Real-Time Vehicle Classification Using CNN

1st Nusrat Jahan
Sr. Lecturer, CSE
Daffodil International University
Dhaka, Bangladesh
nusratjahan.cse@diu.edu.bd

2nd Saiful Islam
Sr. Lecturer, CSE
Daffodil International University
Dhaka, Bangladesh
saiful.cse@diu.edu.bd

3rd Md. Ferdouse Ahmed Foysal
CSE, Daffodil International University
Dhaka, Bangladesh
ferdouse15-5274@diu.edu.bd

Abstract— **Convolutional Neural Network (CNN) is a model of artificial neural networks that has grown to be most well known in computer vision assignment. In this paper work, we presented convolutional neural network for classifying four types of common vehicle in our country. Vehicle classification plays a vital role of various application such as surveillance security system, traffic control system. We addressed these issues and fixed an aim to find a solution to reduce road accident due to traffic related cases. The greatest challenge of computer vision is to achieve effective results to implement a system due to variation in shapes and colors of data. To classify the vehicle we used two methods feature extraction and classification. These two methods can straightforwardly performed by convolutional neural network. The method shows quite good performance on real-time standard dataset. Our mentioned method able to reach 97% accuracy in case of vehicle classification.**

Keywords— Vehicle, Machine learning, Image processing, CNN, Computer Vision

I. INTRODUCTION

Human Life is moving at a lightning speed. To keep up with this lightning lifestyle we are using faster and faster vehicles to transport everything including us from one place to another. But at the same time human life meets an unwanted end in a blink of an eye due to this radical speed of vehicles. More specifically due to our mind set to compete in the road and perform multiple unnecessary lane switching. According to the Transport ministry of Bangladesh compared to 2018 the vehicle crashes have increased by more than 51% and the death rate has increased by more than 17%. In 2019 a total of 5227 souls were lost in road accidents in Bangladesh which is almost a thousand more dead than 2018 [1]. Over the last ten years more than 25000 plus people were dead and more than 20000 people were more than seriously injured in road accidents [2]. The effects of these road accidents are more furious than themselves, economically and most importantly the mental pressure on the family. Irresponsible driving, over speed and lane switching are some of the common reasons. Most problematic scenario is not being able to detect or identify the vehicle behind the road accidents. Mayank Singh Chauhan and et al., presented a CNN based object detection model to identify the different classes of vehicles such as 2 wheelers, 3 wheelers and 4 wheelers with a 75% model performance accuracy [3].

Object classification is colossal field of inquire about in computer vision as well as machine learning that points to classify objects show in pictures into significant categories [4, 5]. Within the setting of brilliantly activity frameworks, object classification of expect a crucial part and have a wide run of work such as traffic surveillance, course optimization, and peculiarity location. Automatic analysis of activities is a significant and crucial issue due to a large number of traffic rule violations and their adverse effects on regular traffic management system [6]. Detection is used on roads, highways, parking or any other place to detect or track the number of vehicle present on the spot. It helps the surveillance to judge the traffic, average speed and category of vehicle. Matched with the conventional machine learning assignments, profound learning based strategies that have made incredible developments in activity reconnaissance methods, and have accomplished great execution in commonsense applications like object detection, feature extraction along with track identification. In these ranges of research areas, accurate and real time object detection and introductory classification are the most principal and essential work to get clear idea on this field. In machine learning models image [7] preparing and pattern acknowledgment include extraction starts from a fundamental set of measured information and develops advanced values aiming to be instructive and non-redundant, encouraging the taking after learning and generalization moves [8], and in a few cases driving to superior human translations.

Now, this is the time of technological revolution. In today's digital age we are using technology everywhere. Our main objective is to classify our proposed vehicle. There are different types of vehicle. We want to classify different type of vehicles by using deep convolutional network. Our goal is to study how to classify or recognize different type of vehicles. To improve a platform that will be capable to identify all kind of vehicle. To visualize some analytical analysis of vehicle like Car, Cng, Rickshaw, Cycle classified by classifier algorithms.

We were interested to do a system to reduce our traffic congestion as well as road accident for this. So, we decided earlier that research on Artificial Intelligence (AI) and Machine Learning (ML) field can help us and then we started to search for doing this research work. Then we want to improve our traffic system for that we decided to classify different types of vehicle. We have reached on an idea and worked on vehicle classification using Convolutional Neural Network. Our work is completely related with machine learning strategies as machine learning is necessary to make a system with knowledge.

