IIETA International Information and Engineering Technology Association
Advancing the World of Information and Engineering

Traitement du Signal
Vol. 41, No. 5, October, 2024, pp. 2573-2584
Journal homepage: http://iieta.org/journals/ts

Efficient Deep Learning Fusion-Based Approach for Brain Tumor Diagnosis

Ajay Krishan Gairola[1], Vidit Kumar[1], Gagan Deep Singh[1], Mohit Bajaj[2,3,4*], Rajkumar Singh Rathore[5], Mohamed H. Mahmoud[6], Walid El-Shafai[7]

[1] Department of Computer Science and Engineering, Graphic Era (Deemed to be University), Dehradun 248002, India
[2] Department of Electrical Engineering, Graphic Era (Deemed to be University), Dehradun 248002, India
[3] Hourani Center for Applied Scientific Research, Al-Ahliyya Amman University, Amman 19111, Jordan
[4] Department of Electrical Engineering, Graphic Era Hill University, Dehradun 248002, India
[5] Cardiff School of Technologies, Cardiff Metropolitan University, Cardiff CF5 2YB, United Kingdom
[6] Department of Biochemistry, College of Science, King Saud University, Riyadh 11451, Saudi Arabia
[7] Department of Electronics and Electrical Communications Engineering, Faculty of Electronic Engineering, Menoufia University, Menouf 32952, Egypt

Corresponding Author Email: thebestbajaj@gmail.com

https://doi.org/10.18280/ts.410530

Received: 16 January 2024
Revised: 10 April 2024
Accepted: 20 August 2024
Available online: 31 October 2024

Keywords:
brain images, smart healthcare, deep learning, brain tumors, transfer learning, deep neural

ABSTRACT

Technology has advanced to the point where it can influence every facet of human existence. Here, we look at how technology can help treat brain tumors, one of the most frequent malignancies and a leading cause of death. Many people lose their lives each year because of brain tumors. In the United States, roughly 85,000 new cases are diagnosed each year, bringing the total number of people with primary brain tumors to an estimated 700,000. Artificial intelligence has helped medicine and people overcome this challenge. Most brain cancers are detected via magnetic resonance imaging. Medical imaging and image processing make extensive use of magnetic resonance imaging for diagnosing anatomical differences. In this paper, we investigate the performance of various convolutional neural network (CNN) models like AlexNet, GoogleNet, VGGnet11, VGGnet13, VGGnet16, VGGnet19, ResNet18, ResNet34, ResNet50, ResNet101, ResNet152, DenseNet121, DenseNet161, DenseNet169, and DenseNet201 for brain tumor diagnosis tasks. On a dataset of 3264 MRI images, we perform experiments for healthy meningioma, glioma, and pituitary brain tumor classification. Our tests reveal that the ResNet and DenseNet models yield the highest accuracy (82%). Furthermore, we investigate the potential of a fusion-based approach where we test for different combinations of fusion of CNN models. The results show that fusing many CNN features improves accuracy even more. Classification accuracy is improved to 86% when ResNet50 and ResNet101 are fused and improves to 84% when DenseNet161 and DenseNet169 are fused.

1. INTRODUCTION

A brain tumor is a potentially fatal condition characterized by aberrant tissue structure in the brain as a result of several cerebrovascular disorders. Since 1970, several researchers have worked to improve brain tumor detection by adding more interaction with medication to ensure proper and speedy treatment, but the goal of medical imaging is to offer adequate observation to neurologists in order to pinpoint the tumor's precise position [1, 2]. Today's computational power and availability of high-quality medical imagery like MRI Dicom images have greatly improved the accuracy of computer-assisted brain tumor identification [3].

The brain, the human body's most vulnerable and heaviest organ, controls most bodily functions, including movement, taste, hearing, vision, sensation, and smell. Brain tumors, resulting from abnormal cell growth, can strain adjacent tissues and disrupt brain function, leading to symptoms such as seizures, severe headaches, and changes in smell, sight, and hearing. These tumors can vary in size and location within the brain, complicating detection and treatment. Symptoms depend on factors like tumor size, location, and origin, and may also include personality changes, insomnia, memory loss, fatigue, nausea, and vomiting. The safest way to diagnose brain tumors is through a medical professional.

The spinal cord and brain make up the human central nervous system (CNS), which is the control center of the nervous system [4, 5]. The brain is responsible for a variety of bodily functions, including data processing, integration, coordination, decision making, and the transmission of commands to different areas of the body. Physiologically speaking, the human brain is extraordinarily complex. Brain tumors, TBI, developmental defects, MS, stroke, dementia, infections, and headaches are some examples of CNS diseases. Magnetic resonance imaging (MRI) and Computed tomography (CT) can diagnose brain problems. Brain MRI diagnosis most CNS diseases non-invasively [6]. Brain MRI allows the general practitioner to view brain slices. It detects

CNS diseases and provides information for patient diagnosis. MRI provides better image contrast than CT and allows living brain structure identification [7].

Medical experts are learning to read medical imaging probes. With its many imaging sequences for different purposes, MRI takes longer to learn. It takes time and effort for medical experts to accurately diagnose CNS problems [8]. The most common MRI sequences, T1 and T2, reveal tissue characteristics and information. Weighted MRI-MRI-T2 is associated with fluid-attenuated inversion retrieval, while T1 is associated with contrast enhancement, tensor eigenvalues, and image texture, and local histograms are used with MRI in brain tumor segmentation research. Human interaction has been associated with manual, semi-automated, and fully automated MRI [9]. Using deep learning in the healthcare industry will allow for more accurate diagnoses, better treatments, and better overall decision making.

This study makes three key contributions:

(i) A comparison of several state-of-the-art CNN models for detecting tumors in brain MRI scans.
(ii) An investigation into the impact of a late fusion-based approach on classification metrics.
(iii) An experiment using a dataset of 3264 images to evaluate the performance improvement achieved by late fusion of ResNet and DenseNet.

This paper follows the structure described below. Related work is discussed in Section 2. The methodology follows in Section 3. In Section 4, we provide a brief overview of the experiments, like the dataset, hardware-software configurations, and results, and in Section 5, we draw a conclusion and talk about possible future works.

2. RELATED WORK

In this section, we review the current medical state of tumor detection research. Existing surveys on artistic methods are reviewed, and the literature review of deep learning-based approaches is discussed.

2.1 Machine learning based methods

Image segmentation improves the detection of brain abnormalities. Researchers found strategies to accurately localize brain malformations from MRI images. SIENA and SIENAX are fully automated longitudinal and cross-sectional (single time point) analysis methods by Smith et al. [10]. BET took out the skull and brain. Registration techniques were developed to isolate the region of interest by separating the brain from the skull. The "Binary Oread" operation reduced brain transformation artifacts. For change analysis, brain edge motion was tracked in time-sequential images. Both are precise and robust, with 0.15% longitudinal brain volume change inaccuracy and 0.5-1% single-time point accuracy (cross-sectional). They are particularly insensitive to the scanning parameters. Vovk et al. [11] studied the MRI intensity inhomogeneous approach. Papers were ranked by their correction techniques, optimization, modeling, forward processing, and prior knowledge contribution. Makropoulos et al. [12] explored several segmentation strategies for automated prenatal and neonatal brain MRI imaging. Deformation models, unsupervised, classification, parametric, and atlas fusions have been created to search for brain, tissue, or more specific regions of interest. The neonatal tissue segmentation algorithms were then evaluated using NeoBrainS12, a benchmark dataset. In 2021, Fawzi et al. [13] provided a more thorough evaluation of recent brain imaging segmentation algorithms. This article discusses the 2019-2020 technologies. This includes most brain tumor segmentation deep learning models. This review paper found that hybrid and deep learning metaheuristic approaches perform better to segment brain tumors [14, 15].

2.2 Deep learning based methods

Artificial intelligence (AI) imaging methods provide a wide range of recommendations to facilitate the transfer of image data between systems [16]. The ability of radiologists to differentiate between malignant and benign breast lesions from MRI scans has been enhanced by artificial intelligence [17]. AI helps radiologists make more accurate diagnoses, which can lead to more effective treatment and ultimately a faster patient recovery. In paper [18], we build a high-efficiency algorithm for the detection and recognition of meniscus tears in an MRI of the knee. In paper [19], the authors perform a systematic literature analysis on the topic of radiomics and artificial intelligence (AI), covering all medical imaging methods with a focus on their non-oncological and oncological uses for general medical applications. In paper [20], we get an overview of MRI image analysis and processing using deep learning. In paper [21], a deep learning algorithm that uses an end-to-end training technique is designed to accurately diagnose breast cancer on screening mammograms.

The difficulties of convolutional neural networks (CNNs) have been the subject of many published studies and tutorials [22, 23]. In their book, Zhu et al. [24] and colleagues explain the layers that make up a CNN, as well as their applications and design. The forward and backward propagation of CNNs was also discussed. Pereira et al. [25], using CNN, employed normalization as a pre-processing step. He contributed to BraTS in 2013 and 2015, and his contributions included a proposal for a CNN-based automatic segmentation approach. MBINet model classifies RMB tumor pictures [26]. Self-organized operational neural networks power MBINet. The SMBI system collected 920 raw microwave brain (RMB) picture samples using a tiny 3D stacked wideband nine antenna array sensor. In their survey [27], Bernal et al. [27] provided an overview of CNN techniques in MRI image analysis, discussing architecture, data preparation, pre-processing, and post-processing methodologies. He details the development of several CNN designs and explores state-of-the-art methods in the field.

