Improved Spatial Invariance for Vehicle Platoon Application using New Pooling Method in Convolution Neural Network

M S Sunitha Patel[1], Srinath S[2]

Dept. of Computer Science & Engineering, ATME College of Engineering, Mysuru, Karnataka, India[1]
Dept. of Computer Science & Engineering, SJCE College of Engineering, Mysuru, Karnataka, India[2]

***Abstract*—The imbalanced dataset is a prominent concern for automotive deep learning researchers. The proposed work provides a new mixed pooling strategy with enhanced performance for imbalanced vehicle dataset based on Convolution Neural Network (CNN). Pooling is crucial for improving spatial invariance, processing time, and overfitting in CNN architecture. Max and average pooling are often utilized in contemporary research articles. Both techniques of pooling have their own advantages and disadvantages. In this study, the advantages of both pooling algorithms are evaluated for the classification of three vehicles: car, bus, and truck for imbalanced datasets. For each epoch, the performance of max pooling, average pooling, and the new mixed pooling method was assessed using ROC, F1-score, and error rate. Comparing the performance of the max-pooling method to that of the average pooling method, it has been found that the max-pooling method is superior. The performance of the proposed mixed pooling approach is superior to that of the maximum pooling and average pooling methods. In terms of Receiver Operating Characteristics (ROC), the proposed mixed pooling technique is approximately 2 per cent better than the maximum pooling method and 8 per cent better than the mixed pooling method. Using a new pooling technique, the classification performance with an imbalanced dataset is improved, and also a novel mixed pooling method is proposed for the classification of vehicles.**

Keywords*—Average pooling; convolution neural network; imbalance dataset; max pooling; mixed pooling

I. Introduction

Automotive companies are conducting cutting-edge research on vehicle platoon management to improve 'vehicle-to-vehicle' communication for enhanced performance. Vehicle platoon management increases customer and societal benefits by attaining greater fuel efficiency, less pollution, less road congestion, and fewer road accidents [1] to [2]. To improve performance, a key feature of vehicle platoon management is categorizing vehicles based on their size and grouping them suitably to achieve reduced aerodynamic drag when driving in the longitudinal direction, as illustrated in Fig. 1. To achieve the requirement, vehicles having Advanced Driving Assistance System (ADAS) features such as lane-keeping, automated cruise control, pedestrian safety, platoon management, and others must have at least two cameras installed. Vehicle platoon management is also categorized as ADAS [3], which falls within the L2 and L3 levels of vehicle automation. As vehicle automation level increases from L1 to L5 then vehicle intelligence should also increase to improve safety. Considering the features given by automotive manufacturers for modern vehicles, it is clear that camera utilization is expanding day by day, necessitating further study into image processing algorithms. Strict safety requirement calls for the need for more cameras in vehicles along with other sensors. As the number of cameras on a vehicle increases, image processing output will become more vital for controlling the vehicle according to safety requirements.

CNN has been widely used in vehicle image processing applications like pedestrian safety [4], vehicle classification [5,6,7], and many more applications. The focus of current research is on categorizing vehicles by type, which is the initial stage in vehicle platoon management before combining them for greater efficiency. To begin with car, bus, and truck are three classes of vehicles that are regularly seen on highways, and these three have been taken into consideration for experimental work using CNN. To train the CNN algorithms, high-quality and sufficient volumes of images are required. Unlike other applications where image datasets are publicly available, such as medical imaging, sign language identification, pattern recognition, and so on, the availability of vehicle image datasets is limited. Open repository datasets may not contain all of the images needed to do experimental work on the specified topic. In another instance, researchers have created datasets for vehicle rear parts in [8], and a few vehicle datasets are published in open source PKU-VD [9], VeRi-776 [10], VehicleID [11], as well as different vehicle datasets are available for smart city study [12]. But, to get dataset access from open source, researchers need to get approval from the owners. However, researchers must obtain permission from the owners of open-source datasets to access them. Still, sometimes delays or no response can be expected when seeking approval from the owners. There are few open access datasets for automotive applications [14,15,16]. While developing required image dataset, all this leads to an imbalanced image dataset collection, which is a prevalent problem in the automotive arena. In multiclass analysis, imbalance datasets indicate that the dataset's distribution includes unequal amounts of images for each class. Uneven quantities of images will impact the accuracy of the learning algorithms. In comparison to the amount of car images, the availability of images in open source for bus and truck images are minimal. As a result, various image augmentation techniques such as flipping, rotation, cropping, and so on are used to increase the quantity of bus and truck images. Therefore, to increase the classification performance even with an imbalanced dataset a mixed pooling approach has been proposed in this research. In the further sections deal with the works carried out related to vehicle classification using CNN and the motivation behind the mixed pooling implementation to improve the performance of the CNN learning algorithm for the imbalance vehicle datasets has been explained.

