Project Proposal
Handwritten Text Detection and Recognition

Janaksinh Ven (2023702018) Evani Lalitha (2024701023)

1 Introduction

Handwritten Text Recognition (HTR) is a subdomain of Optical Character Recognition (OCR) that focuses on automatically converting handwritten text into digital text. HTR is a challenging problem that intersects with image processing, requiring the handling of various complexities inherent in handwritten documents. Some of the key challenges include the uniqueness of handwriting, noise, blurring, smudges, incomplete characters, variations in ink density, the orientation of characters, and the scaling of characters. In this project, we aim to leverage traditional image processing techniques for line segmentation and word segmentation, which are critical steps in the overall process of recognizing handwritten text.

2 Objectives

The primary objective of this project is to utilize the image processing techniques to perform segmentation of a handwritten document image and subsequently move on to recognition. Specifically, we aim to:

- Use basic image processing techniques to preprocess the handwritten documents at hand.
- Use traditional image processing algorithms like piece-wise painting algorithm etc., to perform line segmentation and word segmentation.
- Perform word recognition of the resultant images and state the resultant metrics.
- If the results are satisfactory we plan to extend it to multiple languages in addition to English.

3 Methodology

- The Pipeline is mainly divided into three parts: pre-processing, line segmentation, word segmentation, and word recognition.
- **Pre-processing:** The first step in the pipeline involves the use of basic image processing techniques like gray-scale conversion, image enhancement, basic augmentations for alignment to pre-process the handwritten documents.
- **Line Segmentation:** The next step in the pipeline involves segmenting the processed document into individual lines. This is done using the method introduced by Alaei *et al.* [1], which is a combination of several traditional image processing techniques with main focus on piece-wise painting algorithm.
- **Word Segmentation:** Once the lines are segmented, the next step is to segment each line into individual words. This will be achieved using vertical projection profiles. For each line segment, the sum of pixel values along each column is calculated. The resulting profile is analyzed to identify valleys that correspond to spaces between words. By identifying these spaces, the line can be divided into separate word segments. An alternative approach will involve connected component analysis, where bounding boxes are drawn around clusters of connected pixels to identify word boundaries.
- **Word Recognition:** After successfully segmenting the words, the final step is word recognition. For word recognition we plan to use existing deep learning based word recognizers such as CRNNs [3](handwritten) or PARSeq (printed)[2]. We will use a pretrained model for this step. If the project is completed within the time period we will try

to use classical image processing techniques for the same.

- **Evaluation:** The performance of the pipeline will be evaluated using established metrics such as Character Error Rate (CER) and Word Error Rate (WER). Benchmark datasets for handwritten text, such as the IAM Handwriting Database[4], will be used to assess the effectiveness of the proposed methodology. If possible we also plan to perform recognition on printed data with benchmark datasets like the ICDAR printed text datasets.

4 Expected Outcomes

- Qualitative analysis of each step in the pipeline.
- Quantitative analysis by using appropriate metrics for each step.
- Insights into the strengths and weaknesses of classic image processing techniques for segmentation.

5 Timeline

- Week 1-4: Explore and implement the models for line and word segmentation.
- Week 5-7: Implement the word and character recognition model.
- Week 8-9: Analyze the performance of the different models on different datasets.
- Week 10-11: Compile results, write the project report, and prepare for presentation video or publication.

6 Resources Required

- Access to computational resources for model training (if required) and experimentation (Note: Computational resources available to us are sufficient).
- Necessary software tools and libraries for image processing and data analysis.
- Collaboration with domain experts (from CVIT OCR team) for validation and feedback.

7 Conclusion

In this project, we aim to study handwritten text detection and recognition on a paragraph level by combining traditional image processing for segmentation with deep learning for recognition. We will evaluate our approach using benchmarks, providing insights into the effectiveness of classic techniques.

References

[1] Alireza Alaei, Umapada Pal, and P Nagabhushan. A new scheme for unconstrained handwritten text-line segmentation. *Pattern Recognition*, 44(4):917–928, 2011.

[2] Darwin Bautista and Rowel Atienza. Scene text recognition with permuted autoregressive sequence models. In *European conference on computer vision*, pages 178–196. Springer, 2022.

[3] Santhoshini Gongidi and CV Jawahar. iiit-indic-hw-words: A dataset for indic handwritten text recognition. In *Document Analysis and Recognition–ICDAR 2021: 16th International Conference, Lausanne, Switzerland, September 5–10, 2021, Proceedings, Part IV 16*, pages 444–459. Springer, 2021.

[4] U-V Marti and Horst Bunke. The iam-database: an english sentence database for offline handwriting recognition. *International journal on document analysis and recognition*, 5:39–46, 2002.