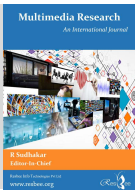
Vol.4 No.2 April 2021

Multimedia Research

Received 04 December, Revised 11 February, Accepted 26 March



Tamil Character Recognition Using K-Nearest-Neighbouring Classifier based on Grey Wolf Optimization Algorithm

Vishal Gali
University of Georgia, Georgia
Ishgali139@gmail.com

Abstract: Optical character recognition (OCR) systems are well-knownand very effectivein the area of the majorityof trendy language recognitions in present data. Not like other languages, the recognition of the Tamil language is highly difficult and thereforesignificantendeavorshave been put in state-of-the-art. Nevertheless, in order to recognize the Tamil characters in an accurate manner, the techniques are not so far developed. Hence, this paper presents a new Tamil Handwritten Character recognition model using 2 important procedures such as recognition as well as pre-processing. The conversion of RGB to grayscale is performed by the pre-processing stage, morphological operations image complementation, binarization with thresholding, as well as linearization. Subsequent to the linearization, the pre-processed image is fed to the recognition through an optimally configured K-Nearest Neighbour. Moreover, the Grey Wolf Optimization (GWO) algorithm is exploited to fine-tune the weights. The developed model performance is evaluated over the conventional techniques regarding various metrics.

Keywords: Handwritten,Preprocessing,Recognition, Tamil, Thresholding

Nomenclature

Abbreviations	Descriptions
CNN	Convolutional Neural Network
OCR	Optical Character Recognition
HTCR	Handwritten Tamil Character Recognition
MMCNN	multi-scale CNN
SVM	Support Vector Machine

1.Introduction

In multi-script documents, script recognition is considered as the important role of OCR [1]. The script is represented as a writing model that consists of a graphical shape and definite symbols. Every feature of script definite attributes which are differing from each other. In an automatic manner, to understand the handwritten documents, the identification of the script is very important. Handwritten script identification received huge attention in the image processing field. The main reason for global digitization is various handwritten books as well as scriptures. The distinctive spatial relation is the foundation of script identification between the strokes of a specific script from each other [3].

Using imaging devices, a machine can transform the handwritten documents that are detained to their corresponding machine-readable as well as searchable documents format frequently indicated as recognition of offline handwritten text. In pattern identification, handwritten text identification is considered the most important active research domain. In diverse real-world applications, the recognition of offline handwritten document text is considered as a vital model, the applications namely bank cheque book processing, postal mail sorting as well as data entry applications. In document images, by recognizing the text contents, the document text identification effortless the automatic data entry applications and therefore it highly minimizes the manual attempt. In implementing a document identifier for any specific language script, the primary step is chosen of appropriate document image database [5]. Here, the initial task is making a novel document image database for the script if the database is not present. The benchmark databases are frequently necessities to implement, evaluate as

https://doi.org/10.46253/j.mr.v4i2.a1

well as analyzing diverse handwritten document identification systems. In handwritten document recognition, numerous benchmark databases for research and advancements are present such as Latin, Malayalam, Tamil, and Arabic [2].

Tamil is considered as the ancient language that is mostly exploited in southern India, Malaysia as well as Srilanka. In the Tamil language, the specialty is each sound pronounced possesses a syllable. The characters' economy to point out a word is minimum in the Tamil language.Althoughnumerous isolated research analysesare performed in Tamil Handwritten document recognition, the main problem is the lack of a document image database in the Tamil language.

The main objective of this research is to present a K-nearest neighbor model which is optimally configured for the Tamil characters recognition and also tunes the weight. Moreover, the GWO Algorithm is introduced. Here, the performance of the developed method shows better performance with the conventional techniques.

