Proceedings of the 8th International Conference on Inventive Material Science and Applications (ICIMA-2026)
IEEE Xplore Part Number: CFP264B2-ART; ISBN: 979-8-3315-7814-5

Aquila Optimization Driven Deep Learning Framework for Emergency Vehicle Alerting and Recognition in Dense Indian Traffic

K. Sasikala
Department of Electrical and Electronics Engineering,
Vel's Institute of Science, Technology & Advanced Studies,
Chennai, Tamil Nadu, India.
skala.se@vistas.ac.in

Kiran Kumar R
Department of Electrical and Electronics Engineering,
Vel's Institute of Science, Technology & Advanced Studies,
Chennai, Tamil Nadu, India.
kumarkiran0787@gmail.com

Abishek A
Department of Electrical and Electronics Engineering,
Vel's Institute of Science, Technology & Advanced Studies,
Chennai, Tamil Nadu, India.
abishekabishek15966@gmail.com

Sanjay Ragul C
Department of Electrical and Electronics Engineering,
Vel's Institute of Science, Technology & Advanced Studies,
Chennai, Tamil Nadu, India.
novah0014@gmail.com

***Abstract*— Emergency Vehicles (EV) in congested traffic is important for assuring rapid response and reducing important interruptions an effective appreciation and prioritizing. For accurate and real-time identification the structure uses image processing and Deep Learning (DL) algorithms. The raw input data is fed to pre-process with an Adaptive Bilateral Filter (ABF), which efficiently reduces noise whereas protecting critical edge features. The improved images are then fed to Fuzzy Cluster-Based Thresholding (FCBT) for precise segmentation of EV in contrast to difficult traffic backgrounds. To demonstrate discriminative visual patterns, Discrete Wavelet Transformation (DWT) is utilized as a multi-level feature extraction. The extracted features are then recognized using a unique Deep Attention Dilated Residual Aquila Convolutional Neural Network (DADR-AquilaNet), which combines dilated convolutions, residual learning, and attention processes to improve feature sensitivity as well as contextual awareness. Further to fine tune DADR, Aquila Optimization Algorithm (AOA) is to regulate the network limits, resulting in better convergence and classification resilience. An experimental evaluation is implemented by Indian Emergency Vehicles Dataset displayed an enhanced detection accuracy of 98% and occlusion situations. The structure offers an intelligent and scalable solution for real-time EV alerts in congested urban traffic.**

Keywords*— *Emergency Vehicles, Adaptive Bilateral Filter, Fuzzy Cluster-Based Thresholding, Discrete Wavelet Transformation, Deep Attention Dilated Residual Aquila Convolutional Neural Network, Aquila Optimization Algorithm

I. Introduction

An EV forms a measure in confirming public safety and must be given maximum importance and space to traverse over traffic professionally [1]. But, traffic congestion remains one of the most pervasive global challenges [2]. Consequently, traffic encompasses pedestrians, vehicles, bicycles, livestock, trains, and other transport modes sharing the roadway network [3-4]. Furthermore, the rapid population growths in urban regions have led to a significant rise in the number of vehicles, thereby intensifying traffic congestion and related issues [5]. However, with ongoing economic expansion, highway traffic volumes are increasing daily, resulting in more severe congestion, particularly during peak hours and holidays [6-7]. Moreover, persistent traffic congestion not only disrupts daily life but also leads to considerable economic losses, increased air pollution, and elevated fuel consumption, all of which negatively affect living standards and productivity [8]. Moreover, traffic congestion critically delays the rapid response of emergency services, including ambulances and fire trucks, compromising their ability to reach emergencies in time [9-10]. To overcome this issue some of the following techniques are used.

