2022 26th International Conference on Pattern Recognition (ICPR)
August 21-25, 2022, Montréal, Québec, Canada

DPAM: A New Deep Parallel Attention Model for Multiple License Plate Number Recognition

Amish Kumar[1], Palaiahnakote Shivakumara[2], Pinaki Nath Chowdhury[1], Umapada Pal[1] and Cheng-Lin Liu[3]

[1]Computer Vision and Pattern Recognition Unit, Indian Statistical Institute, Kolkata, India.

[2]Facutlty of Computer Science and Information Technology, University of Malaya, Kuala Lumpur, Malaysia.

[3]National Laboratory of Pattern Recognition, Institute of Automation of Chinese Academy of Sciences, and also School of Artificial Intelligence, University of Chinese Academy, Beijing, China

Email: amishkumar562@gmail.com, shiva@um.edu.my, pinakinathc@gmail.com, umapada@isical.ac.in, liucl@nlpr.ia.ac.cn.

2022 26th International Conference on Pattern Recognition (ICPR) | 978-1-6654-9062-7/22/$31.00 ©2022 IEEE | DOI: 10.1109/ICPR56361.2022.9956285

Abstract— **License plate number recognition is challenging for complex scenes containing multiple vehicles of different types, shapes, distances etc. To recognize multiple license plate numbers in an image, we propose a new model, called Deep Parallel Attention Model (DPAM), which simultaneously extracts unique features at character levels. The proposed model exploits the observation that the combination of alphanumeric characters does not have correlation at semantic level for extracting the features. This led to the introduction of parallelism for feature extraction at character levels to make it efficient in terms of time to fit in a real time environment. To test the proposed model, we consider our own dataset consisting of Indian license plate numbers and other standard datasets to show the superiority of the proposed model over the existing methods in terms of recognition rate. Furthermore, the proposed method is tested on scene text dataset to show its ability to detect text in natural scene images.**

Keywords—Deep learning, Attention models, Visual Attention model, Parallel Attention models, License plate recognition.

I. INTRODUCTION

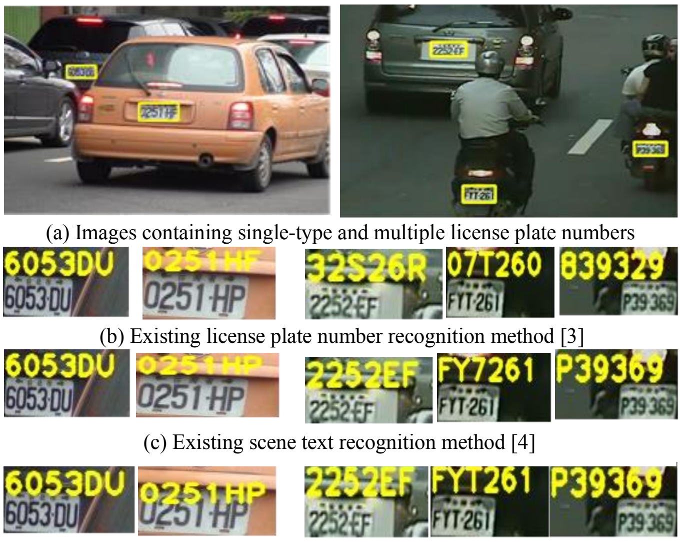
Despite of the availability of several deep learning-based models for license plate number recognition, it is still challenging to achieve reliable and stable recognition results which are essential for real time applications [1] such as automatic vehicle driving, which is part of an intelligent transportation system. The key challenges of such real time applications are (i) the effect of perspective distortion, (ii) Multiple license plates in the same spot, (iii) the effect of distance between camera and vehicles, (iv) the effect of low light, blur, low resolution, low contrast and other degradations due to illumination environment etc. The same challenges can be seen in the case of vehicle tracking and re-identification [2]. Therefore, obtaining stable and reliable recognition results for the above applications has been pursued by many researchers.

The license plate number recognition methods developed in the past address several challenges like arbitrary orientation, different scripts, low quality images etc. [1]. But efficient and accurate multiple license plate number recognition has not been addressed properly. The examples in Fig.1 show that for the input images containing single type vehicles (two license plate numbers are visible) and multiple vehicles (defocusing effect due to distance variation), the existing methods [3, 4], which are state-of-the-art methods for license plate number recognition and scene text recognition, respectively, do not report consistent results, as shown in Fig. 1(b) and Fig. 1(c). On the other hand, the proposed model recognizes license plate numbers accurately for all the images as shown in Fig. 1(d). This indicates that the existing method [3] is not robust to different situations especially for images with multiple vehicles and multiple license plate numbers. In the same way, the results of scene text recognition shown in Fig. 1(c) imply that the scene text recognition methods [4] are not effective enough. This is because most of the scene text recognition methods exploit the context and semantic correlation between characters for achieving the results. However, the alphanumeric characters in license plate numbers are rather semantically independent, so as to facilitate parallel feature extraction and recognition.

