10.48047/jocaaa.2024.32.01.69

Cross-Script Translation of Numerals from Kannada Epigraphy using Machine Learning Algorithms

[1]Mallamma V. Reddy, [2]Sachhidanand Sidramappa
[1,2]Department of Computer Science, Rani Channamma University, Vidyasangam, Belagavi, India.
[2]Residential Govt. First Grade College Ghodampalli, Bidar, India.
E-mail: [1]mvreddy@rcub.ac.in, [2]sachhi.r@gmail.com

Abstract

Kannada epigraphy, which includes inscriptions dating back to the 5th century CE, poses distinct challenges due to the evolution of the script, its degradation over time, and regional variations. One notable example is the Siribhoovalaya, a work penned by Jain monk Kumudendu Muni in the 9th century. This text is entirely written in Kannada numerals and utilizes a cipher system that encodes content across various scripts and languages. Interpreting such inscriptions requires the use of advanced computational techniques. Systematic archaeological research in Karnataka began during the British colonial period, with significant efforts to document and study inscriptions found across the region. One of the earliest references to Kannada inscriptions can be traced back to 1809. In that year, Francis Buchanan, a British officer, conducted a comprehensive survey of various districts in Karnataka. Buchanan meticulously examined numerous inscriptions on stone and copper plates, providing detailed accounts of the rulers, cultural practices, and historical events of the region at the time. In particular, the study of Cross-Script Epigraphy of numerals in Karnataka provides an insightful lens through which to understand the evolution of writing systems and numerical representation in the region**.** This research focus on Cross-Script Epigraphy of numerals of Karnataka, which helps to provide valuable information about kings, ministers, Soldiers, donations, land grants, taxes, and various aspects of social and cultural life.

Keywords: Cross-Script, Inscriptions, CRNN,CNN,

I. INTRODUCTION

In the context of Karnataka, inscriptions were often written in various ancient Kannada scripts witnessed from various dynasties to name Ashoka, Brahmi, Kadamba, Badami Chalukya, Kalyan Chalukya, Hoysala, Shatavahana, Rashtrakuta, and Mysore Wodeyar dynasty scripts, which are engraved, carved, or etched on durable materials such as stone, metal, clay, temple walls, pillars or copper plates and monuments [1] are shown in figure 1.

10.48047/jocaaa.2024.32.01.69

Fig 1: Various forms of inscriptions available in Karnataka

These inscriptions are valuable historical sources that help scholars understand the political history, social life, economy, religion, and culture of the region. Historically, inscriptions were used to [1]:

- Record royal orders, decrees, and proclamations.
- Document donations to temples or communities**.**
- Preserve religious texts, prayers, or verses**.**
- Mark important events, victories, or treaties**.**

The way these numbers were written varied over time and across dynasties, including:

- Unique numeral symbols used in dynasties like Kadamba, Badami Chalukya, and Hoysala.
- Word-based numbers**,** where numerals were written out in full words in Old Kannada.
- Hybrid forms**,** mixing symbols and words.

Ancient Kannada numerals vary significantly across historical periods such as Ashoka (3rd century BC), Kadamba, Badami Chalukya, Kalyana Chalukya, Hoysala, and Mysore Wodeyar, making manual recognition a time-consuming and expertise-driven task for historians and epigraphists is elaborated in this section. This research works on “Translation of Kannada Inscription to Number: Hindi, Kannada, English" is designed to identify and translate numerals found in ancient Kannada inscriptions from different dynasties into their modern

10.48047/jocaaa.2024.32.01.69

equivalents in three numeral systems — English (0–9), Modern Kannada (೦–೯), and Hindi/Devanagari (०–९) as shown in figure 2.

1	೧
2	೨
3	೩
4	೪
5	೫
6	೬
7	೭
8	೮
9	೯
10	೧೦

0	०
1	१
2	२
3	३
4	४
5	५
6	६
7	७
8	८
9	९
10	१०

Inscriptions from different dynasties Modern Kannada Hindi/Devanagari

Fig 2: Kannada and Hindi Numerals (Old and Modern)[2]

Kadamba Dynasty (4th–5th Century CE):The Kadamba dynasty played a key role in the early growth of Kannada inscriptions. Notable among these is the Hamidi inscription (circa 450 CE), which is one of the first full-length Kannada inscriptions. This period saw the divergence of Kannada from Brahmi, establishing its own writing style.

