iJRASET
International Journal For Research in Applied Science and Engineering Technology

# INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

**Volume:** 9 **Issue:** V **Month of publication:** May 2021

**DOI:** https://doi.org/10.22214/ijraset.2021.34142

www.ijraset.com

Call: 08813907089 | E-mail ID: ijraset@gmail.com

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
*ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.429*
*Volume 9 Issue V May 2021- Available at www.ijraset.com*

# A Review of Optical Character Recognition (OCR) in Healthcare

Aishwarya Bhure[1], Dhanashree Patil[2], Angel Negi[3], Akshita Maskara[4], Madhuri Bhalekar[5]
[1, 2, 3, 4, 5]*School of Computer Engineering & Technology, MIT World Peace University, Pune, Maharashtra, India*

***Abstract: Abstract: Information is present everywhere in newspapers, magazines, documents etc. but healthcare information majorly consisting of medicine labels, drug information, and personal health records etc. is something which is important and confidential at the same time. Today's world is the world of digitization. Technology is advancing day by day and medical healthcare is no exception. Everyday millions of electronic health records (EHR's) are generated, hundreds of invoices are generated, and prescriptions are written. But how to categorize this data and make the best out of it? Optical character recognition (OCR) proves to be of great help in this field. OCR can be used to categorize the EHR's under certain labels. Even most of the time a physician's written prescription is unrecognizable. With the help of OCR, this text can be identified. Knowing one's medicine is highly important , it is easy for normal people but considering visually impaired people, OCR and TTS technologies can be used to get these and other clinical information available to them in the form of an audio. Optical Character Recognition (OCR) is a technique where images or scanned records are perused and converted into OCR recognizable characters, which are then extended for editing and searching purposes. Pattern Recognition techniques and advanced Computer Vision are the main building blocks of working behind an OCR. The paper discusses the role of OCR in various healthcare applications.***

***Keywords: Optical character recognition (OCR), text-to-speech, text extraction, electronic health records (EHR), text recognition.***

## I. INTRODUCTION

Text is present everywhere in magazines, newspapers, bank documents, medical reports, medicine covers, etc. As much as it seems easy for normal people it is extremely difficult for visually impaired people to read this text without the help of a third person.

As a person loses vision, along comes the loss of independence. They struggle in doing day-to-day activities including text identifying. According to a WHO (World Health Organization) report, out of 285 million visually impaired people across the globe, 246 million suffer from low vision while the remaining 39 million are completely blind. Today with the rising advanced digital technologies like Artificial Intelligence, Computer vision, etc. it has become an easy task to build applications that can extract this text present and assist visually impaired people and help them live their lives independently. The systems built using these technologies extract the text present examples- medicine labels, reports, etc. and convert the extracted text into an invoice message which could be easily heard by visually impaired people.

Optical Character Recognition (OCR) is majorly used for text recognition. It was created to help blind or visually impaired people back in 1914. OCR allowed people with visual impairments to have written text read out to them by the system. Over the years there have been tremendous advancements in OCR technology and also overcome the issues like speed, accuracy, font size, etc. Many OCR engines available these days are Tesseract, Google OCR, Transym , OmniPage, etc. Few of these are free whereas many of these are paid.

Tesseract is the most widely used OCR these days and is freely available on the internet. It works on a step-by-step process involving four steps. Instead of focusing on accuracy, it focuses on covering a wide range of languages and fonts. But when compared with another freely available OCR Transym, Tesseract offers higher accuracy but is not fast every time.

## II. LITERATURE SURVEY

In [1], Snigdha Kesh proposes a system which can detect text from any documents including medical reports and can also recognize the medicine that can assist visually impaired without any help from third person and encourages them to live long. The system helps these people in terms of mobility and can also verify the medicine details. The whole process is implemented using a smartphone and an android application.

On shaking the phone, the application starts automatically. Implementation involves two major steps pre handling and post-handling. Pre-handling includes strategies like De-skew, Binarization, line evacuation, Script acknowledgment, Character segregation, etc., to enhance the odds of effective acknowledgment. Additionally, the Levenshtein Distance calculation has been utilized as a part of OCR post-handling to advance outcomes from an OCR API. TOMCAT server is used to store the information in the MySQL database. On shaking the phone, the app starts. When the mobile phone is placed above the medicine, scanning starts and the text detector detects the medicine name and sends it to the TOMCAT server. The server checks whether the medicines are matching or not and checks what the user should take at that time. If the medicine needs to be taken at that time, the server sends the notification to the user, once the mobile application receives the information, then converts the text information to speech and plays the audio.

