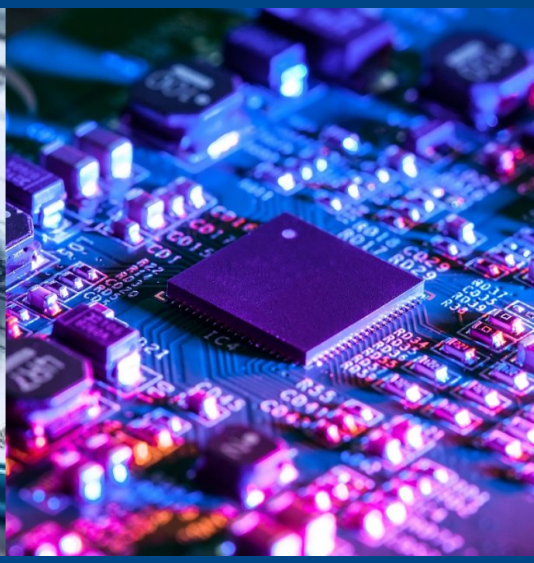
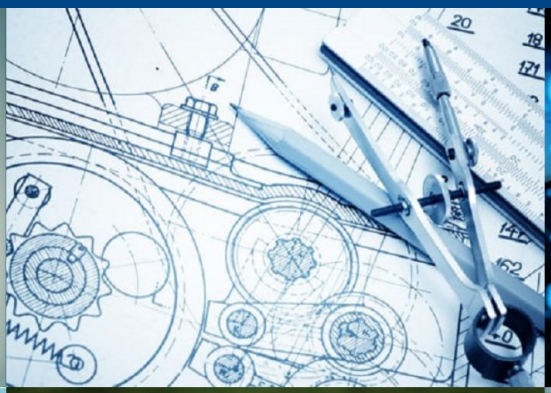
ISSN: 2582-7219

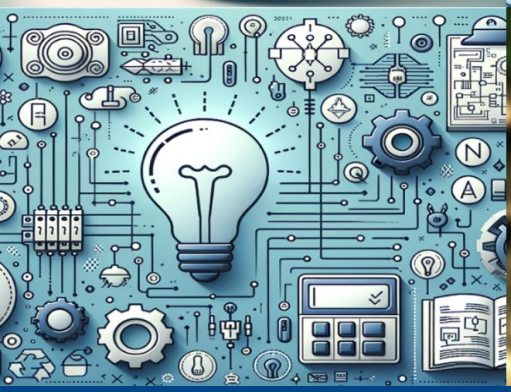




# International Journal of Multidisciplinary Research in Science, Engineering and Technology

*(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)*



**Impact Factor: 8.206** **Volume 9, Issue 4, April 2026**

© 2026 IJMRSET | Volume 9, Issue 4, April 2026 | DOI:10.15680/IJMRSET.2026.0904393

ISSN: 2582-7219 | www.ijmrset.com | Impact Factor: 8.206| ESTD Year: 2018|

International Journal of Multidisciplinary Research in Science, Engineering and Technology (IJMRSET)

(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)

# Tamil Inscription Digitization using Heuristic Optimization Algorithm

**G.David Raj, Bharani Kumar K**

Assistant professor Department of Computer Applications, B.S Abdur Rahman Crescent Institute of Science and Technology, Chennai, Tamil Nadu, India

MCA 2nd Year Department of Computer Applications, B.S Abdur Rahman Crescent Institute of Science and Technology, Chennai, Tamil Nadu, India

**ABSTRACT:** Ancient Tamil inscriptions are valuable sources of historical and cultural information, but interpreting them is often difficult due to surface damage, weathering, and variations in ancient Tamil scripts. Traditional methods rely on manual analysis by historians and epigraphists, which is time-consuming and requires specialized expertise. To address this problem, this project proposes a mobile application for Tamil inscription digitization and multilingual translation using an AI-based Bidirectional Long Short-Term Memory (BiLSTM) model with heuristic optimization. The system captures inscription images using a mobile device camera or dataset input and applies image preprocessing techniques such as noise reduction, contrast enhancement and grayscale conversion to improve image quality. Character segmentation techniques are then used to extract Tamil characters from the inscription images. The extracted characters are recognized using a BiLSTM deep learning model that analyzes sequential patterns and converts them into digital Tamil text. A heuristic optimization module is applied to improve recognition accuracy and reduce prediction errors. Finally, the recognized Tamil text is translated into user-selected languages using an AI-based multilingual translation model. The translated output is displayed through a mobile application interface, making ancient inscription content easier to understand for researchers and the general public. The proposed system provides an efficient and accessible solution for digitizing, preserving, and interpreting Tamil inscriptions using modern artificial intelligence techniques.

