

2023 Fall CSE438 Final Project 1st Draft

Group 05

November 25, 2023

Literature Survey

1. Handwritten Chinese Text Recognition using Separable Multidimensional Recurrent Neural Network

In the research paper titled *Handwritten Chinese Text Recognition Using Separable Multidimensional Recurrent Neural Network*, the study addresses the challenges associated with handwritten Chinese text recognition, including various writing styles, character segmentation difficulties, a vast character set, and an unconstrained language domain. The proposed Separable Multi-Dimensional Long Short-Term Memory Recurrent Neural Network (SMDLSTM-RNN) aims to overcome these challenges by efficiently capturing contextual information in several orientations, reducing computation time, and improving recognition accuracy. The paper demonstrates the effectiveness of the SMDLSTM-RNN on datasets with diverse character class counts, showcasing competitive accuracy in detecting Chinese characters with complex writing styles.

- The research presents a novel approach to handwritten Chinese text recognition utilizing a 7-layer architecture of Separable Multi-Dimensional Long Short-Term Memory Recurrent Neural Network (SMDLSTM-RNN) and Convolutional Neural Networks (CNN) modules.
- The suggested method outperforms a Bidirectional Long Short-Term Memory (BLSTM)-based model in trials done on datasets with varied character class counts.
- Experimental findings on the ICDAR-2013 competition dataset show that the suggested method performs better than earlier LSTM-based techniques and is capable of competing with cutting-edge systems.
- For future work to further improve system performance, the paper suggests tackling the context overfitting issue.

2. Bangla Handwritten Character Recognition using Convolutional Neural Network with Data Augmentation

The research paper, *Bangla Handwritten Character Recognition using Convolutional Neural Network with Data Augmentation*, introduces a novel approach to recognizing individual Bangla characters. The study focuses on the BanglaLekha-Isolated dataset and employs a Convolutional Neural Network (CNN) architecture. The model achieves significant accuracy, especially after data augmentation, showcasing the effectiveness of CNN for Bangla handwritten character recognition. The paper emphasizes the importance of recognizing Bangla characters in handwritten form, highlighting the scarcity of research in this area compared to English handwriting recognition. The proposed CNN architecture offers a promising solution, paving the way for further developments in this domain.

- The methodology involves training the CNN model with the base dataset and then expanding it using data augmentation techniques.
- Data augmentation proves effective in enhancing the model's accuracy.
- The paper provides a comprehensive analysis of the results, including precision, recall, and F1 score.
- The model's versatility is demonstrated by applying it to other datasets, such as MNIST, Ekush Bangla, and CMATERdb 3.1.2, showcasing competitive accuracy.

3. OCR in Bangla: an Indo-Bangladeshi Language

The paper, *OCR in Bangla: an Indo-Bangladeshi Language*, outlines a comprehensive Optical Character Recognition (OCR) system designed for single-font Bangla documents. The approach combines template and feature matching for character recognition without requiring preprocessing steps, contributing to its speed. The paper discusses the system's accuracy in recognizing basic and compound characters, providing insights into the statistical analysis of Bangla text. The research holds significance for OCR applications in the context of the Bangla script, potentially aiding in text entry automation and assisting visually impaired individuals.

- The feature-based tree classifier is employed for recognizing simple characters, achieving an accuracy of approximately 96
- The system's accuracy is demonstrated through experiments, achieving a 97
- The overall recognition accuracy is reported as 96.55

Collected Data

For this paper, we selected a secondary dataset consisting of handwritten Bangla texts. The BanglaWriting dataset includes single-page handwritings of 260 people of various personalities and ages. The dataset comprises 21,234 words, 32,787 characters, and 5,470 unique Bangla vocabulary items. The dataset is suitable for optical character/word recognition, author identification, and handwritten word segmentation.

Dataset: <https://data.mendeley.com/datasets/r43wkvdk4w/1>

Proposed Methodology

As there has been several research regarding handwritten Bangla text recognition using various neural networks such as RNN, CNN, etc., we choose the Separable Multi-Dimensional Long Short-Term Memory Recurrent Neural Network (SMDLSTM-RNN) as our flagship model. This model, having shown good results in detecting Chinese handwritten characters, is expected to provide efficiency in detecting Bangla handwritten characters. We will apply the model for testing on the datasets we have collected and, if needed, create custom datasets for further testing.

Reference

1. Y.-C. Wu, F. Yin, Z. Chen and C.-L. Liu, "Handwritten Chinese Text Recognition Using Separable Multi-Dimensional Recurrent Neural Network," 2017 14th IAPR International Conference on Document Analysis and Recognition (ICDAR), Kyoto, Japan, 2017, pp. 79-84, doi: 10.1109/ICDAR.2017.22.
2. R. R. Chowdhury, M. S. Hossain, R. ul Islam, K. Andersson and S. Hossain, "Bangla Handwritten Character Recognition using Convolutional Neural Network with Data Augmentation," 2019 Joint 8th International Conference on Informatics, Electronics & Vision (ICIEV) and 2019 3rd International Conference on Imaging, Vision & Pattern Recognition (icIVPR), Spokane, WA, USA, 2019, pp. 318-323, doi: 10.1109/ICIEV.2019.8858545.
3. U. Pal and B. B. Chaudhuri, "OCR in Bangla: an Indo-Bangladeshi language," Proceedings of the 12th IAPR International Conference on Pattern Recognition, Vol. 3 - Conference C: Signal Processing (Cat. No.94CH3440-5), Jerusalem, Israel, 1994, pp. 269-273 vol.2, doi: 10.1109/ICPR.1994.576917.