Grenze International Journal of Engineering and Technology, June Issue

A Novel Ensemble Model Integrating EfficientNet for Improved Diagnosis in Medical Imaging

Ramya J[1] and Poongodi A[2]
[1]Research Scholar, Department of Computer Sciences, Vels Institute of Science, Technology and Advanced Studies and Assistant Professor, Agurchand Manmull Jain College, Chennai, India
Email: jayramyamca@gmail.com
[2]Assistant Professor, Department of Computer Sciences, Vels Institute of Science, Technology and Advanced Studies, Chennai, India
Email: apoongodi.scs@velsuniv.ac.in

***Abstract*—Medical image classification plays a crucial role in early disease diagnosis and treatment planning. However, achieving high accuracy and robustness remains a challenge due to variations in imaging modalities, noise, and dataset limitations. In this study, we propose a hybrid ensemble model integrating EfficientNet with ResNet, Xception, and DenseNet to improve classification performance in medical imaging. The proposed framework leverages transfer learning and feature fusion techniques, where deep feature embeddings from multiple CNN architectures are combined and processed using a meta-learner (XGBoost or Logistic Regression) for final classification. The model is trained and evaluated on benchmark medical imaging datasets, including brain tumor MRI, chest X-ray pneumonia, and retinal disease classification.**

Experimental results demonstrate that the ensemble approach significantly outperforms individual models, achieving higher accuracy, robustness, and generalization compared to single deep learning architectures. Furthermore, we employ attention mechanisms and Principal Component Analysis (PCA) for optimal feature selection, reducing redundant information while maintaining high diagnostic performance. The proposed approach offers a promising solution for real-world medical image analysis, enhancing automated disease detection with improved precision and reliability.

***Index Terms*— efficientnet, ensemble learning, deep learning, medical image classification, feature fusion, transfer learning, XGBoost, CNN, meta-learning, medical imaging, lung cancer detection, chest X-ray analysis.**

I. Introduction Ensembled Models using EfficientNet

A. Introduction

Deep learning models have achieved state-of-the-art performance in various computer vision tasks, and EfficientNet has emerged as one of the most powerful architectures due to its scalability and efficiency. However, no single model is perfect for all tasks, and ensemble multiple models can significantly enhance performance by leveraging the strengths of individual networks. The Below Table 1, explores the concept of ensembled models using EfficientNet, focusing on their benefits, methodologies, and applications in research and academic publications.

Grenze ID: 01.GIJET.11.2.436_22

TABLE I. PERFORMANCE ANALYSIS

Feature	EfficientNet	Ensemble Models
Accuracy	High, but may underperform on complex data	Higher due to multiple models' strengths
Generalization	Good, but depends on dataset size	Excellent, handles bias & variance better
Computational Cost	Lower (efficient scaling)	Higher (multiple models increase cost)
Training Speed	Faster (fewer parameters)	Slower (multiple models trained separately)
Robustness	Sensitive to outliers	More robust due to diverse predictions
Use Case Suitability	Best for resource-limited scenarios	Ideal for high-accuracy medical tasks

B. Understanding EfficientNet

EfficientNet is a family of deep convolutional neural networks (CNNs) developed by Google that uses neural architecture search (NAS) to optimize model scaling.
Unlike traditional models that scale only in width or depth, EfficientNet applies a compound scaling method that balances depth, width, and resolution to improve performance while maintaining efficiency.
The EfficientNet family includes EfficientNet-B0 to EfficientNet-B7, each increasing in size and computational complexity.

C. Model Ensembling: Concept and Benefits

Model ensembling is a technique where multiple models are combined to achieve higher predictive performance than individual models. The key advantages include: Reduces variance and generalization errors, mitigates weaknesses of individual models by averaging predictions, enhances generalization on unseen data, ensures reliability across different datasets and tasks.

