

Automated Telugu Printed and Handwritten Character Recognition in Single Image using Aquila Optimizer based Deep Learning Model

Vijaya Krishna Sonthi¹
Research Scholar
Department of CSE, FEAT
Annamalai University
Chidambaram, India

Dr. S. Nagarajan²
Associate Professor
Department of CSE, FEAT
Annamalai University
Chidambaram, India

Dr. N. Krishnaraj³
Associate Professor
School of Computing, SRM Institute
of Science & Technology
Kattankulathur, India

Abstract—Machine printed or handwritten character recognition becomes a major research topic in several real time applications. The recent advancements of deep learning and image processing techniques can be employed for printed and handwritten character recognition. Telugu character Recognition (TCR) remains a difficult task in optical character recognition (OCR), which transforms the printed and handwritten characters into respective text formats. In this aspect, this study introduces an effective deep learning based TCR model for printed and handwritten characters (DLTCR-PHWC). The proposed DLTCR-PHWC technique aims to detect and recognize the printed as well as handwritten characters that exist in the same image. Primarily, image pre-processing is performed using the adaptive fuzzy filtering technique. Next, line and character segmentation processes are performed to derive useful regions. In addition, the fusion of EfficientNet and CapsuleNet models is used for feature extraction. Finally, the Aquila optimizer (AO) with bi-directional long short-term memory (BiLSTM) model is utilized for recognition process. A detailed experimentation of the proposed DLTCR-PHWC technique is investigated using Telugu character dataset and the simulation outcome portrayed the supremacy of the proposed DLTCR-PHWC technique over the recent state of art approaches.

Keywords—Optical character recognition; Telugu; deep learning; Aquila optimizer; BiLSTM; handwritten characters; printed characters

I. INTRODUCTION

Recognition of hand printed or machine printed documents plays a significant role in applications such as text-speech converter, automated language-language translator, and intelligent scanning machines [1]. The aim of document image analyses is to identify the graphics and text modules in the paper documents and for extracting the proposed data, as human beings do. The two modules of document image analyses are textual and graphical processing. The former textual processing handles the text element of the document images and the graphical processing handles symbol and non-textual line elements which form line diagrams, delimit straight lines among company logos and text sections, and so on [2]. The image processing method is utilized for recognizing the handwritten Telugu character. Telugu is widely speaking in Telangana and Andhra Pradesh in India.

Optical character recognition (OCR) is a procedure which transforms the handwritten or printed files to their equivalent text format [3]. The OCR can be separated into two classes: online and offline character identification. Offline character identification is additionally separated into two components i.e., machine and handwritten printed character identification. In hand-written character recognition, there is a large number of problems in comparison with machine printed documents [4]. It is a challenging and fascinating field of pattern detection using many real-time applications. Many commercial OCR systems are accessible for printed Arabic characters however they have a lot of technological issues, particularly in the segmentation phase the result isn't enviable [5]. During the past years, OCR has become increasingly important since the need for translating the scanned image into computer identifiable forms like text documents has improved application. The OCR challenging problems such as distortion, lighting variations, variance in font size, and blurring of the printed character images, have improved the demands for OCR in the study work. The significant disadvantage witnessed in OCR is that the infra-class variation is larger because of the huge amount of images accessibility for the processing [6]. The OCR system is currently under development for most of the common languages and Telugu is no exemption for it. The OCR process has become really hard for Telugu and possess its individual challenges to the developer of Telugu character recognition (TCR) system. In recent times, significant studies have been made toward the development of an effective TCR scheme [7].

This study focuses on the design of effective deep learning based TCR technique for printed and handwritten characters (DLTCR-PHWC) in the same image. The DLTCR-PHWC involves adaptive fuzzy filtering (AFF) technique for image pre-processing. Besides, line and character segmentation processes are performed to derive useful regions. Moreover, the fusion of EfficientNet and CapsuleNet models is used for feature extraction. Furthermore, the Aquila optimizer (AO). With bidirectional long short-term memory (BiLSTM) model is applied for recognition process. A comprehensive experimental analysis takes place using Telugu character dataset a detailed comparative analysis is carried out in terms of different evaluation parameters.

The rest of the paper in Section 2 discusses literature review, Section 3 briefs proposed model and Section 4 discusses results and discussion and Section 5 ends up with conclusion.

II. LITERATURE REVIEW

Prameela et al. [8] proposed an OCR method for Telugu documents that consists of 2 steps, i.e. classification, preprocessing, and feature extraction. In preprocessing phase, they used median filtering on the input character and employed skeletonization and normalization methods over character for extracting boundary edge pixel points. In feature extraction phase, first, all the characters are separated into 3x3 grids and the equivalent centroid for every 9 regions is calculated. Cheekati and Rajeti [9] deal in emerging a reliable, fast Telugu hand-written ResNet for offline and online character identification and improve classification accuracy. The method is estimated by IIITS-Telugu Handwriting Dataset; HP Lab databases (Telugu) India and attained a remarkable result. The presented residual net (ResNet-50) attains 2.37% error on ResNet-18 & 34 test set.