Badami Chalukya Dynasty (6th–8th Century CE):The Badami Chalukyas, known for their bilingual inscriptions (Kannada and Sanskrit), were instrumental in further refining the script. Inscriptions from this era, found in places like Aihole, Pattadakal, and Badami, show a move toward a more standardized form of Kannada script. The numerals and writing styles during this time became more consistent, and there was a noticeable shift toward regional stylistic preferences.

Rashtrakuta Dynasty (8th–10th Century CE):During the rule of the Rashtrakutas, Kannada continued to grow as a language of administration, royal decrees, and religious documentation. Inscriptions from this period, such as those in Ellora and the Kailasa temple, demonstrate the use of Kannada alongside Sanskrit for both political and religious purposes.

Hoysala Dynasty (12th–14th Century CE):The Hoysala dynasty is remembered for its highly decorative and ornate Kannada inscriptions, often carved into stone and accompanied by

10.48047/jocaaa.2024.32.01.69

elaborate temple architecture in places like Belur, Halebidu, and Somanathapura. These inscriptions, which were often bilingual, showcased the grandeur of Hoysala art and architecture. The Hoysala script became increasingly refined, with numerals taking on artistic flourishes, matching the intricate temple carvings of the time.

Vijayanagara Empire (14th–17th Century CE):Under the Vijayanagara Empire, Kannada was increasingly used for royal edicts, military proclamations, and temple patronage. The bilingual inscriptions of the Vijayanagara period, such as those found in Hampi, marked a flourishing of Kannada as both an administrative and cultural language. These inscriptions document the empire's conquests and its dedication to religious and cultural causes.

Mysore Kingdom (18th–19th Century CE):By the 18th century, under the Mysore Wodeyars, Kannada inscriptions had fully transitioned to the modern script. The inscriptions from this period, primarily from places like Srirangapatna, document legal matters, royal decrees, and social organization. The Mysore kingdom's inscriptions are significant for marking the modern development of the Kannada language, both in terms of script and its role in governance.

Key Features of Kannada Inscriptions

1. Script Evolution: Kannada inscriptions evolved from the Brahmi script and developed into a distinctive writing system over centuries, with regional variations reflecting local influences.
2. Language Use: While Kannada was the dominant language for royal and religious inscriptions, Sanskrit was often used alongside Kannada in the Hoysala, Chalukya, and Vijayanagara periods, reflecting the dual linguistic traditions of the time.
3. Numerals: Early Kannada inscriptions used Brahmi numerals, which gradually transformed into the familiar Kannada numerals seen today. These numerals were primarily used for recording dates and royal decrees.
4. Material: Most Kannada inscriptions were carved into stone, though some were written on copper plates or manuscripts. Stone inscriptions, often highly formal, typically served as important historical documents.
5. Purpose: Kannada inscriptions were used to document political decisions, royal orders, land grants, religious donations, and military conquests. They provide valuable insight into the governance, culture, and history of the region.

10.48047/jocaaa.2024.32.01.69

Translating ancient Kannada numeral inscriptions into modern languages such as Kannada, Hindi, and English is a complex task, due to several challenges. One of the main obstacles is the scarcity of digitized datasets for ancient Kannada numerals, which hampers the development of precise translation models. Furthermore, the representation of inscriptions from different dynasties like Ashoka, Kadamba, Badami Chalukya, and Wodeyar is uneven, making the task even more intricate. Additionally, many historical inscriptions suffer from erosion, cracks, and general wear, reducing their legibility and complicating the translation process. Lastly, the ambiguity in language adds to the difficulty, as the ancient Kannada language often contains words with multiple meanings depending on context, which complicates accurate translation into modern languages.

II. LITERATURE REVIEW

Efficient recognition of Kannada numerals has been a significant area of research. Most studies emphasize the scarcity of annotated training data, the necessity of synthetic augmentation to simulate wear and stylistic variation, and the integration of historical context for correct numeral mapping. Evaluation typically involves character accuracy, numeral sequence accuracy, and expert validation.

