REAL-TILE MULTIPAGE DOCUMENT OCR FOR MOBILE APPLICATIONS

Submitted in partial fulfillment of the requirements for the award of
Bachelor of Engineering degree in Computer Science and Engineering

By

CSE-037

ARUNKUMAR V (Reg.No -- 39110086)
MOHAMED ASRAF (Reg.No – 39110626)

DEPARTMENT OF COMPUTER SCIENCE AND ENGINEERING

SCHOOL OF COMPUTING

SATHYABAMA

INSTITUTE OF SCIENCE AND TECHNOLOGY
(DEEMED TO BE UNIVERSITY)
Accredited with Grade "A" by NAAC
JEPPIAAR NAGAR, RAJIV GANDHISALAI,
CHENNAI - 600119

APRIL - 2023

SATHYABAMA

INSTITUTE OF SCIENCE AND TECHNOLOGY

(DEEMED TO BE UNIVERSITY)

Accredited with ¯A" grade by NAAC

Jeppiaar Nagar, Rajiv Gandhi Salai, Chennai - 600 119

www.sathyabama.ac.in

DEPARTMENT OF COMPUTER SCIENCE AND ENGINEERING

BONAFIDE CERTIFICATE

This is to certify that this Project Report is the bonafide work **ARUNKUMAR V (39110086) and MOHAMED ASRAF (39110626)**who carried out the Project Phase-2 entitled **"REAL-TILE MULTIPAGE DOCUMENT OCR FOR MOBILE APPLICATIONS"** under my supervision from December 2022 to April 2023.

Internal Guide
Dr. ROOBINI M.S, M.E., Ph.D.

Head of the Department

Dr. L. LAKSHMANAN, M.E., Ph.D.

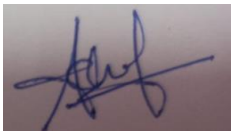
Submitted for Viva voce Examination held on 24/04/2023

Internal Examiner **External Examiner**

DECLARATION

I, **ARUNKUMAR V (Reg.No- 39110086),** hereby declare that the Project Phase-2 Report entitled "**REAL-TILE MULTIPAGE DOCUMENT OCR FOR MOBILE APPLICATIONS"** done by me under the guidance of **Dr. M.S.ROOBINI, M.E., Ph.D.** is submitted in partial fulfillment of the requirements for the award of Bachelor of Engineering degree in **Computer Science and Engineering**.

DATE: 24.04.2023



PLACE: Chennai **SIGNATURE OF THE CANDIDATE**

ACKNOWLEDGEMENT

I am pleased to acknowledge my sincere thanks to **Board of Management** of **SATHYABAMA** for their kind encouragement in doing this project and for completing it successfully. I am grateful to them.

I convey my thanks to **Dr. T. Sasikala M.E., Ph. D**, **Dean**, School of Computing, **Dr. L. Lakshmanan M.E., Ph.D.,** Head of the Department of Computer Science and Engineering for providing me necessary support and details at the right time during the progressive reviews.

I would like to express my sincere and deep sense of gratitude to my Project Guide **Dr. ROOBINI M.S, M.E., Ph.D.** for her valuable guidance, suggestions and constant encouragement paved way for the successful completion of my phase-1 project work.

I wish to express my thanks to all Teaching and Non-teaching staff members of the **Department of Computer Science and Engineering** who were helpful in many ways for the completion of the project.

ABSTRACT

Many businesses, including document digitization, information retrieval, and machine translation, use OCR, a method for extracting text from photographs. Since mobile technology has advanced so quickly, it has become more important to investigate OCR on mobile devices. OCR performance on mobile devices is still hampered by their CPU and memory capacities. In this study, we present a real-time multipage document OCR solution for mobile apps. The recommended method is designed to process large documents with many pages on mobile devices fast and effectively. We provide a special technique for streamlining the OCR process by cutting the page into smaller pieces and processing them simultaneously on the mobile device. translation. We also offer a novel approach that increases recognition accuracy by utilizing deep learning. The proposed system is put to the test against a variety of multipage texts, and the results show that it can compete with traditional desktop-based OCR systems in terms of performance. Furthermore, we evaluate the system's operation on a range of mobile devices and find that the vast majority of them can successfully use the suggested system. The proposed system offers a wide variety of applications, including document management, information retrieval, and machine translation. It is a promising method for real-time multiple pages document OCR on portable devices.

