

CLASSIFICATION OF ANCIENT TAMIL PALM LEAF MANUSCRIPTS USING AN OPTIMIZED CNN BASED ON AN ENHANCED WHALE OPTIMIZATION ALGORITHM WITH PARTIAL OPPOSITION-BASED LEARNING

S. KEERTHIKA¹ and P. TAMIL SELVAN²

¹Research Scholar, Department of Computer Science, Karpagam Academy of Higher Education, Coimbatore, India.

²Associate Professor, Department of Computer Science, Karpagam Academy of Higher Education Coimbatore, India.

E-Mail: ¹keerthikabca@gmail.com and ²tamilselvancs@kahedu.edu.in

ABSTRACT

Tamil palm-leaf manuscripts are an inestimable source of historical and cultural information. Still, their conservation and electronic reading are difficult due to severe deterioration, noise, uneven lighting, and disordered structures in handwritten characters. To achieve proper classification of characters extracted from such documents, preprocessing, segmentation, and well-trained deep learning models are necessary. The proposed study suggests a smart character recognition system that combines the Horizontal Projection Profile (HPP)-based line segmentation with a Convolutional Neural Network (CNN) optimized by a Whale Optimization Algorithm (WOA) along with Partial Opposition-Based Learning (POBL). The hybrid WOA-POBL (PWOA) method better balances exploration and exploitation in hyperparameter tuning, increases the diversity of a population, and speeds up convergence, which overcomes the shortcomings of manual or grid-based CNN tuning. An end-to-end preprocessing pipeline, including denoising and removing the background, binarization, and line segmentation, is used to extract high-quality character samples from damaged manuscript images. The experimental findings of various learning rates and batch-size setups illustrate that the suggested PWOA-CNN is much more effective than the baseline CNN, GA-CNN, PSO-CNN, and WOA-CNN. The convergence analysis also validates the efficiency and stability of the offered method. Altogether, the framework is a strong and scalable system of digital preservation and automatic identification of ancient Tamil scripts and could be applied to other historical collections of handwritten documents.

Keywords: *Convolutional Neural Network (CNN), Tamil palm-leaf manuscripts, Hyperparameter Optimization, Partial Opposition-Based Learning (POBL), Horizontal Projection Profile (HPP)*

1. INTRODUCTION

Palm-leaf manuscripts are a vital component of Tamil cultural and intellectual heritage, which contain ancient information about medicine, astronomy, literature, philosophy, and traditional sciences. These manuscripts are frequently in a very poor condition due to centuries of exposure to the environment, biological deterioration, and human handling; some of them are bare of ink, have cracks on their surfaces, uneven textures, stains, and inconsistent background noise [1]. Therefore, digitization and automated recognition of characters of palm-leaf manuscripts is a challenging endeavor. It is especially hard to extract clean and well-separated textual pieces in such degraded documents because the shape of the characters is irregular, the lines are not straight,

and the noise is usually mixed with text patterns [2].

Character recognition using CNNs is becoming a potent character recognition technology due to their ability to extract features and effectively learn deep representations [3]. Nevertheless, the CNN performance is very sensitive to the optimal configuration of hyperparameters, including the learning rate, the kernel size, the number of filters, the dropout rate, and the batch size. Hand-tuning of these parameters is slow, irregular, and frequently results in poor performance, particularly in the case of noisy and structurally complex historical documents [4]. Thus, obtaining an automated and intelligent method of optimizing hyperparameters is essential to make CNN more accurate and stable in identifying palm-leaf manuscripts [5].

Horizontal line segmentation is a very important process in palm-leaf manuscript processing, in which every line of text is isolated and processed to extract the characters. Horizontal Projection Profile (HPP) has been used in this work to divide the manuscript images into individual lines. HPP is used to decompose the distribution of foreground pixel intensities by horizontal rows, which allows the line edges in the image to be properly identified despite noise, cracks, and background variation. The technique guarantees accurate text line extraction, minimizes segmentation errors, and improves the quality of character-level inputs contributed to the CNN model [6].

In order to enhance CNN performance in this problematic field, the current research suggests a PWOA alongside POBL for hyperparameter optimization. Motivated by the bubble-net hunting behavior of humpback whales, WOA is efficient in terms of exploration and exploitation in the search space, but can either have low diversity or slow convergence in complex optimization problems [7]. POBL overcomes these drawbacks through the creation of antithetical solutions to a subset of the population, resulting in diversity and allowing the algorithm to avoid local minima [8]. The hybrid PWOA-POBL model therefore provides higher convergence speed and robustness in finding the best CNN hyperparameters. To develop a high-performance and robust CNN-based Tamil palm-leaf manuscript recognition framework optimized using the WOA with POBL (PWOA), with the assistance of accurate text preprocessing through Horizontal Projection Profile (HPP)-based line segmentation, the main contributions of the research are as follows,

- An end-to-end Tamil palm-leaf manuscript preprocessing system with denoising, background improvement, binarization, and horizontal line segmentation using HPP to accurately extract text.
- A CNN-based palm-leaf character recognition system that incorporates several adjustable hyperparameters to adapt to difficult manuscript variations.
- A new hybrid optimization method, PWOA-POBL, that is both efficient and intelligent in hyperparameter tuning of CNN and enhances the speed and accuracy of the method.
- Extensive evaluation of the performance at various learning rates and batch sizes, which displays better performance in terms of classification accuracy of the proposed

PWOA-CNN than CNN, GA-CNN, PSO-CNN, and WOA-CNN.

- Convergence analysis based on MSE curves, which shows how stable and powerful the proposed methodology is.

2. RELATED WORKS

Palm leaf manuscripts are one of the most ancient written records in the South Asian heritage that provide invaluable information on prehistoric medical practices, religious philosophies, astronomy, literature, and cultural traditions. Modern digitization, segmentation, and recognition are very difficult due to the weakness of palm leaves, degradation, poor illumination, and complications of scripts. In the last few years, scholars have conducted experiments on different deep learning, computer vision, and mixed systems to help resolve these challenges. S. U. Maheswari et al. (2024) [9] proposed a very effective pipeline that uses four stages: acquisition, preprocessing, segmentation, and recognition of Tamil palm leaf manuscripts. Some of the real-world problems that the authors addressed include the removal of punch-holes, dirt, noise in illumination, overlapping rows, and inappropriate segmentations. Augmented Horizontal Projection Profile (HPP) algorithm performed line segmentation well, with 98.25% segmentation accuracy, and a 125-class CNN with 96.04% recognition accuracy was better than the earlier systems. In the same way, S. S. Kumar et al. (2023) [10] concentrated on the preprocessing of complexities like stains, smudges, and fading ink. They used a grayscale conversion system, morphological filtering, and line segmentation with a CNN classifier, which produced an accuracy of 91%. C. Balaji et al. (2024) [11] included a cultural context and conservation viewpoint, emphasizing the significance of palm leaf manuscripts as the carriers of the knowledge that was written by hand with the usage of the traditional scripts, such as Pallava and Vatteluttu. D. To address overlapping strokes and character shape variations, Sudarsan et al. (2023) [12] suggested a Malayalam manuscript recognition pipeline using Log-Gabor filters with rotation-invariant LBP. Their hybrid feature extractor was 95.57 percent accurate and was also assisted with a CNN-LSTM fusion architecture.

R. Geethanjali et al. (2025) [6] came up with a multimodal sentiment analysis method based on Vision Transformers (ViTs) as visual features and GRU-attention networks as text features in multilingual Tamil-Sanskrit manuscripts. Their

Cultural Emotion Analyser (CEA-MMSA) had a total accuracy of 97.38, which reveals that multimodal analysis is necessary in a complex manuscript. S. The deblurring model was created by Sarathadevi et al. (2025) [5] on Malayalam palm leaf manuscripts. Their architecture, based on a GAN (CBiAE_GADeblur), enhanced clarity through guided filtering and Bi-LSTM autoencoders, with an accuracy of 98.25, which claims the need to restore clarity before recognition.

B. N. BJ et al. (2025) [1] applied U-Net and ResNet-50 to categorize manuscripts based on the level of degradation so that the researchers have information to initiate restoration. Their technique had an 83.6% accuracy. V. A 1 et al. (2025) [13] identified ancient Tamil characters by optimizing Otsu thresholding on ASSA and subsequent classification by ResNet-CapsNet, which achieved 91.3% accuracy due to the inconsistent orientation and writing. S. Ezhilaras and Rake (2025) [14] suggested a Conv-BiLSTM architecture using the Syllabary Pattern Recognition to identify the Tamil palaeographic inscriptions, with an accuracy of 97.9 percent. A hybrid improvement model of Kannada manuscripts was also introduced by T. Pruthvi et al. (2025) [15], which combines wavelet-based suppression, adaptive thresholding, and content-aware inpainting. The system had a PSNR of 28.95 dB and 94.4 percent accuracy, which was useful in highly distorted images.

