MobileNet architecture -Based Pill Detection System for Accurate Drug Identification in Healthcare

[1]Nonika Nissy.Kariketi, [2]Mr. Suresh Tiruvalluru
[1]*M-Tech, Dept. of CSE Gokula Krishna College of Engineering, Sullurpet*
[2]*Associate Professor, M-Tech., (Ph.D), CSE Gokula Krishna College of Engineering, Sullurpet*

Abstract: **Accurate pill identification is vital for patient safety, especially in environments like hospitals, pharmacies, and eldercare, where medication errors can be life-threatening. Manual verification methods are prone to error, particularly with visually similar pills. This paper presents a lightweight, efficient, and accurate pill detection and classification system using the MobileNet architecture, optimized for deployment on mobile and edge devices. Leveraging transfer learning and a curated pill image dataset, the system is capable of identifying pills based on shape, color, imprint, and size. Experimental results show that the MobileNet-based model achieves high classification accuracy with minimal latency, making it ideal for real-time, on-device healthcare applications. The system improves drug safety and supports healthcare professionals in minimizing medication errors.**

Keywords: **Medication Identification, Pill Detection, MobileNet Architecture, Drug Safety, Healthcare Automation.**

1. INTRODUCTION

Medication errors remain a critical issue in the healthcare sector, leading to severe health consequences and even fatalities. Many pills share similar appearances, making visual identification challenging. Traditional methods like manual cross-checking with databases or relying on barcode scanning are often inefficient or unavailable in emergency or rural settings.

With advancements in deep learning, computer vision systems can now assist in drug identification by analyzing pill images. MobileNet, a lightweight convolutional neural network designed for mobile and embedded vision applications, offers an efficient and scalable solution. This paper explores the use of MobileNet for pill classification, focusing on accuracy, speed, and real-world deployment potential. The proposed system demonstrates that deep learning models can significantly enhance the accuracy and usability of drug recognition tools, ultimately reducing the risk of medication errors.

Over time, deficiencies in medical care have supplanted illnesses as the leading cause of mortality, with an estimated 400,000 or more deaths annually. Medication errors are the main type of curable medical error, according to reports of the medical error epidemic presented by the data from EHRs and medical institutions. Additionally, lowering pharmaceutical errors—which result in significant financial losses—is covered in the Institute of Medicine's 2006 study. From prescribing to tracking the patient's response, implementation is possible [1]. From prescribing to tracking the patient's response, implementation is possible. It's possible that the general public is unaware of the consequences of medical errors, such as the fact that unknown pills (in terms of their name or shape) might cause patients to misuse them and end up with undesired medicine and medical poisoning. Data extraction from clinical text can be useful for a variety of applications, including data mining, research subject identification, automatic vocabulary management, elimination of identification of clinical text, study of the illness medication and its adverse effects, and more.

Due to dictated transcriptions, the majority of biomedical data is usually available in an unstructured format. Without the pill's outer cover, it is quite impossible to distinguish or even identify the chemical composition and medical name of the specific medication, unless one is educated to do so. Most tablets don't have any visible marks that would identify their name or make up. Elderly individuals, kids, and anybody else unfamiliar with the pill find it nearly difficult to recognise, which leads to their taking the incorrect prescription, receiving the incorrect drug at the wrong time, or receiving the incorrect medication altogether. When medical poisoning or physical side effects result from this negligence, a patient may need to be admitted to the hospital and get extensive treatment before they pass away.

This research employs deep learning techniques, such as MobileNet and TensorFlow, to develop a model that recognizes pills. The model receives image descriptions of the pills through image sensing, which trains it to accurately identify each pill. This enables the creation of an application where patients can capture images of their medication using a camera. The application uses visual recognition to compare captured images with a database of numerous pills, thereby providing the patient with the pill's name, substance, and dosage information.

