ABSTRACT:

An OCR (Optical Character Recognition) system which is a branch of computer vision and in turn a sub-class of Artificial Intelligence. Optical character recognition is the translation of optically scanned bitmaps of printed or hand written text into audio output by using of Raspberry Pi. OCRs developed for many world languages are already under efficient use. Thismethod extracts moving object region by a mixture-of-Gaussians-based background subtraction method. A text localization and recognition are conducted to acquire text information. To automatically localize the text regions from the object, a text localization and Tesseract algorithm by learning gradient features of stroke orientations and distributions of edge pixels in an Ada boost model. Text characters in the localized text regions are then binaries and recognized by off-the-shelf optical character recognition software. The recognized text codes are output to blind users in speech. Performance of the proposed text localization algorithm. As the recognition process is completed, the character codes in the text file are processed using Raspberry pi device on which recognize character using Tesseract algorithm and python programming, the audio output is listened.

TABLE OF CONTENTS

CHAPTER NO.	TITLE	PAGE NO.
	ABSTRACT	**V**
	LIST OF FIGURES	**VIII**
	LIST OF ABBREVIATIONS	**IX**
1	**INTRODUCTION**	**1**
	• OUTLINE OF THE PROJECT	**1**
	• APPLICATION SECURITY	**2**

	• Attack Model	**2**
	• Forgery Attack	**2**
	• Man-in-the-Middle Attack	**2**
	• Identity Revealing Attack	**3**
	• SECURITY FEATURES	**3**
	• Data Integrity	**3**
	• Privacy Preservation	**3**
	• Data Confidentiality	**3**
	• Availability	**3**
	• Single Point of Failure	**4**
	• PROBLEM STATEMENT	**4**
	• PROBLEM DESCRIPTION	**5**
2	**LITERATURE SURVEY**	**6**
	• STATE OF ART	**6**
	• Blockchain in IOT	**6**
	• Blockchain in Healthcare	**6**
	• Blockchain in WSN	**7**
	• INFERENCE FROM LITRATURE	**8**
3	**SYSTEM ANALYSIS**	**9**
	• EXISTING SYSTEM	**9**
	• PROPOSED SYSTEM	**9**
	• Architecture Overview	**10**
	• Workflow of SCIT	**12**
	• KEY REQUIREMENTS	**12**
	• Android Platform	**12**
	• Challenges of App Development	**13**
	• Building Multi Screens	**13**
	• Maximization App Performance	**13**
	• Development Process	**13**
	• Android Emulator	**14**
	• SYSTEM REQUIREMENTS	**14**
4	**METHODOLOGIES**	**15**
	• MOBILE APPLICATION DEPLOYMEN	**15**
	• SIMULATION ENVIRONMENT	**16**

	• Remix IDE	**16**
	• Ganache	**16**
	• Metamask	**16**
	• System Specification	**16**
	• ENCRYPTED SYSTEM	**17**
	• Symmetric Encryption	**17**
	• Asymmetric Encryption	**17**
5	**RESULT AND CONCLUSION**	**19**
	5.1 RESULT	**19**
6	**CONCLUSION AND FUTURE WORK**	**20**
	6.1 CONCLUSION	**20**
	6.2 FUTURE ENHANCEMENT	**20**
	REFERENCES	**21**
	APPENDICES	**22**
	• **SAMPLE CODE**	**22**
	• **PUBLICATION WITH PLAGARISM REPORT**	**-**
	• **PUBLICATION ACCEPTANCE LETTER AND COPYRIGHT FORM**	**-**

LIST OF FIGURES

FIGURE NO.	NAME OF THE FIGURE	PAGE NO.
1.1	OVERVIEW APPLICATION FLOW OF SCIT	**1**
1.2	MAN-IN-THE-MIDDLE ATTACK	**2**
3.1	ARCHITECTURE OF SCIT	**10**
3.2	ARCHITECTURE OF BLOCKCHAIN	**11**
3.3	DEVELOPMENT OF WORKFLOW OF SCIT	**12**
3.4	ANDROID APP DEVELOPMENT FLOW	**13**
4.1	REPRESENTATION OF THE ANDROID APPLICATION EXECUTION	**15**
4.2	SYMMETRIC ENCRYPTION EPRESENTATION	**17**

4.3	ASYMMETRIC ENCRYPTION REPRESENTATION	**18**
5.1	SCIT DASHBOARD	**19**
5.2	DATA INPUT SCREEN	**19**
5.3	MINIG PROCESS	**19**
5.4	SCIT BLOCKCHAIN NETWORK	**19**

