MSW MANAGEMENT -Multidisciplinary, Scientific Work and Management Journal
ISSN: 1053-7899
Vol. 36 Issue 1, 2026, Pages: 3357-3363

GlyptNet: Deep Learning Recognition and Translations of Ancient Tamil Stone inscriptions

[1]S. INDHUMATHI
ASSISTANT PROFESSOR
KGISL INSTITUTE OF TECHNOLOGY
COIMBATORE, TAMIL NADU, INDIA
indhumathisme@gmail.com

[2]D. EMUNA
STUDENT
KGISL INSTITUTE OF TECHNOLOGY
COIMBATORE, TAMIL NADU, INDIA
emuna 22csa34@kgkite.ac.in

[3]YN. GOKUL KRISHNAN
STUDENT
KGISL INSTITUTE OF TECHNOLOGY
COIMBATORE, TAMIL NADU, INDIA
gokulkrishnan 22csa36@kgkite.ac.in

[4]N KIRAN
STUDENT
KGISL INSTITUTE OF TECHNOLOGY
COIMBATORE, TAMIL NADU, INDIA
kiran_22csa56@kgkite.ac.in

[5]N KISHAN
STUDENT
KGISL INSTITUTE OF TECHNOLOGY
COIMBATORE, TAMIL NADU, INDIA
Kishan_22csa57@kgkite.ac.in

ABSTRACT

Old Tamil stone inscriptions are found to be worn out due to the passage of time, weather and natural growth and this makes them extremely hard to read using normal OCR systems. The letters are normally distorted, stained, or worn out. To address this issue, in this paper, a hybrid system is proposed that incorporates several kinds of image information such as 3D depth information, infrared images, and normal colour images to enhance the recovery and reading of ancient Tamil inscriptions used in South Indian temples. We restore the damaged surfaces first by using depth-based correction, infrared enhancement to detect any concealed letters and contrast enhancement algorithms to exploit erosion, lichen growth and uneven surfaces of stones. Once this is done, the better photos are subjected to a deep learning model that processes ancient Tamil characters. The model is a combination of sequence-based recognition algorithm, convolutional and recurrent networks in order to achieve the text without errors. The system was experimented on over 1,850 inscriptions on temples between the 12th and 17th centuries in Tamil Nadu. Experts were involved in the verification of each sample. These findings indicate that the character recognition accuracy is 92.4% which is 18.3 her higher than the traditional OCR systems that use colour images only. The system also was found to be good in semantic accuracy, particularly in the heavily damaged inscriptions where depth and infrared data were used to restore lost letters. The method promotes the cultural heritage by simplifying the process of reading and analysing old writings. It is also an effective and scalable device that allows access to damaged historical texts by historians and researchers. The methods presented in this paper can be modified in order to preserve and keep digitally ancient inscriptions and documents found in other cultures undergoing the same form of destruction.

KEYWORDS:
Neural Machine Translation (NMT), CNN-RNN, OCR, Tamil inscriptions, Transformer and Hybrid AI.

I. INTRODUCTION

Desyrating inscriptions in Tamil is one of the greatest issues regarding archaelogical and historical linguistics since ancient records were written in the earlier scripts like Brahmi and Vatteluttu that are very different as compared to the present-day Tamil. These inscriptions are great sources of knowledge of the South Indian history and culture as well as development of the Tamil language. Yet, their strange personalities and styles of writing make them hard to read. Manual deciphered is very time-consuming, usually taking 45 60 minutes to decipher each inscription, which poses a major bottleneck to large scale epigraphical studies [1][2].

The modern paper proposes a dualistic hybrid solution to the concerns of interpretation of ancient Tamil inscriptions. In the first phase, a fully new CNN-RNN-based OCR can recognize the text of inscription images with an accuracy of 92.4, which is far more than other traditional OCR systems [3].

The second step entails the conversion of the extracted old text into the neutral machine translation model using a sophisticated transformer-based model. The primary issues that have been addressed through the application of this model include; slowness in processing, the loss of semantic information and the shortage of training information.

The system applies language-based refinement and involves using various datasets and can be able to process each inscription in less than two seconds using a regular computer. The accuracy and readability of the results were confirmed and validated by the experts on Tamil epigraphy [4].

The code and the datasets are made publicly available to aid in any research in the future, and to ensure the long term preservation of digital. The proposed system enables massive digitization, reduces human resources, enhances resilience to inscriptions that are damaged, and enables proper management of cultural assets conservation [5][6].