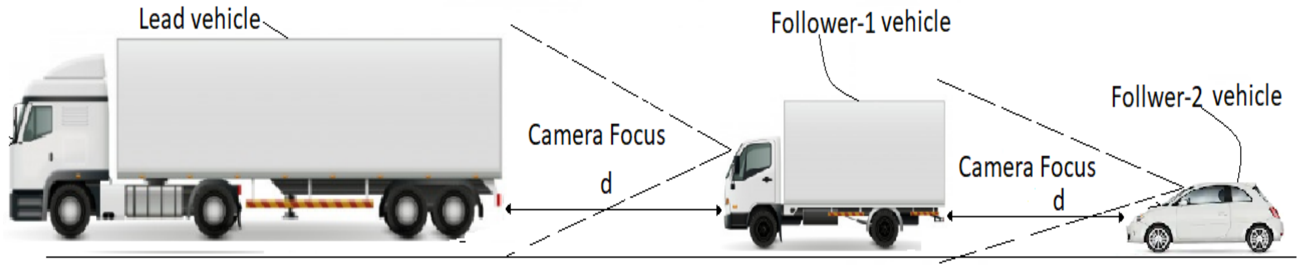

Fig. 1. Example for Three Vehicles in Platoon Management, Where the First Vehicle is called as Lead Vehicle and subsequent Two Vehicles are called as Follower-1 and Follower-2. Vehicles Separated from each with Distance - d.

II. Related Work

Deep Learning methods for image processing currently dominate computer vision, particularly for image recognition related applications. Several works utilising Deep Learning for vehicle detection are documented in the literature.

Using the CNN deep learning technique, [23] proposed a daytime vehicle detection system. Using appearance-based feature extraction techniques, the experiment yielded improved results. Fast R–CNN-based vehicle classification algorithm for real-time traffic surveillance was developed by Wang et al. [24]. A dataset of 60,000 images depicting traffic junction was compiled and divided into training and tested data, on which the suggested approach achieved an accuracy of 80.051%. Chauhan et al. [25] have developed a CNN-based framework for the classification and counting of vehicles on highways. On the collected dataset of 5,562 CCTV camera videos of highway traffic, the suggested framework achieved 75 percent MAP. Jo et al. [13] have suggested a GoogLeNet framework for vehicle categorization based on transfer learning. The authors demonstrated that the provided classifier achieved a 0.983% accuracy rate on the ILSVRC-2012 data set. Chia-Chi Tsai et al., 2018 [17], developed an improved Convolutional Neural Network architecture based on deep learning methods for intelligent transportation applications. Improved Spatio-Temporal Sample Consensus is the name of a method developed by Yu Wang et al. [26] for detecting and classifying moving vehicles. First, the moving vehicles are detected using the Spatio Temporal Sample Consensus technique, based on the brightness variation and shadow of the vehicles. In addition, using feature fusion algorithms, the objects are categorised based on area, face, licence plate, and vehicle symmetry characteristics. In the study proposed by Kaiming He [18], the deep networks are equipped with the pooling approach called spatial pyramid pooling. This eliminates the requirement of fixed size image being given as input to the network. The novel network structure, termed SPP-net, may provide a fixed-length representation independent of image size/scale. Pyramid pooling is also robust to object deformations. With these features, SPP-net improves all CNN-based image classification algorithms.

Deep feature-based techniques can effectively improve the accuracy of vehicle classification, but they require an enormous quantity of data to attain considerable accuracy. Hence, in this current work an attempt has been made to achieve good accuracy with moderate dataset size.

III. Motivation

CNN's components include a convolution layer, a pooling layer, and a flattening layer that performs feature extraction and uses the output of the flattening vector for classification. The convolution layer performs a linear operation that extracts a feature map from an image. For the given image, the i^{th} convolution layer generates i^{th} output feature map and it can be represented as mentioned in (1) where ‘w_i’ is kernel or filter and ‘x’ is the input image with 2D convolution operator and y is an output.

$$y_i = f(w_i * x) \tag{1}$$

The kernel used in convolution will travel across the image left to right and top to bottom, the amount of movement for the given input image depends on the input image size and kernel size. The activation function in the convolution layer produces a feature map, which is fed into the pooling layer. Pooling is a nonlinear function that generates output by summing the net of certain pixel positions. The pooling layer has numerous advantages in CNN architecture, including improved processing time, noise invariance, and overfitting. In the following sections below some of the most commonly utilized pooling strategies are discussed.

A. Max Pooling

Max Pooling selects the maximum value in the feature map for the selected filter dimension. Here computation has made faster [19, 20] because of the elimination of non- max values and the output of the max pool will reduce from X to Y as mentioned in Fig. 2. By selecting the max value in the feature map, max-pooling picks the brightest pixel from the given image. Max pooling can be represented mathematically as given in (2). Where for given i^{th} feature map x is the input element at (p,q) within pooling region R_{jk} which represents local neighbourhood position (j,k)with output y.

$$y_{ijk} = \max_{(p,q)\in R_{jk}} (x_{ipq}) \tag{2}$$

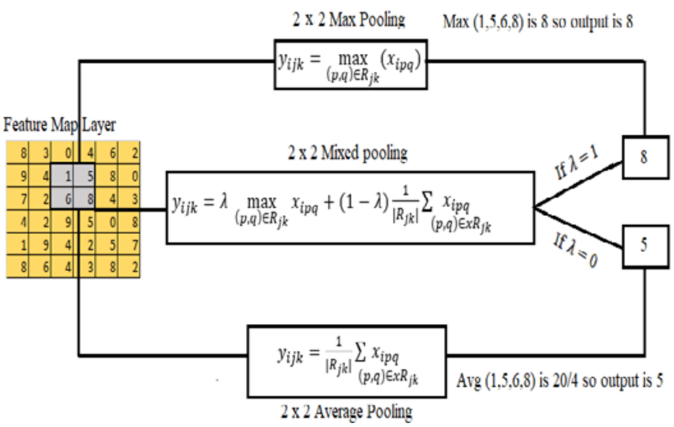

Fig. 2. Toy Example to Illustrate 2x2 Max, Average and Mixed Pooling Principle.

