This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/ACCESS.2020.2993008, IEEE Access

IEEE Access
Multidisciplinary ⋮ Rapid Review ⋮ Open Access Journal

Date of publication xxxx 00, 0000, date of current version xxxx 00, 0000.

Digital Object Identifier 10.1109/ACCESS.2017.Doi Number

Automatic Vehicle License Plate Recognition using Optimal K-Means with Convolutional Neural Network for Intelligent Transportation Systems

Irina Valeryevna Pustokhina[1], Denis Alexandrovich Pustokhin[2], Joel J. P. C. Rodrigues[3], Deepak Gupta[4], Ashish Khanna[4], K.Shankar[5], IEEE Member, Changho Seo[6*], Gyanendra Prasad Joshi[7]*

[1] Plekhanov Russian University of Economics, Moscow, Russian Federation, Email: pustohina.iv@rea.ru

[2] State University of Management, Moscow, Russian Federation. Email: dpustokhin@yandex.ru

[3] Federal University of Piauí (UFPI), Teresina - PI, Brazil. Email: joeljr@ieee.org

[4] Department of Computer Science & Engineering, Maharaja Agrasen Institute of Technology, Delhi, India. Email: deepakgupta@mait.ac.in , ashishkhanna@mait.ac.in

[5] Department of Computer Applications, Alagappa University, Karaikudi, India. Email: drkshankar@ieee.org

[6] Department of Convergence Science, Kongju National University, Gongju 32588, South Korea: chseo@kongju.ac.kr

[7] Department of Computer Science and Engineering, Sejong University, Seoul, South Korea. Email: joshi@sejong.ac.kr

*Corresponding authors: Gyanendra Prasad Joshi (e-mail: joshi@sejong.ac.kr); Changho Seo (email: chseo@kongju.ac.kr).

ABSTRACT Due to recent developments of highway and the increased utilization of vehicles, significant interest has been paid on the latest, effective, and precise intelligent transportation system (ITS). The process of identifying particular objects in an image plays a crucial part in the fields of computer vision or digital image processing. Vehicle license plate recognition (VLPR) process is difficult because of variations in viewpoint, shape, color, multiple formats and non-uniform illumination conditions while acquiring images. This paper presents effective deep learning based VLPR model using optimal K-means clustering based segmentation and convolutional neural network (CNN) based recognition, called OKM-CNN model. The proposed OKM-CNN model operates on three main stages, namely license plate (LP) detection, segmentation using (OKM) clustering technique, license plate number recognition using the CNN model. For LP detection and localization, Improved Bernsen Algorithm (IBA) and Connected component analysis (CCA) models are employed. An extensive experimental investigation takes place using three datasets namely Stanford Cars, FZU Cars and HumAIn 2019 Challenge dataset. The attained simulation outcome ensured the effective performance of the OKM-CNN model over the compared methods in a considerable way.

INDEX TERMS Intelligent transportation system, CNN, Traffic Management, Vehicle license plate recognition, Character recognition.

I. INTRODUCTION

Recently, the speedy development in ITS and Graphical Processing Units (GPU), Automatic Vehicle license plate recognition (VLPR) has acquired major attention from several research domains. LPR is considered to be more significant in various applications like unmanned parking fields, security management of unattended regions, as well as traffic safety administration [1]. Unfortunately, these operations are very tedious, since the distinct format of plates as well as dynamic outdoor illumination constraints at the time of image acquisition, namely, background, brightness, vehicle's speed, and distance among the camera and vehicles. Hence, many techniques can be operated at restricted rules like permanent illumination, lower vehicle speed, allocated paths, and static background.

A common method for LPR has been comprised with 4 blocks such as acquiring a vehicle image, LP localization, segmentation, character classification and standardization, as well as character analysis. The process of location procedure is considered to be more complex throughout the mechanism, due to the fact that it has a direct impact on accuracy and efficiency of the consecutive procedure. Therefore, it becomes more complex to resolve the issues as the illumination conditions and some other tedious backgrounds existed. Many developers have been presented with massive approaches to placing the LP, like the edge prediction model, line sensitive filters used for extracting plate regions, window scheme, and arithmetic morphology approach [2]. Though the predefined models are capable of processing the LP's position, it is comprised of formidable demerits like sensitivity to illumination, higher computation time, and absence versatility while applying for diverse platform.

Character segmentation was attained under the application of projection [3], morphology, relaxation labeling, as well as linked components. Additionally, it has been composed with maximum count of character analyzing methodologies as reported in [4], Baye's classification, Artificial Neural Networks (ANN), fuzzy c-means (FCM), support vector machine (SVM), Markov chain model, and K-nearest neighbor (kNN) classifier. Even though these methods are able to compute the task of placing an LP segmentation and analysis, several models perform only on individual line character segmentation and 2 kinds of character analysis were established namely, English and numerals. The highly tedious LP recognition techniques and various types of character analysis have not been explained.

In the case of LPR study, users are inspired by LPR which works on LP localization as well as segmentation and character recognition. Therefore, effective placement of LP system is meticulous and extensive dissection of the single part requires the way of performing a task in combined manner images. This paper presents an effective DL based VLPR model using optimal K-means clustering based segmentation and CNN based recognition, called OKM-CNN model. The proposed OKM-CNN model operates on three main stages. In the first stage, LP localization and detection process take place using

Improved Bernsen algorithm (IBA) and Connected component analysis (CCA) model. Subsequently, optimal K-means (OKM) clustering with Krill Herd (KH) algorithm gets executed to segment the LP image and finally, characters in LP recognition takes place using CNN model. An extensive experimental investigation takes place using three datasets namely Stanford Cars, FZU Cars and HumAIn 2019 Challenge dataset.

The upcoming portions of the study are formulated as follows. Section 2 elaborates the existing works related to LP detection, character segmentation, and recognition. Section 3 introduces the presented OKM-CNN model. Section 4 provides the experimentation part and the conclusion is drawn n section 5.

II. RELATED WORKS

In this section, the survey of the existing works takes place in a three-folded way, namely LP detection, character segmentation, and recognition.