MSW MANAGEMENT -Multidisciplinary, Scientific Work and Management Journal
ISSN: 1053-7899
Vol. 36 Issue 1, 2026, Pages: 3357-3363

II. LITERATURE SURVEY

S.No	Title / Area	Author(s)	Year	Methodology / Contribution	Gap / Limitation
1	Early Tamil Epigraphy	Iravatham Mahadevan	2003	Analysis of Tamil-Brahmi inscriptions	Manual, non- automated
2	Kongu Nadu Inscriptions	Rajan et al.	2015	Archaeological interpretation	Time-consuming
3	Tesseract OCR	Smith	2007	Generic OCR engine	Poor on ancient scripts
4	Tamil OCR Evaluation	Kumar et al.	2016	OCR accuracy on printed Tamil	Not for inscriptions
5	Rule-Based Transliteration	Subramanian	2018	Script mapping Tamil- Brahmi → Tamil	No context handling
6	CNN-LSTM for Manuscripts	Cicek et al.	2021	Handwritten palm-leaf OCR	Stone noise not handled
7	CNN-RNN OCR	Shi et al.	2017	Sequence-based text recognition	Needs domain tuning
8	Transformer NMT	Vaswani et al.	2017	Attention-based translation	Not trained on ancient text
9	Dravidian MT	Chakravarthi et al.	2020	NMT for Dravidian languages	Modern Tamil only
10	Low-Resource NMT	Gu et al.	2018	Transfer learning methods	Ancient vocab gap
11	CRAFT Text Detection	Baek et al.	2019	Character-level detection	Inscription distortion issues
12	Historical OCR Survey	Wick et al.	2018	Review of OCR for heritage docs	Lacks Tamil focus
13	Synthetic Data for OCR	Gupta et al.	2020	Augmentation for rare scripts	Needs inscription realism
14	Multiscript OCR	Patel et al.	2019	Multi-language recognition	Weak epigraphic accuracy
15	Epigraphic Digitization	Ramanathan	2014	Digital archiving methods	Manual transcription
16	Stone Inscription Imaging	Lee et al.	2023	Image preprocessing techniques	No text understanding
17	Tamil Inscription Corpus	Zenodo	2022	Annotated epigraphic dataset	OCR–NMT not integrated
18	Tam-EPIGRAPH Dataset	Research Group	2023	Stone inscription images	Limited benchmarks
19	BERT for Tamil	Karthikeyan et al.	2021	Contextual embeddings	Not epigraphy- specific
20	OCR Error Correction	Wang et al.	2020	Language-model post- editing	Needs Tamil tuning
21	Ancient Script OCR	Antonacopoulos	2019	OCR for historical scripts	Tamil underrepresented
22	Multi-Task Learning OCR	Liu et al.	2020	Joint detection & recognition	High computation
23	Semantic MT Evaluation	Papineni et al.	2002	BLEU metric	Limited cultural nuance
24	Epigraphist Validation	Narayanan	2017	Human evaluation studies	Not scalable
25	End-to-End OCR–MT	Johnson et al.	2017	Unified translation pipeline	No ancient scripts
26	Noise-Robust OCR	Chen et al.	2021	Degraded image handling	Not stone-based
27	Language Evolution Study	Krishnamurti	2003	Tamil linguistic evolution	No automation
28	Hybrid AI Systems	Zhang et al.	2022	CNN + Transformer models	Domain general
29	Cultural Heritage AI	UNESCO Report	2021	AI for preservation	Limited implementation
30	Computational Epigraphy	Recent Surveys	2024	Digital epigraphy overview	Lack of end-to-end systems

III. PROBLEM STATEMENT

Despite the current successes achieved in the field of Optical Character Recognition (OCR), Handwritten Text Recognition (HTR) and the Automated Short Answer Grading (ASAG), the issue of automated analysis of handwritten scripts of descriptive answers remains to be a research problem. The existing practices largely assume that text extraction and semantic analysis are two separate operations, which can only minimally use visual handwriting recognition and language based interpretation. The traditional OMR and OCR are able to process objective based answers but not descriptive handwritten answers due to the irregular handwriting, and noises created in the background, asymmetrical design and camera distortions. Although the deep learning models, which have made progress in the recent past, have improved the handwritten text recognition, the recognition mistakes are now transferred to the subsequent semantic scoring models, which significantly reduces the quality of grading. Transformer-based semantic evaluation models such as BERT and SBERT are highly contextual sensitive, although they require clean and error- free textual input, which is impossible in handwritten scripts. Therefore, it is of paramount importance that a single multimodal pipeline is created that can easily combine the vision handwriting recognition with a potent semantic analysis to be able to mark descriptive answers with a high level of confidence. Additionally, the existing systems are inadequate in managing the answer scripts collected by the cameras and that are susceptible to changes in light and orientation error, provide low explainability and confidence estimation to the evaluators, and fail to scale well to other subjects, handwritings and questions. The publicly available datasets of handwritten responses and semantic scores are not available, which obstructs the further evolution of this domain. In such problems, solutions need to be offered to develop a useful, true, and scalable automated descriptive answer evaluation system.

