Improving the performance for automated brain tumor classification on magnetic resonance imaging deep learning-based

Muhammad Fachrurrozi[1], Annisa Darmawahyuni[2], Samsuryadi Samsuryadi[1], Rossi Passarella[3], Jerrel Adriel Archibald Hutahaean[1]

[1]Department of Informatic Engineering, Faculty of Computer Science, Universitas Sriwijaya, Palembang, Indonesia
[2]Intelligent System Research Group, Faculty of Computer Science, Universitas Sriwijaya, Palembang, Indonesia
[3]Department of Computer Engineering, Faculty of Computer Science, Universitas Sriwijaya, Palembang, Indonesia

Article Info

Article history:

Received Jun 12, 2023
Revised Oct 23, 2023
Accepted Nov 15, 2023

Keywords:

Brain tumor
Convolutional neural networks
Magnetic resonance image
Supervised learning

ABSTRACT

Brain tumor is an uncontrolled growth of abnormal cell in the brain. Early diagnosis of brain tumor has a crucial step in this type of cancer, which is fatal. Magnetic resonance imaging (MRI) is one of the examination tools to examine brain anatomy in clinical practice. The high resolution and clear separation of the tissue enable medical experts to identify brain tumor. The earlier of brain tumor is detected, the wider of treatment options. However, manually analysed of brain anatomy on MRI images are time-consuming. Computer-aided diagnosis with automated way is helpful solution to help management with unreliable degrees of automation to trace various tissue boundaries. This study proposes convolutional neural network (CNN) with its excellences to automated features extraction in convolution layer. The popular architectures of CNN, i.e., visual geometry group16 (VGG16), residual network-50 (resNet-50), inceptionV3, mobileNet, and efficientNetB7 in medical image processing are compared to brain tumor classification task. As the results, VGG16 outperformed other architectures of CNN in this study. VGG16 yields 100% accuracy, precision, sensitivity, specificity, and F1-score for testing set data. The results show the excellent performance in classifying brain tumor and no tumor from MRI images that demonstrate the efficiency of system suggested.

Corresponding Author:

Annisa Darmawahyuni
Intelligent System Research Group, Faculty of Computer Science, Universitas Sriwijaya
Palembang 30139, Indonesia
Email: annisadarmawahyuni@unsri.ac.id

1. INTRODUCTION

Brain tumor is deadly cancer results from an uncontrolled division of abnormal cells. Cancer pertaining to the brain in leading causes of mortality [1]. There is no age limit for the occurrence of brain tumor. It is considered as the third most prevalent cancer among teenagers and adults [1]. The earlier a brain tumor is identified, the better the odds of survival and the more therapeutic choices there are. Hence, the concern of brain tumor object is still under research. Medical image processing has essential point in the field of medicine, especially in noinvasive treatment. It is one of the most important tools to identify as well as diagnose brain tumor. Commonly, the anatomy of brain tumor can be examined by magnetic resonance imaging (MRI), computed tomography (CT) scan, and X-ray. Unlike X-ray, MRI, and CT scan are not contain any radiation. However, MRI more provide accurate anatomical structure of tissues which it is providing clear images of most tumors [2]. MRI is a widely accepted due to it is the most efficient technique for high quality medical imaging in brain. Therefore, MRI is better compared to CT scan analysis. Unfortunately, the MRI analysis manually is time consuming, tedious, and innacurate sometimes. MRI is examined based on visual interpretation of the films to identify the presence of abnormal tissue. Its complex structure of brain tissue such as white and gray matter in the brain images is also being a concerned task [3].

