Recognition of road in bad weather using deep learning

Yu Nakagawa, Huimin Lu, Kenta Kagemoto, Nobuhiro Hirano, Shiyuan Yang, Seiichi Serikawa

Kyushu Institute of Technology, 1-1 Sensui-cho Tobata-ku Kitakyushu-shi Fukuoka-ken, 804-8550, Japan

*Corresponding Author: nkyu1077@gmail.com

Abstract

Although the development and research of automatic driving has advanced in recent years, it has not yet been realized. We think that recognition of objects is necessary to make automated driving practical. It is necessary to grasp the course regardless of the weather. Use deep learning to recognize objects. We use CNN which is suitable for image recognition even in depth learning. Recognition of objects during rainy weather has not yet been realized, and it is necessary to create that database. Creating databases photographs the roads themselves and collects many databases. Make sure that the road in various situations can be recognized as a road. Then, it learns the database by deep learning and identifies whether it can be recognized as a road.

Keywords: deep learning, road, recognition.

1. Introduction

In recent years, artificial intelligence (AI) has been used in various fields such as cleaning robot, pepper, google's automatic driving car and so on. The development and research of artificial intelligence (AI) is done day by day, and it is a technology which is constantly drawing attention. It is expected that artificial intelligence (AI) will be used more closely with the development of automatic driving technology in the future. While driver assistance systems are being introduced in automatic operation, perfect automatic driving vehicles are not yet practical. Recently, there are many elderly accidents in automobiles, which is also a problem. In recent years the number of traffic accidents of elderly people is also on the rise, and with the aging of the population the number of elderly people will increase. Therefore, if the automatic driving technique develops, the number of accidents will be reduced. Although there is support for drivers such as automatic braking, although it is decreasing as the number of accidents, it can not be said to be zero at all. Realization of automatic driving vehicles is

expected to eliminate accidents. For example, since accidents have occurred in various cases, such as vehicles that put too much speed, ignoring temporary stops, dozing driving, ignoring signals, cars running in opposite lanes, if automatic driving becomes a reality in the future, I think that things should not be done. In order to realize automatic driving, grasping of objects is important as a first step. We must grasp all kinds of objects such as people, cars, roads and signals. Machine learning that can recognize people, cars are present many times but the image where noise occurred such as road recognition at the time of rain is not yet accurately made. Advancement of artificial intelligence (AI) technology has made it possible to extract feature quantities, and it has become increasingly possible to machine the optimum work using the search data and behavior data of individual features. If it can be used and grasping the road at the time of rain, automatic operation becomes possible even in fine weather and rainy weather. In this research, we use the neural network in CNN, which is widely used for image processing, to recognize roads in rainy weather. In recognition, when we compare rainy weather and sunny weather, noise is generated, the sky and the road become similar colors, and we think that recognition is difficult. It is deep learning by Caffe and has a fast speed feature which attention now is made to recognize an image with CNN. Therefore, in this research, we aim to recognize more accurate rainy roads by Caffe. I believe that deep learning by Caffe is still indispensable for accurately recognizing images.

2. Principle of deep learning

2.1 Deep learning

Describe the principle of deep learning. In a word, deep learning is a generic term for "algorithms using multilayered neural networks with multiple hidden layers with deep hierarchies". In recent years, it was confirmed that the problem using artificial intelligence shows that the method using this deep net shows a performance far superior to the

conventional method. Deep learning is a technology that is one step ahead of machine learning, and allows machines to have more difficult recognition, cognition and discrimination functions. The neural network is a computer processing mechanism modeling the neural network of the brain's brain. In deep learning, depending on the object to be applied, it often shows better performance than machine learning, and its scope of application is spreading gradually. A diagram of depth learning is shown in Fig. 1

Fig. 1: Deep learning

Bag of features uses "image features" for image recognition. Whether or not the image feature points are "similar points" among a plurality of images is compared. By comparing the distances between the image feature quantities, it is possible to measure the closeness of image feature points and the similarity of image feature points. The basic concept of Bag of features is the idea that "images with the same type of objects have many similar parts". Compare parts and parts of objects and take out images partly. The disadvantages are that it is influenced by variations in density of image feature points, that the positional relationship is ignored in the structure of the data to be handled, and that it is affected by changes in appearance. All three weaknesses are derived from image feature points and image feature values. In other words, the properties of image feature points and image feature quantities will affect the accuracy of image recognition.

2.2 Image recognition using CNN

Various algorithms using deep hierarchical neural networks as deep learning algorithms have been proposed. Among them, Convolution Neural Network, Auto Encoder, Recurrent Neural Network and others are commonly used. Since CNN has high performance in image recognition, the algorithm of CNN is described below. Convolution Neural

