A deep learning based system for accurate diagnosis of brain tumors using T1-w MRI

Mona Ahmed[1], Fahmi Khalifa[1], Hossam El-Din Moustafa[1], Gehad Ahmed Saleh[2], Eman AbdElhalim[1]
[1]Electronics and Communications Engineering Department, Faculty of Engineering, Mansoura University, Mansoura, Egypt
[2]Diagnosti Radiology Department, Faculty of Medicine, Mansoura University, Mansoura, Egypt

Article Info

ABSTRACT

Detection and classification of brain tumors are of formidable importance in neuroscience. Deep learning (DL), specifically convolution neural networks (CNN), has demonstrated breakthroughs in the field of brain image analysis and brain tumors classification. This work proposes a novel CNN based model for brain tumor classification. Our pipeline starts with prepossessing and data augmentation techniques. Then, a CNN classification step is developed and utilizes ResNet50 architecture as its core. Particularly, our design modified the ResNet50 output with a global average pooling (GAP) layer to avoid over-fitting. The proposed model is trained and tested using different optimization algorithms. The final classification is achieved using a sigmoid layer. We tested the proposed structure on T1 weighted contrast-enhanced magnetic resonance images (T1-w MRI) that are collected from three datasets. A total of 3586 images containing two classes (i.e., bengin, and malignant) were used in our experiments. The proposed model reach highest accuracy 99.8%, and optimal error 0.005 using Adam when compared with other six well-known CNN architectures.

Corresponding Author:

Mona Ahmed
Electronics and Communications Engineering Department, Faculty of Engineering, Mansoura University
Elgomhouria St., Mansoura City, Egypt
Email: Monagaffer@std.mans.edu.eg

1. INTRODUCTION

Brain tumors [1] affect all ages, races, and ethnicities. Acorrding to National Brain Tumor Society (NBTS), an estimated 700,000 americans are living with a primary brain tumor approximately 71% of all brain tumors are benign and 29% of all brain tumors are malignant. In 2022, more than 18,200 people are estimated to lose their life because of a malignant brain tumor. There are more than 120 different types of brain and central nervous system (CNS) tumors. Almost one third (29.7%) of the brain and central nervous system tumors are malignant [2]. Survival after diagnosis with a primary brain tumor varies widely by age, geographical location, type of tumor, tumor location and molecular markers [3]. Clinically, several imaging techniques can be used to detect and classify brain tumors. For example, magnetic resonance imaging (MRI) [4] is one of the most common non-invasive techniques. MRI is primarily a non-ionizing radiograph and is considered the most common technique for diagnosing brain tumors as compared to computed tomography or ultrasound imaging [5]. Human evaluation of the MR images is the most common and accurate method for detecting and grading brain tumors from MRI images. However, this is faced by multiple challenges. First it relies heavily on the radiologists' extensive analysis of image characteristics, i.e., it is a heavy workload, time-consuming and error-prone. Secondly, the classification of tumors using MRI images is a difficult task due to overlap-

ping intensities, inconsistency in size, shape, and orientation. Other difficulties include ambient disturbances and low image contrast [6]. As a result, radiologists have demonstrated an interest in modern computer-aided diagnostic (CAD) systems to reduce time and effort. Therefore, there is a continuous need to create computerized diagnostic tool to help clinicians automatically identify and diagnose brain tumors from medical images. In recent years, the use of various MRI-based non-invasive methods, or CAD systems, for diagnosing brain tumors has greatly increased to provide fast and improved diagnoses. Recent advances in machine learning (ML) [7] and deep learning (DL) combined with advanced computing techniques, enabled efficient MRI-CAD systems for tumor diagnosis with minimal, or even without, human intervention. DL [8], [9] techniques have been widely used in the automatic analysis of radiological images [10]. In the last few years, various research works have been proposed for detecting and classifying brain tumors using MR images. Those studies employ different algorithms ranging from traditional ML algorithms to advanced DL models [11]. Next, we'll show you the latest work on brain tumor classification and detection focusing on the same dataset of *figshare* [12]. Cheng *et al.* [13] *figshare* was used for the experiment on the brain tumor dataset. They used the augmenting tumor region as a region of interest and applied an adaptive spatial-division approach to subdivide these regions into sub regions. Model-based features such as intensity graph, gray-level Co-occurrence matrix, and a bag of words were extracted, on features extracted using the circular division approach, they claim the highest accuracy are of 87.54%, 89.72%, and 91.28%. Ismael and Abdel-Qader [14] introduced 2D discrete wavelet transform (DWT) and Gabor filter techniques, statistical characteristics from MRI slices were extracted. They used a back-propagation multi-layer perceptron neural network to classify the brain tumors and got the greatest accuracy of 91.9%. Afshar *et al.* [15] proposed a brain tumor classification using capsule network (CapsNet). The spatial relationship between the tumor and its surrounding tissues was used to classify the tumor in CapsNet. The proposed CapsNet achieves an overall accuracy of 90.89%.

