International Journal of Innovative Technology and Exploring Engineering (IJITEE)
ISSN: 2278-3075, Volume-8 Issue-4S2 March, 2019

Classification of the Loading Type of Trucks using Convolutional Neural Network

DongGyu Lee

Abstract--- This paper proposes a classification method using Convolutional Neural Network(CNN) to classify the types of a truck. The images of the vehicle from the camera are classified according to the vehicle type and the cargo compartment. Those data are used as training data. To training the neural networks with supervised learning, the appropriate CNN structure is designed and classified images and correct output results are presented to train the weights of neural networks. When the actual image is input, the output of CNN can be used to distinguish whether the loading part of a truck is the covered or not. Experimental results show that images can be classified according to car type and loading type of cargo and it can be used for the pre-classification of loading defect inspection.

Index Terms: Convolutional Neural Network, Classification of Vehicle Type, Type of Trucks

1. INTRODUCTION

With the movement of large-scale logistics, the risk of damage to the road surface or accidents caused by falling objects due to the loading defect of a truck is increasing. So it is important to prevent loading defect and subsequent accidents through crackdown on loading defect. In order to judge whether or not the loading status of a truck is poor, many vehicle images are searched and judged by persons. However, it need to check the images of all vehicles, there are a lot of images to search and it includes passenger cars, refrigerated Truck and special vehicles that are not included in crackdown on the load defects. In order to improve the search speed by reducing the number of search targets, the pre-classification system is required, which gathers the vehicle images that are subject to load defect check.

Typical image classification uses a method to detect the position of objects from taken by the camera, to extract geometric and statistical features that can classify objects such as their outline, size and shape and to compare with the reference value and distinguish them [1-3]. However, since there are various types of vehicles in the images used for classifying the vehicles, the size and shape of the vehicle used as a feature are shown in many ways. It is difficult to determine appropriate criteria for classification and to construct an appropriate algorithm. Also most vehicle Image processing algorithm use the front and side view images to detect vehicle and to classify vehicle type [4-8]. In order to crackdown the loading defect, it is necessary to use images taken from the top of the vehicle so that the status of the loading box can be shown and to classify the vehicle type according to the shape of the load box.

Recently, as deep neural network technology has developed, the methods using the deep neural network technique have achieved good results in areas such as character recognition, image processing, natural language processing, and speech recognition. Among the deep neural network techniques, since Convolutional Neural Network(CNN) can extract the features by itself that are suitable for the object classification through the learning process and can distinguish the results, it is used in various image understanding fields [9,10].

In this paper, we propose an algorithm which can distinguish the vehicle type and the loading condition using the CNN. The top view images of vehicle are used as input, which is taken from a camera installed on the highway access road. This can automatically classify vehicle images subject to loading defect inspection

2. IMAGE CLASSIFICATION USING CONVOLUTIONAL NEURAL NETWORK

CNN has a structure that can extract a feature and classify itself through training process. CNN can be divided into a convolutional network for extracting feature of image and a full-connection network for classifying extracted features. First, the input image is input to the Convolutional network, convolution layer and Pooling Layer are repeated to extract the features of the image. Classification is carried out through the multi-layer perceptron in the succeeding full-connection network.

Convolutional network

Convolution operation is performed on the input image to generate the feature map. Convolution layer calculates the result by multiplying the coefficient of filter h of its size HxH by the pixel value of the input image as shown in (1). For all pixels, the results obtained through convolution operations are called feature maps and represent the characteristics extracted by filters with specific coefficients.

$$u_{ij} = \sum_{p=0}^{H-1} \sum_{q=0}^{H-1} I(i+p, j+q) h(p,q) + bias \qquad (1)$$

In order to extract various features appeared in the image, generally, several filters are used to obtain convolution result, so that feature maps output from the convolution layer are generated by the number of filters. The coefficients of the filter used in the convolution operation are obtained from the learning process. The coefficients of the filter are initially set to a random value and updated to values that can extract a feature through adjustment of coefficients in the learning process.