Lakshmi and Babu [10] perform a new Telugu script identification and retrieval method named HCH model. Hash coding would be utilized as a feature extractor and the hamming distance would be used as a replacement for traditional Euclidean distance for measuring the similarity among database and query images. Experimental analyses exhibit that the presented model has outstanding performances to the traditional methods proposed in the survey. Burra et al. [11] proposed 2 methods for improving the glyph/symbol segmentations in a Telugu OCR scheme. The main features having an effect on the entire performance of a Telugu OCR scheme can be able to divide/segment scanned document images into identifiable units. In Telugu, this unit is generally interconnected component and is known as glyphs. Once a document is removed, interconnected components-based algorithm for segmentation fails. They provide malformed glyphs (a) partial and results of break in the character because of uneven dispersal of ink on the noise /page, and (b) are an integration of more than two glyphs due to smudging in noise/print. The previous one is labeled broken and the last one is merged character. Hebhi et al. [12] introduce a cross language environment for recognizing the words and characters of lower resources script that is script which doesn't have typical database and the dataset isn't accessible for public access. Indic script comes from a popular origin and few scripts have a standard 3 regional structures. Identification of this script could be performed using another script having same structures. In order to identify this character, the models are trained using source language Kannada with zone wise testing and training is made by Kannada and the targeting language Telugu.

Rani et al. [13] proposed a method for feature classification and extraction of Telugu hand-written script based personalized template matching method through caching method for achieving better results. The method of caching is executed by main databases with a cache database maintain the frequently employed character template for a set of each character template. The XML databases are utilized in

the class for different character templates and the class representation is given by new class structures depending on XML tags. In Sarika and Sirisala [14], OCR methods like preprocessing, digitization, feature extraction, recognition, and segmentation were discussed. HWCR with distinct ML methods like Bayesian decision theory, SVM, and Bayesian classifier was discussed. The current technique of HWCR is examined for native language and related their features and functionality.

Madhavi et al. [15] proposed an effective model named TR for correcting and detecting slant angles of MTW. Telugu language is India's common language speaking around 80 million people. The difficult character is attached with further marks called "vatthus" and "maatras" it is stimulating for detecting a slant angle. The presented TR model performs preprocessing and identifies interconnected components within the provided MTW. Later, estimate the slant angle of every interconnected component by acquiring interconnected slant lines on the boundary of every interconnected component. BJ et al. [16] handle classification and recognition of confusing Kannada characters. The RF and SVM as classifiers to categorize the confusing character. The presented method attained 78% of classifiers accuracy. Lastly, this method identifies the confusing characters with template matching and feature value outcomes-based classifier.

III. THE PROPOSED TCR MODEL

In this study, a new DLTCR-PHWC technique is derived to detect and recognize the printed as well as handwritten characters that exist in the same image. The DLTCR-PHWC technique encompasses different operations such as AFF based pre-processing, line/paragraph segmentation, fusion-based feature extraction, BiLSTM based recognition, and AO based hyperparameter optimization. Fig. 1 showcases the overall block diagram of proposed DLTCR-PHWC model. The detailed working of every module is offered in the following subsections.

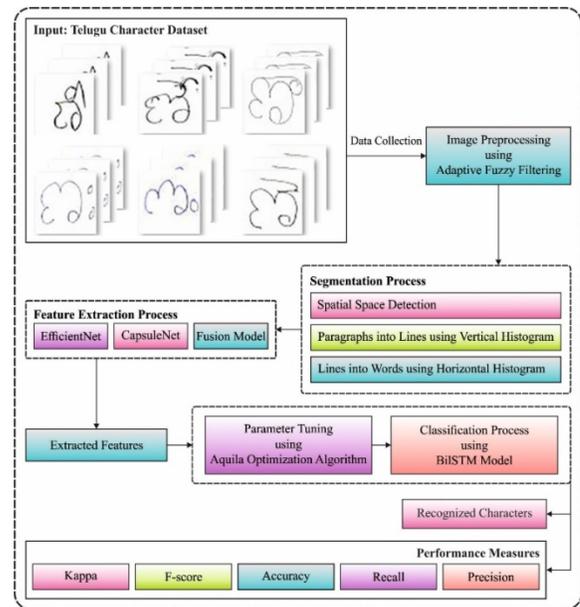


Fig. 1. Overall Process of DLTCR-PHWC Model.

A. Pre-Processing

In this model, the input receives image as input for removing the presence of noise. Once the image is reduced by high noise and the whole pixels in filter window are similar to the maximum values, it increases the window size up to a predefined maximal size W_{max} , it changes $I(x,y)$ by the median values i.e., estimated from filter window. In the process of finding a median value, each pixel value is similar to the maximum values in filter window, through evaluated pixel value in filter window except the maximum values to obtain median values.

B. Segmentation Process

Segmentation is a major phase in detection method since it extracts significant fields for additional analyses. Typically, it is employed for verifying the boundary and objects such as lines, curves, and so on, in an image. The scanned images are segmented to the paragraph via spatial space recognition method, paragraph to lines through vertical histogram, lines to words through horizontal histogram [17]. Accurateness of character identification is based largely on segmentation performance. Mainly, this procedure has the subsequent steps:

- Recognize the text line on the page.
- Recognize single character in all the words.

The widely employed technique for line segment of the grayscale images is the projection profile model. By adding it to the horizontal course of the documents, gaps among the text line could be recognized through detecting the projection value. They employ horizontal projection profiles analyses since the text in most document images is arranged with the horizontal line. This method determines horizontal projection histogram, the amount of black pixel for every column of the raster images. Once these profiles are employed on $M \times N$ images, a column vector of $M \times 1$ sizes were attained. Element of these column vectors is the amount of pixel value in every row of the document images. Afterward, the line segmentation the resulting outputs are fed into the character segments. Finally, the extracted line is segmented to the character. In order to detect the boundaries among the characters, they employ threshold values on the length of the space amongst the characters. Afterward detecting the position of the space among the characters they remove the part of the line segmentation. As per the abovementioned procedure, the character and lines are segmented from the pre-processed documents. Afterward the segmentation procedure the resulting outputs are fed into the feature extraction.