It is observed that for recognizing alphanumeric characters in license plate numbers, one should extract features at character level and then combine them to recognize the whole license plate number. This observation motivates us to propose a new Deep Parallel Attention Model (DPAM) that simultaneously extracts unique features for each character in the license plate number. This leads to time efficiency for license plate number recognition. In addition, since the DPAM works at character level, it can be used for scene text recognition successfully compared to existing recognition methods.

Therefore, the main contributions of the proposed model are as follows. (i) Addressing the challenges of multiple license plate number recognition in a single video frame is the first of its kind. (ii) Exploring Deep Parallel Attention Model (DPAM) for recognizing multiple license plate numbers is new compared to the existing methods. (iii) The proposed DPAM works well for the image of license plate number and scene text. (iv) Achieving better accuracy and time efficiency for recognizing multiple license plate numbers of multiple vehicles in the images are the two most important contributions of this work.

The rest of the paper is organized as follows. The review of the methods related to scene text recognition and license plate number recognition is discussed in Section II. Section III presents the Parallel Deep Attention Model for recognizing license plate numbers. Experimental results on our own license plate number dataset, standard datasets and benchmark of natural scene text datasets are discussed in Section IV. Section V summarizes the proposed model.



(a) Images containing single-type and multiple license plate numbers

(b) Existing license plate number recognition method [3]

(c) Existing scene text recognition method [4]

(d) Recognition results of the proposed model

Fig. 1. Recognition performance of the proposed and state-of-the-art methods for the images containing multiple vehicle license plate.

II. Related Work

We consider text recognition methods in natural scene images, license plate number recognition and the methods for recognizing both license plate number and natural scene text. for reviewing here.

The methods developed for natural scene text detection are ineffective for license plate number recognition [4-7]. The reason is that most of the recent deep learning-based methods use the degree of similarity between characters to extract semantic information. Since the words in the natural scene image provide clear meaning, semantic feature can be extracted to recognize the words. However, the alphanumeric characters in the license plate number are independent of semantics and hence the scene text recognition methods may not work for license plate number recognition.

There are methods for license plate recognition in the literature. Chen et al. [1] developed a model for license plate recognition based on hybrid learning system and convolutional neural network. Zherzdev et al. [3] used deep neural networks for license plate recognition. He et al. [8] proposed a robust method for automatic recognition of Chinese license plate in natural scenes. However, the scope of the method is limited to Chinese license plate numbers. Henry et al. [9] used character sequence detection for multinational license plate recognition. Huang et al. [10] proposed a method for multi-style vehicle license plate number recognition based on character recognition. Kong et al. [11] introduced a federated learning-based license plate recognition for 5G-enabled internet of vehicles. Liu et al. [12] used ABCNet for license plate recognition where the ABCNet is used for license plate detection and then CRNN has been used for recognition.

Sung et al. [13] proposed a method for real-time license plate recognition using YOLOv4. Zhang et al. [14] used a robust attentional framework for license plate recognition in the wild. Zhang et al. [15] proposed a unified framework for license plate detection, tracking and recognition. The scope of the method is limited to traffic videos. Hsu et al. [16] proposed a method for recognizing the license plate number based on the combination of hand-crafted features and deep neural network. Laroca et al. [17] developed a method for real time automatic license plate recognition using YOLO detector. Li et al. [18] used deep neural networks for developing end-to-end models for license plate detection and recognition. Li et al. [19] explored deep neural network and LSTM for reading license plate numbers. Wu et al. [20] proposed generative adversarial networks for recognizing license plate numbers. The reason to explore GAN is to increase the number of samples so that it improves the performance of the method.

Recently, Connectionist Temporal Classification based models [21], Hybrid learning system-based models [22], Automatic perspective alignment-based models [23], and Segmentation-free network-based models [24] are proposed for improving scene license plate number and scene text recognition. However, it is not clear whether the models work for the image having the license plate number of multiple vehicles in a single image.

In summary, the review of the above methods shows that none of the methods focuses on images containing multiple vehicles and license plate numbers. In addition, there is no discussion on the license plate number which suffers from defocus caused by long distance from the camera. Besides, the methods are confined to license plate number recognition.