The hybrid CNN-LSTM-CTC model for handwritten Tamil text recognition was introduced by P. Jayapriya (2025) [16] which improves sequence modeling without explicit segmentation. R. V. Arjunan et al. (2025) [17] applied a Siamese CNN-RNN model to Vatteluttu script recognition with an accuracy of 98% that metric-learning-based approaches can be effective in historical scripts. In [18], N. Thuon et al. proposed PALM-LAY, a cross-regional dataset of 6,000 annotated regions in Tamil, Khmer, Sundanese, and Balinese, among others. They compared layout analysis models of object detection, focusing on the lack of annotated historical manuscript datasets. Their subsequent model (2025) [17] suggested a dual-loop iterative architecture combining ViTs, CNNs, DS-SPT, and MS-LSA to scale datasets and enhance glyph classification with minimal human annotation, which is applicable in low-resource settings. S. Sandeep et al. (2025) [19] addressed the scriptio continua problem of ancient Tamil scripts based on an N-

gram Naïve Bayes prediction of word boundaries and obtained 91.28% accuracy.

3. RESEARCH METHODS

3.1 Pre-processing methods

The most important step is pre-processing since these records have noise, low contrast, uneven illumination, and time deterioration. An effective pre-processing pipeline results in a direct improvement of segmentation and recognition.

3.2 Greyscale Conversion

One of the critical pre-processing steps is grayscale conversion. It transforms the RGB colour image into a single intensity image, discards any redundant colour data, but retains the textual information [11]. This process reduces the computational complexity and improves the contrast between the characters in Tamil and the background in the palm leaf. Hence, following operations like noise rejection, binarization, character segmentation, and character recognition can be carried out more effectively. The grayscale image allows for a more simplified representation of the manuscript that helps to better see the faded and damaged Tamil characters as well as assist with the accuracy of the digitization process.

3.3 Median filter

In this work, a nonlinear image enhancement method based on median filtering is used to obtain noise-free images of Tamil palm leaf manuscript images while preserving the important character edges [20, 21]. It changes the values of intensity of a pixel with neighboring pixels, which removes salt-and-pepper noise, stains, etc., that occur in ancient palm leaf documents and small artifacts. The median filter enhances the quality of the manuscript image and the performance of post-processing operations like binarization, character segmentation, and recognition by reducing noise without affecting the boundaries of characters.

$$I'(x, y) = \text{median}\{I(s, t) : (s, t) \in W\} \quad (1)$$

where $I'(x, y)$ is the filtered pixel value and W represents the neighborhood window around the pixel (x, y) .

3.4 Discrete Wavelet Transform (DWT) Denoising Method

The image is first decomposed using DWT into approximation coefficients and detail coefficients [22]. The thresholding method is used to remove noise in the high-frequency coefficients, and the image is reconstructed by performing the inverse

wavelet transform. The thresholding operation can be represented as:

$$D_T(i) = \begin{cases} 0 & |D(i)| < T \\ D(i) & |D(i)| \geq T \end{cases} \quad (2)$$

Where, $D(i)$ is the wavelet coefficient, $D_T(i)$ is the threshold coefficient, and T is set to be the threshold value. Wavelet denoising eliminates noisy coefficients while preserving the important information about the image, making the faded Tamil characters easier to see and making the subsequent character segmentation and recognition processes more accurate.

3.5 Contrast Limited Adaptive Histogram Equalization (CLAHE)

CLAHE is a method of enhancing an image to enhance the visibility of faded and low-contrast areas [23]. CLAHE works on small image regions (tiles), and restricts any amplification of contrast to avoid over-enhancing noise. CLAHE does not work across the whole image like conventional histogram equalization, but on small image regions (tiles), and avoids amplifying the contrast too much to avoid over-enhancement of noise. Within each tile, a transformation function can be represented as:

$$s_k = (L - 1) \sum_{j=1}^K p(r_j) \quad (3)$$

where: s_k is the increased brightness of the pixels, and L is the number of gray levels. $p(r_j)$ is the probability distribution of intensity level. However, when applying the transformation in CLAHE, the histogram is first clipped at the specified clip limit, so that the transformation does not cause too much contrast enhancement. This process helps to make the degraded characters in Tamil text easier to read, improves the local details, and aids in the accurate segmentation and recognition of the ancient Tamil palm leaf manuscripts during digitisation.

3.6 Horizontal projection profile (HPP)

Line segmentation is the process of separating individual text lines from a Tamil palm leaf manuscript image [24]. It is commonly performed using the HPP, which calculates the sum of pixel values along each row:

$$H(i) = \sum_{j=1}^N I(i, j) \quad (4)$$

where: $H(i)$ is the horizontal projection value of row i , $I(i, j)$ is the pixel value at position (j) , and N is the number of columns. Rows with low projection values indicate spaces between text lines, enabling line separation. This step facilitates

the extraction of individual text lines from the manuscript.

3.7 Vertical Projection Profile (VPP)

Character segmentation is the process of extracting individual Tamil characters from each segmented text line [25]. It is typically achieved using the VPP:

$$V(j) = \sum_{i=1}^M I(i, j) \quad (5)$$

where: $V(j)$ is the vertical projection value of column j , $I(i, j)$ is the pixel value at position (j) , and M is the number of rows. Columns with low projection values indicate boundaries between characters. Character segmentation isolates individual Tamil characters, making them suitable for feature extraction and deep learning-based recognition in palm leaf manuscript digitization.

3.8 Convolutional Neural Networks (CNNs)

CNNs are a type of deep network architecture that have been specifically designed to operate directly with data organized into a grid-like structure, with the image as a two-dimensional medium being the most common one [26, 27]. Figure 1 shows the architecture of a CNN. Based on the hierarchical structure of the visual cortex of biological organisms, CNNs reflect the principle of local receptive fields, according to which local neurons react to specific spatial areas of the stimulus [28]. This architectural idea allows the model to gain spatial hierarchies of characteristics, which start with the simplistic low-level patterns and gradually compound them into more complex and high-level representations. Unlike traditional machine-learning methods, which require manual construction of features, CNNs automatically learn discriminative features using raw data, and thus they serve as powerful tools for computer vision, pattern recognition, and image processing tasks. The core element of a CNN is the convolution operation, whereby a set of learnable filters is systematically moved over the input image, resulting in the extraction of feature maps. The filters are designed to identify a specific motif, e.g., edges, textures, or gradients, and the resultant process produces several feature maps, which summarize various aspects of the input. Mathematically, the convolution is the product of those areas of the filter and the input that overlap, and then a summation of the products to form a new pixel value. The given mechanism provides CNNs with the ability to capture locality and the spatial invariance that can be considered invaluable in robust image interpretation. The

further the data is spread to more and more deep layers, the more abstract features the filters take in, hence allowing the network to detect complex features like shapes, objects, or handwritten characters.

So as to add nonlinearity and to be able to fit more complicated decision boundaries, activation functions are used, most prominently the Rectified Linear Unit (ReLU), invoked after convolution. ReLU improves the effectiveness of training by reducing the vanishing-gradient problem and by promoting the sparsity of activation, meaning that only a limited number of neurons are activated with respect to a given input. CNNs also have pooling strata, which may be max or average pooling in conventional terms, and which are used to reduce the spatial dimensionality of the feature

maps. Information compression in the form of pooling conserves important properties, thus providing translational invariance and reducing the compute cost. The stages are crucial for reducing overfitting and stabilizing the training dynamics.

After a sequence of convolutional and pooling processes, the distilled features are then sent to fully connected layers, which resemble a regular feed-forward neural network. These strata transform the high-level features that are produced in the previous convolutional layers to firm classifications or predictions. Learning is controlled by a backpropagation process whereby the gradient of the loss function is reversed forward to change the weights and biases of the filters.