B. Average Pooling

Average pooling selects the average value in the feature map for the selected filter dimension. The computation is faster [21] since it selects the average value of each region and the output value is lowered from X to Y, as illustrated in Fig. 2. By selecting the average value in the feature map, average pooling picks the average pixel from the given image. Average pooling is represented mathematically as given in (3). Where for given i^{th} feature map x is the input element at (p,q) within pooling region R_{jk} which represents local neighbourhood position (j,k)with output y.

$$y_{ijk} = \frac{1}{|R_{jk}|} \sum_{(p,q) \in xR_{jk}} x_{ipq} \quad (3)$$

C. Mixed Pooling

Max pooling performance will degrade when the feature map has a low pixel value and average pooling performance will degrade when the feature map has high pixel values. To overcome this problem, mixed pooling approach has been implemented [22]. As illustrated in Fig. 2, the mixed pooling approach combines both max and average pooling by picking any one method during execution. Mixed pooling can be represented as in (4). Where λ can be 1 or 0 and is chosen at random; pooling strategy is average if λ is 0 and pooling strategy is to max if λ is 1.

$$y_{ijk} = \lambda \max_{(p,q) \in R_{jk}} x_{ipq} + (1-\lambda) \frac{1}{|R_{jk}|} \sum_{(p,q) \in xR_{jk}} x_{ipq} \quad (4)$$

D. Proposed Novel Method

In the current work, novel pooling method has been proposed to get the benefits of mixed and max pooling. Because mixed pooling has the advantage of taking into account both the max and average value of the pixels, it has been implemented after the first layer of convolution, and max pooling has been implemented in the second and third layer pooling, as shown in Fig. 3 for a three-layer CNN. The proposed method has the advantage that once the first convolution layer captures the maximum needed feature using mixed pooling; the subsequent layer using max pooling will reinforce the features.

IV. Methodology

The present work begins with gathering bus, car, and truck images and creating the requirement dataset for the vehicle platoon environment. With the developed dataset training, testing and validation has been performed for three different pooling scenarios as presented in Fig. 3. Performance evaluation for all three pooling scenarios has been visualized using ROC Curve.

A. Dataset Creation

To undertake experimental work in deep learning, the first step is to develop a suitable image dataset to train the deep learning algorithm. The car images were acquired from an open-source repository [14], the bus images were acquired from an open-source repository [15], and truck images were acquired from the open-source repository [16]. There were approximately 3000 images available for truck and bus, and nearly 8000 images were acquired for the car vehicle class. The image data size for bus and truck has been increased from 3000 images to 5000 images using image augmentation techniques. Different image augmentation techniques like padding, flip, rotate, tilt, blur, crop, adding noise, etc. have been incorporated based on the available images. Therefore, the total number of images used to conduct the experiment is 5000 for the bus and truck classes and 8000 for the vehicle class.

B. Training, Testing and Validation Dataset

As demonstrated in Fig. 4 by utilizing open source and different augmentation techniques final dataset has been created. The training and testing ratio are set at 80% and 20%, respectively. To optimize the CNN algorithm's learning performance, 20% validation datasets were used.

C. CNN Architecture

As mentioned earlier, a three-layer Convolution Neural Network has been implemented as shown in Fig. 5. The experiment has been carried out using Python 3.7.9, with IDE PyCharm and OpenCV Library. The CNN architecture has taken care of feature extraction, including filtering, image segmentation, and image enhancement for images that are 64 by 64 in size.

CNN model consists of convolution layer, pooling layer, fully connected layer, and output layer. The first convolution layer consists of 32 filters, followed by 64 filters in the second and third layers. Padding has been considered to maintain input and output image size. Three alternative pooling strategies have been studied in terms of pooling techniques. In Scenario-1, all three pooling configurations are max pooling. In Scenario-2, all three pooling configurations are mixed pooling and in Scenario-3 the first pooling configuration is mixed pooling and the second two pooling configurations are max pooling. The output of pooling with window size 2 and stride size 2 is fed to the size 2 of the fully connected layer. The fully connected layer is the deep layer which is used for classification. Finally, the fully connected layer passes its input to softmax with 2 layers for the final classified image output. Each scenario's output is executed in three distinct steps, and the outcomes are summarized and discussed in the following section.

