INDIAN JOURNAL OF SCIENCE AND TECHNOLOGY

RESEARCH ARTICLE

Check for updates

OPEN ACCESS

Received: 07-09-2021

Accepted: 02-06-2023

Published: 19-06-2023

Citation: Harilakshmi VS, Rani PAJ (2023) Deep Learning based Artificial Intelligent Systems in Road Traffic Density Estimation and Congestion Classification. Indian Journal of Science and Technology 16(24): 1768-1776. https://doi.org/10.17485/IJST/v16i24.1660

* **Corresponding author**.

harilakshmi.vs@gmail.com

Funding: None

Competing Interests: None

Copyright: © 2023 Harilakshmi & Rani. This is an open access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Published By Indian Society for Education and Environment (iSee)

ISSN
Print: 0974-6846
Electronic: 0974-5645

Deep Learning based Artificial Intelligent Systems in Road Traffic Density Estimation and Congestion Classification

V S Harilakshmi[1]*, P Arockia Jansi Rani[2]

1 Research Scholar, Department of Computer Science and Engineering, Manonmaniam Sundaranar University, 627012, India
2 Associate Professor, Department of Computer Science and Engineering, Manonmaniam Sundaranar University, 627012, India

Abstract

Objectives: The main objective of this paper is to employ the subset of artificial intelligence, namely, deep learning to estimate road traffic density and thus mitigate the undesirable effects caused by traffic congestion and improve the quality of life of people. **Methods:** This work presents a method of classification of road traffic conditions based on video surveillance data obtained from CCTV cameras mounted on highways. A simple, basic architecture of deep convolutional neural network (DCNN) based method is introduced that learns traffic density from pre-labeled images in order to estimate the traffic flow density in highways. **Findings:** The standard publicly available UCSD dataset of real videos is used for experimental verification. The experimental results obtained shows that the proposed model outperformed all the existing conventional methods in the literature by reaching the highest accuracy and classifies the test video in less computational time. **Novelty:** The proposed methodology employs Matlab deep learning network designer with hyper parameter tuning, cross validation and activation maps to classify the road traffic density into three different states namely light, medium and heavy.

Keywords: Deep Learning; Artificial Intelligent Systems; Density Estimation; Intelligent Transportation Management; Deep CNN

1 Introduction

High traffic congestion creates a deep negative impact in the urban environment. Time and fuel waste, accidents and loss of human life's, additional cost on the economy, environmental air pollution resulting in impaired quality of life are amongst the few ravaging effects of congestion. Hence to alleviate traffic congestion problem accurate estimation of traffic density is important. But in real road scenario, there exists many challenges in traffic density estimation. Various climatic changes, fog, mist or CCTVs uninterrupted service relying on power supply all pose challenge in accurate and improved estimation of traffic density.

Traditional density estimation techniques involve magnetic loops, sensors that incurs heavy maintenance cost and limited range of operations respectively. With CCTV cameras installed in road networks, the use of cctv camera feeds for traffic monitoring has enabled gathering of useful information in real time including traffic speed, lane occupancy, traffic density, etc., from very large areas. This provides more flexible solutions compared to the traditional devices like, magnetic loop radars, microwave, infrared detectors which are limited on a single point, high installation cost, and are difficult to install and maintain [1,2].

This paper proposes a robust traffic-density classification model based on a deep neural network architecture, which overcomes all the issues and challenges faced by existing techniques. A DCNN is used in the estimation of the road traffic density and classify the traffic flow into three congestion classes, light, medium and heavy in different challenging environmental conditions, like fog, mist and rain and in overlapping of vehicles during heavy congestion.

The organization of the remainder of this paper is, Section II briefs the related works in the literature. Section III provides an in-depth description of our Deep CNN architecture. Section IV discusses the experimental results by presenting an evaluation of our proposed method's performance and a comparison with different approaches used on the same publicly available UCSD dataset. Concluding remarks and further suggestions as extension of the proposed methodology are given in section V.

1.1 Related Works

This section briefly explains the traffic density estimation methods, right from the traditional loop detectors installed on roads, then the vision system with CCTV feeds to the very recent deep neural network technology that is being used for density estimation. The pros and cons of these methods, the research gap between these are also discussed in the below section.

