IJNRD.ORG ISSN : 2456-4184
INTERNATIONAL JOURNAL OF NOVEL RESEARCH AND DEVELOPMENT (IJNRD) | IJNRD.ORG
An International Open Access, Peer-reviewed, Refereed Journal

INTELLIGENT MEDICAL PRESCRIPTION ANALYSIS USING EASYOCR AND NER

[1]Kanisshka U P , [2]Jayachandiran U , [3]Ananya G S

[1]Student, [2]Associate Professor, [3]Student.
[1]BTech Artificial Intelligence and Data Science,
[1]Sri Sairam Engineering College, Chennai, India

Abstract : This project presents an innovative approach to automating the extraction and classification of medical terms from prescription images using Optical Character Recognition (OCR) and Named Entity Recognition (NER) techniques. Leveraging EasyOCR for text extraction, the system identifies and processes textual information from prescription documents, ensuring accuracy and efficiency. The extracted text is subsequently analyzed using a fine-tuned BioBERT model designed for token classification, enabling the categorization of essential medical entities such as drugs, dosages, durations, times, and patient details. This integration of advanced machine learning methods facilitates the swift interpretation of medical prescriptions, supporting healthcare professionals in enhancing patient care and minimizing errors. By automating these processes, the project aims to improve accessibility and reliability in medical documentation, paving the way for future developments in health informatics and digital health solutions.

Keywords—EasyOCR, Named Entity Recognition(NER), BioBERT, Machine Learning, Medical Terminology Classification

INTRODUCTION

The integration of technology in healthcare has revolutionized the way medical information is processed and utilized. [1]In particular, the automation of extracting and classifying medical terms from prescription images presents a significant advancement in enhancing patient care. This research paper introduces a novel approach that combines Optical Character Recognition (OCR) and Named Entity Recognition (NER) techniques to streamline the interpretation of medical prescriptions.

[2] Despite the proliferation of digital writing tools,many individuals still prefer traditional note-taking methods using pen and paper. This practice poses several challenges, including difficulties in efficiently storing, accessing, and sharing physical documents. As a result, significant knowledge may remain unreviewed or lost due to the lack of digital conversion.

So [3]To enable humanists, historians, genealogists as well as ordinary people to efficiently work with these documents, it is subject to current research and scientific discussion to make the content of these documents digitally available. [4]Thus, images are the source of information to offline text recognition, which can be applied for transcriptions of historical manuscripts , medical prescriptions [5], forms , [6]Image acquisition, License plate detection and Character recognition and so on. This emphasizes the need for research into the area of building large scale systems for many languages and scripts.

As a result [7]deep learning was integrated on image and text embeddings for recognition of words that explores how a combined embedding model improves recognition accuracy compared to using image or text data alone. Deep learning based methods that recognize text in images of medical reports[8] are developed in order to enhance the accuracy and efficiency of text recognition in medical reports to improve diagnostic decision-making and reduce manual errors.[9]Moreover, deep learning models like Convolutional Neural Networks (CNN) and Recurrent Neural Networks (RNN) have gained prominence for their ability to learn complex patterns and structures in textual data[10].Traditional OCR methods rely on pattern matching and feature extraction to recognize characters, while[11] machine learning algorithms such as[12] Support Vector Machines (SVM) and Random Forests have shown improved performance in text recognition tasks. [13]Optical character recognition (OCR) with natural language processing (NLP)' presents a method that combines OCR and NLP for extracting information from unstructured data and structured medical information extraction. [14] Optical Character Recognition (OCR), utilization of Neural Networks, and various Feature Extraction techniques. Furthermore, the document highlights the hurdles faced in HCR, as well as its applications across diverse domains such as document analysis, mail sorting, and enhancing computer security. Methods such as [15] Semi Incremental Recognition method,Incremental recognition method,Line and Word segmentation,zoning method,slope and slant correction helps in recognizing the strokes that are present in a script.These gradually get ensemble with NLP for further more recognition of the text.

[16]A new language represents an ion model called BERT, which stands for Bidirectional Encoder Representations from Transformers. Unlike recent language representation models), BERT is designed to pre train deep bidirectional representations from unlabeled text by jointly conditioning on both left and right context in all layers. As a result, the pre-trained BERT model can be fine tuned with just one additional output layer to create state-of-the-art models for a wide range of tasks, such as question answering and language inference, without substantial task specific architecture modifications. Basically ,For medical purposes,
[18]BioBERT (Bidirectional Encoder Representations from Transformers for Biomedical Text Mining), which is a domain-specific language representation model pre-trained on large-scale biomedical corpora. With almost the same architecture across tasks, BioBERT largely outperforms BERT and previous state-of-the-art models in a variety of biomedical text mining tasks when pre-trained on biomedical corpora.

EXISTING SYSTEM

An existing system for automating the extraction and classification of medical terms from prescription images typically integrates multiple advanced technologies to achieve high accuracy and efficiency.The automation of extracting and classifying medical terms from prescription images involves a complex interplay of OCR, NLP , and machine learning technologies. By integrating these components, existing systems can provide accurate and efficient processing of prescription information, significantly enhancing the workflow in healthcare settings and ensuring better patient care.

i)Optical Character Recognition (OCR):
This project utilizes Optical Character Recognition (OCR) to automate the extraction and classification of medical terms from prescription images. Key preprocessing techniques, such as binarization, noise reduction, and skew correction, enhance text recognition accuracy. Once prepared, the OCR engine analyzes pixel data to identify characters and words, ultimately producing machine-readable text for further processing and analysis.

ii)Natural Language Processing (NLP)
Processing prescriptions relies on Natural Language Processing techniques for accurate classification of medical terms. Named Entity Recognition (NER) identifies entities like drug names and dosages using libraries such as spaCy and NLTK. Additionally, text classification categorizes terms, while machine learning algorithms enhance accuracy through pattern recognition, ultimately improving healthcare delivery and patient safety by extracting meaningful insights from prescription texts.

iii)Machine Learning Integration:
Machine learning integration enhances the system's ability to learn from data and improve performance over time through supervised training on annotated prescription data. Deep Learning techniques, like Convolutional Neural Networks, analyze image features, while RNNs and Transformers process sequential text, capturing contextual relationships. This combination enables accurate extraction and classification, allowing the system to adapt to evolving medical language and data.

PROPOSED SYSTEM

The proposed solution for the research project involves a two-step process to extract and classify medical terms fromprescription images. First, EasyOCR is utilized to extract text from images, allowing for the recognition of words without bounding box details. This is achieved through the extract_text_from_image function, which initializes an EasyOCR reader and compiles recognized words into a single string. Next, the extracted text is processed using a pre-trained BioBERT model for Named Entity Recognition(NER), implemented through the classify_medical_terms function. This function categorizes the recognized terms into predefined groups such as DRUG, DOSAGE, DURA TION, TIME, and PA TIENT by leveraging BioBERT's capabilities. The entire workflow is encapsulated in the process_medical_prescription function, which seamlessly integrates text extraction and classification, ultimately providing a structured output of classified medical terms.This approach not only enhances the efficiency of processing medical prescriptions but also contributes significantly to the accuracy and reliability of information extraction in healthcare applications.

i)BioBERT:
BioBERT is a fine-tuned language model built on the BERT architecture, specifically designed for named entity recognition (NER) in the biomedical domain. Unlike general BERT, it is trained on large biomedical datasets, including PubMed articles and clinical notes, allowing it to effectively understand medical terminology and context. BioBERT undergoes further task-specific fine-tuning on annotated datasets, significantly enhancing its accuracy in identifying entities such as diseases and drugs. Its applications include analyzing clinical notes and mining biomedical literature for relevant entities. BioBERT consistently demonstrates superior performance on biomedical NER benchmarks,making it a valuable tool for researchers and practitioners.

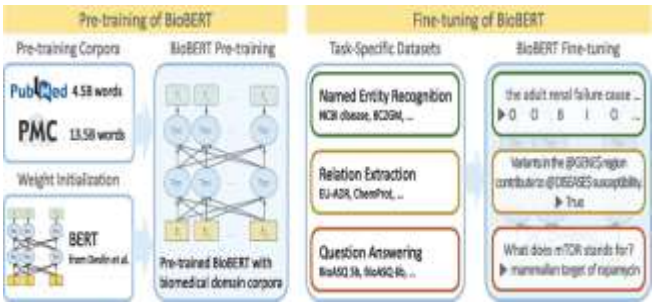

Fig 1. Overview of the pre-training and fine-tuning of BioBERT

ii)EasyOCR:

EasyOCR serves as the primary tool for text extraction from prescription images. It initializes a reader specifically configured for English language recognition, enabling it to accurately identify and extract words from visual data. The *readtext* function is employed to process the image, yielding results without the need for bounding box details, which simplifies the output. By joining the recognized words into a single string, EasyOCR facilitates the subsequent classification of medical terms. Its effectiveness in recognizing diverse fonts and layouts makes it particularly suitable for handling the varied presentations of prescription texts, enhancing the overall data extraction process.

IMPLEMENTATION FRAMEWORK

The workflow automates the extraction and classification of medical information from prescription images using EasyOCR for text recognition and a fine-tuned BioBERT model for Named Entity Recognition, facilitating efficient data management and analysis in healthcare applications.

Start by taking a clear picture of the prescription.Ensure that all details are legible. This image will serve as the foundation for processing and extraction of crucial medical data.

- Leverage EasyOCR to swiftly extract text from the captured image. This powerful tool identifies text accurately,setting the stage for further analysis and classification of medical terms.
- Employ a pre-trained NER model like BioBERT to classify identified medical terms. This process helps in distinguishing between various categories like medications, dosages, and durations essential for patient care.
- Carefully review the extracted data for accuracy. Ensure that all terms have been classified and that nothing critical is overlooked, as this information is vital for patient safety and care continuity.
- Categorize the extracted information by compiling lists of drugs, dosages, durations, times, and patient details, creating a structured and accessible database for future reference.

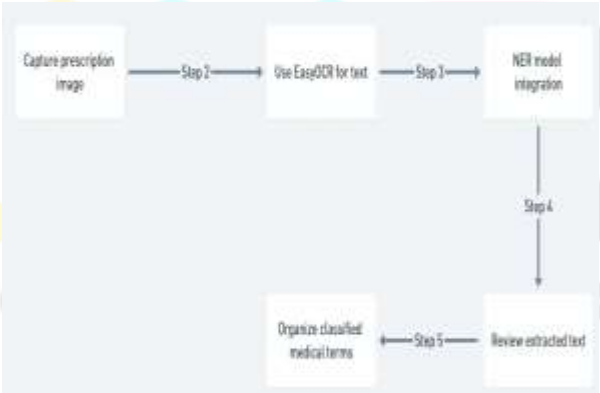

Fig 2. Project Workflow

ARCHITECTURE DESIGN

The diagram showcases an advanced "Medical Transcription Image-to-Text Architecture" aimed at automating the transcription of medical documents or images into structured, machine-readable text. This architecture leverages both Optical Character Recognition (OCR) and Natural Language Processing (NLP) techniques to extract and process medical information from scanned documents or images, such as patient reports, prescriptions, and clinical notes.

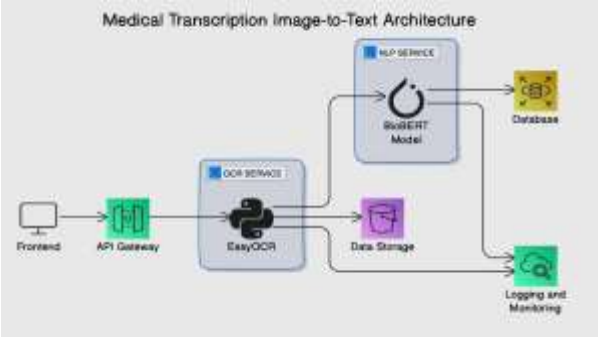

Fig 3. Architecture Design

The process begins at the Frontend, where users—such as medical practitioners, healthcare staff, or transcription teams—interact with the system. They upload images or documents containing medical text (e.g., handwritten notes, printed medical records) through a web interface or another application. These uploads are directed through an API Gateway, which acts as an intermediary, routing user requests and managing traffic between different services within the architecture. This API Gateway ensures that the system scales well and handles multiple requests efficiently.

Once the medical image is received, it moves into the OCR Service, powered by EasyOCR, an open-source OCR tool known for its ability to recognize text from a variety of languages and fonts, including handwritten and printed text.EasyOCR extracts the raw text from the medical document images, which might include patient information, medical terminologies, prescriptions, or other relevant data. This raw text is then stored temporarily in a Data Storage bucket, likely a cloud-based storage service, allowing for flexible access by subsequent processes. After the text is extracted by the OCR component, it is processed further by the NLP Service. This service utilizes a BioBERT Model, a specialized version of the BERT (Bidirectional Encoder Representations from Transformers) language model. BioBERT has been fine-tuned specifically for biomedical and healthcare-related tasks. It processes the extracted text, parsing it to understand the medical terminology, extract meaningful entities (e.g., diagnoses,treatments, medications), and structure the data for analysis.

This step is crucial for converting unstructured raw text into a format that can be used for clinical decision-making, research, or reporting. The structured output is then stored in a Database for easy retrieval and use in downstream applications.
To ensure reliability and robustness, the entire process is equipped with Logging and Monitoring mechanisms. These tools track system performance, record errors or warnings, and provide insights into the operations. Logging and monitoring ensure that the system operates smoothly and allows developers or operators to quickly identify and resolve any issues, such as OCR inaccuracies, failed API requests, or NLP model errors.

Overall, this architecture provides a seamless workflow for converting medical images into structured text. It combines OCR and NLP technologies to automate the extraction of information from medical documents, reducing the need for manual transcription, improving accuracy, and enabling easier analysis of medical data for healthcare providers.

PERFORMANCE METRICS

In our investigation into the effects of varying corpus sizes during the pre-training of BioBERT v1.0, particularly when augmented with the PubMed dataset, we conducted a comprehensive analysis aimed at understanding how these factors influence performance on Named Entity Recognition(NER) tasks.

For our experiments, we set the number of pre-training steps to a total of 200,000. We meticulously varied the size of the PubMed corpus, allowing us to examine the impact of using different volumes of textual data on the model's effectiveness.The results are illustrated in Figure 2(a), which displays the performance metrics of BioBERT v1.0 when tested across three distinct NER datasets: NCBI Disease, BC2GM, and BC4CHEMD. Our findings indicate a clear trend:
Pre-training on a corpus containing 1 billion words yields notable improvements in performance. Moreover, as the size of the PubMed corpus increases, up to a size of 4.5 billion words, we observe a consistent enhancement in the model's performance on each of the NER datasets. This suggests that a larger corpus can provide richer contextual information, which is critical for effectively identifying and classifying biomedical entities.

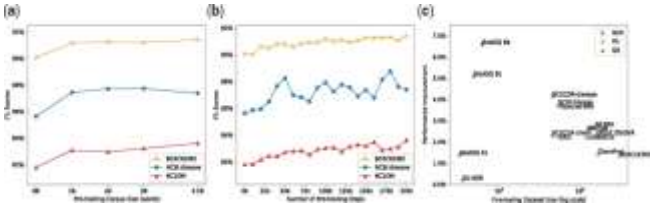

Fig.4.(a) Effects of varying the size of the PubMed corpus for pre-training.
(b) NER performance of BioBERT at different checkpoints.
(c) Performance improvement of BioBERT v1.0 (+ PubMed + PMC) over BERT

In addition to analyzing the impact of corpus size, we also saved the pre-trained weights of BioBERT v1.0 at various intervals during the pre-training process. This approach allowed us to assess how the number of pre-training steps correlates with performance in subsequent fine-tuning tasks. The results of this investigation are depicted in Figure 2(b). The data clearly demonstrates that as the number of pre-training steps increases, the performance of BioBERT v1.0 on each of the three NER datasets continues to improve.

This trend emphasizes the importance of extensive pre-training; more steps allow the model to better capture the nuances of language and context, leading to more accurate predictions during fine-tuning.
Finally, we expanded our analysis to compare the absolute performance gains of BioBERT v1.0 when combined with both the PubMed and PMC corpora against the baseline BERT model across a comprehensive set of 15 datasets.
Figure 2(c) summarizes these findings, illustrating that BioBERT consistently outperforms BERT in terms of F1 scores for NER and Relation Extraction (RE) tasks, as well as Mean Reciprocal Rank (MRR) scores for Question Answering (QA) tasks. These results underscore the significant advancements achieved by BioBERT, indicating that the incorporation of specialized biomedical corpora, coupled with rigorous pre-training, leads to substantial improvements in performance across various biomedical NLP applications.

RESULT AND DISCUSSION

In this project, we developed a comprehensive pipeline for extracting and classifying medical information from prescription images using EasyOCR for text extraction and a fine-tuned BioBERT model for Named Entity Recognition (NER). The pipeline effectively processes images by first utilizing EasyOCR to convert the visual text into a structured format. Subsequently, the extracted text is analyzed by the BioBERT model to identify and categorize key medical terms, including drugs, dosages, durations, times, and patient identifiers. Our implementation demonstrates the ability to convert unstructured data from images into actionable medical insights, significantly aiding healthcare professionals in managing and interpreting prescription information.

The results of the pipeline indicate high efficacy in recognizing and classifying medical terms, which is crucial for ensuring accurate medication administration and patient care. The integration of EasyOCR and BioBERT allows for a seamless transition from image to information, minimizing human error in interpreting prescriptions. Furthermore, the use of a specialized model trained on biomedical texts enhances the accuracy of entity recognition, making the pipeline a valuable tool in clinical settings. Future research could explore further optimization of the model and pipeline to improve classification precision and expand its applicability across diverse medical documentation types, ultimately contributing to enhanced patient safety and healthcare efficiency.

REFERENCES

[1] Hassan, E., Tarek, H., Hazem, M., Bahnacy, S., Shaheen, L., & Elashmwai, W. H. (2021). “Medical prescription recognition using machine learning.” Annual Computing and Communication Workshop and Conference : 0973-0979.

[2] Balci, Batuhan, Dan Saadati, and Dan Shiferaw. "Handwritten text recognition using deep learning." CS231n: Convolutional Neural Networks for Visual Recognition, Stanford University, Course Project Report, Spring (2017): 752-759.

[3] Michael, Johannes, et al. "Evaluating sequence-to-sequence models for handwritten text recognition." 2019 International Conference on Document Analysis and Recognition (ICDAR). IEEE, 2019.

[4] de Sousa Neto, Arthur Flor, et al. "HTR-Flor: A deep learning system for offline handwritten text recognition." 2020 33rd SIBGRAPI Conference on Graphics, Patterns and Images (SIBGRAPI). IEEE, 2020.

[5] Sethi, P . S., Gupta, M., Kumar, P ., & Kaur, G. (2023). " Simplifying Handwritten Medical Prescription: OCR Approach Check for updates." Machine Intelligence and Data Science Applications: Proceedings of MIDAS 2022 : 1-47.

[6] D.R. V edhaviyassh; R. Sudhan; G. Saranya; M. Safa; D. Arun "Comparative Analysis of EasyOCR and TesseractOCR for Automatic License Plate Recognition using Deep Learning Algorithm." Published in:(2022)6th International Conference on Electronics, Communication and Aerospace Technology. DOI: 10.1109/ICECA55336.2022.10009215

[7] Mhiri, Mohamed, Christian Desrosiers and Mohamed Cheriet, "Word recognition and recognition through deep integrated image and text embedding." Pattern View 88(2019), 312-320.

[8] Xue, Wenyuan, Qingyong Li and Qiyuan Xue. (2020)"Text Detection and Recognition for Images of Medical Laboratory Reports with a Deep Learning Approach." IEEE Access, (8):407–416.

[9] A. M. Sabu and A. S. Das, "A Survey on various Optical Character Recognition Techniques", in: Proceedings of the 2018 Conference on Emerging Devices and Smart Systems (ICEDSS), Tiruchengode, 2018, pp. 152-155. doi: 10.1109/ICEDSS.2018.8544323.

[10] Lily Rojabiyati Mursari , and Antoni Wibowo (2021) "The Effectiveness of Image Preprocessing on Digital Handwritten Scripts Recognition with The Implementation oF OCR Tesseract" Computer Engineering and Applications Journal 10(3):177-186

[11] S. M. Shamim, "Handwritten digital recognition using machine learning algorithms" .Global Journal of Computer Science and Technology, 2018.

[12] Fanany, Mohamad Ivan and others. (2017) "Handwriting recognition on form document using convolutional neural network and support vector machines (CNN-SVM). 5th International Conference on Information and Communication Technology (ICoIC7), IEEE:16

[13] Dash, B. (2021). "A hybrid solution for extracting information from unstructured data using optical character recognition (OCR) with natural language processing (NLP)."Academia.

[14] Surya Nath RS, and S. Afseena. "Handwritten Character Recognition–A Review." International Journal of Scientific and Research Publications (2015).

[15] Salma Shofia Rosyda, and Tito Waluyo Purboyo. "A Review of Various Handwriting Recognition Methods." International Journal of Applied Engineering Research 13.2
(2018): 1155-1164

[16] Devlin J. et al. "Bert: pre-training of deep bidirectional transformers for language understanding."(2019) In: Proceedings of the 2019 Conference of the North American Chapter of the Association for Computational Linguistics:Human Language Technologies, V olume 1 (Long and Short Papers), Minneapolis, MN, USA. pp. 4171–4186. Association for Computational Linguistics.

[17] Giorgi J.M. , Bader G.D. "Transfer learning for biomedical named entity recognition with neural networks."(2018) Bioinformatics, 34, 4087.

[18]Jinhyuk Lee, Wonjin Yoon, Sungdong Kim, Donghyeon Kim, Sunkyu Kim, Chan Ho So, Jaewoo Kang,"BioBERT: a pre-trained biomedical language representation model for biomedical text mining." Bioinformatics, Volume 36, Issue 4, February 2020, Pages 1234–1240.

[19]Habibi M. et al. "Deep learning with word embeddings improves biomedical named entity recognition." (2017) Bioinformatics, 33, i37–i48.

[20]Lin C. et al. "A bert-based universal model for both within-and cross-sentence clinical temporal relation extraction."(2019) In: Proceedings of the 2nd Clinical Natural Language Processing Workshop, Minneapolis, MN, USA. pp. 65–71. Association for Computational Linguistics.

[21]Yang, Tianjiao, Ying He and Ning Yang. (2022) "Named Entity Recognition of Medical Text Based on the Deep Neural Network." Journal of Healthcare Engineering (2022):1–10.

[22]Landolsi, Mohamed Yassine, Lobna Hlaoua and Lotfi Ben Romdhane.(2022) "Information Extraction from Electronic Medical Documents: State of the Art and Future Research Directions." Knowledge and Information Systems (65):463-516.

[23]Peters M.E. et al. "Deep contextualized word representations."(2018) In: Proceedings of the 2018 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, Volume 1 (Long Papers), New Orleans, LA. pp. 2227–2237. Association for Computational Linguistics.