

# Automatic Detection of Stationary Fronts around Japan Using a Deep Convolutional Neural Network

Daisuke Matsuoka<sup>1,2</sup>, Shiori Sugimoto<sup>1</sup>, Yujin Nakagawa<sup>1</sup>, Shintaro Kawahara<sup>1</sup>,  
Fumiaki Araki<sup>1</sup>, Yosuke Onoue<sup>3</sup>, Masaaki Iiyama<sup>4</sup>, and Koji Koyamada<sup>4</sup>

<sup>1</sup>Japan Agency for Marine-Earth Science and Technology (JAMSTEC), Kanagawa, Japan

<sup>2</sup>Japan Science and Technology Agency (JST), Saitama, Japan

<sup>3</sup>Nihon University, Tokyo, Japan

<sup>4</sup>Kyoto University, Kyoto, Japan

## Abstract

In this study, a stationary front is automatically detected from weather data using a U-Net deep convolutional neural network. The U-Net trained the transformation process from single/multiple physical quantities of weather data to detect stationary fronts using a 10-year data set. As a result of applying the trained U-Net to a 1-year untrained data set, the proposed approach succeeded in detecting the approximate shape of seasonal fronts with the exception of typhoons. In addition, the wind velocity (zonal and meridional components), wind direction, horizontal temperature gradient at 1000 hPa, relative humidity at 925 hPa, and water vapor at 850 hPa yielded high detection performance. Because the shape of the front extracted from each physical quantity is occasionally different, it is important to comprehensively analyze the results to make a final determination.

(Citation: Matsuoka, D., S. Sugimoto, Y. Nakagawa, S. Kawahara, F. Araki, Y. Onoue, M. Iiyama, and K. Koyamada, 2019: Automatic detection of stationary fronts around Japan using a deep convolutional neural network. *SOLA*, **15**, 154–159, doi:10.2151/sola.2019-028.)

## 1. Introduction

It is widely known that stationary fronts such as the Baiu front or an autumnal rain front, occur at the equal contact surface of warm and cold air, occasionally resulting in long-term and, sometimes, heavy rainfall. Therefore, accurate estimation of the position and shape of a weather front is important for analyzing rainfall events. If a weather front is automatically detected using weather data, a statistical analysis of a rainfall event using large-ensemble simulation data for future climate change, such as d4PDF (Mizuta et al. 2017), becomes possible. There is a demand among potential users who make decisions in agriculture, forestry, fisheries, and local governments to search for cases similar to past disasters from the data, such as “a stationary front brought heavy rain” or “a typhoon approached a stationary front and brought heavy rain”.

There are many front detection methods based on the thermal front parameter (TFP) at single/multiple atmospheric pressure levels (500–1000 hPa) (Renard and Clarke 1965; Clarke and Renard 1966; JMA 1988; Zwatz-Meise and Mahringer 1988; Huber-Pock and Kress 1989; Berry et al. 2011a; Berry et al. 2011b; Catto and Pfahl 2013). Other front detection approaches are based on temporal changes in the wind at 10 m (Simmonds et al. 2012) and the relative vorticity and horizontal temperature gradient at 900 and 600 hPa (Parfitt et al. 2017). As previously described, multiple physical quantities are used for front detection but the shapes of the frontal surface are different. Although the details are not published, the Japan Meteorological Agency (JMA)

subjectively extracts weather front data according to the weather around Japan using these various indicators from daily forecasting operations and announces the information as a weather map. If front data equivalent to weather map can be extracted automatically from the data, then, it would be useful for non-experts who make decisions based on simulation data.

However, deep learning—a machine learning method based on a neural network—has attracted attention in recent years for its application in image pattern recognition, including image classification, object detection, and image conversion (Krizhevsky et al. 2012; Simonyan and Zisserman 2015). Deep learning techniques related to meteorology include super resolution and image segmentation. Super resolution, an enhancement technique that converts an image from a lower to higher resolution, is utilized for downscaling climatic data (Dong et al. 2014). In addition, image segmentation, a pixel-level extraction of an arbitrary region, is utilized for detecting atmospheric structures such as tropical cyclones and atmospheric rivers (Kurth et al. 2018).

In the present study, we adopt a deep-learning-based image conversion to automatically detect stationary fronts from multiple quantities of weather data. Our deep neural networks train the conversion process from weather data to a front from a 10-year set of analysis and front data extracted from weather map. As deep learning is a statistical method, variations in pattern characteristics due to subjective extraction are reduced. In the present study, we examine the proposed methodology and provide preliminary results of stationary front detection for each physical quantity.