The paper [2] discusses how a machine can help people with their disobedience, whether it be intentional or simply forgetfulness. Using OCR technology and decision tree with psychology, the module starts by taking an image of the medicine strip and applying multiple pre-processing image enhancements on the jpeg, so it is easier for the module to detect the right medicine. Next, OCR performs its task to recognize the characters in the processed image. Upon recognizing the name of the medicine, a database search is done to extract the record of taking medicine which is saved in the server. Now, the app sets reminders according to the record extracted. After each reminder, it asks a series of questions to confirm whether the patient has taken the medicine or not. It goes as far as asking them to check whether the ring is taken off.

The questions designed by the authors are psychology supportive and help the system to give better results. Having tried this application on a group of people, authors found out that 80% of the users were satisfied with the results and 90% of them were willing to use the app in their daily lives.

The paper[3] aims to explore the task of classifying handwritten text & invoices and updating the fields of the same in the database .The invoices mainly processed manually ,with the system automation done can reduce manual labor .Their proposed system architecture contains modules: preprocessing, text detection & text recognition .The input image is preprocessed by applying threshold the unwanted data is removed & segmentation is performed where lines are extracted .The text detection module used in the paper is EAST(Efficient & Accurate Scene Text detector algorithm).It is implemented using CNN .The text recognition module then uses RNN for recognizing the text from the output feature map of the EAST algorithm .The text extracted with the help of OpenCV & Python packages is compared with the trained dataset using Tensorflow & recognized then updated to the database .The Nanonet OCR is used which brings accuracy and plays an important role in building a Deep Learning model for finding the right data .An android application has been developed using Android studio to do the above process .The paper suggests for finding similar solution for unformatted bills as future work.

The purpose of [4] paper is to feed in medical test reports and discern diseases using OCR technology with natural language processing and machine learning algorithms. Optical Character Recognition is used to convert images into textual data, wherein the model takes scanned medical check-up test reports as input to produce text output. This textual data is pre-processed to formulate it for feature selection. Once these features are extracted, bags have to be created for each disease type using the bag-of-words technique. To enhance the performance of classification on these medical test reports, an ensemble method is used, more specifically Adaptable Boosting. This machine learning technique will provide the final disease that the model predicts, the name for the doctor, whereas an image for the patient for better understanding.

Paper[5] aims to help eradicate the errors committed to reading prescriptions by using a mobile application called Medipic which is developed utilizing a Tesseract OCR. The OCR helps in converting the handwriting of the doctor to digital text. The extracted text using Regular Expressions is matched with the database entry, Repres is used to match the characters with the possible results in the database of drug names. The paper suggests that a more accurate engine must be used to recognize cursive handwriting to improve accuracy. The paper [6] proposes a system to group the clinical and nonclinical documents into suitable categories which are again subclassified. .Electronic Health Records have also known as (EHR's) contain a large number of scanned documents such as radiology reports, clinical correspondence, identification cards, etc. Documents are combined into six pair types which are mentioned in Table 1.Specifically, authors have rated optical character recognition (OCR) and text classification models trained on documents. These documents were earlier physically classified to decide if this approach could accurately recognize scanned documents within an EHR. In image pre-processing, the first image grayscale is done to escalate its OCR execution with tesseract. The contract is increased by 20% and then one iteration of erosion transformation is done. Three supervised classification models are used by the system which are multinomial naive Bayes, Logistic regression, and Random forest. Data is trained using these models. And finally, with the help of machine learning and OCR, the documents are correctly grouped into clinically meaningful categories.

| | | |
|---|---|---|
| 1. | Disclosure Authoriza | Disclosure Authorizations |
| 2. | External Medical Rec | External Medical Records |
| 3. | H&P for Surgeries | History Physical |
| 4. | Questionaire /Consent | Questionnaire/Consents |
| 5. | Referrals/ Authoriza | Referrals/Authorzations |
| 6. | Drivers License | Drivers Liscense |

Table 1: Six pairs of document types

The paper[7] includes the development of a mobile application namely 'Pharma Guide' which is intended to retrieve the name, use, price, and side effects of the medicines. The app also allows the deliverance of medicine at the doorstep. Android Studio platform and MySQL database have been used for the development purpose. The user can register and login into the account where he can view information on all medicines. The users can scan a particular medicine using MedLens which uses Google Vision API for OCR. It allows developers to create vision-based Machine Learning applications based on OCR without having any knowledge in ML. The Google Cloud Vision API takes complex ML models having image recognition and formats them into simple REST API. It uses model which is trained on a large image dataset so there is no need to develop and train your custom model. The text from the image of the medicine(name) is extracted and given to the user along with its use and side effects if found in the trained dataset. The paper suggests training the system for other font styles for name extraction as well as the addition of more medicines in the database. For future work, it also mentions an additional security module for certain medicines in terms of quantity that can be ordered to prevent the risk of high dosage.