2. LITERATURE SURVEY

The Random Forest Classifier prioritises accuracy and performs best when used correctly. Random selection in RFC decision trees can capture complicated feature patterns for improved accuracy. The Random Forest Classifier is ideal for datasets with high outliers and categorical data.[2] used an RF Classifier to extract complex feature patterns from existing datasets, including the National Library of Medicine's Pillbox and Pill Image Recognition Dataset (PIRD). RFC uses an ensemble of decision trees to identify numerous feature combos by training each tree on a random selection of features.

SVM is a popular machine learning technique for classification. The authors use SVM in conjunction with a linear, poly and radial basis functions (RBF) kernels to address the classification problem. The linear SVM classifier produced the highest results with an accuracy of 99.57% in [3]. In the oval and round classes, the SVM/LIN classifier produces the best average precision results. SVM/LIN and SVM/POL classifiers yield the greatest results for the oval and capsule classes.

Artificial neural networks of the MLP (multilayer perceptron) type are often used in machine learning tasks including pattern recognition, regression, and classification. During training, the MLP uses the backpropagation method in two phases. The sample is sent to the input layer in the first step. The output layer is reached once the outcome of this input passes through the network layers. The outcome is compared to what was anticipated for that sample in the second phase. Although its performance varied over the course of the 30 trials, the MLP classifier's accuracy of 99.34% and standard deviation of around 0.49% in [4] indicate that MLP successfully spotted 99.34% of the pill pictures.

A deep learning system called a Convolutional Neural Network (CNN) is made to receive and interpret visual input, including images and videos. Convolutional layers are used by CNNs to identify local patterns in incoming data. These layers use learnable filters or kernels to conduct convolution operations on the input picture. As the filters go through the input, they calculate dot products to create feature maps that depict various aspects of the picture, including shapes, edges, and textures. ConvNet/CNN is a deep learning technique described in [5] that can recognise an input image, weight distinct features based on their importance, and then differentiate between them with an accuracy of greater than 90%.

In a new hospital, Craswell et al. installed dispersed automated drug distribution devices. The system uses a variety of image processing techniques. With 568 of the most popular pill kinds in the US, the proposed technique has shown approximately 91% accuracy in automatically identifying the correct prescription. In addition to this, a voice-based software system was created for the blind. This device sets an alarm and then uses a voice output to tell the user when to take the tablets.

The main topic of this study is a suggested system that identifies new tablets using an effective database extension approach and a deep learning algorithm that may improve the accuracy of detection based on sparse training data. The suggested algorithm aims to distinguish certain pills from a range of medications. To minimise the quantity of data required for learning, a single tablet is photographed instead of many pills while producing the training data. For pill detection, a two-step methodology is suggested that can be used for both multiple classes pill detection. Post-processing for detection and single pill extraction are applied to further boost the detection ratio.

3. IMPLEMENTATION

The suggested method for pill identification employs a MobileNet architecture to utilise its robust image processing capabilities. First, the MobileNet framework preprocesses input images, focusing on regions of interest for potential pill areas. Convolutional layers subsequently examine these areas, categorising them according to pill attributes (e.g., colour, shape) for accurate identification. This approach facilitates improved detection with the application of MobileNet framework feature extraction, minimising false positives and augmenting overall accuracy. Figure 1 depicts the suggested structure employed in this methodology, delineating each detection phase.

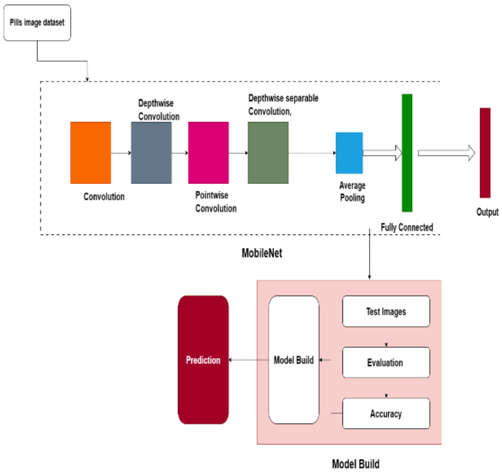


Figure 1. Proposed Architecture

3.1 Dataset collection

Jaya Prakash Pondy's (2023) pill dataset includes images of a variety of pills, along with metadata descriptions such as the colour, shape, imprint, and score of each pill. Healthcare applications can use this dataset to accomplish tasks such as visual recognition, picture categorisation, and pill identification. This dataset can train machine learning or deep learning models to recognise drugs based on their visual characteristics. The sample of the dataset in represented in Figure 2.