A technique based on a combination of CNN and genetic algorithms was implemented by Anaraki et al. [28]. To reduce variation in prediction errors, they proposed the use of bagging as an ensemble technique to non-invasively classify glioma grades from MRI scans. According to Tandel et al. [29], the goal of the initiative is to develop a non-invasive, quicker-than-a-biopsy automated MRI-based method for classifying brain tumors. This instrument can be used as a primary or secondary method for grading brain tumors. The authors used a fully connected conditional random field in three dimensions to reduce false positives [30]. Furthermore, a 3D CNN architecture was also described for automatic lesion segmentation.

The dual-force training method is introduced in study by Chen et al. [31] to motivate deeper models to acquire excellent multiple scales advantages. In paper [32], we zero in on a completely convolutional network that learns quickly, takes in data of any size, and returns data of the same size. For segmenting and evaluating images slice by slice, Havaei et al. [33] suggested FCNN. A two-phase training program was also suggested as a solution to the problem of social stratification. Menze et al. [34] suggested an approach for multi-modality image segmentation of brain tumors. This tactic can be viewed as either creative or selective. For 3D-based deep learning components, Frischer et al. [35] presented a three-pass CNN. Asad et al. [36] created a generic model for identifying tumors in brain cells. The study's primary objective was to improve medical professionals' ability to identify one of the worst diseases afflicting human beings. In his research, tumor types were determined using the CNN deep-learning system. The model was trained using a 2D CNN and ResNet-50. The SGD optimizer was used to improve the effectiveness of the model.

3. METHODOLOGY

The diagnostic framework for brain tumor diagnosis is shown in Figure 1. There were 15 different types of CNNs used in this work [37, 38]. We create a late fusion model to increase brain tumor categorization accuracy. We built our model using 15 unique convolutional neural network (CNN) layers. In order to acquire multi-granularity and multiscale global features, we combine the networks' features and employ the multi receptive field. The purpose of the late fusion module is to improve the system's capacity to identify pathogenic and nonpathological regions. In late fusion, we independently run models on all available modalities, extracting their features individually, and then combine these features for classification. To be more precise, our network architecture features a preprocessing dataset, a two-stream network, a late-stage fusion module, and multiple classifications. The two-stream network employs a pair of networks in tandem to collect and profit from a greater variety of disease characteristics.

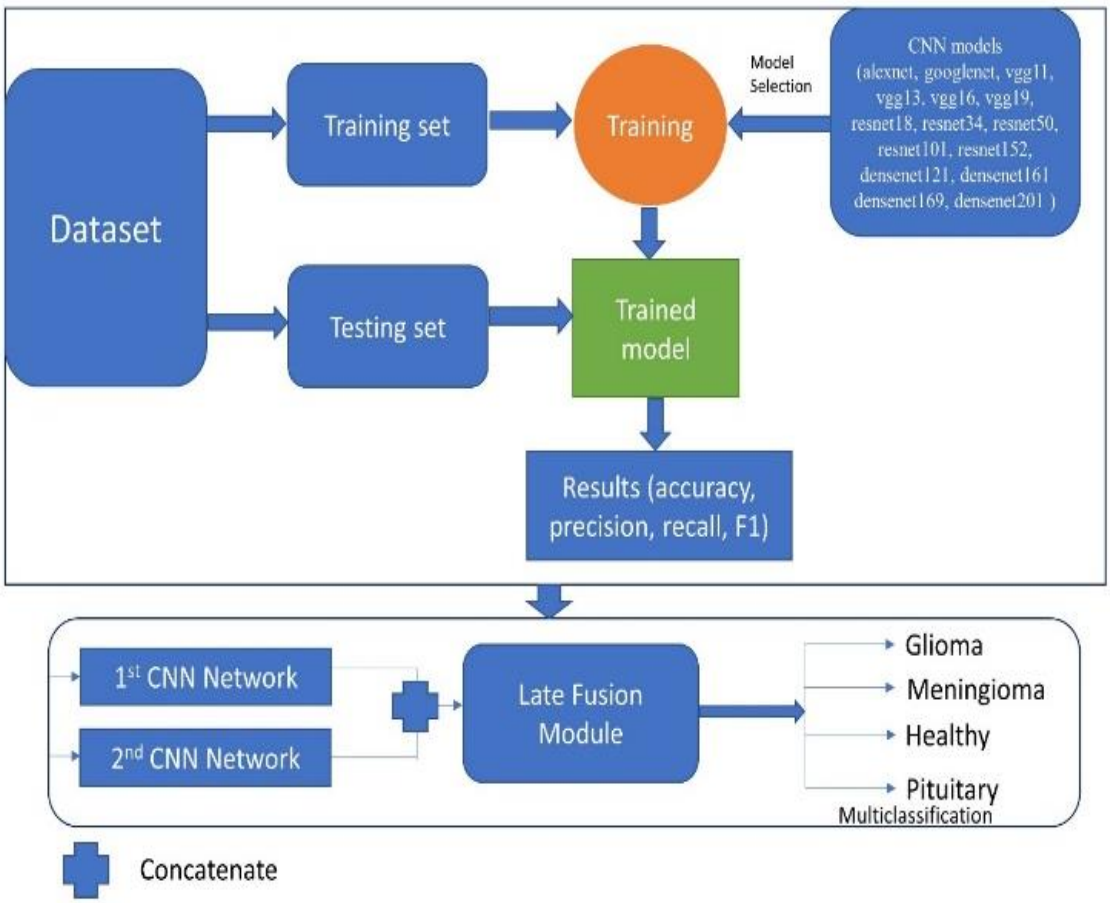


Figure 1. Deep learning framework for brain tumor diagnosis

4. EXPERIMENTAL ANALYSIS

4.1 Dataset

This study made use of data from a publicly available Kaggle dataset [39]. Meningioma, Glioma, healthy, and pituitary MRIs make up the whole set of 3264 pictures. Gliomas, the most prevalent type of malignant brain tumor, occur in spinal cord and brain glial cells. Even though meningiomas are mostly harmless, they can progress to more dangerous tumors if left untreated. Most cases of meningioma develop in the membranes that surround the brain (the meninges). Pituitary tumors, which originate in the pituitary gland located at the base of the brain, are also considered benign. The lack of symptoms shared by meningiomas and pituitary tumors makes them challenging to identify. Physicians should double-check the accuracy of the classifications used. Input images are 64×64 pixels in size. There are a total of 2870 pictures in the dataset's training set, and only 394 in the test set. In Table 1, we can see the breakdown of the data, and Figure 2 shows dataset images from training and testing datasets.

Table 1. Brain tumor distribution into 4 categories

Class	Training	Testing
Meningioma	822	115
Glioma	826	100
Pituitary	827	74
Healthy	395	105

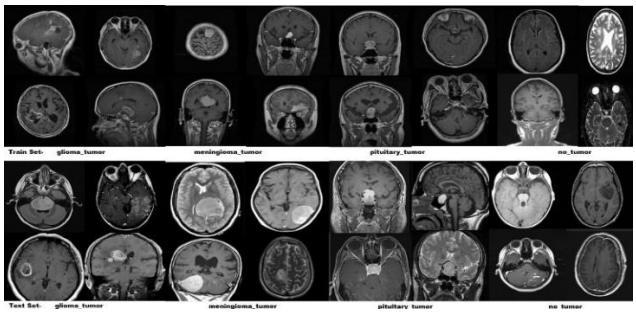


Figure 2. Sample images from dataset [39]

4.2 Implementation

15 pre-trained CNNs are used. We used K80 GPU in combination with a Google Colab notebook running models created with the PyTorch toolkit to conduct our tests. The CNN is trained by decreasing the size of an image from its original 256×226 at random, and then randomly flipping the image horizontally. The learning rate was set at 0.001, the mini-batch size was 16, and the maximum number of epochs for fusion was 60. Their goal was to lay the groundwork for using fusion in computing to facilitate transfer learning. The batch size, maximum epochs, momentum, initial learning rate, and step size employed in the experiments (single CNN and fusion) are summarized in Tables 2 and 3.

Table 2. Hyperparameters for single CNN

Parameter	Value
Epochs	40
Learning Rate	0.001
Minibatch	16 images
Momentum	0.9
Step size	20

Table 3. Hyperparameters for late fusion

Parameter	Value
Epochs	60
Learning Rate	0.001
Minibatch	16 images
Momentum	0.9
Step size	20

A standard metric that is focused on accuracy can be used to quantify the success of a model. The fact that there were enough false positives (FP), and true positives (TP) as well as enough false negatives (FN), and true negatives (TN) gave additional evidence that the quality assessment model was accurate. For the purposes of calculating precision, accuracy, weighted average, F1-score, and recall, the formulas (1), (2), (3), (4), and (5) were utilized:

$$Precision = TP/(TP + FP) \tag{1}$$

$$Accuracy = \frac{(TP + TN)}{(TP + FP + TN + FN)} \tag{2}$$

$$\text{Weighted Avg.} = \frac{\sum_{i=1}^{n} w_i x_i}{\sum_{i=1}^{n} w_i} \tag{3}$$

$$F1 - score = 2 \times (Precision \times Recall) /(Precision + Recall) \tag{4}$$

$$Recall = TP/(TP + FN) \tag{5}$$

4.3 Results

4.3.1 Experiment with transfer learning

The results of 15 deep CNN models being applied to a classification job are presented in Table 4. The confusion matrices for cross-validation are depicted in Figure 3 for Alex-Net, Google-Net, VGG11, VGG13, VGG16, VGG19, ResNet18, ResNet34, ResNet50, ResNet101, ResNet152, densnet121, DenseNet161, DenseNet169, and DenseNet201. We show the following classes in the experiment table for glioma tumor i.e., GT, meningioma tumor i.e., MT, pituitary tumor i.e. PT and no tumor i.e. NT. Here is a detailed examination of each CNN model:

