

Advancing Armenian Inscription Recognition

Gevorg Nersesian¹
Suren Khachatryan¹
Narine Sarvazyan^{1,2,3}

¹American University of Armenia
²The George Washington University
³L. A. Orbeli Institute of Physiology NAS RA

This research focuses on the detection and recognition of Armenian epigraphic inscriptions carved in stone, using a combination of classical computer vision and machine learning techniques. The goal is to develop tools for automating the documentation and analysis of Armenian cultural heritage.

1 Introduction

Armenian stone inscriptions, carved over centuries onto khachkars and monuments, represent a vital cultural and historical archive. Despite their significance, the field remains almost entirely unexplored, with only a single academic paper addressing the use of modern methods for their analysis. The deterioration of stone surfaces and the complexity of ancient scripts pose major challenges to traditional documentation efforts.

Problem: Manual transcription is slow, error-prone, and often impossible for inaccessible or damaged sites.

3 Methodology

1. Data Collection

- High-resolution photographs of carved inscriptions and projections from 3D scans of khachkars and tapanakars
- Manual annotation of symbol masks to serve as ground truth

2. Preprocessing Pipeline

- Adaptive Thresholding: Adjusts to varying stone surface conditions
- Morphological Operations: Enhances the structure of carved characters
- Connected Components Analysis: Identifies candidate regions

