Bulletin of Electrical Engineering and Informatics
Vol. 15, No. 6, December 2026, pp. 5654~5663
ISSN: 2302-9285, DOI: 10.11591/eei.v15i6.13430

Deep learning assisted automated vehicle detection and classification for intelligent transportation systems

Abel Jopaul Vettilkattil Paulose[1], Prabhu Arthanareeswaran[2], Mohammad Shabbir Alam[3], Barkathulla Asan Mohamed[4], Anjaiah Talamala[5], Manoj Senthil Kailasam[6]

[1]Department of Computer Applications, LEAD College (Autonomous), Palakkad, India
[2]School of Computer Science and Information Technology, Jain (Deemed-to-be University), Bengaluru, India
[3]Department of Computer Science, College of Engineering and Computer Science, Jazan University, Jizan, Saudi Arabia
[4]Department of Computer Science and Engineering, Vel Tech Rangarajan Dr. Sagunthala R&D Institute of Science and Technology, Chennai, India
[5]Aditya University, Surampalem, India
[6]Department of Electronics and Communication Engineering, Kongu Engineering College, Perundurai, India

Article Info

Article history:

Received May 21, 2026
Revised Aug 27, 2026
Accepted Sep 12, 2026

Keywords:

Crayfish optimization algorithm
Deep belief network
Intelligent transportation system
Transfer learning
Vehicle classification

ABSTRACT

Vehicle classification is one of the key elements of intelligent transportation systems (ITS). Various ITS applications, such as surveillance structures, security systems, fleet monitoring, traffic protection, and automatic parking employs vehicle classification. Current developments in artificial intelligence (AI) have greatly impacted the study related to vehicle detection and classification. This study presents a fusion-based transfer learning model with a crayfish optimization algorithm for robust vehicle classification in intelligent transportation system (FTLMCOA-VCITS). The FTLMCOA-VCITS model aims to enhancing vehicle classification using advanced deep learning (ADL) models. In the initial stage, the image preprocessing utilizes Wiener filtering (WF) to remove the noise. For the feature representation process, the fusion of the DenseNet-201, EfficientNet-B7, and CapsNet is employed. For the classification process, COA with deep belief network (DBN) approach is used. The comprehensive comparative results demonstrate an increased accuracy of 99.08% over existing methods under the vehicle classification dataset.

This is an open access article under the CC BY-SA license.

Corresponding Author:

Abel Jopaul Vettilkattil Paulose
Department of Computer Applications, LEAD College (Autonomous)
Palakkad, Kerala – 678009, India
Email: abel.jp@lead.ac.in

1. INTRODUCTION

With the growing network of roads and rising traffic worldwide, utilizing modern technology to monitor and manage traffic has become a necessity. Because of this, intelligent transportation systems (ITS) have gained more interest in the past years [1]. The ITS is a possible mechanism for overcoming traffic problems. Traffic movement is determined as a significant and basic factor in ITSs while detecting and counting the number of vehicles with traffic videos accurately, as well as rapidly [2]. Vehicle recognition is an important task in the traffic monitoring system. Once the vehicle is identified, further applications will be used effortlessly. In the past, active sensors, namely light detection and ranging (LIDAR), millimeter-wave lasers, or radars, were employed in detecting vehicles. Vehicle detection and classification are important stages of planning ITS. After a vehicle is identified, data is then employed for activity detection, congestion analysis, speed analysis, and anomaly detection. Conventional procedures have an extremely restricted visual

field and are also expensive [3]. Images and videos are the most commonly leveraged methods to identify vehicles. These are cost-effective solutions as the data will be recovered with low-cost traffic cameras fixed on the roads. The analysis and processing period is much less when compared with conventional procedures. Recently, many studies have been done in the deep networks domain for building smart transport systems, and it is noted that this model produces outstanding results for image and audio/video classifications. The actual inspiration for utilizing deep learning (DL) is the quick processing of data in the training process [4]. Due to this, the method trains on its own and pulls the essential characteristics needed for classification and detection.

Khang *et al.* [5] examined the wireless sensor network (WSN) application in ITS. Smart transport uses information technology (IT) to realize real-time automated gathering of traffic data. The incorporation of traffic direction, intervention models, and traffic signal control safeguards systematic traffic processes through real-time management. Wazid *et al.* [6] designed a secure DL-based malware attack recognition for internet of things-based ITS (SDLMA-IITS). The internet of thing (IoT)-based ITS devices and systems need to be secured from cyberattacks for several motives. Singh *et al.* [7] explored this kind of service securely through an intelligent automatic transportation method utilizing a vehicle delay-tolerant system. To answer this, the authors proposed an ITS for automatic medical service for protecting healthcare staff from getting infections in the collection and testing of health information by cooperating with delay-tolerant networking vehicles in intelligent transport systems. Nagarajan *et al.* [8] addressed this challenge and introduced a fresh Dumpster–Shafer data fusion-driven adversarial deep learning (ADL) for ITS. The projected module concentrates on 3 confrontational threat levels: feature level, decision level, and original image level. Confrontational samples are produced at every level to evaluate the vulnerability of the system widely. To improve the system's ability, numerous important modules are used.