Now, this is the time of technological revolution. In this digital age we are using technology everywhere. In this paper work our leading objective is to classify vehicle using our proposed vehicle. There are different types of vehicle. We want to classify different type of vehicles by using deep convolutional network. So, we can describe our objectives in a list-

- To classify or recognize vehicle we considered common four types of vehicles- Car, CNG, Rickshaw, Cycle.
- To construct a platform to detect different lanes and vehicles to automate the whole process of traffic control system.

The main challenges of this work is to collect and process the dataset, dealing with the data set was too much hard in the field of machine learning research to achieve quiet good results. To clean and normalize the collected data (images) we used several steps and methods. Training the network consists of many layers with different size of epoch will take long time to train the machine, if the data is not processed or cleaned properly. We did find previous work based on his data set so we started this paper work from data collection. Fig. 1 to demonstrate our planned system outline.

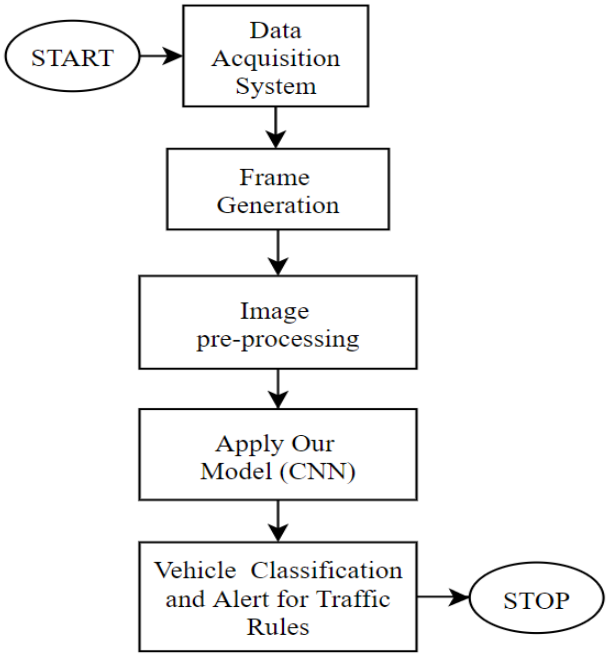

Fig. 1. Proposed System Outline

II. LITERATURE REVIEW

In this section, we will discuss about related works, research summary and challenges about this research. In related works section, we will discuss other research paper and their works, their methods, and accuracy which are related to our work.

In 2013, R.Sindoori and et al. prescribed a methodology to find out vehicle disclosure structure from groveling pictures. Vehicle disclosure is finished by pixel clever grouping procedure. The important point of this research was to remove the incorporate and arrange the vehicle shading system. Extraction of incorporate incorporates edge and corner area. To edge revelation, the Canny edge locator strategy was applied. The Harris corner discoverer handle was associated with corner identification. Adaboost model used for the extraction of vehicle shading to isolate vehicle and non-vehicle hues. Totally, morphological activities are associated with the improvement of the vehicle revelation purpose [9]. In 2013, Youporn Hu and et al., displayed an approach where they combined the Haar features extraction approach with Histograms of Oriented Gradients (HOG) as strong features for the area of the vehicle independently. They projected the system to distinguish vehicles in accounts and order them into 2 sorts relying upon consolidated the Haar features along with HOG. Because of this system, it can arrange and distinguish the vehicles in multidirection approach with extraordinary grouping [10].

In 2015 Chiman Kwan, Bryan Chou claimed a system for tracking Vehicle Using Deep Learning. They show a couple of preliminary vehicle following comes about using compressive estimations from the principal infrared video. In addition, the compressive estimations are implying video traces with discretionarily lost pixels. CNN's have outlined predominant execution in picture classification [11], and have been, all things considered, associated for question area, video classification [12], and division. These impels have too added to an unused focus in asking about on tall execution CNN frameworks. The plans of these frameworks have too observed their execution pushed ahead by using increasingly significant and progressively broad structures [13]. Simonson et al. proposed VGGNet [14], which empowered the task about by using significant structure. Szegedy and et al., displayed GoogleNet that contains the inception modules, defining a cutting edge condition of the craftsmanship for ImageNet Challenge.