a) Transfer learning with different CNN models

We first tried AlexNet out in a series of experiments. Table 4 displays the experiment results, while Figure 3 depicts the confusion matrix. The AlexNet model has a 77% success rate on the accuracy test. Additionally, the experiment was run through GoogleNet. There are visual representations of both the confusion matrix (Figure 3) and the experimental outcomes (Table 4). Googlenet's testing accuracy was 80%. Tests with vgg11 were performed. Table 4 displays the experimental data, and Figure 3 depicts the confusion matrix in graphic form. On the accuracy test, the VGG11 model performs at 81%. The vgg13 was used in the study. There are visual representations of both the confusion matrix (Figure 3) and the experimental outcomes (Table 4). A total of 80% accuracy was achieved in vgg13 testing. The obtained findings are shown in Table 4, with an accuracy of 81% from vgg16. In Figure 3, we see the confusion matrix. The vgg19 serves as the experimental medium. There are visual representations of both the confusion matrix (Figure 3) and the experimental outcomes (Table 4). Testing accuracy in the vgg19 averaged 81%. Next, we put the ResNet18 to work in the experiment. Figure 3 depicts the confusion matrix, and Table 4 provides an overview of the experiment's findings. An 81% accuracy is reached in ResNet18. The ResNet34 was then used to conduct the experiment. Figure 3 depicts the confusion matrix, and Table 4 provides an overview of the experiment's results. The accuracy of ResNet34 is 82%. The ResNet50 was then used to conduct the experiment. Figure 3 depicts the confusion matrix, and Table 4 provides an overview of the experiment's results. The accuracy of ResNet50 is 82%. The experiment was then conducted using ResNet101. Figure 3 depicts the confusion matrix, and Table 4 provides an overview of the experiment's results. The ResNet101 pass rate is 82%. The ResNet152 was used in the experiment. Figure 3 depicts the confusion matrix, and Table 4 summarizes the experiment's results. In ResNet152, accuracy is at 82%. The DenseNet121 played a key role in the execution of the experiment. Figure 3 depicts the confusion matrix, and Table 4 summarizes the experiment's results. DenseNet121 has the potential to achieve an accuracy of around 80%. The experiment was then conducted using DenseNet161. Figure 3 depicts the confusion matrix, and Table 4 summarizes the experiment's results. To

be more specific, DenseNet161 achieves 81 percent accuracy. Finally, the experiment was carried out using DenseNet169. Figure 3 depicts the confusion matrix, and Table 4 summarizes the experiment's results. The accuracy of DenseNet169 is 82%. DenseNet201 was used to conduct the experiment. Figure 3 depicts the confusion matrix, and Table 4 summarizes the experiment's results. DenseNet201 achieves an accuracy of 82%.

Table 4. Classification results with different CNN models

	Class	Pre.	Rec.	F1	ACC
AlexNet	GT	1	0.25	0.41	77
	MT	0.65	1	0.79	
	NT	0.8	1	0.89	
	PT	0.95	0.76	0.84	
	Weighted avg.	0.83	0.77	0.73	
Googlenet	GT	1	0.27	0.43	80
	MT	0.7	1	0.82	
	NT	0.8	1	0.89	
	PT	0.96	0.92	0.94	
	Weighted avg.	0.85	0.8	0.76	
vgg11	GT	1	0.35	0.52	81
	MT	0.71	1	0.83	
	NT	0.81	1	0.89	
	PT	1	0.89	0.94	
	Weighted avg.	0.86	0.81	0.79	
vgg13	GT	1	0.32	0.48	80
	MT	0.71	1	0.83	
	NT	0.77	1	0.87	
	PT	0.98	0.85	0.91	
	Weighted avg.	0.85	0.8	0.77	
vgg16	GT	0.9	0.46	0.61	81
	MT	0.73	0.97	0.83	
	NT	0.8	1	0.89	
	PT	1	0.8	0.89	
	Weighted avg.	0.84	0.81	0.8	
vgg19	GT	0.93	0.39	0.55	81
	MT	0.72	1	0.84	
	NT	0.81	1	0.9	
	PT	0.98	84	0.91	
	Weighted avg.	0.85	0.81	0.79	
ResNet18	GT	0.81	0.44	0.57	81
	MT	0.8	0.95	0.87	
	NT	0.75	1	0.86	
	PT	0.97	0.84	0.9	
	Weighted avg.	82	81	0.8	
ResNet34	GT	1	0.4	0.57	82
	MT	0.69	1	0.82	
	NT	0.84	1	0.91	
	PT	1	0.85	0.92	
	Weighted avg.	0.87	0.82	0.8	
ResNet50	GT	1	0.38	0.55	82
	MT	0.71	1	0.83	
	NT	0.82	1	0.9	
	PT	1	0.88	0.94	
	Weighted avg.	0.87	0.82	0.8	
ResNet101	GT	1	0.34	0.51	82
	MT	0.71	1	0.83	
	NT	0.8	1	0.89	
	PT	1	0.92	0.96	
	Weighted avg.	0.86	0.82	0.79	
ResNet152	GT	0.96	0.45	0.61	82
	MT	0.74	0.98	0.84	
	NT	0.78	1	0.88	
	PT	1	0.81	0.9	
	Weighted avg.	0.86	0.82	0.8	
DenseNet121	GT	1	0.39	0.56	80
	MT	0.71	1	0.83	
	NT	0.78	1	0.88	
	PT	1	0.78	0.88	
	Weighted avg.	0.86	0.8	0.78	
DenseNet161	GT	1	0.32	0.48	81
	MT	0.68	1	0.81	
	NT	0.83	1	0.91	
	PT	1	0.89	0.94	
	Weighted avg.	0.86	0.81	0.78	
DenseNet169	GT	1	0.34	0.51	82
	MT	0.67	1	0.8	
	NT	0.87	1	0.93	
	PT	1	0.92	0.96	
	Weighted avg.	0.87	0.82	0.79	
DenseNet201	GT	1	0.36	0.53	82
	MT	0.68	1	0.81	
	NT	0.86	1	0.93	
	PT	1	0.92	0.96	
	Weighted avg.	0.87	0.82	0.8	

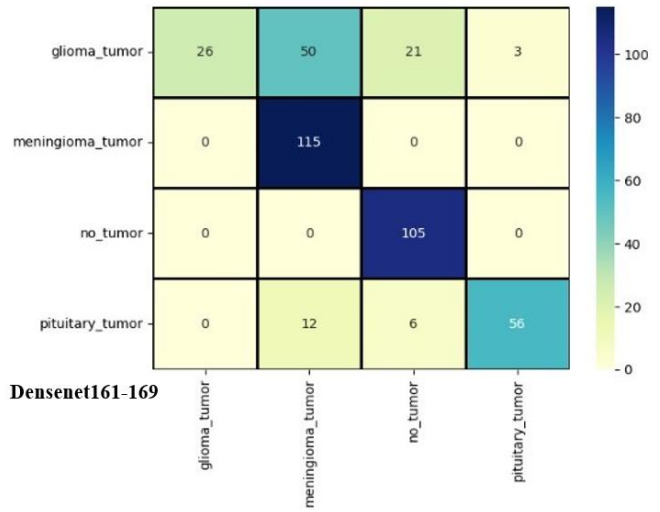
Table 5. Classification results with fusion of different CNN models

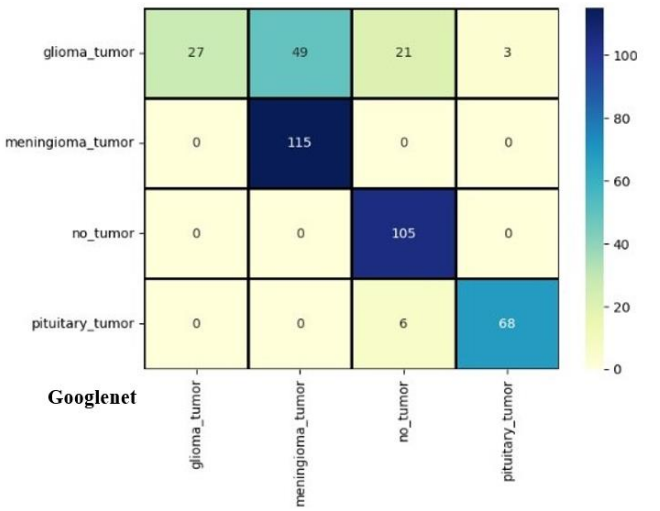
	Class	Pre.	Rec.	F1	ACC
ResNet34+ ResNet50	GT	0.97	0.39	0.56	83
	MT	0.79	0.99	0.88	
	NT	0.75	1	0.86	
	PT	0.99	0.92	0.95	
	Weighted avg.	0.86	0.83	0.8	
ResNet34+ ResNet101	GT	0.98	0.42	0.59	83
	MT	0.77	0.99	0.86	
	NT	0.77	1	0.87	
	PT	1	0.88	0.94	
	Weighted avg.	0.86	0.83	0.81	
ResNet34+ ResNet152	GT	0.97	0.34	0.5	80
	MT	0.76	0.98	0.86	
	NT	0.73	1	0.85	
	PT	0.97	0.88	0.92	
	Weighted avg.	0.85	0.8	0.78	
ResNet50+ ResNet101	GT	0.88	0.59	0.71	86
	MT	0.89	0.91	0.9	
	NT	0.76	1	0.86	
	PT	0.97	0.92	0.94	
	Weighted avg.	0.87	0.86	0.85	
ResNet50+ ResNet152	GT	1	0.35	0.52	82
	MT	0.69	1	0.82	
	NT	0.84	1	0.91	
	PT	1	0.92	0.96	
	Weighted avg.	0.87	0.82	0.79	
ResNet101 +ResNet152	GT	1	0.35	0.52	82
	MT	0.72	1	0.84	
	NT	0.8	1	0.89	
	PT	1	0.92	0.96	
	Weighted avg.	0.87	0.82	0.79	
DenseNet161+DenseNet169	GT	1	0.43	0.6	84
	MT	0.73	1	0.84	
	NT	0.84	1	0.91	
	PT	1	0.92	0.96	
	Weighted avg.	0.88	0.84	0.82	
DenseNet161+DenseNet201	GT	1	0.35	0.52	82
	MT	0.74	1	0.85	
	NT	0.78	1	0.88	
	PT	1	0.92	0.96	
	Weighted avg.	0.86	0.82	0.79	
DenseNet169+DenseNet201	GT	1	0.4	0.57	83
	MT	0.76	1	0.86	
	NT	0.78	1	0.88	
	PT	1	0.92	0.96	
	Weighted avg.	0.87	0.83	0.81	

