Asian Journal of Science, Engineering and Technology (AJSET)
eISSN: 3006-6980 (online)
https://ideapublishers.org/index.php/ajset
Vol. 5, No. 1 (January-December), 2026, pp. 83-100
https://doi.org/10.47264/idea.ajset/5.1.7
Original Research Article

EfficientNet-driven compound scaling for performance enhancement of CNNs in multimodal cancer imaging

Muhammad Hassan Farid*, Muhammad Adnan Nasim, Muhammad Abdullah Umar, Usman Javed

Faculty of Engineering Sciences and Technology, Hamdard University, Islamabad Campus, Pakistan.

*Corresponding Author Email: muhammadhassanfarid04@gmail.com

Article History

Received:
10-Jan-2026

Revised:
20-Feb-2026

Re-revised:
18-Mar-2026

Accepted:
23-Mar-2026

Published:
31-Mar-2026

Abstract:

Recent developments in deep learning-based medical image analysis have paved the way forward in improving the accuracy of cancer detection. In this study, we focused on enhancing the performance of different variants of existing Convolutional Neural Network (CNN) architectures by leveraging the compound scaling feature of EfficientNet. The EfficientNet architecture is known to scale the network width, resolution, and depth using a specific scaling factor. We integrated the compound scaling factor of EfficientNet-B2 to scale the network width and resolution of conventional models while preserving the default depth of all architectures. The proposed approach is evaluated on two distinct medical imaging tasks: lung cancer detection using Computed Tomography (CT) scans and brain tumor classification using Magnetic Resonance Imaging (MRI). Multiple CNN architectures, including different variants of VGG, ResNet, and DenseNet, are assessed in both baseline and scaled configurations. On the CT dataset, the proposed method achieves an accuracy improvement of up to 8.30%, while on the MRI dataset, improvements reach 12.95%. These improvements suggest that compound scaling is a suitable strategy to enhance the utilization of the existing conventional CNN architecture instead of replacing it.

Keywords: Convolutional neural networks, Deep learning, Magnetic resonance imaging, Computed tomography, Brain tumor detection, Lung cancer detection.

How to Cite: Farid, M. H., Nasim, M. A., Umar, M. A., & Javed, U. (2026). EfficientNet-driven compound scaling for performance enhancement of CNNs in multimodal cancer imaging. *Asian Journal of Science, Engineering and Technology (AJSET), 5*(1), 83-100. https://doi.org/10.47264/idea.ajset/5.1.7

1. Introduction

Cancer is proven to be one of the most lethal health complications, costing millions of lives each year. The precise and early detection of cancer is crucial to increase the patient's chances of survival. The medical imaging has served as an effective tool regarding early diagnosis in cancer detection by providing deep insights into critical physiological information (Kaur, *et al.* 2023). By capturing vital structural and functional information the medical image analysis serves as a reliable and comprehensive source of diagnostic data, essential for accurate disease identification (Rana, *et al.* 2023). Cancer detection and treatment have been transformed by a wide range of imaging techniques including Positron Emission Tomography (PET), Magnetic Resonance Imaging (MRI), Computed Tomography (CT), and chest radiographs (X rays) (Narvekar, *et al.* 2022). The X-rays offers limited two-dimensional viewing whereas CT scans generate three-dimensional cross-sectional images with rich details, offering comprehensive insights of the human body (Patel, *et al.* 2023) and hence providing better cancer detection in comparison to chest X-rays. The detailed anatomical information embedded in CT images facilitates the acquisition of some crucial information regarding tumor location, size and shape that enables deeper investigation and effective treatment of diseases. It even helps identify enlarged lymph nodes, which are a critical indicator of cancer metastasis (Guo, *et al.* 2024). In addition to computer-aided diagnosis (CAD) of cancer the CT images are also utilized in biopsy (Yang, *et al.* 2024).

In context of brain tumor detection, MRI has proven to be the most powerful and effective imaging modality (Martucci, *et al.* 2023), (Yan, *et al.* 2024). MRI is a non-invasive imaging modality that utilizes magnetic fields and radio waves, providing exceptional imaging clarity while being safer for patients (Teixeira, *et al.* 2025). The superior soft-tissue contrast in MR images enables precise delineation of brain structures, thereby facilitating early identification of tumors, edema, and other intracranial pathologies (Falk, 2025). This better diagnostic accuracy and safe operation, make MRI indispensable for evaluating complex neurological conditions. The medical image analysis has been revolutionized with the introduction of CNNs which demonstrates exceptional accuracy in cancer classification (Panayides, *et al.* 2020). In literature, mostly the studies focus on single imaging modality (CT or MRI) and there is a need to understand that how CNNs perform across different medical imaging types because each modality presents unique challenges, such as noise in CT scans and intensity variations in MRI (Li, *et al.* 2024). Addressing these crucial challenges is essential to reveal the full potential of deep learning in medical diagnosis. In medical image classification, the CNNs including: VGG (Simonyan, K., & Zisserman, A. 2015), (Alnaggar, *et al.* 2024), ResNet (He, *et al.* 2016), (Xu, *et al.* 2023), and DenseNet (Huang, *et al.* 2017), are established as benchmarks.

In this study, we employed EfficientNet's (Tan & Le, 2019) compound scaling principle on different variants of conventional CNNs which balances network width, and input resolution while keeping the default depths constant for each particular architecture. We have applied the EfficientNet-B2 scaling factor with increased filter width and higher input resolution on VGG-16, VGG-19, ResNet-50, ResNet-101, DenseNet-121 and DenseNet-169 while their default depths were kept constant. The performance is compared in terms of accuracy and other standard metrics for each model with its unscaled version. Overall, this study performs a head-to-head comparison of scaled vs unscaled CNN models to detect lung cancer from CT scans and brain tumors from MRI scans. The main objectives of this work are summarized as follows:

- An application of compound scaling strategy that jointly increases network width and input resolution for improved feature representation in medical imaging tasks.
- A comprehensive evaluation across multiple CNN architectures (VGG, ResNet, DenseNet) demonstrating architecture-sensitive performance improvements.
- Extensive experimental validation on both CT and MRI datasets, including multi class classification scenarios.
- Detailed analysis through convergence behavior, confusion matrices, and qualitative predictions to assess model robustness.

By addressing these objectives, this study provides valuable insights into optimizing existing CNN architectures for multimodal cancer image classification and demonstrates that significant performance gains can be achieved without designing entirely new network models.

2. Related work

Cancer detection and classification in medical imaging have undergone remarkable advancements due to the emergence of advanced deep learning techniques (Chen, *et al.* 2022), (Shamshad, *et al.* 2023). The CAD systems that leverage various CNNs for processing medical images, have shown optimal performance, leading to accurate diagnosis and classification results (Iqbal, *et al.* 2023). In contrast to the traditional machine learning techniques, CNNs are able to learn the intricate relationships between neighboring pixels, enabling remarkable precision in feature extraction (Alzubaidi, *et al.* 2021). In literature there are multiple implementations of CNN architectures for lung cancer detection using CT images. In (Pandian, *et al.* 2022), a VGG-16 based classifier was implemented for lung cancer diagnosis using image texture features. The proposed classifier achieved the classification accuracy of 83% for CT images. Likewise, in (Tian, *et al.* 2024), a ResNet-18 architecture is employed to address the challenge of enhancement in diagnostic accuracy for lung cancer detection achieving 98.84% accuracy. Similarly, DenseNet has set new benchmarks in cancer classification, delivering exceptional precision in medical imaging. In (Mohandass, *et al.* 2024), the lung nodule detection is demonstrated by using DenseNet-201 and Namib Beetle Optimization. The feature computation for the LUNA16 images is done using the Empirical Wavelet Transform, achieving 24.8% accuracy enhancement in comparison to deep belief networks. In recent years, researchers have worked out to push the limits of CNN performance, using EfficientNet scaling techniques to achieve better accuracy. EfficientNet, introduced by Tan & Le, (2019), revolutionized deep learning with compound scaling method that optimally balances three critical parameters: depth (layer count), width (channels per layer) and resolution (input image dimension). The EfficientNet's strategic scaling has been adapted in medical imaging to deliver improved classification accuracy. For instance, an EfficientNet for lung cancer classification, named Lung-EffNet is built by Raza, *et al.* (2023), based on transfer learning-based predictor. The model achieved 99.10% accuracy and trained using IQ-OTH/NCCD dataset.

