International Journal of Scientific Research in Engineering and Management (IJSREM)
Volume: 08 Issue: 10 | Oct - 2024 SJIF Rating: 8.448 ISSN: 2582-3930

Image-to-Text Scanning and Language Translation Android Application Using Machine Learning and Kotlin

Tanaya Sawant
IT Student
HSNC University
Mumbai, India
geeknerdy383@gmail.com

Dr. Rakhi O. Gupta
Head of the Department of IT
HSNC University
Mumbai, India
rakhi.gupta@kccollge.edu.in

Ms. Nashrah Gowalker
Assistant Professor of IT,
HSNC University
Mumbai, India
nashrah.gowalker@kccollege.edu.in

Abstract –

Mobile phones are the most used devices in today's world. Applications that help us solve real-world problems make our lives easier. With hardware installed within them, they are no longer just transmission devices. Still, they can perform multiple complex tasks such as capturing images through a camera lens, recording high-quality videos, detecting live locations, and much more. Hence, it makes our lives simpler and much easier. Using advanced technology, we can scan text and recognize it. A key point here is that we can translate the extracted text into any language we choose. People from foreign countries find it difficult to communicate in a different to communicate in a different language. Keeping all this in mind, the proposed solution to this is an application that can scan an image, then detect and extract all the text from that image and provide it to the user using ICR technology which is a type of OCR technology and then also convert it into the language of the user's choice, and this android application is implemented using Kotlin and Java programming language Therefore, proving it effective and useful for the user.

Keywords – OCR technology, ICR technology, detect, extract, scan, image,translate, Kotlin, Java, recognize,live-location, transmission, real-world.

I. INTRODUCTION

The world contains too many significant messages and useful information but unfortunately, most of them are written in different official languages depending on the host country [1]. It makes it difficult for people belonging to different countries to read and understand text from a signboard or any important message from documents. Images carrying important information in different languages are often hard to understand and might cause the important text to be missed. Hence text recognition comes into play.

Text recognition is the process of detection, extraction, and identification of text in images [3]. After extracting the text, the user will be able to translate the text into whichever language he/she wants. Thus, making it easier, less time-consuming, and overcoming the language barrier problem. Google Translate can translate texts from any language to any. other language, but the limitation is that we have to feed the text manually to the Google Translate application [4]. This proposed solution can recognize and extract the text automatically from an image already stored in a folder or an image just recently clicked and will be translated according to the user's requirements. To implement this technique of text detection and extraction, we will have to make use of OCR. OCR stands for optical character recognition. With the use of OCR technology, camera-captured or scanned documents can be converted into machine-editable soft copies that can be reproduced, edited, searched, and transported with ease [2]. OCR helps machines recognize characters automatically using optical tools similar to a human being using eyes to see or detect an object. OCR extracts text from documents or images within a fraction of a second, making the work easier for the user. For example, if the user is a university student and has PDFs or such stuff and wants to extract text out of it to use the text from those files or those PDFs or images of books or notes, then this Android application will scan the image or file using machine learning and will detect the text and will present it to the user in the form of reusable text. In addition, with the feature of language translation. This proposed solution will be implemented using Kotlin technology.

Kotlin is a perfect programming language for Android applications without presenting errors like compatibility, performance, interoperability, and compilation time. This proposed solution is a perfect blend of machine learning and Android application development.

II. LITERATURE REVIEW

A huge number of proposals and research papers have been published on text detection from images using OCR and language translation. The purpose of a literature review is to disclose any gaps existing in the literature. Text recognition and extraction from real-world images is a complex task due to the multi-orientation background issues, misaligned text, or blurred images. If the above issues are fixed, then the OCR system's accuracy can be enhanced. In the translation phase, the system's manageability and case of use help to recognize the different

languages and help to translate them into the selected language by the user.

The main limitation of OCR technology is that it provides high accuracy with printed text. OCR is mostly focused on converting the printed text into digital text. On the contrary, ICR, which stands for Intelligent Character Recognition, excels in machine-printed text as well as handwritten text and converts both texts into digital, readable, and reusable formats. That's where we get the "intelligent" part of the system. ICR is slightly more specific compared to OCR. The proposed solution is implemented using ICR technology, thus reducing time and providing more accuracy to the user. All the articles and research papers based on OCR text detection and translation are reviewed to reach an appropriate solution on how the above technologies can play a crucial role in user experience.

III. METHODOLOGY

A. ICR

OCR is a technology that allows users to convert text or documents captured by an input device into an Editable, searchable, and reusable data types for further image processing [1]. **Intelligent Character Recognition is a type of OCR, but it is more flexible, reliable, and specific than OCR.** It proves to be more efficient than OCR because it not only converts handwritten texts or image-captured texts but also works with different fonts. That's where the "intelligent" part comes into the picture. Over time, it has proven to be smarter and more accurate than OCR.