Rani and Sharma [9] presented mmWave propagation capacities outdoors and indoors on 66 GHz frequency for internet of vehicles setups. The anticipated method inspected the corresponding route losses utilizing free-space path loss (FSPL), depending up the receiver as well as the transmitter distance for outside and inside vehicle communication. In the indoor setup, path loss propagation possesses less diffusion loss; however, it is ineffective in the outside setup due to the distance rising as FSPL rises. Elassy *et al.* [10] studied important modules of ITSs, underlining their role in improving transport safety, sustainability, and efficacy. It discovered current progressions in communication networks and enables real-time ITS processes, presenting the understanding of ecologically approachable smart cities. Zhang *et al.* [11] studied the widespread difficulties in ITS and explained the drivers in using FL from numerous perspectives. Waqar *et al.* [12] implemented the you only live once (YOLO) procedure. It is an NN-based method to detect vehicles on streets and roads. This method identifies vehicles precisely and quickly.

Yuan *et al.* [13] introduced finetuned versions of YOLOv5, YOLOv7, YOLOv8, YOLOv10, DETR-ResNet50, and Faster region-based convolutional neural network (Faster R-CNN). This study incorporated methods depending on YOLOv7, YOLOv8, and YOLOv10 further enhances the algorithm performance utilizing non-maximum suppression (NMS) and dynamic weighting models. Ashari *et al.* [14] presented the impact of multiple image augmentation models on YOLOv9t method for traffic vehicle detection. The methods evaluated include blur, brightness adjustment, contrast adjustment, color jitter, cropping, flipping, noise injection, rotation, scaling, and zoom-in. Results reveal that scaling and brightness adjustment significantly enhance recognize performance. Hasan *et al.* [15] suggested the BDNet combines three corresponding architectural modules: i) HyDASE, a hybrid detail-preserving downsampling model that moderates data loss; ii) C3k2_MogaBlock, which supports long-rang contextual interactions, and iii) A2C2f_FRFN neck, which refine multi-scale attributes by containing redundancy and highlighting discriminative responses.

Moreover, existing approaches frequently rely on single-feature extraction models, which may limit classification efficacy and generality. Thus, there is a need for an intelligent system that integrates multiple deep feature representations with optimization strategies to enhance vehicle classification accuracy in intelligent transportation environments. To deal with these limitations, this manuscript presents a fusion-based transfer learning model with a crayfish optimization algorithm for robust vehicle classification in intelligent transportation system (FTLMCOA-VCITS) model. The significant contributions of this work are summarized as given below:

a. Proposed a novel hybrid transfer learning feature extraction approach, which integrates DenseNet-201, EfficientNet-B7, and CapsNet for capturing both deep semantic features and spatial relationships, thereby offering more informative vehicle representations than single-network-based methods.
b. Employ a deep belief network (DBN) as the classifier to accurately categorize vehicles into different classes by efficiently capturing difficult feature dependencies extracted from the feature extraction method.
c. Utilize the COA to automatically optimize the DBN parameters, enhancing classification outcome, decreasing the model difficulty, and diminishing the need for manual parameter tuning.
d. Conduct extensive experiments on the vehicle classification dataset to evaluate the efficiency of the FTLMCOA-VCITS model. The study outcome illustrates that the proposed model reliably outperforms current state-of-the-art vehicle recognition models in terms of accuracy, precision, recall, and F1-score.

2. PROPOSED METHOD

This work introduces a FTLMCOA-VCITS approach for increasing robustness of vehicle classification in intelligent transportation environments using advanced DL techniques. The proposed model encompasses multiple phases, including preprocessing, feature representation, parameter tuning, and classification. Initially, image preprocessing is accomplished to increase image clarity and remove noise. Next, discriminative features are extracted and utilized for vehicle classification. Finally, parameter tuning is carried out to optimize the model performance. Figure 1 shows the overall workflow of the proposed FTLMCOA-VCITS methodology.

