



An automatic trough line identification method based on improved UNet

Yali Cai, Qian Li*, Yin Fan, Liang Zhang, Hong Huang, Xinya Ding

College of Meteorology and Oceanography, National University of Defense Technology, Changsha, China



ARTICLE INFO

Keywords:
Trough line
Automatic recognition
Weather analysis
UNet

ABSTRACT

A trough is an elongated region of relatively low atmospheric pressure. Automatic analysis and recognition of trough lines in upper weather charts are challenging and of great significance to weather analysis. The existing automatic identification methods mainly depend on manually setting rules which cannot cover all trough line types and have low generalization ability. This paper proposes an automatic trough line identification method based on an improved model which can extract the trough line from meteorological element data at 500 hPa. The model adopts the UNet of a U-shaped encoder and decoder as basic structure, which is designed to enable precise localization by continuously combining low-level and high-level features. To extract abstract semantic features of the trough, the Xception, which takes depthwise separable convolution as basic unit, is adopted to replace the encoder of the original UNet. In addition, the Squeeze and Excitation (SE) module with an attention mechanism is added after each ordinary convolution in the decoder part to improve the recognition accuracy by increasing the weighting of the trough area. The experiments are conducted on a meteorological dataset and the results show that the recognition accuracy with our proposed method on the testing dataset can reach over 80%. We also compare our results to several other types of networks and traditional automatic identification methods, which demonstrates that the performance of the proposed network is superior to other methods.

1. Introduction

In atmospheric science, a trough is defined as a long narrow area extending from a low-pressure area (Sawyer, 1956), in which the pressure value is lower than that on both sides, the horizontal convergence of airflow, the strong shear wind direction and the temperature variation are usually more intense (Wallace and Hobbs, 2006). The ability of trough to greatly affect forecasters' predictions makes it essential for weather analysis.

In the upper-air weather map, the trough is usually represented by a line that connects the maximum points of cyclonic curvature in the low-pressure trough of isolines (Allen et al., 2006). Automatic trough line analysis can objectively assist forecasters in diagnosing weather system structure and improve the efficiency of weather analysis. In addition, it can provide data support for historical statistics, retrieval, and comparative analysis of related forecast quantities, such as precipitation, thunderstorms, and gales (Juckes and Smith, 2000; Feteris, 1973; Bandara et al., 2021; Mahbod and Rafiee, 2021). Objective analysis of the trough line, therefore, plays a key role in weather analysis (Neu et al., 2013; Ulbrich et al., 2013). However, it is still undertaken manually, which prolongs the weather forecast and is prone to

introduction of errors, omissions, inconsistent results, and other issues (Tan, 2006; Yu and Hu, 2011). As a result, automatic recognition of trough lines has become one of the key issues requiring solution in the field of automated weather analysis.

Existing automatic trough line analysis methods have disadvantages such as incomplete coverage of trough line types caused by manual setting of rules and low generalization ability (Mou, 2010; Dai et al., 2016; Li et al., 2018). In order to address this problem, this paper proposes a new method which is primarily based on the convolutional neural network (CNN) to identify the trough line from the data grid of three meteorological elements, i.e., geopotential height, wind, and temperature. In the designed network, the UNet which adopts the structure of encoder-decoder is employed as a backbone, the Xception network is employed in the encoder part of the structure and an attention mechanism module named Squeeze and Excitation (SE) is fused into the decoder part. Firstly, a training dataset composed of the aforementioned three meteorological elements and trough labels is used to train the network, this generates a model which can calculate the probability of each point in the grid being a trough point. Points with probability values exceeding 0.5 are determined as trough points and a map showing the detected trough will be output. Finally, the trough line can

* Corresponding author.

E-mail addresses: public_caiyali@163.com (Y. Cai), public_liqian@163.com (Q. Li).

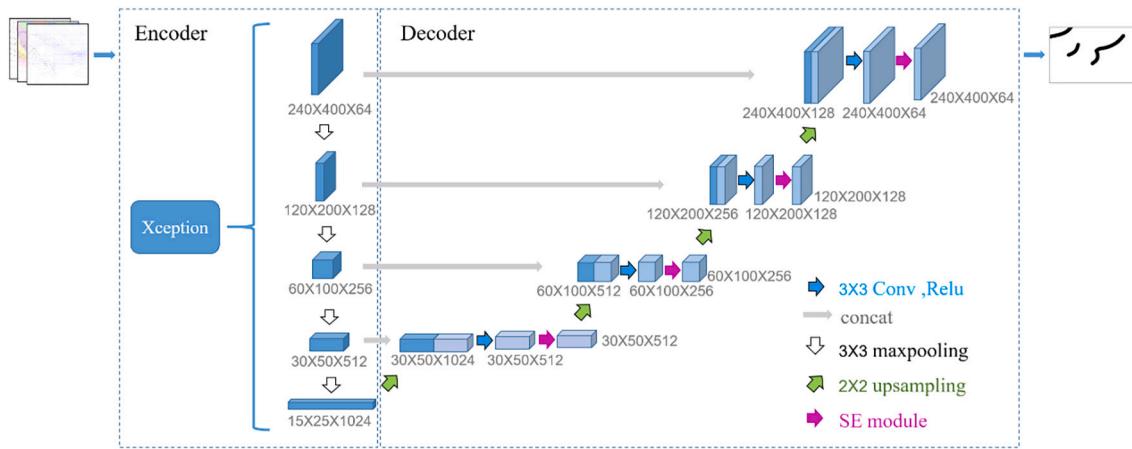


Fig. 1. The framework of the proposed network contains two components: an encoder and a decoder. The encoder is used to extract trough features from the meteorological data, in which the input samples are downsampled 4 times. The decoder is used to transform the parsed information into the final feature maps whose shape is the same as the initial features through 4 upsampling stages. Each blue box in the figure corresponds to a multi-channel feature map. The output feature shape is at bottom of each layer, in which the first two represent the height and width of the feature maps, and the last represents the number of feature maps. The arrows denote different operations. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

be extracted from the trough marked map using a skeleton extraction algorithm.

In the next section, we provide a brief literature review of relevant work. [Section 3](#) introduces the proposed method, and [Section 4](#) shows the experiment's dataset, results, and analysis, in which the results clearly indicate superiority over traditional methods and other networks. Finally, in [Section 5](#) we discuss our results and outline key conclusions.

2. Related work

The existing automatic trough line identification methods identify discrete trough points in the data and connect them mainly depending on the graphic characteristics of the trough line. Based on previous studies, these methods can be split into three categories: those based on wind data, geopotential height data, and a combination of the first two data types.

In the methods based on wind data, the location of the trough line is determined by detecting wind shear. For example, [Bluestein and Speheger \(1995\)](#) firstly determined the trough points by the wind speed and height between stations, and then connected them according to the predefined rules. [Fröhlich and Knippertz \(2008\)](#) used the vertical averages of the upper-level Potential Vorticity (PV) to identify low-altitude disturbances from the extratropics with a certain vertical depth. The above methods based on the wind generally have poor performance because the formation of trough is affected by multiple factors such as geopotential height, temperature, and so on.

In the methods based on geopotential height data, the position of the trough line is obtained by calculating the curvature of the isobaric line in data grid according to various algorithms. [Huang and Zhao \(2000\)](#) identified the features of the trough line by establishing a relational framework on the basis of spatial aggregations. [Wong et al. \(2008\)](#) used a genetic algorithm to define the fitness function for low-voltage area identification, and then used the adjacent arc segments of the isobars to analyze trough lines. [Mou \(2010\)](#) first identified the isobars through an edge detection algorithm, then analyzed the resulting trough points and drew trough lines based on the ant colony algorithm. The identification performance of these methods based on calculating isobaric curvature is slightly improved compared to the original methods based on a single wind data type; however, these approaches didn't integrate all of the data relevant to trough formation and were very likely to lead to errors and omissions.

Since the formation and evolution of trough are related to various

factors such as wind, geopotential height, etc., some researchers ([Dai et al., 2016; Li et al., 2018](#)) considered fusing wind and geopotential height data to automatically identify the trough line. In their work, the initial position of the trough line was determined on the basis of the geopotential height data, and wind data was used to modify the preliminary position in order to calculate the final position of the trough line. However, like the previous two methods, some rules and thresholds in the recognition process needed to be manually specified on the basis of historical data, so their application to different trough line types and geographic regions was limited by data availability. In recent years, deep learning methods have been applied to identify weather systems. [Lagerquist et al. \(2019\)](#) used CNN to identify fronts in data grid of wind velocity, temperature, specific humidity and wet-bulb potential temperature. [Biard and Kunkel \(2019\)](#) employed an algorithm for the automatic detection of fronts, which was developed by training a two-dimensional CNN with manually analyzed fronts and surface fields of five atmospheric variables.

To achieve integration of multiple diverse meteorological elements and model multi-scale features of trough lines, inspired by the similarity between meteorological data grid and image data, we regard the meteorological element data grid as multi-channel image data and apply the Deep Convolutional Neural Network (DCNN) ([Hu et al., 2015; Long et al., 2015; Badrinarayanan et al., 2017; Lin et al., 2017](#)) for semantic segmentation ([Jiang et al., 2019; Liu et al., 2019; Sultana et al., 2020; Taghanaki et al., 2021](#)) to extract trough from three meteorological elements related to the formation of the trough lines. When data are input into the network model, the features of the trough can be extracted through continuous convolutions and pooling. The probability of each point being a trough point will be recorded through parameter learning. Generally, the skeleton center line of the trough is determined as a trough line. Through experimental verification, we demonstrate that the proposed method performs better than traditional recognition methods and other semantic segmentation networks.