Abiwinanda *et al.* [16] introduced different CNN architectures for brain tumor diagnosis. The CNN with two layers of convolution, activation (RELU) and max pool, followed by one hidden layer of 64 neurons, produced the highest accuracy among the approach. They achieved 98.51% and 84.19% on training and validation sets, respectively. However, the validation accuracy was not high and a color balancing step was required. Anaraki *et al.* [17] employed CNN using a genetic approach to classify brain tumor images. It has six convolutional and max-pooling layers, along with one fully connected layer. The authors achieved an accuracy of 94%. In sum, literature work is rich with many automated brain tumor detection systems. However, existing work is still not producing good results and there is a high demand for reliable CAD tools for brain tumor identification. ML methods and models need domain knowledge and experience. Those methods involve the segmentation and extraction of hand-crafted structural or statistical features, which may lead to a decrease in system accuracy and efficiency. In the proposed model to overcome this limitation, we adopted a DLS based on deep transfer learning [18] for automated extraction of rich and robust visual and discriminative characteristics of brain MRI tissue to classify brain MR images. In this work, we propose a novel CAD system for brain tumor classifications using an enhanced DLS applied to T1w MRI. Unlike the majority of the literature approaches that use public dataset (*figshare*), the proposed DLS is evaluated on the T1w MRI *figshare* dataset and two other*locally-acquired* and *publicly-available* brain tumor datasets. Additionally, to solve the vanishing gradient problem affecting deep neural networks, we employed residual network (ResNet50) as the backbone of our deep learning system (DLS), in addition to transferring learning and GAP. The proposed pipeline is extensively evaluated using several performance metrics, including, (A_c), (P_r), (R_r), and F1-Score. The experimental results show that the proposed DLS is competitive when compared to the other approaches. The rest of the paper is sectioned as follows. The proposed methodology is discussed in detail in Section 2. Section 3 delves into the evaluation and discussion of experimental results. Finally, work conclusions are given in Section 4.

2.   METHOD

The proposed DLS for brain tumor diagnosis is shown in Figure 1. The analysis starts with data preprocessing and augmentation techniques. A pre-processing step is conducted before the images are fed into network. The next step in our pipeline is CNN classification. The data sets and CNN network hyper-parameter settings, optimization algorithm, training and performance computations are fully described in the following sub section.
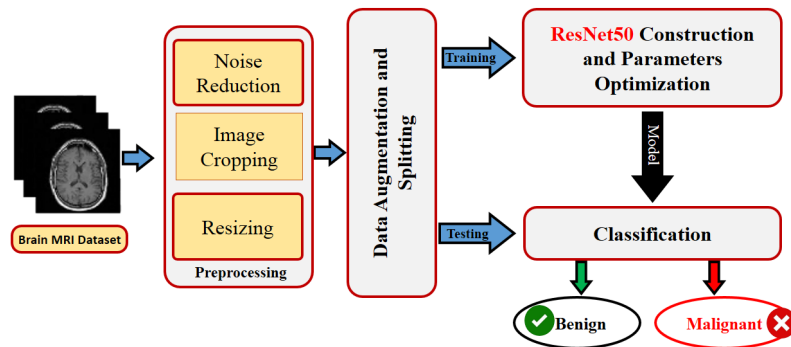
Figure 1. Block diagram of the proposed methodology

2.1. Data Sets

For the evaluation of the proposed DLS, we used a locally-acquired T1-w MRI dataset and two publicly-accessible datasets. In the locally-acquired dataset, a total of 30 patients' data images were collected using a 1.5 Tesla ingenia MR scanner (Philips medical systems, best, Netherlands) using standard head coils for brain studies at the radiology department, Mansoura University hospitals, Mansoura, Egypt. All participants were given complete information about the study's objectives and they gave their informed consent. T1-weighted, T2-weighted, and FLuid-attenuated inversion recovery (FLAIR) sequences were first acquired using the following parameters: T1-weighted (repetition time (TR)/echo time (TE) =580/15 ms), T2-weighted (TR/TE = 4432/100 ms), and FLAIR (TR/TE/inversion time (TI) = 10,000/115/2700 ms). The scanning parameters were an 80 x 80 matrix, a 250 x 170 mm2 field of view (FOV), and a 5 mm slice thickness. After intravenous administration of gadolinium-based contrast agent (adulterate meglumine) at a dose of 0.1 mmoL/kg at a flow rate of 2mL/sec using an automated injector with a maximum dose of 10 ml using a 20–22 G venous cannula, post-contrast T1-weighted axial, sagittal, and coronal images were obtained. In addition, the first public dataset is collected from nan fang hospital and a general hospital, Medical University of Tianjin, China. This dataset was released in 2015, upgraded in 2017 [12] and are accessible from *figshare* website. Axial, coronal, and sagittal T1-CE images were visible. The second public dataset is accessible from the Kaggle website [19]. All data sets have binary labels or classes for benign ("0") and malignant ("1"). The total number of cases in each category is 165 and 3421, respectively. Examples of T1w MRI images for each class from three datasets are shown in Figure 2.
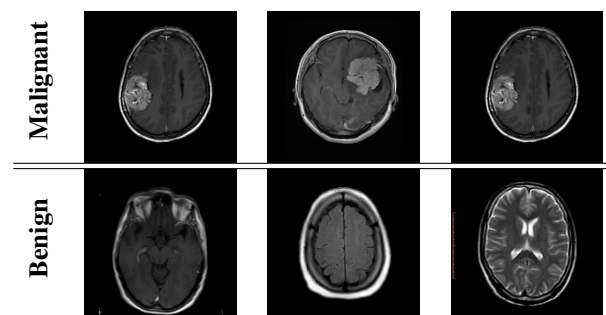