Earlier estimation of traffic density was done by collection devices that are installed beneath the roads. This incurs huge installation costs and also enormous processing time and human ability [3]. To this drawback came the surveillance cameras as a solution. The CCTV cameras initially installed for surveillance, later became a boon for solving many traffic related problems With CCTV cameras getting popular and deployed in traffic surveillance and monitoring, computer vision-based systems for density estimation became popular. The main advantage with these systems is that, these can be laid without damaging the road infrastructures. Computer vision-based systems with their unique capabilities solve many problems in intelligently monitoring and managing the road traffic and its smooth functioning [4]. Even with many advantages, vision systems suffer from poor performance with low resolution camera feeds and high occlusion videos in real time. The direct background subtraction module in vision system takes less computational time, but it suffers from non-identification of dark vehicles, thus posing vehicle color to be a problem in finding the traffic density [5]. Most vision - based system for density estimation starts with a background subtraction module. These modules suffer miserably under high occluded and cluttered traffic scenes.

Ever since the concept of deep learning, a subset of machine learning has been introduced it has been widely used and has brought in revolutions in image processing and computer vision processing technologies [5–9]. Artificial Intelligence, AI's contribution in intelligent transportation management has opened new arenas in this field. Many researchers in recent years explore and apply deep learning framework in intelligent transportation management system, and proved its effectiveness in traffic incident detection and traffic flow prediction [10]. In [11] the use of artificial neural networks (ANN) proves to be a good choice in the tasks of collecting, interpreting, and analyzing huge data coming from video cameras. CCTV camera feeds that give low-resolution surveillance data that uses deep neural networks to count vehicles on the road and estimates traffic density are proposed in [8,12]. Another interesting application of deep learning on road traffic networks is an analysis on traffic accidents have been proposed by number of researchers [13]. Research that focuses on hot spot detection of traffic accidents is proposed in [14].

All these research works, works well on clear, smooth roadways which is not the actual road scenario [15–17]. Also, non-lane, cluttered, chaotic roads of developing countries is still a challenge in smart transportation systems, that are yet to be addressed, living it as a research gap even today [18–20]. Still there are research possibilities that exist in deep learning based intelligent transportation management system [21] as there is no single framework that meets all the requirements in this field. Thus, in this paper research on dynamic estimation of traffic density based on the deep learning algorithms automatic feature learning capability is explored in congestion classification for smooth mobility of commuters.

1.2 Proposed Approach

In this section, we present a background about deep learning and deepCNN architecture and describe the proposed deep CNN architecture in more detail and the techniques performed to learn the optimal architecture in order to increase the road traffic density classification accuracy.

1.3 Deep Learning

Deep learning is a promising technology in Artificial Intelligence. It is a subset of machine learning with automatic feature extraction. It automatically learns features and tasks directly from the data. More data when used for feature extraction and training yields better model.

Deep learning basically uses multi-layer neural network architecture. It consists of input, output and n number of hidden layers. The below Figure 1 explains how a deep learning architecture works. The input layer accepts the inputs. The hidden layer consists of activation, bias elements. The output layer gives the output. The hidden layer is named so, as it is not accessible outside the architecture. Deep learning can be applied on different data types like image, video, numeric, signal-1 D, speech, audio and text data types. Based on the data the right deep neural network has to be chosen. For image and video frame data CNN or DCNN is the best suitable architecture, like LSTM for time series or text series.

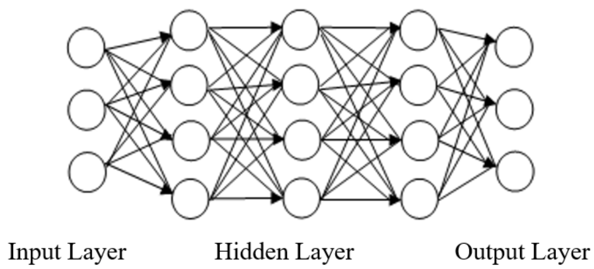

Fig 1. Deep Network Architecture

1.3.1 Overview of layers in DCNN Architecture

DCNNs are a special kind of neural network. In deep learning, a convolutional neural network (CNN) is defined as a class of deep neural networks. Based on the application and on the complexity of the data the network can be formed by a certain number of convolution and sub sampling layers that are stacked on top of each other layers.