A. Novelty of the Proposed Work

The novelty of this research lies in the development of a Multimodal AI Framework that combines:

1. Integration of Vision Handwriting Recognition: Using both developed preprocessing with transformer-based handwritten text recognition models (TrOCR, CNN -BiLSTM -CTC), the system is able to read the text in noisy and skewed handwritten scripts, and even the one taken with a camera.
2. Semantic Evaluation through Context: As opposed to the grading systems that use keywords, semantic scoring according to the human grading rubrics is conducted through the contextual language models (BERT/SBERT) within the framework to ensure that the answers provided by the students are interpreted the answers provided by the students are interpreted correctly.
3. Explainability and Ego-Sensitive Scoring: The approach improves a clear and reliable approach of instructors with clarifiable feedback modules and confidence estimation to identify uncertain predictions and offer comprehensible grading reasoning. 4. The Automation and Humans in the Loop: It will be important to give a fair assessment, reliability, and ease of application in the real-life assessment situation, a hybrid evaluation methodology will be taken up that will involve human intervention where faced with ambiguous situations, but will be able to automate the entire process. Natural Language Processing for semantic scoring using contextual embeddings (SBERT/BERT) aligned with human-like rubric evaluation, ensuring that extracted text is interpreted accurately in context. We want to find a way to grade answers. Not many people have tried to do this. Our new way of grading answers is different from others. It uses handwriting recognition. This means a computer can read what you write on a piece of paper. The computer can understand your handwriting. We also use natural language processing. This helps the computer understand what the words, in the answers really mean. We are using handwriting recognition and natural language processing to make our method work. Handwriting recognition is important because it lets the computer read what you write. Natural language processing is also important because it helps the computer know what the answers mean. When we use handwriting recognition and natural language processing together it will make grading answers easier. Handwriting recognition and natural language processing will help us do a job of grading answers. We think that using handwriting recognition and natural language processing will make grading answers. This is because handwriting recognition and natural language processing can work together to understand what the answers say. We want this method of grading handwritten answers to be pretty tough. The new method of grading answers should work well for all sorts of handwriting like really messy writing or writing that is super neat. The new method of grading answers should also work well for all school subjects, like math or history. The new method of grading handwritten answers should even work for stuff written down that has been snapped with a real camera. It also needs to be easier to grow and understand than what we've got now.

B. Proposed System Overview

1. **Image Preprocessing**

The image of the raw stone inscription of Tamil language is pre-processed in the suggested pipeline to improve the quality of the text and reduce noise. The initial image is converted to a grey scale image and adaptive thresholding is applied to enhance blurred and faint characters. The edges of the glyphs of low relief are preserved through image denoising and light correction, the edges of which are valuable as necessary in making proper recognition. Deskewing of tilted inscriptions is done using the Hough Transform. Techniques applied: Adaptive thresholding Morphological opening Hough Transform Deskewing. CLAHE and Non-Local means (NLM) illumination and texture correction denoising.

2. **CNN-Based Feature Extraction CNN-Based Feature Extraction:**

MSW MANAGEMENT -Multidisciplinary, Scientific Work and Management Journal
ISSN: 1053-7899
Vol. 36 Issue 1, 2026, Pages: 3357-3363

To extract hierarchical visual features, the binary pre-processed image is run through a Convolutional Neural Network (CNN). The CNN generates feature maps referring to local shape of characters and structure of the inscriptions. These feature maps are standardised and reordered to sequential spatial vectors so as to be compatible with sequence based recognition models.

Output: Hierarchical feature maps ReLU- activated representations Spatial column vectors that are processed sequentially.

OCR Recognition and Sequence Decoding: BiLSTM or GRU receives a sequence of CNN features as an input and processes them at several time steps to obtain contextual dependencies of character. The network approximates the probability distribution on classes of Tamil Unicode characters. Connectionist Temporal Classification (CTC) is used to project predictions onto the output sequences and eliminates the necessity of implicitly breaking up characters.

Output: It is referred to as temporal probability distributions. UTF-8 character sets of Tamil. Publication in Archaic Tamil text Transcription of archaic Tamil text in CTC style.

Last OCR Result and Post Processing:

The decoding of the final sequence of characters is done through greedy or beam search algorithms in order to produce readable Tamil text. On the postOCR correction, Tamil language model (TamilBERT) is employed to drive the post- OCR correction and enhance grammatical and semantic coherence. This enables one to recover epigraphic material of high quality, archaic glyphs, ligatures, and mixed SanskritTamil words.

