Global ocean mesoscale vortex recognition based on DeeplabV3plus model

Taotao Wang1, Hongchang He1, Donglin Fan1, Bolin Fu1, Shundan Dong 1
1 College of Geomatics and Geoinformation Guilin University of Technology, Guilin, Guangxi, China
*Corresponding author’s e-mail: hhe@glut.edu.cn

Abstract. Ocean mesoscale vortices play a very important role in global energy and material transport, and to a large extent affect the distribution of nutrients and phytoplankton. Traditional mesoscale vortex extraction methods have the problems of dependence on threshold and sensitivity to noise. In recent years, the machine learning methods that have emerged in recent years have also made the generalization ability of the model poor due to the limited coverage of the training data set, resulting in most methods only suitable for extraction in specific sea areas. This paper uses daily global sea level data from 2008 to 2017, combined with the py-eddy algorithm, to complete the construction of a label data set covering the entire sea. At the same time, based on the DeeplabV3plus model, by adjusting the loss function, a recognition model conforming to the mesoscale vortex characteristics is realized. Experimental results show that the verification accuracy rate of the model reaches 80.5%, and the Kappa coefficient is 0.758. Compared with the previous extraction method, the accuracy of this model is increased by 13.8%, and the Kappa coefficient is increased by 0.255. Experiments show that the mesoscale vortex recognition method proposed in this paper has good recognition accuracy for different sea areas and different scale vortices.

1. Introduction
Ocean mesoscale vortex refers to the rotating vortex in the ocean with a diameter between tens to hundreds of kilometers and a duration of several days to several months [1]. The formation of mesoscale vortices is accompanied by vertical and horizontal movement, which affects the global nutrient transmission, and is accompanied by local upwelling currents. For example, the cyclonic cold vortex formed in the northern hemisphere will transport nutrients on the seafloor upward, making the vortex area often good fishing grounds can be formed, thus affecting the development of the marine economy. The research on the extraction of mesoscale vortices is of great significance to the research of marine ecosystems, fishery resource assessment, fishing grounds forecasting and military research [2].

At present, the extraction methods of mesoscale vortices are mainly divided into three categories based on physical parameters, geometric parameters and machine learning. Common identification methods are: Sea Surface Height (SSH) method, Okubo-Weiss (OW) method, Vector Geometry (VG) method and Hybrid Detection (HD) method. The geometry-based SSH method uses the closed contour of the outermost circle as the boundary of the vortex, and takes the SSH local extremum within the vortex boundary as the center of the vortex. Although the geometric-based SSH method increases the automaticity of the algorithm, the calculation amount. The physics-based OW method uses the W value as a physical parameter, and the W value is calculated by SSH. Use -0.2σw (σw is the standard deviation of the entire W field) as the threshold to determine whether the detection area is a vortex. The
OW method defines the criterion from the physical properties of the mesoscale vortex, which can better reveal the physical nature of the mesoscale vortex, but it has obvious defects [3]. First, the Coriolis force tends to zero near the equator. Therefore, the OW parameters calculated by SLA, Coriolis force and gravitational acceleration are incorrect. Second, the W field produces many noise points, which are often misjudged as vortices. Third, the results obtained by different thresholds fluctuate greatly. The vortex center obtained by the VG method is the grid point with the smallest vortex velocity, and the vortex edge is the closed streamline corresponding to the maximum average rotation speed. This method has an advantage over the OW method in terms of accuracy, but its detection process is more complicated [5]. HD method is a combination of OW method and SSH method. First, determine the vortex center with the local extreme value of SSH and the "vortex core" (W<−0.2gw), and then use the closed SSH contour of the innermost circle containing the "vortex core" as the vortex edge, and finally determine the vortex boundary. When the above methods are used for vortex extraction, the problems of large computational complexity and low mesoscale vortex recognition rate will occur, which makes it impossible to perform rapid extraction of mesoscale vortices in large areas. With the rapid development of machine learning, many new mesoscale vortex extraction methods have emerged. These methods can not only improve the accuracy and speed of mesoscale vortex extraction, but also integrate various features of mesoscale vortices [6]. Eddynet, a mesoscale vortex extraction method based on U-net, applies deep learning to the extraction of mesoscale vortices for the first time. However, because the data set used in this method is made based on the data of the South Atlantic region, it is very complex and changeable. The global ocean is not universal. Compared with traditional mesoscale vortex extraction methods, deep learning can create a multi-layer convolutional neural network model, which can identify and extract mesoscale vortices based on remote sensing data, which is more conducive to improving the accuracy and speed of mesoscale vortex extraction.