Various computer-aided diagnosis for brain tumor classification using MRI images have many explored. They are consisted of as traditional or non-autonomous techniques and techniques pertaining to deep learning (DL). Machine learning, as traditional feature engineering has been investigated to solve the brain tumor classification task [2], [4]–[7]. However, machine learning approach is still required a human intervention to determine the features as input. In addition, the obtained performance by the previous results is still need to be improved. Therefore, this study focuses on to build a diagnosis model to brain tumor by incorporating automated feature engineering manner. DL has the power for the automated feature engineering without human intervention [8]. This study proposes a convolutional neural network (CNN), as one of DL algorithm to examine the brain MRI images in classification way. CNN is a frequently used for medical image classification task [9]–[12]. We have compared numerous architectures of CNN, i.e., visual geometry group16 (VGG16), residual network-50 (ResNet-50), inceptionV3, MobileNet, and efficientNetB7. Those architectures of CNN are popularly used for medical image processing task [13]–[17]. This study aims to improve the performance of brain tumor classification using MRI images. The significant contributions of this study is:

- Comparing and modifying the popular architecture of CNN (VGG16, resNet-50, inceptionV3, mobileNet, and efficientNetB7) with varying hyperparameter tuning.
- Obtaining a high performance (yields 100% accuracy, precision, sensitivity, specificity, and F1-score) for brain MRI images classification task.

2. MATERIAL AND METHOD

The brain tumor classification using MRI images in this study, has consisted of following research methodology; i) data acquisition of Kaggle: brain MRI images for brain tumor detection [18]; ii) the brain MRI images preprocessing (resize and grayscale); iii) the classification based on CNN architectures; and iv) the performance evaluation of brain tumor classification. The visualization of research methodology of brain tumor classification can be presented in Figure 1.

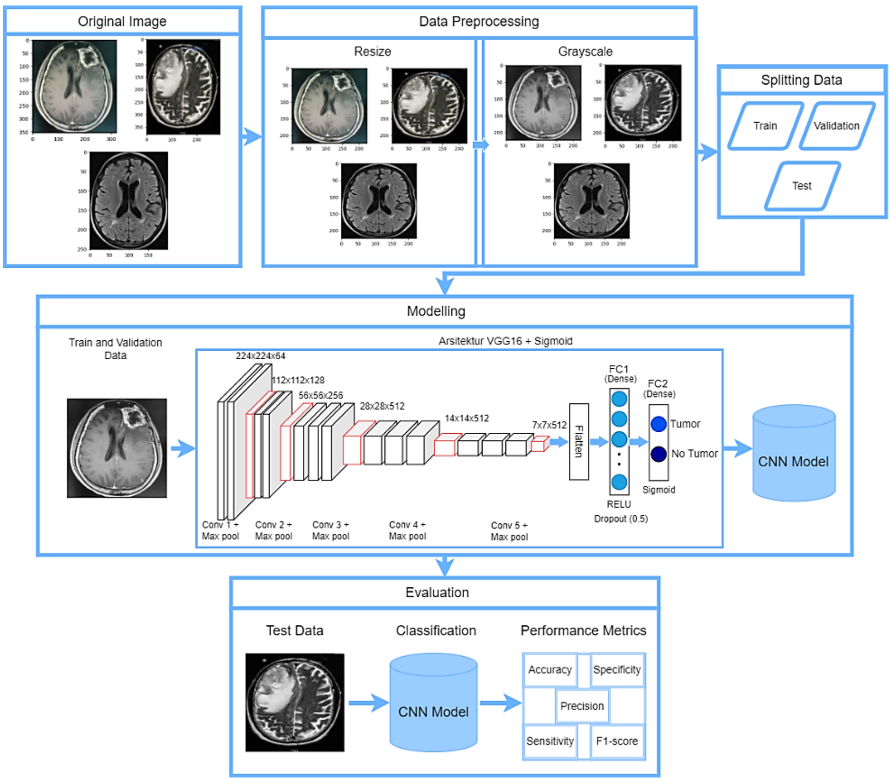
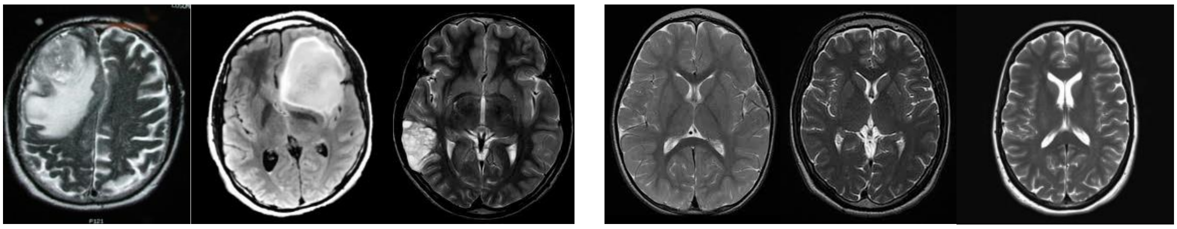