3. Clustering and Filtering

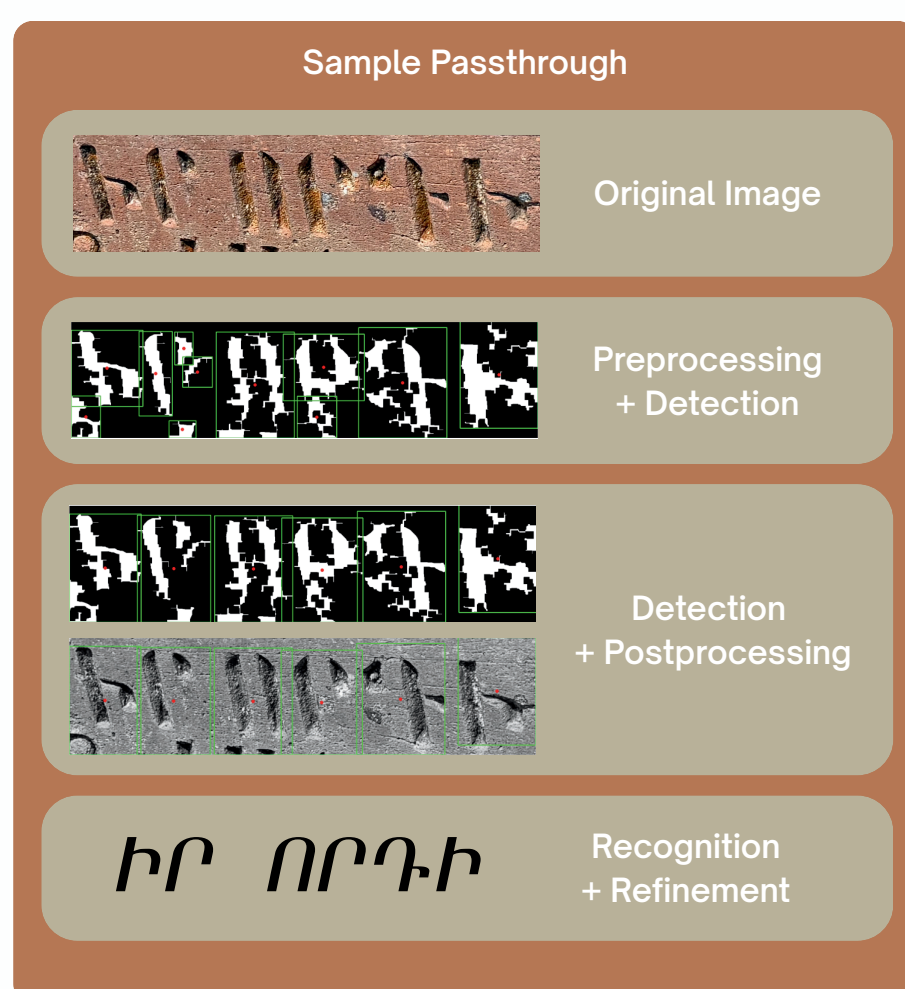
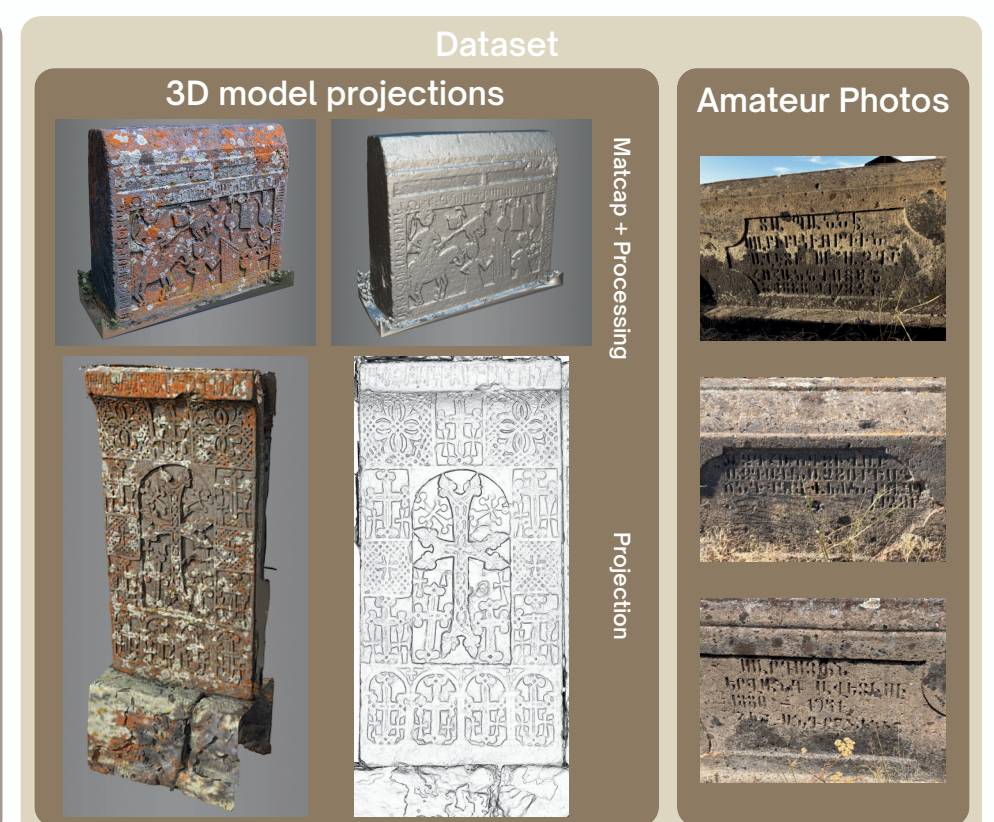
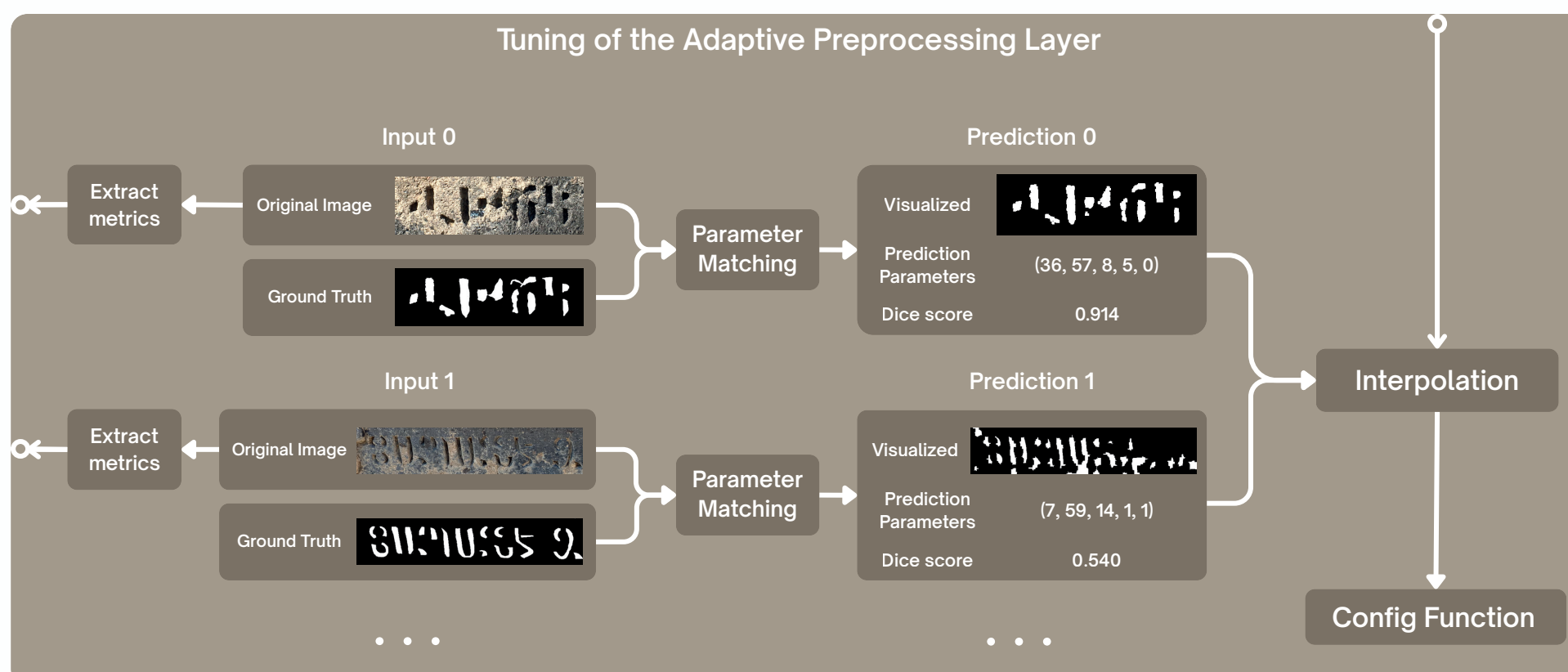