The quality and quantity of the produced data set greatly affect the quality of the model. According to a large number of experiments and work, increasing the number of data sets can effectively avoid over-fitting and improve the performance of the model [7]. In this study, the global sea level anomaly data provided by Copernicus was used as the research data, the py-eddy algorithm was used to make deep learning label data, the convolutional neural network DeeplabV3plus was used as the method and the dice loss function was used to construct the ocean mesoscale vortex extraction model. The experiment mainly evaluates the accuracy of the Deeplabv3plus model and the Eddynet model in extracting mesoscale vortices from the ocean, and explores a high-precision and suitable method for mesoscale vortex extraction in global oceans. This experiment provides an important data basis for the development of marine resources.

2. Data and methods

2.1. Data

2.1.1. Data source. In this study, the daily global sea level anomaly data fused by multiple satellite altimeters obtained from the Copernicus Marine Environmental Monitoring Service (CMEMS, http://marine.copernicus.eu) is shown in Figure 1. The product is processed by the SL-TAC multi-task altimeter data processing system, and the data fusion comes from Jason-3, Sentinel-3A, HY-2A, Saral/Altika, Cryosat-2, Jason-2, Jason-1, T/P, ENVISAT, GFO, ERS1/2 data, multiple satellite altimeter fusion data significantly improve the accuracy and spatial resolution of sea surface height changes, the spatial resolution of the fusion data is 1/4°×1/4°, and the time resolution is 1d, the range is from 2008 to 2017, from January 1st to December 31st of each year, a total of 3653d. In this study, three variables, sea level anomaly (SLA), absolute geostrophic velocity: zonal component (Ugos), and absolute geostrophic velocity: meridional component (Vgos) were selected as the parameters when making labels. Since the sea level anomaly (SLA, sea level anomaly) is the difference between the instantaneous sea level and the multi-year average sea level, the value is small, so it does not need to be normalized, just change the land value to a fixed value.
SLA = SSH – MSS – TE – IB - HF  

The sea level anomaly SLA is the difference between sea level SSH and mean sea level MSS. MSS is the average value of sea level for many years, and the abnormal value of sea level can be calculated by formula (1). Among them, the calculation of SSH also considers high-frequency oscillation correction, tide correction and atmospheric inverse correction. MSS is the average value of SSH in period N, and TE is the tide correction value, including correction values such as extreme tide, ocean tide and solid tide. HF It is the high-frequency oscillation of the sea surface caused by changes in air pressure and wind.

2.1.2. Label Data Set. This research uses the py-eddy algorithm [8] to produce label data, which is based on closed contours for identification. First, the SLA data is subjected to spatial high-pass filtering with a radius of 5°, and then the contour line is downward (upward) at 2mm intervals. Perform a search, where each adjacent closed contour needs to meet the following relevant criteria:

- The ratio of the difference between the closed contour and the area of the fitted circle to the area of the fitted circle ≤ 55%;
- The number of pixels included I satisfy 8 ≤ I ≤ 1000;
- Only pixels whose SLA value of the cyclone (anticyclone) vortex is lower (higher) than the current SLA interval value are included.
- For a cyclone (anticyclone), it contains at most one SLA minimum (maximum).
- The vortex amplitude (A) satisfies 1 ≤ A ≤ 150 cm, and the amplitude A is defined as the absolute value of the difference between the SLA extreme value of the closed contour and the SLA value of the closed contour.

When the closed contour meets the above criteria, it is regarded as a mesoscale vortex. The vortex radius is the same as the area of the closed contour. The centroid of the closed contour is the center of the mesoscale vortex. This article divides the label data into three categories, namely (land "0", anti-air vortex "1", air vortex "2"). Because the shape and size of the vortices vary greatly, they are usually asymmetrical, and the translation and rotation speeds vary greatly [9]. Therefore, this study chose not to perform data enhancement processing, to better maintain the eddy current characteristics of the original data.

