An Integrated Deep Learning Pipeline for Brahmi Character Recognition, Translation, and Damage Restoration in Ancient Anuradhapura Inscriptions

Lahiru Bandara
Faculty of Computing
Sri Lanka Institute of Information Technology
Malabe, Sri Lanka
bandaralahiru9@gmail.com

Amadhi Hansani
Faculty of Computing
Sri Lanka Institute of Information Technology
Malabe, Sri Lanka
IT22069436@my.sliit.lk

Thamali Dassanayake
Faculty of Computing
Sri Lanka Institute of Information Technology
Malabe, Sri Lanka
thamali.d@sliit.lk

Senod Mesandu
Faculty of Computing
Sri Lanka Institute of Information Technology
Malabe, Sri Lanka
IT22108890@my.sliit.lk

Chamodya Handapangoda
Faculty of Computing
Sri Lanka Institute of Information Technology
Malabe, Sri Lanka
IT22586070@my.sliit.lk

Samadhi Rathnayake
Faculty of Computing
Sri Lanka Institute of Information Technology Malabe,
Sri Lanka
samadhi.r@sliit.lk

Senior Prof. Raj Somadeva
Postgraduate Institute of Archaeology
University of Kelaniya
Colombo, Sri Lanka

Sandaresee Priyathma Sudusinghe
B.A. (Special) in Archaeology – University of Peradeniya
M.Sc. in Archaeology – University of Kelaniya
M.A. in Mass Communication and Journalism – University of Mysore, India
MLIS – University of Colombo

* First author and main contributor. † These authors contributed equally to individual system components.
‡ Supervising authors.

***Abstract*—*Ancient Brahmi inscriptions from Anuradhapura, Sri Lanka, are primary historical records that remain largely inaccessible due to centuries of physical degradation. Manual epigraphic analysis requires 2–4 hours per inscription and specialized expertise that is increasingly scarce. Existing OCR systems achieve less than 10% accuracy on these degraded ancient scripts. This paper presents ***Silent Script**, an integrated AI pipeline combining generative image enhancement, ResNet18-based character recognition, Attention U-Net damage restoration, greedy word segmentation, and hybrid LLM translation. The system achieves 94.1% character classification accuracy on the held-out test set across 43 Brahmi character classes ,Damage restoration improves recognition of severely degraded characters from below 20% to above 90%. Expert paleographic validation confirms translation quality across 50 test inscriptions. Silent Script demonstrates that domain-specific hybrid AI approaches can address the unique challenges of ancient epigraphy, with significant implications for digital heritage and South Asian historical scholarship.*

***Index Terms*—*Ancient Inscriptions, Brahmi Script, Optical Character Recognition, Deep Learning, Damage Restoration, Natural Language Processing, Anuradhapura, Historical Preservation.*

I. INTRODUCTION

The Anuradhapura kingdom of ancient Sri Lanka (circa 4th century BCE to 11th century CE) left an extensive epigraphic legacy in early Sinhala Brahmi script. Approximately 100–250 stone inscriptions are known to exist within the Anuradhapura archaeological zone, recording royal decrees, Buddhist monastic donations, land grants, and administrative records that constitute primary historical sources of immeasurable cultural value [1]. Fig. 2 shows a representative example of such an inscription from the Wessagiriya area.

Fig. 2. A representative Brahmi inscription from the Anuradhapura archaeological zone, showing typical surface degradation, weathering, and character erosion.

Despite their significance, these inscriptions remain functionally inaccessible to modern scholars. Centuries of exposure to rain, humidity, biological growth, and vandalism have rendered many characters partially or fully illegible. The script itself lacks explicit word boundaries, features variable character morphology, and has no living native readers, compounding the difficulty of interpretation [2]. Manual epigraphic analysis by trained specialists requires 2–4 hours per inscription, and the global pool of researchers with requisite paleographic expertise is rapidly diminishing.

