CNN-GRU Based Fusion Architecture For Bengali License Plate Recognition With Explainable AI

Abstract—**Because of recent improvements to Bangladesh's roads and highways, Automatic Number Plate Recognition (ALPR) has become a crucial component. Numerous crimes, including kidnapping, failure to pay tolls, and harassment of women, occur on both public and private transportation. The security forces will be able to locate offenders more quickly with the earlier and more accurate detection of license plates. The authors of this research proposing a deep learning-based fusion model for ALPR that integrates CNN and GRU on the basis of these circumstances. A total of 4753 images from various Bangladeshi roads and highways have been collected for training, validation, and testing purposes. The dataset consists of three classes of data namely Private cars, Public buses, and Trucks where all the images are in RGB format. To get precise and reliable findings, a variety of preprocessing approaches have been applied. After passing the images to the proposed architecture all the necessary parameters have been fine-tuned that causes a lesser amount of trainable parameters and more accuracy. The research demonstrates that the suggested CNN-GRU based fusion architecture, with a 98.97% F1-score, outperforms the leading models. Both static photos and CCTV video material can be used to accomplish ALPR tasks with comparable efficiency. Later, Explainable Artificial Intelligence (XAI) model SHAP has been used in order to interpret the outstanding result with a region of features.**

Index Terms—**ALPR, fusion, CNN, GRU, SHAP, license-plate, recognition**

I. Introduction

In areas where complex patterns are difficult for humans to interpret, computer vision techniques are improving the comprehension of intelligent systems. In 2022, there were more than 6700 traffic collisions and accidents, and 45% of the vehicles involved were unfit for the road [1]. Approximately 5.35 million automobiles are currently on the road in Bangladesh. Unquestionably, the primary feature of a car that identifies a criminal is its license plate [2]. Numerous studies have been conducted to identify license plates written in English and other languages [3-6]. Lexical structure of Bengali language is ambiguous for machines to understand. Bengali license plate detection and recognition with discernment is more difficult for machines because of the ambiguity in the language [7,8]. We will be able to instantaneously track automobiles from anywhere by establishing an intelligent system. The intelligent system must be capable of reliably detecting license plates under a variety of circumstances. Long-running vehicles may have license plates that are positioned differently and may have numbers or letters that are illegible [9]. Even from video footage, it is difficult for robots to recognize license plates at dusk. Multiple automobiles frequently appear in photos or videos, necessitating the need for sophisticated algorithms to distinguish between them. To get precise findings from video footage, each frame must be processed appropriately [10]. Parking lots, in addition to various other locations like toll collection zones, can benefit from automatic license plate recognition (ALPR). ALPR can have a significant influence in areas where security is a top priority. Authors of this study focus on developing a novel, finely tuned Convolutional Neural Network-Gated Recurrent Unit (CNN-GRU) based deep learning architecture to detect and identify Bengali license plates in various ambient situations after taking all the relevant factors into consideration. Motivations and contributions of this research are stated in Section I. Section II recognizes all the recently attempted researches in this sector along with their advantages and disadvantages. Description of the gathered data and necessary methodologies are stated in Section III. All the necessary preprocessing techniques of the collected data and description of the proposed model are presented also in this mentioned section. The achieved results are analyzed and compared with the state-of-the-art architectures in Section IV. Finally the future direction of this research is discussed in Section V.

A. Research Motivation

In recent years, Bangladesh has undergone numerous advancements, including the building of numerous new highways. In order to prevent any unfavorable circumstances like unpaid tolls, car accidents, and parking allotments, security considerations must also be taken into account. To find and identify automobiles that conduct such crimes, an intelligent system must be implemented in those regions. The authors' main goal is to present a fine-tuned fusion model that can quickly detect license plates. Another issue to be concerned about in areas with problematic weather is accuracy. Finally, in order to comprehend the output data, the authors used an Explainable Artificial Intelligence (XAI) algorithm. This fusion model using static images and video is used to carry out ALPR tasks.

B. Contributions of this Research

Major contributions of this research can be encapsulated as follows:

I) Proposing a fine-tuned CNN-GRU based fusion model to perform ALPR tasks faster with improved accuracy using cutting-edge architectures.