A. LP Detection Techniques

Lin *et al.* [5] devised a new technique for detecting LPs, which is mainly applicable to predict the vehicles and finds LP for vehicles for minimizing the false positives on plate prediction. The scope of character recognition value has been increased for blurred and noisy images under the application of CNN. Ullah *et al.* [6] concentrated on predicting the LP based on the mathematical morphological attributes. The newly presented model is capable of working for every English LP that differs in shapes and structure. Omran *et al.* [7] projected an automated LP analyzing mechanism by applying Optical Character Recognition (OCR) along with templates mapping, and correlation technique for plate detection.

Babu *et al.* [8] implied a set of 4 major steps for LP analysis. Initially, in preprocessing, images were gathered by cameras, modified proper brightness, eliminate the noise, and transformed to a grayscale image. Then, the edges in an image are used in extracting the LP location. Furthermore, characters were segmented in LP. Consequently, it has been used with template matching technique to analyze every character in LP image. Rana *et al.* [9] have defined numerous detecting approaches for LP and relate the function on same metrics. Also, it is applied with signature analysis along with CCA, and Euclidean distance transformation. Hence, the above-mentioned models are used to attain better accuracy and failure because of improper illumination as well as blurring.

B. Character Segmentation

Liang *et al.* [10] employed a novel wavelet Laplacian technique for randomly segmenting the characters from the video text lines. It searches the zero crossing points to explore space among words as well as characters. The function of this model attains minimum performance while an image is filled with noisy background. Also, few approaches were projected for character segmentation present in LP images. Khare *et al.* [11] developed a novel sharpness-relied model for segmenting characters of LP images. It encounters the gradient vector and accurateness for segmenting operation. Therefore, the approach is

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/ACCESS.2020.2993008, IEEE Access

referred to be more responsive in improving the point selection as well as blur existence.

Kim *et al.* [12] deployed an effective model for LP detection at different illumination platforms. Then it has been utilized with binarization and super pixel paradigm to segment the characters in LP. The model mainly aims on a particular reason; however, does not exist in many reasons. Meanwhile, Dhar *et al.* [13] projected a system deployment for LP recognition under the application of edge detection and CNN. It has been consumed with character segmentation in the form of pre-processing phase for LP analysis. In case of character segmentation, the newly developed technique finds an edge prediction, morphological task and regional features. Thus, it is more effective for images with elegant backgrounds, but the images were not affected by the comprised complexities.

Ingole and Gundre [14] employed character feature-relied vehicle LP prediction and recognition. Initially, the model performs the process of segmenting characters from LP regions. In case of character segmentation, the technique presents vertical as well as horizontal projection profile-centric features. The presented features could be ineffective for images with difficult backdrops. Radchenko *et al.* [15] applied a character segmentation technique according to the CCA. The CCA performs quite-well if the input image has been binarized in absence of character shapes and presence among the characters. Hence, the images which are composed with complex backgrounds are very difficult to deploy a binarization model which divides foreground and background data. Finally, it has been revealed that a greater number of approaches have been tried to resolve the issues with less illumination effects, however, does not enclose with alternate complications like blur, touching and difficult backgrounds. Also, no models were existed with redevelopment in for character segmentation from LP images.

C. Character Recognition

Raghunandan *et al.* [16] projected a Riesz fractional-centric approach to enhance the LP detection and recognition. It is used to report the reasons that affect LP discovery and identification. According to the experimental outcome, it is pointed that improvement in LP images might enhance the recognizing simulation outcome which is unsuitable for real-time domains. Al-Shemarry *et al.* [17] applied an ensemble of Adaboost cascades for 3L-LBPs classification in LP recognition particularly for low quality images. In this model, it has been identified with texture attributed which depends upon the LBP task and applies a classification model for LP analysis from the images influenced by diverse factors. Therefore, the function of a technique is based on learning and count of modelled instances. Additionally, the value has been restricted in text prediction; however, the recognition is not terminated from the developed approach. Text detection is simpler when compared with recognition, because of the detection process which does not acquire complete shapes of characters. In recent times, a robust ability as well as discriminating energy of DL methods, few technologies have found with various DL approach for LP recognition.

Dong *et al.* [18] implied a CNN-oriented technique for automated LP recognition. It is explored with R-CNN for the purpose of LP analysis. Bulan *et al.* [19] established a segmentation and annotation-free LP recognition along with deep localization as well as error identification. In line with this, it has been found that CNNs detects a collection of candidate regions. Followed by, it extracts the false positive from candidate regions according to the robust CNNs.

Yang *et al.* [20] introduced a Chinese vehicle LP recognition with the application of kernel-based extreme learning machines (ELM) with deep convolutional features. Then, it explores the integration of CNN and ELM which is applied in LP recognition. These features were identified from DL modules which perform well in case of having massive number of predetermined samples. But, it becomes more difficult to select the predefined instance which shows the feasible differences from LP recognition, only for the images influenced by several adverse factors. Also, DL method is constrained with the shortcomings like parameter optimization for various databases as well as retaining the reliability of DNN. It is clear from the predefined definition that, there exists a gap from earlier models and the recent requirements. This observation has acquired higher attention to project a novel approach for LP recognition with no dependency of classification models and more count of labelled samples, as present in the previous approaches.

