OPEN ACCESS JOURNALS

Indiana Journal of Multidisciplinary Research

Abbriviate Tittle- Ind J Multi Res

ISSN *(Online)*- 2583-3820

Journal Homepage Link- https://indianapublications.com/Journals/IJMR

DOI: https://doi.org/10.5281/zenodo.12674988

INDIANA PUBLICATIONS
PRODUCTIVE AND QUALITY RESEARCH

Research Article

Volume-04|Issue-03|2024

Using Machine Learning Models to Detect and Identify Pills

Kavitha N[1], Alina Stefi[2], P M Varsha[3], Sanvika Sumana[4], Vani Vandana Simha*[5]

[1]Assistant Professor, Department of Electronics Communication & Engineering, RV Institute of Technology and Management, Bengaluru, Karnataka, India.

[2,3.4,5] Student, Department of Electronics Communication & Engineering, R V Institute of Technology and Management, Bengaluru, Karnataka, India.

Article History

Received: 20.05.2024

Accepted: 05.06.2024

Published: 30.06.2024

Citation

Kavitha, N., Stefi, A., Varsha, P. M., Sumana, S., Simha, V. V. (2024). Using Machine Learning Models to Detect and Identify Pills. *Indiana Journal of Multidisciplinary Research*, *4*(3), 155-161.

Abstract: For the pharmaceutical industry, healthcare systems, and law enforcement agencies, pill detection and identification are critical tasks. Traditional methods of pill identification often rely on manual inspection, which is time-consuming and prone to human error. The Deep learning methods have demonstrated incredible promise in automating this process in recent years, offering efficient and accurate solutions. We provide a thorough analysis of the most recent deep learning methods for pill identification and detection in this paper. The study discusses modern learning architectures, particularly focusing on the YOLOv5 architecture customized for pill recognition tasks. Furthermore, we delve into the challenges associated with dataset preparation, model training, and deployment in real-world scenarios. We highlight important performance indicators including F1-score, recall, accuracy, and precision and examine how they relate to real-world applications. Additionally, we explore the integration of auxiliary technologies such as image preprocessing techniques and data augmentation strategies to enhance detection accuracy. Through a thorough examination of recent literature and experimental results, we provide insights into the strengths and limitations of existing methodologies and offer directions for upcoming studies. Overall, our research emphasizes the transformative deep learning's potential in reshaping pill identifying and detecting pills, offering a robust framework that holds practical promise for enhancing healthcare systems and pharmaceutical quality control measures.

Keywords: Pill detection, Pill identification, Deep learning, YOLO v5Architecture, Accuracy, Pharmaceutical industry.

INTRODUCTION

Precise pill identification and detection are essential in the pharmaceutical and healthcare industries, ensuring patient safety and effective treatment. With advancements in technology, particularly in deep learning, innovative methods for identifying and detecting pills have surfaced, offering promising solutions to healthcare professionals and patients. In relation to pill detection, the YOLO architecture tailored for such tasks stands out. YOLO, renowned for its efficiency and effectiveness in image classification, proves to be a valuable tool in analyzing pill images and extracting relevant features. This architecture, designed for mobile and embedded vision applications, offers lightweight yet powerful solutions for pill detection and identification. By leveraging YOLO capabilities, researchers and developers can enhance the accuracy and efficiency of pill detection systems. The integration of YOLO with computer vision technologies enables the development of automated pill detection systems that streamline processes in healthcare facilities. Intelligent image processing algorithms, coupled with YOLO extraction capabilities, swiftly analyze images from cameras or scanners, accurately identifying pills and providing essential information on composition, dosage, and usage instructions. This integration lowers errors related to manual identification methods and increases efficiency. Furthermore, YOLO versatility extends beyond pill detection, as its lightweight architecture it can be implemented in settings with limited resources, such as mobile devices and edge computing platforms.

This adaptability opens up new possibilities for extending identifying and detecting pills capabilities remote or underserved areas, where access to traditional healthcare infrastructure may be limited. Moreover, YOLO's efficiency makes it appropriate for use in wearable technology or mobile applications aimed at improving medication adherence and patient outcomes. By incorporating algorithms into these platforms, patients can conveniently verify their medication intake by capturing images of pills. These applications cross-reference captured images with a comprehensive database, ensuring patients take the correct medication at the prescribed dosage and schedule. Additionally, personalized reminders generated by these applications promote adherence to treatment regimens, ultimately improving health results. The utilization of YOLO architecture in pill detection underscores the significance of interdisciplinary collaboration between healthcare professionals, computer scientists.