In [11] Zohaib *et al* (2024) have proposed attention-based temporal spectrum network (ATSN) designed for ambulance siren sound detection. ATSN effectively captures both temporal variations (siren patterns over time) and spectral characteristics. Nevertheless, requires a diverse of siren and non-siren sounds for effective training. Consequently, Louati *et al* (2021) [12] have presented case-based reasoning algorithm (CBRA) exploited to react to detected and forecasted the traffic. CBRA quickly retrieve and reuse past solutions, allowing timely response to congestion or incident detection. Nonetheless, as the number of cases grows, memory usage and retrieval time increase, potentially limiting scalability for large-scale urban traffic network. Whereas, in [13] Gupta et *al* (2023) have developed EfficientDet and Tensor Flow lite for the concentration of vehicles on the road and the assigned priority levels for specific EV. Both vehicle density and EV presence supports adaptive signal control for faster EV navigation through congested routes. Nevertheless, quantization and pruning for TensorFlow Lite require expertise to maintain accuracy although reducing model size. Additionally, [14] Parada *et al* (2025) have invented Multi-Agent Proximal Policy Optimisation (MAPPO) for traffic efficiency and safety. MAPPO reduces collision risks and supports safe maneuvers for EV and pedestrians in traffic situations. Yet, it requires significant computational resources for training, especially in high-dimensional traffic networks with many agents. Moreover in [15] Hosseinzadeh *et al* (2022) have introduced multi-objective genetic algorithm (MOGA) for EV detection and classification. MOGA effectively handles high-dimensional, nonlinear, and noisy datasets typical of traffic images and acoustic signals. Nevertheless, MOGA reduces reproducibility without sufficient control mechanisms.

A. Research contribution

- Indian Emergency Vehicles Dataset is fed to pre-processing based ABF, it effectively reduce noise whereas preserving essential edge details, ensuring superior image quality for subsequent analysis.
- FCBT based segmentation is used to separate EV accurately, even in difficult, cluttered, as well as dynamically varying traffic divisions.
- Feature extraction based on DWT is utilized to seizure both spatial as well as frequency-domain features through multiple resolutions, refining the discriminative power of visual representations.
- Proposed a new DADR-AquilaNet design incorporating dilated convolutions, residual learning, and attention mechanisms to develop feature sensitivity, contextual understanding, and classification accuracy.
- AOA is suggested for parameter tuning, allowing faster convergence, better generalization, and strong performance under varied illumination and occlusion conditions.

B. Research Layout:

Introduction and literature survey presents in section I. Section II describes proposed system. Discusses the experimental results, insights, and potential challenges associated with the proposed method in section III. Lastly, section IV completes the article.

II. Proposed Work

Proposed architecture is designed for the efficient recognition and prioritisation of EV in dense urban traffic conditions as shown in figure 1. In the pre-processing stage the raw Indian Emergency Vehicles Dataset is fed as an input. At first, a raw traffic image is fed to ABF to destroy noise whereas preserving edge and texture parts for precise object identification. The processed enhanced images are then fed through FCBT, which adaptively segments EV from complex and vibrant traffic environments. Subsequent segmentation, DWT is suitable for multi-level feature extraction, seizing both spatial as well as frequency-domain characteristics crucial for differentiating EV based on visual patterns. These features are subsequently classified using a novel DADR-AquilaNet, which integrates attention mechanisms, dilated convolutions, and residual learning to enhance the method. To further improve performance, the AOA is employed for optimal hyper parameter tuning, ensuring faster convergence, reduced over fitting, and improved detection accuracy.