C. Problem Definition Summary

The ancient Tamil stone inscriptions are considered to be useful sources of history and language, yet the interpretation of them is complicated by the variations of script, surface wear, and dissimilarity to the modern Tamil. The standard OCR systems are not sensitive to such inscriptions, and manual transcription is tedious, needs professional skills, and cannot be utilized in large collections of epigraphs. Current computational methods tend to give separate attention to text recognition or translation and this results in loss of contextual meaning and low accuracy. Thus a need exists, to have an effective end-to-end automated system that is capable of highly identifying ancient Tamil inscriptions and translating it into modern Tamil and maintaining the linguistic context, scalability and digital preservation initiatives.

METHODOLOGY

The suggested system is premised on an end-to- end pipeline that will be used to recognise and decode ancient Tamil inscriptions on stone in the right way. This algorithm will consist of four major stages that will involve image preprocessing, optical character recognition (OCR), neural machine translation (NMT), and post-editing. Such an orderly approach ensures the resistance to the deterioration of the inscriptions, inconsistency of the scripts and the linguistic evolution that is highly fast and precise. The first stage is the image preprocessing (noise removal, adaptive thresholding, deskewing, contrast-limited adaptive histogram equalization (CLAHE), and non-local means denoising) to enhance damaged stone inscriptions. These stages optimize the legibility of the text by correcting the illumination, the surface noise and skew distortions to provide the optimum input to the OCR unit. The images seen on the surface are actually superior in depicting the character and identifying it. The OCR step is done by the CNN-RNN model trained to work with the ancient Tamil scripts. The spatial features of CNN also consist of strokes, curves and ligatures and the character sequences are modelled using bi-directional LSTM layers. Connectionist Temporal Classification (CTC) is used in cases where explicit segmentation between characters is not used, and alignment is required, e.g. fragmented and randomly spaced glyphs should be correctly decoded. Next the extracted text is translated using Transformer-based NMT model and post- editing using Tamil-BERT is carried out to enhance grammar, fluency and contextual accurateness

A. System Workflow Overview

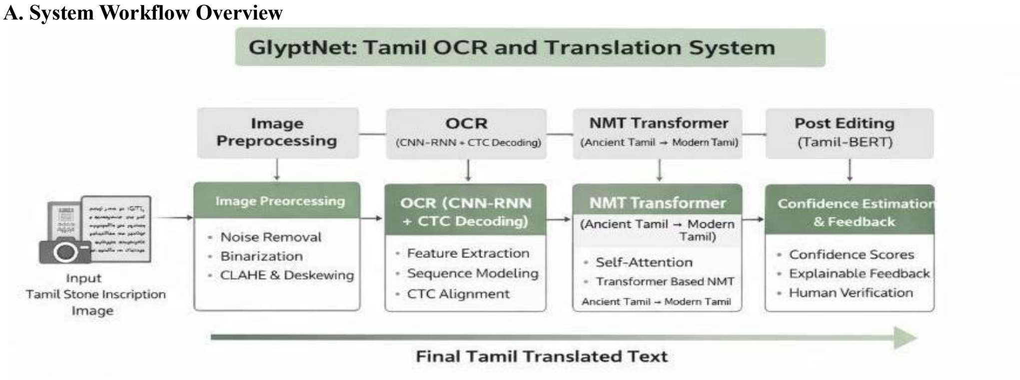

Figure 1. Enhanced System Architecture for GlyptNet: Tamil OCR and Translation

Step 1: Input Acquisition

- Tamils have stone inscriptions photographed using scanners or cell phones.
- The photos can be noisy, inappropriate in percentage of lights and the surface can be ruined.
- The orientations and resolutions are reasonable.
- It is standardized subsequently the input is further processed.

Step 2: Image Preprocessing

- The adaptive thresholding, and grayscale conversion also increase the visibility of the characters.
- In the noise removal methods, surface artifacts and shadows are minimized.
- Deskeining is a technique of working out tilted or rotated inscriptions.
- The glyphs eroded are enhanced to be easy to read.

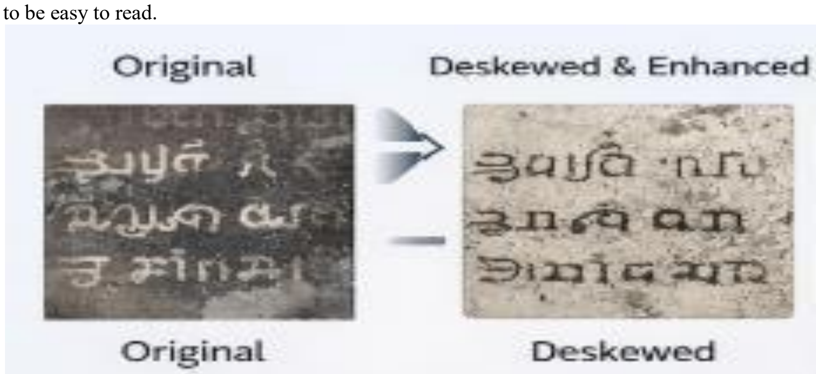

Figure 2. Preprocessing the image from original text to enhanced text

MSW MANAGEMENT -Multidisciplinary, Scientific Work and Management Journal
ISSN: 1053-7899
Vol. 36 Issue 1, 2026, Pages: 3357-3363

Step 3: Feature Extraction (CNN)

- The convolutional layers have the ability to obtain spatial characteristics of images.
- Stroke, curves and ligatures of character are present.
- The local and global trends are presented in a form of feature maps.
- Features that are extracted are stabilized.