3. Method

The network presented in this paper firstly downsamples the data to extract abstract features from meteorological elements and then continuously upsamples these features to their original size. After a series of convolution and pooling operations, a probability grid which records the probability of each point in the grid to be a trough point is output, from which the trough segmentation results can be determined. At the end, the skeleton center lines of the trough area will be identified

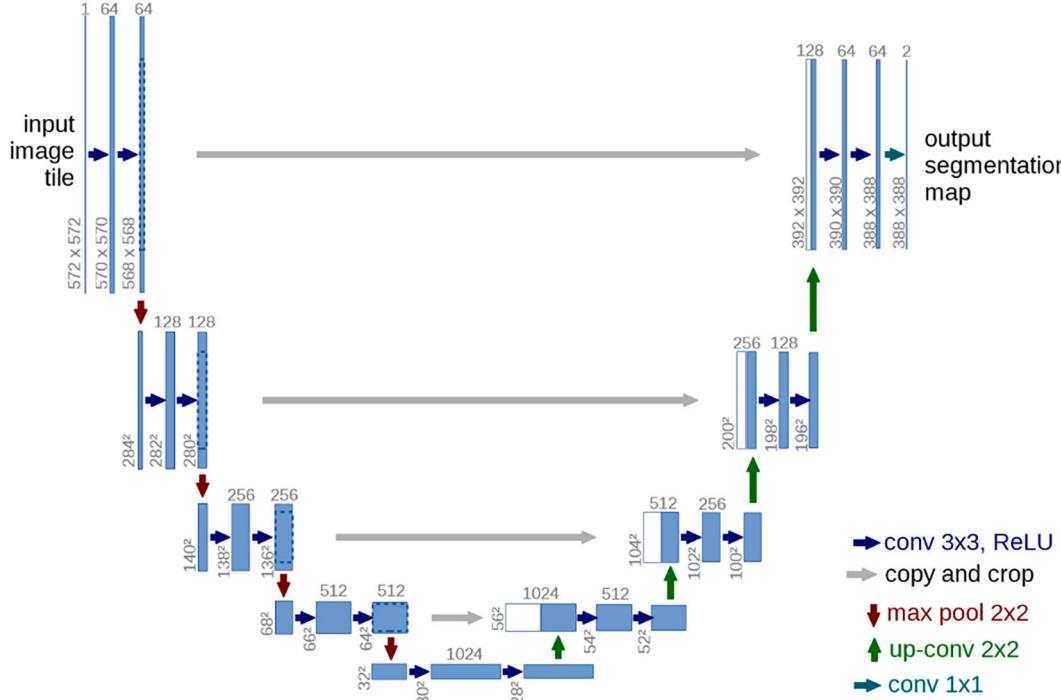


Fig. 2. The UNet consists of three parts: downsampling, upsampling, and skip connection. On the left is the compression process, namely the encoder, in which convolution and downsampling are used to reduce the feature shape and extract shallower features. The right part is the decoding process, namely the decoder, in which convolution and upsampling are used to identify deep features. In the middle, the features obtained in the encoder are combined with the features obtained in the decoder by skip connection and the features are refined by combining them.

as trough lines.

3.1. The network

3.1.1. Proposed network architecture: XSU-Net

We propose an improved network based on UNet (Ronneberger et al., 2015) to perform automatic identification of trough. As shown in Fig. 1, the network consists of two modules, that is, an encoder and a decoder. The encoder module performs multiple pooling and depthwise separable convolution operations on the entered meteorological element data to enlarge the receptive field while the feature shape is changed from $240 \times 400 \times 64$ to $15 \times 25 \times 1024$. The decoder module performs multiple convolution and upsampling operations on the feature maps output by the encoder and finally outputs a set of highly abstract features while the feature shape is changed from $15 \times 25 \times 1024$ to $240 \times 400 \times 64$.

The encoder module is shown in the dashed window on the left, which continuously reduces the feature shape. After multiple depthwise separable convolutions and four downsampling operations, the feature maps are entered into the decoder module in the dashed window on the right and upsampled. A skip connection is performed between the features obtained from each upsampling step and the corresponding features of the same shape in the encoder module, which can fuse the generalized low-level features and the obvious characteristics of high-level features. The feature maps obtained by skip connection are input to the attention mechanism SE (Hu et al., 2017) module after an ordinary convolution layer. The SE module calculates the importance of each channel mainly through the Squeeze and Excitation operation which can focus on the more relevant data channels to produce good performance. Finally, the obtained features are input into the sigmoid activation function (Finney, 1952) and the probability of trough points for each point is mapped to [0,1]. The points with a probability of more than 0.5 are defined as trough points. Finally, the data grid marked with the trough area is output.

3.1.2. UNet

UNet utilizes a symmetric structure which can fuse feature maps between different levels. The architecture of the UNet is shown in Fig. 2. The UNet consists of three parts: downsampling, upsampling, and skip connection. On the left is the compression process, namely the encoder, in which convolution and downsampling are used to reduce the feature shape and extract shallower features. The right part is the decoding process, namely the decoder, in which convolution and upsampling are used to identify deep features. In the middle, the features obtained in the encoder are combined with the features obtained in the decoder by skip connection and the features are refined by combining them. The network proposed in this paper adopts the U-shaped symmetric UNet as the backbone to supplement the usual shrinking network with continuous layers; this consists of a contraction path (on the left in Fig. 2) and an expansion path (on the right in Fig. 2). The contraction path implements repeated convolution and downsampling on input samples, whereas in the expansion path the pooling operations are replaced by upsampling to increase the resolution of the output. The skip connection between the contraction path and the expansion path can effectively achieve partial localization of useful features. The continuous convolutional layers learn to assemble a more accurate output from this information by combining the high-resolution characteristics of the contraction path with the upsampled output.

3.1.3. The Xception

One of the challenges in automatic trough identification is to model trough areas of different scales. The encoder module reduces the shape of input features constantly and increases the number of feature channels to model the trough areas by performing multiple pooling and downsampling operations. However, the encoder in the original UNet is a simple structure containing only four pooling layers and is mainly aimed at the relatively simple biomedical semantic segmentation problem. The comprehensive mathematical relationship between the formation of trough lines and meteorological elements is very

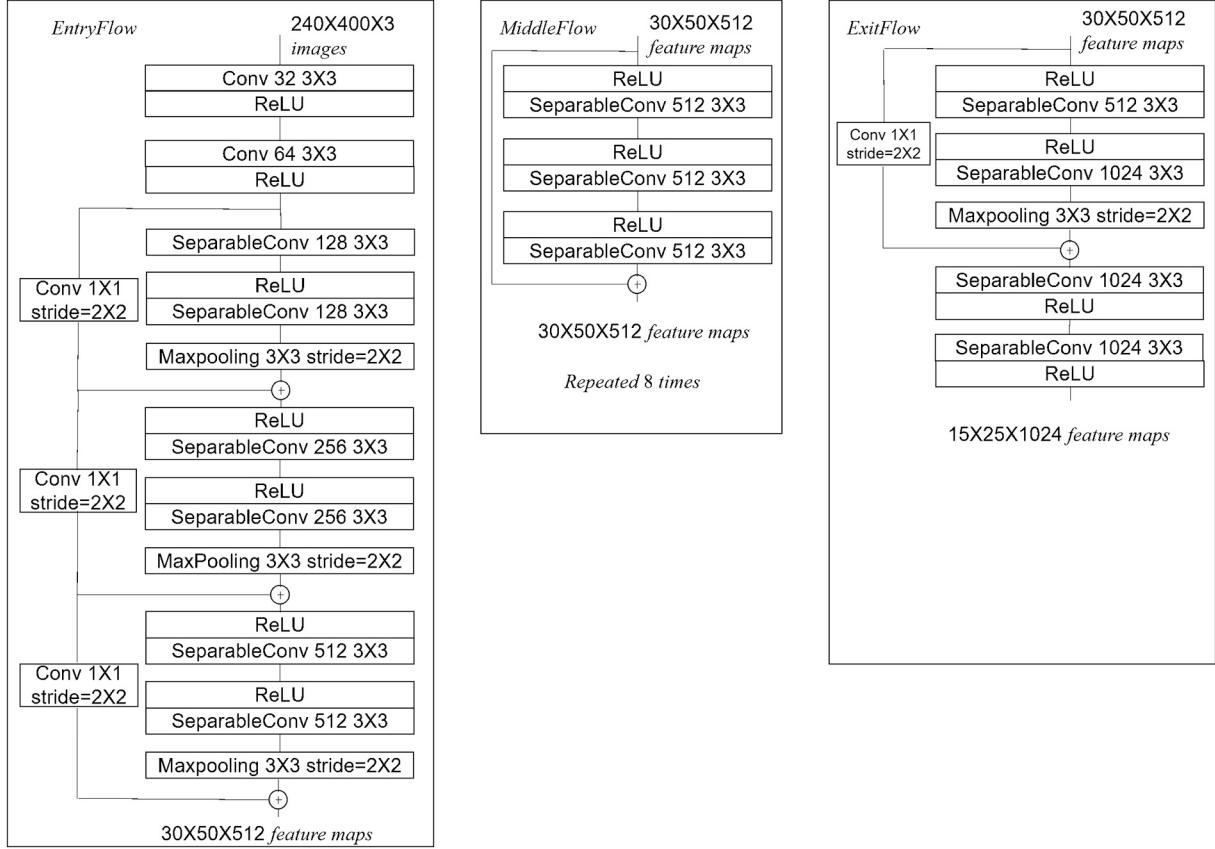


Fig. 3. The encoder architecture: the encoder is composed of three modules, i.e., an entry flow, a middle flow which repeats 8 times, and an exit flow. There are 14 blocks, including 4 in the entry flow, 8 in the middle flow, and 2 in the exit flow respectively.

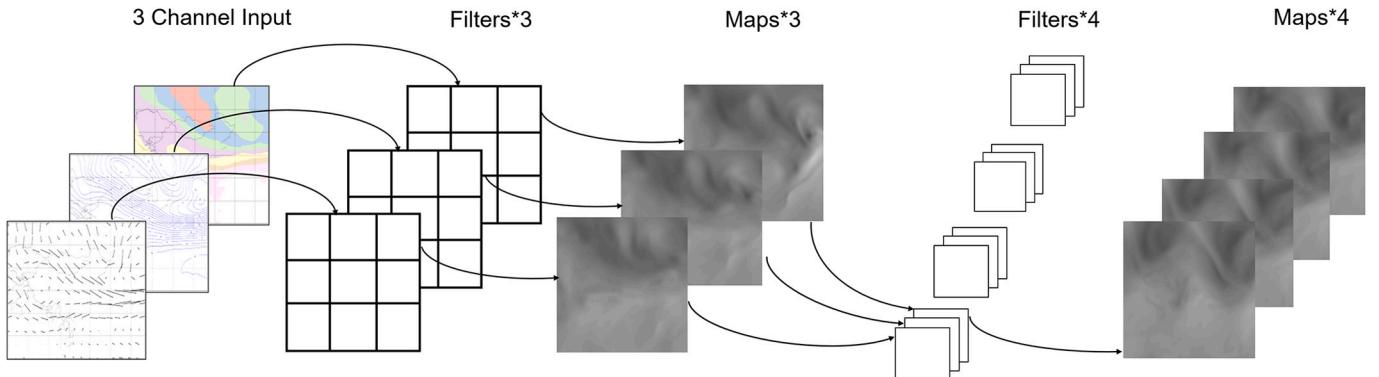


Fig. 4. In depthwise separable convolution, depthwise convolution is carried out first, in which a convolution kernel is only responsible for one channel, and then separable convolution is performed, which is similar to conventional convolution.

complicated, it will be difficult to extract features of trough from meteorological elements data with a simple encoder module, so the encoder module should be optimized to extract more semantic information. Considering that the Xception (Chollet, 2017) network has better performance and fewer parameters compared to other low-level feature extraction networks, we replaced the encoder module of the UNet with Xception in our proposed network. The encoder module has 36 convolution layers, which constitute the basis for feature extraction of the proposed network. The 36 convolution layers are structured into 14 modules, all of which have linear residual connections except for the first and last modules. The structure diagram of the encoder module is shown in Fig. 3. The data grid first goes through the entry flow in which the feature shape changes from the original 240 × 400 × 3 to 30 × 50 ×

512, then through the middle flow in which the feature shape remains the same, and finally through the exit flow where the feature shape changes to 15 × 25 × 1024. Fig. 4 also demonstrates that depthwise separable convolution can effectively reduce the computation.