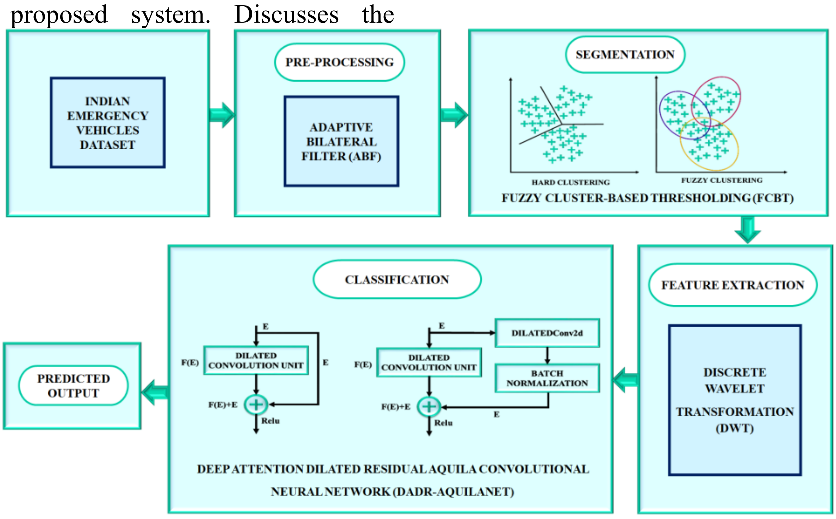

Fig. 1. Block diagram for proposed work

A. Dataset

In this paper a Indian Emergency Vehicles Dataset, comprising images of ambulances, fire-trucks and police vehicles operating under dense Indian urban traffic conditions. The data were annotated with bounding boxes and class labels for emergency vehicles vs non-emergency vehicles, under varied illumination, occlusion and traffic density.

B. Pre-processing using Adaptive Bilateral Filter

In dense urban traffic scenes, the captured images often suffer from noise, illumination variation, and background clutter. For this an ABF is used, which preserves edges whereas suppressing noise. Initially, the raw traffic images undergo enhancement using an ABF, which serves to effectively reduce noise whereas maintaining important structural information within the image. This balance between noise suppression and edge retention is essential for ensuring that fine details such as the contours and textures of vehicles. For a pixel at location p, the output of the ABF is:

$$I_{ABF}(p) = \frac{1}{W_p} \sum_{q\in\Omega} I(q).f_s(\|p-q\|).f_r\big(I(p) - I(q)\big).\alpha(p,q) \quad (1)$$

$$f_s(\|p-q\|) = exp\,\left(-\frac{\|p-q\|^2}{2\sigma_s^2}\right) \quad (2)$$

$$f_r\big(I(p)-I(q)\big) = exp\;\frac{\big(I(p)-I(q)\big)^2}{2\sigma_r^2} \quad (3)$$

Where, $I(p)$ denotes intensity value at pixel p, $q \in \Omega$ is the neighbouring pixels nearby p, $f_s(\|p-q\|)$ denotes spatial Gaussian kernel, $f_r\big(I(p)-I(q)\big)$ represents range Gaussian kernel, $\alpha(p,q)$ is the adaptive weight function that modifies filtering strength based on local contrast or texture. W_p indicates normalization factor to ensure weights sum to 1. Pre-processed images provide a cleaner and more detailed input for subsequent stages, finally improving the accuracy of traffic scene analysis. The processed image is fed to segmentation stage.

C. Segmentation using Fuzzy Cluster-Based Thresholding

The FCBT technique is employed to achieve accurate segmentation of EV from complex and dynamic traffic backgrounds. This approach, each pixel in the image is evaluated based on its intensity and spatial features, and a fuzzy membership function determines the likelihood of that pixel belonging to a specific cluster such as the vehicle region or the background as in fig. 2.

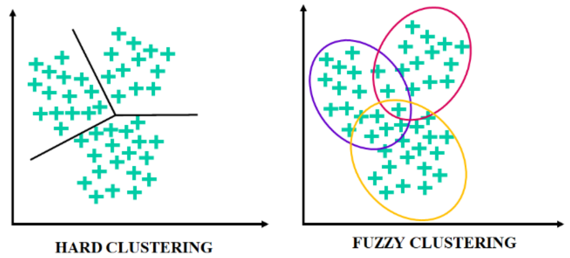

Fig. 2. FCBT method

To reduce the intra-cluster variance the method iteratively refines these memberships whereas exploiting inter-cluster separation. This procedure generates an adaptive threshold that effectively distinguishes EV even in challenging traffic scenes with complex textures or illumination variations. The FCBT can be expressed as:

$$J = \sum_{i=1}^{c}\sum_{j=1}^{N}\mu_{ij}^{m}\left\|x_j - v_i\right\|^2 \tag{4}$$

Here, c indicates number of clusters, N represents number of pixels, x_j feature vector of the j^{th} pixel, m designates fuzziness parameter. The segmented image for EV is clearly insulated from their surroundings, forming a reliable foundation for subsequent feature extraction and recognition stages in intelligent traffic monitoring systems.