C. Fusion based Feature Extraction Technique

At this stage, the segmented images undergo feature extraction process using CapsNet and EfficientNet models. The underlying framework of the CapsNet, adapted in this study, is made up of 1 fully connected layer and two convolution layers. The 1st convolution layer is generated with 256 9×9 kernels using a stride of one and ReLU activation, deliver feature map which is additionally fed into the initial set of capsules in the abovementioned layers. The 2nd layers represent the initial capsule and accommodate thirty-two channels, therefore eight-dimensional convolution capsules of 9×9 kernels and a stride of two, in which every

capsule captures each unit in the 1st convolution layer that receptive field overlaps with the center of capsules. Therefore, the set of initial capsules output an overall of $32 \times 6 \times 6$ eight dimensional vectors, and the capsule of the similar grid share their weight with one another. The final layer of the CapsNet is an FC layer of 2 sixteen dimensional units which is interconnected to each capsule in the prior layer [18]. As the output of 1st convolution layer is 1D, it doesn't transfer the similar quality of data as capsule in the abovementioned layers (viz., the output of 1st layer doesn't offer orientation attribute to agree on), no routing is proposed with the initial capsule. It is notable that each logit B_{ij} are initiated as zero that indicates the primary capsule output is transmitted to each possible parent capsule with equivalent probabilities C_{ij} . In another word, previous to knowledge optimization, the initial capsule assumes an equivalent agreement with parent capsule (for example each entity tied with initial capsule is related to the entity tied to above parent capsule). Fig. 2 illustrates the framework of CapsNet model.

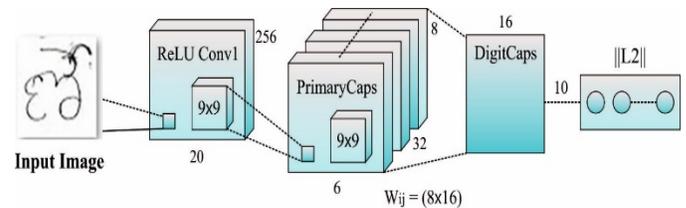


Fig. 2. Structure of CapsNet.

With respect to the loss function for the networks to learn, it consists in extending the length of instantiation vectors V_k for an entity class (car and solar panel in these cases) given that certain entity is witnessed in an image. This could be expanded for several entities through a single margin loss L_k for all entities k capsule :

$$L_k = T_k \max(0, m^+ - \|V_k\|)^2 + \lambda(1 - T_k) \max(0, \|V_k\| - m^-)^2 \quad (1)$$

Whereas T_k is fixed to 1 when the entity is existing.

In recent years, the rapid growth of DL method has spawned many excellent CNN models. From the initial simple network to the current complex network, the performance of the model is getting better and better in every aspect. EfficientNet combines the advantage of previous network, which summarize the development of network performances into 3D: (1) Deepen the network, i.e., use the skip connection to increase the depth of the neural networks, and attain feature extraction via deeper layer; (2) Widen the network, i.e., increase the amount of convolution layers to attain more functions and features; (3) By increasing the input image resolution, the network could express and learn more things, which is beneficial to improve accuracy. Then, use a compound coefficient ϕ to uniformly balance and scale the resolution, depth, and width of the networks, as well maximize the network performance on limited resources. Estimation of the compound coefficient is given in Eq. (1):

$$\begin{aligned} \text{depth: } d &= \alpha^\phi & \text{width: } w &= \beta^\phi & \text{resolution: } r &= \gamma^\phi \\ \text{s.t. } \alpha \cdot \beta^2 \cdot \gamma^2 &\approx 2 \end{aligned} \quad (2)$$

$$\alpha \geq 1, \beta \geq 1, \gamma \geq 1$$

Whereas d , w , and r represent the coefficient used to scale the resolution, depth, and width of the networks [19]. The α , β , & γ denotes resource allocation for network depth, width, and resolution. EfficientNet largely consists of Stem, 16 Blocks, Conv2D, GlobalAveragePooling2D, and Dense layer. The design of Block is based mainly on the attention mechanism and residual structure, also other structures are similar to traditional CNN model.

The extracted features from both DL models are fused together. During the fusion process, it is needed to decide the value of λ , which calls the R feature fusion. The fusion of features can be defined as follows.

$$NF = \lambda \cdot LF + (1 - \lambda) \cdot HF, \quad (3)$$

where NF is the fusion feature, and LF and HF indicate the features derived by the CapsNet and EfficientNet models.

D. Design of AO based BiLSTM Model for Recognition Process

Finally, the BiLSTM model receives the feature vectors as input and carried out the recognition process. The LSTM refers to a special RNN model that resolves the problems of gradient vanishing of the RNN by presenting a threshold mechanism and memory unit [20]. But x represent the network input at distinct times, y indicates the network output, h represents the hidden layer, u signifies the weight from input to the hidden layers, w denote the weights of prior node hidden layer to the present node hidden layer, and v represents the weight from hidden to the output layers.

In the actual execution of the LSTM, the LSTM units are upgraded at time t as follows:

$$i_t = \sigma(w_i h_{t-1} + U_i x_t + b_t) \quad (4)$$

$$f_t = \sigma(w_f h_{t-1} + U_f x_t + b_f) \quad (5)$$

$$\tilde{c}_t = \tanh(W_c h_{t-1} + U_c x_t + b_c) \quad (6)$$

$$c_t = f_t \odot c_{t-1} + i_t \odot \tilde{c}_t \quad (7)$$

$$o_t = \sigma(w_o h_{t-1} + U_o x_t + b_o) \quad (8)$$

$$h_t = o_{t-1} \odot \tanh(c_t) \quad (9)$$

Now, σ represents the sigmoid function and \odot indicates the equivalent product of the element. x_t denotes the input vector at time t . h_t signifies the hidden state vector, that is called output vector, and store each data at time t and the preceding time. U_i, U_f, U_c, U_o represent the weight of input vector x_t for the input, forgotten, unit, and output gates, correspondingly. W_i, W_f, W_c, W_o denote the weight of dissimilar gates to the hidden state vector h_t . b_t, b_f, b_c, b_o denotes the off-set vector. Using the 3 gates structure, the LSTM allows the recurrent network to maintain the beneficial data for the task in the memory unit at the time of the training model, thus evading the problems of the RNN disappearing while obtaining long range data.

