Attention-based MobileNet for Effective Recognition of Brain Tumors

Tianyi Du*

Electronic Information Engineering Beijing Information Science and Technology University Beijing, China

*Corresponding author email: 2020010976@bistu.edu.cn

Abstract. Brain tumor is a very dangerous disease, which often causes human death. When the consulting doctor is inexperienced, it is easy to misdiagnose and cause serious consequences. Therefore, accurate and timely identification of brain tumors from Magnetic Resonance Imaging (MRI) images is crucial due to the life-threatening nature of this disease and the potential consequences of misdiagnosis. To address this challenge, this research employs a neural network model for brain tumor identification, incorporating model fusion techniques to enhance accuracy. Specifically, the study compares the performance of two models, namely MobileNet and SE+MobileNet Block, in recognizing MRI images. The SE+MobileNet Block model integrates a section of SENet into MobileNet, introducing a 3 x 3 convolutional layer, two Fully Connected (FC) layers, and the last two activation functions (Rule and Sigmoid), effectively incorporating an attention mechanism. Comparative analysis using accuracy and loss values demonstrates that the inclusion of the attention mechanism in the MobileNet model (i.e., the SE+MobileNet Block model) improves the recognition accuracy of brain tumor MRI.

Keywords: component; brain tumor; attention; SENet; MobileNet.

1. Introduction

A brain tumor denotes an anomalous accumulation of tissue characterized by uncontrolled cellular proliferation, exhibiting an apparent evasion of regulatory mechanisms governing physiological cellular behavior. Brain tumors are a deadly cancer that kills thousands of people every year. Between 2014 and 2018, 83,029 people died from malignant brain and other Central Nervous System (CNS) tumors, according to the U.S. Central Brain Tumor Registry [1]. This translates to an average annual death rate of 4.43 per 100,000 people, an average of 16,606 deaths per 100,000 people per year. In general, the five-year relative survival rate for malignant brain tumors and other CNS tumors was 35.6%, compared with 91.8% for non-malignant brain tumors and other CNS tumors. Traditionally, the diagnostic approach for assessing tumor malignancy relies on manual inspection of Magnetic Resonance Imaging (MRI) images, leveraging expertise, subsequent scrutiny, and laboratory analysis to ascertain malignancy. However, this process often takes a lot of time. For some brain tumor patients, waiting too long will undoubtedly aggravate the condition and affect the follow-up treatment. Moreover, the delay in receiving follow-up examination results significantly increases the likelihood of missing critical diagnostic and therapeutic opportunities. Hence, the integration of Artificial Intelligence (AI) with medical imaging data becomes imperative.

Computer-aided diagnosis (CAD) technology is considered a useful way to assist neuro-oncologists in several ways. The use of CAD in neuro-oncology includes applications such as tumor detection, classification, and grading. The classification of brain tumors into benign and malignant tumors based on CAD is currently the subject of extensive research. The CAD systems described above rely on MRI images of the brain. This is because MRI provides higher contrast to the soft tissues in the brain than Computed Tomography (CT) images [2, 3]. Convolutional Neural Network (CNN) is an artificial intelligence algorithm that learns from images and performs tasks such as classification, segmentation, and detection Attention module. Another model SENet used in this article is a lightweight, simple-minded new network, which is also derived from the attention mechanism and has strong performance and portability. The target of this study is to propose a brain tumor detection system in the hospital, so both portability and transplant ability are necessary. Hu et al. used CNN for tumor prediction with the highest accuracy of 70% [4], and the TD-CNN-LSTM model proposed by Montaha [5]. In general, the effect of multi-model fusion is usually better than that of a single model, because a single model may pay attention to some useless features during training, resulting in an increase in the amount of computation.

In this study, the convolutional neural network and attention mechanism will be considered for combining together to improve the prediction performance. Firstly, CNN will be used to select the target area, and then perform subsequent analysis. In the process of analysis, the SE Block algorithm will be combined to test whether the overall detection rate can be improved. The results demonstrated that the combined model shows the better result of detection. Therefore, the method adopted in this study is reasonable.

2. Method

2.1. Dataset Description

In this study, the data set used in the study sourced from the Kaggle website [6]. The data set of this study has a total of 3, 000 MRI images, and the main structure of the dataset is divided into two folders, one with brain tumor and one without brain tumor, and the number between them is 1, 500 MRI images each. In this study, since the size of the photos provided by the data set is not consistent, in the follow-up study, this study adopted the unified image format to 156×156 for reducing the computational memory. The Figure 1 is an example image of the dataset for this study.