The basic unit block of the encoder is the depthwise separable convolution layer, which is completed in two steps. As shown in Fig. 4, the first step is to convolve each channel separately. The second step is to expand the depth and convolve between channels, which can simultaneously map cross-channel and spatial correlation and greatly reduce the number of network parameters. Convolution operations occupy the vast majority of computation time in neural networks; the main reason for this is the large computational resource demand involved in matrix multiplication operations. When the stride = 1, the calculation quantity

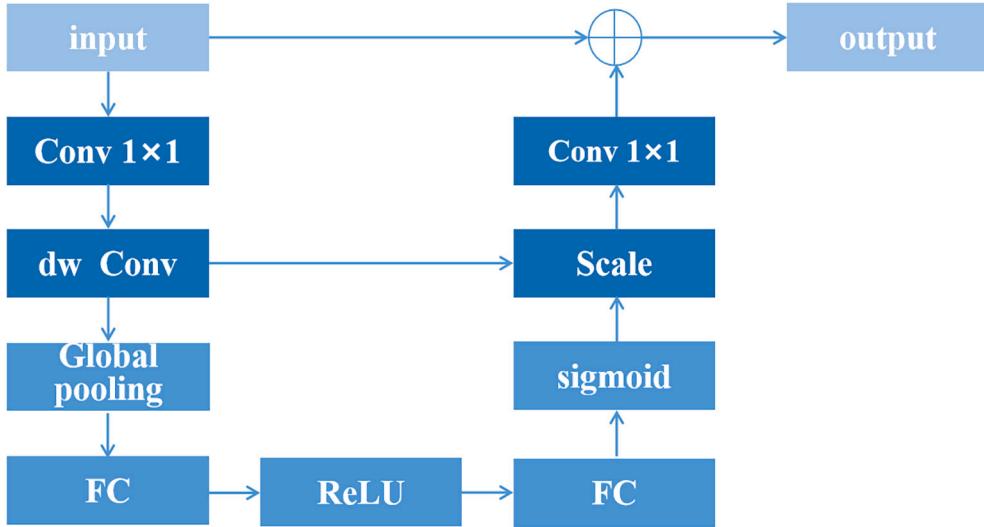


Fig. 5. The schema of the SE module; by processing the convolved feature maps, a one-dimensional vector with the same number of channels is obtained as the evaluation score of each channel, and the modified score is then applied to the corresponding channel to get the result.

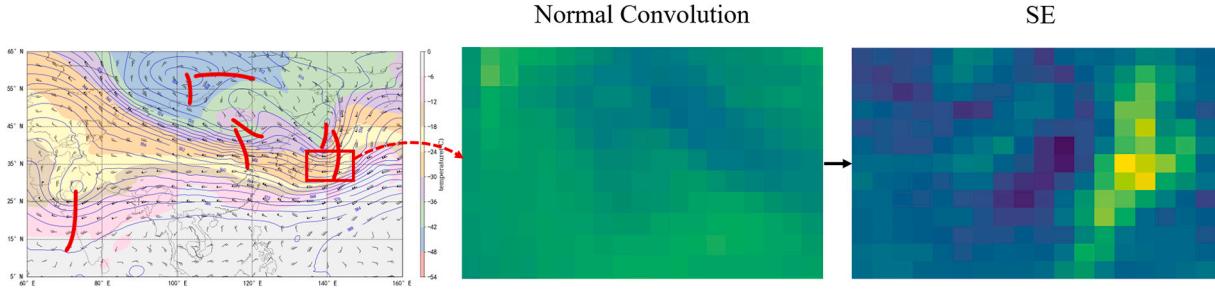


Fig. 6. The left image is an upper-air weather map at 0000 UTC 1 January 2001, which contains three meteorological elements of geopotential height, wind, and temperature. The trough areas are marked in red. After ordinary convolution, the feature map of the trough area is shown in the center. After applying the attention module, the feature map on the right side is obtained, in which the features of the trough area are extracted more effectively. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

in conventional convolution is:

$$F_{conv} = C_i \times K_H \times K_W \times C_o \times H_o \times W_o \quad (1)$$

where K_H and K_W represent the height and width of the convolution kernel respectively, C is the channel number, H and W represent the height and width of the feature map respectively, i and o represent the input and output.

Compared with conventional convolution, the calculation quantity in depthwise separable convolution is:

$$F_{DS_conv} = C_i \times K_H \times K_W \times H_o \times W_o + C_i \times C_o \times H_o \times W_o \quad (2)$$

The ratio of the two is as follows, generally, the value of this ratio is less than 1.

$$\frac{F_{DS_conv}}{F_{conv}} = \frac{1}{C_o} + \frac{1}{K_H \times K_W} \quad (3)$$

3.1.4. SE module with attention mechanism

Considering the proportion of trough areas in weather maps is so small, a considerable amount of computational resources are wasted in non-trough areas when using regular convolution. Therefore, an SE module which can adaptively recalibrate channel-wise feature responses by explicitly modeling interdependencies between channels is added in the decoder, so that the representational power of the network can be strengthened by enhancing the quality of spatial encoding throughout its feature hierarchy. As shown in Fig. 5, the SE module contains a

Squeeze operation and an Excitation operation. The Squeeze operation carries out feature compression along the spatial dimension through a global average pooling and turns each two-dimensional feature channel into a real number. The result has a global receptive field and the operation matches the output dimension with the number of input feature channels. This represents the global distribution of responses on the feature channel and enables the layer close to the input to obtain the global receptive field, which is a useful operation for many tasks. After the pooling operation, the numerical distribution of each feature map, i.e., the global information, can be obtained. The Excitation operation consists of two fully connected layers and a ReLU layer, which is similar to a gate in the recurrent neural network. The weighting is generated for each feature channel through the parameters that have been learned to explicitly model the correlation between feature channels. As shown in Fig. 6, embedding the SE module into the upsampling part of the UNet can effectively increase the weight of the trough area and improve the overall segmentation accuracy of the network.

3.2. The loss function

In the network model, a loss function needs to be designed to optimize the parameters of the network model. When the classifier in the network uses a sigmoid function and is applied to biomedical segmentation problems, the cross-entropy loss function (Farahnak-Ghazani and Baghshah, 2016) is usually used to achieve faster convergence. Since the trough area accounts for only about 5% of the whole weather map, we

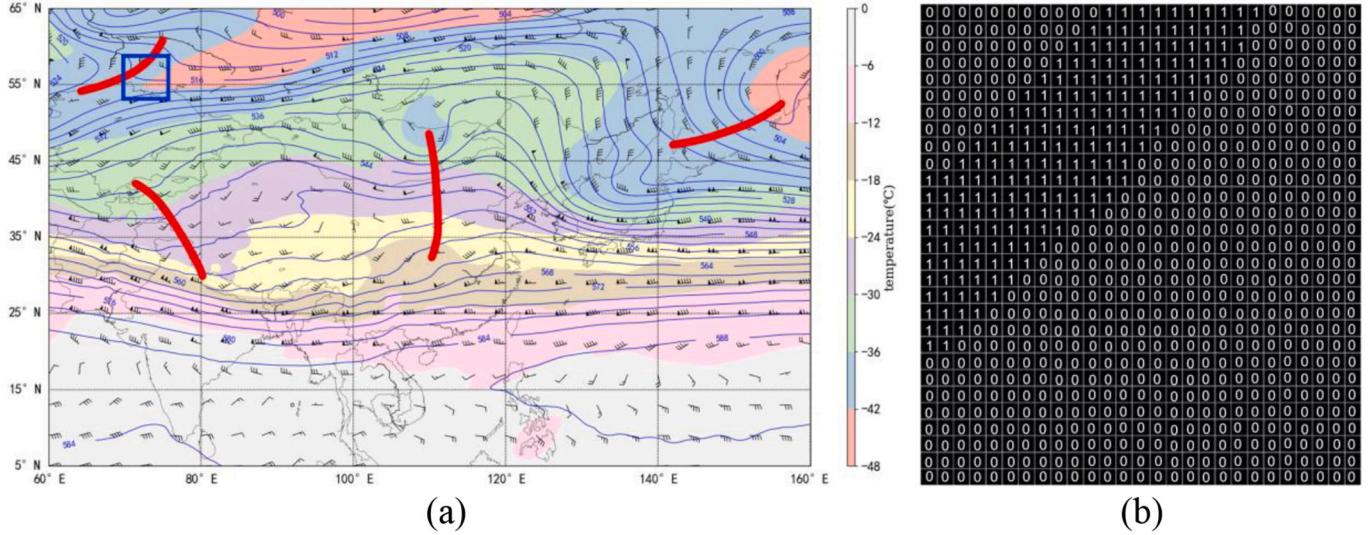


Fig. 7. (a) shows the weather map at 0000 UTC 23 January 2001. The map includes isobaric line, temperature, wind and regional contours. The red curves are the manually marked trough lines. After expanding the trough label, the data grid generated from the area highlighted in the blue box is shown in (b), in which the value of points in the trough area is 1, and the value of other areas is 0.

used the Class-Balanced cross-entropy loss function (Cui et al., 2019) to improve the imbalance between positive and negative categories. Assuming that the predicted output of the model is $\mathbf{z} = [z_1, z_2, \dots, z_c]^T$, where c is the total number of classes and y represents a given class, this is expressed as (4):

$$CB_{\text{sigmoid}}(\mathbf{z}, y) = \frac{1 - \beta}{1 - \beta^{n_y}} \sum_{i=1}^c \log \left(\frac{1}{1 + \exp(-Z_i^t)} \right) \quad (4)$$

In the formula, β represents the proportion of labeled trough points relative to all points in the grid, and n_y represents the number of samples in the truth value of class y . Where Z_i^t is defined as (5):

$$Z_i^t = \begin{cases} Z_i, & \text{if } i = y. \\ -Z_i, & \text{otherwise.} \end{cases} \quad (5)$$

3.3. Datasets and preprocessing

The meteorological elements data can be downloaded from the website of the European Centre for Medium-Range Weather Forecasts (ECMWF) (Simmons et al., 2007), which include the meteorological elements of geopotential height, wind (u and v), and temperature. The data coverage ranges from 60° E to 160° E longitude, from 5° N to 65° N latitude at 500 hPa, and data resolution is 0.25 degrees. Labels for these data were produced by professional meteorologists. We selected a total of 7122 samples from the period 2001 to 2010, which were then divided into three sets: 5000 samples as training set, 1500 samples as validation set, and 622 samples as testing set. The training set is used to fit the model and train the segmentation model under a fixed setting of hyper-parameters. The validation set records the accuracy of the model and adjust its hyper-parameters in order to optimize the model after training it multiple times. The testing set is used to measure the model's performance and segmentation ability.