There are methods that focus on both natural scene text recognition and license plate number recognition. For example, Shivakumara et al. [25] proposed a method for keywords spotting from video, natural scene images and license plate images. However, the scope of the method is limited to keyword spotting but not recognition. Khare et al. [26] used character segmentation-reconstruction approach for license plate and scene text recognition. Mokayed et al. [27] developed a method for license plate number and scene text detection based on the combination of Discrete cosine transform and phase congruency model.

Overall, the review on the methods of license plate number recognition and the methods that focus on both license plate number and natural scene text show that none of the approaches considers images containing multiple license plate numbers, which suffer from defocusing caused by long distance from camera. In addition, most of the methods either focus on achieving better accuracy or time efficiency but not both. Thus, this work presents a new Deep Parallel Attention Model (DPAM) for multiple license plate number recognition which shows better accuracy as well as computational efficient.

III. Proposed Method

It is noted from the license plate number that there is no strong semantic correlation between alphanumeric characters, rather each character has its own meaning. In addition, there are methods to achieve either accuracy or time efficiency but not both. This observation motivated us to explore parallel attention

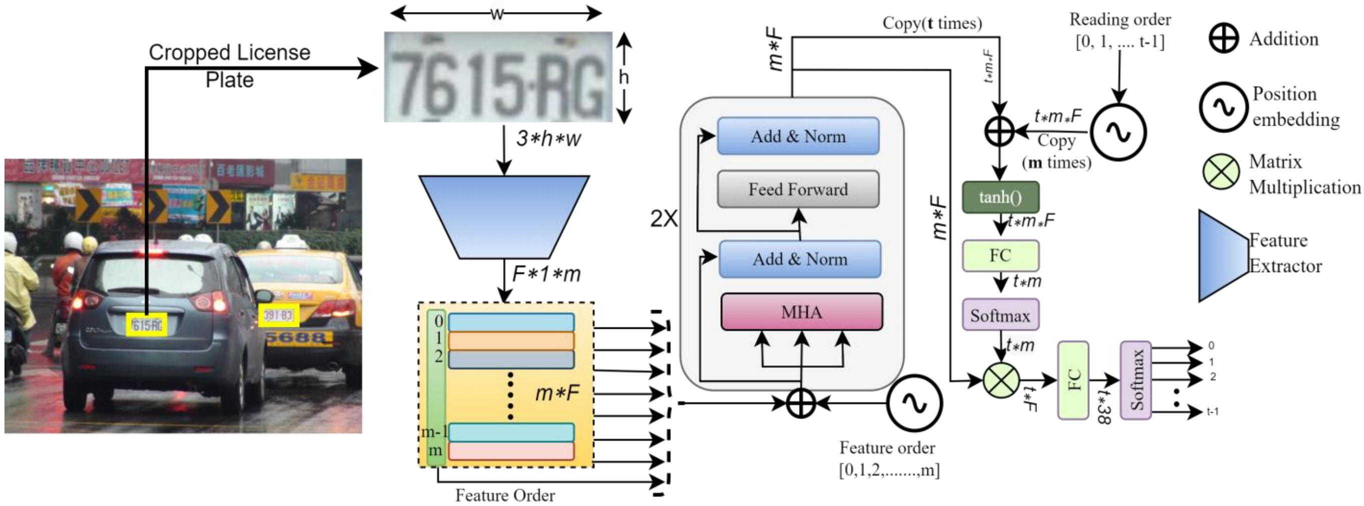


Fig. 2. Detailed architecture of the proposed methodology. Here *w* and *h* represent the width and height of input license plate. Features extractor block extracts *F* features with *m* sequence, and pass-through transformer unit, which generates (m×F) features. Attention mechanism is then applied to generate separate features corresponding to each target sequence. Next, it passed through a fully connected layer, which convers *F* features into *C* classes.

model to extract local information at character level such that the method can recognize the whole license plate number accurately with a smaller number of computations. Inspired by the attention models, which are good in extracting dominant information in the images, we explore the same for license plate recognition in this work.

The proposed DPAM is an end-to-end trainable network, which extracts features parallelly at character levels using transformer-based decoder. As a result, the proposed model achieves time efficiency compared to existing models [14, 28, 29, 30]. The proposed architecture consists of two parts: Feature extractor and Transformer-based parallel decoders are presented in subsequent sub-sections, (A) and (B), respectively. Since the main objective of the proposed work is to develop a model for recognizing multiple license plate numbers in the images, the license plate regions are cropped manually, and cropped license plate number images are fed to the proposed model for recognition as shown in Fig. 2. Let the feature extractor network extract features F from cropped license plate X and then the extracted features are fed to decoders to generate enhanced aligned features G. The detailed structure of the proposed architecture can be seen in Fig. 2.