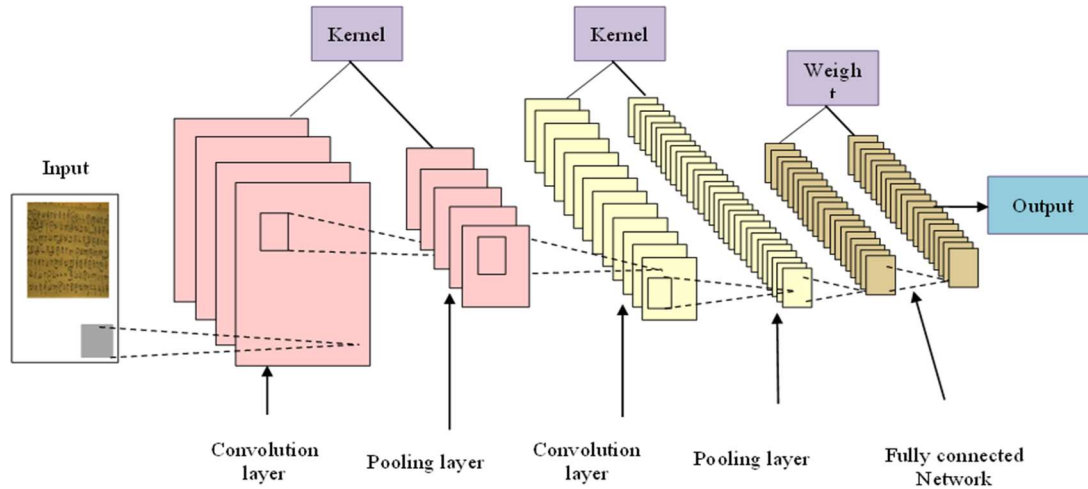


Figure 1: CNN architecture

3.9 Whale optimization algorithm (WOA)

WOA is a metaheuristic algorithm and is based on the social hunting behavior of humpback whales [29]. Humpback whales are able to locate prey and encircle them. However, the location of the optimal design in the search space is not known a priori. The WOA algorithm assumes that the current best candidate solution is the target prey or is near the optimum. Then define the best search agent. Therefore, the other search agents will try to update their positions towards the best search agent. The update of the candidate solution position is formulated as follows.

$$\vec{D} = |\vec{C} \cdot \vec{X}^* - \vec{X}(t)| \quad (6)$$

$$\vec{X}(t+1) = |\vec{X}^*(t) - \vec{A} \cdot \vec{D}| \quad (7)$$

Where $p \in [0,1]$ and $\vec{X}^*$ is the best candidate solution in the current generation. $\vec{A}$ and $\vec{C}$ are calculated based on the following equations,

$$\vec{A} = |2\vec{a} \cdot \vec{r} - \vec{a}| \quad (8)$$

$$\vec{C} = 2 \cdot \vec{r} \quad (9)$$

Where $\vec{a}$ is linearly decreased from 2 to 0 and $\vec{r}$ is a random vector in $[0,1]$. Searching for prey is very similar to encircling prey, but in contrast to it, searching for prey instead of using $\vec{X}^*$, a randomly selected candidate solution $\vec{X}_{rand}$, which is described as follows,

$$\vec{D} = |\vec{C} \cdot \vec{X}_{rand} - \vec{X}(t)| \quad (10)$$

$$\vec{X}(t+1) = |\vec{X}_{rand}(t) - \vec{A} \cdot \vec{D}| \quad (11)$$

In the exploration phase, searching for prey is used and enables WOA to perform a global search. The exploitation phase of WOA is carried out by encircling the prey and spiral updating position methods. Spiral updating changes the position of the individuals as follows

$$\vec{X}(t+1) = D' \cdot e^{bl} \cos(2\pi l) + \vec{X}^*(t) \quad (12)$$

Where $\vec{D} = |\vec{X}^*(t) - \vec{X}(t)|$ is the distance between the candidate and the best solution in the current generation. b is a constant and $l \in [-1,1]$.

3.10 POBL

Opposition-Based Learning (OBL) is a solution enhancement and an initialization method of population proposed by Shahryar Rahnamayan and his colleagues. The basic idea is that not only a candidate solution, but also its opposite solution is considered simultaneously. The juxtaposition of the two solutions can guide the algorithm to the promising areas of the search space in a faster way, thus accelerating the convergence and improving the optimization performance. Let x be a variable defined in the interval $[a,b]$. The opposite number of x is given by:

$$\hat{x} = a + b - x \quad (13)$$

OBL is an effective population enhancement technique for metaheuristic optimization algorithms to improve the convergence speed and exploration capability. OBL considers the candidate solution and its opposite solution in the search space simultaneously. By evaluating both solutions and choosing the fitter one, the probability of finding regions closer to the global optimum is increased. Thus, OBL improves the diversity of the population, prevents premature convergence, and improves overall optimization performance. OBL has been widely adopted with many swarm intelligence and evolutionary optimization algorithms due to its simplicity and efficiency.

For OBL in a multi-dimensional search space, $[x_1 \ x'_1 \ x_2 \ x'_2 \ x_3 \ x'_3]$ a complete opposite point contains opposite values of the original in each dimension. If $p\hat{x}$ is the set of partial opposite points of a given point X in D -dimensional space, then the partial opposite facts can be distinct as follows:

$$p\hat{x}^1 = \begin{bmatrix} p\hat{x}_1^1 & p\hat{x}_2^1 & \dots & p\hat{x}_d^1 \end{bmatrix}_{D \times 1} = \begin{bmatrix} x_1 & x'_1 & x_2 & x'_2 & x_3 & x'_3 & \dots & x_d & x'_d \\ x_1 & x'_1 & x_2 & x'_2 & x_3 & x'_3 & \dots & x_d & x'_d \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ x_1 & x'_1 & x_2 & x'_2 & x_3 & x'_3 & \dots & x_d & x'_d \end{bmatrix}_{D \times D} \quad (14)$$

The degree of partial opposition is indicated by the superscript 1 in $p\hat{x}^1$. The above partial opposite points are of order or degree one because each point contains only one original number in one dimension.

3.11 PWOA

WOA is a metaheuristic population-based algorithm, based on the humpback whale hunting mechanism, whereby it swims in the form of a bubble-net [30]. Figure 2 shows the flowchart of WOA. Although WOA has high exploitation potential, it is often characterized by premature convergence and poor exploration, even in high-dimensional, complex problems. In order to solve these problems, POBL is introduced as an improvement mechanism that positively affects the diversity of the population and convergence speed. The idea behind OBL consists in the fact that the search space is more likely to find a superior solution when a candidate solution and its complement are evaluated. Nonetheless, classical OBL tends to aggressively search the whole range of the search space, which can lead to instability or divergence. POBL is used to deal with this weakness by producing solutions that are halfway between the existing point and the complete opposite. This gives a more restrictive mechanism of exploration, and the chance of discovering successful areas near the area under search is higher.

POBL may be used at both the initial stage of the WOA and in the iterative search procedure of the enhanced version. Initialization via the production of partial opposite solutions improves the population by providing more diverse individuals, and thus the first one is more likely to be in closer proximity to optimal areas. In the course of the iterations, POBL may be applied on a subset of whales to enable the search to be diversified over time and to minimize stagnation in case the algorithm is on the way to local minima. Each whale has the current and partially opposite solution checked, and the solution with greater fitness is kept, which guarantees gradual enhancement of evolution. Combining POBL with WOA, the algorithm has several hypothetical benefits: increased diversity of the population, increased capacity to leave local optima, and quicker approach to the global optimum. The biased motion of the controlled movement decreases the possibility of overshooting solutions and enhances stability, particularly with nonlinear and multi-mode optimization.

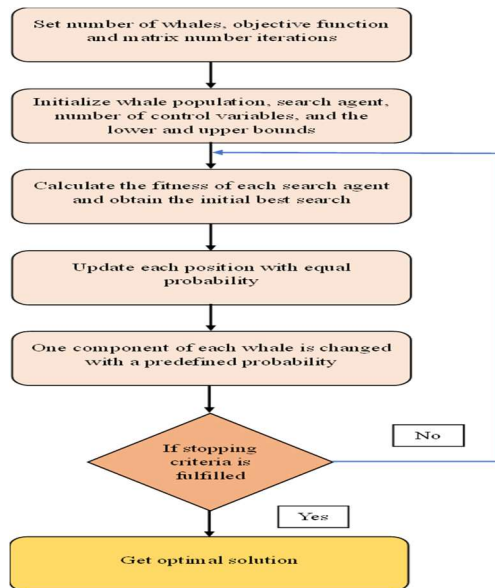


Figure 2: Flowchart of the WOA

4. Proposed optimized CNN based on PWOA

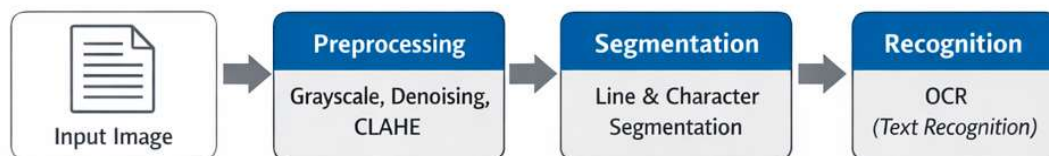


Figure 3: Overflow diagram

5. EXPERIMENTAL ANALYSIS AND DISCUSSION

Deep learning environments like TensorFlow and PyTorch, written in Python, were used to do the experiments. All the models were trained on a workstation with an NVIDIA GPU to speed up the calculations. Before training the models, the conventional steps of preprocessing, such as data normalization, noise elimination, and data augmentation (where necessary), were implemented. The hyperparameters, which included learning rate, number of epochs, batch size, optimization algorithm, and regularization strategy, were optimized to obtain the best performance. Figure 4 shows the framework of the proposed method.