**KEYWORDS:** Tamil Inscription Recognition, Image Processing, BiLSTM, Heuristic Optimization, Character Segmentation, Multilingual Translation, Deep Learning, Artificial Intelligence.

## I. INTRODUCTION

Tamil inscriptions are an important part of the cultural and historical heritage of South India. These inscriptions are commonly found on temple walls, stone monuments, and ancient structures. They provide valuable information about the history, culture, religion and administration of ancient Tamil societies. Many historical events, donations to temples and details about rulers were recorded in the form of inscriptions. However, understanding these inscriptions is difficult because they are written in ancient Tamil scripts and are often damaged due to weather conditions and aging.

Traditionally, historians and researchers study these inscriptions manually by visiting historical sites and examining the engraved text. This process involves taking photographs of inscriptions, manually identifying characters and translating them into modern languages. Although this method helps preserve historical information, it requires expert knowledge in ancient Tamil scripts and takes a significant amount of time. In addition, many inscriptions are partially damaged, making manual interpretation even more difficult.

With the advancement of digital technologies, image processing and artificial intelligence techniques can be used to automate the process of inscription recognition. Image preprocessing techniques can improve the clarity of inscription images by reducing noise and enhancing contrast. After preprocessing, segmentation techniques can be used to extract individual characters from the inscription image. These extracted characters can then be analyzed using machine learning and deep learning models for recognition.Deep learning models such as Bidirectional Long Short-Term Memory (BiLSTM) networks are widely used for recognizing sequential patterns in text data. BiLSTM models analyze character sequences in both forward and backward directions, which improves recognition accuracy. In thisproject,

ISSN: 2582-7219 | www.ijmrset.com | Impact Factor: 8.206| ESTD Year: 2018|

International Journal of Multidisciplinary Research in Science, Engineering and Technology (IJMRSET)

(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)

the BiLSTM model is used to recognize Tamil characters from segmented inscription images and convert them into digital Tamil text. To further improve the recognition accuracy, heuristic optimization techniques are applied to reduce errors caused by damaged or unclear characters.

The proposed system develops a mobile application that captures Tamil inscription images and processes them using image preprocessing, character segmentation, and deep learning-based recognition techniques. The recognized Tamil text is then translated into user-selected languages using an AI-based multilingual translation model. This system helps digitize and preserve ancient Tamil inscriptions while making their meaning easier to understand for researchers, students and the general public.

## II. RELATED WORKS

Research on Tamil character recognition and inscription analysis has gained significant attention in recent years due to the need to preserve historical and cultural heritage. Several researchers have proposed different techniques using image processing, machine learning and deep learning methods to recognize characters from ancient Tamil inscriptions and manuscripts.

Bhuvaneswari and Kathiravan (2024) proposed a deep learning-based approach for recognizing and analyzing ancient Tamil inscriptions. Their method combines advanced preprocessing techniques with Convolutional Neural Networks (CNN) and Vision Transformers to improve the recognition accuracy of Tamil characters from temple inscriptions. The study demonstrated high recognition accuracy and showed that deep learning techniques can effectively handle complex character patterns in ancient scripts.Priya et al. (2022) introduced a self-adaptive lion optimization algorithm combined with transfer learning for Tamil character recognition from stone inscriptions. In their approach, image preprocessing techniques such as brightness and contrast enhancement were applied to improve the quality of inscription images. The optimized transfer learning model was used to recognize Tamil characters from the processed images. However, the system's performance was affected by poor lighting conditions and unclear inscriptions.

Suganya and Murugavalli (2021) proposed a feature selection-based approach using shape and Hough transform techniques to extract important features from Tamil characters. The extracted features were classified using machine learning algorithms such as k-Nearest Neighbor and Naïve Bayes classifiers. Although the method improved feature extraction accuracy, it lacked an effective parameter tuning mechanism to optimize classification performance.Manigandan et al. (2021) presented a recognition system that combines Optical Character Recognition (OCR) and Natural Language Processing (NLP) techniques for identifying Tamil characters from stone inscriptions. Their method used feature extraction algorithms such as Scale Invariant Feature Transform (SIFT) and classification techniques like Support Vector Machine (SVM). While the system showed promising results, the recognition accuracy depended heavily on the quality and resolution of the input images.

