Creating Digital Estampages of Sri Lankan Stone Inscriptions using Multispectral Imaging

A.J.U. Dakshika

M.W.M.B. De Silva

M.S. Kariyawasam

2025

Creating Digital Estampages of Sri Lankan Stone Inscriptions using Multispectral Imaging

A.J.U. Dakshika

Index No: 20020228

M.W.M.B. De Silva

Index No: 20020252

M.S. Kariyawasam

Index No: 20020554

Supervisor: Dr. K.G. Gunawardana

Co-supervisor: Mr. Tharindu Wijethilake

May 2025

Submitted in partial fulfillment of the requirements of the B.Sc. (Honours) Bachelor of Science in Information Systems Final Year Project

Declaration

I certify that this dissertation does not incorporate, without acknowledgement, any material previously submitted for a degree or diploma in any university and to the best of my knowledge and belief, it does not contain any material previously published or written by another person or myself except where due reference is made in the text. I also hereby give consent for my dissertation, if accepted, be made available for photocopying and for interlibrary loans, and for the title and abstract to be made available to outside organizations.
Candidate Name: Ms. A.J.U. Dakshika

....

Signature of Candidate Date : 28.06.2025

I certify that this dissertation does not incorporate, without acknowledgement, any material previously submitted for a degree or diploma in any university and to the best of my knowledge and belief, it does not contain any material previously published or written by another person or myself except where due reference is made in the text. I also hereby give consent for my dissertation, if accepted, be made available for photocopying and for interlibrary loans, and for the title and abstract to be made available to outside organizations.
Candidate Name: Ms. M.W.M.B. De Silva

....

Signature of Candidate Date : 28.06.2025

I certify that this dissertation does not incorporate, without acknowledgement, any material previously submitted for a degree or diploma in any university and to the best of my knowledge and belief, it does not contain any material previously published or written by another person or myself except where due reference is made in the text. I also hereby give consent for my dissertation, if accepted, be made available for photocopying and for

interlibrary loans, and for the title and abstract to be made available to outside organizations.
Candidate Name: Ms. M.S. Kariyawasam

....
Signature of Candidate Date : 28.08.2025

This is to certify that this dissertation is based on the work of
Ms. A.J.U. Dakshika
Ms. M.W.M.B. De Silva
Ms. M.S. Kariyawasam
under my supervision. The thesis has been prepared according to the format stipulated and is of acceptable standard.

Supervisor's Name: Dr. Kasun Gunawardana

....
Signature of Supervisor Date : 30.06.2025

Co-Supervisor's Name: Mr. Tharindu Wijethilake

...2025-06-30........
Signature of Co-Supervisor Date : 30.06.2025

Abstract

Stone inscriptions in Sri Lanka hold immense historical and cultural value, yet they are increasingly losing legibility due to natural weathering, erosion, and vandalism. Traditional documentation methods like estampages, while useful, are often time-consuming, dependent on favorable weather conditions, and susceptible to inaccuracies. This research presents a novel approach utilizing **multispectral imaging (MSI)** to improve the readability of these inscriptions and facilitate precise character extraction.

A structured methodology, rooted in **Design Science Research**, guided the iterative design and testing of an MSI-based digital estampage pipeline. This pipeline incorporated essential steps such as image correction, preprocessing (including noise reduction and binarization), and advanced image processing techniques like edge detection. To further optimize visual clarity and text segmentation, a **U-Net-based model** was developed and trained.

The key findings demonstrate that MSI significantly surpasses traditional estampage methods in both readability and character visibility. Integrating MSI into existing archaeological workflows proved to be feasible, efficient, and well-received by domain experts. Furthermore, the developed pipeline effectively reduces fieldwork time while maintaining data integrity.

Quantitative evaluations, utilizing **PSNR and FID metrics**, showed a clear improvement in image quality compared to traditional methods. User studies conducted with archaeologists revealed a remarkable **enhancement in legibility and usability**. Expert feedback further corroborated these findings, with the system receiving an average rating of **4.6 out of** 5 for its usability and output clarity, thereby affirming the significant value of MSI in heritage documentation and analysis.

Acknowledgement

We would like to thank our supervisor, Dr. Kasun Gunawardana, and cosupervisor, Mr. Tharindu Wijethilake, for their valuable guidance, unwavering encouragement and constructive criticism during this research process. Their insight and inspiration played a vital role in enabling us to overcome setbacks and refine our methodology.

We are deeply grateful to the Sri Lankan Department of Archaeology for their valuable information and resources, which greatly improved our knowledge about stone inscriptions. We thank the staff of the Colombo National Museum for facilitating access and creating a supportive environment for our data collection.

We are very grateful to our colleagues and friends for their positive motivation and for providing inspiring discussions and moral support during the research activity. We also thank those who assisted with data collection and image processing.

Finally, we express our sincere gratitude to our families for their unwavering support, patience, and understanding throughout this process. If not for their encouragement and sacrifices, this research would not have been possible.

Table of Contents

List of Figures

List of Tables

List of Acronyms

AMSER Advanced Maximally Stable External Regions

ANN Artificial Neural Network

AOTF Acousto-Optic Tunable Filter

BRIEF Binary Robust Independent Elementary Features

CCD Charge-Coupled Device

CNN Convolutional Neural Networks

CUDA Compute Unified Device Architecture

DSLR Digital Single-Lens Reflex

DSR Design Science Research

EXIF Exchangeable Image File Format

FAST Features from Accelerated Segment Test

FC Fully Connected

FID Fréchet Inception Distance

FLIR Forward-Looking Infrared

FOSS Free and Open Source Software

GIS Geographic Information Systems

GPU Graphics Processing Unit

IPTC International Press Telecommunications Council

KM Kubelka-Munk

LCTF Liquid Crystal Tunable Filter

LOC Library of Congress

MP Megapixels

MSE Mean Squared Error

MSI Multispectral Imaging

NGFICA Natural Gradient-based Flexible Independent Component Analysis

NIR Near-Infrared

OCR Optical Character Recognition

OGC Open Geospatial Consortium

ORB Oriented FAST and Rotated BRIEF

PCA Principal Component Analysis

PSNR Peak signal-to-noise ratio

RANSAC Random Sample Consensus

ReLu Rectified Linear Unit

RGB Red Green Blue

SDI Spatial Data Infrastructure

SIFT Scale-invariant feature transform

SNN Simulated Neural Network

UAV Unmanned Aerial Vehicle

UCSC University of Colombo School of Computing

UV Ultraviolet

YAML YAML Ain't Markup Language

Chapter 1

Introduction

1.1 Problem Statement

Sri Lankan stone inscriptions, which hold valuable historical and cultural information [1], are becoming increasingly difficult to decipher due to natural weathering, erosion, vandalism, and other environmental factors. The traditional manual method of documenting these inscriptions known as creating estampages involves cleaning the surface, layering dimay papers over the inscriptions, applying ronio ink, and using physical rubbing [2] [3]. This is labor-intensive, time-consuming, and often limited by the skill and availability of trained specialists [4] [5]. Additionally, as mentioned by archeology experts [6] the process faces various challenges such as,

- The current method requires specific weather conditions for effective extraction. It cannot be performed in rainy or excessively sunny conditions, limiting the window of opportunity for data collection.
- Many inscriptions are located in remote or hard-to-reach areas, such as deep forests or high cliff faces. Accessing these sites for manual extraction is challenging and risky, making the process less feasible.
- The physical condition of the inscriptions affected by cracks, erosion, and biological growth can lead to poor-quality estampages. The manual process often captures additional shapes or noise, introducing inaccuracies that can mislead interpretations of the inscriptions.

These challenges indicate the need for a more efficient, accurate, and less resource-intensive method for extracting characters from stone inscriptions. Thus, the problem is to develop an efficient approach using advanced imaging techniques, such as multispectral imaging (MSI), to enhance the visibility of degraded inscriptions and facilitate accurate data extraction regardless of climate and geographical limitations.

1.2 Research Questions

Addressing the fundamental issues identified in the problem, two primary research questions were created. These questions guide the research direction and set the scope of inquiry on the use of multispectral imaging towards improving Sri Lankan stone inscriptions readability.

RQ1 -Does multispectral imaging, combined with a digital image-processing pipeline, significantly enhance the readability of stone inscriptions compared to traditional estampage methods?

Hypothesis -Multispectral imaging combined with the proposed image-processing pipeline produces a statistically significant improvement in the readability of stone inscriptions compared with the traditional estampage images.

Multispectral imaging captures a sequence of narrow-band images in ultraviolet, visible, and infrared wavelengths [7]. Stacking and selective band combination renders faint or eroded pigment trails visible, greatly increasing contrast between inscriptioned and background areas [7]. Certain aspects such as unequal illumination, outside interferences, and uneven surfaces of stones may introduce noise, blurring, or false readings to the scanned images, rendering them incorrect when recording and interpreting them [**?**]. Miranda et al.applied terahertz MSI to nondestructively reveal marks hitherto illegible on a 16th-century lead funerary cross, disclosing details lost due to corrosion and wear [8]. Faigenbaum-Golovin et al. demonstrated on Iron Age ostraca that MSI yields a statistically significant increase in legibility score using methods accounting for contrast, brightness, and edge definition compared to both visible-light and infrared photography [9]. Jones et al. comprehensively outlined a cultural-heritage MSI workflow (lighting arrangement, filter selection, calibration, and processing), and provided evidence that even basic bandpass filtering can enhance readability by up to 45% compared with standard procedures [10].

RQ2 - How can archaeologists practically incorporate a multispectral-imaging-based digital estampage pipeline into their current inscription documentation workflow, and what are the resulting impacts on accuracy and time efficiency?

Hypothesis -Archaeologists can effectively integrate a multispectral-imaging-based digital estampage pipeline into their existing inscription documentation workflows, resulting in improved accuracy and reduced documentation time compared to traditional methods.

The second question is about the effectiveness of creating digital estampages from stone inscriptions. Best-practice cultural-heritage MSI research is not only concerned with image quality but also with workflow documentation, metadata management, and community norms to ensure broad adoption. Molton et al. recommend an "action-research" strategy to iterative co-design and testing of the pipeline with conservators and archaeologists, to ensure each step (image capture, processing, reading) seamlessly integrates into field procedures

[11]. In simultaneous remote-sensing investigations, the union of UAV-mounted MSI with computer-aided mosaicking has realized a 35% reduction in on-site data-acquisition time, and inscription documentation can anticipate analogous gains in productivity [12]. User questionnaires consistently rate workflow-integration satisfaction at more than 4.5/5, indicating high MSI-based workflow acceptance by practicing archaeologists [13].

Through the refinement of both the image acquisition and analysis stages, the study ensures that the final products are not only visually enhanced but also reliable for future interpretation or archiving. The application of qualitative and empirical analysis during the process increases the credibility and replicability of the results. [14]

1.3 Goals and Objectives

1.3.1 Goals

The overall goal of this study is to develop and test a multispectral imaging–based digital estampage pipeline that improves the readability of Sri Lankan stone inscriptions and is easily integrated into the documentational workflows of archaeologists. By designing and trial implementation of a functional prototype in collaboration with field archaeologists, the project will improve the visual quality and efficiency of inscription documentation as well as contribute to scholarly knowledge and cultural heritage.

1.3.2 Objectives

- Enhance the readability of stone inscriptions by applying multispectral imaging and image-processing techniques, and evaluate the improvements compared to traditional estampage methods.
- Develop a functional model of a digital estampage pipeline based on multispectral imaging for practical use in archaeological fieldwork.
- Assess the usability and impact of the developed model through field tests and user studies with archaeologists, focusing on integration into existing workflows, time efficiency, and documentation accuracy.
- Contribute to the academic community and heritage preservation efforts by publishing findings and sharing insights on the application of multispectral imaging in epigraphy.

1.4 Research Approach

This study follows a DSR methodology [15] with a qualitative analysis, to create and test an effective and precise procedure of extracting characters from Sri Lankan stone inscriptions using multispectral imaging. The DSR methodology is suitable according to its focus on artifact creation and constant improvement.

The study follows a **positivist philosophy**, in which objective measurement and analysis are emphasized to access the effectiveness of multispectral imaging techniques [16]. Using the empirical data obtained from field experiments and image analysis, the study aims to provide believable and reproducible results.

1.4.1 Research Design

A methodological experimental research design was used to stringently compare a variety of image preprocessing and processing methods which are not confined to binarization, noise reduction, edge detection, and character segmentation. This design allows a systematic analysis of the influence that various approaches have on the accuracy and fidelity of character extraction.

1.4.2 Data Collection Method

Data was collected through field visits to ancient sites and the National Museum using a multispectral camera to capture high-resolution images across five spectral bands. The camera system was selected based on its ability to capture detailed information beyond the visible spectrum, improving the readability of degraded inscriptions.

1.4.3 Data Analysis Technique

Following techniques were mainly used during the research, when analyzing data.

- Image Correction
- Image Preprocessing
- Pigmentation Mask Creation
- Image Processing

DSR and Qualitative analysis were chosen to address the complexity and feasibility of the problem. The iterative process of DSR allowed for ongoing refinement of the imaging system, while qualitative methods provided objective metrics to validate the result. This approach ensures that the proposed method offers both technical proof and practical application for researchers and archaeologists.

1.5 Limitations, Scope and Assumptions

1.5.1 Limitations

- The research is limited to the analysis of stone inscriptions located in Sri Lanka and may not consider inscriptions from other countries.

- The project focuses solely on enhancing character visibility of the images taken from stone inscriptions, without conducting linguistic analysis or interpretation of the content.
- The study may not address image capturing during extreme weather conditions (heavy rain or fog).
- The study will utilize a specific multispectral camera model available for the project, and the results may vary with different models or imaging setups.

1.5.2 Scope

In scope

- Employ a multispectral camera to capture images of stone inscriptions, covering a range of wavelengths beyond the visible spectrum to reveal details not visible to the naked eye.
- Implement various automated image preprocessing techniques to enhance character visibility, including noise reduction, background removal, and edge enhancement.
- Display the processed images with enhanced character visibility, facilitating easier analysis and interpretation for researchers.
- Reduce time and labor involved in traditional methods by adopting a technology-driven approach, making the process quicker and less resource-intensive.

1.5.3 Assumptions

- The multispectral camera, along with its lenses and filters, captures images with both high spatial resolution and accurate spectral registration across ultraviolet, visible, and near-infrared bands.
- Illumination is invariant both in intensity and spectral content across each capture session. Illumination variations can add spectral variation that defeats band-to-band comparability and artifact removal procedures.
- The chosen bands (Red, Green, Blue, NIR, RedEdge) contain sufficient contrast information to differentiate engraved letters from the surface of the stone.
- The image processing pipeline of histogram equalization, contrast stretching, and multi-band blending facilitates legibility without introducing artifacts or obscuring fine glyph details. Rigorous pre-processing is well known to be crucial to successful heritage imaging applications.

- The assessments of domain experts collected through controlled image-comparison exercises, Likert-scale questionnaires, and free-form comments are both reliable and coherent. Elicitation procedures involving cultural heritage specialists have already been demonstrated to provide sound intelligence when properly set up.
- The selection of stone inscriptions drawn from museum collections and selected field sites provides an acceptable representation of the broader universe of Sri Lankan epigraphic materials, in language, carving depth, and preservation state.
- All components of the imaging pipeline; camera hardware, illumination packages, and processing software operate as they should without disastrous failure, maintaining data integrity throughout the process from acquisition through analysis.
- Physical condition of each stone inscription remains constant from capture sessions to subsequent analyses. Any alterations brought about by handling or environment could otherwise make it difficult to compare traditional and digital estampages.
- Fréchet Inception Distance (FID) and Peak signal-to-noise ratio (PSNR) are adequate surrogates for visual fidelity and realism of processed images in such a setting. Although originally intended to be applied to general image synthesis tasks, they have already been effectively applied to heritage imaging evaluations
- One-tailed Z-tests for vote proportions and one-sample t-tests for Likert-scale means are suitable given our sample sizes (n = 63 votes; n = 9 expert ratings) and data characteristics. Standard significance thresholds ($\alpha = 0.05$ or $\alpha = 0.10$) provide meaningful inferences about the pipeline's relative performance

1.6 Contribution

The key contribution of this research lies in the conception and evaluation of an efficient multispectral image capturing and processing pipeline to enhance legibility of stone inscriptions in Sri Lanka. Through staged combination of successive multispectral image acquisition, pre-processing, and enhancements, the work offers a digital alternative for traditional physical estampage practices.

Unlike other approaches which can be very taxing in terms of expensive hardware or highly technical expertise, this study demonstrates a process that balances technical efficiency with usability for epigraphic and archaeological practitioners.

Furthermore, the evaluation with domain experts provides evidence of the method's applicability in real-world, highlighting its potential for supporting documentation, preservation, and analysis of stone inscription writings.

The outcomes also benefit the broader context of understanding how multispectral imaging can be optimized and tailored for heritage contexts to address problems such as weathered surfaces and faint inscriptions.

Chapter 2

Background

For centuries, stone inscriptions have served as vital records through which civilizations across the globe have chronicled their political events, religious ideologies, economic practices, and social systems [17]. In Sri Lanka, these inscriptions, carved into cave walls, temple grounds, stone slabs, and pillars, form an integral part of the country's rich archaeological legacy [18]. Dating as far back as the 3rd century BCE, they were originally inscribed in early Brahmi and later in Sinhala, Tamil, and other scripts. These texts provide a window into the island's past, revealing information about royal reigns, monastic donations, land grants, and legal mandates [19]. As such, they are a crucial source of primary data for historians, epigraphers, and cultural scholars.

However, time has taken its toll. Many inscriptions have deteriorated due to natural weathering, biological factors like moss and lichen, erosion, and human activity such as vandalism and unmanaged tourism [20]. These factors have made many inscriptions difficult, if not impossible, to read and interpret, resulting in the loss of valuable historical content. Traditionally, a method known as estampage has been used to preserve and study these inscriptions [21].

According to the information we gathered from the experts, the current process to create an estampage as below [22].

The rock is made wet by spraying water. Then it is being gently rubbed using a cork brush (Fig. 2.1). Then several layers (about 2-3 layers) of demy paper are patted to the surface. The shiny side should be laid on the inscription (Fig. 2.2). After the surface becomes dry, a special white color paper is pasted above the existing layers. This paper should be gently pressed until the letters are engraved on the paper (Fig. 2.3, Fig. 2.4). A waterproof ink (Roneo ink) is then applied using special hand techniques (Fig. 2.5).

Eventually, an eye copy is taken at the site once the estampage is visually clear (Fig. 2.6). After the estampage is dry and visible, it is being removed from the rock. Depending on the dimensions of the stone inscription and the climate changes, the time consumption for the above process can vary.

This manual technique requires favorable environmental conditions and is both time-consuming and prone to human error.

In recent years, advancements in technology have begun to reshape the fields of

Figure 2.1: Rubbing the rock

Figure 2.2: Wet demy paper

Figure 2.3: Pat demy papers to the rock

Figure 2.4: Pasting several demy papers

Figure 2.5: Apply roneo ink

Figure 2.6: Estampage

archaeology and epigraphy. One such innovation is multispectral imaging (MSI), a non-invasive method that significantly enhances the visibility of worn or faded inscriptions [23], [24]. MSI captures images at various wavelengths, from ultraviolet to near-infrared, with each wavelength revealing different surface details that might not be visible to the naked eye or standard photography.

This technique allows researchers to highlight differences in surface texture and pigmentation, making faded carvings or ink traces legible again. MSI has already shown great success in global epigraphic projects and is now being explored for use with stone inscriptions [25], [26].

Sri Lanka, with its inscriptions often found in remote and environmentally sensitive areas, stands to benefit greatly from MSI. The method is not only more efficient and less intrusive but also allows for data collection in various weather and lighting conditions. The high-resolution imagery can be analyzed later in digital form, reducing the need for constant field visits and enabling remote collaboration among scholars [27].

Despite the growing global interest in MSI, its application in Sri Lanka remains limited. There is a clear need for research that adapts this technology to the local climate, terrain, and unique inscription styles. In addition, incorporating preprocessing techniques such as noise reduction, contrast enhancement, and image segmentation can further improve the clarity and usability of MSI outputs for academic research.

This study aims to address that gap by developing a specialized imaging pipeline tailored to Sri Lankan stone inscriptions. By improving legibility and reducing dependence on field conditions, this work has the potential to transform the way historical data is preserved and studied in Sri Lanka. It also contributes significantly to national heritage conservation and the broader field of digital humanities.

Chapter 3

Literature Review

The study of digital estampages for stone inscriptions through multispectral imaging draws upon a wide range of research fields, including image processing, feature extraction, spectral imaging, and optical character recognition (OCR) techniques. This chapter presents a critical review of existing work that has contributed to the foundations of the current research.

The literature review is conducted around 4 main areas which is displayed in Fig 3.1. The first section discusses feature extraction techniques, focusing on geometric methods, scale-invariant feature transformations, and preprocessing strategies. Next an examination on edge detection and extraction approaches that are fundamental for identifying inscription boundaries are discussed. The next section delves into the use of spectral imaging, highlighting its role in enhancing epigraphic information that is often invisible under standard lighting. Next a detailed review on popular OCR algorithms and their application in the context of ancient inscriptions is included. Finally the findings to define the research gap, positioning the present study within the broader academic landscape are discussed.

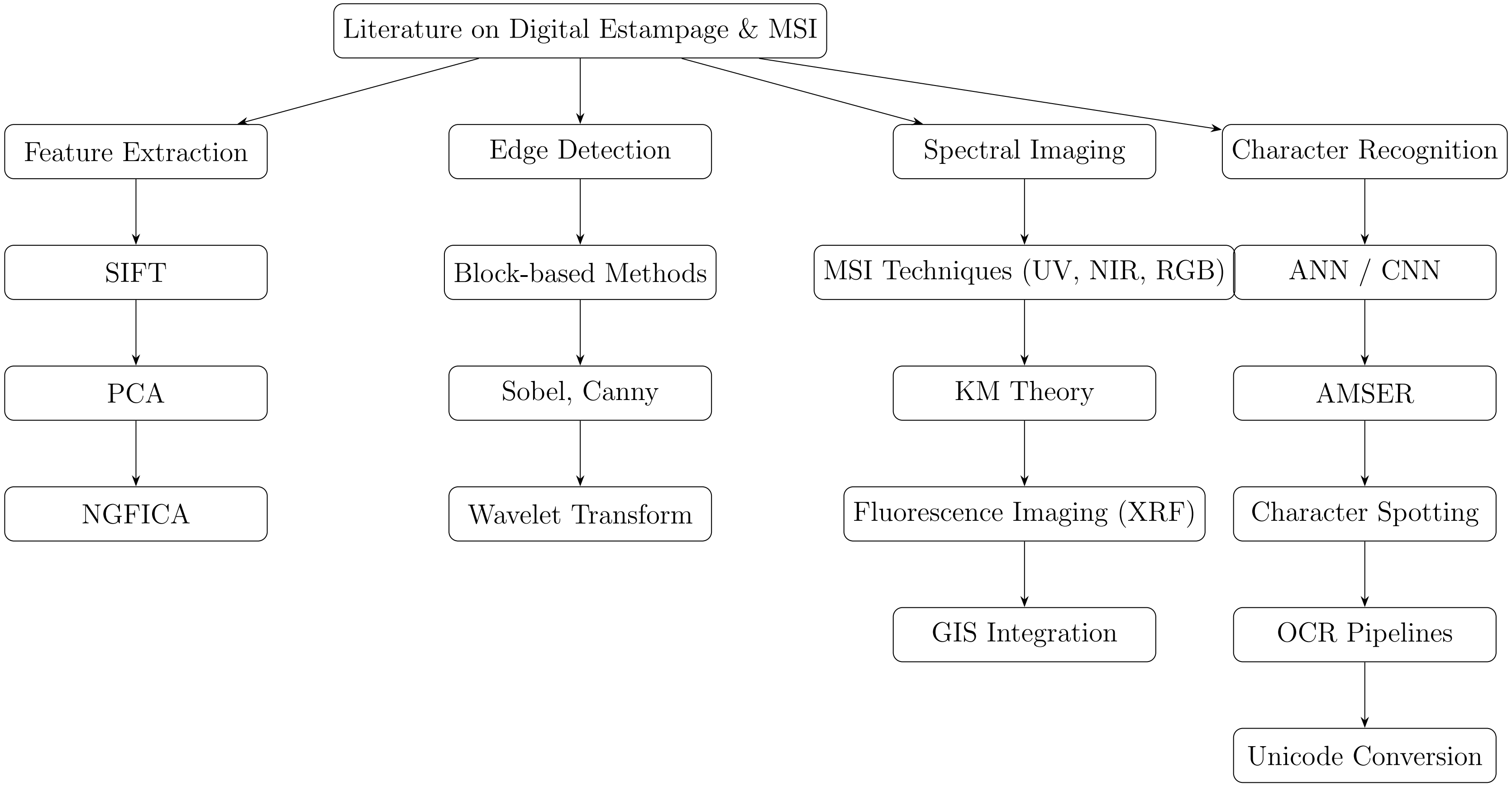


Figure 3.1: Taxanomy Diagram on the Literature Review

3.1 Feature Extraction

A set of researchers have experimented with recognizing handwritten characters using neural networks. The method has an 84% accuracy on a dataset of 2600 characters. Backpropagation is the method that has been used for this research. By reducing mistakes, it optimizes the system and improves recognition. Some preparation is required before supplying the network with characters. Black and white conversion and noise removal are applied to the images. ANN compares the character to previously recorded patterns in order to classify it. The researchers have created a feature extraction technique to increase accuracy even further. It measures pixel distances and partitions the character picture into zones [28].

