Handwritten Text Recognition using Deep Learning: A Case Study with IAM Dataset

Adarsh Jha*, Vinayak Aggarwal† *Roll Number 2022024 †Roll Number 2022574

***Abstract*—This project aims to address the challenge of Handwritten Text Recognition (HTR) by leveraging state-of-the-art deep learning methodologies, with a focus on the IAM Handwriting Database for model training and evaluation. HTR involves accurately transcribing handwritten text from images into machine-readable format, with applications in document digitization, historical text analysis, and more. The complexity arises from variations in writing styles, diverse fonts, and contextual intricacies. Our approach integrates Convolutional Neural Networks (CNNs) and Recurrent Neural Networks (RNNs), including Long Short-Term Memory (LSTM) units, to capture spatial and sequential dependencies in handwritten text images. The IAM dataset offers a diverse collection of handwritten samples with ground truth transcriptions, facilitating accurate model evaluation. Evaluation metrics encompass character accuracy, word accuracy, and recognition time, supplemented by visualization of model predictions against ground truth for performance analysis. This project endeavors to advance Handwritten Text Recognition through deep learning techniques, contributing to the development of more accurate and robust models for automated transcription from handwritten documents.**

I. Introduction

Handwritten Text Recognition (HTR) is a significant challenge in the domain of pattern recognition and document analysis. The task involves transcribing handwritten text from images into machine-readable format, enabling various applications such as digitization of historical documents, automatic data entry, and text analysis. However, the inherent variability in handwriting styles, diverse fonts, and contextual complexities pose significant challenges to accurate recognition.

Supervised learning techniques, including deep learning approaches such as Convolutional Neural Networks (CNNs) and Recurrent Neural Networks (RNNs), have shown promising results in addressing the challenges of Handwritten Text Recognition. Additionally, traditional machine learning algorithms like K-Nearest Neighbors (KNN) have also been applied in this domain to classify handwritten characters or words based on their similarity to training examples.

II. Problem Statement

Accurate transcription of handwritten documents necessitates the development of robust models capable of understanding and interpreting various forms of handwriting across different languages and writing styles. The challenge lies in designing a system that can effectively capture the spatial and sequential dependencies inherent in handwritten text images.

III. Methodology

Our proposed approach utilizes a combination of Convolutional Neural Networks (CNNs) and Recurrent Neural Networks (RNNs), specifically Long Short-Term Memory (LSTM) units, to address the challenges of Handwritten Text Recognition. CNNs are adept at capturing spatial features from input images, while RNNs, particularly LSTM units, excel at modeling sequential data, enabling the recognition of long-range dependencies within the text.

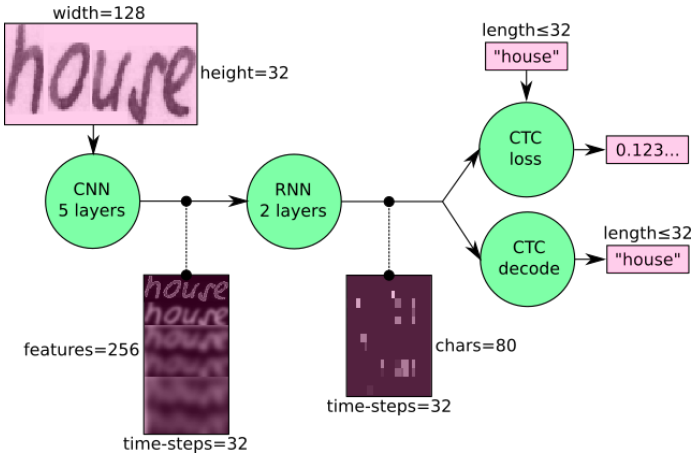

Fig. 1. Illustration of the proposed methodology

To further enhance the decoding capability of our model, we have integrated word beam search decoding. This decoder allows recognition of words constrained to those contained in a dictionary while still recognizing arbitrary non-word character strings (numbers, punctuation marks).

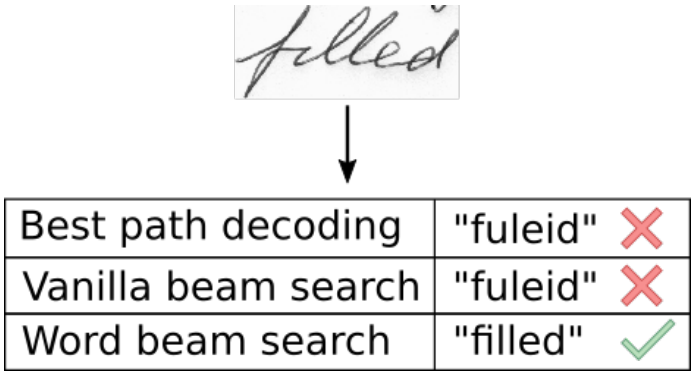

Fig. 2. Illustration of the proposed methodology

IV. Dataset

The IAM Handwriting Database is selected as the primary dataset for model training and evaluation. This dataset provides a diverse collection of handwritten text samples, covering various languages, writing styles, and document types. Ground truth transcriptions at both word and line levels are available, facilitating accurate evaluation of the model's performance. The IAM Handwritten Forms Dataset used in this project can be accessed at the following link: Click this