Ezhilarasi and Mahesari (2021) developed a Bidirectional Long Short-Term Memory (BiLSTM) model for recognizing Tamil words from ancient inscription scripts. The model analyzed linguistic patterns and part-of-speech structures to improve word prediction accuracy. Although the BiLSTM approach demonstrated the potential of deep learning for Tamil script recognition, the system required further improvements to achieve higher accuracy when dealing with damaged or incomplete inscriptions.These existing research works highlight the importance of applying image processing and deep learning techniques for inscription recognition. However, many of the existing methods face challenges such as low recognition accuracy, dependency on high-quality images, and limited support for multilingual translation. Therefore, the proposed system focuses on improving recognition accuracy using a BiLSTM-based model with heuristic optimization and integrating a multilingual translation module to make ancient Tamil inscriptions more accessible to users.

## III. PROPOSED SYSTEM

The system is designed to provide an intelligent solution for digitizing and translating ancient Tamil inscriptions using modern artificial intelligence techniques. It focuses on automating the process of capturing inscription images, recognizing Tamil characters, and converting them into understandable multilingual text. This approach helps reduce the dependency on manual interpretation by historians and provides a faster and more reliable way of understanding historical inscriptions.

ISSN: 2582-7219 | www.ijmrset.com | Impact Factor: 8.206| ESTD Year: 2018|

International Journal of Multidisciplinary Research in Science, Engineering and Technology (IJMRSET)

(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)

Initially, inscription images are captured using a mobile device camera or uploaded from existing historical datasets. These images are then processed using image preprocessing techniques such as grayscale conversion, noise reduction and contrast enhancement. These operations improve the clarity and readability of the inscription characters, which may be affected by surface erosion, uneven lighting conditions, and environmental damage.

After enhancing the image quality, character segmentation techniques are applied to identify the text region and extract individual Tamil characters from the inscription image. Segmentation separates characters from the stone surface background and prepares them for the recognition stage. The extracted characters are then analyzed using a deep learning model based on Bidirectional Long Short-Term Memory (BiLSTM). This model is capable of learning sequential patterns of Tamil characters and converting them into digital Tamil text with improved recognition accuracy.

To further enhance recognition performance, heuristic optimization techniques are incorporated. The optimization module refines the predicted results by correcting possible recognition errors that may occur due to damaged or unclear characters. This process improves the reliability and accuracy of the character recognition process.

After the Tamil text is successfully generated, it is processed by an AI-based multilingual translation module. This module translates the recognized Tamil text into user-selected languages such as English, Hindi, or other languages. The translated output is then displayed through the mobile application interface, allowing users to easily understand the meaning of ancient inscriptions. This framework supports the digital preservation of Tamil heritage and enables easier access to historical information through modern technology.

## IV. ADVANTAGES

The developed framework provides an automated approach for digitizing and interpreting ancient Tamil inscriptions using modern artificial intelligence techniques. By integrating image processing and deep learning models, the system reduces the dependency on manual interpretation, which traditionally requires expert knowledge in ancient Tamil scripts and epigraphy. This automation significantly improves the efficiency of inscription analysis and reduces the time required for transcription and translation.

Another important advantage is the use of advanced preprocessing techniques that enhance the quality of inscription images. These techniques reduce noise, improve contrast and highlight engraved characters that may be affected by erosion or environmental damage. As a result, the visibility of inscription characters is improved, which contributes to more accurate character segmentation and recognition.

The use of the BiLSTM deep learning model improves the recognition performance by analyzing the sequential structure of Tamil characters. BiLSTM networks process data in both forward and backward directions, allowing the system to better understand character patterns and contextual relationships. This capability enhances the accuracy of character recognition, particularly when dealing with complex or partially damaged inscriptions.

In addition, heuristic optimization techniques are incorporated to further refine recognition results and minimize prediction errors. This optimization process helps correct misclassified characters and improves the reliability of the extracted text. Consequently, the overall accuracy and robustness of the recognition process are significantly improved.

Another key benefit of the framework is its ability to translate recognized Tamil text into multiple languages using an AI-based translation module. This feature allows users to understand ancient inscriptions without requiring specialized knowledge of the Tamil language. By supporting multilingual translation, the system increases the accessibility of historical information for researchers, students and the general public.

