atmosphere

MDPI

Article

Multi-Level Circulation Pattern Classification Based on the Transfer Learning CNN Network

Yanzhang Liu [1], Jinqi Cai [2] and Guirong Tan [2,*]

[1] Jiangxi Vocational and Technical College of Information Application/Department of Meteorology, Nanchang 330043, China
[2] Key Laboratory of Meteorological Disaster, Ministry of Education (KLME), Joint International Research Laboratory of Climate and Environment Change (ILCEC), Collaborative Innovation Center on Forecast and Evaluation of Meteorological Disasters (CIC-FEMD), Nanjing University of Information Science and Technology, Nanjing 210044, China
* Correspondence: tanguirong@nuist.edu.cn

Abstract: Deep learning artificial intelligence technology, which has the advantages of nonlinear mapping ability, massive information extraction ability, spatial-temporal modeling ability, and so on, provides new ideas and methods for further improving the accuracy of weather and climate extreme event prediction. A transfer learning CNN (Convolutional Neural Networks) classification model is established to classify the circulation patterns, along with the newly reconstructed dataset of regional persistent historical heavy rain events, daily rainfall data of 2474 observational stations, and the NCEP/NCAR global reanalysis data of daily geopotential height field in 1981–2018. Different from previous classifications, usually with one level variable, here, in addition to 500 hPa heights, 200 hPa zonal winds and 850 hPa meridional winds over the key areas are also considered in the model. The results show that the multi-level circulation pattern classification based on the transfer learning CNN network has a higher accuracy in the independent test than the single-level model, with the accuracy reaching 92.5% (while only 85% for the single-level model). The spatial correlation coefficient of precipitation between each typical mode and related patterns obtained by the multi-level transfer learning CNN classification is greater than that obtained by the single-level transfer learning CNN, and the variance of 500 hPa heights between each typical mode and the associated patterns is also greater than that obtained by the single-level transfer learning CNN. These results show that the performance of the classification by the multi-level transfer learning CNN model is better than that by the single-level transfer learning CNN. The study is helpful to develop circulation classifications related to large-scale weather or climate disaster events and then to provide a physical basis for further improving the forecast effect and extending the valid time of the forecast through combining the numerical model products.

Keywords: multi-level circulation pattern classification; deep learning; transfer learning

Citation: Liu, Y.; Cai, J.; Tan, G. Multi-Level Circulation Pattern Classification Based on the Transfer Learning CNN Network. *Atmosphere* **2022**, *13*, 1861. https://doi.org/10.3390/atmos13111861

Academic Editor: Gianfranco Vulpiani

Received: 13 October 2022
Accepted: 5 November 2022
Published: 9 November 2022

1. Introduction

Artificial neural networks have long been applied to meteorology [1–4]. As denoted by Xu [5], the application of Artificial Intelligence (AI) in meteorology has shown great potential value and broad aspects. In recent years, deep learning technology has risen and is being widely used [6–9]. In the field of weather forecasting, AI is not only used for short-term weather forecasting and nowcasting [6] but also for the diagnosis of extreme weather events [10–13]. There are also a lot of technologies based on deep learning combined with a dynamic model [14–18]. For example, AI technology is used to optimize or replace part of the parameterization scheme in the numerical model to improve the efficiency of the running model [19]. An AI algorithm is also employed to improve the performance of the model forecast in model forecast product interpretation, the statistical post-processing of ensemble forecast, statistical downscaling, and other aspects, etc. [20–29].

AI technology mainly includes traditional machine learning and deep learning. Common traditional machine learning algorithms include random forest, logical regression, etc. Deep learning is able to automatically learn the effective features from a dataset, which is usually used to solve the problem of insufficient fitting ability by traditional machine learning in high-dimensional space [30,31]. The models involved mainly CNN, Long Short Term Memory (LSTM), and the combination of both [32–36]. Each model can be divided into different types. For example, CNN mainly includes conventional CNN, including classification, regression models, as well as FCN (Fully Convolutional Networks) [37] and U-Net [32,38–40] based on semantic segmentation. CNN, as a neural network with a deep learning function, has the advantages of a flexible range of applicability, simultaneous feature extraction with classification, strong generalization ability, fewer parameters used in global optimization training, etc. Compared with the traditional methods, classification no longer requires manual feature description and image extraction for the target. Instead, CNN neural networks learn features from training samples independently. The algorithm has achieved great success in image recognition, data mining, and other fields in recent years due to its excellent spatial feature extraction ability [31,41–43]. At present, this method has been well applied in the forecast correction of weather model products [44,45] and the nowcast based on image recognition with better results achieved [46]. Cai et al. [47] established an objective circulation classification through the step-by-step training of migration learning by the residual neural network (CNN). Additionally, the circulation patterns of persistent heavy rainfall cases in the Jianghuai Basin over China from 1981 to 2018 are objectively classified by the transfer learning CNN model. The results show that the multi-step transfer learning CNN model is better than the other two analog methods with R (pattern classification based on a new analog quantity R) [48] and COS (pattern classification based on cosine similarity) [47,49] as computing parameters. In the objective classification of persistent heavy precipitation, the precipitation correlation coefficients between each base pattern and the related pattern obtained by CNN are much higher than those obtained by the R and COS methods [47]. Additionally, a certain ability of the CNN is also shown to identify the circulation classification for non-persistent heavy rainfall.

The formation of regional weather and climate is very complex, such as the heavy rainfall in Jianghuai Basin; its occurrence and development will be affected by varied weather systems at the low, middle, and high levels of the troposphere, respectively. In addition to the weather systems at 500 hPa, such as the blocking high at middle and high latitudes, the westerly trough, the subtropical high over the western Pacific at low latitudes, the upper jet stream, and the South Asian high at 200 hPa, the southwest jet stream at lower altitudes can also be very important [47,49]. Therefore, this paper intends to take advantage of the transfer learning CNN method. On the basis of the single-level variable transfer learning CNN classification with single 500 hPa heights, the upper-level jet stream and the low-level southwest jet stream are also considered to be added to the transfer learning CNN classification model. That is, considering the changes in the location and intensity of the upper-level jet stream and the impact of the warm and humid air brought by the low-level southwest jet stream on regional extreme weather at lower levels, the circulation classification related to large-scale persistent meteorological disaster events will be developed with multi-level variables, which will provide a physical basis for further improving the forecast effect and extending the valid forecast time through combining with the model products.

The structure of the paper is organized as follows. Following the introduction, the data and methods are described in Section 2. Section 3 gives the spatial distribution of the heavy rainfall and related circulation patterns obtained by the objective analysis. Then, the section explores the circulation classification by multi-level variable transfer learning CNNs, with the performance explained in comparison with those obtained by single-level transfer learning CNN classification. Additionally, conclusions are summarized in Section 4, with some existing important problems also discussed.

2. Data and Methodology

2.1. Data

The data used in this paper mainly include: (1) the daily precipitation of 2474 stations in China from 1981 to 2019 provided by the National Meteorological Center; (2) Historical cases of persistent heavy rainfall events in China from 1981 to 2018 [50]; (3) daily reanalysis data from NCEP/NCAR, with a horizontal resolution of 2.5° × 2.5° [51].