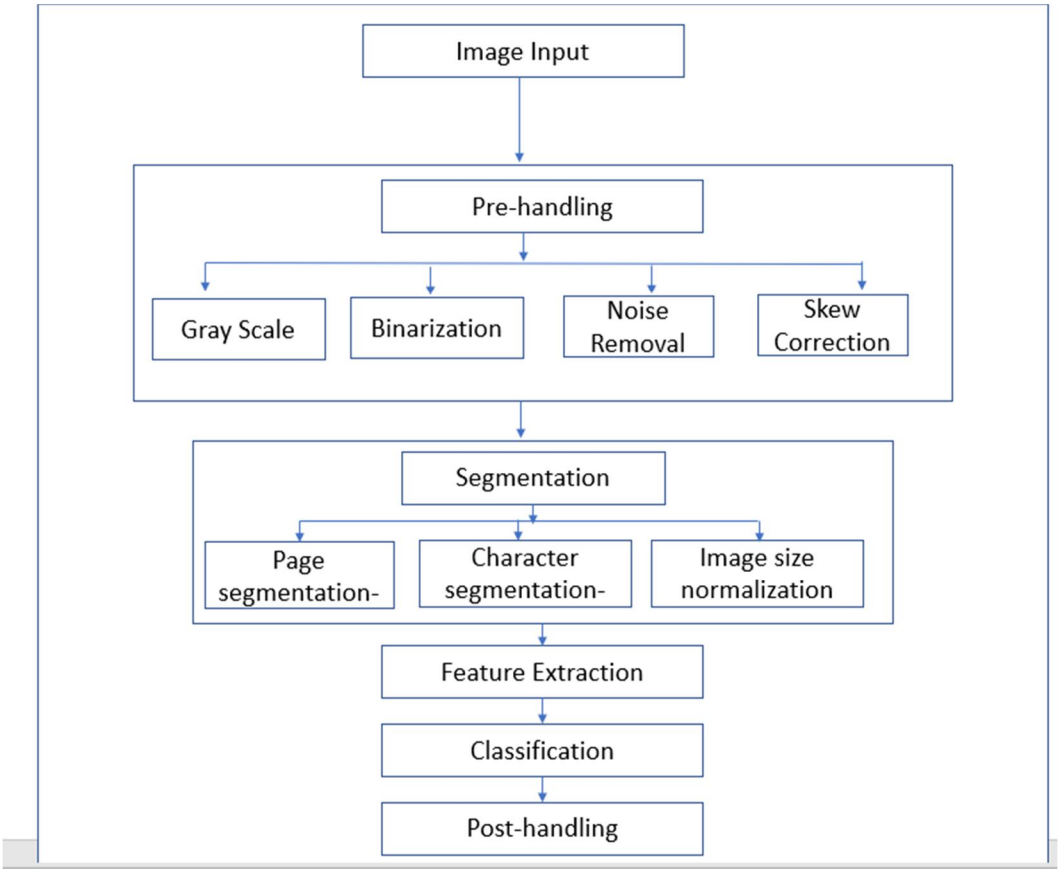



Figure 1: Steps in OCR towards text Recognition

Paper [8], discusses the issue of how illegible handwriting of the physician can cause problems to people. It says, for the patients, unreadable handwriting often results in incorrect dosage or slows down the treatment or test. To overcome this problem, the authors proposed a system that consists of a tool for noting down input. The input device is nothing but a pen which a physician uses to put down the medical prescription and at the same time, the handwritten text is converted into a printed form, which is presented online. The form is lined up. Also, it is styled by making use of some formatting tools. Apart from this, it also comprehends a dictionary, the credentials of the physician, and will put a proper prescription before the patient. Two major processes contribute the most to the system which is styling and recognizing the handwriting of the physician using a Formatting process. An Optical Character Recognition System is used to recognize printed or handwritten text. It is in digital pictures of physical documents which can be scanned copies. The characters written by the physician are scanned one by one relying on the input pen-like when it is lifted and the node lets the user scan without performing OCR. Sentence detection, abbreviation remover, Parts of speech, Tokenization, etc. are done in an online formatting process. Accuracy is evaluated using a confusion matrix. The authors tried to excel the social health of the territory by lowering the issues caused due to unreadable physician's handwriting by utilizing a pen to translate handwritten prescriptions into printed output. Prescription is presented to the patient using a styling method.