Prasanna Kumar and Prasad (2019) applied Convolutional Neural Networks (CNNs) to classify ancient Kannada characters from inscription datasets, demonstrating the efficacy of deep learning in recognizing historical scripts[3] .

Shashidhar and Rajendra (2020) focused on the automated recognition of ancient Kannada scripts, emphasizing the challenges posed by the evolution of the script over time[4] .

Reddy and Subrahmanyam (2018) addressed the recognition of handwritten Kannada numerals using CNNs, achieving high accuracy and contributing to the broader field of handwritten character recognition[5] .

Bhat and Sreenivasa (2013) discussed statistical and rule-based approaches for transliterating Kannada to English, providing insights into cross-language mapping techniques. Giridhar[6]

10.48047/jocaaa.2024.32.01.69

Dharani, and Guruviah (2019) proposed an OCR approach using image recognition-based classification for ancient Tamil inscriptions, drawing parallels to Kannada script recognition challenges [7] .

Bhuvaneswari and Bharathi (2015) presented a positional algorithm for recognizing ancient stone inscription characters, contributing to the understanding of ancient script recognition methodologies [8] .

Sridhar and Rajasekaran (2019) explored rule-based mapping techniques to transliterate Kannada inscriptions to modern scripts, highlighting the importance of preprocessing in script recognition [9] .

Chandrakala, Thippeswamy, and Martis (2021) investigated the impact of total variation regularization on character segmentation from historical stone inscriptions, focusing on enhancing image quality for better recognition [10].

Saini et al. (2021) proposed EffKannadaRes-NeXt, a residual network designed for efficient Kannada numeral recognition, utilizing advanced deep learning techniques for better accuracy and robustness [11] .

Mamatha et al. (2013) introduced a classifier fusion method for handwritten Kannada numeral recognition, combining multiple classifiers for improved performance [12] .

Sree Shankar et al. (2021) extended machine learning to the recognition of Tamil ancient characters and information retrieval from temple epigraphy, which shares similarities with Kannada inscription recognition in terms of historical text processing [13] .

Prabhu et al. (2019) proposed Fonts-2-Handwriting, a framework for universal digit classification using a seed-augment-train method to tackle issues in handwriting recognition across multiple fonts [14] .

Rajithkumar B. K. and Dr. H. S. Mohana (2014) employed template matching methods based on correlation analysis to recognize stone-inscribed Kannada characters from various historical periods, facilitating the conversion of ancient scripts to modern Kannada [15].

10.48047/jocaaa.2024.32.01.69

Vishweshwrayya C. Hallur and R. S. Hegadi (2020) utilized deep convolutional neural networks (DCNNs) for handwritten Kannada numeral recognition, applying preprocessing techniques like noise removal and binarization to enhance accuracy[16].

Hallur et al. (2012) conducted a survey on handwritten and printed Kannada numeral recognition techniques, discussing challenges and solutions in OCR for Kannada numerals [17].

H. R. Mamatha, S. Karthik, and K. S. Murthy (2013) implemented a classifier fusion method combining K-Nearest Neighbour (KNN) and Linear Classifier (LC) for recognizing handwritten Kannada numerals. Utilizing features like run length count and directional chain code, the system achieved an average recognition rate of 96%[18].

Basappa B. Kodada (2013) proposed the use of projection distance metrics and General Regression Neural Networks (GRNN) for recognizing unconstrained handwritten Kannada numerals. This method aimed to address challenges in character segmentation and classification [19].

Shashika Ruwanmini (2023) study focuses on the recognition of ancient Sinhala characters from the period between 10 A.D. and 12 A.D. To identify the most effective method for character recognition, three different Optical Character Recognition (OCR) techniques were tested: template matching, artificial neural networks (ANN), and convolutional neural networks (CNN). The process involved preprocessing, feature extraction, and character recognition. Based on the results, the study concluded that CNN-based OCR provided the most accurate recognition rates for ancient characters. This suggests that CNNs, due to their advanced capabilities in image processing, are the optimal choice for improving character recognition in historical Sinhala inscriptions [20].