Jianghuai Basin denotes the place over the (110°–126° E, 22°–36° N) area, the region between the Yangtze River and the Huaihe River (Figure 1) [52]. Jianghuai Basin is in the transition zone between south China (with abundant rainfall) and north China (with little rainfall). The interannual and seasonal variability of the rainfall is very large usually accompanying frequent disastrous weather or climate events, especially floods or heavy rainfall events during summer with the onsetting of plum rain. Additionally, the resulting factors are very complicated. Due to the developed economy and large population density in the region, disaster events will bring huge losses and impacts to the economy and people's lives. Therefore, it is of great significance to study heavy rainfall events and their prediction to give early warnings of flood disasters.

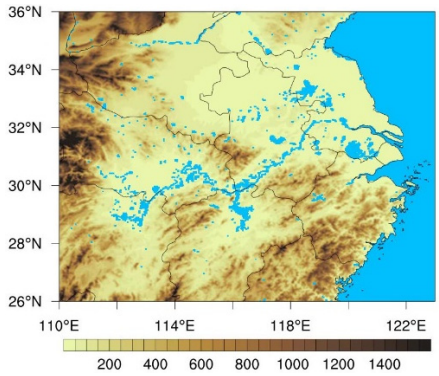


Figure 1. Area of study with topographical elevation (m) in color.

2.2. Methodology

2.2.1. Residual Neural Network

CNN emerged in recent years and is a pattern recognition method combining artificial neural networks and deep learning theory, which has become one of the research focuses in the field of image classification. Different from the traditional image classification method, CNN does not need to extract specific features manually from the original image for specific tasks but performs hierarchical abstract processing on the original image to produce classification results by simulating the human visual system, which has a higher recognition rate and wider practicability. This algorithm applies to local receptive fields, weight sharing, and spatial sub-sampling techniques, which greatly reduces the number of training parameters of the network in comparison with other artificial neural networks. It has been widely used in speech recognition, face recognition, handwriting recognition, pedestrian detection, and other fields. A residual neural network is a special kind of convolutional neural network, which can be trained to achieve a task by fitting residual mapping. The residual neural network (ResNet18) is used here with 18 layers [52]. Here, 18 refers to 18 layers with weight, including convolution layers and fully connected layers apart from the pooling layers and BN (Batch Normalization) layers. If a deep network is built on a shallow network and all the added layers are direct copies of the previous layer (i.e., y = x), the training error of the deep network should be equal to that of the shallow network under extreme conditions. Therefore, the root way to resolve network degradation is optimization. In the residual network, it is not the original mapping allowed to fit directly but the residual mapping. Assuming that the original mapping is H (x), then the mapping of residual network fitting is F (x) = H (x) − x. The output H (x) = F (x) + x of the residual unit is divided into two parts. One is obtained by the input of the network x through

directly connecting the identity mapping, and the other part is obtained by the network training of the residual F (x). In this way, the integrity of the information is protected to some extent, with learning difficulty decreased because of the residual goal of learning.

The ResNet18 here used has the same architecture as that by He et al. [49], which is shown in Table 1. Additionally, the numbers of neurons and layers are kept the same as the ResNet18 used in the reference paper, and the input data are also treated into images. For example, for a single CNN model, the input data are color maps of 500 hPa heights, while for multi-level CNN is a map including 500 hPa heights, 200 hPa zonal winds, and 850 hPa winds. Additionally, the sizes and color scales for different days are the same for the same values. The target output of the CNN model during the training is a value of 1 for Pattern I, 2 for Pattern II, and 3 for Pattern III.

Table 1. Architectures for CNN.

Layer Name	Output Size	18-Layer
conv1	112 × 112	7 × 7, 64, stride 2
		3 × 3 max pool, stride 2
conv2_x	56 × 56	$\begin{bmatrix} 3\times3, 64 \\ 3\times3, 64 \end{bmatrix} \times 2$
conv3_x	28 × 28	$\begin{bmatrix} 3\times3, 128 \\ 3\times3, 128 \end{bmatrix} \times 2$
conv4_x	14 × 14	$\begin{bmatrix} 3\times3, 256 \\ 3\times3, 256 \end{bmatrix} \times 2$
conv5_x	7 × 7	$\begin{bmatrix} 3\times3, 512 \\ 3\times3, 512 \end{bmatrix} \times 2$
	1 × 1	average pool, 1000-d fc, softmax
FLOPs		1.8 × 109

2.2.2. Steps of Model Establishing

The establishment of the classification model is mainly divided into three steps. Firstly, the basic patterns of objective classification from historical data are constructed. Here, the typical modes and related circulation patterns of heavy rainfall in the Jianghuai Basin are extracted by EOF (empirical orthogonal function) using the historical case dataset.

Second, the circulation classification model related to the persistent heavy rainfall in Jianghuai Basin is established by the transfer learning CNN (ResNet18), with the dataset of the typical rainfall modes and related circulations. The transfer learning here mainly lies in the three training processes. That is, a part of the samples used for training are put into the network for training in batches. The samples put into the training for the first time are 284 case days. After the training, the parameter values obtained from the network training are retained and brought into the next network training as the initial values. A second training with another 30 sample days was added in. Similarly, the results after the second training are brought into the third network training. Adding another 30 daily samples for the third training, the corresponding parameter values will be retained at the end of the network training, with the minimum values (threshold values) of each pattern output and counted.

Finally, the circulation patterns for independent samples (40 sample days in the test set) are to be classified objectively by the established transfer learning CNN network model. The corresponding output can be obtained by the CNN network with the data input for the classification. When the maximum value of the output is larger than the threshold value, the corresponding pattern of the day can be obtained (here, three patterns are divided, not limited to the three patterns in the actual application). Otherwise, it is outside the previous

basic patterns (that is, the pattern is not included in the basic three patterns for persistent precipitation, such as the one without precipitation).

3. Results

3.1. Typical Modes of the Heavy Rainfall

In order to obtain the samples for the training and test of the transfer learning CNN network, the typical modes are refined from 72 extreme persistent heavy rainfall cases that occurred in Jianghuai Basin in summer by EOF (empirical orthogonal expansion). In total, 296 days (excluding duplicate dates) are selected. The typical modes of the extreme persistent heavy rainfall events can well reflect the characteristics of the heavy rainfall in the Jianghuai Basin and ensure the typicality and accuracy of each pattern to a certain extent. The main modes of EOF analysis are obtained according to the variance contribution (Figure 2). It can be seen from the figure that the first mode shows a consistent variation of rainfall in the whole region, and the center of the maximum rainfall is located slightly south of the middle part of the Jianghui Basin. The spatial distribution of the second mode appears as a reverse variation over the north to that over the south of the area. The third mode reflects the variation in the middle region opposite to that over the north and south. The variance of the three modes accounts for about 43.8% of the total variance, with 25.8%, 10.9%, and 7.1% for each, respectively. There are 83 day, 79 day, and 53 day samples (215 days in total) selected according to the time coefficients of the three modes, and the rainfall and circulation distribution of the typical mode can be extracted by composite analysis (Figure 3), which will be the basis for further objective classification.