In order to adapt the samples to the network, the original data must be refined before network training, including preprocessing of meteorological elements and labels. Firstly, in the meteorological element variables, the u wind and v wind components need to be converted into wind direction which has a higher correlation with trough formation. Moreover, loss calculation of the network model in deep learning assumes that all features of the data have a zero mean value and have the same order of variance, however, extreme values at both ends of the original data differ greatly from the mean value and it is therefore

necessary to normalize the data to improve the convergence rate of the model. All three types of data were normalized to the range [0, 255] according to the size of the data and distribution interval. Secondly, as the original labels were vector data indicating trough lines, these were rasterized to data grid showing trough locations for matching the labels with meteorological elements. However, the proportion of trough lines by area in the weather map data is too small for sufficient accuracy and it will unduly punish slight offsets between the predictions and the ground truth when using trough line labels, so the linear trough line features were interpolated into trough label grid; the resulting labels generated are shown in Fig. 7.

3.4. Evaluation metrics

In order to measure the automatic trough line recognition performance of the proposed method more comprehensively, this study considers the accuracy of the results from two perspectives. The first type of metric is used to evaluate the identification performance of the network in recognition of the trough area. To measure the performance of recognition results more comprehensively, we also adopt a second type of metric which considers the overall prediction precision of the extracted trough lines.

The first metric type includes ACC , $PTOP$, $PTOT$ and $PNON$, in which ACC is the ratio of correctly classified points to total points in a dataset, $PTOP$ refers to the proportion of correct trough points among the predicted trough points, $PTOT$ refers to the ratio of the predicted correct trough points to all marked trough points and $PNON$ is the ratio of all correctly predicted non-trough points to all non-trough points in labels. Finally, P_1 is used to comprehensively measure the accuracy of the predicted trough area. The above indicators are calculated as follows:

$$ACC = \frac{TT + NN}{TT + NN + NT + TN} \quad (6)$$

$$PTOP = \frac{TT}{TT + NT} \quad (7)$$

$$PTOT = \frac{TT}{TT + TN} \quad (8)$$

$$PNON = \frac{NN}{NN + NT} \quad (9)$$

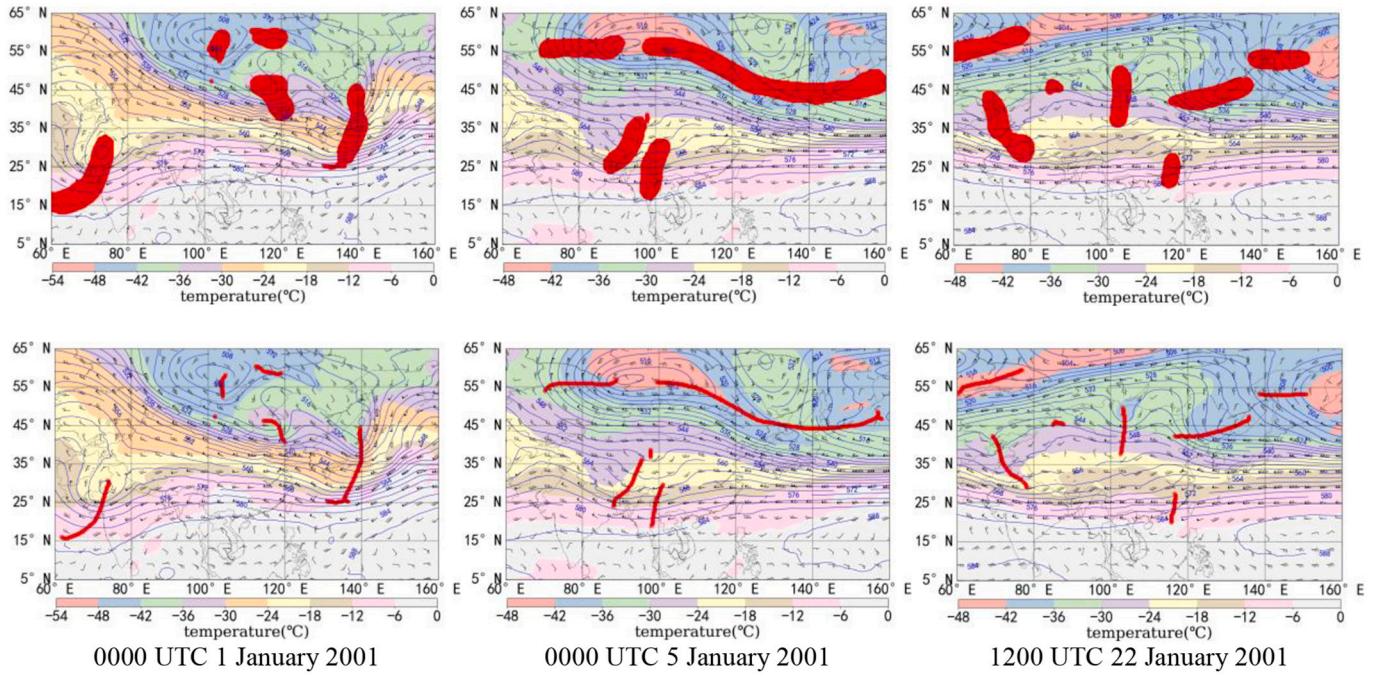


Fig. 8. The upper images show weather maps marked with troughs as automatically recognized by the network, and the lower images show weather maps marked with trough lines as extracted by the skeleton line extraction algorithm.

$$P_1 = 2 \cdot \frac{PTOP \cdot PTOT}{PTOP + PTOT} \quad (10)$$

where TT is the number of correctly predicted trough points, TN is the number of predicted non-trough points which are trough points in the

manually analyzed data, NT is the number of predicted trough points which are non-trough points in the manually analyzed data, and NN is the number of correctly predicted non-trough points.

The second type of evaluation metrics evaluate the results mainly based on the matching degree of each pair of trough lines in the auto-

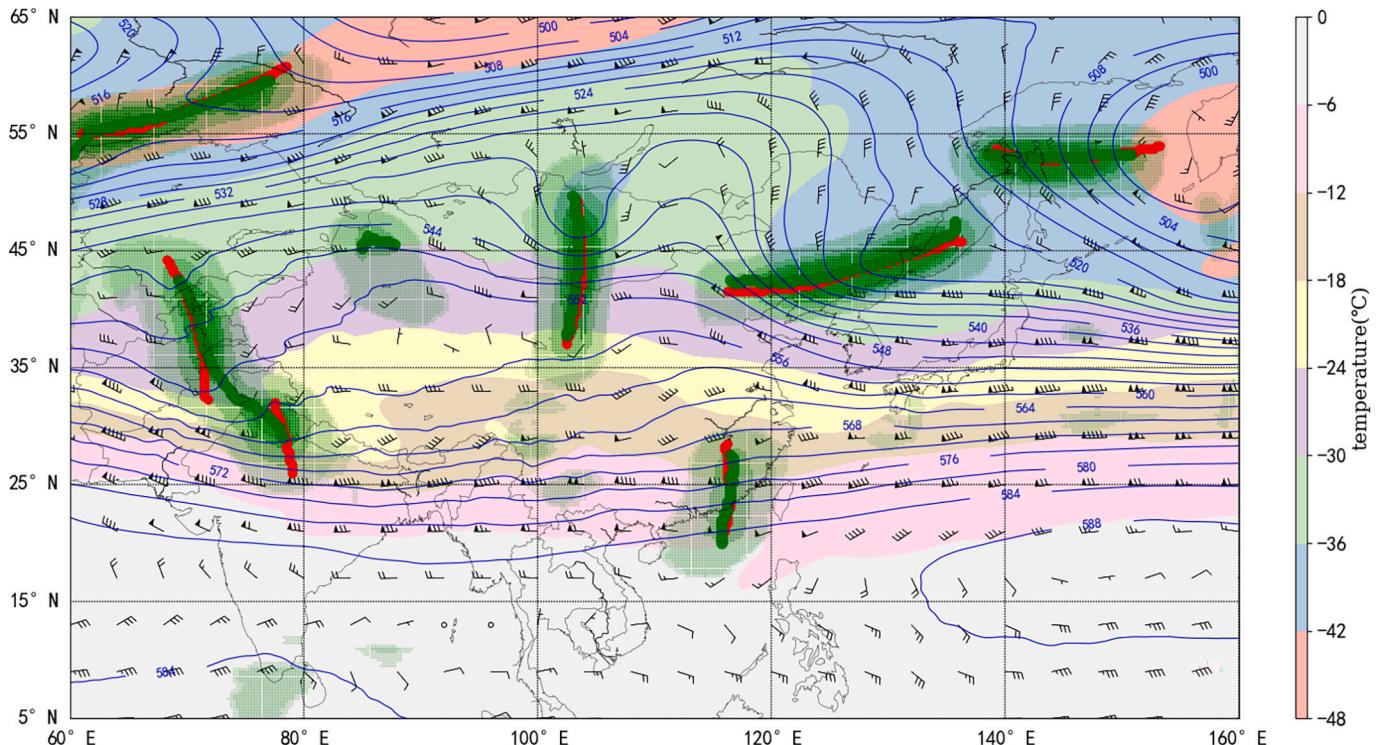


Fig. 9. Automatic trough identification probability map showing weather conditions at 1200 UTC on 22 January 2001. The transparency of the green area of the weather map represents the probability that the area is in trough area, the greater the opacity, the greater the probability that the area is in a trough. The dark green lines are trough lines automatically recognized by proposed method and the red lines are drawn by meteorologists. The temperature colorbar is shown at the right side of the weather map. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