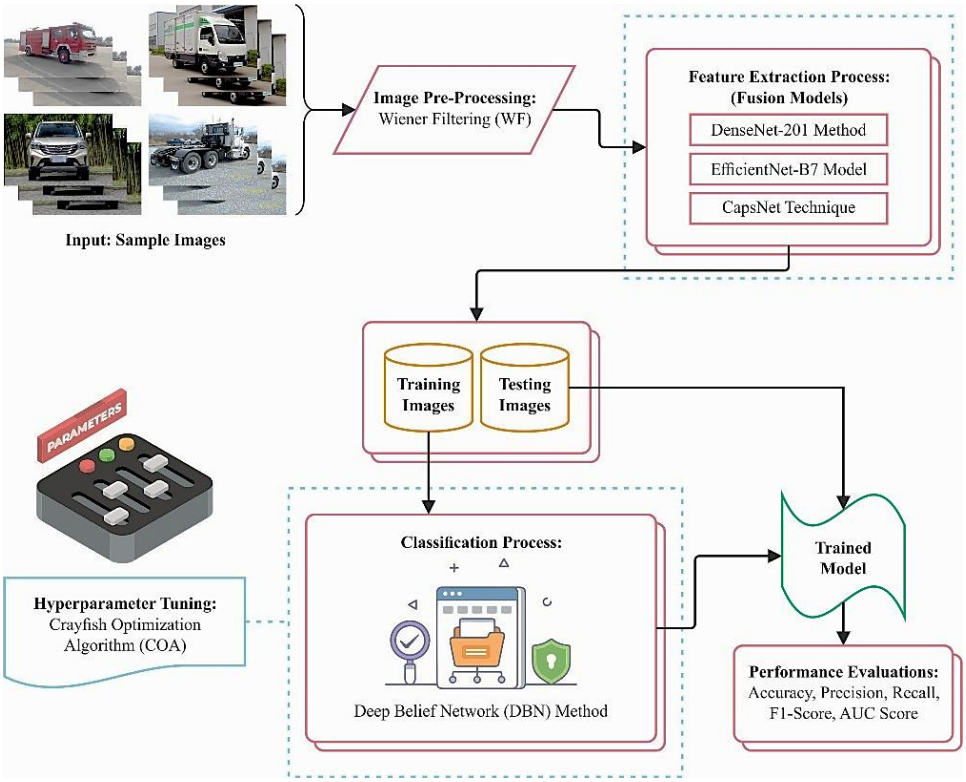


Figure 1. Overall working process of the FTLMCOA-VCITS approach

2.1. Image preprocessing

Initially, the image preprocessing stage employs Wiener filtering (WF) for removing the noise to enhance the image quality. WF is a signal image preprocessing system, which is exploited to eliminate noise and improve significant features in images. In the classification of vehicles, it aids in increasing the quality of images by lessening noise and blurring, making the extraction of features more efficient for DL techniques. The filter functions by minimizing the error of the mean square among the actual and estimated signals. It is mainly beneficial in states with lower-quality or noisy surveillance records. WF is adaptive, meaning it regulates depending on local signal and noise features. When combined with ML methods, it improves classification accuracy by delivering clearer input data.

2.2. Fusion-based extraction of feature method

For the feature representation procedure, the fusion of DenseNet-201, EfficientNet-B7, and CapsNet is employed. The DenseNet-201 structure consists of 2 transition blocks, 3 dense blocks, and a primary convolutional layer [16]. Particularly efficient representation and feature extraction are advanced by the densely connected convolutional layer, which includes all dense blocks. By inserting average pooling layers and transition blocks, small channels, in contrast, produce navigation among dense blocks at ease. Enhancing the DenseNet-201 method and making 4 variants by combining deep networking layers. The improved blocks are made with many layers linked by longer relationships. By stride 2, a 7×7 convolution layer is used in the primary stage. The Max-pool layer was applied before transferring the dense block. From beginning to end, the improved layer was comprised of dense blocks. The dense network model executes well in maximal

image classification applications and permits internal feature recycling that decreases the gradient death rate of the problem.

The EfficientNet-B7 method executes superior image tasks [17]. To enhance the precision and improve the classification approach, the transformed structures of the EfficientNet-B7 method were used. Select this method and improve it by inserting many layers. The block of depth-wise has 3×3 conv, smaller parameters, and captures the complete benefits of a current accelerator. The primary block made Conv 1×1 for the expansion Conv 1×1 in MB Conv and the depth-wise 3×3 conv. Afterward, an average pooling layer was included to increase the model's robustness. The method has a composite model that incorporates numerous sections. Balancing these parameters is an approximation method that results in accurate and efficient neural network topologies. EfficientNet-B7 implementation depends on compound scaling that achieves the sum of channels and layers, whereas H and W signify the height and the width, and C characterizes the channel counts. During this depth-wise convolution, this model used 4 channels, which enhanced the precision of this method.

CapsNet network comprises convolutional layer for the extraction of the features, standardized by activation functions, namely rectified linear unit (ReLU), followed by a primary capsule layer with numerous capsules [18]. One additional advantage of CapsNet is the overview of the digit capsule layer that comes next to the primary capsule layer and performs two functions: recognizing the existence of a particular feature and its possessions. Firstly, the feature vector of attention is connected to the convolution layers using dimension $1xd$, whereas d characterizes the lower dimension. Post convolution is dimensioned yxd, caused by y convolutional filters. These layers transform the matrix to the kxz size, by all of the k primary capsules. Subsequently, digit capsules improve these elementary features into more subtle descriptions of particular states or errors with the help of dynamic routing, dimensioned nxm, where n refers to the number of classes and m signifies the output length of a vector.

2.3. Classification process via the deep belief network model

After that, the DBN method is mainly designed the classification process. Deep neural network (DNN) derives from artificial neural networks [19]. DBN is stacked frequently by restricted Boltzmann machine (RBM); RBM is a type of distinctive NN. It is comprised through the connection of the hidden layer (HL) and visible layer (VL), but there is no connection among VL, VL HL, and HL. In training, the distinctive value of VL maps to HL; VL is recreated over the HL. This novel layer's visible characteristic value maps to HL again, afterward capturing an innovative HL. The major objective is to attain the generative value of power. Therefore, the major features of RBM are the activation characteristics of the layer input to another layer as training data; thus, the learning speed is quicker. This is a layer-by-layer and effective learning approach theory.