A. Encoder: Feature Extractor

For feature extraction, the proposed model explores a ResNet50 architecture as shown in Fig. 3. The details of ResNet50 are as follows. The ResNet50 architecture consists of 5 blocks and each block contains one convolutional layer of 32 filters, which outputs 32, 64, 128, 256 and 512 features for the input image. The feature extractor block generates 512 features with dimension $\left(\frac{w}{4} \times \frac{h}{32}\right)$, where w and h represents the width and height of the input license plate. To ease the implementation, the height of input data, X is normalized to $(1 \times m)$, where, $m = \frac{w}{4}$, which can be considered as a sequence vector of 512 features. The whole feature extractor involves massive parallelism for extracting features to achieve an efficiency inspired by the model [31] where a parallel visual attention model is used for scene text recognition. This is the advantage of the proposed model. For feature extraction, we consider the input image, X of dimension (100×32) and it is fed to feature extractor which outputs 512 feature vector of dimensions (25×1).

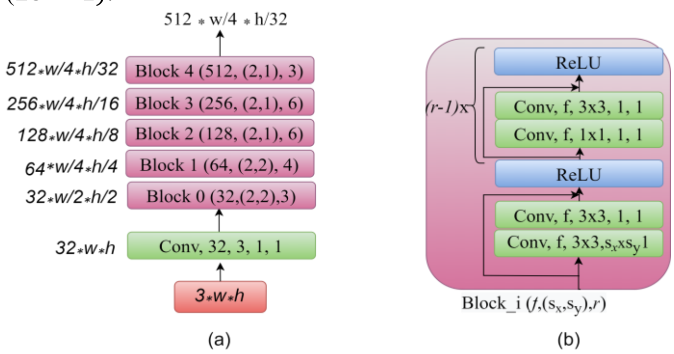


Fig. 3. (a) Basic block diagram of the utilized ResNet architecture (b) detailed representation of each block like Block 0, Block 1, etc. *w* and *h* represent the width and height of input license plate, *f* represents the no of filters utilized in each convolutional layers, s_x and s_y represent the stride in x and y directions.

B. Decoder: Feature Attention and Decoding

The extracted features from ResNet50 are passed through transformer unit [32] followed by parallel attention and a fully connected layer for license plate number recognition. Transformer unit is a stack of two transformers with 8 number of heads and the output of feed forward multi-head attention of 512 dimensions. Attention mechanism in multi-head is defined by key (K), value (V) and query (Q), which measures the similarity between Q and K, and generates attention feature map as defined in Equation (1)

$$Attention(Q, K, V) = softmax\left(\frac{K^T Q}{\sqrt{d_k}}\right)V \tag{1}$$

Where d_k represents the dimension of key (k). In the case of parallel visual attention key (K)-value (V) paired is considered as (f_{ij}, f_{ij}), where f_{ij} represents the feature value at an index (i, j), $0 \leq i \leq 25, 0 \leq j \leq 512$ and reading order is used as a query (Q). Parallel attention map is generated as defined in Equation (2).

$$\alpha_{t,ij} = softmax\big(\tanh\big(R_0 f_0(q_t) + R_v f_{ij}\big)\big) \quad (2)$$

Where q_t is character reading order query, f_0 is embedding function and f_{ij} is generated feature value from transformer unit. Whereas R_0 and R_v can be defined as a repeat vector which converts $f_0(q_t)$ from $(t \times 512)$dimension to $(t \times 25 \times 512)$ and f_{ij} from (25×512) dimension to $(t \times 25 \times 512)$ and aligned feature for corresponding index t is generated as defined in Equation (3).

$$G_t = \sum_{\forall ij} \alpha_{t,ij} v_{ij} \quad (3)$$

The aligned visual feature G for all reading index results in a feature matrix of dimension $(l \times 512)$, where l represents the maximum length of decoded labels on license plate. Furthermore, G is fed to a Fully Connected (FC) layers for character recognition. Since the aim of the proposed work is to recognize license plate number in English, on can expect 0-9 number and A to Z alphabets and hence the architecture uses 36 nodes FC layers for license plate number recognition. We have selected maximum length of decoded string from license plate as 15 for Indian dataset and 9 for AOLP dataset therefore value of l will vary according to our choice

IV. Experimental Results

In this study, we created our own dataset for evaluating the proposed method because none of the standard datasets provided Indian license plate number images for license plate recognition. It is noted that Indian license plate numbers are too complex compared to other datasets because there are no rigid constraints for license plate number creation in India. As a result, it is expected that license plate number will have a large variation. To test the proposed method's robustness, we consider a standard dataset called AOLP [15] for experimentation. On top of that, since the proposed method does not depend on the semantic relationship between the characters, it can work for any text dataset. To validate the effectiveness of the proposed method, the method is also tested on standard dataset of natural scene text detection [35, 36]. For recognition experiments, we manually crop the license plate region from the input images because the scope of the proposed work is to recognize the license plate numbers of multiple vehicles.

