ndrecognized by off-the-shelf optical character recognition software. The recognized text codes areoutput to blind users in speech. Performance of the proposed text localization algorithm. As therecognitionprocessiscompleted,thecharactercodesinthetextfileareprocessedusingRaspberrypideviceonwhichrecognizecharacterusingTesseractalgorithmandpythonprogramming,theaudiooutputislistened.

TABLE OFCONTENTS

LISTOFFIGURES

LISTOFABBREVIATIONS

SCIT	SecureConfidentialInformationTransfer
DApps	DecentralizedApplication
IoT	Internetof Things

WSN	WirelessSensorNetwork
PDP	ProvableDataPossession
D2D	Device-to-Device
CSI	Channel StateInformation
EMR	ElectronicMedical Records
PHI	PersonalHealthInformation
DDoS	DistributedDenialofService
PoW	ProofofWork
UI	User Interface
APK	AndroidApplicationPackage
P2P	PeertoPeer
XRP	ACrypto Currency
NFT	Non-FungibleToken
CIA	Central IntelligenceAgency

CHAPTER1

- ## INTRODUCTION

Intherunningworld,thereisgrowingdemandforthesoftwaresystemstorecognize charactersin computer system when information is scanned through paper documents as we know that wehave number of newspapers and books which are in printed format related to different subjects.Thesedaysthereisahugedemandin“storingtheinformationavailableint hesepaperdocuments in to a computer storage disk and then later reusing this information by searchingprocess”.Onesimplewaytostoreinformationinthesepaperdocumentsi ntocomputersystemis to first scan the documents and then store them as IMAGES. But to reuse this information it isvery difficult to read the individual contents and searching the contents form these documentsline-by-line and word-by-word. The reason for this difficulty is the font characteristics of thecharacters in paper documents are different to font of the characters in computer system. As aresult, computer is unable to recognize the characters while reading them. This concept of storingthe contents of paper documents in computer storage place and then reading and searching thecontent is called DOCUMENT PROCESSING. Sometimes in this document processing we needto process the information thatis related tolanguages other than the English in the world. Forthis document processing we need a software system called **CHARACTER RECOGNITIONSYSTEM**.Thisprocessisalsocalled DOCUMENTIMAGEANALYSIS(DIA).

Thus our need is to develop character recognition software system to perform Document ImageAnalysis which transforms documents in paper format to electronic format. For this process therearevarioustechniquesintheworld.Amongallthosetechniqueswehavechose nOpticalCharacter Recognition as main fundamental technique to recognize characters. The conversion ofpaperdocumentsin toelectronicformatisan on-going taskinmany of the organizationsparticularlyinResearchandDevelopment(R&D)area,inlargebusi

nessenterprises,ingovernment institutions, so on. From our problem statement we can introduce the necessity ofOpticalCharacterRecognitioninmobileelectronicdevicessuchascellphones,digitalcamerasto acquireimagesandrecognize themasa partofface recognitionandvalidation.

To effectively use Optical Character Recognition for character recognition in-order to performDocument Image Analysis (DIA), we are using the information in Grid format. . This system isthuseffective andusefulin***VirtualDigitalLibrary's***designandconstruction.

- **PURPOSE**

Themainpurposeof**OpticalCharacterRecognition(OCR)**systembasedon agridinfrastructureistoperformDocumentImageAnalysis,documentprocessing ofelectronic

document formats converted from paper formats more effectively and efficiently. This improvesthe accuracy of recognizing the characters during document processing compared to variousexisting available characterrecognition methods.Here OCR technique derives themeaning ofthe characters,theirfontpropertiesfromtheirbit-mappedimages.

- The primary objectiveis tospeedup the process of characterrecognitionindocumentprocessing. As a result the system can process huge number of documents with-in less timeandhencesaves thetime.
- Since our character recognition is based on a grid infrastructure, it aims to recognize multipleheterogeneous characters that belong to different universal languages with different fontpropertiesandalignments.

- **PROJECTSCOPE**

- The scope of our product Optical Character Recognition on a grid infrastructure is toprovide an efficientand enhanced software tool forthe users toperform DocumentImageAnalysis, document processing by reading and recognizing the characters in research, academic,governmental and business organizations that are having large pool of documented,

scannedimages.Irrespectiveof thesizeof documents andthetypeof charactersin documents, theproductisrecognizingthem,searchingthemandprocessingthemfasteraccordingtotheneedsoftheenvironment.

1.3EXISTINGSYSTEM

In the running world there is a growing demand for the users to convert the printed documents into electronic documents for maintaining the security of their data. Hence the basic OCR systemwas invented to convert the data available on papers in to computer process able documents, Sothat the documents can be editable and reusable.The existing system/the previous system ofOCR on a grid infrastructure is just OCR without grid functionality. That is the existing systemdealswiththehomogeneouscharacterrecognitionorcharacterrecognition ofsinglelanguages.

- ## 1.4DRAWBACKOFEXISTING SYSTEM

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

- ## 1.5PROPOSEDSYSTEM

Our proposed system is OCR on a grid infrastructure which is a character recognition system thatsupports recognition of the characters of multiple languages. This feature is what we call gridinfrastructure which eliminates the problem of heterogeneous character recognition and supportsmultiple functionalities to be performed on the document. The multiple functionalities includeediting and searching too where as the existing system supports only editing of the document. Inthis context, Grid infrastructure means the infrastructure that supports group of specific set oflanguages.ThusOCR ona gridinfrastructure ismulti-lingual.