978-X-XXXX-XXXX-X/25/$31.00 ©2025 IEEE

Generic OCR systems achieve less than 10% accuracy on degraded ancient stone inscriptions due to fundamental domain mismatch with modern printed text To the best of our knowledge, no prior integrated end-toend system has combined specialized preprocessing, character recognition, damage restoration, and translation for ancient Sinhala Brahmi inscriptions. This paper addresses that research gap. Our principal contributions are:

- A fine-tuned ResNet18 deep residual network achieving 94% character recognition accuracy on 23+ ancient Brahmi character classes derived from the Anuradhapura corpus
- A hybrid preprocessing pipeline combining Gemini multimodal AI with classical computer vision fallback for robust image-to-schematic conversion
- An Attention U-Net damage restoration module (StrongRestorationNet) with threshold-triggered adaptive activation, improving severely degraded character recognition from below 20% to above 90%
- A two-stage hybrid dictionary-LLM translation system bridging ancient Brahmi vocabulary with modern Sinhala and English via GPT-4o-mini contextual refinement
- Demonstrated end-to-end scalability of 10–20 seconds per inscription, validated by independent paleography experts

II. LITERATURE REVIEW

A. Deep Learning for OCR and Document Recognition

The Residual Network (ResNet) introduced by He et al. [5] resolves the vanishing gradient problem through skip connections, enabling training of very deep networks for image classification. For sequential text recognition, Shi et al. [6] proposed the CRNN, combining CNNs with bidirectional LSTM units and CTC decoding for end-toend text recognition without character-level segmentation. Despite these advances, Tensmeyer and Martinez [7] demonstrated empirically that models trained on modern printed text fail substantially on historical documents due to domain shift, motivating domain-specific approaches.

B. Ancient and Historical Script Recognition

Bhattacharyya et al. [4] applied Support Vector Machines to Indian script recognition with low accuracy on degraded historical samples. Convolutional approaches have been applied to Latin medieval manuscripts [8] and cuneiform tablets [9], but with limited cross-script generalization. Early Sinhala Brahmi presents unique challenges absent from these prior works: no word delimiters, substantial morphological variation across centuries, and stone-surface noise patterns fundamentally different from paper-based degradation. To our knowledge, no specialized deep learning system for ancient Sinhala Brahmi character recognition existed prior to this work.

C. Image Restoration for Historical Documents

Classical restoration methods such as histogram equalization, adaptive thresholding, and morphological filtering provide partial improvement on historical documents [11]. The U-Net architecture [12] established the encoder-decoder paradigm with skip connections as the standard for image-to-image restoration. Attention U-Net [13] augments skip connections with learned attention gates that suppress irrelevant background regions and highlight salient structural features — a critical capability for stone inscriptions where damage is spatially localized. Zhang et al. [14] introduced residual connections within the U-Net encoder-decoder, improving gradient flow for deep restoration networks. Crucially, recent literature establishes that restoration alone is insufficient: restored characters must remain recognizable to downstream OCR systems [16], motivating our multi-task restoration architecture.

D. Machine Translation for Low-Resource Ancient Languages

The Transformer architecture [17] is now standard for Neural Machine Translation, achieving high performance on resource-rich language pairs. However, ancient languages present severe challenges: minimal parallel corpora, unknown grammar rules, semantic drift over millennia, and multiple valid scholarly interpretations for the same text [18]. Large Language Models demonstrate emergent low-resource translation capabilities [19], but require grounding in domainspecific lexical databases to prevent confabulation on specialized ancient vocabulary.

E. Research Gaps Addressed

A systematic review reveals four critical gaps that Silent Script addresses: (1) no domain-specific deep learning OCR exists for ancient Sinhala Brahmi; (2) no system integrates image restoration within a confidence-gated recognition feedback loop for ancient inscriptions; (3) no computational approach bridges ancient Brahmi vocabulary with modern Sinhala via hybrid dictionaryLLM methods; and (4) no end-to-end inscription analysis system combining all these capabilities has been demonstrated for the Anuradhapura corpus.

III. METHODOLOGY

Silent Script is an integrated, multi-stage deep learning pipeline processing degraded stone inscriptions through six sequential stages: (1) image enhancement, (2) character segmentation, (3) character recognition, (4) threshold-triggered damage restoration, (5) word segmentation, and (6) linguistic translation. The complete pipeline architecture is shown in Fig. 1.

