**EMITTER International Journal of Engineering Technology**
Vol. 12, No. 2, December 2024, pp. 93~111
DOI : 10.24003/emitter.v12i2.853

# Lite-FBCN: Lightweight Fast Bilinear Convolutional Network for Brain Disease Classification from MRI Image

**Dewinda Julianensi Rumala**[1], **Reza Fuad Rachmadi**[1,2], **Anggraini Dwi Sensusiati**[3], **I Ketut Eddy Purnama**[1,2]

[1]Department of Electrical Engineering, Institut Teknologi Sepuluh Nopember, Surabaya, Indonesia
[2]Department of Computer Engineering, Institut Teknologi Sepuluh Nopember, Surabaya, Indonesia
[3]Department of Radiology, Universitas Airlangga, Surabaya, Indonesia
Corresponding Author: ketut@te.its.ac.id

*Received July 15, 2024; Revised August 26, 2024; Accepted October 29, 2024*

## Abstract

Achieving high accuracy with computational efficiency in brain disease classification from Magnetic Resonance Imaging (MRI) scans is challenging, particularly when both coarse and fine-grained distinctions are crucial. Current deep learning methods often struggle to balance accuracy with computational demands. We propose Lite-FBCN, a novel Lightweight Fast Bilinear Convolutional Network designed to address this issue. Unlike traditional dual-network bilinear models, Lite-FBCN utilizes a single-network architecture, significantly reducing computational load. Lite-FBCN leverages lightweight, pre-trained CNNs fine-tuned to extract relevant features and incorporates a channel reducer layer before bilinear pooling, minimizing feature map dimensionality and resulting in a compact bilinear vector. Extensive evaluations on cross-validation and hold-out data demonstrate that Lite-FBCN not only surpasses baseline CNNs but also outperforms existing bilinear models. Lite-FBCN with MobileNetV1 attains 98.10% accuracy in cross-validation and 69.37% on hold-out data (a 3% improvement over the baseline). UMAP visualizations further confirm its effectiveness in distinguishing closely related brain disease classes. Moreover, its optimal trade-off between performance and computational efficiency positions Lite-FBCN as a promising solution for enhancing diagnostic capabilities in resource-constrained and or real-time clinical environments.

**Keywords**: Brain Image Classification, Bilinear Pooling, Deep Learning, Lightweight CNNs, Magnetic Resonance Imaging.

## 1. INTRODUCTION

Brain disease recognition is a pivotal aspect of clinical diagnosis and treatment planning, where timely and accurate identification of conditions can significantly impact patient outcomes [1]. Magnetic Resonance Imaging (MRI) serves as a non-invasive imaging modality that provides detailed views of brain structures, making it an essential tool in detecting and monitoring brain diseases [2]. However, interpreting these complex images requires advanced computational techniques to assist radiologists and clinicians in making accurate diagnosis.

In recent years, Deep Learning techniques, particularly Convolutional Neural Networks (CNNs), have emerged as powerful tools for automated analysis of medical images, including MRI scans [3]–[5]. CNNs have shown remarkable success in various medical imaging tasks, including brain disease recognition from MR images. Despite their effectiveness, CNNs typically require large annotated datasets to generalize well to new, unseen data. This dependency can be problematic in clinical settings where data may be scarce, diverse, and costly to obtain. Furthermore, the standard CNN architectures often struggle to capture complex feature interactions [6], [7], which are crucial for distinguishing subtle variations in medical images.

To address the limitations of conventional CNNs, researchers have developed Bilinear CNNs (BCNNs) that enhance the model's ability to capture higher-order feature interactions by combining feature vectors from two separate networks through bilinear pooling method [8]. BCNNs have demonstrated improved performance in various image recognition tasks [9]–[12]. However, the dual-network design of BCNNs results in high computational complexity and significant memory usage [13], making them less feasible for deployment in resource-constrained environments such as outpatient clinics, rural healthcare facilities, and other settings with limited computational resources typically found in clinical practice.

In response to these challenges, we propose Lite-FBCN, a Lightweight Fast Bilinear Convolutional Network tailored for brain disease recognition from MRI images. Lite-FBCN addresses the computational inefficiencies of the existing bilinear models by incorporating a single lightweight CNN backbone for feature extraction, thereby reducing the overall computational load and memory footprint. Additionally, Lite-FBCN introduces a channel reducer before the bilinear pooling layer, which reduces the dimensionality of the final feature map. This results in a more compact bilinear vector and a smaller model size, thus accelerating inference time.

Moreover, the complexity and subtlety of brain MRI features necessitate advanced methods capable of capturing fine-grained details. This need becomes especially critical in inter-class brain condition recognition tasks, such as distinguishing between cognitive normal (CN), neoplastic (NEO), cerebrovascular (CVA), degenerative disease (NDD), and inflammatory/infectious disease (INF). While conventional CNNs may struggle with subtle distinctions due to their limited ability to model higher-order

feature interactions, existing bilinear models may be computationally prohibitively expensive. By leveraging the bilinear pooling mechanism more efficiently, Lite-FBCN enhances the recognition of subtle patterns and interactions within MRI scans, all while reducing computational load. This efficiency is crucial for accurate diagnosis in resource-constrained clinics.

## 2. RELATED WORKS

Researchers have proposed CNN-based models for a wide range of medical imaging tasks, such as lesion detection, organ segmentation, and disease classification [1], [14]–[17]. These studies have demonstrated the potential of CNNs to achieve state-of-the-art performance in diagnosing various medical conditions from different imaging modalities, including MRI, CT, and PET. However, the growing demand for efficient and deployable medical image analysis systems has led to the development of lightweight CNN architectures. These models are designed to balance computational efficiency and classification performance, making them well-suited for resource-constrained environments. Examples of lightweight CNN architectures include MobileNets [18], [19] and EfficientNet [20], which achieve compact network structures without compromising accuracy.

Bilinear CNNs [8], which model pairwise feature interactions using bilinear pooling, have shown promising results in various computer vision tasks, including fine-grained image classification and scene understanding [9]–[12]. In the context of medical image analysis, researchers have explored the application of bilinear CNNs, especially for disease diagnosis using from various imaging modalities [21]–[24]. These studies have demonstrated the ability of bilinear CNNs to capture spatial relationships and discriminative features essential for accurate diagnosis from medical images.

However, BCNNs require high computational resource due to the utilization of two different backbone networks. To address this issue, Fast Bilinear Convolutional Networks (Fast BCNNs) was proposed [23] to reduce computational complexity. In Fast BCNNs, bilinear interactions are derived from the single backbone network, which contrasts with the dual-network design of conventional BCNNs. Fast BCNNs has been applied to various image classification tasks, such as for remote sensing image scene classification [10] and breast cancer classification [23], demonstrating its effectiveness in reducing computational overhead while preserving model performance.

In the case of brain disease recognition from MRI images, several studies have investigated the use of CNNs for this task by focusing not only on binary classification [2], [25]–[27], but also multi-class brain disease classification [2], [28]–[33]. Brain diseases often exhibit complex and subtle features that are challenging to distinguish, necessitating advanced feature extraction methods. Thus, different methods have also been proposed for this task, including transfer learning [27], [30], [33]–[36], ensemble learning [14], [30], [35], [37], [38], and even BCNNs [13]. While CNNs with various enhancement-based approaches have shown promising results in these tasks, there remains

a need for lightweight and efficient models capable of accurately classifying brain diseases in resource-constrained clinical settings.

