Available at www.ijsred.com

RESEARCH ARTICLE OPEN ACCESS

Brain Tumor Classification Using Deep CNN Features Via Transfer Learning

DR.Ms.R.Sangeetha Rajendran[1], M.Sakthibala[2]

[1](Assistant professor, Department of computer science, Mangayarkarasi college of arts and science for women, Paravai,Madurai.
E-mail:marysangeetha137@gmail.com)

[2](PG Student, Mangayarkarasicollege of arts and science for women, Paravai,Madurai.
E-mail:queensakthi2000@gmail.com)

Abstract:

Brain tumor computer-aided diagnosis transfer learning convolutional neural network support vector machine Brain tumor classification is an important problem in computer-aided diagnosis (CAD) for medical applications. This paper focuses on a 3-class classification problem to differentiate among glioma, meningioma and pituitary tumors, which form three prominent types of brain tumor. The proposed classification system adopts the concept of deep transfer learning and uses a pre-trained Google Net to extract features from brain MRI images. Proven classifier models are integrated to classify the extracted features. The experiment follows a patient-level five-fold cross-validation process, on MRI dataset from figshare. The proposed system records a mean classification accuracy of 98%, outperforming all state-of-the-art methods. Other performance measures used in the study are the area under the curve (AUC), precision, recall, F-score and specificity. In addition, the paper addresses a practical aspect by evaluating the system with fewer training samples. The observations of the study imply that transfer learning is a useful technique when the availability of medical images is limited. The paper provides an analytical discussion on misclassifications also.

Keywords: **Brain tumor, Computer-aided diagnosis, Transfer learning, Convolutional Neural Network, Support Vector Machine.**

INTRODUCTION:

Tumor is an uncontrolled growth of many cells in any part of the body. Tumors are of different types and have different characteristics and different treatments. At present, brain tumors are classified as primary brain tumors and metastatic brain tumors. The former begin in the brain and tend to stay in the brain, the latter begin as a cancer elsewhere in the body and spreading to the brain. Brain tumor segmentation is one of the crucial procedures in surgical and treatment planning. Brain tumor segmentation using MRI has been an intense research area. Brain tumors can have various sizes and shapes and may appear at different locations. Varying intensity of tumors in brain magnetic resonance images (MRI) makes the automatic segmentation of tumors extremely challenging. There are various intensity-based techniques which have been proposed to segment tumors on magnetic resonance images. Texture is one of most popular feature for image classification and retrieval. From the MRI images of brain, the optimal texture features of brain tumor are extracted by utilizing FCM and JAYA algorithm process. Then using these methods, such an algorithm classifies the tumor and non-tumor tissues and tumor is segmented. This method provides more efficient brain tumor segmentation compared to the segmentation

technique based on existing procedure and will provide more accurate result. Tumor is the abnormal growth of the tissues. A brain tumor is a mass of unnecessary cells growing in the brain or central spine canal. Today, tools and methods to analyse tumors and their behaviour are becoming more prevalent. Clearly, efforts over the past century have yielded real advances. However, we have also come to realize that gains in survival must be enhanced by better diagnosis tools. Although we haveyet to cure brain tumours, clear steps forward have been taken toward reaching this ultimate goal, more and more researchers have incorporated measures into clinical trials each advance injects hope to the team of caregivers and more importantly, to those who live with this diagnosis. Magnetic Resonance Imaging (MRI) has become a widely-used method of high-quality medical imaging, especially in brain imaging where MRI's soft tissue contrast and non-invasiveness are clear advantages. An important use of MRI data is tracking the size of brain tumor as it responds treatment. Therefore, an automatic and reliable method for segmenting tumor would be a useful tool. MRI provides a digital representation of tissue characteristics that can be obtained in any tissue plane. The images produced by an MRI scanner are best described as slices through the brain. MRI has the added advantage of being able to produce images which slice through the brain in both horizontal and vertical planes. This makes the MRI-scan images an ideal source for detecting, identifying and classifying the right infected regions of the brain. Most of the current conventional diagnosis techniques are based on human experience in interpreting the MRI-scan for judgment; certainly this increases the possibility to false detection and identification of the brain tumor. On the other hand, applying digital image processing ensures the quick and precise detection of the tumor. One of the most effective techniques to extract information from complex medical images that has wide application in medical field is the segmentation process. The main objective of the image segmentation is to partition an image into mutually exclusive and exhausted regions such that each region of interest is spatially contiguous and the pixels within the region are homogenous with respect to a predefined criterion. The cause of most cases is unknown. Risk factors that may occasionally be involved include: a number of genetic syndrome such as neurofibromatosis as well as exposure to the chemical vinyl chloride, Epstein-Barr virus, and ionizing radiation.