3. THE PROPOSED OKM-CNN MODEL

The overall working principle of the OKM-CNN model is depicted in Fig. 1. At the earlier stage, the LP localization and recognition process take place using IBA and CCA model. Followed by, characters in the LP are segmented by the use of OKM algorithm, which has incorporated the K-means clustering with KH algorithm. Lastly, the CNN based character recognition process takes place to recognize the characters present in the LP.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/ACCESS.2020.2993008, IEEE Access

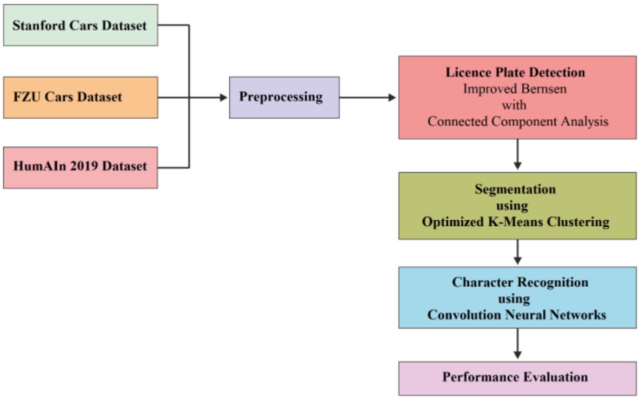

FIGURE 1. The Overall Framework of Proposed OKM-CNN Method.

A. LP Localization and Recognition Process

The dissemination of diverse locations on a LP image varies as the rule of a plate and the impact of the lighting platform. As the binary model which has global threshold is not capable of producing convinced outcome from adaptive local binary technique has been employed. The local binary techniques are referred to as an image would be classified into $m \times n$ blocks, and every block is computed using binary model. In this study, 2 local binary methodologies are applied namely, local Otsu and enhanced BA. One of the conventional binary approaches that can be employed on all sub-blocks. Therefore, the function of Otsu is a contingent on illumination constraints that has drastic variation. I order to overcome the irregular illumination barrier, especially for shadow images, a new binary technique such that the enhanced BA.

IBA based LP Localization

As the computed LP is attained by different illumination cases and difficult backgrounds, shadows, irregular illumination could not be eliminated from LP. Therefore, shadow or uneven illumination removal is an essential procedure in this presented model. The binary results states that, the conventional binary models are not capable to eliminate the shadow, and LP could not predict and segmented. This problem can be resolved, a novel binary technique was presented, and that refers the IBA.

In BA, the target and background have to be divided using a histogram that shows a bi-modal pattern of images, global thresholding binary techniques attain optimal results like Otsu and average grayscale value. Therefore, the real-world images are composed with noise and alternate causes, the image histogram could not produce bimodal pattern. At this point, conventional binary models are not capable to accomplish desired final outcome. Local threshold approaches are typically employed in finding critical interference of an image, namely the Bernsen algorithm (BA) and Niblack technique. Generally, the optimal result of local binary model, the BA is a best solution to resolve the issue of poor illumination.

Assume that $f(u,v)$ is a gray value of point (u,v). The block is composed with a point (u,v) and size is $(2w+1)\times(2w+1)$. The threshold $T(u,v)$ of $f(u,v)$ is determined by

$$T_1(u,v) = \frac{\underset{-w \le z,l \le w}{\mathrm{mx}} f(u+l,v+z) + \underset{-w \le z,l \le w}{\mathrm{mn}} f(u+l,v+z)}{2}. \quad (1)$$

Then, the binary image can be obtained by,

$$b(u,v) = \begin{cases} 0, & \text{if} f(u,v) < T_1(u,v) \\ 255, & \text{else.} \end{cases} \quad (2)$$

Elimination of noise as well as conservation of these characters is highly significant in this mechanism. Assume that $\hat{f}(u,v)$ is a gray value reached through Gaussian filter, σ implies the scale of Gaussian filter, and z and l signifies the metrics of a window.

- Determine the threshold $T_1(u,v)$ of $f(u,v)$ based on (1).
- Develop the Gaussian filter for window $s = (2w+1) \times (2w+1)$ of (u,v), i.e.,

$$\hat{f}(u,v) = \frac{1}{(2w+1)^2}\sum_{u,v\in S} f(u,v) \times \exp\{-\frac{1}{2}[(\frac{u}{\sigma})^2 + (\frac{v}{\sigma})^2] \quad (3)$$

- Determine the threshold $T_2(u,v)$ of $f(u,v)\Lambda$ as

$$T_2(u,v) = \frac{\mathrm{mx}\, f(u+l,v+z) + \mathrm{mn}\, f(u+l,v+z)\, \Lambda\Lambda}{2} \quad (4)$$

- Get a binary image by

$$b(u,v) = \begin{cases} 0, & \text{if} f(u,v) < \beta\big((1-\alpha)T_1(u,v) + \alpha T_2(u,v)\big); \beta \in (0,1) \\ 255, \text{else} & \end{cases} \quad (5)$$

  where α refers the variable to modify the trade-off among the BA with Gaussian filter ($\alpha \in [0,1$ When α is similar to 0, the projected model is a BA. When α is similar to 1, the deployed technique is BA using a Gaussian filter. Under the application of proper α, the shadow has to be avoided in an effectively, as a result the characters could be identified profitably.
- Use the Wiener filtering technique to eliminate the noise.

This model could accomplish a better binary outcome and pointed that the strokes of Kana characters were maintained with some pixels.

LP Detection

After computing the LP using binary technique, the system emerges into LP prediction. The simulation outcome of a prediction is the key to follow the overall performance. The CCA is a popular model in image processing which screens an image and labels the pixels to units based on the connection between the pixels that has been linked with one another. After the determination of groups is computed, every pixel is designated with a value based on the component. According to the LP data, 2 kinds of LP were projected as given in the following:

- Black characters in a white backdrop

- White characters in a black backdrop

Here, 2 detection models are used: initially, the recognition of a white frame using CCA, and second one was detecting black characters by applying CCA. When the LP type is unknown, then LP detection models tends to produce 2 kinds of patterns:

Procedure 1: LP location with a frame. Candidate frames were analyzed on the basis of advanced knowledge of LP. Under the application of procedure 1, few candidate areas could be forecasted.

Procedure 2: LP location with no frame. If the of the LP could not be predicted, then procedure 2 has been applied. It is used in predicting the plate by massive extraction. When the LP with white characters or black backdrop is unpredicted, the reverse image has been examined.

At the initial stage, to reach the linked components of same size, few parameters were applied according to characters' previous data like pixel value of linked component, width is higher than 10, height is maximum than 20, ratio of height to width is lower than 2.5 or maximum than 1.5, etc. Similarly, the maintained with linked units that are of same size. Then, to eliminate few non-character connected components, alternate limitations were generated by character position on LP namely distance among 2 characters angle.