## 3. ORIGINALITY

This paper introduces Lite-FBCN, a novel Lightweight Fast Bilinear Convolutional Network specifically designed to address the computational and performance limitations of conventional Bilinear CNNs (BCNNs) for brain disease recognition from MRI scans. The originality of Lite-FBCN lies in its innovative approach to combining high accuracy with significantly reduced computational complexity. Unlike traditional BCNNs that rely on bilinear interaction from dual-network architectures, Lite-FBCN implements a single-network design where the captured information from this backbone alone is processed through the bilinear pooling layer.

A key design aspect of Lite-FBCN is the introduction of a channel reducer before the bilinear pooling layer to reduce the dimensionality of the backbone's feature maps, thus lowering the computational complexity of the bilinear pooling operation. This results in a smaller, more efficient model with reduced inference time that maintains high accuracy in brain disease recognition. The channel reducer ensures that the feature maps are condensed without losing critical information, thereby enhancing the model's ability to capture intricate feature interactions necessary for distinguishing subtle variations in MRI images.

The versatility of Lite-FBCN is evident in its ability to achieve high accuracy with lightweight backbone networks through fine-tuning. This adaptability allows Lite-FBCN to be tailored to different deployment scenarios, accommodating a range of resource constraints without compromising performance. The combination of high accuracy, reduced computational complexity, and flexibility in backbone selection underscores the originality and potential impact of Lite-FBCN in advancing the state of the art in brain disease recognition from MRI images. This makes Lite-FBCN a promising tool for enhancing diagnostic capabilities in clinical practice.

## 4. SYSTEM DESIGN

### 4.1 Proposed Lite-FBCN

The proposed Lite-FBCN (Lightweight Fast Bilinear Convolutional Network) aims to enhance brain disease recognition from MRI scans by addressing the computational and performance limitations of the existing bilinear CNN models. Lite-FBCN introduces a streamlined, single-network design that leverages transfer learning for the feature extraction backbone. This approach significantly reduces the computational cost compared to conventional BCNNs, which rely on dual-network architectures. Additionally, Lite-FBCN incorporates a channel reducer layer before the bilinear pooling layer, compressing the feature map dimensionality and further decreasing computational complexity. This combination of techniques not only maintains high accuracy but also ensures faster inference times, making Lite-FBCN

suitable for deployment in both resource-constrained and real-time clinical environments. The schematic of the proposed Lite-FBCN design is illustrated in Figure 1. In this research, we focus on the classification task involving the differentiation between multiple categories of brain conditions. This classification task is critical for providing accurate and timely diagnosis, which can impact patient outcomes.

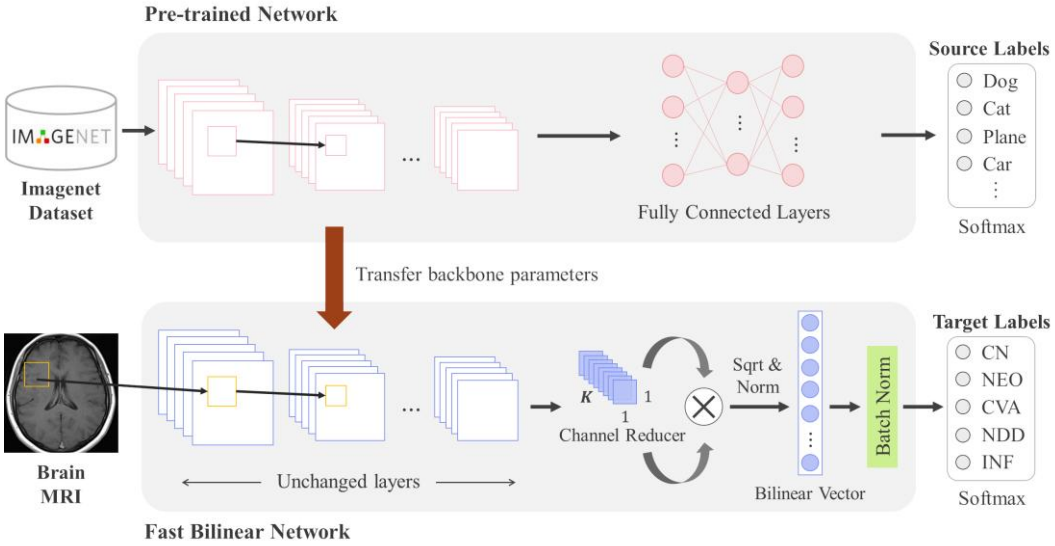


**Figure 1.** Schematic diagram of the proposed Lite-Fast Bilinear Convolutional Network (Lite-FBCN).

### 4.1.1 Backbone Networks

The models used in this study are based on the well-established lightweight pre-trained CNN models, including MobileNetV1 [18], MobileNetV2 [19], and EfficientNetB0 [20]. We employ these CNN models as the backbone network in this study due to its capability to extract low-level features using its smaller kernel. To enhance the performance of the backbone networks and mitigate distributinoal data problem, transfer learning is applied in this work. The data source for the transfer learning application that we used in this study is natural images, derived from ImageNet. During the training, we fine-tuned all layers of the CNN models, so that all weights are optimized according to the learned features starting from the low-level to the high-level, as there are high distinctions between the natural images and medical images [1], [39].

### 4.1.2 Channel Reducer Layer

The channel reducer layer, placed before the bilinear pooling layer, is implemented as 1×1 convolutional layer to effectively combine the information across the input channels and reduce their number according to the specified number of filters prior to engaging in bilinear interactions.

Let $\boldsymbol{F} \in \mathbb{R}^{\boldsymbol{H}\times\boldsymbol{W}\times\boldsymbol{C}}$ be the feature map output of the backbone network, where $\boldsymbol{H}$ is the height, $\boldsymbol{W}$ is the width, and $\boldsymbol{C}$ denotes the depth (number of channels) of the feature map. The channel reducer layer, a 1x1 convolutional

layer with $\boldsymbol{K}$ filters, has weights denoted by $\boldsymbol{\Theta_{cr}} \in \mathbb{R}^{\boldsymbol{1\times1\times C\times K}}$, where the dimensions correspond to the 1×1 spatial size, the input channels $\boldsymbol{C}$, and the output channels $\boldsymbol{K}$. The convolutional operation to generate each element of the output feature map $\boldsymbol{F_{cr}}$ at spatial position $(\boldsymbol{h}, \boldsymbol{w})$ and output channel $\boldsymbol{k}$ is given in equation 1. The resulting output feature map $\boldsymbol{F_{cr}} \in \mathbb{R}^{\boldsymbol{H\times W\times K}}$ retains the original spatial dimensions H and W, while the depth is reduced to K.

$$\boldsymbol{F_{cr}}(\boldsymbol{h}, \boldsymbol{w}, \boldsymbol{k}) = \sum_{\boldsymbol{c=1}}^{\boldsymbol{C}} \boldsymbol{F}(\boldsymbol{h}, \boldsymbol{w}, \boldsymbol{c}) \cdot \boldsymbol{\Theta_{cr}}(\boldsymbol{1}, \boldsymbol{1}, \boldsymbol{c}, \boldsymbol{k}) \quad (1)$$

The transformation of $\boldsymbol{F}$ into $\boldsymbol{F_{cr}}$ is governed by the channel reduction factor (CRF), denoted as $\boldsymbol{\gamma}$. The CRF is defined as the quotient of the initial number of channels $\boldsymbol{C}$ in the original feature maps divided by the number of output channels $\boldsymbol{K}$ in the channel reducer layer, as expressed in equation 2.

$$\boldsymbol{\gamma} = \frac{\mathrm{C}}{K} \quad (2)$$

The channel reducer layer plays a crucial role in making the bilinear pooling method practical and effective. By reducing the dimensionality of the last feature maps $\boldsymbol{F}$ of the backbone network into $\boldsymbol{F_{cr}}$, the channel reducer layer lowers computational and memory costs, focuses on the most relevant features, helps in preventing overfitting, and facilitates a more efficient and stable training process. Applying this layer before the bilinear pooling layer helps the subsequent interactions captured to be both manageable and meaningful, ultimately leading to better model performance.

