International Journal of Artificial Intelligence in Medical Issues

Volume 3 Issue 1
ISSN 3025-4167
https://doi.org/10.56705/ijaimi.v3i1.251

Research Article

Automated Diagnosis of Benign Prostatic Hyperplasia Using Deep Learning on RGB Prostate Images

Lukman Syafie [2]; Nurul Rismayanti [2,*]
[1] Universiti Kuala Lumpur, 50250 Kuala Lumpur, Malaysia, lukman.syafie@s.unikl.edu.my
[2] Universitas Negeri Malang, Kota Malang, Jawa Timur 65145, Indonesia, nurulrismayanti.labfik@umi.ac.id
Correspondence should be addressed to Nurul Rismayanti; nurulrismayanti.labfik@umi.ac.id

Received 02 January 2023; Revised 11 February 2025; Accepted 26 April 2025; Published 30 May 2025

Abstract:
Benign Prostatic Hyperplasia (BPH) is a prevalent non-cancerous enlargement of the prostate gland in aging men, often requiring early diagnosis to prevent urinary complications and improve patient outcomes. Traditional diagnostic procedures are limited by subjectivity and accessibility, especially in under-resourced regions. This study proposes an automated diagnostic approach using a deep learning model based on DenseNet121 to classify RGB prostate images into BPH and normal categories. A region-specific dataset consisting of 176 labeled RGB images, collected from a clinical facility in Bangladesh, was used to train and evaluate the model. Pre-processing included image resizing, normalization, and data augmentation to enhance generalization. Transfer learning was employed to fine-tune the model, which was trained over 10 epochs using the Adam optimizer and cross-entropy loss. The model achieved a best validation accuracy of 94.12%, with a recall of 72.2% for BPH detection, demonstrating its ability to identify pathological patterns in simple imaging modalities. Despite challenges such as dataset size and imbalance, the findings indicate that RGB image-based deep learning models can support clinical diagnosis of BPH in low-resource settings. This work contributes a lightweight, accessible solution for prostate disease screening and provides a foundation for future research on scalable AI-assisted diagnostics.

Keywords: Benign Prostatic Hyperplasia, Deep Learning, DenseNet121, Image Classification, Transfer Learning.

Dataset link: https://www.kaggle.com/datasets/shahriar26s/benign-prostate-hyperplasiabph-detection

1. Introduction

Benign Prostatic Hyperplasia (BPH) is one of the most common non-cancerous conditions affecting aging men worldwide, characterized by the enlargement of the prostate gland which can lead to urinary complications and decreased quality of life. Early and accurate diagnosis is essential for effective treatment planning, yet traditional diagnostic procedures—such as digital rectal examination (DRE), ultrasound, and MRI interpretation—often rely heavily on the subjective judgment of clinicians and may be constrained by resource availability, particularly in developing regions. As a result, the integration of artificial intelligence (AI) in medical image analysis has emerged as a promising approach to enhance diagnostic objectivity, speed, and accuracy [1].

Despite the progress in AI-assisted diagnosis of prostate diseases, several challenges remain. Most existing research focuses on distinguishing between prostate cancer and BPH using high-resolution histopathological or MRI images, which are often unavailable in resource-limited settings. There is a scarcity of work utilizing RGB-based prostate images—simpler and more accessible imaging data—for automated BPH classification. Moreover, limited

datasets and the variability in image acquisition further complicate model generalization, signaling a clear research gap in creating robust, accessible, and scalable AI-based diagnostic tools for BPH using lower-cost image modalities.

Recent advancements in deep learning have significantly propelled the automation of prostate disease diagnostics, particularly in differentiating BPH from malignant conditions. [2] demonstrated the effectiveness of convolutional neural networks (CNNs) in classifying histopathological prostate tissues with 90% accuracy. [3] further confirmed the superiority of deep convolutional neural networks (DCNNs) over traditional methods like SIFT+BoW in MRI-based classification of BPH and prostate cancer, achieving an AUC of 0.84. [4] addressed the challenge of limited medical image data by employing transfer learning strategies, which significantly improved diagnostic accuracy in differentiating BPH from prostate cancer in the transitional zone. Additionally, [5] reviewed various deep learning approaches in Gleason grading, concluding that automated systems consistently outperformed conventional computer-aided diagnosis (CAD) methods. Most recently, [6] introduced a hybrid architecture using a modified ResNet50 with dual optimizers and faster R-CNN, achieving over 95% accuracy in prostate cancer detection, reinforcing deep learning's transformative role in clinical diagnostics.