Overall, the framework contributes to the digital preservation of Tamil cultural heritage by providing an efficient and intelligent solution for inscription digitization and interpretation. The integration of image processing, deep learning, and multilingual translation technologies enables the development of a comprehensive system that facilitates historical research and promotes better understanding of ancient Tamil inscriptions.

© 2026 IJMRSET | Volume 9, Issue 4, April 2026 | DOI:10.15680/IJMRSET.2026.0904393

ISSN: 2582-7219 | www.ijmrset.com | Impact Factor: 8.206| ESTD Year: 2018|

International Journal of Multidisciplinary Research in Science, Engineering and Technology (IJMRSET)

(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)

## V. SYSTEM ARCHITECTURE

### TECHNIQUES USED

**Image Preprocessing**

Image preprocessing techniques are applied to improve the quality of inscription images before performing character recognition. Ancient Tamil inscriptions are often affected by noise, surface erosion, and uneven lighting conditions. To address these challenges, preprocessing operations such as grayscale conversion, noise reduction, contrast enhancement and image normalization are performed. These techniques help enhance the visibility of engraved characters and prepare the images for further processing.

**Character Segmentation**

Character segmentation is used to identify text regions and separate individual Tamil characters from the inscription image. This process involves detecting the boundaries of characters and isolating them from the stone background. Edge detection and region identification techniques are used to extract characters as individual segments, which are then used as input for the recognition model.

**BiLSTM-based Character Recognition**

A Bidirectional Long Short-Term Memory (BiLSTM) model is used for recognizing Tamil characters from segmented images. BiLSTM networks analyze sequential patterns in both forward and backward directions, allowing the model to understand contextual relationships between characters. This capability improves recognition accuracy and enables the system to convert inscription characters into digital Tamil text.

**Heuristic Optimization**

Heuristic optimization techniques are incorporated to improve recognition accuracy and reduce prediction errors. These techniques refine the recognition results by correcting misclassified characters and enhancing the reliability of the output generated by the deep learning model.

**Multilingual Translation**

An AI-based multilingual translation technique is used to convert the recognized Tamil text into other languages. Neural machine translation models analyze the linguistic structure of the extracted text and generate translations in languages such as English and Hindi. This module enables users to understand ancient inscriptions without requiring knowledge of Tamil scripts.

## VI. MODULE DESCRIPTION

**Inscription Image Acquisition**

The inscription image acquisition module is responsible for collecting Tamil inscription images that serve as the primary input for the system. Images can be captured using a mobile device camera or uploaded from historical datasets containing temple inscriptions. The acquired images are stored in the system database for further processing. This module ensures that inscription images are collected in a suitable format for analysis and forwarded to the preprocessing stage for quality enhancement.Image Preprocessing and Enhancement

The image preprocessing and enhancement module improves the quality of the captured inscription images. Ancient inscriptions are often affected by environmental conditions such as erosion, noise, and uneven lighting, which can reduce the visibility of engraved characters. To address these challenges, preprocessing operations such as grayscale conversion, noise reduction, contrast enhancement and image normalization are applied. These techniques enhance the clarity of inscription characters and prepare the images for accurate segmentation.

**Character Segmentation and Extraction**

The character segmentation and extraction module identifies the text region in the inscription image and separates individual Tamil characters from the stone surface background. Segmentation techniques are applied to detect the boundaries of characters and isolate them as individual image segments. Edge detection and region identification methods are used to accurately locate the character structures within the inscription. The extracted character segments are then provided as input to the recognition module.

ISSN: 2582-7219 | www.ijmrset.com | Impact Factor: 8.206| ESTD Year: 2018|

International Journal of Multidisciplinary Research in Science, Engineering and Technology (IJMRSET)

(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)

**Tamil Character Recognition using BiLSTM**

The character recognition module converts segmented inscription characters into digital Tamil text using a deep learning model based on Bidirectional Long Short-Term Memory (BiLSTM). BiLSTM networks are capable of learning sequential patterns in textual data by processing information in both forward and backward directions. This capability allows the model to effectively recognize complex Tamil script structures and contextual relationships between characters. The recognized characters are then combined to generate digital Tamil text.

**Multilingual Translation and Output Display**

The multilingual translation and output display module converts the recognized Tamil text into different user-selected languages. An AI-based neural machine translation model is used to analyze the linguistic structure of the extracted Tamil text and generate translations in languages such as English and Hindi. The translated output is then displayed through the mobile application interface, allowing users to easily understand the content of ancient inscriptions and access historical information in their preferred language.

