IJARSCT

ISSN (Online) 2581-9429

International Journal of Advanced Research in Science, Communication and Technology (IJARSCT)

International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal

Impact Factor: 7.53

Volume 4, Issue 4, May 2024

Handwritten Text Recognition

Tanishq Rampure, Harsh Damahe, Kundan Bharati, Darshil Raj, Namrata Naikwade

MIT Arts, Design and Technology University, Pune, India

Abstract: *Handwritten text detection refers to the capacity of a computer system to interpret and understand handwritten input from various sources such as paper documents, touch displays, and photographs. Within the realm of pattern recognition, one specific area is handwritten text recognition, which involves categorizing and interpreting handwritten text. In this study, we propose a novel approach to offline handwritten text recognition utilizing deep neural networks. The availability of vast amounts of data and continuous algorithmic enhancements have made training neural networks more accessible than ever before. This handwritten character recognition system relies on image segmentation to accurately identify and interpret handwritten characters. Through this approach, we aim to improve the accuracy and efficiency of recognizing handwritten text across various applications and contexts.*

Keywords: Handwritten text detection, deep neural network, image segmentation

I. INTRODUCTION

Handwriting remains a significant mode of daily communication, with many still opting for pen and paper despite the availability of modern writing tools. However, handwritten notes pose certain challenges, including difficulties in storage, retrieval, and sharing. The manual nature of updating and organizing handwritten data also leads to potential data loss and inefficiencies in analysis. Recognizing the need to address these issues, our project aims to bridge the gap between handwritten and digital text, making data more accessible and manageable.

Our project focuses on handwritten text recognition systems, which come in two forms: offline and online. Offline systems utilize images of handwritten text, employing optical character recognition (OCR) to convert them into digital format. Online systems, on the other hand, interpret pen tip motions recorded as a sequence to represent handwritten characters.

The core objective of our project is to delve deeper into classifying handwritten text and transforming it into a digital form. Given the broad concept of handwritten text, we have defined specific parameters to narrow down our project scope. We aim to classify images containing handwritten text, aiming to provide a comprehensive solution to the challenges associated with handwritten data.

By identifying and addressing the nuances of handwritten text recognition, our research endeavors to enhance accessibility and usability, bridging the gap between traditional and digital modes of communication and data management.

II. LITERATURE REVIEW

[1] Andrew Smith introduced a method for Offline handwritten text recognition, which involves extracting text from a photograph and converting it into a digital format. Previous approaches relied on lexical segmentation and complex feature extraction methods along with linguistic knowledge. The proposed method combines a convolutional recurrent neural network with connectionist temporal classification to recognize offline handwritten text. This approach is character- independent, making it suitable for various languages and globally trainable. Optical character recognition (OCR) is employed to convert both printed and handwritten texts into machine-readable language. The recognition process involves utilizing a convolutional recurrent neural network with three neural network blocks: one for image feature extraction, another for sequence learning, and a final block for labeling with connectionist temporal classification linked to a scoring function.

[2] Thomas Deselaers developed a Handwritten Text Recognition (HTR) system, which is integrated into an Optical Character Recognition (OCR) model. This system incorporates three techniques: the HTR system, the OCR model, and a Dual Head architecture, which merges the HTR into the OCR. Compared to standalone handwritten text recognition

systems and OCR models, the Dual Head model excels in categorizing both printed and handwritten text, achieving the highest accuracy. Initially, a substantial amount of high-quality training data is required, followed by the implementation of a line recognition model based on neural networks but without recurrent connections.

[3] Anshul Gupta proposed a method for identifying offline handwritten English words, wherein individual characters are recognized before the complete word is identified. Handwritten text recognition approaches are typically categorized into holistic and segmentation methods. Holistic approaches are employed for recognizing a limited vocabulary size, where global features extracted from the entire word image are considered. However, as the vocabulary size increases, the complexity of holistic approaches also increases, leading to a decrease in character recognition rates. On the other hand, segmentation-based approaches employ a bottom-up strategy, starting from character level recognition and gradually forming meaningful words. Segmentation reduces the task from recognizing entire words to recognizing individual characters, thus accommodating larger vocabularies. Neural networks are utilized for identifying individual characters in this process.

