# APPLICABILITY OF DETECTING SURFACE WEATHER FRONTS USING CONVOLUTIONAL NEURAL NETWORK BASED DEEP LERANING

Yiwen MAO[1], Tomohito J. YAMADA[2]

[1]Member of JSCE, Graduate School. of Eng., Hokkaido University
(N13-W8, Kita-ku, Sapporo 060-8628, Japan) (Corresponding Author)
E-mail: ymaopanda@gmail.com
[2]Member of JSCE, Professor, Faculty of Eng.., Hokkaido University
(N13-W8, Kita-ku, Sapporo 060-8628, Japan)
E-mail: tomohito@eng.hokudai.ac.jp

A deep learning model using a convolutional neural network (CNN) and U-net is built to study the applicability in detecting surface weather fronts in Japan and surrounding sea. First, a CNN model is used to predict whether there are fronts (1) or not (0) in the region. If CNN outputs 1, frontal locations can be predicted by assembling binary classification of front/no front at each grid point within the region using a U-net. The predictability of CNN/U-net stems from their ability to match locations of outstanding horizontal gradients with observed frontal locations. Our study shows that CNN can achieve higher accuracy in filtering out no front cases by using fewer predictors than the U-net. Overall, the predictability of frontal locations varies with season and differs when stationary fronts are present or not, which suggests that dynamics related to seasonal frontogenesis can influence the spatial distribution of predictability of fronts.

***Key Words :*** *Weather fronts, Machine learning, Deep learning, Convolutional neural network, U-net*

## 1. INTRODUCTION

About 90% of extreme precipitation in the midlatitudes can be associated with frontal boundaries[1)]. It is desirable to achieve objective detection of surface weather fronts without relying on subjective determination by meteorologists[2)]. One common approach is to automatically detect fronts based on diagnostic parameters such as the thermal front parameter (TFP) and the F index[3)]. The shapes of the detected fronts by diagnostic parameters based on different variables often vary[4)].

As artificial intelligence (AI) advances, an alternative methodology based on deep learning for image processing has been explored for the automatic detection of surface fronts. Deep learning can incorporate many variables as inputs, which has the potential to have better predictability in frontal locations than that output by diagnostic parameters which depend on a small number of parameters. In general, deep learning refers to machine learning algorithms that use multiple layers of non-linear transformation to derive features from input fields[5)]. There have been a few successful methodologies for detecting surface weather fronts in this category. For example, a deep learning (DL) front algorithm is built by convolutional neural networks for detecting fronts in North America[6)]. U-net (one type of convolutional neural network) is applied in real-time identification of fronts over United States for operational forecasting[7)].

U-net is primarily developed for medical image segmentation[8)]. Although similarity exists between the identifications of biological features on medical images and surface weather fronts (i.e., both are boundaries), atmospheric data differ from medical images in terms of multidimensionality (i.e., fronts impact multiple weather variables). There has been no consensus on how effective U-net can be in identifying surface weather fronts. The goal of developing AI based detection of fronts is to achieve automation of such tasks, which can be applied to large ensemble climate model datasets. Therefore, it is necessary to know why AI methods, such as U-net can work well in identifying surface weather fronts.

In this study, besides building a deep learning algorithm using convolutional neural network based methods (e.g. CNN and U-net) for detecting surface

weather fronts in Japan and surrounding sea, we also shed some light on where the predictability of CNN/U-net originates, thereby elucidating the applicability of the deep learning based methodology for detecting surface weather fronts.

## 2. DATA AND METHODOLOGY

### (1) Predictands and Predictors

The area within 20N-50N and 120E-150E is chosen as the region for frontal detection. Due to limitation of available data, we only consider the months from June to October of 2003 to 2019 for model training and test. The ground truth (i.e. predictands) is obtained from the gridded locations of daily fronts at 00UTC extracted from weather charts archived by Japan Meteorological Agency (JMA)[9]. Overall, there are a total of 2600 days considered in this study, and out of which 2231 days have fronts. To ensure continuous representation of fronts, the ground truth dataset marks the locations of fronts at a resolution of $1° \times 1°$ [9]. The types of fronts (e.g. cold/warm fronts) cannot be distinguished at each grid point. However, based on detecting if the symbols of stationary fronts are present on the original weather chart, we further identify that stationary fronts are present on 1873 days out of the 2231 days with fronts, and the remaining 358 days have mobile fronts only.

