# ARU2-Net: A Deep Learning Approach for Global-Scale Oceanic Eddy Detection

| | |
|---|---|
| Journal: | *Journal of Selected Topics in Applied Earth Observations and Remote Sensing* |
| Manuscript ID | JSTARS-2024-00522 |
| Manuscript type: | Regular |
| Date Submitted by the Author: | 13-Feb-2024 |
| Complete List of Authors: | Geng, Junmin; Qingdao University, the Department of Computer Science and Technology<br>Gao, He; Qingdao University, College of Computer Science and Technology<br>Huang, Baoxiang; Qingdao University, Colledge of computer science and technology<br>Radenkovic, Milena; University of Nottingham, School of Computer Science<br>Chen, Ge; Department of Marine Technology, College of Information Science and Engineering; Ocean University of China, |
| Keywords: | Artificial intelligence |
| Subject Category<br>Please select the subject category that best reflects the scope of your manuscript: | Methodologies and Applications to: Oceans and Water |
| | |

# ARU$^2$-Net: A Deep Learning Approach for Global-Scale Oceanic Eddy Detection

Junmin Geng, He Gao, Baoxiang Huang, *Member, IEEE,* Milena Radenkovic, *Member, IEEE* and Ge Chen, *Senior Member, IEEE*

***Abstract*—Ocean eddies have a significant impact on marine ecosystems and the climate because they transport essential substances in the ocean. Detection of ocean eddies has become one of the most active topics in physical ocean research. Current eddy detection techniques only identify tiny regions and are mostly based on conventional detection techniques. While the proposed method has achieved satisfactory results in detecting eddy in limited regions, its performance is not optimal in the global domain. In this paper, a high-precision attention residual U$^2$-Net(ARU$^2$-Net) model is proposed to address this problem. The model is based on the U$^2$-Net model in combination with the Convolutional Block Attention Module (CBAM). By refining the attention features, the CBAM makes the model more attentive to the feature information of the eddy surface from both channel and spatial perspectives, improving the performance of the original base model while keeping the computational cost low. In addition, we use two feature information, sea level anomaly and sea surface temperature, to realize the fusion input of multidimensional information. Finally, we demonstrate the effectiveness of our approach on the global eddy dataset, achieving a test performance of 94.926%, significantly exceeding the results of previous models for the global region.**

***Index Terms*—Global ocean eddy; Deep learning; Attention mechanism; Intelligent detection; ARU$^2$-Net Architecture;**

## I. Introduction

OCEANNIC eddies are rotational motions of seawater controlled by the geostrophic eddy conservation equation and have different forms and scales. They are highly variable on horizontal and vertical scales, with propagation speeds on the scale of kilometers[1], [2]. In the Northern Hemisphere, cyclonic eddies are influenced by the Koch force, which causes water to spread outward, creating negative anomalous sea surface heights. Cyclonic eddies drive cold water upwards and are known as cold eddies. On the contrary, anticyclonic eddies cause the sea surface water to gather inside, resulting in positive anomaly sea surface heights. Anticyclonic eddies carry warm water downwards and are called warm eddies. Eddies have important effects on material cycling, energy transport and loop coupling[3], [4]. The strong auxiliary sinking motion within the ocean eddy drives the growth of phytoplankton and reduces the peak chlorophyll concentration in the subsurface layer to a certain extent[5]. The ocean eddy facilitates vertical movement and mixing within the oceanic layers through the exchange of matter and energy, exerting a significant influence on global ocean circulation and air-sea interaction[6], [7]. Hence, research on identifying ocean eddies plays a pivotal role and has emerged as a current focal point in ocean research.

During the initial period, eddies were identified using manual visual methods, but this approach has significant limitations and is susceptible to subjective differences. As a result, individuals commenced employing mathematical knowledge to detect mesoscale eddies. Nichol et al.[8] employed computer-based image analysis to identify regions with similar grayscale values and extract features that resemble eddy structures. Due to the intricate nature of ocean remote sensing images, extracting features for eddy detection poses significant challenges. Peckinpaugh and Holyer et al.[9] presented a method for eddy detection using the Hough transform circle detection operator[10] based on edge detection in remote sensing images[8]. This method relies on edge detection in remote sensing images[11]. Because the shapes of eddies are irregular polygons, this detection method is relatively rough. Eddies' two- or three-dimensional structures were later used by people to recognize them based on the physical properties of the flow field. The main methods include the physical parameter method[12], [13], the flow field geometric method[14], [15], and the machine learning method[16]. The physical parameter method detects eddy regions by calculating the gradient of ocean physical parameters, nevertheless has a problem of threshold selection. The flow field geometric method, based on fluid mechanics theory, detects eddies by analyzing flow field data and geometric shapes, but it requires smooth and interpolate treatment for data quality.

With the widespread application of deep learning in ocean remote sensing images in recent years[17], [18], [19], [20], a large number of researchers have also started attempting to use deep learning for the purpose of identifying and detecting oceanic eddies[21]. This innovative approach has quickly become a hot topic and has led the wave of research in the field of ocean eddies[19]. Inspired by the U-Net[22] architecture, Lguensat et al.[23] proposed the EddyNet network model.

Junmin Geng is with Qingdao University, Qingdao, 266071, China (e-mail: 2022025872@qdu.edu.cn).

He Gao is with Qingdao University, Qingdao, 266071, China (e-mail: gaohe1@qdu.edu.cn).

Corresponding author: Baoxiang Huang is with the College of Computer Science and Technology, Qingdao University, Qingdao 266071, China, and also with the Laboratory for Regional Oceanography and Numerical Modeling, Laoshan Laboratory, Qingdao, 266100, China. (e-mail: hbx3726@163.com, baoxianghuang@qdu.edu.cn).

Milena Radenkovic is with the School of Computer Science and Information Technology, The University of Nottingham, Nottingham, NG8 1BB UK. (e-mail: milena.radenkovic@nottingham.ac.uk)

Ge Chen is with the Frontiers Science Center for Deep Ocean Multispheres and Earth System, School of Marine Technology, Ocean University of China, Qingdao 266100, China, and also with the Laboratory for Regional Oceanography and Numerical Modeling, Laoshan Laboratory, Qingdao, 266100, China. (e-mail: gechen@ouc.edu.cn)

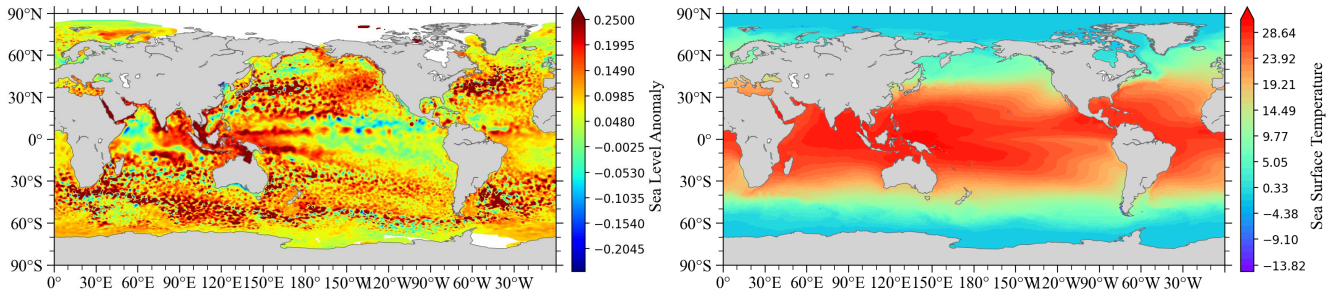


Fig. 1: (a) Map of global sea level anomalies. (b) Global sea surface temperature.