CHAPTER 1

INTRODUCTION

OCR, a method for extracting text from photographs, is widely used in a variety of fields, including document digitization, information retrieval, and machine translation. The rapid advancement of mobile technology has increased the significance of OCR on mobile devices as a research issue. OCR performance on mobile devices is still hampered by their CPU and memory capacities. This is especially true for large and complex manuscripts that have multiple pages. We recommend a multipage real-time OCR solution for mobile apps in this study. The suggested method is designed to process large documents with multiple pages on mobile devices fast and effectively. We offer a special technique for optimising the OCR process by cutting the page into smaller parts and processing them simultaneously on the mobile device. We also provide a novel approach that boosts recognition accuracy by utilising deep learning techniques. The proposed system is put to the test against a variety of multipage texts, and the results show that it can compete with traditional desktop-based OCR systems in terms of performance. The recommended system can operate efficiently on the majority of mobile devices, according to our assessment of the system's usability on various devices. Our recommended method may be applied in a number of fields, such as mobile machine translation, document management, and information retrieval.However, primitive communication technology, handwriting and meaningful and simple collection of data will continue to occupy a leading position until the beginning of a paperless society. The representatives of contemporary communication the methods are examples of Personal Digital Assistants (PDA), Electronic Pen (E-Pen), Electronic Mail (Email), mobile phone, computer, etc. During the last few decades, the efforts have been highly successful in machine-printed digitization, OCR (Optical Character Conversion). Handwriting recognition techniques differ in two aspects: online or vibrating and offline or static depending on the methods used by handwritten information to enter identification schemes. Dynamic or moving writing is often used to digitize the travel paths of a pen, using a single pen on an electronic liquid crystal display floor with a sensor. The trajectories of movements, texts are recorded at points by two dimensional coordinates. However, an overline and static scanner is used to scan and digitize a paper that contains handwritten data and stores only two-dimensional data. The aim of this initiative is to further investigate the

classification of handwritten texts and digital transmission of handwritten texts. We were challenged in this initiative to classify an image from each manuscript that maybe italic or block type. This document can be coupled with algorithms that process concept images into a specified line image, which in turn can be coupled with algorithms to convert the line images into a specific line photo. handwritten picture

CHAPTER 2

LITERATURE SURVEY

T. T. Zin, M. Z. Pwint and S. Thant, "A Mobile Application for Offline Handwritten Character Recognition," 2020 Users' penned letters and words can be recognised by a computer programme called handwritten character recognition. In this study, we provide an offline deep learning architecture for character identification from handwritten text. The Android application for the proposed system also made use of OpenCV and TensorFlow Lite. The proposed system is designed to be utilised as a teaching aid to assist children in kindergarten through primary school, notably forwriting and learning practise. For training and testing purposes, the local handwriting dataset is used. It includes symbols, English letters, and digits gathered from students.

Y. Gao, L. Jin, C. He and G. Zhou, "Handwriting Character Recognition as a Service: A New Handwriting Recognition System Based on Cloud Computing," 2011. In this paper, we propose a novel architecture for providing handwritten character recognition as a service (HCRaaS) via the Internet using cloud computing technologies. Using the suggested Cloud-based recognition platform, it is possible to put many cutting-edge algorithms, such as the modified quadratic discriminant function (MQDF) and SVM classifier used for huge character recognition, writing adaptation technology, and handwriting Chinese word/text line recognition, into practise. Users of various mobile devices are no longer bound by local storage and computing resource restrictions thanks to the advantages and characteristics of HCRaaS, and they can also benefit from increased identification accuracy and personalised service with minimal hardware investment.

