SvedbergOpen
DISSEMINATION OF KNOWLEDGE

International Journal of Artificial Intelligence and Machine Learning

Publisher's Home Page: https://www.svedbergopen.com/

Research Paper Open Access

A Boundary aware Graph attention framework for License Plate Detection under Challenging Traffic Conditions

B. Kalaimathi[1], C.S. Manikandababu[2], Jeevanantham C[3], T. Rajkumar[4]

[1]Department of Electronics and Communication Engineering, Sri Ramakrishna Engineering College, Coimbatore.
E-mail: kalaimathi@srec.ac.in
[2]Department of Electronics and Communication Engineering, Sri Ramakrishna Engineering College, Coimbatore.
E-mail: manikandababu.shelvaraju@srec.ac.in
[3]Department of Information Technology, Dr. Mahalingam College of Engineering and Technology, Coimbatore.
E-mail: jeevananthamc@drmcet.ac.in
[4]Department of Information Technology, Nandha College of Technology, Erode.
E-mail: tsrajkumar87@gmail.com

Abstract:
License Plate Detection (LPD) is one of the most important components of intelligent transportation systems, traffic surveillance systems, vehicular monitoring systems and automated law enforcement systems. LPD is complicated by various factors including lightening condition, occlusion, motion blur, complicated background condition and small scale license plates. In this research work introduced a new framework called Boundary-Aware Graph Attention Network (BA-GAT) to solve above mentioned problems in order to provide robust and accurate license plate detection under unconstrained traffic scenarios. The framework contains a Boundary-Aware Feature Enhancement Module in which Sobel and Scharr operators are utilized to enhance boundaries of license plates prior to deep feature extraction process. An improved CNN backbone using Channel & Spatial Attention mechanism (CBAM) is utilized to select discriminative channels and spatial parts and ignore the unnecessary background information. Moreover, Multi-Scale Feature Fusion Network is used to fuse hierarchical feature representations and detect small and long range license plates. To benefit from contextual relations between candidate regions, a Graph Attention Network (GAT) module is used to aggregate neighboring objects adaptively. Experimental results are reported on Kaggle License Plate Dataset, UFPR-ALPR, and CCPD datasets. The proposed BA-GAT model obtains precision of 97.8%, recall of 96.9%, F1-score.

Keywords: License Plate Detection, Graph Attention Network, CBAM, Boundary-Aware Learning, Multi-Scale Feature Fusion, Intelligent Transportation Systems.

I. Introduction

Automatic license plate detection (LPD) has emerged as a core requirement in any intelligent transportation system (ITS) for various applications like traffic monitoring, vehicle tracking, tolling, parking management and policing. The recent advancement of deep learning has helped license plate detection systems greatly as they enable automated discrimination of useful visual cues from the traffic scene. License plate detection is still a difficult problem because of environmental factors like lighting variations, motion blur, weather effects, occlusion, background clutter and variations in sizes and orientations of license plates. The problem is more prominent when considering real-life scenarios in which license plates constitute a very small fraction of the overall image frame.

Several deep learning-based techniques have been suggested to overcome these difficulties. YOLO-based object detectors, attention-guided neural networks, feature pyramid networks and end-to-end recognition models have shown some effectiveness on benchmark data sets. Even though these techniques can improve object detection efficiency, their main emphasis is on feature extraction and multi-resolution learning, completely ignoring the issue of boundary enhancement and neighboring object relationship. Therefore, the object detection efficiency may be adversely affected in cases where there are small license plates, occlusion and background similarities.

To address the above problems introduced a new technique for license plate detection of Boundary Aware Graph Attention Network (BA-GAT). This method comprises four key modules: Boundary Aware Feature Enhancement Module, CBAM enhanced CNN architecture, Multi-Scale Feature Fusion Network and Graph Attention Network. Through the boundary enhancement module, boundaries around the plates will be highlighted, whereas the CBAM architecture enhances both channel and spatial features. Multi-scale fusion network allows for efficient detection in different plate sizes, while graph attention network helps in detecting the context relationships among candidate license plate areas and nearby objects. The experimental results conducted on three datasets namely Kaggle License Plate Dataset, UFPR-ALPR and CCPD, showed that our proposed BA-GAT method surpasses all the current state-of-the-art techniques.

II. Related Work

The researchers Peng et al. (2024) presented an effective license plate detection approach that solves the problem of small annotations of license plates in traffic datasets. Annotation improvement and robustness of detection can be achieved with appropriate training. According to the experiment, better accuracy of detection was achieved especially for small and distant license plates. Unfortunately, the framework does not utilize any relations between neighboring objects [1]. The authors Wang et al. (2025) developed an advanced YOLOv5s object detector with LPRNet for license plate recognition in difficult circumstances. The authors have achieved better feature extraction and recognition accuracy under illumination changes, occlusions, and motion blur. Superior accuracy in comparison with other YOLO-based detectors was revealed during experiments. Unfortunately, the problem of contextual reasoning of candidate plate regions has not been considered [2].

The authors Xiong et al. (2026) suggested an advanced methodology for license plate recognition based on the CSCM-YOLOv8 and CSM-LPRNet. The authors introduced channel-spatial context modeling to enhance feature extraction and character recognition accuracy. Excellent results have been obtained in difficult traffic situations with complex backgrounds. However, graph-based relation learning has not been used in the detection phase [3]. An information-maximizing license plate detection algorithm was introduced by Wang et al. (2022). This algorithm helps optimize feature distribution to increase the accuracy of localization and detection features under different environmental situations. Experiments have demonstrated the effectiveness of the algorithm on various benchmark datasets. However, the model does not include any explicit attention mechanism for boundary enhancement [4].

An attention-guided dual feature pyramid network (ADFPN) for license plate detection was developed by Xiong et al. (2022). This algorithm uses multi-scale feature fusion and attention mechanism to detect small objects. The method demonstrated much better localization than other feature pyramid architectures. Nonetheless, the long-range contextual dependency was not utilized [5]. Tao et al. (2024) developed a real-time license plate detection and recognition model applicable for unconstrained traffic environments. This algorithm includes efficient object detection and recognition components to allow fast processing. The experiment results have proven the effectiveness of the method in real-world traffic surveillance applications. Nonetheless, the model is not robust enough to handle severe occlusions [6].