It can be seen from Figure 2 that the 500 hPa heights of the first mode (Figure 3a) at middle and high latitudes are dominated by negative anomalies, with weak positive anomalies near the south of the Ural Mountains and negative centers over the north of the Ural Mountains and the Okhotsk Sea, corresponding to the development of troughs and ridges, respectively. Additionally, at low latitudes, the heights are characterized by the western Pacific subtropical high with a weak positive variation. The 500 hPa heights of the second typical mode (Figure 3b) at the middle and high latitudes are almost covered by positive anomalies, especially with a positive center over the Ural Mountain area, denoting block high occurring frequently. The low latitudes are also dominated by positive anomalies over most areas, especially to the northwest edge of the subtropical high with a positive center denoting a little northward extension but a great westward extension of the western Pacific subtropical high, and the strength of the high stronger than normal. The 500 hPa heights of the third typical mode (Figure 3c) are negative at most areas of the high latitudes, with a center in the middle of Siberia. Additionally, at mid- and low-latitudes, almost positive anomalies can be seen, with a positive center located to the north of the western Pacific subtropical high. This positive center is much stronger than that in Figure 2b, appearing greatly northward of the western Pacific subtropical high. However, the character line of the subtropical high with a value of 588 dgpm maintains over the Sea (while extending westward to the eastern continent of China in Figure 3b) denotes mainly the northward movement of the high. The Jianghuai Basin is controlled by a low trough and is located to the west and north of the western Pacific subtropical high in 500 hPa heights for all of the three typical modes. Additionally, the heavy rainfall results from the intersection with the cold air brought by the westerly trough at the middle and high latitudes and the warm/humid air at the northwest edge of the western Pacific subtropical high from the low latitudes.

On 200 hPa winds (Figure 3d–f) of the three typical modes, the location of the South Asia high and the related jet stream is mainly concerned. In the field of the first typical mode (Figure 3d), the South Asia high is mainly located over the land to the south of 30° N, and the upper jet stream is mainly between 32°–42° N. The Jianghuai Basin is located to the south of the South Asia high and the upper-level jet stream. Additionally, the whole region is divergent, though the center is over the sea. The position of South Asia high of the second typical mode deviates slightly to the north, in comparison with that of the

first typical mode roughly between 35°–45° N but weaker with the exit located far west (Figure 3e). Additionally, there is a divergent center over the Jianghuai area. The position of the South Asia high and upper jet stream for the third typical mode is a little northward of that for the second typical mode, and the jet stream is also weaker and thinner than the exit of the jet stream also a little west (Figure 3f). Additionally, there is a divergence over the north of Jianghuai but a convergence over the south.

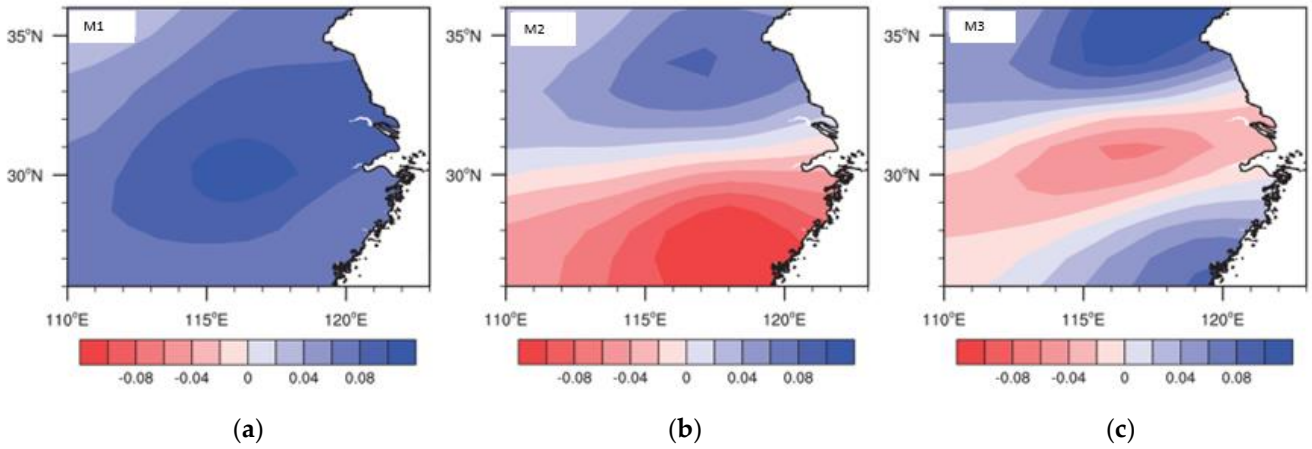


Figure 2. The three lead EOF modes (**a**–**c**) of persistent heavy rain in the Jianghuai Basin.

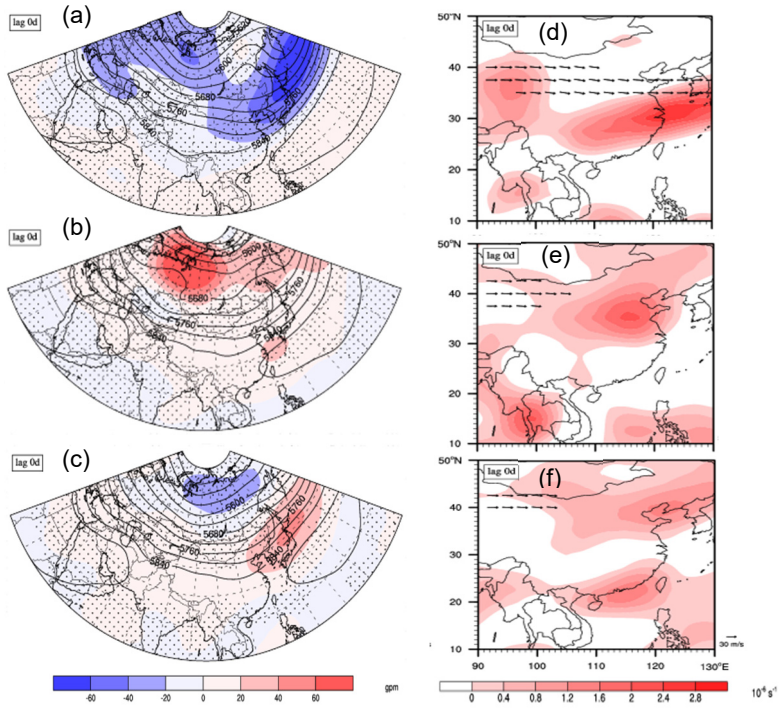


Figure 3. Composite heights at 500 hPa and zonal wind at 200 hPa for Mode1 (**a**,**d**), Mode 2 (**b**,**e**), Mode 3 (**c**,**f**); contour for height, color for height anomaly in (**a**–**c**), the dotted denote the *t*-test exceeding confidence level of 0.1 in (**a**–**c**). Arrows for the mean u $\geq$ 30 m·s^{-1} in (**d**–**f**), color for the anomalous divergence in (**d**–**f**), unit: 10^{-6}·s^{-1}).