Figure 1. The research methodology of brain tumor classification

2.1. Data acquisition

The experiment of this study were taken from the Kaggle: brain MRI images for brain tumor detection [18]. Dataset consisted of 253 brain MRI images, which 155 brain MRI images has tumor, and the rest has no tumor. All images are .jpeg/.jpg. The size of brain MRI images is varying, and the sample of brain MRI images can be presented in Figure 2. Figure 2 presents the brain MRI images with tumor at Figure 2(a) and brain MRI images without tumor at Figure 2(b).

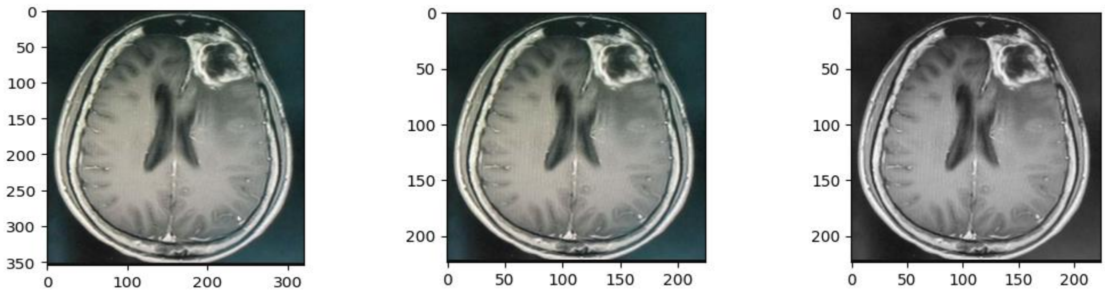


(a) (b)

Figure 2. The sample of brain MRI images: (a) brain MRI images with tumor and (b) brain MRI images without tumor

2.2. Brain magnetic resonance imaging image preprocessing

Medical image processing is extremely important in the area of medicine, particularly in noninvasive therapy and clinical research. The clear brain MRI images has become an essential method of high-quality medical imaging, particularly in brain imaging, where soft tissue contrast and non-invasiveness are obvious advantages. The step of preprocessing of brain MRI images can be seen in Figure 3. Figure 3(a) represents the raw data of brain MRI images. The raw data of brain MRI images have random size, with various color of images i.e., black and white and RGB partially. To obtain a fix size and same color of pixels, we have rescaled the image to the same size, and convert the random color to same color also. The first preprocessing is resizing the brain MRI images to same size of 224×224 pixels that can be seen in Figure 3(b). In next step, we convert all images to grayscale that can be seen in Figure 3(c).



(a) (b) (c)

Figure 3. The step of brain MRI image preprocessing; (a) brain MRI images raw data, (b) resizing the brain MRI images to same size, and (c) converting brain MRI images to grayscale

2.3. Convolutional neural networks

CNN is a one of outstanding DL algorithm, which predominantly trained in supervised manner by a stochastic gradient descent method. CNN consisted of convolutional and pooling layers, that combine the feature extraction and feature classification into a single learning body. CNN composed of four main components, i.e., kernels, a convolution layer, a non-linearity actication function, and pooling layer. The interconnections feeding the convolutional layers are assigned by weighting filters (w) and kernel size of (K_x, K_y). The convolution layer representing the image (I) is convolved with the smaller two-dimension kernel matrix (K), which it is given in (1):

$$S_{i,j} = (I * K)_{i,j} = \sum_m \sum_n I_{i,j} . K_{i-m,j-n} \quad (1)$$

where (*) is the two-dimension discrete convolution operator, to the matrix size mx n.
In this experiment, we have tuned five popular architectures of CNN, i.e., VGG-16, resNet-50, inceptionV3, mobileNet, and efficientNetB7 with varying parameter. The aforementioned CNN models have popular for medical image processing [17], [19]–[22]. In addition, we also tuned the learning rate and batch size as hyperparameter tuning for CNN model. The hyperparameter tuning of CNN model can be listed in Table 1.