Literature also reveals various CNN architectures utilized for classification of brain tumors in MR images. VGG and ResNet have consistently delivered effective performance in cancer diagnosis, redefining the utilization of AI in pathology detection capabilities. For instance, the VGG-16 model depicts promising results on MR datasets i.e., BRATS 2013, OPEN-I and BRATS 2015 for tumor image classification task and achieving 89.3% accuracy (Bairagi, *et al*. 2023). Similarly, Ulli, *et al.* (2024) modified VGG-19 CNN was trained on MRI dataset to learn brain tumor patterns achieving the classification accuracy of 95%, which is superior to VGG-16 and ResNet-50. In reference to ResNet architecture, a modified version of the pre-trained ResNet-18 is employed by El-Feshawy, *et al.* (2023), achieving peak accuracy of 98.67% in brain tumor detection task while outperforming the conventional CNNs. There are also examples of DenseNet based brain tumor detection in literature. For instance, a DenseNet-121 CNN with adaptive learning rates is utilized to address the low accuracy and lack of generalization problems in brain tumor detection. The model achieved a training accuracy of 98.81% and a test accuracy of 98.09%, outperforming current methods (Venkatachalam, *et al.* 2025). Similarly, in (Abraham, *et al.* (2025) employed a DenseNet-169 CNN for brain tumor detection using MRI Kaggle dataset of brain tumor achieving accuracy of 98.78%.

All these previous studies suggest that limited attention has been given to evaluating how architectural scaling strategies influence the performance of conventional CNNs across heterogeneous imaging modalities. Moreover, the compound scaling principle of EfficientNet have rarely been systematically applied to enhance existing CNN architectures without modifying network depth. The focus of our work is to address this gap by integrating the EfficientNet-driven compound scaling into well-known CNN architectures including: VGG, ResNet, and DenseNet for performance improvement.

The remainder of this paper is organized as follows. Section 2 describes the proposed methodology, including the compound scaling strategy and model configurations. Section 3 presents the experimental results and discussion for both CT and MRI datasets. Finally, Section 4 concludes the study.

3. Materials and methods

This section presents the datasets, model scaling and architecture modifications using EfficientNet-driven compound scaling strategy adopted in this study. The proposed framework is designed to systematically evaluate the effect of compound scaling on conventional CNN architectures across heterogeneous medical imaging modalities.

3.1. Dataset

In this study, we used two datasets of different imaging modalities (CT and MRI) for lung cancer and brain tumor classification. For lung cancer classification, the LIDC-IDRI dataset (Armato III, *et al.* 2011) developed by the National Cancer Institute for early cancer detection is used. It contains 1,018 cases with CT images. From this dataset collection, we selected a total of 5,000 CT images, that contains 2,000 benign and 3,000 malignant samples.

For brain tumor detection, we constructed a composite MRI dataset by combining images from three publicly available sources: Figshare (Cheng, J. 2017), Kaggle (Nickparvar, 2021), and Kaggle (Bhuvaji, *et al.* 2020). This resulting composite MRI dataset includes a total of 13,351 images which are categorized into four different classes including: glioma containing 3,973 images, meningioma containing 3,290 images, pituitary tumor containing 3,588 images, and no tumor containing 2,500 images.

Further, in order to meet the scaling factor of EfficientNet-B2, we used uniform input dimensions across both CT and MRI datasets and the dataset was partitioned into 80% training and 20% validation subsets. Additionally, we have normalized both the LIDC-IDRI lung cancer dataset and the brain tumors dataset to standardize the pixel intensities.

3.2. Model scaling and architecture modifications

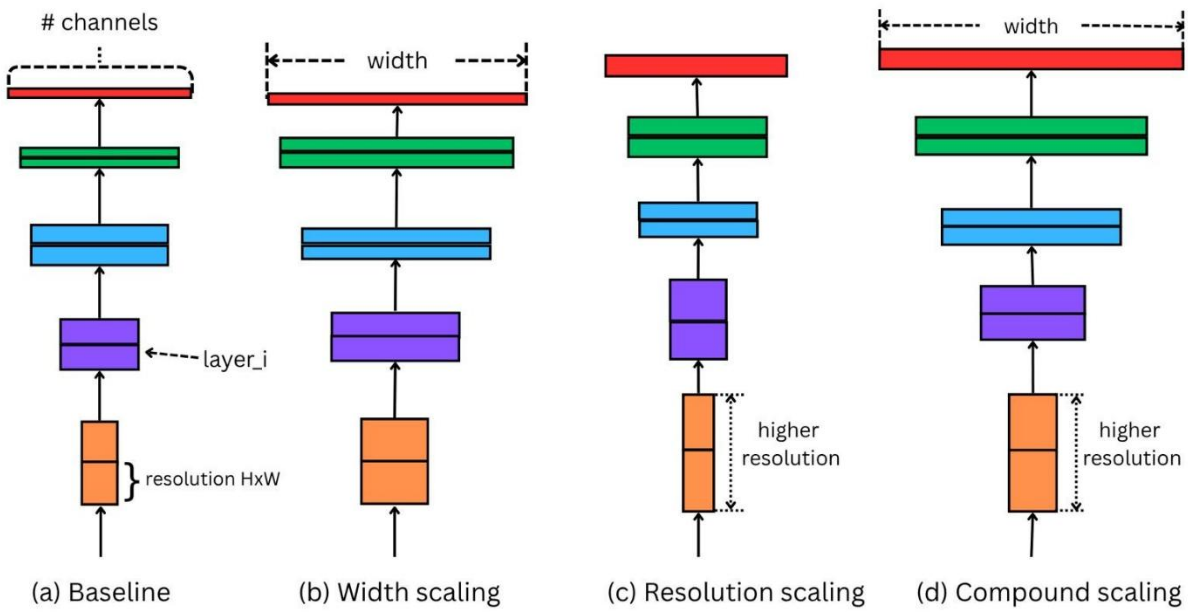
This study compares unscaled and scaled versions using EfficientNet's (Tan & Le, 2019) compound scaling method of conventional CNNs, including VGG-16, VGG-19, ResNet-50, ResNet-101, DenseNet-121, and DenseNet-169. Scaled models are optimized by adjusting width and input resolution as shown in Figure 1; depths of all models are kept constant, as we are using the standard depth models of VGG, ResNet and DenseNet. Width and input resolution scaling factors are defined by Eq.1 and 2.

$$w' = w\,.\,\beta^{\phi} \tag{1}$$

$$r' = r\,.\,\gamma^{\phi} \tag{2}$$

Where w is original width (number of filters), r is original input resolution, β is width scaling factor, γ is resolution scaling factor, ϕ is compound coefficient and w', r' are scaled width and resolution, respectively.

The β = 1.1, γ = 1.15 and ϕ = 2 as explained for EfficentNet-B2. Therefore, width w' grows by factor of ≈ 1.21 and resolution r' grows by factor of ≈ 1.32. Since filter counts must be integers, the scaled width is rounded to the nearest integer after applying Eq. 1. For VGG-16 and VGG-19, this means that convolutional layers originally having 64, 128, 256, and 512 filters are scaled to approximately 77, 155, 310, and 620 filters, respectively across their blocks. In ResNet-50 and ResNet-101, scaling is applied stage-wise: the canonical stage widths of 64, 128, 256, and 512 filters become 77, 155, 310, and 620 filters, which are then used consistently for all residual blocks in each stage while keeping the number of blocks unchanged. For DenseNet-121 and DenseNet-169, scaling is reflected in the growth rate parameter, where the standard growth rate of 32 is scaled to ≈ 39.



