KSÜ Mühendislik Bilimleri Dergisi, 26(Özel Sayı), 2023 KSU J Eng Sci, 26(Special Issue), 2023
9th Uluslararası IFS Çağdaş Matematik ve Mühendislik Konferansı Özel Sayısı
9th International IFS Contemporary Mathematics and Engineering Conference Special Issue
Araştırma Makalesi Research Article

Kahramanmaras Sutcu Imam University
Journal of Engineering Sciences

Geliş Tarihi : 09.08.2023 Received Date : 09.08.2023
Kabul Tarihi : 25.09.2023 Accepted Date : 25.09.2023

CLASSIFICATION OF BRAIN TUMORS ON MRI IMAGES USING DEEP LEARNING ARCHITECTURES

DERİN ÖĞRENME MİMARİLERİ KULLANILARAK MRI GÖRÜNTÜLERİ ÜZERİNDE BEYİN TÜMÖRÜ SINIFLANDIRMASI

Samaneh SARFARAZI[1] (ORCID: 0009-0007-5139-5662)
Önsen TOYGAR[2*] (ORCID: 0000-0001-7402-9058)

Computer Engineering Department, Faculty of Engineering, Eastern Mediterranean University, 99628, Famagusta, North Cyprus, via Mersin 10, Turkey

*Sorumlu Yazar / Corresponding Author: Önsen TOYGAR, onsen.toygar@emu.edu.tr

ABSTRACT

A brain tumor is a dangerous neural illness produced by the strict growth of prison cells in the brain or head. The segmentation, analysis, and separation of unclean tumor parts from Magnetic Resonance Imaging (MRI) images are the main sources of anxiety. To report the segmented MRI images including tumor, the usage of computer-assisted methods is necessary. In this paper, a Convolutional Neural Network (CNN) approach is applied to identify brain cancers in MRI images. Two datasets are used in this study, namely Kaggle Brain MRI database and Figshare Brain MRI database. Models of deep CNN, consisting of VGG16, AlexNet, and ResNet, are utilized to extract deep features. The classification accuracies of the aforementioned Deep Learning (DL) networks are used to measure the efficiencies of the implemented systems. For the Kaggle database, AlexNet achieves 98% accuracy, VGG16 has 97% accuracy and ResNet has 66% accuracy. Among these networks, AlexNet has provided the highest level of accuracy. In the Figshare dataset, AlexNet and VGG16 both achieve 99% accuracy, and ResNet has 96% accuracy. In terms of accuracy, AlexNet and VGG16 outperform ResNet. These performances aid in the early detection of cancers before they cause physical harm such as paralysis and other complications.

Keywords: Brain tumor classification, convolutional neural networks, deep learning

ÖZET

Beyin tümörü, beyindeki veya kafadaki hapishane hücresinin katı bir şekilde büyümesiyle üretilen tehlikeli bir nöral hastalıktır. Manyetik Rezonans Görüntüleme (MRI) görüntülerinden temiz olmayan tümör parçalarının segmentasyonu, analizi ve ayrılması kaygının ana kaynağıdır. Tümör içeren MRI görüntülerinin raporlanabilmesi için bilgisayar destekli yöntemlerin kullanılması gerekli hale gelmiştir. Bu makalede, MRI görüntülerinde beyin tümörlerini tanımlamak için Evrişimli Sinir Ağları (CNN) yaklaşımı kullanılmıştır. Bu çalışma için Kaggle Brain MRI veri kümesi ve Figshare Brain MRI veri kümesi olmak üzere iki veri kümesi kullanılmıştır. Derin öznitelikleri çıkarmak için VGG16, AlexNet ve ResNet'ten oluşan derin CNN modelleri kullanılmıştır. Söz konusu Derin Öğrenme (DL) modellerinin sınıflandırma doğrulukları, uygulanan sistemlerin verimliliklerini ölçmek için kullanılmıştır. Kaggle veri kümesi için AlexNet %98, VGG16 %97 ve ResNet %66 doğruluk elde etmiştir. Bu ağlar arasında AlexNet en yüksek düzeyde doğruluk sağlamıştır. Figshare veri kümesinde ise, AlexNet ve VGG16'nın her ikisi de %99, ResNet ise %96 doğruluk elde etmiştir. Doğruluk açısından AlexNet ve VGG16, ResNet'ten daha iyi performans göstermiştir. Bu performanslar, kanserlerin felç ve diğer komplikasyonlar gibi fiziksel zararlara yol açmadan önce erken teşhis edilmesine yardımcı olacaktır.