LIST OF ABBREVIATIONS

SCIT	Secure Confidential Information Transfer
DApps	Decentralized Application
IoT	Internet of Things

WSN	Wireless Sensor Network
PDP	Provable Data Possession
D2D	Device-to-Device
CSI	Channel State Information
EMR	Electronic Medical Records
PHI	Personal Health Information
DDoS	Distributed Denial of Service
PoW	Proof of Work
UI	User Interface
APK	Android Application Package
P2P	Peer to Peer

XRP	A Crypto Currency
NFT	Non-Fungible Token
CIA	Central Intelligence Agency

CHAPTER 1

- **INTRODUCTION**

In the running world, there is growing demand for the software systems to recognize characters in computer system when information is scanned through paper documents as we know that we have number of newspapers and books which are in printed format related to different subjects. These days there is a huge demand in “storing the information available in these paper documents in to a computer storage disk and then later reusing this information by searching process”. One simple way to store information in these paper documents in to computer systemis to first scan the documents and then store them as IMAGES. But to reuse this information it is very difficult to read the individual contents and searching the contents form these documents line-by-line and word-by-word. The reason for this difficulty is the font characteristics of the characters in paper documents are different to font of the characters in computer system. As a result, computer is unable to recognize the characters while reading them. This concept of

storing the contents of paper documents in computer storage place and then reading and searching the content is called DOCUMENT PROCESSING. Sometimes in this document processing we need to process the information that is related to languages other than the English in the world. For this document processing we need a software system called **CHARACTER RECOGNITION SYSTEM**. This process is also called DOCUMENT IMAGE ANALYSIS (DIA).

Thus our need is to develop character recognition software system to perform Document Image Analysis which transforms documents in paper format to electronic format. For this process there are various techniques in the world. Among all those techniques we have chosen Optical Character Recognition as main fundamental technique to recognize characters. The conversion of paper documents in to electronic format is an on-going task in many of the organizations particularly in Research and Development (R&D) area, in large business enterprises, in government institutions, so on. From our problem statement we can introduce the necessity of Optical Character Recognition in mobile electronic devices such as cell phones, digital cameras to acquire images and recognize them as a part of face recognition and validation.

To effectively use Optical Character Recognition for character recognition in-order to perform Document Image Analysis (DIA), we are using the information in Grid format. . This system is thus effective and useful in ***Virtual Digital Library's*** design and construction.

- **PURPOSE**

The main purpose of **Optical Character Recognition (OCR)** system based on a gridinfrastructure is to perform Document Image Analysis, document processing of electronic

document formats converted from paper formats more effectively and efficiently. This improves the accuracy of recognizing the characters during document processing compared to various existing available character recognition methods. Here OCR technique derives the meaning of the characters, their font properties from their bit-mapped images.

- The primary objective is to speed up the process of character recognition in document processing. As a result the system can process huge number of documents with-in less time and hence saves the time.
- Since our character recognition is based on a grid infrastructure, it aims to recognize multiple heterogeneous characters that belong to different universal languages with different font properties and alignments.

- **PROJECT SCOPE**

- The scope of our product Optical Character Recognition on a grid infrastructure is to provide an efficient and enhanced software tool for the users to perform Document Image Analysis, document processing by reading and recognizing the characters in research, academic, governmental and business organizations that are having large pool of documented, scanned images. Irrespective of the size of documents and the type of characters in documents, the product is recognizing them, searching them and processing them faster according to the needsof the environment.

1.3 EXISTING SYSTEM

In the running world there is a growing demand for the users to convert the printed documents in to electronic documents for maintaining the security of their data. Hence the basic OCR system was invented to convert the data available on papers in to computer process able documents, So that the documents can be editable and reusable. The existing system/the previous system of OCR on a grid infrastructure is just OCR without grid functionality. That is the existing system deals with the homogeneous character recognition or character recognition of single languages.

- **1.4 DRAWBACK OF EXISTING SYSTEM**

The drawback in the early OCR systems is that they only have the capability to convert and recognize only the documents of English or a specific language only. That is, the older OCR system is uni-lingual.

• 1.5 PROPOSED SYSTEM

Our proposed system is OCR on a grid infrastructure which is a character recognition system that supports recognition of the characters of multiple languages. This feature is what we call grid infrastructure which eliminates the problem of heterogeneous character recognition and supports multiple functionalities to be performed on the document. The multiple functionalities include editing and searching too where as the existing system supports only editing of the document. In this context, Grid infrastructure means the infrastructure that supports group of specific set of languages. Thus OCR on a grid infrastructure is multi-lingual.