[4] Peng Ren has developed a handwritten text recognition program using deep learning techniques, primarily focusing on a detection algorithm for objects. The recognition process for offline handwritten text occurs in two main steps: preprocessing and character recognition. Preprocessing begins with a faster R-CNN for initial processing, followed by character recognition using a convolutional neural network. By treating character segmentation as object detection, the segmentation problem is effectively addressed. The process involves preliminary processing, where sentences are segmented into words, followed by meticulous processing, which involves segmenting words into characters and then recognizing those characters.

[5] Francisco Zamora-Martinez has developed a novel technique to remove slants from handwritten text, along with balancing the size of text images using artificial neural networks. The process of normalizing handwritten text from scanned images involves several steps, including image cleaning, page skew correction, and line detection. However, since a skew-corrected lines database is utilized, the steps for page skew correction and line detection can be skipped. Various preprocessing steps are employed to reduce variations in handwriting. Initially, the scanned line is cleaned, followed by slope and slant removal. Once the entire image is free from slopes and slants, text line size normalization is performed to minimize changes in size and position. An artificial neural network (ANN) system is used for recognizing offline handwritten text lines.

[6] Martin Rajnoha has introduced a method for recognizing handwritten text in the Comenia font. This method involves preprocessing and normalization of data, followed by optical character recognition (OCR) based on Support Vector Machine (SVM). The Comenia script is characterized by its simplicity and modernity, resembling block letters. The proposed model utilizes multiple classification models for character recognition, resulting in improved accuracy compared to a single- model approach. The process of recognizing handwritten text encompasses feature extraction, with support vector machine being employed for classification.

[7] Batuhan Balsi developed a method for classifying individual handwritten words for translation into digital form. Two approaches were employed: direct word classification and character segmentation. A convolutional neural network with various topologies was utilized to train the model for word classification. Bounding boxes for single characters were constructed using long short-term memory networks with convolution. Segmented characters were then fed into a convolutional neural network for classification, and words were reconstructed based on the character classification and segmentation results. An alternative strategy was implemented to improve direct word classification results by segmenting characters and reconstructing words based on individual character classification.

[8] Rohan Vaidya has introduced a method for detecting offline handwritten characters utilizing deep neural networks. This method incorporates a handwritten character recognition system based on image segmentation, designed using the Python programming language. Various tools such as Android OpenCV and TensorFlow were employed to develop the offline handwritten character recognition system. The Android application built must enable users to capture photos of handwritten text using the phone's camera for identification. A pretrained neural network model is utilized for predictions, followed by image processing operations. The neural network model is trained using TensorFlow. Image preprocessing is conducted to remove noise from the image, followed by conversion to grayscale. Subsequently, thresholding is applied to distinguish between darker and lighter regions in the image.

IJARSCT

ISSN (Online) 2581-9429

International Journal of Advanced Research in Science, Communication and Technology (IJARSCT)

International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal

Impact Factor: 7.53

Volume 4, Issue 4, May 2024

[9] Vu Pham has introduced a recurrent neural network (RNN) for recognizing handwritten texts. In this method, the input image is divided into blocks and fed into long short-term memory (LSTM) layers, which scan the input. The output of the LSTM layer is then passed to a convolutional layer. The output of the convolutional layer is vertically summed and processed through a softmax layer, followed by connectionist temporal classification. Dropout, a technique involving the removal of some hidden units during training and their random utilization during testing, is employed. Another similar method is DropConnect, where connections are dropped instead of hidden units' values. Both convolutional and recurrent layers are compatible with dropout. Initially, dropout is applied to the top layer of the LSTM. If the model size is large and overfitting is observed, dropout proves beneficial.

III. DATASET

The IAM Handwriting Database comprises handwritten English text samples, providing valuable resources for training and testing handwritten text recognition systems, as well as conducting experiments related to writer identification and verification.