Utilizing Bernoulli-Bernoulli RBM and Gauss-Bernoulli RBM to interconnect, the outcome of the lower layer is the input to the upper layer. In Bernoulli RBM, HL and VL units are dual: $h \in \{0,1\}^K$, $V \in \{0,1\}^D$ K, and D existing unit numbers of HL and VL. In Gaussian RBM, a unit of VL is a real number: $y \in R^D$. The h and V joint possibility is given.

$$P(v,h) = \frac{1}{Z} exp\big(-E\ (v,h)\big) \tag{1}$$

$E(v,h)$ specifies an energy equation, and Z denotes a normalized constant.

$$E(v,h) = -\sum_{i=1j}^{D} \quad \sum_{=1}^{K} \quad w_{ij} v_i h_j - \sum_{i=1}^{D} \quad b_i v_i - \sum_{j=1}^{K} \quad a_j h_j \tag{2}$$

w_{ij} indicates the quality of the un-connected node of VL v_i, and HL nodes h_j, and b and a are implicit biases of hidden and visible units.

$$E(v,h) = \sum_{i=1}^{D} \quad \frac{(v_i - b_i)^2}{2} - \sum_{i=1j}^{D} \quad \sum_{=1}^{K} \quad w_{ij} v_i h_j - \sum_{j=1}^{K} \quad a_j h_j \tag{3}$$

2.4. Parameter tuning using crayfish optimization algorithm

To further enhance method outcome, the COA is leveraged for parameter tuning to assure that best hyperparameters are chosen for enhanced accuracy. COA imitates the crayfish behaviors like searching, competition, and summer resort behaviors [20]. These behaviors are associated with the exploration and exploitation stages of optimization, depending on temperature. High temperatures make crayfish search for shelter for competition or rest, but proper temperatures encourage foraging in exploration. Temperature changes give unpredictability in discovering the best solutions. The important steps of COA are demonstrated:

Initialization: during COA, an optimizer issues by sizes is characterized by every crayfish that acts as the possible outcome in structure of a 1x d vector. All variables $(X_1, X_2, X_3, and\ X_d)$ are characterized within the searching region that is subject to the lower and upper limits. The solutions are compared with a stepwise model, and the most promising option is established and preserved as the final best solution. The COA population's initial distribution is determined by (4):

$$X_i = Lb + (Ub + Lb) \times rand \tag{4}$$

Whereas Ub and Lb characterize the optimizer obstacles' limits. $Temp$ is an important feature in numerous phases of the crayfish and was described by (5). If the $temp$ surpasses 30 degrees, they move to the coolest place as their summer refuge. The $temp$ is formulated (5):

$$temp = rand \times 15 + 20 \tag{5}$$

Then, the performance of looking for a meal is simulated by leveraging a Gaussian distribution that is impacted by the temperature as defined in (6):

$$p = C_1 \times \left(\frac{1}{\sqrt{2\times\pi}\times\sigma} \times exp\right)\left(\frac{(temp-\mu)^2}{2\sigma^2}\right) \tag{6}$$

In the COA, the fitness selection is a substantial part influencing the outcome. The process of parameter range holds the solution-encoding system by assessing the efficacy of the applicant outcomes. The COA imitates precision as the chief standard by projecting fitness function.

$$Fitness\ =\ max\ (P) \tag{7}$$

$$P = \frac{TP}{TP+FP} \tag{8}$$

Here, TP and FP signify the true and false positive.

Table 1 depicts hyperparameter tuning, whereas the Algorithm 1 provides the step-by-step workflow of the proposed FTLMCOA-VCITS method employs DenseNet-201, EfficientNet-B7, and CapsNet. The DBN classifier with 3 HL and 256 neurons is optimized utilizing the COA method (population size=50, 100 iterations), whereas training uses the AdamW optimizer, 0.001 learning rate, batch size of 32, and 50 epochs.

Table 1. Hyperparameter tuning configuration of the proposed FTLMCOA-VCITS method

Module	Hyperparameter	Optimal value
DenseNet-201	Growth rate	32
	Feature dimension	1024
EfficientNet-B7	Input image size	600×600
	Dropout rate	0.3
CapsNet	Capsule dimension	16
	Routing iterations	3
Future fusion	Fusion dimension	512
	Fusion strategy	Concatenation
DBN	Hidden layers	3
	Hidden neurons	256
	Learning rate	0.001
COA	Population size	50
	Maximum iterations	100
	Exploration factor	0.6
Training configuration	Optimizer	AdamW
	Batch size	32
	Epochs	50
	Weight decay	1×10^{-4}

Algorithm 1. FTLMCOA-VCITS vehicle classification
Input: Vehicle image dataset
Output: Predicted vehicle class
Apply Wiener Filtering (WF) to the input images for noise removal.
The DenseNet-201, EfficientNet-B7, and CapsNet, extract deep features.

Initialize the DBN classifier with the identified hyperparameters.
Initialize the COA population with a population size of 50 and maximum iterations of 100.
Evaluate the fitness of each candidate solution utilizing the classification results.
Update the candidate solutions through the COA exploration, competition, and foraging mechanisms.
Select the optimum hyperparameter configuration depending on the best fitness value.
Train the DBN model utilizing the optimal parameters and AdamW optimizer.
Detect the input vehicle image and create the final vehicle class.