Yann Lecun proposed the CNN model. Roused from human brain structures he proposed this model. Truly CNN performs like the regular visual cortex. CNN is one of the chief fruitful models to work with deep learning research field. CNN's image grouping precision is also predominant than some other regular image order estimations. There are assorted sorts of layers that actually focused on CNN model. Convolution layer incorporates a moving channel or bit which can go according to the image. All around it can go through a 2D arrange (portrayal of the picture) and take a specific package and applies bit duplication and accumulates in another lattice.

In 2018, Alpatov and et al., presented a system to detect and count real time vehicle to ensure road safety. They used images from a stationary camera to processing and implement the algorithm to avoid any incident on road [15]. Mallikarjun Anandhalli and Vishwanath P. Baligar proposed an algorithm to detect and track vehicle based on color. They used Kalman filter to track the vehicle [16]. Nirmal Purohit and et al., discussed a system to classify vesicle. They considered category of tanks and trucks to Identify and classify an enemy vehicle to improve military defense system. Here, they used HOG and SVM [17].

Seda Kul in 2017 [18] proposed an analysis paper on vehicle detection and its classification. They mentioned many algorithms for detecting and classifying vehicle such as- Logistic regression, Neural networks, Support vector machines, and Convolutional neural network. Watcharin Maungmai in 2019 proposed a system to classify vehicle based on two characteristics- Vehicle types and color. Here they applied CNN as a deep learning algorithm and got 81.62% and 70.09% for that two case. In this study, they considered 763 seconds video data to produce images [19]. In 2018 Bensedik Hicham and et al., discussed a system to classify vehicle using CNN. Total 2400 images they used to produce their dataset and got average 89% accuracy for 4 types of vehicle. Their aim was to control traffic light based on priority or in emergency basis [20].

After analysis those previous work we observed that for country a smart traffic management system is necessary to reduce a large number of road accidents in every day. So, in our proposed system we want to construct a model that can be able to detect a vehicle it is in wrong lane. To reduce road accident it will be help us in initial level.

III. CONVULITIONAL NEURAL NETWORK

Convolutional Neural Network (CNN) is a decisively proven most popular deep learning algorithms. CNN takes an image, after that process it and classify the image into trained categories. CNN basically a network where we can able to find Kernals, Pooling, after that fully connected layers (FC) and a activation function (Softmax or sigmoid) for classifying an object or any image with probabilistic values that must be among 0 and 1.

The first step of convolutional layers is "kernals". It helps to extract different types of features like edge detection, blur and sharpen.

However, this filtering sometimes does not work to fit image as an input image. So, we can use padding. After that, non-linearity (Relu) function is needed. "Relu" stands for Rectified Linear Unit to perform a non-linear function in CNN network. Here, Equation (1) to perform the non-linear function.

$$f(A) = \max(0, A) \tag{1}$$

Max Pooling is another important layers that helps convolutional network to reduce the number of parameters if we have too large images. . Max pooling is disposed of 75% of the enactments and controlling overfitting. Fig. 2 to depict the max pooling approach.

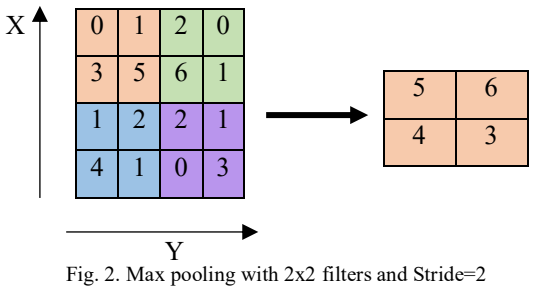

Fig. 2. Max pooling with 2x2 filters and Stride=2

Finally, we flatten our final matrix to fit it into the fully connected layers to implement the whole network system to classify the input image using activation function (Softmax). In Fig. 3 we presented the structure of fully connected neural network.

$$\sigma(x_j) = \frac{e^{x_j}}{\sum_{j=1}^{n} e^{x_i}} \tag{2}$$

Where i =1, 2, 3 ……. m and j =1, 2, 3 …….m

This CNN model takes input dataset as an algorithm and produces a score of each class. Then the selected "Softmax" activation function converts the probability score between 0 and 1.