## 2. Data and methodology

We used analysis data from April 2006 to March 2017, provided as initial data from JMA's meso-scale numerical prediction Grid Point Value Meso Spectral Model (GPV-MSM). We used several physical quantities, including precipitation, sea level pressure, and relative humidity as well as water vapor, air temperature, and wind velocity at 1000, 925, 850 and 700 hPa. We also used the horizontal temperature gradient, equivalent potential temperature ( $\theta_e$ ), and wind direction at the same pressure levels, as calculated using the aforementioned quantities. The analysis domain was a rectangular area of 120°E–150°E and 22°N–48°N. We converted the data into 128 × 128 uniform grids as a matter of convenience for machine learning. We also used weather front data extracted from a weather map based on the RGB values of each pixel and their connectivity in the image. The front data are used for the ground truth. Here, “ground truth” is a machine learning term that means the correct answer for supervised learning. The weather map image was originally provided by the JMA and downloaded through the National Institute of Informatics “Digital Typhoon” (<http://agora.ex.nii.ac.jp/digital-typhoon/weather-chart/>). In this work, the JMA's front data—specialized for Japanese weather—is used as the ground truth; however, it is possible to replace it with other front data according to the purpose.

To detect a weather front from the weather data, we employed a U-Net convolutional neural network (Ronneberger et al. 2015),

Corresponding author: Daisuke Matsuoka, JAMSTEC, 3173-25 Showa-machi, Kanazawa-ku, Yokohama, Kanagawa 266-0001, Japan. E-mail: daisuke@jamstec.go.jp.

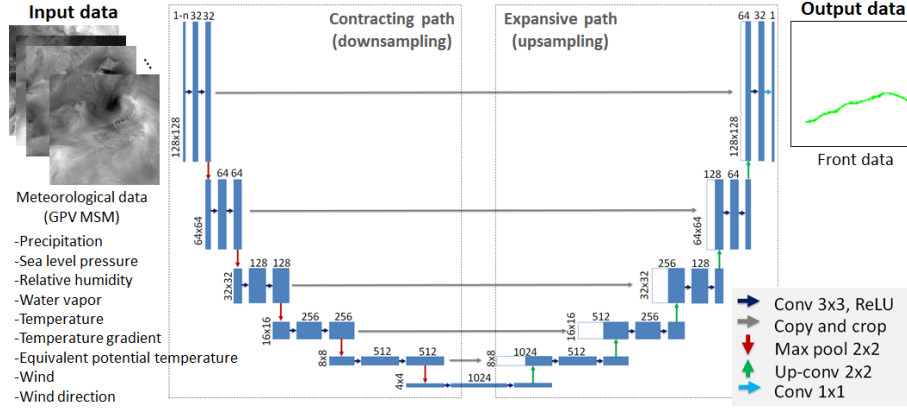
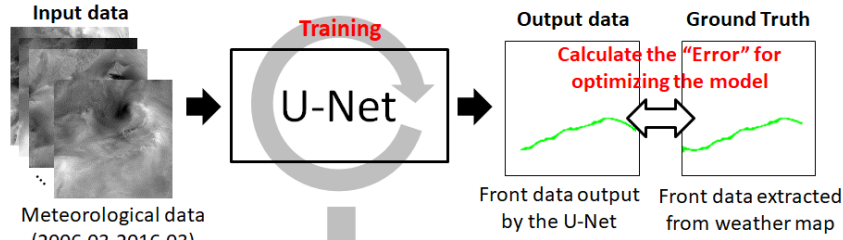


Fig. 1. Architecture of the proposed U-Net convolutional neural network.

## 1. Training phase



## 2. Test phase

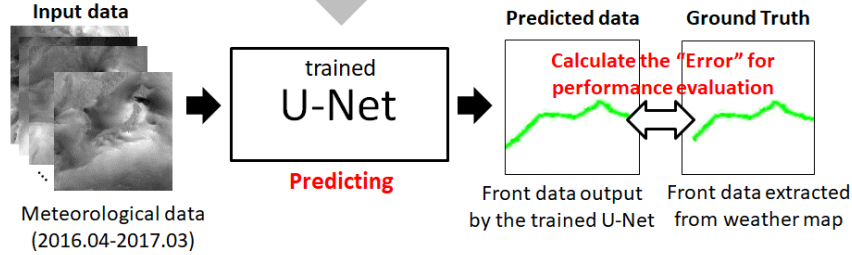


Fig. 2. Conceptual diagram of the training and test phase.