2. Literature Review

In 2019, Kavitha B.R. and Srimathi C [1], worked on the handwritten Tamil characters identificationby exploiting CNN in offline mode. From conventional techniques, CNNs were different from the HTCRto extract the features automatically. Moreover, by training mode, a CNN method was developed with Tamil characters in offline mode from scratch.In 2020, SuganyaAthisayamani et al [2], worked on the B-spline curves recognition, which was exploited to identify the 12 vowels in palm leaf manuscripts. The benefits of the B-spline curve wereforcefulness as well as uniqueness. In the Tamil language, each vowel has numerous curves of diverse angles. To identify vowel, the integration of curves wasexploited. In 2017, RiteshSarkhel et al [3], presented a new multi-column MMCNN basedmodel. For recognition with SVM, a deep quad-tree on the basis of a staggering prediction approachwasaugmented with the technique. To maximizerecognition rate amultiple level tree network was exploited as it votes viasoftmax probabilities of all decahexadrantselsequadrants than a single CNN.In 2019, K. Manjusha et al [4], modeled a handwritten character image database for Malayalam language script. The conventional handwritten document image databases were asignificantobligation for advancementas well as objective assessment of diverse handwritten text recognition systems for any language script.For handwritten Malayalam recognition, substantial research efforts werestated in state-of-the-art.Yet, the Malayalam languagewas presented. In 2020, Ayan Kumar Bhunia et al [5], developed anewtechnique of word-level Indic script recognitionby exploiting merely character-level data in the training phase. A multi-modal deep network was used in this paper that uses both the online as well as offline data modality as input. A new conditional multi-modal integration technique was presented to integrate information from the online as well as an offline modality that uses the original modality of data beingsubjected tothe network as well astherefore it integrates adaptively.

3. Proposed Handwritten Recognition model

3.1 System Model

In this research, a new Tamil handwritten recognition model is proposed by adopting 2 important procedures as pre-processing as well as recognition. Moreover, the handwritten document is produced manually which comprises numerals, consonants, and verbals. The architecture model of the developed model is illustrated in Fig 1. In the primary phase, the pre-processing is performed;the input image (Im) is fed to theRGB to grayscale conversion, morphological operations, image complementation, binarization with thresholding, as well aslinearization. Generally, using the RGB to grayscale thresholding, the input image (Im) is in RGB format that is converted into a grayscale image (Im_{in}). In the binarization with thresholding stage, the grayscale image Im_{in} is transformed into a 2D-image (binary) as well asthe global threshold value is calculated over it. Here, Otsu's technique is employedto calculate the image threshold. The threshold calculated (Im_{in}) istransformed to a binary image (white and black) (Im_{Th}) as well as it is complemented (binary imageinversion). Using morphological operations such as open and close, the pixels below 30 are evaded from the complemented image $\left(\mathrm{Im}^{Co}\right)$. The resultant image from the morphological operation is indicated as Im_{mo}. Then, Im_{mo} image is linearized that is gamma-accurate RGB values are linearized. At last, Im_{Lin} indicates the linearized image and it is fed to identificationthrough an optimally configured KNN technique.

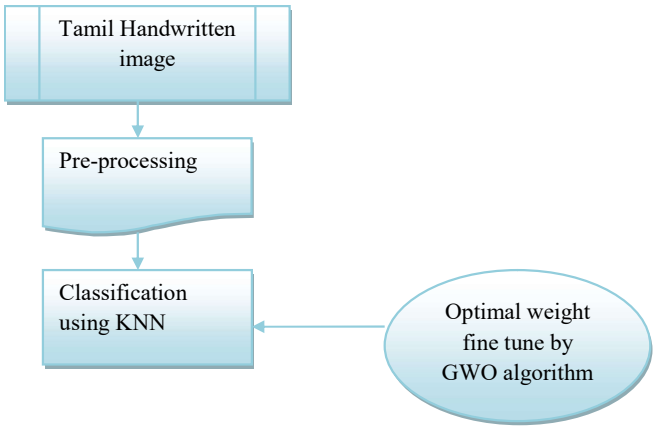


Fig. 1. *Architecture model of Tamil handwritten character recognition*

3.2 Pre-processing

To handle the handwritten document image, the pre-processing phaseinvolvesspecific sub-processes as well as to create it extrasuitablefor the precise character recognition process.

