International Journal of Current Engineering and Technology
E-ISSN 2277 – 4106, P-ISSN 2347 – 5161

Available at http://inpressco.com/category/ijcet

Research Article

Development of Handwritten Text to Editable Text to Audio Converter

Prof. Aniket Bhoyar, Prof. Alka Shrivastav, Satvik Khedkar, Pranav Purkar, Shrutika Mohod, Sanket Gotmare, and Vanshaj Nagdeve

Department of Computer Science and Engineering, S. B. Jain Institute of Technology, Management and Research, Nagpur, India

Received 25 April 2023, Accepted 16 May 2023, Available online 18 May 2023, **Vol.13, No.3 (May/June 2023)**

Abstract

This project attempts to classify any unique handwritten text in order to convert the text material to a digital format. To complete this objective, we used two different methods: direct word classification and separating each character. Converting text to audio using an optical character reader (OCR) with the help of state- of-the-art technology has solved the main problem of convenient communication. We know that listening is more interactive than reading a book. It is good that the system captures images and detects and recognizes text. OCR converts input images into editable text, allowing you to turn paper documents into editable and searchable documents. This can help reduce the physical space required to store document and significantly improve the work flow associated with those documents.

Keywords: *Optical character recognition, translation, audio speech*

1. Introduction

Text Recognition: The app uses machine learning algorithms to recognize and extract the text from the image or video. This involves training the machine learning models on a large dataset of labelled images, where the text regions are labelled with the corresponding text. Post-processing: The app post-processes the extracted text to correct any errors or inconsistencies, such as by spell-checking, grammar-checking, or using language models to improve the accuracy of the recognition. Overall, the development of a text recognition app using machine learning involves a combination of technical expertise in computer vision, natural language processing, and machine learning, as well as access to large datasets of labelled images for training and validation. A text recognition app using machine learning is an application that uses artificial intelligence algorithms to recognize and extract text from images, videos, or other sources of digital media. The main challenge in developing such an app is to train the machine learning models to accurately recognize and extract text from a wide variety of sources, such as different fonts, languages, writing styles, and image qualities. The basic workflow involves data collection, pre-processing, text detection, text recognition, and post-processing. This requires technical expertise in computer vision, natural language processing, and machine learning, as well as access to large datasets of labelled images.

*Corresponding author's ORCID ID: 0000-0000-0000-0000
DOI: https://doi.org/10.14741/ijcet/v.13.3.7

2. Objectives

- To develop an accurate and efficient text recognition system that can accurately extract text from a wide range of sources, including images, videos, and other digital media.
- To train machine learning models using large datasets of labeled images to accurately recognize and extract text from a variety of fonts, languages, writing styles, and image qualities.
- To use computer vision techniques to detect text regions within an image or video, and to pre-process the media to enhance the visibility of text before recognition.
- To post-process the extracted text to correct errors and inconsistencies using spell- checking, grammar-checking, and language models to improve recognition accuracy.
- To integrate the text recognition app with other applications and services, such as translation, transcription, or information retrieval, to enhance the usability and usefulness of the app.

3. Literature Review

A Survey of the research done for Optical Character Recognition and the currently existing system give the following results. Many countries are contributed to the structure and infrastructure of various sectors in order to implement and support the digital process. The road to success, they say, is paved with digitization. Transparency and effectiveness are improved by digitization. OCR is a capable method for

integrating analogue life information into a cyberspace world. This technology has long been used to create digital libraries, recognize natural scenes, and understand handwritten office forms. Using OCR technology, documents scanned or captured by the camera are editable electronic versions that can be readily edited, retrieved, duplicated, and communicated.

Handwriting recognition is a very necessary method in modern in character recognition engine that is extensively considered as one of the correct currently available. To recognize text in an image, tesseract first converts the input image to a binary image by performing an adaptive threshold. It performs an analysis of the connected components to preserve the outlines of each component. Lines of text are split into words based on letter spacing. Text recognition is performed in two steps to increase accuracy. Follow the same sequence of steps for to recognize both printed and handwritten text.