Characteristics	OCR	ICR
Text recognition of printed text format	Easily recognizes text with high accuracy	Processes handwritten text
Text recognition of handwritten text format	Limited contextual knowledge is mostly used for printed text	Works best with detecting handwritten texts as well as printed texts.
Cost	Cost-effective	More expensive than OCR
The result concerning time.	Takes a small amount of time to print the output. Time-consuming compared to ICR.	Print the output on the screen within a fraction of a second. Thus less time-consuming.

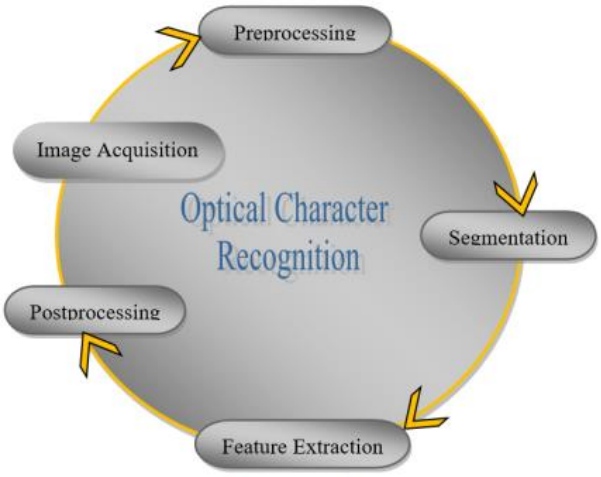


Fig.1. OCR Algorithm

Source: (Vaibhav.V.Mainkar; Jyoti A.Katkar; Ajinkya B.Upade; Poonam R.Pednekar, 2020)

1. Image pre-processing:

The images are captured through a camera or some other device. During image capture and text extraction, the scanned images or the images stored in the phone's memory are easier to detect than those captured by the camera because they may contain issues like low brightness and blurriness. Therefore, some techniques like image enhancement, binarization, and noise reduction are needed in the pre-processing phase to increase the performance and accuracy of a character recognition system [1].

2. Text segmentation:

Text segmentation splits the text into smaller and more meaningful units, such as words, lines, sentences, and topics. It is useful for humans while reading texts and for man-made operations that are installed in computer systems, mostly used for NLP. Natural Language Processing (NLP) is nothing but the process of how to segment a text correctly. Text segmentation can be used in various applications such as object detection, vehicle license plate number identification, document recovery, and much more. There are many well-known segmentation methods available, which are projection, region growth, tracing contour, etc. [1].

3. Text recognition:

In this stage, after the text or character is segmented, it is organized by reducing the noise. After this, the OCR extracts the text, detects it, and further identifies it. Even though the OCR is convenient, it does have its disadvantages. The major ones occur while detecting text from an image that is captured with these issues of low light, minimal brightness level, misalignment of text, formatting errors, and blurriness. ICR can detect the text despite all the issues. Hence, ICR is better in accuracy than OCR.

With the growth of fast and advanced technology, mobile phones have the built-in feature of capturing high-resolution images of approximately 720x1280 pixels in size. This feature enables OCR to detect text faster and proves to be suitable. Despite all this, mobile phones still have their drawbacks. One of the drawbacks is limited power to run complex software like OCR engines due to limited hardware and memory resources, unlike desktops [1]. The major one is sensitive information exposure. This drawback can lead the confidential data to become vulnerable to cyberattacks, and security breaches might take place. Thus, this drawback should also be taken into consideration.

B. Translation

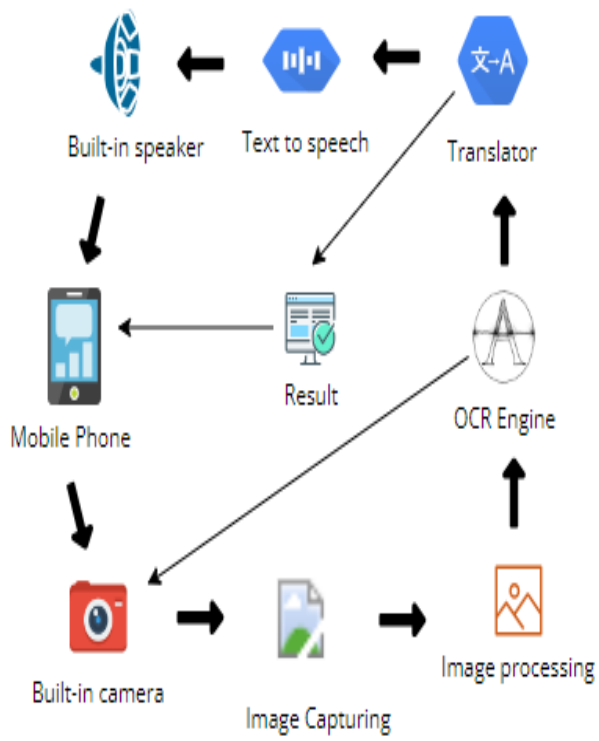
Translation simply means the procedure of conversion by copying text or speech from one source-language text into a different target-language text. In this proposed solution application, there are two ways to add the text inside the text field: by simply writing the text, or we can add it by using the built-in mic functionality. Text detection and language translation are mostly implemented and are available in web-based applications rather than Android-based applications.

