Extreme Weather Recognition using Convolutional Neural Networks

Ziqi Zhu[1], Li Zhuo[1,2], Panling Qu[1], Kailong Zhou[1], Jing Zhang[1]

[1]Signal & Information Processing Laboratory
Beijing University of Technology
Beijing , China
[2]Collaborative Innovation Center of Electric Vehicles in Beijing
Beijing , China
zhuziqi0228@gmail.com, zhuoli@bjut.edu.cn, ursular86@sina.com, zklkey@emails.bjut.edu.cn, zhj@bjut.edu.cn

***Abstract*—Extreme weather always brings potential risk to driving, which leads to people's life and property being put into great dangers. Therefore, the automatic recognition of extreme weather plays an important role in the application of the highway traffic condition warning, automobile auxiliary driving, climate analysis and so on. Generally, multiple sensors are adopted in traditional methods of automatic extreme weather recognition with artificial participation and low accuracy. A new extreme weather recognition method based on images by using computer vision manners has been proposed in this paper. Since the weather is affected by many factors, features that can accurately represent various weather characteristics are difficult to be extracted. Therefore, in this paper, convolutional neural networks (CNNs) are applied to settle this problem. Features of extreme weather and recognition models are generated from big data. Moreover, a large-scale extreme weather dataset, "WeatherDataset", has been collected, in which 16635 extreme weather images are divided into four classes (sunny, rainstorm, blizzard, and fog), and complex scenes are coverd. A recognition model for extreme weather is obtained through two steps: Pre-training and Fine Tuning. In Pre-training step, ILSVRC-2012 Dataset is trained to obtain the model of ILSVRC using GoogLeNet. A more accurate model for extreme weather recognition is obatined by further fine-tuning GoogLeNet on WeatherDataset. The experimental results show that the proposed method is able to achieve a high performance with the recognition accuracy rate of 94.5% and can meet the requirements of some real applications.**

Keywords-extreme weather recognition; weather dataset; convolutional neural networks; GoogLeNet; fine-tuning

I. INTRODUCTION

Driving on highway is highly sensitive to meteorological conditions, and the highway safety problems caused by extreme weather conditions have been the focus of attention. In the field of traffic engineering, rainstorm, blizzard, and fog are three kinds of most studied extreme weather, which will lead to visibility and friction coefficient of road reduced, resulting in tremendous potential dangers. Therefore, automatically recognizing extreme weather is essential for many applications, such as highway traffic condition warning, automobile auxiliary driving, climate analysis and so on.

At present, the methods of weather recognition are mainly based on multiple sensors. However, the installation and maintenance of sensors will consume a lot of manpower and material. Moreover, the recognition accuracy will be affected by environment. In recent years, with the development of the intelligent transportation system, all kinds of monitoring equipments are installed on the roads, thus the weather recognition method based on image processing has been gradually developed. Dynamic features from the image sequence have been extracted by using mixture of Gaussian (MOG) model and Fourier Transform of input sequence have been proposed in [1]. Fog and snow could be recognized, however the recognition error rate of snow is very high, reaching to 19.8%.

In [2], Chen Z *et al.* proposed a weather recognition method based on captured images. Sky region was firstly extracted to avoid disturbance of the non-sky region. Then multiple kernel learning was applied to gather and select an optimal subset of image features from several kinds of features. In order to further improve the recognition performance, the training set was obtained by the method of multi-pass active. Finally, sunny, cloudy and overcast can be recognized by support vector machine (SVM) classifier, but the dataset in [2] is collected from a fixed location, which can not be applied to other scenarios.

In [3], the data are acquired by many sensors, including the Light Detection and Recognition (LIDAR) [4] instruments, cameras, etc. The weather recognition result is obtained by fusing the data of multiple sensors and images. But the recognition accuracy will be greatly affected if one of the sensors is out of work.

The method of image processing was used to recognize weather in [5]. Firstly features of weather images was extracted, such as the power spectrum slope, contrast, noise and saturation and so on. Then SVM classifier based on decision tree was trained by the features. Finally, sunny, fog, rain, overcast can be recognized by the trained SVM

978-1-5090-4571-6/16 $31.00 © 2016 IEEE
DOI 10.1109/ISM.2016.81

classifier. The method has been tested on the Wild [6] Dataset which contains hundreds of images, and the recognition error rates of fog and rain were 15% and 25% respectively, which can not meet the practical requirements. In addition, the images are captured at a fixed location and the number of images is not enough, which leads to the lack of authority.