N. Shorim, T. Ghanim and A. AbdelRaouf, "Implementing Arabic Handwritten Recognition Approach using Cloud Computing and Google APIs on a mobile application," 2019. Apps for cloud computing that automatically recognise Arabic handwriting are becoming more and more in demand, and they become even more important when they are connected with mobile applications. It is a group of applications containing a number of challenging components. a smartphone app that uses Arabic handwriting recognition and is cloud-based. It is imperative to provide this service to non-Arabic speakers, especially if they are visiting countries that speak Arabic. Our approach is the first to create a mobile cloud application that recommendsa multi-phase hybrid classifier for reading Arabic Handwriting text. The current approach's recognition component can handle big databases and challenging calculations because to the multi-stage classifier it uses.

V. V. Mainkar, J. A. Katkar, A. B. Upade and P. R. Pednekar, "Handwritten Character Recognition to Obtain Editable Text," 2020. More and more information from handwritten papers is being archived for use in the future. It is simple to store the data by taking a picture of the handwritten paper and storing it as an image file. The method used to translate handwritten data into electronic representation is called "Optical Character Recognition." Some of the steps include pre-processing, segmentation, feature extraction, and post-processing. Numerous scholars have used OCR to identify characters. In this method, the document is photographed using an Android phone, and OCR does the remaining tasks. It is designed to recognise handwritten data and generate text that may be edited. The information that the writer must enter will determine the system's outcome.

S. Dome and A. P. Sathe, "Optical Charater Recognition using Tesseract and Classification," 2021. Using a technique or technology known as optical character recognition, text contained in a digital image may be found (OCR). The majority of the time, it is employed to convert transcribed, handwritten, or printed text into editable text data. While technology advances quickly, people need tools that are quick, convenient, and reliable so they can fulfil their daily demands. OCR technology may be used to read and verify handwritten documents, even though it was originally developed to read printed text.

B. H. Rao et al., "MTESSERACT: An Application for Form Recognition in Courier Services," 2022. Manually entering data from handwritten forms into a system takes time and is error-prone. To prevent this situation, the suggested approach extracts text from handwritten forms and saves the acquired data in a CSV file. This approach minimises errors and cuts down on pointless work delays. In this research, text is recovered from handwritten form shots by first segmenting the images, and then applying modified tesseract algorithms on the segmented data. Formatted input can be segmented to achieve better results. Pytesseract, an OCR tool, is used for text extraction.

Park, S. S., & Chung, K. S. (2022). Regularized Convolutional Neural Network for Highly Effective Parallel Processing. According to some study, parallel processingtechniques should be used to accelerate OCR on mobile devices and boost productivity. These techniques break the document into little fragments, which are then analysed simultaneously on the mobile device. Other studies have recommended using OCR algorithms that are simple to use to improve mobile phone OCR efficiency.

2.1 Inferences from Literature Survey:

A collected image can be transformed using OCR technology into a machine-editable format[8]. The recommended methodology uses a multi-stage process. The first step is to scan the document pages using the camera on a mobile device or a local storage device. Preprocessing, which involves rotating and cropping the scanned image, is the next stage. The final phase is optical character recognition, which is the third level. The proposed methology is well discussed. The technique, called OCR, makes use of an Android phone. A camera may be used to take pictures of handwritten documents. This is just how scanning functions. A digital replica of the original image can be made by scanning. The text is written in black characters on a white backdrop, with the original images.

2.2 OPEN PROBLEMS IN EXISTING SYSTEM

S.no	Technique	Pros	Cons
1	Character Extraction	• Simple form of technique to recognize the handwriting.	• Problem becomes non-trivial if the character overlaps such that there is no gap for character extraction.
2	Character Recognition: i) Neural Network ii) Feature Extraction	• It can learn and model nonlinear complex relationships from initial inputs and their relationships. • Removes redundancy from data and enhances the class patter variability.	• Fails if the property learnt from previous input is not important for the current data. • Needs more time as the characteristics are needed to be minutely verified.
3	SVM (Support Vector Machine)	• Beneficial is there is no idea of the data. • Risk of overfitting is less. • Works good for unstructured and semi structured data. • With the application of Kernel function complex problem can be solved.	• Long time on training the datasets. • Difficult to understand and interpret final model. • Choosing kernel function is not easy.
4	KNN (K-Nearest Neighbor)	• It is simple technique. • It requires only a	• It requires greater processors and physical memory.

