

Abstract

This project aims to extract and recognize words from a handwritten document and convert into digital form. We use two main steps to accomplish this task: Word segmentation and word recognition. For segmentation we use scale space technique to separate words. Then for recognition we use Neural Networks. It consists of convolutional neural network(CNN) layers, recurrent neural network(RNN) layers, and a final Connectionist Temporal Classification (CTC) layer with various architectures to train a model that can accurately classify words.

Reference:

1. Sri. Yugandhar Manchala, Jayaram Kinthali, Handwritten Text Recognition using Deep Learning, IJERT , 2020
2. Frinken V, Bunke, H.: Continuous handwritten script recognition, in Doermann, D., Tombre, K. (eds.): Handbook of Document Image Processing and Recognition, Springer Verlag, 2014
3. S. GÃ¼nter and H. Bunke. A new combination scheme for HMM-based classifiers and its application to handwriting recognition. In Proc. 16th Int. Conf. on Pattern Recognition, volume 2, pages 332337. IEEE, 2002.
4. U.-V. Marti and H. Bunke. Text line segmentation and word recognition in a system for general writer independent handwriting recognition. In Proc. 6th Int. Conf. on Document Analysis and Recognition, pages 159 163.
5. M. Liwicki and H. Bunke, Iam-ondb – an on-line English sentence database acquired from the handwritten text on a whiteboard, in ICDAR, 2005
6. A. Graves and J. Schmidhuber, Offline handwriting recognition with multidimensional recurrent neural networks, in Advances in neural information processing systems, 2009, pp. 545552.