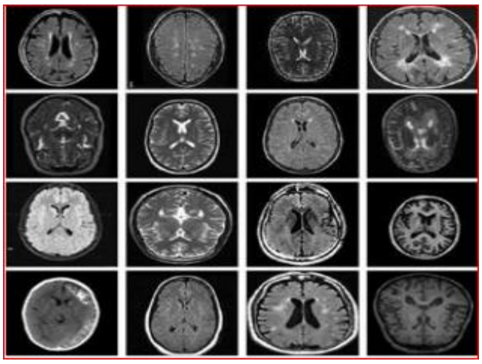
Anahtar Kelimeler: Beyin tümörü sınıflandırması, evrişimli sinir ağları, derin öğrenme

ToCite: SARFARAZI, S., & TOYGAR, Ö., (2023). CLASSIFICATION OF BRAIN TUMORS ON MRI IMAGES USING DEEP LEARNING ARCHITECTURES. *Kahramanmaraş Sütçü İmam University, Journal of Engineering Sciences*, 26(Özel Sayı), 1177-1186.

INTRODUCTION

A brain tumor is an imbalanced type of cell in the human brain. The brain of a human is surrounded by a firm head. Slight development in such a minor part will cause intense problems. Tumors of brain may be malignant and nonmalignant. The gravity inside the head will increase such as nonthreatening or malicious cancers progress. This will result in enduring head damage or death of the person. Experts and investigators have been studying complex methods and approaches aimed at diagnosing tumors of brain. While MRI image depictions and Tomography of Computer (CT) are the two approaches with broad usage in which both aimed at clarifying the anomalies in form, mass, or brain materials place, that help doctors in identifying cancers; MRI image is preferred more than the aforementioned methods by the specialists. Therefore, experts and scientists have used MRI images. However, automatic methods, mostly applied by computer-assisted medicinal image processing methods, exist progressively helping surgeons to notice tumors of brain. Handcrafted methods with Machine Learning (ML) classifiers are developed for training data examples.

Medicinal image processing includes pre-processing and post-processing. These stages may be applied through the Handcrafted methods as perfect as the method of deep learning. In handcrafted methods, features are extracted to get consequences from images of the test and the procedure is quick. In the DL methods, networks are adjusted through properly choosing the sum of layers, activation function, and pooling. But, in both methods, new algorithms are possible to be employed to increase the system's accuracy in a wider viewpoint. DL techniques for identifying brain cancers in MRI scans are the topic of this paper. The main part of this paper focuses on finding tumors of brain through MRI images using methods of DL. Consequently, this paper will offer the anticipated result i.e., an effective DL method to distinguish tumors of brain by MRI images that will contribute to medical experts to run appropriate cures. Sample MRI images on both Kaggle (Kaggle Brain Tumor Dataset, 2020) and Figshare (Figshare Brain Tumor Dataset, 2018) datasets for healthy and unhealthy brains are demonstrated in Figure 1 and Figure 2, respectively.


(a)

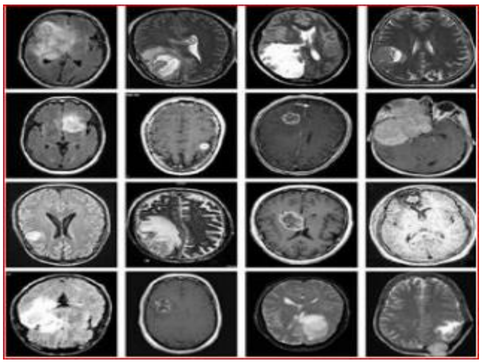
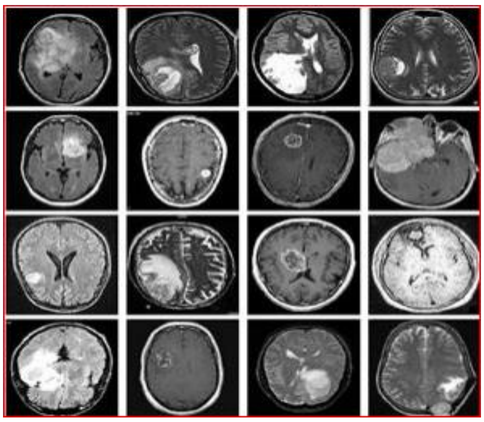

(b)

Figure 1. a. Samples of Kaggle Database Healthy Brain MRI and **b.** Unhealthy Brain MRI Images


(a)

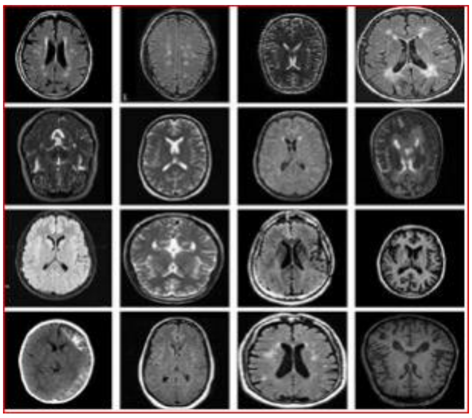

(b)

Figure2. a. Samples of Figshare Database Healthy Brain MRI and **b.** Unhealthy Brain MRI Images

The following is the structure of the paper. Background on brain tumor classification and the works done by other researchers using handcrafted and DL methods are presented in Section 2. In Section 3, the paper methodology and metrics for evaluations are presented. Section 4 presents the experimental results, discussion and comparison with the state-of-the-art. Lastly, Section 5 gives the conclusions of this paper and future directions.