a) RGB to Grayscale conversion

At first, (Im)in RGB format posses its spectral module of red, green as well as blue [6].As exhibited in eq. (1), the conversion of RGB is stated.

$$\mathrm{Im_{in}} = \mathrm{rgb2gray(Im)} \tag{1}$$

The (Im_{in})obtained will range in amidwhiteas well asblack values (i.e. 0 to 255). Hence, it is referred to seizemerelygrayshadesas well as no other color.

b)Binarization with Thresholding

It is a technique of converting $(\mathrm{Im_{in}})$into a digital image that is black as well as white image. At first, the gathered image is smuttybecause of thesmearing, text smudging, and aging [7]. The binarized image is signified as Im_{bi} as well as it is obtainedbased on a threshold value.

Generally, the count of thresholding methods is being used to convert $(\mathrm{Im_{in}})$into a binary image. Amid them, Otsu's technique is highlyefficientas well as it producesbetter binarization outcomeswhilebackground, as well as foreground intensities images, are separated.

Eq. (2) represents the formulation for Optimum thresholdcalculation, in that $H(Th)$ indicates given grayscale image histogram $(\mathrm{Im_{in}})$as well as for complete image, mean intensity is indicated as $m_f(Th)$ land $\mathrm{m_b(Th)}$.

$$\mathrm{Th}^{\mathrm{opt}} = \arg\max \mathrm{H(Th).(1 - H(Th).m_f(Th) - m_b(Th)} \tag{2}$$

Moreover, the global thresholding is carried out by fixing the optimum thresholdamid 0 as well as 1 of a grayscale image. Furthermore, this completesearch for thresholding aids in changingintra-class variancethat is variance occurs within the classas well as itsignified as a summation of weights of two classes stated in Eq. (3).

$$\sigma_{\mathrm{w}}^{2}(\mathrm{Th}) = \sigma_{\mathbf{1}}^{2}(\mathrm{Th}).\omega_{\mathbf{1}}(\mathrm{Th}) + \sigma_{\mathbf{2}}^{2}(\mathrm{Th}).\omega_{\mathbf{2}}(\mathrm{Th}) \tag{3}$$

$\omega_{\mathrm{i}}(\mathrm{t})$ indicates separated 2 classes probability with a threshold Th as well as mean of class [10]. As per Eq. (4), the between-class variance is statedwhere, $\mu_{\mathrm{i}}(\mathrm{Th})$ and $\omega_{\mathrm{i}}(\mathrm{Th})$ represents class meansas well as class probabilities, correspondingly.

$$\begin{aligned}\sigma_{\mathrm{b}}^{2}(\mathrm{Th}) &= \sigma^{2} - \sigma_{\mathrm{w}}^{2}(\mathrm{Th}) \\ &= \omega_{\mathbf{1}}(\mathrm{Th}).\omega_{\mathbf{2}}(\mathrm{Th}).[\mu_{\mathbf{1}}(\mathrm{Th}) - \mu_{\mathbf{2}}(\mathrm{Th})]^{2}\end{aligned} \tag{4}$$

Generally, from the histogram $\mathrm{H(Im_{in})}$, the class probability $\omega_i(Th)$ is calculated, which is stated in Eq. (5). By Eq. (6), the mean class is calculated, where, $\mathrm{x(Im_{in})}$ indicates the value at the histogram bincenter.

$$\omega_{\mathrm{i}}(\mathrm{Th}) = \sum_{\mathbf{0}}^{\mathrm{Th}=\mathbf{1}} \mathrm{H(Im_{in})} \tag{5}$$

$$\mu_1(\mathrm{Th}) = \frac{\left[\sum_{\mathbf{0}}^{\mathrm{Th}=\mathbf{1}} \mathrm{H}(\mathrm{Im}_{\mathrm{in}}).\mathrm{x}(\mathrm{Im}_{\mathrm{in}})\right]}{\omega_1} \quad (6)$$

In between 2 values $\mathrm{Th} = \mathbf{0}$ and $\mathbf{1}$, thresholding operation is attained as:

A pixel turns out to be black if its gray level is $\mathrm{Th} = \mathbf{0}$.