### 4.1.3 Bilinear Pooling

Unlike standard BCNNs that utilize dual-network architectures, and Fast BCNNs that apply bilinear pooling on a single network without a channel reducer, Lite-FBCN incorporates a channel reducer to streamline the feature maps before bilinear pooling. This integration significantly reduces the computational load while preserving the model's capacity to capture detailed feature interactions. The key difference between standard Bilinear Pooling (BCNNs), Fast Bilinear Pooling (FBCNNs), and the proposed Lite-FBCN is illustrated in Figure 2.

Bilinear pooling captures complex feature interactions by computing the outer product of the reduced feature map $\boldsymbol{F_{cr}}$ with itself at each spatial location, followed by summation across all spatial locations. The bilinear feature $\boldsymbol{B} \in \mathbb{R}^{\boldsymbol{K\times K}}$ is calculated according to equation 3, where $F(i, j)$ represents a $\boldsymbol{K}$-dimensional vector at spatial location $(i, j)$, and $\otimes$ denotes the outer product.

$$\boldsymbol{B} = \sum_{\boldsymbol{i=1}}^{\boldsymbol{H}} \sum_{\boldsymbol{j=1}}^{\boldsymbol{W}} \boldsymbol{F_{cr}}(\boldsymbol{i}, \boldsymbol{j}) \otimes \boldsymbol{F_{cr}}(\boldsymbol{i}, \boldsymbol{j})^{\top} \quad (3)$$

To improve the stability and performance of the bilinear features, normalization techniques such as signed square root and $\ell_2$ normalization is applied. After that, we apply batch normalization to enhance the stability and

convergence of the network. Batch normalization normalizes the activations of each layer across the mini-batch, which helps mitigate issues such as internal covariate shift and accelerates training. Following batch normalization, the processed features are fed into a classification layer, where SoftMax activation is applied to produce class probabilities. To regularize the model and prevent overfitting, we apply $\ell_2$ regularization with a regularization strength of 0.01 to the weights of the classification layer.

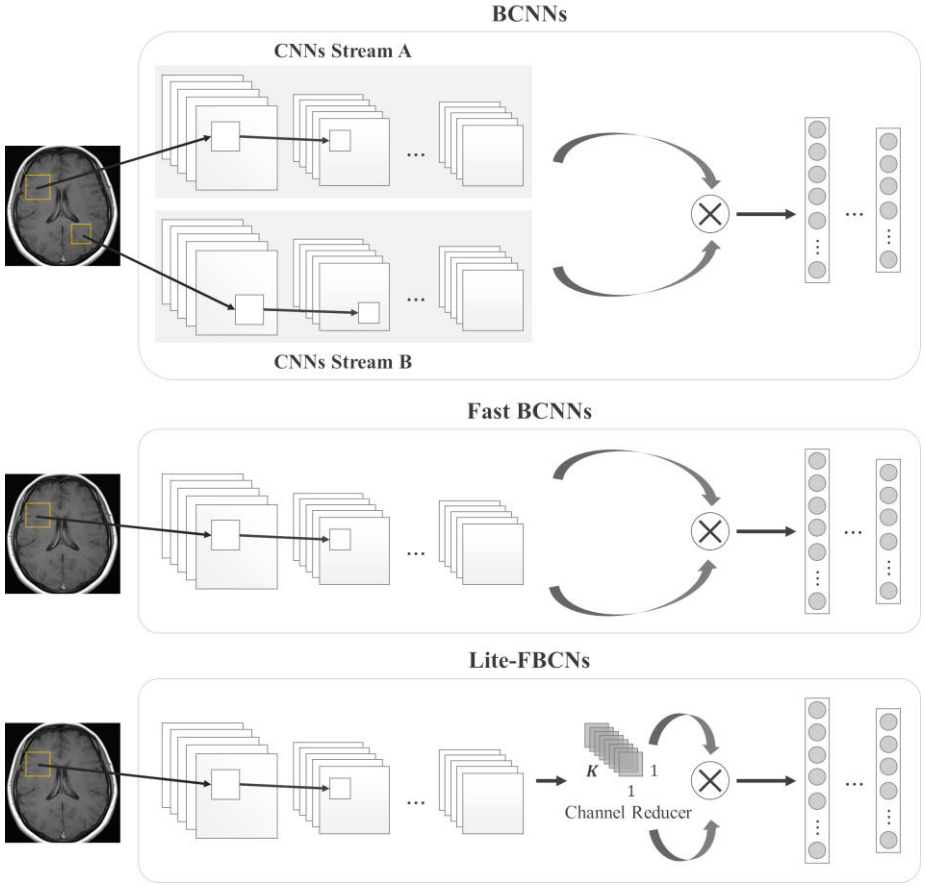


**Figure 2.** The structure differences between the standard BCNNs, Fast BCNNs, and the proposed Lite-FBCN.

### 4.2 Implementation

We used the same dataset of axial T2-weighted brain MR images and preprocessing steps as a previous study on brain disease diagnosis [30]. The dataset includes 1055 images across five classes of human brain conditions: 250 CN (Cognitively Normal), 214 NEO (Neoplasm), 315 CVA (Cerebrovascular Accident), 141 NDD (Neurodegenerative Disease), and 135 INF (Infection). These numbers and class labels are consistent with those specified in the previous study [30]. Our models were also evaluated using five-fold cross-validation for training, validation, and testing.

To enhance the model's robustness, we incorporated real-time data augmentation during training, introducing variability to the training datasets as outlined in [13]. We optimized the model using the Stochastic Gradient

Descent (SGD) optimizer with the following parameters: a learning rate of 0.01, momentum of 0.5, and Nesterov momentum enabled. If the validation loss did not decrease for 50 epochs, the learning rate was reduced by a factor of 10, with a minimum learning rate of 0.0001. The models were trained for 100 epochs using categorical cross-entropy as the loss function. To ensure the best-performing model was retained during training, we implemented a custom model checkpoint callback as proposed in [30].

For a comprehensive insights into the methodologies, supplementary materials, and code supporting the study findings, please refer to https://djrumala.github.io/sup/lite-fbcn.

### 4.3 Performance Metrics and Evaluation

To assess the classifier model's performance, we computed accuracy, precision-recall, and F1 score. Additionally, we conducted a statistical test, repeated-measures ANOVA, to compare the proposed Lite-FCBN, the baseline, and the existing bilinear CNN models based on accuracy, precision-recall, and F1 score. A significance level of $P < 0.05$ was employed to determine statistical significance.