RELATED WORK

This section in brief deliberates the research that is shown to distinguish tumors of brain using the dissimilar advanced know-how. Rehman et al. (2021) recommended a different knowledge-founded technique for mini tumor of brain finding and type of tumor classification. The initial stage of their study focused on using a 3D CNN to abstract tumors in the brain, which were then transferred to a CNN model having already received training to abstract model features. The features that have been extracted are carried out to a correlation-based election procedure, and the greatest characteristics are selected in place of the outcome. In the last classification, with the usage of Feed Forward Neural Network (FFNN), the elected features are tested.

Amin et al. (2018) used a Deep Neural Network (DNN) based design for brain tumor segmentation. The suggested model has seven layers in the classification stage, including three ReLU layers, three convolutional layers, and a SoftMax layer. On the other hand, Kebir & Mekaoui (2018) suggested an approach that is supervised for identifying the anomalies of brain through the MRI images in several phases. The initial stage is to change a DL based CNN model, and after that, a subclass of MRI brain images is completed via the k-mean process conformed by the factor of brain grouping as standard or nonstandard groups in accordance with the advanced CNN model.

Alternatively, Vinoth & Venkatesh (2018) introduced a CNN-based automatic separation technique. At this point, classification was done with kernels, and Support Vector Machine (SVM) classification was done with computed variables. Furthermore, MATLAB is used to extract and recognize malignancies from MRI images of the brain. A CNN founded as a model of DL was effectively connected to the regarded issue of tumor of brain classification. Classifiers based on CNN constructions have the advantage of not requiring bodily sectioned tumor zones.

Talo et al. (2019) classified normal and pathological Brain MRI images with 100% accuracy by means of the ResNet34 pre-trained CNN model through a data augmentation technique of transfer learning. On the other hand, a model of the pre-trained ResNet50 CNN was updated (Çınar & Yıldırım, 2020) by eliminating the preceding five levels and replacing them with eight new layers, matching the accuracy of prior pre-trained patterns as ResNet50, AlexNet, and GoogleNet. The reconstructed ResNet50 pattern achieved 97.2% accuracy indicating real repercussions. There are several ML approaches for brain tumor classification and segmentation using MRI in scientific literature.

Hasan et al. (2019) suggested an image of an MRI brain scan categorization system based on deep and custom features. Preprocessed MRI image is useful for an altered Gray Level Co-occurrence Matrix (GLCM) for extraction of statistical features. CNN extracts features automatically. SVM classification with 10-fold Cross-validation performed 99.30% based on 600 sagittal MRI scans. While likened to new networks of transfer learning, such as GoogleNet and AlexNet, the recommended method performed fine on the other hand by means of combining CNN and GLCM features.

In another study, a Naive Bayes based brain cancer identification approach employed a maximum entropy segmentation (Zaw et al., 2019). The REMBRANDT dataset, which includes 114 MRI images, is used to examine the system. The recommended system has the advantage of detecting tumors anywhere in the brain, such as the temporal lobe. On the other hand, Sert et al. (2019) proposed a different scheme aimed at detecting brain tumors using a combination of CNN and advanced segmentation techniques utilizing maximum fuzzy entropy to improve the resolution of MRI, the super resolution of a single image is employed. Pre-trained ResNet architecture is used to extract features. SVM with binary classification has a 95% accuracy rate.

Edge Adaptive Total Variation (EADTV) (Deepak & Ameer, 2019) applied the mean shift clustering approach in regard to brain tumor categorization segmentation. The proposed technique has 2 advantages: When utilizing the image, EADTV preserves the edges with mean shift clustering, unlike K-mean and fuzzy c-means, and automatically updates cluster centers. In a combined approach of Particle Swarm Optimization (PSO) with fusion features for tumor of brain diagnosis, a fine-tuned Capsule Network feature extraction and Local Binary Patterns are applied. SVM classification accuracy on the BRATS2018 and RIDER databases is 98.3 % and 97.9%, respectively. The new proposal has shown good results by combining handcrafted and deep features. On the CEMRI dataset, SVM and kNN classifiers are used to assess pre-trained GoogleNet for deep feature extraction for 3 class classification into Glioma, Meningioma with accuracy of 97.8% and 98%, respectively. The BRATS 2017 dataset, which contains 48 images, is used to assess the accuracy of a multinomial logistic regression model for brain tumor categorization. The system's functioning, however, should be evaluated on bigger databases.

Narayana & Reddy (2018) proposed a system that works well with the 9 Genetic Algorithm (GA) based SVM classification method. In an effective brain tumor classification optimization technique, GA is employed in order to segment tumors. SVM and GLCM texture characteristics provided 91.23 % performance for high-grade glioma (HGG) and low-grade gliomas (LGG) brain tumor categorization.