A pixel turns out to be white if its gray level is $\mathrm{Th} = \mathbf{1}$.

In obtaining the textual information aforesaidtechniqueaids even from lower quality images. At last, the grayscale image (Im_{in}) is transformed into a white as well as black image, as well as threshold calculated binary image is indicated as $\mathrm{Im}_{\mathrm{Th}}$.

c) Image Complementation

The ensuing $(\mathrm{Im}_{\mathrm{Th}})$ is complemented (inverse of binary orintensity image). "The image complement block estimates a binary or intensity image complement. The complemented image $\left(\mathrm{Im}^{\mathrm{Co}}\right)$ is fed to binary morphological operation [8].

d) Binary Morphology

The input image evaluation is performed based on the shape is referred to as morphology that isMorphology = study of shape[9]. Moreover, the ensuing Im^{Co} is fed to morphological operations such as opening as well as closing.

Opening: "This operation is erosionensuing by dilation" as well asindicatedin Eq. (7). Moreover, it aids in eradicating the lesser object from the complemented binary image $\left(\mathrm{Im}^{Co}\right)$.

$$\mathrm{Im}^{\mathrm{Co}} \circ \mathrm{S} = (\mathrm{Im}^{Co} \ominus S) \oplus S \quad (7)$$

Closing: "This operation is dilation ensuing by erosion". This helps in removinglesserholes as well as gaps from the complemented binary image Im^{Co}, and it is indicated in Eq. (8).

$$\mathrm{Im}^{\mathrm{Co}} \bullet \mathrm{S} = (\mathrm{Im}^{\mathrm{Co}} \oplus S) \ominus \mathrm{S} \quad (8)$$

Erosion is "Structureseradication of particularsize as well as a shape which is stated as S". As per Eq. (9), the erosion (E) of the image Im^{Co} by structuring elements S is indicated.

$$\mathrm{E} = \mathrm{Im}^{\mathrm{Co}} \ominus S \quad (9)$$

Filling of holes of a particular size as well asshape and size stated using S is expressed as Dilation. As per eq. (10), for the image $\left(\mathrm{Im}^{\mathrm{Co}}\right)$, the dilation (D) by the structuring element S is stated.

$$D = \mathrm{Im}^{Co} \oplus S \quad (10)$$

Hence, by applying morphological operations such as open and close, the pixels below 30 are eradicated. Subsequent to the binary morphology the ensuing image is indicated as Im_{mo} , that is furthermoresubjected to the linearization procedure.

e) Linearization

In analyzing the image quality linearization is very important. It is a "process whichtransforms non-linear image data into linear image data, therefore it is made-up to regress non-linear tone curveisused to linear image data". From linearization, the image ensuing is referredto Im_{Lin} .

4. Objective model

As aforesaid, the KNN technique is optimally configured through fine-tuning weights using the proposed optimization technique.

Thestated objective model is signified in Eq. (17), whereas *Acc* indicates recognition accuracy of KNN.

$$\mathrm{Ob} = \mathrm{Max}(\mathrm{Acc}) \quad (11)$$

4.1 K-Nearest Neighbour

k-NN is a kind of classification where the function is merely approximated locally and all computation is deferred until function evaluation [14].

As this technique lies on distance for classification if the features indicate diverse physical units or come in greatly diverse scales then normalizing the training data can enhance its accuratenessnoticeably

[11]. Both for classification and regression, a practicalmodel can be to allocate weights to the neighbor's contributions, hence nearer neighbors give more to average than large distant ones.