Chung et al. (2024) proposed YOLO-SLD – an attention mechanism-aided YOLO architecture for license plate detection. Attention mechanisms were used to improve feature extraction and prevent background noise. The method showed better precision and recall rates than regular YOLO detectors. However, the algorithm does not consider the relationship between the areas [7]. Zhou et al. (2023) considered deep learning approaches for license plate recognition in complex road environments. Various architectures for detection and recognition were analyzed depending on environmental conditions. The authors emphasized the efficiency of deep neural networks in processing complicated license plates. Nevertheless, high computational complexity is a drawback of the algorithm for practical application [8].Zou et al. (2022) suggested using YOLOv3 and ILPRNet together for license plate detection and recognition. YOLOv3 was used for localization, and ILPRNet for recognition. Experimental results revealed good performance of the model for various datasets. Nevertheless, the method is not effective in cases of illumination changes [9].

Seo and Kang (2022) created a layout-invariant attention-based license plate detection and recognition model. This architecture is capable of addressing various types of license plates without the need to provide pre-defined layout. Attention-based models improve feature representations and recognition accuracy. However, the presented model does not involve multi-scale context learning [10]. Fan (2026) created a robust multi-oriented license plate detector to detect license plates in arbitrary orientations. The method also involves an end-to-end recognition module to minimize processing complexity. Experimentally, good robustness to rotation and perspective transformation has been achieved. However, boundary-aware localizations have not been addressed [11].Qin and Liu (2021) presented an end-to-end framework for joint detection and recognition of license plates in unconstrained settings. The detection and recognition tasks are incorporated within one network, thus minimizing error accumulation. Experiments confirmed competitive performance on traffic datasets. However, small-sized license plate detection is still problematic [12].

Tao et al. (2021) have put forth a framework for license plate localization and recognition based on object detection in challenging environments. The framework employs modern-day object detection methodologies to increase the accuracy of localization in different traffic scenarios. Experimental results have shown the framework's applicability in practice. However, the mechanisms involved in feature enhancement are somewhat limited [13]. Zhang et al. (2021) have created a strong attention-based framework for license plate recognition in an unconstrained environment. The attention mechanism has enabled selective attention of character regions, thereby increasing the accuracy of recognition. Nonetheless, the framework is mostly concerned with recognition improvement [14]. Chen et al. (2020) developed a multi-branch attention neural network that could simultaneously detect vehicles and their license plates. The network was able to learn features of the vehicles and plates together, thus improving localization. Experiments proved increased accuracy and efficiency of detection. However, the approach does not provide any graph reasoning abilities [15]. Liang et al. (2022) developed EGSANet, the Edge-Guided Sparse Attention Network, which is designed for license plate detection in difficult conditions. The network makes use of edge information and sparse attention mechanism in order to increase localization accuracy. Superior results were achieved for cluttered and low-quality traffic situations. However, the approach does not take into account dynamic interactions between neighbor objects [16].

Weihong and Jiaoyang (2020) developed the license plate recognition framework based on deep learning. This work examined the effectiveness of using convolutional neural networks in different conditions such as varying lighting, orientation, and occlusions. Experimental results proved high recognition rate. However, the framework is mainly aimed at recognition, not localization of plates [17]. The CNN architecture based license plate detection system has been proposed by Bathirappan and Soundararajan (2026) with the use of channel and spatial attention. Incorporating CBAMs in this network ensures focus on important plate-related features while ignoring the background features. Results have proved better in terms of precision, recall, F1-score, and mAP than the traditional CNN architectures. But still, there is no way of modeling the contextual relations between possible plate regions [18].

Spatial attention based license plate detection system has been developed by Lee et al. (2021). This system ensures the better localization performance due to attention to plate regions while suppressing background information. Results have proved the effectiveness of the system in real-road traffic conditions. However, there is less work done on the interaction of features at different scales [19]. The Deep Learning Boundary Filtering Method (BFM) for detecting license plates from frames of surveillance videos was introduced by Bathirappan and Jayanthy in 2024. The model makes use of boundaries in order to improve the localization process of the license plate. From the experiments conducted, it can be observed that the detection accuracy is improved when compared to traditional methods. However, no attention mechanisms are included in the model [20].

III. Proposed Methodology

The BA-GAT Architecture that is suggested is intended for accurate detection of license plates under complex traffic environment shown in Fig 1. The process starts from getting a traffic image as an input. Then the input goes through the Boundary-Aware Feature Enhancement Module that can find prominent edges and boundaries related to the license plates. The enhanced feature is further passed to CBAM-CNN backbone network, like ResNet50 or EfficientNet, where channel and spatial attention can help emphasize the important part of the images while suppressing the rest of the background. After that the output of the CNN backbone goes through the Multi-Scale Feature Fusion Network (FPN/PAN) that helps to detect the license plates at different scale and sizes. On the basis of these fusion features, the candidate regions that represent the license plate are created along with the contextual objects and vehicles. Further, in the Graph Construction Module these regions get the relationships and the Graph Attention Network (GAT) learns their spatial and contextual relations by assigning the adaptive attention weights to them. After processing the features through GAT they go to the Detection Head to classify the license plate and perform the bounding box regression.