D. Discrete Wavelet Transformation based Feature extraction

To perform multi-level feature extraction the DWT is useful in order to effectively capture discriminative visual patterns of EV within complex traffic scenes. For identifying subtle variations in texture, edges, and shapes that characterize EV it is a useful one. In this process, the traffic image is decomposed into four sub-bands at each level:

Approximation (LL): The low-frequency components representing the coarse structure or general appearance of the vehicle are captured.

Horizontal (LH), Vertical (HL), as well as Diagonal (HH): The high-frequency components corresponding to edge and texture details in different orientations are captured.

The 2D DWT for an image $f(x, y)$ is expressed as:

$$DWT(j,k) = \sum_x\sum_y f(x,y)\psi_{j,k}(x,y) \tag{5}$$

Where, $\psi_{j,k}(x,y)$ signifies wavelet basis function then $f(x,y)$ represents the pixel intensity values of the input image. Next, extracted image is fed to the proposed DADR-AquilaNet classification stage.

E. Deep Attention Dilated Residual Aquila Convolutional Neural Network Based Classification

DADR-AquilaNet is is proposed as a specialized DL architecture designed to improve the detection as well as recognition of EV in compound and active traffic places.

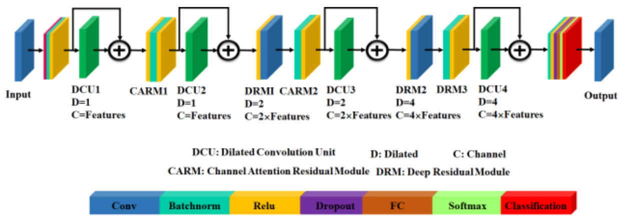

Fig. 3. Deep Attention Dilated Residual Convolutional Neural Network

Figure 3 show the model individually integrates dilated convolutions, residual learning, and attention mechanisms to achieve high feature sensitivity and strong contextual awareness two serious aspects for accurately identifying EVs amid diverse background conditions.

i) Dilated Convolutions: Through inserting gaps among kernel elements, DADR-AquilaNet seizure multi-scale contextual information, allowing it to recognize EV from varying distances and perspectives, even when they are partially occluded or small in size within the frame.

ii) Residual Learning: the residual blocks are incorporated to mitigate the vanishing gradient problem and enable deeper network training. The residual links permit the network to absorb individuality mappings alongside complex transformations, ensuring that essential low-level features such as color, shape, and edge cues of EVs are effectively retained and propagated through deeper layers.

iii) Attention Mechanisms: The attention module supports the model focus selectively on salient regions of the image such as flashing lights, reflective surfaces, or vehicle markings whereas suppressing irrelevant background information like other vehicles, road signs, or shadows. Both spatial attention and channel attention are employed to dynamically refine feature maps and improve the network's discriminative capability.

Precisely, the dilated convolution is signified as:

$$y[i] = \sum_{k=1}^{K} x[i + r.k].w[k] \tag{6}$$

Here $y[i]$ indicates output feature, $x[i]$ signifies input signal, $w[k]$ represents convolutional kernel, and r indicates dilation rate controlling spacing between kernel elements.

F. Aquila Optimization Algorithm

The AOA is employed to fine-tune as well as optimize the parameters of DADR-AquilaNet, ensuring faster convergence, improved stability, and higher classification accuracy for EV detection. Stimulated by the intelligent chasing strategies of the Aquila (eagle), the AOA mimics its dynamic transition between search as well as attacking prey to achieve an optimal balance between global and local search during network training.