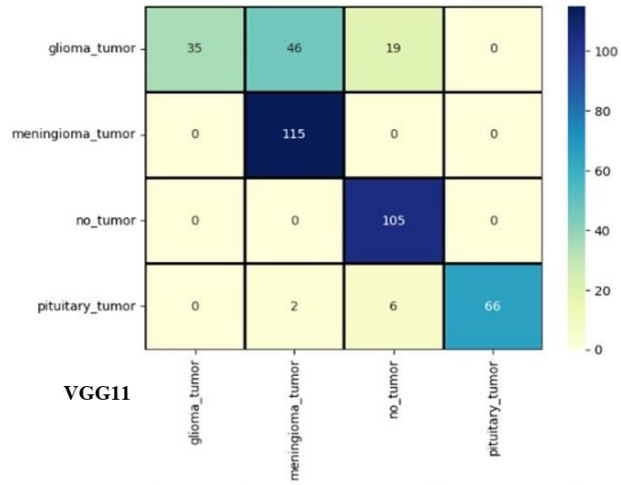
b) Experiment with late fusion of CNN

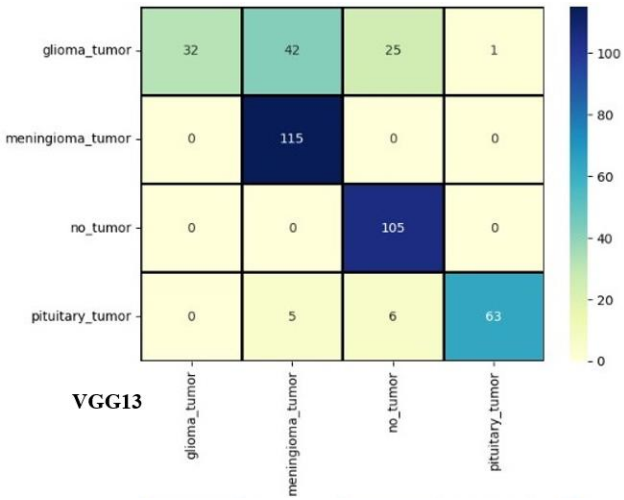
We select the top-performing seven models (ResNet34,

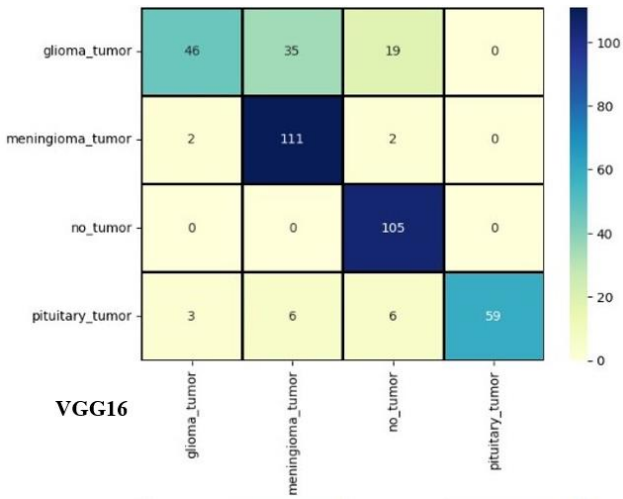
ResNet50, ResNet101, ResNet152, DenseNet161, DenseNet169, and DenseNet201) and employ them in a wide range of late fusion configurations. By fusing ResNet34 and ResNet50, as shown in Table 5, an accuracy of 83% was reached; the related confusion matrix is displayed in Figure 4. ResNet34 and ResNet101 are the two networks chosen for fusion. By fusing ResNet34 with ResNet101, as shown in Table 5, an accuracy of 83% was reached; the related confusion matrix is displayed in Figure 4. We use ResNet34 and ResNet152 for fusion. By fusing ResNet34 with ResNet152, as shown in Table 5, an accuracy of 80% was reached; the related confusion matrix is displayed in Figure 4. ResNet50 and ResNet101 are the two networks chosen for fusion. By combining ResNet50 and ResNet101, as shown in Table 5, an accuracy of 86% was reached; the related confusion matrix is displayed in Figure 4. We use ResNet50 and ResNet152 for fusion. Table 5 shows that by combining ResNet50 and ResNet152, an 82% accuracy was achieved, and Figure 4 depicts the associated confusion matrix. Two networks, ResNet101 and ResNet152, have been selected for fusion. Table 5 shows that by combining ResNet101 and ResNet152, an 82% accuracy was achieved, and Figure 4 depicts the associated confusion matrix. DenseNet161 and DenseNet169 are used in the fusion process. Table 5 shows that by combining DensNet161 and DenseNet169, an 84% accuracy was achieved, and Figure 4 depicts the associated confusion matrix. For the fusion, we use DensNet161 and DenseNet201.