In all cases, depth remains constant, but filter widths and growth rates are expanded according to EfficientNet-B2 scaling factors, ensuring structural consistency with the original networks while increasing representational capacity.

3.3. Models training and evaluation

Each CNN model (VGG-16, VGG-19, ResNet-50, ResNet-101, DenseNet121 and DenseNet-169) in our study is trained for 20 Epochs using Adam optimizer and categorical cross-entropy loss function which is suitable for multi-class classification. The dataset for each modality (CT and MRI) is divided into 80% training and 20% validation.

The Binary Cross Entropy (BCE) was employed as loss function in lung cancer classification problem task due to its binary nature. The average binary cross-entropy loss over N samples is given by Eq. 3:

$$L_{BCE} = -\frac{1}{N}\sum_{i=1}^{N}[\, y_i log(p_i) + (1 - y_i) log(1 - p_i)] \qquad (3)$$

Where N is the Total number of samples, y_i is the True label of the i-th sample ($y_i \in \{0, 1\}$) and p_i is the predicted probability that the i-th sample belongs to the positive class ($0 \le p_i \le 1$). The categorical cross entropy was used as loss function for brain tumor classification which is a multi-class problem. The categorical cross-entropy loss for a dataset with N samples and C classes is given by Eq. 4:

$$L_{CCE} = -\frac{1}{N}\sum_{i=1}^{N}\sum_{k=1}^{C} y_{ik} log(p_{ik}) \qquad (4)$$

Where N is Total number of samples, C is the Total number of classes, y_{ik} is True label for class C of the i-th sample (one-hot encoded: $y_{ik} \in \{0, 1\}$), p_{ik} is Predicted probability that the i-th sample belongs to class C ($0 \le p_{ik} \le 1$).

Model performance was evaluated using accuracy, precision, recall, specificity, and F1 score. For binary classification, these metrics were computed based on true positive (TP), true negative (TN), false positive (FP), and false negative (FN) counts.

4. Results and discussion

This section presents a comprehensive evaluation of the proposed EfficientNet-driven compound scaling strategy across two medical imaging tasks: lung cancer detection using CT scans using LIDC-IDRI dataset and brain tumor classification using MRI merged dataset from multiple sources. The evaluation focuses on quantitative performance comparison, convergence behavior, confusion matrix analysis, and qualitative prediction assessment.

4.1. Lung Cancer Detection (CT scans)

For the lung cancer detection, models were trained using CT scans dataset (LIDC-IDRI) and evaluated based on their ability to distinguish between cancerous and non-cancerous lung tissues. The performance metrics in case of lung cancer detection with all the CNN architectures used in our experimentation with their scaled and unscaled versions are summarized in Table 1.

Table 1:Performance Metrics for Lung Cancer Detection Models

Model	Accuracy	Scaled-Accuracy	Precision	Scaled-Precision	Recall	Scaled-Recall	F1 Score	Scaled-F1 Score
VGG-16	98.56%	99.1 4%	98.64%	99.43%	98.97%	99.13%	98.80%	99.28%
VGG-19	98.56%	99.00%	98.51%	99.36%	99.10%	98.97%	98.80%	99.16%
ResNet-50	90.70%	99.00%	90.64%	99.30%	94.23%	99.03%	92.40%	99.17%
ResNet-101	97.72%	98.56%	97.10%	98.57%	99.17%	99.03%	98.12%	98.80%
DenseNet-121	98.84%	99.42%	98.61%	99.73%	99.47%	99.30%	99.04%	99.52%
DenseNet-169	99.00%	99.56%	99.56%	99.83%	98.77%	99.43%	99.16%	99.63%

The results demonstrate that the scaled DenseNet-169 shows strong performance with 99.56% accuracy, while the scaled Densenet-121 partially lags at 99.42%. The other conventional models used in the experimentation like VGG-16, VGG-19, ResNet-50 and ResNet-101 shows consistent improvement in comparison to their respective unscaled versions by demonstrating accuracy above 98%. The comparison of training and validation accuracies of VGG-16 with scaled and unscaled versions is represented in Figure 2(a). The graph shows that the scaled version of VGG-16 converges gradually and demonstrates comparable performance with its unscaled version with marginal improvement in accuracy. Both the unscaled and scaled versions ultimately achieved higher accuracies. Similarly, for VGG-19, there is minimal improvement in validation accuracy as indicated by the graph in Figure 2(b) and Table 1. Both the scaled and unscaled versions of VGG-19 curves are closely aligned, indicating marginal improvement with compound scaling.

For ResNet-50, the scaled version shows a clear advantage over the unscaled version. The scaled ResNet-50 demonstrated the improvement in accuracy with approximately 8.30% as represented by the graph in Figure 2(c). The validation accuracy of unscaled ResNet-50 is significantly lower than the scaled version, indicating poor generalization. For ResNet-101, the training and validation accuracies of unscaled versions are not consistent specially in the middle epochs as indicated by the graph in Figure 2(d). However, the validation accuracy of the scaled ResNet-101 demonstrates consistent improvement and ultimately showing close alignment with the training curve specially in the later epochs. Overall, the graph of ResNet-101 indicates that with compound scaling the model demonstrates stable training and achieves higher validation accuracy.

For DenseNet-121, the unscaled model shows rapid improvement in training accuracy, but its validation accuracy lags significantly which is due to the overfitting as represented by the graph in Figure 2(e). In contrast, the scaled DenseNet-121 shows steady improvement in both training and validation accuracy with a smaller gap between the two curves. Therefore, the impact of scaling in DenseNet-121 is smoother convergence, reduced overfitting and delivering superior validation performance compared to the unscaled version. For DenseNet-169, both the scaled and unscaled models achieve high accuracy, but with some differences in training behavior as represented in the graph of Figure 2(f). Unlike the performance of DenseNet-121, the unscaled DenseNet-169 already generalizes well, with validation accuracy staying close to training accuracy. Overall, the benefit of scaling is less significant in DenseNet-169 than in DenseNet-121, but it still suggests stronger generalization and consistency.

4.2. Brain Tumor Detection (MRI Scans)

For brain tumor detection we used a composite MRI dataset as mentioned in the dataset section. The models are evaluated based on their ability to classify images into categories of glioma, meningioma, pituitary tumor and non-tumor. Table 2 summarizes the performance metrics of brain tumor detection for all CNN architectures used in our study. The results shown in Table 2, demonstrate that the scaled DenseNet-169 excels with 95.82% accuracy, indicating strong performance, while scaled Densenet-121 lags at 94.86%. Other models like VGG-16, VGG-19, ResNet-50 and ResNet-101 show consistent improvements in comparison with their respective unscaled versions. The ResNet-50 demonstrated the highest improvement from its unscaled version with the difference of 12.95%.

The graph in Figure 3(a) shows the comparison of training and validation accuracies of VGG-16 with scaled and unscaled versions. The scaled VGG-16 converges gradually and demonstrates improved training and validation accuracy in comparison with the unscaled version. However, when we compared the scaled and unscaled versions of VGG-19, the difference between the curves is more noticeable as represented by the graph in Figure 3(b). The scaled version of VGG-19 is able to converge faster and achieved higher accuracy as clearly indicated by the validation curve of scaled VGG-19 that consistently stays above its unscaled version. Hence in comparison to VGG-16, the compound scaling enabled VGG-19 shows significant improvement, ensuring more stable validation performance.