On the 850 hPa winds (Figure 4) of the three typical modes, the low-level jet stream and related water vapor transportation are important for heavy rainfall. For the first typical mode, the low-level jet stream is strong, reaching about 30° N, and the southwest wind is dominant in the south of the Jianghuai basin, with water vapor transported mainly from the Bay of Bengal and the South China Sea. For the second typical mode, the low-level jet stream is similar to that of the first typical mode, but it can reach the area farther north

of the basin. The low-level jet stream of the third typical mode is relatively weak in the Jianghuai basin, mainly affecting the southern basin, while the northern part of the region is influenced by the southeast wind from the western Pacific, with the water vapor mainly transported from the South China Sea and the western Pacific.

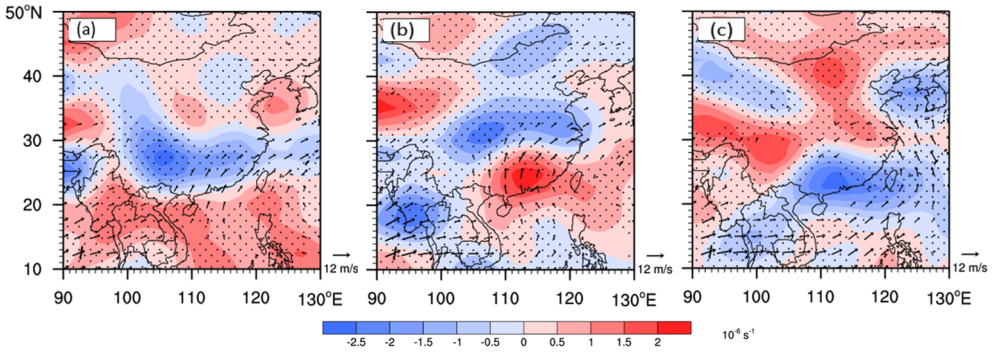


Figure 4. Composite anomaly winds and related divergence at 850 hPa for Mode1 (**a**), Mode2 (**b**), and Mode3 (**c**). (Arrows mean wind speed larger than 3 m/s; shaded areas are divergence, units: 10–6 s^{-1}, dotted areas denote the *t*-test exceeding the confidence level of 0.05).

3.2. Establishing of the CNN for Multi-Level Circulation Pattern Classification

Cai et al. (2022) established a transfer learning CNN model for single-level circulation classification according to the characteristics at 500 hPa heights, and the results show that the transfer learning CNN has better performance in circulation classification than the traditional methods such as the similarity by R and COS [47,49]. However, due to the complex formation of the regional weather and climate, its development will be affected by different key systems at different levels at the same time. In addition to 500 hPa heights with the mid- and high-latitude blocking highs, the westerly trough, the western Pacific subtropical high at low latitudes, etc., there are also weather systems at the upper levels (such as the upper-level jet stream, the South Asian high, and the low-level jet stream) that play important roles in the formation [47,49].

Therefore, considering both the location and intensity of the upper-level jet stream and the warm/humid air brought associated with the lower-level jet stream, multi-level variables are added to the transfer learning CNN to establish the multi-level circulation classification model. According to the evolution characteristics of the three typical modes at 200 hPa in Figure 3 and those at 850 hPa winds in Figure 4, the zonal wind over the area (90°–130° E, 20°–45° N) at 200 hPa and the meridional wind at 850 hPa over the area (90°–130° E, 10°–35° N) are selected as the key areas for the establishment of a multi-level circulation classification model. With the three-level data of 384 d samples, the transfer learning CNN model for three-layer classification is established by using the method of three steps of migration training.

The training data for the pre-trained CNN model to establish the multi-level circulation pattern classification model are the three-level variables of the 284 d samples. When training the CNN network, the transfer learning includes two parts (Figure 5). Firstly, to realize the migration of parameters by initializing the parameters from a pre-trained model before the formal training. The pre-training model not only retains the network's good abilities of image feature extraction and feature learning but also makes the network suitable for specific tasks or problems through freezing the convolution layer parameters and updating the parameters of the full-connection layers. Additionally, the network to be established can also converge faster by loading the parameters of the pre-training model directly, instead of randomly initializing the parameters. Secondly, to train the model three times (three-step transfer learning). The samples used for training (90 case days) are divided into three parts, which are put into the network for training in batches. For the first step of training, there are 284 samples put into the training. Additionally, the corresponding parameter values obtained from the training are retained and brought into the next step of training

as the initial parameter values of the network. In the second step of training, another 30 samples were added (that is, 314 samples), and the parameter values obtained from the second step of training are also retained and brought into the next step of training as the initial parameter values of the network. Similarly, another 30 samples were added (that is, 344 samples) into the network for the third step of training. Additionally, the corresponding parameters should be retained after the three-step network training, with the minimum output values of the network (threshold) for each type counted. The test samples are 40 sample days for all of the three-step transfer learning.

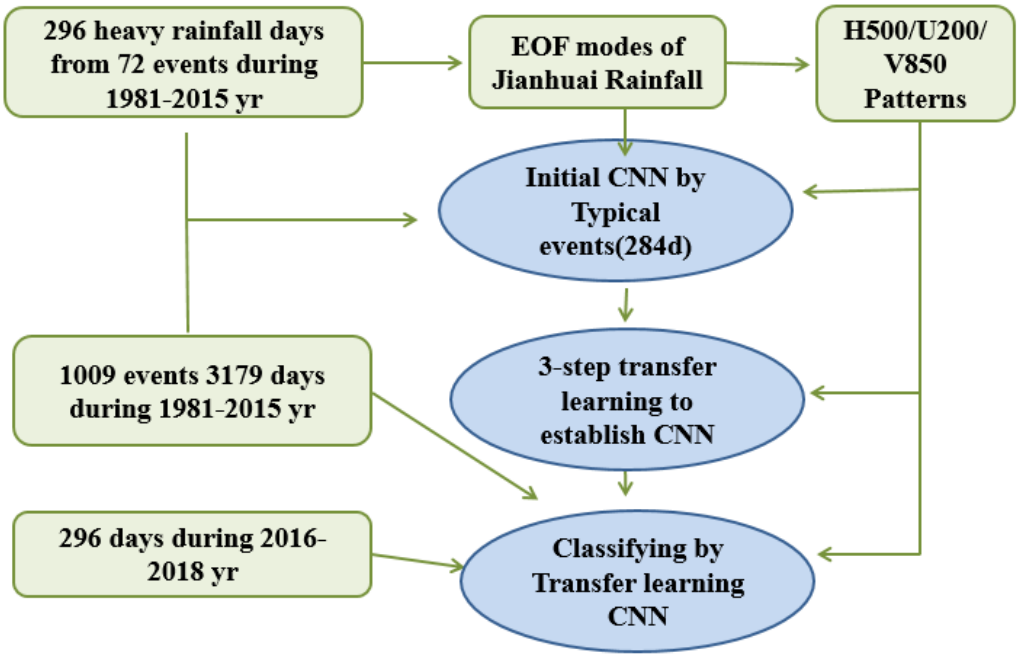
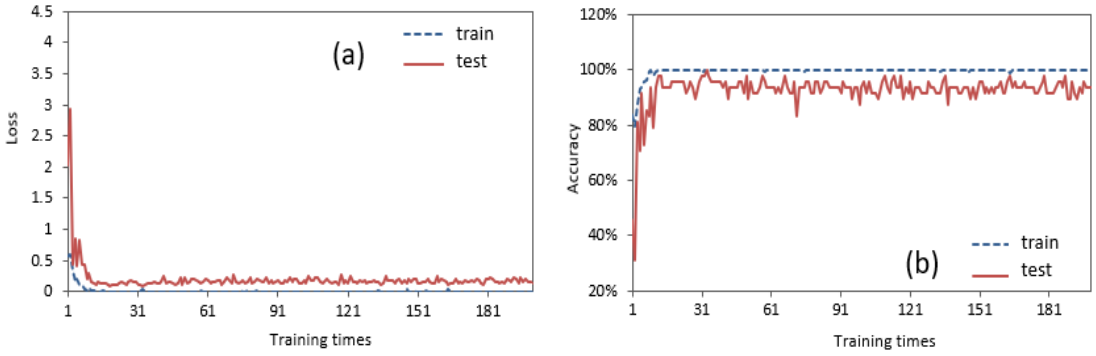