This research addresses these gaps by developing a curated dataset of Kannada inscription numerals, creating a preprocessing pipeline optimized for inscriptions on curved and reflective surfaces, and implementing a hybrid recognition system to translate these numerals into modern Kannada, Hindi, and English representations. Such a system will contribute to epigraphy, digital heritage preservation, and multilingual numeric accessibility.

III. METHODOLOGY

Deep learning models, such as CNNs (Convolutional Neural Networks), CRNNs (Convolutional Recurrent Neural Networks), and transformer-based models, have been successfully applied to historical script recognition, with significant progress in tasks like

10.48047/jocaaa.2024.32.01.69

character-level classification and sequence modelling. However, a notable gap in the literature is the absence of a specialized dataset or recognition system for Kannada numerals found in inscriptions, especially those on metal artifacts like bells.

To address this gap, the proposed system utilizes image processing techniques that compare uploaded images of inscription numerals pixel-by-pixel against a curated reference database. The system employs a unigram statistical approach, where each numeral in the reference dataset is mapped to unique features based on its visual characteristics.

- Its modern Kannada numeral equivalent (೦, ೧, ೨, ೩, etc.)
- Its Hindi/Devanagari numeral equivalent (०, १, २, ३, etc.)
- Its English numeral equivalent (0, 1, 2, 3, etc.)
- The dynasty period it belongs to (Ashoka (3rd century BC), Kadamba, Badami Chalukya, Kalyana Chalukya, Hoysala, and Mysore Wodeyar)

These features might include pixel intensities, edge details, and geometric properties, enabling efficient and accurate matching between the input image and the database. For the sequence modelling part, a Convolutional Neural Network (CNN) is used to extract hierarchical visual features from the images. For more complex inscriptions where numerals may be part of a sequence, a Convolutional Recurrent Neural Network (CRNN) is employed to recognize the ordering and structure of the numerals. Transformer models, utilizing attention mechanisms, further refine the process by focusing on the most significant features of the image, improving recognition accuracy, especially when the inscriptions are degraded or contain noise.

The recognition process is divided into two main stages:

- Preprocessing and Feature Extraction: Various image processing methods such as thresholding, edge detection, and normalization are applied to enhance the image quality and highlight the essential features of the numerals.
- Numeral Matching: The system compares the processed image with the reference dataset using pixel-by-pixel similarity measures and feature matching techniques. It then outputs the most probable match, along with a confidence score indicating the certainty of the match.

This hybrid system leverages both advanced deep learning techniques for feature extraction and traditional image processing methods for precise matching. It is designed to be

10.48047/jocaaa.2024.32.01.69

adaptable to other regional scripts and can be used to digitize and preserve historical inscriptions, offering a robust framework for archaeological studies and historical research.