Network (CNN) is an algorithm using a neural network having layers called "convolution layer" and "pooling layer". CNN is a method of information processing that got hints on the visual system in the brain of an organism. The convolution layer refers to a layer having a structure such that a group of nodes close to each other in the previous layer connect with nodes of the next layer. The convolution layer plays a role of abstracting a local part of an image. The pooling layer refers to the process of collecting the local portions of the nodes of the previous layer. In particular, the process of taking the maximum value of the local part is called "maximum pooling". The pooling layer plays a role of imparting invariance to small translation by lowering sensitivity to position change. This is because the pooling layer returns the local maximum value to the node of the next layer even if there is a slight translation, so the value of the node of the next layer does not change. CNN has the property that even if there are multiple layers basically no prior learning is required. In other words, supervised learning (stochastic gradient descent method) can be executed from the initial state without problems and good performance is achieved. The convolution layer and the pooling layer play a role of giving abstraction and local invariance of local information of images. This structure is said to be close to the structure of the human visual field, and CNN is considered to be one of the reasons for high performance in image recognition. CNN is basically supervised learning, which requires a large amount of learning data given annotations. Prepare a large amount of data that is a set of images and information indicating the type as inputs, and keep all input sizes. We prepare a neural network to be used for CNN, make the input layer the same size as the image size, and make the number of nodes of the output layer the same as the number of kinds. Let the teacher data of each image be an output vector $(0, \dots, 0, 1, 0, \dots, 0)$ that sets the dimension representing the type of the image as 1 and the other dimension as 0. Then, using the input image and the output vector as a set, learning of the neural network is performed. The learning model is a neural network whose edge weights are optimized by this learning. A method of discriminating the type of an unknown image can be obtained through a neural network by using a neural network created as a learning model and assuming the value of the input layer as the value of the input image. Learning processing prepares training data that combines images and types of objects, prepares a neural network, learns the neural network using training data. In the determination process, an unknown image is input to the neural network, and from

which output type it is determined which kind. Details and structure of CNN are shown in Fig. 2 and Fig. 3 respectively.

Fig. 2: Detailed view of CNN

Fig. 3: Structure of CNN

3. Experiments

The database of roads is very few with regard to image collection, and there are particularly few roads on rainy days and evening. Therefore, it is necessary to take pictures and videos of roads in large quantities and collect data with their own hands. Today, this research shoots the road using a smartphone. Approximately 200 images of the image of the road taken are taken as correct images and about 200 images are selected from the public images on the Internet as noncorrect images. This time I created a database of myself, morning, afternoon and evening rain on sunny weather and rainy weather, and let Caffe's CNN recognize the image. Figures 4 and 5 show examples of images of roads in rainy weather and roads in fine weather, respectively.

Fig. 4: Examples of roads in rainy weather

Fig. 5: Examples of roads in fine weather

The image size was converted to $256 * 256$ LMDB. We need to prepare two databases for training and test, and prepared them. I also prepared prototxt and defined my own model. Prototxt prepared three, deploy, train_test, solver. Next, I confirmed whether I surely grasp that the road is a road. We prepared 200 learning data of the road as a correct image as shown in the figure, 200 pieces of incorrect answer images, and 20 test images of road images of various situations. Examples of correct and non-correct images of learning data are shown. These images were learned by error back propagation method of CNN of all eight layers. After the learning, when the test was carried out, it was possible to recognize all 20 test images as roads. From the results it was possible to recognize the image at the time of rain as a road as it was at fine weather. The experimental results are shown in the figure 6.

```

I0123 13:43:01.074151 20643 caffe.cpp:308] Batch 41, accuracy = 1
I0123 13:43:01.074184 20643 caffe.cpp:308] Batch 41, loss = 0
I0123 13:43:09.168079 20643 caffe.cpp:308] Batch 42, accuracy = 1
I0123 13:43:09.168282 20643 caffe.cpp:308] Batch 42, loss = 0
I0123 13:43:17.623674 20643 caffe.cpp:308] Batch 43, accuracy = 1
I0123 13:43:17.623725 20643 caffe.cpp:308] Batch 43, loss = 0
I0123 13:43:25.730238 20643 caffe.cpp:308] Batch 44, accuracy = 1
I0123 13:43:25.730270 20643 caffe.cpp:308] Batch 44, loss = 0
I0123 13:43:34.213171 20643 caffe.cpp:308] Batch 45, accuracy = 1
I0123 13:43:34.213207 20643 caffe.cpp:308] Batch 45, loss = 0
I0123 13:43:42.303679 20643 caffe.cpp:308] Batch 46, accuracy = 1
I0123 13:43:42.303874 20643 caffe.cpp:308] Batch 46, loss = 0
I0123 13:43:50.857201 20643 caffe.cpp:308] Batch 47, accuracy = 1
I0123 13:43:50.857236 20643 caffe.cpp:308] Batch 47, loss = 0
I0123 13:43:58.986466 20643 caffe.cpp:308] Batch 48, accuracy = 1
I0123 13:43:58.986496 20643 caffe.cpp:308] Batch 48, loss = 0
I0123 13:44:07.479311 20643 caffe.cpp:308] Batch 49, accuracy = 1
I0123 13:44:07.479360 20643 caffe.cpp:308] Batch 49, loss = 0
I0123 13:44:07.479368 20643 caffe.cpp:313] Loss: 0
I0123 13:44:07.479380 20643 caffe.cpp:325] accuracy = 1
I0123 13:44:07.479390 20643 caffe.cpp:325] loss = 0 (* 1 = 0 loss)
yu@yu:~/caffe$

```

Fig. 6: Experimental result

4. Conclusions

In this study, we were able to recognize the rainy road using CNN deep learning by Caffe. As the future task is to increase the number of databases is still small, and building databases in various situations such as snow and fog. We also found that it is also necessary to improve the layer so that various objects can be recognized. We plan to expand the range of recognition so that it can run automatically in the future, to conduct more research on making more classified images, more classified into multiple classes, and more practical images.

References

- (1) Takayuki Okatani : “deep learning”, pp. 80-110, 2015
- (2) Hiromasa Takei : “The first deep learning”, pp. 54-95, 2015
- (3) Hideki Asou : “deep learning”, pp 1-250, 2015
- (4) Ian Lenz : “Deep Learning for Detecting Robotic Grasps”
- (5) Chelsea Finn : “Learning Visual Feature Spaces for Robotic Manipulation with Deep Spatial Autoencoders”