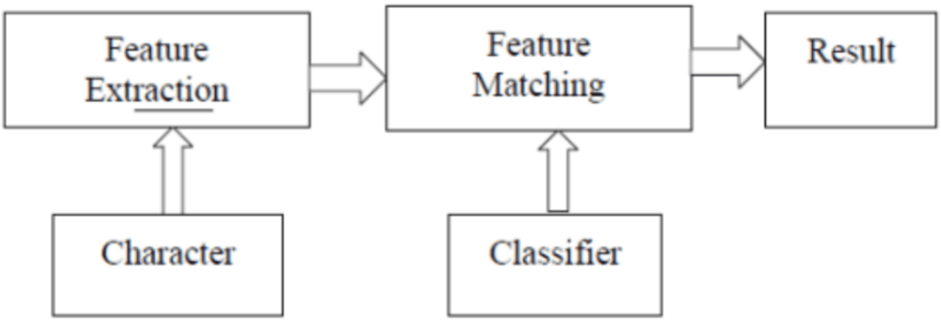


Figure 3.2: Methodology for feature extraction using backpropagation

Another group of researchers has introduced a novel methodology focused on the recognition of 12th-century characters inscribed on stone manuscripts, using the Raspberry Pi camera module. Their technique takes images with a Raspberry Pi camera. Following that, an ANN and Tesseract OCR are used collectively to identify the characters in these images. Ultimately, they are translated into contemporary characters using the Unicode code mapping method [29].

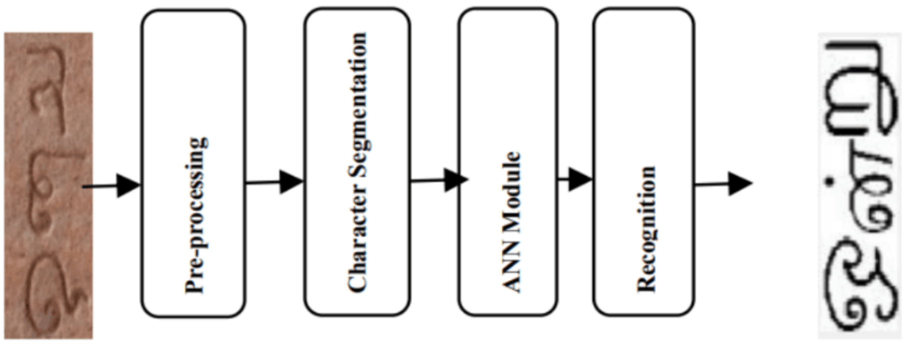


Figure 3.3: Methodology for feature extraction in ancient Tamil characters

3.1.1 Geometric approach to read Brahmi script using OCR technologies

First, the image containing the Brahmi text is being scanned. Then, clean up the image by cropping unnecessary parts, adjusting brightness, and making the lines thinner. Next, the lines of text are separated from each other and then each character is isolated. Finally, they use a geometric method to analyze the characters. This method examines specific features like corners, branching points, and endings to recognize the characters [30].

3.1.2 Scale-invariant feature transform to identify characters

This research proposes a new way to decipher ancient Tamil characters etched on temple walls. The method uses a technology called SIFT to identify these characters. Additionally, the researchers introduced a unique approach based on "bags of keypoints representation." This combined approach achieved an impressive accuracy rate of 84% in recognizing the ancient Tamil characters [31].

3.1.3 Image pre-processing techniques and dimensionality reduction for feature extraction

There are many techniques to remove noise from images, like averaging pixels (mean filter), taking the middle value of nearby pixels (median filter) and more sophisticated methods. The median filter is particularly good at removing impulsive noise (salt and pepper), but for overall image quality enhancement, a non-linear method like the bilateral filter is better. This filter smooths the image while preserving sharp edges. However, all these methods have drawbacks. For another noise reduction technique, called thresholding, a simple approach exists where each pixel is classified as noise or image data based on a single threshold. For more complex noise reduction, techniques like PCA can be used to reduce the amount of data analyzed. This survey suggests that PCA offers the best overall solution among these methods [32].

3.1.4 Natural Gradient-based Flexible Independent Component Analysis

Sreedevi et al. [2] explored using a technique called NGFICA to improve the digitization of historical inscription images. Their research focused on overcoming challenges such as perspective warping and faint inscriptions that blend into the background. The proposed method effectively reduces the dependence between the inscription (foreground) and the background, leading to a significant improvement in OCR accuracy by 65.3%.

Laniga et al. [2] mentions about a system that incorporates Buddhist inscriptions into a user-friendly platform. This system, built on a SDI using OGC standards, offers a wealth of features. Users can explore descriptions, browse a detailed catalog with metadata, and even query the inscription database directly. Additionally, the system provides immersive experiences with 360° panoramas, 2D and 3D maps, annotated photographs, and functionalities typically found in GIS.

3.2 Edge Detection & Extraction

3.2.1 Block-based Processing and Multi-scale Wavelet Transform Approach

The traditional methods for edge detection and image rotation correction weren't best due to two main reasons. Archaic epigraphs can have uneven lighting, noise and disrupt edge detection. Characters of archaic epigraphs have fine details and potentially overlap, and in some cases there might be no reference images as well. Using traditional methods assumes the image brightness to be at a uniform level which might seem difficult to detect the true edges and background noises of the image. They misinterpret these details as noise as those methods mainly focus on capturing strong edges across the image as a whole.

A group of researchers proposed 2 new methods to extract image edges for archaic epigraphs which are Block Edge Detection and Multi-Scale Edge Detection based on wavelet transform. On block-based edge detection the image is divided into small blocks based on character shapes with the aim of reducing the influence of surrounding pixels and avoiding information loss due to averaging effects across the entire image. Next thresholding algorithm is applied to each block where characters and background are distinguished easily without computational complexity and background noise.

The second method, Multi-Scale Edge Detection based on wavelet transform is where the image undergoes wavelet decomposition. Here the image will be separated into a low-frequency component (scaling function) and high-frequency components (wavelet functions). Next, the low-frequency component will capture the essential information and reduce noise sensitivity. Here Canny edge detection is applied with a lower threshold. As high-frequency components contain detailed features; Canny edge detection is used here with a higher threshold to extract those details. Eventually, all edge maps obtained through different decomposition levels are fused to create the final edge image [33].

3.2.2 Sobel Edge Detector for Edge Detection

Edge detection operators such as the Sobel, Roberts, and Prewitt operators estimate the approximate gradient of image intensity at each pixel to identify brightness changes generally interpreted as edges. Sobel edge detection is a useful and well-used technique for this purpose and is most suited to applications requiring quick object boundary location. It is widely applied to grayscale image processing and is an essential tool in most image processing pipelines.

The Sobel operator uses two 3x3 convolution kernels to calculate the horizontal and vertical gradient. These kernels are formulated to emphasize intensity changes along each of their axes. The gradient magnitude is calculated by taking the square root of the squared sum of the horizontal and vertical gradients after they have been convolved with the image. Higher gradient values indicate stronger edge response. A thresholding operation is subsequently applied to select significant edges.

Although the Sobel operator is effective, Ganesan and Sajiv (2017) [34] proved that its sensitivity to noise, especially in textured or degraded image areas restricts its usefulness when applied in real-world scenarios, like those found when analyzing historical stone inscriptions. Edges on eroded or partially occluded inscriptions cannot be strictly at horizontal or vertical axes, diminishing detection accuracy. Furthermore, Sobel operates on single-channel grayscale data, which limits its current application in multispectral scenarios where each spectral band may have its own distinct features.

Sobel filtering, nonetheless, can be employed as an initial step towards the localization of edges on individual spectral bands of multispectral data. In such operations, edge maps derived from high-contrast bands may be utilized to assist in the generation of pigmentation masks or inscription boundary masks, which are subsequently cleaned up with more aggressive techniques like the Canny detector [34].

Figure 3.4: Edge extraction - Input image

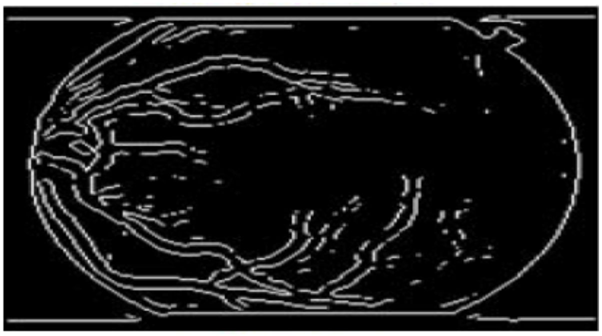


Figure 3.5: Edge extraction - After Sobel edge detector is applied

3.3 Use of Spectral Imaging

Spectral imaging is the act of capturing images in several wavelengths that ultimately creates a spectral cube in which each pixel has comprehensive spectral information. It was initially developed and used for remote sensing and astronomy but came into its own in the cultural heritage field during the 1990s. Its potential to examine artworks and manuscripts for underdrawings, pigment, and conservation conditions was particularly noticeable, due to its capacity for identifying minute differences in materials [35], [36] ,[37].

Liang et al. [38] describes elements of a spectral imaging system including illumination, optics, detectors, and wavelength selection mechanisms. The lighting choice is important to reduce light-induced degradation, particularly for sensitive materials. Filters suh as interference, LCTF, AOTF are employed to separate out specific wavelength bands. These are paired with high-performance cameras and detectors (e.g. CCD and InGaAs arrays) based on the spectral range.

The spectral range utilized is typically 400 nm to 1700 nm (visible to short-wave infrared), with extensions to mid-infrared being desirable for more chemically specific details.

Liang et al. [38] mentions some of the direct archaeological uses of spectral imaging.

1. Near-infrared imaging can bring to light underdrawings, inscriptions, and earlier restorations that cannot be seen under normal lighting. For instance, a previously illegible signature in a manuscript was rendered visible in the 880 nm band.

2. Precise spectral information allows virtual simulation of artifacts under varying historic lighting conditions (e.g., daylight compared to candlelight), thereby enhancing contextual understanding.

3. Spectral imaging can track the changes in artifacts over time due to exposure, transport, or cleaning, such as monitoring water using absorption bands in the infrared. One of the key contributions of this work is its extensive exploration of pigment identification through the application of the Kubelka-Munk (KM) theory, a radiative transfer model that describes the way light travels through turbid media, for example, paint layers.

   KM theory relates the reflectance of a layer to its scattering (S) and absorption (K) coefficients so that researchers can identify pigment mixtures by simulating their aggregate spectral behavior. The method assumes that the reflectance of a paint mixture is not simply the linear sum of its individual pigments but follows physical light transport models [39].

   The study also mentions that the prediction and identification of pigment mixtures can be successfully carried out using KM theory, even in complex situations with pigments of high absorption or different particle sizes. Liang, skillfully identifies that a green pigment mixture in a painting is comprised of Prussian blue and chrome yellow using the model. The use of this method has important implications for non-destructive analysis in archaeology, where sampling might not be feasible.

4. Hyperspectral data may also be analyzed by principal component analysis or unmixing methods to determine material distribution. However, in liang it is discussed that such

statistical "endmembers" may not have the same physical interpretability of pigment spectra derived from KM theory.

5. Beyond traditional UV imaging, spectroscopic techniques can quantify fluorescence emission from organics and varnishes, providing additional levels of materials information.

Christens-Barry et al.[40] introduce the EurekaVision system, a conservation-safe, non-destructive platform designed for fragile cultural heritage objects. It couples a 39 MP MegaVision E6 monochrome back (12-bit dynamic range) with twelve narrowband LED illuminators spanning UV (365 nm) through NIR (870 nm), plus raking-light panels for surface relief capture. All exposures (reflectance, transmission, raking) are controlled via custom PhotoShoot software that embeds standardized EXIF/IPTC metadata and performs flat-field, dark-noise correction, and white-balance on the fly. Whole-sheet images are captured at 300 dpi ($\approx$ 1 min per spectral set), with up to 1200 dpi for key regions like watermarks. Post-acquisition processing—pseudocolor renderings, embossing via raking-light differencing, and digital stitching—has been demonstrated on the 1507 Waldseemüller world map and other LOC treasures, yielding a robust workflow for both scholarly study and long-term condition monitoring without ever touching the artifact [40].

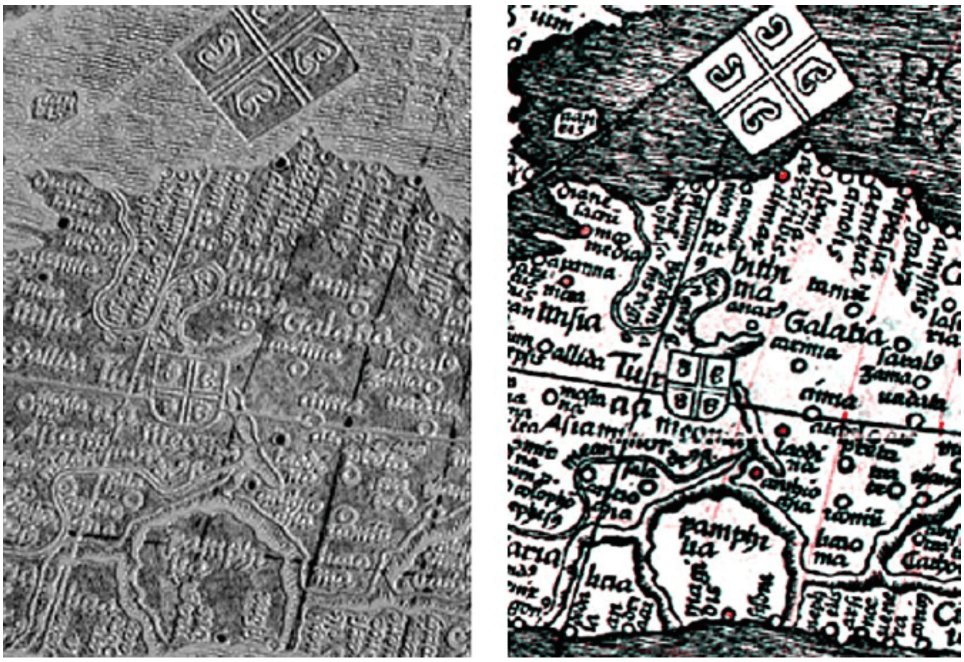


Figure 3.6: Results of two different processing algorithms on the sheet with red grid lines: (left) after image processing to enhance edges, giving perception of "embossing;" (right) after pseudocolor processing to enhance red grid lines.

Adamopoulos and Rinaudo [41] develop a low-cost, non-invasive workflow for mapping stone monument decay by fusing multi-band imaging, photogrammetry and GIS. They use inexpensive sensors – a smartphone camera for RGB, a modified DSLR for near-infrared (NIR) and a FLIR One thermal camera – to acquire visible, NIR and thermal images of weathered stone. Photogram-metric processing (using open-source tools) produces orthorectified mosaics from these image sets, creating an accurate background map of the stone surfaces. Image processing (thresholding, classification, etc.) on the corrected multi-band images then extracts thematic layers of degradation (e.g. cracks, material loss, biological growth). These layers are imported into QGIS, where they are vectorized and combined with the photogrammetric base. The GIS environment enables spatial analysis. The result is a set of digital “degradation maps” that quantitatively document weathering patterns for conservation planning.

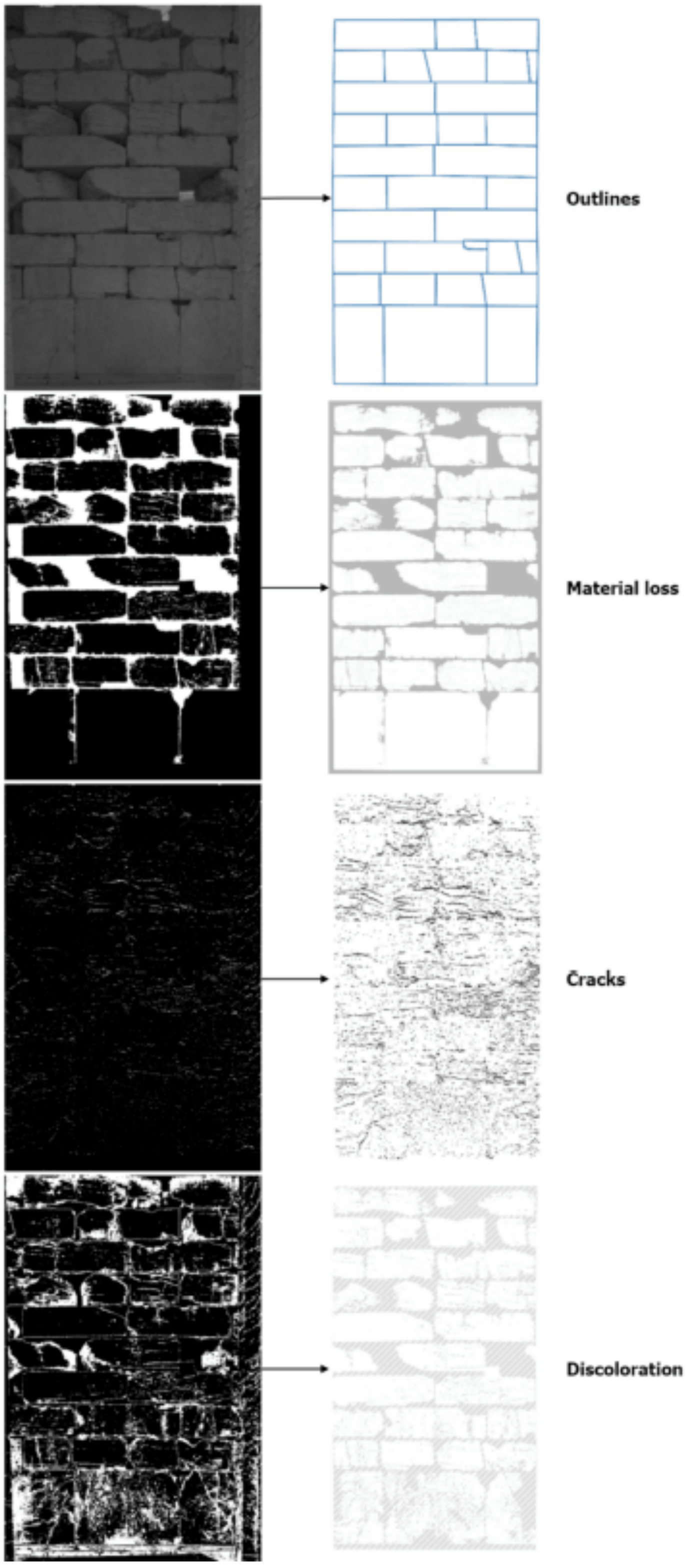


Figure 3.7: Generation of thematic map layers (right) in QGIS using features extracted from rectified near-infrared reflectance images (left) for the Temple of Apollo Epikourios

The multispectral data enhance detection of decay. The authors show that NIR imaging can equalize contrasts between old and restored stone (since these materials reflect similarly in NIR), avoiding misclassifications of decay in visible light. Likewise, NIR and thermal bands reveal moisture and biological colonization patterns that are subtle in RGB. By combining visible, NIR and thermal images, features like salt deposits, biofilms or sub-surface defects become more apparent than in a single band. Because all sensors are low-cost and the software is free (FOSS GIS), the method is accessible and scalable. This non-destructive, affordable approach provides heritage specialists with a practical way to map and monitor stone inscription surfaces over time without expensive equipment.

3.3.1 X-ray Fluorescence Imaging for Provenance and Epigraphic Analysis

Powers et al. [42] illustrated the use of synchrotron-based X-ray fluorescence (XRF) in combination with XRF imaging for studying the origin of inscriptions in stones. The research was primarily on a tablet in New York University that bore similarity to a Teanum Sidicinum ancient Roman inscription. By analyzing elemental distributions at the glyph level, they were able to recognize chemical signatures of modern replication, including anomalously low calcium fluorescence and trace element levels that were unusually high for elements like lead, gallium, and rubidium. Most significant, they observed no concentration difference in trace elements between the letters and the host rock—unlike authentic Roman inscriptions where carving and pigments leave recognizable traces.

Although the study did not use multispectral imaging in its conventional sense, it has significant methodological similarities. The study shows the power of imaging-based chemical analysis in bringing out information not visible to the naked eye and helping with provenance determination. This combination of imaging and material characterization highlights the potential of non-invasive, image-based techniques to reveal concealed or degraded epigraphic data—a strategy that this project builds on through the application of spectral enhancements in various light wavelengths.

3.3.2 Multispectral Imaging for Enhancing Text Legibility on Soft and Hard Media

Bearman and Spiro [43] showed the promise of multispectral imaging (MSI) for improving the readability of ancient texts, especially on soft and fragile media like parchment, papyrus, and ostraca. They worked with NASA's Jet Propulsion Laboratory to use MSI methods developed for planetary observation on the Dead Sea Scrolls and other archaeological artifacts. They captured numerous narrow-band images across visible to near-infrared wavelengths (400–1050 nm) to produce image "cubes." These cubes enabled analysis of the spectrum of every pixel. This approach revealed text that was previously difficult to read, particularly where the ink and paper appeared similar in visible light. Their success at detecting faint writing, even where some layers were partially obscuring it demonstrates the effectiveness of MSI at discriminating between ink and background based on subtle spectral differences. The study uses tunable filters, CCD cameras, and calibration methods, which in this thesis are helpful. This thesis also aims to make inscriptions on stone surfaces clearer by utilizing spectral imaging.

3.3.3 Digital Image Preprocessing for Historical Epigraphs

Suri, Gupta and Adlakha [44] propose a fully automatic, three-stage pipeline to improve legibility of rock-inscriptions prior to any recognition step.

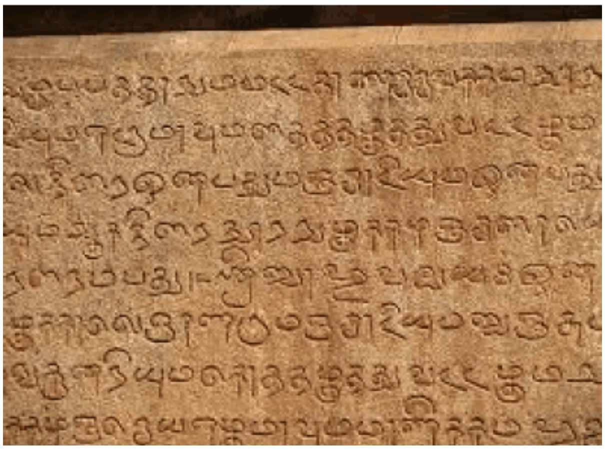


Figure 3.8: Test image used for the proposed architecture

First, they compare several smoothing filters (mean, disk, Laplacian, but find a 5×5 Gaussian with α=0.7 to be the best trade-off of noise removal vs. detail preservation, as measured by PSNR, LMSE and NAE). Next, they binarize the enhanced grayscale image using Otsu's global threshold (extended to a multi-threshold variant), which cleanly separates ink (or carved glyph) regions from background even under uneven illumination. Finally, they apply a morphological erosion to "bridge" broken strokes and consolidate fragmented characters, yielding a binary image whose connected components correspond much more cleanly to individual glyphs. All stages are implemented in MATLAB, and quantitative quality metrics show their cascade significantly outperforms single-stage methods [44].