Figure 5. Frame of training for the multi-level transfer learning CNN model.

From the loss function and accuracy rate of the model training (Figure 6), the loss function and accuracy rate of the training converge quickly and reach stability. For the testing, the loss function gradually decreases as the accuracy rate gradually increases. The loss function of the transfer learning CNN for the three-level circulation classification model is smaller and the accuracy rate is higher in comparison with the training results of the single-level transfer learning CNN classification model. The accuracy of the single-level transfer learning CNN test set is 85%, while it is 92.5% in the three-level transfer learning CNN classification model. The corresponding output of the types on the day can be obtained with related multi-level circulation variables input into the established CNN model for typing. When the maximum output value of the CNN model is larger than the threshold value, the corresponding type included in the basic three types can be obtained; otherwise, its type on the day belongs to types other than the three basic types. The threshold value of the output values from the CNN model for selecting each type is 0.997.

After the three-level transfer learning CNN classification model is built and the related threshold values are obtained for each type, the 200 hPa zonal winds, 500 hPa heights, and 850 hPa meridional winds over the key areas of the persistent heavy rainfall events over China can be inputted into the model for circulation classification objectively, with 3179 d samples in total. Additionally, the three types of heavy rainfall samples classified by the CNN model are 748 d, 327 d, and 310 d, respectively. From the variances of the 500 hPa heights between the different patterns obtained by the multi-level variables of the transfer learning CNN and single-layer transfer learning CNN (Figure 7), it can be seen that the variation of inter-pattern variance obtained by the three-level classification CNN model is consistent with that of the single-level model, and the variances between different the

types for heavy rainfall calculated by the three-level classification model are greater than those calculated by the single-level model. The results suggest that the differences of the 500 hPa heights between each type obtained by the three-level model are more obvious, and the three-level model can better distinguish the circulations of different types for heavy rainfall than the single-level model. Additionally, interestingly, the variances between type I and type II, and type I and type III increase with the leading days, which means that the signals for distinguishing the circulation patterns of type I and type II are stronger with longer days ahead of the onset day. As is shown in Figure 1, the heavy rainfall of type I is mostly disastrous. Therefore, the better performance of the multi-level classification is helpful for the further forecasting of extreme events with a longer valid time.



Loss (error) Accuracy

Figure 6. The loss and accuracy of training dataset and test dataset for transfer learning CNN model. The red real lines denote the loss (**a**) and accuracy (**b**) of the test, and the dashed blue lines denote the loss (**a**) and accuracy (**b**) of the training.

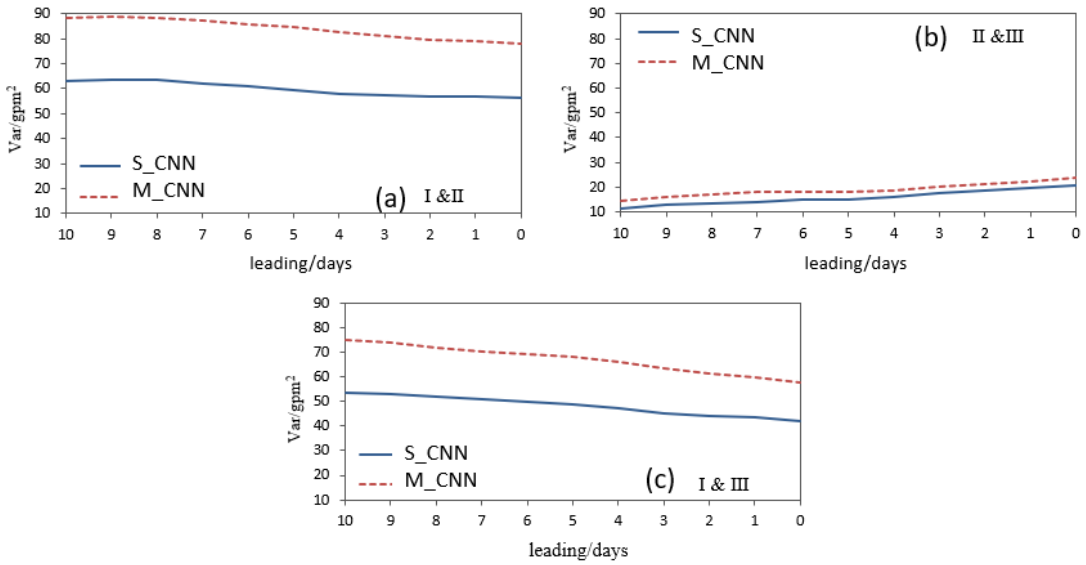


Figure 7. The variance between the 500 hPa height fields of different rainfall types by the CNN model with single-level (S_CNN) and three-level (M_CNN) circulation data (unit: gpm^2). (**a**) Patterns I and II; (**b**) Patterns II and III; (**c**) Patterns I and III.

The spatial distribution of the three types of heavy rainfall is obtained by compositing the rainfall of each type of day as classified by the single-level CNN model (Figure 7a–c) and the three-level CNN model (Figure 8d–f). It can be seen from Figure 8a that the center of the maximum rainfall for Pattern I is located near the middle and a little south of the Basin, the lower reaches of the Yangtze River. For Pattern II (Figure 8b), the composite rainfall is characterized by more rainfall in the north than that in the south, and the center

is located in the north of the basin. Pattern III (Figure 8c) shows more rainfall in the south than that in the north, and the least in the middle, with two rainfall centers in the south and north located in the southwest and northeast of the basin, respectively. Additionally, the rainfall patterns classified by the multi-level model are similar but with larger center values and range, denoting a higher focus of the samples for each type. It suggests that the multi-level model is better. The rainfall pattern distribution from both the single-level and multi-level models is alike to the distribution of the basic modes (Figure 2), suggesting a good effect is able to be achieved by the single-level and multi-level models.

