

FEATURE EXTRACTION AND RECOGNITION ON TRAFFIC SIGN IMAGES

Ilkay CINAR^{1,+}, Yavuz Selim TASPINAR¹, Mucahid Mustafa SARITAS¹, Murat
KOKLU¹

¹Selçuk University, Faculty of Technology, Department of Computer Engineering,
Konya Türkiye

{ilkay.cinar, ytaspinar, mustafa.saritas, mkoklu}@selcuk.edu.tr

Abstract

It is vital that the traffic signs used to ensure the order of the traffic are perceived by the drivers. Traffic signs have international standards that allow the driver to learn about the road and the environment while driving. Traffic sign recognition systems have recently started to be used in vehicles in order to improve traffic safety. Machine learning methods are used in the field of image recognition. Deep learning methods increase the classification success by extracting the hidden and interesting features in the image. Images contain many features and this situation can affect success in classification problems. It can also reveal the need for high-capacity hardware. In order to solve these problems, convolutional neural networks can be used to extract meaningful features from the image. In this study, we created a dataset containing 1500 images of 14 different traffic signs that are frequently used on Turkey highways. The features of the images in this dataset were extracted using convolutional neural networks from deep learning architectures. The 1000 features obtained were classified using the Random Forest method from machine learning algorithms. 93.7% success was achieved as a result of this classification process.

Keywords: Classification, Convolution neural network, Feature extraction, Random forest, Traffic signs

+ This paper has been presented at the ICAT'20 (9th International Conference on Advanced Technologies) held in Istanbul (Turkey), August 10-12, 2020.

TRAFİK İŞARETİ GÖRÜNTÜLERİNDE ÖZELLİK ÇIKARMA VE TANIMA

Özet

Trafiğin düzenini sağlamak amacıyla kullanılan trafik levhalarını sürücülerin algılaması hayati önem taşımaktadır. Sürüş esnasında sürücünün yol ve çevre hakkında bilgi edinebilmesini sağlayan trafik levhaları uluslararası standartlara sahiptir. Trafik levhası tanıma sistemleri son zamanlarda trafik güvenliğini arttırmak amacıyla araçlarda kullanılmaya başlamıştır. Makine öğrenmesi yöntemleri görüntü tanıma alanında kullanılmaktadır. Derin öğrenme yöntemleri, görüntüde yer alan gizli ve ilginç özellikleri çıkararak sınıflandırma başarısını arttırmaktadır. Görüntüler çok sayıda özellik içermektedir ve bu durum sınıflandırma problemlerinde başarıyı etkileyebilmektedir. Ayrıca yüksek kapasiteli donanım gereksinimini de ortaya çıkarabilmektedir. Bu sorunların çözülebilmesi için görüntüden anlamlı özelliklerin çıkarılmasında konvolüsyonel sinir ağları kullanılabilir. Bu çalışmada Türkiye'deki karayollarında sıklıkla kullanılan 14 farklı trafik levhasına ait 1500 görüntü içeren bir veri seti tarafımızca oluşturulmuştur. Bu veri seti kullanılarak derin öğrenme mimarilerinden konvolüsyonel sinir ağları kullanılarak görüntülerin özellikleri çıkarılmıştır. Elde edilen 1000 özellik makine öğrenmesi algoritmalarından Random Forest yöntemi kullanılarak sınıflandırılmıştır. Bu sınıflandırma işlemi sonucunda %93.7 başarı elde edilmiştir.

Anahtar Kelimeler: Konvolüsyonel sinir ağları, Özellik çıkarma, Random forest, Sınıflandırma, Trafik işaretleri

1. Introduction

Traffic signs provide traffic layout on the roads as well as are used to inform drivers and pedestrians about road condition. For this reason, traffic signs are vital for drivers and pedestrians. However, tracking traffic signs while driving can pose safety risks. For this reason, automatic systems are developed to automatically detect traffic signs and warn drivers while driving.

Machine learning methods are used to identify traffic signs. Hidden and different features can be extracted using deep learning methods, which is a method of machine learning to obtain features from images.

In deep learning, feature extraction is performed using many layers. Each layer uses the output of the previous layer as input [1]. Convolutional neural networks (CNN) are one of the deep learning architectures that include artificial neural networks and layers of feature extraction. CNN is also a type of Multi-Layer Perceptron (MLP). As one of the most well-known algorithms of deep learning architecture, CNN is able to classify a model directly from video, images, text, or sound [2].