This proposed solution is a combination of ICR and translation in various languages. The translation is done for the text, which is obtained after the recognition of an image [3]. However, there are issues in this phase too. The translations made sometimes are not reliable and slow because of translating word by word sequentially [1]. The next main drawback of translation is also that during translation sometimes the meaning of the words might change and may guide a dissimilar meaning. Plus, translation can be time-intensive and expensive. The results might not always be accurate.

Online translation services are more reliable and efficient and only require a minute amount of data to be embedded into the code's backend. An Internet connection is mandatory to keep the translator updated to receive accurate translations.

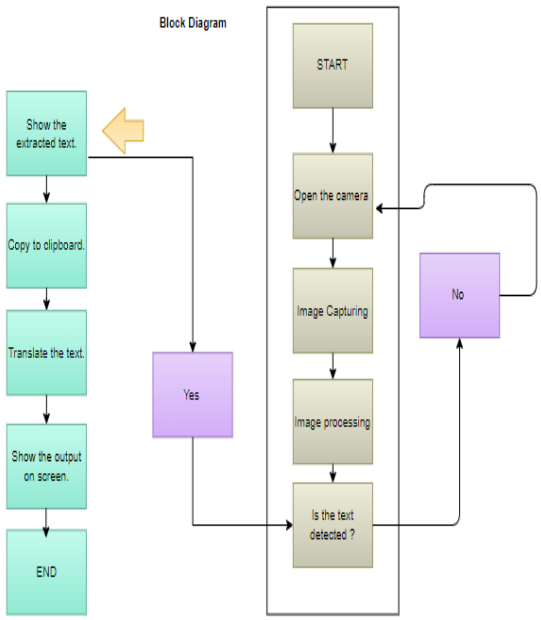
IV. SYSTEM DESIGN

A. Architecture Design



Architecture of the proposed solution

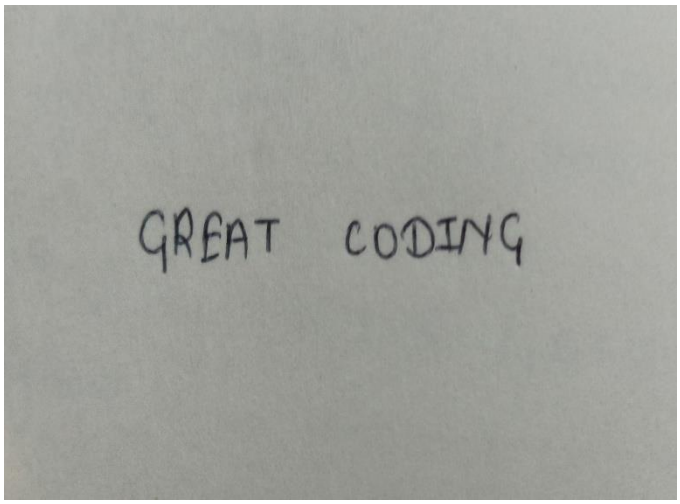
B. Block Diagram



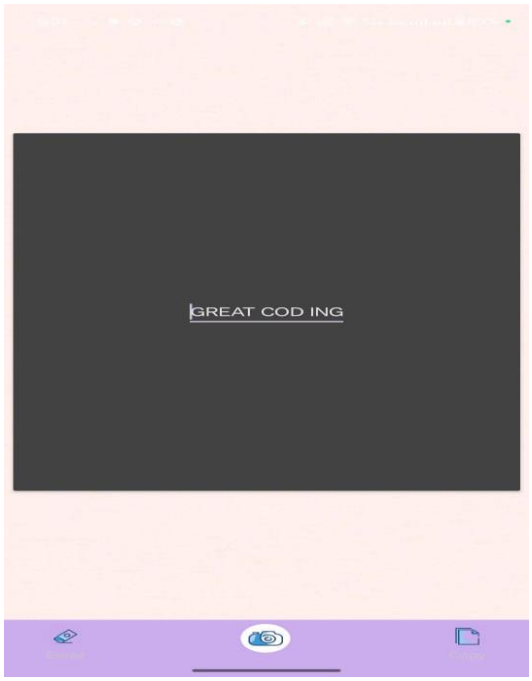
V. TESTING

The final deliverable is tested on an Android device. The following figures show the UI and text extraction:

A. Handwritten text for text detection



B. Extracted text displayed on the output screen.



A study is conducted to understand if the application satisfies the user's needs.