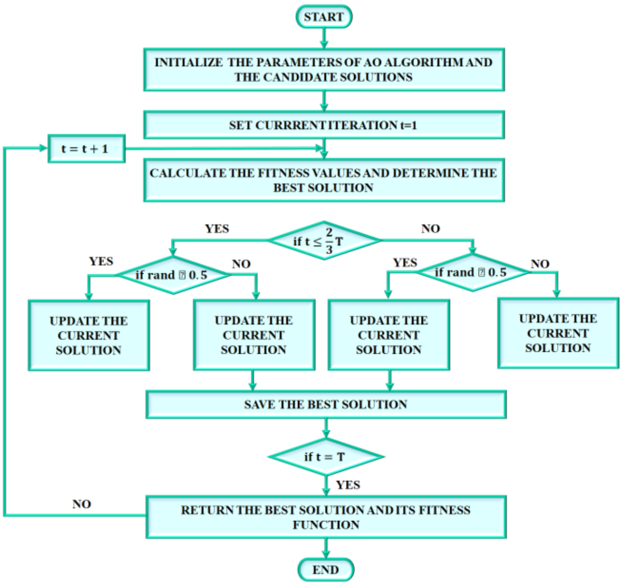

Fig. 4. Flowchart for AOA

The AOA flow chart is shown in figure 4 it is used to optimize the weights and biases, minimizing the classification loss whereas enhancing the network's ability to distinguish EV from other vehicles in varying traffic and lighting conditions.

Method I: Extended Exploration

Initially, Aquila identifies the estimated position of its prey and circles widely, covering a broad area to explore and track it.

$$P_i(t+1) = P_{best}(t) \times \left(1 - \frac{t}{T}\right) + P_M(t) - P_{best}(t) * rand \quad (7)$$

Wherever, $P_i(t+1)$ denotes the location of i^{th} individual at iteration $t+1$, P_{best} signifies finest position identified at the present repetition, whereas $P_M(t)$ signifies the normal location of all entities during the similar repetition

$$P_M(t) = \frac{1}{N}\sum_{i=1}^{N} P_i(t) \quad (8)$$

Here, N denotes Population size of the swarm. T denotes maximum amount of allowed iterations.

Method II: Sharp Exploration

Within the search phase, Aquila identifies its prey; the situation narrows its search by circling above the target area and preparing to land.

$$P_i(t+1) = P_{best}(t) \times Levy(D) + P_R(t) + (q-p) * rand \quad (9)$$

$$Levy(D) = s \times \frac{\mu \times \sigma}{|v|^{\frac{1}{\beta}}} \quad (10)$$

Where,β constant value fixed as 1.5. $P_R(t)$ denotes a randomly particular applicant at the present repetition, whereas p and q define the spring pattern,

$$q = r \times cos\,(\theta) \quad (11)$$

$$p = r \times sin\,(\theta) \quad (12)$$

$$r = r_1 + U \times D_1 \quad (13)$$

$$\theta = -\omega \times D_1 + \theta_1 \quad (14)$$

$$\theta_1 = \frac{3\pi}{2} \quad (15)$$

Here, r_1 is a fixed value, D_1 indicate a digit selected since 1 to the dimensionality of the problem, and $\omega = 0.005$ represents stable constant.

Method III: Extended Exploitation

In the exploitation phase, if the Aquila fails to capture its target, it reinitializes its position. The individuals then update their locations according to the following equation:

$$P_i(t+1) = QF \times P_{best}(t) - G_1 \times P_i(t) \times rand - G_2 \times Levy(D) + rand \times G_1 \quad (16)$$

Here, QF refers the Quality Function, which is used to stable the exploration method then is computed as follows:

$$QF = t^{\frac{2 \times rand - 1}{(1-T)^2}} \quad (17)$$

$$G_1 = 2 \times rand - 1 \quad (18)$$

$$G_2 = 2 \times (1 - \frac{t}{T}) \quad (19)$$

Thus the AOA optimize the DADR technique to get good performance and enhance the classification for the EV.

III. Result and Discussion

Using Python software the result is valuated. By using the Indian Emergency Vehicle dataset the process is done.

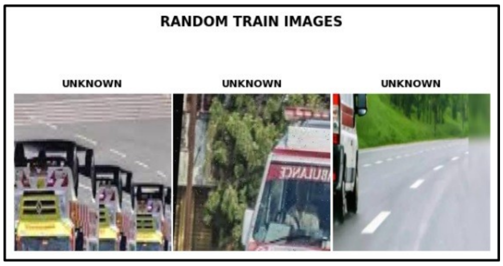

Fig. 5. Random train images

Figure 5 shows the random train images. It is randomly chosen from the Indian emergency vehicle dataset.