While processing sequence data, the BLSTM introduces further backward estimation processes, i.e., unlike normal

LSTM case. This procedure could employ the succeeding data of the sequence. Finally, the reverse and forward evaluations were implemented. The value is output to the output layer concurrently; as a result, each data of a sequence is attained by two directions that is used for multiplying type of natural language processing task.

Followed by, the hyperparameter tuning of the BiLSTM model take place using AO technique. The Aquila is most famous bird of prey. Young Aquila generally attains entire assurance during the fall, subsequent that it can be moved extremely for building territory to themselves. Because of their hunting bravery, Aquila is a most considered bird globally. The Male Aquila become considerably further prey if the solo-hunting. An essential stimulus to the presented AO technique has been resultant in the techniques stated above. The subsequent subsections that define these procedures are modeled in AO. In AO, it can be population-oriented technique, the optimized rules start with population of candidate solution (X) as projected in Eq. (9) that has been created stochastically amongst the upper boundary (UB) and lower boundary (LB) of the provided issues. An optimum solution obtained during all iterations can be defined as follows.

$$X = \begin{bmatrix} x_{1,1} & \cdots & x_{1,j} & x_{1,Dim-1} & x_{1,Dim} \\ x_{2,1} & \cdots & x_{2,j} & \cdots & x_{2,Dim} \\ \cdots & \cdots & x_{i,j} & \cdots & \cdots \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ x_{N-1,1} & \cdots & x_{N-1,j} & \cdots & x_{N-1,Dim} \\ x_{N,1} & \cdots & x_{N,j} & x_{N,Dim-1} & x_{N,Dim} \end{bmatrix} \quad (10)$$

where X refers the group of present candidate solutions that are created arbitrarily with utilizing Eq. (10), X_i indicates the decision values (places) of i^{th} solutions, N represents the entire amount of candidate solution (population), and Dim signifies the dimensional of the issue.

$$X_{ij} = rand \times (UB_j - LB_j) + LB_j, i = 1, 2, \dots, N, j = 1, 2, \dots, Dim \quad (11)$$

where $rand$ denotes the arbitrary number, LB_j demonstrates the j^{th} lower bound, and UB_j implies the j^{th} upper bound of provided issue [21]. The AO technique is transmission in exploration to exploitation stages utilizing distinct performance dependent upon this form if $t \leq \left(\frac{2}{3}\right) * T$ the exploration stages were excited; else, the exploitation stages are implemented. The mathematical process of the AO has been presented. During the primary technique (X_1), the Aquila distinguishes the prey region and elects an optimum hunting region by great soar with vertical stoop. At this point, the AO extremely explorers in great soar for determining the region of search space in which the prey is. This performance has been mathematically projected as in Eq. (12)

$$X_1(t+1) = X_{best}(t) \times \left(1 - \frac{t}{T}\right) + (X_M(t) - X_{best}(t)) * rand, \quad (12)$$

where, $X_1(t+1)$ refers the solution of subsequent round of t that is created by initial searching technique (X_1). $X_{best}(t)$ signifies the optimum attained solution until t^{th}

iteration, this reproduces the estimated prey's position. The formula $\left(\frac{1-t}{T}\right)$ has been utilized for controlling the extended search (exploration) with the amount of iterations. $X_M(t)$ defines the places mean value of present solutions associated at t^{th} iteration that is computed utilizing in Eq. (12). $rand$ indicates the arbitrary value amongst $[0, 1]$. r and T demonstrate the present round and the maximal rounds, correspondingly.

$$X_M(t) = \frac{1}{N} \sum_{i=1}^N X_j(t), \forall j = 1, 2, \dots, Dim \quad (13)$$

where Dim implies the dimensional size of issue and N represents the population size.

During the second technique (X_2), if the prey area has been initiated in a great soar, the Aquila circle on the target, arranges the land, and next attack. This performance was mathematically processed as in Eq. (14).

$$X_2(t+1) = X_{best}(t) \times Levy(D) + X_R(t) + (y-x) \times rand, \quad (14)$$

where $X_2(t+1)$ signifies the solution of next iteration of r that has been created by the next searching technique (X_2). D implies the dimensional space, and $Levy(D)$ demonstrates the levy flight distribution function that was computed utilizing in Eq. (15). $X_R(t)$ represents the arbitrary solution obtained from the range of $[1N]$ at i^{th} iteration.

$$Levy(D) = s \times \frac{u \times \sigma}{|v|^{\beta}} \quad (15)$$

where s indicates the constant values set to 0.01, u and v denotes the arbitrary numbers amongst $[0, 1]$. σ signifies the computed utilizing in Eq. (16).

$$\sigma = \left(\frac{\Gamma(1+\beta) \times \sin\left(\frac{\pi\beta}{2}\right)}{\Gamma\left(\frac{1+\beta}{2}\right) \times \beta \times 2^{\left(\frac{\beta-1}{2}\right)}} \right) \quad (16)$$

In Eq. (14), y and x are utilized for presenting the spiral shape during the searching process that is computed below:

$$y = r \times \cos(\theta) \quad (17)$$

$$x = r \times \sin(\theta) \quad (18)$$

where,

$$r = r_1 + U \times D_1 \quad (19)$$

$$\theta = -\omega \times D_1 + \theta_1 \quad (20)$$

$$\theta_1 = \frac{3 \times \pi}{2} \quad (21)$$

r_1 gets the value amongst $[1, 20]$ to set the amount of search cycles, and U refers the lesser value set to 0.00565. D_1 defines the integer number in 1 to the length of search spaces (Dim), and ω indicates the lesser value set to 0.005. During the third approach (X_3), if the prey area has been identified perfectly, and the Aquila has been prepared to land and attacks, the Aquila inclines vertically with initial attack for discovering the prey reaction. This technique is named minimum flight. At this point, the AO utilizes the chosen region of the target for getting nearby prey as well as attack.