1.3.1.1 Input Layer. Its first layer is the input layer through which the image input is fed into the network. Next is the convolutional layer. The input data is send to the convolution2D layer.

1.3.1.2 Convolutional Layer. The convolutional layer is the core building block in a DeepCNN architecture. It consists of certain number of convolutional filters. The filters are applied on the input image that extracts the local features namely, edges, shapes, etc. The output of each convolutional layer is an activation map. There can be number of CNN layers depending on the number of features we require.

1.3.1.3 Pooling Layer. Next is the pooling layer. After each convolutional layer in the architecture there is a pooling layer. The pooling layers are a form of down-sampling that reduces the resolution of the activation maps. Sum pooling, mean pooling and max pooling are the three main non-linear functions in pooling layer. The conceptual difference between these layers lies in the sort of invariance which they have the ability to catch.

For example, the max-pooling which is the most used, takes the maximum input from a region of the convolutional layer. The sum and mean pooling are set precisely the same as max-pooling but instead using a max function, the sum or the mean function is used. This technique is used to reduce the number of parameters within the model for two reasons. First, is to simplify the computational load and then second, is to reduce the chances of over-fitting.

1.3.1.4 Fully Connected Layer. After including several convolutional and max pooling layers, the high-level features that are found in the images are processed via fully connected layers with the aim of activating only one output per class in the recognition task. In fact, it creates a stochastic likelihood representation of each class based on the activation maps generated by the concatenation of the previous layers and then is the Relu activation layer. The fully connected layer consists of the number of classes that are to be classified.

1.3.1.5 Softmax Layer. Finally the softmax layer gives the loss functions both for target class and for other classes. It gives the probability for the classes. The one with the highest probability is the target class. In our case it is the road traffic density information.

The features that the CNN works upon during the training process can be visualized using the visualization methods namely layer activation, class activation and through deep dream image visualizations.

The Figure 2 below is an overview of a deep CNN architecture which can be deployed to classify the road traffic into low, medium and heavy congestion. The architecture of a typical DCNN is composed of one or more pairs of convolution and pooling layers and finally followed by fully connected layers as seen in Figure 2.

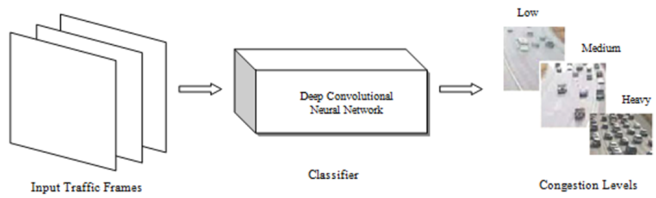

Fig 2. Overview of a general DCNN architecture

2 Significance of Research Work

The main contribution of this research paper is that it develops a simple and an easily testable deep neural architecture. In the following sections, we will describe in detail each module together with the data collected for training and assessment.

2.1 Proposed Model Description

There are two approaches for designing a deep learning network. One approach is to train a deep neural network from scratch and the other approach is to fine tune a pre-trained model, i.e., transfer learning based deep neural network. The biggest advantage of having a DCNN network is that it decreases the human effort in developing a pre-processing functionality module. This paper deals with designing a deep neural network from scratch and then a comparative analysis is done with a pre trained deep networks.

The following are the steps that are carried out in designing a deep cnn learning model.

Step 1: Create a data store and load into the network.

Step 2: Split and load the labeled images into training, validation and test data.

Step 3: Define cnn hyperparameters.

Step 4: Define and construct the cnn architecture and layers.

Step 5: Set and tune the hyperparameters.

Step 6: Train and save the network.

Step 7: Finally evaluate the model by classifying and plotting the confusion matrix.

The proposed deep cnn model architecture is composed of 7 layers. The first layer is the input layer of size (163 174 3), which is the cropped image leaving out the unwanted portions of the frame for ease of computation. Each block contains a convolution module, followed by a relulayer and then max pooling module and a fully connected layer module. The softmax layer and classification layer are the final output layers. Transportation domain specific hyperparameters are chosen and tuned for optimum performance of the proposed system before the training phase begins.