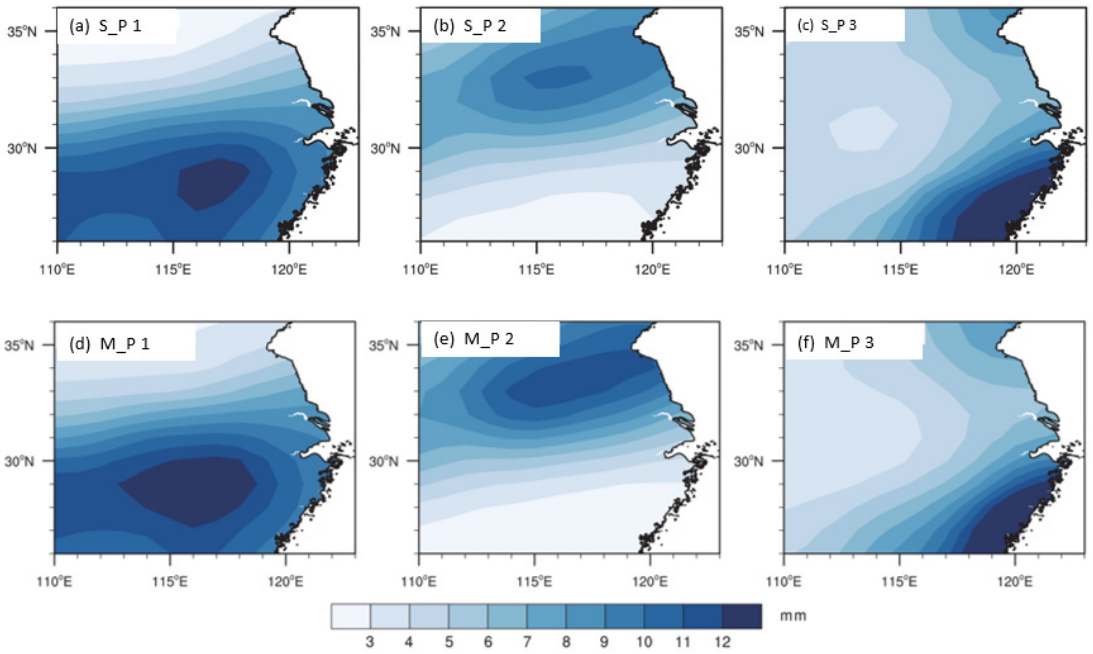


Figure 8. Composite heavy rain patterns in the Jianghuai Region classified by the transfer learning CNN model (unit: mm).

From the rainfall correlation coefficients between the typical modes and each corresponding pattern obtained by the multi-level classification model (Table 2), except that the correlation coefficient between the mean fields of Pattern II (0.970) is slightly smaller than that of the single-level model (0.986), the correlation coefficients of the mean field for both Pattern I and Pattern III by the three-level transfer learning CNN classification model are larger than those by the single-level model. All the correlation coefficients of the three types reached a 0.05 significance level. Additionally, the mean correlation coefficients of all sample days classified by the multi-level model are all greater than those of the single-layer model. The number of days with correlation coefficients reaching a 0.05 significance level is 1093 d in total. In conclusion, compared with the single-level model, the three-level model is better for circulation classification. It indicates that adding multi-level variables such as the 200 hPa zonal winds and 850 hPa meridional winds to the CNN model can improve the classification effect to a certain extent.

Table 2. The correlation coefficients between the transfer learning CNN model and the heavy rainfall of typical mode patterns in samples.

	Methods	**I**	**II**	**III**
Coefficients of mean fields	Single CNN	0.771	0.986	0.913
	3-levels CNN	0.810	0.970	0.916
Mean coefficients of all samples	Single CNN	0.348	0.224	0.382
	3-levels CNN	0.367	0.333	0.417

4. Summary and Discussion

In this paper, the EOF method is used to analyze and refine the typical modes of heavy rainfall in the Jianghuai Basin as the basic type of heavy rainfall. Then, apart from the 500 hPa heights, 200 hPa zonal winds and 850 hPa meridional winds are added to the classification CNN model for transfer learning, and a three-level circulation classification CNN model is established through three-step transfer learning. The main conclusions are as follows:

The basic rainfall and circulation patterns are refined. The EOF decomposition of 72 typical case days from the persistent heavy rainfall events in the Jianghuai Basin is conducted to refine the basic precipitation and circulation pattern datasets for further CNN training. The corresponding time coefficients are obtained by projecting the daily rainfall onto the typical rainfall modes and then used to select the training and test samples for further study.

The three-level transfer learning CNN classification model has higher accuracy than the single-level model in the test. The accuracy of the single-layer transfer learning CNN test set is 85%, while it is 92.5% in the three-layer transfer learning CNN classification model. The rainfall pattern distribution from both the single-level and multi-level CNN is alike to the distribution of the basic modes, suggesting a good effect is able to be achieved by the two CNN models. The rainfall patterns from the patterns classified by the multi-level CNN are similar but with larger center values and range, denoting higher focusing of the samples for each type. All the correlation coefficients of the three types reached a 0.05 significance level. Additionally, the mean correlation coefficients of all sample days classified by the three patterns are all greater than those classified by the single-level model. This suggests that the multi-level model is better.

Circulation is a direct factor responsible for regional weather and climate. Some extreme weather and climate will bring huge losses and impacts to the national economy, social development, and people's lives. If the timeliness and effectiveness of extreme weather forecasting could be improved, it will be better to prevent and mitigate the loss resulting from disasters. At present, the performance of the numerical model on circulation forecasting is much better than that on element forecasting. Therefore, based on the more accurate classification of circulation patterns, the circulation products from the numerical model can be combined for better forecasting [53]. Extreme events usually have close relationships with the circulation such as the height at 500 hPa, the meridional wind component at 850 hPa, the relative humidity at 700 hPa, and the zonal wind at 200 hPa, if they are classified [49]. The examination of the rainfall for the days with daily rainfall of more than 25 mm and 50 mm in the Jianghuai Basin with the analogue method (COS) shows that the TSs in the independent experiments of different lead times from 1 d to 10 d are higher than those from the EC and T639 models for the events with daily rainfall more than 25 mm, and those of 3–10 d for the events with daily rainfall more than 50 mm are also higher than those from the numerical models. That means if we use the classified circulation of the leading time, the corresponding extreme events can be forecasted with higher accuracy and longer valid times than the numerical models. Cai et al. [47] compared the circulation classification for Jianghuai heavy rainfall events obtained by single CNN with those obtained by the R and COS methods; the results show that the single CNN model has the highest classification accuracy on the test dataset. Here, this paper suggests that the multi-level CNN has the best performance in circulation classification. Therefore, the multi-level CNN is a recommendable method in the study of extreme events. On the one hand, the elements forecasted by the numerical model can be corrected in accordance with different types of circulation patterns to improve the prediction performance. On the other hand, it is also possible to extend the valid time of the element forecasting from the numerical model. In this paper, taking the persistent heavy rainfall events in Jianghuai Basin as an example, the multi-level circulation classification of the transfer CNN network is implemented. This method can also be applied to the classification and diagnosis of other disastrous weather or climate events.

Author Contributions: Conceptualization, Y.L. and G.T.; methodology, G.T. and Y.L.; software, J.C.; data curation, Y.L. and J.C.; writing—original draft preparation, Y.L. and J.C.; writing—review and editing G.T. All authors have read and agreed to the published version of the manuscript.

