

Automated Classification of Multiclass Brain Tumor MRI Images using Enhanced Deep Learning Technique

Faiz Ainur Razi¹, Alhadi Bustamam^{2*}, Arnida L. Latifah³, Shandar Ahmad⁴

Department of Mathematics, Universitas Indonesia, Depok, Indonesia^{1,2}

Data Science Center, Universitas Indonesia, Depok, Indonesia²

Research Center for Computing, National Research and Innovation Agency, Bogor, Indonesia³

School of Computational and Integrative Science, Jawaharlal Nehru University, New Delhi, India⁴

Abstract—The brain is a vital organ, and the brain tumor is one of the most dangerous types of tumors in the world. Neuroimaging is an interesting and important discussion in diagnosing central nervous system tumors. Brain tumors have several types, namely meningioma, glioma, pituitary, schwannoma, and neurocytoma. A radiologist uses magnetic resonance imaging (MRI) to detect brain tumors because of its advantages over computed tomography. However, classifying multiclass MRI is difficult and takes a long time. This study proposes an automated classification of multiclass brain tumors using enhanced deep learning techniques. Various models are used in this research, namely VGG16, NasNet-Mobile, InceptionV3, ResNet50, and EfficientNet. For EfficientNet, we applied EfficientNet-B0–B7. From the experiments, EfficientNet-B2 is the superior, with the highest level of training accuracy of 99.90%, testing accuracy of 99.55%, precision of 99.50%, recall of 99.67%, and F1-Score of 99.58% with a training time of 15 minutes. The development of this automatic classification can assist radiologists in classifying brain tumor types more efficiently.

Keywords—Brain tumor; enhanced deep learning; MRI; multiclass; neuroimaging

I. INTRODUCTION

The central nervous system (CNS), composed of the brain and spinal cord, controls all major biological systems. It consists of supporting cells (glial cells) and nerve cells (neurons) that communicate with each other and the rest of the body by sending and receiving impulses through the nerves. Magnetic resonance imaging (MRI) is an imaging technique that shows accurate anatomical images of the human body and provides valuable data for biomedical research and clinical diagnosis. For example, MRI images help diagnose brain tumors, abnormal cell growths that form an odd segment compared with normal cells. As a vital organ in the human body for speaking, thinking, and receiving environmental responses [1], any disturbance in the brain will also affect other organs. Based on the growth speed, brain tumors are classified into benign and malignant. Benign brain tumors can be cured with surgery, but malignant brain tumors are the deadliest of the cancers and can cause instant death [2]–[4]. Meanwhile, brain tumors can either be primary and secondary (i.e., metastatic). Primary tumors originate from the brain or the nerves of the brain. Metastatic brain tumors, conversely,

are caused by cancer cells that spread to the brain from other parts of the body. Clinical studies show that 30%–50% of all patients with brain metastases develop multiple lesions, depending on the type of primary cancer [5], [6].

There are three types of primary brain tumors: meningioma, glioma, and pituitary tumor. Meningiomas arise from arachnoid cells in the brain and account for 37.6% of all adult primary brain tumors. The disease accounts for approximately 35,000 new cases annually, making it the most common type of intracranial tumor in the United States [7], [8]. Gliomas are found in the cerebral pedicle and spinal cord, with symptoms such as vomiting, headache, and discomfort. Glioma tumors represent nearly 30% of primary brain tumors and 80% of all malignant ones. Based on their histopathological appearance, gliomas are traditionally classified by the World Health Organization as grades I and II (low-grade glioma), grade III (anaplastic), and grade IV (glioblastoma) [9], [10]. The pituitary is a complex organ consisting of neuroendocrine cells that secrete hormones from the adenohypophysis; posterior pituitary lobe, which is modified glia; axonal extensions of hypothalamic neurons, which secrete hormones into the bloodstream; and stromal cells, which include blood vessels, nerves, meninges, bones, and other connective tissue elements. Pituitary tumors, which arise from anterior pituitary cells and are called pituitary adenomas, are generally benign and rare (about 0.2%), showing craniospinal or systemic metastases [11]–[13]. Schwannoma (neurilemmoma) is benign neoplasms derived from tumorigenic schwann cells that protect nerve cells [14]. This condition is caused by a loss of function mutation of the neurofibromatosis type 2 (NF2) tumor gene [15]. Then, apart from that, there is also a type of brain tumor called neurocytoma. This tumor is a rare brain tumor according to the world health organization (WHO). Neurocytoma arising from the ventricle accounts for 0.1% - 0.5% of all primary brains. Moreover, these tumors rarely arise from the brain parenchyma [16].

MRI is essential in detecting brain tumors, early tumors, or CNS disorders, simultaneously seeing the response to treatment. MRI has become a part of routine clinical practice, capturing various anatomical and physiological processes [17], [18]. MRI data must be analyzed, considering that brain

*Corresponding Author, alhadi@sci.ui.ac.id

tumors generally consist of distinct structural and functional areas [19]; however, analysis can take quite a long time for accurate results. Moreover, classifying brain images of the normal brain, meningioma, glioma, pituitary, schwannoma, and neurocytoma may take longer since identifying brain images is difficult. Therefore, an automated classification of brain tumors is required to ease and speed up the analysis.

Several deep learning-based models, such as CNN, VGG (visual geometry group)-16, NasNet, and support vector machine, have been used to classify features and show reliable results [20]. Most improvements of the neural networks focused on optimizing network width, depth, and resolution. Meanwhile, the EfficientNet proposed by Tan & Le (2019) combined or collaborated those three factors. Balancing the three factors aims to obtain an optimal model at a certain complexity [21]. This study proposes an enhanced deep learning model, the EfficientNet model, to classify MRI images of brain tumors into six classes: normal brain, meningioma, glioma, pituitary, schwannoma, and neurocytoma. The EfficientNet model is expected to improve classification accuracy with less computation cost. For comparison study, four deep learning-based models namely VGG16, NasNet-Mobile, InceptionV3, and ResNet50 will also be implemented to classify MRI images.

The structure for the rest of the paper is as follows. Section II provides previous works about brain tumor classifications using various models. Section III describes the data and methods used in this study. The results of the proposed brain tumor classifications are presented in Section IV. Section V concludes the paper.

II. RELATED WORK

D. Filatov and G. N. A. H. Yar (2022) classified brain tumors into four classes, with the highest accuracy in the EfficientNet-B0 model of 87.67%, while the ResNet50 model obtained an accuracy of 72.82%. Similarly, M. A. Gómez-Guzmán et al. (2023) used the EfficientNet-B0 model and obtained a higher accuracy of 90.88%. R. Jha, V. Bhattacharjee, and A. Mustafi (2022) used TrFEMNet and obtained an accuracy of 99.39% for two classes and 78.05% for four classes of brain tumors. With the same number of classes, A. Kowshir et al. (2023) used the ResNet50 model and obtained an accuracy of 96.67%.

Using the brain MRI dataset to classify Alzheimer's, the accuracy level obtained in the Hazarika et al. study (2022) was 86.75% and 86.25% for the NasNet-A and NasNet-C models, respectively. Research conducted by [22] classified Alzheimer's using the 3D-Hog feature. Research conducted by S. R. Sowrirajan et al., (2023) using a three-class dataset, obtained accurate results for the VGG16-NADE model with an augmentation of 96.01%, and VGG16 without augmentation of 92.33%. This accuracy is the highest compared with other models in their study. The previous research summary can be seen in Table I, which shows that the highest accuracy rate is 99.39% for two classes of brain tumors. Even though the accuracy is already high, the classification is only implemented for two classes. Meanwhile, the classification of brain tumors with more than four classes has not been investigated in detail yet.