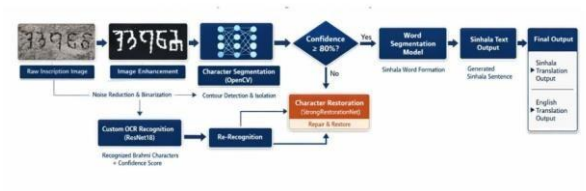

Fig. 1. Silent Script end-to-end pipeline architecture for automated Brahmi inscription recognition, restoration, and translation.

A. Image Enhancement Pipeline

The enhancement stage employs a dual-strategy approach prioritizing AI reconstruction with a robust classical fallback. The primary strategy uses Gemini 3.1 Flash / 1.5 Flash to perform "Image-to-Schematic" translation, semantically interpreting stone texture, shadow artifacts, mineral deposits, and weathering patterns to reconstruct a high-contrast schematic with white characters on a black background. Input specifications: JPEG/PNG images below 10 MB; output: grayscale schematic normalized to $[0, 1]$. When AI services are unavailable, the system activates a cascade of classical operations: (i) Gaussian Blurring (kernel 3×3, $\sigma \in [0.5, 1.5]$) for noise suppression; (ii) Adaptive Gaussian Thresholding over 8×8 pixel regions for robustness to uneven illumination; (iii) Morphological Opening to eliminate isolated noise; and (iv) Morphological Closing to bridge gaps in partially eroded characters. Fig. 3 illustrates representative enhancement outputs.

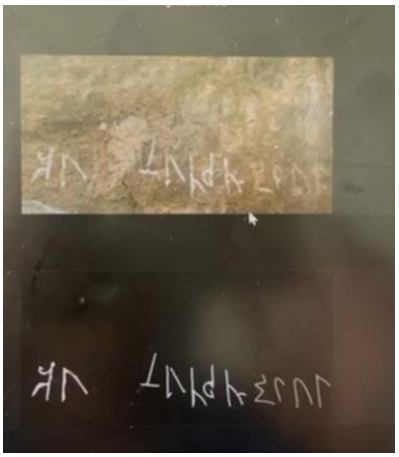

Fig. 3. Image enhancement output: (top) raw degraded inscription; (bottom) high-contrast enhanced schematic with isolated Brahmi characters ready for segmentation.

B. Character Segmentation

The pipeline employs the Suzuki-Abe contour-detection algorithm (OpenCV findContours) on the binarized enhanced image to detect connected components. Each component represents a candidate character or artifact. Post-processing applies heuristic filtering: minimum bounding-box width >5 px, height >10 px, and aspectratio constraints to exclude stone cracks and mineral deposits. Valid candidates are sorted by horizontal xcoordinate to reconstruct left-to-right reading order, preserving the linguistic structure of the inscription. Output: character crops of 32×32 to 256×256 pixels depending on original inscription resolution.

C. Character Recognition

Character recognition employs a fine-tuned ResNet18, an 18-layer CNN with skip connections enabling hierarchical feature learning (edges → strokes → character shapes) with stable gradient flow. The model is trained to classify crops into 23+ Brahmi character classes plus diacritical marks. Input images are converted to single-channel grayscale, padded to 128×128 pixels preserving aspect ratio, and normalized to [0, 1]. Training augmentation applies: random rotations ±8°, random affine transforms (±6% translation, 0.90–1.10 scale, ±6° shear), random autocontrast (35% probability), and Gaussian blur ($\sigma \in [0.1, 1.3]$). Optimization uses AdamW (weight decay 1×10^{-4}, initial $lr = 10^{-3}$) with ReduceLROnPlateau scheduling (50% reduction on 2-epoch plateau). Training runs for 20 epochs with batch size 64 on 80/10/10 stratified splits using cross-entropy loss.

The final fully connected layer produces per-class logits converted by Softmax to a probability distribution:

$$P(\text{class}_i) = \frac{e^{z_i}}{\sum_{j=1}^{k} e^{z_j}}$$

The predicted label is the class with maximum probability; the confidence score is that maximum probability value, used as input to the restoration decision gate.