3. EXPERIMENTAL ANALYSIS

The experimental analysis of FTLMCOA-VCITS system is inspected using the Vehicle Classification dataset [21]. This dataset contains 1,200 total images under 10 types of vehicles. The proposed FTLMCOA-VCITS method was implemented in Python 3.11 through TensorFlow/Keras on a Windows 11 systems with an NVIDIA GPU. The Vehicle Classification dataset was assessed utilizing both 80:20 and 70:30 train-test splits, and the training and testing sets were strictly separated to prevent data leakage. Figure 2 represent the vehicle detection of the FTLMCOA-VCITS method under 80:20. Depending on 80% TRPH, the FTLMCOA-VCITS approach obtains an average accuracy, precision, recall, F1-score, and area under the curve (AUC) score of 98.63%, 93.21%, 93.09%, 93.11%, and 96.16%, respectively. Likewise, by 20% TSPH, the FTLMCOA-VCITS approach secures an average accuracy, precision, recall, F1-score, and AUC score of 99.08%, 95.66%, 95.28%, 95.38%, and 97.38% correspondingly.

Figure 3 exemplify vehicle detection of the FTLMCOA-VCITS method with 70:30 of training phase (TRPH)/testing phase (TSPH). Based on 70% TRPH, the projected FTLMCOA-VCITS system achieves average accuracy, precision, recall, F1-score, and AUC score of 98.45%, 92.40%, 92.25%, 92.24%, 95.69% correspondingly. Furthermore, by 30% TRPH, the anticipated FTLMCOA-VCITS approach gains an average accuracy, precision, recall, F1-score, and AUC score of 98.50%, 92.49%, 92.70%, 92.46%, and 95.93%.

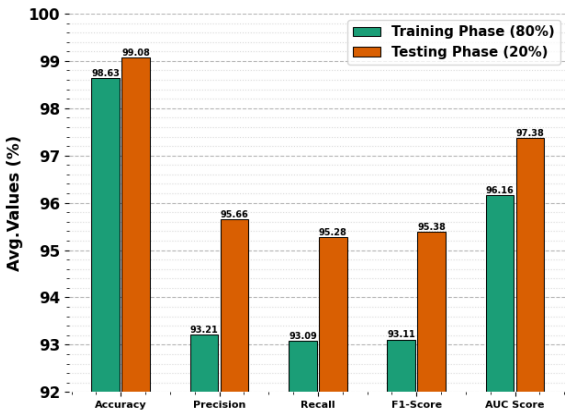


Figure 2. Average of FTLMCOA-VCITS system under 80:20

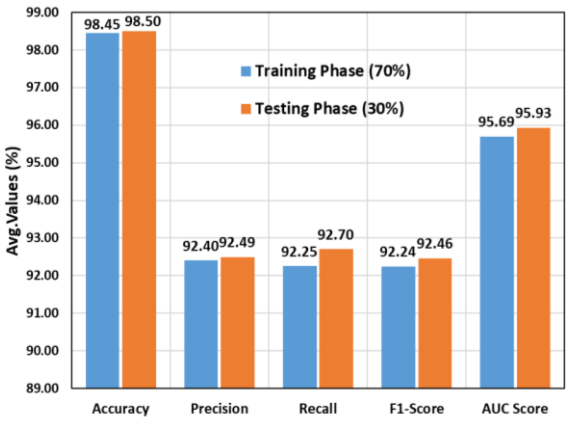


Figure 3. Average of FTLMCOA-VCITS approach under 70:30

Table 2 depict the comparison study of FTLMCOA-VCITS approach by the current method [22]–[25]. The results highlighted that the spotted hyena optimizer with DL-enabled vehicle counting and classification (SHODL-VCC), Faster R-CNN, YOLOv3, YOLOv4, detection with enriched semantics-machine learning (DES-ML), semi-supervised fuzzy c-means (SSFCM), and CenterNet2 models have shown the worst performance. But the FTLMCOA-VCITS method conveyed enriched performance with maximum accuracy, precision, recall, and F1-score of 99.08%, 95.66%, 95.28%, and 95.38%, correspondingly.

An ablation study is performed to understand the significance of multiple elements in the system. It evaluates performance changes when specific parts are removed. Table 3 illustrates the ablation results of the FTLMCOA-VCITS model. The outcome implied that the FTLMCOA-VCITS (WF+DenseNet-201+EfficientNet-B7+CapsNet+DBN+COA) approach has got higher performance with accuracy of 99.08%, precision of 95.66%, recall of 95.28%, and F1-score of 95.38%, whereas the feature fusion+DBN (without COA tuning), DenseNet-201 only+DBN, EfficientNet-B7 only+DBN, CapsNet only+DBN, DenseNet-201+EfficientNet-B7, and fusion features (without DBN) approaches have gained lower performance with various metrics.