In case of ResNet architecture we have performed the experimentation on its two variants ResNet-50 and ResNet-101. The scaled ResNet-50 demonstrated maximum improvement in accuracy with approximately 12.95% as represented by the graph in Figure 3(c). The scaled version of ResNet-50 demonstrates a faster and more stable convergence with both training and validation accuracy improving steeply. The validation curve of the scaled model closely follows its training curve, indicating good generalization and reduced overfitting. However, for unscaled ResNet-50 there is a significantly large gap between the training and validation curves, indicating poor generalization in comparison with the scaled version. For the case of ResNet-101 there is marginal difference between the scaled and unscaled versions as represented by the graph in Figure 3(d). The validation curves of scaled version show marginal improvement specially in the later Epochs in comparison with the unscaled version of ResNet-101. Overall, for ResNet-101 the compound scaling shows a positive impact on performance by ultimately achieving higher accuracies.

Table 2: Performance Metrics for Brain Tumor Detection Models.

Model	Accuracy	Scaled Accuracy	Precision	Scaled Precision	Recall	Scaled Recall	F1 Score	Scaled F1 Score
VGG-16	93.57%	94.22%	93.60%	94.32%	93.57%	94.22%	93.54%	94.19%
VGG-19	90.52%	93.63%	90.60%	93.68%	90.52%	93.63%	90.51 %	93.62%
ResNet-50	79.02%	91.97%	79.10%	92.00%	79.02%	91.97%	79.05%	91.97%
ResNet-101	87.69%	88.28%	87.90%	88.29%	87.69%	88.28%	87.70%	88.25%
DenseNet-121	91.76%	94.86%	91.87%	94.92%	91.76%	94.86%	91.72%	94.84%
DenseNet-169	90.74%	95.82%	90.81 %	95.82%	90.74%	95.82%	90.73%	95.78%

The DenseNet-121 accuracy curves are presented in the graph of Figure 3(e). The graph shows that the validation curve of scaled DenseNet-121 ultimately demonstrates higher accuracy than the validation curve of unscaled version. The unscaled version achieves competitive results but still lags behind in validation accuracy, indicating less efficient learning. The accuracy progression of scaled and unscaled DenseNet-169 models is represented in the graph of Figure 3(f). The scaled DenseNet-169 achieves rapid convergence in both training and validation curves and continue to improve steadily. In contrast, the unscaled DenseNet-169 training starts with lower accuracy and requires more epochs to catch up. Its training accuracy gradually increases, eventually reaching close to the scaled model. The unscaled validation accuracy also improves consistently yet remains below that of the scaled version at the end.

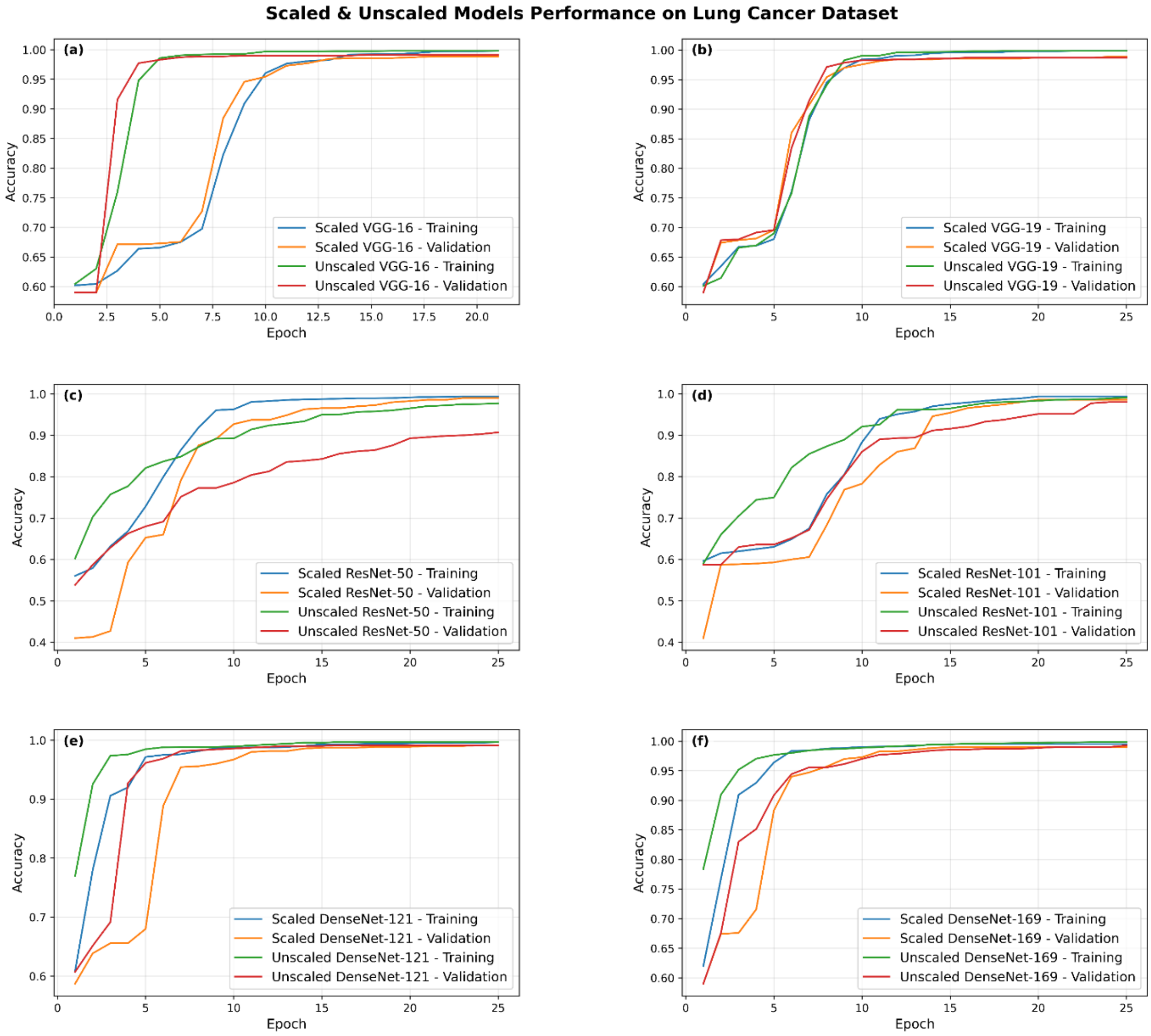


Figure 2: Scaled and Unscaled Model Performance on Lung Cancer Dataset in Terms of Accuracy (a) VGG-16 (b) VGG-19 (c) ResNet-50 (d) ResNet-101 (e) Densenet-121 (f) Densenet-169.

All the experimental results presented in Figure 2 and 3 supports the fact that the compound scaling feature of EfficientNet can be a useful tool to enhance the performance of different variants of well-known CNN architectures. The application of compound scaling in our study for network width and resolution facilitates enhanced feature extraction. In conventional

CNNs, the performance improvement is normally achieved through empirically tuning the hyper parameters. In contrast, the EfficientNet relies on the compound scaling factor for improving the overall performance of the network and hence provides a suitable alternative to enhance the performance of conventional architectures.

4.3. Predictions

Figure 4 and Figure 5 visually show some sample outputs for the scaled and unscaled versions of each CNN model (VGG-16, VGG-19, ResNet-50, ResNet-101, DenseNet-121, DenseNet-169) in case of lung cancer images using CT scans and brain tumor images using MRI, respectively. The lung cancer detection performance of unscaled CNN models in terms of confusion matrices is depicted in Figure 6. The confusion matrices for the scaled CNN models for lung cancer detection are presented in Figure 7. Likewise, the confusion matrices comparison of unscaled and scaled CNN models in brain tumor detection are presented in Figure 8 and 9, respectively.