Generally, procedure 1 requires lower time when compared with procedure 2. The application of procedure 1 or procedure 2, candidate frames might provide maximum frame simultaneously. In case of candidate frames accomplished by procedure 1, CCA is applied to obtain massive numeric characters and process the penetration time at midpoint of the LP to attain few candidate frames. The number of penetration times implies the number modifies from black pixels to white pixels with a midline. The residual frames are not capable to provide final decision. The candidate frames can be forwarded by the given steps and discriminated true LP according to analyzed results.

B. OKM based Character Segmentation Process

Neutrosophic Image

Neutrosophic analysis will be applied to evaluate the indeterminacy or uncertainty of an image dataset. A membership sets are comprised with a certain degree of falsity (F), indeterminacy (I), and truth (TR). The Membership Functions (MF) is applied for mapping the input images to (NS) form, that tends to produce (NS) image (A_{NS}) . Thus, in the image, the pixel $A(u,v)$ is described as $A_{NS}(u,v) =$ A $(t,p,f) = \{TR(u,v), I(u,v), F(u,v)\}$ for (NS) domain providing the true, indeterminate, as well as false falling into a brighter pixel set. Consider that $A(u,v)$ illustrate the intensity measures of pixel (u,v), then $\overline{A}(u,v)$ denotes the local mean value, the MF has been expressed as .

$$TR(u,v) = \frac{\overline{A}(u,v) - \overline{A}}{\overline{A} - \overline{A}}, \tag{6}$$

$$\overline{A}(u,v) = \frac{1}{b*b} \sum_{l=u-\frac{b}{2}}^{u+\frac{b}{2}} \sum_{m=q-\frac{b}{2}}^{v+\frac{b}{2}} A(l,m), \tag{7}$$

$$I(u,v) = \frac{\delta(u,v) - \delta_{mn}}{\delta_{mx} - \delta_{mn}}, \tag{8}$$

$$\delta(u,v) = abs((A(u,v) - \overline{A}(u,v) \tag{9}$$

$$F(u,v) = 1 - TR(u,v), \tag{10}$$

The absolute value is shown by $\delta(u,v)$ where, the measure of $I(u,v)$ used to calculate the indeterminacy of $A_{NS}(u,v)$. The NS image entropy involves totalling of 3 sets as given in the following:

$$E_{NS} = E_{TR} + E_I + E_F, \tag{11}$$

$$E_{TR} = -\sum_{p=mn\{TR\}}^{mx\{TR\}} P_{TR}(p) \ln\left(P_{TR}(p)\right), \tag{12}$$

$$E_F = -\sum_{p=mn\{F\}}^{mx\{F\}} P_F(p) \ln\left(P_F(p)\right), \tag{13}$$

$$E_I = -\sum_{p=mn\{I\}}^{mx\{I\}} P_I(p) \ln\left(P_I(p)\right), \tag{14}$$

The 3 entropy subsets are depicted by (E_{TR}, E_I, E_F).The probabilities of elements present in 3 MF are demonstrated by $(P_{TR}(p), P_I(p), P_F(p))$. Additionally, the deviations of F and TR develop the elements distribution and entropy of I to create F as well as TR associated with I.

The local mean task for a grayscale image A is defined by:

$$\overline{A}(u,v) = \frac{1}{b*b} \sum_{l=u-\frac{b}{2}}^{u+\frac{b}{2}} \sum_{m=q-\frac{b}{2}}^{v+\frac{b}{2}} A(l,m), \tag{15}$$

The α-mean for neutrosophic image A_{NS} is

$$\overline{A}(\alpha) = A(\overline{TR}(\alpha), \overline{I}(\alpha), \overline{F}(\alpha)), \tag{16}$$

where $\overline{TR}(\alpha), \overline{I}(\alpha)$ and $\overline{F}(\alpha)$ are shown as given below:

$$\overline{TR}(\alpha) = \begin{cases} TR, & I < \alpha \\ \overline{TR}_{\alpha}, & I \geq \alpha \end{cases}$$

$$\overline{F}(\alpha) = \begin{cases} F, & I < \alpha \\ \overline{F}, & I \geq \alpha \end{cases} \tag{17}$$

$$\overline{TR}(\alpha)(u,v) = \frac{1}{b*b} \sum_{l=u-\frac{b}{2}}^{u+\frac{b}{2}} \sum_{m=q-\frac{b}{2}}^{b+\frac{b}{2}} TR(l,m), \tag{18}$$

$$\overline{F}(\alpha)(u,v) = \frac{1}{b*b} \sum_{l=u-\frac{b}{2}}^{u+\frac{b}{2}} \sum_{m=q-\frac{b}{2}}^{v+\frac{b}{2}} F(l,m), \tag{19}$$

$$\overline{I}_{\alpha}(u,v) = \frac{\overline{\delta}(u,v) - \overline{\delta}}{\overline{\delta} - \overline{\delta}}, \tag{20}$$

$$\overline{\delta}(u,v) = abs(\overline{TR}(u,v) - \overline{\overline{TR}}(u,v) \tag{21}$$

$$\overline{\overline{TR}}(u,v) = \frac{1}{b*b} \sum_{l=u-\frac{b}{2}}^{u+\frac{b}{2}} \sum_{m=q-\frac{b}{2}}^{v+\frac{b}{2}} \overline{TR}(l,m), \tag{22}$$

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/ACCESS.2020.2993008, IEEE Access

where ‘b’ is the size of average filter. The entropy of I is improved by obtaining a uniform dissemination of elements, where α value has been optimized under the application of KH algorithm.

C. Optimization in (α - Mean) Using KH algorithm

The optimum value of (α) is used computed by applying KH algorithm. KH algorithm is stimulated from the herding nature of krills, which depends upon the individual outcome of krills. The swarm of krills that hunts for foodstuff and commune with swarm members of the technique. A collection of 3 motions in which the position of a krills has been presented as follows.

