Indonesian Journal of Data and Science

Volume 6 Issue 1
ISSN 2715-9936
https://doi.org/10.56705/ijodas.v6i1.236

Research Article

Implementation of Ensemble Deep Learning for Brain MRI Classification in Tumor Detection

Rahmat Fuadi Syam

Universitas Pancasakti Makassar, Makassar, Sulawesi Selatan 90121, Indonesia, rahmat@unpacti.ac.id

Correspondence should be addressed to Rahmat Fuadi Syam; rahmat@unpacti.ac.id

Received 10 November 2024; Accepted 20 March 2025; Published 31 March 2025

Abstract:

Introduction: Brain tumor detection from MRI images is critical for early diagnosis and treatment planning. While individual deep learning models have shown high accuracy in medical image classification, combining multiple models can potentially enhance performance. This study aims to develop an ensemble deep learning framework using ResNet18 and DenseNet121 to improve the accuracy of brain tumor classification. **Methods**: A dataset of 7,023 brain MRI images categorized into four classes—glioma, meningioma, no tumor, and pituitary tumor—was used. Pre-processing included resizing to 224×224 pixels, normalization, and augmentation (random flipping and rotation). ResNet18 and DenseNet121 models were fine-tuned separately using the Adam optimizer with a learning rate of 0.001. The ensemble method was implemented by averaging the softmax outputs of both models to generate final predictions. **Results**: When evaluated individually, ResNet18 and DenseNet121 achieved validation accuracies of 97.72% and 97.79%, respectively. The ensemble model significantly outperformed both, reaching a validation accuracy of 99.36%. This result demonstrates that integrating both architectures effectively reduces misclassification and enhances overall robustness. Confusion matrix analysis confirmed high classification accuracy across all four tumor categories. **Conclusions**: The proposed ensemble deep learning approach successfully leverages the strengths of ResNet18 and DenseNet121, achieving superior classification accuracy for brain tumor detection in MRI images. This method holds promise as a reliable tool in clinical diagnostic workflows. Future research should focus on integrating additional architectures, advanced augmentation strategies, and hyperparameter optimization to further improve performance.

Keywords: Brain MRI, Deep Learning, Ensemble Learning, ResNet18, Tumor Detection.

1. Introduction

Brain tumor detection plays a critical role in medical diagnostics, particularly as advancements in imaging technologies such as MRI continue to enhance our ability to identify subtle abnormalities within the brain. The widespread adoption of deep learning techniques has further accelerated improvements in diagnostic accuracy, with convolutional neural networks (CNNs) becoming instrumental in classifying complex medical images. In this study, we focus on implementing an ensemble deep learning approach that leverages the complementary strengths of ResNet18 and DenseNet121 for classifying 7,023 human brain MRI images into four distinct categories: glioma, meningioma, no tumor, and pituitary tumor [1].

Despite significant progress using individual models, a notable research gap exists in the exploration of ensemble methods that combine multiple deep learning architectures for brain tumor detection. Many previous studies have concentrated on optimizing a single model with various techniques, yet the potential benefits of integrating different models—each with its unique feature extraction capabilities—remain underexplored. This gap highlights the need for innovative approaches that can address challenges such as data variability, class imbalance, and the complex nature of tumor features.

Recent state-of-the-art research has demonstrated the effectiveness of both ResNet18 and DenseNet121 in brain tumor classification. For instance, studies by [2] and [3] have shown that ResNet18, when combined with focal loss and advanced optimization techniques like the Bat Algorithm, can achieve classification accuracies as high as 95.54% and even reach 99% with perfect sensitivity and F1-score in some configurations. Similarly, DenseNet121 has been widely recognized for its robust feature propagation and ability to mitigate the vanishing gradient problem. Research by [4]–[8] underscores the superior performance of DenseNet121, with reported accuracies reaching up to 99% in various hybrid and ensemble settings. These advancements, however, also underline a significant problem: while both models exhibit strong performance individually, there remains an unmet need to develop a more reliable diagnostic tool by effectively combining their predictive strengths. The challenge lies in optimizing the integration of ResNet18 and DenseNet121 to handle the inherent complexities of MRI data and to improve overall classification robustness [9], [10], especially in real-world clinical scenarios [11], [12].

The primary objective of this research is to implement an ensemble deep learning framework that integrates ResNet18 and DenseNet121, aiming to enhance the accuracy and reliability of brain tumor detection from MRI images. By leveraging the complementary feature extraction capabilities of both models, this study aspires to contribute to the development of an automated diagnostic system that not only improves detection performance but also supports more precise clinical decision-making in the treatment of brain tumors.