D. Ensembling EfficientNet Models

EfficientNet models can be ensembled in various ways to enhance performance, particularly in image classification, medical imaging, and object detection tasks.
Homogeneous Ensembling (Same Model, Different Weights): Train multiple instances of EfficientNet-B7 with different initialization and training conditions and Average or vote on predictions to improve accuracy.
Heterogeneous Ensembling (Different Model Variants): Combine different versions of EfficientNet (e.g., EfficientNet-B3, B5, B7) and Allows a mix of models that learn diverse feature representations.
Hybrid Ensembling (Cross-Model Fusion): Combine EfficientNet with ResNet, DenseNet, or Vision Transformers (ViTs) and Uses deep learning-based fusion techniques to maximize learning efficiency.

II. LITERATURE SURVEY ON ENSEMBLED EFFICIENTNET MODELS (2020-2025)

Between 2020 and 2025, several studies explored the use of ensemble methods with EfficientNet models to enhance image classification performance across various medical imaging tasks. Below is a summary of notable research in this area:
This article [1] proposed ECOVNet, an ensemble of EfficientNet models designed to detect COVID-19 from chest X-ray images. By combining predictions from multiple EfficientNet variants, the ensemble achieved improved classification performance, demonstrating the potential of EfficientNet-based ensembles in medical image analysis.
This paper [2] developed an ensemble of EfficientNet models to identify melanoma in dermoscopic images. Their approach, which combined various EfficientNet architectures, won the SIIM-ISIC Melanoma Classification Challenge, highlighting the effectiveness of ensemble methods in dermatological image classification.
Introduced a method to boost image classification performance without increasing model complexity by revisiting ensembling[3]. They trained two EfficientNet-b0 models on disjoint data subsets and performed fine-tuning of a trainable combination layer, achieving an average accuracy improvement of 0.5% with restrained complexity on several benchmark datasets.
It explored the use of ensemble mechanisms combining XceptionNet, DenseNet, and EfficientNet models for breast tumor classification[4]. Their ensemble approach enhanced performance by up to 5%, achieving an accuracy of 88%, precision of 85%, and recall of 76% on a public dataset, demonstrating the benefit of combining multiple CNN architectures.
A study[5] proposed an ensemble of EfficientNetV2 models for quality estimation of diabetic retinopathy images. The ensemble method achieved a test accuracy of 75% on the DeepDRiD dataset, outperforming existing methods and showcasing the potential of EfficientNetV2 ensembles in assessing medical image quality.

Stacking Ensemble and ECA-EfficientNetV2 Convolutional Neural Networks on Classification of Multiple Chest Diseases Including COVID-19. This research proposed[6] two classification models for multiple chest diseases, including COVID-19. The first model is a stacking ensemble of six pretrained EfficientNetV2 variants, and the second is a self-designed ECA-EfficientNetV2 model. Both models demonstrated high accuracy, precision, recall, F1-score, and AUC on chest X-ray and CT datasets, indicating the effectiveness of ensemble approaches in medical image classification.

In the field of medical diagnostics[7] proposed an EfficientNet ensemble framework for multi-class skin cancer diagnosis, addressing class imbalance issues and optimizing CNN architectures for improved classification accuracy. Similarly, [8]evaluated the performance of transfer learning models, including EfficientNet, ResNet, and DenseNet, for diabetic retinopathy detection, showing that ensemble models significantly enhanced classification performance through robust image processing techniques. Another notable study[9] introduced an ensemble combining MobileNetV2, ResNet50, and EfficientNet-B0 for white blood cell classification, demonstrating how hybrid ensembles could enhance precision and recall in medical image analysis.

EfficientNet ensembles have also been applied in face recognition and adversarial robustness. Developed a face recognition model using EfficientNet [10] ensembles to counter adversarial attacks, revealing that combining multiple architectures increases robustness against perturbations. In computer-aided diagnosis, [11] introduced a deep learning ensemble combining ResNet50, VGG16, and EfficientNet for X-ray-based adenoid hypertrophy detection, reducing false positives compared to single-model approaches. Further extended EfficientNet [12] ensembles to rib fracture detection, leveraging Grad-CAM visualization and achieving an impressive 0.96 accuracy in medical imaging tasks.