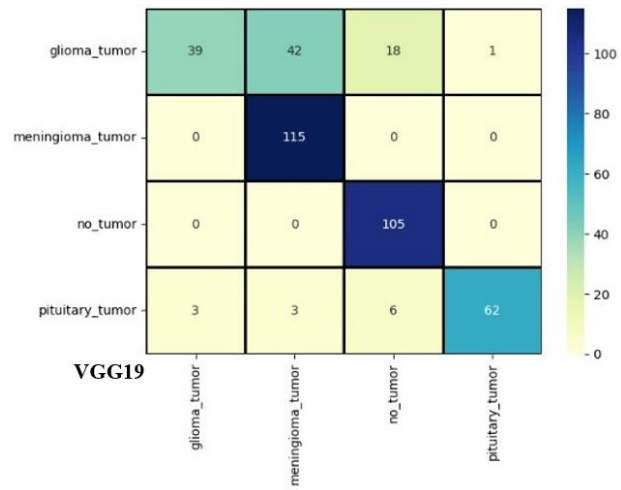


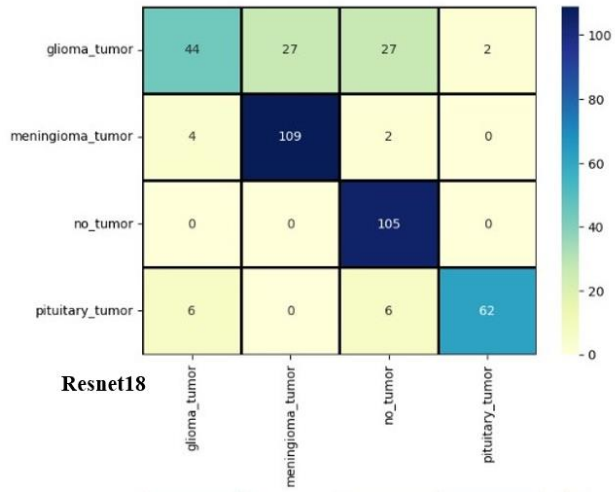




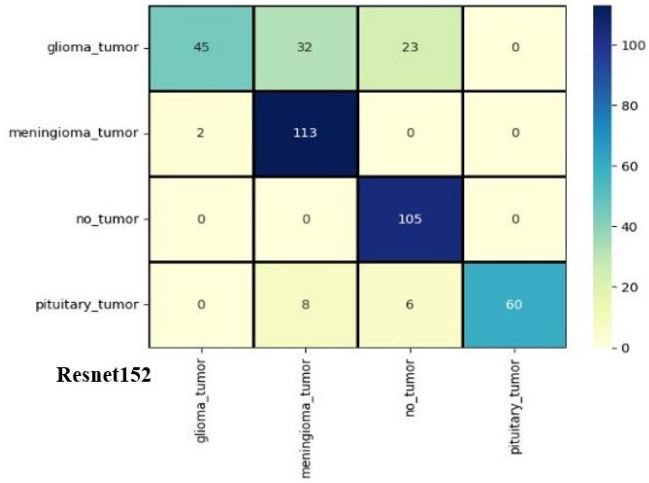




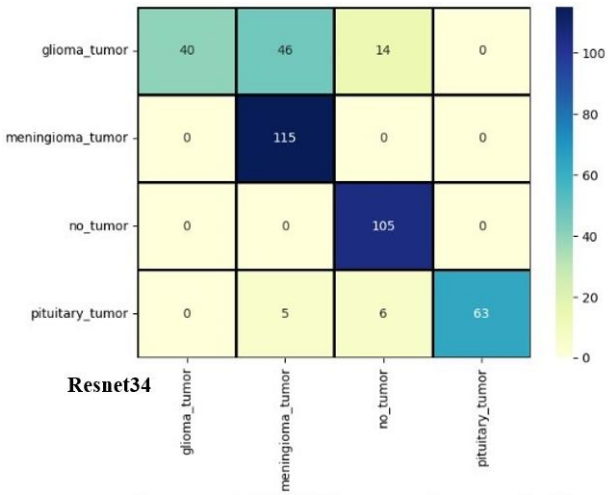




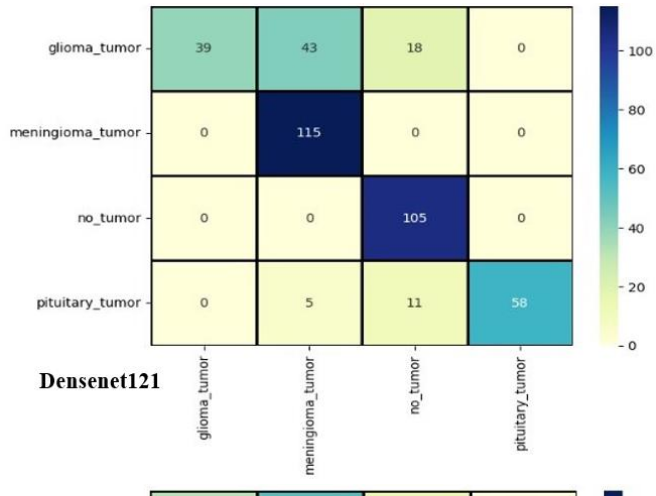
Resnet18



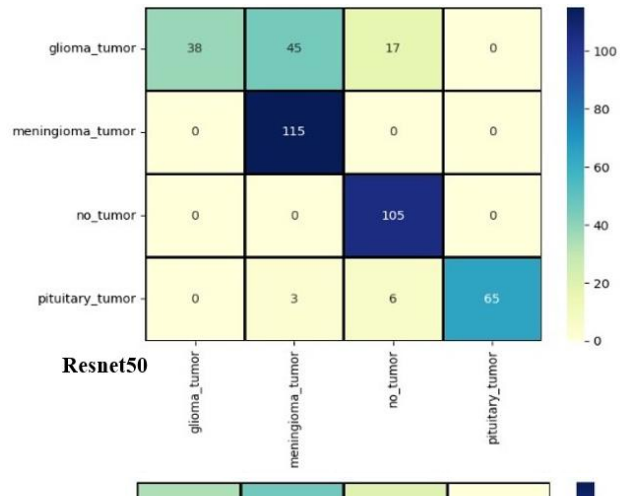
Resnet152



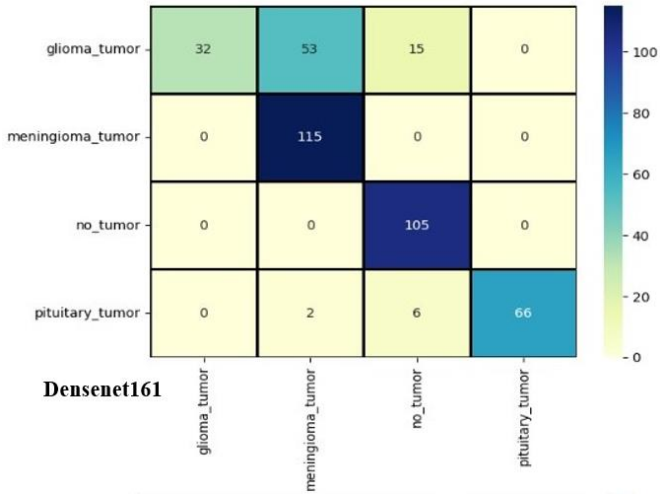
Resnet34



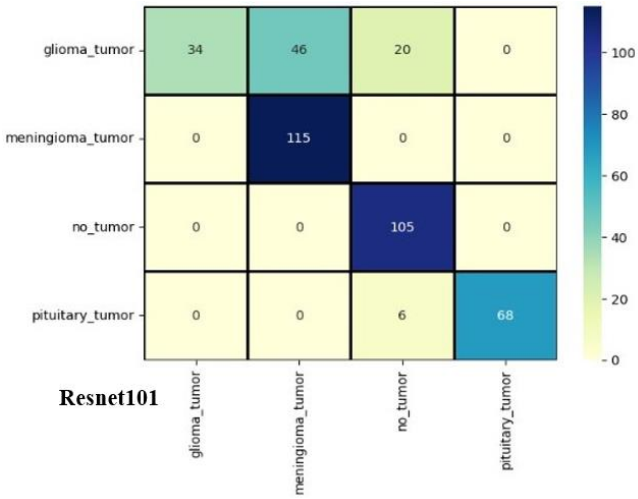
Densenet121



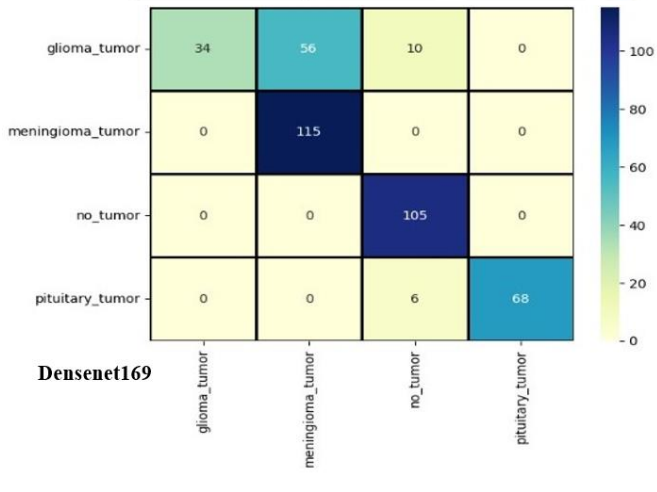
Resnet50



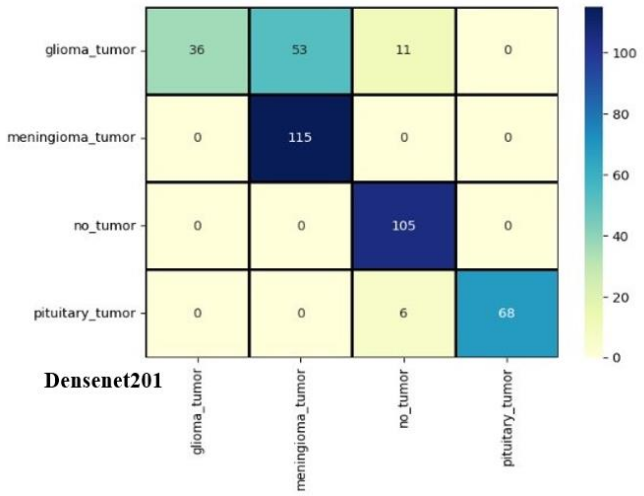
Densenet161



Resnet101



Densenet169



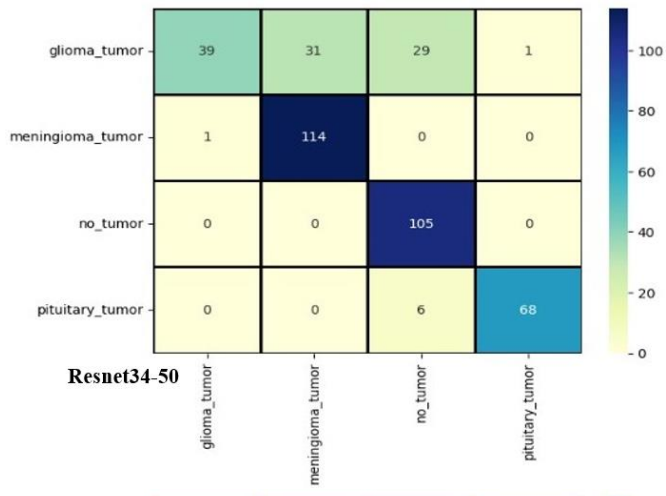
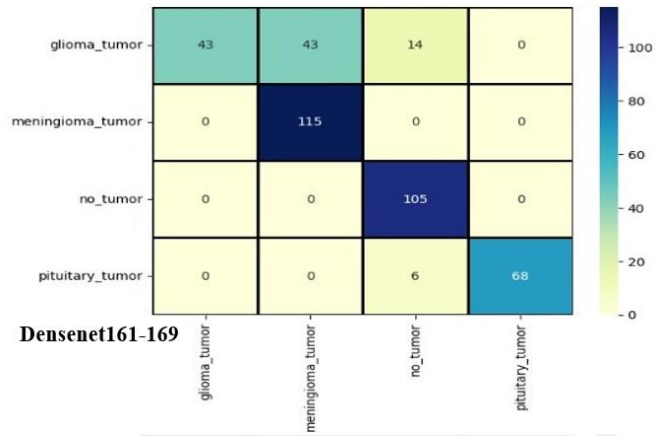
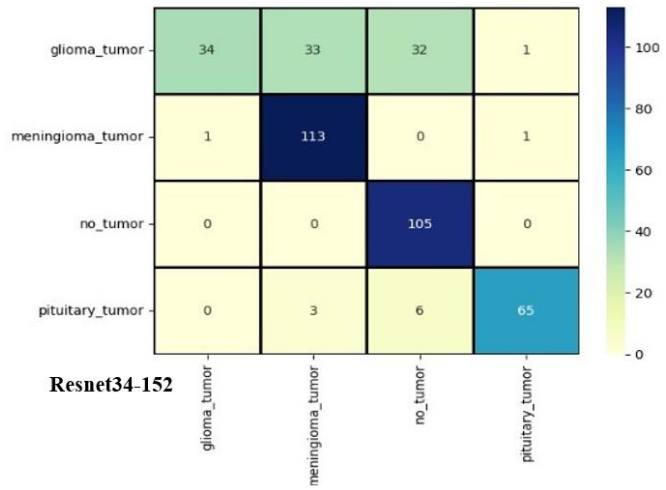
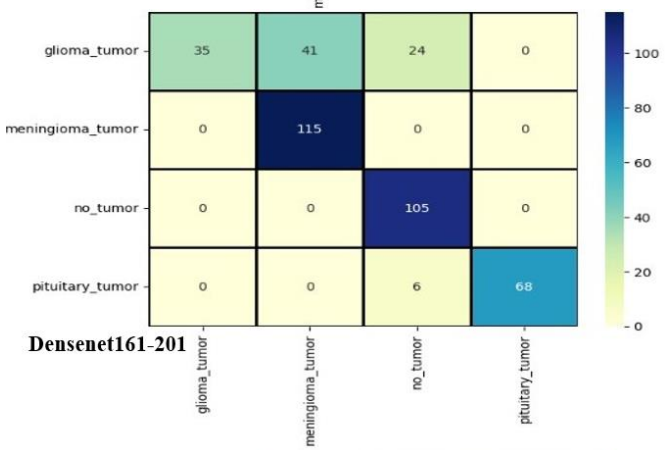
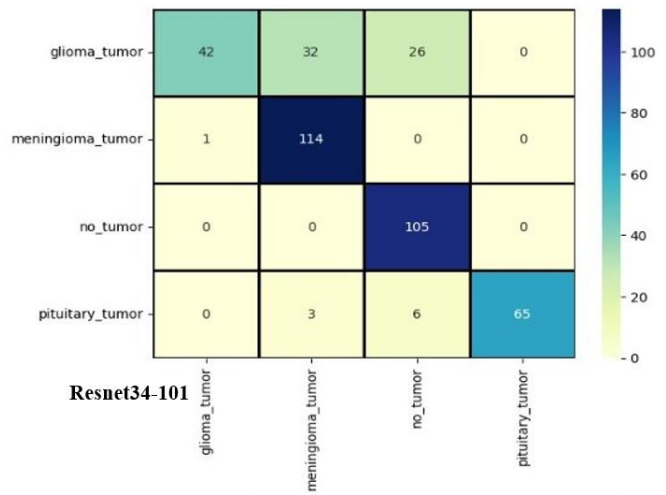