Funding: National Natural Science Foundation of China (Grant No. 42075115) and the Basic science center project of the National Natural Science Foundation of China (42088100).

Institutional Review Board Statement: Not Applicable.

Informed Consent Statement: Not Applicable.

Data Availability Statement: Not Applicable.

Acknowledgments: The authors acknowledge Nanjing University of Information Science and Technology (NUIST) for providing favorable environment and infrastructural needs for conducting research. Special appreciation to all data centers for availing data to use for evaluation studies. The four anonymous reviewers are highly appreciated for the great input that led to the massive improvement of this manuscript.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Sun, Z.; Tan, G.; Zhao, Z. ANN prediction of summer rainfall patterns of east China. *J. Nanjing Inst. Meteor.* **1998**, *21*, 49–52. (In Chinese)
2. Sun, Z.; Tan, G.; Zhao, Z.; Lu, M. Ensemble prediction of summer rainfall patterns over eastern China based on artificial neural networks. *Trans. Atmos. Sci.* **2013**, *36*, 1–6. (In Chinese)
3. Jin, L.; Chen, N.; Lin, Z. Study and comparison of ensemble forecasting based on artificial neural network. *Acta Meteorol. Sin.* **1999**, *57*, 71–80. (In Chinese)
4. Shi, D.P. An application of artificial neural network (ANN) to rainfall mid-range forecasting. *Meteor. Mon.* **2001**, *6*, 40–46. (In Chinese)
5. Xu, X. From Physical Model to Intelligent Analysis: A New Exploration to Reduce the Uncertainty of Weather Forecast. *Meteor. Mon.* **2018**, *44*, 341–350. (In Chinese)
6. Jin, R.H.; Dai, K.; Zhao, R.X.; Cao, Y.; Xue, F.; Liu, C.H. Progress and Challenge of Seamless Fine Gridded Weather Forecasting Technology in China. *Meteor. Mon.* **2019**, *45*, 445–457. (In Chinese)
7. Cervone, G.; Clemente-Harding, L.; Alessandrini, S.; Monache, L.D. Short-term photovoltaic power forecasting using Artificial Neural Networks and an Analog Ensemble. *Renew. Energy* **2017**, *108*, 274–286. [CrossRef]
8. Alessandrini, S.; Sperati, S.; Delle Monache, L. Improving the Analog Ensemble Wind Speed Forecasts for Rare Events. *Mon. Weather Rev.* **2019**, *147*, 2677–2692. [CrossRef]
9. Boukabara, S.-A.; Krasnopolsky, V.; Penny, S.G.; Stewart, J.Q.; McGovern, A.; Hall, D.; Hoeve, J.E.T.; Hickey, J.; Huang, H.-L.A.; Williams, J.K.; et al. Outlook for Exploiting Artificial Intelligence in the Earth and Environmental Sciences. *Bull. Am. Meteor. Soc.* **2021**, *102*, E1016–E1032. [CrossRef]
10. Liu, Y.; Racah, E.; Correa, J.; Khosrowshahi, A.; Lavers, D.; Kunkel, K.; Wehner, M.; Collins, W. Application of deep convolutional neural networks for detecting extreme weather in climate datasets. *arXiv* **2016**, arXiv:1605.01156.
11. Ham, Y.G.; Kim, J.H.; Luo, J.J. Deep learning for Multi-year ENSO Forecasts. *Nature* **2019**, *573*, 568–572. [CrossRef] [PubMed]
12. Franch, G.; Nerini, D.; Pendesini, M.; Coviello, L.; Jurman, G.; Furlanello, C. Precipitation Nowcasting with Orographic Enhanced Stacked Generalization: Improving Deep Learning Predictions on Extreme Events. *Atmosphere* **2020**, *11*, 267. [CrossRef]
13. Chattopadhyay, A.; Nabizadeh, E.; Hassanzadeh, P. Analog Forecasting of Extreme-Causing Weather Patterns Using Deep Learning. *J. Adv. Model. Earth Syst.* **2020**, *12*, e2019MS001958. [CrossRef] [PubMed]
14. Gardner, M.W.; Dorling, S.R. Artificial neural networks (the multilayer perceptron)—A review of applications in the atmospheric sciences. *Atmos. Environ.* **1998**, *32*, 2627–2636. [CrossRef]
15. Marzban, C.; Stumpf, G.J. A Neural Network for Damaging Wind Prediction. *Weather* **1998**, *13*, 151–163. [CrossRef]
16. Marzban, C. Neural networks for postprocessing model output: ARPS. *Mon. Weather Rev.* **2003**, *131*, 1103–1111. [CrossRef]
17. Yuan, H.; Gao, X.; Mullen, S.L.; Sorooshian, S.; Du, J.; Juang, H.M.H. Calibration of probabilistic quantitative precipitation forecasts with an artificial neural network. *Weather* **2007**, *22*, 1287–1303. [CrossRef]
18. McGovern, A.; Lagerquist, R.A.; Gagne, D.; Jergensen, G.E.; Elmore, K.L.; Homeyer, C.R.; Smith, T. Using machine learning and model interpretation and visualization techniques to gain physical insights in atmospheric science. In *AI for Earth Sciences Workshop*; ICLR: Addis Ababa, Ethiopa, 2020.
19. Kasim, M.F.; Watson-Parris, D.; Deaconu, L.; Oliver, S.; Hatfield, P.; Froula, D.H.; Gregori, G.; Jarvis, M.; Khatiwala, S.; Korenaga, J.; et al. Building high accuracy emulators for scientific simulations with deep neural architecture search. *Mach. Learn. Sci. Technol.* **2021**, *3*, 015013. [CrossRef]
20. Dong, C.; Loy, C.C.; He, K.; Tang, X. Learning a Deep Convolutional Network for Image Super-Resolution. In *European Conference on Computer Vision*; Springer: Cham, Switzerland, 2014; pp. 184–199.
21. Wang, X.; Yu, K.; Wu, S.; Gu, J.; Liu, Y.; Dong, C.; Qiao, Y.; Change Loy, C. *ESRGAN: Enhanced Super-Resolution Generative Adversarial Networks*; Springer: Munich, Germany, 2018.

