A Bi-Fold Multi-Classification Scheme for Brain Tumor using Deep Convolutional Neural Network

Muhammad Rizwan[1,3], **Huang Po-Tsang** [1,4] **and Mehboob Ali**[2,5]

[1] Department of Industrial and Systems Engineering, Chung Yuan Christian University, Zhongli, Taiwan.
[2] Department of Applied Mathematics, Chung Yuan Christian University, Zhongli, Taiwan.
[3]shahani753@gmail.com, [4] pthuang@cycu.edu.tw, [5]inform.me.8@gmail.com

Abstract. A brain tumor is an uncontrolled and unstructured growth of brain cells inside the skull. Most of the time, tumors are misclassified due to the complexity of the lesion and as a result, the survival rate of the patient effects adversely. Magnetic resonance imaging (MRI) is usually used to identify various types of brain tumors. Due to the advancement in computer-aided systems and bio-informatics ML and DL algorithms has been applied to assist neurologist in decision-making. However, the current techniques are error-prone and time-consuming. Therefore, a Bi-fold CNN model for tumor detection and classification has been proposed with validation accuracies of 99.13% and 98.10% respectively. The proposed system consists of two ends-to-end connected CNN models, where the first has been used to detect tumors while the second is used to classify them into glioma, meningioma, and pituitary tumors. The respective accuracies on blind testing are 99.36% and 97.35%. The model has been compared with ResNet50, Xception, InsepectionV3, and EfficeintNetV2S. The comparison results showed the novelty and superiority of our proposed system. The publicly available dataset has been used in this research. In addition, due to its structural specifications, it has less computational complexity as compared to the existing methods.
Keywords: Bi-fold CNN model, Machine learning, Deep learning, Ends-to-end.

1. Introduction

The human brain is a vital organ, which controls all intentional and unintentional actions of the body. A brain tumor is considered the one of deadliest diseases in medical science. According to the National Foundation for Cancer Research (NFCR), every year 84,000 brain tumor cases have been diagnosed in the United States (US), out of which only 29.7%is considered malignant [1]. Deaths from cancer can be avoided through early detection, although this is not always practicable. A tumor, unlike cancer, may be benign, pre-cancerous, or malignant. Because they often do not spread to other organs and tissues and can be surgically removed, benign tumors differ from malign ones [2].

Without the help and assistantship of computers, experts can't analyze extensive datasets with reliable accuracy. Image processing and artificial intelligence play a vital role in future prediction and classification, especially in the field of clinical analysis of various diseases, fraud and fault detection, traffic control & management, and many more [3]. ML and DL are the techniques to explore previously unidentified regularities and patterns in large amount of dataset. It has many methods in exposing rules, paradigms, and relationships within the clusters of data as well as in creating hypotheses regarding these commonalities, which may be used in compiling new results [4]. These algorithm-based solutions provide an opinion and assist radiologists and medical professionals in making decisions related to the treatment [4].

The CNN models work like a human brain by watching tens and thousands of images to detect and recognize them and this detection task is done by the neurons inside our brain but for CNN it is done by S-cells and C-cells which are not biological cells but mathematical operations. These cells

work together in a way that S-cells sit in the first layer of the model and C-cells sit in the second layer, being connected with each other. The simple concept behind this is to capture simple to complex concepts and convert them into a computational model for detecting patterns in visual data [5]. A DNN learning architecture is proposed by Mohsen et al. to classify brain tumors using MRIs. First, they segmented the images using Fuzzy C-Means, performed feature extraction with discrete wavelet transform (DWT), and implemented feature reduction through principal component analysis (PCA). They achieved 96.97% classification accuracy and compared the results with KNN at k = (1,3) and with Linear discriminant analysis (LDA) [6].