PROBLEM STATEMENT

***Corresponding Author:** Vani Vandana Simha

The presented project tackles several critical challenges in the realm of medication management. One of the foremost issues addressed is the potential for human error in medication identification, a concern that can result in medication mix-ups and adverse health effects. Furthermore, the project confronts the complexity arising from the diverse shapes and sizes of medical pills, making their identification a challenging task. The variability in pill images adds another layer of complexity, impacting the accuracy of detection and identification. In response to these challenges, the project places a strong emphasis on the development of a user-friendly interface. This proactive approach aims to enhance accessibility and ease of use, particularly for individuals requiring efficient pill identification. By combining technological solutions with user-centric design, the project seeks to significantly improve medication safety and user experience.

MATERIALS AND METHODS

Gathering and Preparing a Labeled Dataset of Pill Images for Classification

The Pills Image Dataset collection began with an examination of publicly available resources, with Kaggle emerging as the key repository for various image datasets[1]. Using Kaggle's large library of picture datasets, we discovered and obtained a dataset including photos of various medications. The dataset was methodically organized into 20 separate classes, each representing a unique sort of medication. To maintain uniformity and consistency, we methodically organized each class into individual files, each dedicated to a certain pill kind. Within each folder, photographs of the particular pill were taken from various angles, under varied lighting situations, and against different backdrops. This thorough organization aided in the production of a labeled dataset in which each image was connected with the appropriate pill type, pharmacological class, and use. By methodically collecting a diversified and well-labeled dataset, we hoped to provide a basic resource for training and assessing deep learning models in pill recognition and categorization.

Data Preprocessing

Images in a dataset is an essential step in preparing them for training a machine learning model. It involves transforming the raw images into a format that is suitable for input into the neural network model. Here's a general overview of image preprocessing and how it's implemented in the provided code: General Image Preprocessing: Resizing: Images in the dataset may have varying sizes. Resizing them to a fixed size ensures consistency and allows the model to process them efficiently. Common target sizes include (224, 224), (128, 128), etc. Normalization: Normalizing pixel values helps in stabilizing the training process and improving convergence. Typically, pixel values are scaled to the range [0, 1] or [-1, 1] by dividing by the maximum pixel value (e.g., 255 for RGB images). Augmentation (Optional): Data augmentation techniques like rotation, flipping, zooming, and shifting can be used to increase dataset variety and improve model generalization. However, augmentation is not always essential and may be determined by the task's unique requirements.

Our dataset was preprocessed using the following steps:

Rescaling: Both training and validation Image Data Generators set the rescale parameter to 1./255. This rescales the photos' pixel values to the range [0, 1], thereby normalizing them. This is a frequent approach in picture preprocessing to guarantee that the input pixel values are within a reasonable range[2].

Target Size: The target size argument resizes images to a set target size of (224, 224). This ensures that the images fed into the model have the same dimensions, which is necessary for compatibility with the Mobile Net architecture used in the code. Batch Size: The batch size option determines how many photos will be included in each batch during training and validation. This value can be modified depending on the computing resources available and the dataset's features.

By performing these preparation processes, the photos in the dataset are converted into a format appropriate for training the Mobile Net model. The Image Data Generators create train generator and validation generator objects, which contain batches of preprocessed pictures and their labels, ready to be fed into the neural network model for training and validation.

Feature Selection

In the context of deep learning and YOLO feature selection is often defined as the process of identifying useful features from raw input data (pictures) in order to properly reflect the dataset's underlying patterns[3]. However, in deep learning, feature selection is frequently accomplished implicitly inside the layers of the neural network rather than as a distinct preprocessing step. Let's talk about feature selection in the context of YOLO and how it relates to the supplied code.

General feature selection in YOLO: In YOLO, convolutional layers are employed automatically learn and extract information from input photos. These layers are made up of filters that convolve over the input image, collecting patterns like edges, textures, and forms. Pooling layers (such as MaxPooling2D) minimize the spatial dimensions of features while keeping important information. They aid in identifying the most significant traits and lowering computational complexity.

