

RESEARCH

Open Access



Comparative study: enhancing legibility of ancient Indian script images from diverse stone background structures using 34 different pre-processing methods

J. Jayanthi^{1*} and P. Uma Maheswari¹

Abstract

In recent times, there has been a proactive effort by various institutions and organizations to preserve historic manuscripts as repositories of traditional knowledge and cultural heritage. Leveraging digital media and emerging technologies has proven to be an efficient way to safeguard these invaluable documents. Such technologies not only facilitate the extraction of knowledge from historic manuscripts but also hold promise for global applications. However, transforming inscribed stone artifacts into binary formats presents significant challenges due to angle distortion, subtle differences between foreground and background, background noise, variations in text size, and related issues. A pivotal aspect of effective image processing in preserving the rich information and wisdom encoded in stone inscriptions lies in employing appropriate pre-processing methods and techniques. This research paper places a special focus on elucidating various preprocessing techniques, encompassing resizing, grayscale conversion, enhancement of brightness and contrast, smoothing, noise removal, morphological operations, and thresholding. To comprehensively assess these techniques, we undertake a study involving stone inscription images extracted from the Tanjore Brihadeeswar Temple, dating back to the eleventh century during the reign of Raja Raja Chola. This choice is informed by the manifold challenges associated with image correction, such as distortion and blurring. We undertake an evaluation encompassing a diverse array of stone background structures, including types like flawless-bright-moderately legible, dark-illegible, flawless-bright-illegible, flawless-dull, flawless-irregular-moderate, highly impaired-dark-legible, highly impaired-irregular-illegible, impaired-dark-moderate, impaired-dull-moderately legible, impaired-dusky dark-moderate, and very impaired-dusky dark-legible. Subsequently, the processed outputs are subjected to character recognition and information extraction, with a focus on comparing the outcomes of various pre-processing methods, including binarization and grayscale conversion. This study seeks to contribute insights into the most effective pre-processing strategies for enhancing the legibility and preservation of ancient Indian script images etched onto diverse stone background structures.

Keywords Ancient Indian scripts, Image preprocessing, Brihadeeswara temple stone inscription, Cultural and heritage, Digitization

Introduction

The ancient technologies of bygone eras are often regarded as advanced and distinct from the technologies we employ today. One remarkable example is the Brihadeeswara Temple, constructed in the eleventh century. The engineering marvel of erecting such a colossal

*Correspondence:

J. Jayanthi
jjay.mtech@gmail.com

¹ Department of Computer Science and Engineering, Anna University, Guindy Campus, Chennai 600025, India



© The Author(s) 2024. **Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>. The Creative Commons Public Domain Dedication waiver (<http://creativecommons.org/publicdomain/zero/1.0/>) applies to the data made available in this article, unless otherwise stated in a credit line to the data.

temple remains a mystery, given its massive size of 216 feet and the use of interlocking stones without any apparent cement or bonding material. To this day, constructing a building or structure of similar magnitude without modern cementing materials remains an unparalleled feat. These structures, along with other artifacts such as inscribed stones, palm leaves, copper plates, wooden artifacts, pillars, temple walls, rock beds, and potsherds, are believed to hold valuable information and knowledge from ancient times. The primary purpose of such inscriptions and repositories is to safeguard and transmit historical information across generations [1–3].

The advent of image processing, data science, machine learning, and artificial intelligence presents a unique opportunity for the modern world to decipher and analyze ancient information. A substantial volume of Tamil historical documents, once stored in libraries, museums, and temples, is now being digitized and made accessible through digital platforms [4–6]. However, this digitization process is met with numerous challenges. Many images and characters are in various stages of degradation and distortion, marred by noise and disruptions. Yet, the extraction of invaluable information from these artifacts can significantly enrich our understanding and be applied to establish literacy, archaeology, and historical context. Ancient Tamil inscriptions can be categorized into three main types: Vatteluthu, Grantha, and Tamil Brahmi. The Brahmi script, originating in the Ashoka era, served as the precursor to nearly all Indian scripts, including Vatteluthu and Grantha.

The pristine Brihadeshwara temple in Tanjore, constructed between 1003 and 1010 AD by Raja Raja Chola I, stands as a repository of valuable information and Vedic technological knowledge. However, the focus of this paper lies in deciphering scripts that have endured

varying degrees of decay and degradation, compounded by a multitude of noises. The central objective of pre-processing techniques is to enhance the image quality by effectively mitigating undesirable distortions while enhancing crucial image features, thus preparing the images for subsequent processing steps. Despite employing advanced photography protocols and sophisticated camera and scanning equipment, historical inscription images often remain unreadable due to the inexorable ravages of time. The deterioration of inscriptions can be attributed to factors such as ink migration, cracking, damages from biological and environmental influences, as well as foreign elements like dirt and discoloration (Fig. 1).

Addressing this challenge entails a series of steps (Fig. 2). The first step involves converting images to grayscale and resizing them for uniformity. The second step encompasses enhancing brightness and contrast using histogram analysis and alpha–beta comma values. The third step centers on smoothing and noise reduction, accomplished through filtering techniques like Gaussian, median, and bilateral filters. Subsequently, thresholding techniques are employed to distinguish between dark and light regions, utilizing methods such as adaptive thresholding, OSTU, Niblack, and Savuola [7, 8]. The final step integrates morphological processing to enhance characters, particularly in cases where thresholding or other preprocessing steps may have eroded or added pixels to characters. There are many languages in India and so the diversity in scripts throughout, much research has been documented about preprocessing of ancient inscriptions; however holistic approaches to understanding these recent technologies and evaluation of all such techniques are very few [9]. For instance; Buzyanov [10] proposed the methodology for improving pixel density of

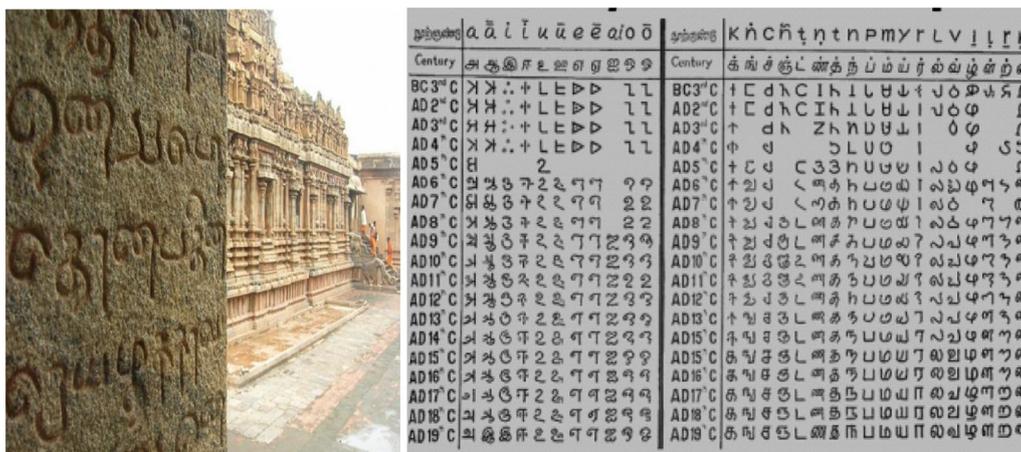


Fig. 1 Brihadeshwara temple of Tanjore Inscription sample and century wise ancient character representation

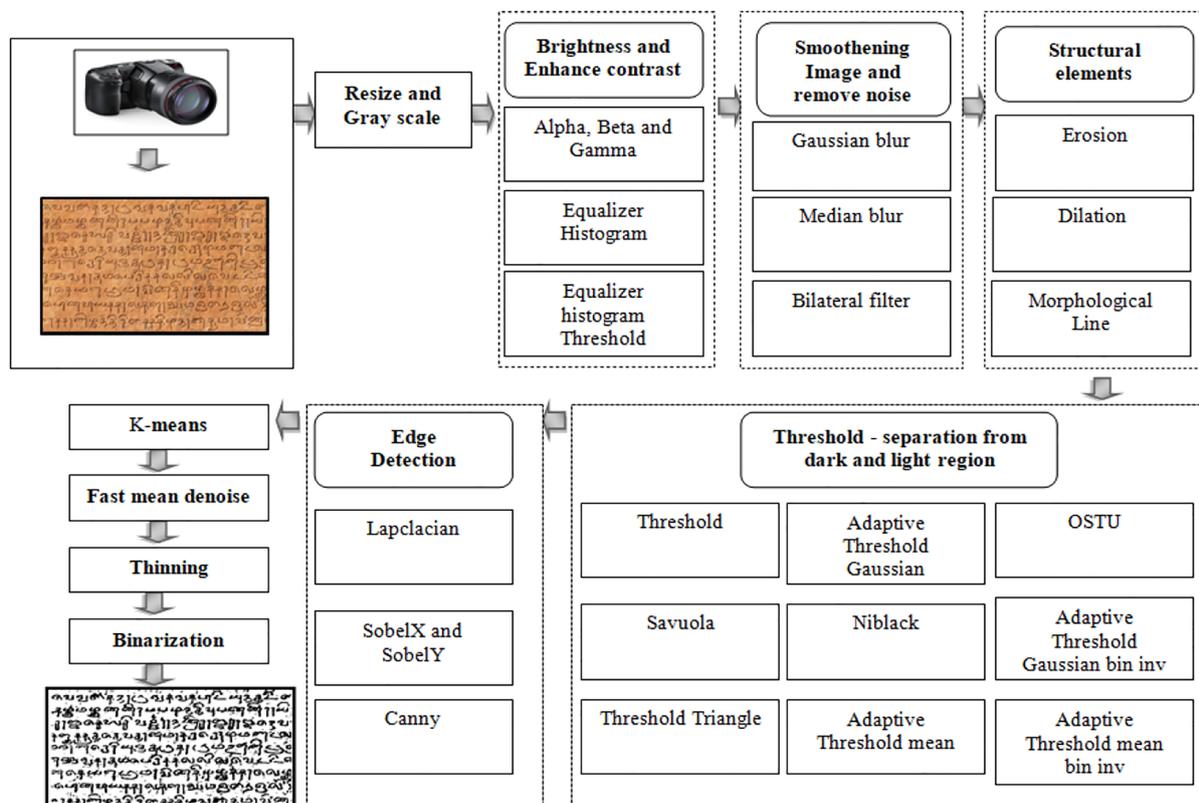


Fig. 2 The sample dataset image undergoes a sequential image pre-processing pipeline