Subsequently, Lguensat[24] developed the EddyResNet based on the ResNet[25] architecture, a residual neural network, and created the EddyVNet, which incorporates eddy evolution during the detection process using a structure similar to V-Net[26]. By adopting more advanced architectural designs, the eddy identification performance has been further improved compared to EddyNet. However, these models only achieve significant classification results in the eddy-rich South Atlantic region, showing only moderate performance in global scale assessments. The reason may be the differences in the characterization and distribution of eddies in other regions compared to the South Atlantic region. In addressing the segmentation challenges posed by multi-scale ocean eddies, the Pyramid Scene Parsing Network (PSPNet)[27] can integrate semantic and detailed features from various layers. This integration enables the effective detection of a multitude of eddies, particularly small-scale ones, within the North Pacific Subtropical Countercurrent[28]. Zhao et al.[29] introduced the Pyramid Separated Attention (PSA) eddy detection structure (PSA-EDUNet). Through non-linear connections, this architecture ensures minimal loss of feature information and incorporates PSA modules to enhance feature extraction. This model achieved excellent detection results in the Kuroshio Extension (KE) and the Southwestern Atlantic region, particularly in the detection of eddy boundaries and small-scale eddies. Subsequently, Huang et al. combined Convolutional Neural Network (CNN) and XGBoost to extract representative vertical features and identify eddies from profile feature information, thereby extending the capability to identify eddies based on altimeter data[30], [31], [32]. But the current vertical profile dataset are not sufficient, placing some limitations on the accuracy and reliability of eddy identification.

These above-mentioned methods have shown satisfactory results in regions with dense eddy distribution, but given the global ubiquity of ocean eddies, these methods have been unable to accurately learn the complex and variable features of the sea surface. The contribution of this paper can be outlined as follows.

1) A high-accuracy deep learning model named ARU$^2$-Net is proposed for global ocean eddy detection. The advantage of this model is that it can capture the global information of the vortex surface at different scales by mixing different receptive fields in the ARU module.
2) The Convolutional Block Attention Module(CBAM)[33] is integrated into the ARU$^2$-Net model. The CBAM has the ability to focus on the characteristics of eddy surfaces from both a channel and spatial perspectives. This addresses complex and diverse eddies while improving the efficiency of the initial foundational model with low rise in computing expenditure.
3) Extensive experimetnal results demonstrate that the proposed method can achieve promising detection results on global scale of eddies.

**Algorithm 1** Eddies labeling algorithm

**Require:** $\mathbf{X_{AE}}$, $\mathbf{X_{CE}}$
**Ensure:** Label matrix E
datas $\leftarrow \mathbf{X_{AE}} + \mathbf{X_{CE}}$
$E[0:720, 0:1440] = 0$;
**for** $n = 0$ **to**len$(datas)$ **do**
  Edge_lon = datas[9:29];
  Edge_lat = datas[29:49];
  minx, maxx, miny, maxy = get_extremum(Edge_lon, Edge_lat);
  X0 = Point_x(minx, 0); X1 = Point_x(maxx, 1);
  Y0 = Point_y(miny, 0); Y1 = Point_y(maxy, 1);
  judge_x, judge_y = get_allPoint(X0, X1, Y0, Y1);
  **for** $L = 0$ **to** len$(judge_x)$ **do**
    Pnpoly(len(Edge_lon), Edge_lon, Edge_lat, judge_x[L], judge_y[L]);
    **if** $Pnpoly = True$ **and** $Edge\_flag =' AE'$ **then**
      $f[L] = 1$;
    **else if** $Pnpoly = True$ **and** $Edge\_flag =' CE'$ **then**
      $f[L] = 2$;
    **else**
      $f[L] = 0$;
    **end if**
    $E[x\_axis(judge\_x[L]), y\_axis(judge\_y[L])] = f[L]$;
    $E[x\_axis(Edge\_lon), y\_axis(Edge\_lat)] = f[L]$;
  **end for**
**end for**
**return** $E$

The remainder of the paper is orgnized as follows. The

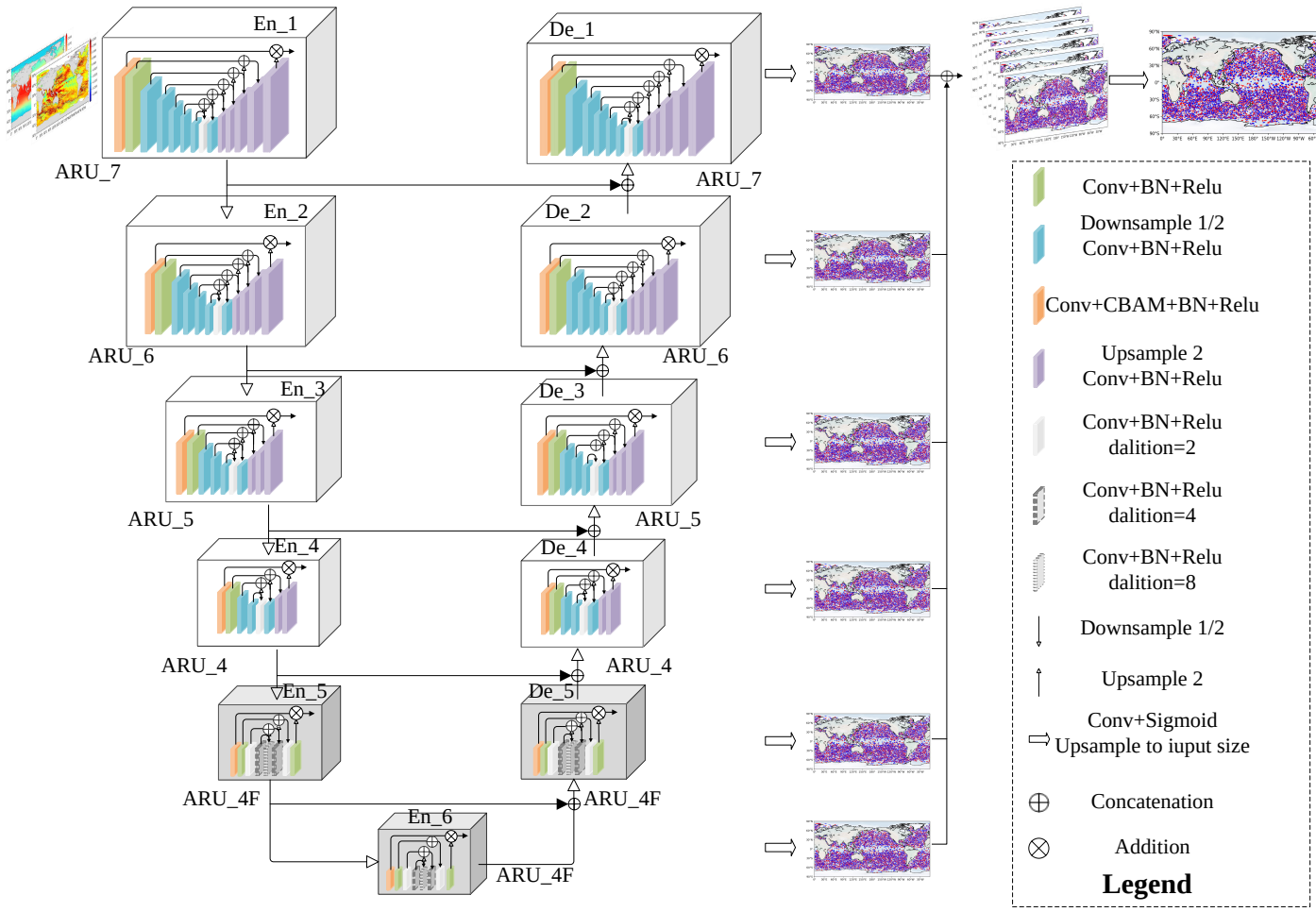


Fig. 2: Architecture of ARU$^2$-Net.

dataset and the preprocessing procedure are discussed in Section II. The architecture of the proposed ARU$^2$-Net is described in section III. Section IV, the results of eddy identification by different models are compared. Finally, Section VI summarizes this study and highlights the future research direction.