Magnetic resonance imaging (MRI) is the prime technique to diagnose brain tumors and monitor their treatment. Different MRI modalities of each patient are acquired and these images are interpreted by computer-based image analysis methods in order to handle the complexity as well as constraints on time and objectiveness. In this thesis, two major novel approaches for analysing tumor-bearing brain images in an automatic way are presented: Multi-modal tissue classification with integrated regularization can segment healthy and pathologic brain tissues including their sub-compartments to provide quantitative volumetric information. The method has been evaluated with good results on a large number of clinical and synthetic images. The fast run-time of the algorithm allows for an easy integration into the clinical work flow. An extension has been proposed for integrated segmentation of longitudinal patient studies, which has been assessed on a small dataset from a multi-center clinical trial with promising results. Atlas-based segmentation with integrated tumor-growth modelling has been shown to be a suitable means for segmenting the healthy brain structures surrounding the tumor. Tumorgrowth modelling offers a way to cope with the missing tumor prior in the atlas during registration. To

this end, two different tumor-growth models have been compared. While a simplistic tumor growth model offered advantages in computation speed, a more sophisticated multi-scale tumor growth model showed better potential to provide a more realistic and meaningful prior for atlas-based segmentation. Both approaches have been combined into a generic framework for analysing tumor-bearing brain images, which makesuse of all the image information generally available in clinics. This segmentation framework paves the way for better diagnosis, treatment planning and monitoring in radiotherapy and neurosurgery of brain tumors.

LITERATURE REVIEW:

Recent works on computer-aided medical diagnosis provide improved performances owing to the advent of deep learning concepts. Deep learning strategies have been extensively used in the medical image analysis of breast cancer studies [14] and lung cancer diagnosis [4]. Zuo et al. [16] developed a deep learning algorithm for human skin detection, which is a part of dermatology diagnostics. Charron et al. [2] used a deep convolutional neural network (CNN) to monitor brain metastases. More recently, a special class of deep learning, known as deep transfer learning, has been dominating the studies on visual categorization, object recognition and image classification problems [10]. Transfer learning allows the use of a pre-trained CNN model, which was actually developed for another related application. Transfer learning has shown its potential in CAD of medical problems also. Zhou et al. [15] used a pre-trained InceptionV3 model for differentiating benign and malignant renal tumors on CT images. Deniz et al. [3] proposed a classifier for breast cancer on histopathologic images. The authors used a pre-trained VGG-16 model and a fine-tuned AlexNet for extracting features, which were then classified using a support vector machine (SVM). Hussein et al. [5] introduced a learning model for lung tumor characterization and for pancreatic tumor characterization. The learning model was based on knowledge transfer and it had a 3D CNN architecture. The accuracy measures reported in the transfer learning-based algorithms were superior to those obtained using handcrafted engineering. Specifically, transfer learning has gathered attention in applications related to neuro-oncology. Studies were conducted to extract deep features from brain MRI images using pre-trained networks[8,1]. The studies showed the capability of transfer learning to work with smaller datasets. Yang et al. [13] used AlexNet and GoogLeNet in their research work on the grading of glioma from MRI images. In terms of the performance measures observed, GoogLeNet proved superior to AlexNet for the task. Talo et al. [12] achieved remarkable classification performance with deep transfer learning in their work on brain abnormality classification. The authors used ResNet-34 and the experiments included training of modified dense layers, training with data augmentation and fine tuning of a transfer learned model. The experimental results concluded that a deep transfer learned model can be adapted to medical image classification, with minimum pre-processing. Jain et al. [6] used a pre-trained VGG-16 network for diagnosis of Alzheimers disease from MRI. Transfer learning was applied to content-based image retrieval (CBIR) for brain tumors [11]. The evaluation was performed on a publicly available dataset and obtained promising results.