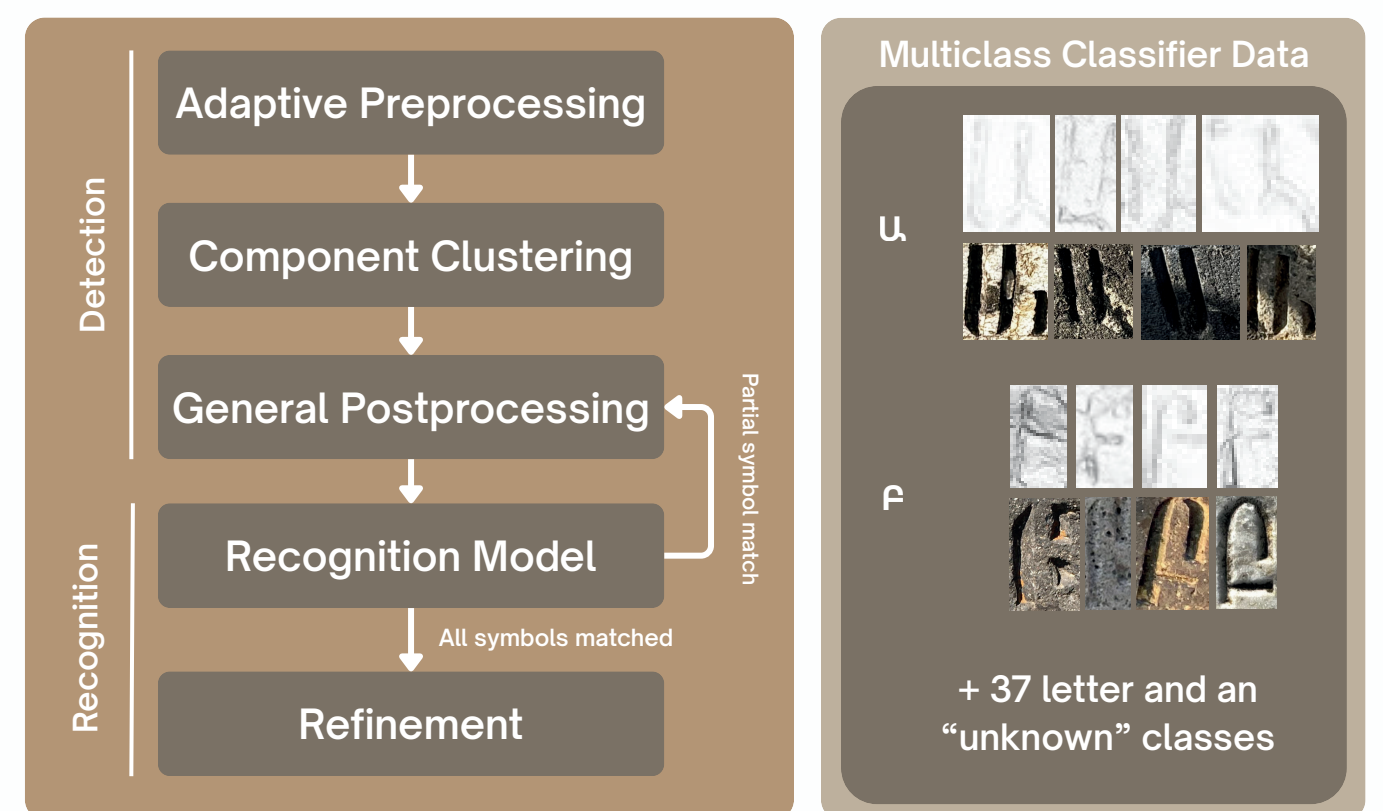
- DBSCAN clustering of components based on spatial and size features
- Filtering of non-symbol artifacts through size, shape, and texture heuristics