This performance has been mathematically processed in Eq. (22).

$$X_3(t+1) = (X_{best}(t) - X_M(t)) \times \alpha - rand + ((UB - LB) \times rand + LB) \times \delta, \quad (22)$$

where $X_3(r+1)$ implies the solution of succeeding round of r that is created by 3rd searching manner (X_3). $X_{best}(t)$ signifies the estimated place of prey till i^{th} iteration (the optimum attained solution), and $X_M(t)$ represents the mean value of present solution at r^{th} iteration that has been computed utilizing in Eq. (12). $rand$ stands for the arbitrary value amongst $[0, 1]$. During the fourth technique (X_4), if the Aquila became nearby the prey, the Aquila attack the prey on the land based on its stochastic movement. This process was mathematically projected as in Eq. (23).

$$X_4(t+1) = QF \times X_{best}(t) - (G_1 \times X(t) \times rand) - G_2 \times Levy(D) + rand \times G_1, \quad (22)$$

where $X_4(t+1)$ implies the solution of next iteration of t that has been created by the fourth search technique (X_4). QF demonstrates the quality function utilized for equilibrium the search approaches that are computed.

IV. PERFORMANCE VALIDATION

The performance validation of the DLTCR-PHWC technique is tested using benchmark Telugu character dataset. The results are investigated under different folds and a comprehensive comparative analysis is also performed. Fig. 1 shows the sample segmented output of the image comprising handwritten and printed characters. The figure stated that the DLTCR-PHWC technique has offered effective identification of characters exist in the applied image shown in Fig. 3.

Table I and Fig. 4 offers the detailed recognition performance of the DLTCR-PHWC technique under ten folds. The results stated that the DLTCR-PHWC technique has gained effectual outcomes on all the applied folds. For example, with F_1, the DLTCR-PHWC approach has obtained a precision of 98.53%, recall of 98.37%, accuracy of 98.24%, F1-measure of 98.43%, and kappa of 97%. Also, with F_2, the DLTCR-PHWC approach has gained a precision of 99.14%, recall of 97.81%, accuracy of 99.30%, F1-measure of 99.24%, and kappa of 96.10%. Moreover, with F_4, the DLTCR-PHWC approach has attained a precision of 98.48%, recall of 98.68%, accuracy of 99.14%, F1-measure of 97.54%, and kappa of 97.29%. Furthermore, with F_6, the DLTCR-PHWC approach has achieved a precision of 99.05%, recall of 99.42%, accuracy of 98.91%, F1-measure of 98.70%, and kappa of 97.89%. Concurrently, with F_8, the DLTCR-PHWC approach has gained a precision of 99.31%, recall of 98.16%, accuracy of 99.57%, F1-measure of 97.97%, and kappa of 98.30%. Lastly, with F_1, the DLTCR-PHWC algorithm has provided a precision of 99.58%, recall of 99.82%, accuracy of 99.78%, F1-measure of 98.03%, and kappa of 99.34%.

మొన్నబు మూలలు దూరంగా పుస్తకాలకు అవేశం తెప్పించాయి. అందులోనూ ఏసయ్యకు కోపమెక్కువ. "మనసులో మాట కక్కేశారు మున్నబుగారూ! మేము కూడా మనుషుల్లా తిరగడం సహించలేకనేగదా ఇన్ని మాట్లాడుతున్నారు...." అంటూ దూసుకొచ్చాడు.

(a)

మొన్నబు మూలలు దూరంగా పుస్తకాలకు అవేశం తెప్పించాయి. అందులోనూ ఏసయ్యకు కోపమెక్కువ. "మనసులో మాట కక్కేశారు మున్నబుగారూ! మేము కూడా మనుషుల్లా తిరగడం సహించలేకనేగదా ఇన్ని మాట్లాడుతున్నారు...." అంటూ దూసుకొచ్చాడు.

(b)

Fig. 3. Sample Segmentation Results (a) Original Handwritten & Printed Characters, (b) Segmented Handwritten & Printed Characters.

Finally, a detailed comparative results analysis of the DLTCR-PHWC technique takes place in Table II [22-26]. Fig. 5 provides a brief precision analysis of the DLTCR-PHWC technique with existing techniques. The figure reported that the MLP-HMM and DNN techniques have obtained worse outcomes with the least precision of 0.8467 and 0.9089 respectively. In addition, the CNN, KNN, MLP, and NN techniques have obtained moderately closer precision of 0.9567, 0.9415, 0.9572, and 0.9634 respectively. However, the DLTCR-PHWC technique has resulted in a higher precision of 0.9889.