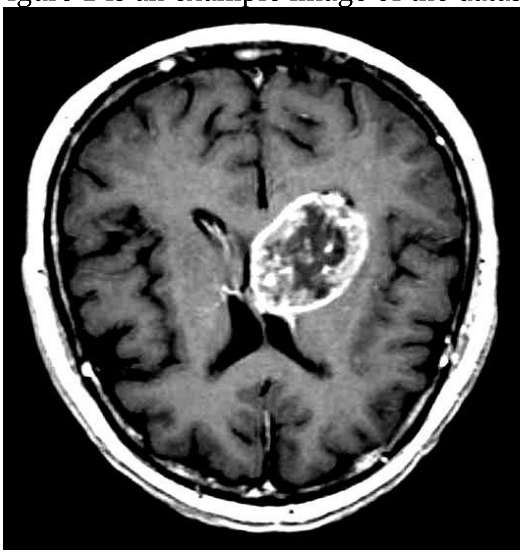

Figure 1. The sample image of data set [6].

2.2. Dataset Preprocessing

In this study, the labeled MRI dataset of brain tumors was utilized. For the processing of the data set, this paper first normalizes the data, which can improve the efficiency of the model operation, and then classifies the training set and the test set at 8:2. A method of image enhancement is adopted, and the method of cropping is used to unify the format of all the pictures, to improve the accuracy of the experimental results.

2.3. Attention mechanism

In this study, a new model related to the attention mechanism - the SE Block model was considered. SE Block is also a kind of attention mechanism [7-9]. Attention Mechanism is a method that imitates the human visual and cognitive system, which allows neural networks to focus on relevant parts when processing input data. By introducing the attention mechanism, the neural network can automatically learn and selectively focus on important information in the input, improving the performance and generalization ability of the model. Attention mechanisms are often applied to the processing of sequential data such as text, speech, or image sequences. Among them, the most typical attention mechanisms include self-attention mechanism, spatial attention mechanism and temporal attention mechanism. The basic idea of the self-attention mechanism is that when processing sequence data, each element can establish associations with other elements in the sequence, rather than just relying on adjacent elements. It adaptively captures long-range dependencies between elements by computing their relative importance.

2.4. CNN model

The CNN model is a deep learning model widely used in image processing and recognition. A CNN model consists of a stack of convolutional, pooling and FC layers. Convolutional layers apply filters to the input image and extract features such as edges, shapes, and patterns. The pooling layer reduces the spatial dimension and noise of the feature map. A FC layer connects all the neurons of the previous layer to the output layer and performs classification or regression tasks. By applying proper preprocessing and architecture, CNN models can also be used on other types of data, such as time series, text, or audio. In this study, CNNs were used for medical impact, i.e. for identifying MRI images of brain tumors.

MobileNet is an efficient convolutional neural network designed for mobile and embedded vision applications [10]. It uses a depth-separable convolution technique that decomposes the standard convolution layer into two parts, namely the depth convolution layer and the point convolution layer, reducing the number of parameters and computation. There are multiple versions of MobileNet, the latest is MobileNet v3, which introduces some new modules based on depth separable convolution, such as linear bottleneck, inverted residual, attention mechanism, etc., to further improve the performance of the model. In simple terms, it describes an efficient network architecture and allows the direct construction of small, low-latency models that meet the needs of embedded devices with two hyperparameters, since only a small number of hyperparameters are used.

This research training is divided into two CNN models, one is a model with only MobileNet and the other is the MobileNet with the SE block In this research, the data first passes through a convolutional layer to activate the activation function, and then the multi-dimensional data is one-dimensionalized through the Flatten layer, and enters the subsequent Dense Layer, a total of 20 rounds were carried out to produce the experimental results.

In this research, the SE Block model combined with MobileNet is output from the activation layer of the MobileNet model shown in Figure 2, then enters the SE Block model, first passes through a global pooling layer, then enters the Full Connection (FC) layer 1, and then enters the activation layer of SE Block, and then enter the FC2 layer. After two fully connected layers, the data is output to the MobileNet model for final calculation. The image below is a flowchart of this model.

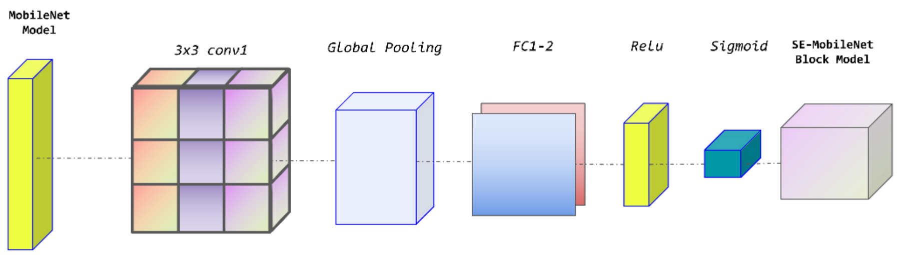

Figure 2. The proposed MobileNet combined with the SE block (Photo/Picture credit: Original).