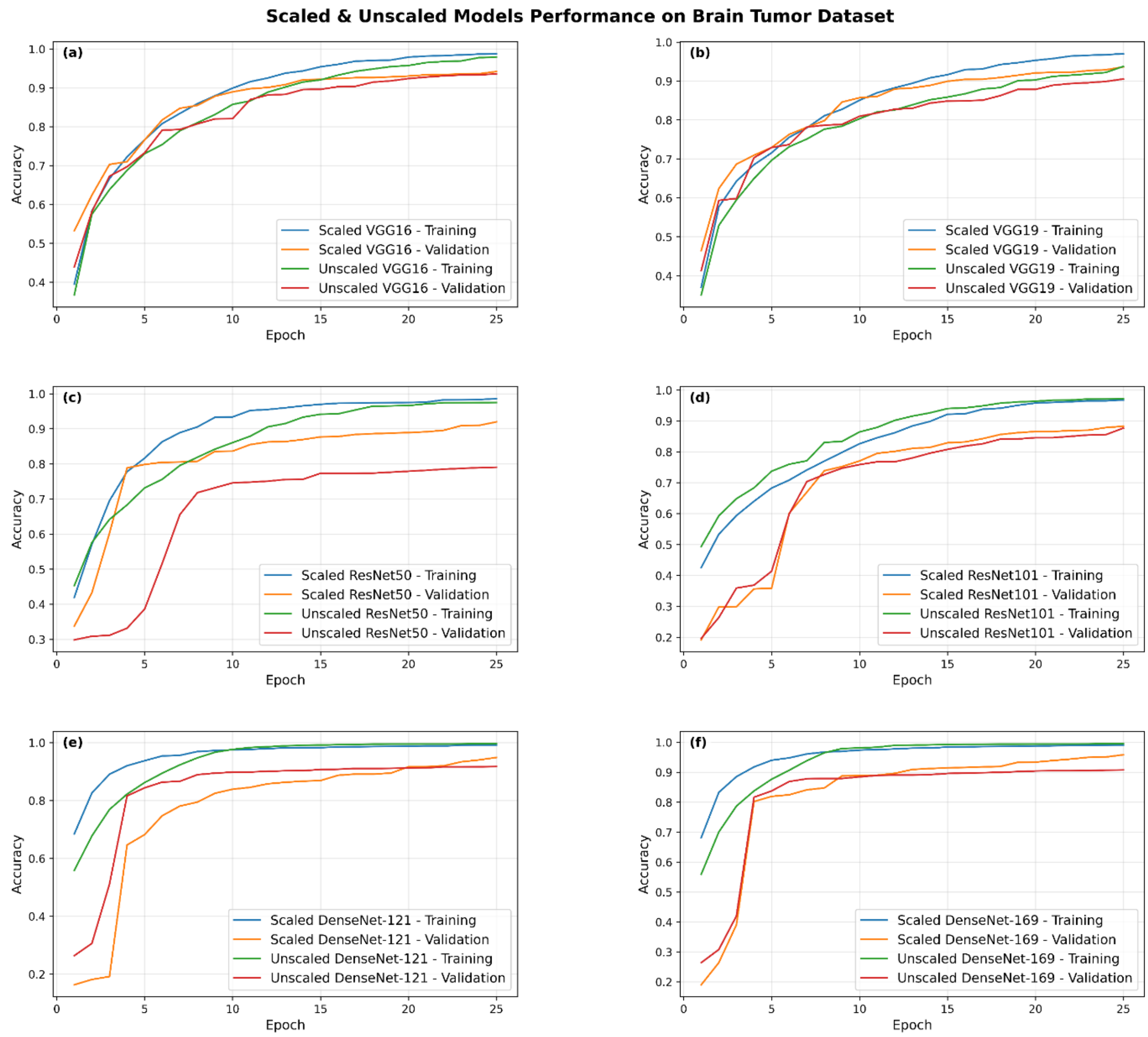


Figure 3: Scaled and Unscaled Model Performance on Brain Tumor Dataset in terms of Accuracy (a) VGG-16 (b) VGG-19 (c) ResNet-50 (d) ResNet-101 (e) DenseNet-121 (f) DenseNet-169.

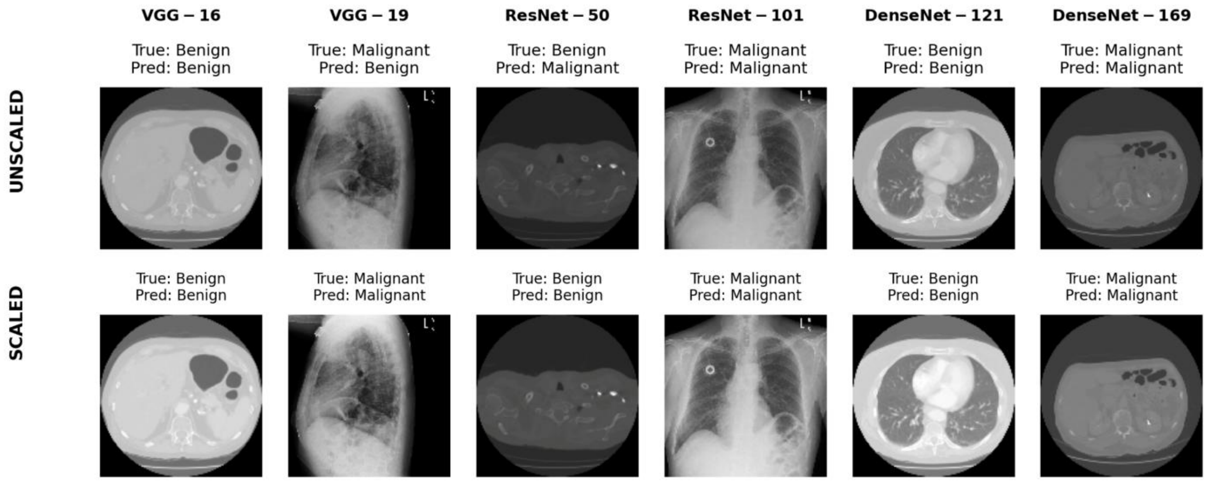


Figure 4: Lung Cancer Detection Predictions for CT Scan Dataset.