Identify applicable funding agency here. If none, delete this.

II) The proposed model will prevent significant crimes from occurring on highways. Culprits in road accidents can be identified easily. The proposed model can be implemented inside a system in parking areas also. License plates can be detected from live footage also. The model performs these tasks with fewer trainable parameters than most deep learning architectures. Traffic controls will also be feasible with the application of the proposed model.

III) 4753 images are gathered by the authors where three classes of vehicles are included. The achieved result is interpreted with an XAI algorithm to understand how features are extracted.

II. Literature Review

An enormous amount of research has taken place for ALPR tasks in the English language, where the amount of research in Bengali license plate detection is insufficient. In [11], authors have proposed a system for detecting and recognizing vehicle number plates using computer vision techniques. The system comprises image preprocessing, plate detection, character segmentation, and recognition stages. The proposed method shows promising results on various datasets, making it a potential solution for automated traffic monitoring systems. Nevertheless, real data evaluation needs to be performed in this project. The authors present a computer vision-based method for detecting and recognizing license plates in parking lots. The proposed process involves image acquisition, preprocessing, plate localization, segmentation, and recognition. The results of [12] indicate high accuracy in detecting and recognizing license plates, making it a promising solution for automated parking management systems. However, where fine-tuning is missing, deep learning models need to be integrated appropriately. Convolutional Neural Network (CNN) is widely used to extract image features. Exploding gradient problems and overfitting are challenging factors while identifying images from CNN. Considering this, this research proposes a method for detecting and recognizing license plates using the YOLOv3 object detection framework and the ILPRNet recognition network. The proposed method [13] achieves high accuracy on various license plate datasets, making it a promising solution for automated traffic monitoring systems. The method can also recognize license plates from different countries.

In [14], the authors identified Bangla number plates using computer vision and convolutional neural networks. Trainable parameters are higher in the proposed model, which is a matter of concern. The researchers review methods for [15] detecting and recognizing car number plates using the particle swarm optimization algorithm. The review analyzes the strengths and weaknesses of different approaches and highlights the potential of combining particle swarm optimization with other techniques for improving the accuracy of car number plate recognition systems. Proposing a lightweight, fully convolutional neural network for license plate detection is the focus of [16]. The proposed method achieves high accuracy in detecting license plates with a few parameters, making it computationally efficient. The method is tested on various license plate datasets and shows promising results, making it a potential solution for real-time license plate detection applications. The method described in [17] proposed a feature extraction-based neural network for performing ALPR tasks. The proposed algorithm is based on the Gated recurrent unit (GRU), where fine-tuning was not performed. However, the sequence information of the images is not appropriately preserved in the proposed architecture. Researchers find an empirical study on different license plate detection and recognition methods. The study compares and evaluates the performance of various techniques on different datasets. The study's results [18] provide insights into the strengths and weaknesses of different methods, making it a valuable resource for researchers and practitioners working in the field of license plate recognition, where no novel architecture has been proposed. The study [19] compares and evaluates the performance of different techniques for license plate detection and recognition. The study uses various datasets to analyze the robustness and frailty of different methods.

The method achieves high accuracy in recognizing license plates in various lighting and weather conditions, making it a potential solution for automated traffic monitoring systems in Iraq. The proposed approach is tested on different datasets and shows promising results. Fusion models in Bengali ALPR tasks still need to be proposed, and these models are more flexible and durable in detection and recognition tasks. The CNN model automatically extracts significant features from images and videos. The problems of CNN, which are overfitting and exploding gradient problems, can be solved using GRU. Vanishing and exploding gradient problems are properly overcome by GRU architecture, where the memory requirements are very truncated. The proposed model can identify all research gaps in previous endeavors. Several evaluation metrics have been judged to justify the effectiveness of the proposed model.

Identifying all the research gaps in the previously mentioned research, authors intend to propose a deep learning based fusion model that can identify license plates from real world data.