Another part of CNN is dense layer. It is actually a general layer of neurons in a CNN model. Each neuron can take input from other neurons in the past layer, and it is densely connected with each other to pass the information from one layer to another. After this dense layer dropout approach is needed in CNN to increase the model performance accuracy. This is a technique where randomly selected neurons are disconnected during training face. So, it helps model to work with high accuracy but with low overfitting.

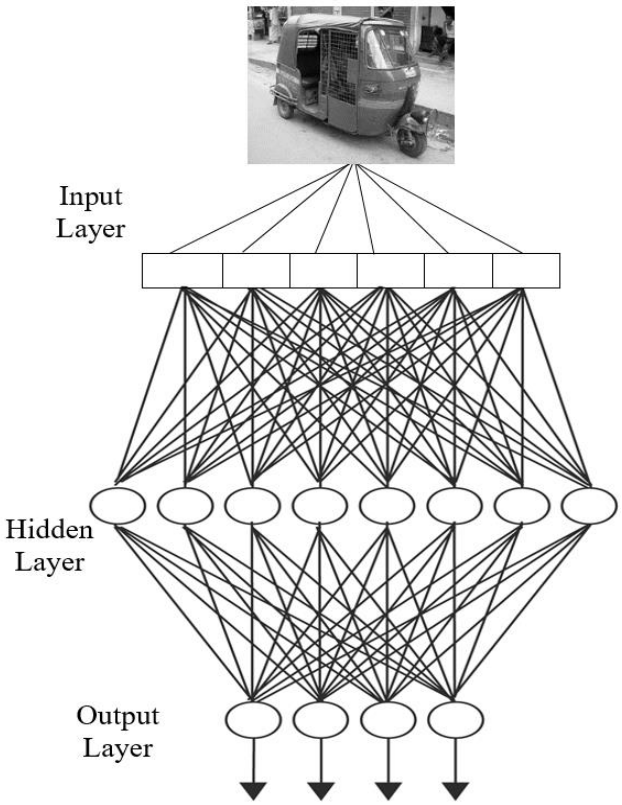

Fig. 3. Convolutional Neural Network

IV. PROPOSED MODEL

This section will intensively discuss the data collection procedure, preprocessing along with proposed model. It has been observed that to achieve better results we need to increase data quality and quantity. So, data is an important part for any machine learning algorithms.

A. Data Collection

An appropriate dataset is required in every steps of machine learning techniques. A total of 2800 images was considered here. 2240 images for training and 560 images for test purpose. As a raw data we collected 100 images of each case. Collecting data from on-site observation as well as images from video was considered here. We collected total 400 images for our selected 4 categories of vehicle- Car, CNG, Rickshaw, and Bicycle. Here, we considered 80% data as a train dataset and remaining 20% data as test purpose. Fig. 3 to present the sample of our dataset.

Fig.4. Sample Dataset

B. Data Preprocessing

Data is the heart of machine learning. Our collected data (images) with different height and width. In next level, we resized all of our collected images into 100 x 100 pixels. Moreover, we have converted every image into gray image. Since, we do not have a way to provide better graphics processing unit in our machine that we have utilized to prepare the show. So, we utilized grayscale images to prepare the model.

We misleadingly extended the dataset to maintain a strategic distance from overfitting. It includes esteem to base information by including data determined from inner and outside sources inside a venture. It makes a difference to extend the amount of significant information within the dataset.

Here, we used 6 following techniques for augmentation purpose to extend the total 400 sample image data:

- Rotation (400)
- Rescale
- Shear
- Vertical Flip
- Zoom
- Horizontal Flip

So, we have total 2800 images. To implement proposed model we considered 2240 as training dataset and 560 as test dataset.

C. Proposed CNN Model

We are now going to discuss about our methodology that was followed for getting expected outcome. For training purpose we considered 10 layered CNN which is showed in Fig. 5. Our proposed CNN model delivered approximately 97% test accuracy.