The IAM Handwriting Database 3.0 is structured as follows:

657 writers contributed samples of their handwriting 1'539 pages of scanned text 5'685 isolated and labeled sentences 13'353 isolated and labeled text lines 115'320 isolated and labeled words

IV. METHODOLOGY

Convolution Neural networks for Text Recognition

In various pattern recognition and image processing tasks, neural networks play a crucial role, with Convolutional Neural Networks (CNNs) emerging as a leading method. Inspired by the connectivity patterns found in the visual cortex of animals, CNNs excel in tasks like image and object recognition. Constructing a CNN model for digit recognition often involves leveraging libraries like KERAS in Python and utilizing TensorFlow as the backend. The Sequential Model, functioning as a classifier, forms a linear stack of layers, allowing for the learning of weights and biases in the network's neurons.

Compared to other methods, CNNs typically require minimal preprocessing, making them efficient for a wide range of applications. Unlike traditional neural networks where the input is a vector, CNNs handle multi-channeled images as input. These networks typically comprise an input layer, hidden layers, and an output layer, with the hidden layers featuring convolutional layers, Rectified Linear Unit (ReLU) activation functions, pooling layers, normalization layers, and fully connected layers.

In the CNN architecture, the input layer often consists of 28 by 28-pixel images, translating to a network of 784 neurons as input data. Following the input layer, the convolutional layer receives the output, subjecting it to convolution operations to reduce the number of free parameters. The ReLU activation function, a linear function known for its simplicity and non-saturation, is applied to enhance performance.

Pooling layers are then utilized to consolidate neuron outputs from one layer into a single neuron in the next, reducing computation complexity. Fully connected layers calculate scores for input digits, followed by the application of a SoftMax classifier to determine the probabilities of output classes, with the class exhibiting the highest probability being selected for final classification.

To work seamlessly with the Keras API, input data is converted into 4-dimensional NumPy arrays, and normalization is typically performed by dividing RGB codes by 255. Adam optimizers are often employed to update neuron weights efficiently, requiring minimal memory resources.

The system comprises two essential components: a deep convolutional network based on regions and a Fast R-CNN detector that leverages these areas. This innovative technology streamlines object detection by significantly reducing the processing time for each image to a remarkable 10 milliseconds. However, due to its high computational demands, this approach may not be suitable for applications running on CPUs with limited processing capabilities.

The administrator initiates the process by loading the handwritten image dataset, which varies in size, color channels, and overall dataset volume. The complexity of processing an image increases with its dimensions and pixel values, highlighting the importance of compressing images into a format that maintains predictive properties while easing computational burden. Document analysis, also known as pre-processing, involves tasks such as background noise reduction, filtering, and image restoration, essential for extracting text from documents.

In the CNN block, seven convolutional layers are employed, with max-pooling layers applied after select layers for downsampling and batch normalization layers for normalization. Max-pooling reduces matrix dimensionality by replacing each sub-region with the maximum value, while Rectified Linear Unit function activates all convolutional layers.

Feature extraction is a crucial step in reducing raw data into smaller groupings, with each character represented as a feature vector. The primary goal of feature extraction is to enhance recognition accuracy by extracting relevant features. The Model Weight parameter in a neural network dynamically alters input data within hidden layers during training. Providing starting values for weight and bias parameters is crucial, as they are adjusted during training, influencing neuron activation functions and facilitating learning. User interaction involves uploading an image, which is then compared with model weights for classification. Input characters are classified by comparing them to templates (prototypes) from each character class. If input characters match the templates closely, the character's identity is assigned to the most similar template. These stages operate in a sequential manner, where the success of each step depends on the preceding one. Furthermore, the accuracy of each phase is interdependent, as the output of one phase serves as input to the next, ensuring a cohesive and effective processing pipeline.