In operational meteorology, fronts are identified visually based on the spatial coincidence of several localized features, such as trough, shift in winds, gradients in air temperature and moisture. The key element of frontal identification is the gradient of thermodynamic parameters in horizontal directions. Therefore, we only consider horizontal gradients of variables as predictors. The predictors are chosen to represent the categories used in operational meteorology (i.e., changes in pressure, wind, heat and moisture).

We consider the horizontal (zonal and meridional) gradients of the following variables as predictors: geopotential height, relative vorticity, equivalent potential temperature, and dew point temperature at 925 mb, 900 mb, 850 mb and 500 mb. In addition, relative vorticity is also included as predictors, and they are derived from the Meso-scale models (MSM) corresponding to the same time of the ground truth (00 UTC) of the predictand. In general, there is no consensus on the most effective variables for predictors. For example, besides dew point temperature, mixing ratio and specific humidity can also be used to represent moisture. However, we choose dew point temperature in this study because the horizontal gradients of dew point temperature are more common to use in detecting surface weather front in operational meteorology than the other two.

The 2D fields of predictors are sampled at $1° \times 1°$. Although coarse, $1° \times 1°$ should be sufficient to determine if fronts are present or not in the region, given that the width of the surface weather fronts in the gridded dataset is 200 km with a resolution of $1° \times 1°$ [9]. Due to the edge effects of calculating the gradients, only $24 \times 24$ grid points are considered for frontal detection. The first and last 3° in both longitude and latitude directions either have no values for gradient or no values for vorticity.

### (2) The model

A two-stage model (Fig.1) is developed to predict the location of fronts. First, CNNs are used to predict if there are fronts in the region (output 1 or 0). If CNN predicts 1 (front in the region), a U-net is used to identify the locations of fronts by output 1 or 0 at each grid point within the region. The key mathematical operation for both CNN and U-net is convolution/up-convolution, which can be expressed as

$$Y(i,j) = \sum_{m=0}^{n_1-1} \sum_{n=0}^{n_2-1} W\, X(Si+m, Sj+n), \quad (1)$$

where $X$ is a $n_1 \times n_2$ input matrix, $W$ is a $m \times n$ kernel matrix (i.e., filter), $S$ is defined as the stride (i.e., the spatial interval of the convolution). The output $Y$ is a matrix (i.e., feature maps), and $(i,j)$ indicates the grid point of $Y$. If $S \geq 1$, Eq (1) represents a downsampling process (i.e., convolution for decreasing spatial resolution, that is, the size of $Y$ is reduced from $X$). If $S < 1$, Eq (1) represents an upsampling process (i.e., up-convolution for increasing spatial resolution, that is, the size of $Y$ is increased from $X$). $W$ is the weights of the model, and $W$ needs to be learned through training if Eq (1) is applied in machine learning (e.g., CNN and U-net).

#### a) CNN

CNNs are used to condense features from input 2D spatial fields to a label[10]. One building block of a CNN consists of convolution (i.e. $S \geq 1$ in Eq.1), activation, and maxpooling (for further downsampling). Multiple kernels ($W$) can be applied in one building block to increase CNN's ability to recognize spatial features. In our study, 25 and 50 $3 \times 3$ kernels are used for the two building blocks respectively as shown in Fig.1a. After convolution, the activation function (rectifier function ReLU) is applied to the output of Eq (1). The rectifier function ReLU is a common activation function, it can be defined as $f(x) = x\ for\ x > 0$ and 0 otherwise. Nonlinearity is