This framework is developed using MATLAB 2021a. The CPU based computation is chosen for training the deep cnn with a dataset of 1272 images, with 424 images in each three classes. The two main tool boxes used are the deep learning tool box and the neural network tool box. These toolboxes enables us to explore the different convolutional training algorithms and other pre-defined layers. This paper works on the development of a deepcnn from scratch. With the deep network designer app the predefined convolutional network algorithms such as AlexNet, GoogleNet and many such can be evaluated for parallel computing and thus to significantly decrease the computation time.

2.1.1 Visualizing Feature Detectors

With its automatic feature learning ability, deep learning algorithms are very complicated to interpret, and hence it is treated as black boxes. However, DCNN algorithms are actually different, in a way that we can visualize various components in it.

This property enables us to give an in depth look into their internal workings and help us understand them much better. These visualizations for each layer help readers understand how DCNN learns features in its intermediate layers. The input frame is passed through the DCNN and a record of the intermediate activations is made as shown in Figure 3. These visualizations serve as a supporting information to help us assess hypotheses about the cause of certain types of errors and understand the relation between the different classes namely, low, medium and heavy.

2.1.2 Visualizing Filters and Feature Maps

Deep learning networks allows us to visualize the filters that a deep CNNs learn during their training for Computer Vision tasks and also allows visualization of feature maps that it produces when applied to an input frame. Visualization of the learned weights gives an idea of how well a network has learned. A lot of zeros means there are many dead filters that are of less or no significance for the network.

The below Figure 7 gives the visualization of deep CNN filters. In deep networks, the term filters refers to the learned weights of the convolutions. For eg., a single 3*3 convolution is referred to as filter and it has a total of 10 weights (9+1 bias).

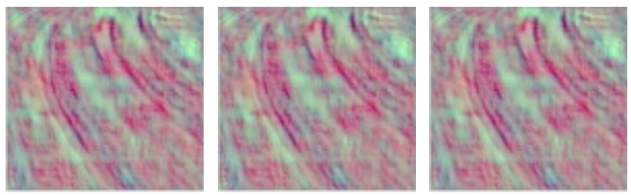

Fig 3. Visualization after convolution layer 1

The feature maps of a deep CNN contain the results obtained after applying the filters to an input frame. The visualization of a feature map enables us get an understanding of what features are detected by the DCNN. In a deep learning network, the early layers detect low-level features like edges, colour, etc., and the later layers detects the higher level features, like shapes and objects.

2.1.3 Hyperparameter Tuning

Hyperparameter tuning is an important step in optimizing the performance of deep CNN model. These parameters control the model's behavior and characteristics for optimum performance. These parameters are not automatically learnt from the input image data. And hence the hyperparameters are to be set initially before the actual network training begins.

The main contribution of the proposed novel system is the choice of transportation domain specific hyperparameters and optimum tuning after several trail and tests. The proposed model is a novel trail and tested transportation domain specific system that works well on different challenging scenarios and different weather conditions. Temporal and spatial resolution, depth and width of network layers, pooling strategies to handle irregular traffic patterns, scheduling of optimum learning rate are the hyperparameters that are tuned for optimal performance of proposed system.

3 Results and Discussion

In this section, we present the results obtained by our approach. Firstly, the data set is described and then the classification results are shown.

3.1 Traffic Video Dataset

The UCSD traffic video dataset consists of 254 video sequences of highway traffic in Seattle, Washington, collected from a single stationary stand still traffic camera over two days[13]. The database contains a number of different traffic patterns and weather conditions, e.g. raining, overcast and sunny. Each video has a resolution of 320*240 pixels and has 42 to 52 frames at 10 frames per second. The database was labeled by hand concerning the level of traffic congestion in each sequence. The database presents a total of 165 sequences of light traffic, 45 of medium traffic and 44 of heavy traffic (very slow), in different environmental conditions, as shown in Figure 4

The below Figure 5 shows a sample of different video frames used for training and testing. The test dataset consists of few real time video frames along with the UCSD video frames used for training. This framework is developed and implemented in Matlab2020a.