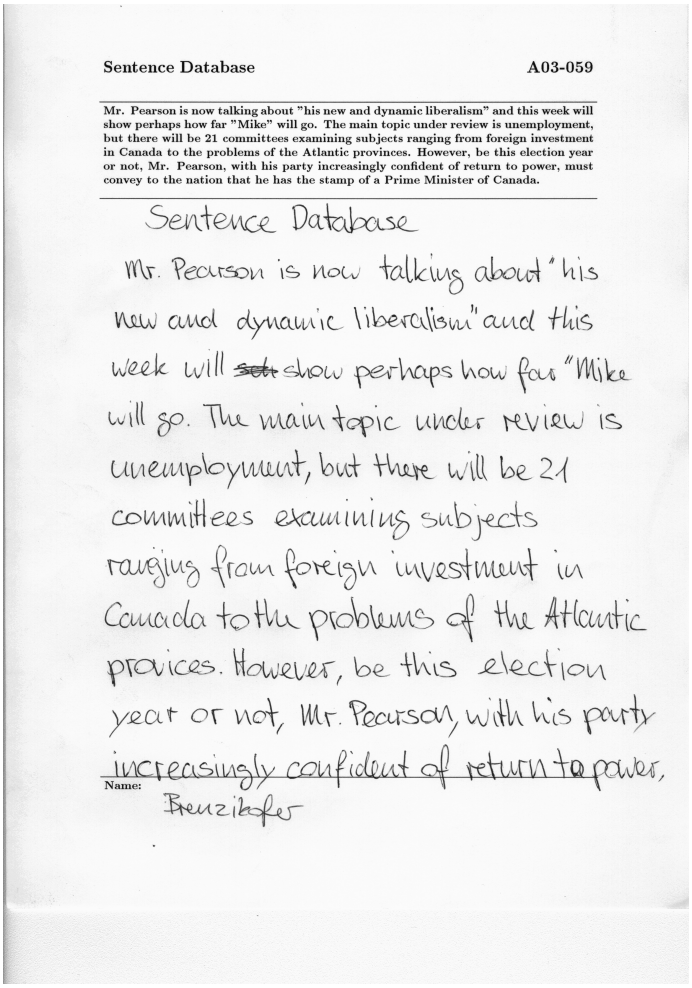

Fig. 3. Sample Dataset

V. More about Dataset

The IAM Handwriting Database is a widely used benchmark dataset for handwritten text recognition research. It contains handwritten forms of English text collected from various sources, including letters, articles, and other documents. The dataset includes over 1,100 pages of handwritten text, with ground truth transcriptions provided at both word and line levels. Each page contains multiple lines of handwritten text, capturing variations in writing styles, fonts, and layouts. The IAM Handwriting Database is often used for training and evaluating handwritten text recognition models due to its size, diversity, and availability of ground truth annotations.

VI. Results

The success of our project will be evaluated based on the model's ability to accurately transcribe handwritten text from unseen samples in the IAM dataset. Performance metrics including character accuracy, word accuracy, and recognition time will be employed for evaluation. Visualization of the model's predictions against ground truth will provide insights into its strengths and areas for improvement.

VII. Conclusion

In summary, this project aims to advance the field of Handwritten Text Recognition through the application of deep learning techniques to a challenging dataset. By developing more accurate and robust models, we aim to contribute to the automation of text transcription from handwritten documents, thereby facilitating various applications in document analysis and information retrieval.

References

1) Graves, A., Schmidhuber, J. (2009). Offline Handwriting Recognition with Multidimensional Recurrent Neural Networks. In Advances in Neural Information Processing Systems (pp. 545–552).
2) Marti, U., Bunke, H. (2002). The IAM-database: an English sentence database for offline handwriting recognition. International Journal on Document Analysis and Recognition, 5(1), 39–46.
3) Lecun, Y., Bottou, L., Bengio, Y., Haffner, P. (1998). Gradient-based learning applied to document recognition. Proceedings of the IEEE, 86(11), 2278–2324.
4) George Retsinas. (2022). HTR-ctc. GitHub Repository. Retrieved from https://github.com/georgeretsi/HTR-ctc
5) Itsishika. (2021). Handwritten-Text-Recognition. GitHub Repository. Retrieved from https://github.com/Itsishika/Handwritten-Text-Recognition