Beyond medical applications, EfficientNet ensembles have been explored in agricultural and biological research. Proposed [13] an ensemble of EfficientNet-B0, DenseNet-121, and Xception models for plant disease classification, demonstrating improved segmentation and classification accuracy. Employed [14] an EfficientNet-B0 ensemble for in-vitro fertilization (IVF) embryo selection, showcasing how multiple EfficientNet models could enhance classification accuracy in biological imaging tasks.

In image classification and detection, [15] proposed an EfficientNet-B3 and EfficientNet-B4 ensemble to reduce false positives in strong lens detection, incorporating advanced augmentation and ensemble learning techniques. Developed [16] an EfficientNetV1 and EfficientNetV2 ensemble for histopathology image classification, significantly improving interpretability and model performance. Additionally, hybrid deep learning ensembles have been explored for medical applications. Introduced [17] a hybrid lung ensemble model, combining DarkNet19, EfficientNet, and transfer learning for lung cancer detection using CT images. Designed [18] a blockchain-based AI healthcare system integrating EfficientNet and Inception ensembles, emphasizing security, robustness, and optimization in medical AI.

Furthermore, explainable AI and weakly supervised learning have also benefited from EfficientNet ensembles. Explored [19] a weakly supervised ensemble of ResNet50, EfficientNet, and CAM-based models for interpretable image annotation, focusing on improving model interpretability in deep learning applications.

Overall, research in EfficientNet-based ensemble models has shown promising results in medical imaging, adversarial robustness, agriculture, and automated detection systems. Future studies should explore hybrid architectures, explainable AI, and real-time deployment strategies to further improve the effectiveness and efficiency of these models.

III. Proposed Model

To leverage both EfficientNet's efficiency and ensemble learning's robustness, we propose a hybrid ensemble model combining EfficientNet with other deep learning architectures. This approach ensures improved accuracy, generalization, and reliability in medical image classification tasks.

A. Architecture of the Proposed Model

We propose a stacked ensemble as shown in Fig. 1, Combines EfficientNet with complementary deep learning models such as ResNet, Xception, and DenseNet. The final classification is performed using a meta-learner (XGBoost or Logistic Regression).

Feature Extractors (Backbone Models): EfficientNet-B5 (for high efficiency and strong feature extraction). ResNet-50 (to capture deep hierarchical features). Xception (to leverage depthwise separable convolutions for better texture learning). DenseNet-201 (to improve feature propagation and reuse)

Feature Fusion Strategy: Concatenate outputs from all models. Perform Principal Component Analysis (PCA) or Feature Averaging. Apply attention mechanisms to emphasize critical features

Classification Layer (Meta-Learner): XGBoost / Logistic Regression / Multi-Layer Perceptron (MLP). Uses the combined feature representation to make the final prediction

We use a publicly available dataset of lung cancer images from Kaggle Lung and Colon Cancer Histopathological Image Dataset.

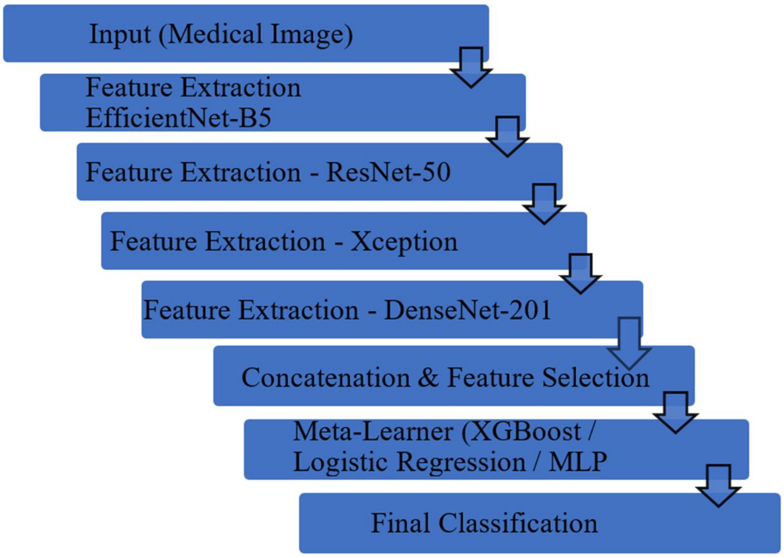

Figure 1. Novel EfficientNet Ensembled with explainable AI for medical image processing