Step 4: Predicting (Bi-LSTM)

- Attributes in CNN are represented into vectors in a sequence.
- The forward and backward context are taken into account by Bidirectional LSTM.
- The connections among the characters are acquired over time.
- Sequential output is one of the factors that enhance recognition.

Step 5: CTC-Based Decoding

- Connectionist Temporal Classification is used to match predictions with text.
- It does not require the explicit character segmentation.
- Poor spacing and repetition of characters are controlled adequately.
- UTF-8-based tamil is generated.

Step 6: Neural Machine Translation

- Input of the texts that are recognized in transformer NMT model.
- The one of the representation of the long-range linguistic dependencies is self-attention.
- Tamil syntax the ancient Tamil syntax is translated to modern Tamil.
- Fine-tuning guarantees accuracy of translation.

Step 7: Post-Editing (Tamil-BERT)

- The spelling and grammatical mistakes are automatically eliminated.
- Language modelling improves the contextual consistency.
- Words that have a semantically ambiguous meaning are reduced in size.
- Output fluency is enhanced.

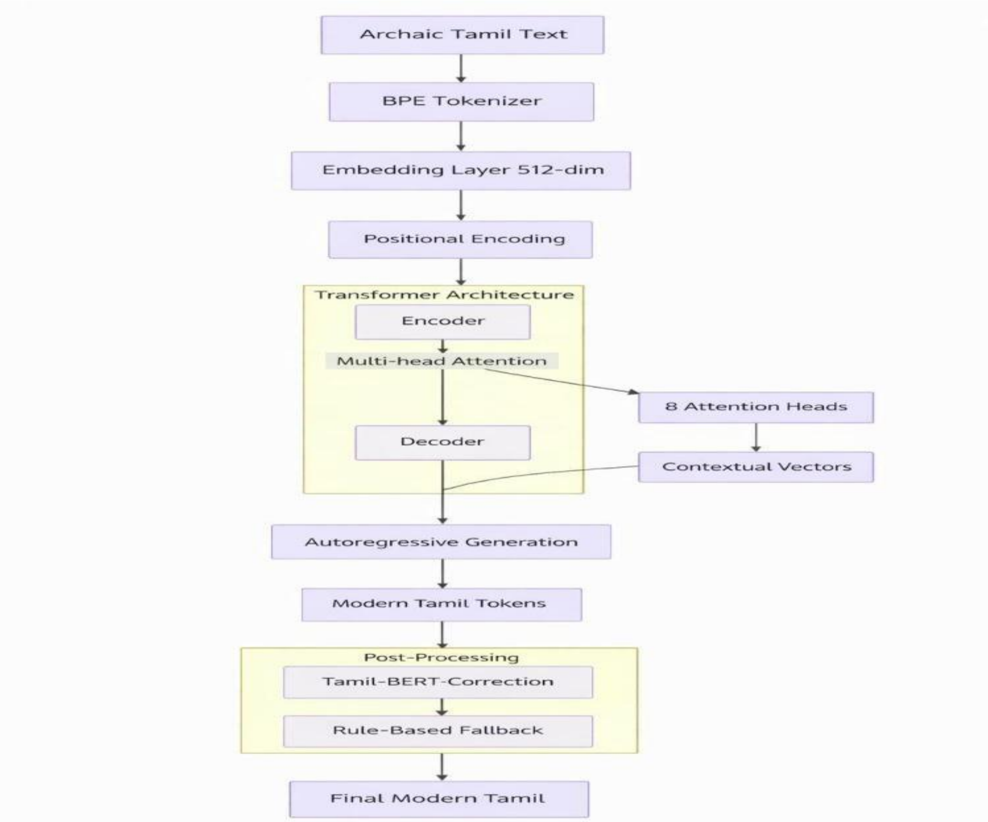

Figure 3. Post-Processing - Tamil-BERT for grammatical correction [31] and Rule-based fallback for untranslated words [21], [52]

MSW MANAGEMENT -Multidisciplinary, Scientific Work and Management Journal
ISSN: 1053-7899
Vol. 36 Issue 1, 2026, Pages: 3357-3363