## 5. EXPERIMENT AND ANALYSIS

### 5.1. Performance of Baseline Models

To establish a reference for evaluating our proposed Lite-FBCN model, we first assessed the performance of several baseline models. These models include standard pre-trained CNNs that have been fine-tuned using brain MRI data for multi-class brain disease classification. Each baseline model was fine-tuned and evaluated on the same 5-fold cross-validation brain MRI data. The average performance of baseline models on the 5-fold cross-validation data is outlined in Table 1. Meanwhile, detailed results for each model's metrics per fold are available in the supplemental table. From Table 1, MobileNetV1 exhibits the highest accuracy and overall performance among these models. However, it is essential to note that our analysis revealed no statistically significant difference ($P = 0.9665$) in accuracy among the baseline models. This suggests that the choice of CNN backbone for the baseline methods may not significantly impact performance.

**Table 1.** Performance of baseline models on the 5-fold cross-validation data

| Backbone | Acc (%) | Prec (%) | Rec (%) | F1-score (%) |
|---|---|---|---|---|
| EfficientNetB0 | 97.43 ± 0.98 | 97.33 ± 0.80 | 97.06 ± 0.78 | 97.11 ± 0.83 |
| MobileNetV1 | 97.62 ± 0.85 | 97.59 ± 1.01 | 97.32 ± 0.89 | 97.37 ± 0.97 |
| MobileNetV2 | 97.43 ± 1.72 | 97.37 ± 1.86 | 97.09 ± 1.60 | 97.20 ± 1.73 |

Average metrics and standard deviations were calculated from 5-fold Cross-Validation.

### 5.2. Performance of Lite-FBCN Models

We evaluated the effectiveness of the proposed Lite-FBCN model in brain disease classification tasks, employing the same training and evaluation

procedures as the baseline models. Table 2 presents the performance metrics of the Lite-FBCN model. The results show that Lite-FBCN models achieved higher accuracy compared to baseline models when using the MobileNets backbone. Conversely, there was a decrease in performance for EfficientNetB0, though we found no statistically significant difference between both methods (P = 0.5659). The choice of backbone architecture significantly influences the performance of Lite-FBCN (P = 0.0047). However, further analysis suggests that differences in performance between MobileNetV1 and MobileNetV2 may not be significant (P = 1). This implies that EfficientNetB0 exhibited the lowest performance when utilized as a backbone in Lite-FBCN models. Meanwhile, the CRF (γ) also did not play a significant role in performance (P = 0.6462).

**Table 2.** Performance of Lite-FBCN on the 5-fold cross-validation data

| $\gamma$ | Backbone | Acc (%) | Prec (%) | Rec (%) | F1-score (%) |
|---|---|---|---|---|---|
| 2 | EfficientNetB0 | 91.05 ± 2.61 | 91.38 ± 2.36 | 91.05 ± 2.61 | 90.64 ± 2.92 |
| | MobileNetV1 | 98.00 ± 0.92 | 98.04 ± 0.94 | 98.00 ± 0.92 | 97.99 ± 0.93 |
| | MobileNetV2 | 98.29 ± 0.49 | 98.33 ± 0.48 | 98.28 ± 0.49 | 98.18 ± 0.56 |
| 4 | EfficientNetB0 | 93.81 ± 1.62 | 93.79 ± 1.66 | 93.81 ± 1.62 | 93.61 ± 1.67 |
| | MobileNetV1 | 98.10 ± 0.90 | 98.21 ± 0.88 | 98.10 ± 0.90 | 98.09 ± 0.91 |
| | MobileNetV2 | 97.81 ± 0.49 | 97.89 ± 0.44 | 97.81 ± 0.49 | 97.77 ± 0.50 |
| 8 | EfficientNetB0 | 93.33 ± 2.48 | 93.49 ± 2.47 | 93.33 ± 2.48 | 93.20 ± 2.58 |
| | MobileNetV1 | 97.52 ± 0.76 | 97.59 ± 0.76 | 97.32 ± 0.57 | 97.52 ± 0.76 |
| | MobileNetV2 | 97.52 ± 0.70 | 97.62 ± 0.67 | 97.52 ± 0.70 | 97.50 ± 0.70 |

Average metrics and standard deviations were calculated from 5-fold Cross-Validation.

### 5.3 Comparison with Existing Bilinear CNN Models

To further validate the performance of the Lite-FBCN model, we assessed the performance of existing bilinear CNN models, including both standard BCNNs and Fast BCNNs. These comparisons were based on the same dataset and evaluation metrics. Due to the observed poor performance of EfficientNetB0 backbone in the Lite-FBCN model, we focused the comparison only on MobileNetV1 and MobileNetV2 backbones. Table 3 summarizes the performance results of the existing bilinear models: BCNNs and Fast BCNN models. Statistical analysis showed no significant performance difference between BCNNs and Lite-FBCN models across all backbones ($P$ = 0.1759). No significant difference was also found between Fast BCNNs and Lite-FBCN models with MobileNetV1 backbone ($P$ = 0.4572). However, a statistically significant difference was observed when comparing both methods using MobileNetV2 backbone ($P$ = 0.0087), suggesting that channel reducer layer in Lite-FBCN model can help maintain performance using this backbone. Although statistically significant difference in classification performance was not much detected between Lite-FBCN and existing bilinear models, computational efficiency should also be considered.

**Table 3.** Performance of the existing conventional Bilinear CNNs and Fast BCNNs on the 5-fold cross-validation data

| Pooling Method | Backbone | Acc (%) | Prec (%) | Rec (%) | F1-score (%) |
|---|---|---|---|---|---|
| BCNNs | MobileNetV1+ MobileNetV2 | 98.19 ± 3.00 | 98.27±17.28 | 98.19 ± 5.68 | 98.19 ± 5.68 |
| Fast BCNNs | MobileNetV1 | 97.62 ± 3.00 | 97.73 ± 2.67 | 97.62 ± 7.71 | 97.63 ± 7.71 |
| | MobileNetV2 | 96.95 ± 3.00 | 97.04±17.00 | 96.95 ± 5.68 | 96.91 ± 5.68 |

Average metrics and standard deviations were calculated from 5-fold Cross-Validation.

### 5.4 Comparison on Computation Efficiency

In addition to classification performance, we evaluated the computational efficiency of the proposed Lite-FBCN. The findings in Table 4 present a comparison of inference time and parameter count between the Lite-FBCN and other methods using various CNN backbones. This investigation allows for insights into the computational efficiency of Lite-FBCN relative to a diverse range of architectural choices. From these results, significant reductions in model size and inference time are achieved by Lite-FBCN, particularly with higher CRF (γ). While the lightest parameter count is attained with γ = 8, implementing γ = 4 results in faster inference times across all backbones. Among the backbones, MobileNetV1 stands out for its impressive inference time, outperforming other models in acceleration. Notably, MobileNetV1 demonstrated the fastest inference time with γ = 4.