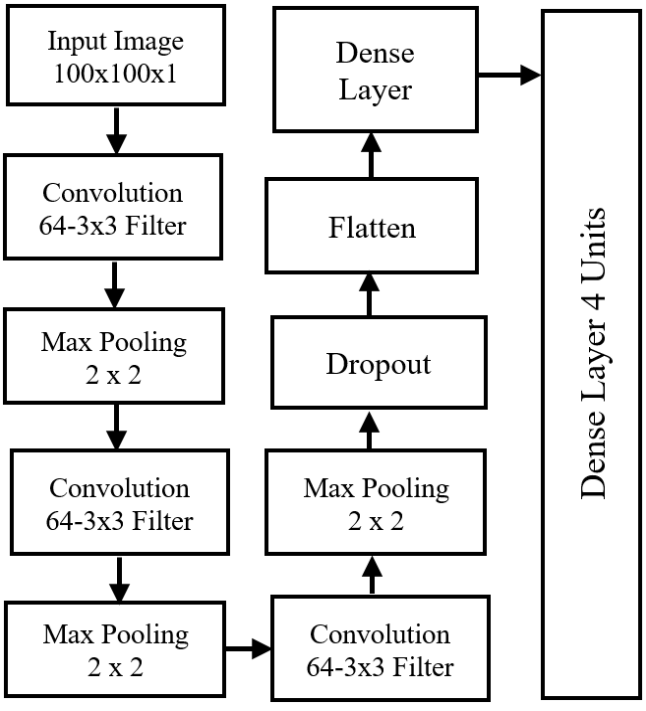

Fig. 5. Proposed CNN Model

Here, we used 3 convolution layer to create the network. In every step relu activation was placed as a non-liner function.

- Network Layer I: 64-3x3 filters with 'Relu' function
- Network Layer II: 64-3x3 filters with 'Relu' function
- Network Layer III: 64-3x3 filters with 'Relu' function

V. Performance of the Model

Here, we are presenting our model performances. We have already discussed about our dataset. To evaluate our model we used some graphical representation.

For getting accuracy from a CNN network where many layers placed one by one we need an optimizer to produce results. Optimizer actually helps the networks to create output layer. To analysis a task with machine learning algorithm training and test results is needed to monitor that model. Optimizer actually calculate the loss and make a solution to reduce it. Though, CNN has several kinds of optimizer procedure, among them we picked "Adam Optimizer" to, reduce the loss.

'Adam' optimizer used here to compile the whole model to get results after creating the whole network. We used batch size 64 and 20 epochs to train the model and observe the results. Fig. 6 to present the test and validation accuracy along with test and validation loss for our model.

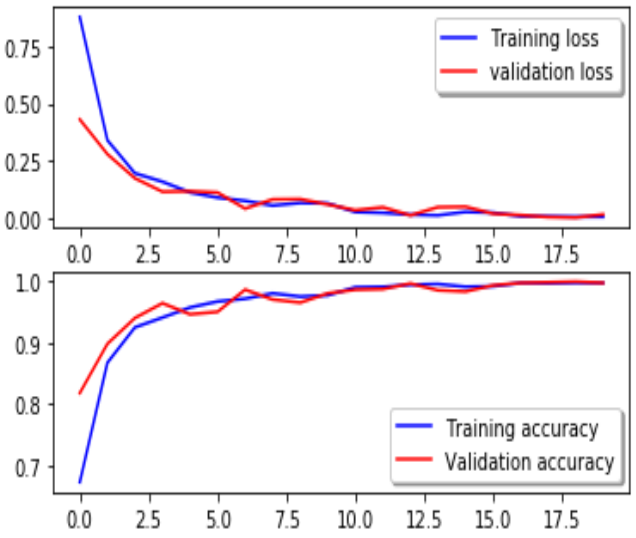

Fig. 6. Accuracy vs. Loss

Precision: Within the field of data recovery, precision is the division of recovered reports that are relevant to the inquiry:

$$\text{Precision}=Tp/(Tp+Fp) \quad (3)$$

Precision is utilized with the review, the percent of all pertinent archives that's returned by the look. Note that the meaning and usage of "precision" within the field of data recovery contrasts from the definition of exactness and exactness inside other branches of science and innovation. Precision helps to find our close relation between data.

Recall: Recall is the piece of pertinent occasions that have been recovered over the full amount of relevant occurrences. Actually it helps to find relevant data correctly.

The tall recall implies that a calculation returned most of the significant results.

$$Recall=Tp/(Tp+Fn) \quad (4)$$

Accuracy: accuracy refers to the familiarity of the measured value to a known value. Accuracy is necessary to understand the accomplishment quality of a model.

$$Accuracy=(Tp+Tn)/(Tp+Tn+Tp+Fn) \quad (5)$$

Here Table I., to provide the performance results of our model. Here, we calculated prediction, recall and f1-score to measure our model.

