Brain Tumor detection from brain MRI using Deep Learning

Abhishek Anil[1], Aditya Raj[2], H Aravind Sarma[3], Naveen Chandran R[4], Deepa P L[5]
Department of Electronics and Communication Engineering, Mar Baselios College of Engineering and Technology, Trivandrum
[1] abhishekanil14@gmail.com, [2] adityaraj1797@gmail.com, [3]aravindsarmah@gmail.com, [4]naveenchandran97@gmail.com, [5]deepa.pl@mbcet.ac.in

Abstract- **Health experts are increasingly taking advantage of the benefits of most modern technologies, thus generating a scalable improvement in the area of health care. Because of this, there is a paradigm shift from manual monitoring towards more accurate virtual monitoring with minimum percentage of error. Advances in artificial intelligence (AI) led to exciting solutions with high accuracy for medical imaging technology and is a key method for enhancing future applications. Detection of brain tumor is a very difficult task in medical field. Detection of brain tumor manually is time consuming and requires large number of mri images for cancer diagnosis. So, there is a need for automatic brain tumor detection from Brain MR images. Deep learning methods can achieve this task. Different deep learning networks can be used for the detection of brain tumors. The proposed method comprises of a classification network which classifies the input MR images into 2 classes: on with tumor and the second without tumor. In this work, detection of brain tumor is done via classification by retraining the classifier using the technique known as transfer learning. The obtained result shows that our method outperforms the existing methods.**

Index terms: Deep Learning, Transfer learning, brain tumor detection, MRI.

1. INTRODUCTION

Cancer can be defined as a disease which causes cells to multiply uncontrollably with a potential to invade or spread to other part of the body. The tumors are swollen mass in part of the body caused by an abnormal growth of cells. The occurrence of tumor in brain led to brain tumors. The tumors can be either of malignant or benign. Malignant tumors are cancerous and they multiply without any control and affects other cells. The benign tumor is a noncancerous tumor which means they will not spread to the nearby tissues and hence less risky.

Brain tumors are one of the deadly cancers which can last longer and can have psychological impact on patients. And the brain tumor needs the most costly care than any other cancers. The brain

tumors can be classified into 120 types. And in this work we are dealing with the astrocytoma which is a part of gliomas. The astrocytoma is the most commonly seen brain tumor which affects the glial cell of the brain. The glial cell are neurological cell that present in the brain for providing nutrition to the neurons. Based on the grade of cancer cell the astrocytoma can be classified into 4 types.

- GRADE 1: These type of tumors grows really slowly and they won't spread to other tissues. They can be cured by surgery.
- GRADE 2: These type of tumors more likely to slow growth and they won't spread to other tissues but there is chance for the return of the tumors.
- GRADE 3: These tumors are more lethal than above types and they are more likely to fast growing and rapid cell dividing.
- GRADE 4: These types are the extreme worse cases of tumors. Here the tumors are actively dividing and they do have blood vessels around the dead tissues.

In the above types, grade1 and grade2 are not malignant but the others are malignant cells. The cause for astrocytoma are unknown even though there are some risk factors they are age, continuous exploration to radiation and family history of gliomas.

The diagnosis methods usually followed by the hospitals are
- NEUROLOGICAL EXAM: Here the doctors check your vision, hearing, muscle strength, reflex etc. and also press the nerves that connect your eyes and brain.
- MRI: The MRI are taken from large machines with strong magnets are connected to computers to take the detail pictures of the brain.
- CT SCAN: The x-ray machines are connected to the computers to take the serial pictures of the brain.
- SPINAL TAP: In spinal tap the doctor may remove some cerebrospinal fluid from the spinal cord .Using thus they can identify the tumor.
- BIOPSY: The removal of tissues for checking the tumor is called biopsy.

Basically MRI are used for detecting the brain tumor. MRI are widely used than the CT scan why because the brain is filled with soft tissues and the MRI are focused on soft tissue but whereas CT are focused on hard tissues. And the accuracy of detection

depends on the radiologist. If the radiologist is inaccurate in his work it will affect the accuracy of the detection so we are introducing deep learning techniques for the efficient detection of the brain tumors and corresponding network for efficient pre-processing and segmentation of the MRI.