Global Average Pooling (GAP) layers combine[4] feature maps across spatial dimensions to summarize the presence of key characteristics. They contribute to

***Corresponding Author:** Vani Vandana Simha

lowering the dimensionality of feature maps and extracting high-level features.

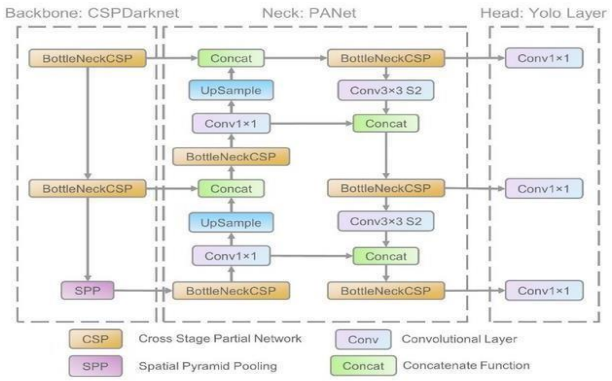

Figure 1: YOLOv5Architecture

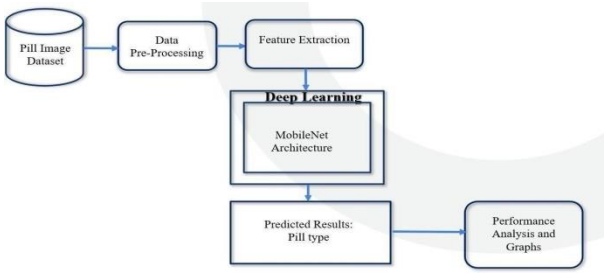

Figure 2: System architecture for Identification and detection of pill using machine learning system

Feature selection in our dataset is done implicitly by the YOLO architecture and subsequent layers. Here's how it's implemented. The foundation model is the YOLO architecture, which has been imported from Tensor Flow Keras applications. MobileNet was pre-trained on the ImageNet dataset and taught to extract key characteristics from photos.

Global Average Pooling Layer: Following the main model's output, a GlobalAveragePooling2D layer is added. This layer combines collected information from the underlying model across spatial dimensions to identify the most essential ones.

Dense Layers: Dense layers are placed on top of the GlobalAveragePooling2D layer to further process and alter the chosen features. These deep layers help us discover complicated patterns and correlations from the dataset.

Finally, a Dense layer with softmax activation is used to get the final predictions. The number of units in this layer is proportional to the number of classes in the dataset (20 in this example), and softmax activation guarantees that the output represents class probabilities. By using YOLO algorithm features and then processing them through thick layers, the model successfully chooses and learns key characteristics from the input photos. This enables it to generate accurate predictions about the provided dataset.

YOLO Architecture

YOLO is a popular deep learning architecture developed for efficient and lightweight convolutional neural networks (CNNs) appropriate for mobile and embedded vision applications. Our technique makes use of the YOLO architecture, a cutting-edge convolutional neural network (CNN) built for efficient picture classification tasks, which is ideal for mobile and embedded vision applications. YOLO was loaded from the Tensor Flow Keras apps library. This lets us to use the pre-defined YOLO architecture and its pre-trained weights. To instantiate the YOLO architecture as the basic model, use base_model=MobileNet(weights = 'ImageNet', include top=False). The weights='ImageNet' option specifies that YOLO's pre-trained weights from the ImageNet dataset should be utilised. These pre-trained weights encapsulate rich feature representations discovered during a large-scale picture classification assignment. The include_top=False option specifies that YOLO's[5] fully connected layers (top layers), which are responsible for ImageNet classification, should be removed. We replace the top layers with custom layers for our specific classification task.

Feature Extraction: To extract features from input photos, use base_model. output to retrieve the Mobile Net base model's output. This result is then passed into the next levels for additional processing and categorization.

Global Average Pooling with Dense Layers A Global Average Pooling 2D layer is used to spatially aggregate the retrieved characteristics. This layer calculates the average of each feature map over all spatial dimensions, producing a fixed-length vector.