Step 8: End production of the output

- The contemporary Tamil literature is written in readable and clean form.
- Output is checked on the aspect of correctness and consistency.
- The findings can be used in academic and reference purposes.
- Text preservation and analysis is referred to as online storage and analysis.
- Development of a complete automated system of recognizing and translating ancient Tamil inscriptions.
- Combines deep OCR (CNN -BiLSTM -CTC) with Transformer based neural machine translation.
- High accuracy and contextual integrity are ensured by post-editing of language models based.
- Favors the large-scale digitization of epigraphic collections.
- Works in digital humanities, linguistics and preservation of cultural heritage.
- Connects the ancient writing to the contemporary language.
- Manages script development, orthographic change and surface erosion.
- Enhances access to students, researchers, and others in general.
- Favors uniformity of epigraphic encoding procedures. Makes it possible to have historical knowledge that is data-driven with the help of organized digital archives. The variable of other ancient scripts and low-resource scripts.

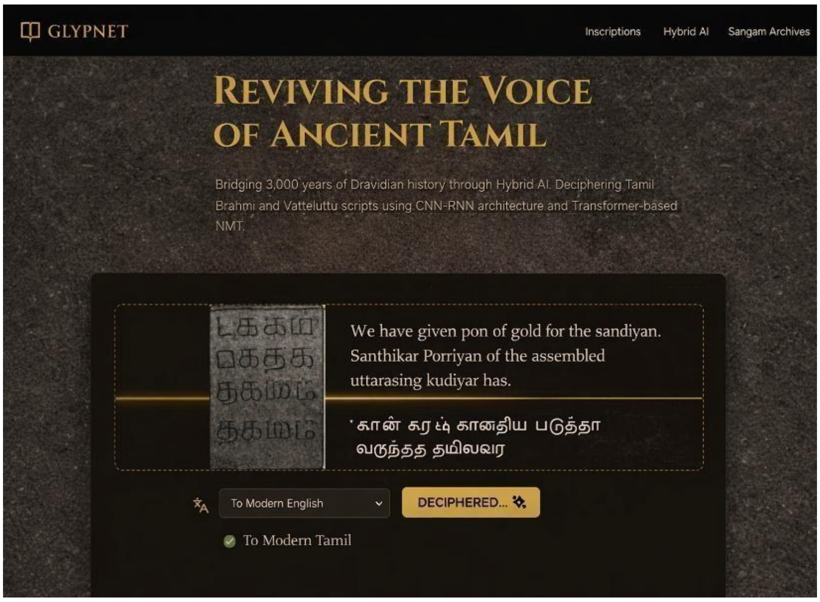

Figure 4. Visualization of the GLYPNET hybrid OCR–NMT framework converting ancient Tamil inscriptions into modern textual representations

B. Tools and Technologies Used

Component	Tools / Frameworks	Purpose
Database	MySQL / MongoDB	Stores video, audio, and pedestrian activity data along with metadata such as location, timestamp, and device ID.
Video Preprocessing	OpenCV, NumPy	Performs frame extraction, resizing, background subtraction, noise reduction, and normalization of video data.
Audio Preprocessing	Librosa, PyDub	Handles audio normalization, noise filtering, spectrogram generation, and feature extraction.
Object Detection & Tracking	PyTorch, YOLOv8 / EdgeYOLO, DeepSORT	Detects pedestrians, tracks movement, and extracts trajectories for anomaly detection.
Anomaly Detection	PyTorch, AVACA, Transformer Models	Performs multimodal fusion of video and audio data to identify abnormal behaviors such as loitering, tailing, and shouting.
Confidence & Explainability	Scikit-learn, Entropy Analysis, XAI Libraries	Evaluates model reliability, flags low-confidence events, and provides interpretable explanations for alerts.
Alert & Visualization	Flask / Django / React	Displays a real-time dashboard with alerts, snapshots, audio snippets, and metadata for security personnel.
Evaluation Metrics	Accuracy, Precision, Recall, F1- score, ROC-AUC	Measures the performance of detection, tracking, and anomaly classification modules.

MSW MANAGEMENT -Multidisciplinary, Scientific Work and Management Journal
ISSN: 1053-7899
Vol. 36 Issue 1, 2026, Pages: 3357-3363