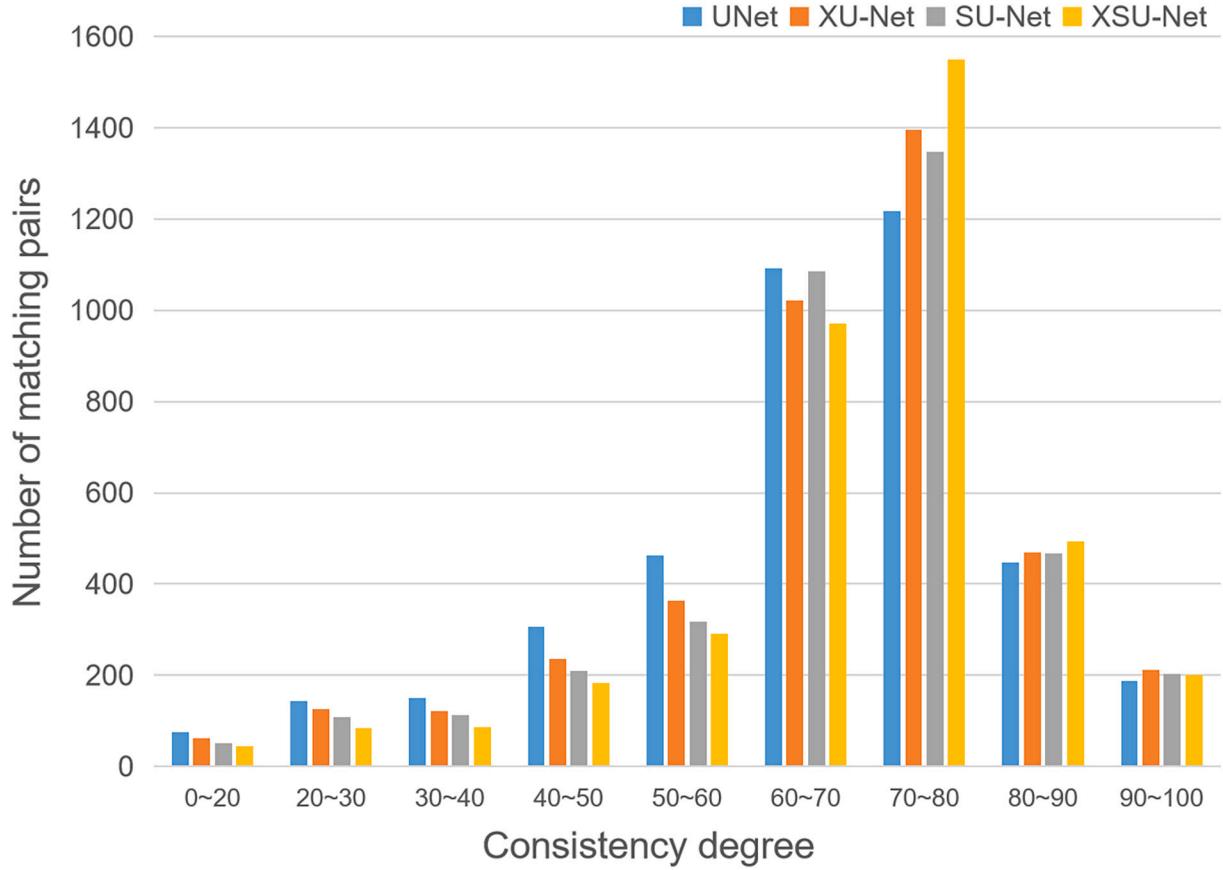


Fig. 10. Statistical results of matching pair consistency when using different network types.

matically recognized data and the manually analyzed data. The quantitative evaluation method (Gong and Li, 2010) is used to evaluate the degree of matching consistency of the two types of trough line:

$$\text{consistency degree} = \left[\frac{N_{os}}{N_s + N_o - N_{os}} \right] \times 100\% \quad (11)$$

where N_{os} is the number of points on the automatically analyzed trough line, N_s is the number of points on the manually analyzed trough line, and N_{os} is the number of points co-located on the two kinds of trough lines simultaneously. If the consistency degree value of the trough line exceeds 60, this is determined as a successful result. In the end, RR (Recall Ratio) and PR (Precision Ratio) are adopted as two quantitative metrics to reflect the effectiveness of trough line automatic extraction:

$$PR = N_s / N_a \quad (12)$$

$$RR = N_s / N_m \quad (13)$$

where N_s represents the number of successfully predicted trough lines, N_m represents the number of all manually marked trough lines, and N_a represents the number of all automatically recognized trough lines.

4. Results

We set a training batch size of 6, and a learning rate of 0.0001. The model was trained for 150 epochs and the Adaptive Moment Estimation (Adam) (Kingma and Ba, 2014) was used for parameter optimization. During the training process, the internal parameters of the model adopt the pre-training model and random definition strategy, and the network model was generated after the training process is completed. According to ablation analysis, element efficiency experiment, and quantitative analysis, we attempt to find the optimum model and demonstrates its

performance improvement over existing models.

4.1. Trough line extraction

After inputting the testing dataset to the network model, a grid marked with the probability that each point is a trough point was generated. When the probability of the point being a trough point is greater than the threshold value of 0.5, it was marked as a trough point in the weather map, at which stage we can extract the trough line using a skeleton extraction algorithm (Au et al., 2008). The comparison before and after the extraction is shown in Fig. 8. Taking the weather map at 1200 UTC 22 January 2001 as an example (Fig. 9), the red lines on the map are manually drawn labels and the green transparency of the overlying grid represents the probability that the point is a trough point; the greater the opacity of the overlay, the greater the trough point probability. The lines marked in dark green are the final extracted trough lines. We observe that there are three low-pressure centers with obvious counter-clockwise circulation and strong cyclonic shear around them. The area with the highest curvature of the isobaric line extending from the low-pressure center and with wind direction change is defined as a trough. The final artificial analysis results confirm seven short wave troughs.

4.2. Ablation study

In order to find a network structure with the best performance on trough recognition, we conducted a series of ablation analysis experiments, namely original UNet, XU-Net in which the encoder is replaced by an Xception network, SU-Net whose decoder adds SE module, and the XSU-Net that combines both of these approaches. Fig. 10 shows the comparison of statistical results of the consistency of matching pairs

Table 1

Comparison of improved identification approaches on the dataset.

Method	ACC	PTOP	PTOT	PNON	P1	RR	PR
UNet	76.21	70.01	68.11	77.90	69.05	76.00	72.11
XU-Net	78.35	72.08	71.04	79.75	71.56	79.91	77.37
SU-Net	79.76	72.82	71.77	81.93	72.29	80.03	79.52
XSU-Net	83.65	76.87	74.58	84.90	75.71	82.96	82.30

when using different network types.

The consistency degree of the trough line identification results obtained by several networks is mainly between 70 and 80%. However, the overall degree of consistency is slightly improved by adding the Xception network and SE module respectively, while the degree of consistency is further improved by incorporating both into the UNet (Fig. 10). As shown in Table 1, for the metrics of trough recognition precision, the XU-Net is around 2% better than UNet, the SU-Net shows around a 3% improvement compared to the base UNet, and when using both Xception

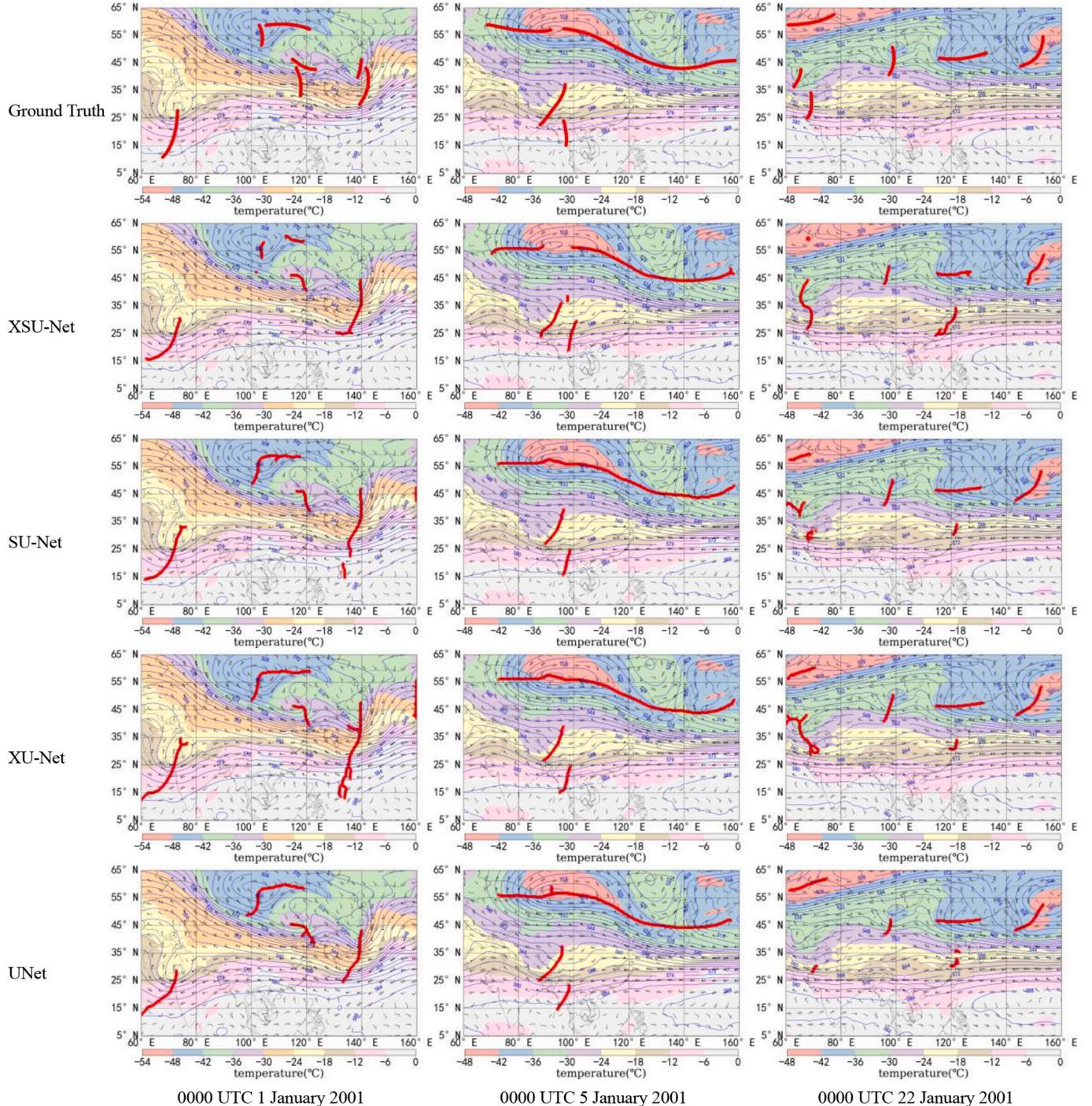


Fig. 11. Comparison of recognition results after adding different modules to UNet. While the red curves in the first row of the weather map are ground truth, and the red curves in the other rows are the trough lines automatically recognized by different networks. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