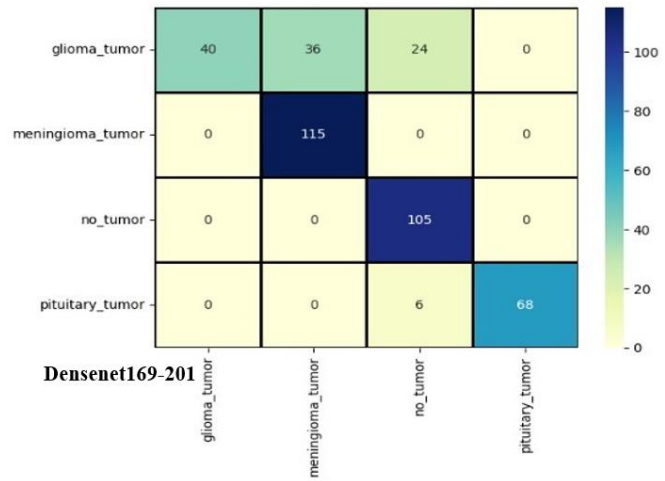
Figure 3. Confusion matrix of all CNN models

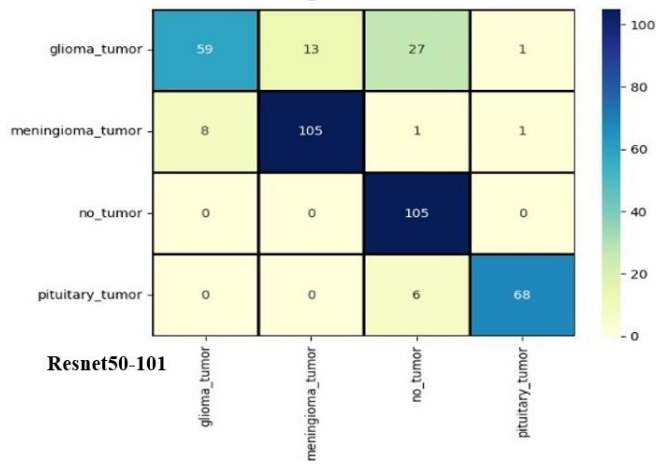












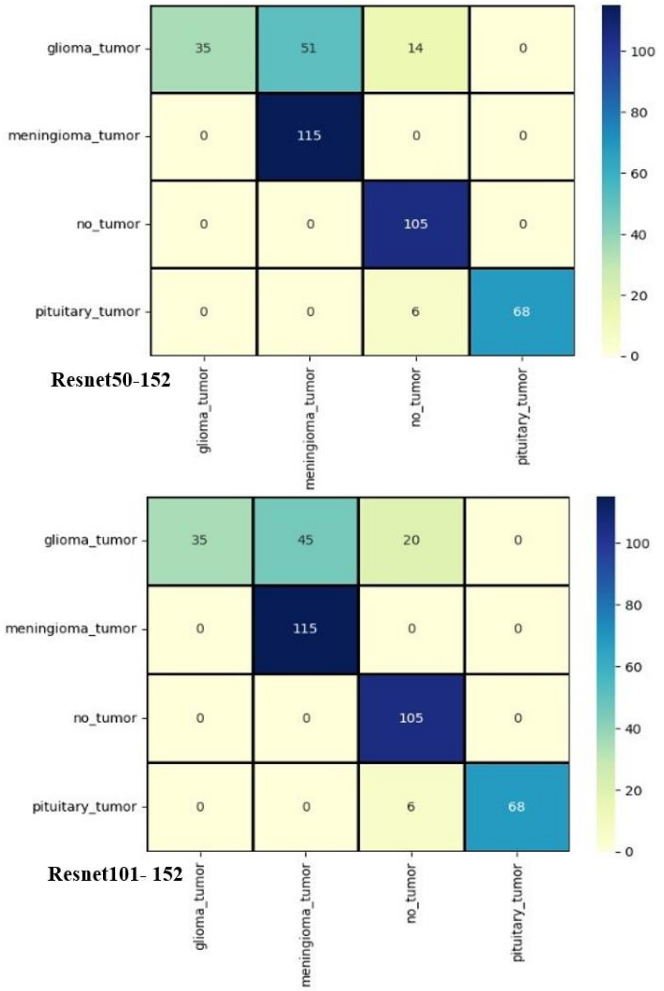


Figure 4. Confusion matrix of all fusion

Table 5 shows that by combining DensNet161 and DenseNet201, an 82% accuracy was achieved, and Figure 4 depicts the associated confusion matrix. DensNet169 and DenseNet201 are used in the fusion process. Table 5 displays the results of using DensNet169 and DenseNet201 together, and Figure 4 shows the related confusion matrix.

c) Analyzing and comparing several CNN Pre-Trained models.

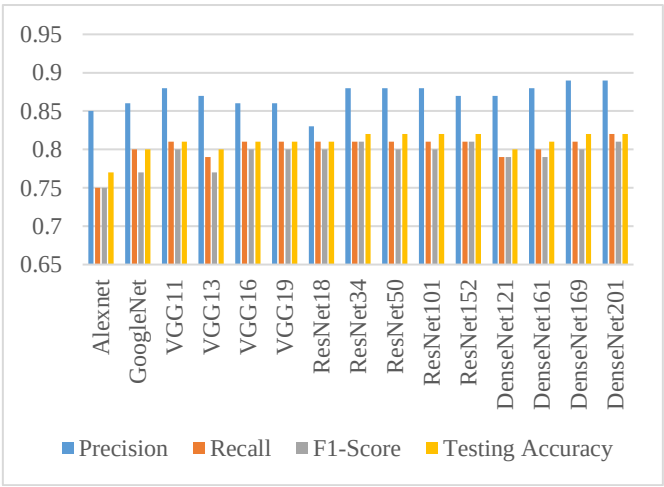


Figure 5. Diagrammatic depiction of classification metrics of fusion models

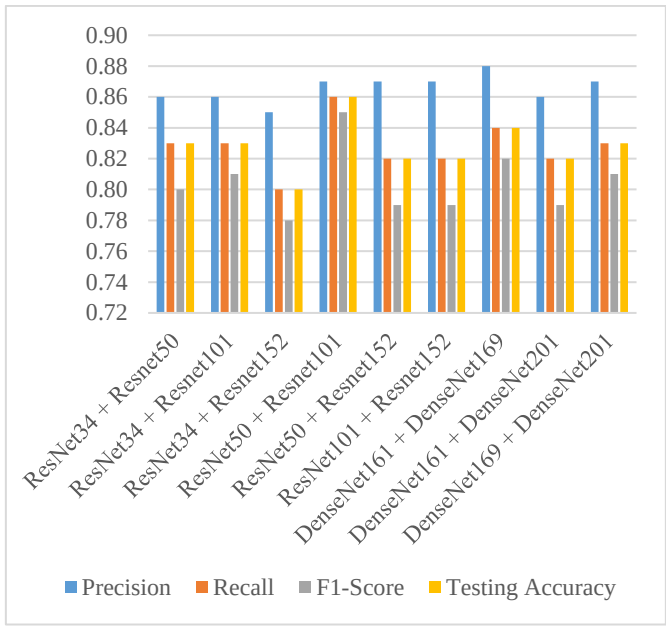


Figure 6. Diagrammatic depiction of classification metrics of fusion models

The findings from different models are compared here, the CNN model is fused, and the differences are explained. Our experiments with numerous CNN models (including AlexNet, GoogleNet, VGGnet11, VGGnet13, VGGnet16, VGGnet19, ResNet18, ResNet34, ResNet50, ResNet101, ResNet152, DenseNet121, DenseNet161, DenseNet169, and DenseNet201) show that ResNet, and DenseNet provide the highest classification accuracy, as shown in Tables 6 and 7. Improvements in precision, recall, and F1-score were especially noticeable when compared to other CNN models. In terms of classification accuracy, fusing ResNet50 and ResNet101 yields an 86% improvement, whereas fusing DenseNet161 and DenseNet169 yields an 84% improvement. There is a graphical representation of fusion methods and pre-trained CNN models in Figures 5 and 6.