III. Research Methodologies

The predominant target of the authors is to identify license plates from numerous vehicles. Three types of vehicles have been considered for detection and recognition purposes. To accomplish this task, data has been gathered from images collected from several busy highways in Bangladesh. In total, 4753 images have been gathered, where 2000 images are from private cars, 1753 images are from public buses, and 1000 images have been collected from truck license plates. After data collection, several preprocessing techniques were applied to achieve better performance from the deep learning-based fusion model. The dataset is split into three sets: the training, validation, and test sets. To avoid overfitting, early stopping, and dropout have been applied. The proposed fusion model is fine-tuned to achieve maximum results after applying the test set. Several evaluation metrics have been observed to understand the efficiency of the model. Later, license plates are detected from CCTV footage where the model has performed tremendously important. Fig 1 represents the whole workflow of the proposed methodology performed in this research.

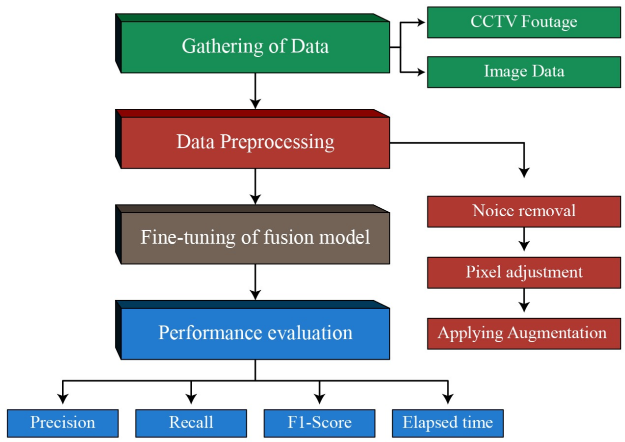

Fig. 1. Workflow diagram of this research

A. Description of Dataset

The dataset used in this research has been gathered by the authors manually. Numerous types of vehicles are usually found on the highways of Bangladesh, where public buses and private cars are the most common. According to [20], 60% of crimes have been performed in these vehicles, where kidnapping is the most common in private cars, and women's harassment is primarily done in public buses. Regarding road accidents, trucks are leading the percentage in such cases. Table 1 represents the detailed value counts of each class of data. Fig 2 shows some of the images gathered by the authors.

TABLE 1
AN OULOOK OF THE GATHERED DATASET OF THIS RESEARCH

Class name	Value counts
Private cars	2000
Public buses	1753
Trucks	1000

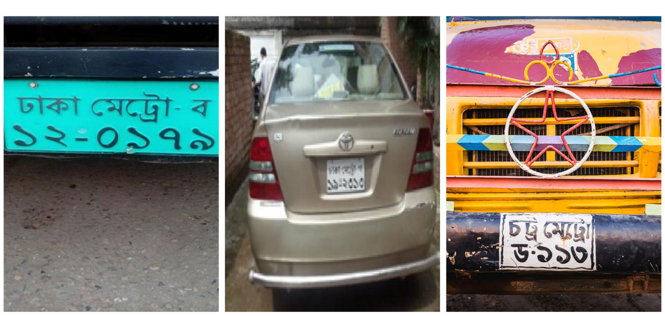

Fig. 2. Images from the utilized dataset in this research

B. Data Preprocessing

Preprocessing data is done to make it more accurate, reliable, and to get rid of discrepancies. The collection includes pictures with various heights and widths. All of the images have been downsized to 128 x 128 pixels, with RGB as the default picture format. The dataset is applied with the Gaussian filter to blur the photos. By obscuring unnecessary information, blurring enables the extraction of necessary features and the reduction of noise. A blurred image has fewer pixels and their associated values, which results in a lesser file size. The next step is data augmentation, which improves the model's capacity for generalization. Another important benefit of data augmentation is that it lowers the likelihood of overfitting. The conversion of images after applying gaussian filter is shown in Fig 3.