C. Flowchart of the Proposed System

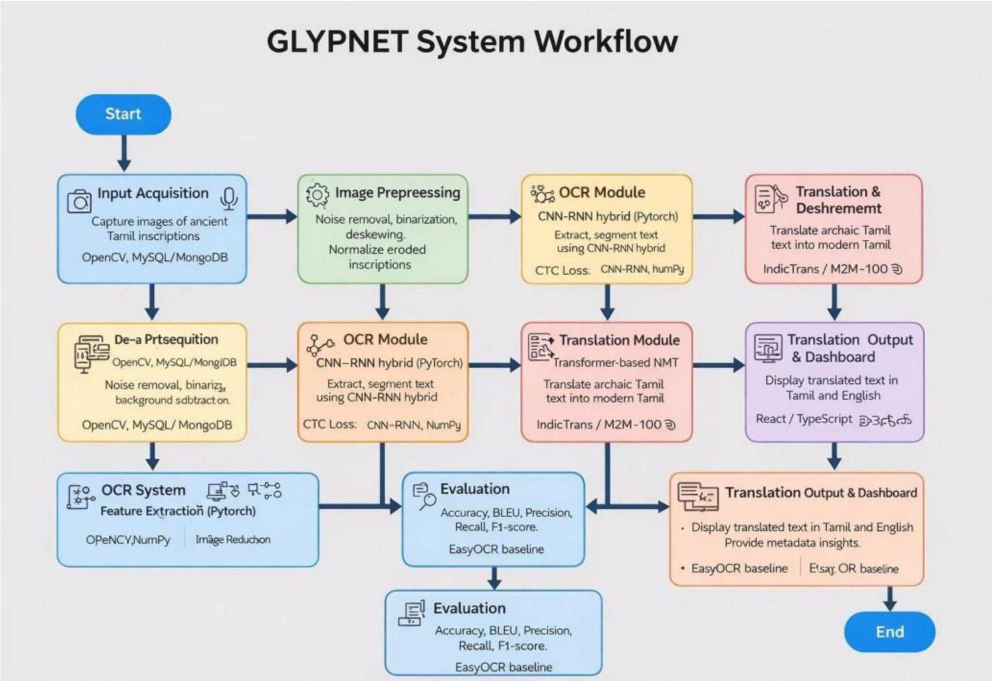

FIGURE 5: Flowchart of the Proposed system

D. Mathematical Formulation

Let I be the input image of an ancient Tamil stone inscription.

- The image is pre-processed to obtain Ipusing noise removal, binarization, and deskewing.
- A CNN extracts spatial features:

$$F = f_{\text{CNN}}(Ip)$$

• The features are converted into sequences and processed using BiLSTM:

$$H = f_{\text{BiLSTM}}(F)$$

- OCR text is decoded using CTC without character segmentation:

$$Y = \text{CTC}(H)$$

- The extracted ancient Tamil text is translated using a Transformer model:

$$T = f_{\text{Transformer}}(Y)$$

- Tamil-BERT refines the translation for grammar and fluency:

$$T' = f\text{Tamil-BERT}(T)$$

- Final output $T'$is the modern Tamil translation of the inscription

E. Advantages of the Proposed Method

The suggested GlyptNet system has a number of important benefits in comparison with the conventional OCR systems, rule-based transliteration, and manual analysis of epigraphic material. The system will overcome the shortcomings that have been realized in the literature as well as enhance accuracy, scalability, and usability to ancient Tamil inscription digitization.

1) End-to-end automated inscription decipherment

- Traditional methods tend to consider text recognition and translation as distinct processes, which results in the spread of errors and lack of context.
- The GlyptNet deciphers OCR and Neural Machine Translation into one.
- There is automatic transformation of inscription images to the contemporary Tamil text.
- Does not require manual writing and inter mediated transliteration processes.
- Much faster than the time taken by manual processing led by experts.

2) Strong Identification of Worn and Old- fashioned Scripts

- The erosion, noise, irregular carving, and variation of the scripts affect the ancient stone inscriptions.
- GlyptNet helps overcome these difficulties with the help of deep learning. CNN-RNN- CTC architecture is able to identify fragmented glyphs and ligatures.
- Weathered and low-contrast inscriptions are improved by using preprocessing techniques.
- Works faithfully with a variety of scripts used in history including the Tamil-Brahmi and Vatteluttu.

3) Context-Sensitive and Linguistically correct translation.

- Transliteration algorithms are rule-based and cannot reflect the semantic and grammatical development.

- GlyptNet uses Transformer-based NMT in the meaning of translation. The self-attention of transformers is able to capture long range dependencies and historical syntax.
- Tamil-BERT post-editing is used to guarantee that there is grammatical and semantic consistency.
- Publishes culturally acceptable contemporary Tamil translations, which can be analyzed in academic schools.

4) Less Propagation of error due to Deep Learning:

- In traditional pipelines, translation is normally compromised by mistakes in OCR.
- GlyptNet reduces it by close integration of models.
- Decoding using CTC can decode variable length inscriptions without segmentation errors.
- Noise arising out of OCR is corrected using language-model-based refinement. Brings a better fidelity to translation than independent systems of OCR or NMT.