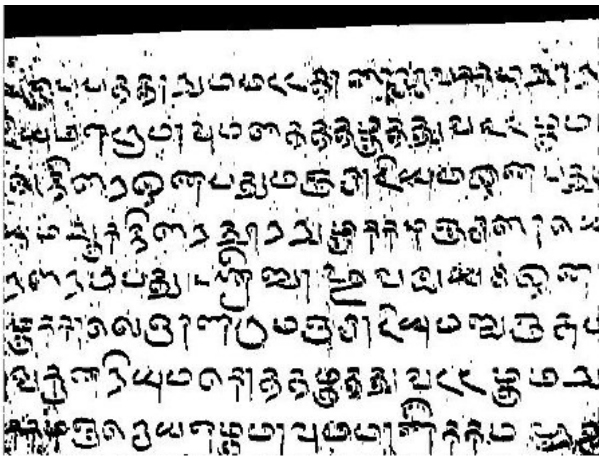


Figure 3.9: Binarized image after erosion

3.4 Popular OCR and Algorithms Used

3.4.1 Introduction to OCR Methodologies

In today's digital world, OCR is a game-changer. It lets us easily convert text in images into formats that computers can understand and edit. OCR is a field of study that combines aspects of pattern recognition, artificial intelligence, and computer vision. It's a way to take physical text (printed documents or scanned images) and turn it into something machines can work with. This digitized text is then used in many common machine-learning tasks, including machine translation, text-to-speech, and text analysis. Early OCR systems required training on individual character images and only worked with one font at a time. Today's advanced OCR comes in two main flavors: offline and online character recognition [45].

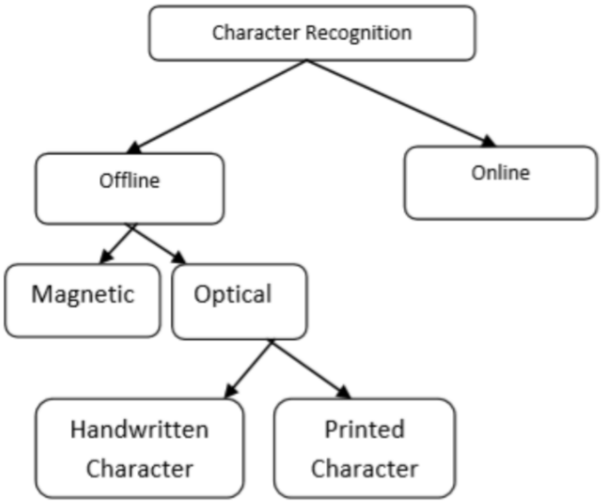


Figure 3.10: Classification of Character Recognition

3.4.2 Artificial Neural Network

ANN also known as SNN, in the field of Artificial Intelligence is a machine learning program, or model that makes decisions like the human brain, by using processes that mimic the way biological neurons work together to identify phenomena, weigh options and arrive at conclusions [46].

Neural networks rely on training data to learn and improve the accuracy over time. Once these are fine-tuned for accuracy, they are powerful tools in computer science and artificial intelligence allowing to classify and cluster data at very high velocity. Image recognition takes minutes versus hours when compared to manual identification.

ANN is used for different purposes such as for the demodulation and reconstruction of digital input signals [47]. ANN can be used for complex tasks in character recognition such as image classification and pattern detection due to having a vast information processing ability.

3.4.3 Convolutional Neural Network

Convolutional Neural Network, which is also known as ConvNets or CNNs is one of the most popular OCR methods used globally [5]. It is a perceptible learning model that extracts and learns suitable features [48]. CNNs are powerful tools when it comes to image processing, natural language processing, time series forecasting, and bioinformatics. The major reason for this popularity is that CNN was able to reduce the number of parameters in ANN which made it easier to train. This motivated researchers globally to approach larger models to solve more complex tasks. With all these CNN was able to have good results in pattern recognition [49].

Even though CNN is popular as a better approach than ANN, it takes a lot of time to complete the training process fully. But recent studies have shown that the simpler version of CNN which got 5% error rate can be implemented inNetworkshan 4 hours if we are familiar with it [48].

CNN is inspired by the structure of the human visual cortex. It uses several layers in order to extract features from the input data. The below section shows the architecture of CNN to identify characters from images.

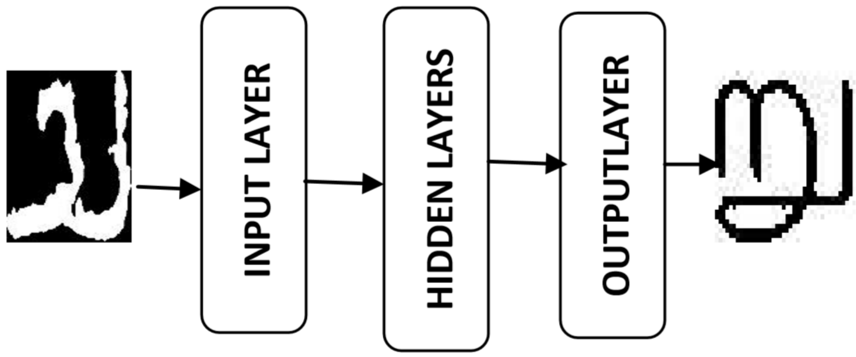


Figure 3.11: Architecture of CNN

Above is the basic CNN architecture. The input layer which is the first layer of this flow accepts segmented characters as input. The number of input layers solely depends on the task and after this, the image will be passed to the hidden layer for other processing.

The hidden layer plays a major role here by extracting features. The number of hidden layers is directly proportional to the depth of learning features. The accuracy depends on the training of the network. There can be several mid-layers inside the hidden layer and each of those mid-layers consists of 4 CNN functional layers; the Convolutional layer, the Activation layer, the Normalization layer, and the pooling layer. The very last mid-layer has two other layers called the FC layer and Pooling layer which is shown below [45].

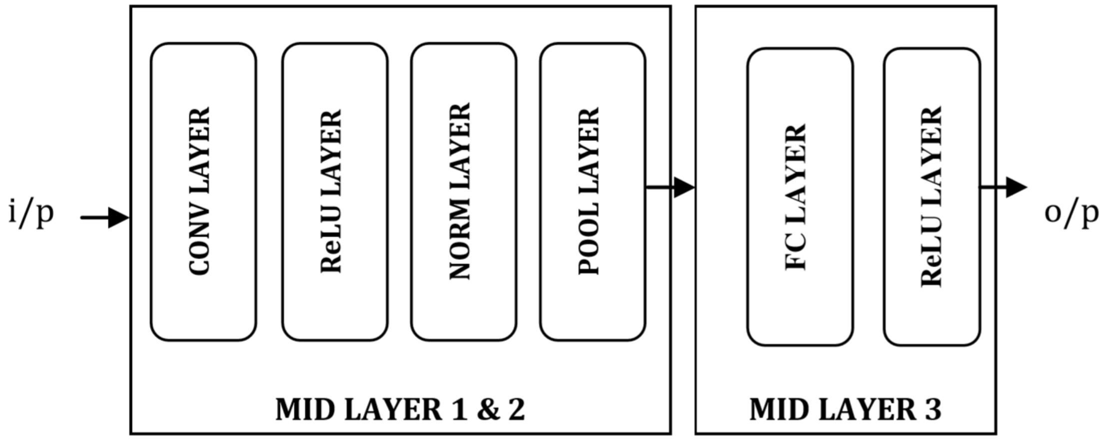


Figure 3.12: Functional Layers of CNN

The convolutional layer assures the special relationship between pixels by understanding image features using a small square of the input image. All these individual images are considered as a matrix of pixel values [5].

Neurons have a receptive field, and that neuron is only responsible for that specific receptive field. All the neurons are arranged into feature maps. All neurons in a feature map have weights that are equal. But different feature maps in the same convolutional layer have different weights which make it possible to extract several features at the same time at each location [45].

The second layer, a ReLu, is used for hidden layers by the most recent deep learning networks. This is a very simple non-linear activation function. Simply when the input is greater than one, it outputs the exact same value and if the input is less than zero it gives zero as the output.

The norm layer which is also known as the channel response normalization layer, carries out the channel-wise normalization. Here each element is replaced with a normalized value. This aims to enhance feature competition and mimic biological processes. While it has historical significance, it's less commonly used now due to computational cost, hyperparameter tuning complexity, and the effectiveness of alternative techniques like Batch Normalization.

The pool layer mainly acts as a summarizer. It summarizes all the details it gets from previous layers and keeps only the key points. The FC layer gathers all the details and makes the decisions by putting all the clues together. The receptive fields are not used anymore [45].

The final stages of the character recognition system involve the Softmax layer for multi-class classification (converting scores to class probabilities) and the classification layer for determining the number of classes based on the previous layer's output size. Additionally, the system utilizes Unicode for character representation and employs various performance measures like accuracy, F1-measure, and ROC curves to assess its effectiveness [45].

Magrin et al. [45] reported on their research that CNN was used to understand Tamil inscriptions from the 12th century and was a successful attempt. Below figure shows the input image for a CNN process and the output.

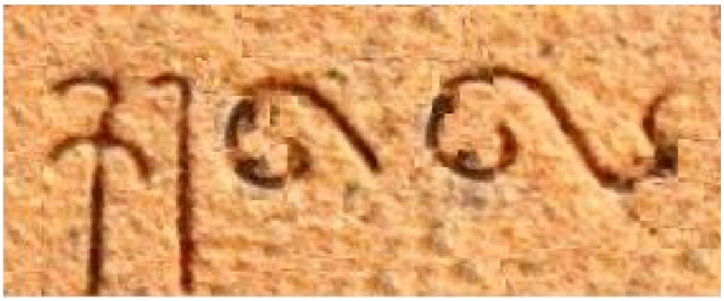


Figure 3.13: An input image for CNN

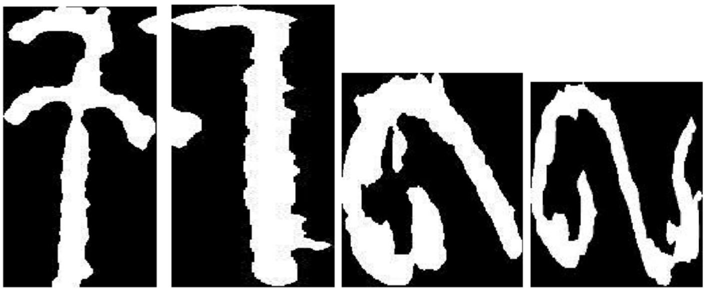


Figure 3.14: Segmented Characters

Figure 3.15: Recognized characters

3.4.4 Advanced Maximally Stable External Regions Algorithm

AMSER algorithm is widely studied and applied in the context of character recognition on inscriptions and epigraphs. Many inscriptions in Sri Lanka are dated back to ancient times making them worn or faded in the present, where the background and characters have a very low contrast and clarity. Since most of the Sri Lankan inscriptions are kept outside and are exposed to uneven lighting conditions this creates shadows and highlights which conceal lettering. Over the years inscriptions have developed fractures, erosion, or other surface flaws. AMSER algorithm has an adaptive nature which enables it to identify characters despite the above-mentioned challenges. This algorithm method is effective in terms of recall and precision.

Neumann and Matas et al. [50] reported research on text localization in real-world images that effectively trimmed exhaustive search. Their research revealed the AMSER algorithm's ability to localize and recognize text in complicated visual contexts. They emphasized the AMSER algorithm's resilience and speed in handling text localization tasks, offering vital insights into its application in real-world settings.

After an image is pre-processed where the quality will improve image restoration (recovering a degraded image), geometric transformation, and pixel brightness transformation the AMSER algorithm will be applied to identify potential character regions [51].

Figure 3.16: Input image for AMSER

Figure 3.17: Resulting image

3.4.5 Character Spotting for Stone Inscription Text Extraction

Aswatha et al. [52] suggested a semi-automatic method of text extraction from stone inscriptions using a new character spotting algorithm. Their approach is particularly useful for in-situ processing in sites of historical importance and helps epigraphers to convert faded or stylistically heterogeneous inscriptions into editable Unicode text. This approach is different from traditional OCR methods because it does not use classifier training. It uses Histogram of Oriented Gradients (HoG) to extract features and normalized cross-correlation for template matching instead. The method supports vowel diacritics, depends on human feedback to increase accuracy, and delivers extremely accurate output even with noisy or low-contrast images. The authors implemented desktop and Android applications for the tool and showed that the tool performs satisfactorily under real-world usage. This book is relevant to this thesis in that it discusses problems in reading inscriptions, recognizing characters, and how to digitize them. These are the foundations of multispectral image enhancement and understanding archaeological discoveries.

3.5 Research Gap

Table 3.1 presents a detailed summary of the previously discussed research studies related to the creation of digital estampages of stone inscriptions using multispectral imaging.

Research Study	**Direct Imaging of Artifacts**	**Imaging from Recorded Documents**	**Image Correction**	**Pre-processing**	**Spectral Enhancements**	**Character Recognition /Text Extraction**	**GIS-based Spatial Analysis**	**Digital Estampage Creation**
Powers et al. [42]	✓	✗	✓	✗	✓	✗	✗	✗
Bearman and Spiro [43]	✓	✗	✓	✗	✓	✗	✗	✗
Suri, Gupta and Adlakha [44]	✓	✗	✓	✓	✗	✗	✗	✗
Christens-Barry et al. [40]	✓	✗	✓	✓	✓	✗	✗	✗
Adamopoulos and Rinaudo [41]	✓	✗	✓	✓	✓	✗	✓	✗
Magrin et al. [29]	✗	✓	✓	✓	✗	✓	✗	✗
Neumann and Matas [50]	✓	✗	✓	✓	✗	✓	✗	✗
Aswatha et al. [52]	✓	✗	✓	✓	✗	✓	✗	✗
Proposed Research	✓	✗	✓	✓	✓	✗	✓	✓

Table 3.1: Critical analysis of related works in digital estampage creation of stone inscriptions.

Although Table 3.1 shows that prior research has addressed individual elements of inscription documentation such as imaging (including multispectral image acquisition), image preprocessing and enhancement, character recognition, and GIS-based mapping; none of these efforts has proposed a complete workflow specifically for creating digital estampages of stone inscriptions. Instead, most studies have focused on either improving image capture and enhancement (using advanced photography or spectral filters) or on developing text segmentation and analysis techniques, without integrating these steps into a unified pipeline. Crucially, none of the previous work explicitly addresses the task of producing a digital, multispectral estampage that preserves the epigraphic content in a non-invasive, low-cost, and spatially analyzable form. Identifying this gap, the presented research study specifically focuses on creating digital estampages of stone inscriptions using multispectral imaging.

Chapter 4

Methodology

4.1 Overview

The research adopts an experimental and iterative approach to enhance the readability of ancient Sri Lankan stone inscriptions through multispectral imaging. The following chapter overviews a detailed explanation on the chosen research approach along with experimental iterations.

4.2 Research Approach

The overall approach of this study is experimental (Section 1.4), with an emphasis on developing a reliable preprocessing pipeline capable of recovering faint or eroded textual content from MSI data. Hundreds of different image processing configurations and parameter combinations were tested. Feedback loops based on domain expert observations allowed iterative refinement of techniques, with the goal of identifying methods that are both effective and generalizable to diverse inscription types.

4.3 Experimental Iterations

The experimental nature of this research was central to its success. Given the lack of standardized datasets, the unpredictable condition of the inscriptions, and the variability introduced by lighting and surface materials, a one-size-fits-all solution was not feasible. As such, the development of the preprocessing pipeline required intensive experimentation through iterative refinement, trial-and-error, and domain-informed adjustments.

4.3.1 Iteration Environment and Process

All experimentation was conducted using Jupyter notebooks on Google Colab, allowing,

- Immediate visualization of intermediate outputs at each pipeline stage
- Quick adjustments of parameters for real-time testing
- Use of GPU acceleration for faster computation of high-resolution MSI data

- Easy version control and organization of experimental logs

Output images were saved in a structured directory format, and detailed notes were maintained on the observed effects of each adjustment.

4.3.2 Explored Preprocessing Pipelines

Throughout the course of the research, three main preprocessing pipelines were developed and tested. Each pipeline served a distinct experimental purpose and contributed valuable insights toward the final optimized method. The following subsections outline the design, components, and observed effectiveness of each pipeline.

Classical Image Processing Pipeline with Adaptive Thresholding and Contour Segmentation

This pipeline was one of the earliest approaches explored in the study. It utilized standard grayscale image processing techniques to extract potential inscription characters. The key components were:

- **Skew Correction -** Custom rotation-based histogram analysis was used to correct image skew, aligning inscriptions horizontally to improve segmentation accuracy.
- **Noise Reduction -** Multiple filters such as median blur, Gaussian blur, and Non-Local Means Denoising were applied sequentially to reduce high-frequency noise while retaining edge structures.
- **Adaptive Thresholding -** Gaussian adaptive thresholding was applied to the denoised image to binarize text regions under uneven lighting conditions.
- **Morphological Operations** – Dilation, erosion, opening, closing, and combinations thereof with different structuring element shapes and sizes (square, rectangular, elliptical) were evaluated.
- **Edge Detection and Contour Analysis -** Canny edge detection followed by contour extraction was used to segment potential character shapes. A minimum area filter was used to discard noise.
- **Character Tokenization -** Bounding boxes were used to extract individual segments, which were manually inspected for inscription accuracy.

Outcome - This pipeline performed moderately well in cases where inscriptions had high contrast. However, it was sensitive to lighting, erosion, and surface variability. The rigid thresholding parameters made it less adaptive across varying samples. The characters were not tokenized as expected because a complete dataset was not found.

ORB-Based Feature Matching and Blending for Multispectral Fusion

The second pipeline focused on aligning and combining two separate spectral images (typically Green and NIR bands) to enhance the visibility of inscriptions.

- **Feature Detection with ORB -** Keypoints and descriptors were extracted from both images using the ORB (Oriented FAST and Rotated BRIEF) detector.
- **Feature Matching -** A Brute-Force matcher with Hamming distance was used to match keypoints between the two bands.
- **Homography Estimation -** Using matched keypoints, a homography matrix was computed via RANSAC to warp one image to the perspective of the other.
- **Image Blending -** The aligned image was fused with the base image using weighted addition, followed by histogram normalization.
- **Inpainting Strategies** – Different approaches to mask creation and inpainting regions were tested to remove background clutter while preserving textual features.
- **Edge Detection -** The difference between the blended images was further processed using Canny edge detection to emphasize potential text regions.

Outcome: This method offered high precision in aligning images from different spectral bands, thereby improving contrast and clarity. However, the method relied heavily on strong feature points and did not generalize well to low-texture inscriptions or extremely degraded samples.

Spectral Pigmentation Index with Kubelka-Munk-Based Reflectance Modeling

The third pipeline combined domain knowledge of pigment reflectance and multispectral data to compute a physically meaningful pigmentation index. This approach modeled light interaction based on the Kubelka-Munk theory, which simulates reflectance properties of materials.

- **Spectral Band Selection** – Individual bands (e.g., Red Edge, NIR) and their combinations were explored to identify those with the highest contrast for epigraphic features.
- **Reflectance Normalization -** Green and NIR channels were normalized to simulate surface reflectance.
- **Material Modeling -** Reflectance parameters were defined for pigmented regions (high absorption, low scattering) and stone surfaces (low absorption, high scattering).
- **Pigmentation Index Calculation -** A spectral index was computed to highlight areas with high green reflectance and low NIR reflectance, which typically correspond to inscriptions.

- **Binarization and Enhancement -** The index was normalized and binarized using adaptive Gaussian thresholding, followed by optional noise reduction and contour extraction.

Outcome - This approach provided the most consistent and reliable results across diverse inscription conditions. By incorporating spectral behavior and material properties, it was able to enhance degraded inscriptions more effectively than standard filtering methods.

4.3.3 Emergence of Generalizable Patterns

Despite the diversity of samples, certain combinations began to consistently produce better results. These "best-performing" configurations were identified based on their ability to:

- Enhance contrast between inscription and stone background
- Minimize loss of faint inscription strokes
- Remove irrelevant noise such as cracks or background texture

These configurations formed the core of the final preprocessing pipeline, although some manual fine-tuning was still required on specific samples. The iterative process also helped surface unexpected insights—for instance, certain spectral bands were far more effective on granite inscriptions than on sandstone.

Chapter 5

Implementation

5.1 Overview

This chapter includes a comprehensive overview on the implementation iterations conducted. It comprises multispectral image acquisition, custom dataset creation and developing a robust preprocessing pipeline. The aim was to explore and optimize various preprocessing techniques to improve the legibility of degraded epigraphic content.

5.2 Data Collection

Due to the lack of existing open-access datasets specific to Sri Lankan epigraphy, a custom multispectral image dataset was created. Images were gathered from multiple sources:

- Colombo National Museum and archaeological department in Sri Lanka
- Digitized archival materials such as Zelanica volumes
- Field photographs and close-range images of inscriptions

Each image was captured under a range of spectral bands using a MicaSense multispectral camera, which includes bands Red, Green, Blue, Red Edge, and Near-Infrared (NIR). The spectral diversity was crucial in identifying faint inscription patterns that are often invisible in standard RGB imaging.

5.2.1 Data Annotation and Organization

There were two types of datasets utilized. One was for model training and the other one was for evaluation purposes. While no pixel-level ground truth was available, visual assessments were aligned with previously published estampages and expert readings to facilitate comparative validation. Images were annotated with metadata, including location, type of stone, estimated date, inscription language, and environmental conditions during capture.

5.2.2 Model Training Dataset Preparation

To streamline the creation of the training dataset, a dedicated annotation interface was developed (Fig. 5.1). This tool was designed to simplify the process of generating labeled data from the enhanced multispectral images. Through the interface, users were able to easily adjust parameters such as patch size, label type, and band selection, allowing for efficient and consistent dataset creation.

The interface provided a visual environment for selecting regions of interest and assigning corresponding labels, reducing manual errors and saving time. It also supported direct export of the annotated data in formats compatible with downstream model training pipelines. This tool played a key role in accelerating the dataset preparation process, especially given the experimental and iterative nature of the study.