text images via low pass signal Gaussian Laplacian filtering algorithm; however, evaluation of such technique and comparing this method with other efficient methods are not been documented since then. Similarly, an adaptive binarization method proposed by Kavelieratou et al. [11]; serve as an iterative mix of international and local thresholding. This method suggests the use of adjacent pixels to measure the average pixel value from the text to distinguish between text and nontext zones. Parul Saber and Sunjay proposed the multilingual character segmentation and recognition schemes for Indian ancient document images; which involve fewer sub-processes like binarisation, resizing, skew correction, and thinning to improve the clarity of images. The proposed methodology uses Otsu’s binarization technique which converts grey scaling concerning image size correction and pixel resolution to a binary image. Such methods imply the importance of backgrounds on the inscriptions; Seeger and Dance [12] estimate the intensity of background region and process binarization as compared to threshold intensity. Their algorithm calculates non-text region intensities at each pixel by which an appropriate threshold surface can be calculated [13]. In such context and advancement of the algorithm to evaluate the

image text; historical handwritten images are also subjected to pre-processing; Lakshmi and Patvardhan [14] proposed pre-processing and classification methodology for Telugu handwritten characters. Their research emphasizes gathering the dataset of handwritten ancient Telugu characters from numerous scripts and printed on high-quality paper [14]. The number of basic handwritten Telugu characters considered in this proposed system is 50 and their system considered 18,000 samples in total (50×360). All the documents are collected from various scribes and they are scanned at 300 dpi and stored as digital images.

The intensity and threshold of an image are crucial in character recognition, recent research suggested projection-based text line segmentation with a variable threshold [15]. To enhance the quality of images, in pre-processing they suggested converting the colour image to a grey scale image, further binarization, and then noise reduction. Greyscale conversion method, images are evaluated by calculating the weight of Red, Green, and Blue (primary colours) components from each colour pixel, and further grey images are converted to binary images using the threshold processing method. Threshold values in binary images were calculated using Otsu’s

Binarization method. And morphological errors in the background are removed by employing the morphological opening and closing operation.

The use of greyscale conversion and detection with skew angle; inscription images input RGB conversion and further smoothing on greyscale images would be an efficient method to reduce the high-frequency noise using Wiener filter. Finally, the resulted grey scale pixel values can have used in detecting skew angle. In addition, text position differences in the image were compared to evaluate whether the document is left-skewed or right-skewed.

Using the various method, the methods are Skew detection and correction, Binarization, noise removal, and morphological operations. These are techniques used in the proposed system. Panyama et al. studied the palm leaf character recognition system using a transform-based technique; for which digitalization and storage of palm leaf manuscripts utilizes a specific 3D function which is proportional to the pressure exerted at that stage by the scribe [16]. The advantage is that in the YZ plane, the precision is higher at 96 per cent and the demerit is smaller than in other planes as reported.

Ptak et al. studied and reported projection-based text line segmentation with variable threshold for which new algorithm were created for text line separation in handwriting using the projection profile. It employs thresholding, but the threshold value is variable [15]. This permits the determination of low overlapping peaks of the graph. The disadvantage of this paper is that the algorithm does not deal well with slanting and curved text lines.

This research work is devoted to presenting an experimental methodology focused on enhancing the inscriptions from the Brihadeeswara Temple in Tamil stone through an array of preprocessing methods. The intent is to unlock and revive the valuable historical information embedded in these artifacts.

Research aim

The aim of this research is to enhance the readability and extract valuable historical and technological information from ancient inscriptions, focusing on the Brihadeeswara Temple in Tanjore. These inscriptions hold crucial insights into advanced Vedic building technologies and historical narratives. The objective is to employ modern digital technologies, including image processing, data science, machine learning, and artificial intelligence, to overcome the challenges posed by the degradation, distortion, and noise present in the aged inscriptions. The primary focus will be on inscriptions in various states of decay and with significant levels of noise. The research will involve a multi-step preprocessing approach to restore the image quality, rectify distortions, and enhance

critical image features. This includes techniques such as grey conversion, resizing, brightness and contrast correction using histogram adjustments, and filtering methods like Gaussian, median, and bilateral filters for noise reduction. The application of adaptive thresholding and morphological processing will also be explored to separate dark and light regions and improve character visibility. By employing these advanced techniques, the study aims to uncover the hidden knowledge encoded in the inscriptions, contributing to a deeper understanding of Vedic technologies and historical contexts. The successful extraction of information from these inscriptions could potentially provide insights into architectural practices and ancient building methods.

Methodology

Dataset

This paper takes a camera-captured Stone inscription as the Input. The inscriptions in Tanjore Brahadeswar Temple, are shown in Figure. These data are different in size and background. The camera used to capture these images are of very high resolution or quality (DSLR). The default format of these images is png or jpeg. These data cannot be used directly for the training purpose since it belongs to several years back, the letter impressions were faded. Hence Archeology department uses white chalks or paints to distinguish the characters and background. Character recognition for this language imposes a challenge as there is a high range of noise detected in stone images. Hence Preprocessing is the foremost step applied to stone images. Figure 1 camera captured original image from Archeology record.

Preprocessing of images

Image enhancement includes mechanisms for enhancing image quality, allowing improved visual and computational analysis. It is commonly used in many applications because of its ability to solve some of the limitations that image acquisition systems pose. Deblurring, removal of noise, and improving contrast are some examples of image enhancement operations. Some of the image quality factors are discussed below.

Thresholding

One of the simplest, most powerful, and most frequently used segmentation algorithms is thresholding-based segmentation. It is useful in discriminating the foreground from the background. Thresholding yields a binary image, which reduces the complexity of data and simplifies the process of recognition and classification [17, 18]. There are three types of thresholding approaches, namely, Global, Local, and Adaptive.

The adaptive mean is the basic method of thresholding in which the current pixel value of the image is replaced with mean or average of all the neighboring pixels and that value is compared with current pixel value. If the value of current pixel is less than the mean value then it is set to black otherwise it is set to white. Adaptive thresholding using a Gaussian filter depends on value of standard deviation. As the value of standard deviation increases more noise is suppressed but the image also gets blurred respectively.

A threshold Flm is a value such that

$$O(i, j) = \begin{cases} 0, & lm < Flm * \frac{(1-t)}{100} \\ 255, & otherwise \end{cases}$$

where $O(i, j)$ is the binarized image and $lm \in [0,1]$ be the intensity of a pixel at location (i, j) of the image.

A global thresholding technique makes use of a single threshold value for the whole image, whereas a local thresholding technique makes use of unique threshold values for the partitioned sub-images obtained from the whole image. In adaptive thresholding, for each pixel in the image, a threshold has to be calculated. However, automatic selection of optimally significant robust values is a difficult challenge. If the pixel value is below the threshold, it is set to the background value; otherwise, it assumes the foreground value [19, 20].

Otsu thresholding method [21] is a highly applicable fully automatic global thresholding algorithm. It is based on separability maximization in the grey level classes. Otsu's method looks for the threshold value which can minimize the intra-class variance. The intra-class variance can be defined as the two classes' weighted sum of variances.

$$g(i, j) = \begin{cases} 1 & \text{iff } (i, j) > T \\ 0 & \text{iff } (i, j) \leq T \end{cases}$$

Assuming an image is represented in $L \{0, 1, 2, \dots, L-1\}$ gray levels, the number of pixels at level i is denoted by n_i , and the total number of pixels is denoted by $i MN = n_0 + n_1 + n_2 + \dots + n_{L-1}$.

select a threshold $T(k) = k, 0 < k < L-$, and use it to.

Assume the image is divided into two categories $C1$ and $C2$ (target and background) with threshold k , then $C1$ and $C2$ respectively correspond to the pixels whose grey levels are $\{0,1,\dots, k\}$ and $\{k+1,k+2,\dots, L-1\}$.

The gray level probability distributions for the two classes are $p_1(k) = \sum_{i=0}^k p_i, p_2(k) = \sum_{i=k+1}^{L-1} p_i = 1 - p_1(k), p_1 m_1 + p_2 m_2 + m_G$

$p_1 + p_2 = 1$. The global variances for the two classes are $\sigma_G^2 = \sum_{i=0}^{L-1} (i - m_G)^2 p_i$

The class variance and measure separability for the two classes are $\sigma_B^2(k) = p_1 p_2 (m_1 - m_2)^2 = \frac{(m_G p_1(k) - m(k))^2}{p_1(k)(1-p_1(k))}$,

$$\eta = \frac{\sigma_B^2}{\sigma_G^2}$$

In the gray range $[0, L]$, the t with maximum $0 \leq \eta(k^*) \leq 1$ is the optimal threshold.