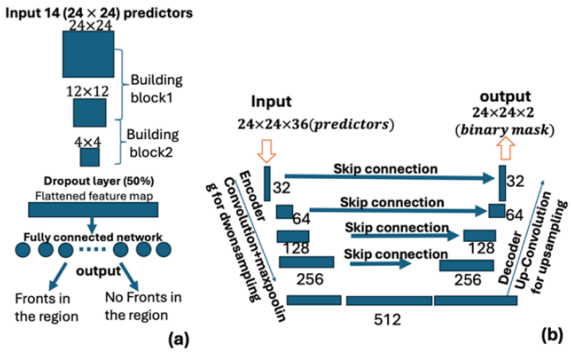


**Fig. 1** (a): The schematic of CNN used for binary classification of fronts or not in the study region. A building block refer to three layers of operation: convolution, activation and maxpooling. (b): The schematic of U-net used for detecting location of fronts in the study region

added to the convolution process by ReLU. Finally, the size of feature maps output by Eq (1) is reduced by maxpooling, that is, resampling before being passed to the next building block. Specifically, in each nonoverlapping subregion (e.g. 2×2) pre-defined by the user, the maximum value is chosen to represent the subregion.

The building block can be repeated several times (i.e., deep learning). The flattened feature maps output from the last building block will be input to fully connected layers for classification. To prevent overfitting, 50% random removal of the output feature map of the last building block is implemented as the dropout layer (Fig.1) before the fully connected layer.

The fully connected layer refers to one or multiple layers of neurons (arranged like 1 dimensional vector), and each neuron (i.e., computational units with weights and bias obtained by training) is connected to every neuron in the previous layer. The output of the layer will also go through an activation function before making the classification of whether there is front or no front in the entire region.

We only select 14 of the initial 36 predictors for the CNN model. Specifically, the importance of each predictor is tested by replacing the horizontal gradients of each variable with random numbers, and the CNN model with the modified predictors is trained again. If the metric of predictability after the modification is better than or approximately the same as the original model, the predictor is not useful, and it will be removed. Eventually, the 14 predictors used for CNN are zonal and meridional gradients of the following quantities: equivalent potential temperature at 850mb and 500 mb, dew point temperature at 850 mb, geopotential height at 925 mb, 850 mb, 900 mb, and 500 mb. The architecture of the model is shown in Fig.1(a).

In general, vorticity at all levels are not useful predictors for the CNN model. However, in some frontal cases, vorticity can still help to indicate the locations of fronts. In other words, large values of gradients of vorticity do not always indicate fronts, but when fronts are present, they often impact vorticity. Due to the noisy nature of atmospheric variables (e.g., vorticity), it would be ideal to use a relatively simpler CNN model to filter out no front cases before applying the U-net (with more predictors) to predict the location of fronts. In this way, we can reduce the number of irrelevant cases (i.e. no front) for training the U-net for frontal boundary identification, in analogous to image segmentation, which is the original purpose of U-net.[8)]

**b) U-net**

The architecture of the U-net consists of two parts: an encoder and a decoder. The encoder part has essentially the same structure as a CNN without the fully connected layer. The feature map resulting from the encoder will undergo decoder layers for restoring the location information. The building block of a decoder consists of up-convolution (for increasing the spatial resolution of feature maps) and skip connection (for passing location information from the corresponding layers of the encoder). The decoder part undergoes the same number of the building blocks and kernels of the encoder. We use 3 building blocks with 32, 64 and 128 3 × 3 kernels represented by $W$ in Eq (1) for the encoder and decoder respectively (Fig.1b). All 36 variables listed in section 2(1) are used as predictors.

**c) Model implementation**

In this study, CNN is used for imbalanced prediction as the ratio of fronts and no fronts in the study region is 2231 to 369. However, we do not consider undersampling or oversampling to address the imbalance because without extra data processing to address the imbalance, the predictability of the minority class is good enough to meet the requirement of effectively filtering out no front cases. For both CNN and U-net, the Adam algorithm is used as the optimizer. Since the final stage for both CNN and U-net is binary classification, the loss function is defined as

$$L = -\sum_{i=1}^{n} y_i \log(p_i), \qquad (2)$$

where $y_i$ denotes the label (0 or 1), $p_i$ denotes the predicted probability for $y_i$, and n=2 for binary classification [8)].