Table 6. Evaluation of various pre-trained CNN models

Models	Prec.	Rec.	F1.	Accu.	Time (Training)
AlexNet	0.85	0.75	0.75	0.77	164m 41s
GoogleNet	0.86	0.80	0.77	0.80	367m 34s
VGG11	0.88	0.81	0.80	0.81	994m 47s
VGG13	0.87	0.79	0.77	0.80	1576m 2s
VGG16	0.86	0.81	0.80	0.81	1974m 59s
VGG19	0.86	0.81	0.80	0.81	2344m 39s
ResNet18	0.83	0.81	0.8	0.81	385m 31s
ResNet34	0.88	0.81	0.81	**0.82**	677m 17s
ResNet50	0.88	0.81	0.8	**0.82**	933m 45s
ResNet101	0.88	0.81	0.8	**0.82**	1535m 9s
ResNet152	0.87	0.81	0.81	**0.82**	2167m 5s
DenseNet121	0.87	0.79	0.79	0.80	990m 54s
DenseNet161	0.88	0.8	0.79	**0.81**	1964m 7s
DenseNet169	0.89	0.81	0.8	**0.82**	1193m 39s
DenseNet201	0.89	0.82	0.81	**0.82**	1507m 55s

d) Statistical analysis of late fusion approach

Using the late fusion method, we employed seven CNN models selected based on their accuracy, achieving an accuracy range of 80% to 86%. In contrast, various single CNN models achieved accuracy ranging from 77% to 82%. Thus, our late fusion strategy enhanced accuracy by 3% to 4%.

Table 7. Evaluation of various fusion methods

Models	Prec.	Rec.	F1.	Accu.	Time (Training)
ResNet34+ResNet50	0.86	0.83	0.80	83.00	2299m 31s
ResNet34+ResNet101	0.86	0.83	0.81	83.00	3287m 24s
ResNet34+ResNet152	0.85	0.80	0.78	80.00	4276m 53s
ResNet50+ResNet101	0.87	0.86	0.85	86.00	3616m 39s
ResNet50+ResNet152	0.87	0.82	0.79	82.00	4909m 59s
ResNet101+ResNet152	0.87	0.82	0.79	82.00	5803m 60s
DenseNet161+DenseNet169	0.88	0.84	0.82	84.00	6018m 55s
DenseNet161+DenseNet201	0.86	0.82	0.79	82.00	7266m 58s
DenseNet169+DenseNet201	0.87	0.83	0.81	83.00	3978m 25s

e) Comparative analysis

In this evaluation, we compare our proposed method with previous studies that employed the same brain tumor types but differed in network topologies and parameters. The results of this comparative study, presented in Table 8, demonstrate that our proposed network outperformed all tested methodologies in terms of both classification accuracy and robustness. These findings affirm the reliability and resilience of our model.

Table 8. Comparative analysis of the proposed approach with previous works

Method	Year	Accuracy
Das et al. [40]	2019	84.19
Abiwinanda et al. [41]	2021	79
Wahid et al. [42]	2019	84.18
Devi and Selvaraju [43]	2020	85
Nhlapho et al. [44]	2024	85
Proposed	2024	86

5. CONCLUSIONS

We compared fifteen distinct deep learning models for classifying brain MRI images: the AlexNet, GoogleNet, VGGnet11, VGGnet13, VGGnet16, VGGnet19, ResNet18, ResNet34, ResNet50, ResNet101, ResNet152, DenseNet121, DenseNet161, DenseNet169, and DenseNet201. We used an MRI dataset containing 3,264 images that consist of four categories: glioma, meningioma, pituitary tumor, and healthy. Tables 4-7 display the results with comparisons obtained using various CNN-based methods. Figure 3 and Figure 4 show the cross-validation confusion matrices. Our results show that the ResNet and DenseNet models provide the best classification accuracy (82 percent). Fusion of ResNet50 and ResNet101 achieves 86% accuracy in classification, while fusion of DenseNet161 and DenseNet169 achieves 84% accuracy. Several variations of deep learning models, such as EfficientNet, and their combinations may allow for further progress in the future. Ultimately, we hope to soon explore the prospect of training the model on a massive labeled dataset.

FUNDING

This work is funded by King Saud University, Riyadh, Saudi Arabia (Grant No.: RSP2024R406).

DATA AVAILABILITY

The datasets analyzed during current study are available from the corresponding author on reasonable request.

REFERENCES

[1] Liu, H., Zhang, S., Gamboa, H., Xue, T., Zhou, C., Schultz, T. (2024). Taxonomy and real-time classification of artifacts during biosignal acquisition: A starter study and dataset of ECG. IEEE Sensors Journal, 24(6): 9162-9171. https://doi.org/10.1109/JSEN.2024.3356651

[2] Liang, D., Guan, Q., Huang, M., He, Y., Ou, Y., Chen, M., Zheng, X., Lin, X. (2023). Changing trends of disease burden of stroke from 1990 to 2019 and its predictions among the Chinese population. Frontiers in Neurology, 14: 1255524. https://doi.org/10.3389/fneur.2023.1255524

[3] Li, M., Wei, R., Zhang, Z., Zhang, P., Xu, G., Liao, W. (2023). CVT-based asynchronous BCI for brain-controlled robot navigation. Cyborg and Bionic Systems, 4: 0024. https://doi.org/10.34133/cbsystems.0024

[4] Lee, H. (2023). Engineering in vitro models: Bioprinting of organoids with artificial intelligence. Cyborg and Bionic Systems, 4: 0018. https://doi.org/10.34133/cbsystems.0018

[5] He, B., Dai, C., Lang, J., Bing, P., Tian, G., Wang, B., Yang, J. (2020). A machine learning framework to trace tumor tissue-of-origin of 13 types of cancer based on DNA somatic mutation. Biochimica et Biophysica Acta (BBA)-Molecular Basis of Disease, 1866(11): 165916. https://doi.org/10.1016/j.bbadis.2020.165916

[6] Wang, X., Chen, G., Zhang, Y., Ghareeb, W.M., Yu, Q., Zhu, H., Lu, X., Huang, Y., Huang, S., Hou, D., Chi, P. (2020). The impact of circumferential tumour location on the clinical outcome of rectal cancer patients managed with neoadjuvant chemoradiotherapy followed by total mesorectal excision. European Journal of Surgical Oncology, 46(6): 1118-1123. https://doi.org/10.1016/j.ejso.2020.02.034

[7] Soomro, T.A., Zheng, L., Afifi, A.J., Ali, A., Soomro, S., Yin, M., Gao, J. (2022). Image segmentation for MR brain tumor detection using machine learning: A review. IEEE Reviews in Biomedical Engineering, 16: 70-90. https://doi.org/10.1109/RBME.2022.3185292

[8] Li, Q., You, T., Chen, J., Zhang, Y., Du, C. (2023). LI-EMRSQL: Linking information enhanced Text2SQL parsing on complex electronic medical records. IEEE Transactions on Reliability, 73(2): 1280-1290. https://doi.org/10.1109/TR.2023.3336330

[9] Chahal, P.K., Pandey, S., Goel, S. (2020). A survey on brain tumor detection techniques for MR images. Multimedia Tools and Applications, 79(29): 21771-21814. https://doi.org/10.1007/s11042-020-08898-3

[10] Smith, S.M., Zhang, Y., Jenkinson, M., Chen, J., Matthews, P.M., Federico, A., De Stefano, N. (2002).

Accurate, robust, and automated longitudinal and cross-sectional brain change analysis. Neuroimage, 17(1): 479-489. https://doi.org/10.1006/nimg.2002.1040

[11] Vovk, U., Pernus, F., Likar, B. (2007). A review of methods for correction of intensity inhomogeneity in MRI. IEEE Transactions on Medical Imaging, 26(3): 405-421. https://doi.org/10.1109/TMI.2006.891486

[12] Makropoulos, A., Counsell, S.J., Rueckert, D. (2018). A review on automatic fetal and neonatal brain MRI segmentation. NeuroImage, 170: 231-248. https://doi.org/10.1016/j.neuroimage.2017.06.074

[13] Fawzi, A., Achuthan, A., Belaton, B. (2021). Brain image segmentation in recent years: A narrative review. Brain Sciences, 11(8): 1055. https://doi.org/10.3390/brainsci11081055

[14] Zhao, X., Zhang, G., Chen, J., Li, Z., Shi, Y., Li, G., Zhai, C., Nie, L. (2024). A rationally designed nuclei-targeting FAPI 04-based molecular probe with enhanced tumor uptake for PET/CT and fluorescence imaging. European Journal of Nuclear Medicine and Molecular Imaging, 51(6): 1593-1604. https://doi.org/10.1007/s00259-024-06691-0

[15] Liu, Z., Chen, L., Cheng, H., Ao, J., Xiong, J., Liu, X., Chen, Y., Mao, Y., Ji, M. (2024). Virtual formalin-fixed and paraffin-embedded staining of fresh brain tissue via stimulated Raman CycleGAN model. Science Advances, 10(13): eadn3426. https://doi.org/10.1126/sciadv.adn3426

[16] Thrall, J.H., Li, X., Li, Q., Cruz, C., Do, S., Dreyer, K., Brink, J. (2018). Artificial intelligence and machine learning in radiology: Opportunities, challenges, pitfalls, and criteria for success. Journal of the American College of Radiology, 15(3): 504-508. https://doi.org/10.1016/j.jacr.2017.12.026