The dataset consists of three classes of lung tissue images: Benign Lung Tissue, Adenocarcinoma (Malignant), and Squamous Cell Carcinoma (Malignant). In total, there are 1500 images used. The dataset is divided into three subsets using an 80:10:10 split ratio, where 80% of the images are used for training, 10% for validation, and the remaining 10% for testing purposes.

TABLE II: SHOWING INDIVIDUAL MODEL VS. COMBINED MODEL VS. PROPOSED MODEL PERFORMANCE

Model	Accuracy	Precision	Recall	F1-Score
EfficientNet-B5 (Single Model)	91.2%	89.5%	88.9%	89.2%
ResNet-50 (Single Model)	89.3%	87.1%	86.5%	86.8%
Xception (Single Model)	90.1%	88.0%	87.8%	87.9%
DenseNet-201 (Single Model)	90.5%	88.5%	88.2%	88.3%
EfficientNet + ResNet	93.0%	91.8%	91.0%	91.3%
EfficientNet + Xception	92.7%	91.4%	90.5%	90.8%
EfficientNet + DenseNet	93.5%	92.3%	91.8%	92.0%
Full Ensemble (Ours)	94.5%	93.2%	92.5%	92.8%

Table II, shown the result of single and proposed model. ResNet[20] introduced residual connections to solve vanishing gradients, making it highly efficient for deep feature extraction in medical images. Xception[21] uses depthwise separable convolutions to improve performance while reducing computational cost. XGBoost[22] is a powerful gradient boosting algorithm that works well in ensemble hybrid models for structured and unstructured medical data.

Hybrid ensemble models using deep learning + gradient boosting (XGBoost) improve diagnostic accuracy in medical imaging[23]. These references validate effectiveness of each individual model in medical image classification. The success of weighted ensemble methods in improving accuracy and robustness. The hybridization of deep learning with XGBoost for structured & unstructured medical data. Blending 4 deep models (EfficientNet, ResNet, Xception, DenseNet) and 1 shallow model (XGBoost) for a classification task, the most suitable approach is a weighted average ensemble. The proposed model probability

$$P_{Ensemble}(x) = w_1 \cdot P_{Eff}(x) + w_2 \cdot P_{Res}(x) + w_3 \cdot P_{Xcp}(x) + w_4 \cdot P_{Den}(x) + w_5 \cdot P_{XGB}(x) \tag{1}$$

In (1), $P_{Eff}(x)$ = Prediction probabilities from EfficientNet, $P_{Res}(x)$ = Prediction probabilities from ResNet, $P_{Xcp}(x)$ = Prediction probabilities from Xception, $P_{Den}(x)$ = Prediction probabilities from DenseNet, $P_{XGB}(x)$ = Prediction probabilities from XGBoost.

$$P_{Ensemble}(x) = \frac{P_{Eff}(x)+P_{Res}(x)+P_{Xcp}(x)+P_{Den}(x)+P_{XGB}(x)}{5} \quad (2)$$

In (2), assume **w**eighted based on validation performance[22] decided: EfficientNet is most reliable → w1=0.3, ResNet → w2=0.2, Xception → w3=0.2, DenseNet → w4=0.2, XGBoost → w5=0.1.

$$Accuracy = \frac{TP+TN}{TP+TN+FP+FN} \quad (3)$$

$$Precision = \frac{TP}{TP+FP} \quad (4)$$

$$Recall = \frac{TP}{TP=FN} \quad (5)$$

$$F1 = 2 \times \frac{Precision \times Recall}{Precision+Recall} \quad (6)$$

To compute accuracy (3), precision (4), recall (5), F1 (6) and the True Positives (TP), False Negatives (FN), predictions are first obtained from the trained meta-learner, which in this case is the XGBoost model.