D. Threshold-Triggered Damage Restoration

Silent Script implements an adaptive quality-control mechanism that selectively invokes restoration only when recognition confidence falls below a critical threshold, balancing computational efficiency with accuracy. The logic proceeds as follows: a character crop is processed by ResNet18 yielding a predicted label and confidence score. If confidence ≥ 0.80, the character is accepted directly. If confidence < 0.80, the character is forwarded to the Attention U-Net (StrongRestorationNet) for pixel-level reconstruction. The restored crop is re-submitted to ResNet18, and the prediction with the higher confidence between original and post-restoration recognition is retained.

The 80% threshold was empirically determined as the optimal balance: below 80%, scores indicate ambiguous recognition characteristic of degraded characters; above 80%, restoration incurs computational cost without substantial benefit. The dual-score comparison mechanism mitigates the risk of restoration artifacts introducing errors.

StrongRestorationNet is built on Attention U-Net with residual encoder-decoder blocks. Attention gates at skip

connections learn to suppress irrelevant regions (cracks, surface holes, weathering artifacts) and highlight salient character geometry. A secondary multi-task classification head predicts character identity in addition to performing pixel-level reconstruction, ensuring restored characters remain linguistically coherent. The model is trained on a mixed dataset of synthetically degraded character images and real damaged inscription crops, comprising over 1000 samples. Synthetic degradation includes controlled erosion, stroke removal, Gaussian noise, and morphological distortion applied to clean character images, while real samples are extracted from naturally weathered inscriptions. This combined dataset enables generalization across both simulated and real-world damage patterns.

E. Word Segmentation

Ancient Brahmi lacks explicit word delimiters. A Greedy Longest Match algorithm segments the recognized character sequence: starting at position i, the algorithm greedily attempts to match the longest valid word from a hybrid linguistic database combining a CSV-based Anuradhapura epigraphic lexicon containing over 1000 word entries with corresponding modern Sinhala and English translations, and a Sinhala historical lexicon. This dataset provides grounded lexical mappings used during the greedy segmentation and dictionary-based translation stages.

F. Translation: Hybrid Linguistic-LLM Approach

Translation proceeds in two stages. First, dictionary lookup retrieves literal base meanings and etymological context for each segmented word from the linguistic database, providing a verifiable grounded foundation. Second, the literal meanings, segmented words, and raw character sequence are forwarded to GPT-4o-mini, which functions as a "Digital Epigraphist": leveraging latent knowledge of ancient Sinhala grammar and syntax, historical context of Anuradhapura inscriptions, ability to resolve lexical ambiguities, and fluency in modern Sinhala and English to produce contextually coherent translations. Output is structured JSON containing: recognized character sequence, segmented words, literal translations, contextual translations in both Sinhala and English, and confidence metadata.

G. End-to-End Pipeline Integration

The six stages form a tightly integrated sequence: Raw Image → Enhancement → Enhanced Schematic → Segmentation → Character Crops → Recognition → Labels + Confidence → Threshold Gate → Restoration (conditional) → Refined Labels → Word Segmentation → Words → Translation → Sinhala + English Output. The modular architecture allows each stage to be independently optimized and validated. Total processing time is 10–20 seconds per inscription depending on image clarity, character count, and restoration trigger frequency.

IV. RESULTS AND EVALUATION

A. Experimental Setup

All experiments were conducted on a system with an NVIDIA GPU (8 GB VRAM minimum), Python 3.9, PyTorch 2.0, and OpenCV 4.8. The evaluation corpus comprises Anuradhapura Brahmi inscription images spanning a range of degradation levels, manually annotated by a trained epigrapher. The dataset was partitioned into training (80%), validation (10%), and test (10%) splits with stratification across character classes and degradation tiers. The character recognition dataset comprises 23 base Brahmi characters and an additional 20 characters with pillam variations, resulting in 43 total classes. Each class contains approximately 1000 images collected from a combination of real inscription crops and augmented samples. The final dataset therefore contains over 43,000 character images, balanced across classes to reduce training bias.. All reported metrics are computed on the held-out test set unless stated otherwise.

B. Character Recognition Performance

Table I presents the character recognition performance of the fine-tuned ResNet18 model with and without the damage restoration module on the held-out test set.