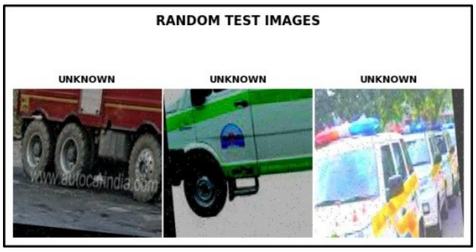

Fig. 6. Random test images

Figure 6 shows the random test images. It is randomly chosen from the Indian emergency vehicle dataset.

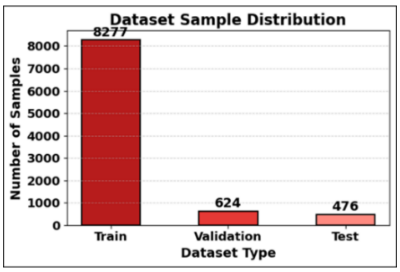

Fig. 7. Dataset Sample Distribution

Figure 7 shows the dataset sample distribution. Here the train have the vaue of 8277, validation have 624 and test have 476

Fig. 8. Input Image

Figure 8 displays the input image. It is taken from the Indian emergency vehicle dataset which is randomly chosen for further analysis.

Fig. 9. Original, Resized and Gray Image

Figure 9 shows the original, resized and gray image. Here the original image is resized using image resizing technique and then the resized image is gray scale using grayscale conversion technique

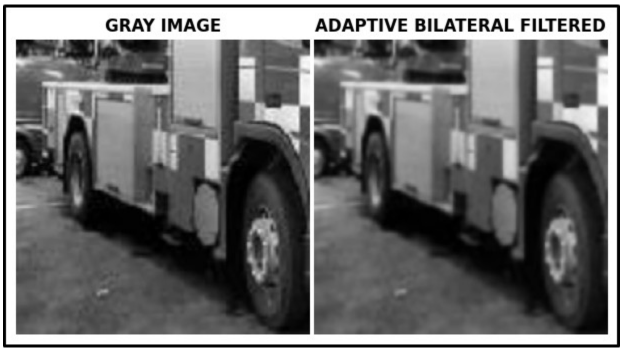

Fig. 10. Gray to ABF

Figure 10 shows the gray image to ABF image. Here the noises are removed and gets the good quality of ABF image.

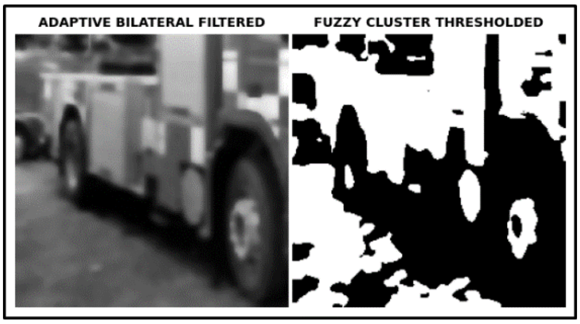

Fig. 11. ABF to Fuzzy Cluster Threshold

Figure 11 displays an ABF to fuzzy cluster thresholding. It produces an adaptive threshold that effectively differentiates EV even in challenging traffic scenes with brightness deviations

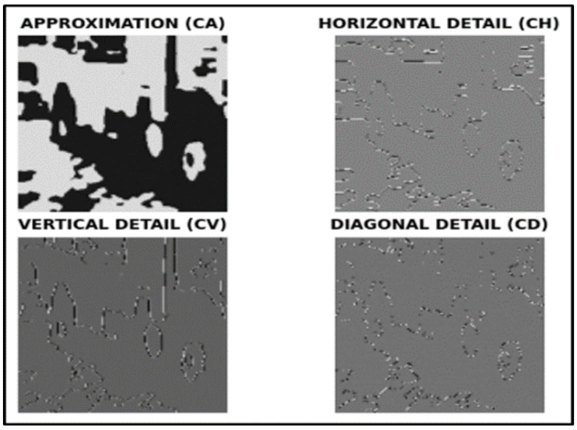

Fig. 12. DWT

Figure 12 shows a DWT. DWT smashes an image into several frequency sub-bands, permitting it to signify both spatial as well as frequency data as CA, CH, CV and CD.