Figure 2. Examples of malignant and benign cases from the three data sets used in our study

2.2. Data preprocessing and augmentation

Before testing, data preparation is essential. In our pipeline, two steps are performed: pre-processing and augmentation. First, to reduce the overall processing time for training and testing, a pre-processing step on the original images is performed. The image dimensionality is scaled-down. Pre-processing of an image are

certain operations of images at the ground level of abstraction. It refers to all of the transformations performed on the raw data before it is fed into DL model. It is implemented in order to improve the image and to remove any erroneous data present in the obtained image. Some of these techniques are used in this paper such as cropping and resizing [20]. Then split data into: (80%) for training, and (20%) for validation and test set. Table 1 summarizes the number of samples in training, validation, and test sets. IN this paper, we use data augmentation techniques to reduce overfitting [21].

Table 1. Data split showing the number of samples in training, validation, and test for the three dataset

	Malignant	Benign
Training	2394	115
Testing	343	17
Validation	684	33

2.3. Proposed CNN architecture

For image classification, a deep CNN has become the most used model. DL's self-learning capabilities have totally changed the field of medical image diagnosis, including brain tumor detection. Many DL systems have been proposed in literature in various application for more accurate disease detection based on increasing (stacking) more layers to extract more image features. However, deeper structures with large number of layers suffer from the vanishing gradient problem. We developed a residual network, namely ResNet50, to increase the performance while reducing the computation time. ResNet50 contains five different types of convolution blocks, each with its own set of layers and filter sizes. CNN networks are built on a foundation of convolutional or transformation layers. The technique of sliding a filter across the entire image is known as convolution. This layer's goal is to generate property mappings in the convolution layer. The ResNet50 model was used with a (224×224×3) input size (color image) and it used filters with variable sizes. A convolution layer and a max-pooling layer compose the first convolution block (CONVC1). This convolution block represents the model's input layer. There are three convolution layers in each of the remaining convolution blocks, each with a different size kernel.

The number of filters employed in each convolution layer varies between these blocks, as shown in Figure 3. The three convolution layers in C2 are with different filters. The three convolution layers, C3, C4, and C5 contain different filters. To avoid the vanishing gradient problem, each convolution block uses a skip relation. Basic ResNet-50 architecture consists of one ConC1 block, three ConC2 blocks, four ConC3 blocks, six ConC4 blocks, and three ConvC5 blocks. Finally, the output layer of basic ResNet50 has been substituted for with GAP and sigmoid layers in the proposed model. The sigmoid activation function converts non-normalized outputs to binary outputs of 0 or 1. As a result, it aids in the final classification of patients with malignant ("1") or benign ("0"). Feature extraction is done by convolution blocks. Instead of a fully connected layer at the input layer of GAP, the proposed approach uses one at the output layer. The GAP layer converts a $(M \times M \times N)$ feature map to a $(1 \times N)$ feature map, where (M ×M) refers to image size and N refers to the number of filters. The approaches of average pooling and maximal pooling are widely employed. The filters in the pooling layer are chosen in $N \times N$ sizes. It uses vectorized feature maps, which can be viewed as perception map categories, to conduct a linear transformation. It is more local to the convolution structure by implementing correspondences between feature maps and categories. The GAP layer also has the following advantages; over-fitting is avoided at this layer because it does not require parameter optimization. It performs as a flatten layer, converting a multidimensional extracted features into a one-dimensional input vector. Furthermore, it takes less time [22]. DL algorithms employ some form of optimization technique to reduce the cost function or the error function. In binary cross-entropy mathematically equation of error is given by:

$$F(\theta) = [z \log P(z) + (1 - z) \log(1 - P(z))] \tag{1}$$

where z is a true label, $p(z)$ is the predicted label. In this paper, we examined multiple optimization algorithms and compared their performance. The first is the adaptive moment estimation (Adam) is a method for calculating adaptive learning rates for individual parameters. In addition to accumulating the mean of the exponential degradation of the squared gradients, x_t also stores the mean of the exponential degradation of the previous gradients u_t in momentum.

$$\left.\begin{array}{l} u_t = \beta_1 u_{t-1} + (1 - \beta_1)\, m_t \\ x_t = \beta_2 x_{t-1} + (1 - \beta_2)\, m_t^2 \end{array}\right\} \tag{2}$$

where u_t and x_t are the first and second moments of gradient, respectively. Secondly, stochastic gradient descent with momentum (SGD) is an optimization methodology that speeds up the descent on a suitable trajectory and reduces its vibration. This is done by adding γ from the previous step update vector to the current update vector.

$$\left.\begin{array}{c} x_t = \gamma x_{t-1} + \eta \nabla_\theta F(\theta) \\ \theta = \theta - x_t \end{array}\right\} \tag{3}$$