**Table 4.** Comparison of inference time and parameter count between models

| Methods | $\gamma$ | Backbone | #Params (mil) | Inference time (ms/img) |
|---|---|---|---|---|
| **Baseline** | N/A | EfficientNetB0 | 4.8381 | 43.48 ± 1.19 |
| | | MobileNetV1 | 3.8863 | 14.50 ± 0.44 |
| | | MobileNetV1 | 3.0465 | 24.94 ± 1.27 |
| **BCNN** | N/A | MobileNetV1+ MobileNetV1 | 17.2833 | 45.54 ± 2.92 |
| **FBCNN** | N/A | MobileNetV1 | 12.6661 | 16.39 ± 0.54 |
| | N/A | MobileNetV2 | 17.0036 | 25.94 ± 1.27 |
| **Lite-FBCN** | 2 | EfficientNetB0 | 8.5558 | 45.28 ± 1.35 |
| | | MobileNetV1 | 7.5713 | 15.89 ± 0.65 |
| | | MobileNetV2 | 6.7642 | 29.72 ± 2.69 |
| | 4 | EfficientNetB0 | 5.3811 | 46.64 ± 1.52 |
| | | MobileNetV1 | 4.4785 | 15.29 ± 0.56 |
| | | MobileNetV2 | 3.5895 | 28.92 ± 1.65 |
| | 8 | EfficientNetB0 | 4.4849 | 45.05 ± 0.39 |
| | | MobileNetV1 | 3.6232 | 16.56 ± 0.51 |
| | | MobileNetV2 | 2.6933 | 29.73 ± 2.47 |

Average metrics and standard deviations were calculated from 5 repeated measurements.

### 5.5 Further Evaluation on Hold-Out Data

In this section, we present an additional evaluation of the proposed Lite-FBCN model and existing bilinear models using the hold-out dataset. This

evaluation provides further insights into the model's performance and robustness in a real-world scenario, outside the cross-validation data. The analysis focuses on the model's ability to correctly classify brain conditions across new subjects, ensuring that the performance improvements were not due to overfitting to specific individuals. For this evaluation, we collected new brain MR images of axial T2-weighted, categorized similarly to the cross-validation data. We obtained 1091 images of CN from 25 subjects from the IXI dataset (https://brain-development.org/ixi-dataset), 3559 images of NEO from 167 subjects from the BraTS dataset [40], [41], 102 images of CVA from the Kurzad Poyraz dataset [31], 716 images from 65 subjects from the ADNI dataset, and 287 images from 35 subjects from Radiopaedia [42], [43].

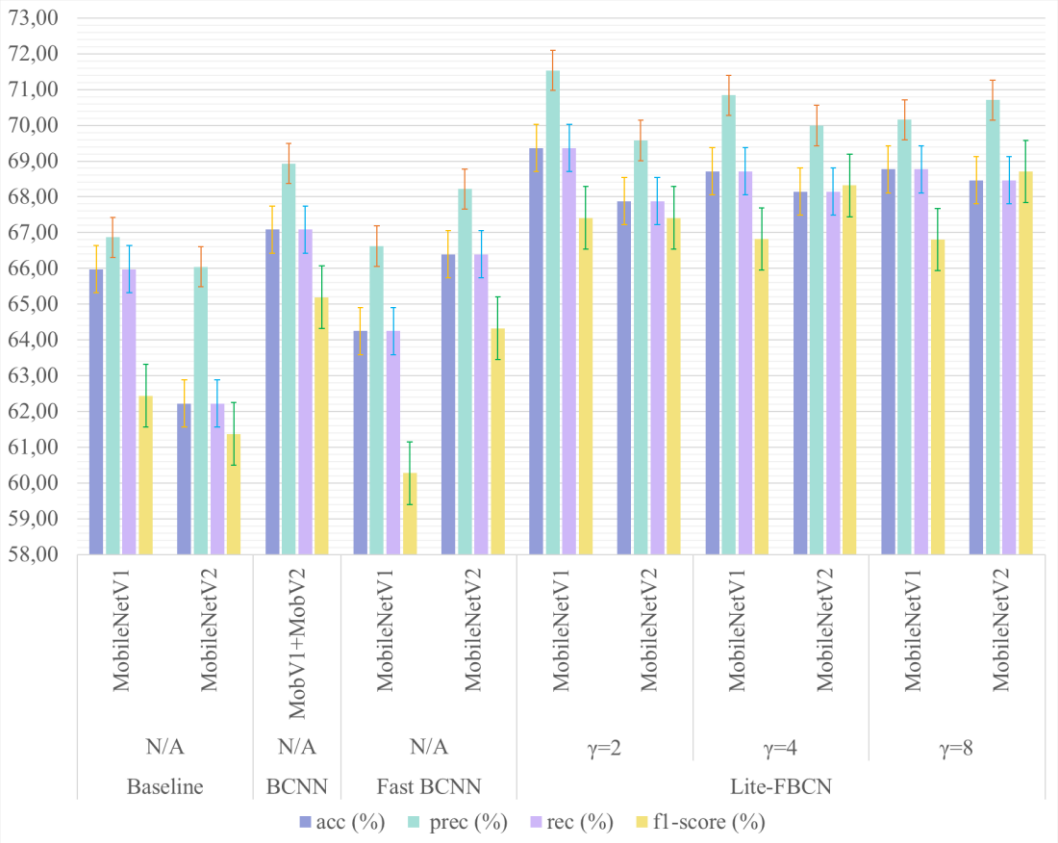


**Figure 3.** Performance comparison of the proposed Lite-FBCN and other methods on the hold-out data.

The results of the evaluation on hold-out data are shown in Figure 3. We observed that Lite-FBCN with MobileNetV1 backbone performed the best compared to other models, particularly when γ = 2. Even with higher γ values, Lite-FBCN with MobileNetV1 continued to outperform other backbones and methods. There was no statistically significant difference in Lite-FBCN performance for MobileNetV1 across different γ values (P = 0.7142). Notably, while MobileNetV1 with Lite-FBCN showed the best performance, there was no significant difference between using MobileNetV1 or MobileNetV2 backbone (P = 0.8876). Additionally, there was no significant performance

difference between Lite-FBCN models and both baseline Fast BCNNs (P = 0.2579) and BCNNs (P = 0.5467).

The confusion matrix comparison in Figure 4 provides a detailed breakdown of classification accuracy across different classes, highlighting the robustness of the Lite-FBCN model, particularly in correctly identifying various brain disease categories compared to the baseline and existing bilinear models. While all models struggled with classifying NDD and INF, especially where these classes were often confused with other closely similar-looking classes, Lite-FBCN achieved better accuracy for these classes. While the choice between MobileNetV1 and MobileNetV2 backbone networks also affects accuracy for each class, it is less significant compared to implementing the bilinear pooling mechanism.