Niblack and savoula

Niblack proposed a simple local adaptive threshold, where a threshold [22–25] is determined for each pixel based on statistics computed from a local window centered on the pixel of interest [26]. Because the threshold is adaptive, it can potentially handle cases of foreground and background intensity distribution overlap (e.g., compare Figs. 2 to 3). Specifically, Niblack thresholding uses the local mean and local standard deviation:

$$\mu(i, j) = \frac{1}{w^2} \sum_{i'=i-w}^{i+w} \sum_{j'=j-w}^{j+w} I(i', j')$$

$$\eta = \frac{\sigma_B^2(k)}{\sigma_G^2}$$

$$(i, j) = \sqrt{\frac{\sum_{i'=i-w}^{i+w} \sum_{j'=j-w}^{j+w} (I(i', j') - \mu(i, j))^2}{w^2}}$$

$$0 \leq \eta(k^*) \leq 1$$

where w is called the window size and controls how much context is used to compute these statistics. The per-pixel Niblack threshold is then

$$T_N(i, j) = \mu(i, j) + k\sigma(i, j)$$

where k is a user-set parameter that controls the trade-off between foreground detection precision and recall. The recommended parameter setting is $k = -0.2$, though the optimal k depends on the image and chosen window size. Binarization is then accomplished with

$$B(i, j) = \begin{cases} 0 & I(i, j) < T_N(i, j) \\ 255 & I(i, j) \geq T_N(i, j) \end{cases}$$

One issue with Niblack is when the window covers only background pixels, it causes the darkest background pixels to be set to foreground (Fig. 3). While this noise is often large, the background immediately around the



Fig. 3 Sample images of stone inscriptions with diverse backgrounds

text is correctly identified, which makes Niblack thresholding useful in combination with other binarization techniques.

Sauvola and Pietikäinen proposed a variant of Niblack to solve the problem with background-only windows.

$$T_S(i, j) = \mu(i, j) \left[1 + k \left(\frac{\sigma(i, j)}{R} - 1 \right) \right]$$

where $\mu(i, j)$ and $\sigma(i, j)$ are computed as in Niblack, $k=0.5$ is the recommended value for the user-set parameter, and R is a constant set to the maximum possible standard deviation, i.e., $R=128$ for 256 Gy levels. While Niblack takes $\mu(i, j)$ and adjusts downward based only on the $\sigma(i, j)$ Sauvola adjusts downward based on $\mu(i, j) \sigma(i, j)$. In windows of only background, $\mu(i, j)$ is relatively large, so $T_S < T_N$, which means fewer of these background pixels are set to foreground.

In analyzing heritage stone inscriptions, selecting a thresholding method hinges on diverse factors: inscription traits, image quality, noise levels, and specific analysis goals. Global thresholding, exemplified by Otsu’s Method, is effective with uniform lighting but falters in varied conditions, struggling to establish a single threshold for the entire image. On the contrary, local adaptive methods like Niblack and Sauvola excel in handling lighting variations and background textures, adjusting thresholds locally for distinct inscription areas with varying contrast or degradation. While they demonstrate

adaptability, precise parameter tuning—like window size (w) and user-defined parameters (such as ‘ k ’ in Niblack or ‘ R ’ in Sauvola)—is essential. These local adaptive methods, owing to their capability to adapt thresholds based on local statistics, can be particularly beneficial in deciphering text from backgrounds in challenging scenarios where lighting or degradation varies across the inscription. Nonetheless, achieving optimal segmentation necessitates careful parameter experimentation and potential combination with other techniques to mitigate specific limitations encountered with individual thresholding methods.

Blur

Median blur

The median filter [27] is a nonlinear signal processing technology based on statistics. The noisy value of the digital image or the sequence is replaced by the median value of the neighborhood (mask). The pixels of the mask are ranked in the order of their grey levels, and the median value of the group is stored to replace the noisy value.

The median filtering output is

$$y[m, n] = median\{x[i, j], |(i, j) \in \omega\}$$

where $x[i, j], y[m, n]$ are the original image and the output image respectively, W is the two-dimensional mask: the mask size is $n \times n$ (where n is commonly odd) such as 3×3 ,

5*5, etc.; the mask shape may be linear, square, circular, cross, etc.

Gaussian blur

The implementation of a 2-D Gaussian filter is widely used for smoothing and noise removal. It requires lots of computational resources and its efficiency in the implementation has been a motivating research area. Convolution operators are the Gaussian operators and the idea of Gaussian smoothing is achieved by convolution (Hypermedia Image Processing Reference 1994). The Gaussian operator in 1-D is given as:

$$G(x) = 1/\sigma\sqrt{2\pi}e^{-\frac{(x-a)^2}{2\sigma^2}}$$

The Gaussian operator in 2D (circularly symmetric) is given as:

$$G(x, y) = 1/2\pi\sigma^2e^{-\frac{x^2+y^2}{2\sigma^2}}$$

where σ (Sigma) indicates the standard deviation of the Gaussian function. If it has a large value, the image smoothing effect will be higher. (x, y) indicates the Cartesian coordinates of the image which show the dimensions of the window.

This filter is composed of addition and multiplication processes between the image and the kernel, where the image is represented by a matrix with a value from 0 to 255 (8 bits). The kernel is a normalized square matrix (values between 0 and 1). The kernel is represented by several bits. For the convolution process, the product of each bit of the kernel and each element of the image is then divided by a power of 2.

In the preprocessing of stone inscriptions, the median filter serves a valuable role in reducing sudden, unwanted noise, often caused by scratches or imperfections on the stone surface. It effectively replaces noisy pixel values with the median value of their neighborhood, contributing to a clearer depiction of the inscribed content. On the other hand, the Gaussian blur is adept at smoothing out minor irregularities and noise present in the image. Importantly, it achieves this without compromising the fundamental integrity of the inscribed text or intricate details. Together, these preprocessing techniques work harmoniously to enhance the overall quality and legibility of stone inscriptions for subsequent analysis or interpretation.

Enhance brightness and contrasts

CLAHE's (Contrast-limited adaptive histogram equalization) [28, 29] basic idea is to perform the histogram equalization of the image's non-overlapping sub-areas,

using interpolation to correct boundary inconsistencies. CLAHE also has two important hyper parameters: Clip Limit (CL) and Tile Number (NT). The first one (CL) is a numeric value that governs the noise amplification. When the histogram of and sub-area is determined, they are redistributed in such a way that its height does not surpass a specified "clip limit." Instead, the total histogram is determined to perform the equalization. The second (NT) is an integer value that governs the sum of non-overlapping sub-areas: the image is divided into many (usually squared) non-overlapping regions of similar sizes, based on its value [30–33].

If numbers of pixels and grayscales, in each region, are respectively M and N , and if $h^{ij}(n)$, for $n=0, 1, 2, \dots, N-1$, is the histogram of (i, j) region, then an estimate of the corresponding CDF, properly scaled by $(N-1)$ for grayscale mapping, is

$$f_{ij}(n) = \frac{N-1}{M} \cdot \sum_{k=0}^n h_{ij}(k);$$

$$n = 1, 2, 3 \dots N - 1$$

This function can be used to convert the given grayscale density function, approximately, to a uniform density function. This procedure is referred to as histogram equalization. In order to limit the contrast to a desired level, the maximum slope of (1) is limited to a desired maximum slope. One approach in limiting the maximum slope is to use a clip limit β to clip all histograms. The clip limit β is obtained by:

$$\beta = \frac{M}{N} \left(1 + \frac{\alpha}{100} (s_{max} - 1) \right)$$

where α is a clip factor, if clip factor is equal to zero the clip limit becomes exactly equal to (M/N) results into an identity mapping by evenly distributing all regional pixels into all possible grayscales. Moreover, if clip limit is equal to 100 the maximum allowable slope is s_{max} .

Gamma, alpha and beta correction: Gamma correction [34–36] is a non-linear adjustment to individual pixel values. While in image normalization we carried out linear operations on individual pixels, such as scalar multiplication and addition/subtraction, gamma correction carries out a non-linear operation on the source image pixels, and can cause saturation of the image being altered.

$$I_{out} = cI_{in}^{\gamma}$$

where I_{in} and I_{out} are the input and output image intensities, respectively. c and γ are two parameters that control the shape of the transformation curve. In gamma

correction, controls the slope of the transformation function. The higher the value of γ is, the steeper the transformation curve becomes. And the steeper the curve is, the more the corresponding intensities are spread, causing more increase of contrast [37, 38].

Adjusting the brightness mean, either increasing the pixel value evenly across all channels for the entire image to increase the brightness, or decreasing the pixel value evenly across all channels for the entire image to decrease the brightness.

$$g(x) = \alpha f(x) + \beta$$

The parameters $\alpha > 0$ and β are often called the gain and bias parameters; sometimes these parameters are said to control contrast and brightness respectively [39, 40].

These methods, particularly CLAHE and gamma correction, when applied to stone inscriptions, work synergistically to optimize contrast, control noise, and enhance the overall quality of the inscribed content. By fine-tuning parameters and employing adaptive techniques, these methods cater specifically to the challenges posed by stone inscriptions, resulting in improved readability and preservation of historical content.

Bilateral filter

Bilateral filtering [41] is a non-linear filtering method, where the weight of each pixel is determined using a Gaussian in the spatial domain, multiplied by an impact function in the intensity domain, which decreases the weight of pixels with large variations in intensities. Pixels which vary greatly in intensity from the central pixel are weighted less, even though they may be near to the central pixel. It is then implemented as two Gaussian filters in a localized pixel neighborhood, one in the space domain, called the domain filter that smoothes homogeneous regions, and one in the strength domain, called the range filter that regulates edge preservation smoothing. The key advantage of using bilateral filters is therefore the development of broad and homogeneous areas [42].

$$g(x) = (f * G^s)(x) = \int_R f(y) G^s(x-y) dy$$

The weight for $f(y)$ is equal to $G^s(x-y)$ and depends only on the distance from space $\|x-y\|$. The bilateral filter introduces a concept of weighting that is dependent on the tonal distance $f(y)-f(x)$. The result:

$$g(x) = \frac{\int_R f(y) G^s(x-y) G^t(f(x)-f(y)) dy}{\int_R G^s(x-y) G^t(f(x)-f(y)) dy}$$

Remember that since the weights are directly dependent on the values of the image, we need explicit normalization so that the 'sum' of all weights equals one [43].