Figure 1 lists out the steps that are used in text identification in OCR. The mentioned steps are the common ones which most of the authors have mentioned in their papers. The first step includes pre-handling which consists of converting the image to grayscale, binarization, noise removal, and skew correction. The next major step is segmentation which is used to separate the required text from the rest. It includes steps like page segmentation, character segmentation, image size normalization, etc. Then comes the feature extraction step which performs the task of extraction of a bunch of features, which will scale up the recognition for a trickle of items and will generate identical features used in different occurrences of the same attribute. Classification is the next step which uses the feature vector from the previous step. Post handling is the last step which includes procedures that can enhance the accuracy of OCR frameworks in case the output is not correct.

Table2 gives a glimpse of a comparative data chart consisting of techniques and parameters used in all the above-mentioned eight papers.

| Research Paper[s] | Techniques used | Parameters used |
|---|---|---|
| [1] | OCR, Computer vision, TOMCAT server, MySQL database, Android studio | Medicine image |
| [2] | OCR, Decision Trees, Psychology | Images taken of medicine strips |
| [3] | Nanonet OCR, EAST text detection algorithm, RNN, CNN, Computer Vision, Tensorflow, Android application | Invoices |
| [4] | OCR, Bag-of-Words, AdaBoost, Naive Bayes Classification Algorithm | Medical check-up test report |
| [5] | Tesseract OCR, Android Studio, Repres, Database | Medical Prescriptions |
| [6] | Tesseract OCR, Random Forest, Logistic Regression, Multinomial naive Bayes | Electronic health records |
| [7] | Android Studio, MySQL, Google Vision API for OCR. | Medicines |
| [8] | OCR, Tokenization, Parts of Speech, Online Formatting | Prescription written using a special pen. |

Table 2: List of techniques and parameters used in the above-mentioned research papers.

## III. CONCLUSION AND FUTURE WORK

OCR is one of the important paradigms in the medical field. All the organizations are in the procedure of making records digitized such as records of patients and employees, client information, billings, health records to reduce costs in storage, decrease manual work, increase the security of data stored, and enable efficient analytics processes. And OCR would make this process easier. OCR can be a virtual eye for visually impaired people and can make crucial medical information available to them with the help of TTS technology.

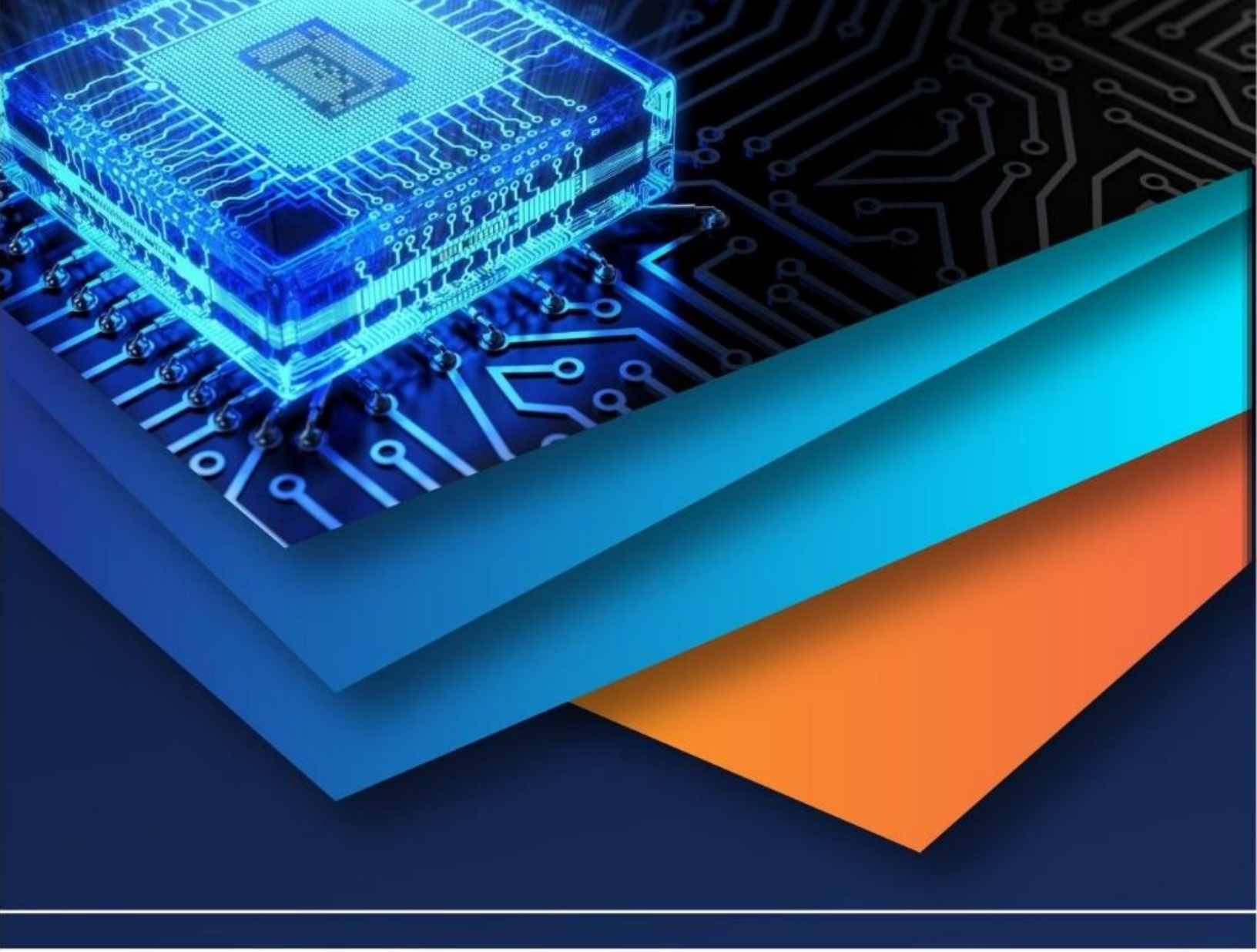
When our documents will be available in a text format, we can search their content easily for specific information without the need to spend time foraging through a large volume of paper documents also it reduces the manual effort and reduces search time. Text detection when subsequently followed by text recognition using any pattern recognizing algorithm not only improves the accuracy but also accelerates the system. The concerned devices in the above-discussed papers achieve short processing time, also the readability tolerance attained is low, which can be further reduced, along with the error rate, if the image is taken with good lighting. As future work, this OCR along with the advanced deep learning algorithms can be used to extract accurate information and convert it to audio output. Considering the development of applications, security modules to restrict a certain amount of dosage for medicines can also be included.