TABLE I. COMPARATIVE RESEARCH FOR BRAIN CLASSIFICATION

Reference	Model	Result
[23] D. Filatov and G. N. A. H. Yar, (2022)	EfficientNet-B7 EfficientNet-B0 ResNet50	84.19% 87.67% 72.82%
[24] R. Jha et al., (2022)	TrFEMNet	2 classes: 99.39% 4 classes: 78.05%
[25] R. A. Hazarika et al., (2022)	NasNet-A NasNet-C	86.75% 86.25%
[26] M. A. Gómez-Guzmán et al., (2023)	InceptionV3 EfficientNet-B0 Generic CNN	97.12% 90.88% 81.08%
[20] A. Kowshir et al., (2023)	InceptionV3 ResNet50 Xception	94.71% 96.67% 91.18%
[27] S. R. Sowrirajan et al., (2023)	VGG16	96.01%

In this study, the classification of brain tumors into six classes, normal brain, meningioma, glioma, pituitary, schwannoma, and neurocytoma will be proposed using five types of models, VGG16, NasNet-Mobile, InceptionV3, ResNet50, and EfficientNet. Classification of six classes will provide a better level of training and validation accuracy with shorter training time using enhanced deep learning techniques.

III. RESEARCH METHOD

A. Dataset

Fig. 1 shows the samples of the MRI images dataset, consisting of the normal brain, meningioma, glioma, pituitary, schwannoma, and neurocytoma tumor. Meningioma is the most common CNS or primary tumor. Meningioma tumors grow from the meninges, the tissues surrounding and protecting the brain just below the skull [28].

Glioma-type tumors arise from glial cells and are intraparenchymal tumors [29], [30]. An example of glioma tumor MRI results can be seen in Fig. 1 (c). The pituitary gland is a complex organ composed of the hormone-secreting neuroendocrine cells of the pituitary gland. The MRI results for this type of tumor can be seen in Fig. 1 (d). In Fig. 1 (e) is an image of a schwannoma-type tumor and the MRI result of a neurocytoma-type tumor shown in Fig. 1 (f).

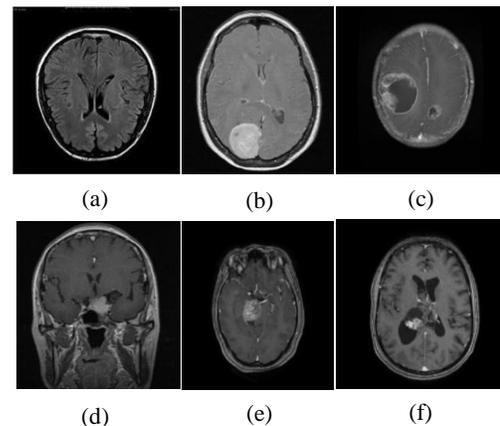


Fig. 1. MRI result (a) normal brain, (b) meningioma, (c) glioma, (d) pituitary, (e) schwannoma, (f) neurocytoma.

The dataset used in this study is retrieved from [31] and [32], and there are six types of images with a total of 7519 images. The number of details of the dataset can be seen in Table II.

TABLE II. DATASET DISTRIBUTION

Phase	Train	Test	Total
Normal	1595	205	1800
Glioma	1321	167	1488
Meningioma	1339	159	1498
Pituitary	1457	163	1620
Schwannoma	463	88	551
Neurocytoma	457	105	562
Total	6632	887	7519
	88.2%	11.8%	

B. VGG16 Architecture

VGG is one of the CNN architectures. VGG16 has five convolution blocks with 13 convolution layers and 3 fully connected. VGG is more capable of processing small datasets and has better recognition efficiency [33], [34]. During the training process, a loss function is used to measure the error between predicted and actual values. The following formula was applied for the cross-entropy loss function for experiments with VGG16:

$$L_{CE} = - \sum_{i=1}^n t_i \log p_i, \quad \text{for } n \text{ classes} \quad (1)$$

Where t_i is truth label and p_i denotes the softmax loss functions for i^{th} class. The VGG16 architecture can be seen in Fig. 2.

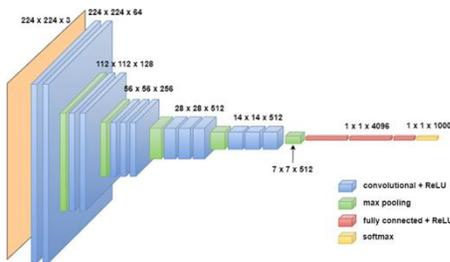


Fig. 2. VGG16 architecture [35].

C. NasNet-Mobile

NasNet is one of the CNN architectures consisting of basic building blocks optimized using reinforcement learning [36]–[38]. There are two blocks that must be considered, namely the child block and the parental block. The child block serves to adjust the network based on changes in effectiveness, while the parental block serves to evaluate the effectiveness of the child block.

NasNet development defines a high-performance building block in image set categorization (CIFAR-10). The block is generalized to a wider dataset so that it can achieve a higher classification capacity [39]. The illustration for this model can be seen in Fig. 3.

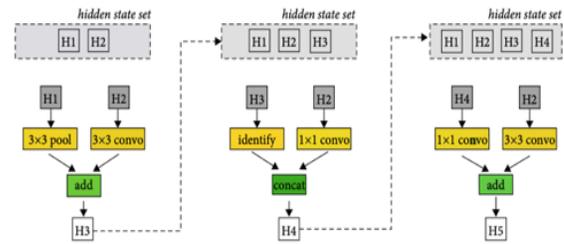


Fig. 3. NasNet architecture [55].

D. InceptionV3

InceptionV3 is the latest version of the InceptionV1 model. The InceptionV3 model has a wider network than InceptionV1 and V2. Training takes longer and is very difficult to complete. This problem is solved using transfer learning techniques [40], [41].

The structure of InceptionV3 can be seen in Fig. 4. In InceptionV3, the probability of each label $k \in \{1, \dots, K\}$ can be determined by

$$Q(k|z) = \frac{\exp(y_k)}{\sum_i^K \exp(y_i)} \quad (2)$$

where, y denotes the nonnormalized log probability. The ground truth distribution on labels $p(k|z)$ is normalized by $\sum_k p(k|z) = 1$.

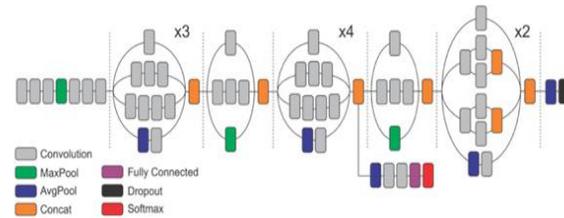


Fig. 4. InceptionV3 structure [54].