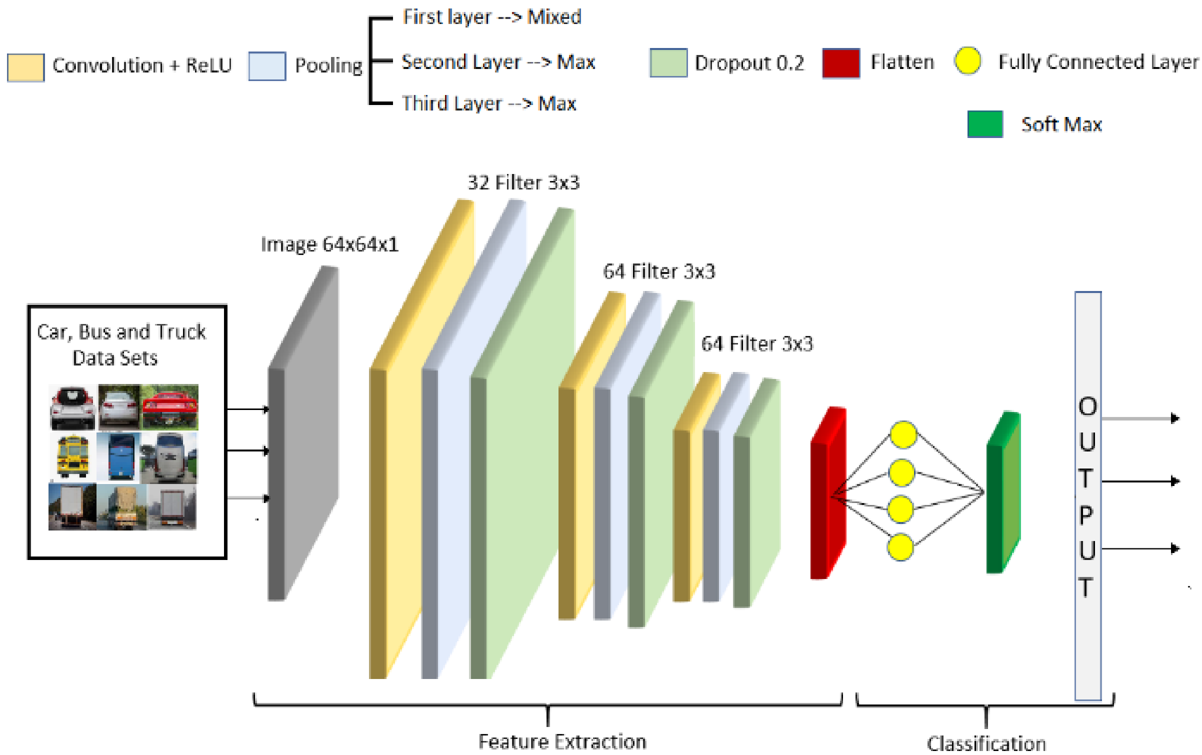

Fig. 3. Proposed Architecture Where the First Pooling Layer will be mixed and the subsequent Pooling Layer will be Max.

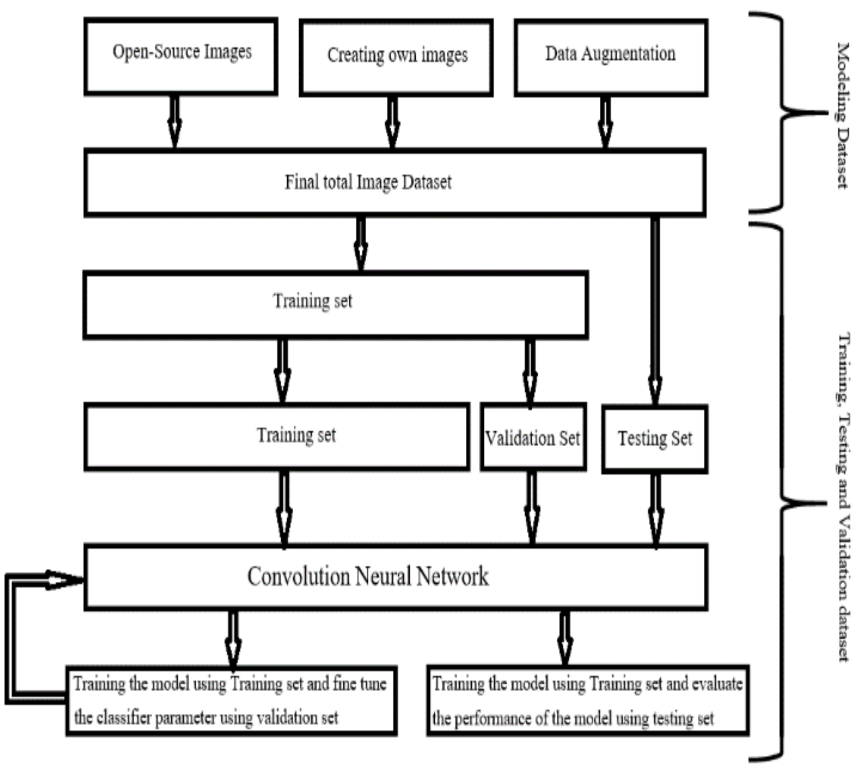

Fig. 4. Training Dataset 80% and Testing Dataset 20 %, for Fine-tuning the Learning Algorithm 20% of the Validation Dataset has been used.