METHODOLOGY:

The system proposed a novel semi-automatic segmentation method based on population and individual statistical information to segment brain lesion in magnetic resonance (MR) images. The probability of each pixel belonging to the

foreground (tumor) and the back ground is estimated by the morphological based FCM is used.

A feature extraction algorithm based on GLCM is constructed followed by these probabilities and it is extracts the features from the image. It can easily be realized that the full or semi-automatic segmentation and classification methods are in fact region segmentation methods.

ADVANTAGES:

- This algorithm cancorrectly separate the regions that have the same properties we define.
- This methods can provide the original images which have clear edges the good segmentation results.

Dataset and pre-processing:

The dataset from fig share is openly available [26] and is commonly used for evaluating classification and retrieval algorithms [16]. It is a collection of 3064 brain MRI images from 233 patients, diagnosed with one of the three brain tumors (meningioma, glioma and pituitary tumors). The images belong to the T1-CE MRI modality and include coronal, sagittal and axial views. It contains 1426 brain MRI images with glioma (corresponding to 89 patients), 708 images for meningioma (corresponding to 82 patients) and the remaining 930 images correspond to cases of pituitary tumor (belonging to 62 patients). The images are available as.mat files and the size of each image is 512x512. GoogLeNet was originally designed for RGB colour images, with an input layer of size 224x224x3. The MRI images in the dataset were pre-processed in the following manner (Fig. 5). They were normalized in intensity values. A min-max normalization technique was followed to scale the intensity values between 0 and 1. They were resized to 224x224. Because MRI images are greyscale images, three channels were then created by replicating the greyscale values three times. The evaluation procedure for the designed system on the figshare dataset followed a patient-level five-fold cross-validation. The entire dataset of 233 patients was divided into five disjoint subsets. The divided subsets were of approximately equal size. One subset was selected as the test set while the rest formed the training set. This process was repeated such that every subset formed the test set once. Such a division of the dataset was to ensure that the data of a particular patient was not simultaneously present in the test set and in the training set.

Classifier settings:

The performance of an image classification system depends on the combination of image features and the classifier model. With regard to the classifier models used in the final stage of the proposed brain tumor classification system, there are three distinct experimental settings.

1. Transfer learned deep CNN model with its softmax classifier, as a stand-alone system. The modified GoogLeNet was trained using the training set (after pre-processing). The hyperparameters of the network were heuristically adjusted so as to facilitate the convergence of the loss function during training (Fig. 6). Adam was the chosen optimizer, considering its good learning rate and the parameter-specific adaptive nature of the learning rates. For Adam, the initial learning rate was chosen as 0.0003. The choice of a high value might prevent the loss function from converging and could cause overshoots. And a very small value of the learning rate increases the training time. The mini-batch size was set to 30. The choice is a compromise between the speed of training (a larger batch size means faster training) and the computational requirements (limit set by

the computer specifications). Also, a very large batch size adversely affects the model quality. Cross-entropy is the loss function used as it gives the measure of the closeness of the predicted and actual distributions. A higher learning rate is desirable at the modified FC layer so as to learn the MRI image specific features. So, a learning factor of 10 is set. The number of epochs was limited to 10, considering the occurrence of overfitting. The hyperparameter settings of our experiment are listed in Table 1.

2. Deep CNN features with SVM classifier. We extracted features from the pooling layer, placed afterthe final inception module of the modified GoogLeNet. The features were then classified using SVM. We useda multi-class SVM with an error-correcting output code (ECOC) model. A one-vs-all strategy was used for multi-class classification. There were three binary SVM learners, each with a linear kernel. Other parameters of SVM are given in Table 1.

3. Deep CNN features with KNN classifier. Features were extracted in a manner similar to the previous setting. A classification experiment was then performed with the KNN classifier. The main parameters of KNN include k, the number of nearest neighbours and the distance metric. We chose the value for k as 49, the square root of the number of samples in the training set as per our cross-validation settings. A lower value of k can make the system susceptible to noise and overfitting. A higher value means more computations. In addition, data imbalances with respect to classes may dominate the results if k is chosen high. We used Euclidean distance as the distance metric.