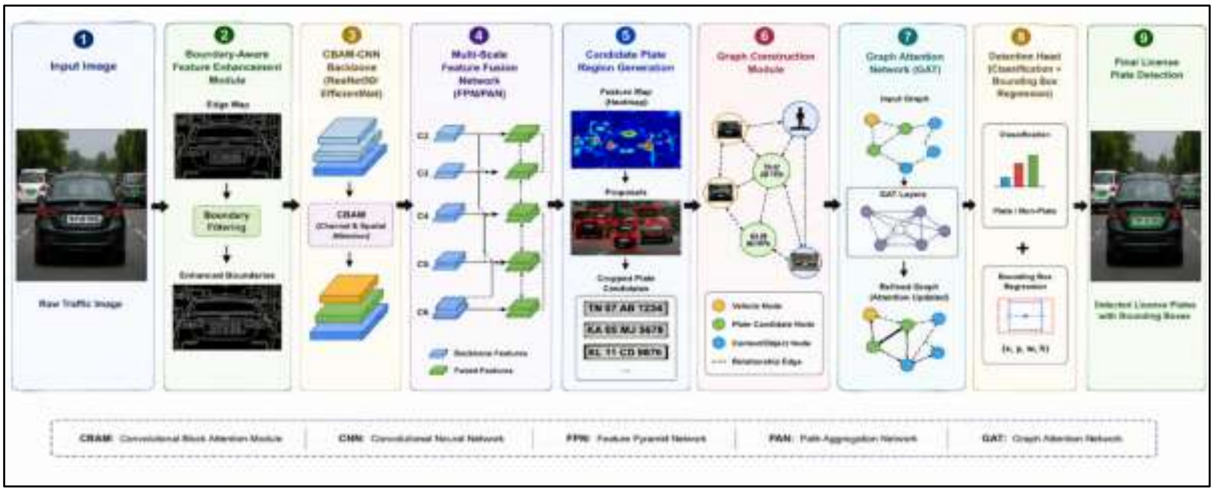

Fig 1. Architecture of Boundary-Aware Graph Attention Network

A. Boundary-Aware Feature Enhancement Module

The Boundary Aware Feature Enhancement module illustrated in fig 2 is intended to assist the license plate localization by stressing out the boundary information before feature extraction using deep networks. License plates in real-world traffic images usually have issues related to poor contrast, motion blurring, shadowing, and complicated backgrounds, hence leading to suboptimal detection results. This challenge is solved by integrating the gradient edge detector with the Boundary Filtering Method (BFM) for producing the edge confidence map, representing the license plates boundaries. Such representation of the boundaries makes it easier for the CNN layers to learn features.

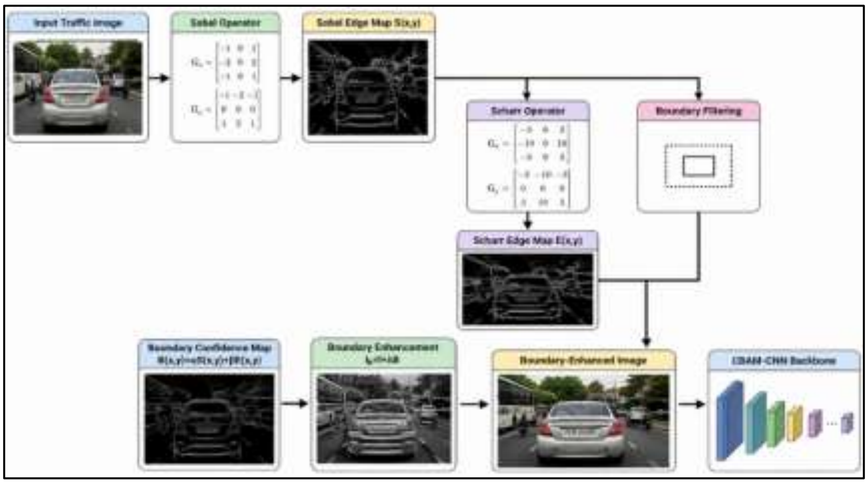

Fig 2. Boundary-Aware Feature Enhancement Module

The input image $\mathrm{I}(x, y)$is processed using Sobel and Scharr operators to estimate horizontal and vertical intensity gradients and the Sobel gradient magnitude is computed as

$S(x, y) = \sqrt{(G_x^2 + G_y^2)}$ ------------- (1)

Where $S(x, y)$ represented as Sobel edge response at pixel location (x,y), Gx is Horizontal intensity gradient and Gy : Vertical intensity gradient. The Sobel kernels are defined as

$$G_x = \begin{bmatrix} -1 & 0 & 1 \\ -2 & 0 & 2 \\ -1 & 0 & 1 \end{bmatrix} ------(2)$$

$$G_y = \begin{bmatrix} -1 & -2 & -1 \\ 0 & 0 & 0 \\ 1 & 2 & 1 \end{bmatrix} \text{------------- (3)}$$

The Scharr operator provides improved rotational symmetry and stronger edge localization and Scharr edge response is calculated as

$$E(x, y) = \sqrt{\left(H_x^2 + H_y^2\right)} ------- (4)$$

Where, $E(x, y)$ is Scharr edge response, H_x is Horizontal Scharr gradient, H_y is Vertical Scharr gradient. It is expressed as

$$H_x = \begin{vmatrix} -3 & 0 & 3 \\ -10 & 0 & 10 \\ -3 & 0 & 3 \end{vmatrix} ------(5)$$

$$H_y = \begin{bmatrix} -3 & -10 & -3 \\ 0 & 0 & 0 \\ 3 & 10 & 3 \end{bmatrix} ------- (6)$$

The Sobel operator captures coarse structural boundaries and the Scharr operator preserves fine edge details. Combining both enables boundary extraction under varying illumination and noise conditions. The outputs of the Sobel and Scharr operators are fused to construct a boundary confidence map. The boundary response is formulated as

$$B(x, y) = \alpha S(x, y) + \beta E(x, y) ------(7)$$

Where $B(x, y)$ Boundary confidence value at pixel (x, y), $S(x, y)$ is Sobel edge magnitude, $E(x, y)$ represented as Scharr edge magnitude, α , β denoted as Sobel weighting coefficient . The coefficients satisfy the following condition.

$$\alpha + \beta = 1 ------(8)$$

where $0 \leq \alpha, \beta \leq 1$. α and β are empirically optimized during training to balance coarse and fine boundary information. A larger value of $B(x, y)$ indicates a higher probability that the corresponding pixel belongs to a license plate boundary. To further strengthen plate contours, the Boundary Filtering Method (BFM) enhances the original image using the generated boundary confidence map. The enhanced image is computed as

$I_b(x,y) = I(x,y) + \lambda\, B(x,y)$ -------- (9)

where

- $I(x,y)$ is Original input image
- $I_b(x,y)$ is Boundary-enhanced image
- $B(x,y)$ is Boundary confidence map
- λ is Boundary enhancement factor

The parameter λ controls the contribution of boundary information. Higher values increase edge prominence, whereas lower values preserve the original image characteristics. Substituting the equation, the enhanced image can be represented as

$I_b(x,y) = I(x,y) + \lambda\,[\alpha S(x,y) + \beta E(x,y)]$ --------- (10)

This formulation enables adaptive enhancement of plate contours while suppressing weak background edges.The boundary-enhanced image is $I_b(x,y)$ subsequently forwarded to the CBAM-CNN backbone for feature extraction. Since license plate boundaries are explicitly emphasized before deep processing, the CNN can focus on structurally meaningful regions and learn more discriminative features.