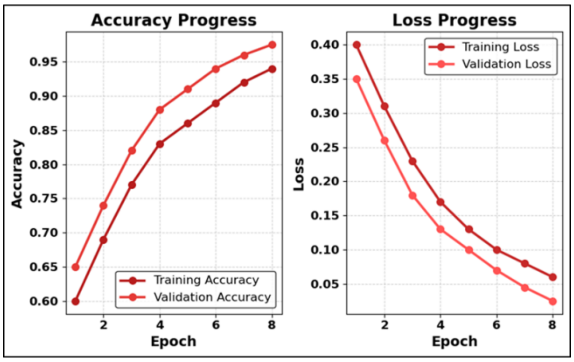

Fig. 1. Accuracy progress as well as loss

Figure 13 shows the accuracy as well as loss. The model's training as well as validation accuracy consistently improve with epochs, reaching over 98%, whereas losses steadily decrease, presenting effective learning.

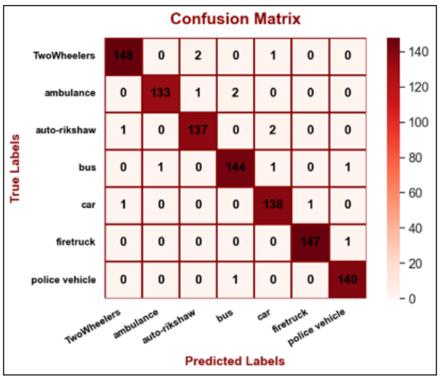

Fig. 14. Confusion Matrix

Figure 14 displays the confusion matrix. The confusion matrix appearances that most vehicle types are accurately classified, with very few misclassifications across classes. High diagonal values indicate excellent model performance and strong discriminative capability among different vehicle groups.

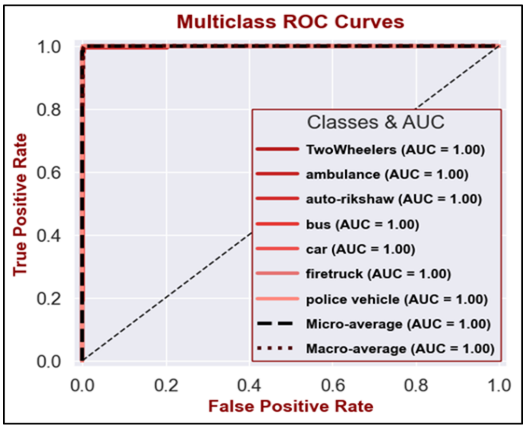

Fig. 15. ROC Curve

Figure 15 shows the ROC curve. In the proposed work all the classes have the value of 1 which is have good performance.

A. Comparison for Accuracy

TABLE I. COMPARISON FOR ACCURACY.

Author	Method	Accuracy (%)
Zohaib *et al* [11]	ATSN	96.19
Hosseinzadeh, *et al* [15]	MOGA	96.65
Proposed work	DADR-AquilaNet	**98**

In table I comparison of the proposed and the existing method is valuated. ASTN have the value of 96.19% [11], MOGA have 96.65% [15] and proposed DADR-AquilaNet have the value of 98%.

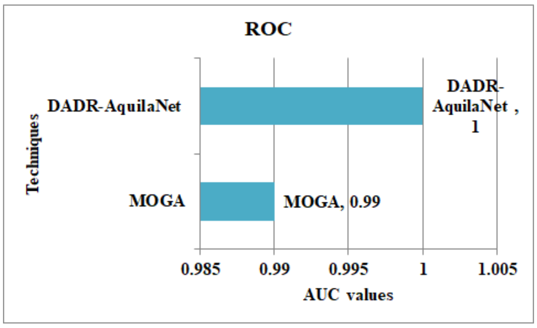

Fig. 16. ROC Comparison

Figure 16 shows the comparison of ROC. MOGA have 0.99 [15] and proposed DADR-AquilaNet have the value of 1.