Smaller training data:

The original training set, used for validation of one test set, constitutes 80% of the images in the fig share dataset. The theory of transfer learning is recognized as a solution to the problem of scarcity of data to train a deep CNN [13,23]. We tested the capability of our system to perform under the condition when the availability of training data is limited: 1. By randomly selecting 70% of the original training set (i.e. 56% of the total images in fig share) 2. Using 50% of the original training set (i.e. 40% of the total images in fig share) 3. Using 25% of the original training set (i.e. 20% of the total images in fig share) We trained the transfer learned model with the smaller amount of data. The test set was then provided to the model. The features were extracted and then classified using SVM. Table 4 presents the corresponding performance measures. In addition to the overall classification accuracy, we used receiver operating characteristic (ROC) curves during analysis. Fig. 7 represents the ROC curves obtained for the case when 50% of the training data is used. Ideally, the area under the curve (AUC) of ROC is unity. AUC values for the three classes of tumors are shown in Table 4. The values indicate that the reduction in size of the training data has not impaired the system performance significantly. The observation has an advantage in practical scenarios. This is because the number of training samples available could be limited due to the unavailability of more medical data. Another obvious advantage of the smaller training set is a shorter training time. In our experiments, the training time for deep CNN was decreased from 2hr 40min to 55min when the number of training samples was reduced to 25% of the original.

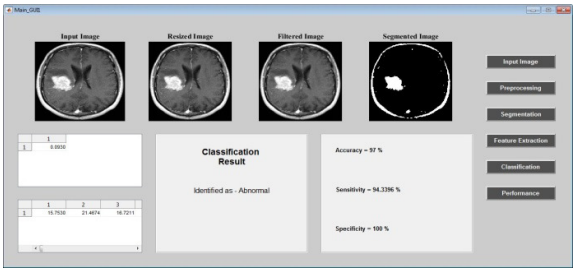

FIGURE-1 Performance analysis

RESULTS AND DISCUSSION:

We implemented the proposed classification model in MATLAB 2018b on a computer having specifications of 32GB RAM and Intel E3-1245v6 @3.70GHz CPU.

We conducted the experiments five times, and each experiment followed a five-fold cross-validation process.The average of the results after five trials is presented in a mean±standard deviation format.

Model	Parameters	Settings (values)
Transfer learned deep network	Initial learning rate mini-batch size algorithm loss function maximum epochs learning fator @ FC layer	0.0003 30 Adam cross-entropy 10 10
SVM	Model sub-type loss function coding learner kernel regularisation solver	ECOC Hinge One-vs-all SVM Linear L2 BFGS
KVM	Number of neighbours, k distance	49 Eucliedean

TABLE 1: Experimental parameters

Performance metrics and evaluation:

Several performance measures are defined for the standard evaluation of a classifier. Classification accuracy is the most extensively used quality index. Accuracy, in classification, is defined as the ratio of the number of correctly classified samples to the total number of data samples. The classification accuracies obtained in our experiments are the following.

- The classification accuracy ofthe deep transfer learned (stand-alone) model is 92.3±0.7%.
- The accuracy with SVM on deep CNN features is 97.8±0.2%.
- The accuracy with KNN on deep CNN features is 98.0±0.4%.

The data reveals that a superior performance is achieved when SVM or KNN is used to classify the deep CNN features. Classification accuracy is an effective measure to characterize the performance when the test dataset contains an equal number of samples from each class. However, the dataset considered for the discussed classification problem is an unbalanced dataset. This necessitates further evaluation of the proposed system with more performance indices. We used confusion matrices to study the performance of our tumor classification system. A confusion matrix summarizes correct and incorrect classifications in a tabular form. Table 2 shows a sample confusion matrix for the SVM classifier, obtained during our experimentation. (M, G and P refer to meningioma, glioma and pituitary tumor, respectively.)

ACTUAL	Predicted		
	M	G	P
M	684	11	13
G	30	1394	2
P	9	0	921

TABLE 2: Confusion matrix for SVM classifier on deep CNN feature.

From a confusion matrix, different metrics can be derived to indicate the classifier's performance, specific to each tumor class. Essential metrics are precision, recall (or sensitivity) and specificity and are calculated using the relations given below.

Precision = T P/ T P + FP

Recall = TP /T P + FN

Specificity = TN /TN + FP

where, TP, FP, TN and FN are the number of classified cases of true positives, false positives, true negatives and false negatives, respectively. Table 3 presents the category-specific performance of the proposed system when the SVM classifier was used with deep CNN features.