- Endeavor persuaded by other krills,
- Foraging act and
- Physical diffusion.

KH is treated as the Lagrangian model as given in Eq (23)

$$\frac{dU_p}{dt} = N_p + F_p + D_p \tag{23}$$

where N_p indicates the movement of other krills, Fp is the searching movement and D_p is the physical distribution.

The optimization Fitness Function (FF) is jaccard (JAC), that is said to be statistical values which estimate the union ∪’ as well as intersection ∩’ operators of 2 sets. The fitness JACs is expressed by:

$$JAC(f, q) = \frac{A_{rf} \cap A_{rq}}{A_{rf} \cup A_{rq}}, \tag{24}$$

where, A_{rf} denotes the computerized segment area by applying presented ONKM model, and A_{rq} refers the ground truth region.The ONKM (LP) character segmentation method to attain (α optimal). In order to attain the higher JAC coefficient using KH,
Then the application of Eq. (25)

$$F(f, q) = 1 - JAC(f, q), \tag{25}$$

D. k - means clustering using optimized (α - mean)

K-means is defined as clustering method that consolidates the objects into K groups. The arithmetical function establishes the k-means:

$$O = \sum_{q=1}^{q} \sum_{p=1}^{d_q} \| W_p - Z_q \|, \tag{26}$$

where, q implies the overall cluster count, Z_q denotes the center of q^{th} cluster, and d_q represents number of pixels of q^{th} cluster. From k-means algorithm, it is essential to reduce O by using the given constraint:

$$Z_q = \frac{1}{d_q} \sum_{W_p \in C_q} W_p, \tag{27}$$

where the dataset, $W = \{w_p, p = 1,2, \dots, n\}, w_p$ signifies sample in d-dimensional space and $C = \{C_1, C_2, \dots, C_q\}$ refers the partition $W = \cup_{p=1}^{q} C_p$. Once the optimization is completed α, (T and I) subsets is named as new value while the consequence of indeterminacy is computed as:

$$W(u, v) = \begin{cases} TR\,(u, v), I(u, v) < \alpha_{optimal} \\ TR_\alpha(u, v), I(u, v) \geq \alpha_{optimal} \end{cases} \tag{28}$$

Then, it is applied with k-means clustering for optimal NS to a subset (TR).

E. CNN based Recognition Process

CNN is a familiar DL model used to recognize the characters present in the segmented LPs. The CNN includes a set of conv, pooling and fully connected (FC) layers as shown in Fig. 2. They are employed for the construction of a CNN model with diverse number of blocks, addition or deletion of blocks.

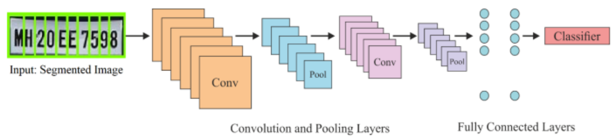

FIGURE 2. The Architecture of CNN for character recognition.

Conv layer

It varies from a NN in such a way that not each pixel is linked to the subsequent layer with weights and biases. However, the whole image gets partitioned into smaller parts and weights/biases are employed into it. They are called as filters or a kernel which are convoluted with each smaller region in the input image and offers the feature maps as output. The filters are considered as simpler ‘features’ which have undergone searching in the input image in the conv layer. The parameters needed to perform the convolution function is less since the similar filter get traversal over the whole image over the whole image in an individual feature. The filter count, local region size, stride, and padding are the hyper parameters of convolution layer. With respect to the sizes and genres of the input image, these hyper parameters undergo tuning for achieving effective performance.

Pooling layer

For reducing the spatial dimension of the image and parameter count, pooling layer is utilized to minimize the processing cost. It carries out a predefined function over the input, therefore no parameters are devised. Various kinds of pooling layer exist namely average pooling, stochastic pooling, max pooling. Here, Max pooling is utilized where the nxn window is slided over the input with a stride value s and for every location, the maximum value in the nxn region is considered and consequently input size gets reduced. It offers translational invariance so that a minor difference in location can also be recognized.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/ACCESS.2020.2993008, IEEE Access

FC layer

Here, the flattened output of the final pooling layer is provided as input to a FC layer. It acts as a classical NN where each neuron of the earlier layer is linked to the current layer. Therefore, the parameter count in the layer is higher than the conv layered. It is linked to the output layer, commonly known as classifier.

Activation function

Various activation functions are employed along different architectural models of CNN. The nonlinear activation function called ReLU, LReLU, PReLU, and Swish are available. The nonlinear activation function assisted to speed up the training process. In this paper, ReLUs function is found to be effective over the other ones.

4. PERFORMANCE VALIDATION

A. Implementation details

The presented OKM-CNN method is accelerated by applying a PC i5, 8th generation, 16 GBRAM. The OKM-CNN approach is processed under the application of Python language with TensorFlow, Pillow, OpenCV and PyTesseract. Fig. 3 depicts the sample test images. For experimental analysis, a collection of 3 dataset has been employed. The primary Stanford Cars dataset is comprised with a set of 297 model cars with 43615 images. The second Stanford Cars dataset encloses a set of 196 model cars with 16,185 images. At last, images from HumAIn 2019 Challenge dataset were utilized.

FIGURE 3. a) Images in Stanford Cars dataset b) Images in FZU Cars dataset c) HumAIn 2019 Challenge.

B. Results analysis

Fig. 4 depicts the visualization result of OKM-CNN system. It is implied that the OKM-CNN approach analyzes the LP number on the images. Fig. 4a displays the input image, the LP detected image is shown in Fig. 4b. Then, the segmented characters in the LP is displayed in Fig. 4c and the characters are recognized in Fig. 4d.

FIGURE 4. a) Original Image b) Detected Image c) Segmented Image d) Recognized Image.