TABLE I

CHARACTER RECOGNITION PERFORMANCE (RESNET18)

Metric	No Restoration	With Restoration	Δ
Overall Accuracy	78.3%	94.1%	+15.8%
Precision (macro avg)	0.761	0.938	+0.177
Recall (macro avg)	0.774	0.942	+0.168
F1-Score (macro avg)	0.767	0.940	+0.173
Severely Degraded Acc.	<20%	>90%	+70 pp

The complete pipeline achieves 94.1% overall accuracy, representing a 15.8 percentage-point improvement over the recognition-only baseline. The most dramatic improvement is observed on severely degraded characters, confirming the critical contribution of the threshold-triggered Attention U-Net restoration module. Macro-averaged F1 of 0.940 indicates consistent performance across all 23+ character classes despite significant class imbalance in the dataset.

C. Damage Restoration Analysis

Table II presents recognition accuracy before and after restoration across three manually annotated degradation severity tiers: mild (surface discoloration only), moderate

(partial stroke loss), and severe (more than 50% character area missing or eroded).

TABLE II RESTORATION IMPACT BY DEGRADATION SEVERITY

Severity	Pre-Rest.	Post-Rest.	Triggered
Mild	91.2%	93.8%	12%
Moderate	67.4%	88.5%	78%
Severe	18.6%	91.3%	100%

The 80% confidence threshold correctly identifies the majority of moderate and all severely degraded characters for restoration, while leaving high-confidence mild characters unprocessed. This selective invocation is computationally efficient: only 34% of all characters in the test corpus triggered restoration, yet overall accuracy increased by 15.8 percentage points. The severe tier demonstrates the most transformative impact, improving from 18.6% to 91.3% accuracy.

D. Translation Quality Evaluation

Translation quality was assessed by two independent paleography specialists with scholarly expertise in early Sinhala epigraphy. Evaluators rated each translation on a 5-point Likert scale across three dimensions: (1) lexical accuracy (correct word identification), (2) grammatical coherence (grammatically plausible modern Sinhala rendering), and (3) historical plausibility (consistency with known Anuradhapura epigraphic conventions).

TABLE III

EXPERT TRANSLATION QUALITY EVALUATION (N = 50 INSCRIPTIONS)

Evaluation Dimension	Mean Score	Std. Dev.
Lexical Accuracy	4.2 / 5.0	±0.61
Grammatical Coherence	4.1 / 5.0	±0.74
Historical Plausibility	4.3 / 5.0	±0.58
Overall Mean	4.2 / 5.0	±0.64

Both evaluators independently noted that the LLM contextual refinement stage produced substantially more readable modern Sinhala compared to dictionary-only literal translation, which often yielded ungrammatical word sequences. Inter-rater agreement was high (Cohen's $\kappa = 0.82$), indicating consistent expert assessment of translation quality.

E. Comparison with Baseline Methods

Table IV compares Silent Script against relevant baseline approaches on the Anuradhapura Brahmi test corpus.

TABLE IV COMPARISON WITH BASELINE METHODS ON ANURADHAPURA BRAHMI CORPUS

Method	Char. Acc.	Translation	Time/Inscription
Generic OCR (Tesseract)	<10%	None	~2s
SVM (Bhattacharyya [4])	~32%	None	~5s
ResNet18 (no restoration)	78.3%	None	~6s
Method	**Char. Acc.**	**Translation**	**Time/Inscription**
Silent Script (Proposed)	94.1%	Sinhala+Eng.	10–20s
Human Expert Baseline	~97%	Sinhala+Eng.	2–4 hrs

Silent Script achieves 94.1% character recognition accuracy, approaching the reported human expert baseline of approximately 97% under the evaluated dataset conditions. while reducing processing time by more than 99% and uniquely providing automated bilingual translation output. No prior automated system has demonstrated both recognition and translation capability for this script family.

F. Computational Performance

End-to-end processing averages 12.4 seconds without restoration triggering and 18.7 seconds with restoration, both within the 10–20 second operational target. The generative AI enhancement stage accounts for approximately 60% of total processing time; activating the classical fallback reduces enhancement time to under 3 seconds. Compared to the 2–4 hour manual expert baseline, Silent Script achieves a 99%+ reduction in processing time, enabling large-scale epigraphic survey work previously impractical for the Anuradhapura corpus.