E. ResNet50

The ResNet50 model is a CNN model with a 50-layer residual network partitioned into five parts. The first part contains a convolutional layer for input preprocessing. Part 2–5 contain the bottleneck components. This model was first introduced by Microsoft in 2015 [42]. Residual building blocks can be shown in the following formula:

$$y = F(x) + x \quad (3)$$

where, $F(x)$ is the residual function, x is the input, and y is the output parameter of the residual function. ResNet50 architecture can be seen in Table III [43].

F. EfficientNet

EfficientNet has eight different architectures, namely EfficientNet-B0–B7 with the basic model being B0 obtained from neural architecture search (NAS) and B1–B7, which is an additional model with an extension of the basic model [44]. In NAS [45] to get the optimal architecture, a controller is needed to maximize the expected results represented by $J(\theta_c)$:

$$J(\theta_c) = E_{P(\alpha_{1:T}; \theta_c)}[R] \quad (4)$$

TABLE III. RESNET50 ARCHITECTURE

Layer Name	Output Size	Layer
Conv1	112 x 112	7 x 7, 64, stride 2
Conv2_x	56 x 56	3 x 3 max pool, stride 2
		$\begin{bmatrix} 1 \times 1, 64 \\ 3 \times 3, 64 \\ 1 \times 1, 256 \end{bmatrix} \times 3$
Conv3_x	28 x 28	$\begin{bmatrix} 1 \times 1, 128 \\ 3 \times 3, 128 \\ 1 \times 1, 512 \end{bmatrix} \times 4$
Conv4_x	14 x 14	$\begin{bmatrix} 1 \times 1, 256 \\ 3 \times 3, 256 \\ 1 \times 1, 1024 \end{bmatrix} \times 6$
Conv5_x	7 x 7	$\begin{bmatrix} 1 \times 1, 512 \\ 3 \times 3, 512 \\ 1 \times 1, 2048 \end{bmatrix} \times 3$
	1 x 1	Average pool, 1000-d fc, softmax

List of action $\alpha_{1:T}$ is used to design the child network architecture. This child network will achieve accuracy R at the time of dataset convergence. By taking advantage of the R accuracy, it can be used as a reward signal for training controllers or as reinforcement learning. The reward signal R is not differentiable, so a gradient method is required to iteratively update θ_c by using the reinforce rule [46]:

$$\nabla_{\theta_c} J(\theta_c) = \sum_{t=1}^T E_{P(\alpha_{1:T}; \theta_c)} [\nabla_{\theta_c} \log P(a_t | a_{(t-1):1}; \theta_c) R] \quad (5)$$

The empirical approximation of the quantity (4) is:

$$\frac{1}{m} \sum_{k=1}^m \sum_{t=1}^T \nabla_{\theta_c} \log P(a_t | a_{(t-1):1}; \theta_c) R_k \quad (6)$$

where, m is the number of different architectures containing example controllers in the stack. On the other hand, T is the number of hyperparameters expected by the controller to design the neural network architecture.

The validation accuracy obtained by the k^{th} neural network architecture after being trained on the training dataset is R_k . To reduce variance, the baseline function is used:

$$\frac{1}{m} \sum_{k=1}^m \sum_{t=1}^T \nabla_{\theta_c} \log P(a_t | a_{(t-1):1}; \theta_c) (R_k - b) \quad (7)$$

EfficientNet is based on NAS technology as a simple, scalable, and generalizable benchmark network. In increasing the resolution and complexity of the network structure, the architecture of the EfficientNet model can be seen in Table IV [47].

By combining the three factors in the architecture, the coefficient calculation formula is as follows [48], [49]:

$$\begin{cases} \text{depth: } d = \alpha^\phi \\ \text{width: } w = \beta^\phi \\ \text{resolution: } r = \gamma^\phi \end{cases} \quad (8)$$

TABLE IV. ARCHITECTURE OF EFFICIENT NET

Description	Layer	Input Resolution	Channel
EfficientNet-B0	240	224x224	1280
EfficientNet-B1	342	240x240	1280
EfficientNet-B2	342	260x260	1408
EfficientNet-B3	387	300x300	1536
EfficientNet-B4	477	380x380	1792
EfficientNet-B5	579	456x456	2048
EfficientNet-B6	669	528x528	2304
EfficientNet-B7	816	600x600	2560

where, $\alpha \geq 1, \beta \geq 1, \gamma \geq 1$ and $\alpha \cdot \beta^2 \cdot \gamma^2 \approx 2$ with $w, d,$ and r can be used to scale network width, depth, and resolution coefficients. While value ϕ can be used to determine the number of effective resource extension models. The constants $\alpha, \beta,$ and γ are used to allocate these resources into three-dimensional network depth, width, and resolutions.

In the proposed model, the preprocessing stage will carry out a cropping process with an image size of 150×150 . This aims to eliminate noise or delete unnecessary image information. The results of the cropping process can be seen in Fig. 5. After the cropping process is carried out, the new data will be saved in a new directory. Then the data will be augmented. This data augmentation functions to suppress overfitting when data is run through artificial data augmentation techniques [50].

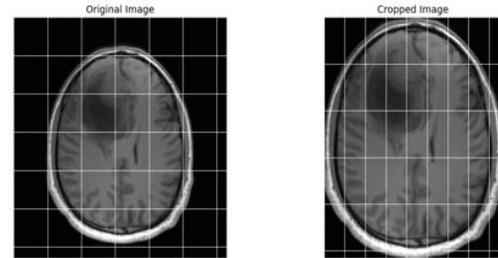


Fig. 5. Original and cropped image.

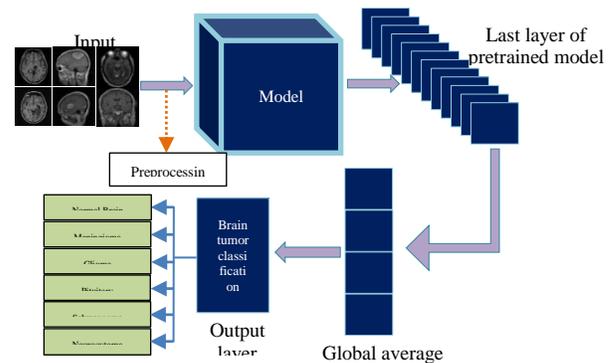


Fig. 6. Proposed model for brain tumor classification.

The complete model proposed in this study is shown in Fig. 6. The input model is the MRI image, which results in the image's class; it is detected as a normal brain or one of the brain tumors: meningioma, glioma, pituitary, schwannoma, or neurocytoma. Furthermore, the structure of the EfficientNet model is illustrated in Fig. 7.

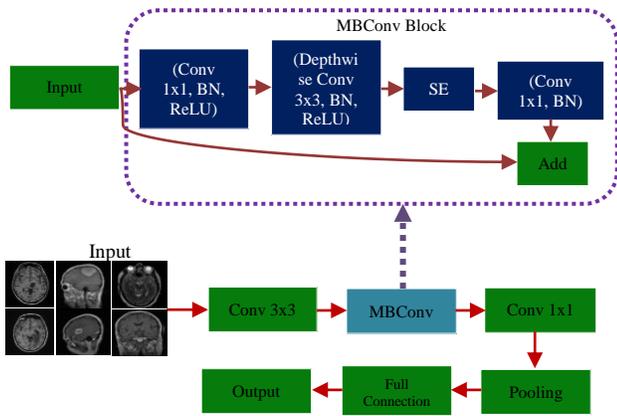


Fig. 7. The structure of the efficientnet model.