Furthermore, adaptive gradient (AdaGrad) is a randomized optimization method that adapts the learning rate to parameters. It performs updates to parameters associated with frequently occurring features. In its update rule, AdaGrad adjusts the general learning rate η at each time t step for each parameter θ based on previous gradients against θ:

$$\boldsymbol{\theta}_{t+1} = \boldsymbol{\theta}_t - \frac{\eta}{\sqrt{\mathbf{D}_t + \epsilon}} \mathbf{m}_t \tag{4}$$

here D_t represents a diagonal matrix, where each of its diagonal elements represents the sum of squared gradient estimated with respect to θ at time step t. In addition to Adam, SGD, and AdaGrad, the RootMeans square propagation (RMSProp) is an adaptive learning technology, developed to answer the problem of the significantly weak learning rate. It is using the average exponential deterioration of the date so that it can converge quickly once a convex bowl is found. Given as (5):

$$\boldsymbol{\theta}_{t+1} = \boldsymbol{\theta}_t - \frac{\eta}{\sqrt{E\left[\mathbf{m}_t^2\right] + \epsilon}} \mathbf{m}_t \tag{5}$$

finally, AdaptiveDelta (AdaDelta) is a stochastic optimization technique that enables the learning rate method for each dimension of SGD. It is an extension of AdaGrad that seeks to reduce the rate of aggressively decreasing monotonously. Instead of aggregating all prior square gradients, AdaDelta constrains the accumulated prior gradients window to a fixed size w.

$$\Delta\theta_t = -\frac{RMS[\Delta\theta]_{t-1}}{RMS[m]_t} m_t$$
$$\theta_{t+1} = \theta_t + \Delta\theta_t \tag{6}$$

Then, we have investigated algorithms that are most commonly used for optimizing AdaGrad, AdaDelta, RMSProp, Adam and SGD.

In conclusion, RMSProp is an AdaGrad enhancement that addresses AdaGrad's rapidly declining learning rates. It's the same as AdaDelta, except that in the numinator update rule, AdaDelta employs the RMS of parameter updates. Also, it is possible that SGD will take much longer than some of the other optimizers. Finally, Adam is too fast, converges rapidly, and has less variance. We practiced our model with 180 epochs and used different optimizers with the defined learning rate till (1e-6). The proposed model was also implemented using Python 3.6, with the Keras 2.2.4 library and Tensorflow 1.13 as the backend. The values of the hyper parameters are shown in Table 2.
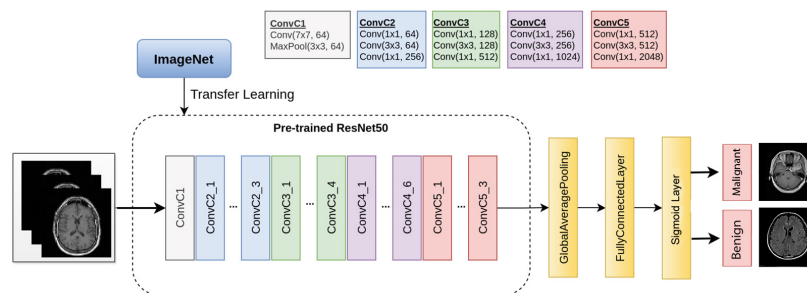

Figure 3. Block diagram of the proposed CNN architecture

Table 2. Summary of the proposed structure parameter settings

Layer	Out shape	No. of parameters
Input	(224, 224, 3)	0
ResNet50 (Model)	(None, 7, 7, 2048)	23587712
Global average pooling	(Gl (None, 2048)	0
Dense	(None, 1)	2049
Batch size	32,16	
Learning rate	1e-6	
No of epochs	180	
Total parameters	23,589,761	
Trainable parameters	23,536,641	
Non-trainable parameters	53,120	

3.   RESULTS AND DISCUSSION

The proposed framework was trained and tested on the T1w-MRIs subjects described in Section 2.1. Transfer learning-based DL algorithms are assessed in this paper for the precise classification of tumor types into benign and malignant. The proposed DLS evaluation is based on four metrics A_c, P_r, R_r and F1-Score. The A_c is the ratio between the number of correctly classified cases and the total number of cases is defined as (7).

$$A_c = \frac{TP + TN}{TP + TN + FP + FN} \tag{7}$$

Here TP (TN) is the number of correctly classified malignant (benign) cases, and FN (FP) is the number of incorrectly classified malignant (benign) cases. Generally, accuracy is a commonly-used metric to measure the model's performance and comparison. Nevertheless, using it alone can be misleading as it lacks the sensitivity to imbalanced data and increases the performance of certain classes [23]. Thus, we also used P_r, R_r, F1-score metrics. So over the imbalanced MRI image dataset, we should have a good predictor of model generalization. Since later metrics are affected by class inequality, they are used to show the overall performance of the model irrespective of the count of the individual class. P_r is used to determine how accurate classifiers are. The presence of a large number of FPs in the classifier is indicated by a low P_r value. Also, the lower the R_r value, the more likely the classifier has problems with huge FP values. Finally, the F1-score is used to accurately handle the distribution issue. When there are unbalance classes in the dataset, it is useful. The harmonic mean of P_r and R_r is frequently used to describe it. The mathematical formulations for P_r, R_r, and F1-score are defined, respectively, as (8) to (10).