which is one of the convolutional neural network architectures used for image segmentation (LeCun et al. 1999; Krizhevsky et al. 2012). Image segmentation involves the partitioning of a digital image into multiple segments (or sets of pixels) for better analysis. More precisely, it requires the labeling of all pixels in an image such that pixels with the same labels share similar characteristics. In a convolutional neural network, local image features are automatically recognized without prior knowledge using a convolutional operation with convolutional and pooling layers. The U-Net has a contracting path (left side of Fig. 1) for downsampling and an expansive path for upsampling (right side of Fig. 1) and consists of 22 convolutional layers and 10 pooling layers. Input data are weather data with multiple channels (i.e. the color picture has three RGB channels). While general image pattern recognition techniques use RGB values as the channels, the proposed method uses physical quantities of weather data on a uniform grid as the channels by assuming grid points are pixels. Convolutional layers (i.e. “conv  $3 \times 3$ ,” dark blue arrows in Fig. 1) extract local features (feature maps) from the input data using convolutional filters  $3 \times 3$  pixels in size. Pooling layers have two types of functions: downsampling for reducing the resolution of input data (“max pool  $2 \times 2$ ,” red arrows in Fig. 1) and upsampling for recovering the resolution (“up-conv  $2 \times 2$ ,” green arrows in Fig. 1), both using filters  $2 \times 2$  pixels in size. Local and global features are

trained in shallow (upper side of Fig. 1) and deep (bottom side of Fig. 1) layers, respectively. In addition, by merging the layers of a contracting path and an expansive path at the same depth, global information is converted while retaining local features. Based on the architecture of the original U-Net, we examined a correct detection ratio (described later) of 27 combinations of architecture settings: the number of convolutional layers (14, 18, 22), size of the convolutional filter ( $3 \times 3$ ,  $5 \times 5$ ,  $7 \times 7$ ), and number of feature maps in the first convolutional layer (32, 64, 128). Accordingly, we adopted the architecture with the highest level of accuracy.

A conceptual diagram of the proposed method, consisting of a training phase and a test phase, is shown in Fig. 2. During the training phase, the training data, including input data (weather data) and ground truth (front data extracted from weather map), were provided to U-Net as shown in Fig. 2. One or more physical quantities of weather data can be input through the multiple input channel. Here, the loss function, which is an error between the ground truth and output data by U-Net, is defined using the Tanimoto coefficient (Tanimoto 1957) as follows:

$$Loss = 1 \times Tc(F_{GT}, F_{PRED}), \quad (1)$$

$$Tc(F_{GT}, F_{PRED}) = F_{GT} \cdot F_{PRED} / (|F_{GT}|^2 + |F_{PRED}|^2 - F_{GT} \cdot F_{PRED}), \quad (2)$$

where  $F_{GT} \cdot F_{PRED} = \sum_i F_{GTi} F_{PREDi}$  and  $|F_{GT}|^2 = \sum_i F_{GTi}^2$ , and  $|F_{PRED}|^2 = \sum_i F_{PREDi}^2$ .  $F_{GTi}$  and  $F_{PREDi}$  are the existence probability of the stationary front of the  $i$ -th grid in the ground truth and prediction results, respectively; their values range from 0 to 100%.  $T_c$  is the correct detection ratio ( $CDR$ ) based on the similarity between the sets  $F_{GT}$  and  $F_{PRED}$ . When the sets  $F_{GT}$  and  $F_{PRED}$  are completely matched, the  $CDR$  is 100%. The transformational process repeatedly learns in U-Net from training data until the loss function converges to a sufficiently low value. If the sets  $F_{GT}$  and  $F_{PRED}$  are both zero, this data set is not used for training, as the  $T_c$  ( $F_{GT}$ ,  $F_{PRED}$ ) cannot be defined. The Adam optimizer (Kingma and Ba 2015) was applied to U-Net to minimize the loss function. During the test phase, the trained U-Net inputs untrained data and predicts the front data as shown in Fig. 2. The detection performance is evaluated by calculating the similarity between the prediction result and the ground truth based on the Tanimoto coefficient.

We used 10-year analysis and front data from 04/01/2006 0000 to 03/31/2016 1200 UTC (time interval of 12 h and a total time step of 7306) for the training data and 1-year data from 04/01/2016 0000 to 03/31/2017 1200 UTC (730 step) for the test data. The physical quantities used for the input data included 30 types of single variables at a single altitude (ground level, at 1000, 925, 850, and 700 hPa), 7 types of single variables at multiple pressure levels (1000, 925, 850, and 700 hPa), and 4 types of multiple variables at a single level. Single quantities of a single altitude contain precipitation and sea level pressure (SLP) at the ground, relative humidity, water vapor, air temperature, horizontal temperature gradient,  $\theta_e$ , horizontal wind vector, and wind direction at 1000, 925, 850 and 700 hPa. Here, for the vector quantities, such as temperature gradient and wind, both zonal and meridional components were independently input as two channels. Single quantities at multiple levels include relative humidity, water vapor, temperature, temperature gradient,  $\theta_e$ , wind vector, and wind direction at 1000, 925, 850, and 700 hPa. Similarly, both temperature gradient and wind vector were input as eight channels while the other quantities were input as four channels. As for multiple variables at single level, we chose all quantities except for precipitation, SLP, and temperature at 1000, 925, 850, and 700 hPa (8 input channels).