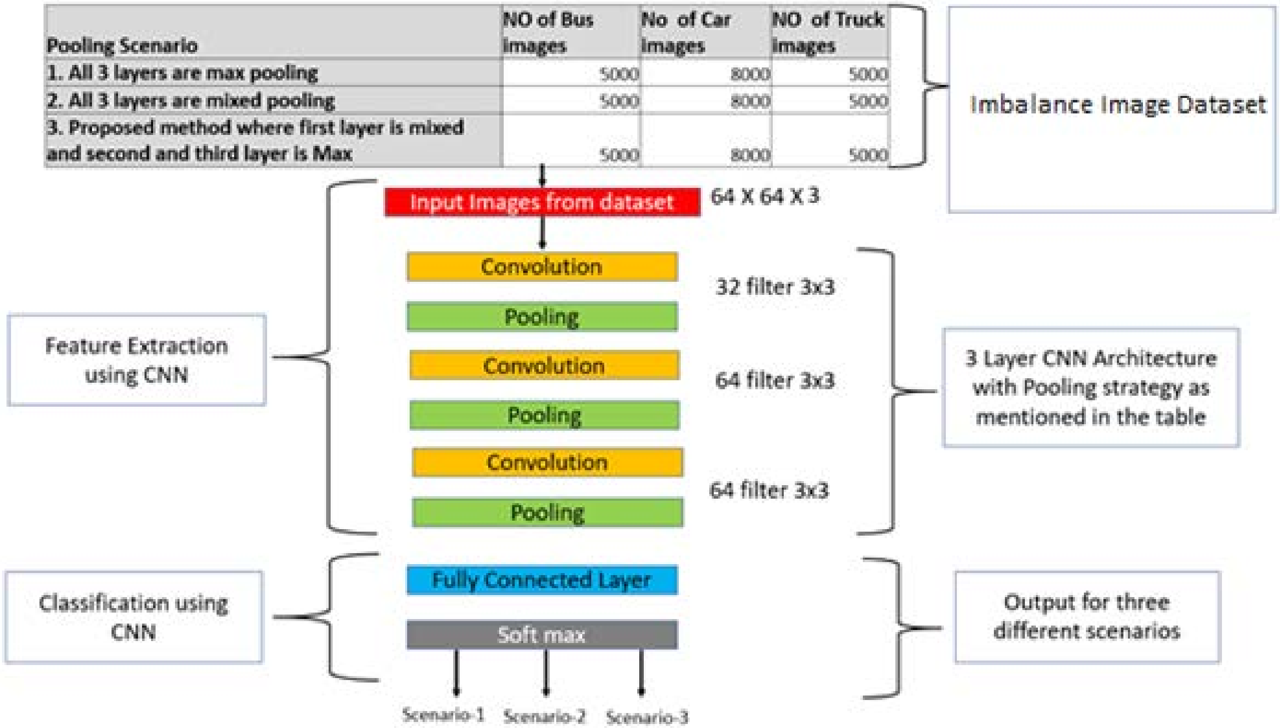

Fig. 5. Implementation Methodology.

D. Performance Evaluation

Standard formulae of accuracy, precision, recall, and F1 score, as indicated (5) through (8), were used to evaluate the performance of three distinct pooling procedures. Here, True Positive (TP) is a correctly predicted class, False Positive (FP) is a label that does not belong to class but is predicted as positive, True Negative (TN) is the correctly predicted for class that does not belong to the class, False Negative (FN) is wrongly predicted for class that does not belong to the class.

$$\text{Accuracy} = \frac{(TP + TN)}{(TP + TN + FP + FN)} \tag{5}$$

$$\text{Precision} = \frac{TP}{(TP + FP)} \tag{6}$$

$$\text{Recall} = \frac{TP}{(TP + FN)} \tag{7}$$

$$\text{F1 Score} = \frac{2 \times TP}{((2 \times TP) + FP + FN)} \tag{8}$$

V. Results and Discussion

As depicted in Fig. 5. imbalance image datasets were fed into the CNN model and run individually for each of the three cases and the results obtained are as shown in Fig. 6. In Fig. 6, label 0 denotes the bus class, label 1denotes the car class, and label 2 denoted the truck class. ROC graphs are also being used to visualize the results of Scenario-1, Scenario-2, and Scenario-3 as illustrated in Fig. 7, Fig. 8, and Fig. 9, respectively.

A. Scenario-1(All Three Max Pooling)

Fig. 7 illustrates the accuracy attained for all three classes using the ROC curve for Scenario-1. The acquired findings show that the achieved average accuracy is 98 % with 100 % accuracy for vehicle classification and 95 % and 97 %accuracy for bus and truck classification. After execution, false-positives for bus and truck are 150 and 250 images, respectively, and false-negatives are 250 and 150 images for the 20% overall dataset during testing.

B. Scenario -2 (All Three Mixed Pooling)

The ROC curve for Scenario-2 is shown in Fig. 8. The average accuracy achieved is 98%, with 100% classification for car class and 94 % and 97 % for bus and truck class, respectively. The false-positive rate for bus and truck is 150 and 300 images, respectively, whereas the false-negative rate for the 20% overall dataset during testing is 300 and 150 images.