Jemimma et al. introduced a Local Directional Pattern (LDP) based deep learning classification technique that extracts features from image data segments using probabilistic fuzzy C-means clustering (PFC) algorithm. Their proposed model achieved an accuracy of 95.78% [7]. Irmak proposed three CNN models for classifying firstly into tumor and no-tumor with an accuracy of 99.33%. Secondly, they categorize tumors into normal, meningioma, glioma, pituitary, and metastatic with an accuracy of 92.66%. Finally, their last model grades tumors into II, III, and IV with an accuracy of 98.14%, and all the parameters of CNN are automatically tuned with grid-search optimization [8]. A brain tumor grading system is being proposed by Sajjad et al., which first do segmentation of the tumor region by using deep learning-based architecture namely Input Cascade CNN and then implemented augmentation (rotation, filliping, skewness, and shears for geometrical changes) to increase the dataset size to feed neural architecture to get satisfactory results. This augmented dataset is trained on the transfer learning architecture of VGG-19 and gained 90.67% accuracy [9].

As the existing methods either only detect a tumor or just classify the tumor into its types and are complex in terms of time, cost, and efficiency. So, considering all these facts in mind, the focus of this research is to develop a deep learning-based automated system that can help early detection with reliable and trustable accuracy to make the precise decision for the treatment of brain tumors. In this research, we proposed a CNN-based solution for identifying and classifying the tumors via a Bi-fold multi-classification approach. To construct our proposed scheme, we integrated two models in such a way that only tumor MRIs identified by the first model go into the second model for the classification into its respective types.

The remainder of this paper has been organized in the following sequence. Section 2 details material and methods including dataset, experimental tools, our proposed automated system for the identification and classification of brain tumors and its evaluation matrices. In addition, we also demonstrate the specifications of the scheme and the mathematical philosophy behind it. A comparative analysis of results is discussed in section 3 to demonstrate the novelty and superiority of our models, using accuracies and losses for both the training and validation processes. We also presented the confusion matrix to evaluate the validation accuracy of our proposed models and their performances vide blind testing. We finalized this research article with conclusions in section 5.

2. Methods

In this section, we demonstrate our proposed scheme for brain tumor detection and classification. We also introduce the datasets used and experimental tools along with its mathematical philosophy at its back end.

2.1 Dataset

To conduct this research, we used two datasets (D^1 and D^2) of MRIs publicly available on Kaggle. D^1 consists of two folders named training and validation, which are further divided into glioma, meningioma, pituitary, and no-tumor. This dataset consists of 1621 gliomas, 1645 meningiomas, 1757 pituitaries, and 2099 no-tumor MRIs, and this dataset was uploaded in the year 2021[10]. While dataset D^2 consists of 926 gliomas, 937 meningiomas, 896 pituitaries, and 500 no-tumors MRIs. The final

version of the dataset was uploaded and made available in 2020 [11]. Figure 1 shows the samples taken from the dataset to clarify no-tumor, glioma, meningioma, and pituitary tumors.

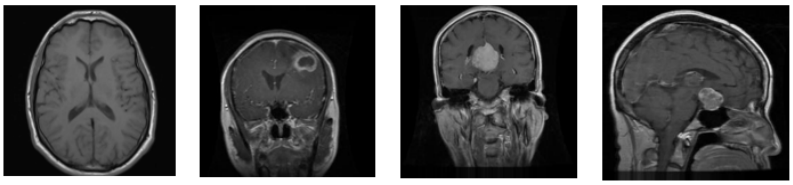

Figure 1. No_tumor, Glioma_tumor, Meningioma_tumor, and Pituitary_tumor

2.3 Experimental Tools

This sub-section explains the proposed experimental setup for the classification of tumors. The CNN-based models have been developed using tensor-flow and Keras with versions 2.9.1 and 2.9 with python version 3.9.6. The experiment has been executed on Central Processing Unit (CPU) core i-5 with 8GB RAM. The model development and implementation tool are Anaconda with Jupiter notebook having version 6.4.8.

2.4 Proposed Scheme

The proposed system consists of two ends-to-end deep learning models named M_1 and M_2. M_1 is a weighted file of binary classification model for detecting the tumor and no-tumor. Whereas M_2 is a multi-classification model for identifying the exact nature of the tumor to categorize it into glioma, meningioma, and pituitary. Figure 2 shows the basic structure of the system and its processing mechanism. The proposed system is a Bi-fold tumor detection framework. It consists of three layers, first one is the pre-processing layer which loads input image data and weighted file of the trained model and then performs the standardization and normalization of the image data. The second layer performs binary classification of the MRIs into tumor and no-tumor by passing through M_1 and forwarding the tumored MRIs to M_2 for their further classification into their exact types, known as glioma, meningioma, and pituitary. Furthermore, the last layer has been designed to demonstrate the results and diagnosis of the exact type of tumor in the form of an excel file.