### 3. Results and discussion

Detection results of the Baiu front at 06/28/2016 0000 UTC are shown in Fig. 3. Figure 3a is the ground truth and (b) to (t) are the results when using typical 19 input patterns because of space limitations. Except for temperature at 1000 and 850 hPa and  $\theta_e$  at 1000 hPa, the shape of the detected front using each input pattern seems qualitatively similar to ground truth. Physical quantities with particularly high  $CDR$  are relative humidity at 850 hPa (67.93%) and 1000 hPa (63.35%),  $\theta_e$  at 850 hPa (62.37%), and all quantities at 1000 hPa (59.49%). In addition, only precipitation at the ground can detect the front and is relatively high (57.15%) in this case. The  $CDR$  of the wind at 1000 hPa (52.24%), 850 hPa (40.50%), the wind direction at 1000 hPa (57.70%), and 850 hPa (56.11%) are not high compared to the aforementioned values. However, their resulting shapes seem qualitatively good.

The results of other case studies are shown in the supplementary materials. Supplement 1 depicts the detection results of the Baiu front at 06/02/2016 1200 UTC. It is notable that the relative humidity at 1000 hPa detects the east side of the Baiu front (approximately east of 135°E), despite the same quantities at 850 hPa detecting the opposite side of the front line. Similar trends can be seen in the results using the temperature gradient. Supplement 2 depicts the detection results of the autumnal rain front at 09/29/2016 0000 UTC. The extracted front is qualitatively similar to the ground truth except for the results using temperature. Supplement 3 depicts the detection results of the Baiu front at 06/05/2016 0000 UTC. In the case of using the relative humidity at 850 hPa or the water vapor at 850 hPa, a cold front and a warm front are misdetected. Supplement 4 depicts the detection results of a stationary front at 08/27/2016 0000 UTC. In this example, a

typhoon exists on the south side of Japan, and results from almost all input patterns do not extract fronts well.

The  $CDR$  for each input pattern and for each time step of the 1-year test data are shown in Fig. 4. The vertical and horizontal axes represent the typical input pattern and time step from 04/01/2016 0000 to 03/31/2017 1200 UTC, respectively. The magenta, green, and red boxes on the monthly notations indicate the occurrence of the Baiu front, autumn front, and typhoon, respectively. The Baiu and autumnal front seasons are based on the announcement by the JMA (e.g. <https://www.data.jma.go.jp/fcd/yoho/hibiten/index.html> and <https://www.data.jma.go.jp/fcd/yoho/baiu/index.html>). The typhoon season is the period when typhoons exist in the analysis area. The temporal changes of each input pattern seem to approximately show seasonality.

Figure 5 shows  $CDR$  for the 1-year total, the Baiu front season, autumnal rain front season, and the approaching typhoon using the major input patterns at each pressure level (1000, 925, 850, and 700 hPa). These statistics are only for the period in which the stationary front exists in the ground truth. Focusing on the median of each box plot, the detection performance of the autumnal rain front and Baiu front are relatively higher than the other seasons for all input patterns. In contrast, detection performance decreases during the typhoon season. For example, the median  $CDR$ s using all quantities at 1000 hPa are 35.61%, 51.72%, 11.00%, and 27.82% during the Baiu front, autumnal rain front, typhoon, and other seasons, respectively.

In almost all seasons, the input patterns with a high median  $CDR$  among the single physical quantities at a single level (e.g., during the autumnal rain front) are wind direction at 1000 hPa (43.69%), wind vector at 1000 hPa (41.73%), temperature gradient at 1000 hPa (43.15%), relative humidity at 925 hPa (40.20%), and water vapor at 850 hPa (41.57%). For any physical quantities, the pressure levels of the input data and the median  $CDR$ s are related to each other. The temperature gradient, wind, and wind direction at 1000 hPa have good detectability compared to those at other pressure levels. In contrast, the pressure level with the highest mean  $CDR$  is 925 hPa for relative humidity and 850 hPa for water vapor and  $\theta_e$ . These tendencies are common in every season except for the typhoon season. In addition, when simultaneously using single quantities at all pressure levels, the detection performance is not significantly higher than those at only a single level. For example, the median  $CDR$  during the autumnal rain front season using the temperature gradient at all levels is 40.34% and there is no significant improvement compared to the same quantity at 1000 hPa (43.15%) and 925 hPa (36.49%). These tendencies are based on the median  $CDR$ , and the input pattern for obtaining the highest  $CDR$  may differ from day-to-day as shown in Fig. 4.