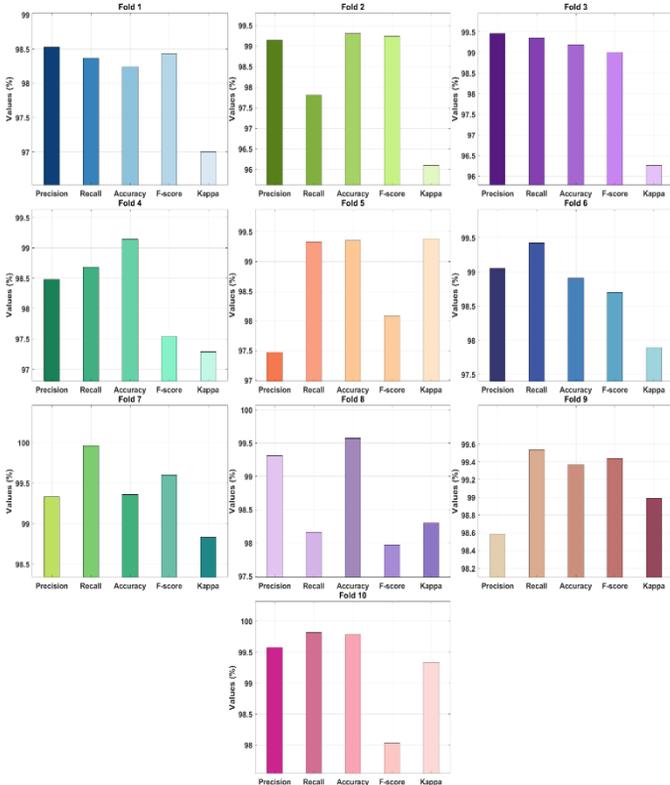


Fig. 4. Result Analysis of DLTCR-PHWC Model under Different Folds.

TABLE I. RESULT ANALYSIS OF PROPOSED MODEL IN TERMS OF DIFFERENT MEASURES

No. of Folds	Precision	Recall	Accuracy	F-score	Kappa
Fold 1	98.53	98.37	98.24	98.43	97.00
Fold 2	99.14	97.81	99.30	99.24	96.10
Fold 3	99.46	99.35	99.19	99.00	96.26
Fold 4	98.48	98.68	99.14	97.54	97.29
Fold 5	97.47	99.33	99.35	98.09	99.38
Fold 6	99.05	99.42	98.91	98.70	97.89
Fold 7	99.33	99.96	99.36	99.60	98.83
Fold 8	99.31	98.16	99.57	97.97	98.30
Fold 9	98.59	99.54	99.37	99.44	98.99
Fold 10	99.58	99.82	99.78	98.03	99.34
Average	98.89	99.04	99.22	98.60	97.94

TABLE II. RESULT ANALYSIS OF EXISTING WITH PROPOSED METHOD IN TERMS OF DIFFERENT MEASURES

Methods	Precision	Recall	F-score	Accuracy
DLTCR-PHWC	0.9889	0.9904	0.9860	0.9922
MLP-HMM	0.8467	0.8689	0.8512	0.8500
CNN	0.9567	0.9582	0.9572	0.9632
KNN	0.9415	0.9543	0.9512	0.9585
MLP	0.9572	0.9590	0.9578	0.9645
NN	0.9634	0.9685	0.9680	0.9750
DNN	0.9089	0.9144	0.9132	0.9210
CNN-RF	-	-	-	0.7140
CNN-MLP	-	-	-	0.7750
CNN-KNN	-	-	-	0.8160

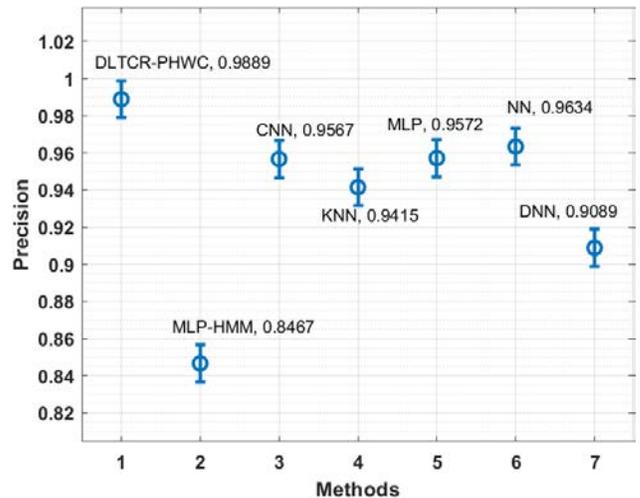


Fig. 5. Comparative Analysis of DLTCR-PHWC Model with respect to Precision.

Fig. 6 offers a detailed recall analysis of the DLTCR-PHWC algorithm with recent algorithms. The figure outperformed that the MLP-HMM and DNN manners have attained least outcome with the minimum recall of 0.8689 and 0.9144 correspondingly. Besides, the CNN, KNN, MLP, and NN approaches have gained moderately closer recall of 0.9582, 0.9543, 0.9590, and 0.9685 correspondingly. Finally, the DLTCR-PHWC methodology has resulted in an increased recall of 0.9904.

Fig. 7 showcases a brief F-score analysis of the DLTCR-PHWC manner with existing algorithms. The figure described that the MLP-HMM and DNN methods have reached minimum results with the reduced F-score of 0.8512 and 0.9132 correspondingly. Followed by, the CNN, KNN, MLP, and NN manners have achieved moderately closer F-score of 0.9572, 0.9512, 0.9578, and 0.9680 correspondingly. Eventually, the DLTCR-PHWC method has resulted in a superior F-score of 0.9860.