The problems of the existing weather recognition methods can be summarized as follows:

1、In the most methods, only two kinds of weather can be recognized, which is hard to be applied in practical applications;

2、The installation and maintenance of sensors will consume a lot of manpower and material, the accuracy may be affected by environment;

3、Only a small data set is used for training and recognition, which results in the lack of robustness.

In order to address the above problems, a large-scale weather dataset “WeatherDataset” is collected in this paper, in which 16635 images are divided into four types (sunny, rainstorm, blizzard, and fog), and complex scenes such as urban roads and highways are coverd. Features that can represent various weather characteristics accurately are difficult to be extracted due to complex scenes. In recent years, convolutional neural networks (CNNs) [7] have achieved state-of-the-art results on object detection [8, 9], human face recognition, segmentation [10] and retrieval [11]. It also shows excellent performance in many challenging tasks of recognition. Effective features can be automatically extracted by CNNs from the big data without mannually selected. Therefore, CNNs are adopted and trained on WeatherDataset in this paper, by which extreme weather can be automatically recognized and the high recognition accuracy rate of 94.5% can be obtained.

II. THE PROPOSED METHOD

CNNs have a high degree of invariance to displacement, scale and deformation, and the recognition task of natural images especially in the complex real-world environment has shown excellent performance. In this paper, CNNs are applied for extreme weather recognition.

The implementation of CNNs is showed in Fig.1. Firstly, the operations of convolution and downsampling are alternately done on the input images, resulting in a hierarchy of increasingly complex features. Then one or more fully connected layers are connected to the former part. With the increasing number of layers, features learned are more and more abstract.

It is an important part for CNNs to get feature maps from each layer. Feature maps in different layers can be calculated as follows, the value at position (x, y) in the n th feature map of the m th layer, denoted as v_{mn}^{xy}, is given by

$$v_{mn}^{xy} = \max(\sum_{k}\sum_{i=0}^{I_m-1}\sum_{j=0}^{J_m-1} w_{mnk}^{ij} v_{(m-1)k}^{(x+i)(y+j)} + b_{mn}, 0) \qquad (1)$$

where k indexes over the set of feature maps in the previous layer connected to the current feature map, w_{mnk}^{ij} is the value at the position (i, j) of the kernel connected to the k th feature map, I_m and J_m are the height and width of the kernel respectively, and b_{mn} *is* the bias for this feature map.

At present, there are a lot of CNN structures. In this paper, GoogLeNet [12] is adopted to achieve weather recognition. Compared with state-of-art network structures, GoogLeNet is deeper, which has 22 layers.

The flowchart of training process is showed in Fig.2. The whole training process consists of two key steps: Pre-training and Fine Tuning. In Pre-training step, ILSVRC-2012 [13] Dataset is trained to obtain the model of ILSVRC using GoogLeNet. The model is further fine-tuned on WeatherDataset to obatin more accurate model for the extreme weather recognition. The images randomly selected from WeatherDataset will be scaled to 224×224, and then, be inputted to GoogLeNet as the training set and the recognition model is obtained through a large number of iterations.

III. THE CONSTRUCTED DATASET

Since there is lack of large-scale weather data set, a data set “WeatherDataset” with 16635 images is constructed in this paper. The images are collected from the web and selected according to the requirements.

The weather in WeatherDataset can be devided into four classes: sunny, rainstorm, blizzard, and fog. Conmplex scenes such as city roads and highway are contained. The image modes cover aerial photography, camera, news, screenshot of movie, traffic accidents, automobile data recorder, etc. Since the images are captured in multiple angles, it is a chanllenging task to recognize weather images in WeatherDataset.

Four kinds of sample weather images in WeatherDataset are shown in Fig.3 (a) ~ Fig.3 (d).

IV. EXPERIMENTAL RESULTS AND ANALYSIS

In order to validate the effectiveness of the method proposed in this paper, the algorithm is implemented based on Caffe [14] open source code. Experiments have been carried out on three network structures: AlexNet [15], modified AlexNet and GoogLeNet. The experiment are as follows:

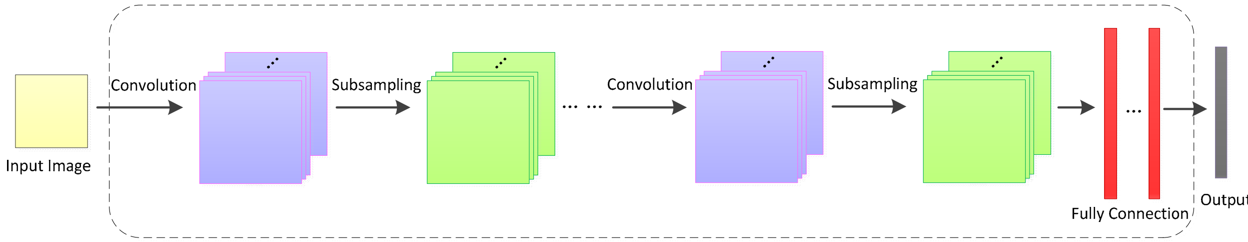

Fig.1 CNN training process

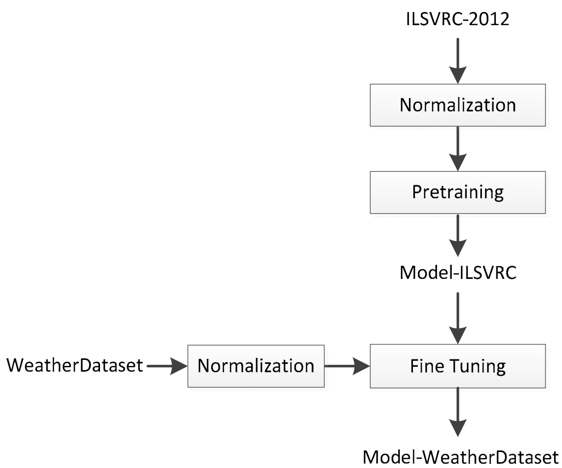

Fig.2 Training process

A. *Without fine-tuning*

The images in WeatherDataset are directly input and the recognition result is obtained through CNN. The momentum and the weight decay are set to 0.9 and 0.0002 respectively. All the models are initialized with learning rate of 0.01 and this value is reduced by 10% after each two thousand iterations.

B. *Fine-tuning*

Firstly, CNN is pre-trained on the dataset of ILSVRC-2012, then the network is fine-tuned by WeatherDataset and the weights are updated. The momentum and the weight decay are set to 0.9 and 0.0002 respectively. All the models are initialized with learning rate of 0.001 and this value is reduced by 4% after each two thousand iterations. To evaluate the effectiveness of GoogLeNet, the recognition result of GoogLeNet is compared with those of AlexNet and modified AlexNet. These three kinds of networks are all pre-trained on the ILSVRC-2012 data set.

WeatherDataset is split by assigning 80% of the images as the training set and 20% as the test set. Both of the training set and the test set are selected randomly.

The images in WeatherDataset are recognized by AlexNet, modified AlexNet and GoogLeNet respectively. And the number of iterations for the three networks without fine-tuning is set to 55000. The recognition results are shown in Table I.

From Table I, it can be seen that :

1、No matter what kind of network structure is, the network with fine-tuning will achieve higher recognition accuracy compared with those without fine-tuning. Due to the lack of sufficient data, it is difficult for CNN to learn effective representations directly from WeatherDataset. Therefore, CNN is pre-trained on the dataset of ILSVRC-2012 to overcome the impact of lack of data.

2、GoogLeNet can obtain the best recognition accuracy of 94.5% with fine-tuning.

The optimal recognition result can be achieved at forty thousand iterations for GoogLeNet with fine-tuning. In the case that the number of iterations is twenty thsousand, the recognition accuracy of the modified AlexNet and AlexNet can reach to 93.7% and 94.2% respectively, and both of them outperform GoogLeNet. The number of iterations that can obtain highest recognition accuracy for modified AlexNet and AlexNet is set to forty thousand, and the recognition accuracy can reach to 94.3% and 93.7% respectively, however, the recognition accuracy of GoogLeNet with forty thousand iterations exceeds them. The recognition accuracy of GoogLeNet with twenty thousand iterations and forty thousand iterations can reach to 93.3% and 94.5% respectively. This indicates that GoogLeNet is deeper than other networks in this paper and more iterations needed to achieve excellent performance.

The experimental results show that high recognition accuracy can be obtained by the proposed method based on CNN in this paper, which can be used into the practical applications.

V. CONCLUSION

In this paper, CNN is applied to realize the extreme weather recognition, and the recognition performance can meet the demand of practical application. In the following work, WeatherDataset constructed in this paper will be extended with more categories, such as freeze, sand storm. Moreover, the category should be subdivided, such as light rain, moderate rain and heavy rain, etc.