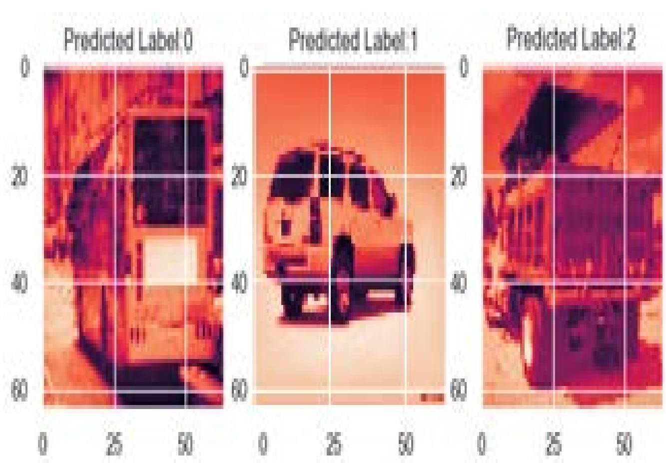

Fig. 6. Predicted Output (Label Convention: Label 0 -Bus, Label 1- Car, Label 2-truck).

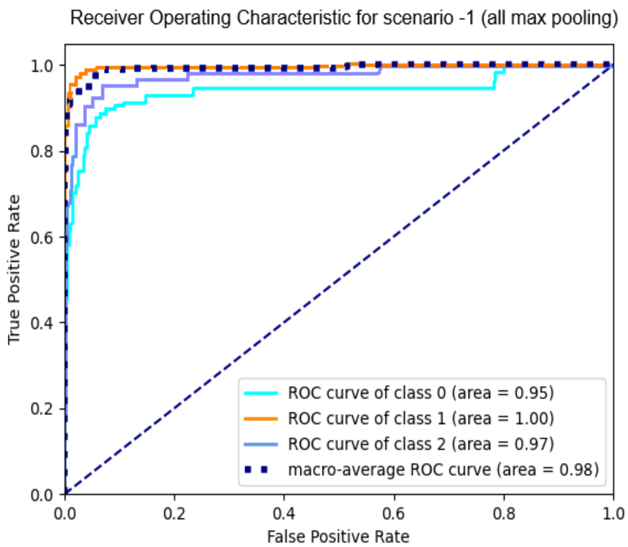

Fig. 7. ROC for Scenario-1, Where Class 0 is for Bus, Class 1 is for Car and Class 2 is for Truck.

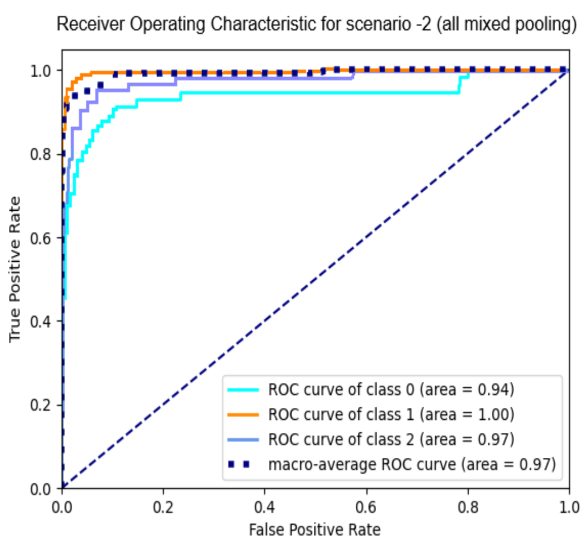

Fig. 8. ROC for Scenario-2, where Class 0 is for Bus, Class 1 is for Car and Class 2 is for Truck.

C. Scenario-3 (New Mixed Pooling)

Fig. 9 depicts the ROC curve for Scenario-3. While 100% of cars are correctly classified, 94% and 97% of buses and trucks are correctly classified. The number of false-positive images for the bus and truck is 150 and 50 respectively and the number of false-negative images is 50 and 150, respectively for the 20% overall dataset during testing.

D. Epoch Error Rate

From Fig. 10, it can be observed that the learning cycle has been completed nearly at 50 epochs for all three scenarios. However, in the proposed methodology, the pooling error rate is decreased to 0.7 percent, whereas the error rate in max-pooling is 1% and 1.4 per cent in mixed pooling.

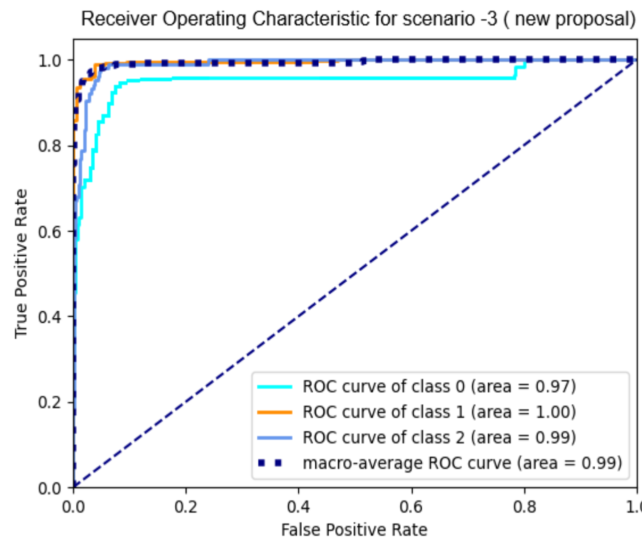

Fig. 9. ROC for Scenario-3, Where Class 0 is for Bus, Class 1 is for Car and Class 2 is for Truck.