22. Singh, A.; Albert, A.; White, B. Downscaling Numerical Weather Models with GANs. In Proceedings of the Climate Change AI Workshop at the 33rd Conference on Neural Information Processing Systems, Vancouver, BC, Canada, 13 December 2019.
23. Baño-Medina, J.; Manzanas, R.; Gutierrez, J.M. Configuration and intercomparison of deep learning neural models for statistical downscaling. *Geosci. Model Dev.* **2020**, *13*, 2109–2124. [CrossRef]
24. Bremnes, J.B. Ensemble Postprocessing Using Quantile Function Regression Based on Neural Networks and Bernstein Polynomials. *Mon. Weather Rev.* **2020**, *148*, 403–414. [CrossRef]
25. Candido, S.; Singh, A.; Delle Monache, L. Improving Wind Forecasts in the Lower Stratosphere by Distilling an Analog Ensemble into a Deep Neural Network. *Geophys. Res. Lett.* **2020**, *47*, e2020GL089098. [CrossRef]
26. Manepalli, A. Generalization Properties of Machine Learning Based Weather Model Downscaling. In Proceedings of the International Conference on Learning Representations, Online, 26–30 April 2020.
27. Kumar, B.; Chattopadhyay, R.; Singh, M.; Chaudhari, N.; Kodari, K.; Barve, A. Deep learning–based downscaling of summer monsoon rainfall data over Indian region. *Theor. Appl. Climatol.* **2021**, *143*, 1145–1156. [CrossRef]
28. Vannitsem, S.; Bremnes, J.B.; Demaeyer, J.; Evans, G.R.; Flowerdew, J.; Hemri, S.; Lerch, S.; Roberts, N.; Theis, S.; Atencia, A.; et al. Statistical Postprocessing for Weather Forecasts: Review, Challenges, and Avenues in a Big Data World. *Bull. Am. Meteor. Soc.* **2021**, *102*, E681–E699. [CrossRef]
29. Hinton, G.E.; Osindero, S.; Teh, Y.W. A fast learning algorithm for deep belief nets. *Neural Comput.* **2006**, *18*, 1527–1554. [CrossRef]
30. LeCun, Y.; Bengio, Y.; Hinton, G. Deep learning. *Nature* **2015**, *521*, 436–444. [CrossRef]
31. Shi, X.; Chen, Z.; Wang, H.; Yeung, D.Y.; Wong, W.K.; Woo, W.C. Convolutional LSTM Network: A Machine Learning Approach for Precipitation Nowcasting. *arXiv* **2015**, arXiv:1506.04214.
32. Zhang, P.; Zhang, L.; Leung, H.; Wang, J. A Deep-Learning Based Precipitation Forecasting Approach Using Multiple Environmental Factors. In Proceedings of the 2017 IEEE International Congress on Big Data (BigData Congress), Honolulu, HI, USA, 25–30 June 2017.
33. Hu, W.; Cervone, G.; Young, G.; Monache, L.D. Weather Analogs with a Machine Learning Similarity Metric for Renewable Resource Forecasting. *arXiv* **2021**, arXiv:2103.04530.
34. Guastavino, S.; Piana, M.; Tizzi, M.; Cassola, F.; Iengo, A.; Sacchetti, D.; Solazzo, E.; Benvenuto, F. Prediction of severe thunderstorm events with ensemble deep learning and radar data. *arXiv* **2021**, arXiv:2109.09791.
35. Ravuri, S.; Lenc, K.; Willson, M.; Kangin, D.; Lam, R.; Mirowski, P.; Fitzsimons, M.; Athanassiadou, M.; Kashem, S.; Madge, S.; et al. Skilful precipitation nowcasting using deep generative models of radar. *Nature* **2021**, *597*, 672–677. [CrossRef]
36. Long, J.; Shelhamer, E.; Darrell, T. Fully Convolutional Networks for Semantic Segmentation. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, San Diego, CA, USA, 7–9 May 2015; pp. 3431–3440. [CrossRef]
37. Ronneberger, O.; Fischer, P.; Brox, T. U-Net: Convolutional Networks for Biomedical Image Segmentation. *arXiv* **2015**, arXiv:1505.04597.
38. Dupuy, F. ARPEGE Cloud Cover Forecast Post-Processing with Convolutional Neural Network. *arXiv* **2020**, arXiv:2006.16678.
39. Grönquist, P.; Yao, C.; Ben-Nun, T.; Dryden, N.; Dueben, P.; Li, S.; Hoefler, T. Deep learning for post-processing ensemble weather forecasts. *Philos. Trans. A Math Phys. Eng. Sci.* **2021**, *379*, 20200092. [CrossRef] [PubMed]
40. Krizhevsky, A.; Sutskever, I.; Hinton, G.E. ImageNet Classification with Deep Convolutional Neural Networks. *Commun. ACM* **2017**, *60*, 84–90. [CrossRef]
41. Scher, S.; Messori, G. Predicting weather forecast uncertainty with machine learning. *Q. J. R. Meteor. Soc.* **2018**, *144*, 2830–2841. [CrossRef]
42. Ayzel, G.; Scheffer, T.; Heistermann, M. RainNet v1.0: A convolutional neural network for radar-based precipitation nowcasting. *Geosci. Model Dev.* **2020**, *13*, 2631–2644. [CrossRef]
43. Men, X.L.; Jiao, R.L.; Wang, D.; Zhao, C.G.; Liu, Y.K.; Xia, J.J.; Li, H.C.; Yan, Z.W.; Sun, J.H.; Wang, L.Z. A Temperature Correction Method for Multi-model Ensemble Forecast in North China Based on Machine Learning. *Clim. Environ. Res.* **2019**, *24*, 116–124. (In Chinese)
44. Chen, J.P.; Feng, Y.D.; Meng, W.G. A Correction Method of Hourly Precipitation Forecast Based on Convolutional Neural Network. *Meteor. Mon.* **2021**, *47*, 60–70. (In Chinese)
45. Chen, C. *Application of Convolution Neural Network in Weather Nowcasting*; South China University of Technology: Guangzhou, China, 2018.
46. Cai, J.; Tan, G.; Niu, R. Circulation pattern classification of persistent heavy rainfall in Jianghuai Region based on the transfer learning CNN model. *J. Appl. Meteor. Sci.* **2021**, *32*, 233–244. [CrossRef]
47. Luo, Y.; Nie, X.W.; Wang, G.S. An exploration on the applicability of similarity parameter in similarity forecasting. *Meteor. Mon.* **2011**, *37*, 1443–1447. (In Chinese)
48. Tan, G.R.; Fan, Y.Y.; Niu, R.Y. 2018. Pattern classification of heavy rainfall in Jianghuai region and associated circulations. *J. Appl. Meteor. Sci.* **2018**, *29*, 396–409.
49. Niu, R.; Liu, C.; Liu, W.; Zhao, X. Characteristics of temporal and spatial distribution of regional rainstorm processes to the east of 95°E in China during 1981–2015. *Acta Meteorol. Sin.* **2018**, *76*, 182–195. [CrossRef]
50. Kalnay, E.; Kanamitsu, M.; Kistler, R.; Collins, W.; Deaven, D.; Gandin, L.; Iredell, M.; Saha, S.; White, G.; Woollen, J.; et al. The NCEP/NCAR 40-year reanalysis project. *Bull. Am. Meteorol. Soc.* **1996**, *77*, 437–470. [CrossRef]

51. Sun, J.; Wang, H.; Wei, J.; Qi, L. The sources and transportation of water vapor in persistent heavy rainfall events in the Yangtze-Huaihe River Valley. *Acta Meteorol. Sin.* **2016**, *74*, 542–555. (In Chinese)
52. He, K.; Zhang, X.; Ren, S.; Sun, J. Deep residual learning for image recognition. In Proceedings of the 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Las Vegas, NV, USA, 26 June–1 July 2016; pp. 770–778.
53. Tan, G.R.; Duan, H.; Ren, H.L. Statistical correction for dynamical prediction of 500hPa height field in mid-high latitudes. *J. Appl. Meteor. Sci.* **2012**, *23*, 304–311. (In Chinese)