Deep learning has been used in many studies in recent years, including visual recognition, speech recognition, and natural language processing. In the literature, some studies with traffic signs are; Huang et al., proposed a method for recognizing traffic signs. In this method, which consists of two modules, the Histogram Oriented Gradient variant (HOGv) property is presented to represent the distinguishing features. In addition, the connection between the input and hidden layers performs random feature mapping, while only the weights between hidden and output layers are trained using the extreme learning machine. The results of the proposed method indicate high recognition accuracy as well as extremely high computational efficiency [3]. Mannan et al. in their study, they proposed a purely data-driven segmentation technique that adaptively selects an optimized color field based on available training data to be able to make a complete distinction between pixels corresponding to traffic signs and background objects. They were able to distinguish traffic signs from other objects with 0.81 precision in experimental results [4]. Ghica et al. they have proposed an artificial neural network system for recognizing traffic signs. First, the input image is processed for the extraction of color and geometric information. A morphological filter is applied to eliminate smaller objects and noise. The coordinates of the resulting objects are determined and the objects are isolated from the original image according to these coordinates. The objects are then normalized and sent to the neural network that performs the recognition process. They tested the system by simulation on a large amount of data obtained by a video camera attached to the vehicle. They stated that they achieved excellent results except where the lighting was poor [5]. Han et al. they have proposed a system of recognizing traffic signs for autonomous vehicles. There are two basic structures in the system, detection and classification of traffic signs. Detection of signs includes Color detection, morphological filter and labeling elements. The k-nearest neighbors (kNN) algorithm was used for classification. Speeded up Robust features (SURF) algorithm is used in the training

features are obtained. They achieved 97.54% accuracy when looking at the results [6]. Kiran et al., carried out work on the detection and recognition of traffic signs using color information. They used color-based segmentation techniques to detect traffic signs. They also used the linear support vector machine for classification [7].

In other parts of the study, convolutional neural networks used in trait inference and the random forest algorithm used in classification were discussed. The following sections provide information about dataset and the results obtained.

2. Convolutional Neural Network (CNN)

Convolutional neural networks are a type of multilayer perception. It derives its name from the convolution operator. It is frequently used in areas such as image processing, audio processing, natural language processing and biomedical. It provides the most successful results in the areas of image processing/recognition and classification. Convolution main purpose is to extract the features of the input image. It uses small squares of convolution input data. In this way, it preserves the spatial relationship between pixels by learning the features of images. The layers that make up the structure of convolution neural networks are given as subheadings [8, 9].

1.1. Convolution layer

Convolution layer on the basis of deep neural networks are located. In this layer, small size filters such as 2×2 , 3×3 and 5×5 are hovered over the entire image. In this way, more specific attributes are extracted from the image, resulting in a new image [10, 11]. In Figure 2.1 shows the process of the convolution filter on the image.

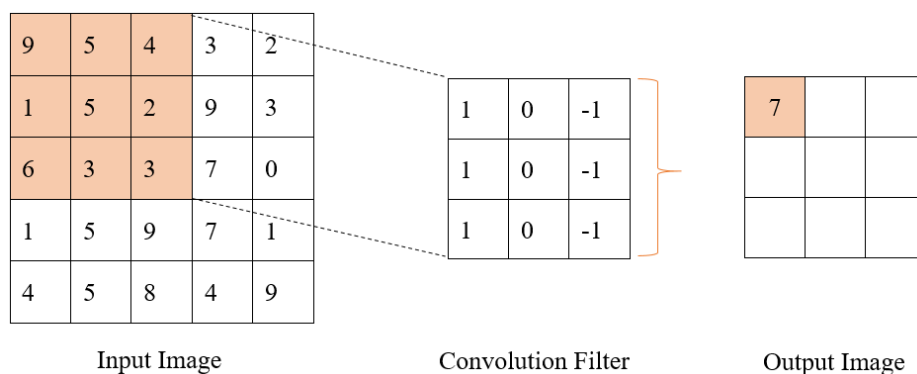


Figure 2.1. Convolution process

In the convolution process given in Figure 1, the output image obtained after the process of the 3×3 convolution filter on a single channel 5×5 size image matrix is given. Shifting the 3×3 convolution filter right-to-left and down-to-up on the input image is applied on the entire image. The filter coefficients are calculated by multiplying and then aggregating with equal-sized windows inside the input image. These operations result in a new output image based on apparent high-level attributes.