A study by Wang et al. (2018) developed a text recognition app for mobile devices using deep learning- based methods. The app was designed to recognize text from images and videos, and it achieved high accuracy in recognizing both printed and handwritten text.

In a study by Wu et al. (2020), a text recognition app was developed using an attention-based encoder-decoder network. The app was able to recognize text from a variety of sources, including scanned documents, photographs, and real-time video. The results showed that the app achieved high accuracy in recognizing both printed and handwritten text.

Overall, these studies demonstrate the potential of text recognition apps that use ML in accurately recognizing text from various sources, including handwritten text. Further research in this area could focus on improving the accuracy and speed of these apps and expanding their capabilities to recognize text in different languages and scripts.

In this paper author has proposed system is to efficiently recognize the offline handwritten digits with a higher accuracy than previous works done. Also, previous handwritten number recognition systems are based on only recognizing single digits and they are not capable of recognizing multiple numbers at one time. So, the author has focused on efficiently performing segmentation for isolating the digits.

Intelligent Systems for Off-Line Handwritten Character Recognition:

A Review Handwritten character recognition is always a frontier area of research in the field of pattern recognition and image processing and there is a large demand for Optical Character 4 Recognition on hand written documents. This paper provides a comprehensive review of existing works in handwritten character recognition based on soft computing technique during the past decade.

4. Requirement Analysis

Data requirements: Large and diverse datasets of labeled images are required to train machine learning models for text recognition. The dataset should cover a wide range of fonts, languages, writing styles, and image qualities. Accuracy requirements: The text recognition system must be accurate and efficient, achieving high accuracy rates for text extraction from various sources, such as images, videos, and other digital media. Speed requirements: The app must be able to process large volumes of images or videos in real-time, and provide fast and efficient text recognition capabilities. Usability requirements: The app must have a user-friendly interface that is easy to use, and it must be accessible to a wide range of users, including those with disabilities. Security requirements: Appropriate security measures must be implemented to ensure the security and privacy of user data. Integration requirements: The app must be able to integrate with other applications and services, such as translation, transcription, and information retrieval, to enhance its usability and usefulness. Scalability requirements: The app must be scalable to handle large volumes of images or videos, and it must be able to handle multiple users simultaneously. Platform requirements: The app must be able to run on various platforms, including mobile devices and web browsers, and it should support multiple languages. Testing requirements: The app must be thoroughly tested using various metrics such as precision, recall, and F1 score to ensure that it is accurate and efficient. Maintenance requirements: The app must be regularly maintained and updated to ensure that it remains functional and up-to-date with the latest technologies and data protection regulations.

5. Functional Requirements

Image and video recognition: The app must be able to recognize and extract text from images and videos. Text extraction: The app must be able to extract text from images and videos in various languages and fonts. Reprocessing: The app must use computer vision techniques to detect text regions within the image or video and pre-process the media to enhance text visibility before recognition. Post-processing: The app must correct errors and inconsistencies in the extracted text using spell-checking, grammar- checking, and language models to improve recognition accuracy. Translation: The app should be able to translate the extracted text into different languages. Transcription: The app should be able to transcribe the extracted text into speech. Information retrieval: The app should be able to search for information related to the extracted text. User authentication: The app should have a user authentication system to ensure that only authorized users can access the app and its features. User feedback: The app should

allow users to provide feedback on the accuracy and performance of the text recognition system, which can be used to improve the system over time. Compatibility: The app should be compatible with different platforms, including mobile devices and web browsers, and should be able to support multiple languages. Performance metrics: The app should be able to measure its performance using metrics such as precision, recall, and F1 score, and display these metrics to users. User interface: The app should have a simple and intuitive user interface, with clear instructions on how to use the app and its features.