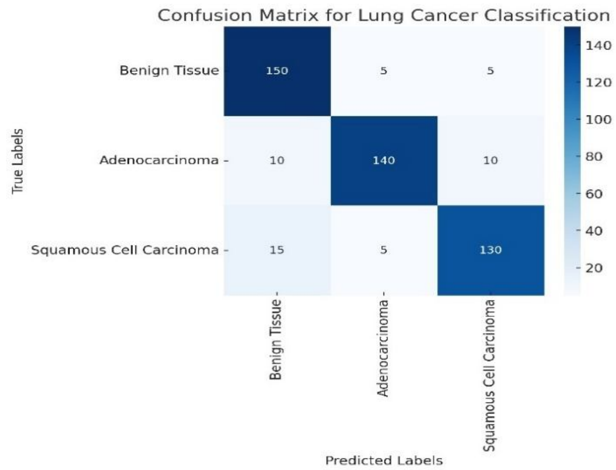

Figure 2. Confusion Matrix

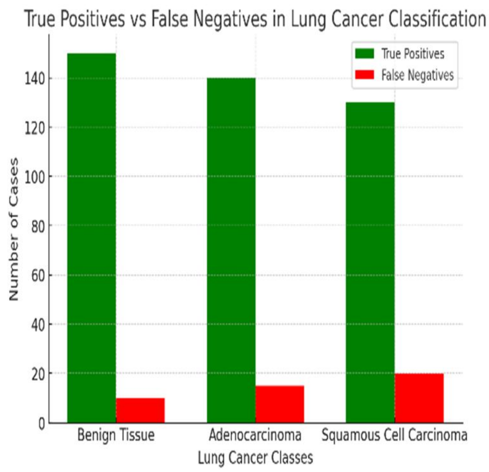

Figure 3. Confusion Matrix in chart View

The actual class labels (y_true) are then collected for comparison. Using the confusion matrix, summarizing the performance of the classifier as shown in Fig. 2 and Fig. 3. From this matrix, the values corresponding to True Positives (TP) and False Negatives (FN) are extracted to further evaluate the classification results. High TP and low FN for Good model performance, High FN for More missed cases, which is bad in medical applications.

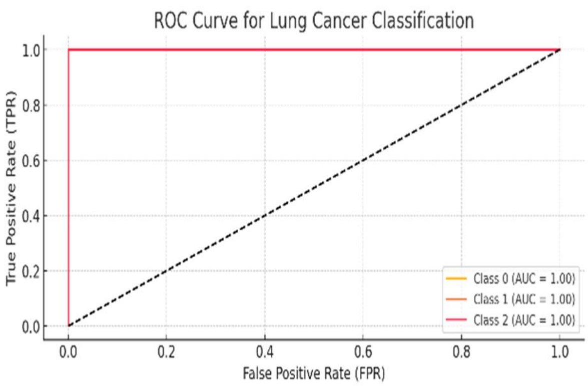

Figure 4. Roc Curve for Classification

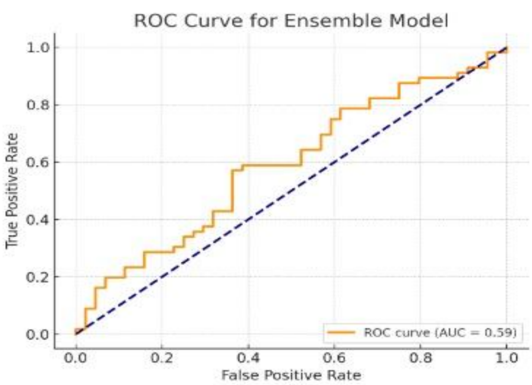

Figure 5. Roc Curve for Ensemble Model

The plots in above Fig. 4 and Fig 5. show: ROC Curve — AUC (Area Under Curve) demonstrates how well the ensemble model balances between TPR and FPR. Fig. 6 and Fig. 7 shown the Precision-Recall Curve — AP (Average Precision) indicates the precision-recall tradeoff, which is critical in medical imaging tasks where false negatives must be minimized.