G. Model Implementation

A depth wise separable convolution (DSC) layer is used to build the base of the MobileNet block. Therefore, this hierarchical structure is also called mobile convolution (MB Conv) [51]. DSC consists of two parts: depth convolution (DWC) and point convolution (PWC). The combined convolution process between DWC and PWC is shown in Fig. 8. The goal of revealing model parameters while preserving output quality was achieved.

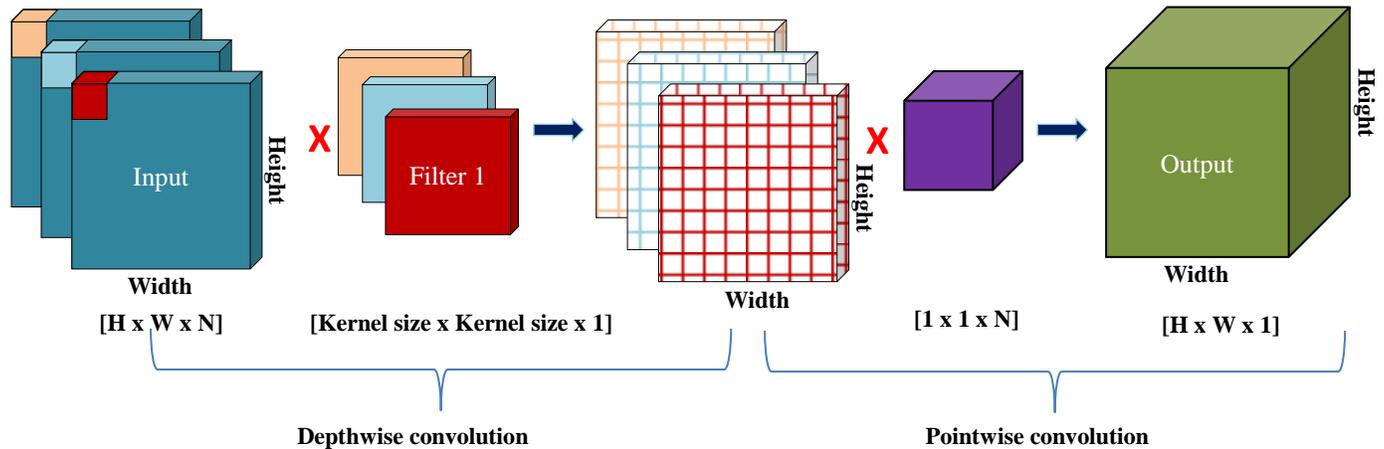


Fig. 8. Depthwise and pointwise convolution illustration.

		Predicted Class			
		C_1	C_2	...	C_N
Actual Class	C_1	$C_{1,1}$	FP	...	$C_{1,N}$
	C_2	FN	TP	...	FN

	C_N	$C_{N,1}$	FP	...	$C_{N,N}$

Fig. 9. Multiclass classification.

IV. RESULT

This research conducted all experiments in the Google Colab application, with the graphic processing unit backend

H. Model Evaluation

Confusion matrix provides a combination of class and actual predictions. This makes it possible to define various multiclass performance metrics as shown in Fig. 6 [52], [53]. The multiclass confusion matrix presented in Fig. 9 has dimensions of $N \times N$, where N is the number of different class labels C_0, C_1, \dots, C_N . From the confusion matrix, we can compute the classification metrics: accuracy, recall, precision, and the F1-score by formulas presented in Table V.

TABLE V. PERFORMANCE METRICS FOR MULTICLASS CLASSIFICATION

Metric	Formula
Accuracy	$\frac{\sum_{i=1}^N TP(C_i)}{\sum_{i=1}^N \sum_{j=1}^N C_{i,j}}$
Recall of class C_i ($TPR(C_i)$)	$\frac{TP(C_i)}{TP(C_i) + FN(C_i)}$
Precision of class C_i ($PPV(C_i)$)	$\frac{TP(C_i)}{TP(C_i) + FP(C_i)}$
F_1 - Score of class C_i	$2 \cdot \frac{TPR(C_i) \cdot PPV(C_i)}{TPR(C_i) + PPV(C_i)}$

Google Compute Engine Python 3 A100. The RAM was 83.5 GB, disk 166.8 GB, model name Intel(R) Xeon(R) CPU@2.20GHz. The detailed parameters that have been optimized of the model experiments are shown in Table VI.

Except for the number of epochs, all models used the same parameters and loss function.

TABLE VI. PARAMETERS USED IN THE PROPOSED BRAIN TUMOR CLASSIFICATION FRAMEWORK

Hyper-parameter	VGG16	NasNet-Mobile	InceptionV3	ResNet50	EfficientNet
Optimizer	Adam	Adam	Adam	Adam	Adam
Batch Size	32	32	32	32	32
Epoch	80	40	20	20	20
Learning Rate	0.0001	0.0001	0.0001	0.0001	0.0001
Verbose	1	1	1	1	1
Loss Function	Flatten Cross-Entropy				

First, we will show the results of four models' experiments: VGG16, NasNet-Mobile, InceptionV3 and ResNet50. Initially, the performance of the training and validation of four models are presented through the loss and accuracy functions in Fig. 10. Meanwhile, Fig. 11 shows the corresponding confusion matrix that we can see that InceptionV3 only gives few false results.

Except in the VGG16 model, there is a jump in both loss and accuracy graphs of the models (see Fig. 10) (b-d). Nevertheless, as the epoch is larger the accuracy tends to improve, and the loss progressively reduces. In the accuracy graph, the initial validation accuracy is very low, 0.2 in VGG16 and NasNet-Mobile, and 0.4 in ResNet50. However, it increases to approximately 95% after the epoch is larger than 20 in the model of VGG16 and NasNet-Mobile. In InceptionV3 and ResNet50, the accuracy is already consistently more than 95% after 10 epochs. The highest accuracy of each model as shown in Table VII is achieved with epoch 24, 40, 13, and 7 for VGG16, NasNet-Mobile, InceptionV3 and ResNet50, respectively. Based on the performance results of the four models presented in Table VII, InceptionV3 achieves the best with only 13 epochs. Further, we will investigate how EfficientNet performs compared to InceptionV3, in particular.

Like the four other models, the experiments using eight varieties of EfficientNet (B0-B7) will also be compared through the loss and accuracy trend functions, confusion matrix, and the model performance metrics. Fig. 12 depicts the loss and accuracy graph of all EfficientNet models over 20 epochs. As expected, the performance of the loss and accuracy of the training are slightly better than the validation. All EfficientNet models tend to increase their performance as the epoch grows. Of the eight EfficientNet models, their performances do not differ significantly. All accuracies are mostly perfect, approximately 99%.