Polly et al. (2018) developed a k-means segmentation algorithm. From wavelet features, Principal Component Analysis (PCA) is employed to determine ten relevant features. On the way to discriminate between images that are normal and abnormal, SVM algorithm is utilized. Once again, SVM classification method is employed on the way to classify LGG and HGG tumors in aberrant images. On 440 images, the suggested technique achieves 99 %, but it needs to be evaluated on a larger database using added important data.

Amin et al. (2018) suggested a new technique for identifying brain tumors using MRI. To reduce noise and smooth MRI, skull stripping and Gaussian filtering were used. Following K-means segmentation, GLCM texture characteristics were extracted. The system was tested on three datasets: local, AANLIB, and RIDER. They used linear, RBF, and cubic SVM kernels. The linear kernel with 5-fold cross validation was found to have 98% accuracy. On the other hand, Minz & Mahobiya (2017) provided a study that uses the AdaBoost classifier to classify brain tumors. Following Median filtering, threshold-based segmentation is used to reduce noise. Using GLCM characteristics, the system proposed texture-based classification.

Shankar et al. (2016) proposed exploiting texture features to classify brain tumors using Gustafson-Kessel(G-K) fuzzy clustering. A histogram-based approach is used to segment preprocessed images with the Wiener filter. G-K fuzzy manner was given GLCM texture features for binary classification with 95% accuracy. The detection of brain tumors using systems of DL was a cutting-edge subject of study. Researchers use a variety of DL architectures to automatically segment and classify brain tumors. For brain classification, Regularized Extreme Learning Technique with Mix Features was also recommended via Gumaei et al. (2019).

A feature extractor of Hybrid PCA-NGIST can be utilized for 3-D feature extraction. The NGIST feature descriptor is a descriptor of standardized feature that is utilized to address image illumination and shadowing issues. RELM is a single hidden layer, input, and output FFNN. The suggested technique is examined for three kinds of tumors: neuroendocrine tumor, glioma, and pituitary tumor using CE-MRI database with 94.33 % after 5 fold cross-validation. Link Net is a small DNN design that is employed to group brain cancers (Hemanth et al., 2019). On a freely released UCI repository database, binary classification achieved 91% accuracy. The Multi-Layer Perceptron (MLP) classification system has a 96 % accuracy rate and a 0.65 Kappa Statistic. However, sparse auto encoder could be examined in forthcoming years when DNN is integrated with other auto encoder versions like denoising auto encoder.

Latif et al. (2018) proposed a brain tumor classification technique derived from transfer learning. To suit the VGG19 network with MRI images that are scaled to 224×224 pixels. To update the weights, fine tweaking of parameters such as learning rate, scheduling rate, and momentum is done block by block. The system has 94.82% accuracy on the CE-MRI database. The disadvantage of that approach is that fine-tuning settings block by block takes 20-30 minutes to train CNN classifier. MLP uses statistical and wavelet features to classify brain tumors (Mohsen et al., 2018).

METHODOLOGY FOR BRAIN TUMOR CLASSIFICATION

This section reviews three CNN based DL architectures implemented for brain tumor classification in this study. The methodology used in this paper for the implementation of brain tumor classification employs AlexNet, VGG16, and ResNet architectures. The main contribution of this paper is to present the implementation of deep learning based AlexNet, VGG16, and ResNet architectures for brain tumor classification separately. The evaluation metrics for the presentation of the experimental results are then discussed in the following section. CNN is the most well-known and extensively utilized approach in the field of DL. CNN's key benefit over its predecessors is that it accurately characterizes relevant characteristics with almost no human intervention. Face recognition, computer vision, audio processing, and other applications have all benefited from CNN. The development of CNN was invigorated by neurons in humans. A typical type of CNN has many convolutions pooling layers, similar to an MLP, except the end layers are Fully Connected layers. In this paper, three of the aforementioned architectures, namely AlexNet (Huafeng et al., 2015), VGG16 (Simonyan & Zisserman, 2014), and ResNet (Wang & Gong, 2020) are implemented for brain tumor classification.

A general block diagram of the implemented system with a general CNN architecture is shown in Figure 3. The methodology used in this study for the implementation of brain tumor classification starts with the brain MRI images that are used as the input of the system, and then the system employs AlexNet, VGG16 and ResNet architectures separately to classify the input images with or without tumors. Feature extraction and classification are performed by the DL architecture automatically and the output is produced at the end of this process. The output of the system for images with tumor is “YES” and for healthy brain images, the output produced is “NO” as shown in Figure 3.