## II. Data Preparation

The global Sea Level Anomaly (SLA) data used in this study were estimated by Copernicus Maritime Services (CMEMS) using an optimal interpolation method, combining L3 along orbit measurements from different altimeter missions available. Global Sea Surface Temperature (SST) data comes from the National Oceanic and Atmospheric Administration (NOAA)[34], [35], [36]. The altimetric Mesoscale Eddy Trajectory Atlas product META3.2 DT allsat[37] was produced by SSALTO/DUACS and distributed by AVISO+ with support from CNES, in collaboration with IMEDEA. This atlas was downloaded the February 2023, and covers the period from January 1993 to February 2022[38]. Because of the huge amount of global eddy data, the data from January 2018 to December 2020 is selected as the training set for this experiment, and the data from 2017, and 2021 is used as the test set. The distribution of SLA and SST in the global ocean is shown in **Fig.1** .

The input data is preprocessed before Eddy detection, and this preprocessing mostly entails the following steps:

1) For the global SLA and SST data, the value of land area is set to a specific value as a mask, so as to exclude the influence of land on the detection of eddies.
2) To ensure the comparability of data, the SLA and SST data are normalized respectively to avoid the impact of too large a span.
3) The normalized SLA and SST data are combined into a set of two-channel feature data, which is used as input for eddy detection.

To meet the requirement of deep learning models for real eddy dataset, the construction process of real eddy data labels can be summarized in **Algorithm 1**. The edge position information of AE and CE eddies is used as input to determine the threshold value of the longitude and latitude of the eddies, and the eddy region is restricted in a rectangular box. Then, the labeling process is summarized in all test points on the edge and inside of the rectangular box are obtained with the spatial resolution of 0.25°. Since the eddy can be regarded as an irregular polygon, number of intersections plus one when the point to be tested is above the edge and the ray has an

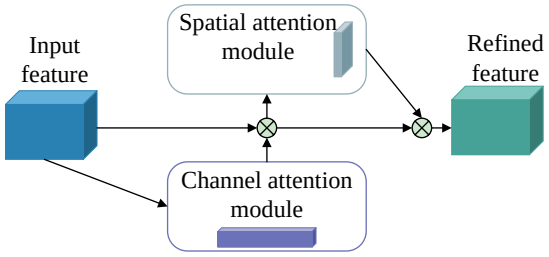


Fig. 3: Architecture of CBAM block

intersection with the edge. Next, if the number of intersections is odd, the points are inside the polygon, otherwise the points are outside the polygon. Finally, the contours and internal points of each eddy are labeled by category. Where "1" is labeled as anticycloncic eddy, "2" is labeled as cycloncic eddy, and "0" is labeled noneddy.

## III. Methodology

Despite the U$^2$-Net[39] model's outstanding performance in the area of picture semantic segmentation, the model's overall segmentation performance suffers when it encounters complex multiple category images because of segmentation errors. The CBAM can weight the feature map in both spatial and channel dimensions, which better accommodates targets of different scales and shapes, and also allows the model to better focus on regions of interest. So to solve the problem of model performance decrease when the U$^2$-Net model segmentation target is complex, this paper proposes a high-precision network model ARU$^2$-Net based on the U$^2$-Net network, which is a double-nested U-Net network structure, as shown in **Fig.2**. The ARU_L and ARU_4F modules are located in the inner layer, which has a U-shaped structure on its outer layer. Where the ARU_L module (L denotes the number of layers encoded) captures contextual eddy surface feature information at multiple scales by adding residual connectivity and cascade pooling[40]. Thus, the model enables efficient sense-field expansion with a small number of parameter additions. The CBAM integrated into the ARU_L module is an attention mechanism operating along two dimensions, channel and spatial, designed to direct the model's focus towards crucial features within the input data (SST and SLA), thereby enhancing detection efficiency. Different with the ARU_L module, the ARU_4F module replaces ordinary convolution with dilated convolution[41], which ensures that the feature map resolution remains the same to continue to obtain richer information about the eddy context features. The final prediction's results are fulfiled by fusing the features from one encoding stage and five decoding stages through a saliency map fusion module[42]. Additionally, the following section briefly explains all the substructures.

### A. ARU_L module

The SST and SLA data are fused into a two-channel feature data $\mathbf{x}^{(H\times W\times C_{in})}$ as input. This input undergoes transformation through common convolution layers, normalization

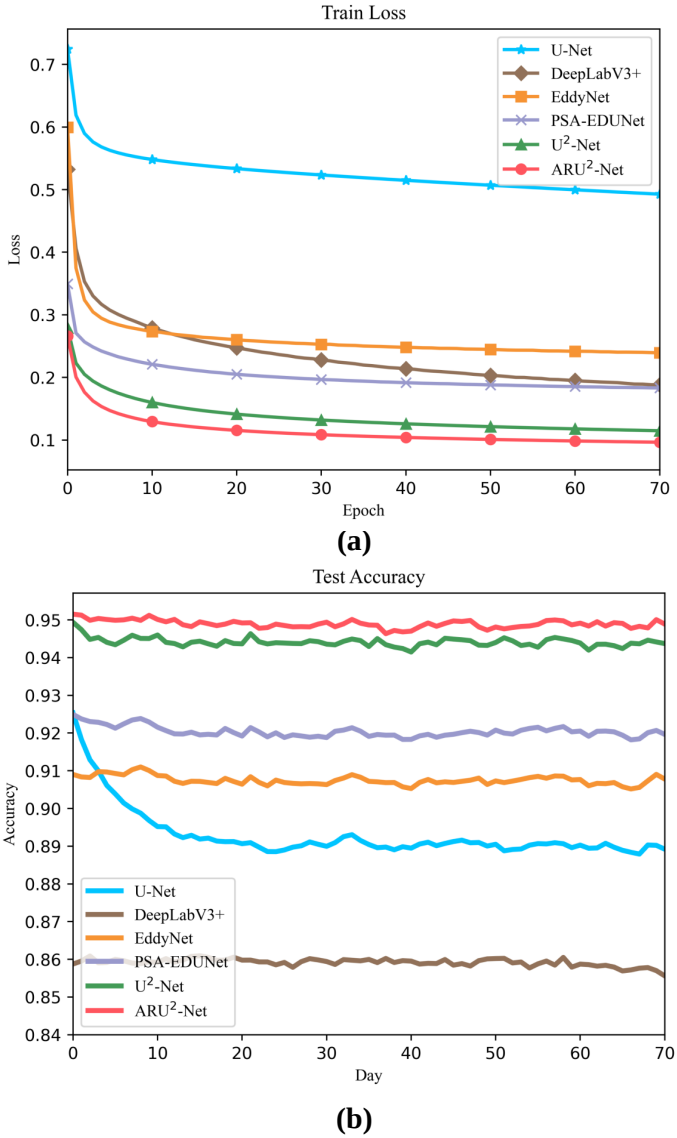


Fig. 4: Performance of the six models: (a) Changes in training losses; (b) Accuracy curves for the six models in the 2021 test set.

layers, and RELU activation function layers to create a feature map $\mathcal{F}(\mathbf{x})$ with a middle channel of $C_{mid}$. Subsequently, $\mathcal{F}(\mathbf{x})$ is inputed the CBAM to obtain the feature map $\mathcal{F}^{'}(\mathbf{x})$ focusing on important features. Then, a symmetric encoding and decoding structure with an L-layer height is used to extract multi-scale eddy context information $\mathcal{U}(F^{'}(\mathbf{x}))$. A larger L implies more pooling, expanding the range of the receptive field, and enriching local and global eddy surface feature information. Finally, the feature map $\mathcal{F}(\mathbf{x})$ and $\mathcal{U}(F^{'}(\mathbf{x}))$ are fused together through a residual connection, as shown in the following equation:

$$\mathbf{x} = \mathcal{F}(\mathbf{x}) + \mathcal{U}(F^{'}(\mathbf{x})) \tag{1}$$

### B. ARU_4F module

It is similar to the overall structure of ARU_L module. The

difference is that ARU_4F module no longer carries out up and downsampling. Following the downsampling to this module, the input feature map has a size of $(\frac{H}{8} \times \frac{H}{8} \times C_{mid})$ with a relatively low resolution. Upsampling and downsampling thereafter would result in an abundant loss of information, thereby affecting the model's training effectiveness. Therefore, ARU_4F module has four layers of encoding and decoding, and the first layer still uses ordinary convolution. The latter three layers use dilated convolution, the size of the convolution kernel is $3 \times 3$ and the dilate rate is multiplied by 2, 4 and 8, respectively.