The specificity values for all the classes are high. This is an indication of correctly identifying samples without a particular disease. The harmonic mean of precision and recall gives another important statistical measure of classification called the F-score, for each class. As there is an imbalance among the threeclasses, the metric called the average F-score (F puted by the relation given below.

Tumor type	Precision	Recall	Specificity
Meningioma	94.7±0.8	96.0±0.5	98.4±0.2
Glioma	99.2±0.3	97.9±0.2	99.4±0.3
Pituitary tumor	98.0±0.7	98.9±0.2	99. ±10.1

TABLE 3: Class-specific evaluation of brain tumor classifier.

The calculated value of Favg for our system with SVM classifier is 0.97.

Comparison with related works:

We compared the performance of our method with all the existing methods on the specific 3-class classification problem of brain tumors. Table 5 provides a broad comparison based on classification accuracy as a metric. The comparison shows that our method surpasses all the state-of-the art methods. The third column in the table defines the portion of the entire dataset used in training. The proposed method recorded the best result when 80% of the data samples was used for training. Instead, we present results for the case when 56% of the dataset is used for training. This is to illustrate the performance of our method with much smaller training data in comparison to related works. The table contains only accuracy as a performance metric because it is the common metric that is used in all the related works. In fact, the proposed work is better than the state-of-the-art methods in terms of all the metrics. Table 6 provides a more detailed comparison. Based on sensitivity and specificity measures, the proposed method is superior to the works [17,19] that used these measures. The proposed method shows an improvement over another method [21] in terms of the Favg score also.

Regarding misclassifications:

Based on the performance evaluation and detailed analysis, the following inferences about the system are made. The accuracy of the system improved when SVM or KNN was used instead of the classification layer within the transfer learned model. This meant that some of the classifications, which went wrong with the softmax-based classifier (of the deep learning model), were correctly classified by the SVM- and KNN-based classifiers. Table 7 shows a few of the sample instances. From the confusion matrix (Table 2) and the calculated class-specific metrics (Table 3), we find that most of the misclassifications pertain to the class meningioma. This finding can be attributed to the fact that there were fewer samples from this class in the dataset and that no class-specific data augmentation was used to balance the dataset. Another aspect that our work concentrated on was the handling of smaller amounts of training data. We noticed a relative drop in performance (Table 4) with the reduction in training samples. This suggests that the discriminative power of features extracted using the transfer learned deep CNN is affected. This effect was further analysed by studying the variation of training and validation losses against iterations. Fig. 8 shows a sample instance where 25% of the training data was used. In this case, we find that the training loss decreases, whereas the validation loss increases after 60 iterations. This behaviour indicates that the phenomenon of overfitting took place. This is the reason for the lower classification accuracy because the model has learned specific training images without gaining a good generalization capability. This, in turn, means that the model

complexity is greater compared to the available training samples. Table 8 represents the confusion matrix for the SVM classifier which was provided with features from the deep transfer learning model trained with 25% of the training set. It is used to study misclassifications as a consequence of overfitting. The class meningioma had the smallest number of samples in the training set and was the most affected class in terms of misclassifications. Overfitting could be avoided either by stopping the training at an earlier stage or by data augmentation. This aspect suggests scope for work in the future.

Training data	Accuracy	AUC(%)		
		Meningioma	Glioma	Pituitary
Full	97.8±0.2	99.5	99.9	99.7
70%	97.1±0.2	99.4	99.7	99.8
50%	95.7±0.5	98.7	99.1	99.7
25%	93.3±0.6	97.8	98.9	99.2

TABLE 4: Performance with reduced training data

Work	method	Training data	Accuracy
Jun cheng[17]	BoW-SVM	80%	91.28%
Ismael[19]	DWT-Gabor-NN	70%	91.90%
Pashaei[21]	CNN-ELM	70%	93.68%
Nyoman[20]	CNN	-	84.19%
Afshar	CapsNet	-	90.89%
proposed	Deep CNN-SVM	56%	97.1%

TABLE 5: Related & comparison using figshare dataset

CONCLUSION:

This paper presents an accurate and fully automatic system, with minimum preprocessing, for brain tumor classification. The proposed system applied the concept of deep transfer learning to extract features from brain MRI images. The features were used with proven classifier models for an improved performance. The system recorded the best classification accuracy compared to all the related works. The performance was evaluated using other metrics also, to ascertain the robustness of the system. Moreover, the system showed a good performance with a smaller number of training samples.