When applied to stone inscriptions, bilateral filtering proves invaluable in refining image quality by smoothing homogeneous regions while maintaining the integrity of intricate details and edges within the inscriptions. Its ability to selectively smooth and preserve edges makes it a powerful tool for enhancing the clarity and readability of the inscribed content.

Resize and gray scale

Images with a large size also have to be resized to a smaller size that is adequate for appropriate distinguishing, because an increase in the size of the input image results in an increase in the parameter to be measured, the computing power needed, and memory. Gray scale conversion process removes all color information, leaving only the luminance of each pixel [43, 44].

In the context of stone inscriptions, these methods—resizing large images and converting to grayscale—are crucial for optimizing computational resources, memory usage, and focusing on the inherent content details essential for analysis and interpretation. They ensure efficient processing while retaining critical information necessary for further analysis.

Dilation and erosion

Dilation is a morphological operation which extends the image to bright structures. To this end, the new gray-value of each pixel in the structuring factor centered at this pixel is defined as the sum of the old gray-values of all pixels. Filtering the dilation can also be iterated. Multiple dilation steps with a given structuring element are equivalent to a single dilation step with a larger structuring element (only an up scaled version of the small structuring element for convex structuring elements). Similarly, erosion extends dark structures in the image by replacing the gray-value of each pixel with the minimum of old gray-values of all pixels within the structuring element [44–47].

Erosion: The erosion transformation of X by B is defined as the set of points x such that the translated B_x is contained in X and is expressed as below (1)

$$X \ominus B = \{x | B_x \subseteq X\}$$

Erosion operation shrinks the image object. The degree of required shrinking and the direction of the shrinking can be controlled by defining the characteristics of the structuring element. If we reduce the size of the structuring element, the "harshness" of the erosion process reduces. Alternatively, various degrees of shrinking can be achieved

by performing iterative erosions by considering the primitive element [48–50].

Dilation: The dilation transformation of X by B is defined as the set of points x such that the intersection of the translated structuring element Bx and image X is not a null set [51, 52]

$$X \oplus B = \{X|Bx \cap X \neq \emptyset\}$$

Opening: The multiscale opening of X by a structuring element B of size n is the combination of erosion followed by dilation by the structuring element nB [53].

$$X \circ nB = (X \ominus nB) \oplus nB, n = 1, 2, \dots, N.$$

Closing: Closing of X by a structuring element B is dilation by nB followed by erosion by nB . Opened image is the subset of main image but main image is the subset of closed image [54, 55].

$$X \cdot nB = (X \oplus nB) \ominus nB, n = 1, 2, \dots, N.$$

Morphological operations—erosion, dilation, opening, and closing—serve as fundamental tools in processing stone inscription images. Erosion shrinks image structures, controlled by the structuring element, while dilation extends and highlights structures. Both erosion and dilation can be iteratively adjusted for fine-tuning. Opening, combining erosion and dilation, removes noise while preserving essential features, whereas closing fills gaps while maintaining overall integrity. These operations are pivotal in manipulating and refining image structures, enhancing the interpretation and analysis of crucial details within stone inscriptions.

Edge detection-laplacian, sobex, sobely and canny

Canny edge detection

It uses the first derivative of Gaussian to detect the edges of an images [56, 57]. The approach is based on convolution of the image function($f(x,y)$) with the following Gaussian operator:

$$G(x, y; \sigma) = \frac{1}{2\pi\sigma^2} e^{-(x^2+y^2)/2\sigma^2}$$

where σ is the spread of the Gaussian which controls the degree of smoothing. A new function $f'(x, y)$ is computed as

$$f'(x, y) = G(x, y; \sigma) \times f(x, y)$$

Then, using the gradient of a pixel(x,y) in the $f'(x, y)$, the edges of the $f(x,y)$ image can be detected.

Laplacian

LoG edge detector is based on the second order derivative of a Gaussian function. Consider the Gaussian function [58, 59].

$$G(x, y) = e^{-\frac{(x^2+y^2)}{2\sigma^2}}$$

The Laplacian of this function is

$$\begin{aligned} \nabla^2 G(x, y) &= \frac{\partial^2 G(x, y)}{\partial x^2} + \frac{\partial^2 G(x, y)}{\partial y^2} \\ &= \left[\frac{x^2+y^2-2\sigma^2}{\sigma^4} \right] e^{-\frac{(x^2+y^2)}{2\sigma^2}} \end{aligned}$$

This function is called Laplacian of Gaussian (LoG). So, the image is convolved with this function and produces two effects- smoothing, and computing the Laplacian which yields a double-edge image. The edges are then detected by finding the zero crossings between the double edges.

Sobel

Functional derivative reflects the marked extent of image gradation variety. The Local Maximum of first derivative reflect the max extent of image gradation variety. The derivative value can be used as edge intensity value, so the edges can be detected by setting threshold [59].

$$G(x, y) = \begin{bmatrix} G_x \\ G_y \end{bmatrix} = \begin{bmatrix} \frac{\partial f}{\partial x} \\ \frac{\partial f}{\partial y} \end{bmatrix}$$

Sobel [60] operator utilizes two convolution kernels to calculate first-order derivative. The convolution kernels are

$$G_x = \begin{bmatrix} -1 & 0 & 1 \\ -2 & 0 & 2 \\ -1 & 0 & 1 \end{bmatrix} \quad G_y = \begin{bmatrix} 1 & 2 & 1 \\ 0 & 0 & 0 \\ -1 & -2 & -1 \end{bmatrix}$$

In stone inscription analysis, these methods are applied to detect and highlight edges crucial for interpreting inscribed content. Canny Edge Detection's ability to preserve edges while smoothing the image aids in capturing intricate details. LoG's double-edge image and zero-crossing detection enhance edge detection accuracy. Sobel's use of derivative values helps discern edge intensities, facilitating edge detection by thresholding. Together, these methods contribute significantly to unveiling and analyzing crucial details within stone inscriptions.

K-means

Clustering or data grouping is a key initial procedure in image processing. K-means is typically used to locate

objects and boundaries in images. It is used to find natural clusters within given data based upon varying input parameters. Clusters can be formed for images based on pixel intensity, color, texture, location, or some combination of these. The membership for each data point belongs to its nearest center, depending on the minimum distance [61].

$$c_j = \frac{\sum_{i=1}^n m(c_j|x_i)w(x_i)x_i}{\sum_{i=1}^n m(c_j|x_i)w(x_i)}$$

$$A(X, C) = \sum_{i=1}^n \min_{j=(1..i)} |x_i - c_j|^2$$

where centers is C and for each data point x_i , compute its minimum distance with each center c_j . For each center c_j , recomputed the new center from all data points x_i belong to this cluster.

In stone inscription analysis, K-means clustering can be applied to segment the inscribed content from the background or distinguish various elements within the images. It helps in identifying patterns, delineating different components, or separating text from non-text regions. By grouping pixels based on their characteristics, this technique facilitates subsequent analysis, interpretation, or feature extraction from stone inscription images.

Fast means denoising

Fast means denoising is replacing the color of a pixel with an average of the colors of similar pixels. NLM, neighborhood weightages are computed using the window similarity technique [62]. In this filter method, each pixel $I'(x_i, y_i)$ is estimated as weighted mean of all the pixels in the image as shown in below equation

$$I'(x_i, y_i) = \sum_{(x_j, y_j) \in \Omega} w(i, j) I(x_j, y_j)$$

where the weight $w(i, j)$ between two pixels (x_i, y_i) and (x_j, y_j) depends on their similarity as shown in below equation

$$w(i, j) = \frac{1}{Z(i)} \sum_{l \in \text{levels}} e^{-g_{l|l} \left[\frac{(I_{l|l}(x_i, y_i) - I_{l|l}(x_j, y_j))^2}{\alpha_{l|l} \sigma^2} \right]}$$

where $(I_{l|l}(x_i, y_i))$ is the pixel mean value of window of level $l|l$ size which is centered at (x_i, y_i) . σ is the standard deviation of the Gaussian noise of the input image, $Z(i)$ is a normalized constant, $g_{l|l}$ is a gaussian weightage for image level $l|l$ and $\alpha_{l|l}$ value is a scale factor to the computed image noise, which maps the noise variance to the corresponding image level $l|l$.

In stone inscription analysis, these denoising techniques—Fast Means and NLM—are instrumental in reducing noise, enhancing image clarity, and preserving critical details within the inscriptions. By averaging similar pixels or considering their weighted similarity, these methods effectively mitigate noise, allowing for clearer interpretation and analysis of the inscribed content. Table 1 displays a summarized comparison of various pre-processing techniques.

MSE

The mean square error (MSE) [62, 63] is the cumulative square error between the restored image and the original image defined as:

$$MSE = \frac{1}{MN} \sum_{y=1}^M \sum_{x=1}^N [I(x, y) - I'(x, y)]^2$$

where, $M \times N$ is the Image size, $I(x, y)$ is an original image and $I'(x, y)$ the restored image.

The peak signal-to-noise ratio (PSNR) is the peak value of signal-to-noise ratio (SNR) and in another words it is defined as the ratio of the maximum possible power of a pixel value and the power of distorting noise. As it known that, it affected the original image quality. It is defined as:

$$PSNR = 20 \times \log_{10} \left(\frac{255}{\sqrt{MSE}} \right)$$

where, 255×255 is the maximum value of pixel present in an image and MSE is calculated for original and restored image with $M \times N$ size.