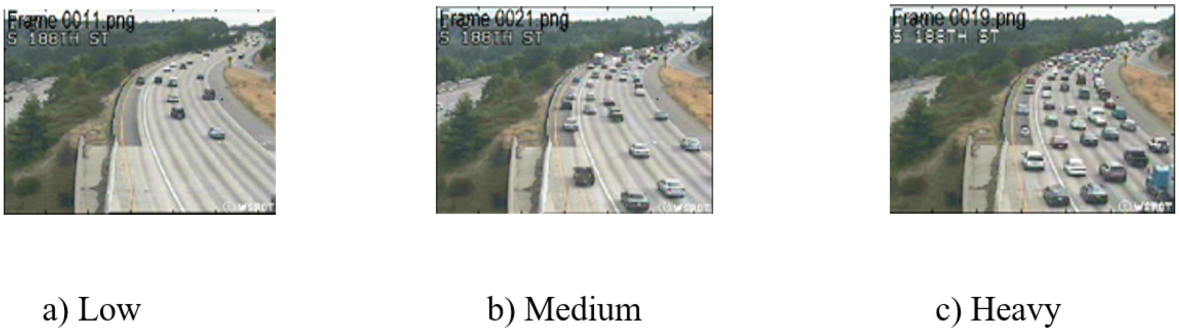

Fig 4. Different Traffic Load Congestion Levels

A total of 1272 images is used for the training of the deep cnn model, with three classes namely heavy, medium and light congestion, with 424 images in each class folder. The UCSD dataset contains images with varying illumination conditions and with varying weather conditions like fog, mist and rain.

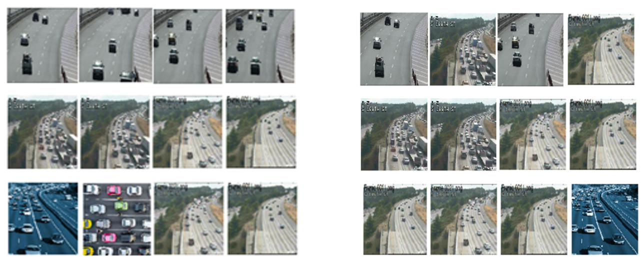

Fig 5. Training and Testing Dataset

Fig 6. Experimental Results (a) Input Test Image (b) Input Test Image (c) Input Test Image

Predicted Class Output=1 Predicted Class Output=2 Predicted Class Output=3, a) Input Test Image b) Input Test Image c) Input Test Image

The above Figure 5 shows the sample database storage consisting of training and testing images and Figure 6 gives the correctly predicted class output of the deepcnn network when a query test image is given to a trained dcnn network.

3.2 Performance Analysis

3.2.1 Recognition Rate

Recognition rate is obtained based on the below formula.

$$Recognition\ rate = \frac{Number\ of\ correctly\ identified\ images}{Total\ number\ of\ images} * 100$$

Table 1. Recognition Rate of the proposed method

Congestion Levels	Recognition Rate(%)
Low	100
Medium	100
Heavy	100

3.2.2 Confusion Matrix

The confusion matrix allows the visualization of the performance of an algorithm. Below Figure 7 gives the confusion matrix with 70 percent of images for training, 15 percent for validation and rest 15 percent for testing.

The confusion matrix of N * N matrix is used to evaluate the performance of the classification model. Here N is the number of target classes. This matrix compares the actual target values with those predicted by the deep cnn learning model. This gives us an understanding of how well our classification model works. Below figure gives the confusion matrix obtained with 7 layers of deep CNN.