Figure 2. Dataset sample

3.2 MobileNet Architecture

Convolution: Convolution is a mathematical procedure crucial for deep learning models in feature extraction from pictures. A tiny filter, or kernel, traverses the input picture, catching local patterns such as edges, corners, and textures. Each filter is activated upon detecting a certain pattern, generating a feature map that emphasises this pattern throughout the picture. This method allows convolutional neural networks (CNNs) to acquire hierarchical features, ranging from basic edges in initial layers to intricate structures and objects in subsequent layers.

Depthwise convolution: Instead of mixing all of the input channels, depthwise convolution applies a distinct filter on each one separately, optimising classic convolution. Because this method avoids channel merging during the first convolution, it uses fewer parameters and requires less computing power. For example, a depthwise convolution would capture channel-specific information in a three-channel RGB picture by individually applying a different filter to each colour channel. In lightweight models like MobileNet, where speed and efficiency are top concerns, this kind of convolution is essential.

Pointwise convolution: Following the use of depthwise convolution, pointwise convolution combines and recombines information across channels using a 1x1 kernel. Pointwise convolution blends information across channels by executing a linear change on each pixel, in contrast to depthwise convolution, which mixes the separate feature maps. Since this step only modifies the depth and not the spatial resolution, it produces richer, more abstract features and permits more expressiveness in the model without appreciably raising the computing burden.

Depthwise separable convolution: CNNs' computational cost is greatly decreased by depthwise separable convolution, which combines depthwise and pointwise convolutions. In order to capture channel-specific information, depthwise convolution first processes each channel independently. The final result is then obtained by recombining these channel-specific feature maps using pointwise convolution. In order to produce effective and high-performing models, depthwise separable convolutions are frequently utilised in architectures such as MobileNet.

Average pooling: By calculating the average of values inside a specified window (such as 2x2 or 3x3) throughout the feature map, average pooling is a downsampling approach that shrinks the spatial dimensions of feature maps. This approach produces a smoothed representation that is less susceptible to little changes or distortions in the picture, making it especially useful for summarising the existence of features over a region. By making it easier for the model to comprehend the input, average pooling helps minimise overfitting. It also reduces computing needs by making feature maps smaller.

Output/ Prediction: This step gives the final judgement and is crucial for jobs where the model must correctly identify and categorise items, such as object detection.

4. RESULTS

Both the physical structure and chemical composition of a pill have been taken into consideration in order to get the desired findings for the full research of pill detections. These results have been successful in meeting the requirements of the task. Each image was used as the training set for evaluation, and the experiment was carried out 10 times to ensure comprehensive coverage. Following a total of ten iterations, the MobileNet algorithm achieved a 99% accuracy rate during the test. The accuracy of testing evaluation of the MobileNet experiment are presented in Figure 3, which can be found here.