		small training dataset with small number of training samples. • It is instance-based learning approach, therefore, requires less processing time	• Higher time is consumed during the classification of images.
5	MLP (Multilayer Perceptron Neural Network)	• Sufficient hidden nodes with the two-layer backpropagation network prove to be a universal approximation. • Gives the required decision based on training the datasets.	• It gets struck in local minima as the minima where it stops; it cannot guarantee that it is a global minimum.
6	Back Propagation	• It is simple to compute. • It descends stochastically in weight space.	• Slow and Inefficient. • A large amount of inputs and outputs are there but relating it to the output is difficult.
7	GRU (Gated Recurrent Unit)	• GRU use internal memory functionality to store and filter the data using update and reset gates. • The last output can easily be replicated by saturating the input gate to zero.	• At each step, the output is the same and the hidden state can cause problems to learn latent sequence feature that is not directly linked to sequence elements.

CHAPTER 3
REQUIREMENT ANALYSIS

3.1 SOFTWARE REQUIREMENTS SPECIFICATION DOCUMENT

REQUIREMENT	SPECIFICATION
Anaconda Navigator	You must have anaconda installed in your device prior to begin.
Spyder, Jupyter Notebook, Flask Framework	1. One should have Spyder and Jupyter notebook. 2. One should install flask framework through anaconda prompt for running their web application 3. We need to build the model using Jupyter notebook with all the imported packages.
Web browser	For all Web browsers, the following must be Enabled: • cookies • JavaScript
VS code	You must have VScode installed in your device prior to begin.
Flutter SDK	One should have installed flutter SDK to deploy model.
Dart SDK	One should have installed Dart SDK to deploy model.

3.2 HARDWARE REQUIREMENTS SPECIFICATION DOCUMENT

REQUIREMENT	SPECIFICATIONS
Operating system	Microsoft Windows UNIX Linux®
Processing	Minimum: 4 CPU cores for one user. For each deployment, a sizing exercise is highly recommended.
RAM	Minimum 8 GB.
Operating system specifications	File descriptor limit set to 8192 on UNIX and Linux
Disk space	A minimum of 7 GB of free space is required to install the software.

CHAPTER 4

DESCRIPTION OF PROPOSED SYSTEM

OCR technology may be used to convert a captured picture into a machine-editable format[8]. The suggested methodology entails a multi-stage approach. The first step is scanning the document pages using a mobile device's camera or a local storage device's picture. The second step then entails preprocessing, which includes rotating and cropping the scanned picture. Optical Character Recognition is the third level. the last phase. The suggested methology is well discussed. The method, known as OCR, utilises an Android phone. Papers written by hand can be photographed using a camera. This is just how scanning works. Scanning allows for the creation of a digital copy of the original image. The text has the original photos in black letters on a white background. This technique turns the digital photo into a grayscale version. Pre-processing is one of the most important steps in character recognition. Systems can read grayscale photographs more easily as a result. By using a filter, the flaws in the images are eliminated. Pre-processing is essential for handwritten images as they are more vulnerable to noise. Greyscale conversion, noise reduction, binarization, thinning, skewing, and normalising are just a few of the pre-processing procedures.