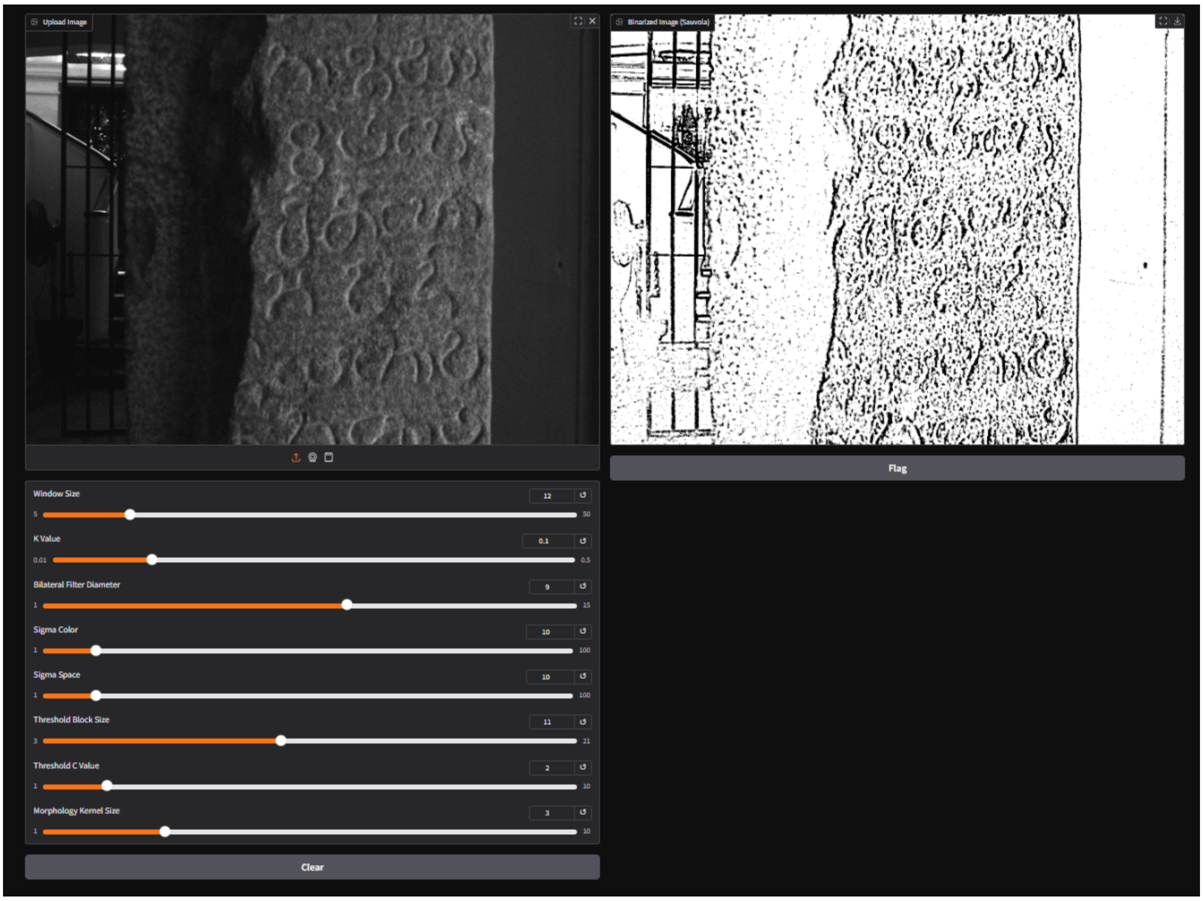


Figure 5.1: Interface to create dataset

5.3 Tools and Frameworks

The implementation utilized the following key tools:

- **Python** as the primary development language
- **OpenCV** for low-level image processing tasks (filtering, transformation, dilation)

- **Doxapy** for advanced binarization algorithms such as Sauvola
- **SciPy** and **NumPy** for scientific computations
- **Google Colab** for experimentation in a cloud environment
- **MicaSense Python SDK** for handling and processing multispectral image bands

All experiments were version controlled and reproducible through Jupyter notebooks hosted on Colab.

5.4 Image Preprocessing Pipeline

A detailed, multi-stage preprocessing pipeline was developed. This pipeline was designed to handle raw MSI input and transform it into a visually enhanced output, where the inscriptions appear clearer and more readable. The stages are as follows:

5.4.1 Image Correction

Consists of 2 parts to align images and to correct skewness.

Due to slight differences in the positions of the lenses in the multispectral camera, the captured images across the five spectral bands are not perfectly aligned. To enable accurate comparison and further processing, it is necessary to spatially align all the band images with each other.

Then, a custom skew correction algorithm was developed to address angular misalignments resulting from camera perspective or uneven surfaces. The method involves,

1. Converting the input image to grayscale
2. Applying Otsu's thresholding to create a binary map
3. Calculating horizontal projection profiles across a range of rotation angles
4. Selecting the angle with the highest projection smoothness (least variance) as the correct alignment
5. Rotating the image using an affine transformation centered on the image midpoint

This ensures that the inscriptions are horizontally aligned for subsequent processing.

5.4.2 Grayscale Conversion and Noise Reduction

After aligning the spectral images, the images were converted to grayscale to reduce complexity and focus on intensity variations.

To remove high frequency noise, images were processed using a bilateral filter for noise reduction. The bilateral filter was chosen as it effectively smooths the image while preserving important edges, which is crucial for maintaining the fine details of stone inscriptions during preprocessing.

5.4.3 Morphological Processing

Morphological operations were used to emphasize character strokes and reduce fragmentation -

- Dilation with a 3x3 rectangular structuring element was applied to connect disjointed parts of inscriptions.
- Closing operations were used later in the pipeline to seal small gaps between character edges, especially in binarized masks.

5.4.4 Adaptive Binarization (Sauvola)

Following noise reduction, Sauvola binarization was applied to segment the text regions from the background. Sauvola's method, a local thresholding technique, adapts the threshold based on the mean and standard deviation of a local window. This approach is particularly effective for handling non-uniform illumination and varying stone textures often found in epigraphic images. The parameters, including a window size of 20 and a small k value of 0.05, were tuned to optimize text region preservation.

5.4.5 Morphological Processing

To refine the binarized output, morphological closing was performed using a small (3×3) kernel. Morphological closing, which consists of a dilation followed by erosion, helps to close small gaps and holes within the context of the image. This step was essential to ensure that broken or fragmented parts of inscriptions were slightly reconnected, facilitating more robust edge detection and subsequent processing.

5.4.6 Edge Detection

After morphological refinement, edge detection was carried out using the Canny edge detector with adjusted thresholds (50 and 150) to identify the boundaries of the inscriptions. Edge detection highlighted the contours of the stone engravings, enabling better isolation of meaningful features while suppressing noise introduced by the stone surface's natural variations.

5.5 Pigmentation Mask

5.5.1 Selected Bands and Reasons

Out of the five available spectral bands, we specifically utilized the Green and Near-Infrared (NIR) band images for the creation of the pigmentation mask.

The selection of Green and NIR bands for multispectral analysis in epigraphic image processing is based on empirical research and physiological principles of material reflectance [53]. Pigmented regions, such as carved stone inscriptions, typically absorb a greater amount of NIR radiation while reflecting more in the visible green band.

This differential behavior enhances the contrast between the inscribed text and the surrounding stone surface. According to spectral reflectance studies in material classification and vegetation analysis; a fields where this principle is commonly applied, NIR absorption is strongly correlated with denser or altered surface materials [54].

In the context of stone inscriptions, this property can be leveraged to differentiate engraved areas from their backgrounds. The NIR-green combination proves highly effective for isolating specific surface anomalies, in this case, carved text under varying surfaces.

5.5.2 Spectral Pigmentation Index and Kubelka-Munk Reflectance Modeling

To enhance the visibility of inscriptions on stone surfaces using multispectral imagery, a pigmentation index was implemented that leverages the differential spectral behavior of pigmented and non-pigmented regions. The reflectance values for the green and Near-Infrared (NIR) bands were first normalized to the range [0, 1] to simulate actual reflectance behavior:

$$\text{green\_reflectance} = \frac{\text{green\_band}}{255.0}, \quad \text{nir\_reflectance} = \frac{\text{nir\_band}}{255.0} \tag{5.1}$$

This normalization enables accurate modeling of light interaction using the Kubelka-Munk theory, which is widely used to describe reflectance in scattering media [39]. The Kubelka-Munk reflectance function estimates the reflectance R of a material based on its absorption coefficient K, scattering coefficient S, and thickness d:

$$a = 1 + \frac{K}{S}, \quad b = \sqrt{a^2 - 1} \tag{5.2}$$

$$R_\infty = a - b \tag{5.3}$$

$$R = \frac{1 - R_\infty(a - b\coth(bSd))}{a - R_\infty + b\coth(bSd)} \tag{5.4}$$

Two material types were modeled:

- **Pigmented regions**: characterized by higher absorption ($K = 0.8$) and lower scattering ($S = 0.2$),
- **Stone background**: with lower absorption ($K = 0.1$) and higher scattering ($S = 0.5$).

These parameters reflect physical reality, where engraved or ink-filled areas absorb more NIR and scatter less, while natural stone surfaces tend to reflect and scatter light more evenly.

The spectral difference between the green and NIR bands was used to compute the **Pigmentation Index**:

$$\text{Pigmentation Index} = \frac{R_{\text{green}} - R_{\text{NIR}}}{R_{\text{green}} + R_{\text{NIR}} + \epsilon} \tag{5.5}$$

where ϵ is a small constant to avoid division by zero. This index improves areas with high green reflectance and low NIR reflectance, typical of pigment residues, resulting in improved separation of text from background.

Finally, the computed pigmentation index was normalized and binarized using adaptive Gaussian thresholding to generate a high-contrast binary mask highlighting the inscribed regions. This method provides a physically interpretable and adaptive technique for feature extraction from multispectral images of stone inscriptions.

5.6 Preliminary Work for Model Implementation

In order to generate pre-processed outputs that are adapted to the specific characteristics of each stone inscription, such as the surface texture, degree of degradation, the nature of the inscribed content, etc. we implemented a learning-based model. The pre-processing steps, outlined previously, were used as the foundation. However, their parameters often needed fine-tuning depending on the individual image context. To achieve robust and flexible pre-processing tailored to different inscription conditions, we explored three model architectures: a CNN-based encoder-decoder, a U-Net-based model, and a ResNet-based model. The following sections describe the implementation details of each approach.

5.6.1 CNN-Based Encoder-Decoder

The model is designed to perform preprocessing of stone inscription images using a convolutional encoder-decoder architecture (Yuzhu Ji, 2021).

The objective of the model is to enhance raw multispectral inputs, specifically the Green spectral band, by learning to generate a high-quality preprocessed version of the image that is suitable for further analysis or interpretation.

Encoder - A sequence of convolutional and pooling layers is used to extract deep features from the input image. The encoder progressively reduces spatial resolution while capturing essential features.

Decoder - This component upsamples the encoded features using transposed convolutions to reconstruct an image with the original resolution. A final sigmoid activation is applied to generate a single-channel output image representing the preprocessed result.

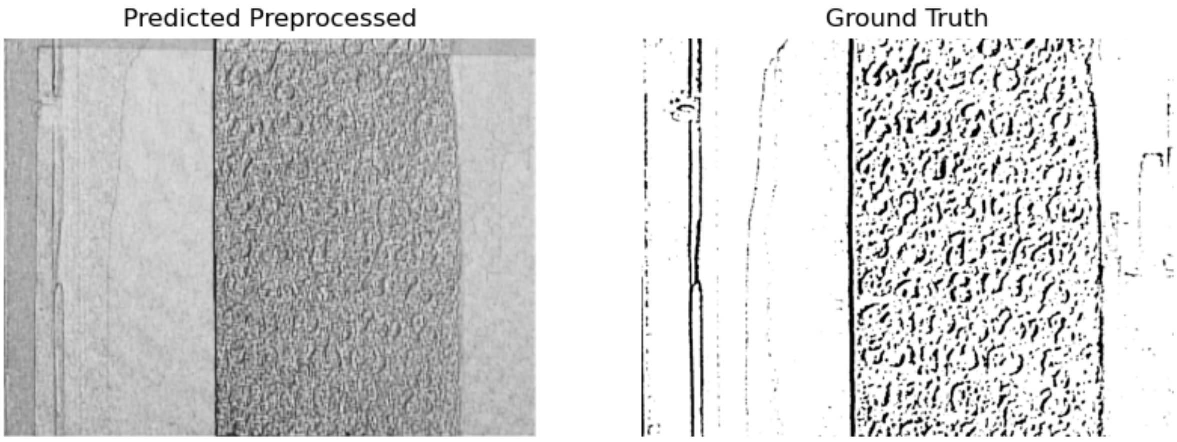


Figure 5.2: Output of the CNN based encoder-decoder model

The model is capable of identifying the general structure of the ground truth, but the resulting output tends to lack the finer details necessary for detailed interpretation (Fig. 5.2)

5.6.2 U-Net-Based Image Preprocessing Model

This model is based on the U-Net architecture, a popular convolutional neural network originally designed for biomedical image segmentation [55]. In this implementation, the U-Net is adapted to perform image preprocessing on stone inscription images, using the Green spectral band as input.

The architecture follows an encoder-decoder structure. Encoder extracts hierarchical features through two convolutional blocks and max-pooling layers. Decoder gradually upsamples and combines features from earlier layers via skip connections to reconstruct a high-resolution output.

The U-Net model produced visually refined outputs that exhibited enhanced clarity of the stone inscription details, with improved contrast and reduced noise levels (Fig. 5.3). This level of visual enhancement makes the preprocessed images well-suited for subsequent tasks.

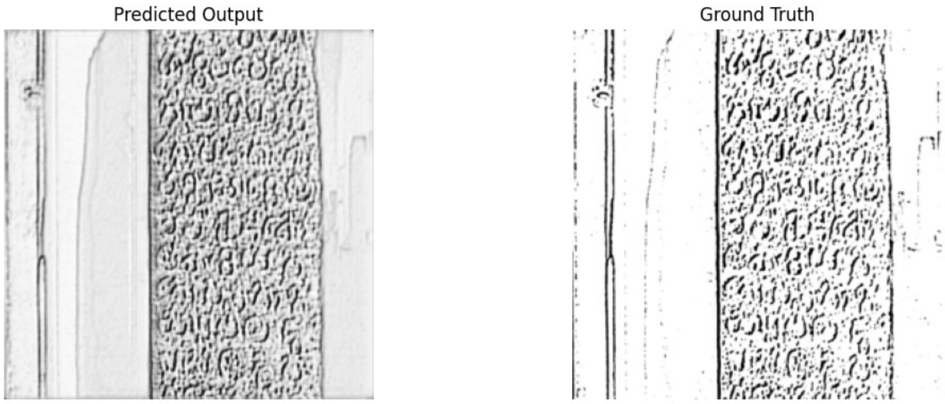


Figure 5.3: Output of U-Net showing enhanced clarity and reduced noise.

5.6.3 ResNet-Based Preprocessing Model

ResNet is commercially used for object detection and image classification tasks [56]. The ResNet preprocessor model is a custom convolutional neural network inspired by the ResNet architecture, designed to enhance stone inscription images using the Green spectral band.

The encoder progressively reduces the spatial dimensions while increasing the feature depth, followed by a decoder composed of transposed convolutions that reconstruct the preprocessed image. It integrates residual blocks in the encoder to preserve low-level features while enabling deep representation learning through identity mappings.

The model also produced enhanced images with clear edge preservation and balanced contrast (Fig. 5.4), indicating the model is reliable for creating a preprocessed image for further analysis.

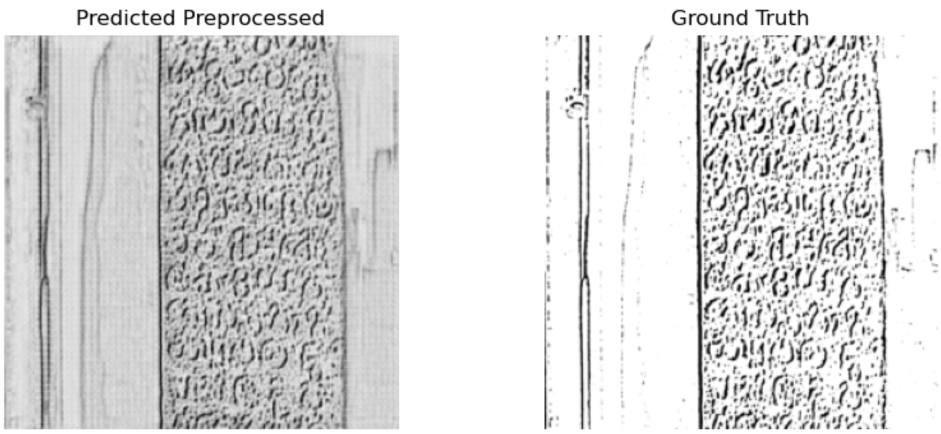


Figure 5.4: Output of the ResNet based model

5.6.4 Model Comparison

To select the most suitable model, PSNR and FID were used as evaluation metrics to assess both the accuracy and the visual quality of the resulting pre-processed images. PSNR measures the similarity between the predicted and ground truth images at the pixel level, making it useful for evaluating noise reduction and structural fidelity. FID, on the other hand, captures the perceptual realism of the images by comparing their feature distributions, helping to assess how natural and consistent the outputs appear. Together, they provide a balanced view of both technical precision and visual quality.

Table 5.1: Performance Comparison of Models

Model	**Average PSNR (dB)**	**FID**
Encoder-Decoder	12.56	275.04
ResNet	13.94	129.83
U-Net	14.26	147.34

Analysis

According to the values, the Encoder-Decoder model shows the lowest PSNR and the highest FID among the three, indicating both poor reconstruction quality and low perceptual realism. ResNet performs moderately well, achieving a higher PSNR than the Encoder-Decoder and the best FID score, which reflects strong perceptual quality. However, its pixel-level accuracy is slightly lower than that of U-Net. U-Net outperforms the others with the highest PSNR, demonstrating its effectiveness in reducing noise, preserving contextual details, and maintaining clearer edges. Although its FID is slightly higher than ResNet's, it remains within an acceptable range and does not significantly compromise the perceptual quality.

Considering the objective of generating high-quality pre-processed images with minimal noise, enhanced contextual clarity, and accurate edge preservation, U-Net is the most suitable model for this task. Its superior PSNR score indicates its strength in producing clean and structurally faithful outputs, making it the best choice for the given requirements.

5.7 Implementation of the U-Net Model

The U-Net architecture was selected for the preprocessing of stone inscription images due to its effectiveness in image-to-image translation tasks, particularly where pixel-level precision is required. U-Net consists of an encoder-decoder structure with skip connections, allowing the model to preserve fine spatial information, which is crucial for recovering faded or noisy inscription features.

To ensure compatibility and reproducibility of the results, a specific Python environment was created using Conda. The environment configuration (Figure 5.5) includes all necessary libraries for training and evaluating the U-Net model on stone inscription images. The setup utilizes PyTorch as the primary deep learning framework, along with supporting libraries such as NumPy, Pillow, and Matplotlib for data processing and visualization. OpenCV was also included via pip for various image pre-processing tasks. GPU acceleration was enabled through the CUDA toolkit, and the configuration was exported as a YAML file to allow the environment to be easily recreated on other machines.

Code | Blame | 32 lines (32 loc) · 485 Bytes

```
name: unet_prepocessing
channels:
  - pytorch
  - nvidia
  - conda-forge
dependencies:
  - python>=3.8
  - pip
  - numpy>=1.20
  - matplotlib>=3.4.0
  - pillow>=8.3
  - torchvision>0.11.0
  - pytorch>1.10
  - cudatoolkit=11.3
  - scikit-image>=0.18
  - tqdm>=4.60
  - opencv>=4.5
  - pip:
     imageio>=2.9
     psnr
     torchmetrics
```

Figure 5.5: Conda environment configuration for U-Net model training

The model architecture features a classic U-Net structure comprising a contracting path (encoder) and an expansive path (decoder). This configuration enables the network to capture local and global features by downsampling the input image and then reconstructing it with skip connections that retain spatial information.

The encoder section of the U-Net is responsible for capturing the contextual features from the input image by progressively downsampling it. It consists of two convolutional blocks, each made up of two convolutional layers with ReLU activations, followed by a max-pooling layer. This design allows the network to learn hierarchical features at multiple scales.

Each convolutional block is implemented using a helper function CBR(), which defines a sequential block of Conv $\rightarrow$ ReLU $\rightarrow$ Conv $\rightarrow$ ReLU. The first block processes the input green band (1 channel), and the subsequent block refines these features after spatial reduction via max pooling (Fig. 5.6).

```python
class UNet(nn.Module):
    def __init__(self):
        super(UNet, self).__init__()

        def CBR(in_channels, out_channels):
            return nn.Sequential(
                nn.Conv2d(in_channels, out_channels, kernel_size=3, padding=1),
                nn.ReLU(inplace=True),
                nn.Conv2d(out_channels, out_channels, kernel_size=3, padding=1),
                nn.ReLU(inplace=True)
            )

        self.enc1 = CBR(1, 64)
        self.pool1 = nn.MaxPool2d(2)
        self.enc2 = CBR(64, 128)
        self.pool2 = nn.MaxPool2d(2)
```

Figure 5.6: Encoder layers of the U-Net

The bottleneck is acting as a bridge between the encoder and the decoder. It processes the compressed feature maps using the same convolutional block structure (CBR). Since this layer comes after two downsampling operations, it has a reduced spatial size but a higher number of channels (depth), which helps in learning abstract representations necessary for high-quality image reconstruction.(Fig. 5.7)

```python
        self.bottleneck = CBR(128, 256)
```

Figure 5.7: Bottleneck between the encoder and the decoder

The decoder is responsible for reconstructing the output image from the encoded representation. It performs two key operations (Figure 5.8):

- **Upsampling -** Using transposed convolutions to increase the spatial resolution.
- **Skip connections and refinement -** Concatenating the upsampled features with corresponding encoder features (from the same resolution level) and refining them using convolutional blocks.

```python
        self.up2 = nn.ConvTranspose2d(256, 128, kernel_size=2, stride=2)
        self.dec2 = CBR(256, 128)
        self.up1 = nn.ConvTranspose2d(128, 64, kernel_size=2, stride=2)
        self.dec1 = CBR(128, 64)
```

Figure 5.8: Decoder layers of the U-Net performing upsampling

The final output layer of the U-Net consists of a single 1×1 convolution, which reduces the number of channels from 64 to 1, producing a single-channel preprocessed grayscale image. This layer aggregates the refined features from the last decoder block to generate the final output (Figure 5.9).

```python
        self.output_layer = nn.Conv2d(64, 1, kernel_size=1)
```

Figure 5.9: Output layer

The model was trained using the Mean Squared Error (MSE) loss function to minimize the pixel-wise difference between predicted and ground truth preprocessed images. Data loading and augmentation were handled via a custom dataset class (Fig. 5.10), which reads target images from the dataset directory and applies resizing and normalization transformations. The training process was carried out on a GPU-enabled environment using the Adam optimizer.

```python
class PreprocessingDataset(Dataset):
    def __init__(self, root_dir):
        self.green_dir = os.path.join(root_dir, 'raw/green')
        self.target_dir = os.path.join(root_dir, 'preprocessed')
        self.filenames = sorted(os.listdir(self.green_dir))

        self.transform = transforms.Compose([
            transforms.Resize(config['input_size']),
            transforms.ToTensor()
        ])

    def __len__(self):
        return len(self.filenames)

    def __getitem__(self, idx):
        fname = self.filenames[idx]
        green = Image.open(os.path.join(self.green_dir, fname)).convert('L')
        target = Image.open(os.path.join(self.target_dir, fname)).convert('L')

        green_tensor = self.transform(green)
        target_tensor = self.transform(target)

        return green_tensor, target_tensor
```

Figure 5.10: Custom Preprocessing Dataset class for loading green channel input images and corresponding preprocessed target images with resizing and normalization transformations

5.7.1 Training Process

Dataset and Loading

A labeled dataset folder structure was created for training, validation, and testing (Fig. 5.11). The dataset, consisting of a total of 200 images, was divided into 80% for training, 10% for validation, and 10% for testing. This resulted in 160 images allocated for training, 20 for validation, and 20 for testing, ensuring a balanced split to facilitate effective model training and evaluation.