B. CBAM-Enhanced CNN Backbone

The CNN backbone enhanced by CBAM is used for the extraction of very discriminating license plate features from the enhanced boundary image obtained from the previous stage shown in Fig 3. Traditional CNNs usually assign equal importance to all channels of feature maps and spatial locations, which might lead to the introduction of unwanted information in the form of background data. The proposed approach solves this problem by including the CBAM layer within the backbone model. The CBAM block performs channel-wise and spatial attention successively to enhance useful license plate features and reduce redundant background responses.

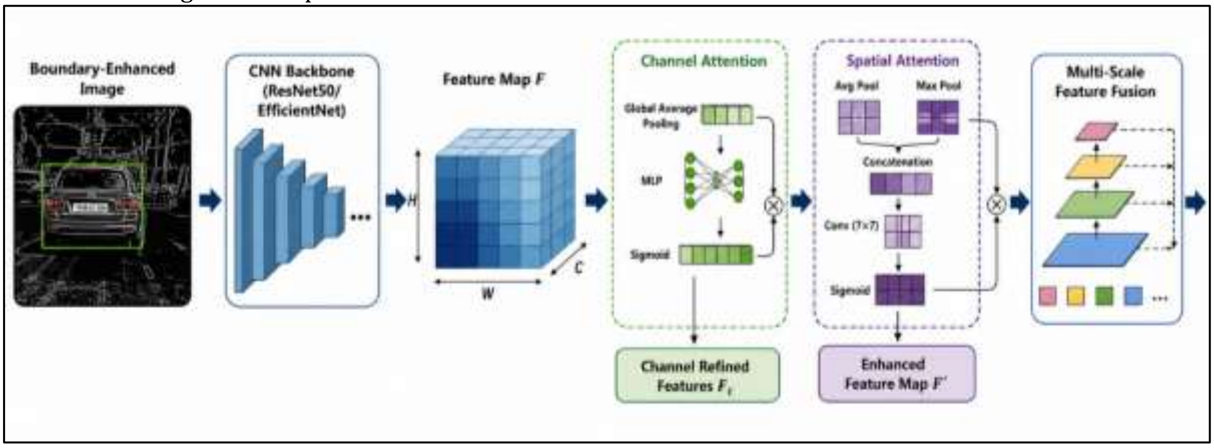

Fig 3. CBAM-Enhanced CNN Backbone Module

The boundary enhanced image I_b is then processed using deep CNN backbones like ResNet50, EfficientNet-B3 or CSPDarkNet in order to derive hierarchical feature representation. The derived feature map can be mathematically expressed as $F \in R(C\times H\times W)$ where F stands for the input feature map, C denotes the number of channels in the feature map and H and W stand for the height and width of the feature map respectively. The feature map F has rich spatial and semantic information that would help in accurately detecting license plates. In order to select the most important feature channels, a channel attention mechanism is utilized. Global Average Pooling and Global Max Pooling operations are performed on the feature map F in order to derive channel descriptor vectors F_{avgc} and F_{maxc}. These channel descriptors are fed into a shared Multi-Layer Perceptron (MLP) defined as MLP(x) = W1 (W0(x)) where W0 and W1 are the first and second fully-connected layers.

$Mc(F) = \sigma\Big(MLP\big(AvgPool(F)\big) + MLP\big(MaxPool(F)\big)\Big)$ ------ (11)

Where $M_c(F)$ is the channel attention map and $\sigma(.)$ denotes the sigmoid activation function defined as

$$\sigma(x) = \frac{1}{(1+e^{-x})} - - - - - - - - (12)$$

The refined channel feature map is then obtained as,

$$Fc = Mc(F) \otimes F - - - - - - (13)$$

Where ⊗ is an element-by-element product. Channel attention highlights those channels that have characteristics of the license plate like edges, corners, textures, characters and suppresses other feature channels. Though channel attention can detect important feature channels, it does not identify the exact position of license plates. Thus, spatial attention is required. The average poolings and max poolings are conducted in the channel direction to get F_{avgs} and F_{maxs} respectively. The two feature maps are concatenated and then convoluted with a 7×7 filter and the result is a spatial attention map defined as:

$$Ms(F) = \sigma\left(f^{7x7}([Favg^{s}\ ;\ Fmax^{s}])\right) - - - - - - - (14)$$

Where $M_s(F)$ represents the spatial attention map, f_{7x7} denotes a convolution operation with a 7x7 kernal and [;] indicated channel – wise concatenation. The spatially refined feature map is then calculated as,

$$Fs' = Ms(F) \otimes Fc - - - - - - - - (15)$$

This operation highlights informative license plate regions while suppressing irrelevant background structure such as roads, buildings and surrounding vehicles roads, buildings and surrounding vehicles. Finally the complete CBAM module sequentially applies channel and spatial attention to generate the enhanced feature represented as

$$F' = Ms(F) \otimes (Mc(F) \otimes F) - - - - - (16)$$

where F' denotes the final CBAM-enhanced feature map. The resulting representation contains enriched license plate information with reduced background interference, thereby improving localization and detection accuracy. The proposed CBAM-enhanced backbone enhances license plate-specific feature channels, accurately localizes salient plate regions, suppresses irrelevant background responses, improves robustness against occlusion and illumination variations, and generates discriminative features for subsequent Multi-Scale Feature Fusion and Graph Attention Network modules.