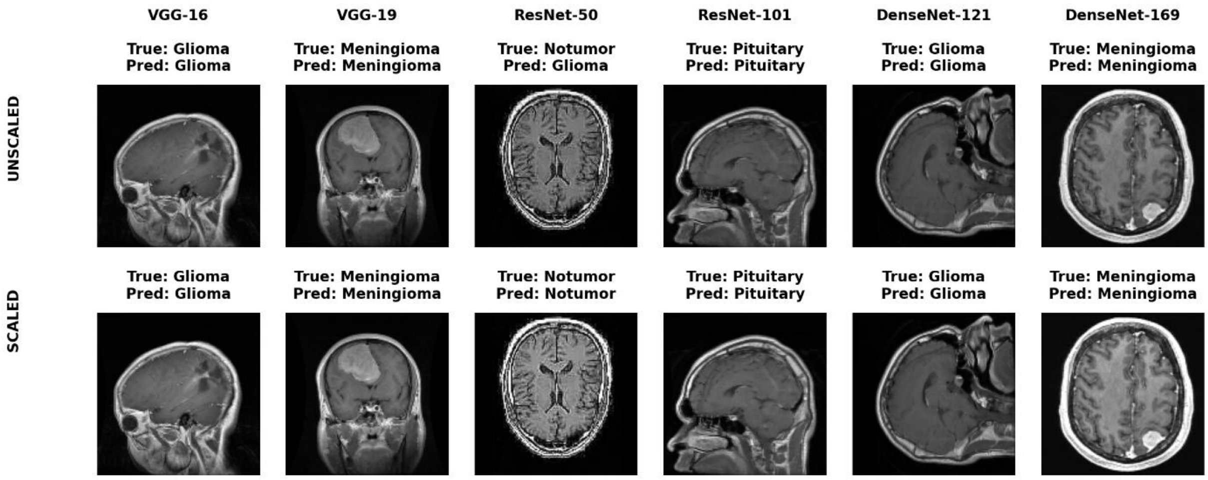


Figure 5: Brain Tumor Detection Predictions for MRI Dataset

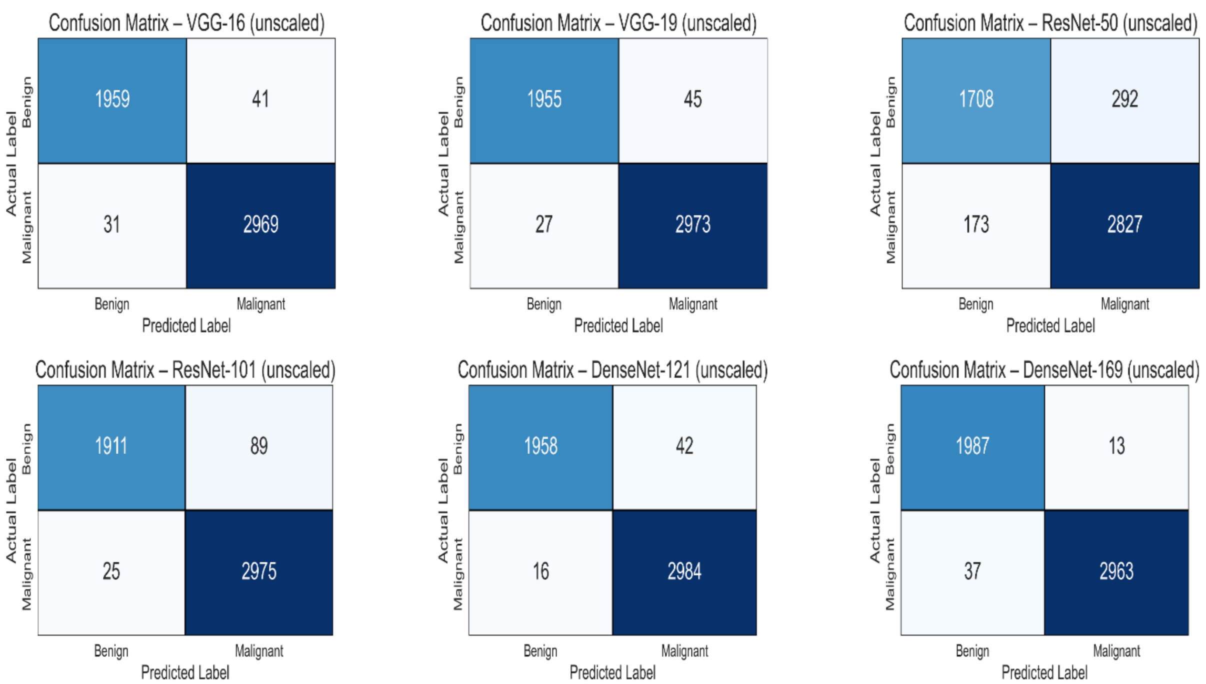


Figure 6: Confusion Matrices for Lung Cancer Detection Using Unscaled CNN.

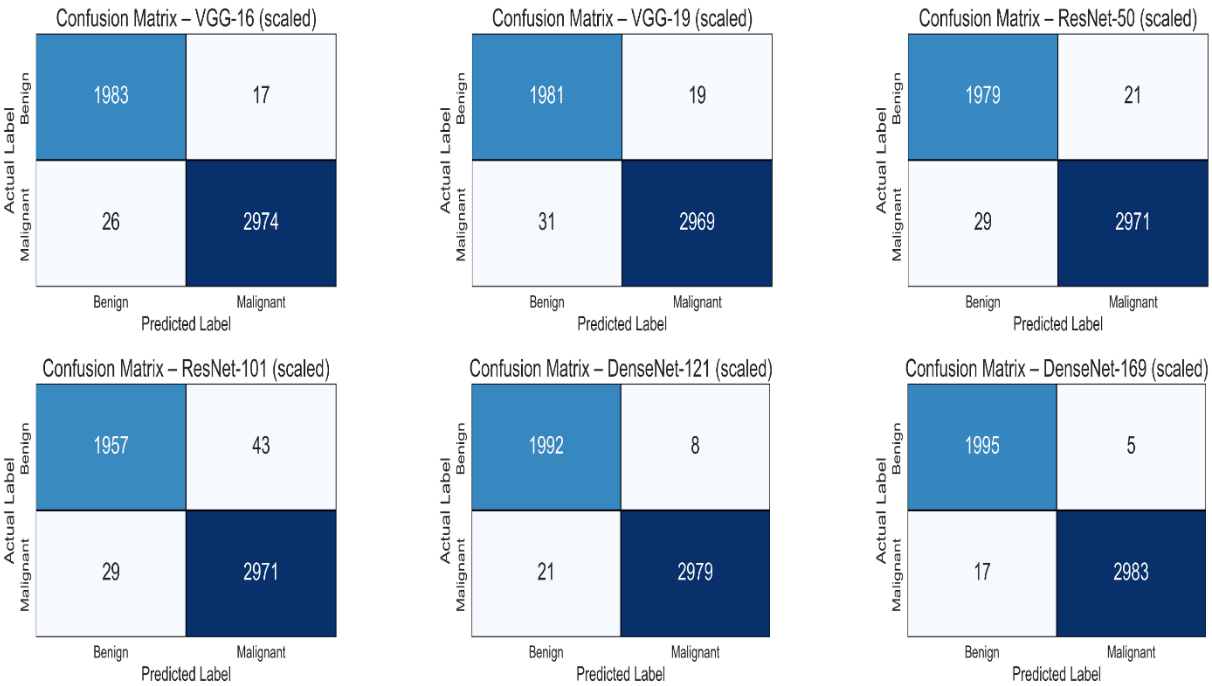


Figure 7: Confusion Matrices for Lung Cancer Detection Using Scaled CNN.

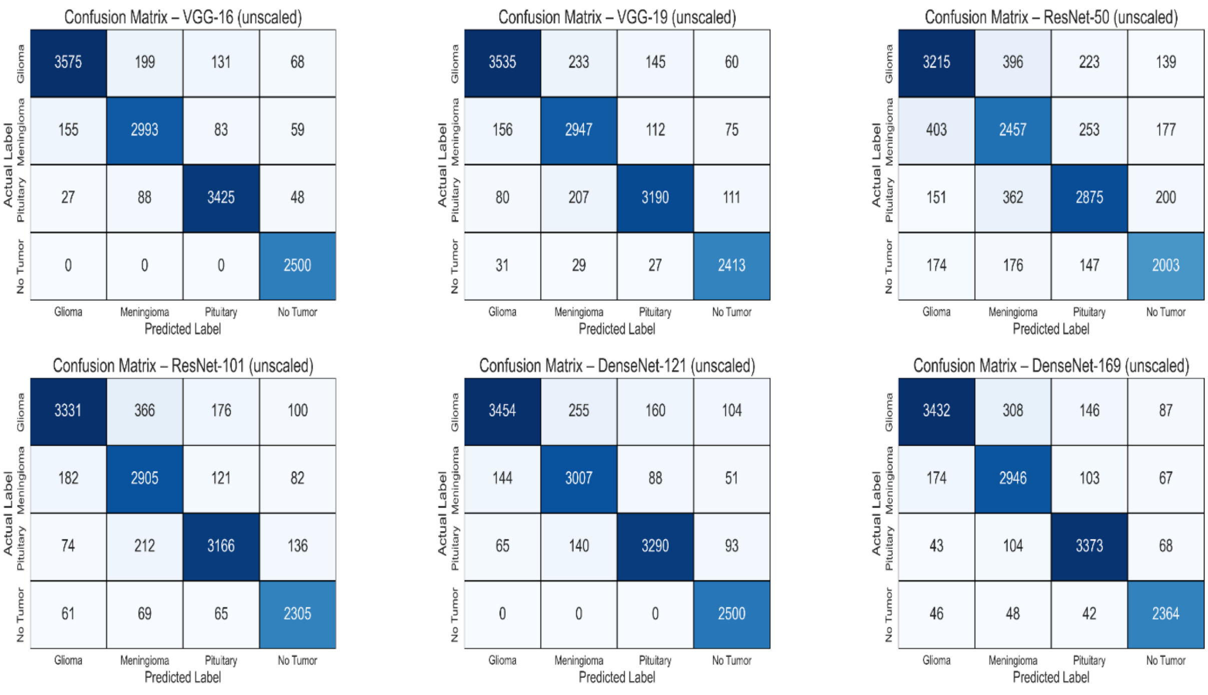


Figure 8: Confusion Matrices for Brain Tumor Detection Using Unscaled CNN.