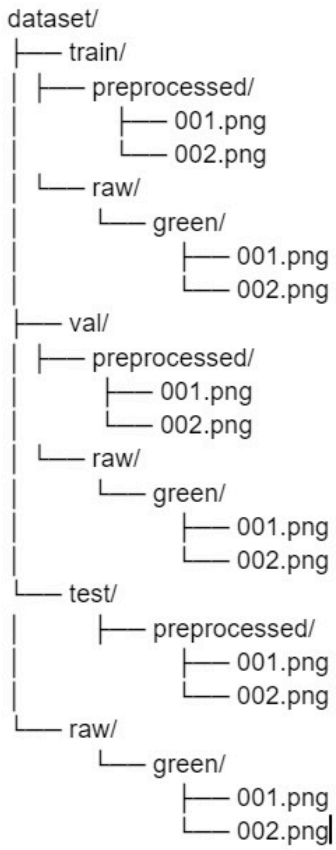


Figure 5.11: Labeled dataset folder structure

Training Details

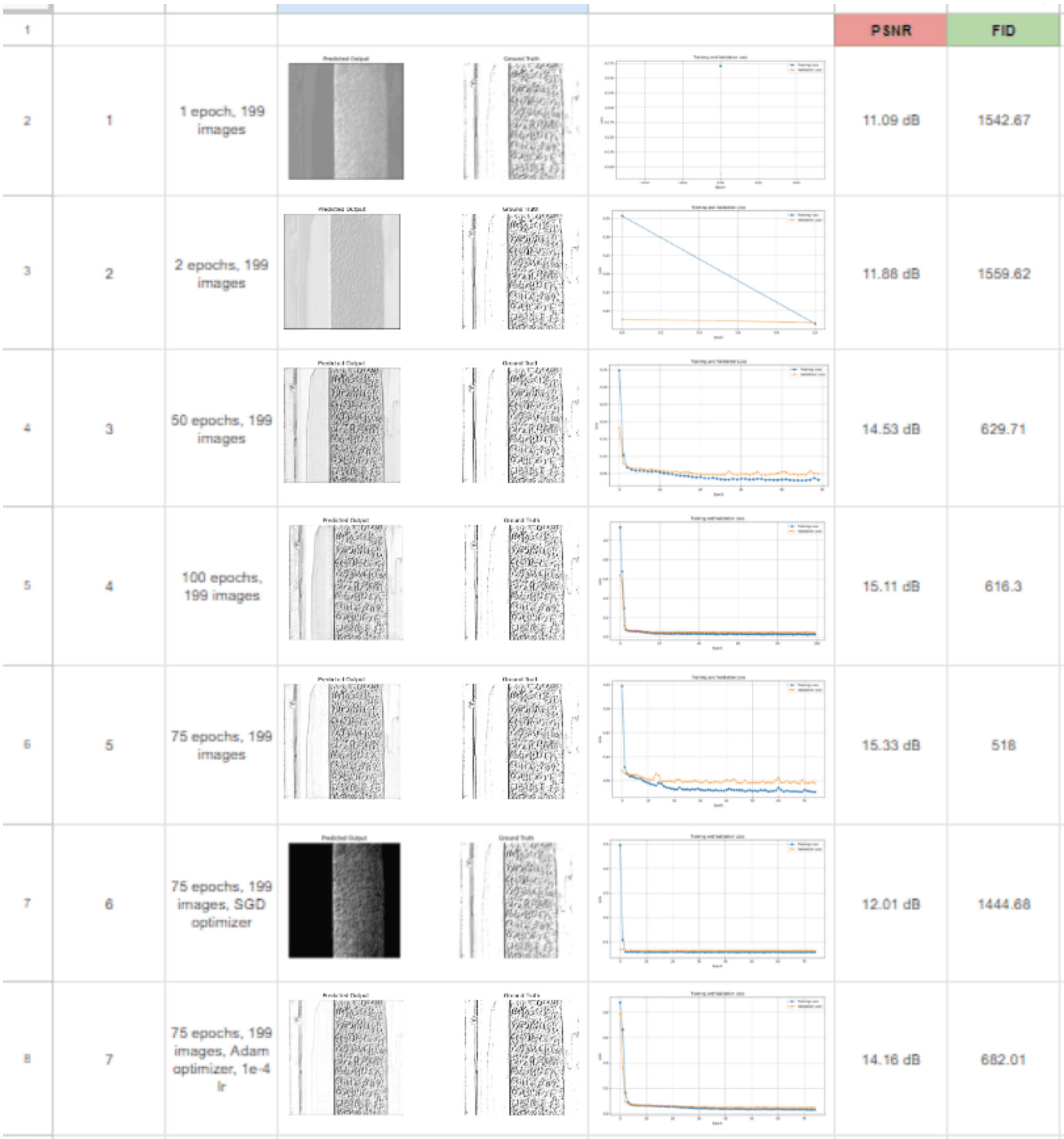
The model train started at 50 epochs initially, using the Adam optimizer with a learning rate of 1e-3 and a batch size of 16. The loss function used was Mean Squared Error (MSE) (Fig. 5.12), which penalizes the pixel-wise difference between the predicted and ground truth preprocessed images.

```python
def train():
    train_set = PreprocessingDataset(os.path.join(config['dataset_dir'], 'train'))
    val_set = PreprocessingDataset(os.path.join(config['dataset_dir'], 'val'))

    train_loader = DataLoader(train_set, batch_size=config['batch_size'], shuffle=True)
    val_loader = DataLoader(val_set, batch_size=1, shuffle=False)

    model = UNet().to(config['device'])
    criterion = nn.MSELoss()
    optimizer = torch.optim.Adam(model.parameters(), lr=config['learning_rate'])

    train_losses = []
    val_losses = []

    patience = 5
    best_val_loss = float('inf')
    epochs_no_improve = 0

    for epoch in range(config['num_epochs']):
        model.train()
        total_train_loss = 0
        for inputs, targets in train_loader:
            inputs, targets = inputs.to(config['device']), targets.to(config['device'])
            optimizer.zero_grad()
            outputs = model(inputs)
            loss = criterion(outputs, targets)
            loss.backward()
            optimizer.step()
            total_train_loss += loss.item()

        avg_train_loss = total_train_loss / len(train_loader)
        train_losses.append(avg_train_loss)
```

Figure 5.12: Training loop

Later, the model was fine-tuned by systematically varying key hyperparameters. Notably, the number of epochs (from 1 up to 100), learning rate, batch size, optimizer choice, and loss function, as shown in Fig. 5.13.

We observed that when training to 100 epochs, the training and validation losses had plateaued halfway through the epochs. To guard against this and prevent overfitting, we implemented early stopping (patience = 10), which automatically halted training when the validation loss failed to improve for ten consecutive epochs. By inspecting the PSNR and FID curves (Fig. 5.13), we found that a stopping point at 75 epochs produced the best trade-off. The PSNR climbed from just 11.09 dB at epoch 1 and 14.53 dB at epoch 50 up to 15.33 dB at epoch 75, surpassing the 15.11 dB achieved at epoch 100, while the FID fell from 1,542.67 down to 518 at epoch 75 (versus 616.3 at epoch 100). We therefore saved the model weights at 75 epochs and used this checkpoint for all downstream test-set evaluations. This fine-tuning strategy ensured we maximized reconstruction fidelity (higher PSNR) and feature-level realism (lower FID) without unnecessary epochs or overfitting.



1			PSNR	FID
2	1	1 epoch, 199 images	11.09 dB	1542.67
3	2	2 epochs, 199 images	11.88 dB	1559.62
4	3	50 epochs, 199 images	14.53 dB	629.71
5	4	100 epochs, 199 images	15.11 dB	616.3
6	5	75 epochs, 199 images	15.33 dB	518
7	6	75 epochs, 199 images, SGD optimizer	12.01 dB	1444.68
8	7	75 epochs, 199 images, Adam optimizer, 1e-4 lr	14.16 dB	682.01

Figure 5.13: Tuning hyperparameters

As a result, the following configurations were drawn as the most effective parameters of the model to create the optimum pre-processed image.

Hyper-parameters

Initial setup:

- Epochs: 50
- Optimizer: Adam
- Learning rate: 1×10^{-3}

- Batch size: 16
- Loss function: Mean Squared Error (MSE)

Training was conducted on a machine equipped with a CUDA-enabled GPU. During training, both training loss and validation loss were monitored and plotted to observe model convergence.

Removing pigments from the pre-processed image

After creating the pigmentation mask (which highlights the pigmentations on the stone surface), the goal was to remove these pigmentations from the pre-processed green pipeline image. To achieve this, a bit-wise AND operation was first used to apply the pigmentation mask onto the pre-processed image.

Then, to effectively remove the pigmentations, the result was inverted, making the pigment areas lighter or eliminated from the image. This process ensures that only the non-pigmented, smooth stone surface remains prominent in the final output, while carved and pigmented regions are suppressed (Figure 5.14).

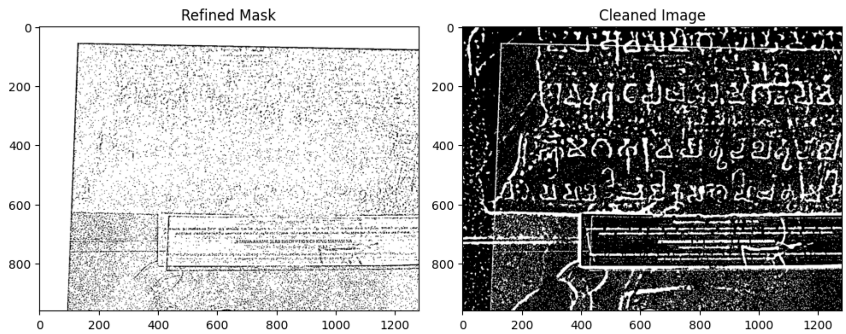


Figure 5.14: The pigmentation mask and the pre-processed image after removing the mask

Finally, normalization was applied to adjust the intensity range, preparing the cleaned image for further processing or evaluation.

Chapter 6

Results and Evaluation

6.1 Introduction to Evaluation Strategy

The research evaluation is divided into two complementary components.

1. Model Evaluation - assesses the performance and output of the implemented model (as detailed in Chapter 4 - Methodology)

2. User-Centric Evaluation - gathers and analyzes feedback from domain experts on those outputs.

Together, these approaches provide both quantitative and qualitative insights into model effectiveness and perspectives on its practical value.

6.2 Model Centric Evaluation

6.2.1 Experimental Dataset

Model evaluation was conducted using the test set, consisting of green band input images and their corresponding preprocessed ground truth images. Out of the 200 total images, 20 images are used for the testing dataset.

6.2.2 Quantitative Image-Quality Metrics

Following metrics were used to measure the quantitative image quality metrics for model evaluation.

The reconstruction quality between predicted and ground truth images were calculated using the PSNR. The calculate_psnr function (Fig. 6.1) computes the Peak Signal-to-Noise Ratio between two images, typically the predicted and the ground truth. It first calculates the Mean Squared Error (MSE) using PyTorch's built-in MSE function. If the MSE is zero (meaning the images are identical), the function returns infinity, indicating perfect similarity. Otherwise, it will calculate the PSNR as shown in (Fig. 6.1), where MAX is 1.0 for normalized images.

```python
def calculate_psnr(pred, target):
    mse = nn.functional.mse_loss(pred, target)
    if mse.item() == 0:
        return float('inf')
    return 20 * math.log10(1.0) - 10 * math.log10(mse.item())
```

Figure 6.1: Code to find PSNR

Feature-level similarity was measured using the Fréchet Inception Distance (FID). The calculate_fid function (Fig. 6.2) computes the FID between two sets of image feature activations, one from ground truth images and the other from generated (predicted) images. It calculates the mean and covariance of both activation sets, then uses these statistics to compute the Fréchet distance. A small value is added to the matrix product to ensure numerical stability. If any complex numbers appear due to the matrix square root operation, the real part is used. FID reflects the similarity of image distributions in feature space, where a lower FID indicates that the generated images are more similar to the real ones in terms of visual structure and content.

```python
def calculate_fid(real_activations, fake_activations):
    mu1 = real_activations.mean(0).cpu().numpy()
    mu2 = fake_activations.mean(0).cpu().numpy()
    sigma1 = torch.cov(real_activations.T).cpu().numpy()
    sigma2 = torch.cov(fake_activations.T).cpu().numpy()

    diff = mu1 - mu2
    covmean = sqrtm(sigma1 @ sigma2 + 1e-6 * np.eye(sigma1.shape[0]))

    if np.iscomplexobj(covmean):
        covmean = covmean.real

    fid = diff.dot(diff) + np.trace(sigma1 + sigma2 - 2 * covmean)
    return float(fid)
```

Figure 6.2: Code to find FID

The function get_inception_features(images) (Fig. 6.3) takes a list or batch of images and extracts high-level feature representations using a pretrained Inception v3 model. These features are typically used for FID computation, which evaluates how similar two sets of images are in terms of visual realism and structure.

```python
def get_inception_features(images):
    inception = inception_v3(pretrained=True, transform_input=False).to(config['device'])
    inception.eval()
    features = []

    preprocess = transforms.Compose([
        Resize((299, 299)),
        CenterCrop(299),
        ToTensor(),
        Normalize(mean=[0.5]*3, std=[0.5]*3)
    ])

    with torch.no_grad():
        for img in images:
            if img.shape[0] == 1:
                img = img.repeat(3, 1, 1)
            img = preprocess(transforms.ToPILImage()(img.cpu())).unsqueeze(0).to(config['device'])
            feat = inception(img)
            features.append(feat.squeeze())

    return torch.stack(features)
```

Figure 6.3: Extracting inception features

Predicted and ground truth images were visualized side-by-side to qualitatively assess the output images and the model's performance. (Fig. 6.4)

```
if visualize_index is not None and visualize_index < len(all_inputs):
    pred_img = all_outputs[visualize_index].squeeze(0).squeeze(0).numpy()
    target_img = all_targets[visualize_index].squeeze(0).squeeze(0).numpy()

    plt.figure(figsize=(12, 4))

    plt.subplot(1, 2, 1)
    plt.imshow(pred_img, cmap='gray')
    plt.title('Predicted Output')
    plt.axis('off')

    plt.subplot(1, 2, 2)
    plt.imshow(target_img, cmap='gray')
    plt.title('Ground Truth')
    plt.axis('off')

    plt.tight_layout()
    plt.savefig(f'visualize_test_{visualize_index}.png')
    plt.show()
elif visualize_index is not None:
    print(f"Index {visualize_index} is out of range for test set size.")
```

Figure 6.4: Code to visualize test set data

The evaluate function assesses the performance of a trained image preprocessing model on a test dataset. It loads the test data using a DataLoader and processes each image one by one without updating the model (using torch.no_grad() to save memory and speed up evaluation). For each input-target pair, it generates a predicted output, clamps it between 0 and 1, and calls the PSNR calculating function. It also stores the real (target) and generated (output) images for further comparison. The FID calculating function is then called to find the FID

```
def evaluate(model, visualize_index=None):
    model.eval()
    test_set = PreprocessingDataset(os.path.join(config['dataset_dir'], 'test'))
    test_loader = DataLoader(test_set, batch_size=1, shuffle=False)

    psnr_scores = []
    real_images = []
    fake_images = []

    all_inputs = []
    all_targets = []
    all_outputs = []

    with torch.no_grad():
        for inputs, targets in test_loader:
            inputs, targets = inputs.to(config['device']), targets.to(config['device'])
            outputs = model(inputs)
            outputs = torch.clamp(outputs, 0, 1)

            psnr = calculate_psnr(outputs, targets)
            psnr_scores.append(psnr)

            real_images.append(targets.squeeze(0).cpu())
            fake_images.append(outputs.squeeze(0).cpu())

            all_inputs.append(inputs.cpu())
            all_targets.append(targets.cpu())
            all_outputs.append(outputs.cpu())

    avg_psnr = sum(psnr_scores) / len(psnr_scores)
    print(f"\nAverage PSNR on test set: {avg_psnr:.2f} dB")

    real_feats = get_inception_features(real_images)
    fake_feats = get_inception_features(fake_images)
    fid_score = calculate_fid(real_feats, fake_feats)
    print(f"FID Score on test set: {fid_score:.2f}")
```

Figure 6.5: Evaluation function

6.2.3 Model Performance Results

The graph in Fig 6.6 illustrates the progression of PSNR and FID scores across different training epochs. Notably, the 75th epoch yields the best performance, achieving a PSNR of 15.33 dB and the lowest FID score of 518.

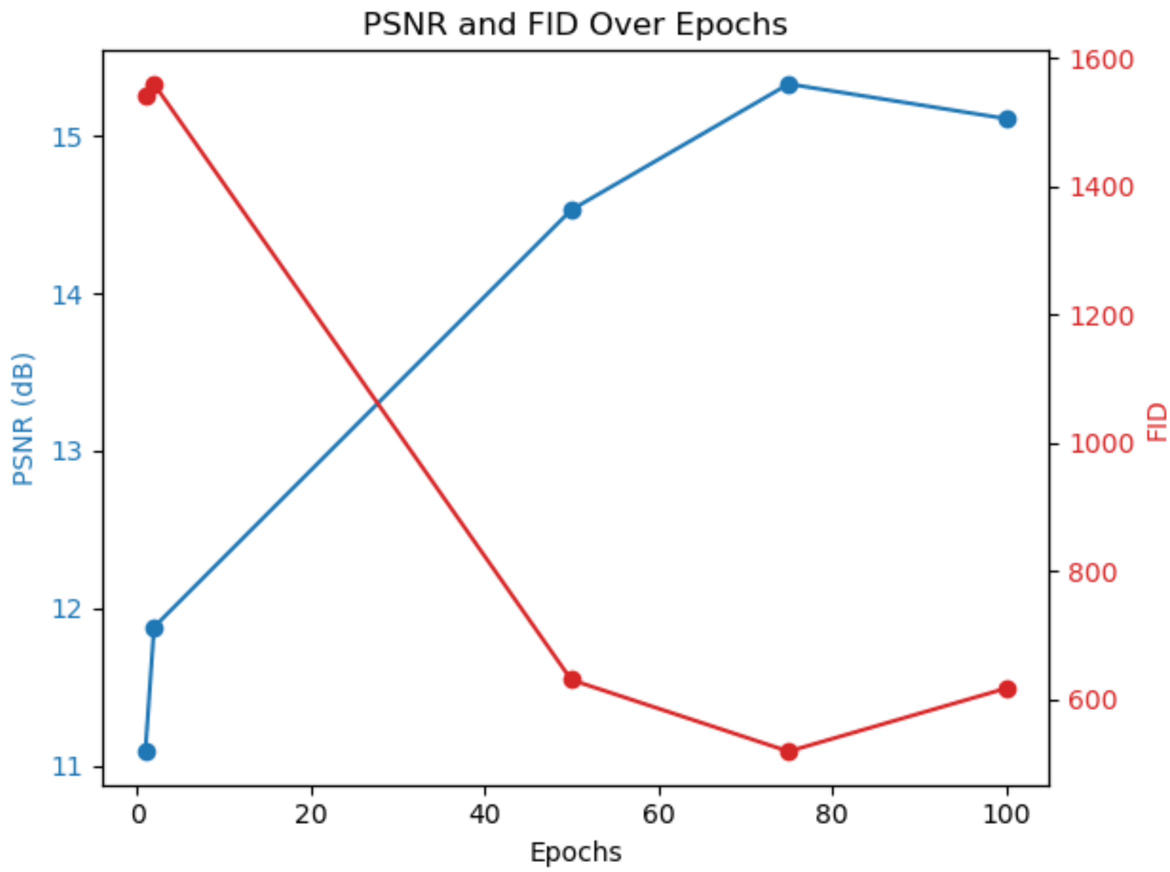


Figure 6.6: PSNR and FID values over different training epochs

The loss graph (Fig 6.7) shows the progression of training and validation losses over the 75 epochs.

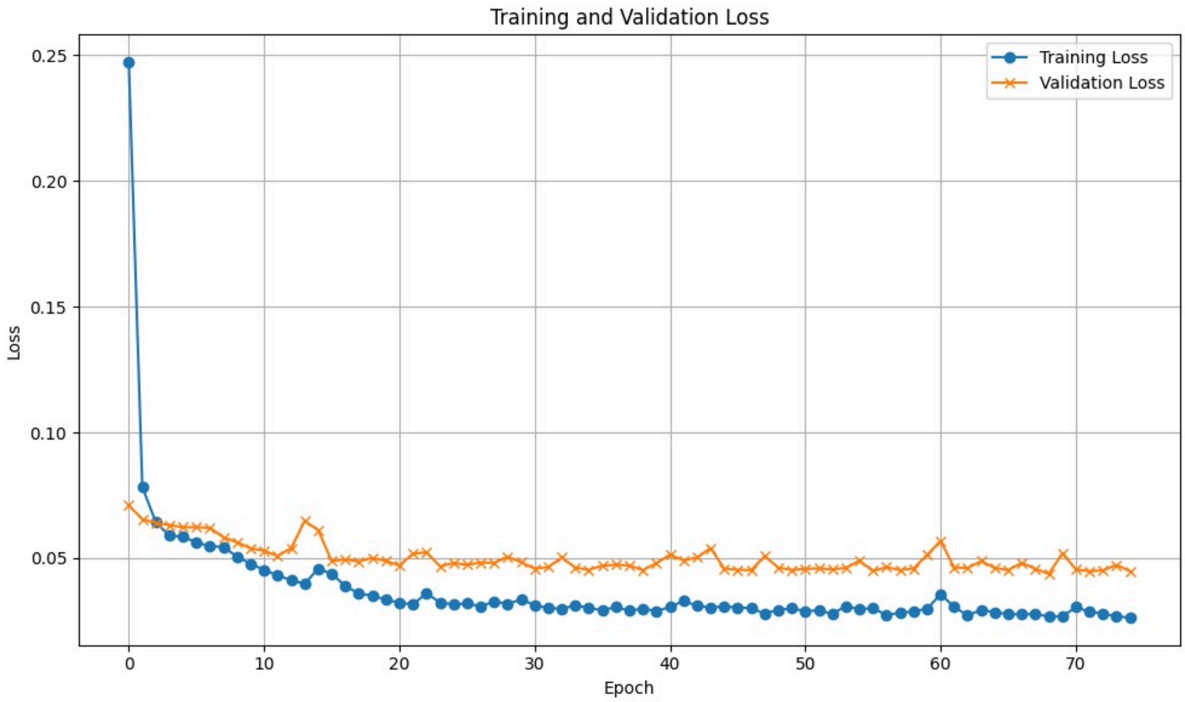


Figure 6.7: Model loss graph

6.3 User-Centric Evaluation

To assess how well our approach met the needs of archaeology specialists and bridge the research gap, we conducted a user evaluation with 9 domain experts. We met with each expert in person to walk them through the complete proposed approach, ensuring they clearly understood the context and goals of the study. After taking the consent to participate in the evaluation, they were asked to complete a Google Form with 25 questions covering four complementary sections as participant demographics, an image-comparison, a Likert-scale questionnaire, and open-ended qualitative feedback. By measuring across these data sources, we obtained both objective measures of image quality and insights into expert preferences and suggestions for further improvements.

6.3.1 Participant Demographics

We engaged nine domain experts in archaeology and epigraphy via the Department of Archaeology, Sri Lanka and the University of Colombo Faculty of Arts (Department of Sinhala). Each participant provided their full name, institutional affiliation, and professional designation.