Table 1. The hyperparameter tuning of CNN model

Model	Architecture	Learning rate	Batch size	Loss function	Epoch	Optimizer
1	VGG-16	10^{-2}	8	Binary-cross entropy	50	Adam
2	VGG-16	10^{-3}	8			
3	VGG-16	10^{-4}	8			
4	VGG-16	10^{-4}	16			
5	VGG-16	10^{-4}	32			
6	ResNet-50	10^{-4}	16			
7	InceptionV3	10^{-4}	16			
8	MobileNet	10^{-4}	16			
9	EfficientNetB7	10^{-4}	16			

2.4. Performance metrics

Classification metrics are implemented to measure and monitor model performance in training, validation, and testing sets. Classification metrics analyse good or bad the classification is, but each of them evaluates it in a different way. There are common classification metrics as in (1)-(5):

$$Accuracy = \frac{TP+TN}{TP+TN+FP+FN} \quad (1)$$

$$Precision = \frac{TP}{TP+FP} \quad (2)$$

$$Sensitivity(recall) = \frac{TP}{TP+FN} \quad (3)$$

$$Specificity = \frac{TN}{TN+FP} \quad (4)$$

$$F1 - score = \frac{2*(Recall*Precision)}{(Recall+Precision)} \quad (5)$$

where TP represents the number of positive samples correctly classified by classifier, TN denotes number of negative examples correctly classified by the classifier, FP denotes the number of the negative samples wrongly classified by the classifier, and FN denotes the number of the positive samples wrongly classified by the classifier.

3. RESULTS AND DISCUSSION

To generate the nine model of CNN, the brain MRI images were splitted to 80% training, 10% validation, and the rest for testing set data. The total is 202, 25, and 26 brain MRI images, for training, validation, and testing set, respectively. The performance metrics are accuracy, sensitivity, specificity, precision, and F1-score. The performance results of nine models of CNN using testing set can be presented in Table 2. As listed in Table 2, VGG16 was firstly experimented to hyperparameter tuning (models 1-5). The results of all metrics in models 1-3 that used the same value of batch size (8) and varying learning rates (10^{-2}, 10^{-3}, and 10^{-4}) have increased. The results of accuracy, precision, sensitivity, specificity, and F1-score show higher when decreased learning rate. From the results, the use of 10^{-4} learning rate presents the highest performance among 10^{-2} and 10^{-3}. Therefore, we have chosen the 10^{-4} learning rate as the optimal parameter. Besides learning rate, we have tuned the batch size from 8, 16, and 32 (models 3–5). Models 3 and 4 perform well, but model 4 better with 100% accuracy, precision, sensitivity, specificity, and F1-score. With the parameter of model 4, we have used batch size of 16 and learning rate of 10^{-4}.

Table 2. The performance results of nine models of CNN

Model	Performance (%)				
	Accuracy	Precision	Sensitivity	Specificity	F1-score
1	84.62	88.67	86.67	86.67	84.62
2	88.46	88.75	87.58	87.58	88.02
3	92.31	92.31	93.33	93.33	92.26
4	100	100	100	100	100
5	92.31	92.31	93.33	93.33	92.26
6	96.15	95.83	96.67	96.67	96.1
7	88.46	88.75	87.58	87.58	88.02
8	88.46	89.29	90	90	88.44
9	92.31	92.31	93.33	93.33	92.26

The others architecture of CNN (resNet-50, inceptionV3, mobileNet, and efficientNetB7) have also analysed with the best parameter of model 4. We have compared the VGG16 and others architecture of CNN based on their performance results. The results tend to stable, but not significant. VGG16 has outperformed the others architecture of CNN. VGG16 consisted of 13 convolution layers, five pooling layers, and three dense layers, but it has only 16 weight layers (learnable parameters layer). VGG could use very small receptive fields instead of massive fields like alexNet. VGG16 improved on alexNet and replaced the large filters with sequences of smaller 3×3 filters. We have concluded that the VGG16 is the best architecture of CNN (model 4) based on the comparison for brain MRI images classification task in this study.