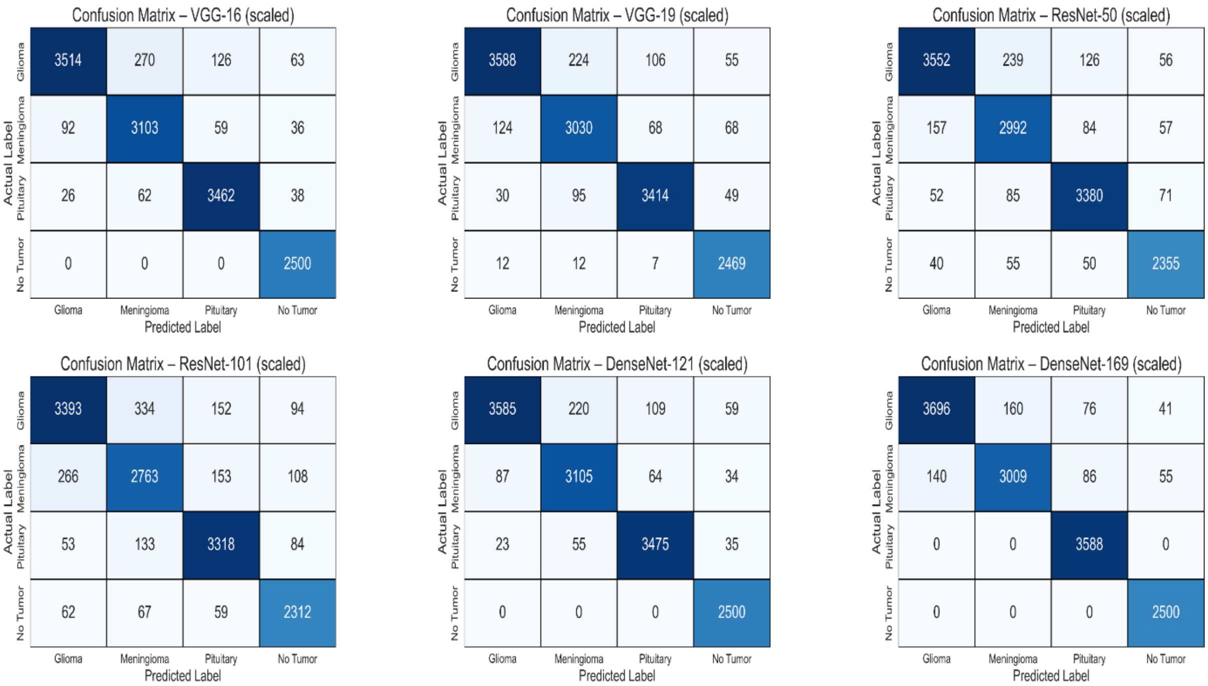


Figure 9: Confusion Matrices for Brain Tumor Detection Using Scaled CNN.

5. Conclusion

This study investigated the effectiveness of EfficientNet-driven compound scaling for improving the performance of conventional CNN architectures in medical image classification. Instead of designing entirely new deep learning architectures, the proposed approach enhances existing CNN models by proportionally scaling network width and input resolution based on EfficientNet scaling principles, this strategy allows conventional networks to achieve improved performance while preserving the original depths of CNN architectures. The proposed approach is evaluated using two heterogeneous medical imaging datasets: lung cancer detection using CT scans and brain tumor detection using MRI images. The commonly used CNN architectures including: VGG-16, VGG-19, ResNet-50, ResNet-101, DenseNet-121 and DenseNet-169 were analyzed in both their baseline and scaled configurations. The experimental results demonstrate that on the lung CT dataset, the scaled models achieved substantial gains in classification accuracy with the scaled DenseNet-169 model attaining the highest performance of 99.56%. Similarly, in the MRI-based brain tumor detection task, compound scaling significantly improved performance in the more challenging multi-class setting, where the scaled DenseNet-169 achieved an accuracy of 95.82%. The results further showed that scaled architectures exhibit improved convergence stability, reduced misclassification rates and stronger generalization behavior across both datasets. The overall performance improvements confirm that the EfficientNet inspired scaling enhances feature extraction capability without introducing optimization instability, enabling a simple yet effective mechanism for enhancing the performance of conventional CNN architectures in medical image analysis. As a future work, the proposed framework can be extended to additional medical imaging modalities and investigating hybrid scaling strategies that incorporate attention mechanisms to further improve diagnostic performance.

Declaration of conflict of interest

The author(s) declared no potential conflicts of interest(s) with respect to the research, authorship, and/or publication of this article.

Funding

The author(s) received no financial support for the research, authorship and/or publication of this article.

Publisher’s Note

IDEA PUBLISHERS (AJSET) stands neutral with regard to the jurisdictional claims in the published maps and the institutional affiliations.

References

Abraham, L. A., Palanisamy, G., & Veerapu, G. (2025). Transparent brain tumor detection using DenseNet169 and LIME. *Scientific Reports, 15*(1), 28185. https://doi.org/10.1038/s41598-025-13233-7

Alnaggar, O. A. M. F., Jagadale, B. N., Saif, M. A. N., Ghaleb, O. A., Ahmed, A. A., Aqlan, H. A. A., & Al-Ariki, H. D. E. (2024). Efficient artificial intelligence approaches for medical image processing in healthcare: Comprehensive review, taxonomy, and analysis. *Artificial Intelligence Review, 57*(8), 221. https://doi.org/10.1007/s10462-024-10814-2

Alzubaidi, L., Zhang, J., Humaidi, A. J., Al-Dujaili, A., Duan, Y., Al-Shamma, O., ... Farhan, L. (2021). Review of deep learning: Concepts, CNN architectures, challenges, applications, future directions. *Journal of Big Data, 8*(1), 53. https://doi.org/10.1186/s40537-021-00444-8

Armato, S. G., III, McLennan, G., Bidaut, L., McNitt-Gray, M. F., Meyer, C. R., Reeves, A. P., ... Clarke, L. P. (2011). The Lung Image Database Consortium (LIDC) and Image Database Resource Initiative (IDRI): A completed reference database of lung nodules on CT scans. *Medical Physics, 38*(2), 915–931. https://doi.org/10.1118/1.3528204

Bairagi, V. K., Gumaste, P. P., Rajput, S. H., & Chethan, K. S. (2023). Automatic brain tumor detection using CNN transfer learning approach. *Medical & Biological Engineering & Computing, 61*(7), 1821–1836. https://doi.org/10.1007/s11517-023-02820-3

Bhuvaji, S., Kadam, A., Bhumkar, P., Dedge, S., & Kanchan, S. (2020). *Brain tumor classification (MRI)* [Dataset]. Kaggle. https://doi.org/10.34740/KAGGLE/DSV/1183165

Chen, X., Wang, X., Zhang, K., Fung, K. M., Thai, T. C., Moore, K., ... Qiu, Y. (2022). Recent advances and clinical applications of deep learning in medical image analysis. *Medical Image Analysis, 79*, 102444. https://doi.org/10.1016/j.media.2022.102444

Cheng, J. (2017). *Brain tumor dataset* [Dataset]. figshare. https://doi.org/10.6084/m9.figshare.1512427.v8