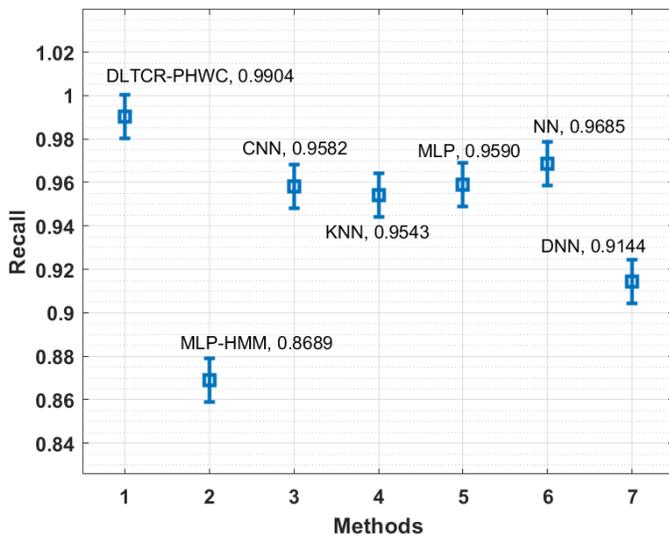


Fig. 6. Comparative Analysis of DLTCR-PHWC Model with respect to Recall.

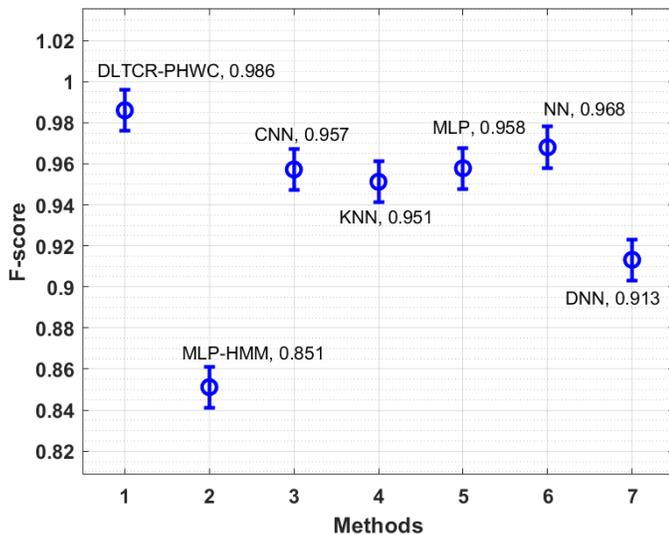


Fig. 7. Comparative Analysis of DLTCR-PHWC Model with respect to F-Score.

Fig. 8 demonstrates a brief accuracy analysis of the DLTCR-PHWC method with existing approaches. The figure stated that the CNN-RF, CNN-MLP, and CNN-KNN algorithms have gained worst outcome with the minimum accuracy of 0.7140, 0.7750, and 0.8160 respectively. At the same time, MLP-HMM and DNN techniques demonstrated a somewhat higher accuracy of 0.8500 and 0.9210 correspondingly. Along with that, the CNN, KNN, MLP, and NN techniques have obtained moderately closer accuracy of 0.9632, 0.9585, 0.9645, and 0.9750 respectively. At last, the DLTCR-PHWC methodology has resulted in a maximal accuracy of 0.9922.

After examining the above results and discussion, it is apparent that the DLTCR-PHWC technique has been found to be a proficient tool for effective TCR.

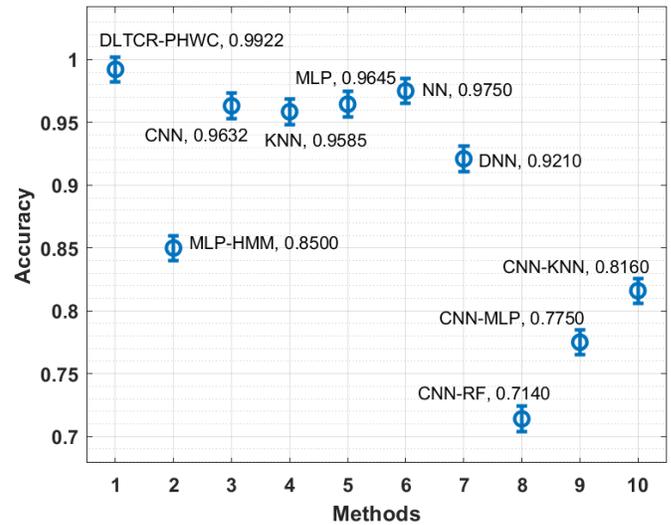


Fig. 8. Comparative Analysis of DLTCR-PHWC Model with respect to Accuracy.

V. CONCLUSION

In this study, a new DLTCR-PHWC technique is derived to detect and recognize the printed as well as handwritten characters that exist in the same image. The DLTCR-PHWC technique encompasses different operations such as AFF based preprocessing, line/paragraph segmentation, fusion based feature extraction, BiLSTM based recognition, and AO based hyperparameter optimization. The design of AO technique to fine tune the hyperparameters involved in the BiLSTM model helps to accomplish enhanced TCR outcomes. In order to showcase the supremacy of the DLTCR-PHWC technique, a wide range of simulations were performed against Telugu character dataset and the results have ensured the betterment of the DLTCR-PHWC technique. Therefore, the DLTCR-PHWC technique finds it useful to recognize both the handwritten and Telugu characters present in same image. In future, advanced DL models can be utilized for recognition process with hybrid metaheuristic-based parameter tuning process.

REFERENCES

- [1] Lin D, Lin F, Lv Y, Cai F, Cao D (2017) Chinese character CAPTCHA recognition and performance estimation via deep neural network. *Neuro Computing* 17:1–40.