The *k*-NN classifier can be represented as allocating the *k* nearest neighbors a weight $1/k$ and all others *"0"* weight. This can be generalized to weighted nearest neighbor classifiers.

$$\sum_{i=1}^{n} w_{ni} = 1 \tag{12}$$

4.2 Proposed GWO Algorithm

GWO is an optimization algorithm that uses atechnique that is enthused using predatory nature of grey wolves in nature by mimicking their hunting behavior [12].

Numerous swarm intelligence methods are mimic the hunting and searching behaviors of few animals. Nevertheless, GWO imitates the internal leadership hierarchy of wolves, hence, in the searching procedure the best solution position can be comprehensively used by 3 solutions. However, for other swarm intelligence approaches, the optimal solution is explored only leaded by a single solution. Therefore GWO can highly minimize the premature and falling into the local optimum probability.

It is taken into consideration of4 kinds of grey wolves known as $\alpha, \beta, \delta \; and \; \omega$ to understand the leadership hierarchy [13].

$$\vec{X}_g(t+1) = \vec{X}_0(t) - \vec{P} \cdot \vec{R} \tag{13}$$

$$\vec{R} = |\vec{Q} \cdot \vec{X}_0(t) - \vec{X}_g(t)| \tag{14}$$

$$\vec{P} = 2\vec{v}.\vec{y}_1 - \vec{v}$$

$$\vec{Q} = 2.\vec{y}_2 \tag{15}$$

(16)

The values $\vec{y}_1$ and $\vec{y}_2$ possess arbitrary values in the [0,1]

$$\vec{v} = 2 - \left[\frac{2t}{\max\_iter}\right] \tag{17}$$

$$\vec{X}_g(t+1) = \frac{\vec{X}_1 + \vec{X}_2 + \vec{X}_3}{3} \tag{18}$$

5. Result and Discussion

The developed Tamil Character recognition with an optimization technique was developed and discussed in this paper.Here, the outcomes of the proposed and conventional modelsobtainedwereshown. These samples consist of 169 folders with 1000's images. In these folders, there are0 to 155 character images written in the Tamil language. The developed modelwas evaluatedwithconventionaltechniques such as the Genetic Algorithm (GA) and Artificial Bee Colony (ABC) algorithms.

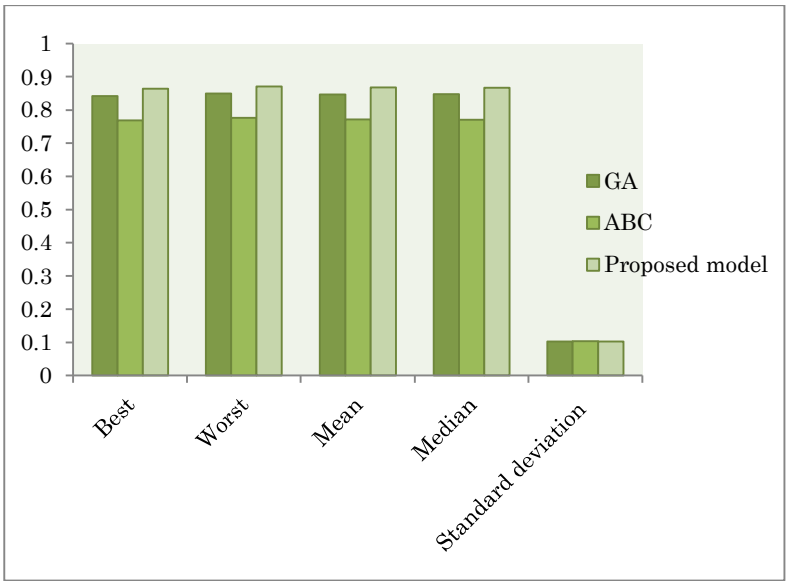


Fig. 2. *Statistical analysis of the adopted as well as existing techniques regarding the accuracy*

Fig 2 illustrates statistical evaluation of developed and other existingtechniques. As the optimizationtechnique is considered in this paper is stochastic in nature, it is very important to execute it 5 times, as well optimal value attained regarding the accuracy is examined. In the best-case performance scenario, the developed modelpossesses the maximum value which is superior to the GA and ABC algorithmscorrespondingly. Here, the proposed method 33% better than the conventional GA, 29% better than the conventional ABC algorithms.From, this analysis, it is evident that the proposed modelshows highlyeffective to the conventional techniques and therefore it is examined to be better in Tamil Character Recognition.