A. Multi-Scale Feature Fusion Module

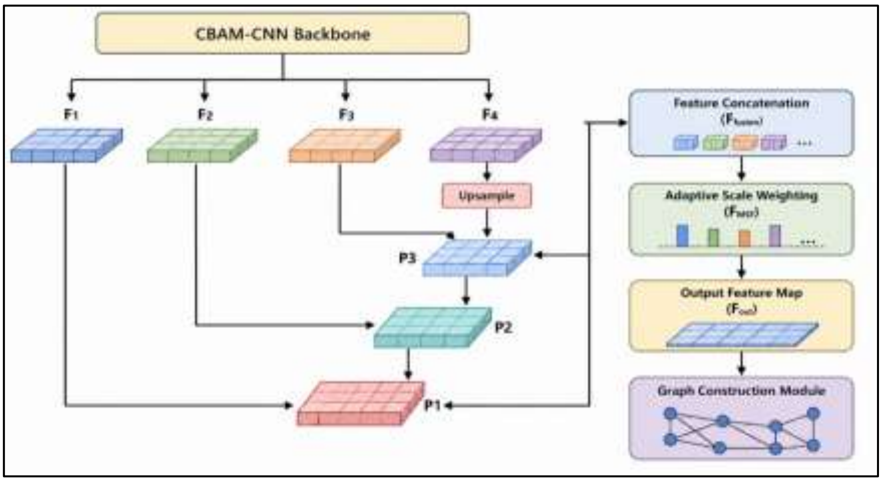

Fig 4. Multi-Scale Feature Fusion Module

Multi-Scale Feature Fusion Module is illustrated in Fig 4 employed to boost the accuracy in detecting tiny, far-off, and occluded license plates that can be seen in the images taken by traffic surveillance cameras. Traditional CNNs mainly use high-level semantic information, which could result in a loss of small-scale details necessary for object detection. In order to deal with this problem, the introduced approach uses the FPN to combine low-level texture information and high-level semantic information together. Through combining feature maps from different layers in the CNN backbone network, the detector could easily extract both boundary information and context information of license plates. The CBAM enhanced CNN backbone generated a feature map from different stages such as F_1, F_2, F_3, F_4.. The spatial resolution satisfy the following condition

$H_1 > H_2 > H_3 > H_4$ and $W_1 > W_2 > W_3 > W_4$. Where H and W represents the height and weight of the feature map.

To recover spatial details lost during convolution and pooling operations, deep feature maps are progressively up sampled and merged with shallow feature maps. The feature propagation process is expressed as

$P_4 = Conv(F_4)$ ------------ (17)

$P_3 = Conv(F_3) + Up(P_3)$ -------------------- (18)
$P_2 = Conv(F_2) + Up(P_3)$ -------------------- (19)
$P_1 = Conv(F_1) + Up(P_2)$ ----------------- (20)
Where P_i denoted as Pyramid Feature map at level i, $Conv$ (.) :1x1 convolution for channel alignment. Up(.) denoted as up sampling operation. The fusion process is represented as,
$F_{fusion} = P_1 \oplus P_2 \oplus P_3 \oplus P_4$ ------------------------ (21)
In adaptive scale weighted fusion is expressed as
$F_{MSF} = \omega_1 P_1 + \omega_2 P_2 + \omega_3 P_3 + \omega_4 P_4$ ----------------------- (22)
Here, ω_i is a learnable scale attention weight, F_{MSF} gives the final multi scale feature representation. The final fused feature map is given as, $F_{out} = F_{fusion} + F_{MSF}$. In candidate region proposal the potential license plate regions are generated using Anchor based detector and YOLO style prediction head. It is represented as $R_i = (x_i, y_i, \omega_i, h_i)$. Graph Construction module is represented by $G = V, E$. V and E are the candidate plate nodes and Spatial Connections. Computation done by using Euclidean distance and similarity.
$e_{ij} = Similarity(v_i, v_j)$ ------------------------------ (23)
The Graph Attention Network (GAT) coefficient is defined as,
$\alpha_{ij} = \frac{\exp(LeakyReLU(a^T[Wv_i||Wv_j]))}{\sum_{k \in N_i} \exp(\exp(LeakyReLU(a^T[Wv_i||Wv_k]))}$ --------------- (24)
Where a denotes the attention vector, W is Learnable weight matrix.
Finally loss function is represented as
$L_{total} = L_{cls} + \lambda_1 L_{box} + \lambda_2 L_{boundary} + \lambda_3 L_{graph}$ ----------------- (25)

IV. Experimental Results

Kaggle License Plate dataset is about 1,732 annotated images of vehicles together with their corresponding license plate bounding boxes. The Kaggle License Plate dataset is mostly used to measure the ability of the proposed model to detect license plates in surveillance images that are in moderate traffic. The images are resized into two scales namely 256×256 and 512×512 pixels in order to examine how scaling of images affects detection results. UFPR-ALPR dataset is composed of more than 4,500 images of vehicles that are collected in real traffic. The UFPR-ALPR dataset gives annotation for not only license plate localization but also for character recognition. The dataset has diverse lighting condition and plate orientation which makes it ideal for measuring the generalization power of the proposed framework. Chinese City Parking (CCPD) dataset is one of the largest license plate datasets with over 250,000 annotated images. Some of the challenges contained in CCPD dataset include tilted license plates, motion blur, partially occluded plates, complex background, weather variation and varied viewing distance. Given its size and diversity the CCPD dataset is used to test the scalability of BA-GAT framework in intelligent transportation application.

Table 1. Dataset Description

Kaggle License Plate Dataset	1732	1039 (60%)	346 (20%)	347 (20%)	256×256 / 512×512
UFPR-ALPR	4500+	2700	900	900	Various
CCPD	250,000+	150,000	50,000	50,000	Various

The datasets employed in evaluating the proposed framework have been illustrated in Table 1. The Kaggle License Plate Dataset, UFPR-ALPR and CCPD datasets consist of different traffic images captured in different environmental settings and lighting conditions. They have been partitioned into training, validation and testing subsets. Out of all the aforementioned datasets, CCPD consists of a large number of images, i.e., more than 250,000 images.

Table 2. Ablation Study of BA-GAT

Method	Boundary Module	CBAM	FPN	GAT	mAP@0.5 (%)	F1 (%)
Baseline CNN	✗	✗	✗	✗	88.3	89.1
CNN + CBAM	✗	✓	✗	✗	90.8	91.4
CNN + CBAM + FPN	✗	✓	✓	✗	93.5	94.2
CNN + CBAM + FPN + Boundary	✓	✓	✓	✗	95.1	95.8
Proposed BA-GAT	✓	✓	✓	✓	97.4	97.1

The performance of each component in the proposed BA-GAT framework is presented in Table 2. It can be seen from the table below that as each module is added in the process, there is a consistent increase in both mAP@0.5 and F1-score. While the Baseline CNN model scores 88.3% in mAP, the complete BA-GAT model scores 97.4%.