### C. Saliency Map Fusion module

This module mainly plays the role of feature fusion. By collecting the output of the last layer of the encoder and each layer of the decoder, we can get a series of feature result maps. Then these feature maps are fed into a regular $3 \times 3$ convolutional layer with an output channel number set to 3. The feature map is then resized to the original feature map's size using the bilinear interpolation approach. Ultimately, the fused eddy probability outcome map was generated by splicing and processing the six feature maps using a $1 \times 1$ convolution layer and softmax activation function.

### D. Convolutional Block Attention Module (CBAM)

The effect is more effective than only focusing on the channel attention mechanism[43]. The CBAM is a kind of attention mechanism module for visual recognition tasks, which can remarkablely map important features and suppress irrelevant noise by weighting channel attention and spatial attention to feature maps, as shown in **Fig.3**. Specific details are described below.

Firstly, to improve the sensitivity to different features, the feature correspondences of the SST and SLA channels are adjusted using the channel attention mechanism. Specifically, the input feature maps $\mathbf{F} \in \mathbb{R}^{C \times H \times W}$ undergo max-pooling and average-pooling to aggregate global features across each channel, generating feature maps $\mathbf{F^c_{max}}$ and $\mathbf{F^c_{avg}}$. Following that, they get connected to a shared network with a single Multilayer Perceptron (MLP) hidden layer at its core. Next, the MLP features are weighted and the sigmoid function activates them to produce the channel attention feature $\mathbf{M_c} \in \mathbb{R}^{C \times 1 \times 1}$. Finally, the feature map needed by the spatial attention module can be produced by multiplying the channel attention feature $\mathbf{M_c}$ by the input feature map $F$. In short, the channel attention is computed as:

$$\mathbf{F_{avg}} = \mathrm{AvgPool}(\mathbf{F}) \tag{2}$$

$$\mathbf{F_{max}} = \mathrm{MaxPool}(\mathbf{F}) \tag{3}$$

$$\begin{aligned} \mathbf{M_c}(\mathbf{F}) &= \sigma(MLP(\mathbf{F_{avg}}) + MLP(\mathbf{F_{max}}) \\ &= \sigma\left(\mathbf{W_1}\left(\mathbf{W_0}\left(\mathbf{F^c_{avg}}\right)\right) + \mathbf{W_1}\left(\mathbf{W_0}\left(\mathbf{F^c_{max}}\right)\right)\right) \end{aligned} \tag{4}$$

$$\mathbf{F}^{'} = \mathbf{M_c}(\mathbf{F}) \otimes \mathbf{F} \tag{5}$$

Where $\sigma$ denotes the sigmoid function, $\mathbf{W_0} \in \mathbb{R}^{C/r \times C}$, and $\mathbf{W_1} \in \mathbb{R}^{C \times C/r}$. The MLP weights $\mathbf{W_0}$ and $\mathbf{W_1}$ of the two inputs are shared. $\otimes$ denotes element-wise multiplication.

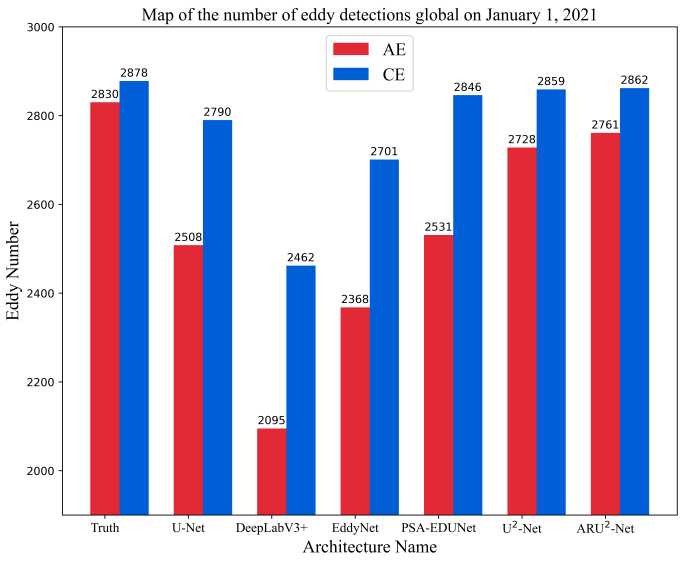


Fig. 5: Map of the number of eddy detection global on January 1, 2021. AE for anticyclone, CE for cyclone.

Secondly, spatial attention mechanisms are employed to adaptively adjust the feature responses at different spatial locations of the eddies, enhancing sensitivity to various positions. More specifically, the input feature map for the module is the output feature map from the channel attention module. The first phase creates a 2D feature map $\mathbf{F^s_{avg}} \in \mathbb{R}^{1 \times H \times W}$ and $\mathbf{F^s_{max}} \in \mathbb{R}^{1 \times H \times W}$ using a channel based on the average pooling and maximum pooling operation. We then create our 2D spatial attention map by concatenating and convolving those using a standard convolution layer. Finally, multiplying the spatial attention feature and the input feature map of this module to obtain the final generated feature map $F'$. In short, the spatial attention is computed as:

$$\begin{aligned} \mathbf{M_s}(\mathbf{F}) &= \sigma\left(f^{7\times7}([\mathbf{F_{avg}}; \mathbf{F_{max}}])\right) \\ &= \sigma\left(f^{7\times7}\left(\left[\mathbf{F^s_{avg}}; \mathbf{F^s_{max}}\right]\right)\right) \end{aligned} \tag{6}$$

$$\mathbf{F}^{''} = \mathbf{M_s}(\mathbf{F}') \otimes \mathbf{F}' \tag{7}$$

Where f $f^{7\times7}$ denotes a convolution operation using a $7 \times 7$ filter size.

### E. Dilated convolution

The resolution of the feature map decreases after multiple downsampling and further downsampling will result in the abundant loss of information about the eddy surface features. Hence, we use dilated convolutions to increase the receptive field while preserving the feature map resolution without adding extra parameters. By introducing the dilate rate based on standard convolution, dilated convolution expands the convolutional kernel's receptive field. If the standard convolution kernel is $K_0$ and the dilate rate is $R$, then the dilated convolution kernel is:

$$K_c = R \times (k_0 - 1) + 1 \tag{8}$$

Naturally, a larger receptive field leads to richer spatial information, which improves experimental results.

## IV. Experimental Results

### A. Experiment Setting

All the experiments in this study were carried out on the dataset described in the second part. These programs are run on Ncidia GeForce RTX 3090 GPU using the PyTorch framework. We used the cross entropy function as a loss function and set the learning rate to 1e-3 in the RMSprop optimizer. When the loss of the validation set stops decreasing for 10 consecutive epoches, the early stop strategy will be adopted to avoid overfitting and improve the model's generalization ability.