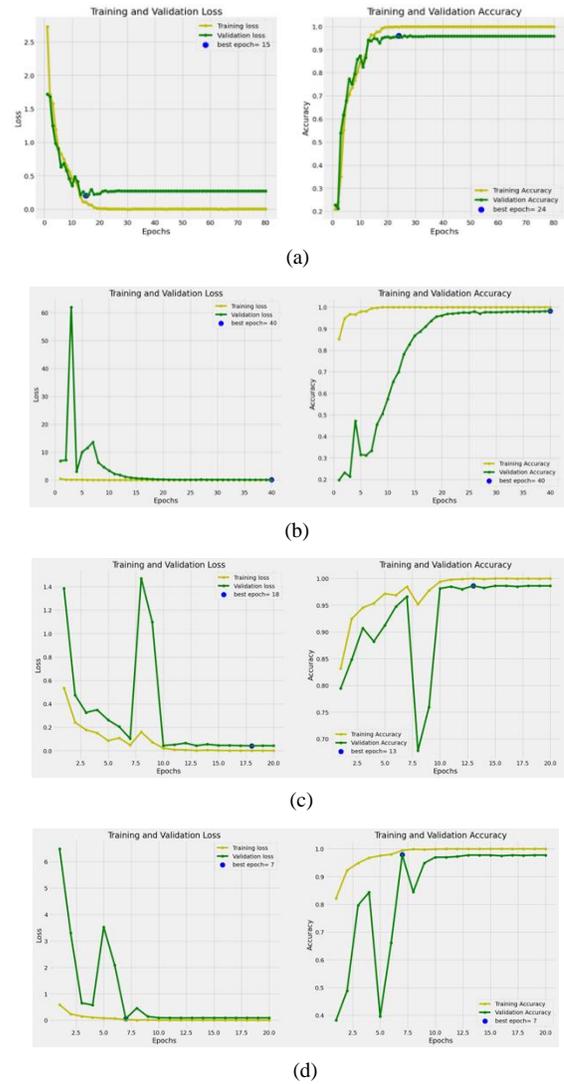


Fig. 10. The loss function and accuracy of the experiment using (a) VGG16, (b) NasNet-Mobile, (c) InceptionV3, and (d) ResNet50.

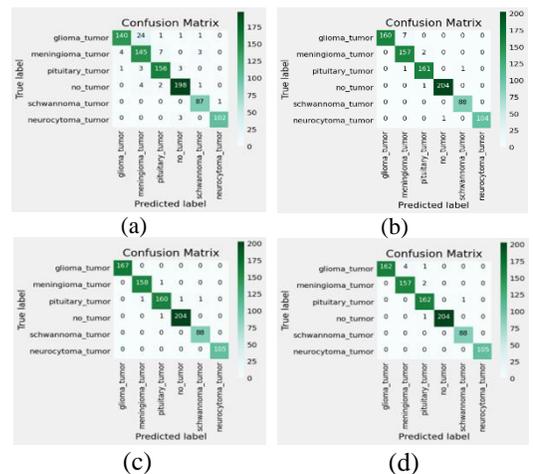


Fig. 11. The confusion matrix of experiment (a) VGG16, (b) NasNet-Mobile, (c) InceptionV3, and (d) ResNet50.

TABLE VII. PERFORMANCE MODEL

Model	Class	Precision (%)	Recall (%)	F1-Score (%)
VGG16	Glioma	97.00	84.00	90.03
	Meningioma	82.00	91.00	86.27
	Pituitary	94.00	96.00	94.99
	Normal	97.00	97.00	97.00
	Schwannoma	95.00	99.00	96.96
	Neurocytoma	99.00	97.00	97.99
	Average	94.00	94.00	94.00
NasNet-Mobile	Glioma	100.00	96.00	97.96
	Meningioma	95.00	99.00	96.96
	Pituitary	98.00	99.00	98.50
	Normal	100.00	100.00	100.00
	Schwannoma	99.00	100.00	99.50
	Neurocytoma	100.00	99.00	99.50
	Average	98.67	98.83	98.75
InceptionV3	Glioma	100.00	100.00	100.00
	Meningioma	99.00	99.00	99.00
	Pituitary	99.00	98.00	98.50
	Normal	100.00	100.00	100.00
	Schwannoma	99.00	100.00	99.50
	Neurocytoma	100.00	100.00	100.00
	Average	99.50	99.50	99.50
ResNet50	Glioma	100.00	97.00	98.48
	Meningioma	98.00	99.00	98.50
	Pituitary	98.00	99.00	98.50
	Normal	100.00	100.00	100.00
	Schwannoma	99.00	100.00	99.50
	Neurocytoma	100.00	100.00	100.00
	Average	99.17	99.17	99.16

From the confusion matrix described in Fig. 13, we observe that all models only deliver a few faulty classifications. The false positives seem to be a little more than the false negatives results. There is no false negative result in EfficientNet-B1 and there is only one false negative in EfficientNet-B2, B3, and B5 while there are two false negative results in InceptionV3.

Further detail of the model performance of EfficientNet can be seen from the performance metrics of precision, recall, F1-score in Table VIII. The metrics of all EfficientNet are shown to be more than 99%.

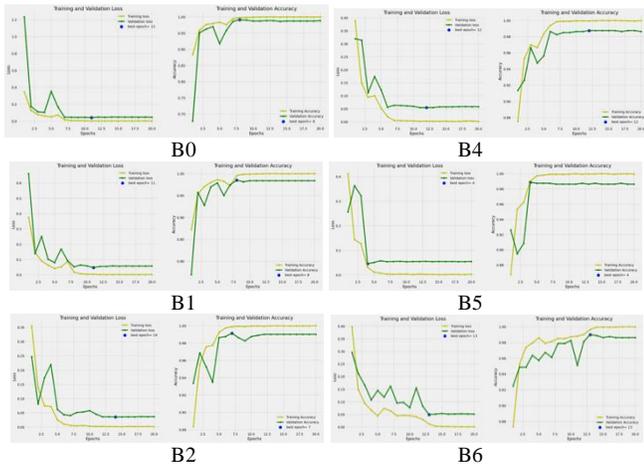


Fig. 12. The loss function and accuracy of the experiment using EfficientNet-B0-B7.

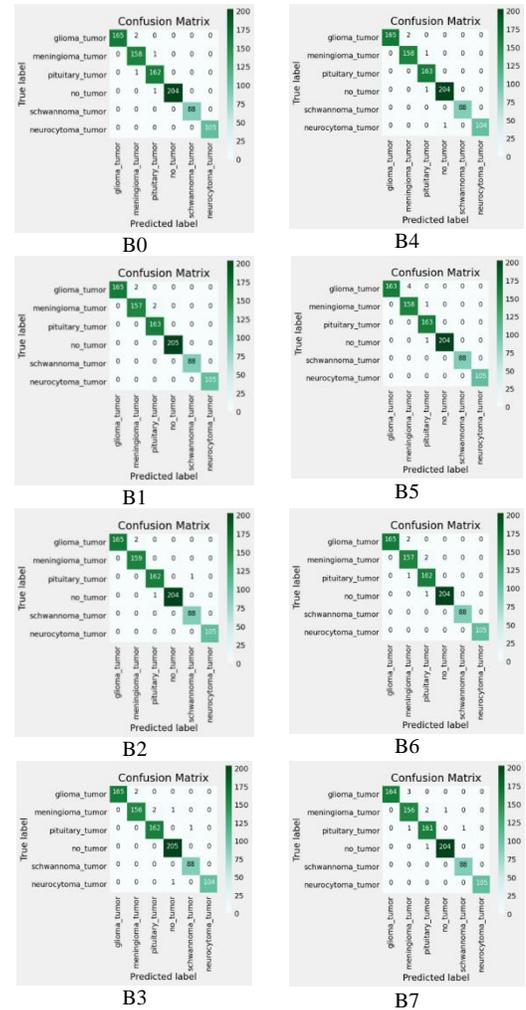


Fig. 13. The confusion matrix of the experiment using EfficientNet-B0-B7.