IV. CONCLUSION

In this paper, a proposed DADR-AquilaNet recognises and prioritization of EV in congested traffic environments using the Indian Emergency Vehicles Dataset. The ABF effectively denoised the traffic images whereas preserving essential edge and FCBT enhanced the segmentation quality under complex urban conditions. Furthermore, DWT enabled multi-level feature extraction to capture fine-grained discriminative patterns. The proposed DADR-AquilaNet validated higher feature learning through the synergistic combination of dilated convolutions, residual learning, as well as attention mechanisms, resulting in improved contextual awareness as well as robustness to occlusions and illumination variants. The AOA effectively optimized network parameters, accelerating convergence and enhancing classification performance. Have the accuracy value of 98% which is high compared to the existing method.

REFERENCES

[1] A. Shabbir, "Smart city traffic management: Acoustic-based vehicle detection using stacking-based ensemble deep learning approach," IEEE access, vol. 12, pp. 35947-35956, 2024.
[2] M. Shatnawi, "An enhanced model for detecting and classifying emergency vehicles using a generative adversarial network (GAN)," Vehicles, vol. 6, no. 3, pp. 1114-1139, 2024.
[3] M. Shamsi, 'Reinforcement learning for traffic light control with emphasis on emergency vehicles," The Journal of Supercomputing, vol. 78, no. 4, pp. 4911-4937, 2022.
[4] M. Cao, "A gain with no pain: Exploring intelligent traffic signal control for emergency vehicles," IEEE Transactions on Intelligent Transportation Systems, vol. 23, no. 10, pp. 17899-17909, 2022.
[5] N. Pathik, "AI enabled accident detection and alert system using IoT and deep learning for smart cities," Sustainability, vol. 14, no. 13, pp. 7701, 2022.
[6] Y. Cui, "Optimizing Internet of Things-based intelligent transportation system's information acquisition using deep learning," IEEE Access, vol. 11, pp. 11804-11810, 2023.
[7] A. Shabbir, "Smart city traffic management: Acoustic-based vehicle detection using stacking-based ensemble deep learning approach," IEEE access, vol. 12, pp. 35947-35956, 2024.
[8] T. Maurya, "Real Time Vehicle Classification Using Deep Learning—Smart Traffic Management," Engineering Reports, vol. 7, no. 3, pp. e70082, 2025.
[9] A.T.R. Priya "Integrating Whale Optimization Algorithm and Machine Learning for Efficient Traffic Management," International Journal of Advanced Trends in Engineering and Management (IJATEM), vol. 02, no. 10, pp. 12-23, 2023.
[10] D. Jayakumar, "Emergency Vehicle Classification Using Combined Temporal and Spectral Audio Features with Machine Learning Algorithms," Electronics, vol. 13, no. 19, pp. 3873, 2024.
[11] M. Zohaib, "Enhancing emergency vehicle detection: A deep learning approach with multimodal fusion," Mathematics, vol. 12, no. 10, pp. 1514, 2024.
[12] A. Louati, "Deep learning and case-based reasoning for predictive and adaptive traffic emergency management," The Journal of Supercomputing, vol. 77, no. 5, pp. 4389-4418, 2021.
[13] M. Gupta, "Real-time traffic control and monitoring," e-Prime-Advances in Electrical Engineering, Electronics and Energy, vol. 5, pp. 100211, 2023.
[14] L. Parada, "Safe and efficient manoeuvring for emergency vehicles in autonomous traffic using multi-agent proximal policy optimisation," Transportmetrica A: Transport Science, vol. 21, no. 2, pp. 2246586, 2025.
[15] M. Hosseinzadeh, "MPC-based emergency vehicle-centered multi-intersection traffic control," IEEE Transactions on Control Systems Technology, vol. 31, no. 1, pp. 166-178, 2022.