In this study investigation, we presented the confusion matrix (CM) to determine the classification model performance for given set of testing data. The features of CM are given for the two prediction classes of classifiers (tumor and no tumor), which divided into two dimensions (predicted values and actual values). The CM makes predictions on testing set and performs how good the proposed model is, and also it calculates the different parameters for the models, such as of accuracy, precision, sensitivity, specificity, and F1-score. Figure 4 presented the CM for nine models of CNN (Figures 4(a)–(i)). Model 4 as the proposed model, obtained the excellent performance where all the classes have successfully classified (15 brain tumor and 11 no brain tumor).

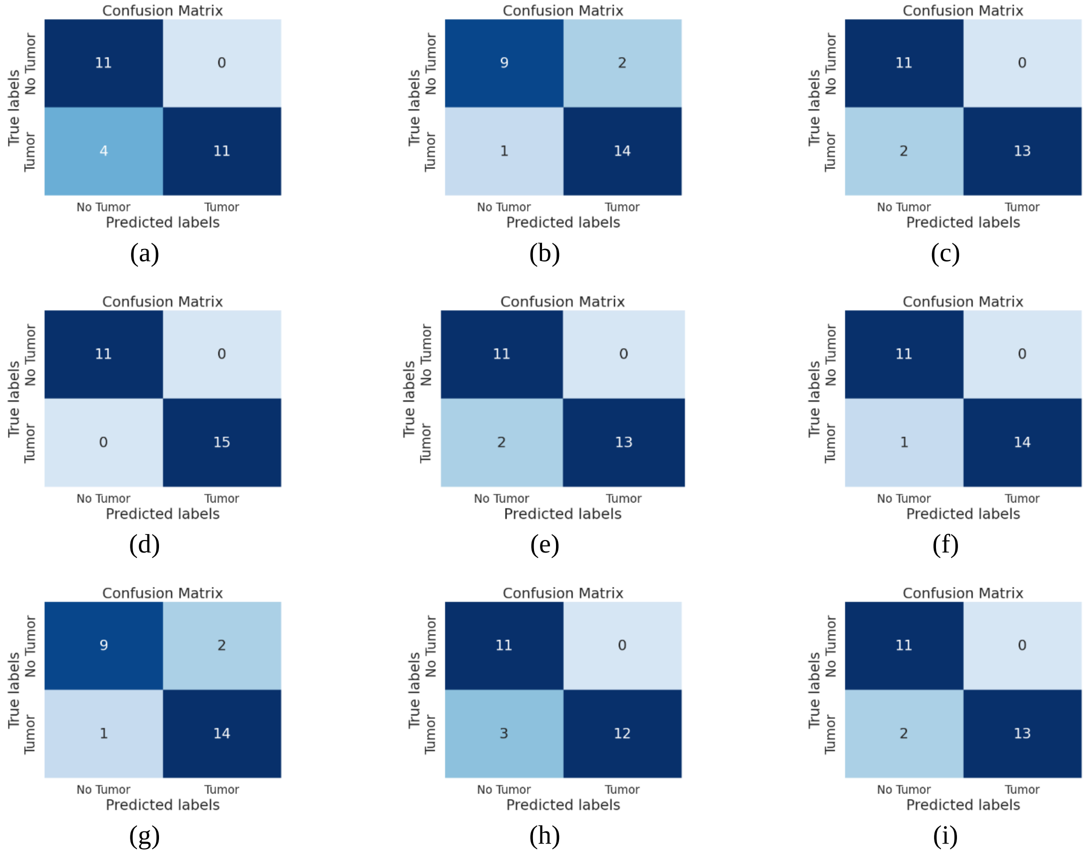


Figure 4. The CM for nine models of CNN based on testing set; (a) model 1, (b) model 2, (c) model 3, (d) model 4, (e) model 5, (f) model 6, (g) model 7, (h) model 8, and (i) model 9