TABLE VIII. PERFORMANCE OF THE EFFICIENTNET MODEL

Model	Class	Precision (%)	Recall (%)	F1-Score (%)
EfficientNet-B0	Glioma	100.00	99.00	99.50
	Meningioma	98.00	99.00	98.50
	Pituitary	99.00	99.00	99.00
	Normal	100.00	100.00	100.00
	Schwannoma	100.00	100.00	100.00
	Neurocytoma	100.00	100.00	100.00

Model	Class	Precision (%)	Recall (%)	F1-Score (%)
	Average	99.50	99.50	99.50
EfficientNet-B1	Glioma	100.00	99.00	99.50
	Meningioma	99.00	99.00	99.00
	Pituitary	99.00	100.00	99.50
	Normal	100.00	100.00	100.00
	Schwannoma	100.00	100.00	100.00
	Neurocytoma	100.00	100.00	100.00
	Average	99.67	99.67	99.67
	EfficientNet-B2	Glioma	100.00	99.00
Meningioma		99.00	100.00	99.50
Pituitary		99.00	99.00	99.00
Normal		100.00	100.00	100.00
Schwannoma		99.00	100.00	99.50
Neurocytoma		100.00	100.00	100.00
Average		99.50	99.67	99.58
EfficientNet-B3	Glioma	100.00	99.00	99.50
	Meningioma	99.00	98.00	98.50
	Pituitary	99.00	99.00	99.00
	Normal	99.00	100.00	99.50
	Schwannoma	99.00	100.00	99.50
	Neurocytoma	100.00	99.00	99.50
	Average	99.33	99.17	99.25
EfficientNet-B4	Glioma	100.00	99.00	99.50
	Meningioma	99.00	99.00	99.00
	Pituitary	99.00	100.00	99.50
	Normal	100.00	100.00	100.00
	Schwannoma	100.00	100.00	100.00
	Neurocytoma	100.00	99.00	99.50
	Average	99.67	99.50	99.58
EfficientNet-B5	Glioma	100.00	98.00	98.99
	Meningioma	98.00	99.00	98.50
	Pituitary	99.00	100.00	99.50
	Normal	100.00	100.00	100.00
	Schwannoma	100.00	100.00	100.00
	Neurocytoma	100.00	100.00	100.00
	Average	99.50	99.50	99.50
EfficientNet-B6	Glioma	100.00	99.00	99.50
	Meningioma	98.00	99.00	98.50
	Pituitary	98.00	99.00	98.50
	Normal	100.00	100.00	100.00
	Schwannoma	100.00	100.00	100.00
	Neurocytoma	100.00	100.00	100.00
	Average	99.33	99.50	99.42
EfficientNet-	Glioma	100.00	98.00	98.99

Model	Class	Precision (%)	Recall (%)	F1-Score (%)
B7	Meningioma	97.00	98.00	97.50
	Pituitary	98.00	99.00	98.50
	Normal	100.00	100.00	100.00
	Schwannoma	99.00	100.00	99.50
	Neurocytoma	100.00	100.00	100.00
	Average	99.00	99.17	99.08

Next, we compare the EfficientNet model with four previous models. The models' performances are evaluated by considering the model accuracy and the computational time, the time consuming for training. The model comparisons are presented in Table IX. Based on the model accuracy in both the training and the validation data, the EfficientNet can outperform NasNet-Mobile, VGG16, InceptionV3, and ResNet50. EfficientNet-B2 achieves the highest accuracy of 99.9% in training and 99.55% in validation. The model with the least accuracy is VGG16, with a training accuracy of 97.2% and validation accuracy of 93.35%. This model also had the longest training time of 84 minutes with 80 epochs. The selection of more epochs is due to the stability of the chart. Compared with other proposed models, VGG16 requires a long epoch to be stable. In terms of the computation time, the proposed EfficientNet model requires vary time, from 10 minute in EfficientNet-B0 up to 60 minutes in EfficientNet-B7. Meanwhile, the best model, EfficientNet-B2 requires 15 minutes for the training. It is almost double the time of InceptionV3 model which only takes eight minutes.

TABLE IX. ACCURACY AND COMPUTATIONAL TIME OF MODEL

Model	Train (%)	Validation (%)	Time (m)
EfficientNet-B0	99.89	99.44	10
EfficientNet-B1	99.84	99.55	14
EfficientNet-B2	99.90	99.55	15
EfficientNet-B3	99.80	99.21	19
EfficientNet-B4	99.86	99.44	25
EfficientNet-B5	99.86	99.32	35
EfficientNet-B6	99.86	99.32	46
EfficientNet-B7	99.85	98.99	60
VGG16	97.20	93.35	84
NasNet-Mobile	99.82	98.53	26
ResNet50	99.77	98.99	14
InceptionV3	99.86	99.44	8

V. CONCLUSION

This study applies eight enhanced EfficientNet models, namely EfficientNet-B0-B7, which is based on the concept of CNN. According to existing literature, EfficientNet is a deep learning model that modifies the model so that computational efficiency produces the best results. With its efficiency advantage, we use the model to build an automated classification of brain MRI images into six classes: normal, meningioma, glioma, pituitary, schwannoma, and

neurocytoma. In this study, the EfficientNet models are also compared to the previous models, namely VGG16, NasNet-Mobile, InceptionV3, and ResNet50. All varieties of EfficientNet perform high accuracy and the EfficientNet-B2 model is superior. The EfficientNet-B2 model achieves the highest training accuracy of 99.9% and validation accuracy of 99.55%. However, it takes a slightly longer time to do the training. It requires 15 minutes, while InceptionV3 only needs eight minutes to achieve a training accuracy of 99.86%. Both EfficientNet-B2 and InceptionV3 models are best options in classifying brain MRI images efficiently and accurately.

ACKNOWLEDGMENT

This research is funded by Directorate of Research and Development, Universitas Indonesia under research grant PUTI 2023-2024 with contract number NKB-723/UN2.RST/HKP.05.00/2023.