	Functionality		Accuracy	Performance	Quality
User 1	YES		YES	YES	BEST
User 2	YES		YES	-	GOOD
User 3	NO		Not exactly accurate	NO	NOT GOOD
User 4	YES		YES	YES	GREAT

VI. CONCLUSION

In this paper, an Android application that detects text from images and translates them into a target language for the user has been developed. This proposed solution will provide better accuracy with translation. This also provides higher efficiency than other recognition tools. Further on, it also resolves the issue of language barriers for the users traveling into different countries. Built-in mic functionality will ensure that visually impaired users can also use this application to access the printed text and not miss out on listening to important messages. This application will be useful for travelers, as they will no longer face the problem of a different language while visiting foreign countries, which represents information in their native language.

VII. LIMITATIONS

Limitations are the shortcomings in a research paper that are not in the researcher's control. This proposed solution contains the following limitations:

- Time constraints: The time available to study the research problem was less and the changes happening over time were not possible for the researcher to meet the paper's deadlines.
- Cost issues: This solution is implemented in Android Studio and hence accessible for Android users only and not for other operating systems.
- Sample size: Limited sample size therefore there was difficulty in predicting the outcome. The larger the data size the more precise and accurate the outcome.

VIII. FUTURE SCOPE

In the future, this application can be improved by enhancing the OCR technology and by using better translation services, which will significantly improve the user experience of the application as well as the performance and quality of this application.

ACKNOWLEDGEMENT

The author is very grateful to everyone who has provided expertise and has helped in all kinds of aspects of this study. Thankful for all the university management support and resources provided throughout this research. A very special thanks to our head of the department, Dr. Rakhi Gupta, and assistant professor, Ms. Nashrah Gowalker, for the guidance and motivation and for allowing working on this paper.

REFERENCES

[1] Kartik Bhasin, Arjun Goel, Gaurav Gupta, Sanjay Kumar Singh, and Abhijit Bhowmick,'A secure mobile application for speech to text conversion using AI techniques',2023.

[2] Utsav Poudel, Aayush Man Regmi, Z. Stamenkovic, S.P. Raja,' Applicability of OCR Engines for Text Recognition in Vehicle Number Plates, Receipts and Handwriting', 2023.

[3] Roseline Oluwaseun Ogundokun, Joseph Bamidele Awotunde, Sanjay Misra, Tobe Segun-Owolabi, Emmanuel Abidemi Adeniyi, and Vivek Jaglan,'An android-based language translator application, 2021.

[4] Akinbowale Nathaniel Babatunde, Christiana Oluwakemi Abikoye, Abdulkarim Ayopo Oloyede, Roseline Oluwaseun Ogundokun, Afeez Adeshina Oke, and Hafsat Omolola Olawuyi,' English to Yoruba short message service speech and text translator for android phones',2021.

[5] Dr. S. Revathy and Sakhayadeep Nath,' Android live text recognition and translation application using Tesseract', IEEE, ICICCS,2020.

[6] Bambang Purnomosidi Dwi Putranto, Robertus Saptoto, Ovandry Chandra Jakaria, and Widyastuti Andriyani, 'A comparative study of Java and Kotlin for Android mobile application development', ISRITI 2020.

[7] Dr.Vasudev T, Chaithanya K N, Madhushree R, Apoorva B P and Arpitha N,' Translation of Recognized Text from Images', IJAEM 2020.

[8] Prof. Vaibhav.V.Mainkar, Ms.Jyoti A.Katkar, Mr.Ajinkya B.Upade and Ms.Poonam R.Pednekar, 'Handwritten Character Recognition to obtain Editable Text', ICESC,1EEE,2020.

[9] B M Vinjit, Mohit Kumar Bhojak, Sujit Kumar, Gitanjali Chalak, 'A Review on Handwritten Character Recognition Methods and Techniques',2020.

[10] Thi Thi Zin, Moe Zet Pwint, and Shin Thant,' A Mobile Application for Offline Handwritten Character Recognition',2020.

[11] Rahul B. Panchal, Rutuja G.Sonawane, Himayat Shaikh and Piyush P.Gawali, 'Design of Text detection and Translation system for camera-based android smartphone', IJSRD 2015.

[12] Sathiapriya Ramia, TanYu Liong and Manoj Jayabalan, 'Detecting Text Based Image with Optical Character Recognition for English Translation and Speech using Android', IEEE 2015.