Table 1 and Fig. 5 Demonstrates the relative analyzes of the OKM-CNN with traditional approaches with respect to precision, recall, F-score and mAP on applied FZU Cars Dataset. The table measures denoted that the ZF technique leads to poor LP detection outcomes by reaching least precision of 0.916, recall of 0.948, F-score of 0.932 and mAP of 0.908. Meanwhile, it is clear that the VGG16 method has concluded to a slightly better result to a finite limit with the precision of 0.925, recall of 0.955, F-score of 0.940 and mAP of 0.912. Then, the ResNet50 scheme has generated a moderate precision of 0.938, recall of 0.951, F-score of 0.944 and mAP of 0.916. Concurrently, the ResNet101 model has resulted to a better precision of 0.945, recall of 0.958, F-score of 0.951 and mAP of 0.922.

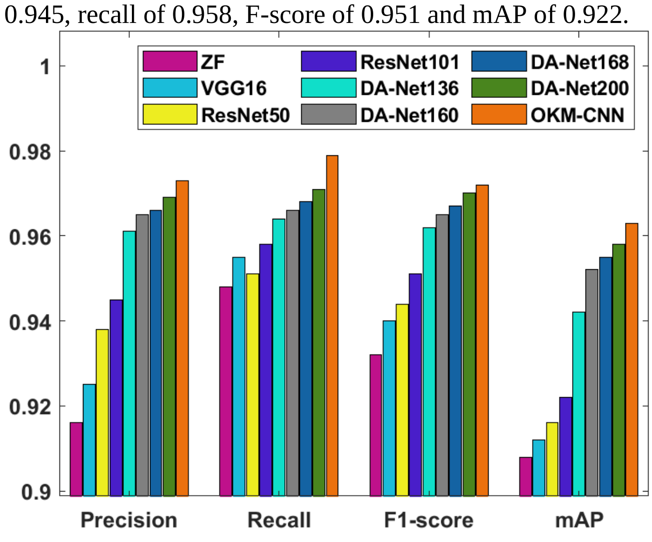

FIGURE 5. Results analysis of FZU Cars dataset

In line with this, the DA-Net136, DA-NET160, DA-168 and DA-Net200 methodologies have reached identical and reasonable performance. The DA-Net136 technology has provided a precision of 0.961, recall of 0.964, F-score of 0.962 and mAP of 0.942. Simultaneously, the DA-Net160 method has accomplished a precision of 0.965, recall of 0.966, F-score of 0.965 and mAP of 0.952. Along with that, the DA-Net168 model has achieved a precision of 0.966, recall of 0.968, F-score of 0.967 and mAP of 0.955. At the same time, the DA-Net200 model has provided a precision of 0.969, recall of 0.971, F-score of 0.970 and mAP of 0.958. At last, the OKM-CNN technology performs well than previous approaches with higher precision of 0.973, recall of 0.979, F-score of 0.972 and mAP of 0.963.

Fig. 6 depicts the relative study of OKM-CNN with conventional methods by means of precision, recall, F-score and mAP on the employs Stanford Cars Dataset. The table values showed that the ZF model performs ineffectively under the LP detection outcome by achieving lower precision of 0.856, recall of 0.897, F-score of 0.876 and mAP of 0.872. Concurrently, it is evident that the VGG16 approach has attained to slighter gradual results with the precision of 0.911, recall of 0.954, F-score of 0.932 and mAP of 0.907. Then, the ResNet50 model has offered a better precision of 0.912, recall of 0.946, F-score of 0.929 and mAP of 0.918. In line with this, the ResNet101 model has provided somewhat moderate precision of 0.941, recall of 0.952, F-score of 0.946 and mAP of 0.909.

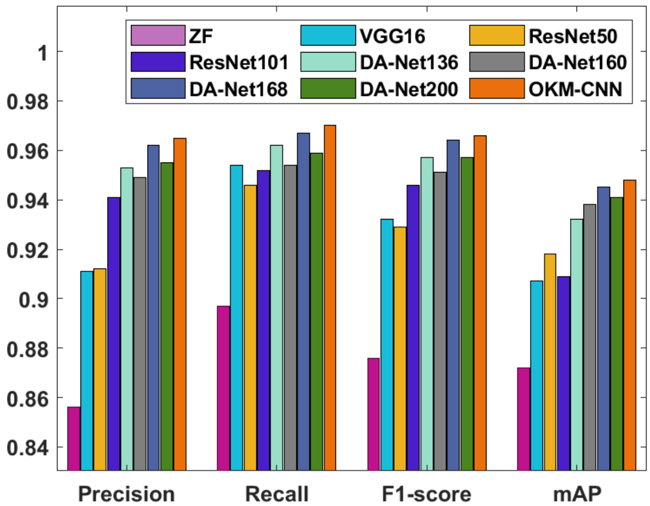

FIGURE 6. Results analysis of Stanford Cars Dataset

Along with that, the DA-Net136, DA-NET160, DA-168 and DA-Net200 models have accomplished nearer and considerable performance. The DA-Net136 model has obtained a precision of 0.953, recall of 0.962, F-score of 0.957 and mAP of 0.932. On the same way, the DA-Net160 model has originated a precision of 0.949, recall of 0.954, F-score of 0.951 and mAP of 0.938. Simultaneously, the DA-Net168 model has reached a precision of 0.962, recall of 0.967, F-score of 0.964 and mAP of 0.945. Meanwhile, the DA-Net200 model has achieved a precision of 0.955, recall of 0.959, F-score of 0.957 and mAP of 0.941. Consequently, the OKM-CNN model has attained superior outcome than traditional approaches with the optimal precision of 0.965, recall of 0.970, F-score of 0.966 and mAP of 0.948.

Fig. 7 showcases the comparative analysis of the OKM-CNN with conventional methods with respect to precision, recall, F-score and mAP on the presented HumAIn 2019 Dataset. The table values represented that the ZF technique tends to produce poor LP detection results by reaching lower precision of 0.863, recall of 0.873, F-score of 0.864 and mAP of 0.869. In line with this, it is clear that the VGG16 scheme has concluded to slighter moderate results with the precision of 0.869, recall of 0.889, F-score of 0.874 and mAP of 0.876. Besides, the ResNet50 model has generated better precision of 0.871, recall of 0.892, F-score of 0.887 and mAP of 0.892.

