [[CrossRef](#)]
58. Jadon, S. A survey of loss functions for semantic segmentation. In Proceedings of the 2020 IEEE Conference on Computational Intelligence in Bioinformatics and Computational Biology (CIBCB), online, 27–29 October 2020; pp. 1–7.
59. Torralba, A.; Russell, B.C.; Yuen, J. Labelme: Online image annotation and applications. *Proc. IEEE* **2010**, *98*, 1467–1484. [[CrossRef](#)]
60. Zhou, Y.; Yang, X.; Zhang, G.; Wang, J.; Liu, Y.; Hou, L.; Jiang, X.; Liu, X.; Yan, J.; Lyu, C.; et al. MMRotate: A Rotated Object Detection Benchmark using PyTorch. In Proceedings of the 30th ACM International Conference on Multimedia, Lisbon, Portugal, 10–14 October 2022.
61. Contributors, M. OpenMMLab Pose Estimation Toolbox and Benchmark. 2020. Available online: <https://github.com/open-mmlab/mmpose> (accessed on 23 February 2023).
62. Deng, J.; Dong, W.; Socher, R.; Li, L.J.; Li, K.; Fei-Fei, L. ImageNet: A large-scale hierarchical image database. In Proceedings of the 2009 IEEE Conference on Computer Vision and Pattern Recognition, Miami, FL, USA, 20–1 June 2009; pp. 248–255. [[CrossRef](#)]
63. Marcel, S.; Rodriguez, Y. Torchvision the Machine-Vision Package of Torch. In Proceedings of the 18th ACM International Conference on Multimedia, Florence, Italy, 25–29 October 2010; pp. 1485–1488.
64. Zhan, P.; Krokos, G.; Guo, D.; Hoteit, I. Three-dimensional signature of the Red Sea eddies and eddy-induced transport. *Geophys. Res. Lett.* **2019**, *46*, 2167–2177. [[CrossRef](#)]
65. Yao, F.; Hoteit, I.; Pratt, L.J.; Bower, A.S.; Köhl, A.; Gopalakrishnan, G.; Rivas, D. Seasonal overturning circulation in the Red Sea: 2. Winter circulation. *J. Geophys. Res. Ocean.* **2014**, *119*, 2263–2289. [[CrossRef](#)]
66. Yao, F.; Hoteit, I.; Pratt, L.J.; Bower, A.S.; Zhai, P.; Köhl, A.; Gopalakrishnan, G. Seasonal overturning circulation in the Red Sea: 1. Model validation and summer circulation. *J. Geophys. Res. Ocean.* **2014**, *119*, 2238–2262. [[CrossRef](#)]
67. Zhan, P.; Gopalakrishnan, G.; Subramanian, A.C.; Guo, D.; Hoteit, I. Sensitivity Studies of the Red Sea Eddies Using Adjoint Method. *J. Geophys. Res. Ocean.* **2018**, *123*, 8329–8345. [[CrossRef](#)]
68. Sun, B.; Feng, J.; Saenko, K. Return of Frustratingly Easy Domain Adaptation. In Proceedings of the 30th AAAI Conference on Artificial Intelligence, Phoenix, AR, USA, 12–17 February 2016.
69. Sun, B.; Saenko, K. Deep CORAL: Correlation Alignment for Deep Domain Adaptation. In Proceedings of the 14th European Conference on Computer Vision, Amsterdam, The Netherlands, 8–16 October 2016.

Disclaimer/Publisher’s Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.