2.5. Implementation details

In the two models, the optimizer used in this study is Adam, and the loss function is cross entropy. Furthermore, the evaluation index adopts is the accuracy. The epoch of this study is 20, and the batch_size is 8.

3. Results and discussion

3.1. The performance of various models

Table 1. The performance of the models

Model	Accuracy	Loss
MobileNet	96.00%	12.59
MobileNet+SENet Block	97.33%	8.57

Table 1 presents the accuracy and the corresponding loss of the MobileNet and MobileNet combined with SENet Block. Among them, the Accuracy of MobileNet is 1.33% lower than the model with the SENet module, and the Loss value of Mobile is 4.02 higher than the model with the attention mechanism module. Through comparison, it can be clearly found that the proposed model with the attention mechanism is obviously better than the traditional MobileNet model. It demonstrated that the method of MobileNet+SENet Block used in this study is more effective. Additionally, the training and vailaition curve were shown in Figure 3, Figure 4, Figure 5 and Figure 6.

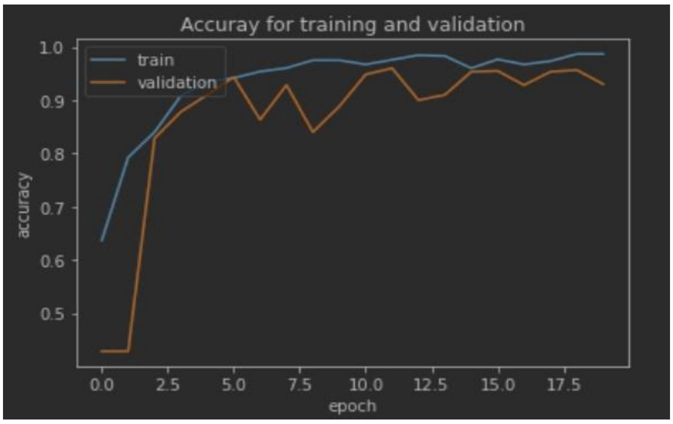

Figure 3: The accuracy curve of MobileNet (Photo/Picture credit: Original).

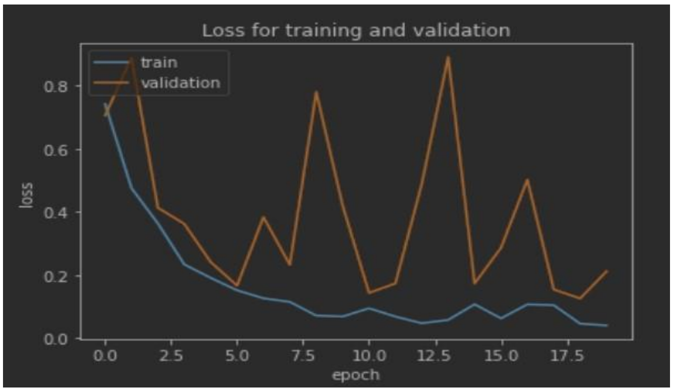

Figure 4: The loss curve of MobileNet (Photo/Picture credit: Original).

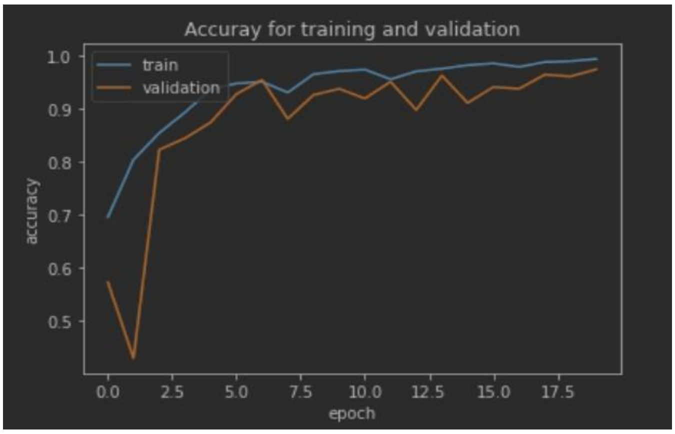

Figure 5: The accuracy curve of MobileNet + SE Block (Photo/Picture credit: Original).

Figure 6. The loss curve of MobileNet + SE Block (Photo/Picture credit: Original).