The hyperparameters used to tune the CNN and U-net are number of filters, learning rate, batch size, and

regularization rate. Grid search is used to find the optimal parameter for tuning models. In general, adding more building blocks (i.e., network layers) to either CNN and U-net cannot increase their predictive skills, and both CNN and U-net are not sensitive to the number of filters. Starting the first layer of a U-net with 16 filters, 32 filters and 64 filters (Fig.1b) generally have no evident influences on its final predictability. Moreover, L2 regularization is used for CNN and U-net to prevent overfitting. The learning rate, batch size and regularization rate adopted for the CNN model are 0.001, 64, and 0.01 respectively, and for U-net they are 0.0005, 32, and 0.0001.

**(3) Model training, test and evaluation**

All predictors for both CNN and U-net are preprocessed by normalizing them into the range of -1 to 1. We use leaving-one-year-out cross validation for model training and test. For each round of training, we also reserve 2 years of data for early stopping as a validation method to prevent overfitting. There are a total of 2600 cases for training and test. The model architectures for CNN and U-net in Fig.1 require 213552 and 1935138 parameters respectively. On average, the number of cases for each round of the CNN model training is 2150 cases after removing 3 years of data for test and validation, and 1837 cases for the U-net which is trained using only days of fronts. Therefore, the ratio of the size of training data to model parameters is $2150 \times 24\ (Lat) \times 24(lon) \times 14(predictors)/231552 \approx 81$ for CNN, and $1837 \times 24\ (Lat) \times 24(lon) \times 36(predictors)/1935138 \approx 19$ for U-net. The values suggest the training data can be considered sufficient because as a rule of thumb, the ratio of training data size to number of parameters should be at least 10.

The predictive skills of both CNN and U-net are evaluated using metrics based on a $2 \times 2$ contingency table for binary classifications (Table 1). For the CNN, the evaluation is on the output of 1 (Front) or 0 (No Front) for the entire region, and for the U-net, the evaluation is on every grid point with some adjustments. Specifically, a hit is counted if the grid point predicted as 1(Front) is within a $1° \times 1°$ box centered at an observed frontal grid point. In this way, the shape of the front can be obtained by assembling the binary classification at all grid points.

Since both the CNN and U-net are used to classify imbalanced classes, accuracy $ACC = \frac{A+D}{A+B+C+D}$ is not sufficient to evaluate the predictive skills of the model. The following metrics are considered in this study: probability of detection $POD = \frac{A}{A+C}$, false alarm ratio $FAR = \frac{B}{A+B}$, success ratio $SR = \frac{A}{A+B}$, critical success index $CSI = \frac{A}{A+B+C}$, and $bias = \frac{POD}{SR}$. POD quantifies the fractional success of predicting the target class out of all occasions when the target class is observed. FAR quantifies the number of times the target class is falsely predicted out of all predictions of the target class. SR and FAR are linked by SR = 1-FAR. CSI quantifies the fractional success of the target class prediction out of all target class prediction and missed prediction of the target. CSI is a more balanced metric than POD and SR[11]. Bias quantifies whether the classification model tends to predict the target class more (bias>1) or less (bias<1) often than it is observed.

Finally, to investigate the efficiency of the two-stage methodology of involving both CNN and U-net for predicting frontal locations, we compare the two-stage model with just using one U-net model to predict the location of fronts without using CNN to filter out no front cases first.