Table 3. Comparison with State-of-the-Art Methods

Method	Precision	Recall	F1	mAP@0.5
YOLOv5s-LPRNet [2]	92.4	91.1	91.7	90.6
CSCM-YOLOv8 [3]	93.2	92.7	92.9	92.4
ADFPN [5]	94.1	93.5	93.8	93.2
YOLO-SLD [7]	94.8	94.2	94.5	94.1
CBAM-CNN [18]	92.4	89.7	91.1	88.9
Proposed BA-GAT	97.8	96.9	97.3	97.4

Table 3 presents a comparison of the suggested BA-GAT model with various state-of-the-art LPD techniques. The suggested framework attains the best precision, recall, F1-score and mean average precision metrics of 97.8%, 96.9%, 97.3% and 97.4%, respectively, compared to other methods being considered. The outstanding performance is because of the use of boundary-aware feature enhancement, multi-scale feature fusion and graph attention learning techniques.

Table 4. Computational Performance

Method	Parameters (M)	FLOPs (G)	FPS	Inference Time (ms)
YOLOv5s	7.2	16.4	55	18
YOLOv8	11.3	28.2	48	21
CBAM-CNN	9.8	22.1	52	19
BA-GAT	12.5	31.5	45	22

Table 4 presents the computational complexity and inference efficiency of different detection models. Although BA-GAT contains slightly higher parameters (12.5M) and FLOPs (31.5G) than existing methods, it maintains real-time processing capability with 45 FPS and an inference time of 22 ms. The modest increase in computational cost is justified by the substantial improvement in detection accuracy. Therefore, BA-GAT provides an effective balance between performance and computational efficiency.

Table 5. Learned Attention Weights

Neighbor Node	Attention Weight
Vehicle	0.42
Candidate Plate	0.31
Traffic Sign	0.12
Background	0.15

Table 5 shows the attention coefficients learned by the Graph Attention Network for neighboring nodes. The vehicle node receives the highest attention weight (0.42), indicating its strong contextual relevance to license plate localization. Candidate plate regions receive a weight of 0.31, while background and traffic sign nodes contribute less significantly. These findings demonstrate that the proposed GAT effectively prioritizes informative contextual relationships during feature aggregation.

Table 6. Statistical Analysis

Method	Mean F1	Std Dev	p-value
YOLOv8	92.9	0.87	<0.05
CBAM-CNN	91.1	0.94	<0.05
BA-GAT	97.3	0.42	Reference

Table 6 presents the statistical significance analysis of the evaluated models based on mean F1-score and standard deviation. The proposed BA-GAT framework achieves the highest mean F1-score of 97.3% with the lowest standard deviation of 0.42, indicating stable and consistent performance across multiple experimental runs. Furthermore, the p-values (<0.05) confirm that the performance improvements are statistically significant. These results validate the reliability and robustness of the proposed approach.

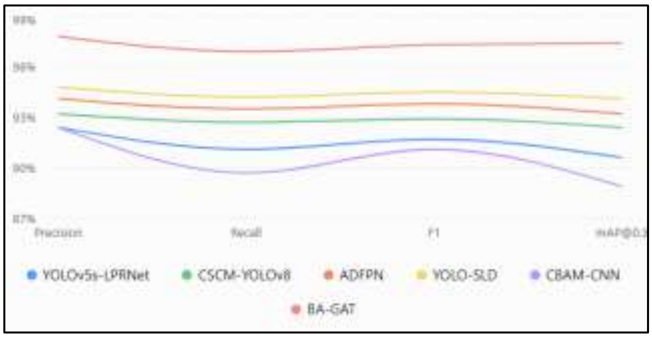

Fig 5. Comparative performance Analysis

Figure 5 compares the detection performance of the proposed BA-GAT model with several state-of-the-art license plate detection methods. The BA-GAT obtains the best values of Precision, Recall, F1-score, and mAP@0.5, indicating its excellent detection ability. The reason for performance gain lies in the combination of boundary-aware enhancement, CBAM attention, and graph attention learning.

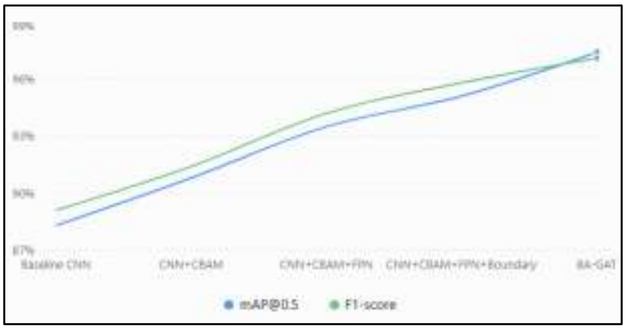

Fig 6. Ablation study performance

This figure 6 shows the contribution of each module used in the proposed BA-GAT model. By adding CBAM, FPN, Boundary Enhancement, and GAT modules step by step, there is a steady increase in the performance of mAP and F1-score. The whole BA-GAT structure obtains the best performance in all structures. It can be concluded that each module plays an important role in the improvement of feature representation and detection.

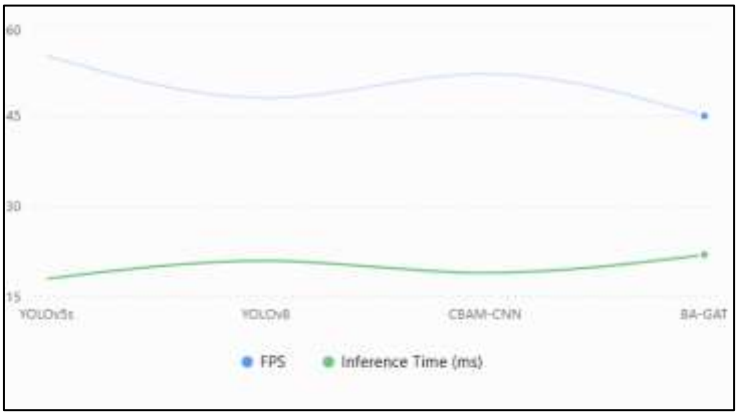

Fig 7. Computational performance comparison

This figure 7 shows a comparison of model complexity and inference efficiency of various detection models. Even though BA-GAT uses more parameters and FLOPs than other methods, it still maintains real-time inference speed of 45 FPS and 22 ms. The extra computational cost of BA-GAT is well worth the detection accuracy. Thus, BA-GAT offers a balance between inference efficiency and detection performance.