The value obtained by adding the bias value to the result of the convolution is input to the activation function. Linear

Revised Manuscript Received on March 02, 2019.
DongGyu Lee, Professor, Division of IT convergence, Shinhan University, South Korea (E-mail: dglee@shinhan.ac.kr)

function, Step function, Sigmoid function, Hyperbolic tangent function is used as the activation function. But in the backpropagation algorithm, which is often used to learn artificial neural networks, the amount of error propagated to the lower layer closer to the input becomes smaller, so that the gradient vanishing problem may occur. It is a state in which weight update does not occur because there is almost no change in error slope. The use of the Rectified Linear Unit (ReLU) such as (2) as activation function solves this problem. Since the ReLU function can deliver the error to a lower layer, in deep neural networks, it is often used as an activation function

$$\mathrm{ReLU}(u) = \max(0, u) \quad (2)$$

The result of the convolution layer is used as an input to the pooling layer and the pooling layer is used to select the representative value of the specified area. It is used to reduce the sensitivity of the position of extracted feature value and decrease the size. As the pooling method, are Max Pooling and Average Pooling are used. When u_{pq} is the feature value in (p,q), R_{ij} is the region of size $P \times Q$ centered on (i, j) and the number of pixels in this area is H, (3) represent the

pooling. By selecting the largest or average value of feature values in each region as a representative value, pooling is used to obtain a similar result even if the position of the object is changed in the image, and also to reduce the amount of the parameter.

$$\text{Max Pooling}: \ R_{ij} = \max_{(p,q) \in R_{ij}} u_{pq}$$

$$\text{Avg Pooling}: \ R_{ij} = \frac{1}{H} \sum_{(p,q)\in R_{ij}} u_{pq} \quad (3)$$

In the convolution layer or the pooling layer, an interval for performing each operation must be set. If the operation interval, stride, is set to 1, the operation is performed while moving the position of 1 pixel. If the stride is s, the operation is performed while moving the s pixels. If the stride is large, the amount of computation is reduced and the size of the output feature map is reduced. However, although it is advantageous when dealing with large images, there is a possibility that the detailed features of the image are lost, so an appropriate stride should be set

In a convolution network, several layers such as convolution layers and pooling layers are used and a network consists of multiple layers with various combination of the number of filter, filter size, stride, and layer connection state. Convolution network is configured to extract the appropriate feature map and to decrease the number of parameters that can be classified in subsequent full-connected network.

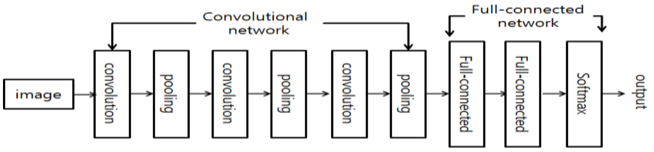
Full-connection network

In the Full-Connected Layer, all nodes in the next layer are connected with current node and are used to classify features extracted from the previous convolution network through multiple layers of Full-Connected network. Generally, it uses the full-connected multi-layered perception and is in the form of passing through the activation function after multiplying the weight by the input. The output of the j-th node is given by (4), when the weight of connecting the i-th input to the j-th node is W_{ij}, the i-th input is x_i, and the activation function is f ().

$$u_j = \sum_{i=0}^{N} W_{ij} \times x_i + b_j \quad (4)$$

$$z_j = f(u_j)$$

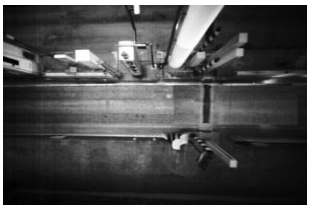
In the neural network for classification, in the final output layer, K nodes are installed as many as the number of categories to be classified, and the SoftMax function is used as the activation function, the probability that input values will be included in each category can be calculated like (5).



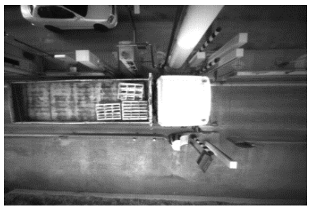
Proposed CNN layer structure and learning data

The CNN structure used in this paper is as shown in Fig. 1. It consists of 4 convolution layers and max-pooling layers and 2 full-connected layers, is composed of 10 layers in total. The last output layer uses the SoftMax function because it is aimed to distinguish the type of the loading part.