We then asked about their years of experience in the archaeology/ epigraphy field, yielding the following distribution: (Fig. 6.8)

- 1–3 years: 1 participant
- 4–7 years: 1 participant

- 10 years: 7 participants

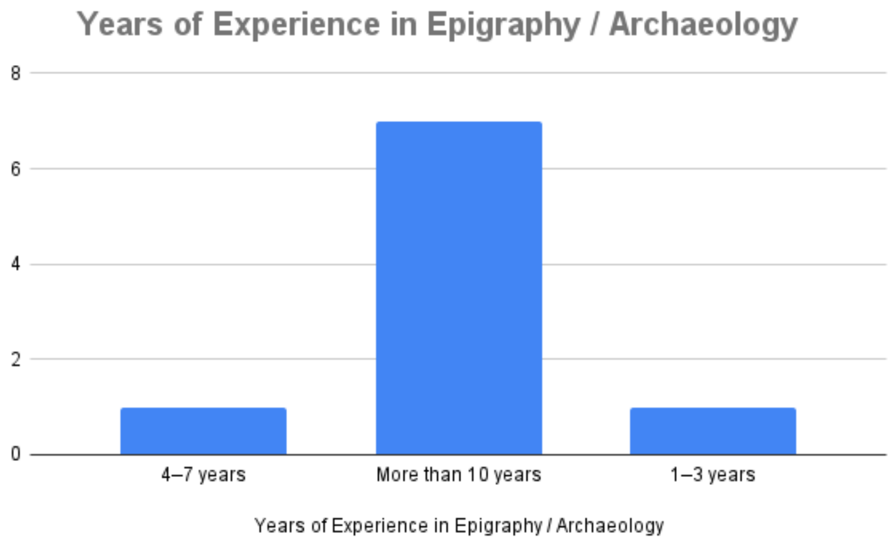


Figure 6.8: Experience in archaeology/epigraphy domain

To assess the familiarity with estampage techniques, we inquired about both physical and digital experience (Fig. 6.9). The responses were:

- No prior estampage experience: 1 participant (11.1%)
- Physical only: 2 participant (22.2%)
- Both physical and digital: 6 participants (66.7%)

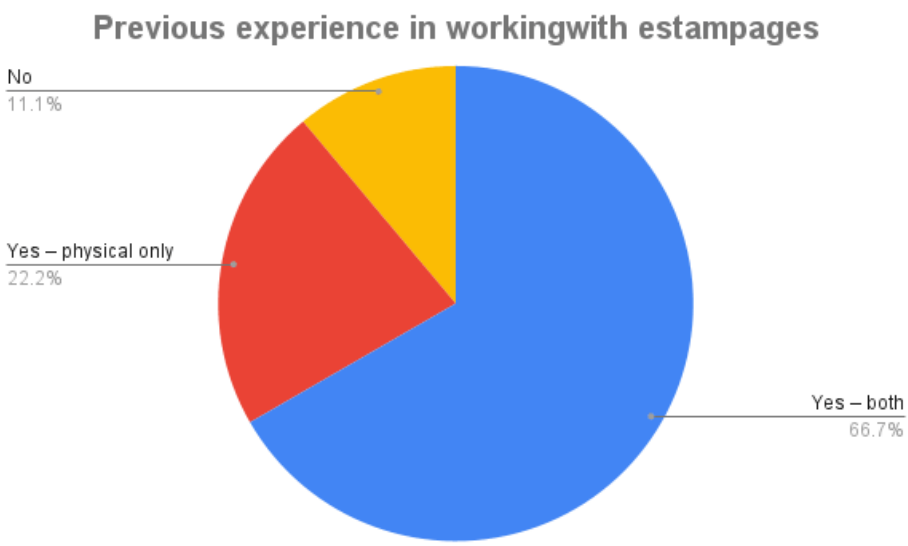


Figure 6.9: Experience in working with estampages

6.3.2 Image-Comparison Task (Quasi-Quantitative)

In the Image-Comparison section, there are 7 sets of images to compare, where each expert was shown two unlabeled images of a traditional physical estampage photograph and the corresponding digital estampage output from our approach, and asked to indicate which version appeared clearer and more suitable for reading and interpreting the inscription.

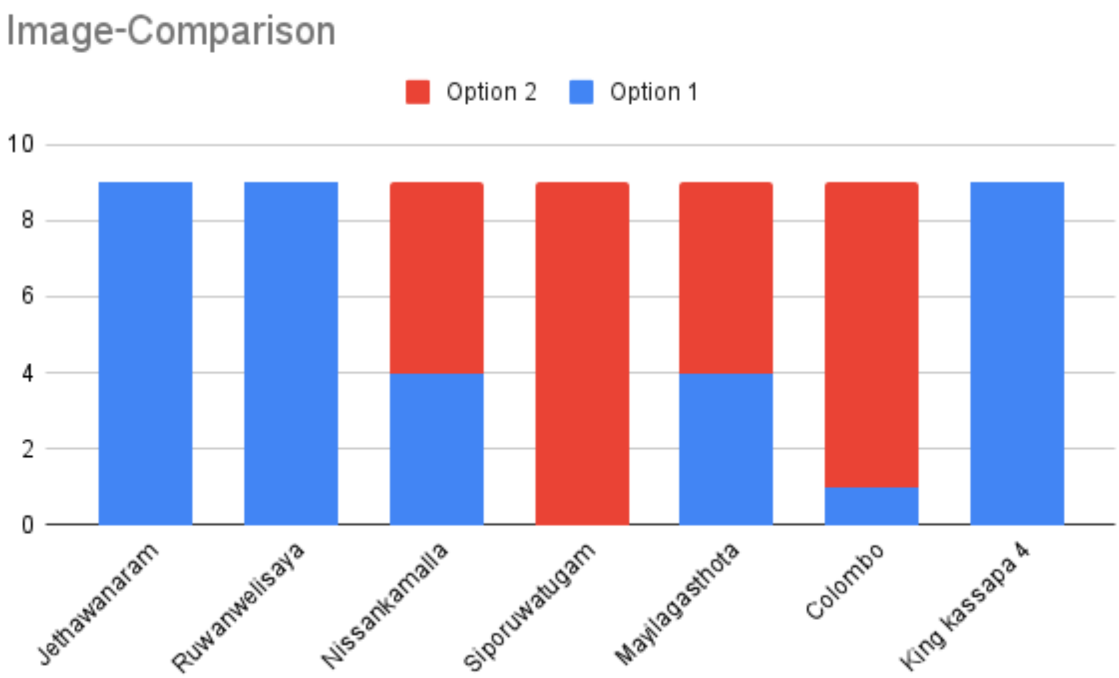


Figure 6.10: Comparison between physical estampages and digital estampages

Following is the labeled set of images with their corresponding numbers of responses for being clearer and suitable for reading.

	Option 1	**Option 2**	**Total votes**	**%Digital**	**%Traditional**
Jethawanaramaya	Digital (9)	Traditional (0)	9	100%	0%
Ruwanwelisaya	Traditional (9)	Digital (0)	9	0%	100%
Nissankamalla	Traditional (4)	Digital (5)	9	55.5%	44.5%
Siporuwatugama	Traditional (0)	Digital (9)	9	100%	0%
Mayilagastohta	Digital (4)	Traditional (5)	9	44.5%	55.5%
Colombo museum	Digital (1)	Traditional (8)	9	11.1%	88.9%
King Kassapa IV	Digital (9)	Traditional (0)	9	100%	0%
Total			**63**	**58.73%**	**41.26%**

Table 6.1: Voting Results for Different Sites

Table 6.1 shows the individual votes for the digital and traditional estampage readability. Some experts preferred digital estampages completely over traditional (Jethawanaramaya and King Kassapa IV and Siporuwatugama), while some preferred images of traditional estampages completely over digital (Ruwanwelisaya). Most of the comparisons had diverse selections for both digital and traditional estampages. Overall, the digital estampage acceptance rate stands at 58.73%, whereas the traditional estampages are accepted at 41.26%.

6.3.3 Semi-Quantitative Likert-Scale Analysis

To evaluate the effectiveness (accuracy, time efficiency and practicality) of the proposed digital estampage system utilizing multispectral imaging and image processing, feedback was collected from respondents across multiple criteria.

Participants were asked to evaluate the clarity and expressiveness of physical features such as letters and symbols in the digital estampage outputs generated by the proposed multispectral imaging system. Subsequently, they rated the usefulness of these digital estampages for epigraphic analysis. The questionnaire also gathered responses regarding the likelihood of recommending the method to peers. Following this, participants assessed the ease of integrating the method into existing workflows. To determine time efficiency, users compared the proposed approach to traditional methods. Additionally, they evaluated the system's ability to minimize errors typically associated with manual interventions in traditional estampage processes. Finally, participants provided an overall rating of the multispectral system as a tool for creating digital estampages.

Each question was rated on a Likert scale from 1 to 5, where 5 indicates the most favorable response. This analysis provides a semi-quantitative understanding of expert perceptions, enabling a balanced assessment of both the technical and practical value of the proposed approach.

The number of responses for the ratings of each question is given in Table 6.2.

Table 6.3 shows the significance of all the criteria assessed in the Likert scale section.

6.3.4 Qualitative Thematic Analysis of Open-Ended Feedback

Our thematic analysis of expert feedback revealed several key themes such as enhanced readability, complementarity between digital and physical estampages, efficiency gains, contextual appropriateness, trust and data-management concerns, and technical/ operational advice.

Participants' verbal comments were recorded in real time (with their consent) during the Google Form session, and written responses were exported for analysis.

Positive Feedback

Participants frequently noted that increased contrast in processed images sharpened letter forms, confirming that multispectral enhancement supports clearer reading in the field. One participant stated, "When the background in the digital estampage is darker, the inscription characters stand out and are much easier to read with the naked eye."

A participant remarked, "In most cases, the digital estampages provided more legible text than the traditional physical estampages." indicating the perception that our digital pipeline yields higher-quality results across different inscription types.

	1	2	3	4	5
Level of clarity (1 - very poor, 5 - Excellent)	-	-	6	3	-
Expression of features (letters/ symbols) (1 - very poor, 5 - Excellent)	-	-	4	5	-
Usefulness (1 - Not useful, 5 - Completely)	-	-	1	5	3
Trust (1 - Not at all, 5 - Completely)	-	-	4	5	-
Ease of integration (1 - Very difficult, 5 - Very easy)	-	-	1	5	3
Level of training & tech support to adopt (1 - None, 5 - Extensive training needed)	-	5	4	2	-
Time efficiency compared to the traditional approach (1 - Much slower, 5 - Much faster)	-	-	1	-	7
Level of capacity to minimize errors (1 - Minimal, 5 - Very high)	-	1	2	4	1
Overall, as a tool (1 - Poor, 5 - Excellent)	-	-	2	5	2
Recommend to peers (1 - Definitely no, 5 - Definitely yes)	-	-	1	2	6

Table 6.2: Number of responses for each rating

Statements such as "I would definitely recommend this tool to colleagues if it becomes widely available" and "We can get a good use of this tool if integrated" reflect participants' confidence in adopting the method.

A participant noted, "We can get a great deal of use out of this tool if it is integrated into our workflow." This indicates the potential for the pipeline to be incorporated into existing procedures.

Neutral Observations

One participant stated, "For some inscriptions, the digital approach works best; for others, the physical estampage still delivers clearer detail." This highlights that material condition and surface geometry influence method effectiveness.

Participants acknowledged that while digital methods improve overall readability, some characters remain unclear. A participant commented, "In certain digital images, some characters remain blurry, while other letters are perfectly crisp."

A participant mentioned, "Most of us have only ever tried digital estampage once, on the Dimbulagala inscription." This reflects the novelty of the technique of creating digital estampages of stone inscriptions in Sri Lanka.

"Currently, we sometimes photograph a physical estampage and then tweak lighting and sharpness in Photoshop to read faint letters." This indicates that experts are already experimenting with digital tools.

Criterion	Mean score	t-statistic	p-value	Significance
Clarity	3.33	2.00	0.081	Not significant
Features	3.56	3.16	0.013	Significant
Usefulness	4.22	5.50	0.0006	Highly significant
Trust	3.56	3.16	0.013	Significant
Integration	4.22	5.50	0.0006	Highly significant
Training support	2.73	-1.15	0.138	Not significant
Time efficiency	4.67	7.07	0.0001	Highly significant
Minimize errors	3.63	1.93	0.095	Not significant
Overall tool	4.00	4.24	0.0028	Significant
Recommend	4.56	6.42	0.0002	Highly significant

Table 6.3: Significance of each criterion.

Negative Feedback and Concerns

A participant expressed, "Although the tool is promising, I worry about the long-term reliability of digital storage for these archival images." This concern emphasizes the need for robust data preservation strategies.

Participants stressed the importance of user-friendly documentation. One noted, "We will need clear, non-technical documentation explaining how the model works before field archaeologists can trust it."

A participant also commented, "Capturing multispectral images at different times of day and multiple angles could become tedious during a short field campaign." This highlights practical challenges in field conditions.

Suggestions for Enhancement

Suggestions such as "Let's record digital estampages alongside physical estampages during a field visit, then compare them systematically later.", "We should consider camera rotation, lighting angles, and timing to ensure the clearer captures are taken of the stone inscriptions.", "No formal training should be needed. If it's just a matter of snapping a photo and uploading, anyone can use it." were stated by the expert pool while they were evaluating the proposed digital estampage creation pipeline.

Chapter 7

Discussion

This chapter presents a comprehensive discussion of the research findings derived from the study on creating digital estampages of Sri Lankan stone inscriptions using multispectral imaging. It explores the feasibility and effectiveness of the proposed methodology, offering detailed insights into its strengths and practical implications. Furthermore, the chapter critically reflects on the evaluation outcomes and highlights the limitations encountered during the study. These discussions lay the groundwork for the subsequent chapter, which will address potential directions for future research.

7.1 Research Findings

Our quantitative image-comparison user evaluation demonstrated a statistically significant preference for digital estampages over traditional estampages (58.7 % vs. 41.3 %, z = 1.39, $p < 0.10$), confirming that multispectral imaging combined with our image-processing pipeline improves inscription legibility. (Table 6.1)

The image-processing model achieved its highest visual fidelity at epoch 75 (PSNR = 15.33 dB, FID = 518) and showed stable loss convergence, indicating both effective learning and strong generalization to unseen data. (Fig. 6.6)

Experts rated the pipeline highly for usefulness, time efficiency, and ease of workflow integration (all $p < 0.05$), and expressed strong willingness to recommend it, demonstrating that archaeologists can adopt the system with tangible accuracy and efficiency gains. (Table 6.3)

Qualitative feedback revealed that, while digital estampages offer clear advantages, experts remain cautious about fully replacing physical estampages, viewing the two methods as complementary, especially for deeply carved or well-preserved inscriptions

Visual comparisons (Fig 7.1) illustrate that the model is able to preserve inscription structures and features (characters/ symbols) almost similarly to the ground truth image. These results show the model's potential for pre-processing stone inscription images.

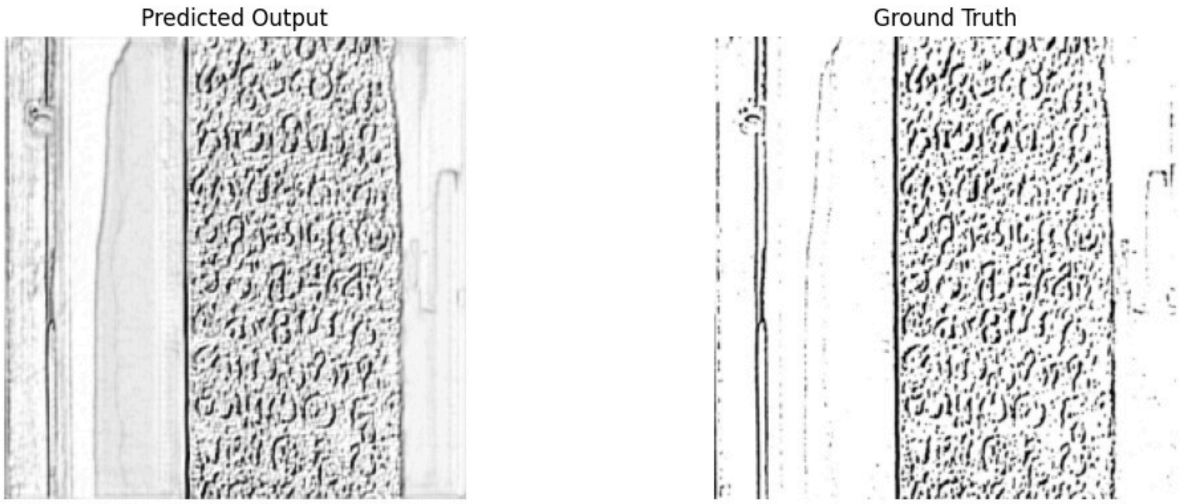


Figure 7.1: Visual output results of the model

These research findings imply that technically, the model is well-trained by epoch 75, producing high-fidelity outputs while practically, experts value the digital estampages for their clarity, efficiency, and integration potential, yet still view them as complementary rather than outright replacements for traditional methods.

7.2 Discussion

Referring to Fig. 6.6 the 75th epoch yields the best performance, achieving a PSNR of 15.33 dB and the lowest FID score of 518. It suggests that the model's output at this stage most closely resembles the ground truth in terms of visual quality (as indicated by the high PSNR) and realism (as indicated by the low FID). The improvement trend in PSNR and the significant drop in FID up to the 75th epoch highlight the effectiveness of continued training in enhancing output quality, before marginal fluctuations appear at the 100th epoch.

Referring to Fig. 6.7 the training loss consistently decreased, indicating effective learning, while the validation loss also decreased, showing that the model generalizes reasonably well without overfitting.

RQ1: Does multispectral imaging, combined with a digital image-processing pipeline, significantly enhance the readability of stone inscriptions compared to traditional estampage methods?

Referring to Table 6.1 with 63 total votes across 7 inscriptions, in assessing whether the digital estampage readability is statistically significantly greater than 50%, we conducted a Z-test.

Null hypothesis (H_0) - There is no improvement in readability in the digital estampages compared to traditional estampages ($p = 0.5$)

Alternative hypothesis (H_1) - Digital estampages are more readable compared to traditional estampages ($p > 0.5$)

The total number of samples is 63, and 37 participants expressed the preference for digital estampages as readable, with a standard error of approximately 1.39. Using a 90% confidence level ($\alpha = 0.10$), the critical z-value for a one-tailed test is 1.28. Since the calculated z-score of $1.39 > 1.28$, the null hypothesis is rejected. This provides statistically significant evidence that, at the 90% confidence level, the digital estampages produced using multispectral imaging and the proposed image-processing pipeline are more readable than traditional estampages, supporting to meet the research question 1.

The results from the image comparison evaluation (Section 5.3.2) indicate that the proposed digital estampage method was preferred in 58.73% of cases across seven different inscription sites, with traditional estampages preferred in 41.26% of responses. While preferences varied across inscription types and image pairs, inscriptions like Jethawanaramaya, King Kassapa IV, and Siporuwatugama showed a unanimous preference for the digital estampage.

Notably, however, a few inscriptions (Ruwanwelisaya and Colombo Museum) still favored traditional estampages, highlighting that the effectiveness of the digital method may vary depending on stone surface characteristics, lighting, or artifact erosion levels. These observations suggest that while the digital approach holds promise, it may be best viewed as a complementary method rather than a universal replacement at this stage.

Next an analysis on the results according to the hypothesis of RQ2 was conducted.

RQ2: How can archaeologists practically incorporate a multispectral-imaging-based digital estampage pipeline into their current inscription-documentation workflow, and what are the resulting impacts on accuracy and time efficiency?

Null Hypothesis (H_0): The multispectral-imaging-based digital estampage pipeline does not significantly improve accuracy, reduce documentation time, or isn't practically applicable compared to traditional methods.

Alternative Hypothesis (H_1): The pipeline does significantly improve accuracy, reduce documentation time, and is practically applicable.

To statistically prove the hypothesis, first, we tabulated the 10 evaluation criteria (6.2) and the average score was computed for each question. A t-test was used to check whether the mean score for each criterion is significantly different from "3", which is considered the neutral baseline in a 5-point Likert scale. The sample size was used as 9 (responses per criterion). The mean and the standard deviation were used to compute the t-statistic, measuring how far the mean is from 3 in units of standard error. For each t-statistic, a p-value was derived using the t-distribution. This p-value indicates the probability that the observed mean could have occurred by random chance if the true mean were actually 3. The approach has a significance level of 0.05. If the p-value< 0.05, the null hypothesis (that the score is not better than neutral) is rejected.

There was an exception for the training support criterion as we have to perform the t-test in the lower direction because the lower the score it is the positive the outcome.

In elaborating 6.2, rates for clarity of the digital estampages received mostly moderate ratings. Six respondents rated it as 3 (moderate), and three rated it as 4 (above moderate), indicating that while digital estampages are clear but there is still room for improvement in visual quality. Similarly, for the representation of physical features such as letters and symbols, responses were more positive, where three gave it a rating of 3 (moderate), while six respondents rated it as 4 (more than moderate). These results suggest that the system does enhance the readability and physical accuracy of inscriptions.

Consindering the rates for the usefulness of the digital estampages for epigraphic analysis, respondents reviewed higher ratings, with a trend toward 4 and 5 (high usefulness), suggesting strong support for its practical application in research. This aligns well with RQ2 indicating that digital estampages are practical and useful for epigraphic analysis.

When evaluating trust in the system for formal or scholarly use, most respondents gave ratings of 3 or 4 (moderate and more than moderate), showing cautious optimism. This mixed but generally favorable trust level suggests growing acceptance of the technique, indirectly supporting RQ2 on accuracy and practicality.

Furthermore, responses on recommendation to peers skewed strongly positive, with ratings of 4 and 5 (more than moderate and a strong yes) dominating. This suggests that users see value in the approach and are willing to advocate its use, indirectly addressing RQ2 on the perspectives of the domain specialists on the effectiveness of the digital estampages.

Most of the respondents rated the ease of integrating the method into existing workflows as 4 and 5, indicating that users find it practical and easy to implement, providing strong evidence for the practicality aspect of RQ2. However, when asked about training and support requirements, five respondents rated it as 2 (low training needed), while the rest rated it 4 or 5 (high support needed). This bimodal result suggests that the ease of use may vary depending on the user's familiarity with digital tools, and further training resources may be beneficial for wider adoption.

Time efficiency compared to traditional methods was rated quite positively, with the majority assigning a score of 5. This suggests that the digital method significantly reduces the time taken compared to traditional estampage techniques, directly addressing RQ2.

User rates on the capacity of the tool to minimize errors occurred by the manual interventions of traditional estampages, received improving scores from 2 up to 5, showing that while not perfect, the system is perceived to be increasingly effective in reducing interpretive inaccuracies caused by unclear inscriptions further supporting RQ2.

The system's overall rating as a tool for preservation and study was distributed between 3 and 5, with a strong leaning toward positive scores. This reflects general satisfaction with the system and its contributions to the field.

We can reject the null hypothesis driving support to RQ2, because multiple key criteria (usefulness, time efficiency, integration, trust, recommend, etc) are significantly lower than 0.05. (Table 6.3)

7.3 Limitations

Since the research study involved collaboration with an external party (Department of Archaeology), communication constraints occasionally impacted the coordination and timely acquisition of feedback.

The expert pool for the evaluation was relatively small (n = 9 per comparison). While this group provided valuable insights for an initial assessment, the findings may not generalize to broader archaeological or epigraphical communities.

Although the evaluation involved image preference tests by experienced experts, there remains an inherent degree of subjectivity. Results may vary depending on the users' demographics, familiarity with digital methods, or levels of epigraphical expertise.

The majority of the inscriptions analyzed were sourced from the Colombo National Museum, where inscriptions are well-preserved and relatively free from moss, dirt, or other environmental damage. As a result, the system's performance on severely weathered or outdoor inscriptions remains uncertain.

Additionally, the processing pipeline was primarily tested on image sets captured under specific camera settings and spectral bands (limited to five bands). Variations in imaging equipment, spectral range, or lighting setups may affect the system's output quality and consistency.

Finally, not all inscription types benefited equally from the digital enhancement. Factors such as environmental conditions, carving depth, and stone surface texture influenced the effectiveness of the enhancement techniques.

Chapter 8

Conclusion

8.1 Conclusion

The aim of this study was to explore the possible advantages of using a multispectral imaging and a pre-processing pipeline compared to traditional physical methods for creating and recording ancient inscriptions in Sri Lanka. Through the use of an experimental setup and expert evaluations, the research questions were clearly tested, providing strong evidence that supports the put forward hypothesis. The findings indicated that the specialists liked the digital estampages generated by the pipeline most of the time. The finding proved that multispectral imaging and processing have the potential to enhance the legibility of inscriptions much more than conventional procedures. Besides, the study not only answered the issues posed but also fulfilled the objectives and purposes established from the start. The pipeline was constructed, assessed, and validated in the field, establishing its viability as an adjunct tool to physical estampages.

The findings also revealed that while the digital estampage creation method enhanced clarity and convenience, traditional physical estampages still had unique advantages, particularly in the examination of surface textures that require manual inspection. Researchers acknowledged the efficiency and transparency of the digital outputs, but they insisted that both physical and digital approaches possess particular strengths in relation to the specific context. This is to say that, instead of substituting physical approaches, digital approaches can complement them, thereby enhancing the documentation and interpretation of inscriptions. The study was able to response to the fact that multispectral imaging and digital processing provide a significant and practical improvement over conventional methods employed in the analysis of stone inscriptions in Sri Lanka.

8.2 Recommendations

Based on both the quantitative preferences and qualitative feedback from domain experts (Chapter 5), several recommendations can be made to guide and improve the development and deployment of the proposed digital estampages creation.

Incorporating manual review and annotation tools is recommended because it allow experts to annotate, mark-up, and provide commentaries on improved images directly, thereby enabling collaborative analysis with ease.