In addition, the median  $CDR$ s using all quantities, except for SLP and temperature, at 1000 hPa during the Baiu front and autumnal rain front season are 35.61% and 51.72%, respectively. Although the detectability using these input patterns are highest, there is also no significant difference compared to the aforementioned results. Theoretically, the detection accuracy should improve as the number of input channels increases; however, more data are required. It might be necessary to strictly select the combination of input quantities most suitable for detecting stationary fronts.

### 4. Conclusions

In this work, we extracted a stationary front from weather data using a U-Net deep convolutional neural network. As a result, we succeeded in extracting the approximate shape of seasonal rain fronts, such as the Baiu front and autumnal rain front, by simultaneously using relative humidity, water vapor, temperature gradient, equivalent potential temperature, wind vector, and wind direction at 1000 hPa as input data. In contrast, detection performance decreased upon the approach of a typhoon. The decrease in performance appears to be caused by the disturbance of the field by the presence of the typhoon. Detailed meteorological analysis should be conducted in the future to deepen understanding of



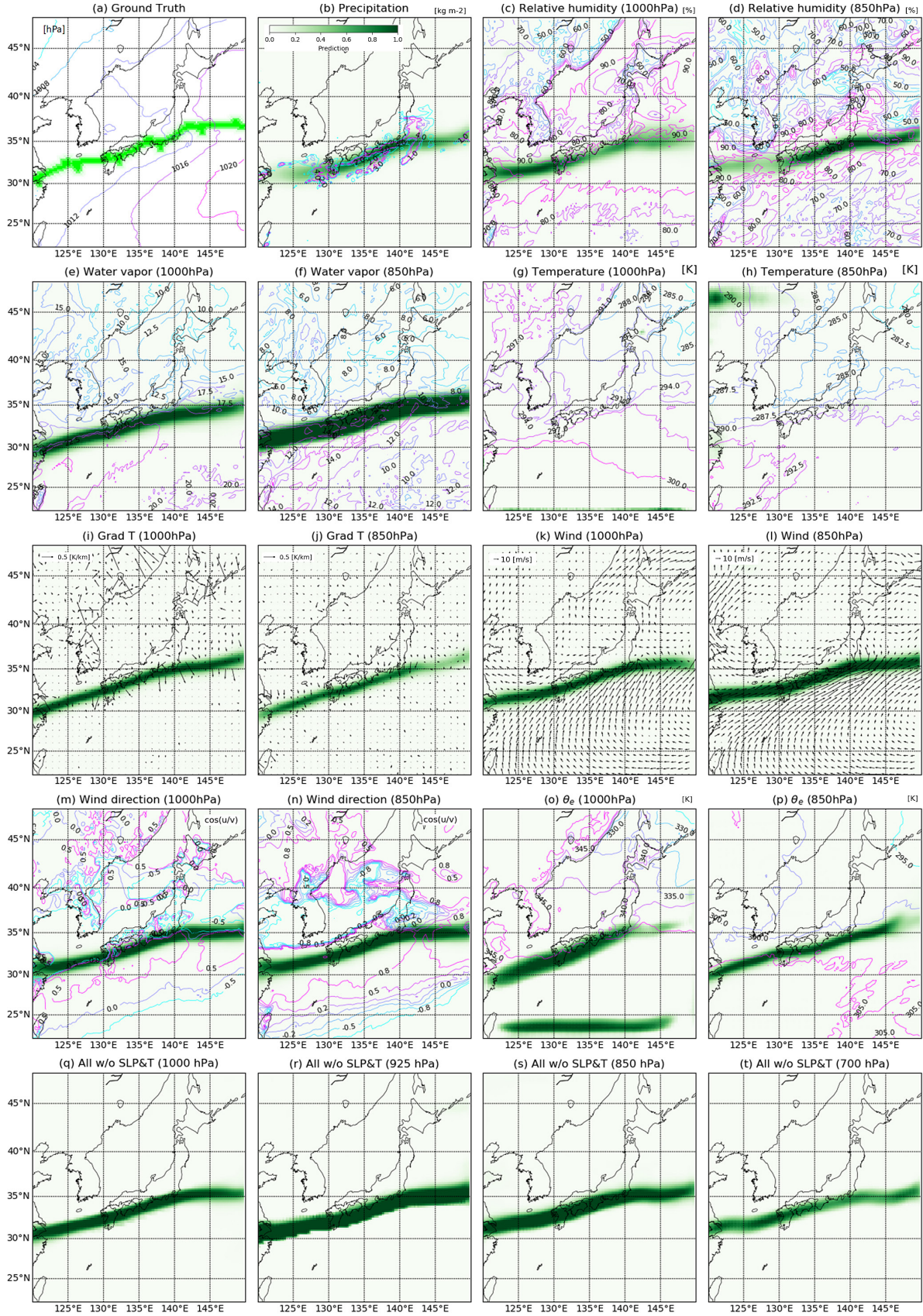


Fig. 3. Detection results of the Baiu front at 2016/06/28 0:00:00 UTC. (a) Ground truth, (b) precipitation, (c) relative humidity at 1000 hPa, (d) relative humidity at 850 hPa, (e) water vapor at 1000 hPa, (f) water vapor at 850 hPa, (g) temperature at 1000 hPa, (h) temperature at 850 hPa, (i) temperature gradient at 1000 hPa, (j) temperature gradient at 850 hPa, (k) wind at 1000 hPa, (l) wind at 850 hPa, (m) wind direction at 1000 hPa, (n) wind direction at 850 hPa, (o)  $\theta_e$  at 1000 hPa, (p)  $\theta_e$  at 850 hPa, (q) all quantities without SLP and temperature at 1000 hPa, (r) all quantities without SLP and temperature at 925 hPa, (s) all quantities without SLP and temperature at 850 hPa, and (t) all quantities without SLP and temperature at 700 hPa.

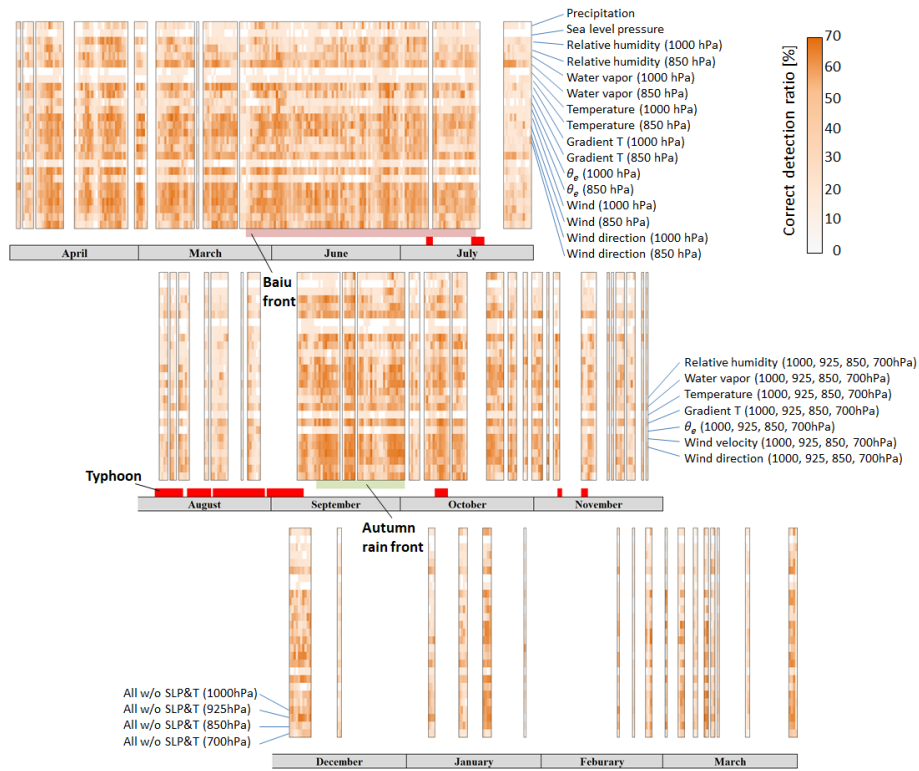


Fig. 4. Detection performance for one-year test data for each input pattern.