A. Dataset Creation and Evaluation

Our dataset, called Indian Vehicles Licence Plate Dataset (IVLPD), has 2000 images. Out of 2000 images, 1600 images are collected by capturing different locations of West Bengal State of India using mobile cameras. Since Indian license plate numbers do not follow regular pattern, we can expect the following challenges for Indian license plate recognition. (i) Color variations in both foreground (characters) and background, (ii) Non-uniform size of the license plates, (iii) Multiple fonts and font size of the characters, (iv) There is no fixed number of characters in the license plate, which varies from length of 8 to 11 characters, (v) License plate can have single line and double line, (vi) Location of license plate changes according to vehicle type. Therefore, IVLPD is a complex dataset compared to other datasets for license plat recognition.

To test the objectiveness of the proposed method, we consider the standard dataset called, Application Oriented License Plate (AOLP) [19], which consists of 2049 images of Taiwan license plate numbers. The dataset can be divided further into the license plates of Access Control (AC), Traffic Law Enforcement (LE) and Road Patrol (RP) applications. AC, LE and RP sets contain 681, 757 and 611 license plates, respectively. In this work, the number of license plates of AC and LE are considered for training and the number of license plate numbers of RP are considered for testing.

To measure the performance of the proposed model, we use standard metric, namely, accuracy as defined in Equation (4).

$$Accuracy = \frac{\#C}{C} \quad (4)$$

where $\#C$ and C represent the number of correctly recognised licence plate number and total number of license plate numbers, respectively. When all the characters in a license plate image are recognized by the proposed method, it is considered as one count. Otherwise, the license plate number is discarded for computing accuracy.

B. Training and Implementation

In this work, we consider a 7:3 ratio for training and testing for all the experiments. The proposed network is implemented using PyTorch open-source deep learning library in conda environment. The hardware specification of the system includes Intel(R) Core (TM) i7-7700K CPU @ 4.20GHz 4.20 GHz, 32 GB ram with NVIDIA TITAN X (Pascals) GPU with a12 GB capacity. For the network training, we use Adam optimizer with Cross Entropy Loss functions without any pre-trained weight. The parameters used during the architecture training are summarized as follows: Batch size: 16, learning rate: 0.0005, epochs no: 300 and maximum length of decoded character: 15. Note that there is no specific training like tuning, pre-processing, masking and post operation during training considered for experimentation.

C. Ablation Study

It is noted from the proposed methodology that the ResNet50 architecture used as a feature extractor is vital for accurate license plate number recognition. To validate the effectiveness of the ResNet50, we conduct recognition experiments using IVLPD and AOLP datasets for different versions of ResNet50. The results reported in Table I show that as the number of layers increases, the accuracy increases gradually. Therefore, the ResNet50 is the best compared to other license plate number recognition architectures. Since the objective of the proposed work is to achieve time efficiency, we choose ResNet50 without increasing the layers further. This is because as the number of layers increases, computations also increase. Therefore, for this study, RestNet50 is the best to achieve both accuracy as well as computational time.

Table I. Analyzing the effectiveness of ResNet50 used as a feature extractor for license plate number recognition using IVLPD and AOLP datasets.

Different architecture	ResNet18+	ResNet34+	ResNet50+
Accuracy	86.45	86.71	**87.24**

D. Experiments on License Plate Number Recognition

Qualitative results of the proposed method for IVLPD and AOLP [33] dataset is shown in Fig. 4 and Fig. 5, respectively, where for the license plate image with different complexities, including orientation, distortion, contrast variation, multiple vehicles, font-font size variations etc, the proposed model recognizes license plate number correctly. This shows that the proposed method is independent of vehicle types and invariant to the above-mentioned challenges. In the same way, the results on standard dataset show that the proposed method is effective, and it works well for different datasets.

Fig. 4. Qualitative results of the proposed model for IVLPD dataset. The length of license plate number in terms of number of characters is varying from one vehicle to another (8-10 characters). For recognition, the license plates are cropped from the input images.

Recognition results of the proposed method on IVLPD dataset are reported in Table II, where it can be seen that the license plates of different lengths in terms of number of characters, the proposed method almost achieves similar accuracy. Therefore, we can assert that the proposed method is independent of the number of characters in the license plate images. In addition, it is also confirmed that the proposed method does not depend on semantic relationship between characters for recognition.