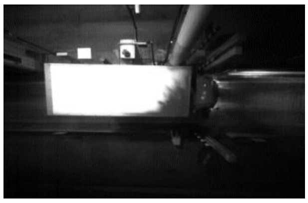
Since the resolution of the input image is high, the stride is 1 or 3, and the area of the pooling layer is also 4x4 or 3x3 so that unnecessary features are filtered out and the size of the output image is reduced. The size of the filter and the computation region are adjusted so that the result of the convolution layer can be reduced to a size that can be input to the full-connected layer



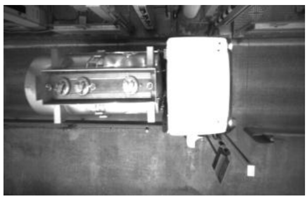
(a)



(b)



(c)



(d)

Fig. 2 Training Images.

(a)Background image (b) truck with uncovered loading part (c) truck with covered loading part (d) specifically-equipped truck

Published By:
Blue Eyes Intelligence Engineering
& Sciences Publication

In order not to need to find the difference between the background and the input image and excluded from feature values such as edges or objects in the background, a background without a vehicle is assigned as one output value. And to reflect the changing background of the day and night, the background images in various conditions are used as learning data. The learning data set for vehicle detection and classification are made using 20 background images without vehicles and 250 vehicle images.

According to the size of the vehicle and the shape of the cargo box, the types of car and loading cargo including the background are classified into 6 categories as shown in Table 1, and the correct vehicle classification result is set as output value together with the image information in the learning data. After the learning step, the image of the vehicle is input to the CNN input layer. The probability of being classified according to the type of the vehicle and the loading box is calculated in the softmax of the output layer. Fig. 2 show examples of the background and vehicle images used as training data in the learning step.

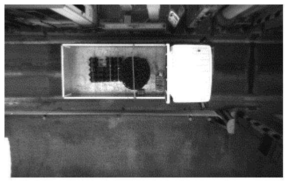
Table 1 Classification category and training data

Image type	The number of training data	Correct output value
background	20	0
car	50	1
small-truck	50	2
large-truck	50	3
small-van, small-specifically-equipped truck	50	4
large-van, large-specifically-equipped truck	50	5

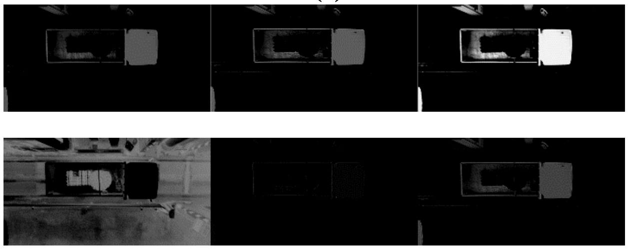
I. EXPERIMENTAL RESULT

Using a camera with a resolution of 1280 * 1024 installed on the highway access road, the top view image of the vehicle is used as the input image so that the cargo area can be seen. When the vehicle appears at a specific region in the image, the image is stored and in this image, we remove the area that is not related to the vehicle. 1180x612 size image is used as CNN input.

In order to generate the training data, the input image was classified into 6 categories including the background, 20 pieces of learning data were selected for each category, and the output value corresponding to the correct classification result of each image was defined as the correct answer. We use the supervised learning process and the error is calculated by comparing the final output with the correct answer. The filter coefficient of the convolution layer and the weight of full-connected layer are updated according to the backpropagation algorithm. The verification data is generated for verification of the learning result, and after the one learning process, the classification result for the verification data is checked to see if it matches the correct answer. This learning process is repeated. The one process of updating the weights for all learning data and verifying the verification data is called an epoch. The learning and verification process is repeated for 100 epochs so that the correct results are obtained.



(a)



(b)

Node 0=4.89354e-13
Node 1=7.92998e-08
Node 2=0.999819
Node 3=0.000180649
Node 4=6.57285e-09
Node 5=5.2241e-11

Fig. 3: Example of Classification.
(a) Input Image b) 6 feature map of Conv1 (c) Output value of softmax

After the learning process, the output values for the real image are calculated using the learned weights and coefficients of filters. Since the soft max function used in the last layer outputs the largest value from the output of the node corresponding to the input image most similar to the learning data, as the actual vehicle image is inputted to the CNN, the number of the node having the largest value in the output layer is determined as the classification result. Fig. 3 shows the actual input image and the feature map at conv1 and the output value of each node in the soft max function of the last layer. In this case, since the output of node 2 is the largest, it is judged to be small-truck.