$$P_r = \frac{TP}{TP + FP} \tag{8}$$

$$R_r = \frac{TP}{TP + FN} \tag{9}$$

$$F_1\text{-}score = \frac{2 \times P_r \times R_c}{P_r + R_r} \tag{10}$$

Using Adam optimizer, the proposed model achieved A_c, P_r, Rc, and F1-score of 99.8% ,99.1%, 99.7%, and 99.4%, respectively. In addition to train/validation/test data split, we also employed cross validation to reduce the bias of data selection. Average training accuracy across the five fold cross validation is 99.48±1.16% illustrated in Table 3. This table reveals that our proposed approach achieves maximum efficiency within five folds. In addition to accuracy metrics, the receiver operating characteristics (ROC) also has been used to verify and validate the robustness and accuracy of our system. ROCs depict the model's classification output based on true positives and false-positive rates and the area under the curve (AUC) of a ROC is used as an additional evaluation metric, as demonstrated in Figure 4. To highlight the advantage of the proposed model, we compared its performance with the base model without GAP. Table 4 shows that the proposed model reaches highest A_c, P_r, R_r, and F1-Score. In addition, Table 5 shows the performance of the proposed model against the number of training epochs. As we can see in this table, when the number of epochs is ≥ 50 there is marginal enhancement in overall system performance. Thus, we set the number of epochs for our experiments to 180 epochs.

Furthermore, additional evaluation using different optimization techniques which was described in Section 2.3 to select the best parameter value is conducted. The experimental result accuracy is shown in

Table 6. The performance measures the proposed model for tumor classification depends on the different optimizer. With the Adam optimizer, the suggested model achieves the highest accuracy (i.e., 99.8%) and optimal loss. The proposed approach achieved an optimum error of 0.005 with optimal performance towards the conclusion of the epoch. Figures 5 and 6 represent the confusion matrices and ROC curves for the different optimizers such as Adam, RMSprop, SGD, AdaGrad, and AdaDelta. As the figure shows, all 2394 images in malignant class were correctly classified.

To obtain the best accuracy, our pipeline is trained using different DL pre-trained networks, including AlexNet, Xception, Inception V3, DenseNet121, ResNet50, and VGG16 to access the system's best accuracy. The image identification is done by using sigmoid layers of pre trained networks. We compared the results in Table 7. The best model is ResNet50 that achieved a high accuracy of 99.8% with the same hyper parameters we made our network on. The confusion matrices and ROC are shown in Figures 4 and 7, respectively, for the different pre-trained models such as proposed model, Xception, InceptionV3, VGG16, DenseNet121, and AlexNet. In addition to DL-based approaches, we compared our approach with different ML approaches using the *figshare* dataset.

Table 8 shows such comparison with support vector machine, k-nearest neighbors, convolution neural network, capsule network, gabor filter-neural network and extreme learning machine. As seen in the comparison in Table 8. This table reveals that the proposed approach outperforms the existing models with 99.8% accuracy. ResNet50 with Adam optimizer architecture has achieved the state-of-the-art quality in accurate classification of the MRI images. This in part can be explained by the fact that DL-based methods, especially ResNet50, can handle the challenges by earning discriminative features at a high level. It has proven the capability of handling diversity, so the quality of diagnostic images in MRIs is greater than any other shallower network. Also, the scientific databases used are limited in size, and hard to access. Previous research employed a traditional approach to ML. However, manual feature extraction was required which was time-consuming. Many corresponding approaches used architectures of the deep CNN have used shallower networks. Hence, they had limitations in learning the high-level dataset features and high accuracy. The following are some of the benefits of our proposed model. Since the proposed model is using a deep neural network, automatic feature extraction has been achieved. ResNet50's computational time is reduced thanks to the use of GAP in the output layer. Due to the use of the Adam optimizer, the proposed model achieves faster convergence.

Table 3. Performance analysis of the proposed approach with five fold

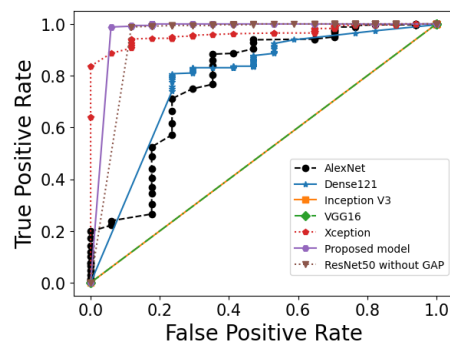| folds	1	2	3	4	5
A_c	97.5%	100%	100%	100%	100%

Figure 4. Receiver operating characteristic curves (ROC) of the proposed model against other DL models such as Xception, InceptionV3, VGG16, DenseNet121, and AlexNet

Table 4. Comparison between proposed model and the baseline ResNet50 model without GAP

Model	Evaluation Metric				
	A_c	F1-score	P_r	R_r	AUC
Proposed model	99.8%	99.4%	99.1%	99.7%	99.8
ResNet50 without GAP	97%	98%	100%	96%	98

Table 5. Over all accuracy (A_c) of the proposed model against the ResNet50 without GAP at different epochs

	Model	
# of epochs	Proposed model	ResNet50 without GAP
50	52.3%	89.3%
100	75.2%	95%
150	90.8%	95.6%
180	99.8%	97%

Table 6. Proposed model evaluation with different type of optimizers Adam, RMSprop, SGD, AdaGrad, and AdaDelta. Here "AUC" stand for area under the curve