4,1 ARCHITECTURE / OVERALL DESIGN OF PROPOSED SYSTEM

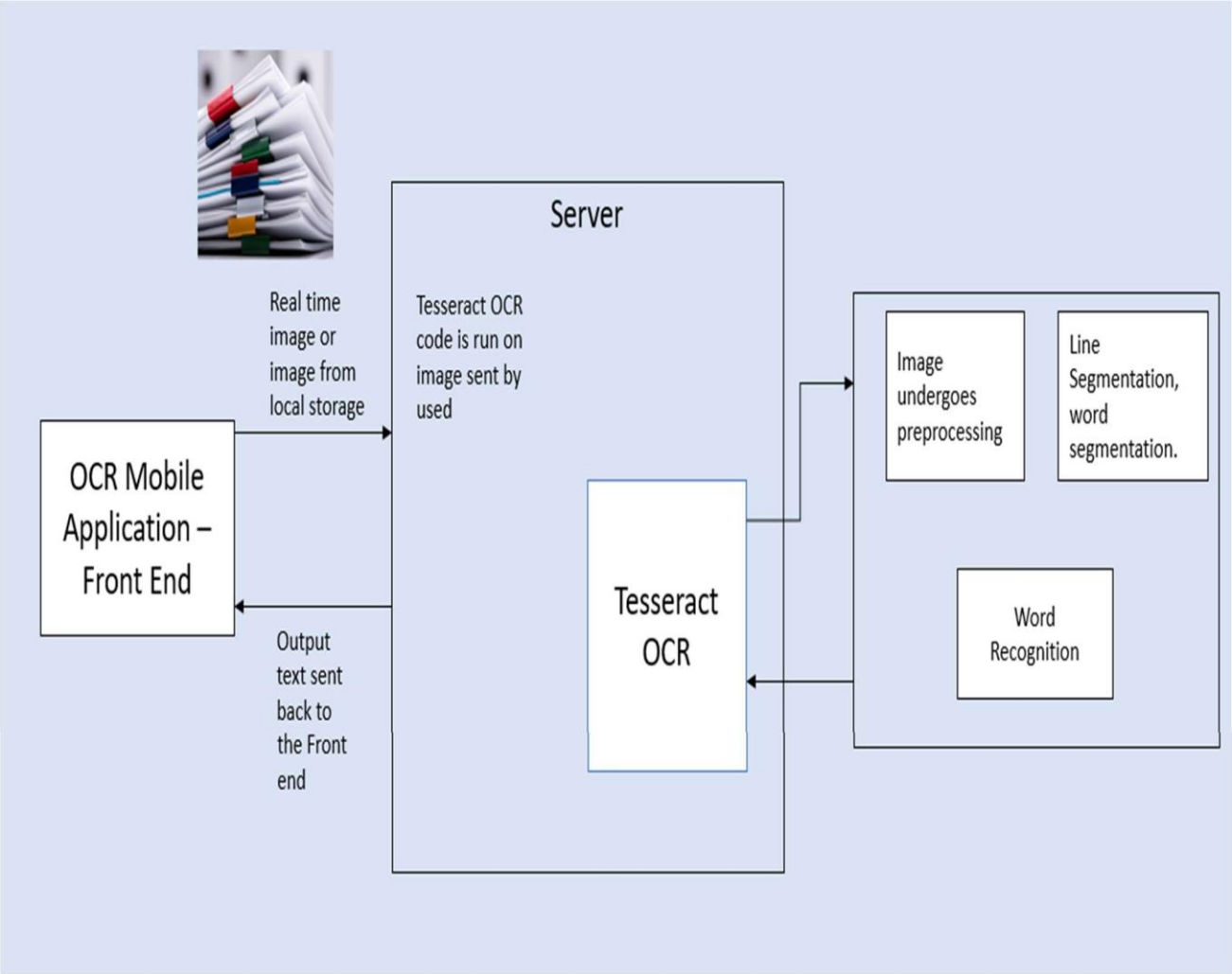
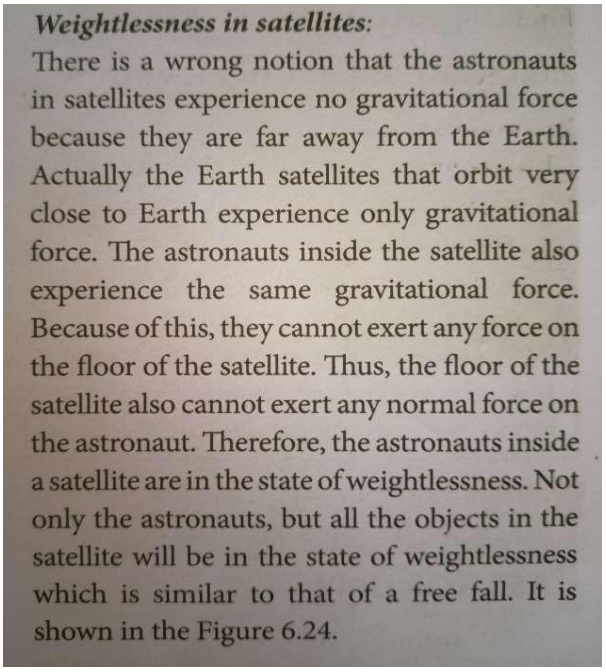