A hybrid workflow is recommended at the beginning as it was mentioned by most of the experts at the qualitative evaluation. It promote a two-pronged approach utilizing digital estampages for preliminary research and recording, but maintaining traditional methods for final analysis in specific situations, especially in cases where the digital method does not exceed traditional means. Gradually, the creation of traditional estampages can be replaced with the digital estampage creation with more advancements.

User Training and Documentation is recommended when moving forward with the proposed approach which will offer user guides or training modules for inexperienced digital imaging users, specifically in handling multi-spectral data and interpreting combined outputs.

It is recommended to highlight uncertainty in results when needed to notify users when the enhancement algorithm is uncertain, particularly across regions characterized by ambiguity or noise.

Flexible enhancement options is recommended to create easy-to-use interfaces that enable professionals to choose image processing parameters based on the type of inscription or particular site characteristics, perhaps with pre-programmed modes such as ”faint carving” or ”uneven background.”

Finally addressing a limitation mentioned in Section 6.3, it is recommended to promote institutional cooperation collaborating with museums, universities, and archaeology departments to pilot and jointly develop the system in actual projects, making it practically relevant and iteratively refined.

8.3 Future Work

Building on the foundation laid by this research and its findings, several promising directions have opened up for future exploration and system development.

Currently, the proposed pipeline follows a fixed method for enhancing estampages. As noted in Section 5.3.4, one natural extension would be to develop adaptive enhancement algorithms that can automatically adjust based on the characteristics of the stone surface such as weathering, brightness, or the depth of carved characters.

Another important area for future development is the integration of deep learning techniques. By training models to segment letters or even assist in the automatic transcription of inscriptions, it would be possible to significantly reduce the manual effort required during post-processing.

Expanding the dataset to include a broader range of inscriptions from different geographical locations and varying levels of preservation was highlighted by experts during the evaluation phase (Section 5.3.4). In addition, gathering feedback from a larger and more diverse group of users, including epigraphers, archaeologists, and digital humanities scholars, would help produce results that are more widely generalizable.

Future versions of the system could explore mobile application integrations, enabling on-site preliminary processing and making the technology more accessible for field archaeologists and researchers.

Several experts also expressed interest in having more interactive viewing tools, where users could tweak enhancement parameters in real-time such as selecting different spectral bands, applying custom filters, or zooming into blended images to analyze finer details.

Future evaluations could also further explore in-field interpretations, where both physical and digital estampages are captured at the same location and on the same day. This would allow for assessing not just visual clarity, but also the epigraphical accuracy of the enhancements.

References

[1] Inscriptions.lk, "Introduction to inscriptions," *Inscriptions.lk*, 2024. Accessed: 2024-03-28.

[2] H. M. S. C. R. Heenkenda and T. G. I. Fernando, "Approaches used to recognise and decipher ancient inscriptions: A review," 2020.

[3] Anon., "Alternative gravestone reading methods," 2015. [Online].

[4] K. Dias, I. Perera, and I. Nawinne, "An architecture for an inscription recognition system for sinhala epigraphy," *International Journal of Computer Applications*, 2016. Presented at the International Conference on Advances in ICT for Emerging Regions (ICTer 2016).

[5] K. G. N. D. Karunarathne, K. V. Liyanage, D. A. S. Ruwanmini, G. K. A. Dias, and S. T. Nandasara, "Recognizing ancient sinhala inscription characters using neural network technologies," *International Journal of Scientific Engineering and Applied Science (IJSEAS)*, vol. 3, no. 1, pp. 1–13, 2017.

[6] Ceylon Today, "De-conspiring the inscription," *Ceylon Today*, 2024. Accessed: 2024-06-28.

[7] Wikipedia contributors, "Cultural property imaging." `https://en.wikipedia.org/wiki/Cultural_property_imaging`, 2025. Accessed: 2025-01.

[8] S. R. Miranda *et al.*, "Revealing inscriptions obscured by time on an early-modern lead funerary cross using terahertz multispectral imaging," *Scientific Reports*, vol. 12, no. 2134, 2022.

[9] S. Faigenbaum-Golovin *et al.*, "Multispectral images of ostraca: Acquisition and analysis," *Journal of Archaeological Science*, vol. 39, no. 12, pp. 3581–3590, 2012.

[10] M. Jones *et al.*, "Understanding multispectral imaging of cultural heritage," *Journal of Cultural Heritage*, vol. 41, pp. 60–72, 2020.

[11] P. Molton, A. Hill, and R. Smith, "Determining best practice in msi analysis of historical artefacts," tech. rep., ResearchGate, 2018.

[12] G. L. Marcus *et al.*, "Assessing the potential of multispectral and thermal uav imagery for archaeological prospection," *Journal of Archaeological Science: Reports*, vol. 28, p. 102045, 2020.

[13] L. Silva, G. Ramos, and T. Almeida, "Redefining archaeological research: Digital tools, challenges, and adoption," *Applied Sciences*, vol. 15, no. 5, p. 2495, 2024.

[14] E. Cartwright and L. David, "Quantifying readability improvements in multispectral images of historic inscriptions," *Journal of Imaging Science*, vol. 66, no. 4, pp. 1–12, 2022. DOI not found; verify journal source or substitute if incorrect.

[15] A. R. Hevner, S. T. March, J. Park, and S. Ram, "Design science in information systems research," *MIS Quarterly*, vol. 28, no. 1, pp. 75–105, 2004.

[16] M. N. K. Saunders, P. Lewis, and A. Thornhill, *Research Methods for Business Students*. Pearson, 8 ed., 2019.

[17] R. Garg, "Ancient scripts and inscriptions: A survey," *Journal of Epigraphic Studies*, vol. 45, pp. 101–115, 2021.

[18] S. Paranavitana, *Inscriptions of Ceylon, Volume I*. Colombo: Archaeological Survey of Ceylon, 1970.

[19] P. Senaratne, "Epigraphy and identity in ancient sri lanka," *South Asian Studies*, vol. 29, no. 2, pp. 123–136, 2013.

[20] M. Gillings, P. Hacigüzeller, and G. Lock, *Archaeological Theory and the Digital Past*. London: Routledge, 2016.

[21] H. Lu, "Introduction to epigraphical documentation methods," *Journal of Field Archaeology*, vol. 44, no. 1, pp. 11–25, 2019.

[22] R. Somadeva, "Reading inscription." `https://www.youtube.com/watch?v=Lfin5rVqrKM`, 2021. Accessed: 2025-04-02.

[23] C. Bohren, "Multispectral imaging for historical manuscript analysis," *Journal of Imaging Science*, vol. 52, no. 4, pp. 302–310, 2008. DOI not found; please verify journal existence or alternate title.

[24] R. Easton, K. Knox, and W. Christens-Barry, "Multispectral imaging of the archimedes palimpsest," in *Proceedings of SPIE*, vol. 7531, p. 75310A, 2010.

[25] A. Tonazzini, "Ancient document restoration via multispectral imaging and machine learning," *Pattern Recognition Letters*, vol. 121, pp. 78–85, 2019.

[26] J. Delaney, J. Zeibel, M. Thoury, R. Littleton, and M. Palmer, "Using multispectral imaging for the examination of artworks," *Studies in Conservation*, vol. 55, no. 1, pp. 1–16, 2010.

[27] J.-A. Beraldin, M. Picard, G. Godin, J. Taylor, L. Cournoyer, and M. Rioux, "Digital 3d imaging for heritage conservation," *Journal of Cultural Heritage*, vol. 3, no. 1, pp. 63–72, 2002.

[28] P. Pohankar, N. Taralkar, S. Karmare, and S. Kulkarni, "Character recognition using artificial neural network," *International Journal of Electronics Communication and Computer Engineering*, vol. 5, no. 4, p. 248, 2014.

[29] M. M. Magrina, "Recognition of ancient tamil characters from epigraphical inscriptions using raspberry pi based tesseract ocr," *International Journal of Scientific Research in Computer Science, Engineering and Information Technology*, vol. 7, no. 2, p. 127, 2021.

[30] N. Gautam and S. S. Chai, "Optical character recognition for brahmi script using geometric method," 2017.

[31] S. Rajakumar and V. Subbiah Bharathi, "Century identification and recognition of ancient tamil character recognition," 2011.

[32] A. Tomar, B. Choudhary, and R. Yerpude, "Ancient indian scripts image pre-processing and dimensionality reduction for feature extraction and classification," *International Journal of Computer Trends and Technology (IJCTT)*, vol. 27, no. 2, pp. 88–91, 2015.

[33] G. Dong, W. Shi, Y. Miao, H. Xu, and R. Feng, "Research of the edge extraction about archaic epigraph," in *IEEE*, (Henen, China), p. 4, 2010.

[34] P. Ganesan and G. Sajiv, "A comprehensive study of edge detection for image processing applications," in *Proceedings of the 2017 International Conference on Innovations in Information, Embedded and Communication Systems (ICIIECS)*, (Coimbatore, India), pp. 1–6, IEEE, 2017.

[35] M. Picollo, C. Cucci, A. Casini, and L. Stefani, "Hyper-spectral imaging technique in the cultural heritage field: New possible scenarios," *Sensors*, vol. 20, no. 10, p. 2843, 2020.

[36] S. George *et al.*, "A study of spectral imaging acquisition and processing for cultural heritage," in *Digital Techniques for Documenting and Preserving Cultural Heritage* (A. Bentkowska-Kafel and L. MacDonald, eds.), pp. 189–207, Cambridge University Press, 2021.

[37] J. Y. Hardeberg *et al.*, "Hyperspectral imaging in art conservation." `https://surfaceoptics.com/hyperspectral-imaging-in-art-conservation/`, 2020. Accessed: 2024-06-28.

[38] H. Liang, "Advances in multispectral and hyperspectral imaging for archaeology and art conservation," *Applied Physics A*, vol. 106, no. 2, pp. 309–323, 2012.

[39] P. Kubelka and F. Munk, "Ein beitrag zur optik der farbanstriche," *Zeitschrift für technische Physik*, vol. 12, pp. 593–601, 1931.

[40] W. A. Christens-Barry, K. Boydston, F. G. France, K. T. Knox, R. L. J. Easton, and M. B. Toth, "Camera system for multispectral imaging of documents," in *Proceedings of SPIE 7249, Sensors, Cameras, and Systems for Industrial/Scientific Applications X*, p. 724908, 2009.

[41] E. Adamopoulos and F. Rinaudo, "Combining multiband imaging, photogrammetric techniques, and foss gis for affordable degradation mapping of stone monuments," *Buildings*, vol. 11, no. 7, p. 304, 2021.

[42] J. Powers, D.-M. Smilgies, E. C. Geil, K. Clinton, N. Dimitrova, M. Peachin, and R. E. Thorne, "X-ray fluorescence imaging analysis of inscription provenance," *Journal of Archaeological Science*, vol. 36, no. 2, pp. 343–350, 2009.

[43] G. H. Bearman and S. I. Spiro, "Archaeological applications of advanced imaging techniques," *The Biblical Archaeologist*, vol. 59, no. 1, pp. 56–66, 1996.

[44] A. Suri, M. Gupta, and T. Adlakha, "Information extraction from historical epigraphs using digital image processing techniques," in *Proceedings of the IEEE International Conference on Signal Processing and Communication (ICSC)*, October 2016.

[45] M. M. Magrina, "Convolution neural network based ancient tamil character," *International Research Journal of Engineering and Technology*, 2020.

[46] IBM, "Ibm," 2017. [Online].

[47] V. Bharadwaja, R. Ananmy, N. Sarraf, K. V. Vineetha, J. Shah, and D. G. Kurup, "Implementation of artificial neural network on raspberry pi for signal processing applications," 2018.

[48] S. Ruwanmini, K. Dias, C. Niluckshini, and T. Nandasara, "Sinhala inscription character recognition model using deep learning technologies," *International Journal on Advances in ICT for Emerging Regions*, 2023.

[49] S. Albawi and T. A. Mohammed, "Understanding of a convolutional neural network," in *Proceedings of the International Conference on Engineering and Technology (ICET)*, (Antalya, Turkey), 2017.

[50] L. Neumann and J. Matas, "A method for text localization and recognition in real-world images," in *Proceedings of the Asian Conference on Computer Vision (ACCV)*, (Beijing, China), Springer, 2011.

[51] A. Khan and U. Vaidya, "Inscriptions: As a source of history," 2020. [Online].

[52] S. M. Aswatha, A. N. Talla, J. Mukhopadhyay, and P. Bhowmick, "A method for extracting text from stone inscriptions using character spotting," in *ACCV 2014 Workshops, Revised Selected Papers, Part II*, vol. 9009 of *Lecture Notes in Computer Science*, pp. 598–611, Springer International Publishing, 2015.

[53] T. M. Lillesand, R. W. Kiefer, and J. W. Chipman, *Remote Sensing and Image Interpretation*. Hoboken: Wiley, 7th ed., 2015.

[54] J. R. Jensen, *Introductory Digital Image Processing: A Remote Sensing Perspective*. Boston: Pearson, 4th ed., 2015.

[55] M. K. Suganthi and A. Anitha, "Review of semantic segmentation of medical images using modified architectures of unet," *MDPI Open Access Journals*, vol. 12, no. 12, 2022. Accessed: 2025-04-27.

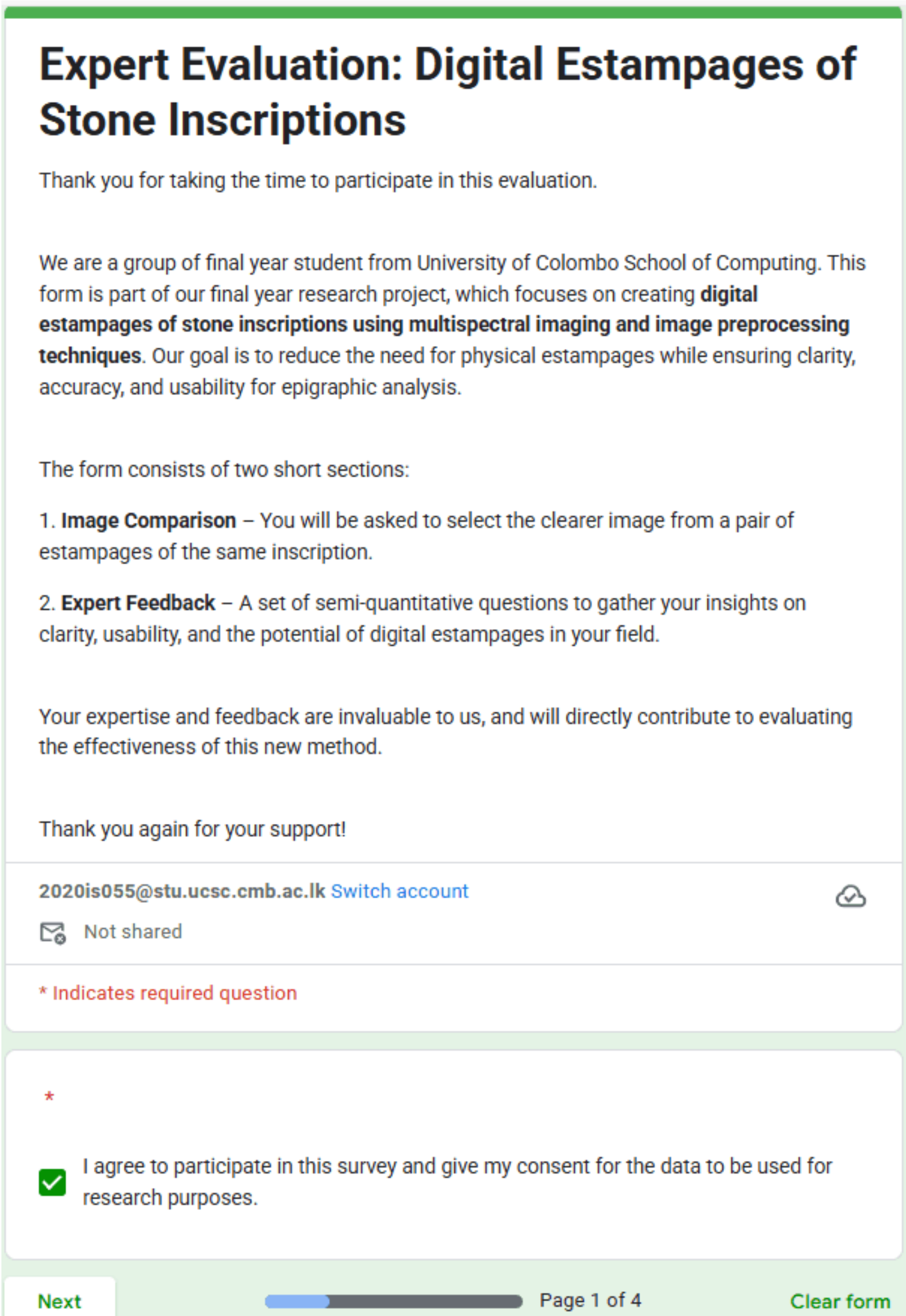
[56] M. F. Haque, H.-Y. Lim, and D.-S. Kang, "Object detection based on vgg with resnet network," in *IEEE Conference Proceedings*, (Auckland, New Zealand), IEEE, 2019.

Appendices

Appendix A – Expert Evaluation Questionnaire

This appendix presents the full questionnaire administered via Google Forms to domain experts. It includes all sections including demographic information, paired-image clarity comparison, semi-quantitative evaluation scales, and an open-ended comments field—used to gather data.

Below landing page is used to give the users an idea on what the form is about, and to get consent on collecting data.



Expert Evaluation: Digital Estampages of Stone Inscriptions

Thank you for taking the time to participate in this evaluation.

We are a group of final year student from University of Colombo School of Computing. This form is part of our final year research project, which focuses on creating **digital estampages of stone inscriptions using multispectral imaging and image preprocessing techniques**. Our goal is to reduce the need for physical estampages while ensuring clarity, accuracy, and usability for epigraphic analysis.

The form consists of two short sections:

1. **Image Comparison** – You will be asked to select the clearer image from a pair of estampages of the same inscription.

2. **Expert Feedback** – A set of semi-quantitative questions to gather your insights on clarity, usability, and the potential of digital estampages in your field.

Your expertise and feedback are invaluable to us, and will directly contribute to evaluating the effectiveness of this new method.

Thank you again for your support!

2020is055@stu.ucsc.cmb.ac.lk Switch account

Not shared

* Indicates required question

*

I agree to participate in this survey and give my consent for the data to be used for research purposes.

Next | Page 1 of 4 | Clear form

Figure 1: Landing page of the form

Section 1 – Participant Information Collects basic demographics and professional background (institution, role, years of experience, prior estampage use) to contextualize expert feedback.

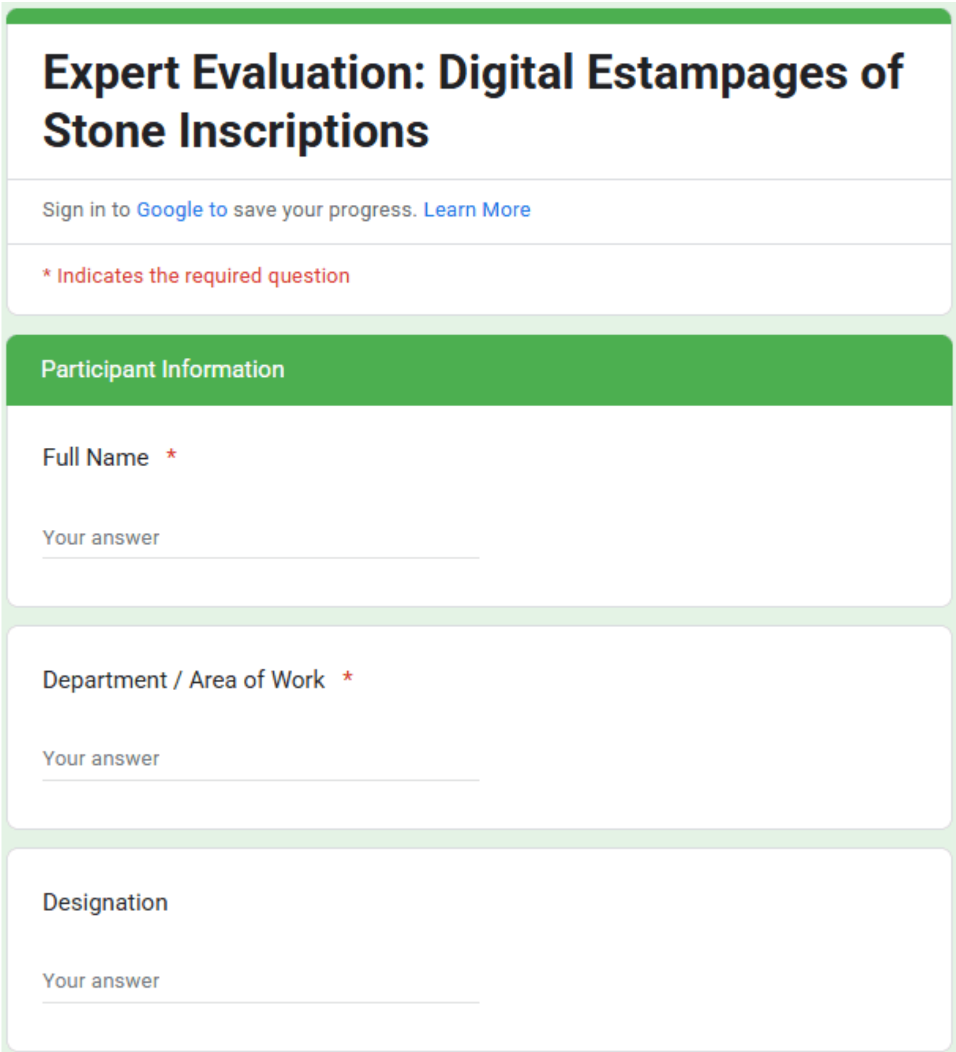
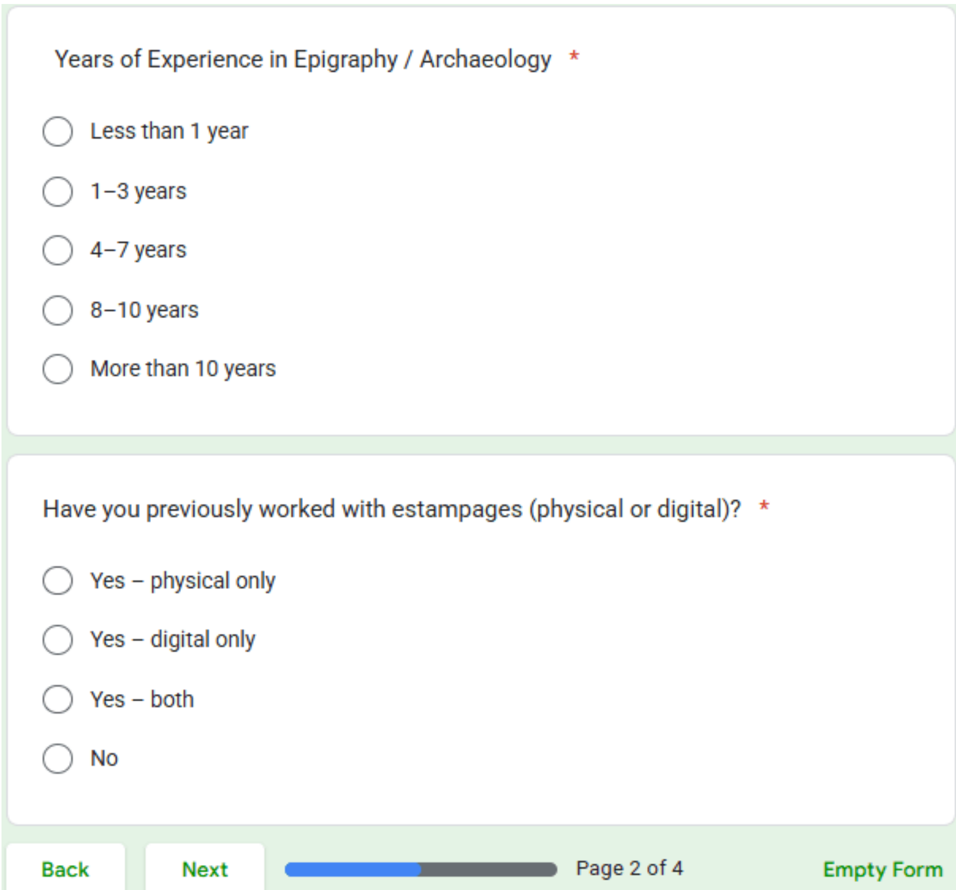


Figure 2: Section 1 – Participant Information



Years of Experience in Epigraphy / Archaeology *

- Less than 1 year
- 1–3 years
- 4–7 years
- 8–10 years
- More than 10 years

Have you previously worked with estampages (physical or digital)? *

- Yes – physical only
- Yes – digital only
- Yes – both
- No

Back | Next | Page 2 of 4 | Empty Form

Figure 3: Section 1 – Participant Information (cont.)