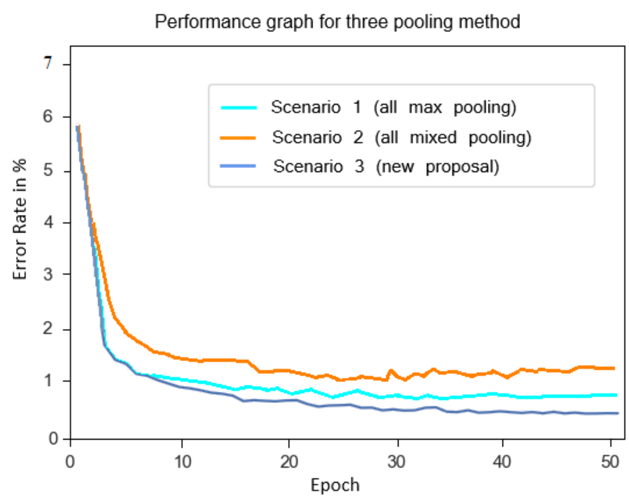

Fig. 10. Performance at different Epochs of for Three different Pooling Scenarios.

E. Overall Performance Evaluation

As the literature review reveals no evidence of deep learning based vehicle grouping to aid in vehicle platoon management, an attempt has been made where the proposed methodology enhances the performance of vehicle classification using an imbalanced image dataset. For the three different situations that were taken into consideration for the experimentation, Table I list the performance metric values for accuracy, precision, recall, and f1-score for multiclass vehicle classification. From the experimental outcomes, it can be inferred that the suggested strategy outperformed other methods while using its own dataset for the platoon management system.

To demonstrate the improved performance of the proposed method, the results obtained from the proposed method is compared with the other methods as shown in Table II. The proposed method also shows a consistent classification performance for all the vehicle classes.

TABLE I. SUMMARY OF THE PERFORMANCE EVALUATION FOR ALL THREE SCENARIOS

Class	Performance Metric	Scenari-1 output in %	Scenari-2 output in %	Scenari-3 output in %
BUS	Accuracy	95	94	97
	Precision	97	97	99
	Recall	95	95	97
	F1 Score	96	95	98
Car	Accuracy	100	100	100
	Precision	100	100	100
	Recall	100	100	100
	F1 Score	100	100	100
Truck	Accuracy	97	97	99
	Precision	95	94	97
	Recall	97	97	97
	F1 Score	96	96	98
Bus, Car & Truck	Average Accuracy in % for all 3 class	98	97	99
Bus, Car & Truck	Average error rate in % at epoch 50	1.4	1	0.7

TABLE II. PERFORMANCE COMPARISON WITH PROPOSED METHOD TO OTHER METHODS

Method	Vehicle class considere d	Accuracy	Precision	Recall	F1 Score
F-RCNN (8)	Bike , Bus and Truck	94.4	88.6	93.5	88
YOLO (7)	Bus, Car, and Truck	99	Not applicable	Not applicable	Not applicable
CNN-Super learner (6)	Bus, Car, Truck, Pedestrian, and Bike	99	98	98	99
Propose d method	Bus, Car, Truck	99	99	99	99

VI. CONCLUSION

A frequent challenge in deep learning-based vehicle classification task is obtaining sufficient visual data for the experimentation. In this regard, a customized dataset for vehicle platoon management has been created combining open source repositories and employing image augmentation techniques. The customized dataset has resulted in an imbalanced dataset consisting of 8000 car images, 5000 bus images, and 5000 truck images. The analyses carried out showed that, the performance of classification algorithm with existing methods considering the imbalance dataset is inconsistent across all vehicle classes, lowering the overall performance of the classification task. Whereas, the proposed novel mixed pooling method with three-layer CNN architecture performs significantly well even for the imbalanced dataset. This also highlights the importance of pooling in CNN. Additionally, the selection of the optimal pooling mechanism plays crucial role in optimising the efficiency of the learning algorithm in a CNN architecture. Employing the proposed method, experimental results show that the suggested strategy outperforms the traditional max and mixed pooling approaches by 2% and 8%, respectively for the imbalanced bus and truck datasets. From the accuracy gained using the proposed method, it is evident that selecting the optimal pooling mechanism is crucial for boosting the performance of CNN architecture. Thus, proposed mixed pooling method outperformed other methods on the imbalance dataset. Further research will focus on the performance of the proposed mixed pooling technique with additional CNN layers and a larger number of epochs.

REFERENCES

[1] Wang, F., Dai, H., Lu, Y., and Han, H., "Management and Control of Connected and Automated Vehicle Platoon in the Process of Variable Speed Driving, Vehicle Cut-Out or Cut-In", SAE Technical Paper 2020-01-5220, 2020.

[2] Zhang, Y., Bai, Y., Hu, J., and Wang, M., "Control Design, Stability Analysis and Traffic Flow Implications of CACC Systems with Compensation of Communication Delay", Transportation Research Record Journal of the Transportation Research Board 0(0):1-15, 202.

[3] C. Peng, M. M. Bonsangue and Z. Xu "Model Checking Longitudinal Vehicle Platoon Systems", in IEEE Access, vol. 7, pp. 112015-112025, 2019.