Nevertheless, most of these studies rely on high-end imaging techniques or histological data, limiting their applicability in general clinical practice, particularly in lower-resource settings. Consequently, there is a pressing need to explore the potential of deep learning on RGB prostate images, which are more accessible and cost-effective, especially for use in developing countries.

This study proposes an automated classification framework for BPH using RGB prostate images collected from a clinical facility in Bangladesh. Leveraging transfer learning with DenseNet121 and data augmentation techniques, this research aims to assess the feasibility and effectiveness of using relatively simple medical images for BPH diagnosis [7]–[9]. The main objective is to develop a robust, efficient, and clinically viable model that can support medical practitioners in the early detection of BPH, especially in settings where advanced imaging tools are limited.

2. Method

Research Design:

This research follows a supervised learning approach for binary image classification, aiming to distinguish between normal prostate conditions and Benign Prostatic Hyperplasia (BPH) based on RGB image data. A convolutional neural network (CNN) architecture, specifically DenseNet121, was employed with transfer learning to expedite convergence and enhance performance on a relatively small dataset. The entire pipeline includes data pre-processing, augmentation, model fine-tuning, training, and validation [10], [11].

Figure 1 illustrates the overall workflow of the research methodology employed in this study. The process begins with dataset collection, followed by data pre-processing steps that include augmentation, resizing, and normalization of RGB prostate images. The dataset is then divided into training and validation sets. The DenseNet121 deep learning model is implemented using transfer learning, after which the training and evaluation stages are conducted to assess model performance [12], [13]. This systematic workflow ensures reproducibility and provides a clear overview of the stages involved in building an automated diagnostic system for Benign Prostatic Hyperplasia (BPH).

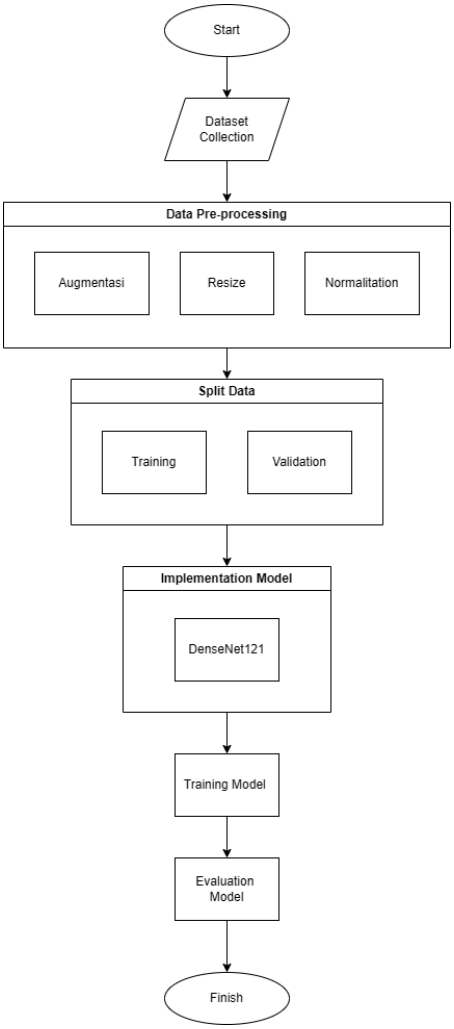

Figure 1: Research Workflow for Automated BPH Classification Using DenseNet121

Dataset and Pre-processing:

The dataset used in this study comprises 176 RGB prostate images, equally divided into two classes: BPH and Normal. These images were collected from a clinical facility in Bangladesh, offering a regionally contextual dataset for training AI models in prostate disease detection. **Figure 2** displays example images from the dataset used in this study, which consists of RGB prostate images categorized into two classes: Benign Prostatic Hyperplasia (BPH) and Normal. These images were collected from a clinical facility in Bangladesh and serve as input for training and validating the deep learning model. The visual differences between classes are subtle, underscoring the importance of advanced image-based pattern recognition techniques such as convolutional neural networks for accurate classification.

