Artificial Intelligence

Dr. Imran Malik

SCHOOL OF ELECTRICAL ENGINEERING AND COMPUTER SCIENCE

PROJECT REPORT

By

Ghulam Hussain Khan Sherwani 366549

Abdul Basit Sajid 386804

Mohammad Irfan 363768

BSCS11-C

December 22, 2023

Contents

1.Introduction

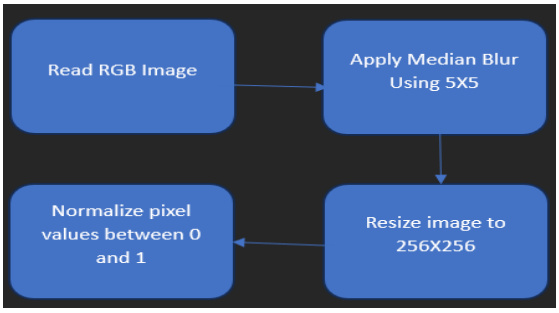
In the face of time's relentless erosion, historical documents stand as cultural touchstones, vulnerable to degradation. The "Binarization of Historical Documents" project addresses this challenge with advanced digital image processing techniques. Focused on enhancing the visibility of foreground text and mitigating image noise, this report outlines the project's significance and methodologies. As custodians of cultural heritage, we delve into the nuances of historical preservation, offering insights for future endeavors in safeguarding our collective history.

2.Flow of Project

The binarization begins by reading an RGB image, followed by the application of a Median Blur for denoising. The image is then resized to dimensions 256x256 and pixel values are normalized. Subsequently, the normalized image is input into a U-Net model, generating a grayscale image. A threshold of 0.5 is applied to convert the grayscale image into a binary representation. The binary image is then unnormalized, and the original size of the image is restored through resizing. This AI-driven process leverages U-Net architecture to enhance image features, contributing to the project's overarching goal of refining historical document representation.

2.1 Preprocessing:

In the preprocessing phase, the RGB image is read as the initial step. Following that, a Median Blur operation is applied to effectively reduce noise and enhance image quality. Subsequently, the image is resized to a standardized 256x256 dimensions, ensuring compatibility with the U-Net model's input requirements. To facilitate stable model training and performance, pixel values are normalized, bringing them within a consistent numerical range. This comprehensive preprocessing prepares the input image for optimal processing by the U-Net architecture in the subsequent phase of the binarization.

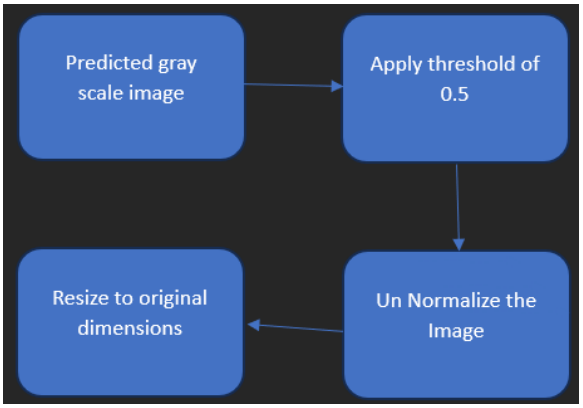


2.2 Model Prediction:

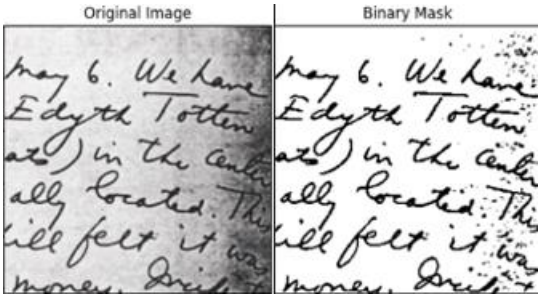
The implemented U-Net model serves as a robust solution for binary image segmentation, particularly tailored for applications in historical document processing. Characterized by an encoder-decoder architecture enriched with skip connections to uphold spatial details, the model leverages convolutional layers and max-pooling in the encoder for efficient feature extraction and downsampling. The decoder, incorporating upsampling and concatenation operations, meticulously reconstructs spatial intricacies. The output layer, governed by a sigmoid activation function, yields a pixel-wise binary prediction. Noteworthy is the utilization of a default threshold of 0.5 for binary classification in the model's output. This meticulously designed U-Net architecture stands as a potent asset for segmenting historical documents, accentuating crucial features and facilitating subsequent analytical endeavors.

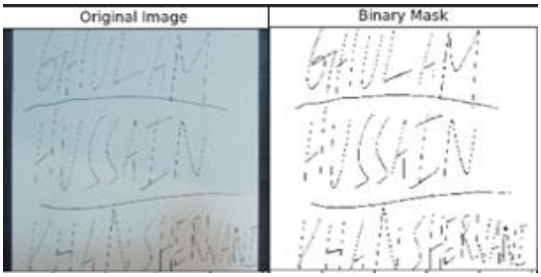
2.3 Post Processing:

In the post-processing phase, the U-Net model's output undergoes key refinements. A threshold of 0.5 is applied to create a binary representation, distinguishing foreground from background. Unnormalization reverses the scaling applied during preprocessing, ensuring the restoration of pixel values to their original scale. Additionally, resizing the binary image to its original dimensions aligns the segmentation output accurately with the structure of the historical documents, contributing to the fidelity and interpretability of the final binarized images.



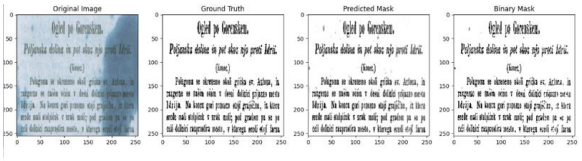
3.Results:

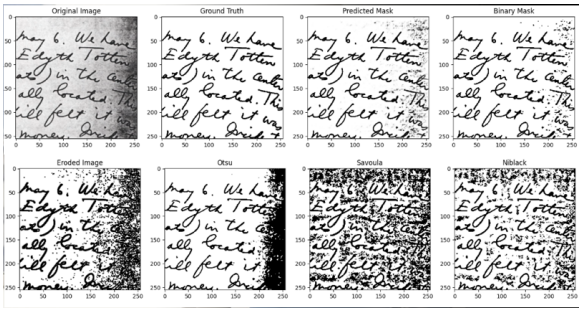




4.Comparison with standard algos:

In this comparison Predicted result is on top right and Eroded image is on Bottom left to the left of Eroded image is Otsu, next is sauvola and then niblack





5.Conclusion:

In summary, our project employs a dual-component approach for historical document binarization. Digital image processing, utilizing techniques like median blur addresses degradation challenges. Simultaneously, the artificial intelligence component, driven by a U-Net model, enhances the binarization process. Though distinct, these components collectively contribute to the preservation and extraction of valuable information from historical documents, effectively mitigating issues related to degradation and noise.