Following the Global Average Pooling 2D layer, Dense layers are used for extra processing and classification. A dense layer of 1024 units with ReLU activation is added, followed by a dense output layer with softmax activation[6].

These dense layers help to understand complicated patterns and correlations in the information, allowing the model to make correct predictions.

Model Compilation and Training: After creating the model architecture, use model. Compile. The Adam optimizer is utilised for gradient descent optimisation, and categorical cross-entropy is chosen as the loss function, which is appropriate for multi-class classification applications. The model is then trained by using the fit function on the training and validation data generators (train generator and validation generator). During training, the model's weights are adjusted iteratively using the calculated loss and back propagation algorithms.

***Corresponding Author:** Vani Vandana Simha

We use the YOLO architecture's efficiency, effectiveness, and transfer learning characteristics to create a CNN model that is suited to our unique categorization problem. The pre-trained YOLO base model functions as a feature extractor, allowing us to focus on fine-tuning the top layers for our specific dataset, resulting in enhanced model performance and efficiency.

Predicted Results

Pill Type Identification The Predicted Results represents the output of our machine learning model, which takes preprocessed images of pills as input and predicts the pill type, specifically identifying the specific pill.

Multi-Class Classification: Our model conducts multi-class classification, with each class representing a separate pill type[7]. For example, our dataset includes photos of 20 distinct types of pills, and our model predicts one of these 20 classes for each input image.

Performance Analysis and Graphs

Accuracy Assessment: The model's predictions for pill type identification are measured using a variety of performance measures, including accuracy, precision, recall, and F1 score. Accuracy evaluates overall prediction accuracy[8], while precision and recall assess the model's ability to properly identify individual pill kinds and prevent misclassification. To better understand and comprehend model performance, we generate graphical representations such as line plots or curves. The graph's x-axis shows the number of training epochs, while the y-axis reflects the related model loss and accuracy. The graph shows two different curves: one for model loss (training and validation loss) and one for accuracy. Analysing the model's loss and accuracy curves throughout training epochs provides insights into its learning dynamics and performance patterns. Ideally, we want to see a drop in both training and validation loss over time, showing that the model is learning and generalizing well. Similarly, a steady improvement[9] in both training and validation accuracy indicates that the model is getting better at properly classifying drugs. Over fitting Detection: Differences between the training and validation curves, such as a decrease in training loss but an increase in validation loss, might suggest over fitting.

Dataset description

The Pills Dataset is a comprehensive collection of data pertaining to various types of pills commonly used in medical treatments. It serves as a valuable resource for pharmaceutical research, medical practitioners, and healthcare analytics[10]. The dataset encompasses detailed information about the physical attributes, composition, and medical indications of different pills. This dataset serves as a fundamental repository of information facilitating comprehensive understanding and analysis of pharmaceutical formulations and their clinical applications. The Pills Dataset amalgamates data from a diverse array of reputable sources spanning pharmaceutical compendia, regulatory databases, peer-reviewed literature, drug monographs, and proprietary drug databases. Rigorous data curation methodologies are employed to ensure data integrity, accuracy, and currency.

RESULTS AND DISCUSSION

In our project, we developed a user-friendly interface that detects both real-time and stored pill images. Through rigorous evaluation, we assessed the performance metrics of our application, including the confusion matrix, F1-confidence curve, precision-recall curve, precision-confidence curve, and recall-confidence curve. These metrics serve as benchmarks for evaluating the accuracy and robustness of our detection system. Our findings demonstrate the effectiveness of our GUI application in accurately identifying pills, providing valuable insights into its practical utility and potential impact in healthcare and pharmaceutical domains.