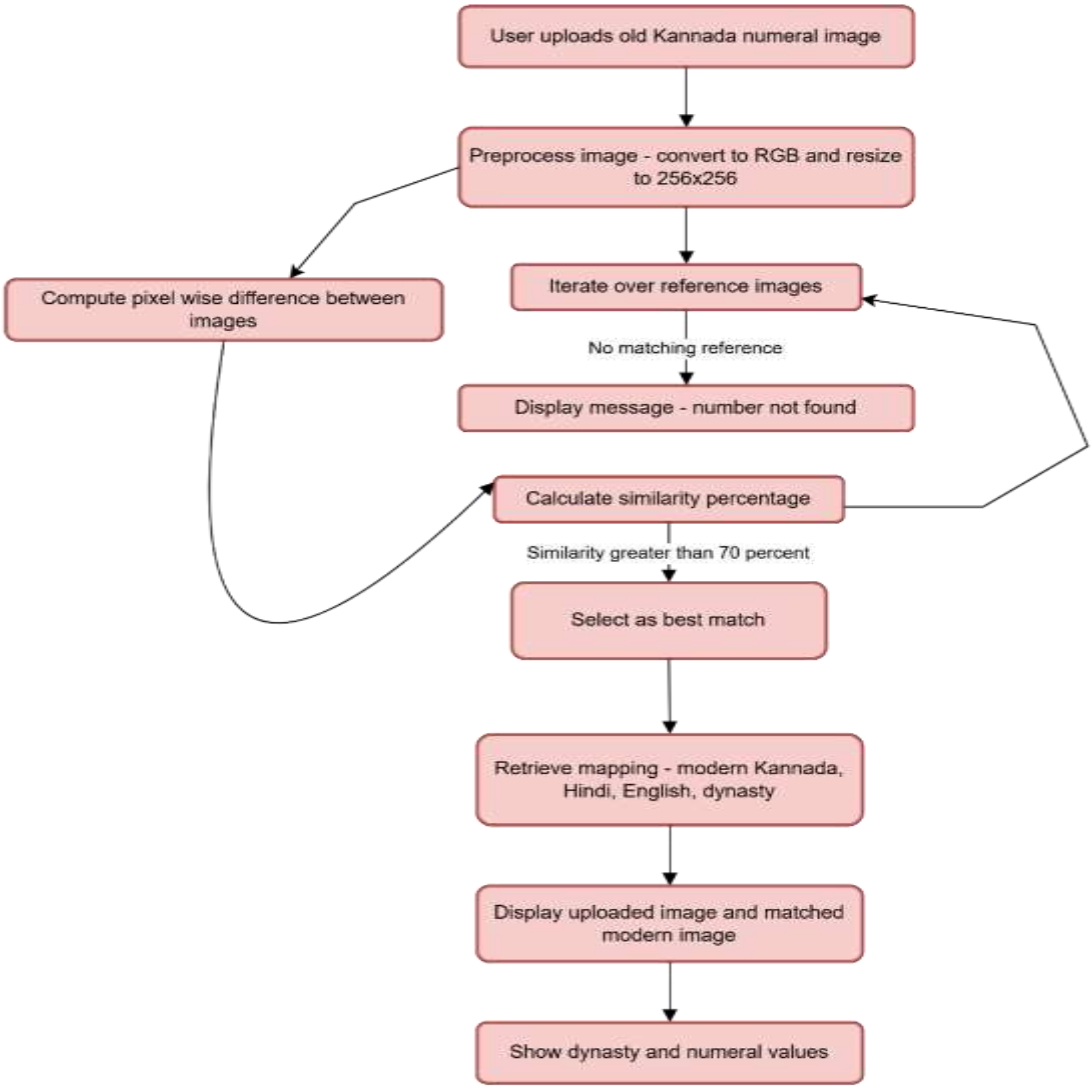

Fig 3: System architecture for numeral translation and dynasty identification

It is expected to accurately recognize ancient Kannada numerals from historical inscriptions and automatically translate them into modern Kannada, Hindi, and English numerals as shown in figure 4 and figure 5. It also provides the historical context by identifying the originating dynasty. The web-based interface offers an interactive and intuitive platform for numeral recognition, paving the way for future AI-based enhancements such as deep learning integration for improved accuracy and support for a wider range of numeral styles.

10.48047/jocaaa.2024.32.01.69

IV. RESULTS

Fig 4: Translation of Number 2

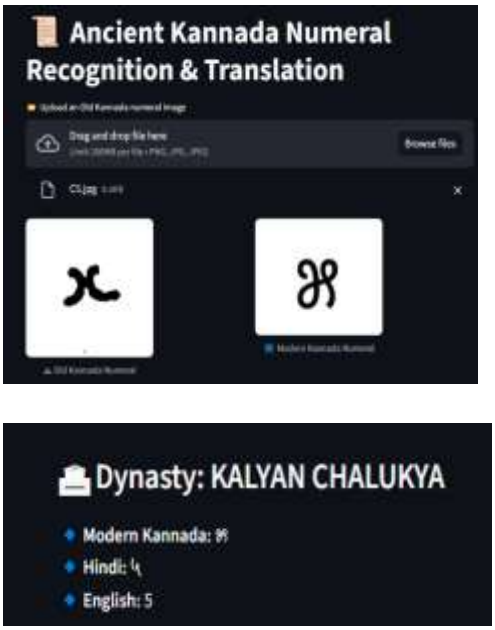

Fig 5: Translation of Number 5

Accuracy: Accuracy is a performance metric that tells us how well a classification model is predicting as shown in figure 6. It is the ratio of correctly predicted observations to the total observations. In simple words:

- Accuracy = How many predictions were correct / Total predictions

  $$\text{Accuracy} = \frac{TP+TN}{TP+TN+FP+FN}$$

Loss: In machine learning, loss is a measure of how well or poorly a model is performing. It represents the difference between the predicted values and the actual values (ground truth)**.** The goal of training is to minimize the loss**,** meaning the model's predictions get closer to the real answers as shown in figure 6

.