2. Method:

In this research design, the process begins with gathering the MRI dataset, followed by essential pre-processing steps—such as augmentation, resizing, and normalization—to standardize and enhance the data. The dataset is then split into training and validation sets, ensuring an effective evaluation of model performance. Two convolutional neural network architectures, ResNet18 [13], [14] and DenseNet121 [15], are independently trained and evaluated to measure their individual capabilities. Subsequently, an ensemble method is implemented by combining these models' predictions to potentially improve classification accuracy [16]. This comprehensive approach aims to optimize the diagnostic performance of brain MRI classification through rigorous model training, evaluation, and ensemble integration.

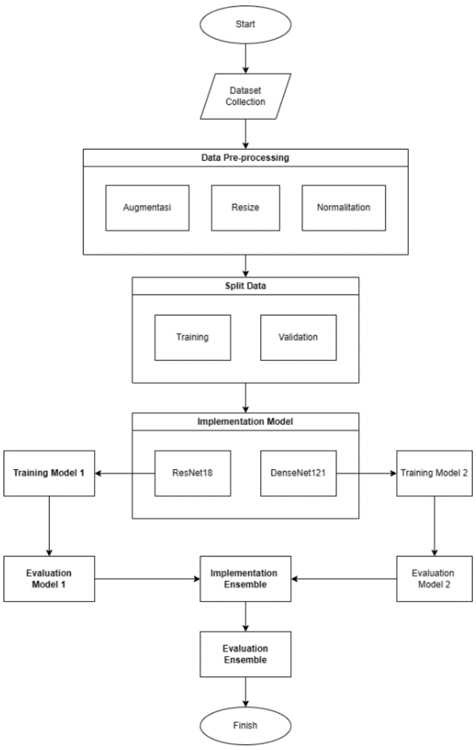

Figure 1. Research Design

Data Collection

The data collection phase involves compiling a dataset consisting of 7,023 human brain MRI images, which have been pre-labeled into four categories: glioma, meningioma, no tumor, and pituitary tumor. The dataset is organized into a structured folder system (e.g., dataset/train and dataset/val) to facilitate efficient loading using tools like datasets.ImageFolder. This systematic organization is crucial to ensure that the training and validation phases can proceed seamlessly.

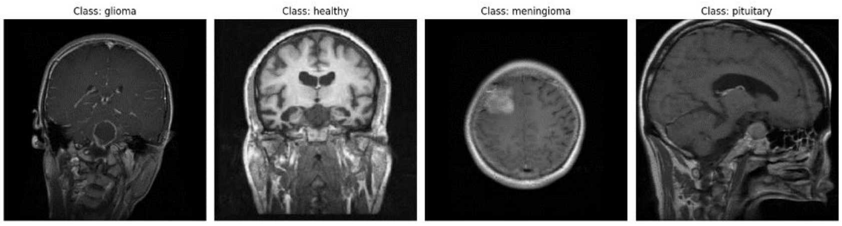

Figure 2. Sample of Dataset

Data Pre-processing

In the data pre-processing stage, the MRI images undergo several transformations to enhance data quality and variability before being input to the deep learning models. Initially, each image is resized to 224×224 pixels, a standard input size for pretrained models. For the training dataset, data augmentation techniques such as random horizontal flipping and random rotation (by 10 degrees) are applied. These augmentations help to increase the diversity of the dataset and reduce the risk of overfitting. After augmentation, images are converted to tensors and normalized using the mean [0.485,0.456,0.406] and standard deviation [0.229,0.224,0.225]. The normalization process is governed by the formula:

$$x_{norm} = \frac{x - \mu}{\sigma} \tag{1}$$

Where $\boldsymbol{x}$ is the original pixel value, $\boldsymbol{\mu}$ is the mean, and $\boldsymbol{\sigma}$ is the standard deviation. For the validation dataset, only resizing and normalization are applied to ensure that the evaluation is performed on data with a consistent distribution.