6. Flow Chart

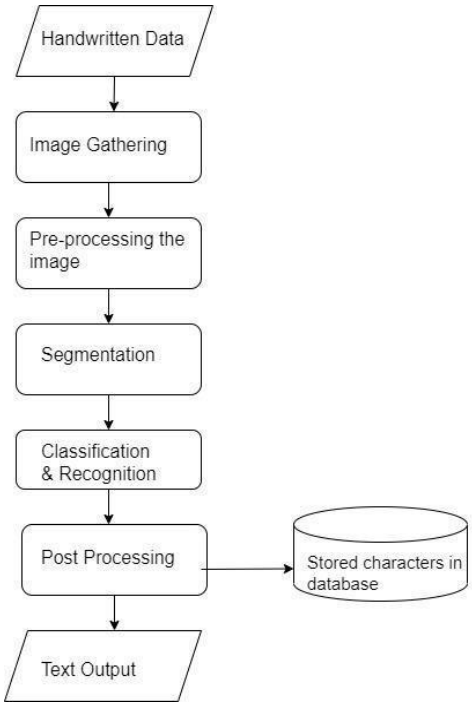

Fig. 5.1 Flow Chart

Above, the user can use the GUI for capturing the handwritten document. The user can load the image from the gallery or capture by camera. After opening the image, changes can be dome like cropping and rotating and clicking on the done button, the cropped image is sent to the OCR engine.

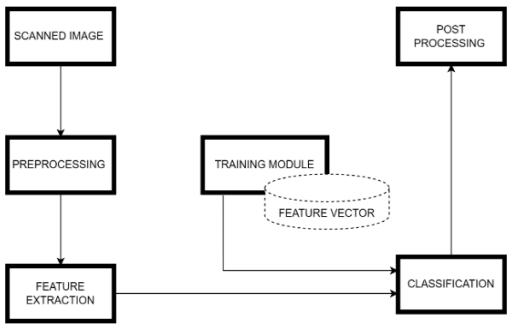

Fig. 6.1 System Architecture

The text converted from the image is displayed on the screen. This text is in editable form. Next, an option to save this converted text as text file(.txt) This text file which consists of converted text will be saved in the OCR folder which would be created in the internal storage of the device. After finishing all the process, can open, edit and save that text file any time in future. This text file can also be saved as a pdf file.

7. Tools / Platform

Software Requirement

- OS – Windows 10
- IDE – Visual Studio code Library/API/Framework: Firebase ML, Text Recognition API.
- Designing tool – Draw.io

Firebase ML Kit

ML Kit is a mobile SDK that brings Google's machine learning expertise to Android and iOS apps in a powerful yet easy-to-use package. Whether you're new or experienced in machine learning, you can implement the functionality you need in just a few lines of code. There's no need to have deep knowledge of neural networks or model optimization to get started. On the other hand, if you are an experienced ML developer, ML Kit provides convenient APIs that help you use your custom TensorFlow Lite models in your mobile apps. ML Kit APIs works both on the device and on the cloud. The on-device APIs are designed to work fast with no internet connection. On the other hand, cloud-based APIs uses Google Cloud Platform's machine learning technology which gives more accurate results but requires an internet connection.

Features:

- Both single-language text and multi-language text are supported.
- Detects the language of text without Internet connection.
- Text recognition.
- Face detection.
- Barcode scanning.
- Image labelling.
- Object detection & tracking

Text Recognition the ML Kit Text Recognition API can recognize text in any Latin-based character set. It can also be used to automate data-entry tasks such as processing credit cards, receipts, and business cards. Key capabilities

- Recognize text across Latin-based languages Supports recognizing text using Latin script.
- Analyze structure of text Supports detection of words/elements, lines and paragraphs.
- Identify language of text Identifies the language of the recognized text.

- Small application footprint On Android, the API is offered as an unbundled library through Google Play Services.
- Real-time recognition Can recognize.