Figure 1. Global sea level anomaly distribution map
2.2. Method

2.2.1. Network structure. In this study, considering the characteristics of mesoscale vortices, a deep convolutional neural network model suitable for the extraction of mesoscale vortices in different sea areas around the world is constructed. After a lot of attempts and parameter adjustments, it was finally decided to use the DeeeplabV3plus network to build the model. DeeeplabV3plus uses ImageNet pre-trained ResNet as its main feature extractor network. A new residual block for multi-scale feature learning is proposed. The last ResNet block uses hole convolution instead of regular convolution. Moreover, each convolution (within this new block) uses a different dilation rate to capture multi-scale context. In addition, at the top of this new block, Atrous Spatial Pyramid Pooling (ASPP) is used. ASPP uses dilated convolutions with different rates to try to classify regions of any scale.

The model is constructed in the keras framework, and the convolutional layer uses a hole convolution algorithm to expand the receptive field to obtain more contextual information. The pooling layer uses porous pyramid pooling as shown in (Figure 3), using a 1×1 convolution and three 3×3 sampling rates in the feature top map, rates={6, 12, 18}. Hole convolution, the hole convolution of different sampling rates can effectively capture multi-scale information, but as the sampling rate increases, the effective weight of the filter will gradually become smaller. When the sampling rate is close to the size of the feature map, the 3×3 filter does not capture the context of the full image, but degenerates into a simple 1×1 filter. At this time, only the center of the filter functions. To overcome this problem, consider applying global averaging to the final feature map of the model, passing the result through a 1×1 convolution, and then obtaining the required spatial dimension through linear upsampling. The encoded features of the input data after porous pyramid pooling and 1×1 convolution are subjected to 4 times upsampling operation, and then the features of the same resolution obtained from the backbone network are spliced, and the final result is obtained through convolution and upsampling result.
2.2.2. **Loss function.** In this study, the dice loss function is used for model training, which greatly accelerates the convergence speed of the model. In the multi-classification problem of deep learning, the classification cross-entropy loss function is usually selected for training, but for segmentation problems, it is better to use overlap-based metric segmentation. In semantic segmentation, the dice coefficient is a loss function commonly used in deep convolutional neural networks. According to the predicted area \( P \) and the real area \( G \), use \(|P|\) and \(|G|\) to denote the sum of pixels in each area respectively. The dice coefficient is the ratio of twice the area of the intersection of the two areas to the sum of the areas of the two areas, as shown in the following formula:

\[
\text{DiceCoeff}(P, G) = \frac{2|P \cap G|}{|P| + |G|} \tag{2}
\]

When the dice coefficient is 1, an ideal segmentation effect can be obtained, and when the dice coefficient is 0, it means a completely wrong segmentation. It can be seen from the above formula that the dice coefficient is the harmonic mean value of the accuracy and recall index. This implementation uses one-hot encoding. The loss function of the convolutional neural network model in this study uses a soft and distinguishable version of the dice coefficient, which takes the output of the softmax layer into account without binarization:

\[
\text{softDiceCoeff}(P, G) = \frac{2 \sum_i p_i \times g_i}{\sum_i p_i + \sum_i g_i} \tag{3}
\]

In the formula (3), \( p_i \) is the probability obtained by the softmax layer \((0 < p_i < 1)\). When \( g_i \) is 1 or 0, it indicates the correct class. Recent studies use another version of the soft dice loss function \([12]\). The study tried to use the average of three pairs of soft dice coefficients for each class to improve the performance of the network. Hope to get the minimized loss function:

\[
\text{Dice loss} = 1 - \text{softMeanDiceCoeff} \tag{4}
\]

3. **Experiment and result analysis**

3.1. **Construction of ocean mesoscale vortex recognition model based on DeeplabV3plus**

According to the spatial heterogeneity of the global mesoscale vortex, the larger the picture, the more relevant information about its texture and surroundings. A lot of experiments have shown that the best training effect is when the input data size is \(720 \times 1440\). When the size becomes larger, the calculation amount will also become larger. Therefore, training can only be performed when the batch-size is set to 1. Use 90% of the data set for training, and use the remaining 10% as validation data. Model training
has high hardware requirements. The experiment chooses to train on a workstation with i9 processor and GV100 graphics card. The model uses dice loss function and Adam optimizer to speed up the convergence speed of the model, thereby improving the classification accuracy of the model and enhancing the generalization ability of the model [11].