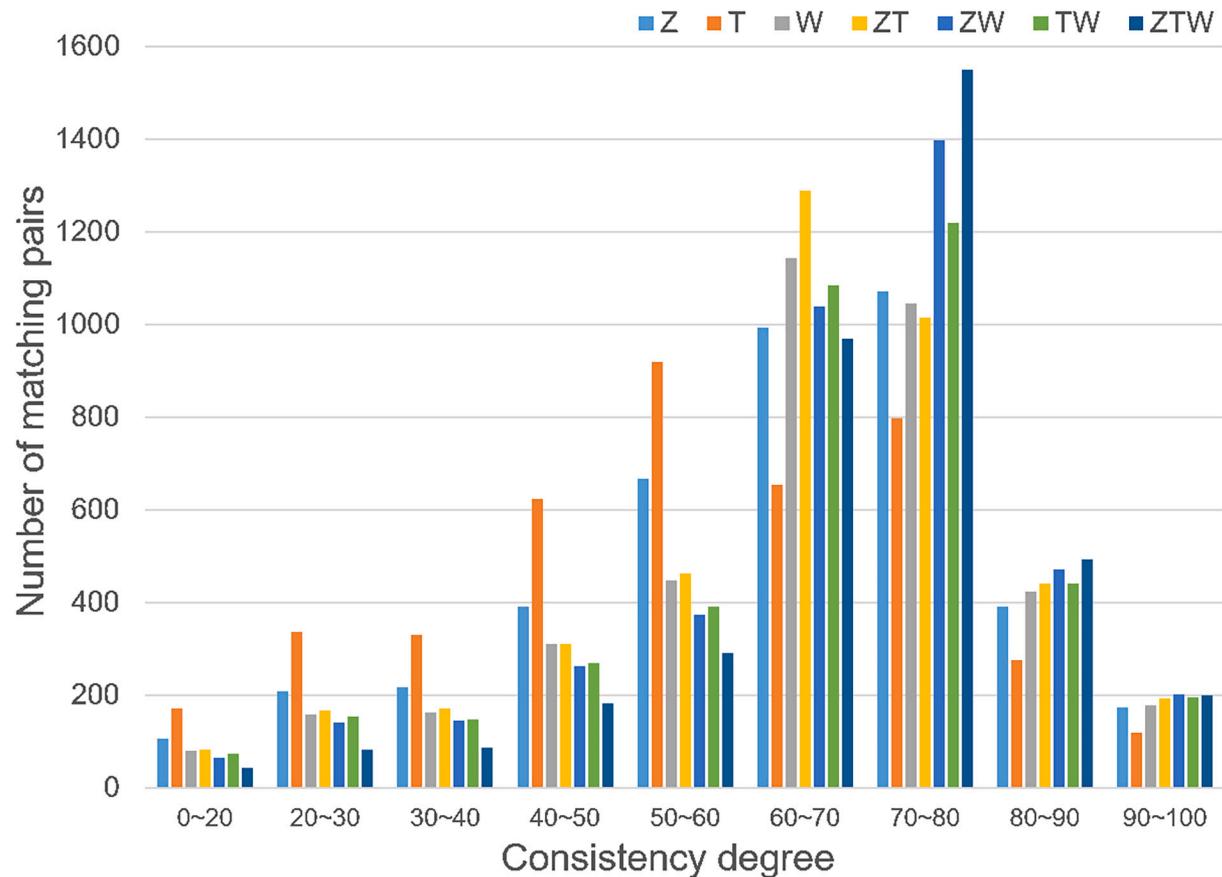


Fig. 12. Matched pair consistency results when using different classes of meteorological elements.

as the encoder and adding the SE module in the decoder, XSU-Net shows around 7% performance improvement over UNet. Comparing the recall rate and the precision rate of the above networks, the network proposed in this paper shows an improvement of ~6–10% over UNet, by about 3–5% compared with the network only using Xception, and by about 2–3% compared with the network with only the SE module. The above experimental results show that the proposed network has the greatest improvement on the performance of trough line recognition compared to the original UNet. The Xception network and the SE module can respectively bring smaller improvements in the recognition accuracy, which proves the rationality of the proposed network structure. Ablation evaluation of the proposed architecture is assessed in Fig. 11, where ground truth and automatic recognition results are visualized on weather maps. We can observe that, the accuracy of the trough lines identified by the original UNet is low, especially when two trough lines are close to each other, there is a phenomenon that two trough lines cross or connect in the automatic identification results like two trough lines in the middle of weather map at 0000 UTC 1 January 2001, which we call crossing phenomenon. The main reason for this phenomenon is that the accuracy of the edge is not high enough when using network to

recognize trough area, resulting in the fusion of the area where the adjacent trough lines are located, and the finally extracted trough lines will also cross. This failure has been improved after the integration of Xception and SE module in UNet, and the overall recognition result is closer to the ground truth.

4.3. Element efficiency

In this section, we evaluated which meteorological element data can improve the efficiency of trough recognition more. We ran a series of experiments using seven combinations of one, two, and three meteorological element data types related to trough formation, namely geopotential height(Z), temperature (T), and wind direction (W). The consistency of statistical results of trough line matching of several different combinations is shown in Fig. 12.

When only the T factor is selected, the consistency of matching pairs is mainly distributed between 50 and 60%. When both Z and T elements are selected or only W elements are selected, the consistency degree of matching pairs mainly concentrates between 60 and 70%, and in other cases, the consistency degree of matching pairs mainly concentrates

Table 2

Comparison of experimental results on datasets composed of different meteorological elements.

Meteorological Elements	ACC	PTOP	PTOT	PNON	P1	RR	PR
Z	68.80	63.62	66.74	67.44	64.99	67.93	63.88
T	42.98	48.01	49.92	42.39	48.95	47.73	43.66
W	71.93	69.65	70.20	73.42	69.92	72.09	70.59
Z, T	73.86	71.72	71.31	76.23	71.51	75.84	71.00
Z, W	80.23	75.89	74.02	81.55	74.94	80.32	75.84
T, W	76.73	74.33	72.07	78.43	73.18	75.96	73.85
Z, T, W	83.65	76.87	74.58	84.90	75.71	82.96	82.30

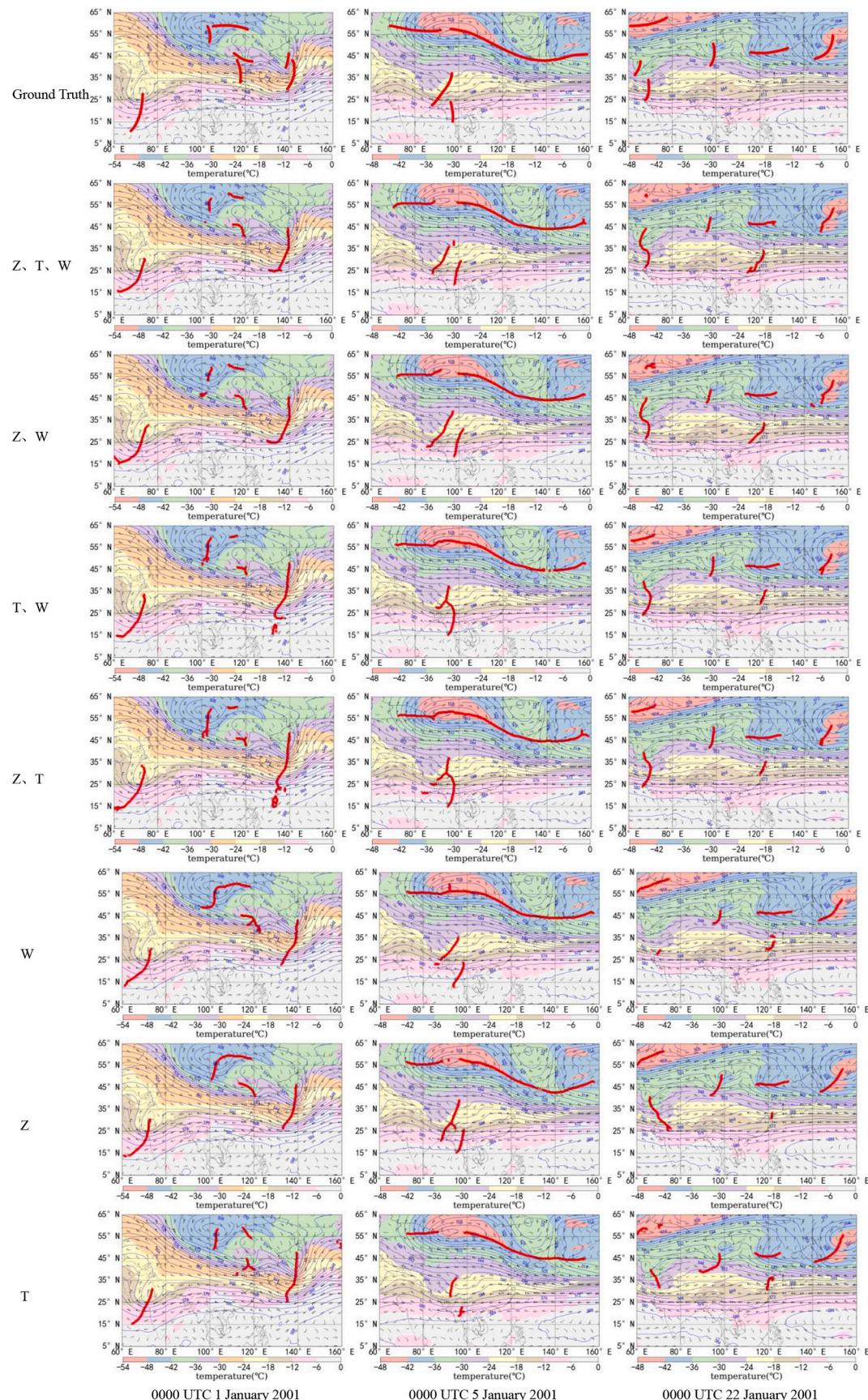


Fig. 13. Segmentation results using the different meteorological elements. While the red curve in the first row of the weather map is the ground truth, and the red curve in the other rows is the trough line automatically recognized by proposed method based on different types of datasets. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