We have compared our proposed model to other previous works that using DL for brain MRI images classification. Çinar and Yildirim [12] have improved the DL model for brain MRI images classification, which the resNet50 model is used as the base and preferred to use a trained model. From the resNet50 as base model, they removed five layers of resNet50 and added ten new layers, which the number of layers from 177 increased to 182. They achieved 97.01% accuracy.

Saxena *et al*. [23] has also implemented and proposed resNet 50 for brain MRI images classification. They compared VGG16 and inceptionV3, however, the resNet50 has outperformed both model with the 95% accuracy. Shahzadi *et al.* [24] has introduced a cascade of CNN (VGG16) with long short-term memory (LSTM). They have also compared to the features extracted from alexNet and resNet, however, VGG16 has the highest accuracy with 84%.

Siddique *et al.* [25] have also proposed VGG16 and achieved 96% accuracy. But, they have modified the general VGG16 architecture. They changed final max-pooling layer to average-pooling layer, also known as global average pooling (GAP) layer.

From the benchmark studies, we have concluded that our proposed model (VGG16) outperformed the previous works (refer to Table 3). With the excellences of VGG16, we obtained excellent performance results with the 100% accuracy, precision, sensitivity, specificity, and F1-score. Though the results look promising to clinical practice, the limitation of this study is the used dataset still not generalized. Add more brain MRI images datasets, can make a DL model robust due to varying features learned.

Table 3. The benchmark studies for Brain MRI images classification using DL

Authors	Method	Performance (%)				
		Accuracy	Precision	Sensitivity	Specificity	F1-score
Çinar and Yildirim [12]	Improved DL model	97.01	-	-	-	-
Saxena *et al*. [23]	ResNet50	95	-	-	-	-
Shahzadi *et al.* [24]	VGGNet - LSTM	84	-	-	-	-
Siddique *et al*. [25]	VGG16	96	93	100	-	97
Proposed	VGG16	100	100	100	100	100

4. CONCLUSION

This study presents the nine model of CNN with five varying architectures, i.e., VGG-16, resNet-50, inceptionV3, mobileNet, and efficientNetB7 for brain MRI images classification. VGG16 with the modified hyperparameters (batch size and learning rate) proved to be a significant way to classify the brain tumor. The performance has outstanding results with 100% accuracy, precision, sensitivity, specificity, and F1-score using testing set. The power of DL provides more efficient result than the conventional algorithm (handcrafted features). The finding of this study lead to inclusion of clinical experts in support of clinical decision systems.

ACKNOWLEDGEMENTS

The research of this article was funded by DIPA of Public Service Agency of Universitas Sriwijaya 2023. Number of SP DIPA-023.17.2.677515/2023, on November 30th 2022, in accordance with the Rector's Decree Number: 0188/UN.9.3.1/SK/2023, on April 18th, 2023.

REFERENCES

[1] J. Cheng *et al*., "Retrieval of Brain Tumors by Adaptive Spatial Pooling and Fisher Vector Representation," *PLoS ONE*, vol. 11, no. 6, pp. 1–15, 2016, doi: 10.1371/journal.pone.0157112.

[2] R. Kumari, "SVM Classification an Approach on Detecting Abnormality in Brain MRI Images," *International Journal of Engineering Research and Applications (IJERA)*, vol. 3, no. 4, pp. 1686–1690, 2013.

[3] J. Patel and K. Doshi, "A Study of Segmentation Methods for Detection of Tumor in Brain MRI," *Advance in Electronic and Electric Engineering*, vol. 4, no. 3, pp. 279–284, 2014.