**Table 1** Contingency table for binary classification

| | Observe1 | Observe 0 |
|---|---|---|
| Predict 1 | A (hit) | B (false alarm) |
| Predict 0 | C (miss) | D (correct rejection) |

## 3. RESULTS

**(1) Overall predictability of frontal detection**

Figure 2 shows the predictability of using CNN to predict if there are fronts or no fronts in the study region in terms of POD, SR, CSI, and bias averaged over all test years from 2003 to 2019. The results by CNN are compared with applying a U-net model to predict frontal locations, and covert the U-net prediction to a binary series of 0 (no front in the region) or 1(fronts in the region). The average predictability by CNN is close to an ideal model, and evidently better than that of U-net as the CSI for predicting 0 (i.e. the minority class) by U-net is around 0.7, about 25% lower from that of CNN. The results indicate that the CNN model is effective in filtering out no front cases in the study region before using U-net to locate fronts, but applying only U-net to predict the locations of fronts can also achieve good predictive skills as suggested by Fig.2. One potential advantage of applying the CNN model is that the CNN model is less complex than U-net. Also, the noise of the training dataset for U-net may be reduced by filtering out data without boundaries (i.e., fronts).

Figures 3 and 4 show the spatial distribution of predictability of frontal locations by CNN and U-net in JJA and SO from 2003 to 2019. In general, the model has considerable skills for days with the presence of

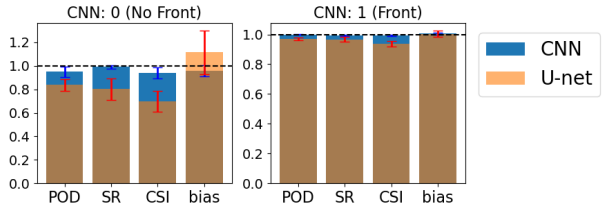


**Fig. 2** predictability of the CNN and U-net for predicting if there is front (1 or 0) in the study region. Bars/errors indicate average values and range of POD, SR, CSI, bias resulting from cross-validation

stationary fronts, especially in JJA. Fronts in JJA occur most frequently between 30N and 36N along the southeastern coast of Japan facing the Pacific Ocean (Fig.3a), and stationary fronts influenced by the Baiu season most likely contribute to the peak values of frontal occurrence in the region. Moreover, when there is no stationary front, mobile fronts tend to align the coast of Japan between 33N and 40N facing the Sea of Japan. There are more areas with low CSI (<0.6) for days without stationary fronts than that with stationary fronts. Fig.3f shows that mobile fronts tend to be overpredicted along the coast of Honshu facing the Pacific Ocean, and in the area further southwest between 30N and 33N. CSI values for days with stationary fronts decrease in Northern Japan, but when there are mobile fronts only, CSI values are relatively high (>0.6) in western Hokkaido and northern Honshu. Figure 4 shows that both stationary and mobile fronts tend to move further southeast in the Pacific Ocean during SO. Comparing to JJA, there are less areas with high CSI values (≥0.8). When stationary fronts are present in SO, fronts tend to be underpredicted (bias≤ 0.9) in most regions in Japan.

Overall, the results suggest that frequency of frontal occurrence can be a factor in influencing the predictability because summer with the presence of stationary fronts tend to be better predicted than other cases. It further suggests that seasonal frontogenesis (e.g., Baiu fronts typically in early summer) may play a role in influencing the predictability of frontal locations.

In addition, the results (Figs. 3g-3h and Figs. 4g-4h) suggest that in most areas, the two-stage model of CNN and U-net is better than using U-net only, and the improvement is generally within 0.2 in terms of the difference in CSI between the two cases. However, in some areas, the opposite is true, and this is most evident for mobile fronts in summer.