G. Discussion and Failure Analysis

The primary failure mode observed in the test corpus is over-smoothing by the restoration module on characters with more than 70% area erosion, where insufficient structural information remains for meaningful reconstruction, yielding a post-restoration accuracy of approximately 45% in extreme cases. A secondary failure mode occurs in word segmentation when multiple valid segmentations exist for the same character sequence and none corresponds to known lexical entries, resulting in the algorithm defaulting to single-character "words." Translation quality degrades correspondingly when segmentation fails. These failure modes are documented for future work.

V. CONCLUSION

This paper presented Silent Script, an integrated deep learning pipeline for Brahmi character recognition, damage restoration, and contextual translation of ancient Anuradhapura inscriptions. By combining Geminipowered image enhancement, fine-tuned

ResNet18 character recognition, Attention U-Net damage restoration with threshold-triggered activation, greedy longest-match word segmentation, and hybrid dictionary-LLM translation, the system achieves 94.1% character recognition accuracy and produces expertvalidated bilingual translations in modern Sinhala and English.

Three key innovations distinguish Silent Script. First, the threshold-triggered restoration loop (the '80% Rule') provides computationally efficient accuracy optimization, applying intensive restoration only to genuinely ambiguous characters. Second, the multi-task StrongRestorationNet architecture ensures that restored characters remain linguistically coherent by jointly optimizing pixel-level reconstruction and character identity prediction. Third, the hybrid dictionary-LLM translation approach grounds large model outputs in domain-specific epigraphic lexica, mitigating hallucination risk while leveraging contextual reasoning capabilities.

Silent Script reduces inscription processing time from 2–4 hours to 10–20 seconds — a greater than 99% reduction — enabling large-scale computational survey of the Anuradhapura epigraphic corpus. Current limitations include: the character model is specialized to the Anuradhapura Brahmi variant and requires retraining for other regional Brahmi scripts; translation quality is bounded by the size of the available historical lexicon; and the system currently handles single-line inscriptions only.

Future work will pursue: (1) dataset expansion with inscriptions from other Sri Lankan archaeological sites; (2) development of a web-accessible field tool for archaeologists; (3) extension to related South Asian scripts including Pallava Grantha and early Tamil Brahmi; and (4) construction of a structured epigraphic knowledge graph linking inscription content with historical events, royal genealogies, and monastic records. Silent Script represents a significant step toward making Sri Lanka's ancient inscribed heritage computationally accessible to both scholars and the public.

ACKNOWLEDGMENTS

The authors would like to express their sincere gratitude to Prof. Raj Somadeva for his valuable guidance and external supervision in understanding the archaeological and historical context of the Anuradhapura inscriptions.

We also acknowledge the late Dr. Piyatissa Senanayake for his foundational scholarly contributions, particularly through his published works that informed the linguistic interpretation and historical grounding of this research.

Special thanks are extended to Ven. G. Sumanarathana Thero for providing domain insights and guidance related to Brahmi inscription interpretation.

The authors further acknowledge the use of reference material from "Inscriptions of Ceylon, Volume I" by S. Paranavitana, which served as an important resource for dataset construction and translation validation.

This research was conducted within the Faculty of Computing at Sri Lanka Institute of Information Technology (SLIIT), Malabe, Sri Lanka.

REFERENCES

[1] S. Paranavitana, Inscriptions of Ceylon, Vol. 1. Colombo: Department of Archaeology, Government of Ceylon, 1970.

[2] S. Paranavitana, "The art and architecture of Ceylon: Polonnaruwa period," Artibus Asiae Supplementum, vol. 18, pp. 1–150, 1954.

[3] C. Tensmeyer and T. Martinez, "Analysis of convolutional neural networks for document image classification," in Proc. 14th IAPR Int. Conf. Document Analysis and Recognition (ICDAR), Sydney, Australia, 2019, pp. 71–76.