The same conclusions can be drawn from the quantitative results of the proposed and existing methods reported in Table III on AOLP dataset. It is observed from Table III that the proposed model is the best at Accuracy compared to the existing methods. Since the existing methods listed in Table III used the same AOLP dataset for evaluation in terms of accuracy, we use the same results for comparative study with the results of the proposed method in this work. The reason for the poor results of the existing methods is that the methods suffer from their own constraints and limitations for recognition. On the other hand, the proposed deep parallel attention model can cope with the challenges posed by adverse effects as we discussed earlier. Hence the proposed model reports the best results for license plate recognition on AOLP dataset.

To show that the proposed method is not only accurate for license plate recognition but also time efficient, we compare the speed in terms of milliseconds with the existing method as the results are reported in Table IV. It is noted from Table IV that the proposed method considers 1.9 MS compared to 3 MS of the method [3] and 11 MS of the method [16]. This makes sense because the proposed method involves massive parallelism for feature extraction. Thus, the results on license plate recognition and speed show that the proposed method is superior to existing methods in terms of accuracy and efficiency.

Fig. 5. Qualitative results of the proposed method for AOLP dataset. The license plates are cropped from the images for recognition.

Table II. Performance of the proposed method in terms of accuracy on IVLPD dataset.

No of characters	Number of license plates	Accuracy
8	26	65.00
9	270	76.34
10	302	75.24
For all	598	75.04

Table III. Performance of the proposed and existing methods on AOLP dataset.

Methods	Li et al. [18]	Li et al. [17]	Wu et al. [19]	Hsu et al. [15]	Laroca et al. [16]	Proposed
Accuracy	88.4	83.7	91.0	88.38	85.45	**91.2**

Table IV. Time efficiency of the proposed method for license plate recognition in milliseconds.

Methods	Zherzdev et al. [3]	R. Laroca et al [16]	Proposed
Speed	3ms	11ms	**1.9ms**

E. Experiments on Natural Scene Text Images

Qualitative results of the proposed method for recognizing text in natural scene images are shown in Fig. 6. The proposed method recognizes scene text images of different complexities. Although the method is developed for license plate recognition, the model works well for natural scene text recognition. This makes sense because the proposed method does not depend on semantic relationship between characters, unlike existing methods that use context information to achieve the results. As a result, the characters of the license plate images and text in the natural scene images are the same for recognition.

To test the proposed method quantitatively for scene text recognition, we conduct recognition experiments on the standard dataset of IIIT5k [34] and Street View Text (SVT) [35] datasets. For training, we use samples of Synth90k [36], SVT and IIIT5K datasets and for testing, 3000 and 567 samples respectively from IIIT5k and SVT are used for testing. In total, we fed 3567 test samples to the proposed model for recognition. The proposed model reports 75.3% Accuracy (recognition rate). Compared to the accuracy of license plate number recognition presented in the previous section, the accuracy is slightly low. This is acceptable because the proposed work's main objective is to develop an accurate and efficient system for license plate number recognition but not for scene text recognition. In addition, the proposed method does not use any linguistic and semantic knowledge of text information for recognition in contrast to existing scene text detection methods, which explore semantic information as the key features for achieving the results. Therefore, achieving 75.3% accuracy for 3567 samples of different datasets is remarkable and it indicates the proposed model is capable of handling challenges of scene text recognition.

Fig. 6. Sample recognition results of the proposed method for the scene text images chosen from the standard datasets of SVT, IIIT5k. Red color indicates wrong recognition results.

Overall, the results on our dataset, on standard license plate number dataset, and on standard natural scene text dataset show that the proposed method is robust to different datasets, it has generalization ability and furthermore, it achieves both accuracy and time efficiency.

Fig. 7. Limitation of the proposed model for license plate number recognition. The red color indicates incorrect recognition results generated by the proposed model.

Sometimes, when the license plate images are affected by severe blur, partial occlusion and other types of degradations, the proposed method may not recognize license plate number correctly as shown in the sample images in Fig. 7 (red color indicates wrong recognition results). The recognition results for the images in the first row in Fig. 7 show that when there is no clear distinction between foreground (characters) and background, the proposed feature extractor fails to extract distinct features for license plate recognition. Therefore, the proposed model does not perform well for such images. Similarly, for the images shown in the second row in Fig. 7, the proposed method fails to recognize the license plate number correctly because of the blur and partial occlusion. However, compared to the results of the images in the first row, the images in the second row are better. This shows that the proposed model can cope with the challenges of blur, noise and distortion but not with the challenges of overlapping foreground and background colors. To address this challenge, it is necessary to extract the spatial relationship between pixels at character level and context information between characters. This is beyond the scope of the proposed work and hence we can conclude that there is a scope for the improvement by extending the proposed approach in future.