[4] M. K. Abd-Ellah, A. I. Awad, A. A. M. Khalaf, and H. F. A. Hamed, "Design and implementation of a computer-aided diagnosis system for brain tumor classification," in *2016 28th International Conference on Microelectronics (ICM)*, 2016, pp. 73–76, doi: 10.1109/ICM.2016.7847911.

[5] M. Soltaninejad, X. Ye, G. Yang, N. Allinson, and T. Lambrou, "Brain tumour grading in different MRI protocols using SVM on statistical features," in *Medical Image Understanding and Analysis*, 2014, pp. 1–6.

[6] R. Ahmmed, A. S. Swakshar, M. F. Hossain, and M. A. Rafiq, "Classification of tumors and it stages in brain MRI using support vector machine and artificial neural network," in *2017 International Conference on Electrical, Computer and Communication Engineering (ECCE)*, 2017, pp. 229–234, doi: 10.1109/ECACE.2017.7912909.

[7] G. Hemanth, M. Janardhan, and L. Sujihelen, "Design and Implementing Brain Tumor Detection Using Machine Learning Approach," in *2019 3rd International Conference on Trends in Electronics and Informatics (ICOEI)*, 2019, pp. 1289–1294, doi:

10.1109/ICOEI.2019.8862553.
[8] J. Patterson and A. Gibson, *Deep Learning: A Practitioner's Approach*. USA: O'Reilly Media, 2017.
[9] F. Milletari, N. Navab, and S. -A. Ahmadi, "V-Net: fully convolutional neural networks for volumetric medical image segmentation," in *2016 Fourth International Conference on 3D Vision (3DV)*, 2016, pp. 565–571, doi: 10.1109/3DV.2016.79.
[10] S. Nurmaini *et al.*, "Deep Learning for Improving the Effectiveness of Routine Prenatal Screening for Major Congenital Heart Diseases," *Journal of Clinical Medicine*, vol. 11, no. 21, pp. 1–15, 2022, doi: 10.3390/jcm11216454.
[11] Z. Jia and D. Chen, "Brain Tumor Identification and Classification of MRI images using deep learning techniques," *IEEE Access*, pp. 1–10, 2020, doi: 10.1109/ACCESS.2020.3016319.
[12] A. Çinar and M. Yildirim, "Detection of tumors on brain MRI images using the hybrid convolutional neural network architecture," *Medical hypotheses*, vol. 139, p. 109684, 2020.
[13] D. G. M., N. P. Pai, A. R., S. G., and K. R. P. M., "Brain Tumor Detection and Segmentation Using VGG16 and Mask R-CNN with Transfer Learning," *Solid State Technology*, vol. 63, no. 5, pp. 9887–9893, 2020.
[14] L. Cai, T. Long, Y. Dai, and Y. Huang, "Mask R-CNN-Based Detection and Segmentation for Pulmonary Nodule 3D Visualization Diagnosis," *IEEE Access*, vol. 8, pp. 44400–44409, 2020, doi: 10.1109/ACCESS.2020.2976432.
[15] Y. Ding *et al.*, "A deep learning model to predict a diagnosis of Alzheimer disease by using 18 F-FDG PET of the brain," *Radiology*, vol. 290, no. 3, pp. 456–464, 2019, doi: 10.1148/radiol.2018180958.
[16] M. A. R. Ratul, M. H. Mozaffari, W.-S. Lee, and E. Parimbelli, "Skin Lesions Classification Using Deep Learning Based on Dilated Convolution," *bioRxiv*, pp. 1–11, 2020, doi: 10.1101/860700.
[17] A. Abirami, S. Bhuvaneswari, B. S. Gowri, S. Narayanan, N. Sharmitha, and M. Sowbarnika, "MRI-based Brain Tumour Classification Using EfficientNetB7 model with transfer learning," *Journal of Survey in Fisheries Sciences*, vol. 10, no. 2S, pp. 1737–1750, 2023.
[18] N. Chakrabarty, "Brain MRI Images for Brain Tumor Detection," *Kaggle*. 2019. [Online]. Available: https://www.kaggle.com/datasets/navoneel/brain-mri-images-for-brain-tumor-detection. Accessed: Jan. 02, 2023.
[19] A. A. Pravitasari *et al.*, "UNet-VGG16 with transfer learning for MRI-based brain tumor segmentation," *Telkomnika (Telecommunication Computing Electronics and Control)*, vol. 18, no. 3, pp. 1310–1318, 2020, doi: 10.12928/TELKOMNIKA.v18i3.14753.
[20] L. V. Fulton, D. Dolezel, J. Harrop, Y. Yan, and C. P. Fulton, "Classification of alzheimer's disease with and without imagery using gradient boosted machines and resnet-50," *Brain Sciences*, vol. 9, no. 9, pp. 1–16, 2019, doi: 10.3390/brainsci9090212.
[21] M. Shoaib and N. Sayed, "YOLO Object Detector and Inception-V3 Convolutional Neural Network for Improved Brain Tumor Segmentation," *Traitement du Signal*, vol. 39, no. 1, pp. 371–380, 2022, doi: 10.18280/ts.390139.
[22] T. H. Arfan, M. Hayaty, and A. Hadinegoro, "Classification of Brain Tumours Types Based On MRI Images Using Mobilenet," in *2021 2nd International Conference on Innovative and Creative Information Technology (ICITech)*, 2021, pp. 69–73, doi: 10.1109/ICITech50181.2021.9590183.
[23] P. Saxena, A. Maheshwari, S. Tayal, and S. Maheshwari, "Predictive modeling of brain tumor: A Deep learning approach," *arXiv preprint arXiv:1911.02265*, 2019.
[24] I. Shahzadi, T. B. Tang, F. Meriadeau, and A. Quyyum, "CNN-LSTM: Cascaded Framework For Brain Tumour Classification," in *2018 IEEE-EMBS Conference on Biomedical Engineering and Sciences (IECBES)*, 2018, pp. 633–637, doi: 10.1109/IECBES.2018.8626704.
[25] M. A. B. Siddique, S. Sakib, M. M. R. Khan, A. K. Tanzeem, M. Chowdhury, and N. Yasmin, "Deep Convolutional Neural Networks Model-based Brain Tumor Detection in Brain MRI Images," in *2020 Fourth International Conference on I-SMAC (IoT in Social, Mobile, Analytics and Cloud) (I-SMAC)*, Oct. 2020, pp. 909–914, doi: 10.1109/I-SMAC49090.2020.9243461.