3.2. Discussion

Through the above experimental results, it can be inferred that the addition of SENet Block can make the final brain tumor recognition accuracy of this experiment higher and the loss value smaller. The attention mechanism can pay more attention to important features, thereby improving the overall performance of the model. In particular, The attention mechanism can more reasonably focus on different brain regions. SEBlock may compresses the feature map of brain tumors, and this Block reduces the input dimension of the fully connected layer, thereby reducing the number of parameters and calculations that need to be learned. This facilitates in reducing the complexity of the model and speeding up the training and inference process of the model.

By learning the weight of each channel, the SENet Block can adaptively select and emphasize important features, such as the key extraction of the features of the brain tumor part, such as the area where the background is black, this model will reduce its presence weight, and weaken the area. This Block will improve the discriminative ability of features. The preceding operations make the network better focus on the most important channel features of information, while suppressing those unimportant or redundant channel features. In addition, the attention mechanism of SENet Block is universal and can be embedded into existing network structures, such as the MobileNet model used in this study, which will not change the output shape of the original network structure, nor will it introduce new spaces Dimensions for feature fusion. Therefore, the accuracy of the original model will be improved.

However, the current model still has some spaces that can be optimized. At present, the weights analyzed by the attention mechanism module have not been visualized. For problems such as brain tumors, the weights of network analysis can be visualized. After the MRI image is obtained, improvements can be added, and the focus of human eye recognition can be narrowed to distinguish the disease more accurately. The overfitting problem demonstrated in this study also requires attention for solving in the future.

4. Conclusion

This study identifies and judges brain tumors through the proposed ensemble CNN network model., The results of this study show that the accuracy of the MobileNet+SE Block model is better than that of the single MobileNet model when faced with the problem of brain tumor recognition. On the data set of 3, 000 pictures, the model with SENet Block has achieved satisfactory results, and the accuracy has increased by 1.33%, which demonstrated that it has good performance. Therefore, when the model faces the problem of brain tumor identification, it has great medical significance, However, future research and extensive testing on larger brain tumor data sets are still needed.

References

[1] CBTRUS, “Central Brain Tumor Registry” https://www.cbtrus.org/www.cbtrus.org/index.html, 2019.

[2] Y. Qiu, C. S. Chang, J. L. Yan, L. Ko and T. S. Chang, "Semantic Segmentation of Intracranial Hemorrhages in Head CT Scans," 2019 IEEE 10th International Conference on Software Engineering and Service Science (ICSESS), Beijing, China, 2019, pp. 112-115, doi: 10.1109/ICSESS47205.2019.9040733.

[3] D. V. Gore and V. Deshpande, "Comparative Study of various techniques using Deep Learning for Brain Tumor Detection," 2020 International Conference for Emerging Technology (INCET), Belgaum, India, 2020, pp. 1-4, doi: 10.1109/INCET49848.2020.9154030.

[4] H. Hu, X. Li, W. Yao and Z. Yao, “Brain Tumor Diagnose Applying CNN through MRI,” 2021 2nd International Conference on Artificial Intelligence and Computer Engineering (ICAICE), Hangzhou, China, 2021, pp. 430-434, doi: 10.1109/ICAICE54393.2021.00090.

[5] S. Montaha, S. Azam, A. K. M. R. H. Rafid, M. Z. Hasan, A. Karim and A. Islam, “TimeDistributed-CNN-LSTM: A Hybrid Approach Combining CNN and LSTM to Classify Brain Tumor on 3D MRI Scans Performing Ablation Study,” in IEEE Access, vol. 10, pp. 60039-60059, 2022.

[6] Kaggle, “Br35H::Brain Tumor Detection 2020”, 2021, https://www.kaggle.com/datasets/ahmedhamada0/brain-tumor-detection?select=Br35H-Mask-RCNN.

[7] Y. Qiu, et al. "Pose-guided matching based on deep learning for assessing quality of action on rehabilitation training," Biomedical Signal Processing and Control, 2022, 72: 103323.

[8] Y. Li, J. Zeng, S. Shan and X. Chen, "Occlusion Aware Facial Expression Recognition Using CNN With Attention Mechanism," in IEEE Transactions on Image Processing, vol. 28, no. 5, pp. 2439-2450, May 2019, doi: 10.1109/TIP.2018.2886767.

[9] M. Abid, P. Compagnon and G. Lefebvre, "Improved CNN-based Magnetic Indoor Positioning System using Attention Mechanism," 2021 International Conference on Indoor Positioning and Indoor Navigation (IPIN), Lloret de Mar, Spain, 2021, pp. 1-8, doi: 10.1109/IPIN51156.2021.9662602.

[10] A. Howard, et al. "Mobilenets: Efficient convolutional neural networks for mobile vision applications," arXiv preprint arXiv:1704.04861, 2017.