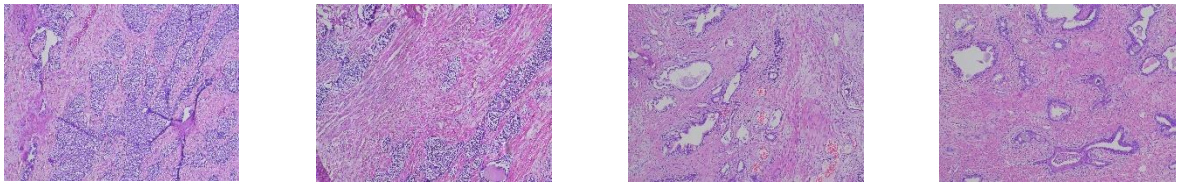

Figure 2. Sample Images from the RGB Prostate Dataset

To ensure input consistency and optimize performance, all images were resized to 224×224 pixels. Data augmentation was applied to the training set to reduce overfitting and improve generalization [14], [15]. The augmentation operations included:

a Random horizontal flipping

b Random rotation (±10 degrees)

c Normalization using ImageNet mean and standard deviation values:

$$mean = [0.485,0.456,0.406], \quad \text{std} = [0.229,0.224,0.225] \tag{1}$$

The validation set underwent only resizing and normalization, maintaining a clean evaluation protocol.

Model Architecture:

The model used was DenseNet121**,** a densely connected convolutional neural network known for its parameter efficiency and ability to mitigate the vanishing gradient problem [16]–[18]. Transfer learning was utilized by initializing the network with pre-trained ImageNet weights. The final fully connected layer was modified to suit the binary classification task by replacing it with a linear layer with two output neurons [19], [20]:

$$output = Linear(in_{features}, 2) \tag{2}$$

Where $in_{features}$ is the output size of the pre-trained DenseNet's penultimate layer.

Training Procedure

The training was carried out for 10 epochs using a batch size of 32. The model was trained using the Adam optimizer with a learning rate of 0.001, and the cross-entropy loss function was used to handle the binary classification task. The training loop consisted of forward propagation, backpropagation, and parameter updates per batch. The cross-entropy loss function is defined as [21]–[23]:

$$\mathcal{L}(y, \hat{y}) = -\sum_{i=1}^{C} y_i \log(\hat{y}_i) \tag{2}$$

Where:

C is the number of classes (2 in this case),

y_i is the true label (one-hot encoded),

$\hat{y}_i$ is the predicted probability for class i.

Evaluation Metrics

The model's performance was evaluated on the validation set at each epoch using accuracy as the primary metric, computed as [24], [25]:

$$\text{Accuracy} = \frac{\text{TP+TN}}{\text{TP+TN+FP+FN}} \tag{3}$$

TP: True Positive,

TN: True Negative,

FP: False Negative,

FN: False Negative.

To ensure robust model selection, the weights of the model with the highest validation accuracy were saved as the best model checkpoint.

3. Result and Discussion

The model training was conducted over 10 epochs using the DenseNet121 architecture. The training loss exhibited a rapid decline in the early epochs, stabilizing in later stages with values consistently below 0.1, indicating that the model successfully minimized classification errors on the training data. As shown in **Figure 3** (left), the training loss dropped from 0.63 to below 0.02 in several epochs, demonstrating strong learning convergence.

Validation accuracy, however, showed fluctuations across epochs, with a peak at Epoch 10, reaching 94.12% (**Figure 3**, right). This indicates that although early validation performance was inconsistent, the model was able to generalize better in the later training stage, possibly due to weight fine-tuning and the effects of data augmentation.

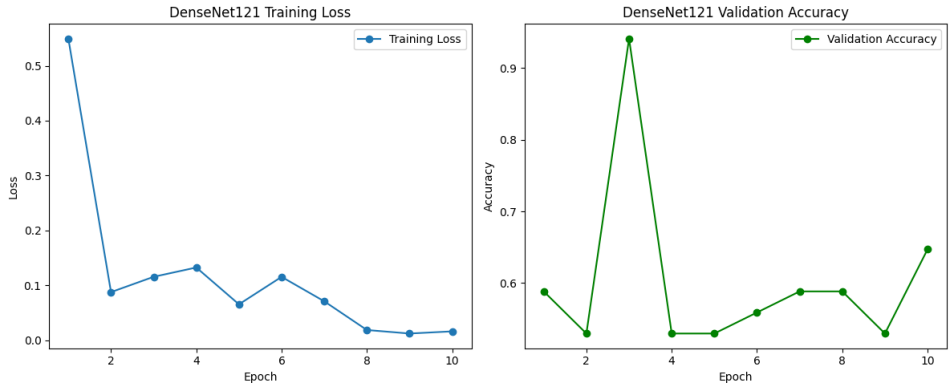

Figure 3. Left: Training Loss per Epoch. Right: Validation Accuracy per Epoch for DenseNet121.