El-Feshawy, S. A., Saad, W., Shokair, M., & Dessouky, M. (2023). IoT framework for brain tumor detection based on optimized modified ResNet 18 (OMRES). *The Journal of Supercomputing, 79*(1), 1081–1110. https://doi.org/10.1007/s11227-022-04678-y

Falk Delgado, A. (2025). Advances of MR imaging in glioma: What the neurosurgeon needs to know. *Acta Neurochirurgica, 167*(1), 174. https://doi.org/10.1007/s00701-025-06593-6

Guo, M., Cao, Z., Huang, Z., Hu, S., Xiao, Y., Ding, Q., ... Zhang, G. (2024). The value of CT shape quantification in predicting pathological classification of lung adenocarcinoma. *BMC Cancer, 24*(1), 35. https://doi.org/10.1186/s12885-023-11802-5

He, K., Zhang, X., Ren, S., & Sun, J. (2016). Deep residual learning for image recognition. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)* (pp. 770–778). IEEE. https://doi.org/10.1109/CVPR.2016.90

Huang, G., Liu, Z., Van Der Maaten, L., & Weinberger, K. Q. (2017). Densely connected convolutional networks. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition* (pp. 4700–4708). IEEE. https://doi.org/10.1109/CVPR.2017.243

Iqbal, S., Qureshi, A. N., Li, J., & Mahmood, T. (2023). On the analyses of medical images using traditional machine learning techniques and convolutional neural networks. *Archives of Computational Methods in Engineering, 30*(5), 3173–3233. https://doi.org/10.1007/s11831-023-09899-9

Kaur, C., & Garg, U. (2023). Artificial intelligence techniques for cancer detection in medical image processing: A review. *Materials Today: Proceedings, 81*(Part 2), 806–809. https://doi.org/10.1016/j.matpr.2021.04.241

Li, Y., Daho, M. E. H., Conze, P. H., Zeghlache, R., Le Boité, H., Tadayoni, R., ... Quellec, G. (2024). A review of deep learning-based information fusion techniques for multimodal medical image classification. *Computers in Biology and Medicine, 177*, 108635. https://doi.org/10.1016/j.compbiomed.2024.108635

Martucci, M., Russo, R., Schimperna, F., D'Apolito, G., Panfili, M., Grimaldi, A., ... Gaudino, S. (2023). Magnetic resonance imaging of primary adult brain tumors: State of the art and future perspectives. *Biomedicines, 11*(2), 364. https://doi.org/10.3390/biomedicines11020364

Mohandass, G. H. K. D. S. C. S. G., Krishnan, G. H., Selvaraj, D., & Sridhathan, C. (2024). Lung cancer classification using optimized attention-based convolutional neural network with DenseNet-201 transfer learning model on CT image. *Biomedical Signal Processing and Control, 95*, 106330. https://doi.org/10.1016/j.bspc.2024.106330

Narvekar, S., Shirodkar, M., Raut, T., Vaingankar, P., Kumar, K. C., & Aswale, S. (2022). A survey on detection of lung cancer using different image processing techniques. In *2022 3rd International Conference on Intelligent Engineering and Management (ICIEM)* (pp. 13–18). IEEE. https://doi.org/10.1109/ICIEM54221.2022.9853190

Nickparvar, M. (2021). *Brain tumor MRI dataset* [Dataset]. Kaggle. https://doi.org/10.34740/KAGGLE/DSV/2645886

Pandian, R., Vedanarayanan, V., Kumar, D. R., & Rajakumar, R. (2022). Detection and classification of lung cancer using CNN and GoogLeNet. *Measurement: Sensors, 24*, 100588. https://doi.org/10.1016/j.measen.2022.100588

Panayides, A. S., Amini, A., Filipovic, N. D., Sharma, A., Tsaftaris, S. A., Young, A., ... Pattichis, C. S. (2020). AI in medical imaging informatics: Current challenges and future directions. *IEEE Journal of Biomedical and Health Informatics, 24*(7), 1837–1857. https://doi.org/10.1109/JBHI.2020.2991043

Patel, P. R., & De Jesus, O. (2023). CT scan. In *StatPearls [Internet]*. StatPearls Publishing.

Rana, M., & Bhushan, M. (2023). Machine learning and deep learning approach for medical image analysis: Diagnosis to detection. *Multimedia Tools and Applications, 82*(17), 26731–26769. https://doi.org/10.1007/s11042-022-14305-w

Raza, R., Zulfiqar, F., Khan, M. O., Arif, M., Alvi, A., Iftikhar, M. A., & Alam, T. (2023). Lung-EffNet: Lung cancer classification using EfficientNet from CT-scan images. *Engineering Applications of Artificial Intelligence, 126*, 106902. https://doi.org/10.1016/j.engappai.2023.106902

Shamshad, F., Khan, S., Zamir, S. W., Khan, M. H., Hayat, M., Khan, F. S., & Fu, H. (2023). Transformers in medical imaging: A survey. *Medical Image Analysis, 88*, 102802. https://doi.org/10.1016/j.media.2023.102802

Simonyan, K., & Zisserman, A. (2015). *Very deep convolutional networks for large-scale image recognition*. International Conference on Learning Representations. https://doi.org/10.48550/arXiv.1409.1556

Tan, M., & Le, Q. V. (2019). EfficientNet: Rethinking model scaling for convolutional neural networks. In *Proceedings of the 36th International Conference on Machine Learning* (Vol. 97, pp. 6105–6114). PMLR. https://doi.org/10.48550/arXiv.1905.11946

Teixeira, P. A. G., Kessler, H., Morbée, L., Douis, N., Boubaker, F., Gillet, R., & Blum, A. (2025). Mineralized tissue visualization with MRI: Practical insights and recommendations for optimized clinical applications. *Diagnostic and Interventional Imaging, 106*(5), 147–156. https://doi.org/10.1016/j.diii.2024.11.001

Tian, L., Wu, J., Song, W., Hong, Q., Liu, D., Ye, F., ... Chen, L. (2024). Precise and automated lung cancer cell classification using deep neural network with multiscale features and model distillation. *Scientific Reports, 14*(1), 10471. https://doi.org/10.1038/s41598-024-61101-7

Ulli, S. S., Akkineni, H., Vuyyuru, U. S., Pathyala, S. R., Nannapaneni, U. K., & Suvarna, B. (2024). Brain tumor detection using modified VGG-19 model. In *2024 International Conference on Electrical Electronics and Computing Technologies (ICEECT)* (Vol. 1, pp. 1–5). IEEE. https://doi.org/10.1109/ICEECT61758.2024.10739175

Venkatachalam, C., Shah, P., Renukadevi, P., John, S., & Venkatachalam, S. (2025). Brain tumor diagnosis using modified DenseNet121 architecture with adaptive learning rate and callback mechanism. *Neural Computing and Applications, 37*(17), 11527–11553. https://doi.org/10.1007/s00521-025-11150-4

Xu, W., Fu, Y. L., & Zhu, D. (2023). ResNet and its application to medical image processing: Research progress and challenges. *Computer Methods and Programs in Biomedicine, 240*, 107660. https://doi.org/10.1016/j.cmpb.2023.107660

Yan, Y., Liu, G., Cai, H., Wu, E. Q., Cai, J., Cheok, A. D., ... Fan, Z. (2024). A review of graph theory-based diagnosis of neurological disorders based on EEG and MRI. *Neurocomputing, 599*, 128098. https://doi.org/10.1016/j.neucom.2024.128098

Yang, H., Zhang, Y., Gong, Y., Zhang, J., He, L., Zhong, J., & Tang, L. (2024). A lung biopsy path planning algorithm based on the double spherical constraint Pareto and indicators' importance-correlation degree. *Computerized Medical Imaging and Graphics, 117*, 102426. https://doi.org/10.1016/j.compmedimag.2024.102426