1.2. Pooling layer

The pooling layer is used for size reduction. Loss of some information in the size reduction process and decrease in performance is seen as the disadvantage of this layer [10]. However, preventing the memorization of the model and reducing the computational burden are also seen as its advantages.

The operations performed in this layer are performed using certain types of filters, such as those performed in the convolution layer. By moving these filters over the image, pooling process is done by taking the maximum, average or minimum values of the pixels in the image [11]. Figure 2.2 provides an example of how to apply maximum, minimum and average pooling operations in window size 2×2 on the image in size 4×4 .

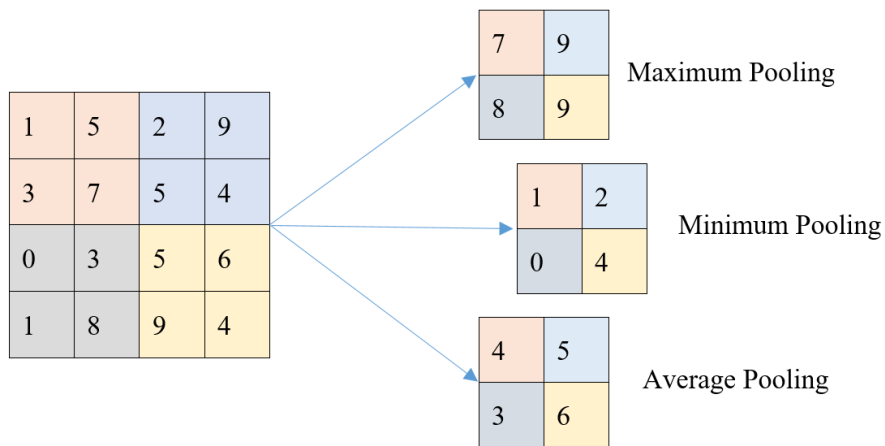


Figure 2.2. Pooling process

1.3. ReLU layer

Relu is one of the most widely used activation functions in models developed based on deep neural networks. Activation functions for convolutional neural networks are an important factor. One of the most important features of the ReLU layer is that the

negative values in the input data are drawn to zero. This allows the network to learn faster by using the Relu function. The Relu function is given in Equation 2.1 [12].

$$f(x) = \begin{cases} 0 & \text{if } x < 0 \\ x & \text{if } x \geq 0 \end{cases}$$

(2.1)

1.4. Normalization layer

The normalization layer is used to regularize data from layers developed based on convolutional neural networks. This process ensures that input data is within a certain range. It also positively affects the performance of the network [13].

1.5. Fully connected layer

This layer is connected to all the neurons of the layers that preceded it. It also consists of a one-dimensional matrix. Fully connected layers are often used towards the end of the CNN architecture and in optimizing class scores.

1.6. Classification layer

In the CNN model, it is the last layer where the classification process is done. The classification layer output values are equal to the number of classes, depending on the number of objects to be recognized. In the classification layer, softmax classifier is used for multiple classification problems (3 classes and above). This classifier produces probabilistic values between 0-1 for each class, and consequently the highest probability value gives the class predicted by the model [10, 13].

3. Random Forest (RF)

Random forest (RF) is a classifier consisting of many Decision Trees (DT). To make a new classification, each DT provides a classification for the entries. Then selects the estimate with the most votes by evaluating the RF classifications [14].

RF has the ability to manage a large number of variables in a dataset. It is also highly successful in predicting missing data. However, the final model and the results after it are difficult to interpret. This is also due to the fact that it contains many independent decision trees. Figure 3.1 shows the general structure of RF [15].

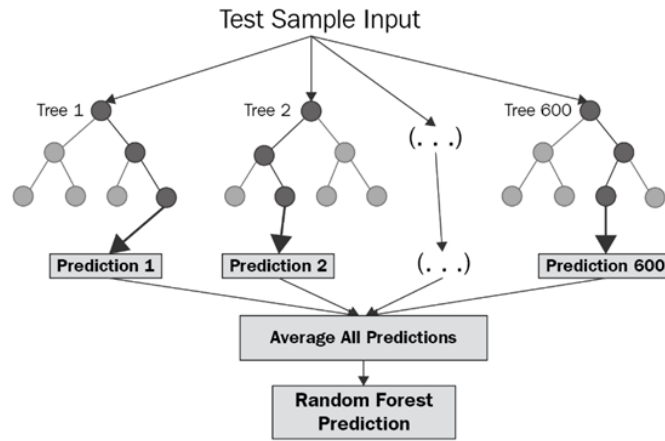


Figure 3.1. RF general structure

4. Dataset

The dataset used in the study consists of 1500 images obtained from 14 different traffic sign used in Turkish highways. The sample images on the dataset are shown in Figure 4.1 and the information on the traffic signs is given below.