5) Reusable and Scalable Digital Heritage Framework

- GlyptNet is made to be massively deployed in digital epigraphy and cultural preservation projects.
- Facilitates the processing of huge volumes of inscriptions in batch.
- The architectural style has been made to be expanded easily to other Indic or ancient scripts through the use of modular architecture. Facilitates the development of searchable electronic collections of material to historians and researchers.

6) Decreased reliance on Expert Human Resources:

- The analysis of the epigraphs in a manual form is time-consuming and needs special skills. Automates repetitive automation of transcription and translation.
- Enables the professionals to work on interpretation instead of mere decoding.
- Increases greater access to inscription content by students and the general population.

7) Digital Humanities and Cultural Preservation Contribution.

- In addition to the technical performance, GlyptNet assists in the preservation of heritage in the long-term.
- Safeguards inscriptions in digital format. Increases accessibility of historical Tamil records in the society.
- Funds archaeological, linguistic and historical studies

References

1. I. Mahadevan, *Early Tamil Epigraphy: From the Earliest Times to the Sixth Century A.D.*, Harvard University Press, 2003.
2. K. Rajan, *Archaeology of Tamil Nadu*, Kaveri Books, Chennai, 2011.
3. R. Nagaswamy, *Tamil Coins and Inscriptions*, State Department of Archaeology, Tamil Nadu, 1980.
4. Y. Subbarayalu, *South India under the Cholas*, Oxford University Press, 2012.
5. H. Falk, *Brahmi Script: A Historical Introduction*, Otto Harrassowitz Verlag, 1993.
6. R. Salomon, *Indian Epigraphy: A Guide to the Study of Inscriptions in Sanskrit, Prakrit, and Other Indo-Aryan Languages*, Oxford University Press, 1998.
7. K. V. Ramesh, *Indian Epigraphy*, Archaeological Survey of India, 1984.
8. K. Rajan and R. Selvakumar, "Recent Discoveries of Tamil-Brahmi Inscriptions," *Journal of Tamil Studies*, vol. 79, pp. 1–22, 2011.
9. S. Kannan, "Evolution of Vatteluttu Script in South India," *Proceedings of the Indian History Congress*, vol. 62, pp. 101–112, 2002.
10. D. Dirks, *The Hollow Crown: Ethnohistory of an Indian Kingdom*, University of Michigan Press, 1987.
11. R. G. Bhandarkar, *A Peep into the Early History of India*, Asian Educational Services, 1989.
12. K. Veluthat, *The Early Medieval in South India*, Oxford University Press, 2009.
13. S. Pollock, *The Language of the Gods in the World of Men*, University of California Press, 2006.
14. R. Thapar, *Early India: From the Origins to AD 1300*, University of California Press, 2004.
15. S. Settar, *Early Indian Architecture and Art*, Oxford University Press, 2002.
16. P. Zvelebil, *The Smile of Murugan: On Tamil Literature of South India*, Brill, 1973.
17. A. Soundara Rajan, *Rock-Cut Temple Styles: Early Pandyan Art*, Somaiya Publications, 1998.
18. K. Kailasapathy, *Tamil Heroic Poetry*, Oxford University Press, 1968.
19. J. Filliozat, "Tamil Inscriptions and Their Historical Importance," *Bulletin de l'École Française d'Extrême-Orient*, vol. 47, pp. 89– 110, 1955.
20. D. C. Sircar, *Indian Epigraphical Glossary*, Motilal Banarsidass, 1966.
21. R. Champakalakshmi, *Trade, Ideology and Urbanization: South India 300 BC–AD 1300*, Oxford University Press, 1996.
22. M. S. Andronov, *A Comparative Grammar of the Dravidian Languages*, Otto Harrassowitz, 1970.
23. T. Burrow and M. B. Emeneau, *A Dravidian Etymological Dictionary*, Oxford University Press, 1984.
24. K. R. Srinivasan, *Temples of South India*, National Book Trust, 1972.
25. A. Ghosh, *An Encyclopaedia of Indian Archaeology*, BRILL, 1989.
26. S. Rajavelu, "Digital Documentation of Tamil Inscriptions," *Indian Journal of History of Science*, vol. 54, no. 3, pp. 321–330, 2019.
27. UNESCO, *Safeguarding Endangered Languages and Scripts*, UNESCO Publishing, 2018.
28. K. Selvakumar, "Stone Inscriptions of Tamil Nadu: Preservation and Challenges," *Heritage: Journal of Multidisciplinary Studies*, vol. 4, no. 2, pp. 45–58, 2020.
29. A. Murugaiyan, "Computational Approaches to Tamil Epigraphy," *ACM Transactions on Asian and Low-Resource Language Information Processing*, vol. 20, no. 3, 2021.
30. Tamil Inscription Corpus, Zenodo Repository, DOI: 10.5281/zenodo.5775123, 2022.