Table 2. Comparative analysis of FTLMCOA-VCITS model with recent approaches

Methods	Accuracy	Precision	Recall	F1-score
FTLMCOA-VCITS	99.08	95.66	95.28	95.38
SHODL-VCC	98.60	94.71	90.63	93.74
Faster R-CNN	96.94	91.36	90.34	93.14
YOLOv3	96.00	94.71	93.97	90.46
YOLOv4	97.74	94.82	94.15	92.57
DES-ML	97.61	93.73	93.70	90.55
SSFCM	97.05	92.78	90.69	93.39
CenterNet2	97.07	91.36	91.73	94.69

Table 3. Ablation study of the FTLMCOA-VCITS method

Methods	Accuracy	Precision	Recall	F1-score
FTLMCOA-VCITS (WF+DenseNet-201+EfficientNet-B7+CapsNet+DBN+COA)	99.08	95.66	95.28	95.38
Feature fusion+DBN (without COA tuning)	98.28	95.07	94.55	94.61
DenseNet-201 only+DBN	97.74	94.35	93.94	94.05
EfficientNet-B7 only+DBN	97.22	93.85	93.16	93.41
CapsNet only+DBN	96.46	93.34	92.60	92.84
DenseNet-201+EfficientNet-B7	95.72	92.59	91.83	92.05
Fusion features (without DBN)	95.07	92.01	91.05	91.43

Table 4 demonstrate the K-fold cross validation results of the FTLMCOA-VCITS model in terms of several metrics. On fold-1, the proposed FTLMCOA-VCITS model attained accuracy of 91.56%, precision of 92.33%, recall of 88.98%, and F1-score of 86.99%. Furthermore, on fold-3, the FTLMCOA-VCITS model showed accuracy of 97.89%, precision of 89.19%, recall of 85.28%, and F1-score of 91.77%. Additionally, on fold-7, the proposed FTLMCOA-VCITS method reaches accuracy of 93.99%, precision of 86.08%, recall of 86.63%, and F1-score of 91.61%. Finally, on fold-10, the proposed MFF-DGERDP approach gained accuracy of 97.96%, precision of 94.84%, recall of 94.13%, and F1-score of 94.22%, respectively.

Table 4. K-fold cross validation of the FTLMCOA-VCITS technique

K-fold cross validation	Accuracy	Precision	Recall	F1-score
Fold-1	91.56	92.33	88.98	86.99
Fold-2	95.04	94.07	93.69	85.35
Fold-3	97.89	89.19	85.28	91.77
Fold-4	93.88	93.69	89.31	91.06
Fold-5	87.49	90.18	86.21	88.76
Fold-6	91.95	91.87	85.73	87.93
Fold-7	93.99	86.08	86.63	91.61
Fold-8	89.45	90.11	90.87	88.45
Fold-9	91.93	85.45	89.80	92.91
Fold-10	97.96	94.84	94.13	94.22

The experimental outcomes demonstrate that the proposed methodology performs efficiently in vehicle classification under diverse train-test splits. The model attained high accuracy, precision, recall, and F1-score, which represents its strong ability in correctly identifying different vehicle classes. The improved performance is mainly because of the combination of WF for noise removal, hybrid feature extraction using DenseNet-201, EfficientNet-B7, and CapsNet, along with DBN-based classification and COA-based parameter tuning. In addition, the proposed model outperformed existing approaches with lower computational time, proving its suitability for real-time intelligent transportation applications. The ablation analysis also verifies that every element of the proposed model contributes to improving the overall classification performance, where removing any module leads to a noticeable drop in accuracy and overall efficiency.

4. CONCLUSION

This paper introduced an FTLMCOA-VCITS model. The primary aim of the proposed approach relies on enhancing vehicle classification methods in transportation systems utilizing innovative DL models. In the initial stage, the image preprocessing applies WF for removing the noise. For the feature representation process, the fusion of the DenseNet-201, EfficientNet-B7, and CapsNet is employed. Followed by, the DBN method is mainly leveraged for the classification process. Moreover, COA is used to optimize hyperparameters for improving model accuracy. The comprehensive experimental analysis of the

FTLMCOA-VCITS approach portrayed a greater accuracy value of 99.08% across current methods in the vehicle classification. Moreover, the comparatively limited dataset may impact the generalizability of the method to larger and more different real-world traffic settings. Further validation on different dataset in therefore need to confirm its robustness. In future, the proposed approach can be extended by testing it on higher and distinct real-world traffic datasets. The model can also be enhanced to handle challenging conditions like poor lighting, weather variations, and occluded vehicles. Additionally, lightweight DL framework can be explored to decrease computation complexity for real-time deployment. Further research may also concentrate on integrating the model with live traffic monitoring systems for practical intelligent transportation applications.

FUNDING INFORMATION

The authors received no funding for this study.

AUTHOR CONTRIBUTIONS STATEMENT

This journal uses the Contributor Roles Taxonomy (CRediT) to recognize individual author contributions, reduce authorship disputes, and facilitate collaboration.