Results and discussion

In this experimental study, the dataset consists of images captured using a camera, featuring stone inscriptions from the Tanjore Brihadeeswar Temple dating back to the eleventh century, during the reign of Raja Raja Chola. A total of 200 stone inscription images were included in the dataset, having been previously scanned. These images are of a resolution of 200×200 dots per inch (dpi) and are in RGB format. Several sample images from the original dataset are illustrated in Fig. 3. The study encompasses various categories of images, such as Dark-illegible, Flawless-bright-illegible, Flawless-bright-legible, Flawless-irregular-moderate, highly impaired-irregular-illegible, Flawless-bright-legible, Flawless-bright-moderately legible, Flawless-dull-moderate, highly impaired-dark-illegible, Impaired-dark-moderate, Very impaired-dusky dark-legible, Impaired-dull-moderately legible, Impaired-dusky dark-moderate. To enhance the images for

Table 1 Tabulated summary of all pre-processing technique

Preprocessing technique	Description	Application in Stone Inscriptions
Thresholding	Techniques like Otsu's Method, Niblack, and Sauvola adaptively establish thresholds, aiding in segmenting text from backgrounds and handling varied lighting conditions	Essential for segmenting inscribed content and adapting to diverse lighting conditions
Median filtering	Replaces pixel values with the median of neighboring pixels, reducing sudden noise caused by imperfections on stone surfaces	Effective in removing unwanted noise from scratches or imperfections on stone surfaces
Gaussian blur	Smooths images without compromising the integrity of inscribed text or details, reducing minor irregularities and noise	Useful for refining stone inscription images without losing essential details
CLAHE and gamma correction	Techniques optimizing contrast, controlling noise, and enhancing image quality, specifically adapted for stone inscriptions	Enhance readability and preservation of historical content in stone inscriptions
Resizing	Adjusts image dimensions for efficient computational processing, optimizing memory usage and retaining essential content details	Manages large image sizes while retaining critical information for analysis
Grayscale conversion	Converts images to grayscale, emphasizing contrast and intensity variations, focusing on the legibility of inscribed content	Helps in emphasizing contrast and clarity, aiding in the interpretation of inscriptions
Bilateral filtering	Smooths homogeneous regions while preserving edges and details within inscriptions, enhancing overall image quality	Improves image clarity while maintaining the integrity of intricate details
Morphological operations	Erosion, dilation, opening, and closing refine image structures, aid in edge detection, and manipulate inscription features	Vital in shaping, refining, and enhancing structural details within stone inscriptions
Edge detection	Techniques like Canny, LoG, and Sobel detect and highlight edges, aiding in capturing intricate details within inscriptions	Essential in uncovering and analyzing crucial features and details within stone inscriptions
K-means clustering	Segments inscribed content or distinguishes various elements within images, aiding in identifying patterns or text from non-text regions	Helps in segmenting and identifying different components or patterns within stone inscriptions
Denoising (Fast means, NLM)	Techniques reducing noise and enhancing image clarity, crucial for preserving critical details within the inscriptions	Mitigates noise interference, allowing for clearer analysis and interpretation of inscriptions

subsequent analysis, a series of preprocessing methods were employed, including Brightness and Contrast adjustment, Image Smoothing, Noise Removal, Structural Elements extraction, Separation of Dark and Light regions, Edge detection, and Fast denoising techniques. Separating the dark and light regions within the stone inscriptions is a critical task, and to achieve this, a range of methods were utilized, including Thresholding, Adaptive Thresholding using Gaussian, OSTU, Threshold Triangle, Adaptive Thresholding using Mean, Adaptive Thresholding using Gaussian (binary inversion), Adaptive Thresholding using Mean (binary inversion), Niblack, Savoula, Niblack Dilation, Niblack Morphology, Niblack Erosion, Savoula Dilation, Savoula Morphology, Savoula Erosion, Nilblack k-means, Nilblack k-means noising, Nilblack k-means noising (binarizations), and Nilblack l-means noising (binarization normalization). For the purpose of edge detection, Sobelx, SobelY, Canny edge detection, and Laplacian methods were employed. The Peak Signal-to-Noise Ratio (PSNR) values were

computed for all image types following preprocessing, and these results are presented in Table 2 (Fig. 4).

The original input images were first transformed into grayscale, after which noise reduction was executed using the Gaussian filtering technique. Subsequently, various binarization techniques were applied. An example of the binarized outcomes obtained from the analysis of ancient stone inscriptions using the diverse preprocessing methods is visualized in Fig. 5. To ascertain the suitability of binarization techniques, a tenfold cross-validation process was executed for the 200 images extracted from the ancient stone inscriptions. The implementation of this experiment utilized OpenCV. Before script recognition, it's crucial to undertake image enhancement and restoration steps. Consequently, this experiment entailed the selection of 20 preprocessing filters. To gauge the image quality following preprocessing, metrics such as PSNR and Mean Squared Error (MSE) were computed. Through human visual evaluation, it was observed that neither Niblack nor Suavola algorithms stood out as superior.

Table 2 Comparative PSNR (Peak Signal-to-Noise Ratio) performance among various pre-processing methods

Methods	A	B	C	D	E	F	G	H	I	J	K	L	M
Brightness and contrast													
Alpha, beta, gamma	28.66	28.95	28.7	28.1	29.23	28.67	28.05	27.82	28.41	28.38	28.33	29.12	28.18
Equalizer histogram	27.84	27.99	28.93	28.02	27.89	27.74	27.91	27.92	27.99	27.93	27.98	28.03	27.91
Equalizer histogram threshold	27.91	27.98	28.14	28.06	27.88	27.71	27.86	27.86	28.13	27.87	28	27.99	27.9
Smoothing image													
Blur	28.04	27.79	27.81	27.88	28.14	27.65	27.83	27.97	27.79	27.97	27.89	28.05	28.05
Gaussian blur	28.02	27.8	27.82	27.81	28.08	27.66	27.81	27.85	27.79	27.89	27.87	28.03	27.93
Median blur	27.89	27.86	27.98	27.9	27.86	27.7	27.78	27.75	27.95	27.78	27.89	28.05	27.87
Smooth and remove noise													
Bilateral filter	28.01	27.79	27.75	27.76	28.02	27.61	27.75	27.85	27.8	27.95	27.9	27.99	28.04
Structure elements morphological													
Erosion	28.12	27.88	27.82	27.83	28.15	27.66	27.82	27.92	27.85	28.06	27.98	27.99	28.1
Dilation	28.01	27.79	27.75	27.76	28.02	27.61	27.75	27.85	27.8	27.95	27.9	27.99	28.04
MorpologyEx	28	27.88	27.75	27.81	28.07	27.59	27.78	27.89	27.8	27.96	27.88	27.96	28.08
Separation of dark and light region													
Threshold	34.9	33.48	39.12	35.7	35.71	34.47	34.23	34.32	35.63	33.77	33.21	38.41	32.75
Adaptive threshold Gaussian	34.83	33.43	38.73	35.53	35.62	34.39	34.21	34.19	35.56	33.68	33.1	38.17	32.67
OSTU	34.5	33.39	38.41	35.21	35.23	34.42	34.08	33.91	35.34	33.46	32.99	38.37	52.52
Threshold triangle	34.89	33.48	39.1	35.13	35.31	34.33	34.21	34	35.65	33.34	33.07	36.95	32.45
Adaptive threshold mean	34.44	33.31	37.99	35.27	35.23	34.3	33.97	33.97	35.25	33.5	33	38.15	32.55
Adaptive threshold Gaussian	34.33	32.98	36.5	34.89	35.2	33.52	33.59	33.64	34.4	33.25	32.49	36.32	32.36
Adaptive threshold mean	34.36	32.96	36.85	35.05	35.1	33.47	33.69	33.63	35.25	33.33	32.48	36.15	32.36
Niblack	33.21	32.67	39.04	31.67	34.38	34.48	31.06	30.88	35.42	32.43	32.89	38.61	31.18
Savuola	28.66	28.95	28.7	28.1	29.23	28.67	28.05	27.82	28.41	28.38	28.33	29.12	28.18
Niblack dilation	34.83	33.43	38.73	35.53	35.63	34.39	34.21	34.19	35.56	33.68	33.1	38.17	32.67
Nilblack morphology	34.89	33.41	38.71	35.57	35.62	34.38	34.17	34.24	35.56	33.71	33.09	38.19	32.69
Nilblack erosion	34.77	33.43	38.49	35.39	35.58	34.38	34.12	34.19	35.56	33.63	33.1	38.14	32.67
Savoula dialtion	34.83	33.43	38.73	35.53	35.62	34.39	34.21	34.19	35.56	33.68	33.1	38.17	32.67
Savoula morphology	34.88	33.42	38.73	35.76	35.66	34.34	34.22	34.3	35.55	33.73	33.06	38.15	32.7
Savoula erosion	34.77	33.41	38.49	35.39	35.58	34.38	34.12	34.19	35.52	33.63	33.06	38.14	32.67
Nilblack k-means	34.83	33.43	38.54	35.52	35.62	34.39	34.21	34.19	35.55	33.69	33.1	38.16	32.66
Nilblack k-means noising	34.86	33.41	38.8	35.62	35.66	34.38	34.18	34.28	35.6	33.71	33.07	38.1	32.69
Nilblack k-means noising binarizations	34.41	33.05	36.59	34.96	35.26	33.57	33.64	33.74	34.51	33.31	32.56	36.25	32.36
Nilblack l-means noising binarization normalization	34.41	33.05	36.59	34.96	35.26	33.57	33.64	33.74	34.51	33.31	32.56	36.52	32.36
Edge detection													
Sobelx	33.69	32.46	34.89	34.28	34.07	32.65	33.12	30.66	31.04	30.59	30.24	31.52	29.86
SobelY	33.75	32.63	35.31	34.21	34.13	32.83	33.07	34.32	33.69	32.61	31.82	33.92	31.92
Canny edge	34.3	33.04	36.63	34.98	34.94	33.53	33.65	34.19	33.64	32.63	31.81	34.24	31.92
Laplacian	30.99	30.42	31.38	31.39	31.54	30.39	30.73	33.91	34.45	33.29	32.55	36.33	32.31
Fast denosing	34	33	36	34	35.2	33.5	33.6	33.7	34.5	33.3	32.5	36.5	32.3