- **1.6BENEFITOFPROPOSEDSYSTEM**

The benefit of proposed system that overcomes the drawback of the existing system is that itsupports multiple functionalities such as editing and searching. It also adds benefit by providingheterogeneouscharacters recognition.

- **1.7ARCHITECTUREOFTHEPROPOSEDSYSTEM**

TheArchitectureoftheopticalcharacterrecognitionsystemonagridinfrastructureconsistsofthethreemaincomponents.Theyare:-

- Scanner
- OCRHardwareorSoftware
- OutputInterface

Figure.1:OCRArchitecture

- **1.8INTENDEDAUDIENCEANDREADINGSUGGESTIONS**

In this section, we identify the audience who are interested with the product and are involved inthe implementation of the product either directly or indirectly. As from our research, the OCRsystem ismainly useful in R&D atvarious scientific organizations,in governmental institutesand in large business organizations, we identify the following as various interested audience inimplementingOCR system:-

- Thescientists,theresearchscholarsandtheresearchfellowsintelecommunicationinstitutions are interested in using OCR system for processing the word document thatcontainsbasepaperfortheirresearch.
- The Librarian to manage the information contents of the older books in building virtualdigitallibraryrequires useofOCR system.
- Various sites thatvendore-books have a huge requirementof this OCR system in-orderto scan all the books in to electronic format and thus make money. The Amazon bookworldislargelyusingthisconcepttobuildtheirdigitallibraries.

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 forthe beginners of the product rather than the regular users such as research scholars, librarians andadministrators of various web-sites. With these suggestions, the user need not waste his time inscrolling the documents up and down, browsing through the web, visiting libraries in search ofdifferent books and … The following are the various reading suggestions that the user can followin-ordertocompletelyunderstandaboutourproductandtosave time:-

- It would help you if you start with Wikipedia.com. It lets you know the basic concept ofevery keyword you require. First learn from it what is OCR? And how does it work basedonaGridinfrastructure?
- Now you can proceedyour further reading with theintroduction of our productweprovided in our documentation. From these two steps you completely get an in-depth ideaoftheuse ofourproductandseveralprocessesinvolvedinit.
- Themoreyouneedistheimplementationoftheproduct.Forthisyoucanvisi tFreeOCR.comwhereyoucan view how thesample OCRworksand youcantryit.

1.9PROBLEMSTATEMENT:

- TraditionalmethodslikeBrailleexistusingwhichtheblindpeoplehavetot raceandreadtext,whichis veryslowandnotverypractical.
- ExistingOCRsystemsarenotautomaticandrequirefull-fledgedcomputerstorunandhence arenoteffective.
- ReaderMobilerunsonacellphoneandallowstheusertoreadmail,receipts, fliers,andmanyotherdocuments

• 2.LITERATUREREVIEW

A feasibility study is a high-level capsule version of the entire System analysis and DesignProcess. The study begins by classifying the problem

definition. Feasibility is to determine if it'sworth doing. Once an acceptance problem definition has been generated, the analyst develops alogical model of the system. A search for alternatives is analyzed carefully. There are 3 parts infeasibilitystudy.

- **TECHNICALFEASIBILITY**

Evaluating the technical feasibility is the trickiest part of a feasibility study. This is because, atthis point in time, not too many detailed design of the system, making it difficult to access issueslike performance, costs on (on account of the kind of technology to be deployed) etc. A numberofissueshavetobeconsideredwhiledoinga technical analysis.Understandthedifferenttechnologies 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 newsystem. Find out whether the organization currently possesses the required technologies. Is therequiredtechnologyavailablewiththeorganization?.

- **OPERATIONALFEASIBILITY**

Proposed project is beneficial only if it can be turned into information systems that will meetthe organizations operating requirements. Simply stated, this test of feasibility asks if thesystemwillworkwhenitisdevelopedandinstalled.Aretheremajorbarriersto Implementation?Herearequestionsthatwillhelp test theoperational feasibilityofaproject:

- Is there sufficientsupport for the projectfrom managementfrom users? If the currentsystem is well liked and used to the extent that persons will not be able to see reasons forchange,theremaybe resistance.
- Arethecurrentbusinessmethodsacceptabletotheuser?Iftheyarenot,Usersm aywelcomea change thatwillbringabouta moreoperationalandusefulsystems.
- Havetheuser beeninvolvedintheplanninganddevelopmentoftheproject?

- Earlyinvolvementreduces the chances of resistance to the system andin general andincreasesthelikelihoodofsuccessfulproject.

Since the proposed system was to help reduce the hardships encountered. In the existing manualsystem,the newsystemwasconsideredtobe operationalfeasible.

- **ECONOMICFEASIBILITY**

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 studygives the top management the economic justification for the new system. A simple economicanalysis which gives the actual comparison of costs and benefits are much more meaningful inthis case. In addition, this proves to be a useful point of reference to compare actual costs as theproject progresses. There could be various types of intangible benefits on account of automation.These couldincludeincreased customer satisfaction,improvementin productquality betterdecisionmakingtimelinessofinformation,expeditingactivities,improved accuracyofoperations,betterdocumentationandrecordkeeping,fasterretrievalo finformation,betteremployeemorale.

- **TRAINING**

Training is a very important process of working with a neural network. As seen from neuralnetworks, there are two forms of training that can be employed with a neural network. They arenamely:-

- Un-SupervisedTraining
- SupervisedTraining

Supervised training provides the neural network with training sets and the anticipated output.Unsupervised training supplies the neural network with training sets, but there is no anticipatedoutputprovided.