10.48047/jocaaa.2024.32.01.69

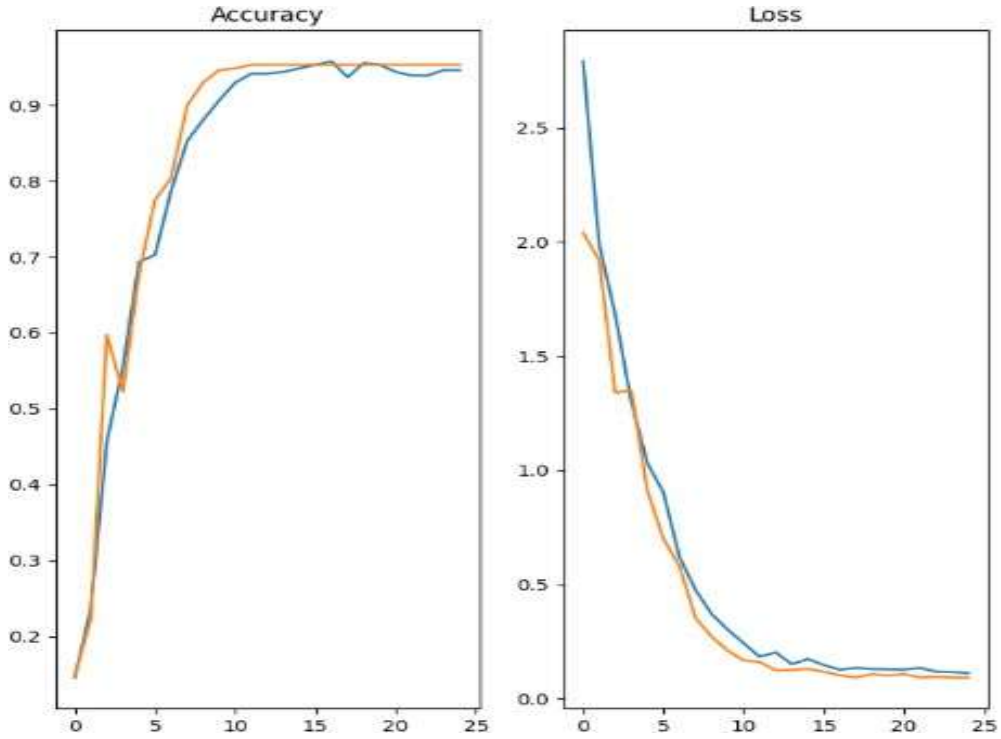

Figure 6: Accuracy and Loss of Trained Model

V. CONCLUSION

The system developed is proficient in recognizing and translating ancient Kannada numerals from various historical dynasties, such as the Ashoka, Badami Chalukya, Hoysala, Kadamba, Kalyan Chalukya, and Mysore Wodeyar periods. Using a pixel-based comparison technique, the system compares the uploaded images of inscriptions to a pre-established reference database, accurately identifying the numeral and translating it into Modern Kannada, Hindi, and English. However, despite the extensive collection of Kannada inscriptions, many have yet to be gathered or published. Numerous inscriptions still remain in villages, temples, and other locations, awaiting proper documentation. It is crucial to continue this effort to ensure the preservation of Karnataka's epigraphic heritage for future generations.

10.48047/jocaaa.2024.32.01.69

REFERENCES:

[1]. https://www.karnataka.com/history/about-history/

[2]. https://karnatakaitihasaacademy.org/evolution-of-kannada-script/

[3]. Prasanna Kumar, M., & Prasad, M. V. N. K. (2019). Recognition of Ancient Kannada Script Using Deep Convolutional Neural Networks. Advances in Intelligent Systems and Computing.Applies CNN models to classify ancient Kannada characters from inscription datasets.