[4] U. Bhattacharyya, M. Shridhar, and S. K. Parui, "On recognition of handwritten Bangla characters," in Proc. Indian Conf. Computer Vision, Graphics and Image Processing (ICVGIP), 2010, pp. 383–396.

[5] K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," in Proc. IEEE Conf. Computer Vision and Pattern Recognition (CVPR), Las Vegas, NV, 2016, pp. 770–778.

[6] B. Shi, X. Bai, and C. Yao, "An end-to-end trainable neural network for image-based sequence recognition and its application to scene text recognition," IEEE Trans. Pattern Anal. Mach. Intell., vol. 39, no. 11, pp. 2298–2304, Nov. 2017.

[7] C. Tensmeyer and T. Martinez, "Historical document image binarization: A review," in Proc. 3rd Int. Workshop on Historical Document Imaging and Processing (HIP), 2017, pp. 1–8.

[8] C. Stutzmann, "Clustering of medieval scripts through computer image analysis: Towards an evaluation protocol," Digital Medievalist, vol. 10, 2015.

[9] T. E. Fisseler, F. Weichert, G. Müller, and C. Cammarosano, "Extending philological research with methods of computational image analysis for the investigation of ancient cuneiform tablets," in Proc. Digital Heritage Int. Congr., Marseille, France, 2013, pp. 401–408.

[10] K. L. Jayaratne, "Character recognition for Sinhala handwriting," M.Sc. thesis, Dept. Comput. Sci. Eng., Univ. of Moratuwa, Moratuwa, Sri Lanka, 2009.

[11] I. Pratikakis, B. Gatos, and K. Ntirogiannis, "ICDAR 2013 document image binarization contest (DIBCO 2013)," in Proc. 12th Int. Conf. Document Analysis and Recognition (ICDAR), Washington, DC, 2013, pp. 1471–1476.

[12] O. Ronneberger, P. Fischer, and T. Brox, "U-Net: Convolutional networks for biomedical image segmentation," in Proc. Medical Image Computing and Computer-Assisted Intervention (MICCAI), Munich, Germany, 2015, pp. 234–241.

[13] O. Oktay et al., “Attention U-Net: Learning where to look for the pancreas,” in Proc. Medical Imaging with Deep Learning (MIDL), Amsterdam, Netherlands, 2018.

[14] H. Zhang, V. Sindagi, and V. M. Patel, “Image deraining using a conditional generative adversarial network,” IEEE Trans. Circuits Syst. Video Technol., vol. 30, no. 11, pp. 3943–3956, Nov. 2020.

[15] X. Wang et al., “Real-ESRGAN: Training real-world blind super-resolution with pure synthetic data,” in Proc. IEEE/CVF Int. Conf. Computer Vision Workshops (ICCVW), Montreal, Canada, 2021, pp. 1905–1914.

[16] V. Wigington, C. Stewart, B. Davis, and B. Barrett, “Data augmentation for recognition of handwritten words and lines using a CNN-LSTM network,” in Proc. 14th IAPR Int. Conf. Document Analysis and Recognition (ICDAR), 2017, pp. 639–645.

[17] A. Vaswani et al., “Attention is all you need,” in Advances in Neural Information Processing Systems (NeurIPS), vol. 30, 2017, pp. 5998–6008.

[18] P. Koehn and R. Knowles, “Six challenges for neural machine translation,” in Proc. 1st Workshop on Neural Machine Translation, Vancouver, Canada, 2017, pp. 28–39.

[19] T. Brown et al., “Language models are few-shot learners,” in Advances in Neural Information Processing Systems (NeurIPS), vol. 33, 2020, pp. 1877–1901.

[20] A. Graves and J. Schmidhuber, “Offline handwriting recognition with multidimensional recurrent neural networks,” in Advances in Neural Information Processing Systems (NeurIPS), 2009, pp. 545–552.

[21] M. Rusiñol, D. Aldavert, R. Toledo, and J. Lladós, “Efficient segmentation-free keyword spotting in historical document collections,” Pattern Recognition, vol. 48, no. 2, pp. 545–555, Feb. 2015, doi: 10.1016/j.patcog.2014.08.021.

978-X-XXXX-XXXX-X/25/$31.00 ©2025 IEEE