7. Results

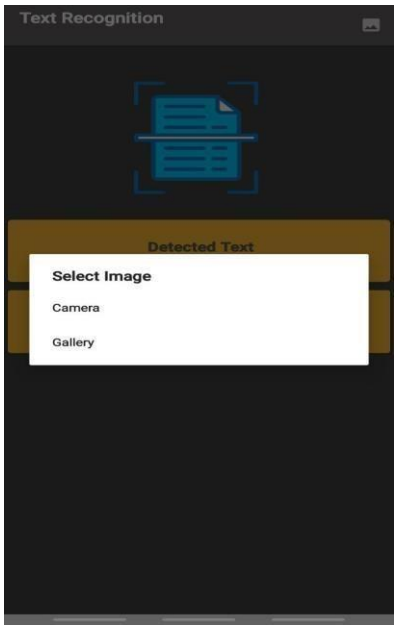

Fig 7.1 Home Page

The sub-window shows three icons. First showing the camera second showing the gallery and third showing photo applications. On clicking any of these three icons, the image is selected to process

Fig 7.2 Image Gathering

After opening the image, changes can be done like cropping and rotating and clicking on the done button, the cropped image is sent to the OCR engine. The text converted from the image is displayed on the screen. This text is in editable form.

Fig 7.3 Image Processing

On clicking the camera icon, the sub-window open. The sub-window shows three icons. First showing the camera second showing the gallery and third showing photo applications. On clicking any of these three icons, the image is selected to process

Conclusion

In our work we designed a mobile application and completed its development by applying engineering knowledge which provides an approach for helping human being. We believe that the application which we developed would be beneficial and efficient in today's world, solves the societal problem and would help in securing important documents. We have used modern tool like Android Studio to implement the project.

Acknowledgement

We would like to convey our sincere gratitude to Mr. Ankit Bhoyar and Mrs. Alka Shrivastava our esteemed project guides, for serving as the foundation of our endeavor. Their persistent support through times of uncertainty and doubt has enabled us to continue with this effort. We appreciate the assistance and support of the Computer Science and Engineering Department.

References

[1]. A Novel Approach to Classify English Handwritten Character Using Convolutional Neural Networks by Rupesh Sharma, Prashant Sharma, Naresh Purohit, International Journal of Advanced Computer Science and Applications, 21- 23 Mar. 2021, West Yorkshire, Malayasia.

[2]. Handwritten character recognition using Convolutional neural networks by Verdha Chaudhary, Mahak Bansal and Guarav Raj, IEEE Seventh International Conference on Intelligent Systems Design and Applications, 09 Dec. 2020, Rio de Janeiro, Brazil.

[3]. CNN Based Common Approach to handwritten character recognition of multiple scripts by Durjoy Sen Maitra, Ujjwal Bhattacharya and Swapan K. Paru, IEEE Proceedings of 2nd International Conference on Document Analysis and Recognition, 20-22 Oct. 2020, Tsukuba, Japan

[4]. Handwritten Character Recognition from Images using CNN-ECOC by Mayur Bhargab Bora, DinthisrangDaimary, Khwairakpam Amitab, Debdatta Kandar., International Conference on Innovative Practices in Technology and Management, 17-19 Feb. 2020, Noida, India

[5]. Building a Text Recognition App with Machine Learning" by Michael Galarnyk: This blog post provides a step-by-step guide to building a text recognition app using machine learning techniques such as Optical Character Recognition (OCR) and Convolutional Neural Networks (CNN).

[6]. "Text Recognition with OCR and Machine Learning in Python" by Prateek Joshi: This article covers the basics of OCR and how to use machine learning algorithms to improve the accuracy of text recognition.

[7]. "Text Recognition using Machine Learning: A Review" by Vikram Singh and Rohit Kumar: This research paper provides a comprehensive review of various machine learning techniques used for text recognition.

[8]. Tesseract OCR: Tesseract is an open-source OCR engine that can be used for text recognition in various programming languages. It is one of the most widely used OCR engines and has been integrated into several text recognition apps.