4. Recognition

- Binary classification model to separate valid glyphs from noise
- Initial experiments with SVMs, Random Forests, and lightweight CNNs

2 Objective

- Develop a computational framework for detecting and recognizing Armenian stone inscriptions
- Explore classical computer vision and machine learning methods to preprocess, segment, and classify glyphs
- Create preliminary datasets and evaluation pipelines for Armenian epigraphic OCR



4 Results

- Introduced a new annotated dataset for Armenian inscription analysis, combining field photographs with projections from a unique 3D digital-twin collection of khachkars and tapanakars.
- Adaptive preprocessing improved segmentation on full-size patches: Dice rose from 0.63 to 0.77, and IoU from 0.48 to 0.64 across different configurations.
- Local tile-wise optimization performed even better: for grid sizes of 100 / 75 / 50 px, mean Dice reached 0.91 / 0.93 / 0.96, mean **IoU 0.84 / 0.87 / 0.92**, and acceptance rates increased **from 62.2% to 78.7%** as tiles became smaller.
- A prototype recognition stage achieved **~85% accuracy** on the visually confusable subset U, U, Վ, demonstrating the feasibility of extending the framework to the full Armenian alphabet.

5 Conclusion

This work presents a novel and highly **adaptable framework for Armenian inscription recognition**, grounded in a unique combination of field photography, 3D cultural-heritage data, and spatially tuned computer-vision methods. The results show that even in a severely low-resource setting, an elegant adaptive pipeline can robustly segment degraded carved text and support promising early recognition performance. As such, the project lays a serious **technical foundation for large-scale Armenian epigraphic analysis and digital preservation**.

Current and Future Directions:

- 3D reconstruction of Armenian inscribed surfaces to separate carved geometry from erosion and improve glyph detection on worn stones
- Semantic post-processing for text recovery to infer and reconstruct partially lost inscription content

Full Paper



Based on Nersesian et al., Advancing Armenian Inscription Recognition, Digital Heritage 2025