PERFORMANCE MEASUREMENT

Class	Precision	Recall	F1-Score
Class 0	0.98	0.94	0.96
Class 1	0.95	0.98	0.95
Class 2	0.96	0.98	0.98
Class 3	0.99	0.96	0.97
Avg Accuracy	0.97	0.96	0.96

Now, we displayed a table (Table II) to present some recent works which was related with our approach. However, many previous works we got which is related to us but not exactly our model. Our dataset is totally a unique in this research area. That is why we prepared this table to enhance the understanding level of our study.

TABLE I. RELATED PREVIOUS WORK

Paper Summary	Algorithm
To detect different types of vehicle used data from delhi roads. Total 5562 video frames was used to perform CNN based vehicle classification with counting [3]	CNN Based Object Detection Model
Proposed an approach to automatic Vehicle Classification using Roadside LiDAR Data [21]	Here, they presented different supervised learning algorithm to classify the vehicle and got 91.98% accuracy.
To work on Small Sample Fine-Grained Vehicle Classification they used 196 datasets and perform 2 experiments [22]	Fine-grained Vehicle Classification with a novel proposed loss function named "Dual Cross-Entropy Loss"
Their aim was to minimize the loss function of CNN is the case of Fine-Grained Vehicle Classification [23]	Claimed a Channel Max Pooling (CMP) modified layer for CNNs.
For vehicle counting and classification was the mail aim in this paper. An automatic lane detection approach was also proposed. [24]	Optimal Classifier with kalman filter, shadow elimination and feature extraction
For Vehicle Tracking and Classification purpose here occlusion, noise, and other perplexing issues were addressed to build a framework [25].	Sparse approximation followed by Bayesian state inference framework

VI. CONCLUSION

It has no doubt that there is a lot of research works on this field. We proposed a model to implement a project to reduce traffic along with the accident which is leading by break the traffic rules. Vehicle recognition and classification are some of the important domains on traffic systems. In a variety of applications, it has become important. We had to collect many data from different places. Then we preprocessed our dataset for training and testing purpose. Data collection and preprocessing were the most challenging parts to us. We used an algorithm which known as convolutional neural network (CNN). Finally, we found our expected results and our accuracy is better than rest of the projects that related with this work.

However, to improve our study we need to collect more data and need to consider other vehicles those are also available on road. To provide a complete system we will have to observe some real events on road. Nowadays there are many kinds of technology used to control traffic, so this approach will invent a new technology which is our main goal to find out and resolve traffic and security based difficulties.

REFERENCES

[1] "Bangladesh: Alarming rise in road crashes". [Online Available: https://www.aa.com.tr/en/asia-pacific/bangladesh-alarming-rise-in-road-crashes/1692643] Acesses Date: 25.05.2020.

[2] "25,526 people killed in road accident in the last 10 years". [Online Available:https://www.dhakatribune.com/bangladesh/nation/2019/06/12/25-526-people-killed-in-road-accident-in-the-last-10-years] Acesses Date: 28.05.2020.

[3] Mayank Singh Chauhan, Arshdeep Singh, Mansi Khemka, Arneish Prateek, Rijurekha Sen, "Embedded CNN based vehicle classification and counting in non-laned road traffic." 10th International Conference on Information and Communication Technologies and Development, pp.1-11, 2019.

[4] Y. Artan, B. Alkan, B. Balci, and A. Elihoş, "Deep Learning Based Vehicle Make, Model and Color Recognition Using License Plate Recognition Camera Images." 27th Signal Processing and Communications Applications Conference, Turkey, pp. 1-4, 2019.

[5] C. M. Bautista, C. A. Dy, M. I. Mañalac, R. A. Orbe and M. Cordel, "Convolutional neural network for vehicle detection in low resolution traffic videos." IEEE Region 10 Symposium (TENSYMP), Bali, pp. 277-281, 2016.

[6] Li Suhao, and et al., "Vehicle type detection based on deep learning in traffic scene." 8th International Congress of Information and Communication Technology (ICICT), vol. 131, pp. 564-572, 2018.

[7] Zhang, F., Li, C., and Yang, F., "Vehicle Detection in Urban Traffic Surveillance Images Based on Convolutional Neural Networks with Feature Concatenation." *Sensors*, no. 3, pp. 594, 2019.