- [2] Guruprasad P, Majumdar J (2016) Multimodal recognition framework: an accurate and powerful Nandinagari handwritten character recognition Model. *Procedia Comput Sci* 89:836–844.
- [3] Sampath AK, Gomathi N (2017) Fuzzy-based multi-kernel spherical support vector machine for effective handwritten character recognition, *Research Article* 10:1-13.
- [4] Naz S, Hayat K, Razzak MI, Anwar MW, Madani SA, Khan SU (2013) The optical character recognition of Urdu-like cursive scripts, *Research Article*, 31:1229-1249.
- [5] Chacko BP, Vimal Krishnan VR, Raju G, Babu Anto P (2012) Handwritten character recognition using wavelet energy and extreme learning machine. *Int J Mach Learn Cybern* 3:149–161.
- [6] Sampath AK, Gomathi N (2017) Decision tree and deep learning based probabilistic model for character recognition. *J Cent S Univ* 24:2862–2876.
- [7] Ajantha Devi V, Santhosh Baboo S (2014) Embedded optical character recognition on Tamil text image using raspberry pi. *International Journal of Computer Science Trends and Technology (IJCSST)* 2(4):127– 132.
- [8] Prameela, N., Anjusha, P. and Karthik, R., (2017), April. Off-line Telugu handwritten characters recognition using optical character recognition. In 2017 International conference of Electronics, Communication and Aerospace Technology (ICECA) (Vol. 2, pp. 223-226). IEEE.
- [9] Cheekati, B.M. and Rajeti, R.S., (2020), October. Telugu handwritten character recognition using deep residual learning. In 2020 Fourth International Conference on I-SMAC (IoT in Social, Mobile, Analytics and Cloud)(I-SMAC) (pp. 788-796). IEEE.
- [10] Lakshmi, K.M. and Babu, T.R., (2018), January. A Novel Telugu Script Recognition and Retrieval Approach Based on Hash Coded Hamming. In *International Conference on Communications and Cyber Physical Engineering 2018* (pp. 571-582). Springer, Singapore.
- [11] Burra, S., Patel, A., Bhagvati, C. and Negi, A., (2017), December. Improved symbol segmentation for telugu optical character recognition. In *International Conference on Intelligent Systems Design and Applications* (pp. 496-507). Springer, Cham.
- [12] Hebbsi, C., Mamatha, H.R., Sahana, Y.S., Dhage, S. and Somayaji, S., (2020). A convolution neural networks-based character and word recognition system for similar script languages Kannada and Telugu. In *Proceedings of ICETIT 2019* (pp. 306-317). Springer, Cham.
- [13] Rani, N.S., Vasudev, T. and Pradeep, C.H., (2017). An Enhanced Template Matching Technique for Recognition of Telugu Script. *International Journal of Signal Processing*, 2.
- [14] Sarika, N. and Sirisala, N., (2021). Deep Learning Techniques for Optical Character Recognition. In *Sustainable Communication Networks and Application* (pp. 339-349). Springer, Singapore.
- [15] Madhavi, G.B., Kumar, V. and Vakula, V.K., (2021). An Effective Slant Detection and Correction Method Based on the Tilted Rectangle Method for Telugu Manuscript Terms. *International Journal of Information Technology Project Management (IJITPM)*, 12(4), pp.25-37.
- [16] BJ, B.N., Athira, M.R. and Prajwal, M.L., (2021), May. Kannada Confusing Character Recognition and Classification Using Random Forest and SVM. In 2021 3rd International Conference on Signal Processing and Communication (ICPSC) (pp. 537-541). IEEE.
- [17] Kowsalya, S. and Periasamy, P.S., (2019). Recognition of Tamil handwritten character using modified neural network with aid of elephant herding optimization. *Multimedia Tools and Applications*, 78(17), pp.25043-25061.
- [18] Mekhalafi, M.L., Bejiga, M.B., Soresina, D., Melgani, F. and Demir, B., (2019). Capsule networks for object detection in UAV imagery. *Remote Sensing*, 11(14), p.1694.
- [19] Wu, L., Ma, J., Zhao, Y. and Liu, H., (2021). Apple Detection in Complex Scene Using the Improved YOLOv4 Model. *Agronomy*, 11(3), p.476.
- [20] Ji, Z., Wang, X., Cai, C. and Sun, H., (2020). Power entity recognition based on bidirectional long short-term memory and conditional random fields. *Global Energy Interconnection*, 3(2), pp.186-192.
- [21] Abualigah, L., Yousri, D., Abd Elaziz, M., Ewees, A.A., Al-qaness, M.A. and Gandomi, A.H., (2021). Aquila Optimizer: A novel meta-heuristic optimization Algorithm. *Computers & Industrial Engineering*, 157, p.107250.
- [22] Kummari, R. and Bhagvati, C., (2018), December. UHTelPCC: A Dataset for Telugu Printed Character Recognition. In *International Conference on Recent Trends in Image Processing and Pattern Recognition* (pp. 24-36). Springer, Singapore.
- [23] Lakshmi, C.V., Jain, R. and Patvardhan, C., (2006). OCR of printed Telugu text with high recognition accuracies. In *Computer Vision, Graphics and Image Processing* (pp. 786-795). Springer, Berlin, Heidelberg.
- [24] Pujari, A.K., Naidu, C.D., Rao, M.S. and Jinaga, B.C., (2004). An intelligent character recognizer for Telugu scripts using multiresolution analysis and associative memory. *Image and Vision Computing*, 22(14), pp.1221-1227.
- [25] Ganji, T., Velpuru, M.S. and Dugyala, R., (2021). Multi Variant Handwritten Telugu Character Recognition Using Transfer Learning. In *IOP Conference Series: Materials Science and Engineering* (Vol. 1042, No. 1, p. 012026). IOP Publishing.
- [26] Sujatha, P. and Lalitha Bhaskari, D., (2019). Telugu and Hindi Script recognition using deep learning techniques. *Int. J. Innov. Technol. Explor. Eng*, 8, pp.2278-3075.