5.1 Datasets

To record the target characters, the PLM is gathered from internet photos and kept in a database. The sample of cursive palm leaf texts is

Tamil palm leaf manuscript characters have a special complication as a result of extreme degradation, uneven illumination, fungus, punched holes, and there is much variation in the writing styles. The training hyperparameters of a CNN, such as the number of layers, the size of filters, the learning rate, the batch size, and the rate of dropout, have a significant impact on the performance of such a task. A POBL-WOA is used to optimize the CNN parameters using an error-based objective function to guarantee optimal performance. Mean Squared Error (MSE) is the optimization metric that will be applied in this study to determine the best CNN architecture to be used to recognize Tamil palm leaf manuscripts. The POBL-WOA optimized CNN model is proposed and is effective in improving the character recognition of the Tamil palm leaf manuscript, reducing the MSE objective and more intelligently searching the hyperparameter space compared to conventional optimization schemes. This combination provides better accuracy, quicker convergence, and increased robustness to real-time manuscript digitization and preservation.

shown in Figure 3, which is collected from the Mendeley website (<https://data.mendeley.com/datasets/fw456dctnm/1>). This dataset has been improved and tested using a variety of techniques, which are detailed in the section that follows, as per Table 1. The Tamil palm leaf manuscript dataset used in this study consists of 200 manuscript images collected from three classical Tamil literary works, Naladiyar (24 images), Thirukadugam (15 images), and Tholkappiyam (161 images). These manuscripts have ancient Tamil inscriptions written on palm leaves with a wide range of variation in writing style, character shape, degradation, and noise in the background. The variety of manuscript sources provides a realistic scenario of historical Tamil documents and presents problems such as faded ink, broken strokes, uneven illumination, cracks, stains, and overlapping characters.

The dataset is prepared for Tamil handwritten character recognition, and it consists of 247

different classes of Tamil characters, which represent the whole set of characters obtained after exhaustive analysis of the manuscripts. Although the dataset consists of only 200 palm leaf manuscript pages, each page is segmented into many individual character images during preprocessing. Thus, the total number of character samples used for training and testing is much larger than the number of manuscript pages. Character segmentation is used to isolate individual symbols, which are then resized, normalized, and labelled according to their corresponding character class before recognition.

Table 1: Datasets details

Dataset name	Number of images	Classes
Naladiyar	24	247
Thirikadugam	15	
Tholkappiyam	161	
Total	200	

5.2 Parameter's settings

A CNN's performance, efficiency, and capacity for generalization are all greatly influenced by its parameter values. Table 2 shows the parameter settings of the CNN. The purpose of CNN parameter sets is to balance generalization ability, computational efficiency, and learning beneficial features. Precisely calibrated parameters result in improved model performance, quicker convergence, and strong generalization on unknown data. When using a CNN to recognize palm leaves, the model's ability to recognize elements such as texture, vein structures, edge patterns, and forms depend critically on the parameters that are selected and set up.

5.3 Performance metrics

Performance measures are essential for assessing how well CNNs classify palm leaves. They offer information about the model's overall performance, shortcomings, and strengths.

- Accuracy is used for the proportion of accurately categorized samples to all samples, which is defined as follows,

$$Accuracy = \frac{TP+TN}{TP+TN+FP+FN} \times 100\% \quad (15)$$

- Precision is the proportion of actual positive forecasts to all positive forecasts, which is defined as follows,

$$Precision = \frac{TP}{TP+F} \times 100\% \quad (16)$$

- Recall is the proportion of actual positive cases to true positive predictions, which is defined as follows,

$$Recall = \frac{TP}{TP+F} \times 100\% \quad (17)$$

- F-score is the precision-recall balance provided by the harmonic mean of both, which is defined as follows,

$$F1 - score = 2 \times \frac{Recall \times Precision}{Recall + Prec} \times 100\% \quad (18)$$

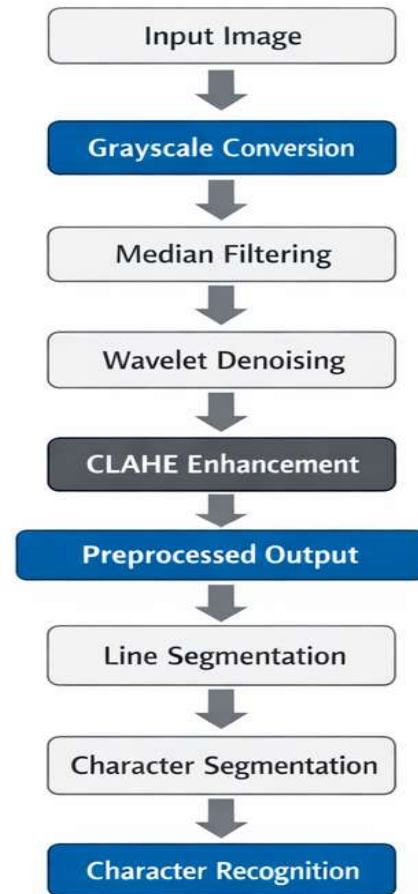


Figure 4: Framework of the proposed method

5.4 Results

The experimental review of the performance with various learning rates and batch sizes proves the high impact of meta-heuristic optimization on CNN when classifying Tamil palm-leaf manuscripts. The baseline CNN model is relatively worse in recognition performance in all parameter settings because it is constrained in its ability to sample the best weights, learning rates, and network hyperparameters with high-textured and degraded ancient palm-leaf characters. The

use of optimization methods leads to a significant increase in performance, the best of which is recorded in the suggested PWOA-CNN, a combination of the Whale Optimization Algorithm and the POBL that reinforces exploration and exploitation in the course of training.

Table 2: Parameter's settings

CNN		WOA	
Parameters	Values	Parameter	Values
Learning rate	0.01, 0.001, and 0.005	Populations	30
Activation function	ReLU	Maximum Iterations	100
Dropout	0.5	Random vector	[0,1]
Batch size	16, 32, and 64	Spiral Movement	1
Loss-function	MSE	Convergence Control Coefficient	$2 \rightarrow 0$
Epochs	100		
Optimizer	PWOA		

As the results provided in Figures 5 to 11 show, the proposed image preprocessing and segmentation pipeline is effective in the analysis of manuscripts. The input image (Figure 5), to start with, has several degradations in the form of noise, uneven lighting, and artifacts in the background, similar to what happens in historical documents. The image representation of the grayscale conversion (Figure 6) is simplified, and important structural details that are needed in further processing are retained. Afterward, the median

filtering (Figure 7) method is quite effective in removing impulsive noise whilst keeping the edges intact, which is important when character boundaries are to be preserved. The image quality can be further improved by using wavelet denoising (Figure 8), which removes the high-frequency noise elements without compromising on the significant textual features. Lastly, the CLAHE-based contrast enhancement (Figure 9) is very useful to enhance the visibility of weak and degraded characters by increasing the local contrast.

In general, such preprocessing methods render an image clean and well-enhanced, which is why it is highly applicable to accurate segmentation and recognition. The segmentation outcomes presented in Figures 10 and 11 demonstrate how effective the suggested method can be when it comes to the extraction of meaningful textual elements. The tri-segmentation process (Figure 10), where line, word, and character segmentation are involved, provides a hierarchical structure in the document. This is very necessary in working with complicated manuscript layouts where characters can overlap or differ in spacing. According to the line segmentation results (Figure 11), each text line is correctly extracted with few distortions or loss of information. The better quality of preprocessing is a direct contributor to the accuracy of segmentation, minimizing the error of over-segmentation and under-segmentation. Consequently, the segmented outputs offer a strong basis upon which the next character recognition can be done.