Table 2 Classification result

Vehicle type of input image	The number of Input Image	The number of correct output	Correct Classification percent(%)
car	26	256	94.8
small-truck	253	235	92.9
large-truck	213	198	93.0
small-van, small-specifically-equipped truck	135	123	91.1

Published By:
Blue Eyes Intelligence Engineering
& Sciences Publication

large-van, large-specifically-equipped truck	175	161	92.0

Table 2 shows the results of classification of actual vehicle images using learned filters and weights from the training data. The input images were stored according to the classified results, and the results were compared with the actual vehicle types. The proposed CNN showed good performance for various vehicle types. Most of the classification errors was due to a misdetection position resulted from the change in illumination at night or at sunset. If the position of the vehicle moves too far or the shape of the cargo of the loading cargo is not clear due to the influence of the lighting, the result is misclassified.

3. CONCLUSIONS

CNN, which is an image processing method using deep learning, performs better than other machine learning methods in image processing. It could be used to classify the status of the loading part from the vehicle image. The features of the image were extracted using several convolutional layers and pooling layers and we used a number of full-connection layers to classify vehicle types and loading types. By using the background image as the learning data in the learning process, it was possible to extract the feature using only the vehicle image without inputting the difference image between the input image and the background image. It is considered that collecting and learning the learning data of various types of vehicles can improve the classification performance. The result of classification of the vehicle type can be used for pre-classification of loading defect inspection and it could be used for automatic crackdown of the loading defect truck that check the poor loading status.

REFERENCES

1. R. Girshick, J. Donahue, T. Darrell & J. Malik, (2014), Rich Feature Hierarchies for Accurate Object Detection and Semantic Segmentation, Proceedings of CVPR 2014, 580-587
2. W. K. Park, Y. Choi, H. K. KIM, G. S. Choi & H. Y. Jung, (2017), Night-time Vehicle Detection Method Using Convolutional Neural Network, The Korean Journal of IEMEK, 12(2), 113-120.
3. H. Huttunen, F. S. Yancheshmeh & K. Chen, (2016), Car type recognition with Deep Neural Networks, Proceedings of the IEEE Intelligent Vehicles Symposium, 1016-1021.
4. Y. O. Adu-Gyamfi & A. Sharma, (2017), Automated Vehicle Recognition with Deep Convolutional Neural Networks, Transportation Research Record Journal of the Transportation Research Board, 2645, 113-122.
5. Peemen, Mesman, Bart, Corporaal & Henk, (2011), Speed sign detection and recognition by convolutional neural networks. Proceedings of the 8th International Automotive Congress, 11(2), 162-170.
6. B. Su, J. Shao, J. Zhou, X. Zhang, L. Mei & C. HuY, (2016), The Precise Vehicle Retrieval in Traffic Surveillance with Deep Convolutional Neural Netwoks, International Journal of Information and Electronics Engine, 6(3), 192-197.
7. Z. Dong, Y. Wu, M. Pei, & Y. Jia, (2015), Vehicle Type Classification Using a Semisupervised Convolutional Neural Network, IEEE Transactions on Intelligent Transportation systems, 16(4), 2247-2256.
8. P. K. Kim & K. T. Lim, (2017), Vehicle Type Classification Using Bagging and Convolutional Neural Network on Multi View Surveillance Image, IEEE Conference on CVPRW, 11(2), 41-46.
9. Mokhtari, E., M. Khamehchian, G. Montazer, and M. Nikudel. "Landfill Site Selection Using Simple Additive Weighting (SAW) Method and Artificial Neural Network Method; A Case Study from Lorestan Province, Iran." International Journal of Geography and Geology 5. 10 (2016): 209-223.
10. Suryanto, T., Haseeb, M., & Hartani, N. H. (2018). The Correlates of Developing Green Supply Chain Management Practices: Firms Level Analysis in Malaysia. Int. J Sup. Chain. Mgt Vol, 7(5), 316.
11. W. S. Shin & C. H. Lee, (2008), Efficient Learning and Classification for Vehicle Type using Moving Cast Shadow Elimination in Vehicle Surveillance Video, The KIPS Transactions, 15B(1), 1-8.

Published By:
Blue Eyes Intelligence Engineering
& Sciences Publication