Implementation Algorithm

The implementation algorithm involves using two state-of-the-art deep learning architectures: ResNet18 and DenseNet121, both of which are pretrained on large-scale datasets like ImageNet [16]–[18]. Each model is modified to accommodate the four-class output required for brain tumor classification by replacing the fully-connected layer in ResNet18 and the classifier layer in DenseNet121 [19], [20]. During the training phase, each model is trained separately using supervised learning [21]–[23]. The forward pass generates predictions which are compared against the true labels using the Cross-Entropy Loss function, mathematically defined as:

$$L = -\sum_{c=1}^{C} y_c \cdot \log(\hat{y}_c) \tag{1}$$

Where y_c is the one-hot encoded true label for class c and $\hat{y}_c$ is the predicted probability for that class. The backpropagation algorithm is then employed to compute gradients, and the model weights are updated using the Adam optimizer with a learning rate of 0.001. The performance is monitored on both the training and validation datasets, and the model achieving the highest validation accuracy is saved as the optimal version. After training, an ensemble evaluation is performed. In this stage, both models produce logits for each sample in the validation set, which are converted into probability distributions using the softmax function:

$$softmax(z)_i = \frac{e^{zi}}{\sum_{j=1}^{C} e^{zj}} \tag{1}$$

The probabilities from both models are then averaged to generate the final prediction, following the equation:

$$\hat{y}_{ensemble} = \frac{1}{M}\sum_{i=1}^{M} \hat{y}_i \tag{1}$$

where M is the number of models (in this case, $M = 2$). This ensemble method aims to leverage the strengths of both models, enhancing overall prediction accuracy and robustness.

Data Analysis Method

The data analysis method involves evaluating the performance of the models primarily through the metric of accuracy [24]. Accuracy is calculated by comparing the number of correct predictions to the total number of samples evaluated:

$$Accuarcy = \frac{Number\ of\ Correct\ Predictions}{Total\ Number\ of\ Samples} \tag{2}$$

Throughout the training process, the models are regularly evaluated on the validation set to monitor their progress. After applying the ensemble technique, accuracy is computed once more based on the averaged probability predictions. This methodical evaluation ensures that the final model not only achieves high accuracy but also demonstrates stable and reliable performance [25], making it suitable for practical clinical applications in brain tumor detection. Each of these stages is designed to ensure high-quality data input, optimal model training, and robust evaluation, collectively supporting the research goal of developing an effective automated diagnostic system for brain tumor detection using MRI images.

3. Results and Discussion

Results

In this study, the training outcomes of two deep learning models—ResNet18 and DenseNet121—were analyzed over a 10-epoch training process, and an ensemble approach was subsequently applied to further enhance performance. The results of each epoch are summarized in **Table 1** and **Table 2**. **Table 1** presents the training summary for ResNet18. The model began with a training loss of 0.3559 and an accuracy of 87.68% in the first epoch, with a corresponding validation accuracy of 77.01%. Notably, the best validation accuracy of 97.72% was achieved during epoch 7. Although the training accuracy continued to improve, reaching 99.02% in the final epoch, the validation accuracy slightly decreased to 96.44% by epoch 10.

Table 1. Training Progress for ResNet18

Epoch	Loss	Training Accuracy	Validation Accuracy
1	0.3559	87.68%	77.01%
2	0.1917	93.29%	82.06%
3	0.1371	95.46%	80.43%
4	0.1316	95.64%	95.66%
5	0.0992	96.51%	95.94%
6	0.0758	97.44%	93.17%
7	0.0684	97.72%	97.72%
8	0.0764	97.51%	96.58%
9	0.0701	97.67%	92.53%
10	0.0339	99.02%	96.44%

Table 2 summarizes the training progress for DenseNet121. This model started with a training loss of 0.3181 and an accuracy of 88.34% in the first epoch, with a high validation accuracy of 94.73% right from the start. Throughout the training, the model demonstrated a steady decrease in training loss—from 0.3181 down to 0.0608—and achieved its best validation accuracy of 97.79% in epoch 10.

Table 2. Training Progress for DenseNet121

Epoch	Loss	Training Accuracy	Validation Accuracy
1	0.3181	88.34%	94.73%
2	0.1873	93.57%	92.46%
3	0.1479	94.75%	95.59%
4	0.1318	95.34%	94.09%
5	0.1	96.42%	97.15%
6	0.0961	96.69%	97.58%

Epoch	Loss	Training Accuracy	Validation Accuracy
7	0.0848	97.22%	94.73%
8	0.0656	97.74%	96.44%
9	0.0773	97.26%	93.52%
10	0.0608	98.08%	97.79%

Following the individual model training, an ensemble evaluation was conducted. This approach involved averaging the probability outputs (after applying the softmax function) from both models. The ensemble method significantly improved the overall performance, achieving a validation accuracy of 99.36%. This result indicates that combining the strengths of ResNet18 and DenseNet121 can substantially enhance the reliability and accuracy of brain tumor classification.