[4] Venkatesh Muravaneni, Amit Babalal Nahar, and Pratima Vishwakarma, "A Modular Approach to Vehicle Management in a Platoon Group", SAE Technical Paper 2021-26-0125, 2021.

[5] Ajitha; Jeyakumar. S; Yadhu Nandha Krishna K; A Sivasangari, "Vehicle Model Classification Using Deep Learning", IEEE 5th International Conference on Trends in Electronics and Informatics (ICOEI), June 2021.

[6] M. A. Hedeya, A. H. Eid and R. F. Abdel-Kader, "A Super-Learner Ensemble of Deep Networks for Vehicle-Type Classification", IEEE Access, vol. 8, pp. 98266-98280, 2020.

[7] Song, H., Liang, H., Li, H. et al "Vision-based vehicle detection and counting system using deep learning in highway scenes", Eur. Transp. Res. Rev. 11, 51 (2019).

[8] V Sowmya, R Radha, "Real-time Vehicle Detection implementing Deep Convolutional Neural Network Features Data Augmentation Technique", Indian Journal of Science and Technology 2021;15(1):44-5.

[9] Y. Lou, Y. Bai, J. Liu, S. Wang, and L. Duan," VERI-wild: A large dataset and a new method for vehicle re-identification in the wild", Proc. EEE/CVF Conf. Computer Vis. Pattern Recognition. (CPR), Jun. 2019, pp. 3235_3243.

[10] M. Naphade, S. Wang, D. C. Anastasiu, Z. Tang, M.C. Chang, X. Yang, Y. Yao, L. Zheng, P. Chakraborty, C. E. Lopez, and A. Shama " The 5th AI city challenge", Proc. IEEE/CVF Conf. Computer. Vis. Pattern Recognition Workshops, Jun. 2021.

[11] T. Ki and Y. Hur, "Deep Scattering Network with Max-Pooling", Data Compression Conference (DCC), 2021, pp. 348-348.

[12] T. Theodoridis, K. Lumponias, N. Vretos, and P. Daras, "Zernike Pooling: Generalizing Average Pooling Using Zernike Moments", IEEE Access, vol. 9, pp. 121128-121136, 2021.

[13] S. Y. Jo, N. Ahn, Y. Lee, and S. J. Kang, "Transfer learning-based vehicle classification," Proceedings of the 2018 International SoC Design Conference (ISOCC), pp. 127-128, IEEE, Daegu, South Korea, 2018.

[14] Jonathan Krause, Michael Stark, Jia Deng, Li Fei-Fei, "3D Object Representations for Fine-Grained Categorization", 4th IEEE Workshop on 3D Representation and Recognition, at ICCV(3dRR-13),Sydney, Australia,2013.

[15] Tabassum, Shaira, Ullah, Md. Sabbir, Al-nur, Nakib Hossain, Shatabda, Swakkhar 2020. "Poribohon-BD". Mendeley Data, V2.

[16] https://github.com/datacluster-labs/Indian-Vehicle-Image-Dataset.

[17] Tsai, C.C., Tseng, C.K., Tang, H.C., Guo, J. "Vehicle detection and classification based on deep neural network for intelligent transportation application", APSIPA Annual Summit and Conference 2018, IEEE.

[18] He K., Zhang X., Ren S., Sun J., "Spatial pyramid pooling in deep convolutional networks for visual recognition", IEEE Transactions on Pattern Analysis and Machine Intelligence, 37, 9, 2015, 1904-1916.

[19] Yu D., Wang H., Chen P., Wei Z., "Mixed pooling for convolutional neural networks", International Conference on Rough Sets and Knowledge Technology, 2014, 364-375.

[20] Zeiler M. D., Fergus R., "Stochastic pooling for regularization of deep convolutional neural networks", arXiv preprint arXiv:1301.3557, 2013, 1-9.

[21] Hongye Liu, Yonghong Tian, Yaowei Yang, Lu Pang, and Tiejun Huang, "Deep relative distance learning: Tell the difference between similar vehicles", Proc. CVPR, pages 2167–2175, 2016.

[22] Ke Yan, Yonghong Tian, Yaowei Wang, Wei Zeng, and Tiejun Huang, "Exploiting multi-grain ranking constraints for precisely searching visually-similar vehicles", Proc. ICCV, pages 562–570, 2017.

[23] Chen, L., Ye, F., Ruan, Y. *et al*. "An algorithm for highway vehicle detection based on convolutional neural network." J Image Video Proc. 2018**,** 109 (2018).

[24] X. Wang, W. Zhang, X. Wu, L. Xiao, Y. Qian, and Z. Fang, "Real-time vehicle type classification with deep convolutional neural networks", Journal of Real-Time Image Processing, vol. 16, no. 1, pp. 5–14, 2019.

[25] M. S. Chauhan, A. Singh, M. Khemka, A. Prateek, and R. Sen, "Embedded CNN based vehicle classification and counting in non-laned road traffic", Proceedings of the Tenth International Conference on Information and Communication Technologies and Development, pp. 1–11, Ahmedabad, India, 2019.

[26] Wang, Y., Ban, X., Wang, H., Wu, D., Wang, H., Yang, S., Liu, S., Lai, J, "Detection and classification of moving vehicle from video using multiple spatio-temporal features", recent advances in video coding and security. IEEE Access 7, 80287–80299 (2019).