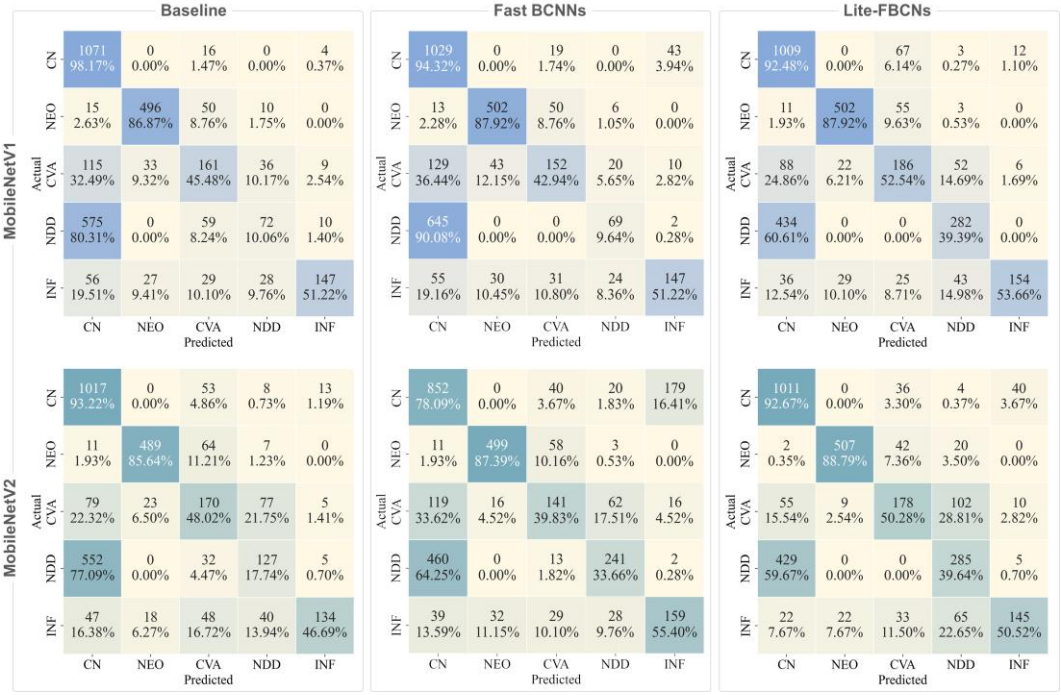


**Figure 4.** Confusion matrix comparison of the proposed Lite-FBCN, baseline, and existing bilinear CNN models on the hold-out data.

The UMAP visualization in Figure 5 provides further insights into model performance by illustrating the clustering and separation of different brain disease categories achieved by the various models. This visualization underscores the superior feature representation capabilities of the Lite-FBCN model. While baseline models exhibit more scattered clusters, the implementation of bilinear pooling methods results in improved clustering, despite some overlapping between closely related brain classes. Notably, classification becomes more distinct in bilinear models, particularly in Lite-FBCN. Similar to the results observed in the confusion matrix, the choice of backbone networks also appears to affect clustering in UMAP, although less significant compared to implementing bilinear pooling mechanism.

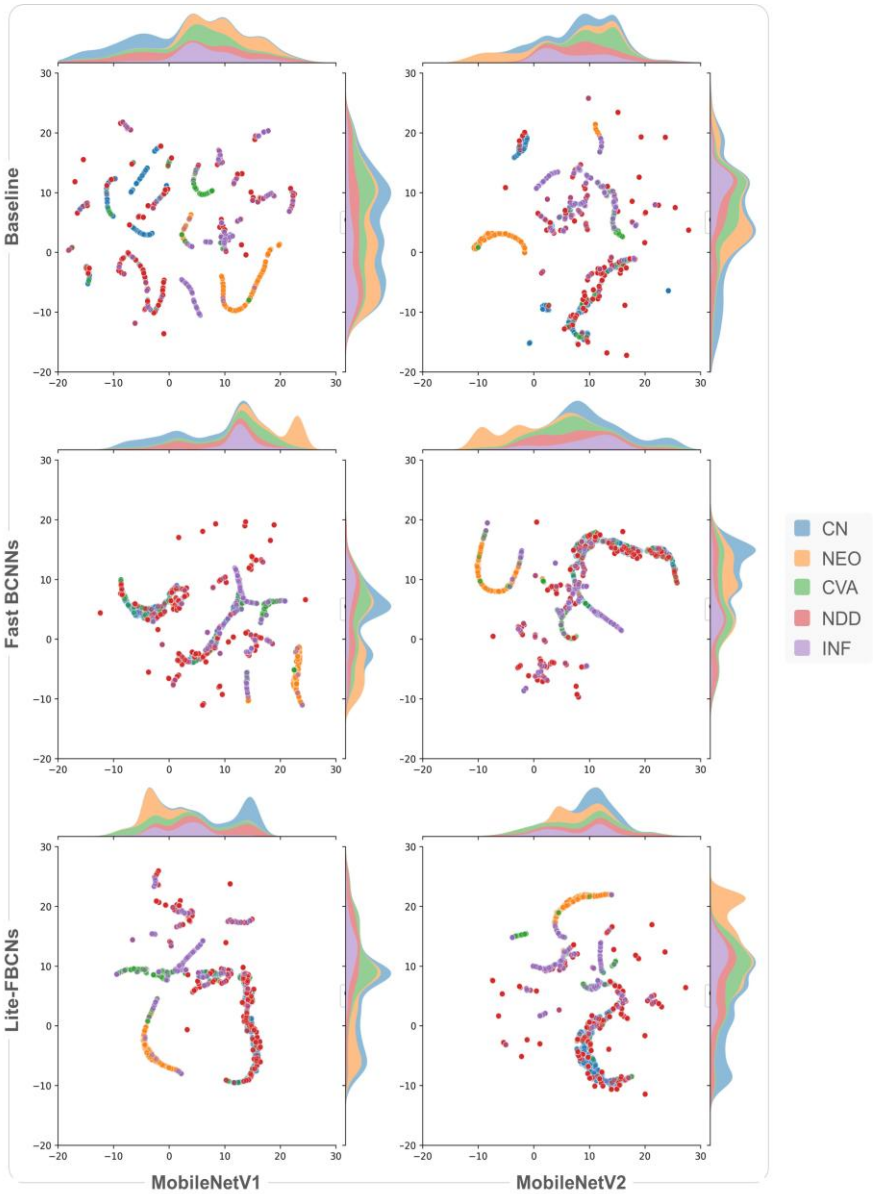


**Figure 5.** UMAP visualization comparison of the proposed Lite-FBCN, baseline, and existing bilinear CNN models on the hold-out data.

## 5.6 Discussion

During both cross-validation and hold-out data evaluations, the choice of backbone networks did not significantly impact different methods. MobileNets, when employed as backbones within bilinear models, demonstrated superior performance compared to their performance in the standard fine-tuned baseline approach, especially evident in the Lite-FBCN model. Conversely, during cross-validation data evaluation, the performance of EfficientNetB0 as the Lite-FBCN backbone was lower, and it incurred higher computational costs. The complex and deeper network design of EfficientNetB0 may lead to overfitting or increased computational overhead, diminishing its effectiveness as a bilinear backbone.

Our evaluation also revealed that Lite-FBCN's performance was unaffected by the channel reduction factor (CRF) of the channel reducer layer, demonstrating its strong design and feature extraction capabilities. It appeared that the addition of this channel reducer layer allowed Lite-FBCN to preserve critical information from brain MR images, despite reducing the dimensionality of the backbone's last feature maps. Among different CRF configurations, MobileNetV1 with a CRF of 4 emerged as the most recommended choice, striking a balance between computational efficiency and robustness in classification performance. A CRF of 4 reduced feature map dimensionality to a lesser extent compared to a CRF of 8, resulting in lower computational load and faster processing time while maintaining sufficient feature richness. Conversely, a CRF of 8 compressed feature maps more aggressively, potentially leading to information loss and increased computational demands.

In terms of class-specific accuracy, our analysis using confusion matrices and UMAP visualizations during hold-out data evaluation reveals that distinguishing between NDD and INF classes presented challenges across all models. Notably, NDD was frequently misclassified as CN, likely due to their similar appearance or shared features. However, Lite-FBCN demonstrated significant improvements in accurately distinguishing these classes. The UMAP visualization underscored Lite-FBCN's superior feature representation capabilities, exhibiting clearer clustering and separation of disease categories compared to baseline and existing bilinear models. The alignment between the UMAP visualization and the findings from the confusion matrices further supports the robustness of our results.