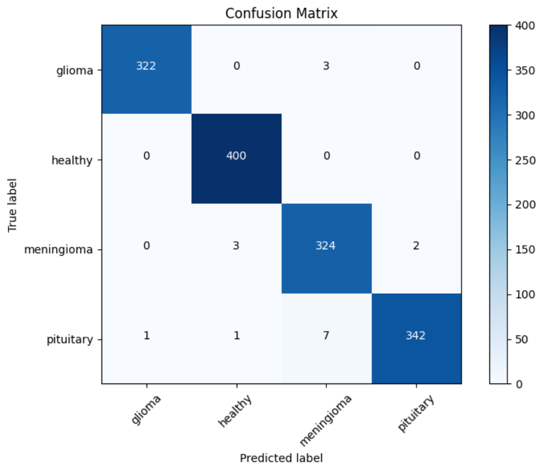

Figure 3. Confusion Matrix

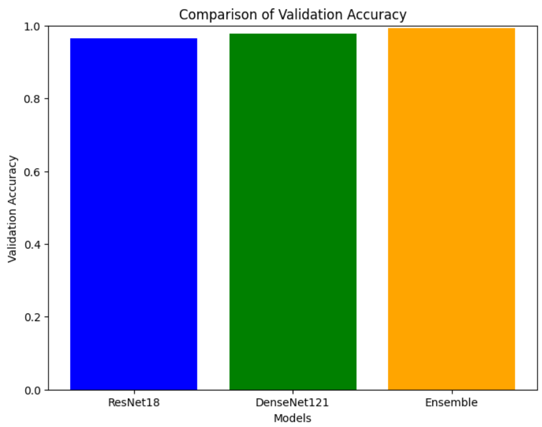

Figure 4. Comparison of Validation Accuracy

The overall performance is further illustrated by the visualizations. For instance, **Figure 3** displays the confusion matrix, where the majority of samples are correctly classified across the four categories. The confusion matrix highlights excellent performance for the healthy class, with 400 samples correctly identified, and similarly high accuracy for the other tumor classes. **Figure 4** provides a comparative analysis of the validation accuracy across ResNet18, DenseNet121, and the ensemble method, clearly showing the superior performance of the ensemble approach. In addition, **Figures 5** and **6** depict the trends of training loss and validation accuracy for DenseNet121 and ResNet18, respectively, confirming that both models demonstrate a consistent decline in training loss and a generally increasing trend in validation accuracy over the epochs.