[17] Jiang, Y., Edwards, A.V., Newstead, G.M. (2021). Artificial intelligence applied to breast MRI for improved diagnosis. Radiology, 298(1): 38-46. https://doi.org/10.1148/radiol.2020200292

[18] Roblot, V., Giret, Y., Antoun, M.B., Morillot, C., Chassin, X., Cotten, A., Zerbib, J., Fournier, L. (2019). Artificial intelligence to diagnose meniscus tears on MRI. Diagnostic and Interventional Imaging, 100(4): 243-249. https://doi.org/10.1016/j.diii.2019.02.007

[19] Sollini, M., Antunovic, L., Chiti, A., Kirienko, M. (2019). Towards clinical application of image mining: A systematic review on artificial intelligence and radiomics. European Journal of Nuclear Medicine and Molecular Imaging, 46: 2656-2672. https://doi.org/10.1007/s00259-019-04372-x

[20] Liu, J., Pan, Y., Li, M., Chen, Z., Tang, L., Lu, C., Wang, J. (2018). Applications of deep learning to MRI images: A survey. Big Data Mining and Analytics, 1(1): 1-18. https://doi.org/10.26599/BDMA.2018.9020001

[21] Shen, L., Margolies, L.R., Rothstein, J.H., Fluder, E., McBride, R., Sieh, W. (2019). Deep learning to improve breast cancer detection on screening mammography. Scientific Reports, 9(1): 12495. https://doi.org/10.1038/s41598-019-48995-4

[22] Bao, C., Hu, X., Zhang, D., Lv, Z., Chen, J. (2023). Predicting moral elevation conveyed in Danmaku comments using EEGs. Cyborg and Bionic Systems, 4: 0028. https://doi.org/10.34133/cbsystems.0028

[23] Lin, Y., Chen, C., Ma, Z., Sabor, N., Wei, Y., Zhang, T., Sawan, M., Wang, G., Zhao, J. (2023). Emulation of brain metabolic activities based on a dynamically controllable optical phantom. Cyborg and Bionic Systems, 4: 0047. https://doi.org/10.34133/cbsystems.0047

[24] Zhu, W., Wang, X., Cui, P. (2020). Deep learning for learning graph representations. Deep Learning: Concepts and Architectures, Springer, Cham, pp. 169-210. https://doi.org/10.1007/978-3-030-31756-0_6

[25] Pereira, S., Pinto, A., Alves, V., Silva, C.A. (2016). Brain tumor segmentation using convolutional neural networks in MRI images. IEEE Transactions on Medical Imaging, 35(5): 1240-1251. https://doi.org/10.1109/TMI.2016.2538465

[26] Hossain, A., Islam, M.T., Abdul Rahim, S.K., Rahman, M.A., Rahman, T., Arshad, H., Khandakar, A., Ayari, M.A., Chowdhury, M.E. (2023). A lightweight deep learning based microwave brain image network model for brain tumor classification using reconstructed microwave brain (RMB) images. Biosensors, 13(2): 238. https://doi.org/10.3390/bios13020238

[27] Bernal, J., Kushibar, K., Asfaw, D.S., Valverde, S., Oliver, A., Martí, R., Lladó, X. (2019). Deep convolutional neural networks for brain image analysis on magnetic resonance imaging: A review. Artificial Intelligence in Medicine, 95: 64-81. https://doi.org/10.1016/j.artmed.2018.08.008

[28] Anaraki, A.K., Ayati, M., Kazemi, F. (2019). Magnetic resonance imaging-based brain tumor grades classification and grading via convolutional neural networks and genetic algorithms. Biocybernetics and Biomedical Engineering, 39(1): 63-74. https://doi.org/10.1016/j.bbe.2018.10.004

[29] Tandel, G.S., Tiwari, A., Kakde, O.G., Gupta, N., Saba, L., Suri, J.S. (2023). Role of ensemble deep learning for brain tumor classification in multiple magnetic resonance imaging sequence data. Diagnostics, 13(3): 481. https://doi.org/10.3390/diagnostics13030481

[30] Kamnitsas, K., Ledig, C., Newcombe, V.F., Simpson, J.P., Kane, A.D., Menon, D.K., Rueckert, D., Glocker, B. (2017). Efficient multi-scale 3D CNN with fully connected CRF for accurate brain lesion segmentation. Medical Image Analysis, 36: 61-78. https://doi.org/10.1016/j.media.2016.10.004

[31] Chen, S., Ding, C., Liu, M. (2019). Dual-force convolutional neural networks for accurate brain tumor segmentation. Pattern Recognition, 88: 90-100. https://doi.org/10.1016/j.patcog.2018.11.009

[32] Long, J., Shelhamer, E., Darrell, T. (2015). Fully convolutional networks for semantic segmentation. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Boston, MA, USA, pp. 3431-3440. https://doi.org/10.1109/CVPR.2015.7298965

[33] Havaei, M., Davy, A., Warde-Farley, D., Biard, A., Courville, A., Bengio, Y., Pal, C., Jodoin, P.M., Larochelle, H. (2017). Brain tumor segmentation with deep neural networks. Medical Image Analysis, 35: 18-31. https://doi.org/10.1016/j.media.2016.05.004

[34] Menze, B.H., Jakab, A., Bauer, S., Kalpathy-Cramer, J., Farahani, K., Kirby, J., Burren, Y., Porz, N., Slotboom, J., Wiest, R., Lanczi, L., Gerstner, E., Weber, M.A., Arbel, T., Avants, B.B., Ayache, N., Buendia, P., Collins, D.L., Cordier, N., Corso, J.J., Criminisi, A., Das, T., Delingette, H., Demiralp, Ç., Durst, C.R., Dojat, M.,

Doyle, S., Festa, J., Forbes, F., Geremia, E., Glocker, B., Golland, P., Guo, X., Hamamci, A., Iftekharuddin, K.M., Jena, R., John, N.M., Konukoglu, E., Lashkari, D., Mariz, J.A., Meier, R., Pereira, S., Precup, D., Price, S.J., Raviv, T.R., Reza, S.M.S., Ryan, M., Sarikaya, D., Schwartz, L., Shin, H.C., Shotton, J., Silva, C.A., Sousa, N., Subbanna, N.K., Szekely, G., Taylor, T.J., Thomas, O.M., Tustison, N.J., Unal, G., Vasseur, F., Wintermark, M., Ye, D.H., Zhao, L., Zhao, B., Zikic, D., Prastawa, M., Reyes, M., Van Leemput, K. (2014). The multimodal brain tumor image segmentation benchmark (BRATS). IEEE Transactions on Medical Imaging, 34(10): 1993-2024. https://doi.org/10.1109/TMI.2014.2377694

[35] Fritscher, K., Raudaschl, P., Zaffino, P., Spadea, M.F., Sharp, G.C., Schubert, R. (2016). Deep neural networks for fast segmentation of 3D medical images. In Medical Image Computing and Computer-Assisted Intervention-MICCAI 2016: 19th International Conference, Athens, Greece, Proceedings, Part II. Springer International Publishing. Springer, Cham, 19: 158-165. https://doi.org/10.1007/978-3-319-46723-8_19

[36] Asad, R., Rehman, S.U., Imran, A., Li, J., Almuhaimeed, A., Alzahrani, A. (2023). Computer-aided early melanoma brain-tumor detection using deep-learning approach. Biomedicines, 11(1): 184. https://doi.org/10.3390/biomedicines11010184

[37] Gairola, A.K., Kumar, V., Sahoo, A.K. (2022). Exploring multiple deep learning models for skin cancer classification. In 2022 International Conference on Computing, Communication, and Intelligent Systems (ICCCIS), Greater Noida, India, pp. 805-810. https://doi.org/10.1109/ICCCIS56430.2022.10037675

[38] Patwal, A., Diwakar, M., Joshi, A., Singh, P. (2022). Facial expression recognition using DenseNet. In 2022 OITS International Conference on Information Technology (OCIT), Bhubaneswar, India, pp. 548-552. https://doi.org/10.1109/OCIT56763.2022.00107

[39] Bhuvaji, S., Kadam, A., Bhumkar, P., Dedge, S., Kanchan S. (2020). Brain tumor classification (MRI) dataset. https://www.kaggle.com/sartajbhuvaji/brain-tumor-classification-mri.

[40] Das, S., Aranya, O.F.M.R.R., Labiba, N.N. (2019). Brain tumor classification using convolutional neural network. In 2019 1st International Conference on Advances in Science, Engineering and Robotics Technology (ICASERT), Dhaka, Bangladesh, pp. 1-5. https://doi.org/10.1109/ICASERT.2019.8934603

[41] Abiwinanda, N., Hanif, M., Hesaputra, S.T., Handayani, A., Mengko, T.R. (2019). Brain tumor classification using convolutional neural network. In World Congress on Medical Physics and Biomedical Engineering 2018, Prague, Czech Republic, Springer Singapore, 1: 183-189. https://doi.org/10.1007/978-981-10-9035-6_33

[42] Wahid, R.R., Anggraeni, F.T., Nugroho, B. (2021). Brain tumor classification with hybrid algorithm convolutional neural network-Extreme learning machine. Ijconsist Journals, 3(1): 29-33. https://doi.org/10.33005/ijconsist.v3i1.53

[43] Devi, S.R., Selvaraju, P. (2020). Classification of brain tumor image based on high grade and low grade using CNN with LSTM. International Journal of Advanced Science and Technology, 29(7): 3008-3017. http://sersc.org/journals/index.php/IJAST/article/view/17373.

[44] Nhlapho, W., Atemkeng, M., Brima, Y., Ndogmo, J.C. (2024). Bridging the Gap: Exploring interpretability in deep learning models for brain tumor detection and diagnosis from MRI images. Information, 15(4): 182. https://doi.org/10.3390/info15040182