TABLE I
RESULT ANALYSIS OF EXISTING WITH PROPOSED OKM-CNN FOR APPLIED DATASET

Models	FZU Cars Dataset				Stanford Cars Dataset				HumAIn 2019 Dataset			
	Prec.	Rec.	F-score	mAP	Prec.	Rec.	F-score	mAP	Prec.	Rec.	F-score	mAP
ZF	0.916	0.948	0.932	0.908	0.856	0.897	0.876	0.872	0.863	0.873	0.864	0.869
VGG16	0.925	0.955	0.940	0.912	0.911	0.954	0.932	0.907	0.869	0.889	0.874	0.876
ResNet50	0.938	0.951	0.944	0.916	0.912	0.946	0.929	0.918	0.871	0.892	0.887	0.892
ResNet101	0.945	0.958	0.951	0.922	0.941	0.952	0.946	0.909	0.913	0.923	0.913	0.925
DA-Net136	0.961	0.964	0.962	0.942	0.953	0.962	0.957	0.932	0.923	0.931	0.926	0.935
DA-Net160	0.965	0.966	0.965	0.952	0.949	0.954	0.951	0.938	0.936	0.942	0.937	0.938
DA-Net168	0.966	0.968	0.967	0.955	0.962	0.967	0.964	0.945	0.932	0.948	0.941	0.942
DA-Net200	0.969	0.971	0.970	0.958	0.955	0.959	0.957	0.941	0.945	0.957	0.949	0.953
OKM-CNN	0.973	0.979	0.972	0.963	0.965	0.970	0.966	0.948	0.954	0.963	0.956	0.961

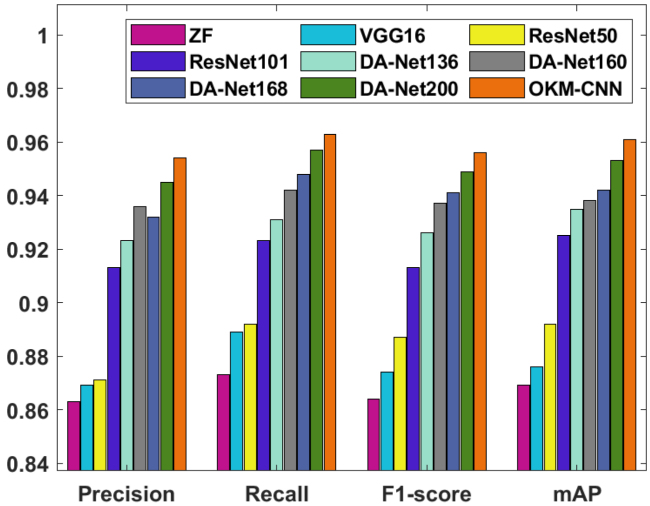

FIGURE 7. Results analysis of HumAIn 2019 Dataset

Concurrently, the ResNet101 model has accomplished gradual precision of 0.913, recall of 0.923, F-score of 0.913 and mAP of 0.925. Meanwhile, the DA-Net136, DA-NET160, DA-168 and DA-Net200 models have reached nearby and identical function. The DA-Net136 model has achieved a precision of 0.923, recall of 0.931, F-score of 0.926 and mAP of 0.935. On the same way, the DA-Net160 model has obtained a precision of 0.936, recall of 0.942, F-score of 0.937 and mAP of 0.938. Likewise, the DA-Net168 model has attained a precision of 0.932, recall of 0.948, F-score of 0.941 and mAP of 0.942. Then, the DA-Net200 model has reached to a precision of 0.945, recall of 0.957, F-score of 0.949 and mAP of 0.953. At last, the OKM-CNN model reached a higher precision of 0.954, recall of 0.963, F-score of 0.956 and mAP of 0.961 when compared with conventional models.

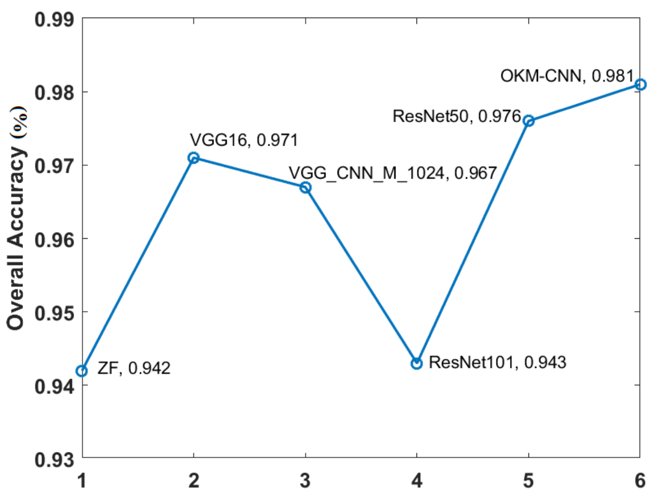

FIGURE 8. Overall accuracy analysis of diverse VLPR models

Table 2 and Fig. 8 tabulated the entire accuracy analysis provided by the OKM-CNN mechanism with previous models on applied dataset. It is depicted that the presented OKM-CNN model provides a best recognition performance by providing a higher accuracy of 0.981. Simultaneously, the VGG16 and ResNet 50 model attains closer and competitive competing results with the accuracy of 0.971 and 0.976.

TABLE II
RESULT ANALYSIS OF EXISTING WITH PROPOSED OKM-CNN IN TERMS OF OVERALL ACCURACY (%)

Model	Overall Accuracy
ZF	0.942
VGG16	0.971
VGG_CNN_M_1024	0.967
ResNet101	0.943
ResNet50	0.976
OKM-CNN	0.981

Likewise, the VGG_CNN_M_1024 model slightly minimum accuracy of 0.967 while even lower accuracy has been reached by the ZF and ResNet 101 models with the minimum accuracy values of 0.942 and 0.943 correspondingly. The projected OKM-CNN model attains efficient recognition on every applied image than the previous techniques.