## VII. RESULT AND DISCUSSION

The developed system was tested to evaluate its ability to recognize Tamil inscription characters and translate the extracted text into multiple languages. Image preprocessing techniques such as noise reduction, grayscale conversion, and contrast enhancement improved the clarity of inscription images. This enhancement helped the system identify characters more accurately, even when the inscriptions were partially damaged or affected by environmental conditions.

The character segmentation module successfully separated individual Tamil characters from the inscription images. These segmented characters were processed by the BiLSTM-based recognition model, which analyzed character patterns and converted them into digital Tamil text. The results show that the deep learning model is capable of recognizing Tamil characters effectively, improving the overall accuracy of the recognition process.

After recognition, the multilingual translation module converted the Tamil text into other languages such as English and Hindi. The translated output was displayed through the application interface, allowing users to easily understand the meaning of the inscriptions. The results demonstrate that the system provides an efficient approach for digitizing and interpreting ancient Tamil inscriptions while improving accessibility to historical information.

## VIII. CONCLUSION

An intelligent framework for digitizing and translating ancient Tamil inscriptions has been developed using image processing and deep learning techniques. The system captures inscription images, enhances their quality through preprocessing operations and extracts individual Tamil characters using segmentation methods. A BiLSTM-based recognition model converts the segmented characters into digital Tamil text with improved accuracy. The recognized text is further processed using an AI-based multilingual translation module, enabling the generation of translated outputs in different languages. The developed framework contributes to the digital preservation of Tamil cultural heritage and improves accessibility to historical inscription information for researchers and the general public.

## REFERENCES

1. S. Bhuvaneswari and K. Kathiravan, “Enhancing Epigraphy: A Deep Learning Approach to Recognize and Analyze Tamil Ancient Inscriptions,” Neural Computing and Applications, Springer, vol. 36, pp. 19839–19861, 2024.
2. R. Devi, S. Kannan, and P. Rajendran, “Recognition of Tamil Characters from Palm Leaf Manuscripts using Convolutional Neural Networks,” Springer Proceedings in Computer Science, 2022.
3. R. Priya, M. Kumar, and S. Ravi, “Self-Adaptive Lion Optimization Algorithm with Transfer Learning for Tamil Character Recognition,” IEEE International Conference on Intelligent Systems, 2022.
4. S. Jayanthi and R. Thenmalar, “ResNet Two-Stage Bottleneck Architecture for Tamil Character Recognition,” Springer Lecture Notes in Networks and Systems, 2022.
5. S. Manigandan, R. Kumar, and S. Murugan, “Tamil Inscription Recognition using OCR and Natural Language Processing Techniques,” IEEE International Conference on Image Processing, 2021.
6. K. Ezhilarasi and R. Mahesari, “Bi-LSTM Based Prediction of Tamil Words from Ancient Inscription Scripts,”

ISSN: 2582-7219 | www.ijmrset.com | Impact Factor: 8.206| ESTD Year: 2018|

International Journal of Multidisciplinary Research in Science, Engineering and Technology (IJMRSET)

(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)

IEEE Conference on Machine Learning Applications, 2021.

7. S. Suganya and S. Murugavalli, "Feature Selection for Tamil Character Recognition using Shape and Hough Transform," Elsevier Journal of Pattern Recognition Letters, 2021.

A. Athisayamani, S. Babu, and K. Kumar, "B-Spline Curve Based Tamil Character Recognition using Neural Networks," Elsevier Journal of Image Processing, 2020.

8. K. Kowsalya and R. Periasamy, "Modified Artificial Neural Network Model for Tamil Character Recognition," Springer Journal of Soft Computing, 2020.
9. M. Vellingiriraj, R. Rajesh, and P. Kumar, "Tamil Character Recognition using Image Zoning and Edge Detection Techniques," IEEE International Conference on Pattern Recognition, 2019.
10. P. Raj, S. Kumar, and V. Anand, "Recognition of Tamil Handwritten Characters using Support Vector Machine," IEEE International Conference on Data Science and Engineering, 2019.

INTERNATIONAL STANDARD SERIAL NUMBER INDIA

# INTERNATIONAL JOURNAL OF MULTIDISCIPLINARY RESEARCH

## IN SCIENCE, ENGINEERING AND TECHNOLOGY

| Mobile No: +91-6381907438 | Whatsapp: +91-6381907438 | ijmrset@gmail.com |

www.ijmrset.com