A few months back news was reported by the New York Times stating "Failure to Diagnose Brain Tumour leads to $1.4 Million Verdict". A child was suffering from headache, parents took him for diagnosis. During the primary check-up, the doctor couldn't find the cause of a headache. So suggested to take MRI scan. Once the result was out, both the radiologist as well as doctor failed to find a growing tumour inside the child's head. So the doctor prescribed the medicine for headache.

Three months later the child had lost his eyesight so sudden. On detailed analysis, it was found that the child was suffering from brain tumour and the child has lost his eyesight as the side effect of this. The family filed a case against the doctor and the radiologist who did the preliminary analysis and the case was charged for medical negligence. Once the court called for hearing, in defence, the doctors argued that the delay itself was so brief that even if a tumour had been timely and properly diagnosed, the child would have still lost his eyesight. The defence also argued that there was no way that the tumours could have grown to the extent that it did between the time of the alleged delay and subsequent discovery.

After hearing, the court comes to a conclusion stating that "if a patient comes for diagnosis it's the doctor's duty to ensure that the patient is properly diagnosed. If the doctor failed to do so then it's a crime". So as a compensation $1.4 million has to be given to the family.

This is not the only case that is happened in and around the world. Each day, the patients suffering from medical negligence is increasing. So, as a remedy, automated system has to developed to diagnose the patient's health from the test results. One way of approach for finding the patient having a Brain tumour or not from the MRI have been discussed here.Heba et. al. [1],Here deep learning technology was used for segmentation of brain tumors. The conducted experiment was to classify the mr images into 4 classes e.g. normal, glioblastoma, sarcoma and metastatic bronchogenic carcinoma tumors. For improving the feature extraction and performance the classifier was mixed with DWT and PCA in the above experiment

Lina Chato Dans Shahram Latifi et. al. [9],used machinelearning techniques for predicting glioma brain tumor. For each slice ROI was located for extracting region histogram. A pretrained CNN approach was used for extracting the features and thereby training the model. The best classification accuracy was achieved by pre - trained and trained AlexNet, best classification accuracy did not exceed 46 percent. Histogram characteristics achieved a 68.5 percent accuracy for two classifications. N. Varuna et. al[2]., focused mainly on grey level co-occurrence matrix (GLCM) and DWT brain for reducing the complexity and for improving the performance. Here morphological methods are used for noise reduction and PNN is used as classifiers. The conducted experiment got an end result of nearly 100% accuracy.

Zahra sobhsninia.et.al[3],they conducted studies about different angles of mr images and used different network for segmentation.The conducted experiment has got a diescore of 0.73 in single network and 0.79 in multiple network. Chinmayi et.al. [5], discussed a method to segment and classify the MRI brain image usingBhattacharyya coefficient into two classes, normal and abnormal. Brain tumor segmentation is done using an unsupervised automatic training method.Nilesh et.al.[6],they had conducted studies about improving the performance and accuracy using BWT and SVM classifiers. The experiment achieved an accuracy of 96.5% and specificity of 94.2% and sensitivity of 97.2% for detecting normal and abnormal tissues from brain MRI.Ali Isin.et.al. [7],discussed about the automatic segmentation of mr images using the recent deep learning methods in the field. Here an assessment of the current state and the future developments needed for standardize mri segmentation to clinical routines are also provided.Athency Antony.et.al[8],they discussed about the effective classification of mr images in to normal and abnormal cells using deep learning CNN. In this method preprocessing, classifying and feature extraction are done using matlab.

In this paper, we are proposing a DL technique for the detection of brain tumor from MR images. Classification networks are retrained using the technique of transfer learning for converting it to detect the tumor. The organization of the paper is as follows. The section 2 describes Deep Learning techniques followed by the proposed system in section 3. The section 4 describes the details of implementation and section 5 results and discussion.

2. DEEP LEARNING

"Artificial intelligence (AI) is the ability of a machine which can replicate the intelligent human behaviour. AI is obtained by a imitating intelligent human behaviour like thinking, learning and decision making when trying to solve a problem. Artificial Intelligence have various subfields. "Machine learning is a sub-field of AI" which uses different algorithms for doing a specific tasks. "Deep learning is a sub-field of Machine Learning" where a big data can be handled in the most effective way where machine learning fails.