V. CONCLUSION

This paper has presented a new OKM-CNN technique for effective detection and recognition of LPs. The proposed OKM-CNN model operates on three main stages. In the first stage, LP localization and detection process take place using IBA and CCA model. Subsequently, OKM based clustering technique gets executed to segment the LP image and finally, characters in LP recognition takes place using CNN model. The proposed OKM-CNN model can be employed as a major element of intelligent infrastructure like toll fee collection, parking management and traffic surveillance. An extensive experimental investigation takes place using three datasets namely Stanford Cars, FZU Cars and HumAIn 2019 Challenge dataset. The proposed OKM-CNN model has offered a maximum overall accuracy of 0.981 on the applied dataset. In future, the performance of the OKM-CNN model can be extended to recognize multilingual LPs.

CONFLICT OF INTEREST

Every author refers that, it is not have any conflict of interest about publication of this paper.

REFERENCES

[1] K. Yamaguchi, Y. Nagaya, K. Ueda, H. Nemoto, and M. Nakagawa, “A method for identifying specific vehicles using template matching,” in Proc. IEEE Int. Conf. Intell. Transp. Syst., 1999, pp. 8–13.

[2] Q. Zuo and Z. Shi, “A real-time algorithm for license plate extraction based on mathematical morphology,” J. Image Graph., vol. 8, no. 3, pp. 281–285, 2003.

[3] L. Salgado, J. M. Menendez, E. Rendon, and N. Garcia, “Automatic car plate detection and recognition through intelligent

vision engineering," in Proc. IEEE Int. Carnahan Conf. Security Technol., 1999, pp. 71–76.

[4] Y. P. Huang, S. Y. Lai, and W. P. Chuang, "A template-based model for license plate recognition," in Proc. IEEE Int. Conf. Netw., Sens. Control, Taipei, Taiwan, Mar. 21–23, 2004, pp. 737–742.

[5] C.-H. Lin, Y.-S. Lin, and W.-C. Liu, ''An efficient license plate recognition system using convolution neural networks,'' in Proc. IEEE Int. Conf. Appl. Syst. Invention (ICASI), Apr. 2018, pp. 224–227

[6] I. Ullah and H. J. Lee, ''An approach of locating Korean vehicle license plate based on mathematical morphology and geometrical features,'' in Proc. Int. Conf. Comput. Sci. Comput. Intell. (CSCI), Dec. 2016, pp. 836–840.

[7] S. S. Omran and J. A. Jarallah, ''Iraqi car license plate recognition using OCR,'' in Proc. Annu. Conf. New Trends Inf. Commun. Technol. Appl. (NTICT), Mar. 2017, pp. 298–303.

[8] K. M. Babu and M. V. Raghunadh, ''Vehicle number plate detection and recognition using bounding box method,'' in Proc. Int. Conf. Adv. Commun. Control Comput. Technol. (ICACCCT), May 2016, pp. 106–110.

[9] N. Rana and P. K. Dahiya, ''Localization techniques in ANPR systems: A-state-of-art,'' Int. J. Adv. Res. Comput. Sci. Softw. Eng., vol. 7, no. 5, pp. 682–686, May 2017.

[10] Liang, G. ,Shivakumara, P. , Lu, T. , & Tan, C. L. (2015). A new wavelet-Laplacian method for arbitrarily-oriented character segmentation in video text lines. In *Proceedings of the 13th international conference on document analysis and recog- nition (ICDAR)* (pp. 926–930). IEEE .

[11] Khare, V. ,Shivakumara, P. , Raveendran, P. , Meng, L. K. , &Woon, H. H. (2015). A new sharpness based approach for character segmentation in License plate im- ages. In *Proceedings of the 3rd IAPR Asian conference on pattern recognition (ACPR)* (pp. 544–548). IEEE .

[12] Kim, D. , Song, T. , Lee, Y. , & Ko, H. (2016). Effective character segmentation for license plate recognition under illumination changing environment. In *Proceedings of the IEEE international conference on consumer electronics (ICCE)* (pp. 532–533). IEEE .

[13] Dhar, P. , Guha, S. , Biswas, T. , & Abedin, M. Z. (2018). A system design for license plate recognition by using edge detection and convolution neural network. In *Proceedings of the IC4ME2* (pp. 1–4) .

[14] Ingole, S. K. ,&Gundre, S. B. (2017). Characters feature based Indian Vehicle license plate detection and recognition. In *Proceedings of the I2C2* (pp. 1–5) .

[15] Radchenko, A. ,Zarovsky, R. , &Kazymyr,V. (2017).Method ofsegmentationand recognition of Ukrainian license plates. In *Proceedings of the YSF* (pp. 62–65) .

[16] Raghunandan, K. S. , Shivakumara, P. , Jalab, H. A. , Ibrahim, R. W. , Kumar, G. H. , Pal, U. , et al. (2017). Riesz fractional based model for enhancing license plate detection and recognition. *IEEE Transactions on Circuits and Systems for Video Technology, 28* (9), 2276–2288 .

[17] Al-Shemarry, M. S. , Li, Y. , & Abdulla, S. (2018). Ensemble of adaboost cascades of 3L-LBPs classifiers for license plates detection with low quality images. *Expert Systems with Applications, 92* , 216–235 .

[18] Dong, M. , He, D. , Luo, C. , Liu, D. , & Zeng, W. (2017). A CNN-based approach for automatic license plate recognition in the wild. In *Proceedings of the BMCV* (pp. 1–12) .

[19] Bulan, O. ,Kozitsky, V. , Ramesh, P. , & Shreve, M. (2017). Segmentation-and annota- tion-free license plate recognition with deep localization and failure identifica- tion. *IEEE Transactions ITS, 18* (9), 2351–2363 .

[20] Yang, Y. , Li, D. , &Duan, Z. (2018). Chinese vehicle license plate recognition using kernel-based extreme learning machine with deep convolutional features. *IET Intelligent Transport System, 12* (3), 213–219 .