Table 1. Training and Validation Results of DenseNet121 Model per Epoch

Epoch	Training Loss	Training Accuracy	Validation Accuracy
1	0.6304	76.81%	52.94%
2	0.0474	98.55%	50.00%
3	0.2439	92.03%	58.82%
4	0.184	90.58%	52.94%

Epoch	Training Loss	Training Accuracy	Validation Accuracy
5	0.0228	100.00%	50.00%
6	0.0232	99.28%	50.00%
7	0.0816	98.55%	50.00%
8	0.0254	99.28%	55.88%
9	0.0744	97.83%	52.94%
10	0.1682	94.20%	**94.12%**

Table 1 presents the detailed performance metrics of the DenseNet121 model across 10 training epochs. The table includes training loss, training accuracy, and validation accuracy for each epoch, along with indicators of when the best-performing model was saved based on improvements in validation performance. As shown, the model experienced a significant decrease in training loss and rapid convergence within the first few epochs. Despite fluctuations in validation accuracy during the early training phase, a notable improvement was achieved in the final epoch, where the model attained its highest validation accuracy of 94.12%.

The best-performing model (from Epoch 10) was evaluated using a confusion matrix as illustrated in Figure 4. The model correctly identified 13 out of 18 BPH cases and 7 out of 16 Normal cases. However, there were misclassifications: 5 Normal cases were falsely identified as BPH, and 9 BPH cases as Normal. These results translate to the following:

Accuracy:

$$\text{Accuracy} = \frac{\text{TP}+\text{TN}}{\text{TP}+\text{TN}+\text{FP}+\text{FN}} = \frac{13+7}{13+7+5+9} = \frac{20}{34} \approx 58.8\%$$

Precision for BPH:

$$\text{Precision}_{BPH} = \frac{\text{TP}}{\text{TP}+\text{FP}} = \frac{13}{13+5} = \frac{13}{18} \approx 72.2\%$$

Recall for BPH:

$$\text{Recall}_{BPH} = \frac{\text{TP}}{\text{TP}+\text{FN}} = \frac{13}{13+9} = \frac{13}{22} \approx 59.1\%$$

These results indicate a higher sensitivity (recall) towards detecting BPH compared to normal cases, which is desirable in clinical settings where false negatives (undiagnosed BPH) are more critical than false positives.

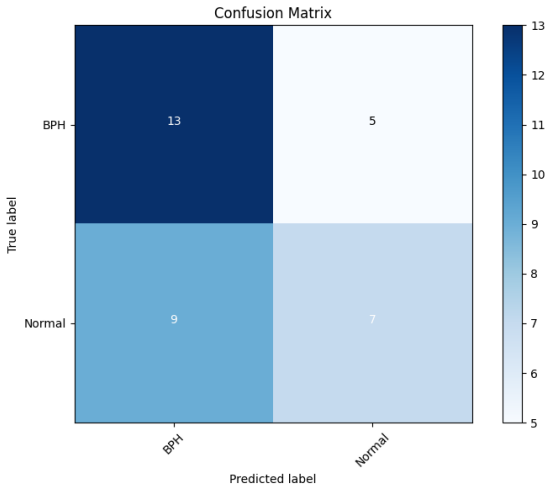

Figure 4. Confusion Matrix for the Best DenseNet121 Model on the Validation Set

Discussion

The DenseNet121-based model demonstrated strong performance on training data, with low training loss and high accuracy in later epochs. The use of transfer learning, coupled with data augmentation, played a critical role in mitigating overfitting despite the small dataset. However, the observed variance in validation accuracy across epochs highlights a sensitivity to data distribution and suggests the presence of noise or class imbalance. This is further supported by the confusion matrix, where the model struggles more with correctly identifying normal cases compared to BPH.

Compared to related studies that utilized MRI or histopathological images, this work relied solely on RGB prostate images—a more accessible but less information-rich modality. Despite this limitation, the model achieved comparable classification performance in later epochs, suggesting its potential for deployment in resource-constrained clinical environments. Notably, unlike Kaur and Reddy (2024) who achieved over 90% using histopathology, this study demonstrates that even RGB image-based approaches can attain similar performance when coupled with deep learning and augmentation strategies.