• 1.6 BENEFIT OF PROPOSED SYSTEM

The benefit of proposed system that overcomes the drawback of the existing system is that it supports multiple functionalities such as editing and searching. It also adds benefit by providing heterogeneous characters recognition.

• 1.7 ARCHITECTURE OF THE PROPOSED SYSTEM

The Architecture of the optical character recognition system on a grid infrastructure consistsof the three main components. They are:-

- Scanner
- OCR Hardware or Software
- Output Interface

Figure.1: OCR Architecture

• 1.8 INTENDED AUDIENCE AND READING SUGGESTIONS

In this section, we identify the audience who are interested with the product and are involved in the implementation of the product either directly or indirectly. As from our research, the OCR system is mainly useful in R&D at various scientific organizations, in governmental institutes and in large

business organizations, we identify the following as various interested audience in implementing OCR system:-

- The scientists, the research scholars and the research fellows in telecommunication institutions are interested in using OCR system for processing the word document that contains base paper for their research.
- The Librarian to manage the information contents of the older books in building virtual digital library requires use of OCR system.
- Various sites that vendor e-books have a huge requirement of this OCR system in-orderto scan all the books in to electronic format and thus make money. The Amazon book world is largely using this concept to build their digital libraries.

Now we present the reading suggestions for the users or clients through which the user can betterunderstand the various phases of the product. These suggestions may be effective and useful for the beginners of the product rather than the regular users such as research scholars, librarians and administrators of various web-sites. With these suggestions, the user need not waste his time in scrolling the documents up and down, browsing through the web, visiting libraries in search of different books and … The following are the various reading suggestions that the user can follow in-order to completely understand about our product and to save time:-

- It would help you if you start with Wikipedia.com. It lets you know the basic concept of every keyword you require. First learn from it what is OCR? And how does it work basedon a Grid infrastructure?
- Now you can proceed your further reading with the introduction of our product we provided in our documentation. From these two steps you completely get an in-depth idea of the use of our product and several processes involved in it.
- The more you need is the implementation of the product. For this you can visit FreeOCR.com where you can view how the sample OCR works and you can try it.

1.9 PROBLEM STATEMENT:

- Traditional methods like Braille exist using which the blind people have to trace and readtext, which is very slow and not very practical.
- Existing OCR systems are not automatic and require full-fledged computers to run andhence are not effective.
- Reader Mobile runs on a cell phone and allows the user to read mail, receipts, fliers, andmany other documents

• 2. LITERATURE REVIEW

A feasibility study is a high-level capsule version of the entire System analysis and Design Process. The study begins by classifying the problem definition. Feasibility is to determine if it's worth doing. Once an acceptance problem definition has been generated, the analyst develops a logical model of the system. A search for alternatives is analyzed carefully. There are 3 parts in feasibility study.

• TECHNICAL FEASIBILITY

Evaluating the technical feasibility is the trickiest part of a feasibility study. This is because, at this point in time, not too many detailed design of the system, making it difficult to access issues like performance, costs on (on account of the kind of technology to be deployed) etc. A number of issues have to be considered while doing a technical analysis. Understand the different technologies involved in the proposed system before commencing the project we have to be veryclear about what are the technologies that are to be required for the development of the new system. Find out whether the organization currently possesses the required technologies. Is the required technology available with the organization?.

• OPERATIONAL FEASIBILITY

Proposed project is beneficial only if it can be turned into information systems that will meet the organizations operating requirements. Simply stated, this test of feasibility asks if the system will work when it is

developed and installed. Are there major barriers to Implementation? Here are questions that will help test the operational feasibility of a project:

- Is there sufficient support for the project from management from users? If the current system is well liked and used to the extent that persons will not be able to see reasons for change, there may be resistance.
- Are the current business methods acceptable to the user? If they are not, Users may welcome a change that will bring about a more operational and useful systems.
- Have the user been involved in the planning and development of the project?
- Early involvement reduces the chances of resistance to the system and in general and increases the likelihood of successful project.

Since the proposed system was to help reduce the hardships encountered. In the existing manual system, the new system was considered to be operational feasible.

- **ECONOMIC FEASIBILITY**

Economic feasibility attempts to weigh the costs of developing and implementing a new system, against the benefits that would accrue from having the new system in place. This feasibility study gives the top management the economic justification for the new system. A simple economic analysis which gives the actual comparison of costs and benefits are much more meaningful in this case. In addition, this proves to be a useful point of reference to compare actual costs as the project progresses. There could be various types of intangible benefits on account of automation. These could include increased customer satisfaction, improvement in product quality better decision making timeliness of information, expediting activities, improved accuracy of operations, better documentation and record keeping, faster retrieval of information, better employee morale.

- **TRAINING**