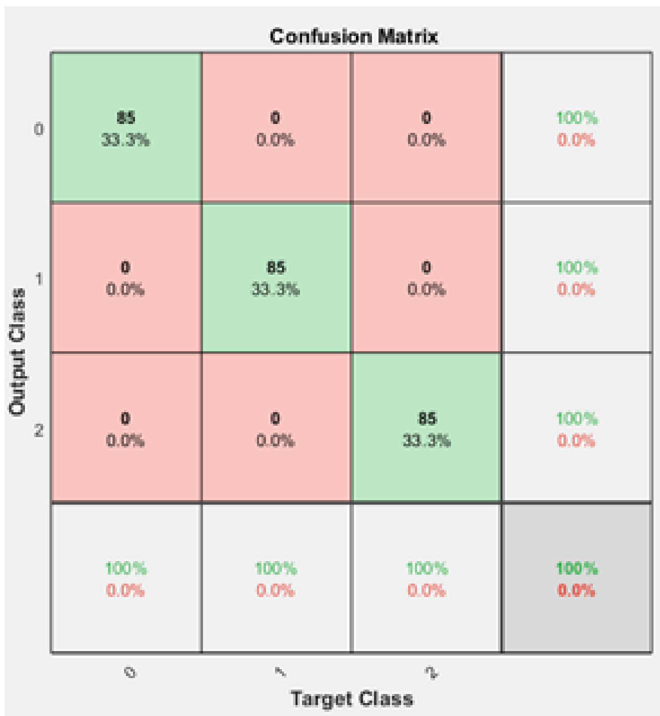

Fig 7. Confusion Matrix

3.2.3 Performance Comparison of the proposed work with other Existing Methods

To prove the effectiveness of the proposed method it has to be compared with recent methodologies that are evaluated on a common dataset. Hence, a comparative analysis of the proposed DCNN with other existing methods on the publicly available UCSD Trafficdb dataset is made and is listed in below Table 2. The below table presents a table of comparison with the other existing deep learning methodologies, J Kurniawan, et al [9], Sabbani Imad, et al [22], D Impedovo, et al [23] and Nguyen L.A, et al.(2021) [24] in the literature towards density estimation and congestion classification techniques.

From the Table 2 above, we say that the proposed deep method performs well compared to other three state of the art methods [9,22–24] and produces exemplary output results.

The pretrained ResNet is used in [23] and yields a better performance in congestion classification task. The ResNet with its residual feature learning capabilities definitely yields effective learning and helps in capturing the more intricate patterns in the input data. But this benefit is more pronounced typically when a large – scale dataset is available for training. The choice between CNN, Resnet or any pretrained network depends on many factors like size of the dataset, available computational

Table 2. Performance Analysis of the Proposed DCNN with existing methods

Authors	Technology / Methodology	Target	Dataset	Accuracy (%)
J Kurniawan, et al.(2018) [9]	CNN with gradient descent with Adam	2 classes (jammed , not jammed) traffic congestion classification,	CCTV images, collected privately	89.50
Sabbani Imad, et al.(2018) [22]	DCNN	3 classes (heavy, medium, light) congestion classification	Publicly available UCSD dataset	99.62
D Impedovo, et al.(2019) [23]	Pretrained ResNet	3 states congestion classification	Publicly available UCSD dataset	98.61
Nguyen L.A, et al.(2021) [24]	CNN with segmented vehicles from input frames	3 states Congestion classification	Publicly available UCSD dataset	98.66
Proposed Methodology	Deep CNN with stochastic gradient descent momentum	3 states congestion classification	Publicly available UCSD dataset	100

resources and the computational time and performance. Here, in the proposed system, a simpler dcnn architecture provides more interpretability and insight into the learned representation and features and performs better than Resnet for the road traffic density estimation and congestion classification task.

The work [24] has a preprocess model where the similar frames are cut an discarded and then passed onto a segmentation model where the non vehicle objects are segmented and discarded. The output of the segmentation module is then fed into cnn model. The proposed model is superior to [24] in a way where the raw road vehicle images are fed into deep cnn and produces a remarkable test result.

Key issue of accurate congestion classification has been addressed in the proposed system with elimination of complex preprocessing and the use of background subtraction for segmenting the foreground vehicle as in many traditional feature-based approaches. The proposed methodology though has attained highest accuracy, still it has to be evaluated with more real time traffic scenarios to prove its efficacy.