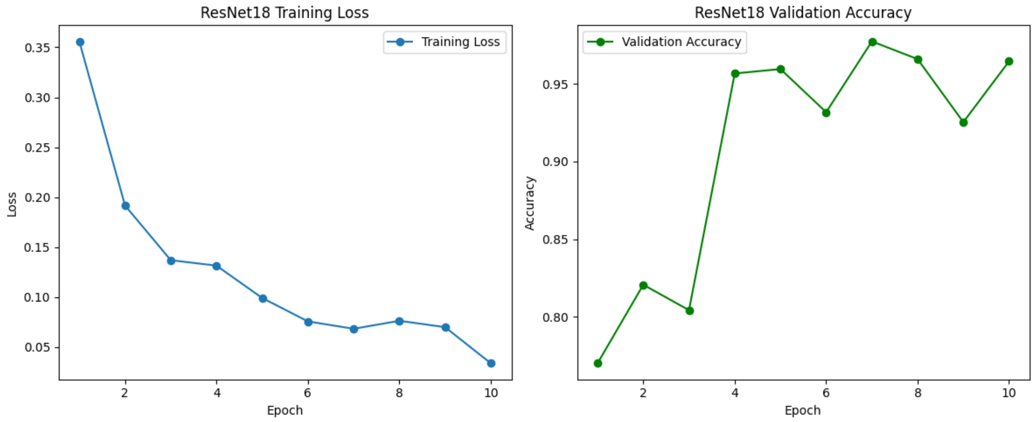

Figure 5. Training Loss and Validation Accuracy for ResNet18

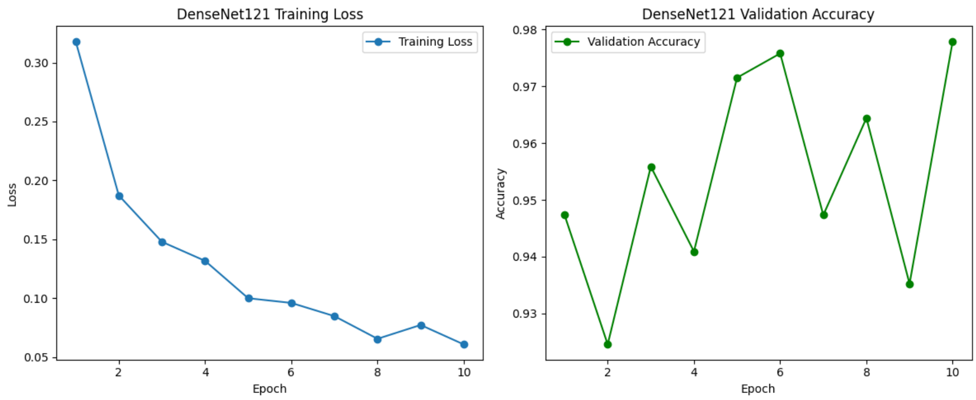

Figure 6. Training Loss and Validation Accuracy for DenseNet121

Discussion

The experimental results indicate that both ResNet18 and DenseNet121 are highly effective in classifying brain MRI images into four categories, with both models achieving validation accuracies above 97%. The steady reduction in training loss throughout the epochs for each model confirms that they are successfully learning and extracting relevant features from the data. However, while both models performed exceptionally well individually, the ensemble approach yielded the highest accuracy (99.36%), thereby illustrating the benefit of combining different model architectures. The improvement observed in the ensemble method is attributed to its ability to harness the complementary strengths of the two models. ResNet18, with its relatively simple feature extraction capabilities, and DenseNet121, known for its densely connected architecture that captures intricate features, when combined, provide a more robust and reliable classification outcome. This ensemble strategy effectively mitigates the misclassification errors that may occur when relying on a single model.

The confusion matrix (**Figure 3**) further supports these findings by demonstrating high accuracy in correctly classifying the samples, particularly for the healthy class and the various tumor categories. Minor misclassifications between similar tumor types were observed; however, the frequency of such errors is minimal compared to the overall sample count. The comparative analysis depicted in **Figure 4** clearly shows that the ensemble method outperforms the individual models, reinforcing the conclusion that ensemble learning is a promising approach for automated brain tumor detection. In summary, the results from this study suggest that an ensemble deep learning framework, which integrates ResNet18 and DenseNet121, can significantly enhance diagnostic accuracy in brain tumor classification using MRI images. Future work may focus on further refining the ensemble approach through additional model combinations, advanced data augmentation techniques, and more extensive hyperparameter tuning to further improve classification performance in clinical applications.

4. Conclusion

In conclusion, this study demonstrates the effectiveness of an ensemble deep learning approach that integrates ResNet18 and DenseNet121 for brain tumor classification from MRI images. Both models individually achieved high validation accuracies—97.72% for ResNet18 and 97.79% for DenseNet121—while the ensemble method further enhanced performance, reaching an impressive 99.36% accuracy. This result confirms that combining the strengths of different deep learning architectures can significantly improve classification reliability and overall diagnostic accuracy. The research also highlights that careful data pre-processing, proper model modification, and systematic training can lead to robust feature extraction and learning, even when dealing with complex medical images. The ensemble strategy, which averages the softmax probabilities from the individual models, effectively minimizes misclassification errors and demonstrates potential for clinical applications in automated brain tumor detection. Future research can build upon these findings by exploring additional model combinations, advanced data augmentation techniques, and extensive hyperparameter tuning to further optimize performance. Overall, the promising results of this study pave the way for the development of more efficient and reliable diagnostic systems in the field of medical imaging.

References:

[1] Jyotismita Chaki, *Brain Tumor MRI Dataset*. 2024, doi: 10.21227/1jny-g144.

[2] S. R. Kempanna *et al.*, "Revolutionizing brain tumor diagnoses: a ResNet18 and focal loss approach to magnetic resonance imaging-based classification in neuro-oncology," *Int. J. Electr. Comput. Eng.*, vol. 14, no. 6, p. 6551, Dec. 2024, doi: 10.11591/ijece.v14i6.pp6551-6559.

[3] Z. Zhu, M. Attique Khan, S.-H. Wang, and Y.-D. Zhang, "RBEBT: A ResNet-Based BA-ELM for Brain Tumor Classification," *Comput. Mater. Contin.*, vol. 74, no. 1, pp. 101–111, 2023, doi: 10.32604/cmc.2023.030790.