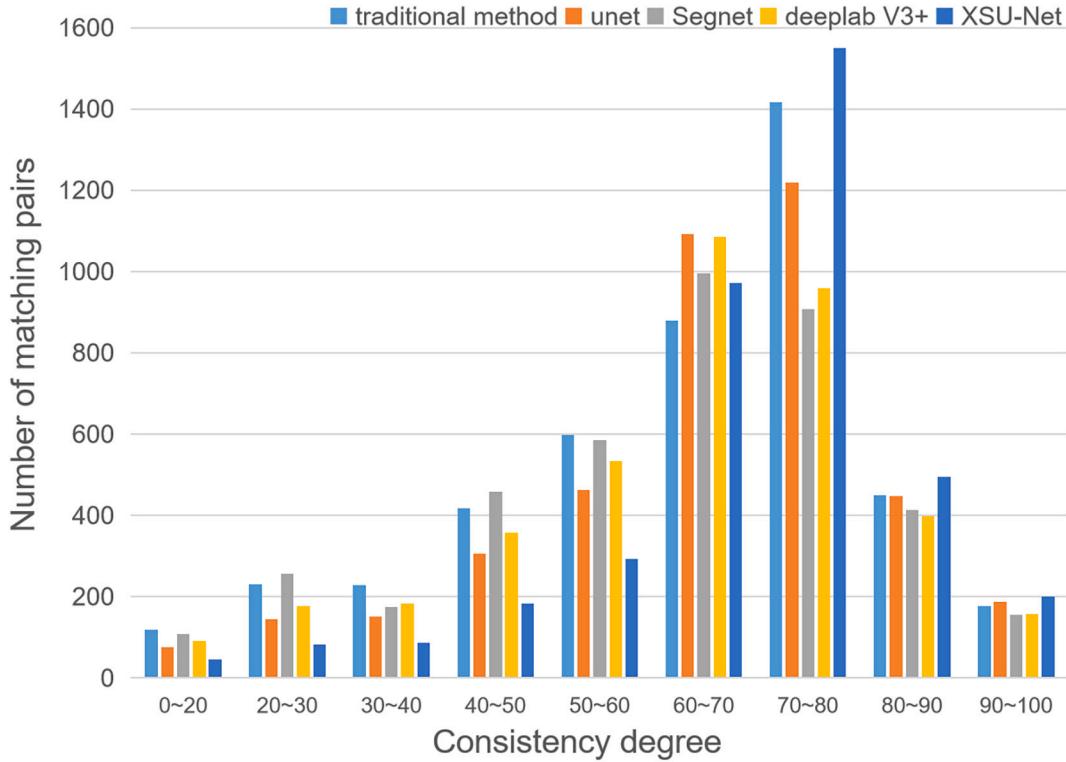


Fig. 14. Matched pair consistency results when using different networks.

between 70 and 80%. By comparing these experimental results, we assume that the T factor has the smallest correlation with trough formation, while the Z and W factors have a greater correlation. When the three factors are integrated, the automatic recognition results and the manual analysis results have the highest similarity (Fig. 12).

Using the test sample sets, we evaluated the recognition results of the network; the comparison of each metric is shown in Table 2. The performance of the trough point recognition is best when using three meteorological elements, with values for all metrics distributed around 80%. This represents an improvement of ~4–40% compared with using only one meteorological element, and a performance improvement of ~2–10% when compared with using two elements. Thus, the performance of the trough point recognition is optimized when using all three data types, i.e., geopotential height, wind, and temperature. In terms of trough line recognition precision, the recall rate of all these methods is generally higher than the precision rate. When using three meteorological elements, the recall rate and precision rate are the highest, the metrics improve by ~3–12% compared with using two elements and ~10–39% compared with using one element. On the basis of these experiments, we suggest that integrating three meteorological elements, namely wind, geopotential height, and temperature, produces the optimum trough line identification performance. We display the ground truth and recognition results in Fig. 13. From figures, it can be concluded that the proposed architecture achieves best on dataset composed of three types of meteorological elements, where it can capture more details of trough lines and the number of trough lines with crossing phenomenon is also the least.

4.4. Quantitative analysis

In order to prove the effectiveness of the proposed network on trough segmentation, we selected several outstanding networks in the field of semantic segmentation, such as UNet, Segnet, deeplabV3+, the network proposed in this paper, as well as the traditional method (Yan et al., 2016) for comparison. The statistical results of trough line matching pair

Table 3
Comparison between existing approaches on our dataset.

Method	ACC	PTOP	PTOT	PNON	P1	RR	PR
Traditional method	–	–	–	–	–	75.39	64.72
UNet	76.21	69.08	67.04	77.75	68.04	74.18	72.71
Segnet	62.36	57.32	54.77	62.93	56.02	64.27	61.03
deeplabV3+	65.20	59.61	57.72	66.38	58.64	67.08	65.97
XSU-Net	83.65	76.87	74.58	84.90	75.71	82.96	82.30

consistency in each group of experiments are shown in Fig. 14.

The network matching pair proposed in this paper outperforms all others in terms of consistency with most consistency results between 70 and 80%. Among the other methods, the matching pair consistency between the traditional method and UNet is relatively high on the whole, with the range is also mainly concentrated in the range of 70 to 80% consistency (Fig. 14).

The experimental results (Table 3) show that the proposed network has the best recognition performance on the dataset, with a 7–8% performance increase over the original UNet network, and 19–22% improvement over Segnet. Compared with the deeplabV3+ network, the metrics are increased by about 16–19%. The data shows that the UNet approach has more advantages in trough point recognition, while the proposed combined network approach has better performance. UNet tends to exceed the performance of both Segnet and deeplabV3+ based on the metrics, while our proposed network shows a further substantial improvement compared to UNet. The results show that the UNet, which uses an encoder-decoder architecture, is more suitable for trough line recognition; we assume that this is mainly due to continuous integration of the characteristics of the upper and lower layers in UNet. Furthermore, our results show that the methods based on DCNN perform better than traditional methods, demonstrating that the proposed approach in this paper is logical from a performance perspective. Quantitative evaluation of the proposed method is shown in Fig. 15. We can see that

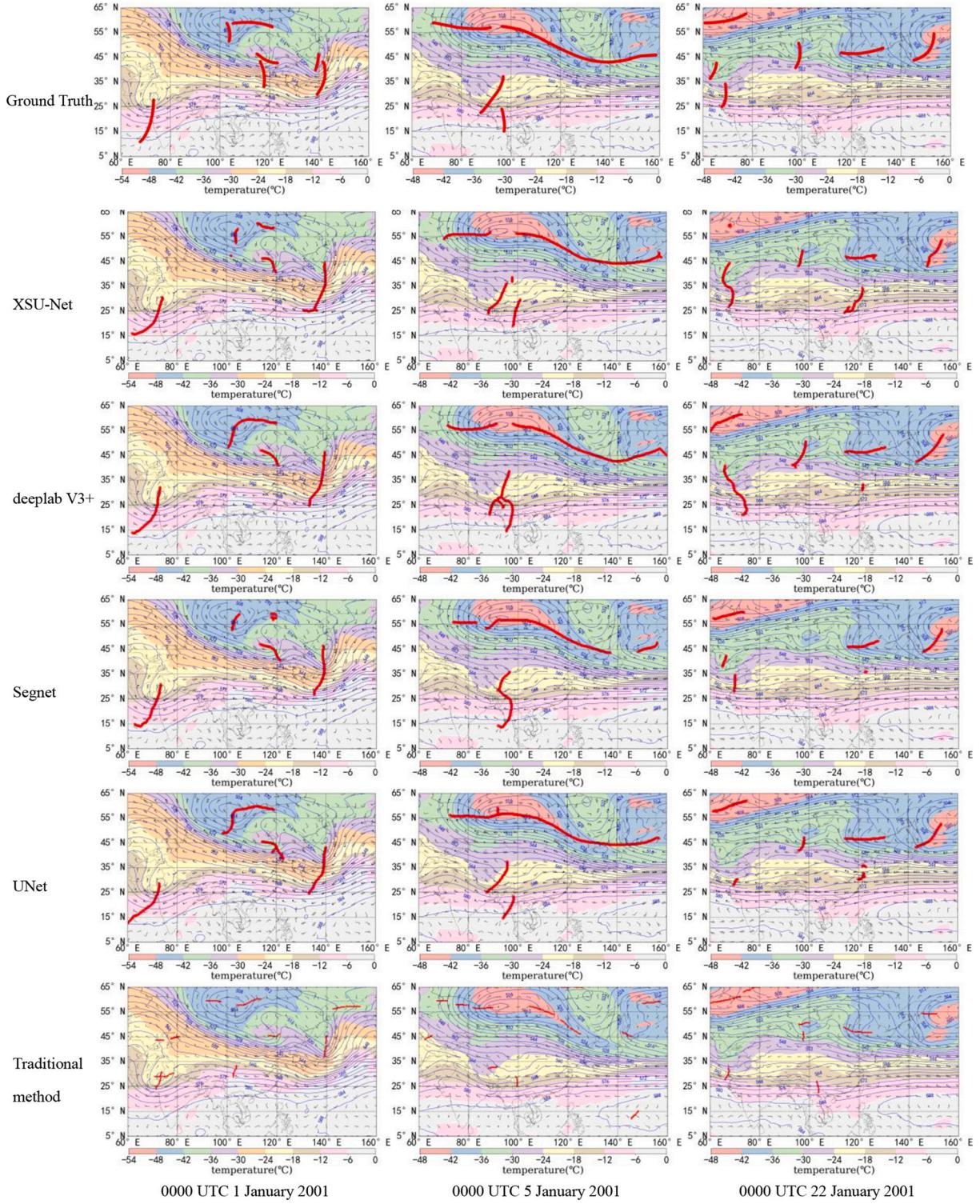


Fig. 15. Comparison of trough recognition results of different methods. While the red curves in the first row of the weather map are ground truth, and the red curve in the other rows is the trough line automatically recognized different methods. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

that the results obtained by the proposed network is more similar to ground truth as a whole. The proposed method can detect weak trough lines or trough lines that are close to each other which may be lost in other methods, thus, it is more powerful to preserve more details. However, the automatic recognition results obtained by our method do not perfectly align with that of ground truth (first row), for example, the crossing phenomenon occurs in weather map at 0000 UTC 1 January

2001 obtained by our XSU-Net, which shows that the recognition accuracy at the edge of the trough line by our method, is still not desirable.

5. Discussion and conclusions

Trough lines can reflect the trend of weather changes and provide an important reference for forecasters. At present, the analysis and

recognition of trough lines in weather maps is still mainly completed by manual analysis, which leads to a waste of resources for weather forecasting. Moreover, it is also difficult for traditional automatic trough line identification methods to set rules covering all the characteristics of trough lines, and their generalization ability is low. This paper proposes a method based on the improved UNet for automatic recognition of trough lines, in which the Xception network structure and attention mechanism are incorporated to effectively reduce the overall network parameters and improve the accuracy of feature extraction. Our study demonstrates that the proposed method based on deep learning can overcome the traditional disadvantages of the rule-based method, and effectively improve the accuracy of segmentation. The validity of the network is verified by a test on meteorological data (500 hPa) from the European Centre for Medium-Range Weather Forecasts.