4 Conclusion

Traffic density estimation plays an important role in transportation management to mitigate congestion. Deep learning, a subset of artificial intelligence, is one of the fastest growing technologies with far-reaching applications in many areas. This paper explores the deep learning-based technique in intelligently managing the road vehicular traffic. This paper proposes a deep learning based frame work to automatically extract features to classify the highway road congestion levels. The proposed simple and basic deep convolutional neural network architecture correctly determine the congestion levels for the given test images. The deep CNN used in this proposed work has the main advantage that with its implementation of convolutional layer and pooling layer employed in automatic feature extraction of spatiotemporal features of the road transportation network reduces the need for manual interpretations of the features extraction. Previous state of the art methodologies with the challenges of video data from cameras that are installed on low posts or at the roadside and the images that contain a significant number of occluding vehicles does not seem to impair the classification task and proves to provide accurate results.

The proposed methodology is validated using standard UCSD video dataset from traffic observation sites of Seattle with different challenges in illumination, lightning and weather conditions. The highest accurate experimental results in classification proves the system can be integrated into an ITS system responsible for traffic monitoring and management. It is also concluded that the proposed method works well on large-scale datasets. In further research extension, recent deep pretrained models such as ResNets can be used for automatic feature extraction. In future with the deep learning architecture's ability to visualize feature maps in the hidden layers makes the fusion of multi branch network architecture i.e., ensemble model architecture feasible.

The devised simple setup of the DCNN proves to be adequate for the defined classification task of traffic density. In an era with Twitter and other social media getting popular and GPS enabled mobile phones, a fusion of the proposed model with these technologies can be deployed into the cloud environment for real time prediction of the congestion levels and in turn help the policy makers in urban centers to develop an effective artificial intelligence-based transportation management system. These congestions estimation and prediction details can serve as an input in vehicle induced air pollution monitoring systems and thus paves way for sustainable pollution free environment.