(a)Fog (b) Rainstorm (c)Sunny (d)Blizzard

Fig.3 Four kinds of sample weather images in WeatherDataset

TABLE I
THE RECOGNITION RESULTS OF THREE NETWORKS WITH FINE-TUNING OR NOT

	Model	**Accuracy**
Without fine-tuning	AlexNet	89.7%
	modified AlexNet	90.6%
	GoogLeNet	90.1%
Fine-tuning	AlexNet(20000 iterations)	93.5%
	AlexNet(40000 iterations)	93.7%
	modified AlexNet(20000 iterations)	94.2%
	modified AlexNet(40000 iterations)	94.3%
	GoogLeNet(20000 iterations)	93.3%
	GoogLeNet(40000 iterations)	**94.5%**

ACKNOWLEDGMENT

The work in this paper is supported by the National Natural Science Foundation of China (No. 61531006, No. 61372149, No. 61370189, and No. 61471013), the Importation and Development of High-Caliber Talents Project of Beijing Municipal Institutions (No.CIT&TCD20150311, CIT&TCD201404043), the Beijing Natural Science Foundation (No.4142009, No.4163071), the Science and Technology Development Program of Beijing Education Committee (No.KM201410005002, No. KM201510005004), Funding Project for Academic Human Resources Development in Institutions of Higher Learning Under the Jurisdiction of Beijing Municipality.

REFERENCES

[1] Lagorio. A, Grosso. E, and Tistarelli. M, "Automatic detection of adverse weather conditions in traffic scenes," in *Advanced Video and Signal Based Surveillance,* 2008, pp. 273-279.

[2] Chen. Z, Yang. F, Lindner. A, Barrenetxea. G, and Vetterli. M, "HOW IS THE WEATHER: AUTOMATIC INFERENCE FROM IMAGES," in *Image Processing, 2012 19th IEEE International Conference on. IEEE,* 2012, pp. 1853-1856.

[3] Dannheim. C, Icking. C, Mader. M, and Sallis. P, "Weather Detection in Vehicles by Means of Camera and LIDAR Systems," in *Computational Intelligence, Communication Systems and Networks, 2014 Sixth International Conference on. IEEE,* 2014, pp. 186-191.

[4] Bluesky International Limited. "How LIDAR works". [Online]. Available: http://www.lidar-uk.com/how-lidar-works.

[5] Li. Q, Kong. Y, and Xia. S. M, "A method of weather recognition based on outdoor images." *Computer Vision Theory and Applications.,* vol. 2, pp. 510-516, Jan. 2014.

[6] Narasimhan. S. G, Wang. C, and Nayar. S. K, “All the images of an outdoor scene,” in *Computer Vision—ECCV 2002*, 2002, pp. 148-162.

[7] LeCun. Y, Boser. B, Denker. J. S, Henderson. D, Howard. R. E, Hubbard. W, and Jackel. L. D, “Backpropagation applied to handwritten zip code recognition,” *Neural computation.*, vol.1, pp. 541-551, 1989.

[8] Girshick. R, “Fast R-CNN,” in *Proceedings of the IEEE International Conference on Computer Vision*, 2015, pp. 1440-1448.

[9] Ren. S, He. K, Girshick. R, and Sun. J, “Faster R-CNN: Towards real-time object detection with region proposal networks,” in *Advances in Neural Information Processing Systems*, 2015, pp. 91-99.

[10] Gacsádi. A, and Szolgay. P, “Variational computing based segmentation methods for medical imaging by using CNN,” in *12th international Workshop on Cellular Neural Networks and Their Applicaiton*, vo1. 1, 2012.

[11] Ku. W. L, Chou. H. C, and Peng. W. H, “Discriminatively-learned global image representation using CNN as a local feature extractor for image retrieval,” in *Visual Communications and Image Processing* , 2015, pp. 1-4.

[12] Szegedy. C, Liu. W, Jia. Y, Sermanet. P, Reed. S, Anguelov. D, and Rabinovich. A, “Going deeper with convolutions,” in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*,2015, pp. 1-9.

[13] O. Russakovsky, J. Deng, H. Su, J. Krause, S. Satheesh, S. Ma, Z. Huang, A. Karpathy, A. Khosla, M. Bernstein, A. C. Berg, and L. Fei-Fei, “ImageNet Large Scale Visual Recognition Challenge,” *International Journal of Computer Vision*, vol.115, pp. 221-252, 2015.

[14] Jia. Y, Shelhamer. E, Donahue. J, Karayev. S, Long. J, Girshick R, Guadarrama. S, and Darrell.T, “Caffe: Convolutional architecture for fast feature embedding,” in *Proceedings of the ACM International Conference on Multimedia*, 2014, pp. 675-678.

[15] Krizhevsky. A, Sutskever. I, Hinton. G. E, “ImageNet classification with deep convolutional neural networks,” in *Advances in neural information processing systems*, 2012, pp. 1097-1105.