V. Conclusion and Future Work

In this work, we have proposed a new model for recognizing the license plate numbers affected by different complexities. To achieve this, we have explored a model called Deep Parallel Attention Model (DPAM) for multiple license plate number recognition. Unlike existing methods that focus on images of single vehicle or multiple vehicles of the same type, the proposed model focuses on the images of multiple types of vehicles and license plate numbers for recognition. Based on the observation that the alphanumeric characters in the license plate do not have semantic relationship, we exploit the same to introduce parallel processing to extract the features at character level. This makes the model generic for any license plate number including text of natural scene images. Experimental results demonstrate that the proposed method is robust to different datasets, types of text and finally it is more efficient than existing methods. However, when the license plate image suffers from severe degradations, such as blurry and too low contrast, the performance of the method degrades. Therefore, we plan to develop an end-to-end model with the help of image enhancement and recognition in the future.

Acknowledgement

This project was funded by Ministry of Higher Education of Malaysia for the generous grant Fundamental Research Grant Scheme (FRGS) with code number FRGS/1/2020/ICT02/UM/02/4. The work received partial support from Technology Innovation Hub, Indian Statistical Institute, Kolkata, India.

References

[1] C. L. P. Chen and B. Wang, "Random-positioned license plate recognition using hybrid broad learning system and convolutional networks", IEEE Trans. ITS, pp 444-456, 2022.

[2] Z. Zheng, T. Ruan, Y. Wei, Y. Yang and T. Mei, "VehicleNet: Learning robust visual representation for vehicle re-identification", IEEE Trans. MM, pp 2683-2693, 2021.

[3] S. Zherzdev and A. Gruzdev, "LPRNet: License plate recognition via deep neural networks", arXiv preprint arXiv:1806.10447, 2018.

[4] D. Yu, X. Li, C. Zhang, T. Liu, J. Han, J. Liu and E. Ding, "Towards accurate scene text recognition with semantic reasoning networks", In Proc. CVPR, pp. 12113-12122, 2020.

[5] G. Meng, T. Dai, S. Wu, B. Chen, J. Lu, Y. Jiang and S. T. Xia, "Sample-aware data augmentation for scene text recognition", In Proc. ICPR, pp 9378-9385, 2021.

[6] L. Nandanwar, P. Shivakumara, R. Raghavendra, T. Lu, U. Pal. D. Lopresti and N. B. Anuar, "Local gradient difference features for classification of 2D-3D natural scene text images", In Proc. ICPR, pp 1112-1119, 2021.

[7] C. Zhang, W. Ding, G. Peng, F. Fu and W. Wang, "Street view text recogniton with deep learning for urban scene understanding in intelligent transportation system", pp 4727-4743, 2021.

[8] M. X. He and A. P. Hao, "Rboust automatic recogniton of Chiense license plates in natural scenes", IEEE Access, pp 173804-173814, 2020.

[9] C. Henry, S .Y. Ahn and S. W. Lee, "Multinational license plate recognition using generalized character sequence detection", IEEE Access, pp 35185-35199, 2020.

[10] Q. Huang, Z. Cai and T. Lan, "A new approach for character recognition of multi-style vehicle license plates", IEEE Trans. MM, pp 3768-3777, 2021.

[11] X. Kong, K. Wang, N. Hou, Z, Hao, G. Shen, X. Chen and F. Xia, "A federated learning-based license plate recognition scheme for 5G-enabled internt of vehicles", IEEE Trans. Inustrial Informatics, pp 8523-8530, 2021.

[12] Y. Liu, J. Yan and Y. Xiang, "Research on license plate recognition algorithm based on ABCNet", In Proc. ICISCAE, pp 465-469, 2020.

[13] J. Y. Sung, S. B. Yu and S. H. Park, "Real-time automatic plate recognition system using YOLOv4", In Proc. ICCE-Asia, 2020.

[14] L. Zhang, P. Wang, H. Li. Z. Li, C. Shen and Y. Zhang, "A robust attentinal framework for license plate recognition in the wild", IEEE Trans. ITS, pp 6867-6976, 2021.

[15] G. S. Hsu, J. C. Chen and Y. Z. Chung, "Application-oriented license plate recognition", IEEE Trans. VT, pp 552-561, 2012.