6.Conclusion

In this research, a new Tamil Character recognition model was proposed by using two important procedures as pre-processing as well as recognition. Here, the pre-processing phase consists of five important stages such as the conversion of RGB to grayscale, image complementation, binarization with thresholding, linearization as well as morphological operations. After the completion of the linearization process, the pre-processed image was attained, which was fed to the recognition through optimally configured KNN.To fine-tune the weights, the GWO algorithm is exploited. The developed model performance is evaluated over the conventional algorithms regarding the positive as well as negative metrics.

Compliance with Ethical Standards

Conflicts of interest: Authors declared that they have no conflict of interest.

Human participants: The conducted research follows the ethical standards and the authors ensured that they have not conducted any studies with human participants or animals.

References

[1] B. R. KavithaC. Srimathi,"Benchmarking on offline Handwritten Tamil Character Recognition using convolutional neural networks",Journal of King Saud University - Computer and Information SciencesAvailable online 15 June 2019.

[2] Suganya AthisayamaniA. Robert SinghT. Athithan," Recognition of Ancient Tamil Palm Leaf Vowel Characters in Historical Documents using B-spline Curve Recognition", Procedia Computer Science4 June 2020.

[3] Ritesh SarkhelNibaran DasMita Nasipuri,"A multi-scale deep quad tree based feature extraction method for the recognition of isolated handwritten characters of popular indic scripts", Pattern Recognition, 26 May 2017.

[4] K. ManjushaM. Anand KumarK. P. Soman,"On developing handwritten character image database for Malayalam language script Engineering Science and Technology", an International Journal5 February 2019.

[5] Ayan Kumar BhuniaSubham MukherjeeUmapada Pal,"Indic handwritten script identification using offline-online multi-modal deep network", Information Fusion, 30 October 2019.

[6] [35] Pushpajit Khaire, Praveen Kumar, Javed Imran, “Combining CNN streams of RGB-D and skeletal data for human activity recognition”, Pattern Recognition Letters, vol. 115, pp. 107-116, 1 November 2018.

[7] Xiaolin Li, Peng Wang, Xin-Jian Xu, Gaoxi Xiao, “Universal behavior of the linear threshold model on weighted networks”, Journal of Parallel and Distributed Computing, vol. 123, pp. 223-229, January 2019.

[8] Thomas H. Sharp, Frank G. A. Faas, Abraham J. Koster, Piet Gros, “Imaging complement by phase-plate cryo-electron tomography from initiation to pore formation”, Journal of Structural Biology, vol. 197, no. 2, pp. 155-162, February 2017.

[9] Cao Yuan, Yaqin Li, “Switching median and morphological filter for impulse noise removal from digital images”, Optik, volume. 126, number. 18, page no. 1598-1601, September 2015.

[10] Bency Jacob and Mr. S.B. Waykar,"Binarization and recognition of characters from historical degraded documents",Recent Advances in Computer Science.

[11] Mashaan AlshammariJohn StavrakakisMasahiro Takatsuka,"Refining a k-nearest neighbor graph for a computationally efficient spectral clustering",Pattern Recognition6 February 2021.

[12] Jun DengWei-Le ChenChi-Min Shu,"Correction model for CO detection in the coal combustion loss process in mines based on GWO-SVM",Journal of Loss Prevention in the Process Industries3 March 2021.

Avinash Gopal," Hybrid classifier: Brain Tumor Classification and Segmentation using Genetic-based Grey Wolf optimization", Multimedia Research, vol 3, no 2, April 2020.

Fatima-ezzahra Lagrari," Image Steganography for Pixel Prediction using K-nearest Neighbor", Multimedia Research, vol 3, no 2, April 2020.

.