A dark-illegible, B flawless-bright-illegible, C flawless-bright-legible, D flawless-irregular-moderate, E highly impaired-irregular-illegible, F flawless-bright-legible, G flawless-bright-moderately legible, H flawless-dull—moderate, I highly impaired-dark-illegible, J impaired-dark-moderate, K very impaired-dusky dark-legible, L impaired-dull-moderately illegible, M impaired-dusky dark-moderate

Therefore, a total of 34 algorithms were selected. The performance of the selection system is outlined in Table 2. The results of this experiment highlight that Adaptive Threshold Gaussian, Otsu’s algorithm, and Nilblack k-means exhibited improved performance. When

comparing outcomes derived from global thresholding methods with those from local adaptive thresholding methods, it was evident that local adaptive thresholding methods yielded a more stable character appearance by fine-tuning outputs at local levels, as opposed to global

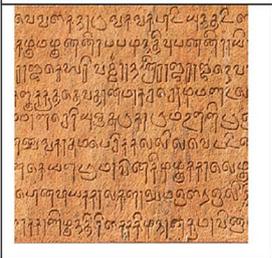
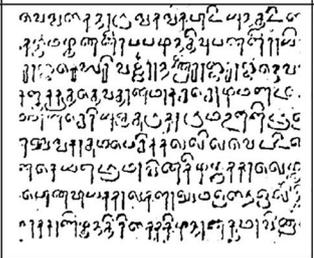
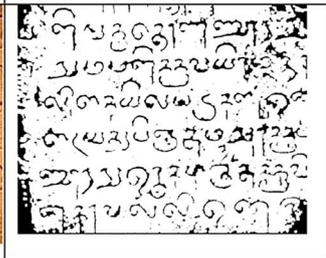
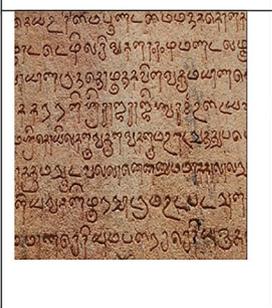
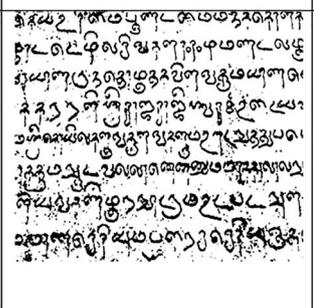
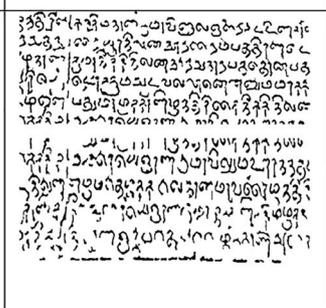
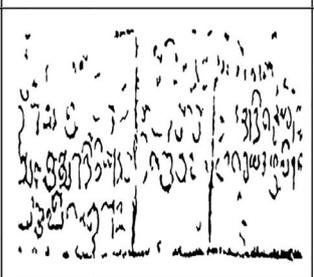
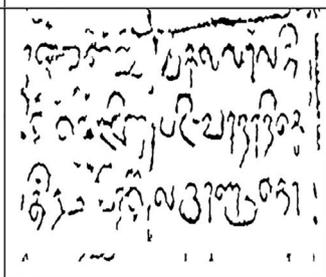
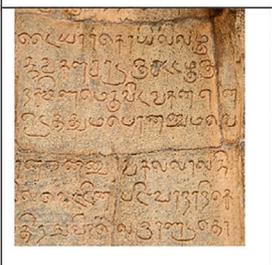
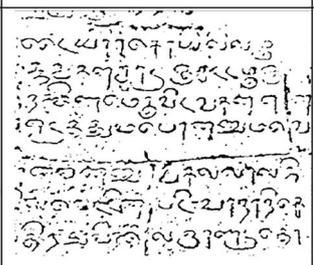
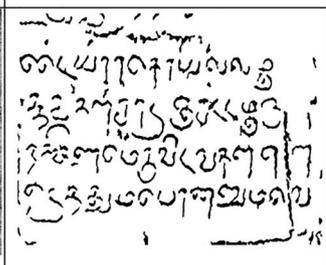
<p>Test case :1 Flawless- bright-legible</p>	<p>(a)</p> <p>MSE: 5.479662754927652, PSNR: 40.743265301537825</p>	<p>Test case: 2 Flawless-bright- legible</p>	<p>(b)</p> <p>MSE: 20.09071525489436, PSNR: 35.100849624111355</p>
			
<p>Test case:3 Flawless- moderately -legible</p>	<p>(c)</p> <p>MSE:20.32192871656873 PSNR:35.0511543721603</p>	<p>Test case: 4 Flawless-bright- moderately legible</p>	<p>(d)</p> <p>MSE:12.942669821551808, PSNR:37.01056488772525</p>
			
<p>Test case: 5 Higly impaired-dark -irregular - llegible</p>	<p>(e)</p> <p>MSE: 23.584614557574454, PSNR: 34.404515778857814</p>	<p>Test case:6 Impaired-dull- moderately lible</p>	<p>(f)</p> <p>MSE: 25.35455590686979, PSNR: 34.09024352723713</p>
			
<p>Test case:7 Flawless-dull - moderate</p>	<p>(g)</p> <p>MSE:16.31160988948388, PSNR:36.005835346151414</p>	<p>Test case:8 dark-legible</p>	<p>(h)</p> <p>MSE:20.460427253905515, PSNR:35.02165662475339</p>
			

Fig. 4 Sample Original Image and Binarized image MSE and PSNR result of diverse type of stone background structure inscriptions (Test case: 1. Flawless-bright-legible-1; 2. Flawless-bright-legible-2; 3. Flawless moderately-legible; 4. Flawless bright-moderately-legible; 5. Highly impaired dark irregular illegible; 6. Impaired-dull-moderately-legible; 7. Flawless dull moderate; 8. Dark legible)

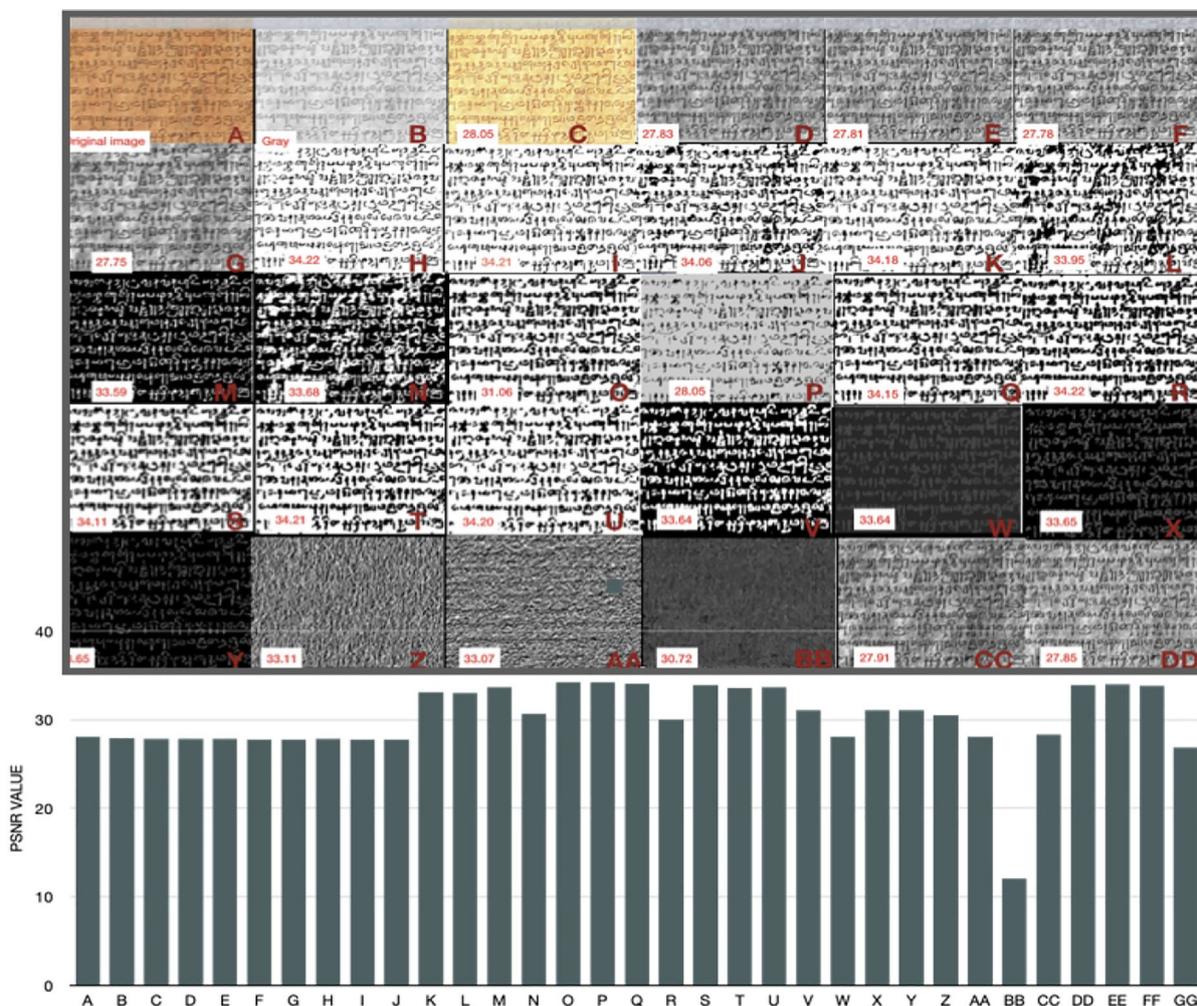


Fig. 5 Results achieved through binarization with the utilization of various preprocessing approaches. **A** Original image, **B** Brightness and contrast, **C** Gray, **D** Blur, **E** Gaussian blur, **F** Median blur, **G** Bilateral, **H** Adaptive Gaussian, **I** Binary, **J** OSTU, **K** Binary triangle, **L** Adaptive mean, **M** Adaptive gaussian bin inv, **N** Adaptive mean bin inv, **O** Niblack, **P** Sauvola, **Q** Adaptive Gaussian morphological line, **R** Erosion, **S** dilation, **T** Adaptive Gaussian k-means, **U** Adaptive Gaussian k-means fast denoising, **V** Binarization, **W** Normalize, **X** Canny edge, **Y** Canny edge Thinning, **Z** Sobel x, **AA** Sobel y, **BB** Laplacian, **CC** Equalizer histogram, **DD** Equalizer histogram threshold