- | | |
|------------------------|---------------------|
| A. Danger | H. School Way |
| B. Landslide Zone | İ. Speed Limit (30) |
| C. No Entry | J. Speed Limit (50) |
| D. No Overtaking | K. Stop |
| E. No U Turn | L. Traffic Lights |
| F. Pedestrian Crossing | M. Uneven Road |
| G. Roundabout | N. Yield |



Figure 4.1. Sample images in dataset

5. Experimental Results

With the obtained images, feature extraction and classification process were performed using Orange data mining application. In feature extraction, 1,000 features were extracted for each image from 1500 images using convolutional neural networks from deep learning architectures. The structure of CNN is given in Figure 5.1.

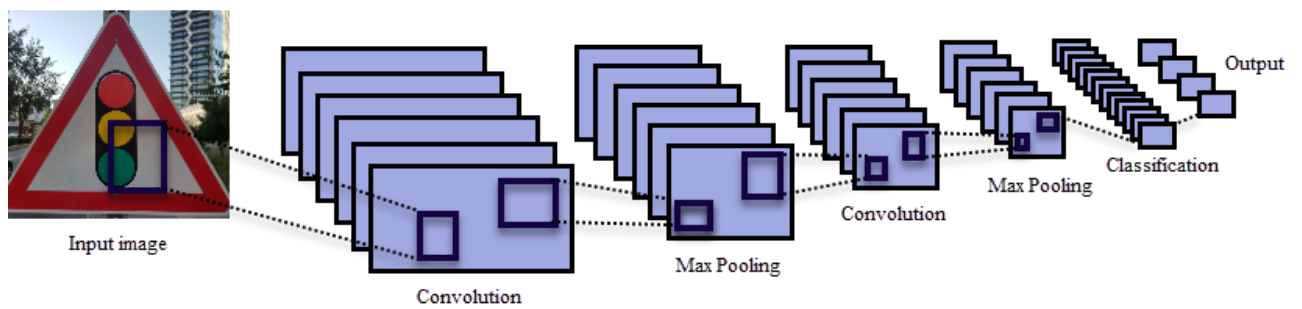


Figure 5.1. Structure of CNN

The resulting features were classified using Random Forest (RF) from machine learning algorithms. The cross validation value for the classification was selected as 10. The confusion matrix of the RF algorithm is given in Table 5.1 and the results of performance criteria Accuracy, F1 score, Precision and Recall are given in Table 5.2.

Table 5.1. Confusion Matrix

	A	B	C	D	E	F	G	H	I	J	K	L	M	N
A	39	0	0	0	0	0	0	0	0	0	0	0	2	0
B	0	40	0	0	0	2	0	1	0	0	0	0	1	0
C	0	0	127	1	0	0	0	0	0	0	0	0	0	0
D	0	0	0	43	0	0	3	0	0	2	0	1	0	0
E	0	0	0	1	43	0	0	0	1	1	0	0	0	0
F	0	1	0	0	0	198	2	8	0	0	0	0	0	1
G	0	1	0	0	0	1	165	2	0	0	0	2	3	1
H	0	0	0	0	0	11	1	84	0	0	0	0	0	1
I	0	0	0	0	0	0	0	0	175	15	1	0	0	0
J	0	0	0	0	1	0	1	0	7	123	0	0	0	1
K	0	0	2	0	0	1	1	0	0	0	92	0	1	0
L	0	1	0	0	0	0	0	1	0	0	0	76	1	0
M	2	1	0	0	0	1	2	0	0	0	1	0	117	1
N	0	0	0	0	0	1	0	0	0	0	0	0	0	84

Table 5.2. Performance Criteria

Model	Accuracy	F1 score	Precision	Recall
Random Forest	0.937	0.937	0.938	0.937

When the tables are examined, it appears that traffic signs have been successfully classified. 93.7% accuracy was achieved as a result of the classification. F1 score, precision and recall values are very close values because the correct and incorrect classification numbers of positive and negative data are close together.