2.1 Machine Learning (ML)

Machine learning is an AI subset which allows computers to learn without being programmed explicitly. Machine is trained on a huge amount of data to create a deep learning model that helps to arrive at conclusions from the learned features. So the machine is trained in such away as to identify unique features from the input dataset, for forming the feature vector and creating a model for the unique identification of a particular object.

Deep Learning (DL) is a subset of Machine Learning based on some specific set of algorithms which try to mimic human neural system using multi-layered neural network. In order to extract the right and unique feature on its own, DL networks need high dimensionality data. DL network will identify and extract the unique features from the given dataset. Therefore the number of inputs required is very large. Since in DL network the amount of the data required is large, dimensionality of the Feature vector cannot be predicted.

2.2 Deep Neural Network (DNN)

The brain consists of several connected neurons known as the neural network. A neural network's basic building block is known as a perceptron. Artificial neurons are a deep neural network. All Deep Neural Networks consist of three different stacks of layers: Input Layer, Hidden Layer and Output Layer

The input layer receives all the inputs the hidden layer extracts unique features and the output layer gives the conclusion The number of neurons in the input layer depends on the size of the input data and on the specific application in the output layer. All layers are termed hidden layers in between input and output layer. Each hidden layer in DL network comprises of a stack of certain layers desired for some particular task. The number of hidden layers in the hidden layer stack depends on the complexity of the input data and the application to be used. In deep learning, the number of hidden layers can be increased for doing different functions. As the network is learning itself by extracting uniqueness from the input data, the number on hidden layers required will be very high compared to an ordinary ML neural network. But as the hidden layer increases the network complexity will increase, but the accuracy will also increase.

Different Deep Learning networks are available in literature. Some of them are given in table I. From the above mentioned DL networks for Brain Tumour detection from MRI, CNN (convolutional Neural Network) is the best.
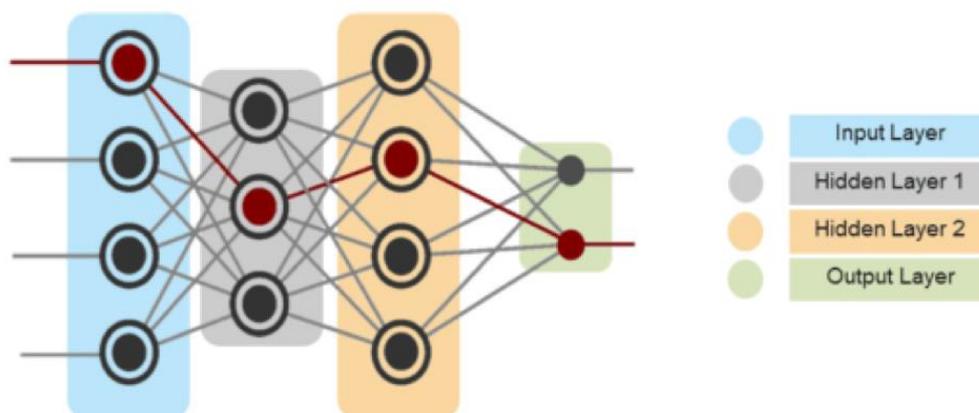

Figure 1:Deep Neural Network

Table 1: Deep Learning Networks

Model	Category	Learning model	Characteristics
Auto-encoder	Generative	Unsupervised	• Suitable for feature extraction • Works with unlabelled data
LSTM	Discriminative	Supervised	• Efficient for data with long time lag • Gates provides necessary protection to the access of the memory cell
CNN	Discriminative	Supervised	• Computations are done by convolutional layers. • Connections are lesser than DNN • Large dataset is needed for training
VAE	Generative	Semi-supervised	• A class of AEs • Suitable for scrutiny of labelled data
GAN	Hybrid	Semi-supervised	• Consist of Generative and Discriminative network • It is capable of noisy data

Different Deep Learning networks are available in literature. Some of them are given in table I. From the above mentioned DL networks for Brain Tumour detection from MRI,CNN (convolutional Neural Network) is the best.

2.3 CNN (Convolutional Neural Network)

Deep CNN consist of several layers which can process large amount of datausing linear as well as non-linear operators that are learned simultaneously to solve a particular problem. In other words, it is a 3d collection of nodes that form several layers that process internally and transform an input to produce an output. Like ordinary neural networks made of neurons with learning weights and biases. CNN's main parts are
1. Convolutional layer
2. Rectified Linear Unit (ReLU)
3. Pooling layer

A stack of these 3 layers are repeated several times periodically to from the Deep CNNs.