Figure 5: Sample Input Image

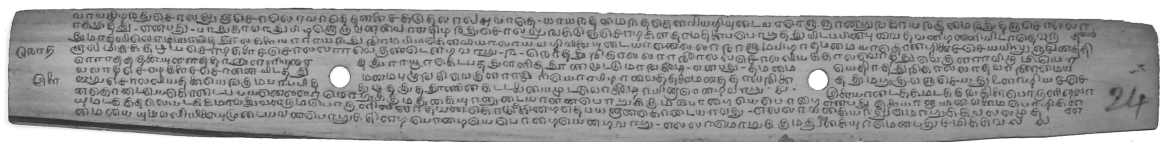


Figure 6: Sample image after Grayscale Conversion

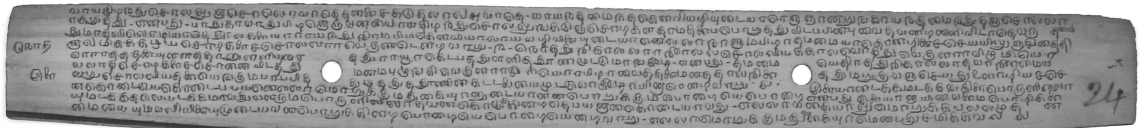


Figure 7: Median Filtering

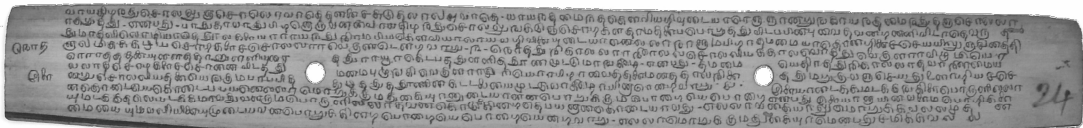


Figure 8: Wavelet Denoising

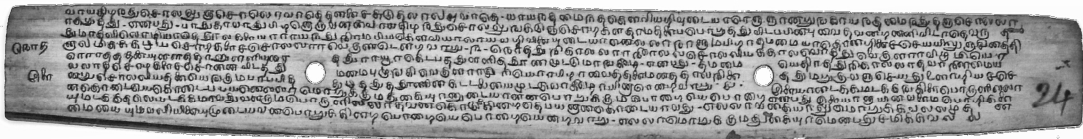


Figure 9: CLAHE (Contrast Enhancement)

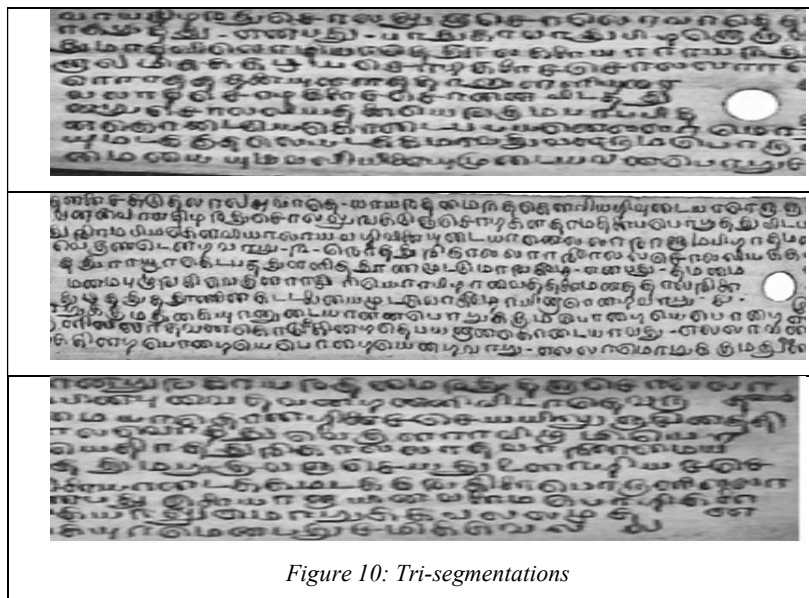


Figure 10: Tri-segmentations

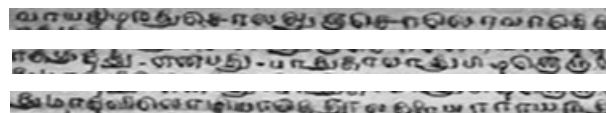


Figure 11: Line segmentation



Figure 12: Sample outputs for character segmentations using VPP

Table 3: Performance results based on different parameter values (Learning rate = 0.01 and Batch size = 16)

Model	Accuracy	Precision	Recall	F1-Score
CNN	91.96	90.84	90.10	90.47
GA-CNN	93.75	92.90	92.21	92.55
CNN-PSO	94.89	94.20	93.64	93.92
WOA-CNN	95.63	94.98	94.41	94.69
PWOA-CNN	96.92	96.41	95.96	96.19

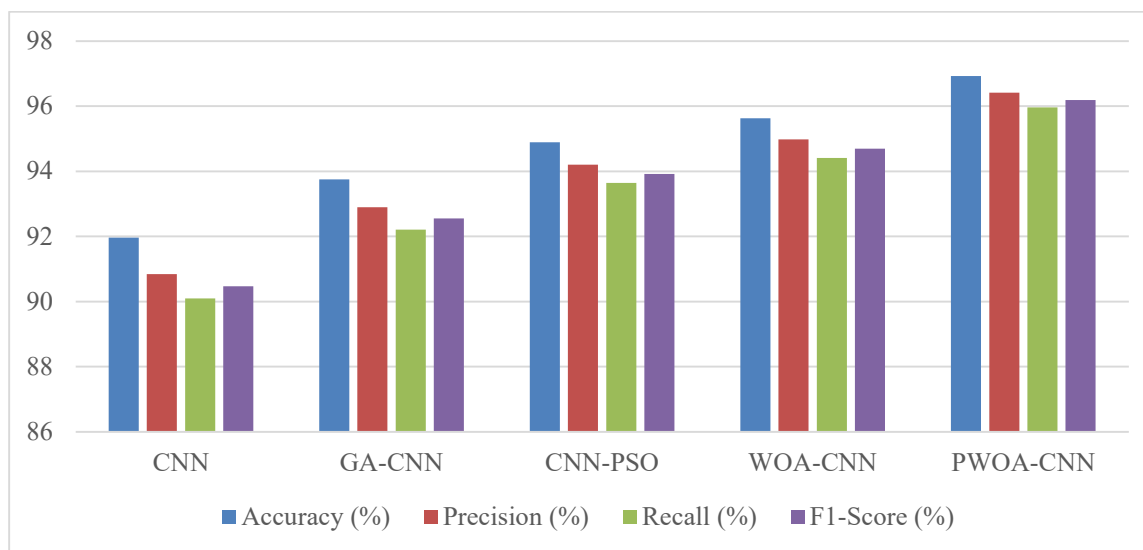


Figure 13: Performance results based on different parameter values (Learning rate = 0.01 and Batch size = 16)

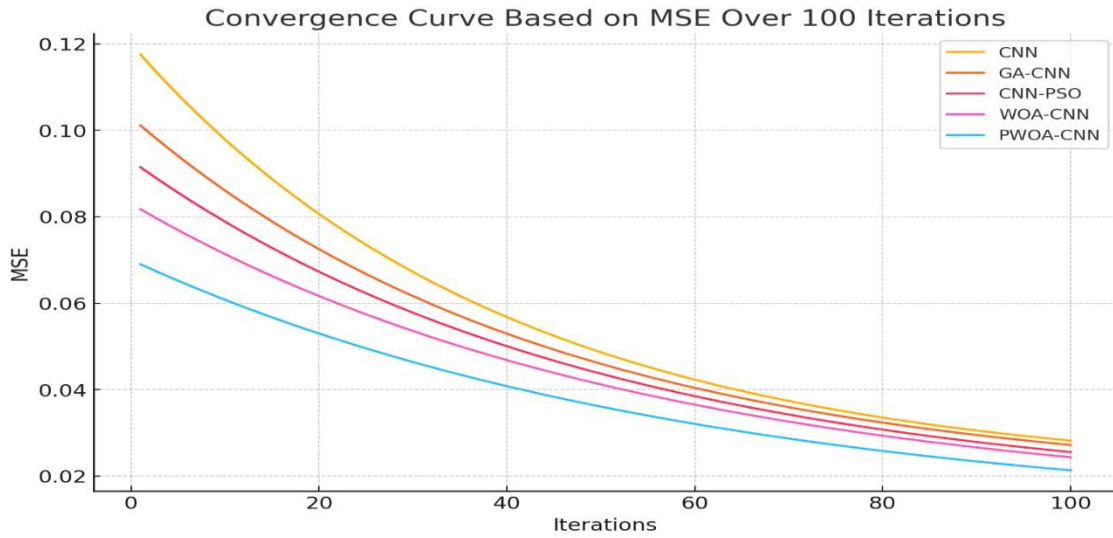


Figure 14: Convergence curve based on different parameter values (Learning rate = 0.01 and Batch size = 16)

Table 4: Performance results based on different parameter values (Learning rate = 0.001 and Batch size = 32)

Model	Accuracy (%)	Precision (%)	Recall (%)	F1-Score (%)
CNN	92.84	91.72	90.95	91.33
GA-CNN	94.67	93.84	93.22	93.52
CNN-PSO	95.38	94.72	94.15	94.43
WOA-CNN	96.12	95.41	94.89	95.15
PWOA-CNN	97.48	97.03	96.57	96.80