[8] Ratanak samrith You, Jang-Woo Kwon, "VoNet: Vehicle Orientation Classification using Convolutional Neural Network." ICCIP 16, pp. 195-199, 2016.

[9] R.Sindoori, K.S .Ravichandran, and B. Santhi, "Adaboost Technique for Vehicle Detection in Aerial Surveillance." International Journal of Engineering and Technology, vol. 5, no. 2, pp. 765-769, 2013.

[10] Y. Hu et al., "Algorithm for vision-based vehicle detection and classification." IEEE International Conference on Robotics and Biomimetics (ROBIO), Shenzhen, pp. 568-572, 2013.

[11] C. Kwan, B. Chou, A. Echavarren, B. Budavari, J. Li and T. Tran, "Compressive Vehicle Tracking Using Deep Learning." 2018 9th IEEE Annual Ubiquitous Computing, Electronics & Mobile Communication Conference (UEMCON), New York City, NY, USA, pp. 51-56, 2018.

[12] R. Girshick and et al., "Rich feature hierarchies for accurate object detection and semantic segmentation." in Proceedings of the IEEE conference on computer vision and pattern recognition, pp. 580–587, 2014.

[13] A. Karpathy and et al., “Large-scale video classification with convolutional neural networks.” in Proceedings of the IEEE conference on Computer Vision and Pattern Recognition, pp. 1725–1732, 2014.

[14] J. Long, E. Shelhamer, and T. Darrell, “Fully convolutional networks for semantic segmentation.” in Proceedings of the IEEE conference on computer vision and pattern recognition, pp. 3431–3440, 2015.

[15] B. A. Alpatov and et al., “Vehicle detection and counting system for real-time traffic surveillance.” 2018 7th Mediterranean Conference on Embedded Computing (MECO), Budva, pp. 1-4, 2018.

[16] M. Anandhalli and V. P. Baligar, “Vehicle Detection and Tracking Based on Color Feature.” 2017 International Conference on Recent Advances in Electronics and Communication Technology (ICRAECT), Bangalore, pp. 240-248, 2017.

[17] Purohit, U Nirmal et al. “Vehicle classification and surveillance using machine learning technique.” 2017 2nd IEEE International Conference on Recent Trends in Electronics, Information & Communication Technology (RTEICT), pp. 910-914, 2017.

[18] S. Kul, S. Eken and A. Sayar, “A concise review on vehicle detection and classification,” 2017 International Conference on Engineering and Technology (ICET), Antalya, pp. 1-4, 2017.

[19] Maungmai, Watcharin and Chaiwat Nuthong, “Vehicle Classification with Deep Learning.” 2019 IEEE 4th International Conference on Computer and Communication Systems (ICCCS), pp. 294-298, 2019.

[20] Hicham, B., Ahmed, A., and Mohammed, M., “Vehicle Type Classification Using Convolutional Neural Network.” 2018 IEEE 5th International Congress on Information Science and Technology (CiSt), 313-316, 2018.

[21] Jianqing Wu and et al., “Automatic Vehicle Classification using Roadside LiDAR Data.” Transportation Research Record: Journal of the Transportation Research Board, vol. 2673, no. 6, pp. 153-164, 2019.

[22] Xiaoxu Li, Liyun Yu, Dongliang Chang, Zhanyu Ma , Jie Cao, “Dual Cross-Entropy Loss for Small-Sample Fine-Grained Vehicle Classification.” IEEE Transactions on Vehicular Technology DOI: 10.1109/TVT.2019.2895651, vol. 68, no. 5, pp. 4204-4212, 2019.

[23] Z. Ma et al., “Fine-Grained Vehicle Classification With Channel Max Pooling Modified CNNs.” in IEEE Transactions on Vehicular Technology, vol. 68, no. 4, pp. 3224-3233, April 2019.

[24] Jun-Wei Hsieh, Shih-Hao Yu, Yung-Sheng Chen and Wen-Fong Hu, “Automatic traffic surveillance system for vehicle tracking and classification.” in IEEE Transactions on Intelligent Transportation Systems, vol. 7, no. 2, pp. 175-187, June 2006.

[25] X. Mei and H. Ling, “Robust Visual Tracking and Vehicle Classification via Sparse Representation.” in IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 33, no. 11, pp. 2259-2272, Nov. 2011.