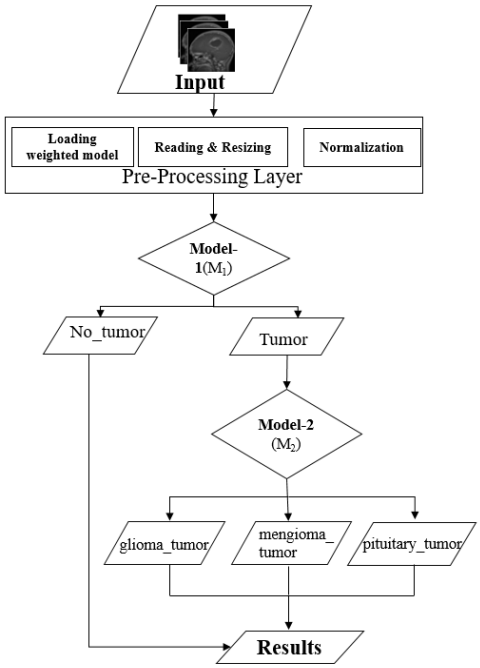

Figure 2. System diagram of the proposed method

For the training of model M_1, two classes having 4526 images of tumors and 1990 images with no tumor, original pulse 2100 augmented (skew, flip, shear, and rotation) images are used. The train validation splits the data into 80:20 ratios, which means 80% data is used for training, and the rest of the 20% is for validation. The original size of the images is 512×512 but its dimension has been changed to 180×180 with RBG channel and the batch size is 32. Our proposed CNN model starts with a convolutional layer of 32 filters, having a 3×3 kernel size. As small filters have a higher capability of detecting small edges, lines, and corners with precision. We then used LeakyReLU as an activation function with a= 0.2. The main reason for selecting LeakyReLU instead of ReLU is to fix the "dying ReLU" problem as well as it enables us to train a deep rectified model from scratch, which is more reliable and effective than many state-of-the-art pre-trained models. Its mathematical representation is furnished as follows:

$$\ell(x) = \begin{cases} x \text{ if } x > 0 \\ ax \text{ if } x \leq 0 \end{cases} \quad (1)$$

Here x is the input variable and a is the constant that controls the angle of the negative slope. Whenever x has a value less than or equal to zero, $\ell(x)$ takes ax as an output otherwise its output equals the input value. Furthermore, the derivative of the LeakyReLU function is described as

$$\ell'(x) = \begin{cases} 1 \text{ if } x > 0 \\ a \text{ if } x \leq 0 \end{cases} \quad (2)$$

The following Figure 3 shows the overall design and the training architecture of the models.

