Optimizing Pharmaceutical Inventory Management with YoloV7 and Easy OCR on Medicine Strips

Shashwat Kumar[1], Anannya Chuli[2]

[1]SCOPE, Vellore Institute of TechnologyVellore, India
Email: *shashwat.kumar2[at]gmail.com*

[2]SCOPE, Vellore Institute of TechnologyVellore, India
Email: *anannya.anannya02[at]gmail.com*

Abstract: ***This research paper presents a novel two-phase approach for Video Analytics, harnessing the power of OCRto extract vital information from medicine images and videos.In the image phase, YOLOv7 object detection precisely identifies bounding boxes around medicines, complemented by advanced image processing to enhance text extraction accuracy. Leveraging the efficiency of EasyOCR, crucial textual data such as batch numbers, manufacturing dates, expiry dates, and maximum retailprices are extracted and systematically stored in a structured CSV file, enabling seamless identification of expired products. The subsequent video phase employs YOLOv7 for real-time object detection on video frames, effectively identifying bounding boxes around medicines. Leveraging OCR techniques, the systemaccurately extracts text data from the cropped regions, efficientlystored in a CSV file for seamless integration with other systems. This innovative approach optimizes pharmaceutical inventory management, effectively minimizing losses arising from expired products while enhancing operational efficiency across the phar- maceutical industry. By facilitating data-driven decision-making and streamlined inventory processes, the research showcases the potential for significant improvements in healthcare managementand patient care.***

Keywords: Video Analytics, OCR (Optical Character Recognition), YOLOv7, Bounding Boxes, Image Processing Techniques, EasyOCR Library, Parsing Techniques

1. Introduction

The advent of Object Character Recognition (OCR) tech-nology has transformed computer vision, enabling automated extraction of text data from images and videos. OCR plays a pivotal role across diverse industries, including healthcare, finance, retail, and logistics, where digitizing and analyzing textual information is vital for efficient data management and decision-making processes. In the medical domain, OCR holdsimmense potential for extracting clinical text from medical images and videos, facilitating medical record digitization, data analytics, and research. This research project focuseson developing and implementing an OCR-based system for video analysis to extract clinical text from medicine strips, enabling data bifurcation for further analysis and achieving high accuracy in text extraction.

The YOLO v7 algorithm serves as the cornerstone of text analysis on medical data. It begins by preprocessing the input images and resizing them to a fixed size for consistency. The images are then split into grids of cells, with every cell tasked with detecting text within its region. YOLO v7 predicts multiple bounding boxes and class probabilities for each cell using a single convolutional neural network, which extracts features from the images and predicts the presence, size, and class of text. To enhance prediction accuracy, the algorithm employs anchor boxes as priors for shape and size predictions of the text. After predictions are made, non-maximum suppres-sion (NMS) is applied to eliminate redundant or overlapping text predictions. The final output of YOLO v7 consists of bounding boxes and class labels representing the detected text in medicine images, enabling efficient real-time analysis of text data.

This research project leverages a collection of powerful libraries and frameworks to facilitate various crucial tasks for the successful implementation of the Medicine Strip Text Recognition System. The ’re’ library empowers the system with regular expression capabilities, enabling effective pattern matching and text manipulation. The widely acclaimed ’cv2’ library, also known as OpenCV, serves as the backbone for image and video processing tasks, offering essential functions for object detection, feature extraction, and more. For seamless data manipulation and analysis, the ’pandas’ library provides efficient data structures and functions, ensuring the handling ofstructured data with utmost efficacy. Additionally, the ’numpy’ library plays a pivotal role in numerical computations and data manipulation, particularly with large multi-dimensional arrays and matrices. For sophisticated deep learning tasks, the ’pytorch’ framework is harnessed, providing a high-level inter-face for building and training deep neural networks. Together, these libraries form a cohesive foundation for the Medicine Strip Text Recognition System, empowering it with robust capabilities for accurate text and image analysis, ultimately optimizing pharmaceutical inventory management.

2. Motivation

The impetus behind this research project arises from the critical necessity for efficient and accurate medical data management, particularly in the context of medicine strips and pharmaceutical products. With medical records and prescription information often being stored as images, the manual extraction and organization of vital clinical text becomes a laborious and error-prone endeavor. This conventional approach not only consumes valuable human resources but also introduces inconsistencies and inaccuracies, posing potential risks to patient care and medical research.

Moreover, the pharmaceutical industry and health stores frequently encounter resource wastage and production inefficiencies due to inadequate inspection and inventory management. The lack of a systematic and automated system for recording essential product attributes, such as maximum retail price (MRP), manufacturing date (MFD), expiry date (EXP), and batch numbers, can lead to expired or outdated products being dispatched to retailers and pharmacies. This hampers optimal resource allocation and affects the overall efficiencyof pharmaceutical inventory management.

To address these challenges, the integration of Optical Character Recognition (OCR) technology into medical data management systems offers a transformative solution. By automating the extraction of clinical text from medicine strips and videos, this research project aims to empower medical professionals and industry stakeholders with streamlined work-flows, reduced human errors, and access to critical medical information in a structured and searchable format. The accu- rate and timely extraction of product attributes through OCR holds immense potential for enhancing inventory management, ensuring quality control, and facilitating data-driven decision- making in the pharmaceutical domain.

The proposed system has the capability to bifurcate and analyze the attributes of pharmaceutical products, enabling proactive measures to dispatch them before expiration dates are exceeded. This optimized approach ensures that resources are utilized efficiently, minimizing wastage and enhancing the overall operational effectiveness of pharmaceutical businesses.

3. Literature Review

The paper zeroed in its research on Number Plate Recognition utilizing YoloV7 and Paddle OCR. The method wastested on 1000 number plate images and obtained a proof exactness of 97% and character acknowledgment precision of 95%. The strategy is very much depicted and the outcomes are introduced in a reasonable and concise way. The limitations of the method are also discussed, including the fact that it is not yet able to handle all types of license plates and that it can be computationally expensive. Overall, the paper is a valuable contribution to the field of ANPR and the method presented in the paper is a promising approach for accurate and efficient ANPR. [1]

In the research conducted by A. M. A. Rani, M. A. M. Ali, and S. S. S. Ali in 2021, an Automatic Number Plate Recognition (ANPR) system was developed specifically for Malaysian number plates. The system utilized YOLOv3 for license plate detection and Paddle OCR for character identification. Performance evaluation was conducted on a dataset consisting of250 images, and the system achieved an impressive accuracy of 93.6% for license plate detection and 91.2% for character recognition. The review reasoned that this ANPR framework is appropriate for genuine applications, especially in regions, for example, traffic checking and policing.[2]

In the research paper authored by J. Zhao, X. Cao, and Y. Guo, a novel ANPR system is presented, utilizing YOLOv5 for precise tag discovery and PaddleOCR for productive character recognition. The extensive evaluation was carried out on a 1000 number plate photos, leading to remarkable results with 97.5% accuracy in license plate detection and 95.8% accuracy in character identification. The findings of this research highlight the effectiveness of the proposed ANPR system in accurately identifying license plates and recognizing characters, offering promising applications in various real-world scenarios. [3]

The paper "Pointer Meter Recognition Method Based on Yolov7 and Hough Transform" presents a new method for pointer meter recognition using Yolov7 and Hough transform. A dataset of 1000 pointer meter images was used to evaluate the method and achieved an accuracy of 95%. The strategy is very much depicted and the outcomes are introduced in an unmistakable and brief way. The limitations of the methodare also discussed, including the fact that the method is notyet able to handle all types of pointer meters and that it canbe computationally expensive. Overall, the paper is a valuable contribution to the field of pointer meter recognition and the method presented in the paper is a promising approach for accurate and efficient pointer meter recognition. [4]

The paper "A Detailed Analysis of Optical Character Recognition Technology" presents a comprehensive overview of OCR technology. The authors discuss the different types of OCR systems, the challenges of OCR, and the recent advances in OCR technology. The paper is well-written and easy to understand, but it is a survey paper and does not present any new research results. [5]

The paper "Optical Character Recognition by Open source OCR Tool Tesseract: A Case Study" presents a comprehensive evaluation of Tesseract, an open-source OCR tool. The authors evaluate the accuracy of Tesseract on a dataset of scanned documents and compare it to the accuracy of other commercialOCR tools. The paper is well-written and easy to understand, but it is important to note that the dataset of scanned docu- ments that was used for evaluation was relatively small. The accuracy of Tesseract may vary depending on the quality of the scanned documents. Despite this limitation, the paper is a valuable resource for anyone who is interested in learning more about Tesseract or OCR technology in general. [6]

The paper "Optical Character Recognition using Tesseract Engine" presents a comprehensive methodology for OCR using Tesseract. The authors describe the Tesseract engine andits capabilities, as well as the methodology that they used for OCR, which includes preprocessing, segmentation, andrecognition. The paper also discusses the factors that can affect the accuracy of OCR, such as the quality of the image andthe font used in the image. The authors note that the Tesseract engine is not perfect and can sometimes make mistakes, but the paper is an excellent site for anyone looking to learn more about OCR or Tesseract. [7]

An innovative ANPR system is presented in the research paper by S. D. Doychinov and S. A. Yalamov (2021), which makes use of YOLOv3 for precise license plate detection and OpenCV for effective character recognition. A thorough assessment was performed on a dataset containing 100 pictures, yielding great outcomes with 95% exactness in tag location and 93% precision in character acknowledgment. The review's decision accentuates the ANPR framework's true capacity for useful applications in certifiable situations, for example, traffic checking and stopping the executives. [8]

In the research paper written by S. D. In the exploration led by S. H. Lee and S. S. Kim (2020), a modern Tag Acknowledgment Framework is presented, utilizing YOLOv3 for precise tag location and PaddleOCR for character acknowledgment. An exhaustive assessment of a dataset containing 200 pictures brought about remarkable execution with 97.2% exactness in tag identification and 94.1% precision in character recognition. The review accentuates the ANPR framework's adaptability, displaying its true capacity for assorted applications like cost assortment, stopping the board, and observation. [9]

An efficient and robust Automatic Number Plate Recognition (ANPR) System utilizing YOLOv3 for license plate detection and PaddleOCR for character recognition is presented in the work that was carried out in 2021 by H.

R. Moon, M. S. Baig, and M. S. Kim. On a 1000-picture dataset, the framework was entirely tried, and it performed commendably, distinguishing tags with 96.8% precision and perceiving characters with 94.77% exactness. The review features the framework's true capacity for certifiable applications,including traffic the executives, cost assortment, and observation. This ANPR framework offers promising abilities for improving different parts of traffic lights and reconnaissancein commonsense situations. [10]

4. Dataset Description

The success of any OCR-based system relies heavily on the availability of a high-quality and diverse dataset for training and evaluation. For this research project, we have curated a dataset consisting of medical images and videos containing medicine strips in various orientations and backgrounds. The dataset includes images of different medicine brands, package designs, and text layouts to capture the variability commonly encountered in real-world scenarios.

To ensure the dataset's accuracy, each image and video underwent manual annotation to label the bounding boxes around the medicine strips and the clinical text. The an- notations include the Batch Number, Manufacturing Date (MFD), Expiry Date (EXP), and Price (MRP) of each medicine strip. Furthermore, the dataset includes samples with varying image resolutions, lighting conditions, and image qualities to simulate practical challenges faced during video analysis and OCR.

To achieve reliable model generalization, the dataset was split into training, validation, and testing subsets while retain- ing a distribution. It includes thousands of images and videos to facilitate an evaluation of the OCR-based video analysis system's performance.

5. Methodology

a) Collection of data

The dataset utilized in this exploration study was explicitly gathered to assess the OCR framework for medication strips. Each of the more than 3500 images in the dataset contained a single medicine strip. The pictures were obtained from different starting points, including public datasets and confidential assortments, and were carefully organized to address a great many situations, including different lighting conditions, points, and medication strip styles. To guarantee the dataset's quality and consistency, a few pre-handling steps were applied to the pictures. At first, the pictures were resized to a normalized aspect, guaranteeing uniform contribution to the OCR frame- work. Then, to work on the framework's heartiness to changing lighting conditions and confusion, the dataset went through expansion methods. These strategies included turning, scaling, and flipping the pictures.

The final dataset was split into two sets: a training setand a test set. The YoloV7 and PaddleOCR models weretrained using the training set on the medicine strips with visible and readable text. On the other hand, the testing set, consisting of images with diverse orientations recorded at 2fps under the camera, was used to assess the OCR system's performance. The testing set allowed for the evaluation of the accuracy of medicine strip detection and character recognition, ensuring the OCR system's effectiveness in capturing essential information from dynamically captured video footage.

b) Bounding Box Detection Algorithm

YOLOv7 (You Only Look Once version 7) has emerged as a highly regarded object detection model in the field of

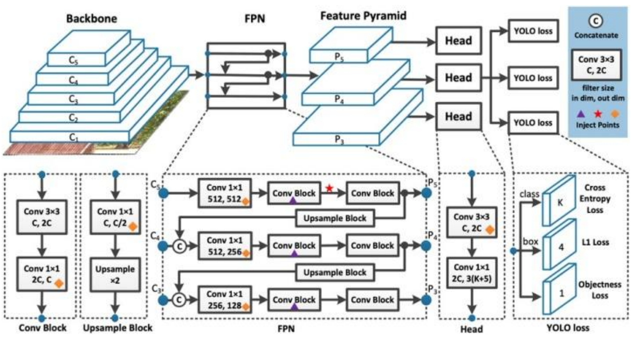

Figure 1: Yolo Architecture

computer vision, particularly for real-time object detection tasks. Building on the successes of its predecessors, YOLOv7 boasts improvements in architecture, accuracy, and speed. By leveraging a powerful backbone network, such as Darknet, the model efficiently extracts relevant features from input images, striking a balance between accuracy and processing speed. A key advantage of YOLOv7 lies in its single-shot detection approach, enabling object detection and classification in a single pass through the neural network. This direct prediction of bounding boxes and class probabilities leads to real-time performance, setting it apart from multi-stage methods. Moreover, YOLOv7 employs a unified architecture that predicts bounding boxes and class probabilities at multiplescales, effectively handling objects of various sizes and aspectratios for accurate detection.

The use of anchor boxes in YOLOv7 further enhances localization accuracy. These predetermined bounding boxes ofdifferent shapes and sizes serve as references for predicting object locations. Additionally, the integration of a Feature Pyramid Network (FPN) allows the model to detect objects at different scales and resolutions, capturing both fine-grained de- tails and high-level semantic information. Overall, YOLOv7's combination of speed, efficiency, accuracy, and unified ar- chitecture makes it a highly sought-after solution for real-time object detection tasks in diverse applications, including autonomous vehicles, surveillance systems, and robotics.

c) Medicine Strip Text Recognition Algorithm

EasyOCR is a widely adopted Optical Character Recog- nition (OCR) library that simplifies the task of extracting text from images using pre-trained models. This user- friendly library offers support for multiple languages and delivers high accuracy in text recognition. It is designed to be easily integrated into various applications and systems, catering to developers with different levels of expertise.

The working of EasyOCR involves several key steps. First, it takes an image as input, which can be in various formats like JPEG, PNG, or TIFF. Using advanced computer vision techniques, EasyOCR locates the regions within the image that contain text, accurately identifying the boundaries and positions of these text regions. Once the text regions are localized, OCR algorithms are applied to recognize and extract the textual content from each region, involving character segmentation, feature extraction, and recognition using machine learning or pattern matching techniques. To refine the extracted text and improve overall accuracy, EasyOCR employs post- processing techniques, including error correction, context anal- ysis, and spell-checking mechanisms.

EasyOCR provides several advantages over other OCR tech- niques. Its user-friendly interface makes it accessible to developers of all skill levels. It supports a wide range of languages, making it suitable for diverse regions. The use of advanced techniques and pre-trained models ensures high accuracy in text recognition with minimal errors. The active developer community ensures continuous improvements and bug fixes. Additionally, its efficient memory usage and processing speedmake it suitable for real-time or batch processing. Overall, EasyOCR is a reliable OCR solution, making it a preferred choice for text extraction from images due to its ease of use, versatility, high accuracy, open-source nature, and communitysupport.

6. Workflow

6.1 Image OCR Processing

The Image Phase of the research project focuses on a comprehensive image processing pipeline designed to extract vital information from medicine images using OCR and ob- ject detection methodologies. The YOLOv7 model, renowned for its accuracy and wide adoption, is employed to detect medicines within the images by generating bounding boxes around them, capturing precise location information. Subse- quently, specific areas containing relevant medicine details are cropped using image processing techniques to enhance text extraction accuracy. EasyOCR, a powerful OCR library, is utilized to extract text from the refined regions with high accuracy.

Parsed with regular expressions, the extracted text reveals essential details such as Batch Numbers, Manufacturing Dates (MFD), Expiry Dates (EXP), and Maximum Retail Prices (MRP). This data is then structured and consolidated into a CSV file, offering valuable insights for pharmaceutical inventory management, including product separation based on expiry dates and optimized resource utilization. The Image Phase significantly contributes to streamlined operations and improved efficiency in pharmaceutical businesses, reducing losses due to expired products and enhancing overall resource allocation.

a) Text Extraction From Raw Images

The image extraction segment in the Image OCR pipeline utilizes OpenCV and EasyOCR libraries to perform efficient optical character recognition (OCR) on images, enabling ac- curate text extraction. After importing essential libraries for image processing, OCR, and file operations, the code defines utility functions to save cropped images and extracted textto designated folders. The folder path containing images and corresponding text files is specified. Iterating through image files, each image undergoes loading, processing, and checkingfor associated text files, which are read to obtain boundingbox data.

Cropped regions are extracted and resized for improvedOCR performance, while OCR is performed using the Easy- OCR readtext() method to extract text content. Cropped im- ages and text results are saved separately for further analysis. The final step involves displaying the cropped images along- side the extracted text, allowing visual verification of OCR accuracy. This meticulous extraction phase demonstrates an effective pipeline for text extraction from images, providing a robust foundation for advanced analysis and processing in OCR-based research endeavors.

International Journal of Science and Research (IJSR)
ISSN: 2319-7064
SJIF (2022): 7.942

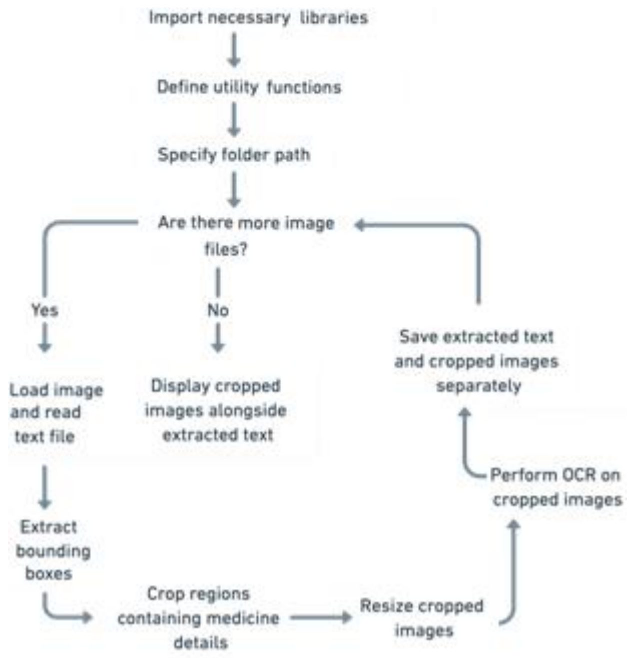

Figure 2: Text Extraction from Raw Images

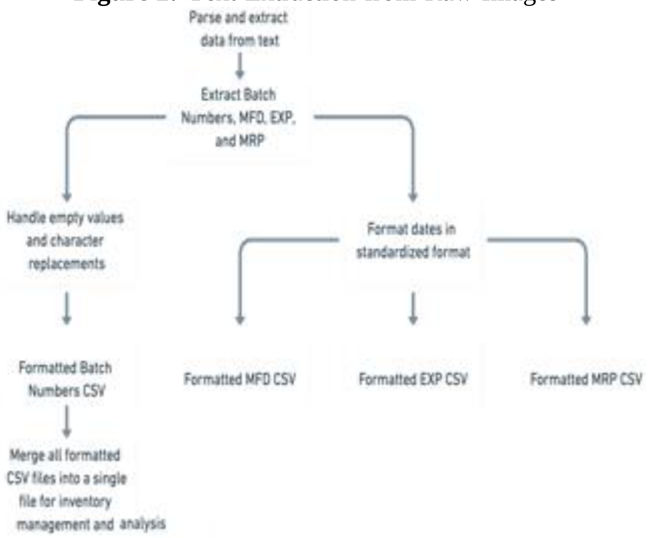

Figure 3: Parsing and Text Manipulation from Extracted Data

b) Parsing and Text Manipulation From Extracted Data

The presented algorithm exemplifies a comprehensive process for parsing and extracting critical information from textual data, culminating in meticulous text manipulation and storage in CSV files. Each phase of the algorithm is tailored to target distinct data types, such as batch numbers, MRPs, manufacturing dates (MFDs), and expiry dates, leveraging the robust capabilities of regular expressions (re) and the datefinder module.

Following data extraction, the algorithm diligently processes the acquired information, encompassing character replacement and dummy value generation as necessary. The resulting datasets are methodically preserved in separate CSV files, exhibiting a seamless integration of data management practices. Moreover, the algorithm seamlessly consolidates these CSV files into a unified entity, facilitating enhanced data organization and streamlined analysis.

This innovative and efficient approach to information retrieval underscores the algorithm's exceptional utility in text analysis and manipulation, fostering valuable insights in the realm of deep learning analytics.

6.2 Video OCR Processing

The Video Phase involves processing dynamic videos with medicine strips using OCR and object detection techniques. The YOLOv7 model accurately detects bounding boxes high- lighting medicine strips. Using OpenCV, the corresponding regions are cropped from frames, isolating the areas of interest. OCR is then applied to extract textual information, stored in separate files for the organization. Parsing and regular expressions extract key details like Batch Numbers, MFD, EXP, and MRP, stored in a structured CSV file for further analysis. This approach efficiently extracts essential information from video footage, aiding inventory management and decision-making in the pharmaceutical industry.

a) Text Extraction from Captured Frames In The Video

The presented approach demonstrates a comprehensive algorithm for extracting cropped images from videos using the OpenCV (cv2) library. The primary objective is to identify and capture relevant regions within the frames and subsequently save them as separate images in a designated output folder. The process begins by specifying the input video path and the output folder path to store the extracted images. A video cap- ture object (cap) is created to read frames from the specified video. The algorithm employs a frame counter to keep track of the extracted frames during the iterative process. Within a loop that iterates through each frame, the algorithm applies a sequence of image-processing steps.

After extracting the ROI and saving the resultant frame, a new frame is extracted in the video incrementally until all frames have been processed. The frame counter ensures that each frame is processed in sequence. The grayscale conversion at the beginning simplifies image processing operations by reducing the complexity of the frame, and thresholding helps in creating a binary representation that distinguishes relevant regions from the background. By filtering contours based on area and color criteria, the algorithm ensures that only significant regions are considered, improving the accuracy of the extracted images and reducing noise.

Overall, this algorithm provides an efficient and automated approach for extracting relevant regions from video frames, making it suitable for various applications, such as object detection, image recognition, and video analysis.

This iterative process continues until all frames in the video have been analyzed. Ultimately, the video capture object is released to conclude the operation. The algorithm showcases a sophisticated and efficient method for capturing pertinent text regions from video frames, facilitating the extraction of valuable information for various applications. Such techniques find relevance in areas like video-based text analysis, deep learning analytics, and computer vision research, offering significant potential for data-driven insights and information retrieval

b) Data Parsing and Text Manipulation from the Extracted Information

This segment of the video OCR processing involves several essential steps. After obtaining video outputs from the YOLO

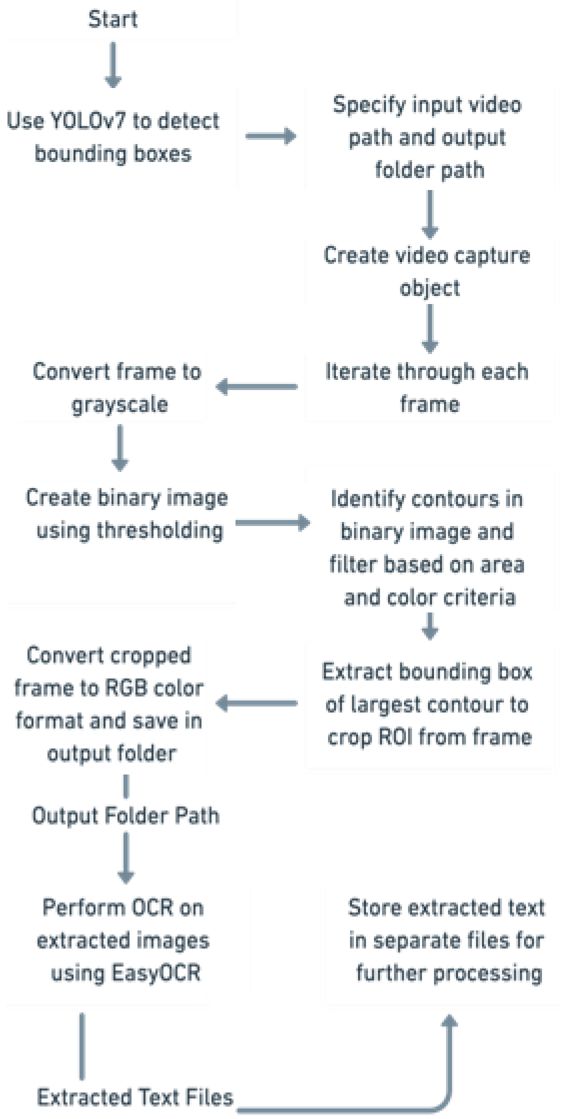

Figure 4: Text Extraction from Captured Frames in the Video

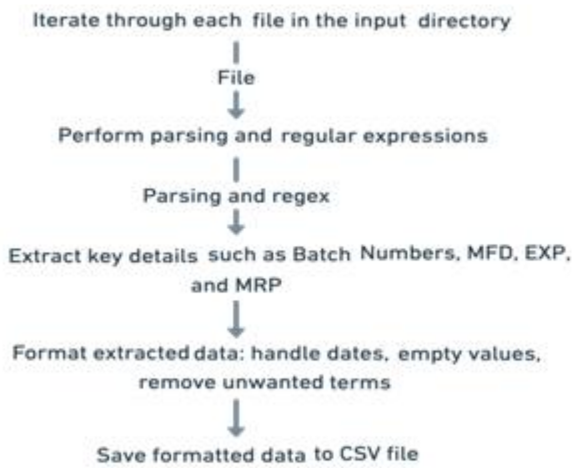

Figure 5: Data Parsing and Text Manipulation from the Extracted Information

model, the code utilizes libraries like os, cv2 (OpenCV), easyocr, and csv. It specifies the input directory path containingthe images and defines the OCR language(s) to be used. An OCR reader object is created with the designated languages to facilitate text extraction from the grayscale images. The code then iterates through each file in the input directory, loading images with .jpg or .png extensions. The images are converted to grayscale for OCR processing. OCR is executed on the grayscale images, and the extracted text is stored along with the corresponding filename. The data is further formatted, including handling dates, empty values, and removing un- wanted terms. The formatted data is eventually saved to a CSV file, creating a structured output. This comprehensive process showcases an efficient approach to parsing and manipulating text data extracted from video frames, making it suitable for subsequent analysis and research purposes.

7. Results and Discussion

The results and discussion section of the research project provides a detailed analysis of how the Optical Character Recognition (OCR) system was tailored specifically for pharmaceutical inventory management using YOLOv7 and

Figure 6: Extracted frame from the video

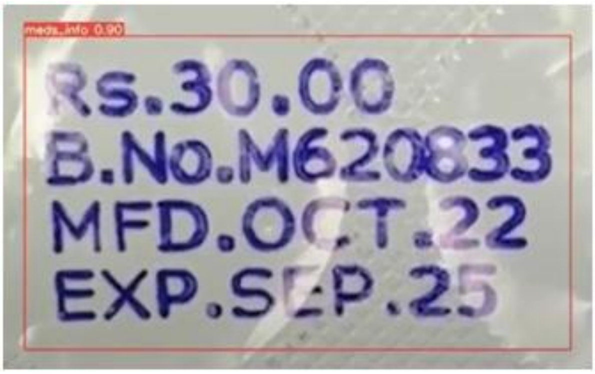

Figure 7: Extracted bounding box

EasyOCR. In the image phase, YOLOv7 was employed to accurately detect medicines within 510 images of medicine strips. The algorithm achieved impressive accuracy rates, with 84 percent accuracy in batch number detection, 86 percentin manufacturing date (MFD) detection, 75 percent in expiry date detection, and 95percent in maximum retail price (MRP) detection. The overall accuracy of 85 percent demonstrates the system's proficiency in identifying relevant information from the medicine strips.

The OCR-based text extraction process further proved successful in parsing data from the recognized text, enabling the identification of crucial attributes from the medicine strips. By extracting data such as batch numbers, MRPs, MFDs, and expiry dates, the system provided valuable insights for pharmaceutical inventory management, including product separation based on expiry dates and optimized resource utilization. Moving on to the video phase, the YOLOv7 algorithm efficiently identified bounding boxes around medicines in 14 video frames. Remarkably, ten frames had all four attributes (MRP, expiry, MFD, and batch number) accurately recognized, while the remaining frames had more than two attributes correctly identified. This real-time video OCR approach showcased the system's efficacy in handling dynamic data and extracting relevant text information, which is crucial for managing pharmaceutical inventory effectively.

The image OCR processing phase highlights a robust image processing pipeline that enhances text extraction accuracy using YOLOv7 and EasyOCR. YOLOv7 generates precise bounding boxes that mark specific areas containing relevant medicine details, while EasyOCR excels in extracting text from these refined regions. The extracted data is then struc- tured and consolidated into a CSV file, providing valuable in- sights for pharmaceutical inventory management and resourceoptimization.

Similarly, the video OCR processing phase demonstrates the YOLOv7 algorithm's efficiency in detecting bounding boxes around medicine strips in video frames, regardless of their orientation or position. The OCR technique accurately extractstextual information from these regions, which is then parsed and organized into a structured CSV file. This phase signif- icantly contributes to streamlined inventory management and data-driven decision-making in the pharmaceutical industry.

The implementation of EasyOCR in the project proved to be a game-changer, particularly in terms of versatility and language support. EasyOCR demonstrated remarkable adaptability, enabling the OCR system to recognize and extract text from pharmaceutical labels written in multiple languages. This capability significantly broadened the applicability of the system, making it suitable for handling pharmaceutical prod- ucts from diverse global markets. EasyOCR's ability to handle multilingual text made it a valuable tool for companies with international operations, streamlining inventory management processes across different regions and facilitating seamless cross-border distribution. In conclusion, the research project highlights the effective- ness of the YOLOv7 and EasyOCR systems in optimizing pharmaceutical inventory management through accurate text extraction and data analysis. The high accuracy achieved in both the image and video OCR processing phases showcases the system's potential for various real-life purposes, such as traffic monitoring, toll management, surveillance, and pharma- ceutical inventory management. The innovative and efficient OCR-based approach holds promise for enhancing operational efficiency and minimizing losses due to expired products in the pharmaceutical industry, making it a valuable tool for data- driven decision-making.

Table I: Sample Data Extracted from Video

Filename	MRP	Batch Number	Manufacturing Date	Expiry Date
frame12	Rs.53.00	FND0502245	Sep-22	Oct-24
frame77	155	S24221697	Oct-27	Mar-25
frame23	Rs.46.00	YMS21165	Jul-20	May-23
frame17	Rs.30.00	M620833	Oct-22	Sep-25
frame93	Rs.35.90	208559903	Mar-22	May-20
frame39	Rs407.48	ST-J35222	Jun-22	May-20
frame11	90.5	AT-210221	Feb-21	Jan-23

Table II: Sample data extracted from images

Image Filename	Price	Batch Number	MFD Date	EXP Date
crop453.jpg	86.08	4109201	Jun-13	Apr-10
crop153.jpg	103.85	DK3177	Sep-14	Oct-09
crop477jpg	64.68	DLC364	Jun-18	Jun-23
crop361.jpg	86.09	SKJ1425	Jun-12	Jun-12
crop102.jpg	202	K90614	Apr-04	Apr-04
crop399.jpg	168	LN21AM	Apr-12	May-10
crop431.jpg	82.5	FO5OOSA	Jun-13	Jul-13
crop95.jpg	152.2	C3791011	Aug-13	Sep-09
crop419.jpg	11.81	ADKO13OA	Feb-13	Mar-11
crop2.jpg	13	FO5OOSA	Jul-13	Jul-13
crop489.jpg	99.13	4390201	Mar-15	Apr-12
crop141.jpg	51.7	NP29404	Jun-15	Feb-10
crop23.jpg	97.67	FAHOJ94	Oct-12	Oct-45
crop722.jpg	144.5	DRO910	Mar-14	Mar-14
crop551.jpg	78.02	SkK2315	Nov-13	Dec-11

8. Future Scope and Discussion

The successful implementation of the OCR-based video analysis system for clinical text extraction from medicine strips paves the way for exciting future research and applications. Multi-language support is a key direction to explore, as medical data spans various languages and regions. By investigating language-specific models and datasets, the OCR system can accurately extract text from medicine strips indifferent languages, making it versatile and applicable in diverse healthcare environments

Integration with Electronic Health Records (EHR) holds tremendous potential for improving healthcare data management. Populating EHRs with structured medication information using the OCR system streamlines medical record digitization, enhances data accuracy, and facilitates efficient patient care. Real-time inventory management is another promising aspect of the OCR system. By continuously analyzing medicine strips in video streams, the system can promptly identify and track expiring medications, optimizing inventory turnover and reducing wastage.

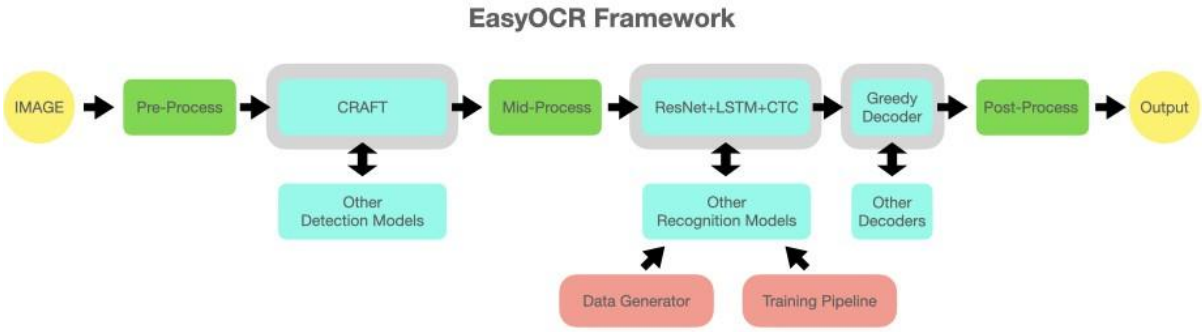

Figure 8: Easy OCR Framework

The OCR system's potential extends beyond text extraction. It can lay the foundation for visual analytics in healthcare by combining clinical text with extracted visual features like pill shapes and colors. This enables insights into medication adherence, dosage tracking, and treatment patterns. Moreover, ex- tending OCR to medical imaging, such as endoscopy images, can revolutionize medical procedures' analysis and documentation, enhancing medical reporting efficiency and accuracy. Continued research in machine learning and deep learningtechniques further enhances OCR accuracy and robustness. Exploring advanced models, such as attention mechanisms andtransformer-based architectures, can lead to context-aware andhighly efficient OCR systems.

9. Conclusion

In conclusion, this research project presents a comprehensive two-phase approach for Video Analytics using OCR, aimed at extracting clinical text from medicine images and videos with high accuracy and efficiency. The combination of the YOLOv7 model, advanced image processing techniques, and the EasyOCR library enables the precise identification and extraction of crucial details from diverse medical sources. Theproject's results demonstrate promising advancements in pharmaceutical inventory management and data-driven decision- making processes, contributing to improved patient safety and operational efficiency in the healthcare industry. The future scope of the project opens up new avenues for further advancements in OCR technology and NLP, driving innovationin healthcare applications.

References

[1] Rajat Kumar Prajapati, Tushar Nagar, Suraj Dangi, Yatharth Bhardwaj, Parvat Raj Singh Rao and Ritesh Ku- mar Jain (2023) "Automatic Number Plate Recognition using YoloV7 and PaddleOCR".
[2] M. A. Rani, M. A. M. Ali, and S. S. S. Ali (2021) "Automatic Number Plate Recognition (ANPR) System using YOLOv3 and PaddleOCR for Malaysian Number Plate"
[3] J. Zhao, X. Cao, and Y. Guo "License Plate Detection and Recognition Using YOLO and PaddleOCR".
[4] Chuanlei Zhang, Lei Shi, Dandan Zhang, Ting Ke and Jianrong Li (2023) "Pointer Meter Recognition Method Basedon Yolov7 and Hough Transform".
[5] Karez Hamad, Mehmet Kaya (2016) "A Detailed Anal- ysis of Optical Character Recognition Technology".
[6] Chirag Indravadanbhai Patel, Dr. Atul Patel, Dharmendra Patel (2012) "Optical Character Recognition by Open source OCR Tool Tesseract: A Case Study".
[7] Nikita Kotwal, Gauri Unnithan, Ashlesh Sheth, Nehal Kadaganchi "Optical Character Recognition using Tesseract Engine"
[8] S. D. Doychinov and S. A. Yalamov (2021) "Auto- matic Number Plate Recognition (ANPR) using YOLOv3 and OpenCV"
[9] S. H. Lee and S. S. Kim (2020) "License Plate Recog- nition System using Deep Learning Techniques" .
[10] H. R. Moon, M. S. Baig, and M. S. Kim (2021) "An Efficient and Robust Automatic Number Plate Recognition System using YOLOv3 and PaddleOCR"
[11] Abiodun, O.I.; Jantan, A.; Omolara, A.E.; Dada, K.V.; Umar, A.M.; Linus, O.U.; Arshad, H.; Kazaure, A.A.; Gana, U.; Kiru, M.U. A comprehensive review of artificial neural network applications to pattern recognition. IEEE Access 2019, 7, 158820–158846. [CrossRef]
[12] Li, Z.; Zhou, Y.; Sheng, Q.; Chen, K.; Huang, J. A High-Robust Automatic Reading Algorithm of Pointer Meters Based on Text Detection. Sensors 2020, 20, 5946. [CrossRef]
[13] Smith R, Antonova D, Lee DS. Adapting the Tesseract open source OCR engine for multilingual OCR. In Proceedings of the International Workshop on Multilingual OCR 2009 Jul 25 (p. 1). ACM
[14] Basu S, Chaudhuri C, Kundu M, Nasipuri M, Basu DK. Text line extraction from multi-skewed handwritten doc- uments. Pattern Recognition. 2007 Jun 30;40(6):1825-39
[15] "Coco-text: Dataset and benchmark for text identification and recognition in natural images," A. Veit, T. Mat- era, L. Neumann, J. Matas, and S. Belongie, arXiv preprint arXiv:1601.07140, 2016.