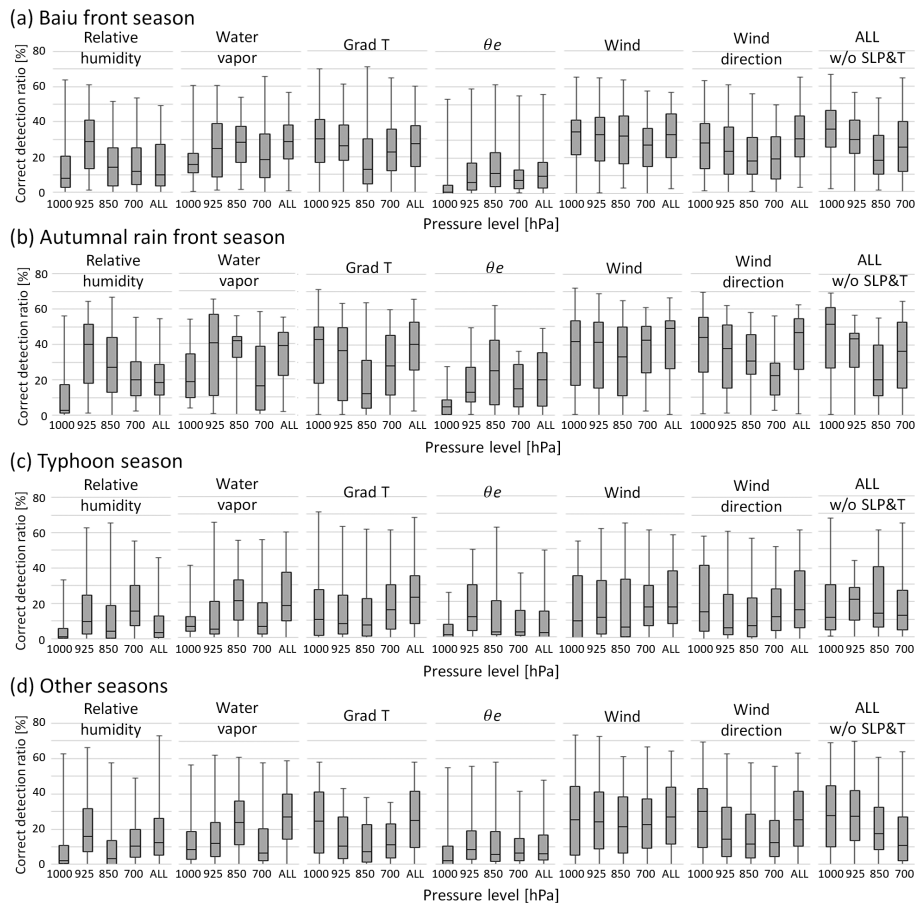


Fig. 5. Seasonal detection performance for each input pattern for each pressure level. (a) Baiu front season, (b) autumnal rain front season, (c) typhoon season, and (d) other seasons.



the effects of typhoons and improve detection. Furthermore, the experiments using single physical quantities show that relatively higher detection performance can be obtained using only single input data, such as horizontal temperature gradient, wind vector, wind direction at 1000 hPa, relative humidity at 925 hPa, and water vapor at 850 hPa. Interestingly, in some cases when using only single physical quantities, different parts of the front line are extracted; their integration results in a shape close to the ground truth. This suggests that it is important to comprehensively determine the front shapes obtained using different physical quantities. Our results only apply to stationary fronts; the detection of cold and warm fronts requires further study.

At present, it is well known that deep learning has a “black box” problem and it is difficult to understand its judgement criteria. Future work should clarify the extraction mechanism of the proposed method and compare it to existing methods. We hoped that these findings will provide an important contribution to automatic front detection in operational organizations.

## Acknowledgements

The authors are grateful to the anonymous reviewers for their constructive comments. This research was performed as a part of the Social Implementation Program on Climate Change Adaptation Technology (SI-CAT). This research was also partially supported by Japan Science and Technology Agency (JST), PRESTO Grant Number JPMJPR1777. We would like to thank Editage (www.editage.com) for English language editing.

Edited by: S. Nishizawa

## Supplement

Supplement 1 depicts the detection results of the Baiu front at 06/02/2016 1200 UTC.

Supplement 2 depicts the detection results of the autumnal rain front at 09/29/2016 0000 UTC.

Supplement 3 depicts the detection results of the Baiu front at 06/05/2016 0000 UTC.

Supplement 4 depicts the detection results of the stationary front at 08/27/2016 0000 UTC.