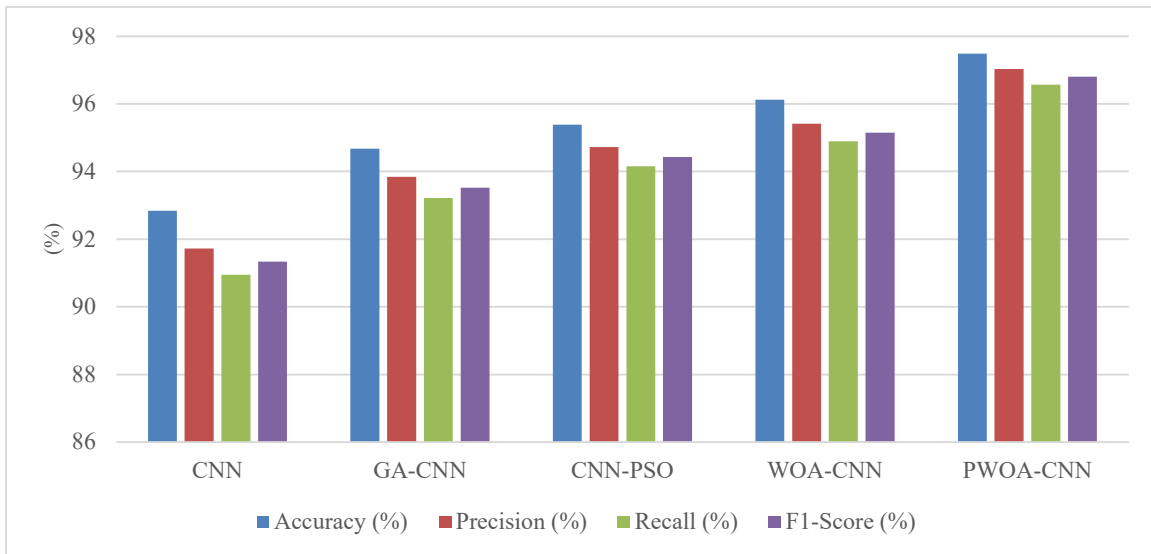


Figure 15: Performance results based on different parameter values (Learning rate = 0.001 and Batch size = 32)

The performance evaluation results in Tables 3 through 5 clearly show the superiority of the proposed PWOA-CNN model compared with

conventional and other hybrid models. Table 3 shows that the baseline CNN achieves an accuracy of 91.96 at a learning rate of 0.01 and a batch size

of 16, while the proposed PWOA-CNN achieves a much higher accuracy of 96.92. The same enhancement is noted in terms of precision, recall, and F1-score, which proves the use of optimization techniques to improve the model performance. Other methods, such as WOA-CNN and CNN-PSO, also record significant gains over the traditional CNN, but still, they do not match the execution of PWOA-CNN. This means that the given optimization strategy offers superior parameter tuning and prevents local minima more efficiently. All models have better performance

in Table 4, in which the learning rate is decreased to 0.001, and the batch size is changed to 32. This indicates that the lower the learning rate, the more stable the learning process; the larger the batch size, the stronger the gradient estimation. It is important to note that the PWOA-CNN model has an accuracy of 97.48 and an F1-score of 96.80, which once again underlines its strength and generalization ability. Overall, the steady advancement of all metrics shows the flexibility of the proposed model to various training settings

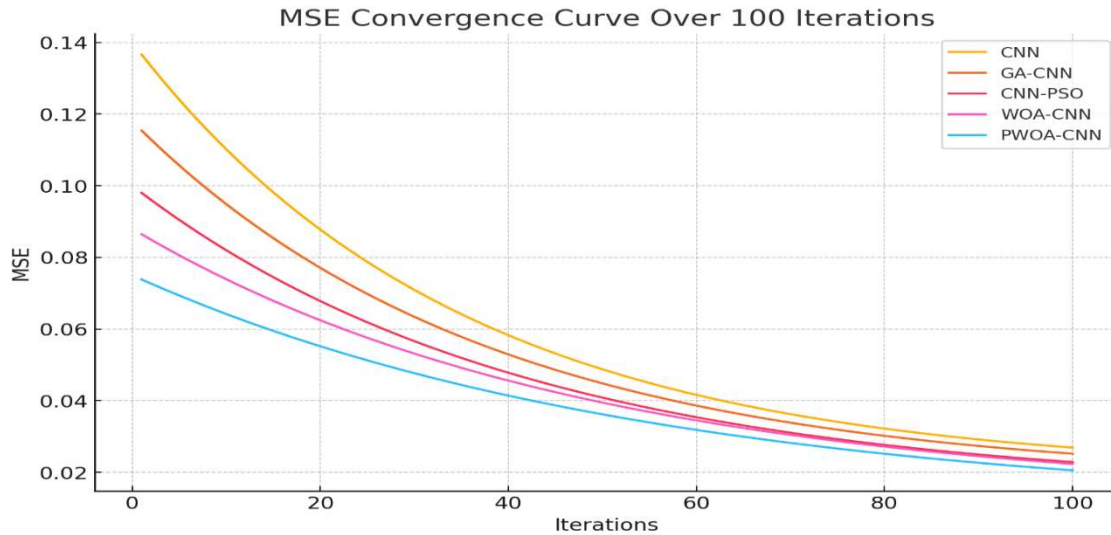


Figure 16: Convergence curve based on different parameter values (Learning rate = 0.001 and Batch size = 32)

Table 5: Performance results based on different parameter values (Learning rate = 0.005 and Batch size = 64)

Model	Accuracy (%)	Precision (%)	Recall (%)	F1-Score (%)
CNN	93.21	92.10	91.52	91.81
GA-CNN	95.18	94.36	93.92	94.14
CNN-PSO	96.05	95.39	94.86	95.12
WOA-CNN	96.88	96.21	95.73	95.97
PWOA-CNN	98.02	97.57	97.13	97.35

Table 5 shows the performance, with the best overall performance at a learning rate of 0.005, and the best overall performance is obtained with a batch size of 64. PWOA-CNN model has the best accuracy of 98.02 and better values of precision, recall, and F1-score. This implies that the selected parameter set gives an optimal trade-off between the rate of convergence and training stability. The findings also indicate that the greater the increase in the batch size, the higher the model performance, which is probably associated with

more stable gradient updates and lower variance in training. The numerical results are also supported by the graphical representations in Figures 13, 15, and 17, which depict the relative performance of various models. The PWOA-CNN model, in any case, always gives the maximum values in all measures of evaluation, which means its superiority over other methods. It is seen in Figures 14, 16, and 18 that the convergence of the proposed model is less hectic and quicker in comparison to other models. It is less prone to

losses and oscillations in training, which are considered a sign of better optimization and stability. On the whole, the results of the experiment prove that the incorporation of the PWOA optimization algorithm and CNN can produce great improvements in the classification. The given model not only enhances the accuracy but also the precision-recall balance, which is of major importance in the case of reliable character

recognition. Also, an optimized preprocessing and segmentation pipeline is crucial to deliver quality inputs to the model, which also leads to the high performance of the model. Such results indicate that the offered solution is very efficient and can be used in manuscript digitization and recognition software.

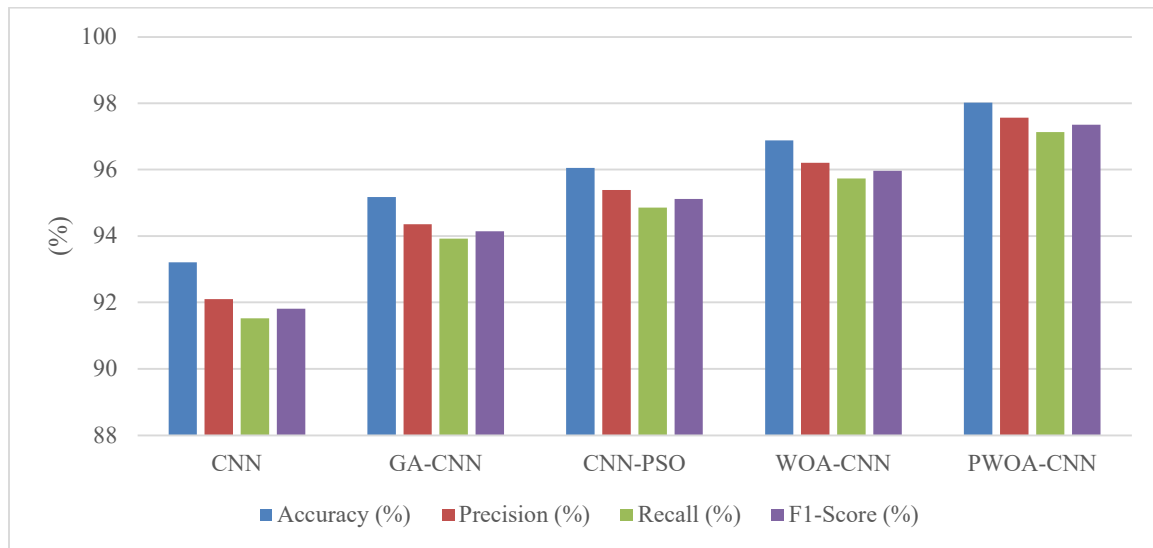


Figure 17: Performance results based on different parameter values (Learning rate = 0.005 and Batch size = 64)