### B. Evaluation metrics

The paper suggested approach is compared to other eddy detection methods. This approach is assessed quantitatively using the mean pixel accuracy (MPA), accuracy, maximal $F\_\beta$, and weighted mean dice coefficient (WMDC) indicators. MPA refers to the average proportion of the number of correctly classified pixels per class; Accuracy is the proportion of correctly classified pixels to the total number of pixels in the entire image.

$$MPA = \frac{1}{N}\sum_i \frac{p_{ii}}{t_i} \tag{9}$$

$$Precision = \frac{TP}{(TP+FP)} \tag{10}$$

$$Recall = \frac{TP}{(TP+FN)} \tag{11}$$

$$Accuracy = \frac{(TP+TN)}{(TP+TN+FP+FN)} \tag{12}$$

The total number of class $i$ pixels in the real image is denoted by $t_i$, and the number of correctly predicted class $i$ pixels is represented by $P_{ii}$. $N$ is the total number of classes in the real image. $TP$ represents the number of matches between accurately predicted Class $i$ eddies and real labeled Class $i$ eddies; $FP$ represents the sum of the number of incorrectly predicted Class $i$ eddies and true non-class $i$ eddies; $FN$ represents the sum of the number of incorrectly predicted non-$i$ eddies and true $i$ eddies; $TN$ indicates the number of accurately predicted non-Class $i$ eddies matched with the real label non-class $i$ eddies. $F_\beta$ is used for comprehensive evaluation of accuracy and recall rate. Set $\beta$ to 1 for a perfect balance of accuracy and recall.

$$F_\beta = \frac{((1+\beta)^2 \times Precision \times Recall)}{(\beta^2 + Precision + Recall)} \tag{13}$$

Furthermore, we assess the degree of similarity between the real labels and the predicted results using the similarity coefficient.

$$DCN(N, G_0) = \frac{(2 \times (N \bigcap G))}{(|N| + |G_0|)} \tag{14}$$

$$DCA(A, G_1) = \frac{(2 \times (A \bigcap G))}{(|A| + |G_1|)} \tag{15}$$

$$DCC(C, G_2) = \frac{(2 \times (C \bigcap G))}{(|C| + |G_2|)} \tag{16}$$

$$WMDC = 0.82 \times DCN + 0.08 \times DCA \tag{17}$$

$$+0.1 \times DCC \tag{18}$$

Where, $G$ is the real data, $N$, $A$ and $C$ represent non-cyclonic, anticyclone and cyclone respectively while $G_0$, $G_1$ and $G_2$ are the corresponding real data. The weights in the $WMDC$ formula correspond to the normalized occurrence frequencies of non-cyclones, anticyclones and cyclones respectively.

TABLE I: Performance in 2021 as a test set on six models(%). Bold indicates the optimal value.

| | U-Net | DeepLabV3+ | EddyNet | PSA-EDUNet | $U^2$-Net | $ARU^2$-Net |
|---|---|---|---|---|---|---|
| Accuracy | 89.215 | 85.674 | 90.916 | 92.141 | 94.473 | **94.946** |
| MPA | 72.874 | 64.398 | 74.379 | 79.318 | 87.069 | **88.147** |
| $F_\beta$ | 75.060 | 66.680 | 78.176 | 81.807 | 87.774 | **88.785** |
| DCN | 93.586 | 91.645 | 94.633 | 95.314 | 96.684 | **96.954** |
| DCA | 61.694 | 51.513 | 65.786 | 71.826 | 81.467 | **82.924** |
| DCC | 69.900 | 56.882 | 74.110 | 78.123 | 85.173 | **86.477** |
| WMDC | 88.667 | 84.958 | 90.373 | 91.742 | 94.315 | **94.784** |

TABLE II: Performance in 2017 as a test set on six models(%). Bold indicates the optimal value.

| | U-Net | DeepLabV3+ | EddyNet | PSA-EDUNet | $U^2$-Net | $ARU^2$-Net |
|---|---|---|---|---|---|---|
| Accuracy | 89.201 | 85.830 | 90.894 | 92.122 | 94.835 | **94.951** |
| MPA | 77.478 | 64.409 | 74.532 | 79.342 | 88.713 | **88.268** |
| $F_\beta$ | 78.292 | 66.687 | 78.178 | 81.754 | 88.689 | **88.826** |
| DCN | 94.147 | 91.645 | 94.622 | 95.314 | 96.894 | **96.971** |
| DCA | 67.726 | 51.519 | 66.234 | 71.825 | 82.901 | **83.116** |
| DCC | 72.977 | 56.896 | 73.680 | 78.123 | 86.270 | **86.392** |
| WMDC | 89.938 | 84.960 | 90.256 | 91.716 | 94.712 | **94.805** |

### C. Result analysis

The proposed method is compared with the existing PSA-EDUNet, EddyNet, U-Net and DeepLabV3+[44] technologies. The effectiveness of the proposed model is fully verified. The results of the comparison experiments are shown in **Table I** and in **Fig.4**(b). In the global region, the accuracy is 94.926%, which is 2.875%, 4.010%, 5.731% and 9.096% higher than PSA-EDUNet, EddyNet, U-Net and DeepLabV3+ models, respectively. We visualize the number of anticyclones and cyclones detected by the different methods in **Fig.5**, where AE represents the number of anticyclones and CE represents the number of cyclonic. The $ARU^2$-Net model proposed in this paper finds more eddy currents than other methods. This shows that the $ARU^2$-Net model has the ability to detect smaller sized eddy currents with high performance. To better illustrate the feasibility of the model, we did further test validation by randomly selecting 2017 SST and SLA data as a new test set, as shown in **Table II**. The results remarkablely show that this set of test data is generally consistent with the