Name of Author	C	M	So	Va	Fo	I	R	D	O	E	Vi	Su	P	Fu
Abel Jopaul Vettilkattil Paulose	✓	✓	✓						✓					
Prabhu Arthanareeswaran			✓	✓	✓				✓					
Mohammad Shabbir Alam		✓			✓	✓			✓	✓		✓		
Barkathulla Asan Mohamed	✓				✓			✓	✓			✓		
Anjaiah Talamala	✓				✓					✓	✓	✓		
Manoj Senthil Kailasam		✓				✓				✓			✓	✓

C : **C**onceptualization
M : **M**ethodology
So : **So**ftware
Va : **Va**lidation
Fo : **Fo**rmal analysis
I : **I**nvestigation
R : **R**esources
D : **D**ata Curation
O : Writing - **O**riginal Draft
E : Writing - Review & **E**diting
Vi : **Vi**sualization
Su : **Su**pervision
P : **P**roject administration
Fu : **Fu**nding acquisition

CONFLICT OF INTEREST STATEMENT

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this manuscript. All authors contributed to the manuscript and approved the final version. The authors did not receive support from any organization for the submitted work.

DATA AVAILABILITY

The data that support the findings of this study are openly available in the Kaggle repository at https://www.kaggle.com/datasets/marquis03/vehicle-classification, reference number [21].

REFERENCES

[1] V. Kocur and M. Ftacnik, "Multi-class multi-movement vehicle counting based on centertrack," in *IEEE Computer Society Conference on Computer Vision and Pattern Recognition Workshops*, IEEE, 2021, pp. 4004–4010, doi: 10.1109/CVPRW53098.2021.00452.

[2] A. Gholamhosseinian and J. Seitz, "Vehicle classification in intelligent transport systems: an overview, methods and software perspective," *IEEE Open Journal of Intelligent Transportation Systems*, vol. 2, pp. 173–194, 2021, doi: 10.1109/OJITS.2021.3096756.

[3] P. Arthurs, L. Gillam, P. Krause, N. Wang, K. Halder, and A. Mouzakitis, "A taxonomy and survey of edge cloud computing for intelligent transportation systems and connected vehicles," *IEEE Transactions on Intelligent Transportation Systems*, vol. 23, no. 7, pp. 6206–6221, Jul. 2022, doi: 10.1109/TITS.2021.3084396.