	Evaluation metrics				
Optimizer	A_c	F1-score	P_r	R_r	AUC
Adam	99.8%	99.4%	99.1%	99.7%	99.8
RMSprop	99%	99.2	% 98.5	% 99	96.6
SGD	94%	96.9%	95.2%	98.8%	73.1
AdaGrad	51%	69.6%	96.3%	54.5%	61.6
AdaDelta	30%	36.2%	95%	22.4%	61.7
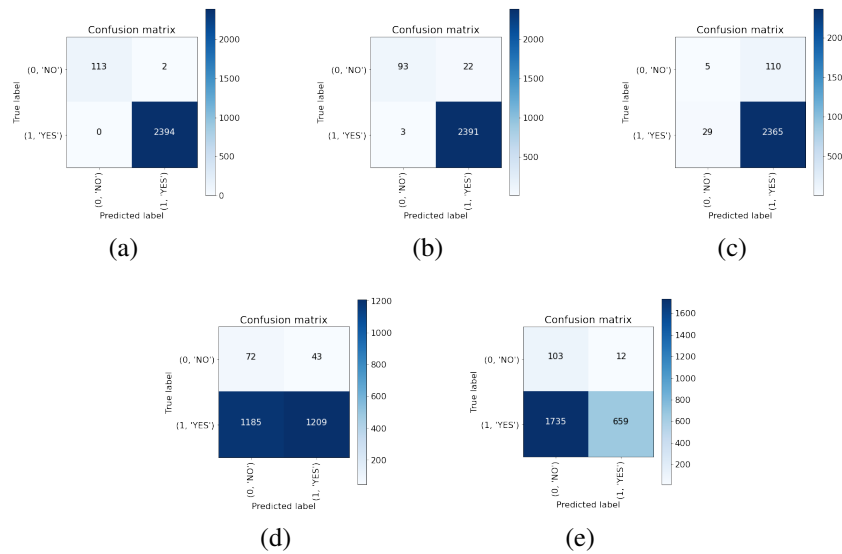
Figure 5. Confusion metrics for classification for different optimizers: (a) Adam, (b) RMSprop, (c) SGD, (d) AdaGrad, and (e) AdaDelta
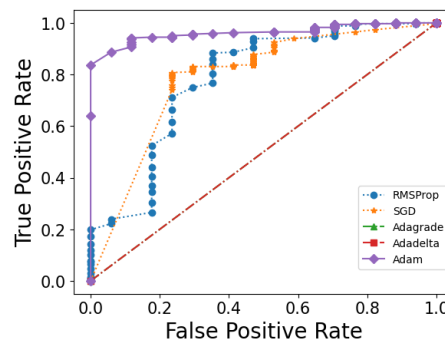
Figure 6. Receiver operating characteristic curves (ROC) for different optimizers such as Adam, RMSprop, SGD, AdaGrad, and AdaDelta

Table 7. Proposed model evaluation against other deep learning models

Model	Evaluation metrics				
	A_c	F1-score	P_r	R_r	AUC
Proposed	99.8%	99.4%	99.1%	99.7%	99.8
Xception	97%	97.7%	96.3%	99.1%	96.2
Inception V3	95%	98 %	100%	99%	50
VGG16	95%	98%	100%	95%	52.8
DenseNet121	84%	90%	84%	97%	80.7
AlexNet	74%	83.7%	94.8%	74.9%	49.8
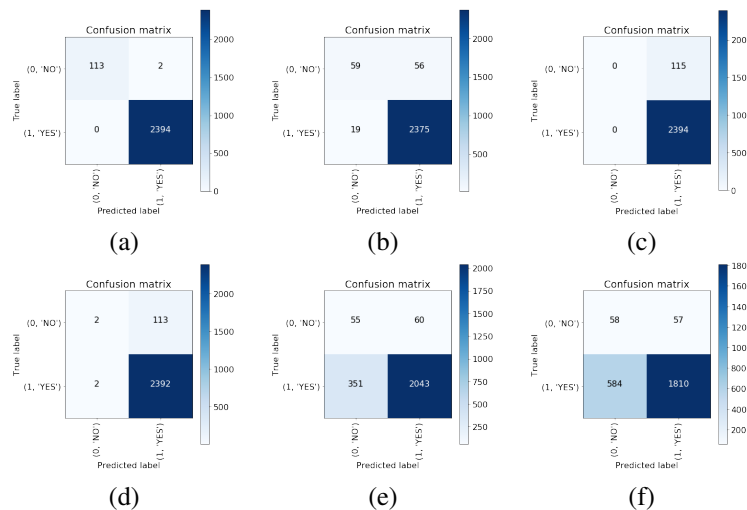
Figure 7. Confusion metrics for (a) proposed model, (b) Xception, (c) InceptionV3, (d) VGG16, (e) DenseNet121, and (f) AlexNet

Table 8. Previous works for brain tumor classification using the public *figshare* dataset. Note that: SVM, KNN, CNN, KELM, and NA stand for support vector machine, k-nearest neighbors, convolution neural network, extreme learning machine, and not applicable, respectively

Method	Model	A_c	F1-score	P_r	R_r
Cheng *et al.*[13]	SVM, KNN	91.28%	NA	NA	NA
J.Paul *et al.*[24]	CNN	91.43 %	NA	NA	NA
Abiwinanda *et al.* [16]	CNN	84.19%	NA	NA	NA
Pashaei *et al.* [25]	KELM + CNN	93.68%	93.00%	94.60%	91.43%
Afshar *et al.* [26]	GoogleNet + CNN	90.89%	NA	NA	NA
Proposed model	ResNet50	99.8%	99.4%	99.1%	99.7%

4. CONCLUSIONS

In this work, we have proposed a transfer learning-based DLS for brain tumor classification from T1-w MRI images using locally and publicly available datasets. The proposed analysis is extensively evaluated on different pre-trained models and different optimization methods. The experiments documented that the ResNet50 with Adam optimizer and GAP is yielded the superior accuracy of 99.8%.