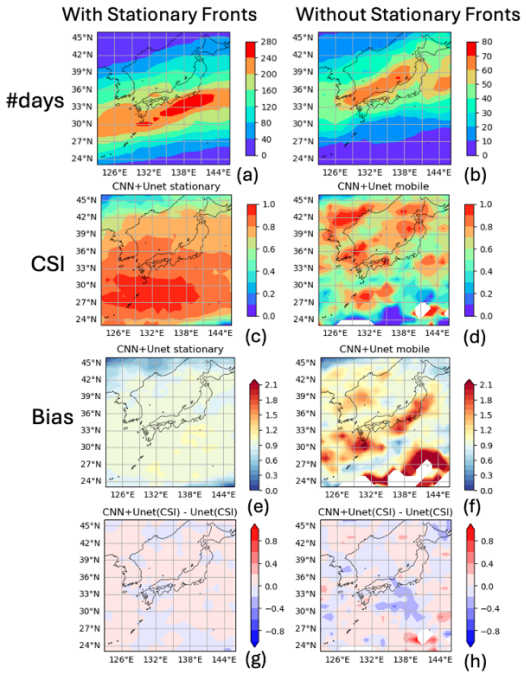


**Fig. 3** (c)-(f):Spatial distribution of JJA predictability of frontal locations (i.e. class 1) in terms of CSI, and Bias from CNN+U-net prediction of all cases using cross-validation. The predictability is calculated for days with and without stationary fronts separately, and number of days in each case is also shown in (a) and (b). (g) and (h): difference between CSI resulted from combining CNN and U-net and U-net only

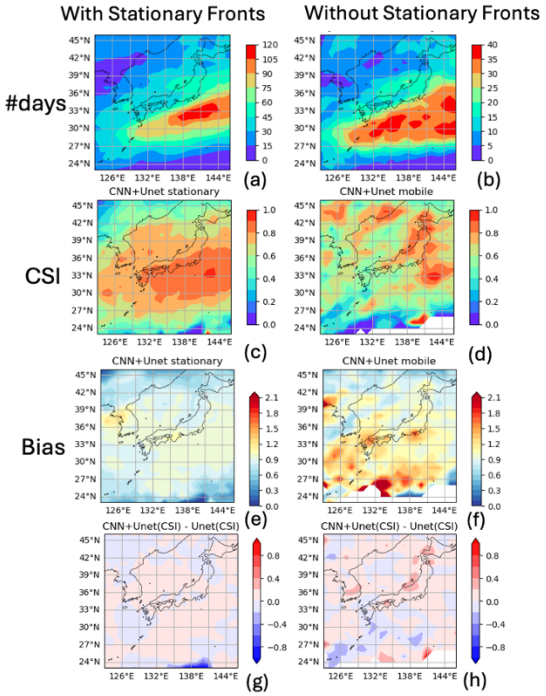


**Fig. 4** Same as in Fig.3 but for September and October (i.e. SO)

## 4. DISCUSSIONS AND CONCLUSIONS

The key element of both CNN and U-net is the convolution/up-convolution process (Eq.1) through which the spatial patterns of the input 2D fields can be recognized by training. In our study, the spatial patterns recognized by CNN/U-net are horizontal gradients of large magnitude for various meteorological variables. If weather fronts exist, they will

leave their imprints in terms of outstanding positive/negative gradients on many meteorological variables related to heat, moisture at multiple atmospheric pressure levels. In other words, the correspondence between the overlapped areas of outstanding horizontal gradients and frontal locations from the ground truth can be regarded as the origin of the predictive signals for CNN/U-net.

Figure 5 shows one example of the predictive signals carried by horizontal gradients of various variables at 850 mb. In this case, we can see the areas of distinctive horizontal gradients often overlap with the locations of fronts, and therefore, the output by the U-net model (Fig.5c) generally matches with the gridded front from the weather chart (Fig.5a-5b). However, not all areas of large magnitude match the frontal locations, which can be regarded as predictive noise. It follows that the predictability of CNN and U-net depends on how well they can recognize the predictive signal for frontal locations while ignoring predictive noise.

Furthermore, the idea that locations of surface fronts can be marked by outstanding positive/negative magnitude of horizontal gradients should hold for all seasons and regions. Therefore, it is ideal to have a universal model for predicting fronts in all areas and seasons. However, the results suggest that predictability varies with season and differ when stationary fronts are present or not, which suggests that dynamics related to seasonal frontogenesis can influence the spatial and temporal distribution of predictability of fronts. Therefore, regional and seasonal models for different types of fronts may be needed.

Overall, our study demonstrates that CNN and U-net are effective in binary detection of frontal locations (with CSI>0.7, bias≅ 1 covering nearly 80% grid points within the study region) during summer when stationary fronts are present.