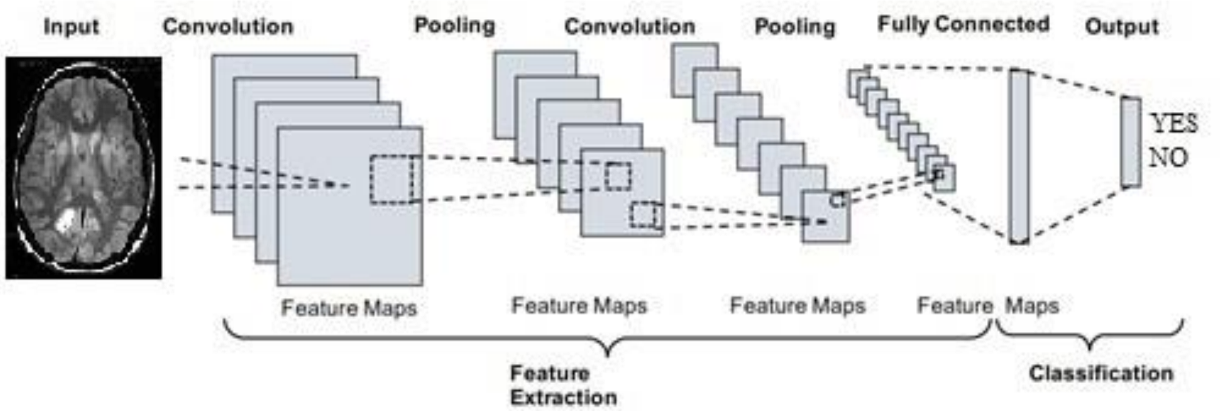


Figure 3. Block Diagram of the Implemented System

Dataset Descriptions

In this study, two datasets, namely Kaggle (Kaggle Brain Tumor Dataset, 2020) and Figshare (Figshare Brain Tumor Dataset, 2018) are employed for the experiments. In the train folder of Kaggle dataset, there are 1200 images labeled as “yes” and 1200 images labeled as “no”, which makes a total of 2400 images in the train folder. In the test folder, there are 300 images labeled as “yes” and 300 images named as “no” which makes a total of 600 images in the test folder. Therefore, Kaggle image dataset in this dataset includes 3000 images totally.

On the other hand, the second dataset, named Figshare, includes two files entitled “test” and “train”. In the train folder, there are 4117 images considered as “yes” and 1588 images termed as “no” and there is a total of 5705 images in the train folder. In the test folder, there are 906 images considered as “yes” and 405 images called “no” and there is an overall of 1311 images in the test folder. However, because of the quality of some images in Figshare dataset, some of the images are not used in the experiments. Consequently, in the experiments, 5600 images are selected for training and 1400 images are selected for testing which makes use of total 7000 images in our Figshare dataset.

Performance Metrics

This section discusses metrics for evaluation, which are accustomed to assess the value of a model statistically. A number of evaluation measures can be employed to show the value of a model such as accuracy, precision, and recall. The evaluation metrics used in DL tasks are critical in determining the best classifier. They are employed in the testing and training stages of a typical data classification process. During the phase of training, it is employed to improve the algorithm of classification. This indicates that the assessment measure is utilized to distinguish between options and select the best one, such as a discriminator, which can yield a more precise estimate of future evaluations when used in conjunction with an exact classifier. In the meantime, the assessment metric is accustomed to analyze the developed classifier's effectiveness, such as a hidden data evaluator during the model test phase. The number of effectively classified negative and positive instances is denoted by TN and TP, respectively. Furthermore, the amounts of misclassified positive and negative cases are defined as FN and FP, respectively. The following are around of the supreme famous evaluation metrics.

1) Accuracy: Computes the percentage of correctly forecast classes in relation to the overall number of samples that were tested. Accuracy can be calculated as follows:

$$\text{Accuracy} = \frac{\text{TP+TN}}{\text{TP+TN+FP+FN}} \tag{1}$$

where TP is the number of true positives that are perfectly recognized. The total of perfectly-identified true negatives is recognized as TN. FP denotes the number of images that are wrongly recognized in place of positive but are really negative. FN is the sum of falsely detected negatives that are truly positive.

Accuracy values are in the range [0,100] percent. If we divide that range evenly, 100-87.5% equals very good, 87.5-75% equals good, 75-62.5% equals satisfactory, and 62.5-50% equals bad. In reality, we regard numbers between 100 and 95% to be excellent, 95 to 85% to be good, 85 to 70% to be satisfactory, and 70 to 50% to be “needs to be improved” for brain tumor recognition.

2) Recall or Sensitivity: The percentage of successfully classified positive patterns is calculated using sensitivity or recall as shown below:

$$\text{Sensitivity} = \frac{\text{TP}}{\text{TP+FN}} \tag{2}$$

where TP is the total number of true positives that are perfectly recognized. FN is the total of falsely detected negatives that are truly positives. The recall is calculated as TP/FN, in which TP represents true positives, and FN represents false negatives. The recall of a classifier refers to its ability to locate all samples that are positive. 1 is the best value while 0 is the worst.

3) Specificity: is used to calculate the percentage of incorrectly classified negative patterns. The formula of specificity is as follows:

$$\text{Specificity} = \frac{\text{TN}}{\text{FP+TN}} \tag{3}$$

where the quantity of perfectly-identified negative images is equal to TN. FP denotes the quantity of positive images that are wrongly known as positive but are truly negative.

In an ideal world, the model would have a high specificity or true negative rate. A greater specificity score would imply a higher real negative rate and a decreased rate of false-positives. A reduced specificity score indicates a lower genuineness score.