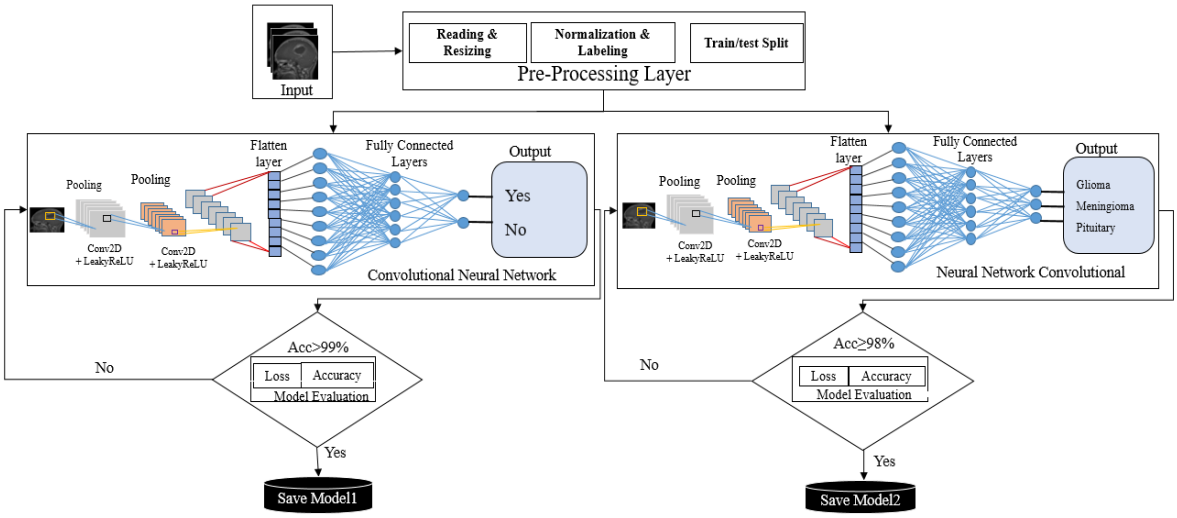

Figure 3. CNN model architecture

Then we added a Max-Pooling layer with a filter size of 2×2 to get the max summary of the given image by accumulating its features and reducing the number of parameters as well as spatial size using a max filter. We then implemented two 64, 128, two 256, and 512 filter convolution layers to our model, followed by Leaky-Relu and Max-Pooling layers with the same parameters having a dropout value of 0.2. Finally, we implemented a fully connected dense layer of 1024 neurons with a sigmoid function as an output function to calculate the probability score for individual classes. Where the mathematical form of the implemented sigmoid function can be expressed as;

$$\delta(x) = \frac{1}{1+e^{-x}} \quad (3)$$

While its derivative function is articulated as;

$$\delta'(x) = \delta(x)(1 - \delta(x)) \quad (4)$$

To calculate the error or difference between the predicted values and the true values, we use the loss function. In order to minimize the error, we use binary cross-entropy in this study with the following algebraic form;

$$\xi(x) = -\frac{1}{n}\sum_{i=1}^{n}\{y_i \, log \, \hat{y}_i + (1 - y_i) \, log(1 - \hat{y}_i)\} \quad (5)$$

Selecting a good optimization function is as much important as the deep learning model because it will affect our learning time as well as accuracy from minutes to hours and days. In this model, we used the Adam optimization algorithm with a learning rate of 0.0001, which is an extension of stochastic gradient descent. It is basically the combination of two gradient descent methods, namely momentum and root mean square propagation. Its mathematical form is presented as follows:

$$\Psi_{t+1} = \Psi_t - \hat{\Lambda}_t \left\langle \frac{\Theta}{\sqrt{\widehat{\Pi}_t + \epsilon}} \right\rangle \quad (6)$$ Where,

$\hat{\Lambda}_t = \frac{\Lambda_t}{1-\lambda_1^t}$ is corrected aggregate of gradients at time t, $\widehat{\Pi}_t = \frac{\Pi_t}{1-\lambda_2^t}$is corrected sum of squares of past gradients, $\Lambda_t = \lambda_1 \Lambda_{t-1} + (1 - \lambda_1) \left\langle \frac{\partial \coprod}{\partial \varpi_t} \right\rangle$is aggregate of gradients at time t,

$\Pi_t = \lambda_2 \Pi_{t-1} + (1 - \lambda_2) \left\langle \frac{\partial \coprod}{\partial \varpi_t} \right\rangle^2$ is sum of squares of past gradients, $\coprod$ is the loss function and ϖ_t represents weights at time t, whereas ΘShows learning rate at time t, ϵ is a small positive constant, and$\lambda_1 = 0.9 \& \lambda_2 = 0.999$are the decay rates of an average of gradients. We then fitted the Model at 20 epochs with a batch size of 32 images resulting in an accuracy rate of 99.13 %.

The realistic identification of the exact nature/type of the tumor is very important to have an accurate diagnosis and treatment. So, we trained M_2 model for a genuine classification of the tumors by taking a dataset consisting of three classes, namely glioma (2147), meningioma (2161), and pituitary (2284). The technical structure of M_2 comprised of six layers, starting with one 64 followed by two 128, 256, and 512 convolutional layers respectively having a 3×3 kernel and LeakyReLU with alpha value of 0.3. To evaluate the performance of the model, we utilized categorical cross entropy as a loss function. We then implemented Adam (lr=0.0003, beta_1=0.9, beta_2=0.999) optimizer to compile the model and it has been fitted at 20 iterations with a batch size of 32, which gives us an accuracy rate of 98.10%.