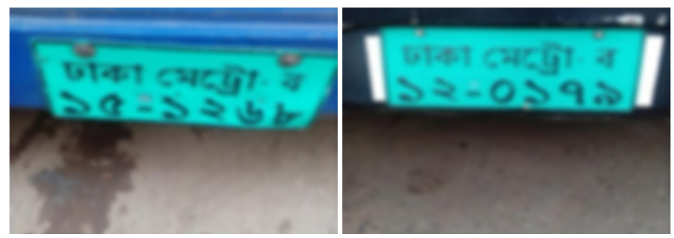

Fig. 3. Transformation of images after applying filter

A number of activities have been carried out to augment data. The operations utilized in the data augmentation process are shown in Table 2, along with their values.

TABLE 2
DATA AUGMENTATION DETAILS

Name of the operation	Value
Range of rotation	25
Range of zooming	0.23
Range of height	0.15
Range of width	0.15
Flip on horizontal side	True
Flip on vertical side	True
Mode	Nearest
Range of shearing	12

After that LabelBinarizer() is applied for converting all the classes into categorical values using the NumPy library. For performing all the calculations NumPy array is utilized. All the images are converted into NumPy arrays for ignoring ambiguity in calculations.

C. Description of the proposed model

The primary purpose of using CNN is to reduce the higher dimensionality where no information will be lost. For pattern recognition and image detection CNN is widely applied. The major disadvantage of CNN is to become overfitted at a specific time. Understanding the orientation of objects is another lacking of CNN. GRU allows an understanding of objects' orientation by preserving the images' sequence information. The applicability of GRU in this fusion model allows the proposed model to be faster and avoid the issue of overfitting.

In the proposed model, three CNN layers have been used. A Max Pooling layer follows each of the CNN layers. To accomplish this task, the Keras library is utilized, whereas in the backend, Tensorflow is used. Researchers have applied dropout after each of the convolution layers. The batch size is set to 128 in the training case. In the validation set, the batch size is 68, whereas, in the test set, the batch size is 128. Adam optimizer is used here as it is suitable for non static data. The input size of the images is 128 X 128, which are placed into the filters. For generating feature maps, filters are passed one by one. The stride value is set to 1 where no padding has been applied. In the Maxpool layer, the stride is set to 2, along with kernel size being set to 2 X2. After performing all the operations in the convolution layers, the images are passed to GRU units. The input and output gates of LSTM are replaced by a single Update gate in GRU. Preserving the sequence information of images is easier in GRU as it controls how much feature needs to be restored. GRU is faster, and the number of trainable parameters is much less. The integration of GRU in this model allows using fewer convolution layers which reduces the total amount of parameters. As a result, the model performs faster and precisely. Fine-tuning of the hyperparameters allows the model to process data more efficiently. After obtaining the image data from GRU, Flatten() is applied in the model. Flatten allows multidimensional data to be converted into one dimensional data. After that data is passed into two dense layers. ReLU activation function is used in order to introduce non-linearity in the data. In the final dense layer, Softmax activation function is applied because of multilabel data. Fig 4 shows the pictorial representation of the proposed model. Tuning of hyperparameters are detailed shown in Table 3 with their values.

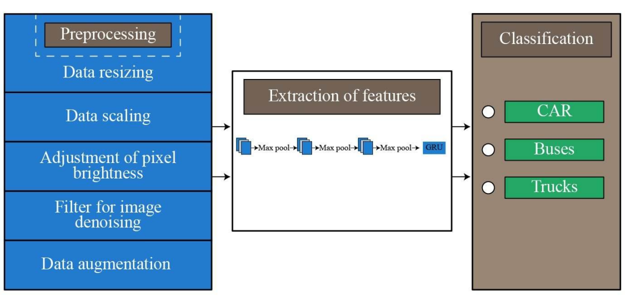

Fig. 4. Procedure of License plate detection

D. SHAP (Shapley Additive Explanations)

To interpret the results shown by the proposed model authors have focused to use SHAP XAI model. By giving each feature or input variable used by the model a relevance value, it offers a mechanism to explain how the model's prediction turned out. The significance value illustrates how each element contributes to the anticipated result. Shapley values are employed in the context of machine learning to equitably share each input variable's contribution to the final prediction provided by a model. SHAP is an effective technique for deciphering machine learning models, and it can help both technical and non-technical audiences have more faith in and understanding of these models.