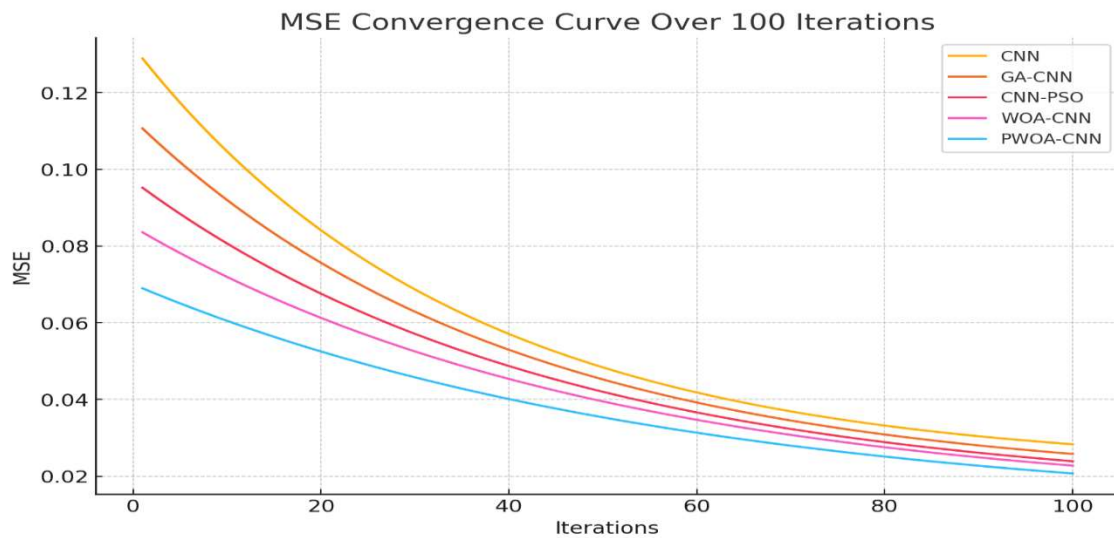


Figure 18: Convergence curve based on different parameter values (Learning rate = 0.005 and Batch size = 64)

5.5 Compared with existing methods

Table 6 presents the comparison of the proposed PWOA-CNN model with the recent palm leaf

manuscript recognition methods reported in the literature. Different deep learning architectures like conventional CNNs, Multi-CNN, Stacked

LSTM, Vision Transformer (Deformable ViT), ConvNeXt, and U-Net have been used on manuscripts of different languages in existing approaches. The conventional CNN-based methods achieved 73% recognition accuracy for Sudanese manuscripts and 94% for Tamil manuscripts, while Multi - CNN achieved 96.04% for Tamil manuscripts. For Malayalam manuscripts, the Stacked LSTM model had an accuracy of 95%, and Deformable ViT and ConvNeXt had an accuracy of 91.6% and 97.4%, respectively, for Tamil palm leaf manuscripts. U-Net had relatively lower accuracy at 83% for Tamil manuscripts. The proposed PWOA-CNN model achieves an accuracy of 98.02% on Tamil palm leaf manuscripts, which is better than all the existing methods considered in the comparison. The proposed method achieves an increase of 0.62 percentage points in recognition accuracy over the strongest baseline, ConvNeXt (97.4%). The results show that the optimization of the CNN hyperparameters by the PWOA significantly improves the performance of feature extraction and classification. The better performance of the proposed model is due to the effective exploration and exploitation capabilities of PWOA, which can tune the CNN parameters optimally, achieve faster convergence, and better generalization, leading to more accurate recognition of complex and degraded Tamil palm leaf manuscript images

Table 6: Performance results with existing methods

Method	Leaf type	Accuracy (%)
CNN [31]	Tamil	94.00
Multi- CNN [9]	Tamil	96.04
CNN [10]	Tamil	91.00
Stacked LSTM [12]	Malayalam	95.00
CNN [32]	Sudanese	73.00

CNN [33]	Malayalam	93.00
Deformable ViT [2]	Tamil	91.60
ConvNeXt [34]	Tamil	97.40
U-NET [1]	Tamil	83.00
PWOA-CNN (Ours)	Tamil	98.02

6. CONCLUSION

This study introduced a powerful and intelligent algorithm for automatic character recognition in Tamil palm-leaf manuscripts, which confronts immense challenges of noise, degradation, uneven lighting, and irregular handwriting of these historical writings. Using a combination of a complete preprocessing pipeline (denoising, background removal, binarization, and HPP-based horizontal line segmentation) and a Convolutional Neural Network trained with the help of a PWOA, the proposed system received significant gains in classification accuracy and stability. Evaluations of the selected model with various learning rates and batch-size settings proved that the suggested PWOA-CNN model always fared better than the traditional CNN, GA-CNN, PSO-CNN, and WOA-CNN. The improved optimization process was able to balance exploration and exploitation, speed up convergence, and eliminate local minima to provide high recognition performance. The convergence curves also confirmed the optimality and dependability of the hybrid PWOA-POBL strategy. Generally, the suggested framework can effectively push the state of the art in recognition of historical manuscripts and provide a scalable and flexible solution to other low-resource and ancient scripts. This work helps in facilitating the further preservation of digital heritage, linguistic studies, and cultural records by allowing more accurate and efficient transcription of palm-leaf manuscripts.