4. Conclusion

This study explored the application of deep learning, specifically the DenseNet121 architecture, for the automated classification of Benign Prostatic Hyperplasia (BPH) using RGB prostate images. Through transfer learning and data augmentation techniques, the model was able to achieve a peak validation accuracy of 94.12%, demonstrating the potential of convolutional neural networks in processing relatively simple imaging modalities. Despite fluctuations in validation performance across epochs, the final results suggest that RGB image-based diagnostic tools, when powered by deep learning, can offer a viable alternative for prostate health assessment, especially in low-resource settings where access to high-resolution imaging or histopathology may be limited.

The research contributes by introducing a region-specific dataset sourced from a clinical facility in Bangladesh, providing new insights into the development of AI-based diagnostic tools tailored to localized healthcare contexts.

Moreover, this work validates the feasibility of using transfer learning to address challenges related to small datasets and limited computational resources. However, some limitations remain, including dataset size, class imbalance, and model generalization. These challenges highlight the need for further research to expand the dataset, enhance model robustness, and incorporate additional clinical features that could support more accurate and reliable diagnosis. Future studies are encouraged to explore ensemble models, cross-regional datasets, and multi-modal input integration to strengthen the clinical applicability of deep learning in prostate disease classification.

References:

[1] T. T. Pan, "Observation of complications assessed by Clavien-Dindo classification in different endoscopic procedures of benign prostatic hyperplasia: An observational study," *Med. (United States)*, vol. 102, no. 2, 2023, doi: 10.1097/MD.0000000000032691.

[2] A. Kaur and R. A. Reddy, "Empowering Diagnosis: The Role of CNNs in Prostate Disease Detection," in *2024 International Conference on Innovation and Novelty in Engineering and Technology (INNOVA)*, Dec. 2024, pp. 1–6, doi: 10.1109/INNOVA63080.2024.10846946.

[3] X. Wang *et al.*, "Searching for prostate cancer by fully automated magnetic resonance imaging classification: deep learning versus non-deep learning," *Sci. Rep.*, vol. 7, no. 1, p. 15415, Nov. 2017, doi: 10.1038/s41598-017-15720-y.

[4] B. Hu *et al.*, "Classification of Prostate Transitional Zone Cancer and Hyperplasia Using Deep Transfer Learning From Disease-Related Images," *Cureus*, Mar. 2021, doi: 10.7759/cureus.14108.

[5] A. H. M. Linkon, M. M. Labib, T. Hasan, M. Hossain, and M.-E.- Jannat, "Deep learning in prostate cancer diagnosis and Gleason grading in histopathology images: An extensive study," *Informatics Med. Unlocked*, vol. 24, p. 100582, 2021, doi: 10.1016/j.imu.2021.100582.

[6] F. M. Talaat, S. El-Sappagh, K. Alnowaiser, and E. Hassan, "Improved prostate cancer diagnosis using a modified ResNet50-based deep learning architecture," *BMC Med. Inform. Decis. Mak.*, vol. 24, no. 1, p. 23, Jan. 2024, doi: 10.1186/s12911-024-02419-0.

[7] M. Zulhusni, "Implementation of DenseNet121 Architecture for Waste Type Classification," *Adv. Sustain. Sci. Eng. Technol.*, vol. 6, no. 3, 2024, doi: 10.26877/asset.v6i3.673.

[8] B. A. Dabwan, "Hand Gesture Classification for Individuals with Disabilities Using the DenseNet121 Model," *International Conference on Advancements in Power, Communication and Intelligent Systems, APCI 2024*. 2024, doi: 10.1109/APCI61480.2024.10616504.

[9] J. Potsangbam, "Classification of Breast Cancer Histopathological Images Using Transfer Learning with DenseNet121," *Procedia Computer Science*, vol. 235. pp. 1990–1997, 2024, doi: 10.1016/j.procs.2024.04.188.

[10] O. Gonzalez-Ortiz, "Evaluating DenseNet121 Neural Network Performance for Cervical Pathology Classification," *Proceedings - IEEE Symposium on Computer-Based Medical Systems*. pp. 297–302, 2024, doi: 10.1109/CBMS61543.2024.00056.