REFERENCES

- [1] K. S. Ayomide, T. Noranis, M. Aris, and M. Zolkepli, "Improving Brain Tumor Segmentation in MRI Images through Enhanced Convolutional Neural Networks," vol. 14, no. 4, pp. 670–678, 2023.
- [2] X. Gu, Z. Shen, J. Xue, Y. Fan, and T. Ni, "Brain Tumor MR Image Classification Using Convolutional Dictionary Learning With Local Constraint," *Front. Neurosci.*, vol. 15, no. May, pp. 1–12, 2021, doi: 10.3389/fnins.2021.679847.
- [3] C. Ge, I. Y. H. Gu, A. S. Jakola, and J. Yang, "Deep semi-supervised learning for brain tumor classification," *BMC Med. Imaging*, vol. 20, no. 1, pp. 1–11, 2020, doi: 10.1186/s12880-020-00485-0.
- [4] F. A. Razi, A. Bustamam, and A. L. Latifah, "Development of Efficient Brain Tumor Classification on MRI Image Results Using EfficientNet," 2023 *Int. Semin. Intell. Technol. Its Appl.*, pp. 575–580, 2023, doi: 10.1109/ISITIA59021.2023.10221186.
- [5] M. A. Proescholdt et al., "The management of brain metastases—systematic review of neurosurgical aspects," *Cancers (Basel)*, vol. 13, no. 7, pp. 1–17, 2021, doi: 10.3390/cancers13071616.
- [6] H. Sung et al., "Global Cancer Statistics 2020: GLOBOCAN Estimates of Incidence and Mortality Worldwide for 36 Cancers in 185 Countries," *CA. Cancer J. Clin.*, vol. 71, no. 3, pp. 209–249, 2021, doi: 10.3322/caac.21660.
- [7] K. Huntoon, A. M. S. Toland, and S. Dahiya, "Meningioma: A Review of Clinicopathological and Molecular Aspects," *Front. Oncol.*, vol. 10, no. October, pp. 1–14, 2020, doi: 10.3389/fonc.2020.579599.
- [8] J. Driver et al., "A molecularly integrated grade for meningioma," *Neuro. Oncol.*, vol. 24, no. 5, pp. 796–808, 2022, doi: 10.1093/neuonc/noab213.
- [9] M. M. Shaver et al., "Optimizing neuro-oncology imaging: A review of deep learning approaches for glioma imaging," *Cancers (Basel)*, vol. 11, no. 6, pp. 1–14, 2019, doi: 10.3390/cancers11060829.
- [10] L. Rong, N. Li, and Z. Zhang, "Emerging therapies for glioblastoma: current state and future directions," *J. Exp. Clin. Cancer Res.*, vol. 41, no. 1, pp. 1–18, 2022, doi: 10.1186/s13046-022-02349-7.
- [11] S. L. Asa, O. Mete, A. Perry, and R. Y. Osamura, "Overview of the 2022 WHO Classification of Pituitary Tumors," *Endocr. Pathol.*, vol. 33, no. 1, pp. 6–26, 2022, doi: 10.1007/s12022-022-09703-7.
- [12] L. Kasuki and G. Raverot, "Definition and diagnosis of aggressive pituitary tumors," *Rev. Endocr. Metab. Disord.*, vol. 21, no. 2, pp. 203–208, 2020, doi: 10.1007/s11154-019-09531-x.
- [13] S. L. Asa et al., "From pituitary adenoma to pituitary neuroendocrine tumor (pitnet): An international pituitary pathology club proposal," *Endocr. Relat. Cancer*, vol. 24, no. 4, pp. C5–C8, 2017, doi: 10.1530/ERC-17-0004.
- [14] J. Gosk, O. Gutkowska, M. Urban, W. Wnukiewicz, P. Reichert, and P. Ziolkowski, "Results of surgical treatment of schwannomas arising from extremities," *Biomed Res. Int.*, vol. 2015, 2015, doi: 10.1155/2015/547926.
- [15] D. L. Helbing, A. Schulz, and H. Morrison, "Pathomechanisms in schwannoma development and progression," *Oncogene*, vol. 39, no. 32, pp. 5421–5429, 2020, doi: 10.1038/s41388-020-1374-5.
- [16] R. Mohamed et al., "Clinicopathological features and treatment outcome of central neurocytoma: a single institute experience," *Egypt. J. Neurol. Psychiatry Neurosurg.*, vol. 58, no. 1, 2022, doi: 10.1186/s41983-022-00540-3.
- [17] J. Kalpathy-Cramer, E. R. Gerstner, K. E. Emblem, O. Andronesi, and B. Rosen, "Advanced Magnetic Resonance Imaging of the Physical Processes in Human Glioblastoma," *NIH-PA Author Manuscr.*, vol. 74, no. 17, pp. 4622–4637, 2015, doi: 10.1158/0008-5472.CAN-14-0383.Advanced.
- [18] A. I. Neugut et al., "Magnetic Resonance Imaging-Based Screening for Asymptomatic Brain Tumors: A Review," *Oncologist*, vol. 24, no. 3, pp. 375–384, 2019, doi: 10.1634/theoncologist.2018-0177.
- [19] D. Aquino, A. Gioppo, G. Finocchiaro, M. G. Bruzzone, and V. Cuccarini, "MRI in Glioma Immunotherapy: Evidence, Pitfalls, and Perspectives," *J. Immunol. Res.*, vol. 2017, 2017, doi: 10.1155/2017/5813951.
- [20] A. Kowshir, H. Imam, S. Yesmin, and I. Mahmud, "Tumor-Net : convolutional neural network modeling for classifying brain tumors from MRI images," vol. 9, no. 2, pp. 148–160, 2023.
- [21] M. Tan and Q. V. Le, "EfficientNet: Rethinking model scaling for convolutional neural networks," 36th *Int. Conf. Mach. Learn. ICML 2019*, vol. 2019-June, pp. 10691–10700, 2019.
- [22] D. Sarwinda and A. Bustamam, "3D-HOG Features-Based Classification using MRI Images to Early Diagnosis of Alzheimer's Disease," *Proc. - 17th IEEE/ACIS Int. Conf. Comput. Inf. Sci. ICIS 2018*, pp. 457–462, 2018, doi: 10.1109/ICIS.2018.8466524.
- [23] D. Filatov and G. N. A. H. Yar, "Brain Tumor Diagnosis and Classification via Pre-Trained Convolutional Neural Networks," pp. 0–5, 2022, [Online]. Available: <http://arxiv.org/abs/2208.00768>.
- [24] R. Jha, V. Bhattacharjee, and A. Mustafi, "Transfer Learning with Feature Extraction Modules for Improved Classifier Performance on Medical Image Data," *Sci. Program.*, vol. 2022, 2022, doi: 10.1155/2022/4983174.
- [25] R. A. Hazarika, D. Kandar, and A. K. Maji, "An experimental analysis of different Deep Learning based Models for Alzheimer's Disease classification using Brain Magnetic Resonance Images," *J. King Saud Univ. - Comput. Inf. Sci.*, vol. 34, no. 10, pp. 8576–8598, 2022, doi: 10.1016/j.jksuci.2021.09.003.
- [26] M. A. Gómez-Guzmán et al., "Classifying Brain Tumors on Magnetic Resonance Imaging by Using Convolutional Neural Networks," *Electron.*, vol. 12, no. 4, pp. 1–22, 2023, doi: 10.3390/electronics12040955.
- [27] S. R. Sowrirajan, S. Balasubramanian, and R. S. P. Raj, "MRI Brain Tumor Classification Using a Hybrid VGG16-NADE Model," *Brazilian Arch. Biol. Technol.*, vol. 66, no. Mcc, 2023, doi: 10.1590/1678-4324-2023220071.
- [28] C. Ogasawara, B. D. Philbrick, and D. C. Adamson, "Meningioma: A review of epidemiology, pathology, diagnosis, treatment, and future directions," *Biomedicine*, vol. 9, no. 3, 2021, doi: 10.3390/biomedicine9030319.
- [29] A. Lasocki, M. Anjari, S. Örs Kokuurcan, and S. C. Thust, "Conventional MRI features of adult diffuse glioma molecular subtypes: a systematic review," *Neuroradiology*, vol. 63, no. 3, pp. 353–362, 2021, doi: 10.1007/s00234-020-02532-7.
- [30] Q. D. Buchlak, N. Esmaili, J. C. Leveque, C. Bennett, F. Farrokhi, and M. Piccardi, "Machine learning applications to neuroimaging for glioma detection and classification: An artificial intelligence augmented systematic review," *J. Clin. Neurosci.*, vol. 89, pp. 177–198, 2021, doi: 10.1016/j.jocn.2021.04.043.
- [31] M. Nickparvar, "Brain Tumor MRI Dataset," 2021. <https://www.kaggle.com/datasets/masoudnickparvar> (accessed May 18, 2023).
- [32] F. Feltrin, "Brain Tumor MRI Images 44 Classes," 2023. <https://www.kaggle.com/datasets/fernando2rad> (accessed Jun. 01, 2023).