## IV. ACKNOWLEDGEMENT

We take this opportunity to express our sincere appreciation for the cooperation and support given by IJRASET for accepting our work and further giving us insights to improve.
We would like to express our gratitude towards the project coordinators of the School of Computer Engineering & Technology. In particular, we would like to thank Prof. Madhuri Bhalekar for her continual support and guidance in the process of research for this project .For all efforts behind the review paper, we would also like to express our sincere appreciation to the staff of the department of School of Computer Engineering & Technology, MIT-WPU Pune, for their extended help and suggestions.

## REFERENCES

[1] Kesh, S., & Ananthanagu, U. (2018). Text Recognition and Medicine Identification by Visually Impaired People. International journal of engineering research and technology, 5.
[2] PoTsun Lai, WeiChen Chen, CiunLi Chin and YuLiang Kuo (2013), Intelligent taking medicine reminding system, Institute of Electrical and Electronics Engineers
[3] Geetha, M & Pooja, R & Swetha, J & Nivedha, N & Daniya, T.. (2020). Implementation of Text Recognition and Text Extraction on Formatted Bills using Deep Learning. International Journal of Control and Automation. 13. 646-651.
[4] Wisam A. Qader and Musa M. Ameen (2019), Diagnosis of Diseases from Medical Check-up Test Reports Using OCR Technology with BoW and AdaBoost algorithms, Fifth International Engineering Conference on Developments in Civil & Computer Engineering Applications.
[5] R. B. Alday and R. M. Pagayon, "MediPic: A mobile application for medical prescriptions," IISA 2013, Piraeus, Greece, 2013, pp. 1-4, doi: 10.1109/IISA.2013.6623682.
[6] Heath Goodrum, Kirk Roberts, Elmer V. Bernstam,Automatic classification of scanned electronic health record documents,International Journal of Medical Informatics,Volume 144, 2020,104302, ISSN 1386-5056,.
[7] Namrata Chetgiri, Arya Gopinath, Aishwarya Sundaresan,Gayatri Hegde,(2020) ”Medicine Information Retrieval Application- PharmaGuide”,International Research Journal of Engineering and Technology (IRJET)
[8] S. Butala, A. Lad, H. Chheda, M. Bhat and A. Nimkar, "Natural Language Parser for Physician’s Handwritten Prescription," 2020 International Conference on Emerging Trends in Information Technology and Engineering (ic-ETITE), Vellore, India, 2020, pp. 1-7, doi: 10.1109/ic-ETITE47903.2020.325.



10.22214/IJRASET

45.98

IMPACT FACTOR: 7.129

IMPACT FACTOR: 7.429

# INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY