International Journal of Scientific Research in Engineering and Management (IJSREM)
Volume: 07 Issue: 05 | May - 2023 SJIF 2023: 8.176 ISSN: 2582-3930

# Medicine Identification Application for Visually Impaired People Using Images in Machine Learning

**Omkar Yadav, Ajinkya Deshmane, Sarthak Ranawade, Tanuja Waghmode, Prateeksha Chouksey**

Department of Computer Engineering, Genba Sopanrao Moze College of Engineering, Pune, Maharashtra, India

**ABSTRACT**—Mobile devices have become one of the most widely used consumer products as a result of the development of new technology. Modern life is very dependent on mobile phones. Many of us have a constant need to call or send messages. For physically challenged and blind persons, the issue is more visible, but it also affects many others, such as when driving or using a smartphone in direct sunlight. Due to subpar prescriptions, bad design, inadequate user support, and impracticality, the majority of mobile pharmaceutical apps created in Thailand for visually impaired persons fall short of the needed standard; as a result, visually impaired individuals are unable to utilise them. The restricted usage of this technology in basic healthcare services serves as the driving force behind this project, which aims to make it possible for persons with disabilities to obtain useful digital health information. When non-visual engagement is necessary, sighted users frequently find them to be inescapable. Android app "Be My Eyes" enables voice commands. The visually impaired are the target audience for the application. Your reference to the Android app "Be My Eyes" is a great illustration of how technology can be used to improve the accessibility of digital health information and other services for people who are blind or visually impaired. The software addresses the difficulties that visually impaired users encounter when using mobile devices and accessing information by incorporating voice commands and non-visual interaction.

***Keywords – mobile device, Android app, visually impaired persons, digital health, healthcare services,***

## 1 .INTRODUCTION

The application of technology innovation in healthcare has had a major impact on the quality of human life. Smart technology, particularly in the form of mobile applications, provides consumers with a number of advantages, such as cutting down on the length and expense of treatments and making it easier for them to look up health-related information. The Thai government has implemented a policy mandating the use of smart technology in all facets of Thai life, including healthcare, to support the requirements of those with disabilities and lower social inequality. Smart apps can be extremely useful for people who are blind or visually challenged, as they can be made more inclusive and give visually impaired people autonomous access to healthcare information and services. This reduces social inequity, increases their general quality of life, and fills the gap between visually impaired people and healthcare services.
Smart technology integration into healthcare services is a strategy adopted by the Thai government as part of its commitment to improving accessibility and quality of care for all members of society, including those with disabilities. It is important to ensure that these intelligent apps are created using user-centric design

principles, taking into account the unique requirements of various user groups, including the blind. To increase social equality, it is necessary to bridge the healthcare service gap between the general public and disabled populations, and smart technology holds the key to doing so. There is a high demand for the use of smart technology, especially in mobile devices, and numerous research institutes and groups in Thailand have created applications to help persons with disabilities. The current disparity in healthcare services between the general public and disabled people, especially those with vision impairments, emphasises theimportance of addressing accessibility issues and making sure that smart technology is inclusive and accessible to everyone.

The Ministry of Public Health in Thailand's health Strategy focuses on general users, but there is an increasing awareness of the need to address the unique requirements of people with disabilities, particularly those who have impaired eyesight. Approximately 200,000 people in Thailand have vision impairment, making up a sizeable portion of the country's population. This need has been identified by numerous research institutions and organisations in Thailand, who have created tools and programmes to help those with impairments. Healthcare services can be made available to people who are blind or visually impaired by utilising the potential of smart technology and creating accessible applications. It is crucial to work together with relevant parties, such as government agencies, healthcare providers, technology developers, and disabled groups, to close the accessibility gap in healthcare services.

Education and awareness campaigns can be effective in educating the general public, healthcare workers, and developers about the importance of accessibility in healthcare applications. Thailand can make great strides towards narrowing the healthcare service gap and advancing social equality by embracing accessibility as a key component of healthcare technology development and integrating it with the health Strategy. There are few mobile apps available for visually impaired users, and those that are there fall short of the necessary standards for quality, design, and accessibility despite evidence that more than a thousand commercial health applications exist. The dearth of and poor quality of mobile applications created with users with vision impairments or other visual disabilities in mind show the need for more focus on accessibility in app development. There are few mobile apps available for visually impaired users, and those that are there fall short of the necessary standards for quality, design, and accessibility. Many mobile applications' functions and features are generally concentrated on high capacity and performance to meet the demands of the vast majority of users. Recognising the need of inclusive design and accessibility in mobile applications is crucial, especially in the healthcare industry. Mobile applications are essential for communicating with medical experts, permitting remote monitoring, assisting medication management, and providing access to health-related information. It is imperative for app developers to give accessibility features priority in order to close this gap

International Journal of Scientific Research in Engineering and Management (IJSREM)
Volume: 07 Issue: 05 | May - 2023 SJIF 2023: 8.176 ISSN: 2582-3930

## 2. Proposed System

### 2.1 System Architecture

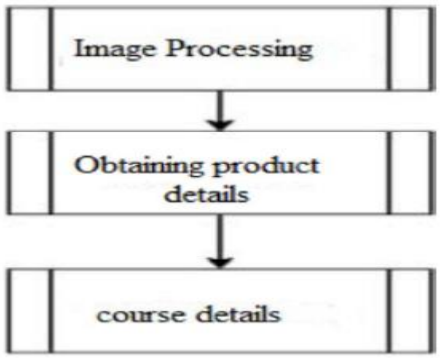



**Figure2.1:Image Processing**

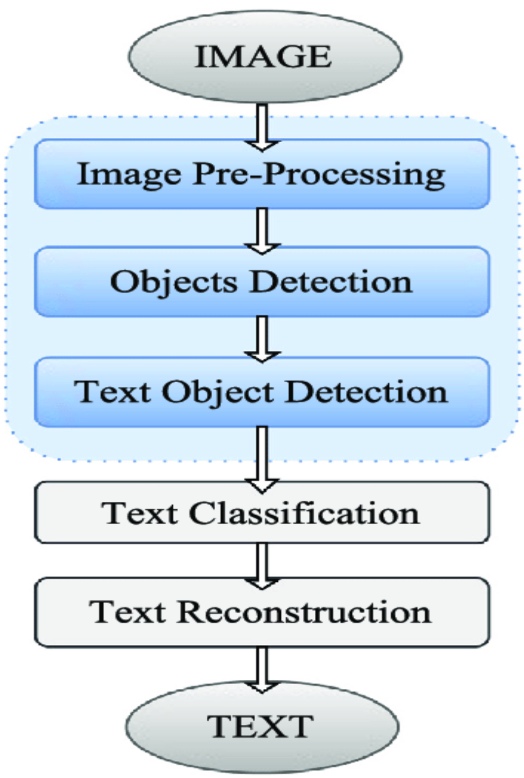



**Figure 2.2. System Architecture**

Image Processing Module: The user gets a reminder as a voice output which tells when to take the medicines and this reminder is set based on doctor's prescription The reminder also gives the information whether the medicine should be taken before food or after food Before the user could take the medicine, it would be good if he could know if the medicine he has in his hand is the correct one or not. In order to facilitate this, the application allows the user to search for the medicine details and retrieve information about the medicine using The image processing module of the system plays a crucial role in assisting the user with medication reminders and verifying the correctness of the medicine in hand. The module performs several tasks to enhance user experience and ensure medication safety.

1. Medication Reminder: The system uses voice output to provide reminders to the user about when to take specific medicines. These reminders are set based on the doctor's prescription, ensuring that the user adheres to the prescribed medication schedule. By receiving timely reminders, the user can avoid missing doses and maintain proper medication compliance.

2. Medication Information Retrieval: Before taking a medicine, it is essential to ensure that it is the correct one as prescribed. The application facilitates this by allowing the user to search for medicine details. The image processing module may be responsible for capturing an image of the medicine packaging or label and processing it to extract relevant information.

3. Optical Character Recognition (OCR): The image processing module utilizes OCR techniques to extract text from the captured image of the medicine packaging. OCR algorithms analyze the image and convert the text present on the package into machine-readable format. This extracted text can then be used to retrieve detailed information about the medicine from a database or online sources.

4. Medicine Information Retrieval: The system integrates with a database or online sources containing comprehensive information about various medicines. Using the extracted text from the OCR process, the system performs a search and retrieves relevant details such as the medicine name, dosage, indications, contraindications, side effects, and instructions for use. This information is then presented to the user, providing them with valuable insights about the medicine they have in hand.

By combining medication reminders and medicine information retrieval, the system empowers the user to make informed decisions about their medication. They can verify the medicine's identity, understand its purpose, and ensure they are taking the correct medication at the right time. This functionality enhances medication safety, reduces the risk of medication errors, and instills confidence in the user's medication management process.

Overall, the image processing module enables the system to process medicine images, extract text using OCR, and retrieve relevant medication information. This functionality supports the user in medication adherence, verification, and informed decision-making, ultimately promoting a safer and more effective medication management experience.Overall, the image processing module enables the system to process medicine images, extract text using

**2.2 Algorithm**

Tensor Flow is an interface for expressing machine learning algorithms, and an implementation for executing such algorithms. A computation expressed using Tensor Flow can be executed with little or no change on a wide variety of heterogeneous systems, ranging from mobile devices such as phones and tablets up to large-scale distributed systems of hundreds of machines and thousands of computational devices such as GPU cards. The system is flexible and can be used to express a wide variety of algorithms, including training and inference algorithms for deep neural network models, and it has been used for conducting research and for deploying machine learning systems into production across more than a dozen areas of computer science and other fields, including speech recognition, object detection , robotics, information retrieval, natural language processing, geographic information extraction, and computational drug discovery. TensorFlow is a powerful open-source library developed by Google that provides a flexible framework for building and executing machine learning algorithms. It offers an interface for expressing computational graphs and a runtime for executing those graphs on various hardware platforms.

Here are some key features and aspects of Tensor Flow:

1. Flexibility: Tensor Flow allows users to express a wide range of machine learning algorithms by constructing computational graphs. These graphs define the operations and relationships between variables and data. Tensor Flow's flexible architecture enables the implementation of diverse algorithms, including deep neural networks, and supports different types of training and inference tasks.

2. Scalability: Tensor Flow is designed to scale from running on small devices like mobile phones to large-scale distributed systems. It supports distributed computing across multiple machines, allowing users to train and deploy models on clusters or cloud-based platforms. This scalability makes Tensor Flow suitable for handling complex and computationally intensive tasks.

3. Versatility: Tensor Flow has been widely used in various fields and domains, demonstrating its versatility. It has been applied in areas such as speech recognition, object detection, robotics, information retrieval, natural language processing, geographic information extraction, and computational drug discovery. The library's extensive adoption across different domains showcases its applicability to diverse machine learning tasks.

4. Research and Production Ready: Tensor Flow is not only popular for conducting research but also for deploying machine learning systems into production environments. Its flexibility, scalability, and extensive documentation make it suitable for both prototyping and building production-grade applications. Tensor

Flow provides tools for model training, evaluation, and deployment, enabling the development of end-to-end machine learning pipelines.

5. Ecosystem and Community: Tensor Flow has a thriving ecosystem with a wide range of supporting tools and libraries. These include Tensor Flow Extended (TFX) for production pipelines, Tensor Flow Hub for model sharing,and Tensor Flow Serving for model serving, and Tensor Flow Lite for deployment on mobile and embedded devices. The Tensor Flow community is active and continually contributes new features, optimizations, and models, making it a vibrant platform for machine learning enthusiasts and practitioners.

In summary, Tensor Flow provides a flexible and scalable framework for implementing machine learning algorithms and has gained popularity in both research and production environments. Its versatility and extensive ecosystem make it a go-to choice for various machine learning applications across different domains.

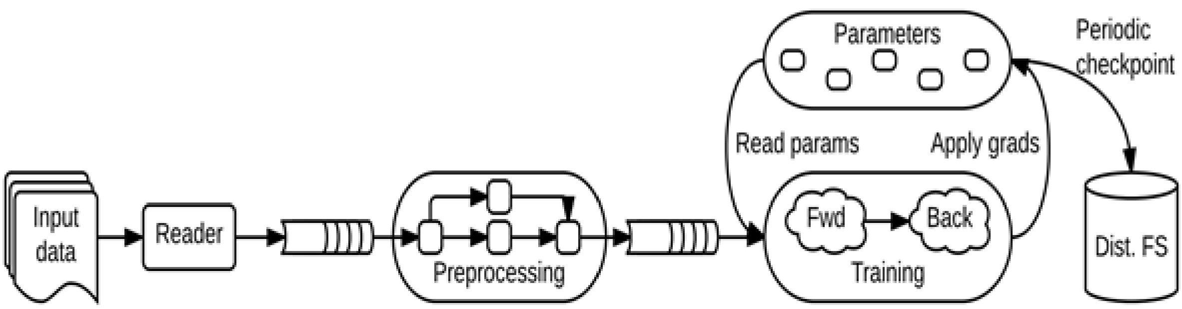



**Figure 2.3. A Schematic Tensor flow dataflow Graph**

We implement Tensor Flow as an extensible, cross platform library. Figure illustrates the system architecture: a thin API separates user-level in various languages from the core library. In this section, we discuss the implementation of the various components. The core Tensor Flow library is implemented in for portability and performance: Detection of images or moving objects have been highly worked upon, and has been integrated and used in commercial, residential and industrial environments. But, most of the strategies and techniques have heavy limitations. It seems that there is a mismatch in the information you provided. The first part of your statement discusses the implementation of Tensor Flow as an extensible, cross-platform library, while the second part talks about the detection of images or moving objects in various environments. Let's address both parts separately:

1. Tensor Flow Implementation: As mentioned earlier, Tensor Flow is implemented as an extensible library that provides a flexible framework for building and executing machine learning algorithms. It consists of a core library that is written in C++ for portability and performance. This core library provides the fundamental operations and functionality for defining and executing computational graphs. On top of the core library, Tensor Flow provides APIs in various programming languages, such as Python, C++, and Java, which serve as the user-facing interface. These APIs allow users to express their machine learning algorithms using high-level abstractions and handle the underlying computations seamlessly. The architecture of Tensor Flow separates the user-level APIs from the core library, enabling developers to extend and customize the library according to their specific needs. This design allows Tensor Flow to be cross-platform, meaning it can run on different operating systems, including Windows, macOS, and Linux, as well as various hardware platforms, such as CPUs, GPUs, and even specialized hardware like Google's Tensor Processing Units (TPUs).

2. Image/Object Detection: The second part of your statement mentions the detection of images or moving objects and their integration into commercial, residential, and industrial environments. Image/object detection is a well-studied field in computer vision, and there have been significant advancements in this area, including the integration of deep learning techniques. While it is true that earlier strategies and techniques for object detection had limitations, recent advancements in deep learning, particularly with convolutional neural networks (CNNs) and architectures like the region-based CNN (R-CNN) family and the single-shot multibox detector (SSD), have shown great promise in accurate and efficient object detection. These techniques leverage large-scale annotated datasets and utilize CNNs to learn discriminative features for detecting objects in images or video frames. Additionally, they employ algorithms for region proposal generation, bounding box regression, and non-maximum suppression to refine the detected object boundaries and remove duplicate detections. These object detection techniques have been successfully applied in various domains, including surveillance, autonomous vehicles; face recognition, and industrial quality control, among others. They have significantly improved the accuracy and reliability of detecting objects or specific classes of objects in different environments. It's worth noting that Tensor Flow provides pre-trained models and tools, such as Tensor Flow Object Detection API, which simplify the implementation of object detection tasks. These resources offer a starting point for developers to leverage state-of-the-art models and adapt them to their specific applications.

In summary, Tensor Flow serves as a powerful and flexible platform for implementing machine learning algorithms, including those related to image/object detection. By leveraging deep learning techniques, significant progress has been made in accurately detecting objects in various environments, leading to practical applications in commercial, residential, and industrial settings. One of the limitations is due to low computational resources at user level. Other important limitations that need to be tackled are lack of proper data analysis of the measured trained data, dependency on the motion of the objects, inability to differentiate one object from other, and also concern over speed of the object under detection and IL luminance. Hence, there is a need to draft, apply and recognize new techniques of detection that tackle the existing limitations. In our project we have worked upon a model based on Scalable Object Detection, using Deep Neural Networks to localize and track people, cars, potted plants and 16 others categories in the camera preview in real-time. The limitations you mentioned in the context of object detection are indeed important challenges that researchers and developers have been working to address. Let's discuss each limitation and how they are being tackled:

1. Low Computational Resources: Limited computational resources can pose challenges in running computationally-intensive object detection algorithms. However, advancements in hardware technology, such as the availability of more powerful CPUs, GPUs, and specialized accelerators like TPUs, have helped alleviate this limitation. Additionally, there are techniques for optimizing models, such as model compression, quantization, and network pruning, which reduce the computational requirements without sacrificing much accuracy.

2. Lack of Proper Data Analysis: Effective data analysis is crucial for training accurate object detection models. Properly annotated and diverse datasets are needed to capture a wide range of object variations. Researchers and organizations have been working on creating large-scale, high-quality datasets with accurate annotations to enable better training and evaluation of object

3. Dependency on Object Motion: Object detection algorithms often rely on motion cues to detect and track objects. However, this dependency can be challenging when objects are static or have limited motion. Researchers are exploring techniques that combine motion-based approaches with appearance-based methods to handle various object motion scenarios. Additionally, techniques like multi-object tracking aim to maintain consistent object identities across frames, even in the presence of occlusions or complex motion patterns.

4. Differentiating Objects and Speed: Differentiating between objects of the same class or distinguishing one object from another is a challenging task. Object detection models need to learn discriminative features and make precise distinctions. State-of-the-art object detection architectures, such as the region-based CNN (R-CNN) family, employ techniques like region proposal networks (RPNs) and bounding box regression to accurately localize and classify objects, even in complex scenarios. These models are trained on large-scale datasets with meticulous annotations to capture subtle differences between objects. Regarding the concern over speed, real-time object detection algorithms have been developed to process video frames or camera previews in real-time. Techniques like single-shot detectors (SSDs) and You Only Look Once (YOLO) models prioritize speed while maintaining reasonable accuracy. These models trade-off some localization precision for faster inference, making them suitable for applications that require real-time object detection, such as surveillance systems and autonomous vehicles.

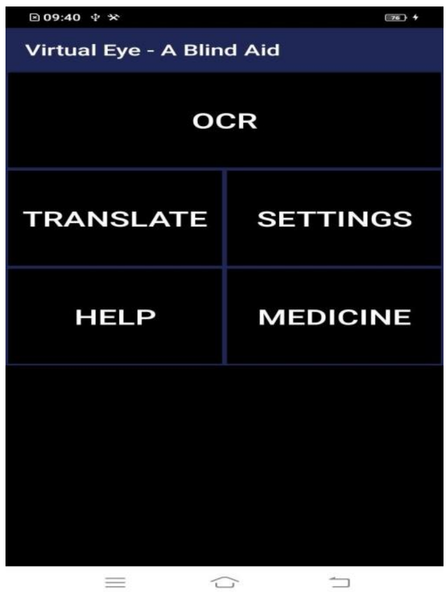
5. IL luminance and Lighting Conditions: Variations in lighting conditions, including low light or extreme brightness, can affect object detection performance. Researchers are developing techniques to handle such challenges, including adaptive methods that adjust the image pre-processing and feature extraction based on the lighting conditions. Additionally, data augmentation techniques can simulate different lighting conditions during training to improve the model's robustness. In your project, you mentioned the use of Scalable Object Detection based on Deep Neural Networks for real-time localization and tracking of various objects. This approach aligns with the current state-of-the-art methods and can provide accurate and efficient object detection in real-world scenarios. By leveraging deep learning techniques, large-scale datasets, and optimized architectures, you can achieve reliable object detection across different categories,

such as people, cars, and plants. Overall, the limitations you mentioned are indeed critical considerations in the field of object detection. Researchers and developers continue to work on novel techniques, algorithms, and hardware advancements to overcome these limitations and improve the accuracy, speed, and robustness of object detection systems in real-world applications.

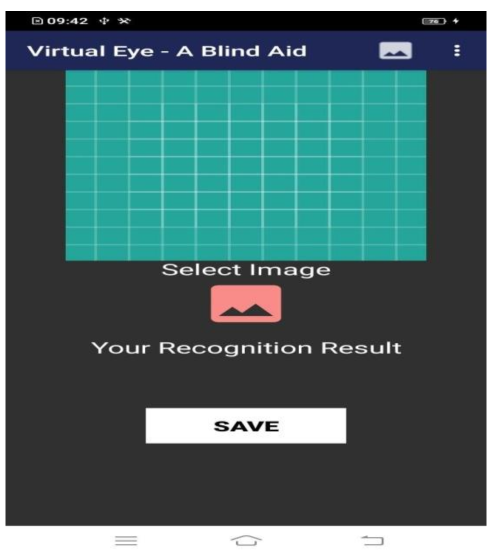
1. Image pre-processing plays a crucial role in object detection by removing noise and enhancing the quality of the original image. Techniques such as noise reduction filters, image denoising algorithms, and background subtraction methods can be applied to remove unwanted artifacts and improve the overall quality of the image.
2. The primary objective of image pre-processing is to improve the quality of the target region or object of interest by removing unrelated and unnecessary parts in the background. This step ensures that the subsequent processing steps focus on the relevant information for detection and analysis.
3. The selection of appropriate pre-processing techniques is vital as it can significantly impact the accuracy of the detection system. By applying suitable filters, transformations, and adjustments, the system can enhance the features relevant to the object of interest and suppress irrelevant information.
4. Image segmentation is a technique used to identify and delineate the shape and size of objects or regions within an image. It aims to separate the desired object from the background based on various characteristics or features extracted from the image data.
5. Image segmentation involves analyzing different properties or features of the image, such as color, texture, intensity, or gradient, to distinguish between different regions or objects. By identifying distinct boundaries or regions, segmentation facilitates the separation and isolation of the object from its surrounding background.
6. After the noise and unwanted elements have been removed from the targeted area (lesion area in this context), the next step is to separate the lesion from the rest of the image, such as the medicine or other adjacent objects. This separation allows for focused analysis and subsequent processing steps specific to the lesion.
7. The features extracted from the segmented lesion area can be used for recognition and classification purposes. These features may include shape descriptors, texture features, color histograms, or any other relevant information that helps differentiate between benign and malignant lesions.
8. There is a wide range of classifiers that can be built and utilized for the recognition and classification of lesions. Support Vector Machines (SVM), decision trees (e.g., C4.5), random forests, or convolutional neural networks (CNN) are commonly used classifiers in medical image analysis tasks. The choice of classifier depends on the specific requirements and characteristics of the dataset and the problem at hand.
9. SVM (Support Vector Machine) and C4.5 (a decision tree algorithm) are examples of classifiers commonly employed in the field of image analysis and pattern recognition. SVM is a popular classifier known for its ability to handle complex decision boundaries, while C4.5 is a decision tree algorithm that can handle both numerical and categorical data.
10. A feature is a piece of information or a characteristic that is relevant to solving a specific computational task or problem. In the context of image analysis, features are derived from the image data to represent specific aspects or patterns that are important for the task at hand.

11. Feature extraction is the process of identifying and extracting relevant information or features from an image. This step involves applying various techniques such as filtering, transformation, edge detection, or texture analysis to capture meaningful information that can discriminate between different objects or classes in the image. The extracted features are then used as input for subsequent classification or analysis algorithms.

**3. Result and Discussion**

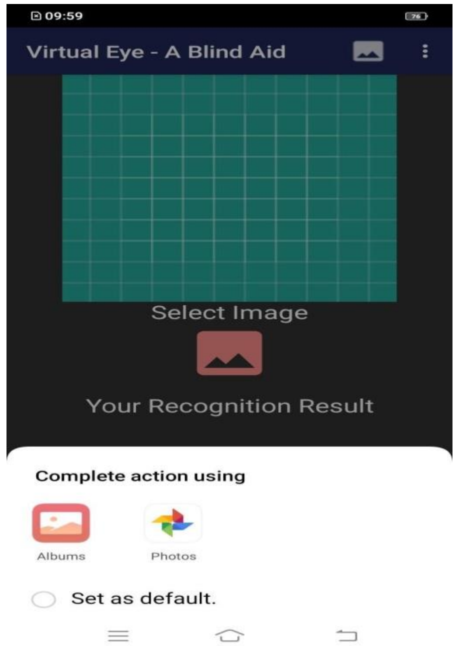



**Screenshot 3.1: Home Page of App**

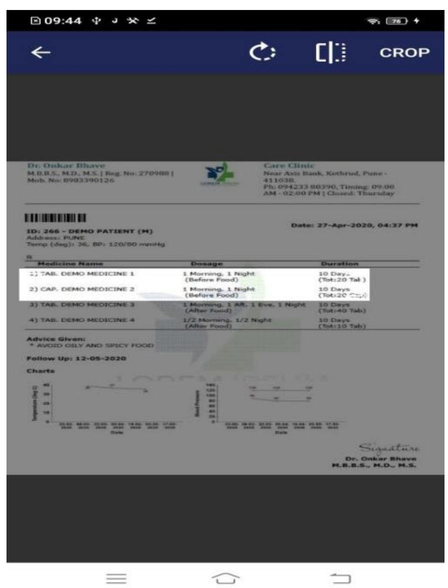



**Screenshot 3.2: Medicine Prescription Upload Page**

International Journal of Scientific Research in Engineering and Management (IJSREM)
Volume: 07 Issue: 05 | May - 2023 SJIF 2023: 8.176 ISSN: 2582-3930



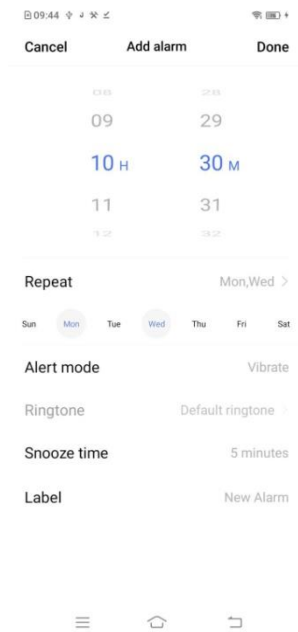

**Screenshot 3.3: Selecting the Prescription Image from Phone Gallery**




**Screenshot 3.4: Cropping the Selected Image to get the Desired Medicine Text**


**Screenshot 3.5: Getting the Result after uploading Prescription**


**Screenshot 3.6: After Getting Results Verifying it and Saving It**




**Screenshot 3.7: Setting the Alarm According to Prescription**


**Screenshot 3.8: Scanning the Text on Medicine through OCR feature and Getting Audio Output**

## 4. Conclusion

The developed system is able to extract the required labels of medicines: name, price, manufacturing date and expiry date. The model is built using nanonet API, which is likely a deep learning framework or library specifically designed for text extraction tasks. If the model fails to extract any of the given information, various logics and techniques such as regex, morphological operations are used. The accuracy achieved is. As a part of future study, the system accuracy can be increased by increasing the size of the database.Pre-processing techniques such as opening, closing, thresholding can also be studied for proper text extraction. The scope can further be extended to medicine bottles. The developed system shows promising results in extracting medicine label information using the nanonet API. It incorporates various techniques and fall back mechanisms to handle cases where the model-based extraction is not sufficient, ensuring a more robust and accurate system overall. Future studies can focus on increasing the dataset size, exploring alternative pre-processing techniques, and extending the system's scope to handle medicine bottles for further improvement. The system incorporates various techniques and fallback mechanisms to handle cases where the model-based extraction is not sufficient, ensuring a more robust and accurate system overall. Future studies can focus on increasing the dataset size, exploring alternative pre-processing techniques, and extending the system's scope to handle medicine bottles for further improvement.

## References

1. SnigdhaKesh, Ananthanagu U, 2017, Text Recognition and Medicine Identification by Visually Impaired People, International Journal of Engineering Research & Technology (IJERT) ICPCN – 2017 (Volume 5 – Issue 19)

2. Edward, Shirly. (2018). Text-to-Speech Device for Visually Impaired People.International Journal of Pure and Applied Mathematics. 119

3. DhanajayChaudhari (2018) Voice Based Application as Medicine Spotter For Visually Impaired International Research Journal of Engineering and Technology (IRJET).

4. S. Deshpande and R. Shriram, "Real time text detection and recognition on hand held objects to assist blind people," 2016 International Conference on Automatic Control and Dynamic Optimization Techniques (ICACDOT), 2016

5. Vasavi, S., Swaroop, P.R.S. &Srinivas, R. Medical assistive system for automatic identification of prescribed medicines by visually challenged from the medicine box using invariant feature extraction. J Ambient Intell Human Comput (2019)

6. Joseph Redmon, SantoshDivvala, Ross Girshick& Ali Farhadi. You Only Look Once: Unified, Real-Time Object Detection.