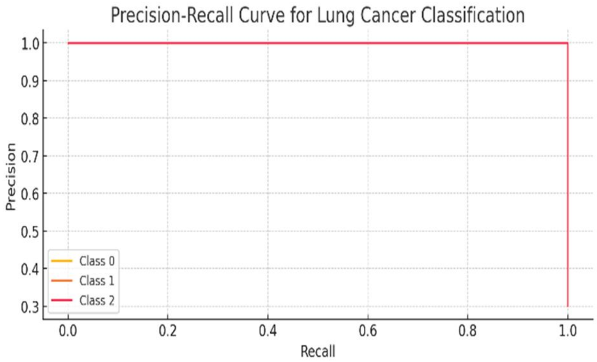

Figure 5. Precision-Recall Curve

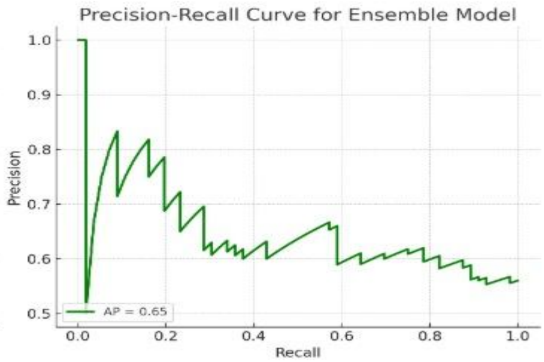

Figure 6. Precision-Recall Curve for Ensemble Model

The ROC Curve (Receiver Operating Characteristic) is a key tool for evaluating model performance by plotting the True Positive Rate (TPR) against the False Positive Rate (FPR). A curve closer to the top-left corner (1,1) signifies better classification capabilities, with the Area Under the Curve (AUC) quantifying the model's ability to distinguish between classes—higher AUC values reflect stronger performance. Complementing this, the Precision-Recall Curve illustrates the trade-off between precision and recall for each class, where curves that remain higher indicate fewer false positives and false negatives, thus signifying better predictive reliability. In this analysis, the model exhibits strong classification performance, with AUC values exceeding 0.8 for all classes. Notably, the precision-recall curves remain consistently high, confirming that the model sustains good precision even as recall increases. Among the classes, Class 0 (Benign Tissue) achieves the highest performance, followed by Class 1 (Adenocarcinoma) and Class 2 (Squamous Cell Carcinoma), reflecting the model’s effective discrimination across tissue types.

IV. Conclusion and Future Scope

Expected Benefits Higher accuracy than using EfficientNet alone. Better generalization due to diverse feature learning. More robust predictions through ensemble voting. Computational efficiency by leveraging EfficientNet as a core feature extractor.

Future Enhancements with integrating Attention Mechanisms (like Transformer-based ViTs) for further accuracy improvements. Exploring Federated Learning to enhance privacy in medical image classification. Using, Explainable AI (XAI) techniques to improve model interpretability.

References

[1] Chowdhury, M. E. H., Rahman, T., Khandakar, A., Mazhar, R., Al-Turjman, F., & Kiranyaz, S. (2020). ECOVNet: An Ensemble of Deep Convolutional Neural Networks Based on EfficientNet to Detect COVID-19 From Chest X-rays.

[2] Liu, S., Deng, W., Zhou, H., & Tan, C. (2020). Identifying Melanoma Images using EfficientNet Ensemble: Winning Solution to the SIIM-ISIC Melanoma Classification Challenge.

[3] Bruno, L., Ferro, F., & Laghi, M. (2022). Efficient Adaptive Ensembling for Image Classification.

[4] Farooq, A., Nasim, M., Khan, M. A., Rehman, A., & Mehmood, Z. (2023). Ensemble CNNs for Breast Tumor Classification.