By training the model, the loss/precision value curve under different epoch values is obtained (Figure 4). It can be observed from the figure that when the epoch value is 10, the training accuracy and verification accuracy of the model have reached more than 0.9, and the training loss value and verification loss value have fallen below 0.1. Studies have shown that when the epoch value is less than 10, as the number of iterations (epoch) increases, the training accuracy value and the verification accuracy value of the deeplabV3plus model increase rapidly and gradually stabilize, and the training loss value and the verification loss value rapidly decrease and gradually stabilize. It tends to be stable, and finally forms a more ideal feature suitable for pattern classification.

Figure 4. Loss/precision value curve under different epochs

### 3.2. Accuracy evaluation

In order to evaluate the extraction effect of the DeeplabV3+ model on mesoscale vortices, accuracy and recall are used as evaluation indicators. The definitions of accuracy and recall are as follows:

\[
\text{Precision} = \frac{TP}{TP + FP} \quad \text{(5)}
\]

\[
\text{Recall} = \frac{TP}{TP + FN} \quad \text{(6)}
\]

Among them, TP is the area of the area where the vortex position is correctly extracted, FP is the area of the non-vortex sea area that is incorrectly identified as a vortex, and FN is the area where the vortex area is identified as other sea areas. The data on the 13th of each month from March to December 2018 is selected as the prediction data. After accuracy verification, the accuracy, recall and F1 score of each prediction result relative to the label data are relatively stable (Figure 5). Therefore, this research model can be well applied to extract mesoscale vortices from different data.
Figure 5. Trends in accuracy of prediction results based on different data

The label making method used in this model is the same as that used in the Eddynet model, so the extraction results of the two models are compared and analyzed. Select the 2018 global sea level data as the forecast data, and use the forecast results of the two models to verify the accuracy with the label data. In order to avoid the contingency of the results, this article will extract one piece of the 2018 data every month for accuracy verification, and then take average data as the result. It can be clearly seen from the table that the DeeplabV3plus model used in this article has generally higher extraction accuracy than the Eddynet model. After calculation, it can be concluded that the Kappa coefficient is increased by 0.2553, and the extraction result parameters of the gas swirl and antigas swirl have been improved. Analysis of the reasons, first, because the label data produced in this experiment is directly used for model training using the entire global data, without data cropping and other enhancement processing, the original texture features and relative spatial position information of the data are better maintained. Second, the DeeplabV3plus model is a convolutional neural network model that performs pixel-by-pixel segmentation. It itself considers the relative laws of surrounding pixels, which is also the main reason for the entire data set to be used.

Table 1. Comparative analysis of evaluation indexes of mesoscale vortex extraction results

| Extract model | DeeplabV3plus model | Eddynet model |
|---------------|---------------------|---------------|
| Parameter category | Anticyclone | cyclone | Anticyclone | cyclone |
| Precision    | 0.818               | 0.791         | 0.673       | 0.681   |
| Recall rate  | 0.759               | 0.819         | 0.507       | 0.493   |
| F1 score     | 0.786               | 0.806         | 0.577       | 0.571   |
| Kappa coefficient | 0.7582            |               | 0.5029      |

3.3. Analysis of results

Using the DeeplabV3plus model to extract the mesoscale vortex distribution results on March 13, 2018 (Figure 6), it can be clearly seen that mesoscale vortices are widely present in the global ocean. Except for the low-latitude equator and Arctic circumpolar current regions, there are a large number of relatively
uniformly distributed mesoscale vortices in other sea areas, especially in the western boundary current sea areas, such as the extension of the Gulf Stream and Kuroshio, and the vicinity of the Brazilian current. The distribution of mesoscale vortices in the equatorial region at low latitudes is relatively sparse, and in the Southern Ocean, they are mainly distributed along the Antarctic Circumpolar Current [13]. Studies have shown that it is the Rossby deformation in low latitudes that causes the mesoscale vortices in the sea area to become larger in scale and smaller in amplitude, making it difficult to extract such vortices.

The DeeplabV3+ model based on semantic segmentation is a convolutional neural network model that performs pixel-by-pixel segmentation, which can fuse high-level abstract semantic features and super-radial detail information, thereby improving the accuracy of mesoscale vortex extraction.