TABLE 3
HYPERPARAMETER DETAILING OF THE PROPOSED ARCHITECTURE

Hyperparameter	Value
Number of convolution layer	3
Epoch	10
Stride	1
Kernel size	2
Activation function	ReLU and Softmax
GRU units	128
Loss function	Categorical cross-entropy
Recurrent dropout	2
Optimizer	Adam
Learning rate	0.00024
Regularizer	L1

IV. EXPERIMENTAL RESULT ANALYSIS

To understand the efficiency of the model, several evaluation metrics have been observed. The model is trained on the image dataset, where the dataset was divided into three categories. The model was run up to ten epochs in a training and validation set where the proposed deep-learning-based model has shown a 98.97% F1-score at epoch 8 in the training phase. Table 4 reflects the confusion matrix shown by the proposed model in the test set, where it has been observed the model has significantly performed well in identifying images of all categories. The model has shown maximum efficiency in terms of identifying Private cars. The reason behind this is the availability of many training images.

TABLE 4
CLASSIFICATION REPORT SHOWN BY THE PROPOSED ARCHITECTURE

Class Name	Precision	Recall	F1-score	Accuracy
Private cars	100%	99.97%	99.98%	99.14%
Public buses	98,65%	99.14%	99.96%	98.45%
Trucks	98.29%	98.17%	98.05%	98.18%

On the other hand in the case of Trucks the model has performed slightly down as the availability of the truck images are comparatively lower in the training and validation set. To understand the efficacy of the proposed model it has been compared with the previous studies [21-23]. To compare between the architecture the authors have focused on the F1-score and trainable parameters. Fig 5 shows, the model has outperformed the state-of-the art architectures in terms of F1-score. In [23] CNN model has been proposed where four convolutional layers have been fine-tuned by the researchers. The proposed fusion model outperforms this model in terms of F1-score. The integration of GRU has made it possible due to preserving sequence information.

Comparing the deep learning architectures in terms of trainable parameters is also another important evaluation criteria.

From Fig 6, it can be seen that the trainable parameters are lesser in the proposed model. The prime reasons for that

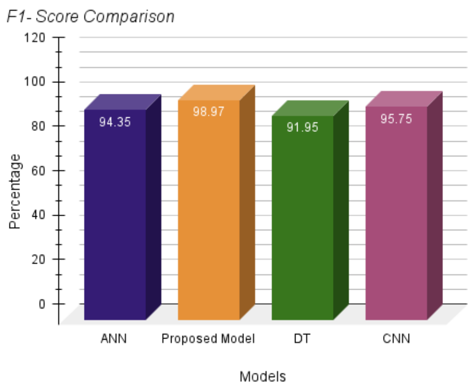

Fig. 5. F1-score comparison with the previous studies

is the availability of a lesser amount of convolution layers and fine-tuning all the essential hyperparameters. GRU has a lesser amount of trainable parameters that allows the model to identify license plates at a faster rate. While comparing with the elapsed time, it has been shown the model can detect license plates faster than other models because of a lesser amount of trainable parameters.

The next step is to identify license plates from live videos from CCTV. To implement this feature authors have used OpenCV. The frames from the videos are first captured then using the HAAR classifiers the videos are processed. In terms of detecting the Viola-Jones algorithm of OpenCV is utilized where multiple cascade layers have been integrated. Fig 7 shows the successful detection of license plates from live CCTV footage.

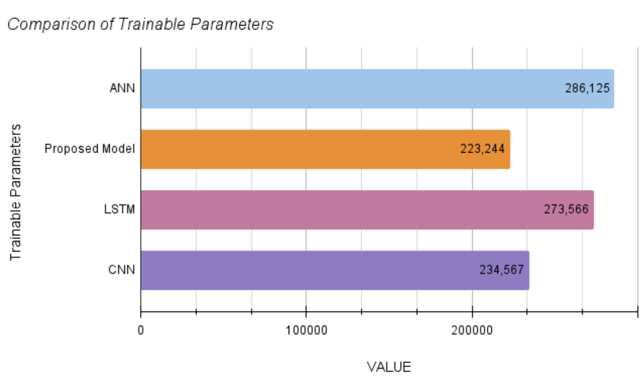

Fig. 6. Trainable parameter comparison with the previous studies

Furthermore, authors are focused to interpret the results by the SHAP XAI algorithm to understand what features are extracted from the images. Fig 8 shows that after providing the images into the XAI algorithm it extracts the region of license plates. The primary reason for this is to apply the Gaussian filter as it blurs all the unnecessary points from the images.