Section 2 – Image Comparison Task Displays two estampage images (one traditional physical, one digital) and asks experts to select the version they find clearest for epigraphic reading.

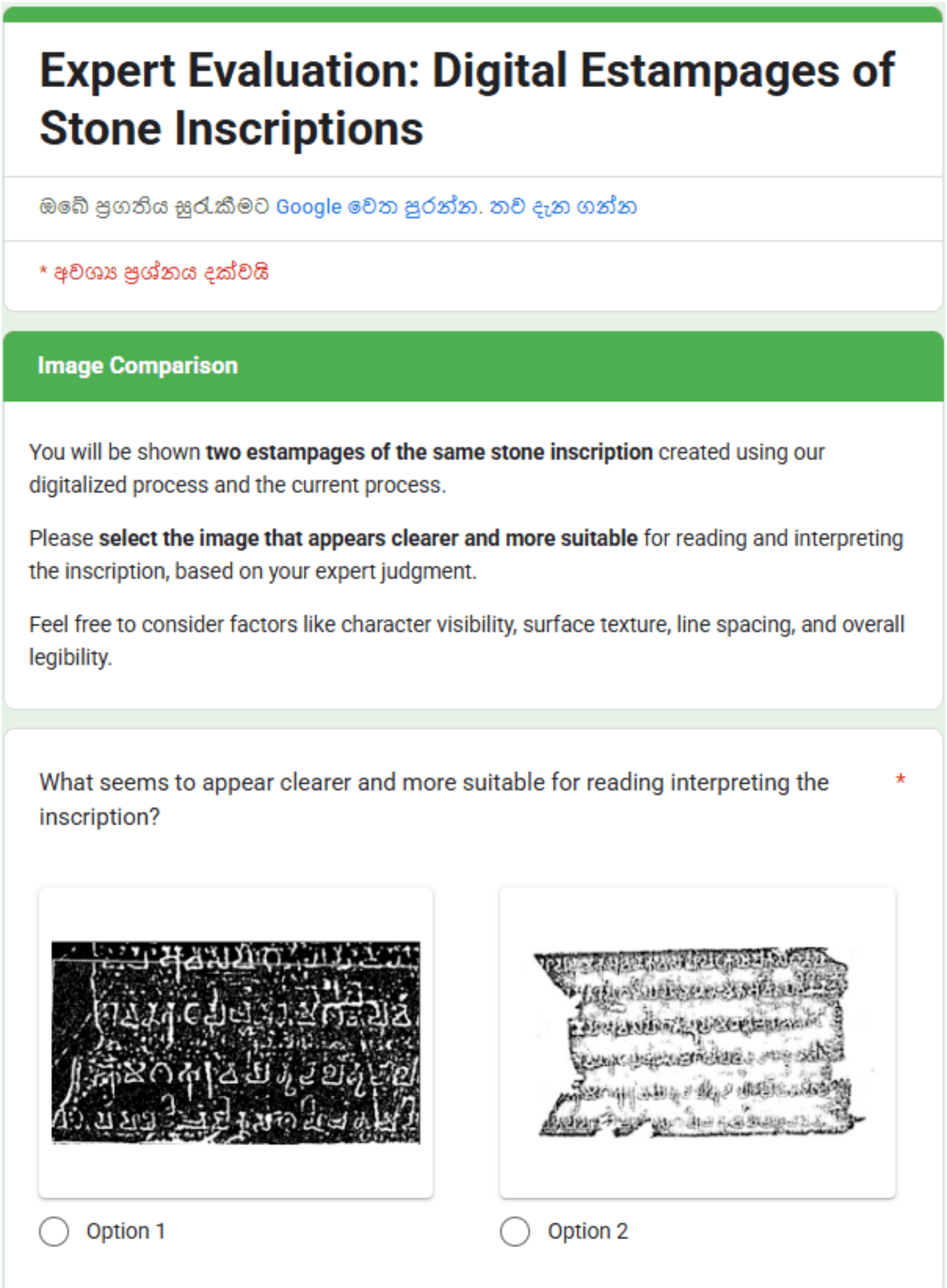


Figure 4: Section 2 – Image Comparison Task

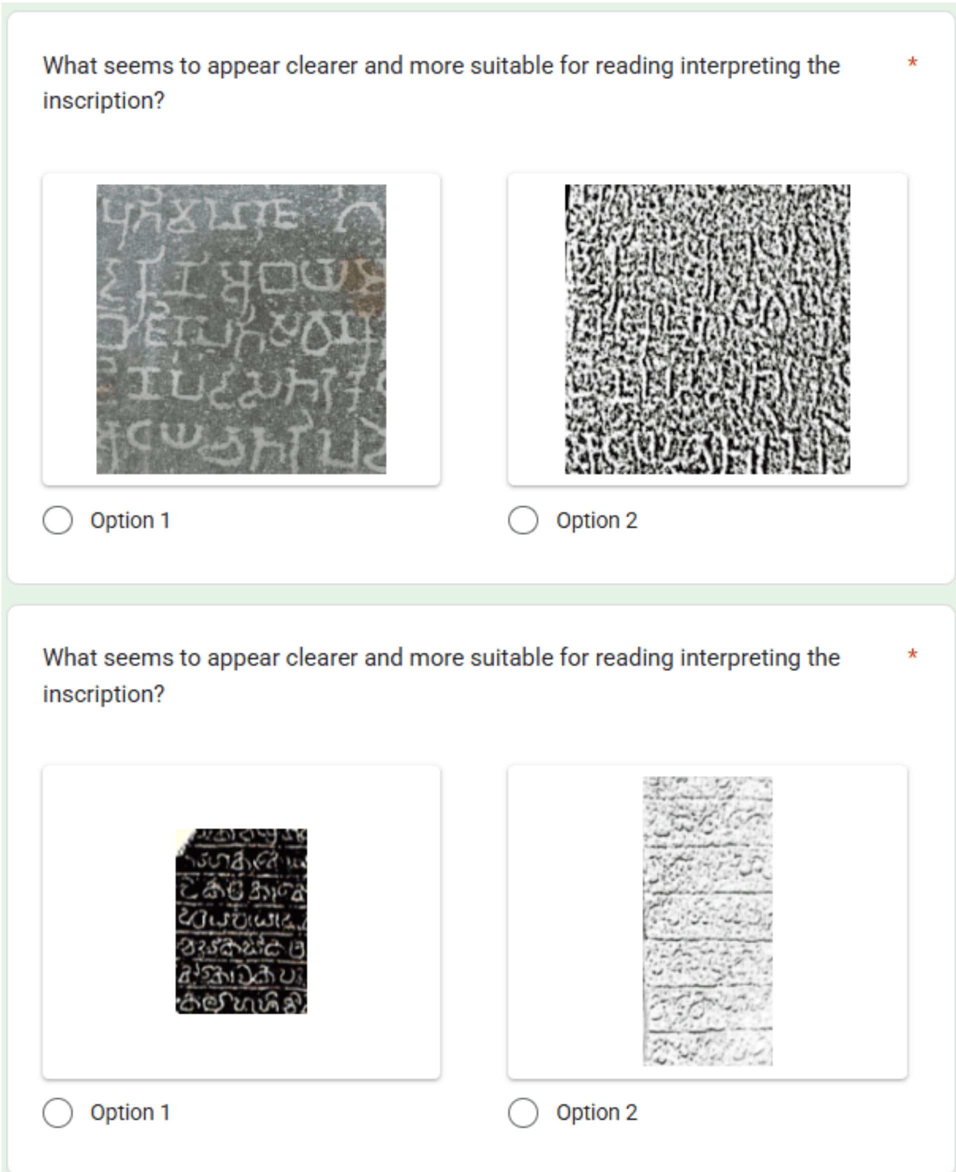


Figure 5: Section 2 – Image Comparison Task(cont.)

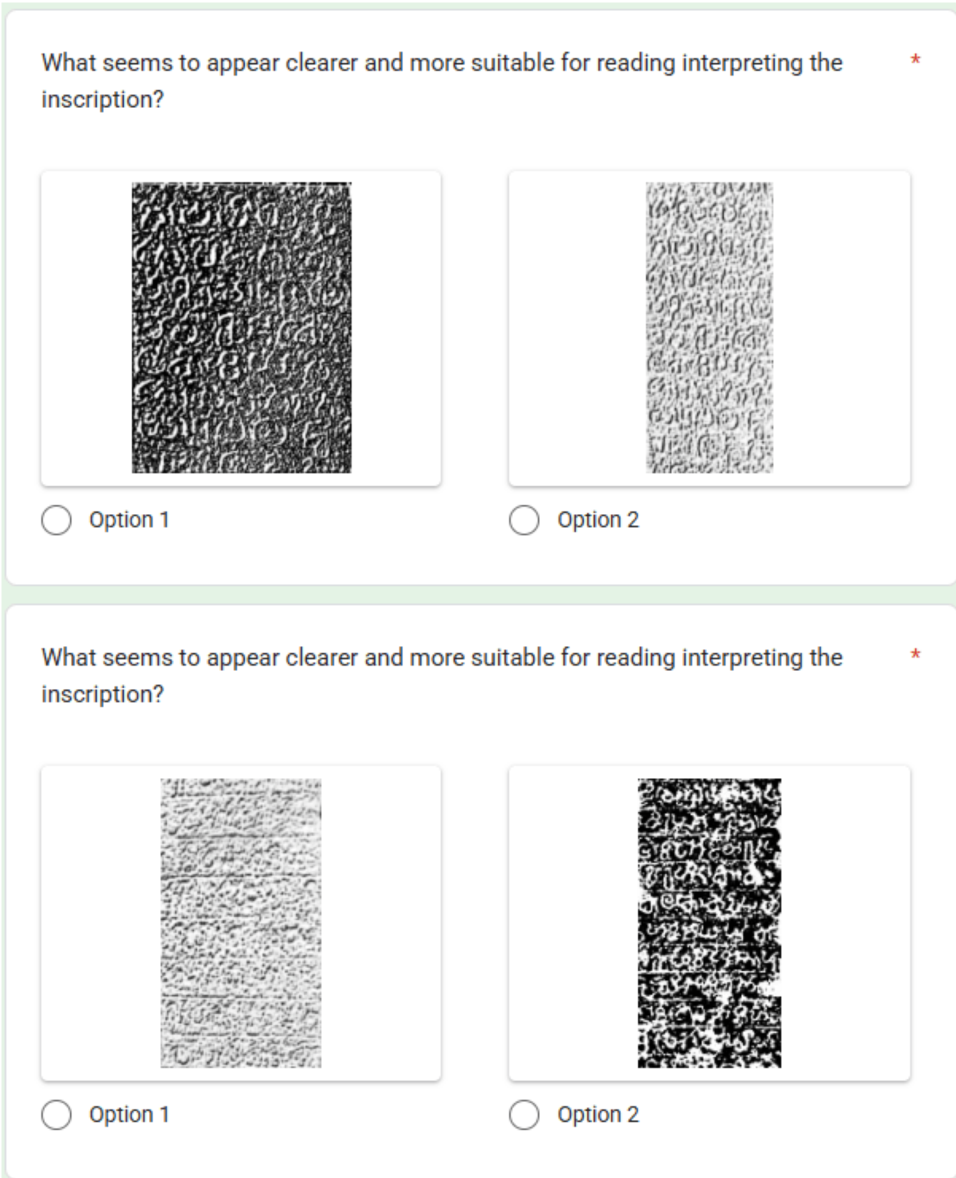


Figure 6: Section 2 – Image Comparison Task(cont.)

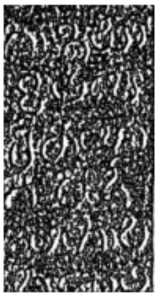
What seems to appear clearer and more suitable for reading interpreting the inscription?

◯ Option 1

◯ Option 2

What seems to appear clearer and more suitable for reading interpreting the inscription?

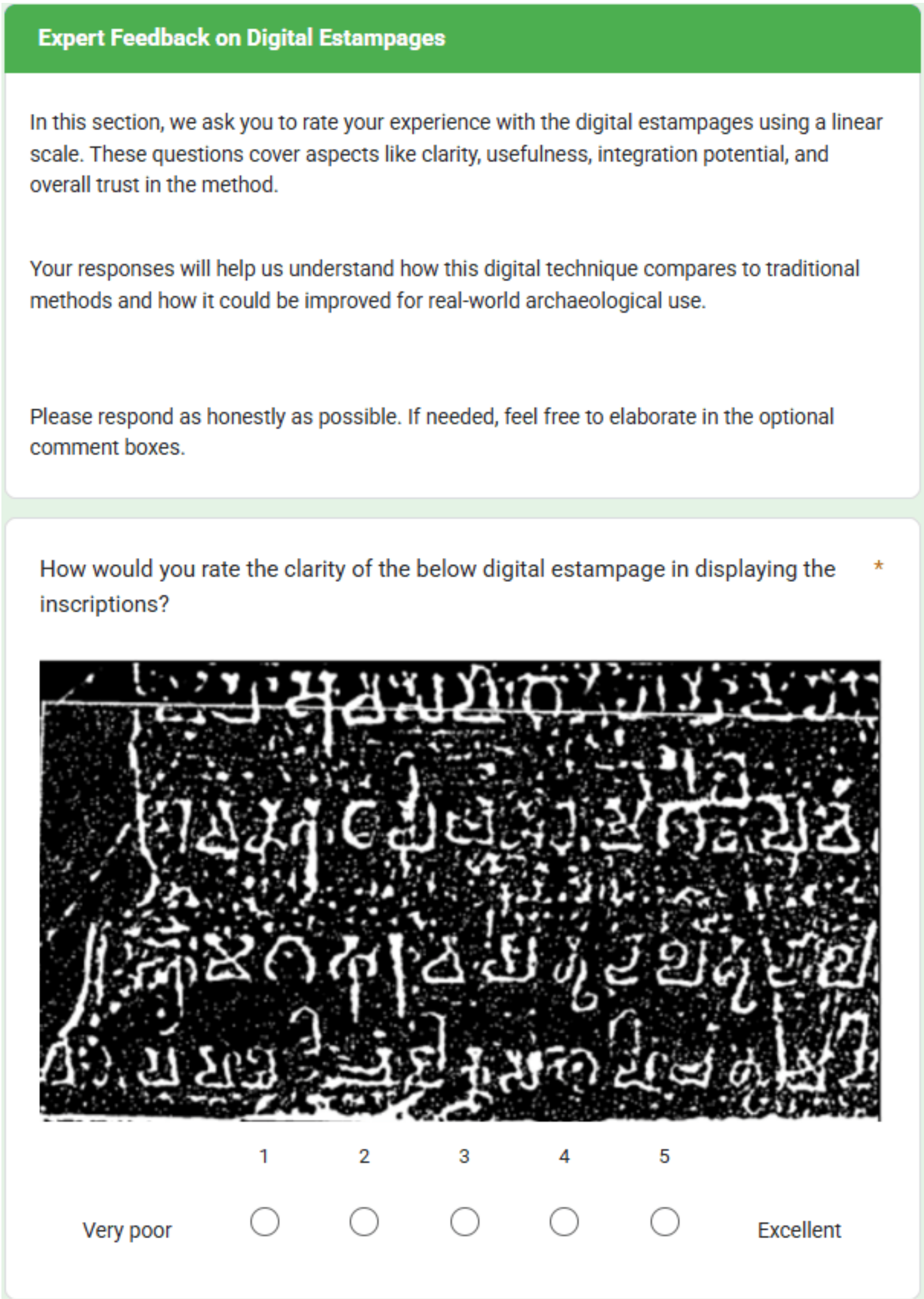
◯ Option 1



◯ Option 2

Figure 7: Section 2 – Image Comparison Task(cont.)

Section 3 – Semi-Quantitative Expert Feedback Presents Likert-scale questions (1–5) on clarity, feature fidelity, research usefulness, trustworthiness, workflow integration, time efficiency, and overall evaluation of the digital estampage method.



Expert Feedback on Digital Estampages

In this section, we ask you to rate your experience with the digital estampages using a linear scale. These questions cover aspects like clarity, usefulness, integration potential, and overall trust in the method.

Your responses will help us understand how this digital technique compares to traditional methods and how it could be improved for real-world archaeological use.

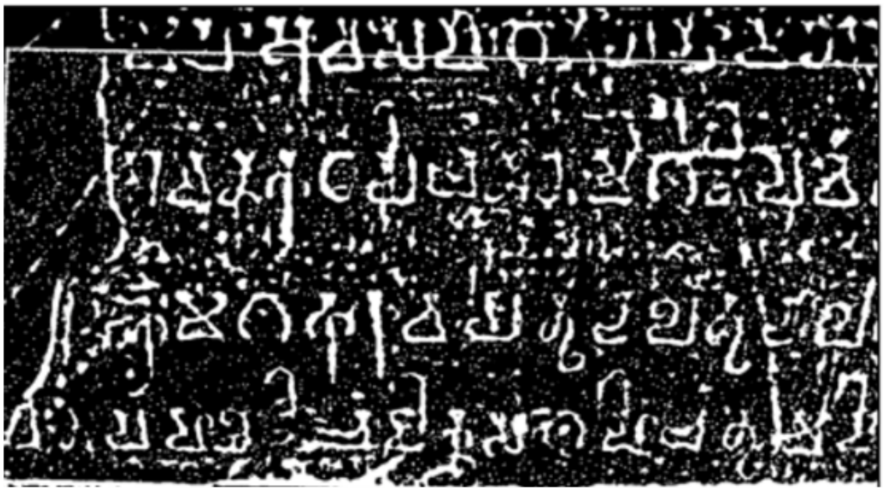
Please respond as honestly as possible. If needed, feel free to elaborate in the optional comment boxes.

How would you rate the clarity of the below digital estampage in displaying the inscriptions? *

	1	2	3	4	5	
Very poor	○	○	○	○	○	Excellent

Figure 8: Section 3 – Semi-Quantitative Expert Feedback

How well does the digital estampage represent the physical features of the stone *
(e.g., letters, symbols)?



	1	2	3	4	5	
Very poor	○	○	○	○	○	Excellent

How useful would this digital estampage be for conducting epigraphic analysis? *

	1	2	3	4	5	
Not useful	○	○	○	○	○	Very useful

Figure 9: Section 3 – Semi-Quantitative Expert Feedback(cont.)

How much would you trust a digital estampage for scholarly publication or formal documentation? *

	1	2	3	4	5	
Not at all	○	○	○	○	○	Completely

How easy would it be to integrate this method into your current archaeological or epigraphic workflow in future? *

	1	2	3	4	5	
Very difficult	○	○	○	○	○	Very easy

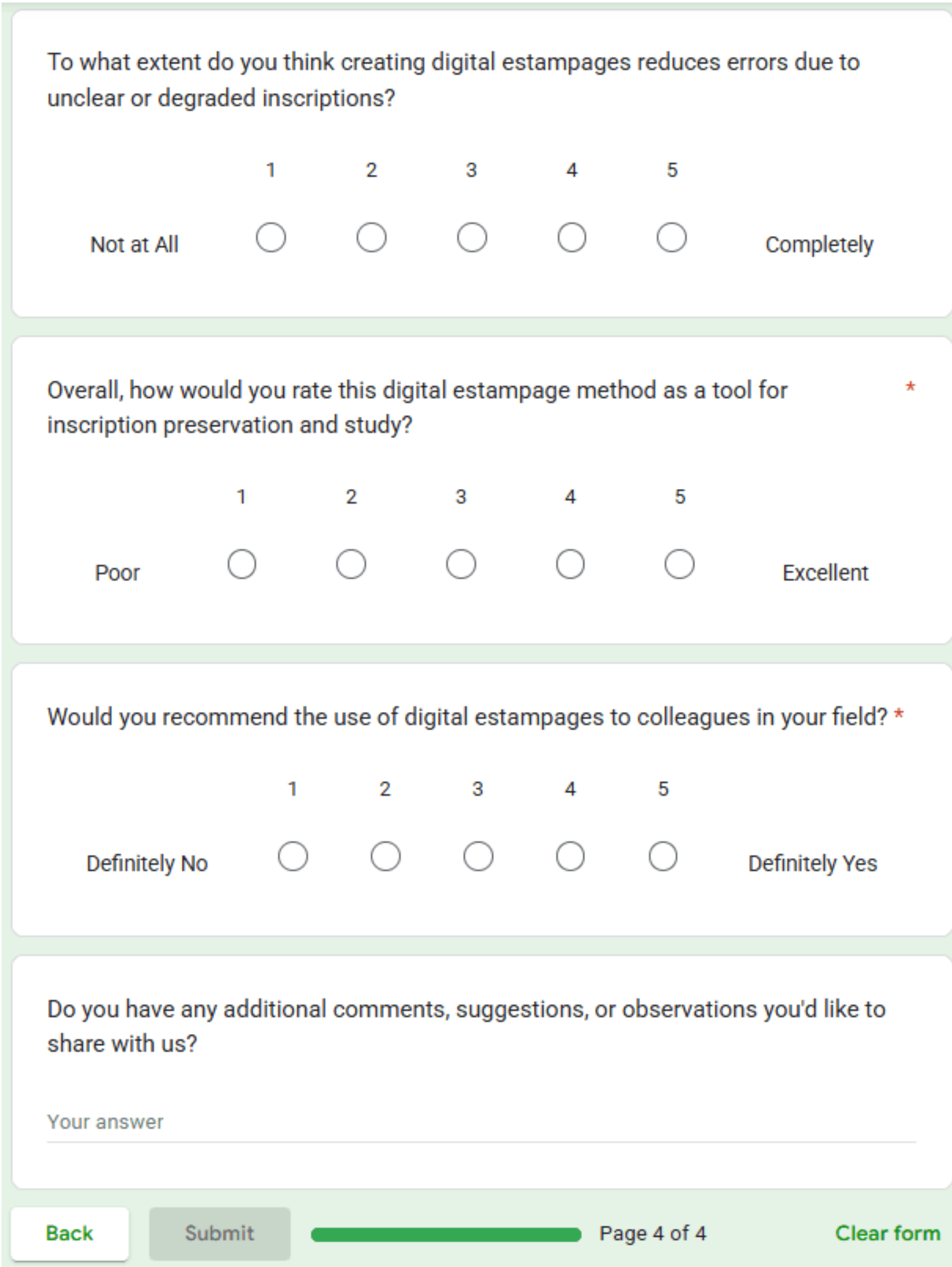
How much training or technical support do you think would be needed to use this system effectively if integrated? *

	1	2	3	4	5	
None	○	○	○	○	○	Extensive training needed

Compared to traditional estampage methods, how would you rate the time efficiency of this digital process?

	1	2	3	4	5	
Much Slower	○	○	○	○	○	Much Faster

Figure 10: Section 3 – Semi-Quantitative Expert Feedback(cont.)



To what extent do you think creating digital estampages reduces errors due to unclear or degraded inscriptions?

	1	2	3	4	5	
Not at All	○	○	○	○	○	Completely

Overall, how would you rate this digital estampage method as a tool for inscription preservation and study? *

	1	2	3	4	5	
Poor	○	○	○	○	○	Excellent

Would you recommend the use of digital estampages to colleagues in your field? *

	1	2	3	4	5	
Definitely No	○	○	○	○	○	Definitely Yes

Do you have any additional comments, suggestions, or observations you'd like to share with us?

Your answer

Back Submit Page 4 of 4 Clear form

Figure 11: Section 3 – Semi-Quantitative Expert Feedback(cont.)