Fig 4.1: System Architecture for OCR

Chapter 5

RESULT AND CONCLUSION

A study on "REAL-TIME MULTIPAGE DOCUMENT OCR FOR MOBILE APPLICATIONS" would likely focus on the ability of the OCR technology to accurately recognize text from images of documents captured by mobile devices in real-time. The results of such a study could include the accuracy rates of the OCR system on different types of documents (such as contracts, receipts, and ID cards) and on different mobile devices. The study may also evaluate the speed of the OCR system, in order to determine the feasibility of real-time processing. Based on these results, the conclusion of the research paper could highlight the strengths and weaknesses of the OCR system, and its suitability for use in mobile applications. The conclusion could also provide recommendations for further research or improvements to the system. It is also important to note that the OCR technology is constantly evolving and new research will be necessary to improve the technology. In this section we will discuss the result from the project. The below figure shows the image captured from the capture which cropped and rotated.

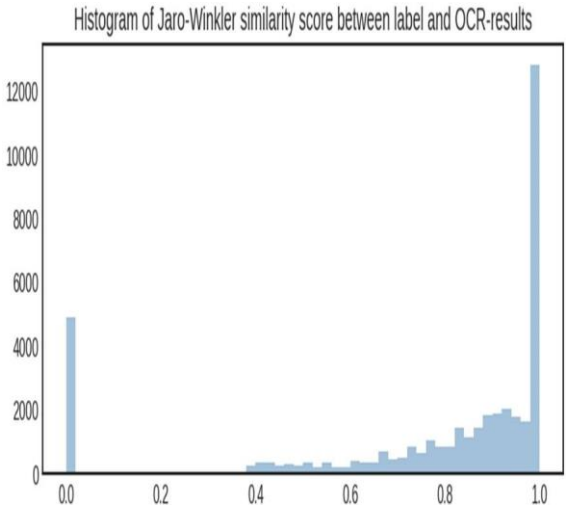


Weightlessness in satellites:
There is a wrong notion that the astronauts in satellites experience no gravitational force because they are far away from the Earth. Actually the Earth satellites that orbit very close to Earth experience only gravitational force. The astronauts inside the satellite also experience the same gravitational force. Because of this, they cannot exert any force on - the floor of the satellite. Thus, the floor of the satellite also cannot exert any normal force on ss es meena Therefore, the astronauts inside a satellite are in the state of weightlessness. Not only the astronauts, but all the objects in the satellite will be in the state of weightlessness which is similar to that of a free fall. It is shown in the Figure 6.24.

And this image is result from from tesseract ocr, the figure shows extracted text from the image. We have analyzed the performance metrics by using jaro winkler similarity score.

This shows in the following figures.

	FILENAME	IDENTITY	CLEAN_TEXT	SIMILARITY_SCORE
0	VALIDATION_0001.jpg	BILEL	BiLeL	1.000000
1	VALIDATION_0002.jpg	LAUMIONIER	LAUMONIER	0.946667
2	VALIDATION_0003.jpg	LEA		0.000000
3	VALIDATION_0004.jpg	JEAN-ROCH	JTEAN-ROCH	0.970000
4	VALIDATION_0005.jpg	RUPP	Ruer	0.733333
5	VALIDATION_0006.jpg	UNREADABLE	CODERNE/UAI;0642494W'neSey2	0.440741
6	VALIDATION_0007.jpg	PICHON		0.000000
7	VALIDATION_0008.jpg	DANIEL	DAMIEL	0.911111
8	VALIDATION_0009.jpg	JEREMY	JEREAMY	0.971429
9	VALIDATION_0010.jpg	JEAN-MICHEL	TEAN-MicHEL	0.939394



Real-time multipage document OCR research and development for mobile applications has the potential to significantly increase productivity and efficiency of people and enterprises in a range of sectors. Document scanning and OCR capabilities may now be used on the move thanks to the integration of this technology into a mobile application. To increase the OCR technology's effectiveness and accuracy and to make sure that it works with a variety of document formats and languages, more study and development is required. The incorporation of this technology in a useful mobile application has the potential to have a substantial influence in the sector given the rising need for mobile OCR solutions.