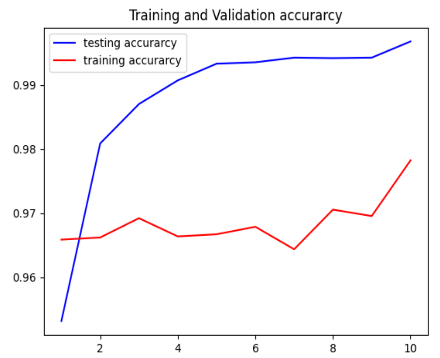
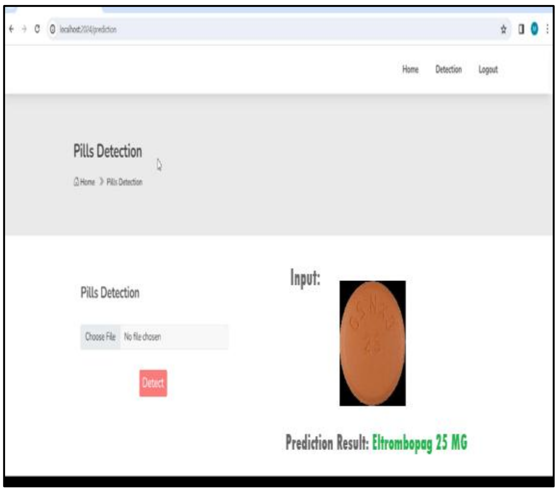


Figure 3. Accuracy of MobileNet

As a consequence of the needed findings for the thorough research of pill detections, the works have achieved the satisfaction of detecting a pill based on the physics structure of the pill. As can be seen in Figure 4, the MobileNet model is able to provide a prediction regarding the name of the pill picture.



5. CONCLUSIONS

With regard to the identification of pharmaceuticals from photographs, the deep learning-based pill detection system that has been suggested exhibits a high level of accuracy and reliability. It is able to achieve a training and validation accuracy of 98% by utilising the MobileNet architecture, which demonstrates its potential to reduce medication mistakes and assist healthcare professionals in confirming prescriptions in an effective manner. Through the implementation of this solution, the pill identification process is greatly automated, hence lowering the reliance on manual techniques and minimising the danger of human mistake, consequently contributing to the protection of patients.

This paper presented a robust and lightweight pill detection and classification system using the MobileNet architecture. The model's ability to accurately identify pills in real time, even with limited computational resources, makes it ideal for mobile healthcare applications. Our experiments show that MobileNet delivers high accuracy with minimal processing overhead, enabling practical deployment in clinical and field settings. Future enhancements could include multi-pill detection, integration with electronic health records (EHR), and using OCR or NLP for imprint interpretation. This system contributes to improving medication safety, reducing manual error, and supporting healthcare providers with reliable drug identification.

REFERENCES

[1] Bates, D. W., & Slight, S. P. (2014). "Medication errors: What is their impact?" British Journal of Clinical Pharmacology, 78(2), 347–357.

[2] Snigdha Sridhar, Akshaya S , Anika C , “Detection and Identification of Pills using Machine Learning Models”, “ 2023 2nd International Conference on Vision towards emerging trends in communication and network technologies”, 979-8-3503-4798-2/23/$31.00 2023 IEEE.

[3] Joyce Saraiva Lima , Luan Sousa Cardeiro , “Pill Image Classification using Machine Learning ’, 2020 , 8th Brazilian Conference of Intelligence Systems (BRACIS), doi:10.1109/BRACIS.2020.00103.

[4] Priyanka Narendra Kadam, Ashwini Dagadu Jagtap, “Drug Pill Recognition using Data

science and Machine Learning'', International Journal of Innovative Research in Science, Engineering and Technology (IJIRSET), Volume 10, Issue 6, June 2021.

[5] Hyuk-Ju Kwon, Hwi-Gang Kim and Sung-Hak Lee, "Pill Detection Model for Medicine Inspection Based on Deep Learning", Chemosensors, vol. 10, no. 1, p. 4, Dec. 2021, doi: 4

[6] A. Craswell, K. Bennett, J. Hanson, B. Dalgliesh, and M. Wallis, "Implementation of distributed automat ed medication dispensing units in a new hospital: Nursing and pharmacy experience," J. Clin. Nurs., vol. 30, pp. 2863–2872, 2021.

[7] Jaya Prakash pondy, Pills dataset. (2023). Kaggle. https://www.kaggle.com/datasets/jayaprakashpondy/pills-dataset