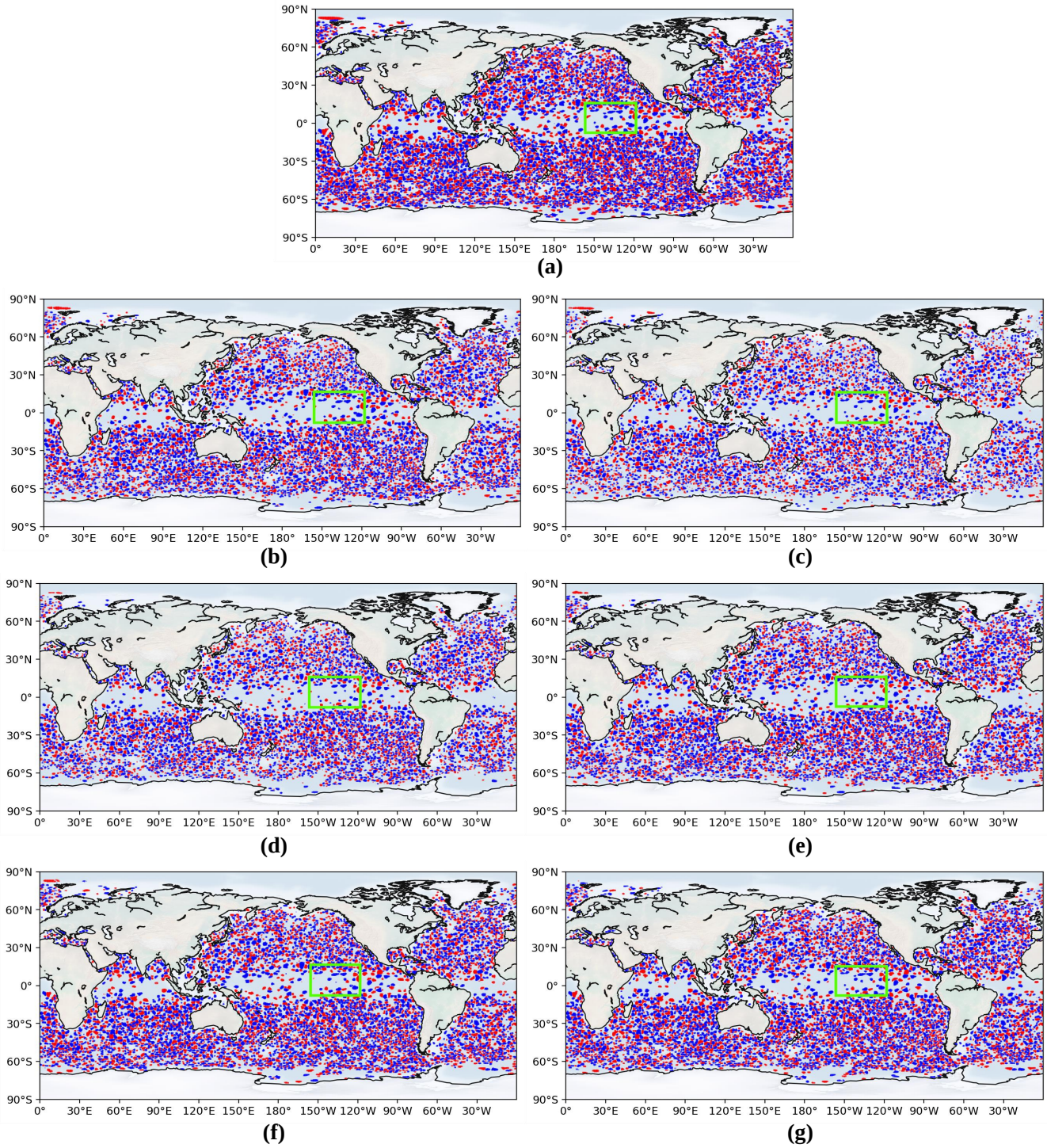


Fig. 6: Visualization of eddy detection results in 2021. (a)Ground truth; (b)U-Net; (c)DeepLabV3+; (d)EddyNet; (e)PSA-EDUNet; (f)U$^2$-Net; (g)ARU$^2$-Net. AEs(red); CEs(blue).

performance of the 2021 test results, further validating the robustness of the model. Meanwhile, we conducted ablation experiments on the CBAM and presented the results in **Table I** and **Table II**, proving the effectiveness of adding the CBAM module. Additionally, **Fig.4(a)** shows the learning curves of six model architectures on the same train set. Blue, brown, orange, purple, green and red represent the learning curves of U-Net, DeepLabV3+, EddyNet, PSA-EDUNet, U$^2$-Net and ARU$^2$-Net, respectively. From **Fig.4(a)**, it is evident that the loss decreases progressively with increasing epochs. The loss value obtained by the ARU$^2$-Net model on this curve is lower than that of the other five models. This observation indicates that the ARU$^2$-Net model has the strongest learning ability to capture the feature information of various eddy surfaces in

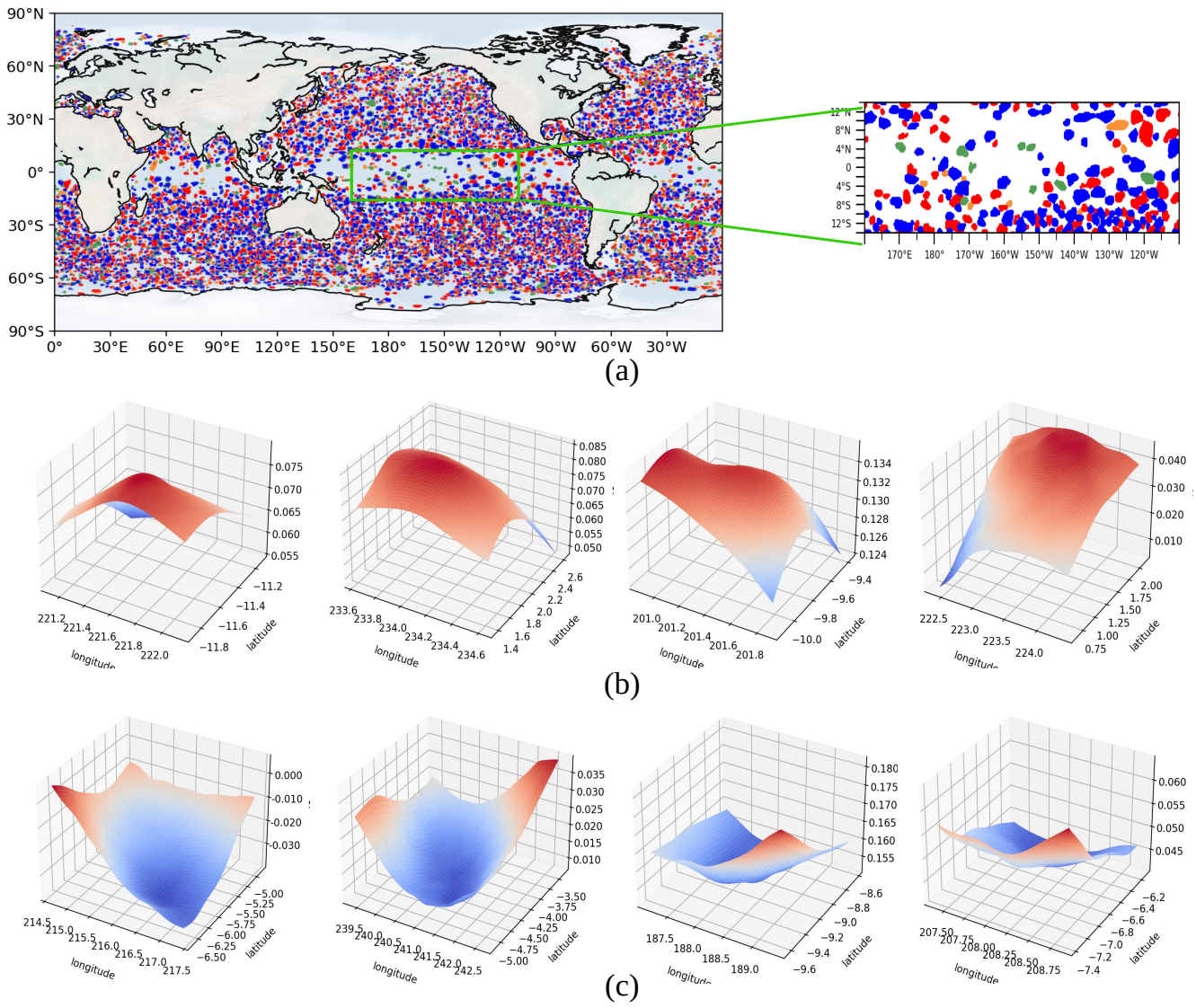


Fig. 7: Distribution of SLA data for newly detected eddies.(a)Detailed map of ARU$^2$-Net network architecture testing results; AEs(red); CEs(blue); new AEs(orange); new CEs(green) (b)New Eddy AE; (c)New Eddy CE.

the dataset. We can also see from **Fig.4(b)** that the ARU$^2$-Net model has little fluctuation in the test results for each day on the test set, and the accuracy is much higher than that of the other models, which further indicates that the model has a strong generalization ability. These results jointly prove the superiority of the ARU$^2$-Net model proposed in this paper.

We visualized the detection results for the first day of 2021. **Fig.6** plots the detection results for all models in the global region. Compared to the true value, it is easy to see that the U-Net and DeepLabV3+ models miss many eddies and falsely detect some additional eddies. Although the detection results are similar on the U$^2$-Net model and the ARU$^2$-Net model, by looking closely at the visual details, you will find that the CBAM added to the basic U$^2$-Net architecture is more accurate than the U$^2$-Net model detection, especially in the detection of small-scale eddies.

Comparison of the real eddies revealed that the ARU$^2$-Net model can detect eddies that had not been detected in real dataset. To determine the correctness of the newly detected eddies, we select a small region with more newly detected eddies from the global region, as shown in **Fig.7(a)**. Then, we match the new eddies with the SLA data, and find that the height of the eddy-centered sea surface of AE is higher than the height of the surrounding sea surfaces. Similarly, the height of the eddy-centered sea surfaces of CE is lower than the height of the surrounding sea surfaces. These observations fully align with the surface characteristics of eddies. The SLA data distribution for newly discovered eddies is shown in **Fig.7(b), (c)**.