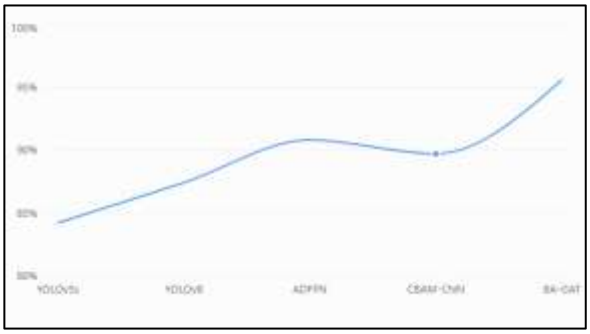

Fig 8. Small license plate detection comparison

In figure 8, recall values of various methods have been evaluated for small-sized license plates. The proposed BA-GAT model yields the highest recall rate for small plates of 95.6% and surpasses other existing techniques. The multi-scale feature integration and graph attention make it possible to better localize tiny and distant license plates. It has been proven by the experiments conducted that BA-GAT is very effective in dealing with difficult real-life scenarios for surveillance where the detection of small-sized license plates is required.

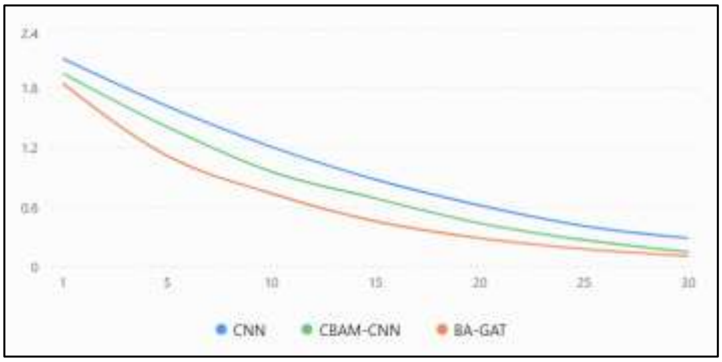

Fig 9. Training Convergence Comparison

As shown in figure 9, the training losses for the proposed BA-GAT model are always lower than CNN and CBAM-CNN for all epochs. In epoch 30, the lowest loss value obtained by BA-GAT model is 0.11 which is much lower than 0.15 by CBAM-CNN and 0.29 by the basic CNN model. This faster convergence is due to the benefits of the combination of boundary-aware feature enhancement, attention-based feature extraction and graph-based context analysis techniques.

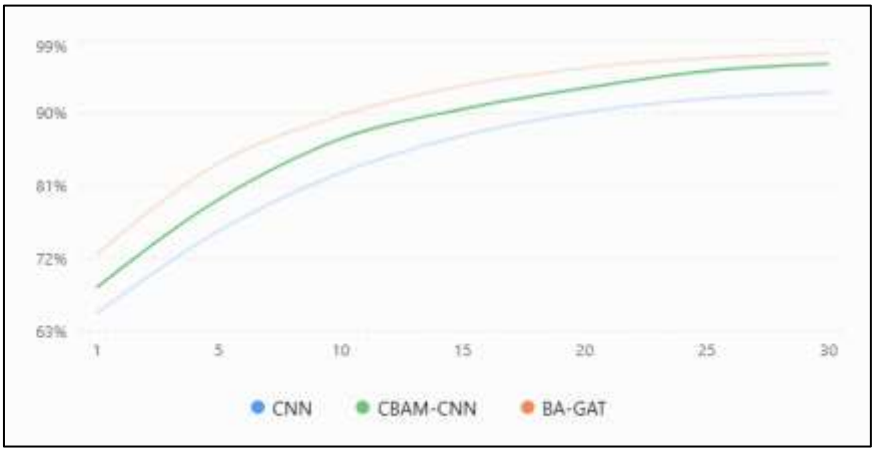

Figure 10. Validation Accuracy Comparison

Figure 10 shows the validation accuracy gained by the three models through 30 training epochs. As can be seen from the figure, the proposed BA-GAT outperforms both CNN and CBAM-CNN at each epoch with 97.4% final validation accuracy, whereas the CBAM-CNN achieves 96.1% final validation accuracy and the baseline CNN 92.6%. The more rapid increase curve during the first few epochs proves the fast convergence rate and efficiency of feature extraction. In addition, the higher accuracy indicates the better generalization ability of the proposed BA-GAT due to its boundary awareness, attention guidance, and graph structure.

Conclusion

An innovative Boundary-Aware Graph Attention Network (BA-GAT) to detect the license plate in a challenging traffic scenario was proposed. This new model incorporates boundary-aware feature enhancement, CBAM-based attention mechanism learning, multi-scale feature fusion, and graph attention network reasoning under one roof. The Boundary-Aware Feature Enhancement Module helps in enhancing the contour by focusing on the boundaries of the license plate, whereas the CBAM-based CNN model highlights informative channels and regions of the license plate. Besides that, the Multi-Scale Feature Fusion Network is helpful to handle the license plates of different scales. The effectiveness of the proposed model is supported by extensive experiments carried out using the Kaggle License Plate Dataset, UFPR-ALPR, and CCPD datasets. In the experiments, the BA-GAT framework has attained Precision of 97.8%, Recall of 96.9%, F1-score of 97.3% and mAP@0.5 of 97.4%, making it better than many other recent state-of-the-art techniques. The ablation study further proves the importance of each of the proposed modules in the framework, and statistical test verifies the significance and robustness of the results obtained. Though there is some rise in computational cost, the framework still supports real-time processing with an inference speed of 45 FPS. Thus, the proposed BA-GAT framework is a robust approach that can be applied for intelligent transportation systems. Future research efforts would include optimization of the model for lightweight architectures, transformer-based contextual learning, cross-domain adaptation, and edge computing applications.