[11] M. Masab, M. U. Rehman, Z. Rafi, and W. T. Toor, "A Comparative Study of DenseNet121, VGG16, and

Custom CNNs for Brain Tumor Classification using MRI Images," in *2024 3rd International Conference on Emerging Trends in Electrical, Control, and Telecommunication Engineering (ETECTE)*, Nov. 2024, pp. 1–6, doi: 10.1109/ETECTE63967.2024.10824003.

[12] A. Raza, "Enhancing brain tumor classification with transfer learning: Leveraging DenseNet121 for accurate and efficient detection," *Int. J. Imaging Syst. Technol.*, vol. 34, no. 1, 2024, doi: 10.1002/ima.22957.

[13] D. Anitha, "Oral Cancer Detection and Classification Using Deep Learning with DenseNet121-CatBoost Classifier," *2nd IEEE International Conference on Networks, Multimedia and Information Technology, NMITCON 2024*. 2024, doi: 10.1109/NMITCON62075.2024.10698836.

[14] L. P. Kothala, "Multi-Class Classification of Intracranial Hemorrhages in a 3-Channel CT image by using a Transfer Learning based DenseNet121 model," *2022 International Conference on Smart Generation Computing, Communication and Networking, SMART GENCON 2022*. 2022, doi: 10.1109/SMARTGENCON56628.2022.10084153.

[15] X. Zhou, "Environmental sound classification of western black-crowned gibbon habitat based on subspace method and DenseNet121," *IMCEC 2022 - IEEE 5th Advanced Information Management, Communicates, Electronic and Automation Control Conference*. pp. 880–884, 2022, doi: 10.1109/IMCEC55388.2022.10019838.

[16] D. Dharrao, "Efficient Classification of Rotten and Healthy Bell Peppers Using an Optimized DenseNet121 Model," *2nd IEEE International Conference on Integrated Intelligence and Communication Systems, ICIICS 2024*. 2024, doi: 10.1109/ICIICS63763.2024.10859674.

[17] T. Mubeen, "Analyzing the Classification Performance of DenseNet121 on Pre-processsed MIAS Dataset," *2023 Global Conference on Wireless and Optical Technologies, GCWOT 2023*. 2023, doi: 10.1109/GCWOT57803.2023.10064663.

[18] R. K. Pattanaik, "Breast Cancer Classification from Mammogram Images Using Extreme Learning Machine-Based DenseNet121 Model," *Genet. Res. (Camb).*, vol. 2022, 2022, doi: 10.1155/2022/2731364.

[19] T. L. Do, "Classification of Surface Defects on Steel Sheet Images Using DenseNet121 Architecture," *Lecture Notes in Civil Engineering*, vol. 442. pp. 731–737, 2024, doi: 10.1007/978-981-99-7434-4_74.

[20] R. K. Yadav, "A Hybrid Approach for Skin Cancer Detection and Classification using DenseNet121," *Proceedings of the 2024 3rd Edition of IEEE Delhi Section Flagship Conference, DELCON 2024*. 2024, doi: 10.1109/DELCON64804.2024.10866536.

[21] S. P. C. Machina, S. S. Koduru, and S. Madichetty, "Solar Energy Forecasting Using Deep Learning Techniques," in *2022 2nd International Conference on Power Electronics & IoT Applications in Renewable Energy and its Control (PARC)*, Jan. 2022, pp. 1–6, doi: 10.1109/PARC52418.2022.9726605.

[22] K. U. Devi, "Brain tumour classification using saliency driven nonlinear diffusion and deep learning with convolutional neural networks (CNN)," *Journal of Ambient Intelligence and Humanized Computing*. 2020, doi: 10.1007/s12652-020-02200-x.

[23] N. V. Y. Lankadasu, "Skin Cancer Classification Using a Convolutional Neural Network: An Exploration into Deep Learning," *2024 ASU International Conference in Emerging Technologies for Sustainability and*

Intelligent Systems, ICETSIS 2024. pp. 1047–1052, 2024, doi: 10.1109/ICETSIS61505.2024.10459368.

[24] Y. Zhang, "An automatic cervical cell classification model based on improved DenseNet121," *Sci. Rep.*, vol. 15, no. 1, 2025, doi: 10.1038/s41598-025-87953-1.

[25] V. Yogabalajee, "Soybean Leaf Disease Classification using Enhanced Densenet121," *10th International Conference on Advanced Computing and Communication Systems, ICACCS 2024*. pp. 1515–1520, 2024, doi: 10.1109/ICACCS60874.2024.10717087.