V. PROPOSED SYSTEM

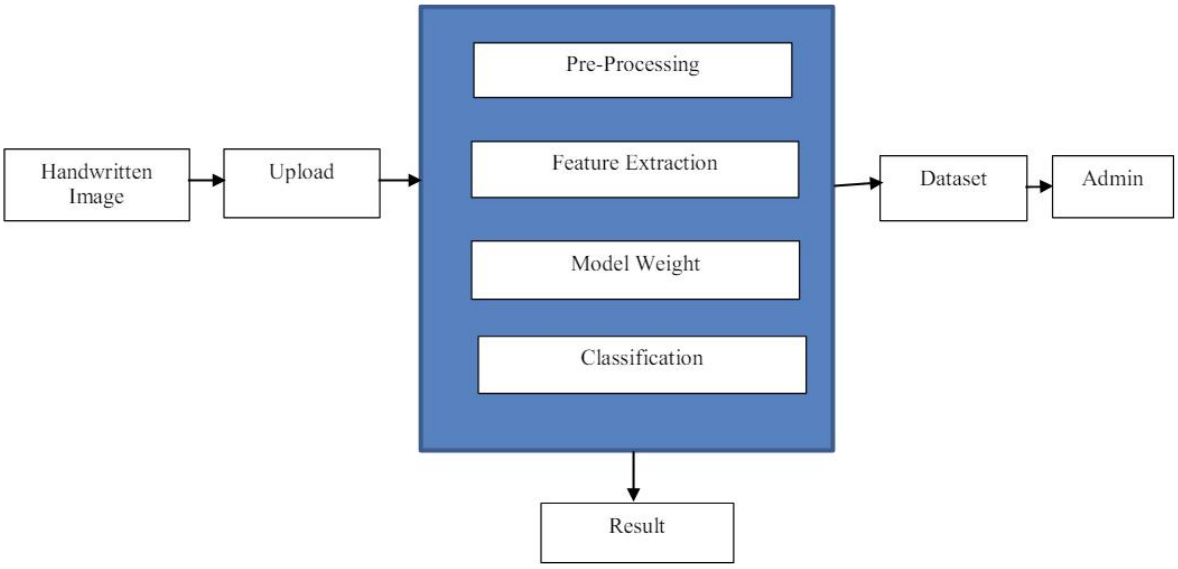


Figure 1. Proposed System

IJARSCT

ISSN (Online) 2581-9429

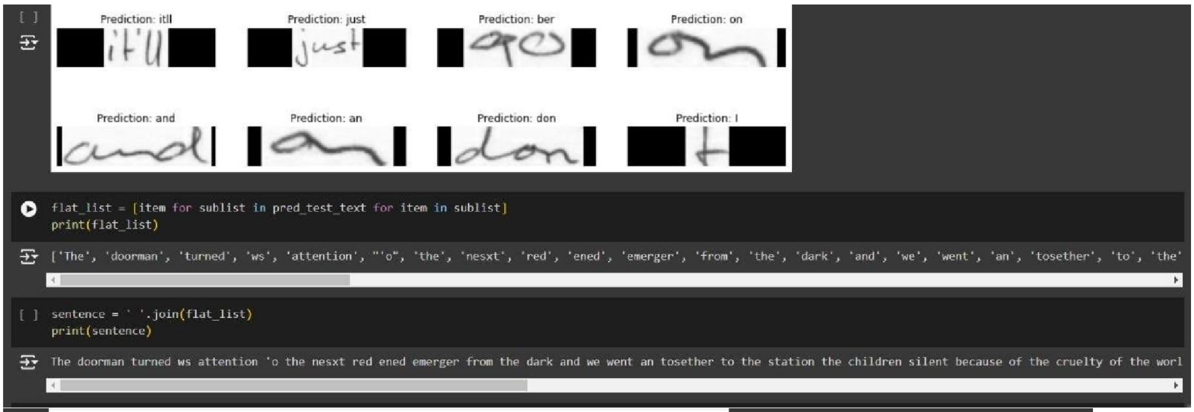
International Journal of Advanced Research in Science, Communication and Technology (IJARSCT)