[4] H. Zhang and X. Lu, "Vehicle communication network in intelligent transportation system based on internet of things," *Computer Communications*, vol. 160, pp. 799–806, Jul. 2020, doi: 10.1016/j.comcom.2020.03.041.
[5] A. Khang, V. Abdullayev, and Y. Niu, "Analysis of wireless sensor networks applications in intelligent transportation system," in *Lecture Notes in Intelligent Transportation and Infrastructure*, Springer Nature Switzerland, 2025, pp. 359–377, doi: 10.1007/978-3-031-72617-0_19.
[6] M. Wazid *et al.*, "Explainable deep learning-enabled malware attack detection for iot-enabled intelligent transportation systems," *IEEE Transactions on Intelligent Transportation Systems*, vol. 26, no. 5, pp. 7231–7244, May 2025, doi: 10.1109/TITS.2025.3525505.
[7] A. K. Singh *et al.*, "Intelligent transportation system for automated medical services during pandemic," *Future Generation Computer Systems*, vol. 163, p. 107515, Feb. 2025, doi: 10.1016/j.future.2024.107515.
[8] S. M. Nagarajan, G. G. Devarajan, T. V. Ramana, J. M. Asha, A. K. Bashir, and Y. D. Al-Otaibi, "Adversarial deep learning based dampster–shafer data fusion model for intelligent transportation system," *Information Fusion*, vol. 102, p. 102050, Feb. 2024, doi: 10.1016/j.inffus.2023.102050.
[9] P. Rani and R. Sharma, "Intelligent transportation system performance analysis of indoor and outdoor internet of vehicle (iov) applications towards 5g," *Tsinghua Science and Technology*, vol. 29, no. 6, pp. 1785–1795, Dec. 2024, doi: 10.26599/TST.2023.9010119.
[10] M. Elassy, M. Al-Hattab, M. Takruri, and S. Badawi, "Intelligent transportation systems for sustainable smart cities," *Transportation Engineering*, vol. 16, p. 100252, Jun. 2024, doi: 10.1016/j.treng.2024.100252.
[11] S. Zhang *et al.*, "Federated learning in intelligent transportation systems: recent applications and open problems," *IEEE Transactions on Intelligent Transportation Systems*, vol. 25, no. 5, pp. 3259–3285, May 2024, doi: 10.1109/TITS.2023.3324962.
[12] M. Waqar, M. Ishaq, M. H. Afzal, and S. Iqbal, "Vehicle detection using artificial intelligence based algorithm in intelligent transportation systems," in *2023 4th International Conference on Computing, Mathematics and Engineering Technologies (iCoMET)*, IEEE, Mar. 2023, pp. 1–6, doi: 10.1109/iCoMET57998.2023.10099381.
[13] Q. Yuan *et al.*, "Ensemble artificial intelligence framework for real-time automatic vehicle detection and classifications," *IEEE Access*, vol. 14, pp. 16034–16050, 2026, doi: 10.1109/ACCESS.2026.3657829.
[14] I. A. Ashari, W. A. Syafei, and A. Wibowo, "Enhancement of YOLOv9 model for traffic vehicle detection using augmentation techniques," *Jurnal Teknik Informatika (JUTIF)*, vol. 7, no. 2, pp. 811–822, Apr. 2026, doi: 10.52436/1.jutif.2026.7.2.5196.
[15] M. M. Hasan *et al.*, "BDNet: a lightweight yolov12-based vehicle detection framework for smart urban traffic monitoring," *Smart Cities*, vol. 9, no. 2, p. 33, Feb. 2026, doi: 10.3390/smartcities9020033.
[16] D. K. Saha *et al.*, "Mpox-xde: an ensemble model utilizing deep CNN and explainable AI for monkeypox detection and classification," *BMC Infectious Diseases*, vol. 25, no. 1, p. 403, Mar. 2025, doi: 10.1186/s12879-025-10811-y.
[17] N. Xiao, D. Kong, and J. Wang, "Ultrasound thyroid nodule segmentation algorithm based on deeplabv3+ with efficientnet," *Journal of Imaging Informatics in Medicine*, vol. 38, no. 6, pp. 3614–3622, Feb. 2025, doi: 10.1007/s10278-025-01436-3.
[18] N. Rezazadeh, F. Annaz, W. A. Jabbar, J. V. Filho, and M. D. Oliveira, "A transfer learning approach for mitigating temperature effects on wind turbine blades damage diagnosis," *Structural Health Monitoring*, vol. 25, no. 3, pp. 1761–1783, May 2026, doi: 10.1177/14759217241313350.
[19] C. Huang, W. Gong, W. Fu, and D. Feng, "A research of speech emotion recognition based on deep belief network and SVM," *Mathematical Problems in Engineering*, no. 1, Jan. 2014, doi: 10.1155/2014/749604.
[20] M. Elhosseny, M. Abdel-Salam, and I. M. El-Hasnony, "Adaptive dynamic crayfish algorithm with multi-enhanced strategy for global high-dimensional optimization and real-engineering problems," *Scientific Reports*, vol. 15, no. 1, Mar. 2025, doi: 10.1038/s41598-024-81144-0.
[21] Marquis, "Vehicle classification," 2023, [Online]. Available: https://www.kaggle.com/datasets/marquis03/vehicle-classification, (Accessed on October 16, 2026).
[22] M. A. Alohali *et al.*, "Spotted hyena optimizer with deep learning enabled vehicle counting and classification model for intelligent transportation systems," *Electronic Research Archive*, vol. 31, no. 7, pp. 3704–3721, 2023, doi: 10.3934/ERA.2023188.
[23] N. Niroomand, C. Bach, and M. Elser, "Robust vehicle classification based on deep features learning," *IEEE Access*, vol. 9, pp. 95675–95685, 2021, doi: 10.1109/access.2021.3094366.
[24] P. Gao, T. Tian, T. Zhao, L. Li, N. Zhang, and J. Tian, "Double fcos: a two-stage model utilizing fcos for vehicle detection in various remote sensing scenes," *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, vol. 15, pp. 4730–4743, 2022, doi: 10.1109/JSTARS.2022.3181594.
[25] H. Song, H. Liang, H. Li, Z. Dai, X. Yun, "Vision-based vehicle detection and counting system using deep learning in highway scenes," *European Transport Research Review*, vol. 11, pp. 1-6, 2019, doi: 10.1186/s12544-019-0390-4.

BIOGRAPHIES OF AUTHORS



Abel Jopaul Vettilkattil Paulose is an academic and researcher in the field of computer applications, affiliated with LEAD College (Autonomous), Palakkad, Kerala, India. His research and teaching interests include computer science, information technology, and emerging computational applications. He can be contacted at email: abel.jp@lead.ac.in.



Dr. Prabhu Arthanareeswaran is an associate professor at the School of Computer Science and Information Technology, Jain (Deemed-to-be University), Bengaluru, Karnataka, India. His academic interests encompass computer science, information technology, and interdisciplinary research. He can be contacted at email: prabhu.a@jainuniversity.ac.in.



Mohammad Shabbir Alam is a faculty member in the Department of Computer Science, College of Engineering and Computer Science, Jazan University, Jizan, Saudi Arabia. His research interests include computer science, computing technologies, and related emerging areas. He can be contacted at email: amushabbir@gmail.com.



Barkathulla Asan Mohamed is an assistant professor (senior grade) in the Department of Computer Science and Engineering at Vel Tech Rangarajan Dr. Sagunthala R&D Institute of Science and Technology, Chennai, India. His academic and research interests focus on computer science and engineering, emerging technologies, and applied computing. He can be contacted at email: barkathullakng@gmail.com.



Anjaiah Talamala is associated with Aditya University, Surampalem, Andhra Pradesh, India. His academic and research interests are centered on computer science, information technology, and contemporary technological applications. He can be contacted at email: anjibabu317@gmail.com.



Manoj Senthil Kailasam is an associate professor in the Department of Electronics and Communication Engineering at Kongu Engineering College, Perundurai, Erode, Tamil Nadu, India. His academic and research interests include electronics and communication engineering, computing applications, and emerging technologies. He can be contacted at email: kmanojsenthil@gmail.com.