Fig. 7. Detection of license plates from CCTV footage videos

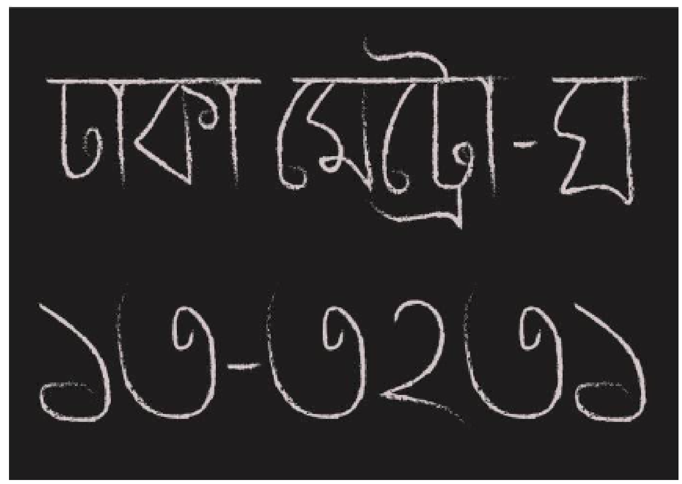

Fig. 8. Feature region detection using XAI

V. Conclusion

In this research, the primary objective is to identify license plates from moving vehicles along with images to avoid any kind of unwanted circumstances. A deep learning-based fusion architecture has been proposed to accelerate the whole process regarding the ALPR. The proposed architecture has shown exceptional performance and outperforms state-of-the-art architectures in numerous evaluations. The result is interpreted using the SHAP XAI algorithm, from which the features have been extracted. In the near future, authors will focus on optimizing the model so that detection can take place faster. Optimizing the hyperparameters will increase the efficiency of the proposed algorithm.

References

[1] Ahmed, R., & Mahmud, K. H. (2022). Potentiality of high-resolution topographic survey using unmanned aerial vehicle in Bangladesh. Remote Sensing Applications: Society and Environment, 26, 100729.

[2] Hasan, M. K., & Younos, T. B. (2020). Safety culture among Bangladeshi university students: A cross-sectional survey. Safety science, 131, 104922

[3] Omar, N., Sengur, A., & Al-Ali, S. G. S. (2020). Cascaded deep learning-based efficient approach for license plate detection and recognition. Expert Systems with Applications, 149, 113280.

[4] Youssef, A. R., Sayed, F. R., & Ali, A. A. (2022, July). A New Benchmark Dataset for Egyptian License Plate Detection and Recognition. In 2022 7th Asia-Pacific Conference on Intelligent Robot Systems (ACIRS) (pp. 106-111). IEEE.

[5] Chakraborty, S., Talukdar, M. B. U., Adib, M. Y. M., Mitra, S., & Alam, M. G. R. (2022, December). LSTM-ANN Based Price Hike Sentiment Analysis from Bangla Social Media Comments. In 2022 25th International Conference on Computer and Information Technology (ICCIT) (pp. 733-738). IEEE.

[6] Guo, Y., Li, Z., Wu, Y., & Xu, C. (2018). Evaluating factors affecting electric bike users' registration of license plate in China using Bayesian approach. Transportation research part F: traffic psychology and behaviour, 59, 212-221.