Convolution layer is performing convolution operation using a set of learnable filters to extract unique features from the input data that can be used as the input for the upcoming layers in a different manner. The selection and formation of learnable filters depends on the application. Every filter or kernel is of one-dimension or two-dimension, depends upon the dimension of the input data and can be small in size but can be extended through the full depth of the input volume. Normally square kernels are used if the input data is a 2D or a 3D one. As the convolutional layer is performing convolution operation, the network is not fully connected and thus it decreases the network complexity. In order to understand the CNN, we need to know the 3 key concepts.

Local Repetitive Fields

In the typical neural network, all neurons in the in input layer are connected to neurons in the first hidden layer, but in CNN, the some neurons in the input layer are connected to the neurons in the hidden layer in small regions of the kernel size. These are called local repetitive fields and are translated from the input layer to the hidden layer through an image to create a feature map. This will reduce the dimension of the parameter space.
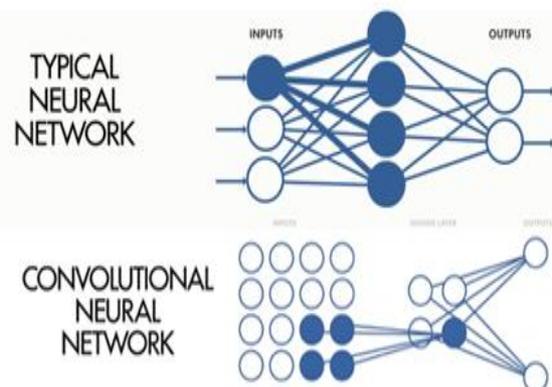

Figure 2: Basic Difference between Typical Neural Network and CNN

Shared Weights and Bias

The weights and biases are known as the trainable filter parameters. Typical neural network have neuron with weight and bias values as parameters, but the value of these parameters changes if there is mismatch in the outputs to make the relation: "Desired output equal to the Actual

output". In convolution layer, the weight and bias can be made same for all the neurons in a particular stack of the hidden layers. That means all neurons in particular layer will detect the same features to identify the object correctly.
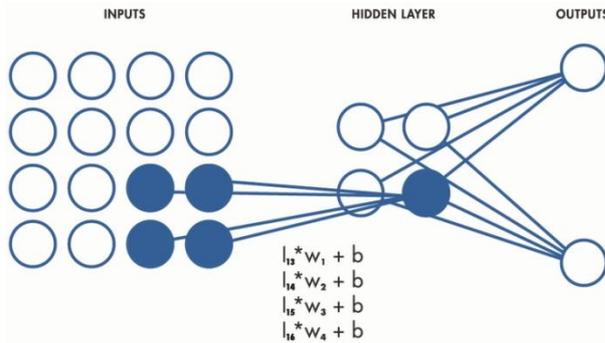

Figure 3: Shared Weights and Bias

Activation and Pooling

In the activation step, the activation function applies a transformation to the output of preceding neuron. The most commonly used activation function in CNN is the rectified linear unit (ReLU). This function works in such a way that if the input is a positive value, the same value is mapped and if the input is a negative value, it is mapped to zero as shown in the figure.
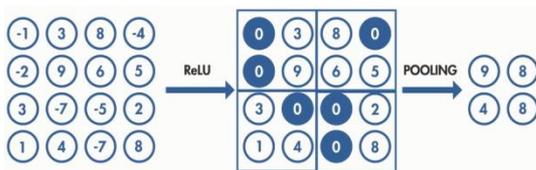

Figure 4: Activation and Pooling

3 BRAIN TUMOR DETECTION SYSTEM

Automated Brain tumor detection system will help the doctor as well as radiologist to diagnose the tumor area correctly and more accurately. This project deals with the detection of brain tumor from MRI. The block schematic of the proposed system is shown in fig 5.
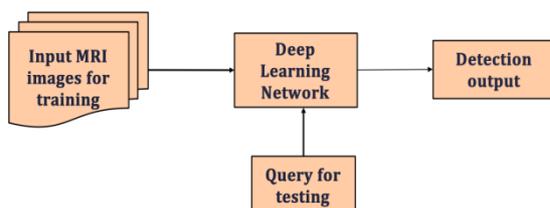

Figure 5: Block Diagram of the proposed system

This system is designed for effective detection and pre-processing of the brain tumor from the MRI. Here, we are going to provide thousands of images from the database to train the network. Through this work, we are proposing a new or modified deep learning network which is capable of effectively pre-processing as well as classifying the input data. Pre-processing comprises of image enhancement, denoising, etc. and it is done by the network itself. After pre-processing, the network will be trained to classify the given MRI into 2 classes, whether tumor present or not, followed by post-processing.The post processing mainly deals with the description or representation of tumor area identified by the deep learning network and it is done by using image processing techniques.