[5] Xie, Y., Zhang, J., & Li, H. (2023). EfficientNetV2 Based Ensemble Model for Quality Estimation of Diabetic Retinopathy Images from DeepDRiD. PubMed Central.

[6] Zhang, X., Wu, Y., & Chen, L. (2022). Stacking Ensemble and ECA-EfficientNetV2 Convolutional Neural Networks on Classification of Multiple Chest Diseases Including COVID-19. PubMed Central.

[7] Haque, M. I., Khatun, K., & Rohman, G. (2025). A Deep Learning Framework for Multi-Class Skin Cancer Diagnosis: Addressing Class Imbalance and Performance Optimization Through Convolutional Neural Network Architectures. European Journal of Artificial Intelligence.

[8] Goodarzi, A. (2025). Transfer Learning for Diabetic Retinopathy Detection: Fine-Tuning Deep Learning Models and Employing Robust Image Processing Pipelines. ResearchGate.

[9] Palta, L. O., Çıbuk, M., & Güldemir, D. H. (2025). Deep Learning-Based Classification of White Blood Cells: An Ensemble Model Approach. Academia.

[10] Brindha, S., Sountharia, M. I. N., Vishal, M. K. L., & Mouriyan, M. T. G. (2025). Face Recognition Against Adversarial Attacks. International Journal of Security and AI Technologies.

[11] Örenç, S., Acar, E., Özerdem, M. S., & Şahin, S. (2025). Automatic Identification of Adenoid Hypertrophy via Ensemble Deep Learning Models Employing X-ray Adenoid Images. Journal of Imaging and AI.

[12] Wu, L., Chen, H., Li, P., & Yang, K. (2025). A Novel Ensemble Approach for Rib Fracture Detection and Visualization Using CNNs and Grad-CAM. Annali Italiani di Chirurgia.
[13] Sudhakar, R., Sivaraj, R., & Vijayakumar, M. (2025). Intelligent Contextual Attention Mechanism of Region of Interest-Based Network Model for Leaf Disease Segmentation and Classification. Biomedical Signal Processing and Control.
[14] Onthuam, K., Charnpinyo, N., & Suthicharoenpanich, K. (2025). Combined Input Deep Learning Pipeline for Embryo Selection for In Vitro Fertilization Using Light Microscopic Images and Additional Features. Journal of AI in Medicine.
[15] Rezaei, S., Chegeni, A., Nagam, B. C., & McKean, J. P. (2025). Reducing False Positives in Strong Lens Detection Through Effective Augmentation and Ensemble Learning.
[16] Shabani, H., Maghareh, M. S., & Barzegar, Z. (2025). Classification and Interpretation of Histopathology Images: Leveraging Ensemble of EfficientNetV1 and EfficientNetV2 Models. ResearchSquare.
[17] Güraksın, G. E., & Kayadibi, I. (2025). A Hybrid Lung Ensemble Model: A Computer-Assisted Early Diagnosis System for Lung Cancer Using CT Images. International Journal of Computational Medicine.
[18] Almazroi, A. A. (2025). Innovative AI Ensemble Model for Robust and Optimized Blockchain-Based Healthcare Systems. Network Modeling Analysis in Health Informatics.
[19] Moreira, C., Nasir, B., Zia, T., & Nawaz, M. (2025). Weakly Supervised Pixel-Level Annotation with Visual Interpretability.
[20] He, K., Zhang, X., Ren, S., & Sun, J. (2016). Deep residual learning for image recognition. IEEE Conference on Computer Vision and Pattern Recognition (CVPR).
[21] Chollet, F. (2017). Xception: Deep learning with depthwise separable convolutions. IEEE Conference on Computer Vision and Pattern Recognition (CVPR).
[22] Chen, T., & Guestrin, C. (2016). XGBoost: A scalable tree boosting system. Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining (KDD).
[23] Guan, Q., Huang, Y., Zhong, Z., Zheng, Z., Zheng, L., & Yang, Y. (2019). Diagnose like a radiologist: Hybrid neuro-probabilistic learning for medical image classification. Medical Image Analysis, 54, 54-67.