REFERENCE:

[1] Robby, G. & Tandra, Antonia & Susanto, Imelda & Harefa, Jeklin & Chowanda, Andry. (2019). Implementation of Optical Character Recognition using Tesseract with the Javanese Script Target in Android Application. Procedia Computer Science. 157. 499-505. 10.1016/j.procs.2019.09.006.

[2] X. -S. Vu, Q. -A. Bui, N. -V. Nguyen, T. T. Hai Nguyen and T. Vu, "MC-OCR Challenge: Mobile-Captured Image Document Recognition for Vietnamese Receipts," 2021 RIVF International Conference on Computing and Communication Technologies

(RIVF), Hanoi, Vietnam, 2021, pp. 1-6, doi: 10.1109/RIVF51545.2021.9642077.

[3] LeCun, Y., Bengio, Y., & Hinton, G. (2015). Deep learning. nature, 521(7553), 436-444.

[4] Kaur, Ravneet & Sharma, Dharam. (2019). Mobile Based OCR Systems State-of-the-art Survey for Indian Scripts. International Journal of Computer Sciences and Engineering. 7. 457-461. 10.26438/ijcse/v7i4.457461.

[5] T. T. Zin, M. Z. Pwint and S. Thant, "A Mobile Application for Offline Handwritten Character Recognition," 2020 IEEE 9th Global Conference on Consumer Electronics (GCCE), Kobe, Japan, 2020, pp. 10-11, doi: 10.1109/GCCE50665.2020.9291735.

[6] Y. Gao, L. Jin, C. He and G. Zhou, "Handwriting Character Recognition as a Service: A New Handwriting Recognition System Based on Cloud Computing," 2011 International Conference on Document Analysis and Recognition, Beijing, China, 2011, pp. 885-889, doi: 10.1109/ICDAR.2011.181.

[7] N. Shorim, T. Ghanim and A. AbdelRaouf, "Implementing Arabic Handwritten Recognition Approach using Cloud Computing and Google APIs on a mobile application," 2019 14th International Conference on Computer Engineering and Systems (ICCES), Cairo, Egypt, 2019, pp. 88-95, doi: 10.1109/ICCES48960.2019.9068124.

[8] V. V. Mainkar, J. A. Katkar, A. B. Upade and P. R. Pednekar, "Handwritten Character Recognition to Obtain Editable Text," 2020 International Conference on Electronics and Sustainable Communication Systems (ICESC), Coimbatore, India, 2020, pp. 599-602, doi: 10.1109/ICESC48915.2020.9155786.

[9] S. Dome and A. P. Sathe, "Optical Charater Recognition using Tesseract and Classification," 2021 International Conference on Emerging Smart Computing and Informatics (ESCI), Pune, India, 2021, pp. 153-158, doi: 10.1109/ESCI50559.2021.9397008.

[10] B. H. Rao et al., "MTESSERACT: An Application for Form Recognition in Courier Services," 2022 3rd International Conference on Smart Electronics and Communication (ICOSEC), Trichy, India, 2022, pp. 848-853, doi: 10.1109/ICOSEC54921.2022.9952031.

[11] Park, S. S., & Chung, K. S. (2022). Regularized Convolutional Neural Network for Highly Effective Parallel Processing. Journal of Computing Science and Engineering, 16(2), 105-112.

[12] R. Smith, "An Overview of the Tesseract OCR Engine," Ninth International Conference on Document Analysis and Recognition (ICDAR 2007), Curitiba, Brazil, 2007, pp. 629-633, doi: 10.1109/ICDAR.2007.4376991.