2.5 Performance Metrics

The selection of an efficient evaluation criterion plays a vital role in effectively assessing the performance of a model. To evaluate our proposed models, we utilized evaluation tools including training and validation accuracies, losses, precision, recall, f1-score and confusion matrix. The mathematical forms for the aforementioned evaluation criterion are demonstrated as follows:

$$A_c = \frac{\Upsilon_p + \Upsilon_n}{\Upsilon_p + \Upsilon_n + \Gamma_p + \Gamma_n} \quad (7)$$

$$R = \frac{\Upsilon_p}{\Upsilon_p + \Gamma_n} \quad (8)$$

$$P = \frac{\Upsilon_p}{\Upsilon_p + \Gamma_p} \quad (9)$$

$$F_1(Sc) = \frac{2 \times (P \times R)}{P + R} \quad (10)$$

Where, Υ_p= True positive, Υ_n= True negative, Γ_p= False positive, Γ_n= False negative, A_c=Accuracy rate, R = Recall, P = Precision, and $F_1(Sc)$ = F_1- Score

3. Result and Discussion

This section presents the comparative study regarding the performance of our proposed models and the existing four state-of-the-art transfer learning techniques available in the literature, namely ResNet50, Inspection V3, Xception, and EfficientNetV2S, using accuracies for both the training and testing data. The testing dataset consists of 510 no-tumor and 906 tumor MRIs, having 300 gliomas, 306 meningioma, and 300 pituitary. Table 1 shows the confusion matrix of Validation and the blind testing results of our proposed model **M_1**.

Table 1. Validation and testing results of M_1

Validation Confusion Matrix			**Blind Testing**		
Tumor	No-tumor	Total	Tumor	No-tumor	Total

Tumor	902	8	910	Tumor	897	9	906
No-tumor	7	807	814	No-tumor	0	510	510

Where it accurately identified a total of 902 images as tumor out of 910 tumor MRIs and miss predicted 8 images as no tumor during the validation process. Furthermore, 807 images were correctly predicted as no-tumor and 7 images were miss-classified as tumor out of 814 no-tumor MRIs. Whereas, during the blind testing it precisely identified 897 MRIs as tumor and 9 images have been miss predicted out of 906. While all of the 510 no-tumor MRIs were correctly predicted.

Table 2. Classification Report for M1

	Precision (%)	**Recall (%)**	**F1-score (%)**
Tumor	99	99	99
No-tumor	99	99	99

The classification report of the model **M1** has been shown in the Table 2, having a percentage of 99 for all of the evaluation tools under consideration.

Table 3. Validation and test results for M2

Validation Confusion Matrix					**Blind Testing**				
	G_t	M_t	P_t	Total		G_t	M_t	P_t	Total
G_t	427	6	2	435	G_t	294	6	0	300
M_t	9	417	3	429	M_t	10	290	6	306
P_t	3	2	450	455	P_t	2	0	298	300

Where, G_t stands for glioma, M_t represents meningioma, and P_t has been used to show pituitary tumor.

The confusion matrix and the blind testing results of our proposed model **M_2** has been presented in Table 3. In which model M_2 classified a total of 427 MRIs as glioma, 6 as meningioma and 2 as a pituitary tumor out of 435 glioma images. Furthermore, for the classification of meningioma, 417 images are correctly predicted while 9 are classified as glioma and 3 as pituitary. For classifying pituitary tumor, 450 MRIs are accurately classified, whereas 3 and 2 were miss-classified as glioma and meningioma respectively out of 455 pituitary MRIs. In the blind testing, out of 300 MRIs of glioma, 294 images were precisely classified, while 6 went into meningioma tumor class. During classification of meningioma tumor, 290 are accurately classified but 10 and 6 were classified as glioma and pituitary tumor respectively out of 306 meningioma images. Lastly, 298 MRIs were accurately classified as pituitary tumors and only 2 went into glioma tumor from a total of 300. Following table 4 shows the classification report of proposed model **M_2** with the measures of precision, recall and F1-score.