Using the recognition result based on the DeeplabV3plus model and the label data based on the py-eddy algorithm for overlay analysis (Figure 7), TP in the figure represents the region where the model recognition result and the label data are both mesoscale vortices, and TN represents the label recognition as a vortex. The area not recognized by the DeeplabV3plus model, FP represents the area recognized by the DeeplabV3plus model as a vortex but not recognized by the label. It can be seen that, compared with the py-eddy algorithm, the DeeplabV3plus model can generally identify more vortices. Most of the TN area in the picture appears around the TP area, indicating that the vortex in the label is still recognized, but the range is reduced on the original basis, making the mesoscale vortex recognition more accurate.

In the figure, there are more FP regions with a smaller area and exist independently, indicating that the DeeplabV3plus model recognizes a larger number of vortices than the py-eddy algorithm. In this study, the recognition results based on the DeeplabV3plus model showed many small-scale vortices, that is, a vortex composed of a small number of pixels, and the number of pixels with the smallest vortex in the label data based on the py-eddy algorithm was greater than 4. Analyze the reason, mainly because the py-eddy algorithm is based on closed contour recognition, at least 4 pixels are required to be recognized as a vortex. In this paper, the recognition based on the DeeplabV3plus model is based on a single pixel and combined with the correlation of surrounding pixels, so there will be more vortices composed of a few pixels.
The method used to make label data in this experiment is based on the recognition of contour closed contours. This recognition method needs to set a threshold to distinguish mesoscale vortices from other background sea areas. The selection of the threshold directly affects the accuracy of mesoscale vortex recognition. According to the complex geographic conditions of different sea areas, different thresholds need to be set to identify mesoscale vortices. The model of this study can avoid the subjectivity of relying on the threshold. This study selects the same sea area to compare and analyze the two mesoscale vortex extraction methods based on deep learning. Since the measured data of mesoscale vortex in the global sea area cannot be obtained, only two The results of the method are compared with the label data (Figure 8). It can be clearly seen that the DeeplabV3+ model is better than the Eddynet model in the identification of the equator. In other sea areas except the equator, the identification results of the two methods are basically the same. This research method reduces a large number of misidentifications, and re-identifies some large-area mesoscale vortices near the equator as smaller-scale vortices, and it takes less time than previous identification methods.

Figure 8. Comparison of the extraction results of label data (a), Deeplabv3+ model (b), and Eddynet model (c) based on py-eddy algorithm in low-latitude equatorial waters
Selecting the Arctic, North Pacific and South Atlantic waters with obvious differences for further comparison (Figure 9), it can be seen that the recognition results of the DeeplabV3plus model are more similar to the label data. In the Arctic region, the extraction results of DeeplabV3plus and Eddynet model are better than the py-eddy algorithm, which avoids the misidentification of long and narrow eddy currents. However, the recognition results of Eddynet model are different from the results of the first two methods. Small-scale eddy currents appear in the sea boundary area. In the North Pacific waters, the Eddynet model recognition results did not better distinguish the details of the vortices, and many smaller mesoscale vortices were not recognized. In the South Atlantic Ocean, the Eddynet model extracts many small-scale eddies.

Figure 9. Tag data based on py-eddy algorithm (a), DeeplabV3plus model extraction result (b), Eddy net extraction result (c)

4. Conclusion
In the traditional mesoscale vortex recognition method, the method based on physical parameters needs to set the appropriate threshold value for the specific area, and the recognition method based on geometry needs to have clear geometric features. All current methods highly rely on expert knowledge and are all aimed at Recognized by a specific area. In this paper, convolutional neural network is applied to ocean remote sensing data processing, namely automatic identification and positioning of mesoscale vortex based on sea level anomaly (SLA). The semantic segmentation network DeeplabV3plus convolutional neural network is applied to the extraction of mesoscale vortices. Through the creation of data sets, network model training, and parameter adjustment, the automatic recognition of mesoscale vortices and semantic segmentation are successfully combined. This model uses ten years of daily global sea level data to produce a label data set. After a large number of experiments, it has been shown that the input data size is the best when maintaining the original data size, so data cutting is not performed. Improve the loss function of the original network, speed up the convergence speed of the model, and improve the training accuracy of the model. Choose the Adam optimizer with adaptive learning rate to improve the generalization ability of the network model. By comparing the extraction results of this model and the Eddynet model, it is concluded that the accuracy of the model recognition accuracy is
increased by 13.8%, and the Kappa coefficient is increased by 0.2553. Near the equator and the north and south poles of the global ocean, this model is significantly better than the Eddynet extraction method in terms of extraction accuracy.