[4]. Shashidhar, H. K., & Rajendra, K. R. (2020). Automated Recognition of Ancient Kannada Script Using Deep Convolutional Neural Networks. International Journal of Intelligent Systems and Applications.Focused on ancient-to-modern Kannada character mapping with deep learning.

[5]. Reddy, P. V., & Subrahmanyam, G. V. B. (2018). Handwritten Kannada Numeral Recognition Using Convolutional Neural Networks. IEEE International Conference on Recent Trends in Electronics, Information & Communication Technology (RTEICT). CNN applied to Kannada numerals recognition with high accuracy.

[6]. K. S. Bhat and H. K. Sreenivasa, "Transliteration of Kannada to English using statistical and rule-based approaches," International Journal of Advanced Research in Computer Science,Mar. 2013.

[7]. Giridhar, L., Dharani, A., & Guruviah, V. (2019). A novel approach to ocr using image recognition based classification for ancient tamil inscriptions in temples. *arXiv preprint arXiv:1907.04917*.

[8]. G. Bhuvaneswari, and V. S. Bharathi, "An efficient positional algorithm for recognition of ancient stone inscription characters", In: Proc. Of 7th International Conf. on Advanced Computing, Chennai, India, pp.1-5, 2015.

[9]. Sridhar, C., & Rajasekaran, S. (2019). Transliteration of Kannada Inscriptions to Modern Script Using Image Processing Techniques. International Journal of Computer Applications. Explores rule-based mapping from inscription text to Hosagannada.

[10]. Chandrakala, H. T., Thippeswamy, G., & Martis, R. J. (2021). Impact of total variation regularization on character segmentation from historical stone inscriptions. *Pattern Recognition and Image Analysis*, *31*(1), 35-48.

[11]. A. Saini, S. Daniel, S. Saini, and A. Mittal, "EffKannadaRes-NeXt: An Efficient Residual Network for Kannada Numeral Recognition," *Multimedia Tools and Applications*, vol. 80, pp. 28391–28417, 2021.

[12]. H. R. Mamatha, S. Karthik, and M. K. Srikanta, "Classifier Fusion Method to Recognize Handwritten Kannada Numerals," *arXiv*, Jan. 2013.

[13]. R. Sree Shankar et al., "Recognition of Tamil Ancient Characters and Information Retrieval from Temple Epigraphy Using Machine Learning," International Journal of Advanced Research in Computer Science and Software Engineering, vol. 11, no. 2, pp. 1–6, 2021.

10.48047/jocaaa.2024.32.01.69

[14]. V. U. Prabhu et al., "Fonts-2-Handwriting: A Seed-Augment-Train Framework for Universal Digit Classification," arXiv, May 2019.

[15]. Rajithkumar, B. K., & Mohana, H. S., "Template Matching Techniques for Recognizing Kannada Inscriptions: A Historical Analysis," *International Journal of Computing and Social Research*, vol. 7, no. 4, pp. 29-38, 2014.

[16]. Hallur, V. C., & Hegadi, R. S., "Convolutional Neural Networks for Kannada Handwritten Numeral Recognition," *Journal of Computational Intelligence and Applications*, vol. 14, no. 2, pp. 155-165, 2020.

[17]. Hallur, V. C., & Hegadi, R. S., "A Survey on Handwritten and Printed Kannada Numeral Recognition Techniques," *Studylib*, 2012.

[18]. H. R. Mamatha, S. Karthik, and K. S. Murthy, "Classifier Fusion Method to Recognize Handwritten Kannada Numerals," *International Journal on Recent Technologies in Mechanical and Electrical Engineering*, vol. 2, no. 6, pp. 77–80, Jun. 2015.

[19]. B.B.Kodada, "Unconstrained Handwritten Kannada Numeral Recognition," *International Journal of Computing and Social Research*, vol. 2, no. 6, pp. 77–80, Jun. 2015.

[20]. Shashika Ruwanmini, Kapila Dias, Clera Niluckshini, Terrance Nandasara, “Sinhala Inscription Character Recognition Model Using Deep Learning Technologies”, International Journal on Advances in ICT for Emerging Regions 2023 16(1).