References

1) Cucchiara R, Piccardi M, Mello P. Image analysis and rule-based reasoning for a traffic monitoring system. *IEEE Transactions on Intelligent Transportation Systems*. 2000;1(2):119–130. Available from: https://doi.org/10.1109/6979.880969.
2) Kastrinaki V, Zervakis M, Kalaitzakis K. A survey of video processing techniques for traffic applications. *Image and Vision Computing*. 2003;21(4):359–381. Available from: https://doi.org/10.1016/S0262-8856(03)00004-0.
3) Vipin JV, Sharma A, Ashlesh, Lakshminarayanan SL. Road traffic congestion in the developing world. *Proceedings of the 2nd ACM Symposium on Computing for Development*. 2012;11:11. Available from: https://doi.org/10.1145/2160601.2160616.
4) Pongpaibool P, Tangamchit P, Noodwong K. Evaluation of road traffic congestion using fuzzy techniques. *TENCON 2007 - 2007 IEEE Region 10 Conference*. 2007;p. 1–4. Available from: https://doi.org/10.1109/TENCON.2007.4429119.
5) Uddin MS. Real-time Area Based Traffic Density Estimation by Image Processing for Traffic Signal Control System. *Bangladesh Perspective International Conference on Electrical Engineering and Information Communication Technology (ICEEICT)*. 2015. Available from: https://doi.org/10.1109/ICEEICT.2015.7307377.
6) Dean, Jeffrey, Corrado, Greg, Monga, Rajat, et al. Large scale distributed deep networks. *Advances in Neural Information Processing Systems*. 2012;p. 1223–1231. Available from: https://proceedings.neurips.cc/paper/2012/file/6aca97005c68f1206823815f66102863-Paper.pdf.
7) Wan J, Yuan Y, Wang Q. Traffic congestion analysis: A new Perspective. In: 2017 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP). IEEE. 2017;p. 1398–1402. Available from: https://crabwq.github.io/pdf/2017%20Traffic%20Congestion%20Analysis%20A%20New%20Perspective.pdf.
8) Cchenge L, Ggregory D, Xxin F, Yyao W. TrackNet: simultaneous object detection and tracking and its application in traffic video analysis. 2019. Available from: https://doi.org/10.48550/arXiv.1902.01466.
9) Kurniawan J, Syahra SGS, Dewa CK, Afiahayati. Traffic Congestion Detection: Learning from CCTV Monitoring Images using Convolutional Neural Network. *Procedia Computer Science*. 2018;144:291–297. Available from: https://doi.org/10.1016/j.procs.2018.10.530.
10) Hinton GE, Osindero S, Teh YW. A Fast Learning Algorithm for Deep Belief Nets. *Neural Computation*. 2006;18(7):1527–1554. Available from: https://doi.org/10.1162/neco.2006.18.7.1527.
11) Hinton GE, Salakhutdinov RR. Reducing the Dimensionality of Data with Neural Networks. *Science*. 2006;313(5786):504–507. Available from: https://doi.org/10.1126/science.1127647.
12) Huang W, Song G, Hong H, Xie K. Deep Architecture for Traffic Flow Prediction: Deep Belief Networks With Multitask Learning. *IEEE Transactions on Intelligent Transportation Systems*. 2014;15(5):2191–2201. Available from: https://doi.org/10.1109/TITS.2014.2311123.
13) Ffukai Z, Cce L, Ffeng Y. Vehicle detection in urban traffic surveillance images based on convolutional neural networks with feature concatenation. *Sensors*. 2019;19(3):594. Available from: https://doi.org/10.3390/s19030594.
14) Zhang S, Shanghang, Wu G, Guanhang, Costeira J, Joãop, et al. FCN-rLSTM: Deep spatio-temporal neural networks for vehicle counting in city cameras. *Proceedings of the IEEE international conference on computer vision*. 2017. Available from: https://doi.org/10.1109/iccv.2017.3966.
15) Xie Z, Yan J. Detecting traffic accident clusters with network kernel density estimation and local spatial statistics: an integrated approach. *Journal of Transport Geography*. 2013;31:64–71. Available from: https://doi.org/10.1016/j.jtrangeo.2013.05.009.
16) Anderson, Tessa K. Kernel density estimation and K-means clustering to profile road accident hotspots. *Accident Analysis & Prevention*. 2009;41(3):359–364. Available from: https://doi.org/10.1016/j.aap.2008.12.014.
17) Akhtar M, Moridpour S. A Review of Traffic Congestion Prediction Using Artificial Intelligence. *Journal of Advanced Transportation*. 2021;2021:1–18. Available from: https://doi.org/10.1155/2021/8878011.
18) Bengio Y. Learning deep architectures for AI. *Foundations and Trends in Machine Learning*. 2009;2(1):1–127. Available from: https://doi.org/10.1561/2200000006G.
19) Chan AB, Vasconcelos N. Classification and retrieval of traffic video using auto-regressive stochastic processes. *IEEE Proceedings Intelligent Vehicles Symposium, 2005*. 2005;p. 771–776. Available from: https://doi.org/10.1109/IVS.2005.1505198.
20) Harilakshmi VS, Jansi PA, Rani. Intelligent vehicle counter - a road to sustainable development and pollution prevention (P2). *International Conference on Energy Efficient Technologies for Sustainability*. 2016;877. Available from: https://doi.org/10.1109/ICEETS.2016.75838711.
21) Talha SUM, Tanzila, Draz H, Umer R, Amjad G, Khan KMU, et al. Intelligent Traffic Signal Automation Based on Computer Vision Techniques Using Deep Learning. *IT Professional*;24(1):27–33. Available from: https://doi.org/10.1109/MITP.2021.3121804.
22) Imad S, Andres BPU, Omar. Deep convolutional neural network architecture for urban traffic flow estimation. *International Journal of Computer Science and Network Security*. 2018;18(7). Available from: https://www.semanticscholar.org/paper/Deep-convolutional-neural-network-architecture-for-imad-Andres/8b85ae148268211d1b364b228bd55f4415d0187f.
23) Donato D, Impedovo F, Balducci V, Dentamaro G, Pirlo. Vehicular Traffic Congestion Classification by Visual Features and Deep Learning Approaches: A Comparison. *Sensors*;2019(23). Available from: https://doi.org/10.3390/s19235213.
24) Luong ATN, Ha T, Xuan. A Novel Approach of Traffic Density Estimation Using CNNs and Computer Vision. *European Journal of Electrical Engineering and Computer Science*. 2021;p. 80–84. Available from: https://doi.org/10.24018/ejece.2021.5.4.353.