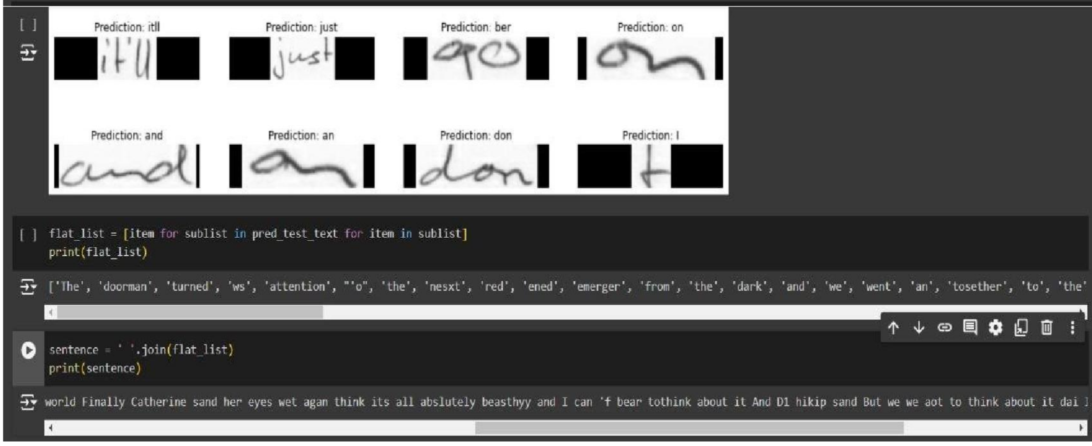
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal

Impact Factor: 7.53

Volume 4, Issue 4, May 2024

VI. OUTPUT





VII. CONCLUSION AND FUTURE ENHANCEMENT

In our endeavor to streamline handwritten text recognition, we've devised an innovative method that significantly reduces the time required for this task. A major hurdle in this process has always been character segmentation, a challenge we've overcome by reframing it as an object detection problem. By adopting this approach, we've achieved remarkable accuracy in character segmentation, thanks to the utilization of Faster-CNN. Following successful segmentation, character recognition is carried out using CNNs, treating each character individually.

Our approach doesn't just stop at solving these challenges; it also revolutionizes the entire recognition process through modular designs. By breaking down the complex task into manageable sub-problems, we've opened up new avenues for effective solutions.

Moreover, our proposed system isn't just limited to recognizing isolated characters; it's capable of handling entire pages of handwritten text. This system ensures not only high accuracy but also remarkable speed, making it a valuable tool for various applications requiring efficient handwritten text recognition.

REFERENCES

[1] Andrew Smith, Hong-Phuong Tran, Eric Dimla "Offline Handwritten Text Recognition using Convolutional Recurrent Neural Network" International Conference on Advanced Computing and Applications 2019 (ACOMP)

[2] R. Reeve Ingle, Yasuhisa Fujii, Thomas Deselaers, Jonathan Baccash, Ashok C. Popat "A Scalable Handwritten Text Recognition System" International Conference on Document Analysis and Recognition 2019 (ICDAR)

[3] Anshul Gupta, Manisha Srivastava, Chitralekha Mahanta "Offline Handwritten Character Recognition Using Neural Network" International Conference on Computer Application and Industrial Electronics 2011 (ICCAIE-2011) [4] Junqing Yang, Peng Ren, Xiaoxiao Kong "Handwriting Text Recognition Based on Faster R-CNN" IEEE xplore 2020

[5] Salvador Espan ã -Boquera, Maria Jose Castro-Bleda, Jorge Gorbe-Moya, and Francisco Zamora-Martinez "Improving Offline Handwritten Text Recognition with Hybrid HMM/ANN Models" IEEE Transactions on Pattern Analysis and Machine Intelligence, VOL. 33, NO. 4, APRIL 2011

[6] Martin Rajnoha, Radim Burget, Malay Kishore Dutta "Offline Handwritten Text Recognition Using Support Vector Machines" International Conference on Signal Processing and Integrated Networks 2017 (SPIN)

[7] Batuhan Balci, Dan Saadati, Dan Shiferaw "Handwritten Text Recognition using Deep Learning"

[8] Rohan Vaidya, Darshan Trivedi, Sagar Satra "Handwritten Character Recognition Using Deep- Learning" 2nd International Conference on Inventive Communication and Computational Technologies 2018 (ICICCT-2018)

[9] Vu Pham, Theodore Bluche ́, Christopher Kermorvant, and Jer ́ome Louradour, "Dropout improves Recurrent Neural Networks for Handwriting Recognition" 14th International Conference on Frontiers in Handwriting Recognition 2014.