## V. Conclusions

This study presents a novel eddy detection model called ARU$^2$-Net, which has been developed specifically for global scale. ARU$^2$-Net is distinct from current eddy detection models as it utilizes a double-nested U-Net structure. Each submodule in this architecture incorporates CBAM to enhance the

features, enabling the model to prioritize the feature information of the eddy. In order to assess the detection capabilities of ARU$^2$-Net, we performed experiment using a publically accessible dataset and compared its performance to that of various eddy detection models. When comparing quantitative and qualitative data, it is evident that the suggested architecture consistently outperforms current deep learning algorithms in terms of accuracy and similarity for eddy detection.

Nevertheless, the ARU$^2$-Net model does possess specific constraints. Initially, while studying the attributes of eddy, the model could unintentionally acquire inaccurate feature data because of the indistinct features found in the periphery of the eddy. This might lead to less than ideal edge detection outcomes. Furthermore, due to the double-nested U-Net structure of the ARU$^2$-Net model, the detection process is relatively time-consuming.

Future research will prioritize the development of streamlined methods for extracting features, followed by the implementation of suitable approaches to enhance the detection effectiveness of the model.

## VI. Acknowledgement

This research was financially supported by the National Natural Science Foundation of China (No.42276203, 42030406), the Laoshan Laboratory (No.LSKJ202204302), and the Natural Science Foundation of Shandong Province (No.ZR2021MD001). The dataset and Python code will be available at GitHub to encourage competing methods.

## References

[1] Allan R Robinson. Overview and summary of eddy science. In *Eddies in marine science*, pages 3–15. Springer, 1983.

[2] Dudley B Chelton, Peter Gaube, Michael G Schlax, Jeffrey J Early, and Roger M Samelson. The influence of nonlinear mesoscale eddies on near-surface oceanic chlorophyll. *Science*, 334(6054):328–332, 2011.

[3] G Chen, J Yang, FL Tian, SG Chen, CF Zhao, JW Tang, YJ Liu, YN Wang, ZH Yuan, Q He, et al. Remote sensing of oceanic eddies: Progresses and challenges. *Natl. Remote Sens. Bull*, 25:302–322, 2021.

[4] Minghan Fu, Changming Dong, Jihai Dong, and Wenjin Sun. Analysis of mesoscale eddy merging in the subtropical northwest pacific using satellite remote sensing data. *Remote Sensing*, 15(17):4307, 2023.

[5] Yingjie Liu and Xiaofeng Li. Impact of surface and subsurface-intensified eddies on sea surface temperature and chlorophyll a in the northern indian ocean utilizing deep learning. *Ocean Science*, 19(6):1579–1593, 2023.

[6] Yign Noh, Bo Young Yim, Sung Hyup You, Jong Hwan Yoon, and Bo Qiu. Seasonal variation of eddy kinetic energy of the north pacific subtropical countercurrent simulated by an eddy-resolving ogcm. *Geophysical Research Letters*, 34(7), 2007.

[7] Changming Dong, James C McWilliams, Yu Liu, and Dake Chen. Global heat and salt transports by eddy movement. *Nature communications*, 5(1):3294, 2014.

[8] David Gawith Nichol. Autonomous extraction of an eddy-like structure from infrared images of the ocean. *IEEE transactions on geoscience and remote sensing*, (1):28–34, 1987.

[9] Sarah H Peckinpaugh and Ronald J Holyer. Circle detection for extracting eddy size and position from satellite imagery of the ocean. *IEEE Transactions on Geoscience and remote sensing*, 32(2):267–273, 1994.

[10] Kaihuai Qin, Haiying Wang, and Jitao Zheng. A unified approach based on hough transform for quick detection of circles and rectangles. *Journal of Image and Graphics*, 15(1):109–115, 2010.

[11] G-R Ji, Xia Chen, Y-Z Huo, and T-J Jia. An automatic detecting method of the marine meso-scale eddy in remote sensing image. *Oceanologia Et Limnologia Sinica*, 33(2):139–144, 2002.

[12] Jordi Isern-Fontanet, Emilio García-Ladona, and Jordi Font. Identification of marine eddies from altimetric maps. *Journal of Atmospheric and Oceanic Technology*, 20(5):772–778, 2003.

[13] Jiya Albert and Prasad K Bhaskaran. Optimal grid resolution for the detection lead time of cyclogenesis in the north indian ocean. *Journal of Atmospheric and Solar-Terrestrial Physics*, 204:105289, 2020.

[14] Danchen Yan, Tianyu Zhang, Shaomei Yu, Yun Li, YanQiang Wang, and Bin Wang. An improved winding-angle method to more accurately identify mesoscale eddies. *Journal of Atmospheric and Oceanic Technology*, 35(2):229–245, 2018.

[15] Changming Dong, Xingliang Jiang, Guangjun Xu, J Ji, X Lin, W Sun, and S Wang. Automated eddy detection using geometric approach, eddy datasets and their application. *Adv. Mar. Sci*, 35(04):439–453, 2017.

[16] Y Hao. *Research on ocean mesoscale eddy detection algorithm based on convolution neural network*. PhD thesis, Shandong University of Science and Technology Shandong, China, 2017.

[17] Yi-Ran Wang and Xiao-Ming Li. Arctic sea ice cover data from spaceborne synthetic aperture radar by deep learning. *Earth System Science Data*, 13(6):2723–2742, 2021.

[18] Yunhe Wang, Xiaojun Yuan, Yibin Ren, Mitchell Bushuk, Qi Shu, Cuihua Li, and Xiaofeng Li. Subseasonal prediction of regional antarctic sea ice by a deep learning model. *Geophysical Research Letters*, 50(17):e2023GL104347, 2023.

[19] Xiaofeng Li, Bin Liu, Gang Zheng, Yibin Ren, Shuangshang Zhang, Yingjie Liu, Le Gao, Yuhai Liu, Bin Zhang, and Fan Wang. Deep-learning-based information mining from ocean remote-sensing imagery. *National Science Review*, 7(10):1584–1605, 2020.

[20] Haoyu Wang and Xiaofeng Li. Deepblue: Advanced convolutional neural network applications for ocean remote sensing. *IEEE Geoscience and Remote Sensing Magazine*, 2023.

[21] Ge Chen, Baoxiang Huang, Xiaoyan Chen, Linyao Ge, Milena Radenkovic, and Ying Ma. Deep blue ai: A new bridge from data to knowledge for the ocean science. *Deep Sea Research Part I: Oceanographic Research*, 190:103886, 2022.

[22] Olaf Ronneberger, Philipp Fischer, and Thomas Brox. U-net: Convolutional networks for biomedical image segmentation. In *Medical Image Computing and Computer-Assisted Intervention–MICCAI 2015: 18th International Conference, Munich, Germany, October 5-9, 2015, Proceedings, Part III 18*, pages 234–241. Springer, 2015.

[23] Redouane Lguensat, Miao Sun, Ronan Fablet, Pierre Tandeo, Evan Mason, and Ge Chen. Eddynet: A deep neural network for pixel-wise classification of oceanic eddies. In *IGARSS 2018-2018 IEEE International Geoscience and Remote Sensing Symposium*, pages 1764–1767. IEEE, 2018.

[24] Redouane Lguensat, Saifeddine Rjiba, Evan Mason, Ronan Fablet, and J Sommer. Convolutional neural networks for the segmentation of oceanic eddies from altimetric maps. *Remote Sens*, 1(1):1–16, 2018.

[25] Kaiming He, Xiangyu Zhang, Shaoqing Ren, and Jian Sun. Deep residual learning for image recognition. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 770–778, 2016.