[4] S. Sal Sabila and H. Peni Agustin Tjahyaningtyas, "Multiple Brain Tumor with Modified DenseNet121 Architecture Using Brain MRI Images," *J. Ilm. Kursor*, vol. 12, no. 3, pp. 147–158, Jul. 2023, doi: 10.21107/kursor.v12i3.379.

[5] S. M. Shuvo *et al.*, "Multi-class Brain Tumor Classification with DenseNet-Based Deep Learning Features and Ensemble of Machine Learning Approaches," 2024, pp. 559–573.

[6] O. N. Belaid, M. Loudini, and A. Nakib, "Brain tumor classification using DenseNet and U-net convolutional neural networks," in *2024 8th International Conference on Image and Signal Processing and their Applications (ISPA)*, Apr. 2024, pp. 1–6, doi: 10.1109/ISPA59904.2024.10536704.

[7] 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.

[8] A. A. S. Ali, "Brain Tumor Classification Using A Hybrid Deep Learning Model: Leveraging Densenet121 And Inceptionv2 Architectures," *Electron. J. Univ. Aden Basic Appl. Sci.*, vol. 5, no. 4, pp. 455–463, Dec. 2024, doi: 10.47372/ejua-ba.2024.4.402.

[9] K. Gayatri, "Classification of Muti-Labeled Retinal Diseases in Retinal Fundus Images Using CNN Model ResNet18," *Lecture Notes in Networks and Systems*, vol. 1165. pp. 163–177, 2025, doi: 10.1007/978-981-97-8336-6_13.

[10] S. Sunkari, "A refined ResNet18 architecture with Swish activation function for Diabetic Retinopathy classification," *Biomed. Signal Process. Control*, vol. 88, 2024, doi: 10.1016/j.bspc.2023.105630.

[11] 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.

[12] 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.

[13] S. Vishwakarma, "Classification of Corn Leaf Disease using Resnet18, Alexnet and VGG16," *15th International Conference on Advances in Computing, Control, and Telecommunication Technologies, ACT 2024*, vol. 2. pp. 1333–1339, 2024.

[14] Z. Zheng, "Assessment of Damage Degree in High-Rise Building Structures Based on Continuous Wavelet Transform and Improved ResNet18," *IEEE International Conference on Signal, Information and Data Processing, ICSIDP 2024*. 2024, doi: 10.1109/ICSIDP62679.2024.10868831.

[15] 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.

[16] S. Tiwari, "Ensemble Deep Learning for DR Identification: Integrating DenseNet121 and VGG19 Architectures," *1st International Conference on Innovative Engineering Sciences and Technological Research, ICIESTR 2024 - Proceedings*. 2024, doi: 10.1109/ICIESTR60916.2024.10798174.

[17] S. S. Sadiq, "Improving CBIR Techniques with Deep Learning Approach: An Ensemble Method Using NASNetMobile, DenseNet121, and VGG12," *J. Robot. Control*, vol. 5, no. 3, pp. 863–874, 2024, doi: 10.18196/jrc.v5i3.21805.

[18] 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.

[19] 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.

[20] 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.

[21] A. Bajpai, "Enhanced Potato Leaf Disease Detection through Multi-Modal Fusion of Graph Neural Networks and ResNet18," *Proceedings - 2024 IEEE 16th International Conference on Communication Systems and Network Technologies, CICN 2024*. pp. 1498–1504, 2024, doi: 10.1109/CICN63059.2024.10847571.

[22] J. Naz, "A Comparative Analysis of Optimization Algorithms for Gastrointestinal Abnormalities Recognition and Classification Based on Ensemble XcepNet23 and ResNet18 Features," *Biomedicines*, vol. 11, no. 6, 2023, doi: 10.3390/biomedicines11061723.

[23] L. Ma, "Multi-Plant Disease Identification Based on Lightweight ResNet18 Model," *Agronomy*, vol. 13, no. 11, 2023, doi: 10.3390/agronomy13112702.

[24] G. Pallavi, "Brain tumor detection with high accuracy using random forest and comparing with thresholding method," *AIP Conference Proceedings*, vol. 2853, no. 1. 2024, doi: 10.1063/5.0198189.

[25] G. Pallavi, "Brain tumor detection with high accuracy using random forest and comparing with thresholding method," AIP Conference Proceedings, vol. 2853, no. 1. 2024, doi: 10.1063/5.0198189.