The limitation of this model is that only one variable of SLA is used in training. In the next step, you can consider combining multiple variables for model training, such as fusing high-resolution sea surface temperature data, sea level height data, and chlorophyll concentration. Improve the recognition accuracy of the model. In addition, you can consider processing the original data set to calculate the gradient value between each pixel in each data, so that the edge features of the mesoscale vortex can be better highlighted, which is more conducive to the training of the convolutional neural network model.

Acknowledgments
Thanks to the National Natural Science Foundation of China (41801071); the ‘Guangxi BaGui Scholars’ Special Fund.

References
[1] Zhang Lin, Li Lianglong, & Lu Xiaoting. (0). Analysis of mesoscale vortex sound field in South China Sea based on parabolic equation model. 2009 National Hydroacoustics Conference.
[2] Li Anzhou, Zhou Weifeng, & Fan Xiumei. (2017). Research progress in remote sensing image mesoscale ocean front and vortex extraction methods. Chinese Journal of Image and Graphics (6).
[3] Okubo, A. . (1970). Horizontal dispersion of floatable particles in vicinity of velocity singularities such as convergences. Deep Sea Research and Oceanographic Abstracts, 17(3), 445-454.
[4] Weiss, J. . (1991). The dynamics of enstrophy transfer in two-dimensional hydrodynamics. Physica D-nonlinear Phenomena, 48(2-3), 273-294.
[5] Nencioli, F. , Dong, C. , Dickey, T. , Washburn, L. , & Mewilliams, J. C. . (2010). A vector geometry-based eddy detection algorithm and its application to a high-resolution numerical model product and high-frequency radar surface velocities in the southern california bight. Journal of Atmospheric & Oceanic Technology, 27(3), 564.
[6] Lugensat, R. , Sun, M. , Fablet, R. , Mason, E. , Tandeo, P. , & Chen, G. . (2017). Eddynet: a deep neural network for pixel-wise classification of oceanic eddies.
[7] Mason, E. , Pascual, A. , & Mcwilliams, J. C. . (2014). A new sea surface height–based code for oceanic mesoscale eddy tracking. Journal of Atmospheric and Oceanic Technology,31(5), 1181–1188.
[8] Chelton, D. B. , Schlax, M. G. , & Samelson, R. M. . (2011). Global observations of nonlinear mesoscale eddies. Progress in Oceanography, 91(2), 167-216.
[9] Early, J. J. , Samelson, R. M. , & Chelton, D. B. . (2011). The evolution and propagation of quasigeostrophic ocean eddies *. Journal of Physical Oceanography, 41(8), 1535-1555.
[10] Yang Jianyu, Zhou Zhenxu, Du Zhenrong, Xu Quanquan, Yin Hang, & Liu Rui. (2019). High resolution remote sensing image rural construction land extraction based on segnet semantic model. Transactions of the Chinese Society of Agricultural Engineering, 35(05), 259-266.
[11] Milletari, F. , Navab, N. , & Ahmadi, S. A. . (2016). V-Net: Fully Convolutional Neural Networks for Volumetric Medical Image Segmentation. 2016 Fourth International Conference on 3D Vision (3DV). IEEE.
[12] Chelton, D. B. , Schlax, M. G. , & Samelson, R. M. . (2011). Global observations of nonlinear mesoscale eddies. Progress in Oceanography, 91(2), 167-216.
[13] Xu, G. , Cheng, C. , Yang, W. , Xie, W. , & Yang, J. . (2019). Oceanic eddy identification using an ai scheme. Remote Sensing, 11(11), 1349-.
[14] Mason, E. , Pascual, A. , & Mcwilliams, J. C. . (2014). A new sea surface height–based code for oceanic mesoscale eddy tracking. Journal of Atmospheric and Oceanic Technology, 31(5), 1181–1188.