4) Precision: is used to figure out which positive patterns in a positive class are the most common. Precision is calculated as follows:

$$\text{Precision} = \frac{\text{TP}}{\text{TP+FP}} \tag{4}$$

where TP is the total of fully recognized positive images. The sum of positive images that are wrongly recognized as positive but are essentially negative is referred to as FP. Precision can be used as a measure of quality. When an algorithm's precision is higher, it produces more relevant outcomes rather than irrelevant ones.

5) F1-Score: also known as F-score and F-measure, is a model's accuracy on a dataset. It is used to assess binary classification systems that categorize examples as positive or negative. F1-Score is calculated as follows:

$$\text{F1 Score} = 2 * \frac{\text{Precision}*\text{Recall}}{\text{Precision} + \text{Recall}} \tag{5}$$

where recall (also called sensitivity) is the percent of related examples identified, and precision for positive predictive value is the proportion of applicable examples found among the improved instances. The greatest rate of an F-score is 1.0, which implies faultless accuracy and recall, while the minimum value is 0 if neither precision nor recall is 0.

RESULTS AND DISCUSSION

This section presents the experimental results obtained using Kaggle and Figshare datasets for brain tumor classification. Afterward, the discussion about the results is given and the comparison with the state-of-the-art demonstrates an analysis of different studies and findings on brain tumor classification and the comparison of these findings with the results obtained in this study.

Results

All evaluation metrics and methods applied for both datasets are described in Table 1 and Table 2. All values and results obtained in these experiments are demonstrated in details. The pathological brain images were taken from the Kaggle dataset and there are two folders called test and train to analyze the performance of each prediction model.

When Figshare dataset is used, the results for three different models (AlexNet, VGG16 and ResNet) are better than when Kaggle dataset is used. The reason for this is that the total images in Figshare dataset is more than the amount of images in Kaggle dataset and also the quality of images in Figshare dataset is better than the quality of images in Kaggle dataset finally, these models are more compatible with Figshare dataset compared to Kaggle dataset.

Table 1. Brain Tumor Classification Results on Kaggle Dataset

Evaluation metric	**Method Used**		
	AlexNet	VGG16	ResNet
Accuracy	0.9883	0.9733	0.6667
Precision	0.9895	0.9767	0.6185
Recall(Sensitivity)	0.9861	0.9703	0.9258
Specificity	0.9904	0.9764	0.3897
F-Measure	0.9878	0.9735	0.7416

Table 2. Brain Tumor Classification Results on Figshare Dataset

Evaluation metric	**Method Used**		
	AlexNet	VGG16	ResNet
Accuracy	0.9943	0.9915	0.9658
Precision	0.9960	0.9928	0.9818
Recall(Sensitivity)	0.9960	0.9948	0.9700
Specificity	0.9900	0.9840	0.9553
F-Measure	0.9960	0.9938	0.9759

Comparison with the State-of-the-Art

Expert radiologists perform the crucial task of brain tumor segmentation and classification. As decision-making aids, radiologists can use ML and DL methods. This paper outlines a number of cutting-edge methodologies for classifying brain tumors automatically. Brain tumor classification results are compared on Kaggle Brain MRI and Figshare Brain MRI datasets in Table 3. In Kaggle dataset, a total of 2400 images exist in the train folder. There is a total of 600 images in the test folder. All the train and test images in Kaggle dataset are used in the experiments. Additionally, there is another dataset named Figshare in which 5600 train images and 1400 test images are selected and used in the experiments. All of the images available in Kaggle dataset are used in this study, however, some of the low quality images in Figshare dataset are not utilized in that dataset.

In recent years, there have been several state-of-the-art studies for the classification of brain MRI images using Kaggle Brain MRI dataset and Figshare Brain MRI dataset. Comparison with the state-of-the-art methods in Table 3 indicates that the results on Kaggle Brain MRI dataset show that most of the DL architectures, such as AlexNet and VGG16, achieve better results compared to handcrafted methods. Similarly, the results on Figshare Brain MRI dataset show that AlexNet achieves the best accuracies for the classification of brain tumors. The best accuracies obtained using AlexNet on Kaggle Brain MRI dataset and Figshare Brain MRI dataset are 98.83% and 99.43%, respectively. It means that using these DL methods can be helpful for achieving better accuracy and results, especially in the medical fields, because of the disease diagnosis in the initial steps is very crucial and vital in these fields.