Figure 3: Pills for stored images

***Corresponding Author:** Vani Vandana Simha

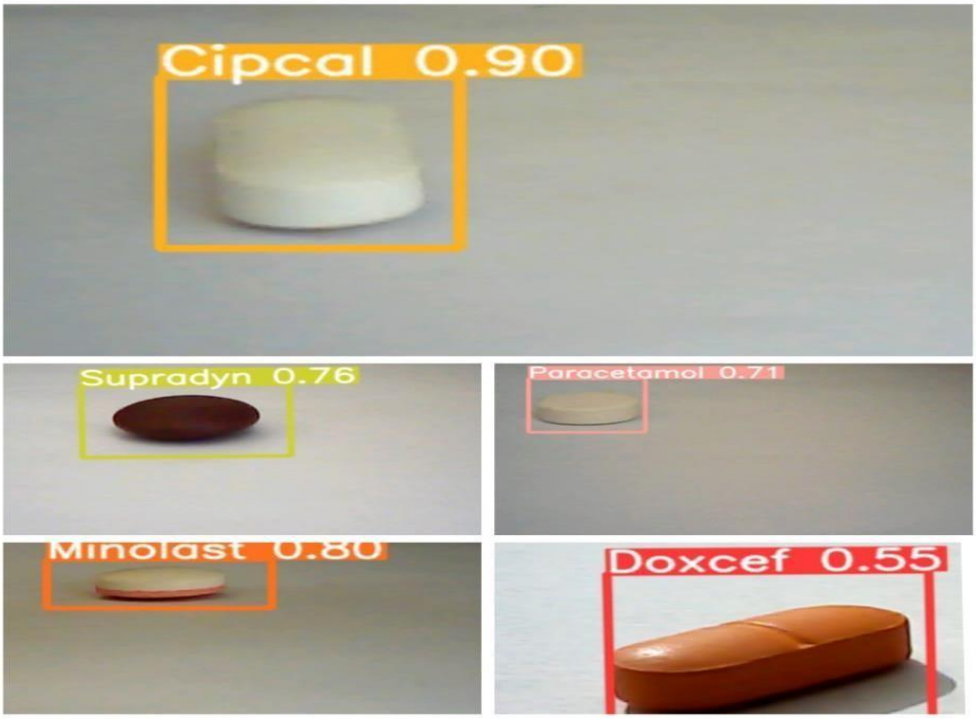

Figure 4:: pills for real-time

We evaluated the performance metrics like confusion matrix, F1-Confidence score, Precision-Recall curve, Precision-Confidence curve and Recall-Confidence curve.

- **Confusion Matrix:** A tabular representation of classification results showing true positives, true negatives, false positives, and false negatives.
- **F1-Confidence Curve:** A graph displaying F1 score variation at different confidence thresholds for optimal precision and recall balance.
- **Precision-Recall Curve:** A plot illustrating the trade-off between precision and recall across various classification thresholds.
- **Precision-Confidence Curve:** A graphical depiction showing how precision changes with confidence thresholds in classification tasks.
- **Recall-Confidence Curve:** A curve demonstrating how recall varies with confidence thresholds in classification scenarios.