However, there are still limitations worth consideration, which are left as possible future work. Firstly, the proposed model occasionally fails when the two trough lines are close to each other, which leads to the situation that the trough lines obtained after skeleton line extraction are connected or crossed, whereas the manually drawn trough lines are not allowed to be crossed. This mainly results from the fact that we use the neural network to identify the trough area first, and then use the skeleton line extraction algorithm to extract the skeleton line, resulting in the error superposition. It remains challenging to determine how to ensure that the trough area automatically identified by the network does not deviate laterally. Secondly, the training network requires a large number of manual labels. As these labels are generated by hand, there are likely to be some errors in these manually labeled data caused by subjectivity, which has a significant impact on network training. Data labeling work also requires a significant amount of human resources, so in future work, consideration should be given to how to use a small number of labels to train the network and further improve the automatic recognition method.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

This research was funded by the National Key Research and Development Program of China (Grant No. 2018YFC1507604) the National Natural Science Foundation of China (Grant No. 42075139, 41305138), the China Postdoctoral Science Foundation (Grant No. 2017M621700) and Hunan Province Natural Science Fund (2021JJ30773).

References

- Allen, G.L., Cowan, C.R.M., Power, H., 2006. Acquiring information from simple weather maps: Influences of domain-specific knowledge and general visual-spatial abilities. *Learn. Individ. Differ.* 16 (4), 337–349. <https://doi.org/10.1016/j.lindif.2007.01.003>.
- Au, O.K.-C., Tai, C.-L., Chu, H.-K., Cohen-Or, D., Lee, T.-Y., 2008. Skeleton extraction by mesh contraction. *ACM Trans. Graph.* 27 (3), 1–10. <https://doi.org/10.1145/1360612.1360643>.
- Badrinarayanan, V., Kendall, A., Cipolla, R., 2017. Segnet: a deep convolutional encoder-decoder architecture for image segmentation. *IEEE Trans. Pattern Anal. Mach. Intell.* 39 (12), 2481–2495. <https://doi.org/10.1109/TPAMI.2016.2644615>.
- Bandara, S., Marshall, T., Karunarathne, S., Stolzenburg, M., 2021. Groups of narrow bipolar events within thunderstorms. *Atmos. Res.* 252, 105450. <https://doi.org/10.1016/j.atmosres.2021.105450>.
- Biard, J.C., Kunkel, K.E., 2019. Automated detection of weather fronts using a deep learning neural network. *Adv. Stat. Climatol. Meteorol. Oceanogr.* 5 (2), 147–160.
- Bluestein, H.B., Speheger, D.A., 1995. The dynamics of an upper-level trough in the baroclinic westerlies: Analysis based upon data from a wind profiler network. *Mon. Weather Rev.* 123 (8), 2369–2383. [https://doi.org/10.1175/1520-0493\(1995\)123<2369:TDOTWF>2.0.CO;2](https://doi.org/10.1175/1520-0493(1995)123<2369:TDOTWF>2.0.CO;2).
- Chollet, F., 2017. Xception: deep learning with depthwise separable convolutions. *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.* 1251–1258.
- Cui, Y., Jia, M., Lin, T.Y., Song, Y., Belongie, S., 2019. Class-balanced loss based on effective number of samples. In: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, pp. 9268–9277.
- Dai, X., Li, Q., Gu, D.Q., 2016. Automatic recognition of meteorological trough lines based on Douglas-Peucker algorithm. *Meteorol. Sci. Technol.* 01 <https://doi.org/10.3969/j.issn.1671-6345.2016.01.007>.
- Farahnak-Ghazani, F., Baghshah, M.S., 2016. Multi-label classification with feature-aware implicit encoding and generalized cross-entropy loss. In: 2016 24th Iranian Conference on Electrical Engineering (ICEE). IEEE, pp. 1574–1579.
- Feteris, P.J., 1973. The role of deep convection and strong winds aloft in triggering gales over the Persian Gulf: comparative case studies. *Mon. Weather Rev.* 101 (5), 455–460. <https://doi.org/10.1021/ja01164a088>.
- Finney, D.J., 1952. *Probit Analysis: A Statistical Treatment of the Sigmoid Response Curve*. Cambridge University Press, Cambridge.
- Fröhlich, L., Knippertz, P., 2008. Identification and global climatology of upper-level troughs at low latitudes. *Meteorol. Z.* 565–573 <https://doi.org/10.1127/0941-2948/2008/0320>.
- Gong, Y., Li, J., 2010. The quantified verification method of synoptic meteorology and the application to AREM. *Scientia Meteorol. Sinica* 30 (6), 763–772. <https://doi.org/10.3788/HPLPB20102207.1462>.
- Hu, F., Xia, G.-S., Hu, J., Zhang, L., 2015. Transferring deep convolutional neural networks for the scene classification of high-resolution remote sensing imagery. *Remote Sens.* 7 (11), 14680–14707. <https://doi.org/10.3390/rs71114680>.
- Hu, J., Shen, L., Sun, G., 2017. Squeeze-and-excitation networks. *IEEE Trans. Pattern Anal. Mach. Intell.* 99 <https://doi.org/10.1109/TPAMI.2019.2913372>.
- Huang, X., Zhao, F., 2000. Relation-based aggregation: finding objects in large spatial datasets. *Intell. Data Anal.* 4 (2), 129–147. <https://doi.org/10.3233/IDA-2000-4204>.
- Jiang, Y., Chen, L., Zhang, H., Xiao, X., 2019. Breast cancer histopathological image classification using convolutional neural networks with small SE-ResNet module. *PLoS One* 14 (3), e0214587. <https://doi.org/10.1371/journal.pone.0214587>.
- Juckles, M., Smith, R.K., 2000. Convective destabilization by upper-level troughs. *Q. J. R. Meteorol. Soc.* 126 (562), 111–123. <https://doi.org/10.1002/qj.49712656206>.
- Kingma, D.P., Ba, J., 2014. Adam: a method for stochastic optimization. *arXiv preprint arXiv:1412.6980*.
- Lagerquist, R., McGovern, A., Gagne II, D.J., 2019. Deep learning for spatially explicit prediction of synoptic-scale fronts. *Weather Forecast.* 34 (4), 1137–1160. <https://doi.org/10.1175/WAF-D-18-0183.1>.
- Li, Q., Huang, Y., Fan, Y., Tang, S., Hu, Y., 2018. Automatic Analysis of Trough Lines based on Curvature Tracing. *Atmosphere* 9 (3), 88. <https://doi.org/10.3390/atmos9030088>.
- Lin, T.-Y., RoyChowdhury, A., Maji, S., 2017. Bilinear convolutional neural networks for fine-grained visual recognition. *IEEE Trans. Pattern Anal. Mach. Intell.* 40 (6), 1309–1322. <https://doi.org/10.1109/TPAMI.2017.2723400>.
- Liu, X., Deng, Z., Yang, Y., 2019. Recent progress in semantic image segmentation. *Artif. Intell. Rev.* 52 (2), 1089–1106. <https://doi.org/10.1007/s10462-018-9641-3>.
- Long, J., Shelhamer, E., Darrell, T., 2015. Fully convolutional networks for semantic segmentation. *IEEE Trans. Pattern Anal. Mach. Intell.* 39 (4), 640–651. <https://doi.org/10.1109/CVPR.2015.7298965>.
- Mahbod, M., Rafiee, M.R., 2021. Trend analysis of extreme precipitation events across Iran using percentile indices. *Int. J. Climatol.* 41 (2), 952–969. <https://doi.org/10.1002/joc.6708>.
- Mou, L., 2010. *Research and Application of Trough Line Identification Based on Ant Colony System Theory*. Wuhan University of Technology.
- Neu, U., Akperov, M.G., Bellenbaum, N., Benestad, R., Blender, R., Caballero, R., Cocozza, A., Dacre, H.F., Feng, Y., Fraedrich, K., 2013. IMILAST: a community effort to intercompare extratropical cyclone detection and tracking algorithms. *Bull. Am. Meteorol. Soc.* 94 (4), 529–547. <https://doi.org/10.1175/BAMS-D-11-00154.1>.
- Ronneberger, O., Fischer, P., Brox, T., 2015. U-net: Convolutional networks for biomedical image segmentation. In: International Conference on Medical Image Computing and Computer-assisted Intervention. Springer, Cham, pp. 234–241.
- Sawyer, J., 1956. Principles of Meteorological Analysis. Walter. J. Royal Aeronaut. Soc. 60 (544), 285–286. <https://doi.org/10.1017/S0368393100135291>.
- Simmons, A.J., Uppala, S.M., Dee, D.P., Kobayashi, S., 2007. ERA-Interim: New ECMWF reanalysis products from 1989 onwards. *ECMWF Newslett.* 110, 25–36.
- Sultana, F., Sufian, A., Dutta, P., 2020. Evolution of image segmentation using deep convolutional neural network: a survey. *Knowl.-Based Syst.* 201, 106062.
- Taghanaki, S.A., Abhishek, K., Cohen, J.P., Cohen-Adad, J., Hamarneh, G., 2021. Deep semantic segmentation of natural and medical images: a review. *Artif. Intell. Rev.* 54 (1), 137–178. <https://doi.org/10.1007/s10462-020-09854-1>.
- Tan, X., 2006. Data warehouse and its potential in weather forecast. *J. Appl. Meteorol. Sci.* 17 (3), 325–330. <https://doi.org/10.1360/aps040087>.
- Ulbrich, U., Leckebusch, G.C., Grieger, J., Schuster, M., Akperov, M., Bardin, M.Y., Feng, Y., Gulev, S., Inatsu, M., Keay, K., 2013. Are greenhouse gas signals of Northern Hemisphere winter extra-tropical cyclone activity dependent on the identification and tracking algorithm? *Meteorol. Z.* 22 (1), 61–68. <https://doi.org/10.1127/0941-2948/2013/0420>.
- Wallace, J.M., Hobbs, P.V., 2006. *Atmospheric Science: An Introductory Survey*. Elsevier.
- Wong, K.Y., Yip, C.L., Li, P.W., 2008. Automatic identification of weather systems from numerical weather prediction data using genetic algorithm. *Expert Syst. Appl.* 35 (1–2), 542–555. <https://doi.org/10.1016/j.eswa.2007.07.032>.
- Yan, C., Hu, W., Zhang, C., Wang, S., Yang, Y., Ren, H., Lv, C., 2016. Auto-analysis of trough system at 500 hPa based on Gradient algorithm. *J. Appl. Meteorol. Sci.* 27 (6), 741–749. <https://doi.org/10.11898/1001-7313.20160611>.
- Yu, L., Hu, Z., 2011. Subsystem for Interactively drawing Synoptic Chart in MICAPS. *J. Appl. Meteorol. Sci.* 22 (3), 375–384. <https://doi.org/10.1007/s11589-011-0776-4>.