Table 4. Classification Report for M_2

	Precision	**Recall**	**F1-score**
Glioma	97	98	98
Meningioma	98	97	98
Pituitary	99	99	99

We also demonstrate the aforementioned comparative analysis in the form of a tabular representation regarding the validation accuracies and the testing accuracies of the models under consideration in the following Table 5.

Table 5. Comparison of validation and testing accuracies

S.No	**Method**	**Validation Accuracy**	**Testing Accuracy**

1	Inspection v3	98.53%	98.02%
2	ResNet50	89.24 %	89.54%
3	EfficientNetV2S	90.41%	89.19%
4	Xception	97.06%	99.15%
5	**Proposed M_1**	**99.13%**	**99.36**

It can be clearly seen in Table 5 that our proposed model has the highest accuracy rate as compared to the existing transfer learning techniques in both validation and testing processes. Hence, on the bases of results achieved from these evaluation measures we can say that our proposed model out-performs the existing models under study.

4. Conclusion

This study proposed a deep learning-based Bi-fold multi-classification system. The initial fold of our proposed system is structured to identify the presence of tumor and the next fold has been architected to classify the tumored images into their respective exact class namely glioma, meningioma, and pituitary. In order to demonstrate the practicality and effectiveness of our proposed system, some existing models like ResNet50, Xception, InsepectionV3 and EfficeintNetV2S have been used for the comparison in terms their validation and blind testing accuracies. It has been observed in the comparison result that our proposed system has outperformed the existing methods by showing highest accuracies and least losses. Furthermore, it is cost effective in terms of computational complexity and reliability. Hence, based on its better performance and results, it can be considered as a reliable and valuable scheme for the early brain tumor detection with its exact nature/type, facilitating the medical professionals for accurate and timely treatment of the patients. Due to unavailability of GPU, we used CPU for our experimentation and drawing the results, which can be considered as the limitation of our study. In the future work, we will extend our research work by doing tumor segmentation to help medical professionals to exactly locate the tumor region.

References

[1]. National foundation for cancer research (2021) *US brain tumor facts and figures*. Available at: https://www.nfcr.org/cancer-types/brain-cancer (Accessed: 15 August 2022).

[2]. Cancer.net (2021) *Brain tumor types and diagnosis.* Available at: https://www.cancer.net/cancer-types/brain-tumor/diagnosis (Accessed: 16 August 2022).

[3]. Alshazly, H., Linse, C., Abdalla, M., Barth, E. and Martinetz, T., 2021. COVID-Nets: deep CNN architectures for detecting COVID-19 using chest CT scans. *PeerJ Computer Science*, **7**, p.e 655.

[4]. Zekiwos, M. and Bruck, A., 2021. Deep learning-based image processing for cotton leaf disease and pest diagnosis. *Journal of Electrical and Computer Engineering*, **2021.**

[5]. Fukushima, K. 1980. Neocognitron: A self-organizing neural network model for a mechanism of pattern recognition unaffected by shift in position. *Biol. Cybernetics* **36**, 193–202.

[6]. Mohsen, H., El-Dahshan, E.S.A., El-Horbaty, E.S.M. and Salem, A.B.M., 2018. Classification using deep learning neural networks for brain tumors. *Future Computing and Informatics Journal*, **3**(1), pp.68-71.

[7]. Jemimma, T.A. and Raj, Y.J.V., 2018, June. Brain tumor segmentation and classification using deep belief network. In *2018 Second International Conference on Intelligent Computing and Control Systems (ICICCS)* (pp. 1390-1394). IEEE.

[8]. Irmak, E., 2021. Multi-classification of brain tumor MRI images using deep convolutional neural network with fully optimized framework. *Iranian Journal of Science and Technology, Transactions of Electrical Engineering*, **45**(3), pp.1015-1036.

[9]. Sajjad, M., Khan, S., Muhammad, K., Wu, W., Ullah, A. and Baik, S.W., 2019. Multi-grade brain tumor classification using deep CNN with extensive data augmentation. *Journal of computational science*, **30**, pp.174-182.