- [33] F. Zhao, B. Zhang, Z. Zhang, X. Zhang, and C. Wei, "Classification and detection method of Blood lancet based on VGG16 network," 2021 IEEE Int. Conf. Mechatronics Autom. ICMA 2021, pp. 849–853, 2021, doi: 10.1109/ICMA52036.2021.9512686.
- [34] K. Simonyan and A. Zisserman, "Very deep convolutional networks for large-scale image recognition," 3rd Int. Conf. Learn. Represent. ICLR 2015 - Conf. Track Proc., pp. 1–14, 2015.
- [35] C. Xia, M. Wang, Y. Fan, Z. Yang, and X. Du, "A hierarchical autoencoder and temporal convolutional neural network reduced-order model for the turbulent wake of a three-dimensional bluff body," no. August 2015, 2023, doi: 10.1063/5.0137285.
- [36] F. Saxe, P. Werner, S. Handrich, E. Othman, L. Dinges, and A. Al-Hamadi, "Face attribute detection with mobilenetv2 and nasnet-mobile," Int. Symp. Image Signal Process. Anal. ISPA, vol. 2019-Sept, pp. 176–180, 2019, doi: 10.1109/ISPA.2019.8868585.
- [37] B. Zoph, V. Vasudevan, J. Shlens, and Q. V. Le, "Learning Transferable Architectures for Scalable Image Recognition," Proc. IEEE Comput. Soc. Conf. Comput. Vis. Pattern Recognit., pp. 8697–8710, 2018, doi: 10.1109/CVPR.2018.00907.
- [38] P. M. Shah et al., "DC-GAN-based synthetic X-ray images augmentation for increasing the performance of EfficientNet for COVID-19 detection," no. August 2021, pp. 1–13, 2022, doi: 10.1111/exsy.12823.
- [39] F. Martínez, F. Martínez, and E. Jacinto, "Performance evaluation of the NASnet convolutional network in the automatic identification of COVID-19," Int. J. Adv. Sci. Eng. Inf. Technol., vol. 10, no. 2, pp. 662–667, 2020, doi: 10.18517/ijaseit.10.2.11446.
- [40] M. Mujahid, F. Rustam, R. Álvarez, J. Luis Vidal Mazón, I. de la T. Diez, and I. Ashraf, "Pneumonia Classification from X-ray Images with Inception-V3 and Convolutional Neural Network," Diagnostics, vol. 12, no. 5, pp. 1–16, 2022, doi: 10.3390/diagnostics12051280.
- [41] C. Szegedy, V. Vanhoucke, S. Ioffe, J. Shlens, and Z. Wojna, "Rethinking the Inception Architecture for Computer Vision," Proc. IEEE Comput. Soc. Conf. Comput. Vis. Pattern Recognit., vol. 2016-Decem, pp. 2818–2826, 2016, doi: 10.1109/CVPR.2016.308.
- [42] K. He, X. Zhang, S. Ren, and J. Sun, "Deep Residual Learning for Image Recognition Kaiming," in Proceedings of the IEEE conference on computer vision and pattern recognition, 2015, vol. 45, no. 8, pp. 770–778, doi: 10.1002/chin.200650130.
- [43] R. Zhang, Y. Zhu, Z. Ge, H. Mu, D. Qi, and H. Ni, "Transfer Learning for Leaf Small Dataset Using Improved ResNet50 Network with Mixed Activation Functions," Forests, vol. 13, no. 12, 2022, doi: 10.3390/f13122072.
- [44] X. Ma, X. Li, A. Luo, J. Zhang, and H. Li, "Galaxy image classification using hierarchical data learning with weighted sampling and label smoothing," Mon. Not. R. Astron. Soc., vol. 519, no. 3, pp. 4765–4779, 2023, doi: 10.1093/mnras/stac3770.
- [45] B. Zoph and Q. V. Le, "Neural Architecture Search With Reinforcement Learning," Iclr 2017, vol. 32, pp. 1–16, 2017.
- [46] R. J. Willia, "Simple Statistical Gradient-Following Algorithms for Connectionist Reinforcement Learning," Mach. Learn., vol. 8, no. 3, pp. 229–256, 1992, doi: 10.1023/A:1022672621406.
- [47] C. M. Bhuma and R. Kongara, "Childhood medulloblastoma classification using efficientnets," 2020 IEEE Bombay Sect. Signal. Conf. IBSSC 2020, pp. 64–68, 2020, doi: 10.1109/IBSSC51096.2020.9332175.
- [48] J. Liu, M. Wang, L. Bao, and X. Li, "EfficientNet based recognition of maize diseases by leaf image classification," J. Phys. Conf. Ser., vol. 1693, no. 1, 2020, doi: 10.1088/1742-6596/1693/1/012148.
- [49] Y. Tadeipalli, M. Kollati, S. Kuraparthi, and P. Kora, "EfficientNet-B0 based monocular dense-depth map estimation," Trait. du Signal, vol. 38, no. 5, pp. 1485–1493, 2021, doi: 10.18280/ts.380524.
- [50] G. S. Tandel, A. Tiwari, and O. G. Kakde, "Performance enhancement of MRI-based brain tumor classification using suitable segmentation method and deep learning-based ensemble algorithm," Biomed. Signal Process. Control, vol. 78, no. March, p. 104018, 2022, doi: 10.1016/j.bspc.2022.104018.
- [51] B. Hu, J. Tang, J. Wu, and J. Qing, "An Attention EfficientNet-Based Strategy for Bearing Fault Diagnosis under Strong Noise," Sensors, vol. 22, no. 17, pp. 1–19, 2022, doi: 10.3390/s22176570.
- [52] P. Machart and L. Ralaivola, "Confusion Matrix Stability Bounds for Multiclass Classification," 2012, [Online]. Available: <http://arxiv.org/abs/1202.6221>.
- [53] I. Markoulidakis, I. Rallis, I. Georgoulas, G. Kopsiaftis, A. Doulamis, and N. Doulamis, "Multiclass Confusion Matrix Reduction Method and Its Application on Net Promoter Score Classification Problem," Technologies, vol. 9, no. 4, 2021, doi: 10.3390/technologies9040081.
- [54] S. F. Stefenon, K. C. Yow, A. Nied, and L. H. Meyer, "Classification of distribution power grid structures using inception v3 deep neural network," Electr. Eng., vol. 104, no. 6, pp. 4557–4569, 2022, doi: 10.1007/s00202-022-01641-1.
- [55] S. V. Kogilavani et al., "COVID-19 Detection Based on Lung Ct Scan Using Deep Learning Techniques," Comput. Math. Methods Med., vol. 2022, 2022, doi: 10.1155/2022/7672196.