REFERENCES

[1] S. Lapointe, A. Perry, and N. A. Butowski, "Primary brain tumours in adults," *The Lancet*, vol. 392, no. 10145, pp. 432–446, 2018, doi: 10.1016/S0140-6736(18)30990-5.

[2] E. Rushing, "Who classification of tumors of the nervous system: preview of the upcoming 5th edition," *memo - Magazine of European Medical Oncology*, vol. 14, pp. 188–191, 2021, doi: 10.1007/s12254-021-00680-x.

[3] A. Behin, K. Hoang-Xuan, A. F. Carpentier, and J.-Y. Delattre, "Primary brain tumours in adults," *The Lancet*, vol. 361, no. 9354, pp. 323–331, 2003, doi: 10.1016/S0140-6736(03)12328-8 .

[4] R. N. Al-Okaili, J. Krejza, S. Wang, J. H. Woo, and E. R. Melhem, "Advanced MR imaging techniques in the diagnosis of intraaxial brain tumors in adults," *Radiographics*, vol. 26, no. suppl_1, pp. S173–S189, 2006, doi: 10.1148/rg.26si065513.

[5] C. S. Rao and K. Karunakara, "Efficient detection and classification of brain tumor using kernel based SVM for MRI," *Multimedia Tools and Applications*, vol. 81, no. 5, pp. 7393–7417, 2022, doi: 10.1007/s11042-021-11821-z.

[6] G. S. Tandel *et al.*, "A review on a deep learning perspective in brain cancer classification," *Cancers*, vol. 11, no. 1, p. 111, 2019, doi: 10.3390/cancers11010111.

[7] M. A. Myszczynska *et al.*, "Applications of machine learning to diagnosis and treatment of neurodegenerative diseases," *Nature Reviews Neurology*, vol. 16, no. 8, pp. 440–456, 2020, doi: 10.1038/s41582-020-0377-8.

[8] D. J. Hemanth, J. Anitha, A. Naaji, O. Geman, D. E. Popescu, and et al., "A modified deep convolutional neural network for abnormal brain image classification," *IEEE Access*, vol. 7, pp. 4275–4283, 2018, doi: 10.1109/ACCESS.2018.2885639.

[9] Y. Liu *et al.*, "A deep convolutional neural network-based automatic delineation strategy for multiple brain metastases stereotactic radiosurgery," *PloS one*, vol. 12, no. 10, p. e0185844, 2017, doi: 10.1371/journal.pone.0185844 .

[10] M. Kaur and D. Singh, "Fusion of medical images using deep belief networks," *Cluster Computing*, vol. 23, no. 2, pp. 1439–1453, 2020, doi: 10.1007/s10586-019-02999-x.

[11] G. Litjens *et al.*, "A survey on deep learning in medical image analysis," *Medical image analysis*, vol. 42, pp. 60–88, 2017, doi: 10.1016/j.media.2017.07.005.

[12] J. Cheng, "Brain tumor dataset," figshare dataset, 2017. [Online] Available: https://figshare.com/articles/dataset/brain_tumor_dataset/1512427.

[13] J. Cheng *et al.*, "Enhanced performance of brain tumor classification via tumor region augmentation and partition," *PloS one*, vol. 10, no. 10, p. e0140381, 2015, doi: 10.1371/journal.pone.0140381.

[14] M. R. Ismael and I. Abdel-Qader, "Brain tumor classification via statistical features and back-propagation neural network," in *2018 IEEE international conference on electro/information technology (EIT)*, 2018, pp. 0252–0257, doi: 10.1109/EIT.2018.8500308.

[15] P. Afshar, A. Mohammadi, and K. N. Plataniotis, "Brain tumor type classification via capsule networks," in *2018 25th IEEE International Conference on Image Processing (ICIP)*, 2018, pp. 3129–3133, doi: 10.1109/ICIP.2018.8451379.

[16] N. Abiwinanda, M. Hanif, S. T. Hesaputra, A. Handayani, and T. R. Mengko, "Brain tumor classification using convolutional neural network," in *World congress on medical physics and biomedical engineering 2018*, 2019, pp. 183–189, doi: 10.1007/978-981-10-9035-6_33.

[17] A. K. Anaraki, M. Ayati, and F. Kazemi, "Magnetic resonance imaging-based brain tumor grades classification and grading via convolutional neural networks and genetic algorithms," *biocybernetics and biomedical engineering*, vol. 39, no. 1, pp. 63–74, 2019, doi: 10.1016/j.bbe.2018.10.004.

[18] S. J. Pan and Q. Yang, "A survey on transfer learning," *IEEE Transactions on knowledge and data engineering*, vol. 22, no. 10, pp. 1345–1359, 2010, doi: 10.1109/TKDE.2009.191.

[19] N. Chakrabarty, "Brain MRI images for brain tumor detection," 2019. [Online]. Available: https://www.kaggle.com/navoneel/brain-mri-images-for-brain-tumor-detection.