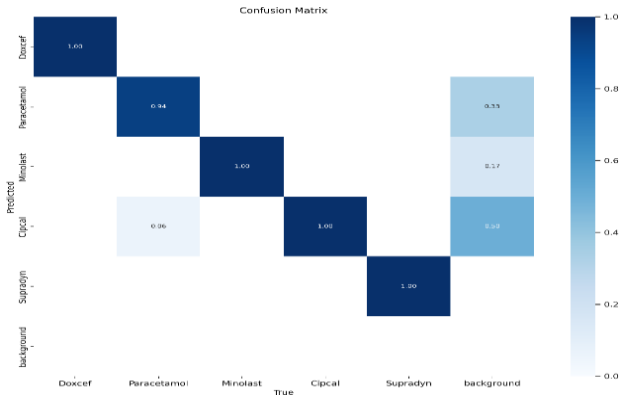

Figure 5: Confusion matrix

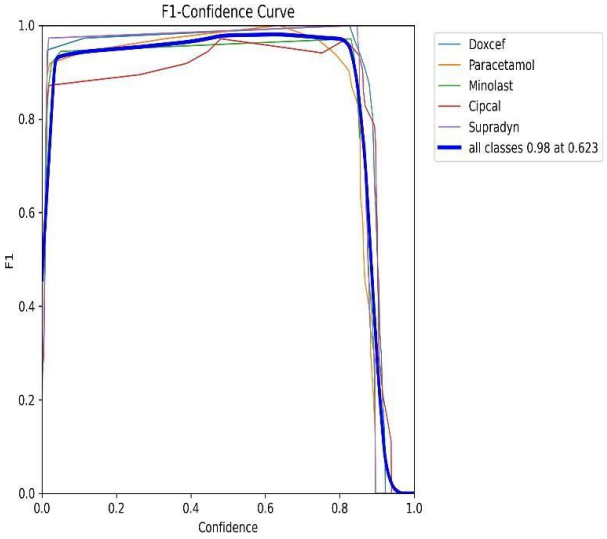

Figure 6: F1-Confidencecurve

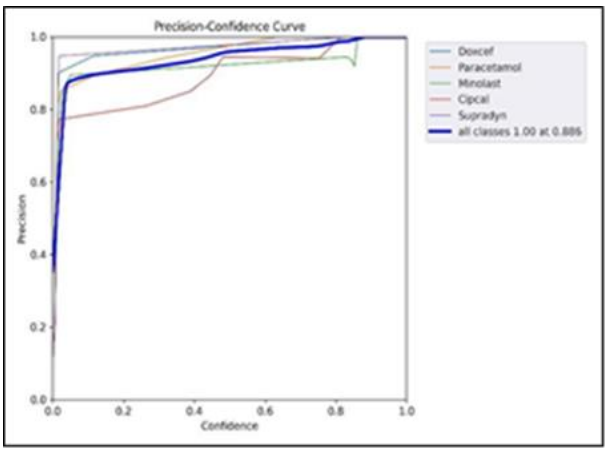

Figure 7: Precision-Confidence curve

***Corresponding Author:** Vani Vandana Simha

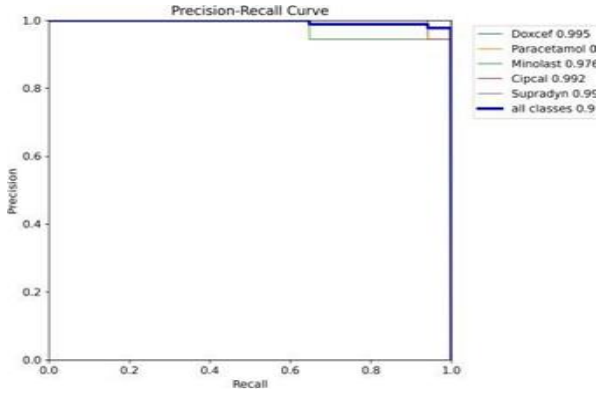

Figure 8: Precision-Recall curve

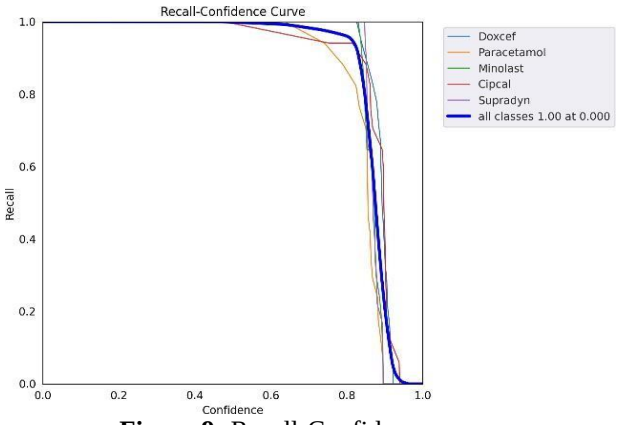

Figure 9: Recall-Confidence curve

CONCLUSION

In conclusion, the presented project represents a comprehensive endeavor aimed at addressing critical challenges in medication management through innovative technological solutions. The persistence of medication errors underscores the urgency of developing accurate and efficient methods for pill detection and identification within healthcare settings. Leveraging advancements in machine learning and artificial intelligence, this project endeavors to enhance patient safety and streamline healthcare processes. The primary objectives of the project include the creation of a precise pill identification system capable of accurately classifying pharmaceutical pills and providing essential information about their properties. By utilizing Python programming language and YOLO architecture, the project achieves significant milestones in model development and training, leveraging a meticulously labeled dataset of pill images. The methodology employed encompasses a systematic approach, encompassing dataset acquisition, preprocessing, model training, and performance evaluation. Each stage is meticulously executed to ensure the reliability, robustness, and accuracy of the pill detection system. Software requirements are meticulously met, with essential libraries and frameworks utilized to facilitate seamless model development and deployment. Anticipated outcomes include the provision of detailed information about detected pills, such as drug class, generic name, and therapeutic uses, thereby empowering healthcare professionals with invaluable insights into medication management. Rigorous evaluation and testing validate the efficacy and reliability of the system, positioning it as a valuable asset in healthcare settings. In summary, the project represents a significant contribution to advancing medication safety and efficiency through the integration of cutting-edge technology and healthcare practices. By harnessing the power of machine learning and artificial intelligence, the project underscores the potential to revolutionize pill detection and identification processes, ultimately improving patient outcomes and advancing the frontier of healthcare technology.