6. Conclusions

Convolutional neural networks from deep neural networks architectures are often used in feature extraction from images. In this study, 93.7% classification success was achieved with features obtained by convolutional neural networks.

By increasing both the variety and number of images of the traffic signs on the dataset, higher classification success can be achieved. In addition, studies can be carried out using other machine learning algorithms that are often used in classification processes.

Features can be extracted with different CNN architectures, as extracting different features from images can increase classification success.

AUTHOR'S NOTE

Abstract version of this paper was presented at 9th International Conference on Advanced Technologies (ICAT'20), 10-12 August 2020, Istanbul, Turkey with the title of “Feature Extraction and Recognition on Traffic Sign Images”.

References

- [1] Smirnov, E.A., D.M. Timoshenko, and S.N. Andrianov, *Comparison of regularization methods for imagenet classification with deep convolutional neural networks*. Aasri Procedia, 2014. **6**: p. 89-94.
- [2] Aytaç, Z.İ., İ. İşeri, and B. Dandil, *Trafik Hız Sınırlama Levhalarının Nesne Tanıma Modeli ile Sınıflandırılması*.
- [3] Huang, Z., et al., *An efficient method for traffic sign recognition based on extreme learning machine*. IEEE transactions on cybernetics, 2016. **47**(4): p. 920-933.
- [4] Mannan, A., et al., *Optimized segmentation and multiscale emphasized feature extraction for traffic sign detection and recognition*. Journal of Intelligent Fuzzy Systems 2019. **36**(1): p. 173-188.
- [5] Ghica, D., S.W. Lu, and X. Yuan. *Recognition of traffic signs by artificial neural network*. in *Proceedings of ICNN'95-International Conference on Neural Networks*. 1995. IEEE.
- [6] Han, Y., K. Virupakshappa, and E. Oruklu. *Robust traffic sign recognition with feature extraction and k-NN classification methods*. in *2015 IEEE International Conference on Electro/Information Technology (EIT)*. 2015. IEEE.
- [7] Kiran, C., L.V. Prabhu, and K. Rajeev. *Traffic sign detection and pattern recognition using support vector machine*. in *2009 Seventh International Conference on Advances in Pattern Recognition*. 2009. IEEE.
- [8] Tashiev, İ., et al., *Konvolüsyonel Sinir Ağı Kullanarak Gerçek Zamanlı Araç Tipi Sınıflandırması Real-Time Vehicle Type Classification Using Convolutional Neural Network*.
- [9] Tüfekçi, M. and F. Karpat, *Derin Öğrenme Mimarilerinden Konvolüsyonel Sinir Ağları (CNN) Üzerinde Görüntü İşleme-Sınıflandırma Kabiliyetininin Arttırılmasına Yönelik Yapılan Çalışmaların İncelenmesi*.

- [10] Kizrak, M.A. and B. BOLAT, *Derin öğrenme ile kalabalık analizi üzerine detaylı bir araştırma*. Bilişim Teknolojileri Dergisi, 2018. **11**(3): p. 263-286.
- [11] Özkan, İ. and E. Ülker, *Derin Öğrenme ve Görüntü Analizinde Kullanılan Derin Öğrenme Modelleri*. Gaziosmanpaşa Bilimsel Araştırma Dergisi, 2017. **6**(3): p. 85-104.
- [12] Türkoğlu, M., et al., *Derin Evrimsel Sinir Ağı Kullanılarak Kayısı Hastalıklarının Sınıflandırılması*. Bitlis Eren Üniversitesi Fen Bilimleri Dergisi. **9**(1): p. 334-345.
- [13] Doğan, F. and İ. Türkoğlu, *Derin öğrenme algoritmalarının yaprak sınıflandırma başarımlarının karşılaştırılması*. Sakarya University Journal of Computer Information Sciences, 2018. **1**(1): p. 10-21.
- [14] Mao, W. and F. Wang, *New advances in intelligence and security informatics*. 2012: Academic Press.
- [15] Chakure, A. *Random Forest Regression*. 2019 [Access Date: 10 July 2020]; Available from: <https://towardsdatascience.com/random-forest-and-its-implementation-71824ced454f>.