[7] Wadud, M. A. H., Kabir, M. M., Mridha, M. F., Ali, M. A., Hamid, M. A., & Monowar, M. M. (2022). How can we manage offensive text in social media-a text classification approach using LSTM-BOOST. International Journal of Information Management Data Insights, 2(2), 100095

[8] Sufian, A., Ghosh, A., Naskar, A., Sultana, F., Sil, J., & Rahman, M. H. (2022). Bdnet: bengali handwritten numeral digit recognition based on densely connected convolutional neural networks. Journal of King Saud University-Computer and Information Sciences, 34(6), 2610-2620.

[9] Susanto, S., Budiarjo, D. D., Hendrawan, A., & Pungkasanti, P. T. (2021). The implementation of intelligent systems in automating vehicle detection on the road. IAES International Journal of Artificial Intelligence, 10(3), 571.

[10] Kim, H. T., Lee, G. H., Park, J. S., & Yu, Y. S. (2012). Vehicle detection in tunnel using Gaussian mixture model and mathematical morphological processing. The Journal of the Korea institute of electronic communication sciences, 7(5), 967-974.

[11] Prabhakar, P., Anupama, P.,& Resmi, S. R. (2014, July). Automatic vehicle number plate detection and recognition. In 2014 International Conference on Control, Instrumentation, Communication and Computational Technologies (ICCICCT) (pp. 185-190). IEEE.

[12] Darapaneni, N., Mogeraya, K., Mandal, S., Narayanan, A., Siva, P., Paduri, A. R., ... & Agadi, P. M. (2020, October). Computer vision based license plate detection for automated vehicle parking management system. In 2020 11th IEEE Annual Ubiquitous Computing, Electronics & Mobile Communication Conference (UEMCON) (pp. 0800-0805). IEEE.

[13] Zou, Y., Zhang, Y., Yan, J., Jiang, X., Huang, T., Fan, H., & Cui, Z. (2022). License plate detection and recognition based on YOLOv3 and ILPRNET. Signal, Image and Video Processing, 16(2), 473-480.

[14] Suvon, M. N. I., Khan, R., & Ferdous, M. (2020, September). Real time bangla number plate recognition using computer vision and convolutional neural network. In 2020 IEEE 2nd International Conference on Artificial Intelligence in Engineering and Technology (IICAIET) (pp. 1-6). IEEE.

[15] Tadas, M. N. C. Review on Detection and Recognition of Car Number Plate using Particle Swarm Optimization Algorithm

[16] Xiang, H., Zhao, Y., Yuan, Y., Zhang, G., & Hu, X. (2019). Lightweight fully convolutional network for license plate detection. Optik, 178, 1185-1194.

[17] Aggarwal, A., Rani, A., & Kumar, M. (2020). A robust method to authenticate car license plates using segmentation and ROI based approach. Smart and Sustainable Built Environment, 9(4), 737-747.

[18] Rahman, M. J., Beauchemin, S. S., & Bauer, M. A. (2020). License plate detection and recognition: An empirical study. In Advances in Computer Vision: Proceedings of the 2019 Computer Vision Conference (CVC), Volume 1 1 (pp. 339-349). Springer International Publishing.

[19] Kumar, J. R., Sujatha, B., & Leelavathi, N. (2021, February). Automatic vehicle number plate recognition system using machine learning. In IOP Conference Series: Materials Science and Engineering (Vol. 1074, No. 1, p. 012012). IOP Publishing.

[20] Faroque, S., & South, N. (2022). Law-enforcement challenges, responses and collaborations concerning environmental crimes and harms in Bangladesh. International Journal of Offender Therapy and Comparative Criminology, 66(4), 389-406.

[21] Zhang, G., Wang, Y., & Wei, H. (2006). Artificial neural network method for length-based vehicle classification using single-loop outputs. Transportation research record, 1945(1), 100-108.

[22] Ashtari, A. H., Nordin, M. J., & Fathy, M. (2014). An Iranian license plate recognition system based on color features. IEEE transactions on intelligent transportation systems, 15(4), 1690-1705.

[23] Al Nasim, M. A., Chowdhury, A. I., Muna, J. N., & Shah, F. M. (2021, September). An automated approach for the recognition of bengali license plates. In 2021 International Conference on Electronics, Communications and Information Technology (ICECIT) (pp. 1-4). IEEE.