Table 3. Comparison with the State-of-the-Art on Kaggle and Figshare Datasets

Autours	Preprocessing and Segmentation	Feature Extraction and Classification	Dataset	Accuracy
Manav et al., (2021)	Segmentation, Image Enhancement	Pixel-based feature extraction and CNN.	Kaggle	97.79%
Prabira & Santi, (2021)	Deep Fusion	PCA and fused deep features and SVM.	Kaggle	97.89%
Arshia et al., (2020)	N/A	VGG16	Figshare	98.69%
Polat & Güngen, (2021)	N/A	DenseNet121, ResNet50.	Figshare	98.91%, 99.02%
This Study	N/A	AlexNet	Kaggle	98.83%
This Study	N/A	VGG16	Kaggle	97.33%
This Study	N/A	ResNet	Kaggle	66.67%
This Study	N/A	AlexNet	Figshare	99.43%
This Study	N/A	VGG16	Figshare	99.15%
This Study	N/A	ResNet	Figshare	96.58%

CONCLUSION

In this paper, three DL models including Alexnet, VGG16 and ResNet are utilized to classify brain tumors by employing MRI images. The performances of these models have been investigated using two datasets, namely Kaggle and Figshare, and five metrics are used to calculate their performances. AlexNet achieves 98% accuracy on the Kaggle dataset, VGG16 has 97% accuracy, and ResNet gets 66% accuracy. AlexNet has offered the highest level of accuracy among these networks. On the other hand, AlexNet achieves 99% accuracy in the Figshare dataset, VGG16 gets 99% accuracy, and ResNet has 96% accuracy. AlexNet and VGG16 outperform ResNet in terms of accuracy. These accuracies allow for the early detection of abnormalities before they create physical harm such as disability or other complications. The experimental results reveal that DL models perform well on Figshare Brain MRI dataset and Kaggle Brain MRI dataset, however, better accuracy is obtained on the Figshare dataset. The reason for this is that we have more images in the train and test sections when Figshare dataset is used. Therefore, the efficiency is increased, and better results are obtained on Figshare dataset. For future work, since identifying the exact location of a brain tumor is very important and the location of the tumor determines the need for surgery to remove malignant tumors, other segmentation methods can be investigated. Additionally, more powerful and efficient DL architectures, such as ResNet50, can be used to increase the accuracy of brain tumor classification.

REFERENCES

Amin, J., Sharif, M., Mussarat,Y., Fernandes, S.L. (2018). Big Data Analysis for Brain Tumor Detection: Deep Convolutiona Neural Networks. Future Gener. Comput. Syst., vol.87. 290–.297. https://doi.org/10.1016/j.future.2018.04.065.

Arshia, R., Saeeda, N., Muhammad, I.R., Faiza, A., Muhammad, I. (2020). A Deep Learning-based Framework for Automatic Brain Tumors Classification Using Transfer Learning, Circuits, Systems and Signal Processing. vol. 39, 757-775.https://doi.org/10.1007/s00034-019-01246-3.

Çınar, A. & Yıldırım, M. (2020). Detection of Tumors on Brain MRI Images Using the Hybrid Convolutional Neural Network Architecture, Med. Hypotheses, 139 (2020), 109684. https://doi.org/10.1016/j.mehy.2020.109684.

Deepak, R. & Ameer, P.M. (2019). Brain Tumor Classification Using Deep CNN Features via Transfer Learning, Computers in Biology and Medicine, 1-7. https://doi.org/10.1016/j.compbiomed.2019.103345.

Figshare Brain Tumor Dataset, https://doi.org/10.6084/m9.figshare.1512427.v5, Accessed: December 2018.

Gumaei, A., Hassan, M.M., Hassan, M.R., Alelaiwi, A. N., Fortino, G. A. (2019). Hybrid Feature Extraction Method with Regularized Extreme Learning Machinefor Brain Tumor Classification. 36266-36273.https://doi.org/ 10.1109/ACCESS.2019.2904145.

Hasan, A.M., Jalab, H.A., Meziane, F., Kahtan, H., Ahmad, A.S. (2019). Combining Deep and Handcrafted Image Features for MRI Brain Scan Classification. IEEE Access. 79959–79967. https://doi.org/10.1109/ACCESS.2019.2922691.

Hemanth, G., Janardhan, M., Sujihelen, L. (2019). Design and Implementing Brain Tumor Detection Using Machine Learning Approach . Third International Conference on Trends in Electronics and Informatics.1-6..https://doi.org/ 10.1109/ICOEI.2019.8862553.

Huafeng, W., Haixia, P., Huafeng,W., Yanxiang , Z., Yehe, C. (2015). Deep Learning for Image Retrieval: What Works and What Doesn't, Conference Paper .1-22.https://doi.org/10.1109/ICDMW.2015.121.

Kaggle Brain Tumor Classification (MRI) dataset, by Bhuvaji, S., Kadam, A., Bhumkar, P., Dedge, S., Kanchan, S., https://doi.org/10.34740/KAGGLE/DSV/1183165, Accessed: December 2020.

Kebir, S.T. & Mekaoui, S. (2018). An Efficient Methodology Of Brain Abnormalities Detection Using CNN Deep Learning Network. in Proc. Int. Conf.Appl. Smart Syst. (ICASS). https://doi.org/10.1109/ICASS.2018.8652054.2018.

Latif, G., Iskandar, D., Alghazo, N.F., Mohammad, J.M.N. (2018). Enhanced MR Image Classification Using Hybrid Statistical and Wavelets Features. IEEE Access .9634- 9644.https://doi.org/10.1109/ACCESS.