## References

- Andreae, M. O., and A. Gelencser, 2006: Black carbon or brown carbon? The nature of light-absorbing carbonaceous aerosols. *Atmos. Chem. Phys.*, **6**, 3131–3148.
- Berry, G., M. J. Reeder, and C. Jakob, 2011a: A global climatology of atmospheric fronts. *Geophys. Res. Lett.*, **38**, L04809, doi:10.1029/2010GL046451.
- Berry, G., C. Jakob, and M. Reeder, 2011b: Recent global trends in atmospheric fronts. *Geophys. Res. Lett.*, **38**, L21812, doi:10.1029/2011GL049481.
- Catto, J. L., and S. Pfahl, 2013: The importance of fronts for extreme precipitation. *J. Geophys. Res. Atmos.*, **118**, 10791–10801, doi:10.1002/jgrd.50852.
- Clarke, L. C., and R. J. Renard, 1966: The U. S. Navy numerical frontal analysis scheme; Further development and a limited evaluation. *J. Appl. Meteor.*, **5**, 764–777.
- Dong, C., C. C. Loy, K. He, and X. Tang, 2014: Learning a deep convolutional network for image super-resolution. *Proc. IV European Conference on Computer Vision (ECCV)*, Zurich, Switzerland, 184–199.
- Huber-Pock, F., and C. Kress, 1989: An operational model of objective frontal analysis based on ECMWF products. *Meteor. Atmos. Phys.*, **40**, 170–180.
- Japan Meteorological Agency (JMA), 1988: On the improvement of the significant weather chart. *Wea. Serv. Bull.*, **55**, 1–16.
- Kurth, T., S. Treichler, J. Romero, M. Mudigonda, N. Luehr, E. Phillips, A. Mahesh, M. Matheson, J. Deslippe, M. Fatica, Prabhat, and M. Houston, 2018: Exascale Deep Learning for Climate Analytics. *Proc. International Conference for High Performance Computing, Networking, Storage, and Analysis (SC18)*, doi:10.1109/SC.2018.00054.
- Krizhevsky, A., I. Sutskever, and G. E. Hinton, 2012: ImageNet classification with deep convolutional neural networks. *Adv. Neural Inf. Process. Syst.*, **25**, 1097–1105.
- LeCun, Y., P. Haffner, L. Bottou, and Y. Bengio, 1999: Object recognition with gradient-based learning. *Shape, Contour and Grouping in Computer Vision, Lecture Notes in Computer Science*, D. A. Forsyth, J. L. Mundy, G. Vd, R. Cipolla, Eds., **1681**, 319–345.
- Mizuta, R., A. Murata, M. Ishii, H. Shiogama, K. Hibino, N. Mori, O. Arakawa, Y. Imada, K. Yoshida, T. Aoyagi, H. Kawase, M. Mori, Y. Okada, T. Shimura, T. Nagatomo, M. Ikeda, H. Endo, M. Nosaka, M. Arai, C. Takahashi, K. Tanaka, T. Takemi, Y. Tachikawa, K. Temur, Y. Kamae, M. Watanabe, H. Sasaki, A. Kitoh, I. Takayabu, E. Nakakita, and M. Kimoto, 2017: Over 5,000 years of ensemble future climate simulations by 60-km global and 20-km regional atmospheric models. *Bull. Amer. Meteor. Soc.*, **98**, 1383–1398.
- Parfitt, R., A. Czaja, and H. Seo, 2017: A simple diagnostic for the detection of atmospheric fronts. *Geophys. Res. Lett.*, **44**, 4351–4358.
- Renard, R. J., and L. C. Clarke, 1965: Experiments in numerical objective frontal analysis. *Mon. Wea. Rev.*, **93**, 547–556.
- Ronneberger, O., P. Fischer, and T. Brox, 2015: U-Net: Convolutional Networks for Biomedical Image Segmentation. *Medical Image Computing and Computer-Assisted Intervention (MICCAI), Lecture Notes in Computer Science*, **9351**, 234–241.
- Simonyan, K., and Z. Zisserman, 2015: Very deep convolutional networks for large-scale image recognition. *Proc. International Conference on Learning Representation (ICLR) 2015*, doi:10.1109/ACPR.2015.7486599.
- Tanimoto, T. T., 1958: *An Elementary Mathematical Theory of Classification and Prediction*, International Business Machines Corporation, **10**.
- Vandal, T., E. Kodra, S. Ganguly, A. Michaelis, R. Nemani, and A. R. Ganguly, 2018: Generating high resolution climate change projections through single image super-resolution: An abridged version. *Proc. 27th Int. Joint Conf. on Artificial Intelligence (IJCAI-18)*, 5389–5393.
- Zwatz-Meise, V., and G. Mahringer, 1988: Use of satellite imagery, combined with numerical model diagnostics, to locate fronts and predict their activity: Methods and examples. *Satellite and Radar Imagery Interpretation, Preprints for a Workshop in a Reading*, M. Mader and T. Waters, Eds., England, 20–24 July 1987, Eumetsat, 143–162.

Manuscript received 11 April 2019, accepted 13 June 2019

SOLA: <https://www.jstage.jst.go.jp/browse/sola/>