## 6. CONCLUSION

The development of Lite-FBCN represents a significant advancement in brain disease classification from MRI scans. With its innovative lightweight architecture and efficient feature extraction mechanisms, Lite-FBCN successfully addresses the challenges of achieving high accuracy while maintaining computational efficiency. By leveraging a single-network design and incorporating a channel reducer layer, Lite-FBCN reduces computational load and model size without compromising performance.

Our evaluations, conducted through both cross-validation and hold-out datasets, demonstrate the superiority of Lite-FBCN over baseline CNNs and existing bilinear models. Notably, its ability to outperform other methods in capturing subtle feature interactions of brain MR images crucial for distinguishing between closely related brain disease classes. The utilization of MobileNetV1 as backbone networks within Lite-FBCN showcases superior performance, especially when combined with a CRF of 4, striking a balance between computational efficiency and classification robustness. Furthermore, our analysis demonstrates how Lite-FBCN effectively tackles class-specific challenges, such as differentiating between classes that have similar features or appearances. The enhanced feature representation capabilities of Lite-

FBCN are demonstrated through confusion matrices and UMAP visualization. The UMAP visualization shows improved clustering and separation of brain condition groups achieved by Lite-FBCN.

The optimal trade-off between performance and computational efficiency exhibited by Lite-FBCN holds promise for enhancing diagnostic abilities in resource-constrained clinical environments. Future research could focus on further validating Lite-FBCN's performance with larger and more diverse datasets, exploring its adaptability to other medical imaging modalities, investigating techniques for continual learning and adaptation to evolving disease patterns. Additionally, efforts could be directed towards enhancing the interpretability of Lite-FBCN's decisions to foster trust and transparency in its clinical use.

**Acknowledgements**

This study was supported by the PMDSU scholarship from the Directorate General of Higher Education and Research Technology, Indonesia, and partially funded by the RISPRO LPDP from the Ministry of Finance, Indonesia.

**REFERENCES**

[1] J. M. Valverde *et al.*, **Transfer Learning in Magnetic Resonance Brain Imaging: A Systematic Review**, *J. Imaging*, vol. 7, no. 4, p. 66, 2021, doi: 10.3390/jimaging7040066.

[2] A. Gudigar, U. Raghavendra, E. J. Ciaccio, N. Arunkumar, E. Abdulhay, and U. R. Acharya, **Automated categorization of multi-class brain abnormalities using decomposition techniques with MRI images: A comparative study**, *IEEE Access*, vol. 7, pp. 28498–28509, 2019, doi: 10.1109/ACCESS.2019.2901055.

[3] M. Narazani, I. Sarasua, S. Pölsterl, A. Lizarraga, I. Yakushev, and C. Wachinger, **Is a PET All You Need? A Multi-modal Study for Alzheimer's Disease Using 3D CNNs**, *Lect. Notes Comput. Sci. (including Subser. Lect. Notes Artif. Intell. Lect. Notes Bioinformatics)*, vol. 13431 LNCS, pp. 66–76, 2022, doi: 10.1007/978-3-031-16431-6_7.

[4] J. Zhang, B. Zheng, A. Gao, X. Feng, D. Liang, and X. Long, **A 3D densely connected convolution neural network with connection-wise attention mechanism for Alzheimer's disease classification**, *Magn. Reson. Imaging*, vol. 78, pp. 119–126, 2021, doi: 10.1016/J.MRI.2021.02.001.

[5] H. P. A. Tjahyaningtijas, A. K. Nugroho, C. V. Angkoso, I. K. E. Purnama, and M. H. Purnomo, **Automatic Segmentation on Glioblastoma Brain Tumor Magnetic Resonance Imaging Using Modified U-Net**, *Emit. Int. J. Eng. Technol.*, vol. 8, no. 1, pp. 161–177, 2020, doi: 10.24003/emitter.v8i1.505.

[6] G. Litjens *et al.*, **A survey on deep learning in medical image analysis**, *Med. Image Anal.*, vol. 42, pp. 60–88, 2017, doi:

https://doi.org/10.1016/j.media.2017.07.005.
[7] M. Rajnoha, R. Burget, and L. Povoda, **Image Background Noise Impact on Convolutional Neural Network Training**, in *2018 10th International Congress on Ultra Modern Telecommunications and Control Systems and Workshops (ICUMT)*, 2018, pp. 1–4, doi: 10.1109/ICUMT.2018.8631242.
[8] T.-Y. Lin, A. RoyChowdhury, and S. Maji, **Bilinear CNN Models for Fine-Grained Visual Recognition**, in *2015 IEEE International Conference on Computer Vision (ICCV)*, 2015, pp. 1449–1457, doi: 10.1109/ICCV.2015.170.
[9] L. C. Neural and S. Classification, **An Efficient and Lightweight Convolutional Neural Network for Remote Sensing Image Scene Classificatio**, 2020.
[10] D. Yu, Q. Xu, H. Guo, C. Zhao, Y. Lin, and D. Li, **An Efficient and Lightweight Convolutional Neural Network for Remote Sensing Image Scene Classification**, *Sensors (Basel).*, vol. 20, no. 7, 2020, doi: 10.3390/s20071999.
[11] M. Lu *et al.*, **Intelligent Grading of Tobacco Leaves Using an Improved Bilinear Convolutional Neural Network**, *IEEE Access*, vol. 11, pp. 68153–68170, 2023, doi: 10.1109/ACCESS.2023.3292340.
[12] A. J. Prakash and P. Prakasam, **An intelligent fruits classification in precision agriculture using bilinear pooling convolutional neural networks**, *Vis. Comput.*, vol. 39, no. 5, pp. 1765–1781, 2023, doi: 10.1007/s00371-022-02443-z.
[13] D. J. Rumala *et al.*, **Bilinear MobileNets for Multi-class Brain Disease Classification Based on Magnetic Resonance Images**, in *2021 IEEE Region 10 Symposium (TENSYMP)*, 2021, pp. 152–157, doi: 10.1109/TENSYMP52854.2021.9550987.
[14] B. Harangi, **Skin lesion classification with ensembles of deep convolutional neural networks**, *J. Biomed. Inform.*, vol. 86, pp. 25–32, Oct. 2018, doi: 10.1016/J.JBI.2018.08.006.
[15] D. S. Luz, T. J. B. Lima, R. R. V. Silva, D. M. V. Magalhães, and F. H. D. Araujo, **Automatic detection metastasis in breast histopathological images based on ensemble learning and color adjustment**, *Biomed. Signal Process. Control*, vol. 75, p. 103564, 2022, doi: 10.1016/J.BSPC.2022.103564.
[16] Y. Oh, S. Park, and J. C. Ye, **Deep Learning COVID-19 Features on CXR Using Limited Training Data Sets**, *IEEE Trans. Med. Imaging*, vol. 39, no. 8, pp. 2688–2700, 2020, doi: 10.1109/TMI.2020.2993291.
[17] G. Varoquaux and V. Cheplygina, **Machine learning for medical imaging: methodological failures and recommendations for the future**, *npj Digit. Med.*, vol. 5, no. 1, p. 48, 2022, doi: 10.1038/s41746-022-00592-y.
[18] A. G. Howard *et al.*, **MobileNets: Efficient convolutional neural networks for mobile vision applications**, *arXiv*, 2017.
[19] M. Sandler, A. Howard, M. Zhu, A. Zhmoginov, and L. C. Chen,

**MobileNetV2: Inverted Residuals and Linear Bottlenecks**, *Proc. IEEE Comput. Soc. Conf. Comput. Vis. Pattern Recognit.*, pp. 4510–4520, 2018, doi: 10.1109/CVPR.2018.00474.