2018.2888488.

Manav, S., Pramanshu, S., Ritik, M., Kamakshi, G. (2021). Brain Tumour Detection Using Machine Learning . Journal of Electronics and Informatics, Volume 3, Issue 4. 298-308, 2021. https://doi.org/10.36548/jei.2021.4.005.

Minz, A. & Mahobiya, C. (2017). MR Image Classification Using Adaboost for Brain Tumor Type. IEEE 7th International Advance Computing Conference .1-5.https://doi.org/10.1109/IACC.2017.0146.

Mohsen, H., Sayed, E., Dahshan, E., Badeeh, A., Salem, M. (2018). Classification Using Deep learning Neural networks for Brain Tumors Future. Computing and Informatics Journal.68-73.https://doi.org/10.1109/ACCESS.2018.2888488.

Narayana, R.L. & Reddy, T.S. (2018). An Efficient Optimization Technique to Detect Brain Tumor from MRI Images. International Conference on Smart Systems and Inventive Technology.1-4.https://doi.org/10.1109/ICSSIT.2018.8748288.

Polat, Ö. & Güngen, C . (2021). Classification of Brain Tumors from MR Images Using Deep Transfer Learning. The Journal of Supercomputing ,volume 77. 7236–7252. https://doi.org/10.1007/s11227-020-03572-9.

Polly, E.P., Shil, S.K., Hossain, M.A., Ayman, A., Jang, Y.M. (2018). Detection and Classification of HGG and LGG BrainTumor Using Machine Learning. International conference on Information Networking. 813-817.https://doi.org/10.1109/ICOIN.2018.8343231.

Prabira, K.S. & Santi, K.B. (2021). A Data Constrained Approach for Brain Tumour Detection Using Fused Deep Features and SVM, Multimedia Tools and Applications, vol. 80, 28745-28760.https://doi.org/10.1007/s11042-021-11098-2.

Rehman, A., Khan, M.A., Saba, T.Z., Mehmood, T.U., Ayesha, N. (2021). Microscopic Brain Tumor Detection And Classification Using 3D CNN And Feature Selection Architecture. Microsc. Res. Techn, vol. 84, no. 1,133–149, 2021. https://doi.org/10.1002/jemt.23597.2021.

Rehman,A., Naz, S., Razzak,M.I., Akram ,F., Imran,M.Amin, J., Sharif, M., Yasmin, M., Fernandis, S.(2017). A Distinctive Approach in Brain Tumor Detection and Classification Using MRI. Pattern Recognition Letters.1-10.https://doi.org/10.1016/j.patrec.2017.10.036.

Sethy, P.K. & Behera, S.K. (2021). A Data Constrained Approach for Brain Tumour Detection Using Fused Deep Features and SVM . Multimedia Tools and Applications (2021) 80.28745–28760.https://doi.org/10.1007/s11042-021-11098-2.

Sert, E., Ozyurt, F., Doğantekin, A. (2019). A New Approach for Brain Tumor Diagnosis System: Single Image Superresolution Based Maximum Fuzzy Entropy Segmentation and Convolutional Neural Network. Medical Hypothesis.1-9.https://doi.org/10.1016/j.mehy.2019.109413.

Shankar, A.S., Asokan, A., Sivakumar, D. (2016). Brain Tumor Classification Using Gustafson–kessel (G-k) Fuzzy Clustering Algorithm. International Journal of Latest Engineering Research and Applications .68-72.

Simonyan, K. & Zisserman, A. (2014). Very Deep Convolutional Networks for Large-Scale Image Recognition. arXiv preprint arXiv:1409.1556.https://doi.org/10.1109/ACPR.2015.7486599.

Talo, M., Baloglu, U.B., Yıldırım, O., Acharya, U.R. (2019). Application Of Deep Transfer Learning For Automated Brain Abnormality Classification Using MRI Images.Cognitive Systems Research, 54(2019), 176–188.https://doi.org/10.1016/j.cogsys.2018.12.007.

Vinoth, R. & Venkatesh, C. (2018). Segmentation And Detection Of Tumor in MRI Images Using CNN And SVM Classification. in Proc. Conf. Emerg. Devices Smart Syst. (ICEDSS). 21–25. https://doi.org/10.1109/ICEDSS.2018.8544306.2018.

Wang, F.& Gong, M. (2020). Single Image Super-Resolution by Residual Recovery Based On an Independent Deep Convolutional Network.VOLUME 4.1–10.https://doi.org/10.1109/ACCESS.2020.2986365.

Wang, H., Pan, H., Zhang, Y., Cai,Y. (2015). Deep Learning for Image Retrieval: What Works and What Doesn't. Conference Paper ·1-22. https://doi.org/10.1109/ICDMW.2015.121.

Zaw, H.T., Maneerat, N., Win, K.Y. (2019). Brain Tumor Detection Based On Naïve Bayes Classification. International Conference on Engineering, Applied Sciences and Technology.1-4. https://doi.org/10.1109/ICEAST.2019.8802562.