[16] R. Laroca, E. Severo, L. A. Zanlorensi, L. S. Oliveira, G. R. Gonçalves, W. R. Schwartz and D. Menotti, D, "A robust real-time automatic license plate recognition based on the YOLO detector", In Proc. IJCNN, pp 1-10, 2018.

[17] H. Li, P. Wang and C. Shen, "Towards end-to-end car license plates detection and recognition with deep neural networks", CoRR abs/1709.08828, 2017.

[18] H. Li and C. Shen, "Reading car license plates using deep convolutional neural networks and LSTMs", arXiv preprint arXiv:1601.05610, 2016.

[19] C. Wu, S. Xu, G. Song andS. Zhang, "How many labeled license plates are needed?", In Proc. PRCV, pp 334-346, 2018.

[20] C. Zhang, Q. Wang and X. Li, "V-LPDR: Towards a unified framework for license plate detection, tracking, and recognition in real-world traffic videos", Neurocomputing, pp 189-206, 2021.

[21] Y. Wang and J.E. Ha, "Scene text recognition with multi-decoders", In Proc. ICCAS, pp 1523-1528, 2021.

[22] C. L. P. Chen and B. Wang, "Random-positioned license plate recogniton using hybrid broad learning system and convolutional networks", IEEE Trans. Intelligent Transportation System, pp 444-456, 2022.

[23] H. Xui, X. D. Zhou, Z. Li. L. Liu, C. Li and Y. Shi, "EILPR: Toward end-to-end irregular license pate recognitin based on automatic perspective alignment", IEEE Trans. Intelligent Transportation System, pp 2586-2595, 2022.

[24] X. Fan and W. Zhao, "Improving robustness of license plates automatic recognition in natural scenes", IEEE Trans. Intelligent Transportation System, 2022. doi: 10.1109/TITS.2022.3151475.

[25] P. Shivakumara, S. Roy, H. A. Jalab, R. W. Ibrahim, U. Pal, V. Khare and A. W. B. A. Wahab, "Fractional means based method for multi-oriented keyword spotting in video/scene/license plate images", Expert Systems with Applications, pp 1-19, 2019.

[26] V. Khare, P. Shivakumara, C. S. Chan, T. Lu, L. K. Meng, H. H. Woon and M. Blumenstein", A novel character segmentation-reconstrution approach for license plate recognition", Expert Systems with Applications, pp 219-239, 2019.

[27] H. Mokayed, P. Shivakumara, H. H. Woon, M. Kankanhalli, T. Lu and U. Pal, "A new DCT-PCM method for license plate number detection I drone images", Pattern Recognition Letters, pp 45-53, 2021.

[28] A. Graves, S. Fernández, F. Gomez and J. Schmidhuber, "Connectionist temporal classification: labelling unsegmented sequence data with recurrent neural networks", In Proc. ICML, pp. 369-376, 2006.

[29] H. Li and C. Shen, "Reading car license plates using deep convolutional neural networks and LSTMs", arXiv preprint arXiv:1601.05610 ,2016.

[30] S. M. Silva and C. R. Jung, "License plate detection and recognition in unconstrained scenarios", In Proc, ECCV, pp. 580-596, 2018.

[31] A. Vaswani, N. Shazeer, N. Parmar, J. Uszkoreit, L. Jones, A. N. Gomez, L. Kaiser and I. Polosukhin, I., "Attention is all you need", Advances in neural information processing systems, pp 5998-6008, 2017.

[32] K. He, X. Zhang, R. Ren and J. Sun, "Deep residual learning for image recognition", In Proc. CVPR, pp. 770-778, 2016.

[33] Hsu, G.S., Chen, J.C. and Chung, Y.Z., 2012. Application-oriented license plate recognition. *IEEE transactions on vehicular technology*, *62*(2), pp.552-561.

[34] A. Mishra, K. Alahari and C.V. Jawahar, "Scene text recognition using higher order language priors", In Proc. BMCV, 2012.

[35] K. Wang, B. Babenko and S. Belongie, "End-to-end scene text recognition", In Proc. ICCV, pp 1457-1464, 2011.

[36] A. Gupta, A. Vedaldi and A. Zisserman, "Synthetic data for text localization in natural images", In Proc. CVPR, pp 2315-2324, 2016.

[37] Raghunandan K. Srinivas, Palaiahnakote Shivakumara, Hamid Abdullah Jalab, Rabha W. Ibrahim, G. Hemantha Kumar, Umapada Pal, Tong Lu:Riesz Fractional Based Model for Enhancing License Plate Detection and Recognition. IEEE Trans. Circuits Syst. Video Technol. 28(9), pp. 2276-2288, 2018.