[20] M. Sajjad, S. Khan, K. Muhammad, W. Wu, A. Ullah, and S. W. Baik, "Multi-grade brain tumor classification using deep cnn with extensive data augmentation," *Journal of computational science*, vol. 30, pp. 174–182, 2019, doi: 10.1016/j.jocs.2018.12.003.

[21] P. Chlap *et al.*, "A review of medical image data augmentation techniques for deep learning applications," *Journal of Medical Imaging and Radiation Oncology*, vol. 65, no. 5, pp. 545–563, 2021, doi: 10.1111/1754-9485.13261 .

[22] S. Wang, Y. Jiang, X. Hou, H. Cheng, and S. Du, "Cerebral micro-bleed detection based on the convolution neural network with rank based average pooling," *IEEE Access*, vol. 5, pp. 16576–16583, 2017, doi: 10.1109/ACCESS.2017.2736558.

[23] T. Tazin *et al.*, "A robust and novel approach for brain tumor classification using convolutional neural network," *Computational Intelligence and Neuroscience*, vol. 2021, p. 2392395, 2021, doi: 10.1155/2021/2392395.

[24] J. S. Paul, A. J. Plassard, B. A. Landman, and D. Fabbri, "Deep learning for brain tumor classification," in *Medical Imaging 2017: Biomedical Applications in Molecular, Structural, and Functional Imaging*, 2017, vol. 10137, p. 1013710, doi: 10.1117/12.2254195.

[25] A. Pashaei, H. Sajedi, and N. Jazayeri, "Brain tumor classification via convolutional neural network and extreme learning machines," *2018 8th International Conference on Computer and Knowledge Engineering (ICCKE)*, 2018, pp. 314-319, doi: 10.1109/ICCKE.2018.8566571.

[26] P. Afshar, K. N. Plataniotis, and A. Mohammadi, "Capsule networks for brain tumor classification based on mri images and coarse tumor boundaries," in *ICASSP 2019-2019 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, 2019, pp. 1368-1372, doi: 10.1109/ICASSP.2019.8683759.

BIOGRAPHIES OF AUTHORS

Mona Ahmed 🆔 📖 sc ℗ was born in Mansoura, Egypt, in 1992. She received he B.Sc. degree in Biomedical Engineering from Helwan University, Helwan, Egypt, in 2016. Currently, she is a Clinical Engineer and master student at Mansoura University, Mansoura, Egypt pursuing her M.Sc.degree. Her research interests are machine learning applications for biomedical data analysis, medical image analysis with main focus on medical diagnostics. She can be contacted at email: onagaffer@std.mans.edu.eg.

Fahmi Khalifa, Ph.D. 🆔 📊 SC P received his B.Sc. and M.Sc. degrees in electronics and electrical communication engineering from Mansoura University (MU), Egypt, in 2003 and 2007. He received his Ph.D. degree from the Electrical and Computer Engineering (ECE) Department at the University of Louisville (UofL) in 2014. Dr. Khalifa has more than 12 years of hands-on experience in conducting interdisciplinary research in the fields of signal and image processing, machine and deep learning, medical data analysis, computer-aided diagnosis, and digital and analog signal processing. He has more than 150 peer-reviewed publications appearing in prestigious journals, top-rank conferences, leading edited books, and U.S. patents. Dr. Khalifa is a senior IEEE Member and an Associate Editor for the IEEE Access and IEEE Journal of Biomedical and Health Informatics (J-BHI). He has received various awards and recognition including the Distinctive Student Award for five consecutive years (MU), Theobald Scholarship award (UofL); the ECE-UofL Outstanding Student award for two years in a row; the John M. Houchens award for the outstanding PhD dissertation (UofL); the second-place Research! Louisville Post-doctoral Fellow; and the EPIC innovation award in 2020 and 2021, UofL. He can be contacted at email: fahmikhalifa@yahoo.com.

Hossam El-Din Moustafa 🆔 📊 SC P Professor at the Department of Electronics and Communications Engineering, the founder and former executive manager of Biomedical Engineering Program (BME) at the Faculty of Engineering, Mansoura University. He is an IEEE senior member. Research interests include biomedical imaging, image processing applications, and bioinformatics. He can be contacted at email: hossam-moustafa@hotmail.com.

Gehad Ahmad Saleh, M.D. 🆔 📊 SC P is a lecturer of radiology & intervention radiology, Faculty of medicine, Mansoura University,Egypt. Dr. Saleh received the B.Sc.degree and M.Sc.degree from Mansoura University, Cairo, Egypt, in 2010 and 2015, respectively. She also received here Ph.D. degree in 2020. Dr. Saleh main research interests include machine learning application for medical diagnostics with application in liver, oncology and female pelvic, head and neck radiology using various imaging techniques. She can be contacted at email: gahmad1988@gmail.com.

Eman AbdElhalim, Ph.D. 🆔 📊 SC P is an Assistant Professor, Electronics & Communications department, Faculty of engineering. Her research interests cover several aspects of communications, cloud computing, big data analytics, and medical imaging using deep learning techniques. Dr. Eman is a reviewer of more than 10 peer-reviewed publications appearing in reputable journals. Eman teaches several courses on artificial intelligence, signal processing, and programming. Eman also organized workshops on applications of artificial intelligence in the industry and production environment. She is training several sessions in digital transformation. She has been working since 2008 on graduation projects in the field of applications of deep learning in communications and medical imaging fields. She can be contacted at email: eman-haleim@mans.edu.eg.