BIOGRAPHIES OF AUTHORS

Muhammad Fachrurrozi is currently is a lecturer and researcher Faculty of Computer Science, Universitas Sriwijaya, Indonesia. His research interest includes biomedical image processing, deep learning, and machine learning. He can be contacted at email: mfachrz@unsri.ac.id.

Annisa Darmawahyuni is currently is a lecturer and researcher in Intelligent System Research Group, Faculty of Computer Science, Universitas Sriwijaya, Indonesia. Her research interest includes biomedical signal/image processing, deep learning, and machine learning. She can be contacted at email: riset.annisadarmawahyuni@gmail.com or annisadarmawahyuni@unsri.ac.id.

Samsuryadi Samsuryadi is currently is a lecturer and researcher in Faculty of Computer Science, Universitas Sriwijaya, Indonesia. His research interest includes image processing, deep learning, and pattern recognition. He can be contacted at email: samsuryadi@unsri.ac.id.

Rossi Passarella is currently is a lecturer and researcher in Faculty of Computer Science, Universitas Sriwijaya, Indonesia. His research interest includes deep learning and big data. He can be contacted at email: passarella.rossi@unsri.ac.id.

Jerrel Adriel Archibald Hutahaean is currently postgraduate student at Faculty of Computer Science, Universitas Sriwijaya, Indonesia. His research interest includes image processing and deep learning. He can be contacted at email: jerreladriel@gmail.com.