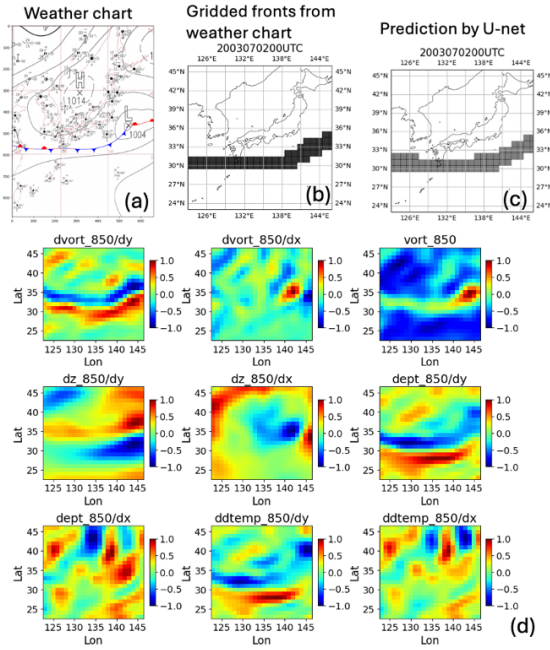


**Fig. 5** (a): one case of frontal observation from the weather chart, (b) gridded datasets of fronts from weather chart and (c) prediction by U-net. (d): normalized horizontal gradients of various meteorological variables and vorticity at 850 mb.

**ACKNOWLEDGMENTS:** This study was supported by the MEXT-Program for the advanced studies of climate change projection (SENTAN) Grant Number JPMXD0722680734 and JSPS KAKENHI grant No 22H01594, Strategic Innovation Promotion Program (SIP) "Development of a Resilient Smart Network System against Natural Disasters" by the Council for Science, Technology and Innovation, Cabinet Office, Government of Japan (Research Promotion Agency: National Research Institute for Earth Science and Disaster Prevention)

## REFERENCES

1) Catto, J L, and Pfahl. S. "The importance of fronts for extreme precipitation." Journal of Geophysical Research: Atmospheres 118.19 (2013): 10-791.
2) Berry, G., Reeder, M. J., & Jakob, C. (2011). A global climatology of atmospheric fronts. *Geophysical Research Letters*, *38*(4).
3) Parfitt, R., Czaja, A., & Seo, H. (2017). A simple diagnostic for the detection of atmospheric fronts. *Geophysical Research Letters*, *44*(9), 4351-4358.
4) Matsuoka, D., Sugimoto, S., Nakagawa, Y., Kawahara, S., Araki, F., Onoue, Y., ... & Koyamada, K. (2019). Automatic detection of stationary fronts around Japan using a deep convolutional neural network. *SOLA*, *15*, 154-159.
5) Goodfellow, I., Bengio, Y., & Courville, A. (2016). Deep learning. MIT press.
6) Biard, James C., and Kenneth E. Kunkel. "Automated detection of weather fronts using a deep learning neural network." Advances in Statistical Climatology, Meteorology and Oceanography 5.2 (2019): 147-160.
7) Justin, A. D., Willingham, C., McGovern, A., & Allen, J. T. (2023). Toward operational real-time identification of frontal boundaries using machine learning. Artificial Intelligence for the Earth Systems, 2(3), e220052.
8) Siddique, N., Paheding, S., Elkin, C. P., & Devabhaktuni, V. (2021). U-net and its variants for medical image segmentation: A review of theory and applications. Ieee Access, 9, 82031-82057.
9) Miyamoto, M., & Yamada, T. J. (2023). Points of consideration on identification of the atmospheric fronts depicted on weather charts. In *IOP Conference Series: Earth and Environmental Science* (Vol. 1136, No. 1, p. 012023). IOP Publishing.
10) Albawi, S., Mohammed, T. A., & Al-Zawi, S. Understanding of a convolutional neural network. In2017 international conference on engineering and technology (ICET) 2017 Aug 21 (pp. 1-6).
11) Nurmi, P. (2003). Recommendations on the verification of local weather forecasts (at ECWMF member states). Consultancy Report (ECMWF Operations Department).

**(Received May 31, 2024)**
**(Accepted September 15, 2024)**