REFERENCE:

[1] K.B. Ahmed, L.O. Hall, D.B. Goldgof, R. Liu, R.A. Gatenby, Finetuning convolutional deep features for MRI based brain tumor classification, in: International Society for Optics and PhotonicsMedical Imaging 2017: Computer-Aided Diagnosis, vol. 10134, 2017, 101342E.

[2] O. Charron, A. Lallement, D. Jarnet, V. Noblet, J.B. Clavier, P. Meyer, Automatic detection and segmentation of brain metastases on multimodal MR images with a deep convolutional neural network, Comput. Biol. Med. 95 (2018) 43-54.

[3] E. Deniz, A. “engür, Z. Kadiro§lu, Y. Guo, V. Bajaj, Ü. Budak, Transfer learning based histopathologic image classification for breast cancer detection, Health Information Science and Systems 6 (1) (2018) 18.

[4] Y. Gu, X. Lu, L. Yang, B. Zhang, D. Yu, Y. Zhao, T. Zhou, Automatic lung nodule detection using a 3D deep convolutional neural network combined with a multi-scale prediction strategy in chest CTs, Comput. Biol. Med. 103 (2018) 220-231

[5] S. Hussein, P. Kandel, C.W. Bolan, M.B. Wallace, U. Bagci, Lung and pancreatic tumor characterization in the deep learning era: novel supervised and unsupervised learning approaches, IEEE Trans. Med. Imaging (2019). https://doi.org/10.1109/TMI.2019.2894349.

[6] R. Jain, N. Jain, A. Aggarwal, D.J. Hemanth, Convolutional neural network based Alzheimer’s disease classification from magnetic resonance brain images, Cogn. Syst. Res. (2019). https://doi.org/10.1016/j.cogsys.2018.12.015

[7] S. Kumar, C. Dabas, S. Godara, Classification of brain MRI tumor images: a hybrid approach, Procedia Comput. Sci. 122 (2017) 510- 517.

[8] R. Liu, L.O. Hall, D.B. Goldgof, M. Zhou, R.A. Gatenby, K.B. Ahmed, Exploring deep features from brain tumor magnetic resonance images via transfer learning, in: IEEE International Joint Conference on Neural Networks (IJCNN), 2016, pp. 235-242.

[9] G. Mohan, M.M. Subashini, MRI based medical image analysis: survey on brain tumor grade classification, Biomed. Signal Proces. 39 (2018) 139-161.

[10] L. Shao, F. Zhu, X. Li, Transfer learning for visual categorization: A survey, IEEE Trans. Neural Netw. Learn. Syst. 26 (5) (2015) 1019-1034.

[11] Z.N.K. Swati, Q. Zhao, M. Kabir, F. Ali, A. Zakir, S. Ahmad, J. Lu, Content-based brain tumor retrieval for MR images using transfer learning, IEEE Access 7 (2019) 17809-17822

[12] M. Talo, U.B. Baloglu, U.R. Acharya, Application of deep transfer learning for automated brain abnormality classification using MR images, Cogn. Syst. Res. 54 (2019) 176-188.

[13] Y. Yang, L.F. Yan, X. Zhang, Y. Han, H.Y. Nan, Y.C. Hu, X.W. Ge, Glioma grading on conventional MR images: a deep learning study with transfer learning, Frontiers in Neuroscience 12 (2018).

[14] M. Yousefi, A. Krzyz`ak, C.Y. Suen, Mass detection in digital breast tomosynthesis data using convolutional neural networks and multiple instance learning, Comput. Biol. Med. 96 (2018) 283-293.

[15] L. Zhou, Z. Zhang, Y.C. Chen, Z.Y. Zhao, X.D. Yin, H.B. Jiang, A deep learning-based radiomics model for differentiating benign and malignant renal tumors, Transl. Oncol. 12 (2) (2019) 292-300.

[16] H. Zuo, H. Fan, E. Blasch, H. Ling, Combining convolutional and recurrent neural networks for human skin detection, IEEE Signal Proc. Let. 24 (3) (2017) 289-293.