[26] F Milletari, N Navab, S Ahmadi Ahmadi, and V Net. Fully convolutional neural networks for volumetric medical image segmentation. In *Proceedings of the 2016 Fourth International Conference on 3D Vision (3DV)*, pages 565–571.

[27] Guangjun Xu, Cheng Cheng, Wenxian Yang, Wenhong Xie, Lingmei Kong, Renlong Hang, Furong Ma, Changming Dong, and Jingsong Yang. Oceanic eddy identification using an ai scheme. *Remote Sensing*, 11(11):1349, 2019.

[28] Guangjun Xu, Wenhong Xie, Changming Dong, and Xiaoqian Gao. Application of three deep learning schemes into oceanic eddy detection. *Frontiers in Marine Science*, 8:672334, 2021.

[29] Nan Zhao, Baoxiang Huang, Jie Yang, Milena Radenkovic, and Ge Chen. Oceanic eddy identification using pyramid split attention u-net with remote sensing imagery. *IEEE Geoscience and Remote Sensing Letters*, 20:1–5, 2023.

[30] Baoxiang Huang, Linyao Ge, Xiaoyan Chen, and Ge Chen. Vertical structure-based classification of oceanic eddy using 3-d convolutional neural network. *IEEE Transactions on Geoscience and Remote Sensing*, 60:1–14, 2021.

[31] Alexis Chaigneau, Marie Le Texier, Gérard Eldin, Carmen Grados, and Oscar Pizarro. Vertical structure of mesoscale eddies in the eastern south pacific ocean: A composite analysis from altimetry and argo profiling floats. *Journal of Geophysical Research: Oceans*, 116(C11), 2011.

[32] Ge Chen, Xiaoyan Chen, and Baoxiang Huang. Independent eddy identification with profiling argo as calibrated by altimetry. *Journal of Geophysical Research: Oceans*, 126(1):e2020JC016729, 2021.

[33] Sanghyun Woo, Jongchan Park, Joon-Young Lee, and In So Kweon. Cbam: Convolutional block attention module. In *Proceedings of the European conference on computer vision (ECCV)*, pages 3–19, 2018.
[34] Arthur D Voorhis, Elizabeth H Schroeder, and Ants Leetmaa. The influence of deep mesoscale eddies on sea surface temperature in the north atlantic subtropical convergence. *Journal of Physical Oceanography*, 6(6):953–961, 1976.
[35] Changming Dong, Francesco Nencioli, Yu Liu, and James C McWilliams. An automated approach to detect oceanic eddies from satellite remotely sensed sea surface temperature data. *IEEE Geoscience and Remote Sensing Letters*, 8(6):1055–1059, 2011.
[36] Gang Zheng, Xiaofeng Li, Rong-Hua Zhang, and Bin Liu. Purely satellite data–driven deep learning forecast of complicated tropical instability waves. *Science advances*, 6(29):eaba1482, 2020.
[37] C Pegliasco, C Busché, and Y Faugère. Mesoscale eddy trajectory atlas meta3. 2 delayed-time all satellites: version meta3. 2 dt allsat. *AVISO https://doi. org/10.24400/527896/A01-2022.005*, 220209, 2022.
[38] Evan Mason, Ananda Pascual, and James C McWilliams. A new sea surface height–based code for oceanic mesoscale eddy tracking. *Journal of Atmospheric and Oceanic Technology*, 31(5):1181–1188, 2014.
[39] Xuebin Qin, Zichen Zhang, Chenyang Huang, Masood Dehghan, Osmar R Zaiane, and Martin Jagersand. U2-net: Going deeper with nested u-structure for salient object detection. *Pattern recognition*, 106:107404, 2020.
[40] Tianyu Zhang, Ying Yang, Mohammed Shokr, Chunlei Mi, Xiao-Ming Li, Xiao Cheng, and Fengming Hui. Deep learning based sea ice classification with gaofen-3 fully polarimetric sar data. *Remote Sensing*, 13(8):1452, 2021.
[41] Panqu Wang, Pengfei Chen, Ye Yuan, Ding Liu, Zehua Huang, Xiaodi Hou, and Garrison Cottrell. Understanding convolution for semantic segmentation. In *2018 IEEE winter conference on applications of computer vision (WACV)*, pages 1451–1460. Ieee, 2018.
[42] Qiong Yan, Li Xu, Jianping Shi, and Jiaya Jia. Hierarchical saliency detection. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 1155–1162, 2013.
[43] Jie Hu, Li Shen, and Gang Sun. Squeeze-and-excitation networks. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 7132–7141, 2018.
[44] Liang-Chieh Chen, Yukun Zhu, George Papandreou, Florian Schroff, and Hartwig Adam. Encoder-decoder with atrous separable convolution for semantic image segmentation. In *Proceedings of the European conference on computer vision (ECCV)*, pages 801–818, 2018.

**Junmin Geng** received a B.S.degree in Software Engineering from Zhongyuan University of Technology in 2022. she is currently pursuing a M.S. degree in Computer Technology at Qingdao University. Her current research interests include Artificial Intelligence Oceanography and Deep Learning.

**Gao He** received a B.S. degree in Computer Science and Technology from Qilu University of Technology in 2019, and a M.S. degree in Computer Technology from the College of Computer Science and Technology at Qingdao University in 2023. He is currently pursuing a Ph.D. in Software Engineering at Qingdao University. His current research interests include Artificial Inte lligence Oceanography and Deep Learning.

**Baoxiang Huang** (Member, IEEE) received the B.S. degree in traffic engineering from the Shandong University of Technology, China, in 2002, the M.S. degree in mechatronic engineering from Shandong University. China, in 2005, and the Ph.D. degree in computer engineering from Ocean University of China, in 2011, China. She was an academic visitor of Nottingham University. Currently, she is a Professor at the College of Computer Science and Technology, Qingdao University, China. Her research interests include remote sensing image processing and analysis, big data oceanography, and artificial intelligence.

**Milena Radenkovic** received the Dipl.−Ing. degree in electric and electronic engineering from the University of Nis, Nis, Serbia, in 1998, and the Ph.D. degree in computer science from the University of Nottingham, Nottingham, U.K., in 2002. She has authored more than 80 papers in premium conferences and journal venues. Her research interests include intelligent mobile and disconnection-tolerant networking, complex temporal graphs, self-organized security, distributed predictive analytics with applications to autonomous vehicles, mobile social networks, smart manufacturing, and predictive telemetry. Dr. Radenkovic was a recipient of multiple EPSRC and EU grants for her research. She has organized and chaired multiple ACM and IEEE conferences and served on many program committees. She is an Editor for premium journals such as the Ad Hoc Networks (Elsevier), the IEEE Transactions on Parallel and Distributed Computing, and ACM Multimedia.

**Ge Chen** received the B.S. degree in marine physics, the M.S. degree in satellite oceanography, and the Ph.D. degree in physical oceanography from the Ocean University of China (OUC), Qingdao, China, in 1988, 1990, and 1993, respectively. He worked as a Post-Doctoral Fellow with The French Research Institute for the Exploitation of the Sea (IFREMER), Brest, France, from 1994 to 1996. Since 1997, he has been a Professor of satellite oceanography and meteorology with OUC. He is the Deputy Dean with the Institute for Advanced Marine Sciences, OUC, and the Chief Scientist for Ocean Science Satellite Missions with the National Laboratory of Ocean Science and Technology, Qingdao. He has authored or coauthored more than 110 peer reviewed scientific articles published in internationally recognized journals. His research interests include satellite remote sensing of the ocean and big data oceanography. Dr. Chen served as the Executive Secretary of the International Pan Ocean Remote Sensing Conference (PORSEC) Association from 1998 to 2002. He was a recipient of the National Science Fund for Outstanding Young Scientists awarded by the Natural Science Foundation of China in 2001. He became the Chair Professor of Cheung Kong Scholars Program nominated by the Chinese Ministry of Education.