[20] M. Tan and Q. Le, **EfficientNet: Rethinking Model Scaling for Convolutional Neural Networks**, in *Proceedings of the 36th International Conference on Machine Learning*, 2019, vol. 97, pp. 6105–6114, [Online]. Available: https://proceedings.mlr.press/v97/tan19a.html.

[21] R. Yu, C. Pan, X. Fei, M. Chen, and D. Shen, **Multi-Graph Attention Networks With Bilinear Convolution for Diagnosis of Schizophrenia**, *IEEE J. Biomed. Heal. Informatics*, vol. 27, no. 3, pp. 1443–1454, 2023, doi: 10.1109/JBHI.2022.3229465.

[22] S. M. Abd-Alhalem, H. S. Marie, W. El-Shafai, T. Altameem, R. S. Rathore, and T. M. Hassan, **Cervical cancer classification based on a bilinear convolutional neural network approach and random projection**, *Eng. Appl. Artif. Intell.*, vol. 127, p. 107261, 2024, doi: https://doi.org/10.1016/j.engappai.2023.107261.

[23] W. Liu, M. Juhas, and Y. Zhang, **Fine-Grained Breast Cancer Classification With Bilinear Convolutional Neural Networks (BCNNs)**, *Front. Genet.*, vol. 11, no. September, pp. 1–12, 2020, doi: 10.3389/fgene.2020.547327.

[24] C. Fan *et al.*, **Bilinear neural network with 3-D attention for brain decoding of motor imagery movements from the human EEG**, *Cogn. Neurodyn.*, vol. 15, no. 1, pp. 181–189, 2021, doi: 10.1007/s11571-020-09649-8.

[25] D. J. Rumala *et al.*, **Activation Functions Evaluation to Improve Performance of Convolutional Neural Network in Brain Disease Classification Based on Magnetic Resonance Images**, in *2020 International Conference on Computer Engineering, Network, and Intelligent Multimedia (CENIM)*, Nov. 2020, pp. 402–407, doi: 10.1109/CENIM51130.2020.9297862.

[26] D. R. Nayak, R. Dash, B. Majhi, and U. R. Acharya, **Application of fast curvelet Tsallis entropy and kernel random vector functional link network for automated detection of multiclass brain abnormalities**, *Comput. Med. Imaging Graph.*, vol. 77, p. 101656, Oct. 2019, doi: 10.1016/J.COMPMEDIMAG.2019.101656.

[27] M. Talo, U. B. Baloglu, Ö. Yıldırım, and U. Rajendra Acharya, **Application of deep transfer learning for automated brain abnormality classification using MR images**, *Cogn. Syst. Res.*, vol. 54, pp. 176–188, 2019, doi: https://doi.org/10.1016/j.cogsys.2018.12.007.

[28] D. R. Nayak, R. Dash, and B. Majhi, **Automated diagnosis of multi-class brain abnormalities using MRI images: A deep convolutional neural network based method**, *Pattern Recognit. Lett.*, vol. 138, pp. 385–391, 2020, doi: 10.1016/j.patrec.2020.04.018.

[29] T. K. Dutta, D. R. Nayak, and Y.-D. Zhang, **ARM-Net: Attention-guided**

**residual multiscale CNN for multiclass brain tumor classification using MR images**, *Biomed. Signal Process. Control*, vol. 87, p. 105421, 2024, doi: https://doi.org/10.1016/j.bspc.2023.105421.

[30] D. J. Rumala, P. van Ooijen, R. F. Rachmadi, A. D. Sensusiati, and I. K. E. Purnama, **Deep-Stacked Convolutional Neural Networks for Brain Abnormality Classification Based on MRI Images**, *J. Digit. Imaging*, 2023, doi: 10.1007/s10278-023-00828-7.

[31] A. Kursad Poyraz, S. Dogan, E. Akbal, and T. Tuncer, **Automated brain disease classification using exemplar deep features,** *Biomed. Signal Process. Control*, vol. 73, p. 103448, Mar. 2022, doi: 10.1016/J.BSPC.2021.103448.

[32] M. F. Siddiqui, G. Mujtaba, A. W. Reza, and L. Shuib, **Multi-class disease classification in brain MRIs using a computer-aided diagnostic system**, *Symmetry (Basel).*, vol. 9, no. 3, pp. 1–14, 2017, doi: 10.3390/sym9030037.

[33] M. Talo, O. Yildirim, U. B. Baloglu, G. Aydin, and U. R. Acharya, **Convolutional neural networks for multi-class brain disease detection using MRI images**, *Comput. Med. Imaging Graph.*, vol. 78, p. 101673, 2019, doi: 10.1016/j.compmedimag.2019.101673.

[34] J. V. Shanmugam, B. Duraisamy, B. C. Simon, and P. Bhaskaran, **Alzheimer's Disease Classification Using Pre-Trained Deep Networks**, *Biomed. Signal Process. Control*, vol. 71, p. 103217, 2022, doi: 10.1016/j.bspc.2021.103217.

[35] R. Solovyev, A. A. Kalinin, and T. Gabruseva, **3D convolutional neural networks for stalled brain capillary detection**, *Comput. Biol. Med.*, vol. 141, p. 105089, 2022, doi: 10.1016/j.compbiomed.2021.105089.

[36] S. Lu, Z. Lu, and Y. D. Zhang, **Pathological brain detection based on AlexNet and transfer learning**, *J. Comput. Sci.*, vol. 30, pp. 41–47, 2019, doi: 10.1016/j.jocs.2018.11.008.

[37] A. Gamal, M. Elattar, and S. Selim, **Automatic Early Diagnosis of Alzheimer's Disease Using 3D Deep Ensemble Approach**, *IEEE Access*, vol. 10, pp. 115974–115987, 2022, doi: 10.1109/ACCESS.2022.3218621.

[38] A. Das, S. K. Mohapatra, and M. N. Mohanty, **Design of deep ensemble classifier with fuzzy decision method for biomedical image classification**, *Appl. Soft Comput.*, vol. 115, p. 108178, Jan. 2022, doi: 10.1016/J.ASOC.2021.108178.

[39] V. Cheplygina, **Cats or CAT Scans: Transfer Learning from Natural or Medical Image Source Data Sets?**, *Curr. Opin. Biomed. Eng.*, vol. 9, pp. 21–27, Mar. 2019, doi: 10.1016/j.cobme.2018.12.005.

[40] B. H. Menze *et al.*, **The Multimodal Brain Tumor Image Segmentation Benchmark (BRATS)**, *IEEE Trans. Med. Imaging*, vol. 34, no. 10, pp. 1993–2024, 2015, doi: 10.1109/TMI.2014.2377694.

[41] S. Bakas *et al.*, **Advancing The Cancer Genome Atlas glioma MRI collections with expert segmentation labels and radiomic features**, *Sci. Data*, vol. 4, no. 1, p. 170117, 2017, doi: 10.1038/sdata.2017.117.

[42] D. Smith and F. Gaillard, **Multiple sclerosis**, *Radiopaedia.org*, May 2008, doi: 10.53347/RID-1700.
[43] D. Bell and A. Stanislavsky, **HIV/AIDS (CNS manifestations)**, *Radiopaedia.org*, Oct. 2010, doi: 10.53347/RID-11079.