techniques [64–66]. While global thresholding methods managed to eliminate noise from certain background regions, they often rendered characters in other areas illegible [67]. The study involved the generation and preprocessing of stone inscription images, encompassing smoothening, noise reduction, filtering, structural element extraction, brightness enhancement, contrast enhancement, denoising, edge detection, and dark–light region separation. The investigation aimed to determine the most effective preprocessing technique by calculating PSNR and MSE values. It was noted that for images with a dark background, the median filter and the Niblack method outperformed other preprocessing techniques. Recent research has highlighted Adaptive Gaussian

thresholding as effective for sharp bright images [40], which aligns with the findings of this study for primitive pillar stone inscriptions that are typically dull. Ancient Tamil stone inscriptions have proven challenging to interpret due to their high number of characters with subtle differences [68]. Research also suggests that performance metrics like PSNR play a pivotal role in assessing existing preprocessing methods [69], with denoising interventions having a positive impact on image resolution and character recognition [70].

It sounds like the preprocessing steps are critical in achieving clear and complete character extraction from the background. From what you’ve described, the order of operations for preprocessing is crucial in determining

the success of separating foreground characters from the background. It appears that the combination of preprocessing steps involving brightness and contrast adjustment, grayscale conversion, resizing, median and Gaussian blur, erosion and dilation, adaptive thresholding (like Gaussian and Sauvola), followed by k-means denoising, leads to successful character extraction without breakage or discontinuity in the flawless-bright-legible/moderate background scripts shown in Fig. 4a–d. However, when the same preprocessing steps are applied to darker or more irregular backgrounds (such as dark/dusky dark –legible/moderate and impaired /illegible /dull/irregular backgrounds), the resulting images show added noise, broken character pixels, discontinuity, and meshed or meaningless character retrieval as shown in Fig. 4e–h. This poor result leads to a lower recognition rate. It seems that the challenge lies in adapting the preprocessing techniques to different types of backgrounds, especially when dealing with darker, irregular, or noisy backgrounds. Further optimization or adjustment of the preprocessing steps might be necessary to handle such variations in backgrounds for better character extraction and recognition, especially in cases with more challenging backgrounds.

Limitations of current work: The study's primary limitation revolves around its focused exploration of preprocessing methods. While it examines numerous techniques, it doesn't encompass the entire array of available algorithms and methodologies. This selective approach might have overlooked potentially effective techniques or synergistic combinations that could offer improved outcomes but were not included in this study. Additionally, while the study successfully applies various preprocessing methods to enhance image quality and facilitate character recognition in ancient Tamil inscriptions, it primarily focuses on the preprocessing aspect rather than directly addressing script recognition. This implies that while the images are improved for analysis, the study does not delve deeply into the accuracy or efficacy of recognizing and interpreting characters within these inscriptions. Moreover, the evaluation metrics, like PSNR and MSE, while informative about image quality improvement, might not completely capture the nuances of text legibility or character recognition accuracy, which are critical for historical script analysis. Lastly, the study doesn't extensively address the optimization of parameters within each preprocessing method. Optimizing these parameters could potentially yield better results in terms of character recognition or text legibility for ancient scripts.

In summary, while the research makes significant strides in enhancing stone inscription images, it is constrained by the limited scope of preprocessing methods

assessed and the lack of direct analysis on character recognition accuracy. This leaves room for future investigations to explore a wider array of preprocessing techniques and to deeply assess the effectiveness of these methods in improving the interpretation of ancient scripts.

Conclusion

Overall 34 pre-processing methods were assessed for best PSNR values for different type of images (Dark-illegible, Flawless-bright-illegible, Flawless-bright-legible, Flawless-irregular-moderate, highly impaired-irregular-illegible, Flawless-bright-legible, Flawless-bright-moderately legible, Flawless-dull-moderate, Highly impaired-dark-illegible, Impaired-dark-moderate, Very impaired-dusky dark-legible, Impaired-dull-moderately legible, Impaired-dusky dark-moderate). In the realm of Tamil literature, the task of identifying characters is of paramount importance, although it remains a challenging and intricate endeavor due to the presence of numerous characters in ancient Tamil scripts that exhibit resemblances or subtle deviations. Within the scope of this investigation, it was revealed that employing a combined approach involving the Nilblack-k-means method effectively handles the inscription processing, ultimately yielding processed images of elevated resolution.

Acknowledgements

The first author is thankful to Anna University Chennai for supporting the research.

Author contributions

Conceptualization, JJ and PUM. Methodology, JJ and PUM. Investigation, JJ; writing original draft preparation, JJ; writing review and editing JJ and PUM.

Funding

Not applicable.

Data availability

The datasets used and/or analysed during the current study available from the corresponding author on reasonable request.

Declarations

Ethical approval and consent to participate

Not applicable.

Competing interests

The authors declare no competing interests.

Received: 13 November 2023 Accepted: 3 February 2024

Published online: 20 February 2024

References

1. Devi KD, Maheswari PU. Insight on character recognition for calligraphy digitization. In 2017 IEEE Technological Innovations in ICT for Agriculture and Rural Development (TIAR). IEEE. 2017. pp. 78–83.
2. Bhuvaneshwari G, Subbiah Bharathi V. An efficient positional algorithm for recognition of ancient stone inscription characters. In: Advanced