Reference

[1] Z. Peng, Y. Gao, S. Mu, and X. Xu, "Toward Reliable License Plate Detection in Varied Contexts: Overcoming the Issue of Undersized Plate Annotations," IEEE Transactions on Intelligent Transportation Systems, vol. 25, no. 9, pp. 10689–10701, 2024, doi: 10.1109/TITS.2024.3424663.

[2] X. Wang, M. Wang, H. Guo, J. Li, X. Wang, and Y. Zhang, "License plate recognition system for complex scenarios based on improved YOLOv5s and LPRNet," Scientific Reports, vol. 15, Art. no. 34741, Oct. 2025, doi: 10.1038/s41598-025-18311-4.

[3] W. Xiong, L. Cao, D. Yan, Y. Jiang, G. Zhang, Y. Q. Wang, X. Huang, and L. Liu, "License plate recognition methodology in complex scenarios based on CSCM-YOLOv8 and CSM-LPRNet," PLOS ONE, vol. 21, no. 1, Art. no. e0339649, Jan. 2026, doi: 10.1371/journal.pone.0339649.

[4] Y. Wang, Z. Wang, Z. Li, J. Wang, and R. Hong, "License Plate Detection via Information Maximization," IEEE Transactions on Intelligent Transportation Systems, vol. 23, no. 9, pp. 15725–15737, Sept. 2022, doi: 10.1109/TITS.2021.3135015.

[5] Y. Xiong, Y.-B. Gao, J.-Q. Zhang, and Y. Wang, "License Plate Detection with Attention-Guided Dual Feature Pyramid Networks in Complex Environments," Electronics, vol. 11, no. 23, Art. no. 3967, Nov. 2022, doi: 10.3390/electronics11233967.

[6] L. Tao, S. Hong, Y. Lin, X. Chen, and J. Huang, "A Real-Time License Plate Detection and Recognition Model in Unconstrained Scenarios," Sensors, vol. 24, no. 9, Art. no. 2791, Apr. 2024, doi: 10.3390/s24092791.

[7] M.A. Chung, Y.-J. Lin, and C.-T. Lin, "YOLO-SLD: An Attention Mechanism-Improved YOLO for License Plate Detection," IEEE Access, vol. 12, pp. 100245–100258, 2024, doi: 10.1109/ACCESS.2024.3419587.

[8] F. Zhou, X. Lu, and C. Zheng, "License Plate Recognition Algorithms for Complex Situations," in Proceedings of the 2023 International Conference on Artificial Intelligence, Computer Vision and Machine Learning (AICVML), Dec. 2023, pp. 1–6, doi: 10.1145/3638884.3638892.

[9] Y. Zou, Y. Zhang, and J. Yan, "License plate detection and recognition based on YOLOv3 and ILPRNET," Signal, Image and Video Processing, vol. 16, no. 9, pp. 2187–2194, Mar. 2022, doi: 10.1007/s11760-021-01981-8.

[10] T.M. Seo and D.-J. Kang, "A Robust Layout-Independent License Plate Detection and Recognition Model Based on Attention Method," IEEE Access, vol. 10, pp. 26077–26089, 2022, doi: 10.1109/ACCESS.2022.3154678.

[11] X. Fan, "A Robust Multi-Oriented License Plate Detector and a Derived End-to-End License Plate Recognizer," IET Image Processing, vol. 20, no. 1, pp. 1–14, 2026, doi: 10.1049/IPR2.70272.

[12] S. Qin and S. Liu, "Towards end-to-end car license plate location and recognition in unconstrained scenarios," Neural Computing and Applications, vol. 33, no. 21, pp. 14585–14599, 2021, doi: 10.1007/S00521-021-06147-8.

[13] T. Tao, D. Dong, S. Huang, Y. Chen, and X. Li, "Object Detection-Based License Plate Localization and Recognition in Complex Environments," Transportation Research Record, vol. 2675, no. 3, pp. 470–482, 2021, doi: 10.1177/0361198120954202.
[14] L. Zhang, P. Wang, H. Li, Z. Li, C. Shen, and Y. Zhang, "A Robust Attentional Framework for License Plate Recognition in the Wild," IEEE Transactions on Intelligent Transportation Systems, vol. 22, no. 11, pp. 6967–6976, Nov. 2021, doi: 10.1109/TITS.2020.3000072.
[15] S.L. Chen, C. Yang, J.-W. Ma, Y.-J. Chen, and S.-Y. Wang, "Simultaneous End-to-End Vehicle and License Plate Detection With Multi-Branch Attention Neural Network," IEEE Transactions on Intelligent Transportation Systems, vol. 21, no. 9, pp. 3946–3954, Sept. 2020, doi: 10.1109/TITS.2019.2931791.
[16] J. Liang, G. Chen, Y. Wang, and X. Zhang, "EGSANet: Edge-Guided Sparse Attention Network for Improving License Plate Detection in the Wild," Applied Intelligence, vol. 52, no. 4, pp. 4021–4037, 2022, doi: 10.1007/S10489-021-02628-4.
[17] W. Weihong and T. Jiaoyang, "Research on License Plate Recognition Algorithms Based on Deep Learning in Complex Environment," IEEE Access, vol. 8, pp. 91661–91675, 2020, doi: 10.1109/ACCESS.2020.2994287.
[18] K. Bathirappan and J. Soundararajan, "A channel and spatial attentions mechanisms with CNN for enhanced license plate detection," Scientific Reports, vol. 16, Art. no. 17874, Apr. 2026, doi: 10.1038/s41598-026-46156-y.
[19] S. W. Lee, B. S. Choi, and S. K. Yoo, "A Vehicle License Plate Detection Scheme Using Spatial Attentions for Improving Detection Accuracy in Real-Road Situations," Journal of the Korea Society of Computer and Information, vol. 26, no. 1, pp. 93–100, Jan. 2021, doi: 10.9708/JKSCI.2021.26.01.093.
[20] K. Bathirappan and S. Jayanthy, "Video-Based License Plate Detection Using Deep Learning Boundary Filtering Method for License Plate Detection from Surveillance Images," Traitement du Signal, vol. 41, no. 6, pp. 3083–3095, 2024, doi: 10.18280/TS.410625.