Abstract

This project aims to extract and recognize words from a handwritten document and convert into digital form. We uses two main steps to accomplish this task: Word segmentation and word recognition. Word segmentation is done using scale space technique and other algorithms. These extracted words are then used for text recognition. For text recognition we use Deep learning using TensorFlow and Keras. Convolutional Neural Network(CNN) layers are used for feature extraction and classification. We additionally add support for data augmentation and spellcheckers for obtaining better recognition.

References:

1 Harald Scheidl, *Build a Handwritten recognition using tensorflow*, Towardsdatascience.com, 2018

2 Sri. Yugandhar Manchala, Jayaram Kinthali, *Handwritten Text Recognition using Deep Learning*, IJERT, 2020

3 Batuhan Balci, Dan Saadati, Dan Shiferaw, *Handwritten Text Recognition using Deep Learning*, Stanford university, 2017

4 R Manumatha and N Srimal, Scale space technique for handwritten manuscripts, CITESEERX, 1999