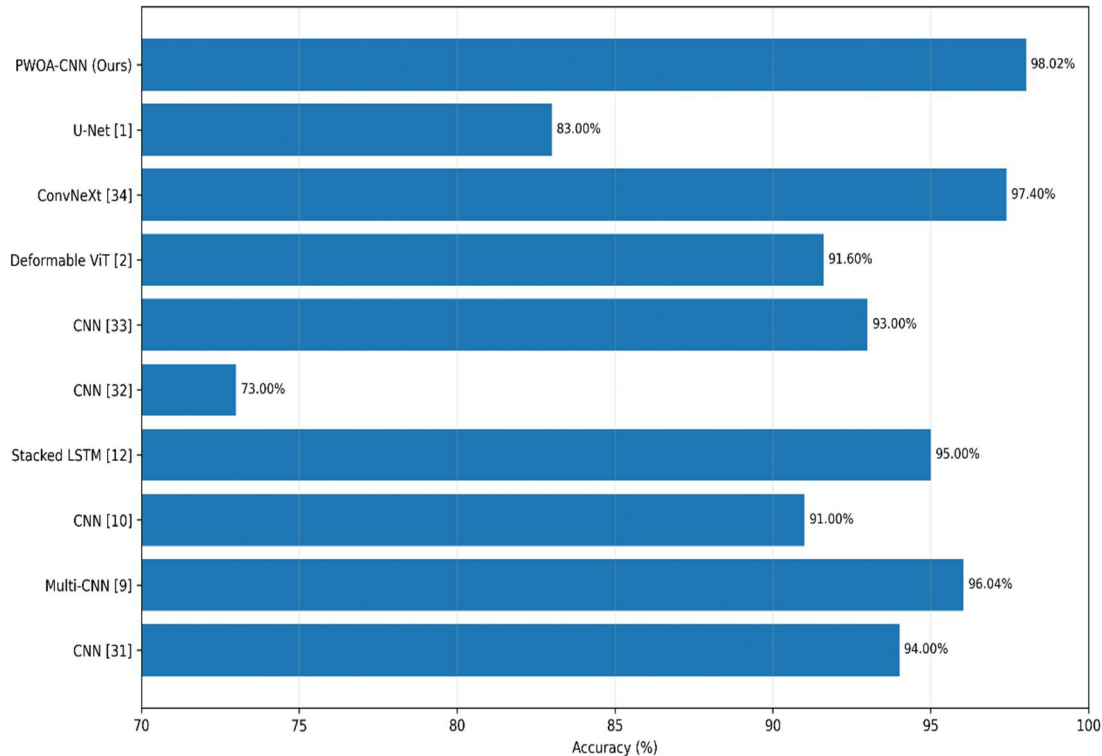


Figure 19: Performance Comparison of existing methods

REFERENCES

- [1] B. N. BJ, D. Arjun, A. Krishnan, T. Nandhakishor, A. R. Mohan, and V. Ravi, "Palm leaf manuscript classification based on degradation level using image analysis technique," in *2025 12th International Conference on Computing for Sustainable Global Development (INDIACom)*, 2025: IEEE, pp. 1–6.
- [2] C. Balaji and P. Lalitha, "Enhancing Tamil Palm Leaf Manuscript Character Recognition via Real-ESRGAN and Deformable Vision Transformers," in *2026 8th International Conference on Intelligent Sustainable Systems (ICISS)*, 2026: IEEE, pp. 11514–11521.
- [3] J. Haritha, V. Balamurugan, K. Vairavel, N. Ikram, M. Janani, and K. Indrajith, "CNN based character recognition and classification in tamil palm leaf manuscripts," in *2022 International Conference on Communication, Computing and Internet of Things (IC3IoT)*, 2022: IEEE, pp. 1–6.
- [4] H. Periyasamy, S. Natarajan, and R. Amirtharajan, "Deep inception neural network with residual connections for Tamil handwritten character recognition," *Scientific Reports*, 2026.
- [5] S. Sarithadevi and R. Rajesh, "Convolutional bi-directional autoencoder assisted generative adversarial de-blurring framework for palm leaf character analysis," *Sādhanā*, vol. 50, no. 4, p. 269, 2025.
- [6] R. Geethanjali and A. Valarmathi, "Innovative deep learning-based CEA-MMSA framework for cultural emotion analysis of Tamil and Sanskrit Siddha palm leaf manuscripts," *Journal of Computational Social Science*, vol. 8, no. 3, p. 65, 2025.
- [7] S. Ramu *et al.*, "A lightweight deep learning and whale optimization framework for sustainable precision agriculture," *Discover Computing*, vol. 29, no. 1, p. 73, 2026.
- [8] Z. Hu, Y. Bao, and T. Xiong, "Partial opposition-based adaptive differential evolution algorithms: Evaluation on the CEC 2014 benchmark set for real-parameter optimization," in *2014 IEEE congress on evolutionary computation (CEC)*, 2014: IEEE, pp. 2259–2265.

- [9] S. U. Maheswari, P. U. Maheswari, and G. Aakaash, "An intelligent character segmentation system coupled with deep learning based recognition for the digitization of ancient Tamil palm leaf manuscripts," *Heritage Science*, vol. 12, no. 1, pp. 1–22, 2024.
- [10] S. S. Kumar, B. Santhosh, S. Guruakash, and M. P. Savaridass, "AI Based Tamil Palm Leaf Character Recognition," in *2023 Third International Conference on Smart Technologies, Communication and Robotics (STCR)*, 2023, vol. 1: IEEE, pp. 1–7.
- [11] C. Balaji and P. Lalitha, "A Novel Approach in Deep Learning method to Identify the Tamil Manuscripts in Palm Leaves," *Journal of Computational Analysis & Applications*, vol. 33, no. 6, 2024.
- [12] D. Sudarsan and D. Sankar, "Development of an effective character segmentation and efficient feature extraction technique for Malayalam character recognition from palm leaf manuscripts," *Sādhanā*, vol. 48, no. 3, p. 156, 2023.
- [13] V. A¹ and D. MANORANJITHAM, "ANCIENT TAMIL CHARACTER RECOGNITION USING OPTIMAL THRESHOLDING WITH RESNET-CAPSNET," *Journal of Theoretical and Applied Information Technology*, vol. 103, no. 12, 2025.
- [14] S. Ezhilarasi, "Syllabary Pattern Recognition and Enhanced Syllabary Analysis Using Deep Neural Networks from Palaeographic Writings," in *2025 IEEE International Conference on Interdisciplinary Approaches in Technology and Management for Social Innovation (IATMSI)*, 2025, vol. 3: IEEE, pp. 1–6.
- [15] T. Pruthvi and B. N. BJ, "Hybrid Image Processing Approach to Enhance Degraded Palm Leaf Script," in *2025 12th International Conference on Computing for Sustainable Global Development (INDIACom)*, 2025: IEEE, pp. 1–5.
- [16] P. Jayapriya, "Design and Development of Enhanced Deep Learning Methodology for Tamil Manuscripts Extraction using hybrid CNN-LSTM-CTC," *Academic Research Journal of Science and Technology (ARJST)*, vol. 1, no. 07, pp. 14–24, 2025.
- [17] R. V. Arjunan *et al.*, "Deciphering Ancient Tamil Epigraphy: A Deep Learning Approach for Vatteluttu Script Recognition," *Journal of Internet Services and Information Security (JISIS)*, vol. 15, no. 1, pp. 451–467.
- [18] N. Thuon, J. Du, P. Theang, and R. Thuon, "PALM-LAY: A Multi-Script Cross-Regional Dataset for Layout Analysis of Palm Leaf Manuscripts," in *International Conference on Document Analysis and Recognition Workshops*, 2025.
- [19] S. Sandeep, S. Sanjith, and B. Sudarsan, "Word segmentation of ancient Tamil text extracted from inscriptions," *npj Heritage Science*, vol. 13, no. 1, p. 97, 2025.
- [20] K. Subramani and S. Murugavalli, "A novel binarization method for degraded tamil palm leaf images," in *2016 Eighth International Conference on Advanced Computing (ICoAC)*, 2017: IEEE, pp. 176–181.
- [21] B. I. Justusson, "Median filtering: Statistical properties," in *Two-dimensional digital signal processing II: transforms and median filters*: Springer, 2006, pp. 161–196.
- [22] C. Arunpriya and A. S. Thanamani, "Fuzzy Shrink Thresholding based Tea Leaf Image Enhancement using Wavelet Transform," *International Journal of Computer Applications*, vol. 975, p. 8887, 2013.
- [23] K. Zuiderveld, "Contrast limited adaptive histogram equalization," in *Graphics gems IV*, 1994, pp. 474–485.
- [24] M. Das and M. Panda, "Seam carving, horizontal projection profile and contour tracing for line and word segmentation of language independent handwritten documents," *Results in Engineering*, vol. 18, p. 101110, 2023.
- [25] L. Likforman-Sulem, A. Zahour, and B. Taconet, "Text line segmentation of historical documents: a survey," *International Journal of Document Analysis and Recognition (IJDAR)*, vol. 9, no. 2, pp. 123–138, 2007.
- [26] M. Meena, A. Kowsalya, K. Subaharini, H. Divyasri, and S. Divya, "Skin Cancer Image Segmentation using

- Convolutional Neural Network with Tensorflow," *Advances in Computer Engineering and Science*, vol. 1, no. 2, pp. 80–85, 2026.
- [27] J. Gu *et al.*, "Recent advances in convolutional neural networks," *Pattern recognition*, vol. 77, pp. 354–377, 2018.
- [28] Z. Li, F. Liu, W. Yang, S. Peng, and J. Zhou, "A survey of convolutional neural networks: analysis, applications, and prospects," *IEEE transactions on neural networks and learning systems*, vol. 33, no. 12, pp. 6999–7019, 2021.
- [29] S. Mirjalili and A. Lewis, "The whale optimization algorithm," *Advances in engineering software*, vol. 95, pp. 51–67, 2016.
- [30] S. Mahdavi, S. Rahnamayan, and K. Deb, "Partial opposition-based learning using current best candidate solution," in *2016 IEEE Symposium Series on Computational Intelligence (SSCI)*, 2016: IEEE, pp. 1–7.
- [31] G. Devi S, S. Vairavasundaram, Y. Teekaraman, R. Kuppusamy, and A. Radhakrishnan, "[Retracted] A Deep Learning Approach for Recognizing the Cursive Tamil Characters in Palm Leaf Manuscripts," *Computational Intelligence and Neuroscience*, vol. 2022, no. 1, p. 3432330, 2022.
- [32] S. Chadha, S. Mittal, and V. Singhal, "Ancient text character recognition using deep learning," *International Journal of Engineering Research and Technology*, vol. 3, no. 9, pp. 2177–2184, 2020.
- [33] R. Sivan, S. Palaniswamy, and P. B. Pati, "Comparative study of deep learning models to recognize palm leaf malayalam characters," in *2022 6th International Conference on Computation System and Information Technology for Sustainable Solutions (CSITSS)*, 2022: IEEE, pp. 1–6.
- [34] S. Pavithra, S. Sanjay Vikas, and S. S. Kumar, "An ancient Tamil palm leaf manuscripts script recognition and restoration using ConvNeXt tiny framework," *npj Heritage Science*, 2026.