ACKNOWLEDGEMENTS

The successful presentation of our paper would be incomplete without the mention of the people who made it possible and whose constant guidance crowned our effort with success. We would like to extend our gratitude to the RV Institute of Technology and Management, Bengaluru, and Dr. Jayapal R, Principal, RV Institute of Technology and Management, Bengaluru for providing all the facilities to carry out this research article. We thank Dr. Nataraj V, Professor and Head, Department of Electronics Communication and Engineering, RV Institute of Technology and Management, Bengaluru, for his initiative and encouragement.

We would also like to thank our guide, Dr. Kavitha N, Assistant Professor, Department of Electronics Communication and Engineering, RV Institute of Technology and Management, Bengaluru, for her constant guidance and input. We would like to thank all the Teaching Staff and Non-Teaching Staff of the college for their cooperation. Finally, we extend our heartfelt gratitude to our families for their encouragement and support without which we would not have come so far. Moreover, we thank all our friends for their invaluable support and cooperation.

REFERENCES

1. Sridhar, S., Akshaya, & Anika. (2023). Detection and identification of pills using machine learning model. In *Vision towards emerging trends in communication and network technologies* (pp. 979-8-3503-4798-2/23/$31.00 ©2023 IEEE).
2. Borude, S., Patil, S., Bhoirkar, R., & Shirsath, I. (2022). Drug pill recognition system using deep learning. *International Research Journal of Engineering (IRJET)*, 9(11), November 2022. e-ISSN: 2395-0056.
3. Kwon, H.-J., Kim, H.-G., & Lee, S.-H. (2022). Pill detection model for medicine inspection based on deep learning. *Chemosensors, 2022*.
4. Ting, H.-W., & Yow-wen. (2022). A drug identification model developed using deep learning technologies. In *2020 IEEE International*

***Corresponding Author:** Vani Vandana Simha

Conference on Bioinformatics and Biomedicine (BIBM) (pp. 1154-1159). IEEE.

5. Eghbali, S., Shakib, S., & Azizi, S. (2018). Deep learning-based pill recognition system using convolutional neural networks. In *2018 Third International Conference on Smart Systems and Inventive Technology (ICSSIT)* (pp. 780-782).
6. Rashid, A., Jaffar, A., & Rafique, M. U. (2020). Medication identification from pill images. In *2020 8th Brazilian Conference on Intelligent Systems (BRACIS)* (pp. 556-561). doi: 10.1109/BRACIS.2019.00103.
7. Xie, Y., Lv, P., Liu, X., & Zheng, Y. (2017). An improved and robust pill recognition system using convolutional neural networks. In *2017 2nd International Conference on Advance Computing and Innovative Technologies in Engineering (ICACITE)* (pp. 773-777).
8. Ozmermer, T. E., Stanislavs, V. R., & Nescerecka, A. (2021). Pill identification using proxy-based deep metric learning and exact solution. In *Proceedings of the IEEE International Conference on Computer Vision, October 2021* (pp. 2980-2988).
9. Srikamdee, S. (2022). An application for pill identification using deep learning and means clustering. In *26th International Computer Science and Engineering Conference (ICSEC)* (pp. 978-1-6654-9198-3/22/$31.00 ©2022 IEEE).
10. Meehan, J. (2021). DLI-IT: A deep learning approach to drug label identification through image and text embedding. In *Proceedings of the 2021 IEEE Conference on Computer Vision and Pattern Recognition, Columbus, OH, USA, 23–28 June 2021* (pp. 580–587).

***Corresponding Author:** Vani Vandana Simha