- Computing (ICoAC) 2015 Seventh International Conference on. IEEE. 2015.
3. Janani G, Vishalini V, Mohan Kumar P. Recognition and analysis of Tamil inscriptions and mapping using image processing techniques. *Science Technology Engineering and Management (ICONSTEM) Second International Conference on. IEEE.* 2016.
 4. Devi K, Maheswari PU. Digital acquisition and character extraction from stone inscription images using modified fuzzy entropy-based adaptive thresholding. *Soft Comput.* 2019;23:2611–26.
 5. Mahalakshmi M, Sharavanan M. Ancient Tamil script recognition and translation using LabVIEW. In *International conference on communication and signal processing.* 2013. pp. 1021–6.
 6. Vellingiriraj EK, Balamurugan M, Balasubramanie P. Text analysis and information retrieval of historical Tamil ancient documents using machine translation in image zoning. *Int J Lang Lit Linguist.* 2016;2(4):164–8.
 7. Chaki N, Shaikh SH, Saeed K. A comprehensive survey on image binarization techniques. In: Chaki N, editor. *Exploring image binarization techniques.* Berlin: Springer; 2014. p. 5–15.
 8. Pal U, Roy PP, Tripathy N, Lladós J. Multi-oriented Bangla and Devanagari text recognition. *Pattern Recogn.* 2010;43(12):4124–36.
 9. Durga Devi K, Maheswari PU, Polasi PK, Preetha R, Vidhyalakshmi M. Pattern matching model for recognition of stone inscription characters. *Comput J.* 2023;66(3):554–64.
 10. Buzzykanov SN. Enhancement of poor resolution text images in the weighted Sobolev space. In *2012 19th International Conference on Systems, Signals and Image Processing (IWSSIP).* IEEE. 2012. pp. 536–9.
 11. Kavallieratou E, Antonopoulou H. Cleaning and enhancing historical document images. *Lect Notes Comput Sci.* 2005;3708:681–8.
 12. Seeger M, Dance C. Binarising camera images for OCR. In: *ICDAR 2001.* 2001. pp. 54–9.
 13. Pal U, Chaudhuri BB. Indian script character recognition: a survey. *Pattern Recogn.* 2004;37:1887–99.
 14. Vasantha Lakshmi C, Patvardhan C. An optical character recognition system for printed Telugu text. *Pattern Anal Appl.* 2004;7(2):190–204.
 15. Ptak R, Żygadło B, Unold O. Projection-based text line segmentation with a variable threshold. *Int J Appl Math Comput Sci.* 2017;27:101–414.
 16. Panyam NS, Vijaya Lakshmi TR, Krishnan R, Koteswara Rao NV. Modeling of palm leaf character recognition system using transform based techniques. *Pattern Recogn Lett.* 2016;84:29–34.
 17. Goh TY, Basah SN, Yazid H, Safar MJA, Ahmad Saad FS. Performance analysis of image thresholding: Otsu technique. *Measurement.* 2018;114:298–307.
 18. Ripon S, Chowdhury L, Ashour AS, Dey N. Machine-learning approach for ribonucleic acid primary and secondary structure prediction from images. In: Dey N, Ashour AS, Shi F, Balas VE, editors. *Soft computing based medical image analysis.* Cambridge: Academic Press; 2018. p. 203–21.
 19. Davies ER. The role of thresholding. In: Davies ER, editor. *Computer vision.* 5th ed. Cambridge: Academic Press; 2018. p. 93–118.
 20. Siddique MAB, Arif RB, Khan MMR. Digital Image Segmentation in Matlab: a Brief Study on OTSU's Image Thresholding. In *2018 International Conference on Innovation in Engineering and Technology (ICIET).* IEEE. 2018. pp. 1–5.
 21. Otsu N. A threshold selection method from gray-level histograms. *IEEE Trans Syst Man Cybern.* 1979;9(1):62–6.
 22. Su B, Lu S, Tan CL. Binarization of historical document images using the local maximum and minimum. In: *Proc Intl Workshop on Document Analysis Systems,* 2010. pp. 159–65.
 23. Khurshid K, Siddiqi I. Comparison of Niblack inspired Binarization methods for ancient documents. In: *Proceedings of SPIE,* 2009. pp. 1–10.
 24. Saxena LP. Niblack's binarization method and its modifications for real-time applications: a review. *Artif Intell Rev.* 2017;47(4):469–98.
 25. Sudarsan D, Sankar D. A Novel complete denoising solution for old Malayalam palm leaf manuscripts. *Pattern Recognit Image Anal.* 2022;32(1):187–204.
 26. Sezgin M, Sankur B. Survey over image thresholding techniques and quantitative performance evaluation. *J Electron Imaging.* 2004;13(1):146–65.
 27. Bovik AC. Streaking in median filtered images. *IEEE Trans Acoust Speech Signal Process.* 1987;35(4):493–503.
 28. Stark JA. Adaptive image contrast enhancement using generalizations of histogram equalization. *IEEE Trans Image Process.* 2000;9(5):889–96.
 29. Yadav G, Maheshwari S, Agarwal A. Contrast limited adaptive histogram equalization based enhancement for real-time video system. In: *2014 International Conference on Advances in Computing, Communications and Informatics (ICACCI),* 2014. 2392–7.
 30. Bedi S, Khandelwal R. Various image enhancement techniques: a critical review. *Int J Adv Res Computer Commun Eng.* 2013;2(3):1605–9.
 31. Cheng H, Shi X, Tan CL. A simple and effective histogram equalization approach to image enhancement. *IEEE Trans Image Process.* 2004;14(2):158–70.
 32. Huang SC, Cheng FC, Chiu YS. Efficient contrast enhancement using adaptive gamma correction with weighting distribution. *IEEE Trans Image Process.* 2013;22(3):1032–41.
 33. Chaki N. *Exploring image Binarization techniques.* Berlin: Springer; 2014.
 34. Park GH, Cho HH, Choi MR. A contrast enhancement method using dynamic range separate histogram equalization. *IEEE Trans Consum Electron.* 2008;54(4):2067–74.
 35. Goh TY, Basah SN, Xue X. Fog removal from video sequences using contrast limited adaptive histogram equalization. *Computational Intelligence and Software Engineering 2009. CISE 2009. International Conference.* 2009. pp. 1–4.
 36. Jin Y, Laura M, Laine Fayad A. Contrast enhancement by multi scale adaptive histogram equalization. *Proc SPIE.* 2001;4478:206–13.
 37. Ntirogiannis K, Gatos B, Pratikakis I. Performance evaluation methodology for historical document image binarization. *IEEE Trans Image Process.* 2013;22(2):595–609.
 38. Pratikakis I, Zagoris K, Kaddas P, Gatos B. ICFHR 2018 competition on handwritten document image binarization (H-DIBCO 2018). In: *2018 16th International Conference on Frontiers in Handwriting Recognition (ICFHR).* 2018. pp. 489–93.
 39. Zimmerman JB, Pizer SM, Staab EV, Perry JR, McCartney W, Brenton BC. An evaluation of the effectiveness of adaptive histogram equalization for contrast enhancement. *IEEE Trans Med Imaging.* 1988;7(4):304–12.
 40. Rahman NA, Haroon F. Adaptive Gaussian and double thresholding for contour detection and character recognition of two-dimensional area using computer vision. *Eng Proc.* 2023;32(1):23.
 41. Durrand F, Dorsey J. Fast bilateral filtering for the display of high dynamic range images. In: *Proceedings of SIGGRAPH 2002.* 2002. pp. 844–7.
 42. Durand F, Dorsey J. Fast bilateral filtering for the display of high dynamic-range images. *ACM Trans Graph.* 2002;21(3):257–66.
 43. Paris S, Durand F. A fast approximation of the bilateral filter using a signal processing approach. *Int J Comput Vision.* 2009;81:24–52.
 44. Tcheslavski GV. *Morphological image processing: grayscale morphology.* ELEN 4304/5365 DIP, Spring 2010. 2010.
 45. Déforges O, Normand N, Babel M. Fast recursive grayscale morphology operators: from the algorithm to the pipeline architecture. *J Real-Time Image Proc.* 2013;8(2):143–52.
 46. Srisha R, Khan A. *Morphological operations for image processing: understanding and its applications.* 2013.
 47. Clienti C, Beucher S, Bilodeau M. A system on chip dedicated to pipeline neighborhood processing for mathematical morphology. In: *IEEE Conference in Signal Processing,* 16th European, 1–5. 2008.
 48. Torres-Huitzil C. Fast hardware architecture for grey level image morphology with flat structuring elements. *IET Image Proc.* 2013;8(2):112–21.
 49. Heijmans H. *Morphological image operators.* In: Marton L, editor. *Advances in electronics and electron physics.* Cambridge: Academic Press; 1994.
 50. Haralick R, Sternberg S, Zhuang X. Image analysis using mathematical morphology. *IEEE Trans Pattern Anal Mach Intell.* 1987;9(4):532–50.
 51. Bartovský J, Dokládál P, Dokládálová E, Georgiev V. Parallel implementation of sequential morphological filters. *J Real-Time Image Proc.* 2014;9(2):315–27.
 52. Gil J, Kimmel R. Efficient dilation, erosion, opening, and closing algorithms. *IEEE Trans Pattern Anal Mach Intell.* 2002;24(12):1606–17.
 53. Gil J, Kimmel R. Efficient dilation, erosion, opening and closing algorithms in mathematical morphology and its applications to image and signal processing. In: Goutsias J, Vincent L, Bloomberg D, editors. *Proceedings of Shape in Picture '92, NATO Workshop, Driebergen, The Netherlands, September 1992.* Springer-Verlag. 2000. pp. 301–10.

54. Vincent L. Morphological area openings and closings for greyscale images. In: Proceedings of Shape in Picture '92, NATO Workshop, Driebergen, The Netherlands. 1992.
55. Dokládal P, Dokladalova E. Computationally efficient, one-pass algorithm for morphological filters. *J Vis Commun Image Represent*. 2011;22(5):411–20.
56. Gonzalez CI, Melin P, Castro JR, Castillo O. Edge detection methods and filters used on digital image processing. In: Gonzalez CI, Melin P, Castro JR, Castillo O, editors. *Edge detection methods based on generalized type-2 fuzzy logic*. Berlin: Springer; 2017.
57. Mutneja V. Methods of image edge detection: a review. *J Electr Electron Syst*. 2015;4:2332–796.
58. Gentsos C, Sotiropoulou C, Nikolaidis S, Vassiliadis N. Realtime canny edge detection parallel implementation for FPGAs. In: Proceedings of the International Conference on Electronics, Circuits and Systems. 2010. pp. 499–502.
59. Chao L, Jiliu Z, Kun H. Adaptive edge-detection method based on canny algorithm. *Comput Eng Design*. 2010;31(18):4036–9.
60. Vincent OR. A descriptive algorithm for Sobel image edge detection. In: Proceedings of Informing Science & IT Education Conference (InSITE). 2009.
61. Jayanthi N, Sharma T, Sharma V, Tyagi S, Indu S. Classification of ancient inscription images on the basis of material of the inscriptions. In: 2021 3rd International Conference on Signal Processing and Communication (ICPSC), 2021. pp. 422–7.
62. Vijayalakshmi R, Gnanasekar JM. A review on character recognition and information retrieval from ancient inscriptions. In: 2022 8th International Conference on Smart Structures and Systems (ICSSS), 2022. pp. 1–7.
63. Dhivya S, Beulah JR. Ancient Tamil character recognition from stone inscriptions—a theoretical analysis. In: 2022 2nd Asian Conference on Innovation in Technology (ASIANCON), 2022. pp. 1–8.
64. Rajithkumar BK, Mohana HS, Uday J, Bhavana MB, Anusha LS. Read and recognition of old Kannada stone inscriptions characters using a novel algorithm. In: 2015 International Conference on Control, Instrumentation, Communication and Computational Technologies (ICCICCT), 2015. pp. 284–88.
65. RajaKumar S, Subbiah Bharathi V. Eighth century Tamil consonants recognition from stone inscriptions. *Int Conf Recent Trends Inform Technol*. 2012;2012:40–3.
66. Rajnish P, Kamath KP, Kumar B, Nishanth M, Preethi P. Improving the quality and readability of ancient Brahmi stone inscriptions. In: 2023 2nd International Conference for Innovation in Technology (INOCON), 2023. pp. 1–8.
67. Rogowska J. Overview and fundamentals of medical image segmentation. In: Bankman IN, editor. *Handbook of medical image processing and analysis*. 2nd ed. Cambridge: Academic Press; 2009. p. 73–90.
68. Priya RD, Karthikeyan S, Indra J, Kirubashankar S, Abraham A, Gabralla LA, Nandhagopal SM. Self-adaptive hybridized lion optimization algorithm with transfer learning for ancient Tamil character recognition in stone inscriptions. *IEEE Access*. 2023. <https://doi.org/10.1109/ACCESS.2023.3268545>.
69. Sukanthi S, Murugan SS, Hanis S. Binarization of stone inscription images by modified bi-level entropy thresholding. *Fluct Noise Lett*. 2021;20(06):2150054.
70. Zhang H, Qi Y, Xue X, Nan Y. Ancient stone inscription image denoising and inpainting methods based on deep neural networks. *Discret Dyn Nat Soc*. 2021;2021:1–11.

Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.