3.1 Proposed Architecture

The high- dimensional input data are transferred to the input layer. The input layer consists of the same no: of neurons as that of no: of elements in the input data. If the input is an image of MxN size, for example, the number of input neurons is MN.

Now, Brain MRI image is sent as an input to the input layer and the network will identify whether brain tumor present or not. Here the input layer is acting as a buffer for passing the input data from the outside world to the first set of hidden layers. The first set of hidden layers known as Hidden layer 1 contain kernels that can extract the low level features of the input data like edges of the brain images based on the difference in contrast levels. The output from the first hidden layer is given to the next set of hidden layers that is Hidden layer 2, where the various mid-level brain features are extracted.
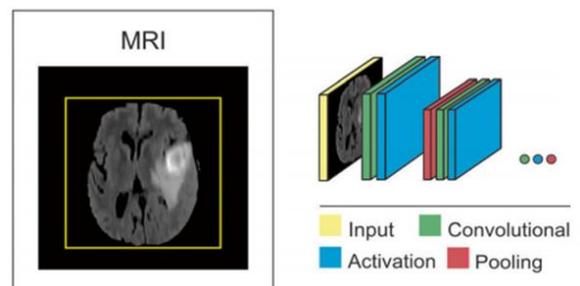

Fig. 6: CNN architecture of the proposed system

These features are now fed to the hidden layer 3 where the entire brain, which is the high-level feature, can be recovered. The layer hidden layer 3 output is then sent to the output layer. The

output layer is a layer that is fully connected. The output layer is therefore classified on the basis of the previous result and predicts whether the tumor is present.

Different DL networks are available for classification and detection purpose. Some of the examples are Alexnet [14] and VGGnet [11, 12]. Alexnet is the name of the CNN designed by Alex Krizhevsky and is trained on more than a million images from the ImageNet database.Alexnet consist of Eight layers: 5 convolutional layers and 3 fully connected layers. Alexnet uses the non-saturation RELU Activation function. The input image size is fixed and is 227*227*3.

Vgg16 [11] is a deep convolutional neural network for object recognition that was developed and trained by visual group geometry. Vgg16 consists of 16 layers: 13 convolutional layers and 3 fully connected layers. Vgg16 is preferred over alexnet because in alexnet the filter size is 11*11. In order to extract fine details of image is very difficult. The accuracy of vgg16 is 92.3%. Disadvantages are: (1) extremely slow to train and consumes a lot of time (2) Network architecture weights are large.

Vgg19 [12] is a convolutional network that is trained on more than one millions images from the image-net database.The network is 19 layers deep and can classify images into 1000 categories such as mammal,animals, keyboard etc.Vgg19 consists of 19 layers: 16 convolutional layers and 3 fully connected layers. To reduce the volume max pooling layers were used as handlers in vgg19.Two fully connected layers were used with 4096 neurons. Convolutional layers were used for feature extraction and max pooling layers associated with convolutional layers for reducing feature dimensionality.

These networks can be retrained or reused for some other applications by the method of transfer learning.

4. TRANSFER LEARNING

Training a deep model may require a lot of data and computational resources but with the help of transfer learning we create a new model by modifying the pre trained model, At First, we have to load the network after loading the network replace the final layers according to result to be obtained after replacing it, train the network with new dataset after the training phase is completed ,test image is given to the trained network and will get the deployed result. Transfer Learning can be done in two types:

Fig. 7: Transfer Learning

At First, we have to load the network after loading the network replace the final layers according to result to be obtained after replacing it,train the network with new dataset after the training phase is completed ,test image is given to the trained network and will get the deployed result. Transfer Learning can be done in two types:
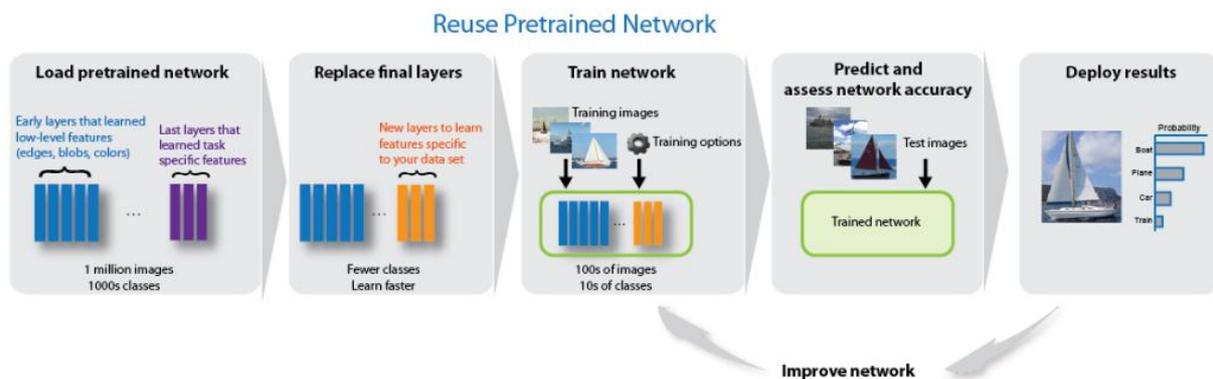
Train the top layers only

Chop off the top layer only and replace it with a randomly initialized one. then train the parameter only on the top layer of the network, while all other parameter are fixed. This method can be used for Feature extraction. Since fixed part act as the feature extractor and top layer act like traditional fully connected neural network.

This approach works the best if the data and the task are similar to the data and the task that the original model was trained on. Having few parameters to train also reduce the risk of over fitting. If we have more data can unfreeze these transferred parameters and train the entire network.
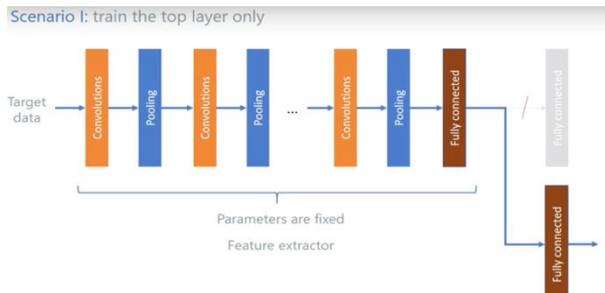
Fig. 8: Transfer learning - train the top layers

Fine Tune all layers

In this setting what actually transfers is basically the initial value of the parameter. Initializing the weight using pretrained model instead of initializing them randomly can give the model a warm start and speed up the convergence. To preserve the initialization from the pretraining its common practice that to lower the learning rate by its magnitude.
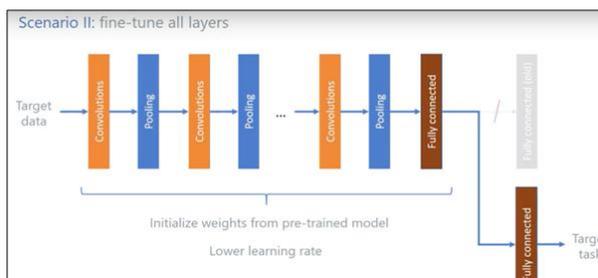

Fig. 9: Transfer learning –fine tune all the layers

To prevent changing the transfer parameter too early, it's common to start with frozen parameter , train only the randomly initialized layer until they

converge then unfreeze all parameters and frequency tune the network.

Basically used when having limited amount of data for one task. But have large volume of data for a similar task or have model already been there.

In this proposed method, we have used transfer learning for detecting brain tumor via classification of input data into one with tumor and the other without tumor. All the three above said networks are trained for classification purposes and the number of classes are 1000. Here, in our method we have only 2 classes: images with tumor and without tumor. Therefore, by using transfer learning we have created another network which can classify the input data into 2 classes. From the classified output we have checked the classification accuracy and detection accuracy.

4. RESULT AND DISCUSSION

We created a dataset consists of several thousands of images from the available datasets like OASIS and BRATS2018 challenge. The dataset consists of 2 classes, one with tumor and second without tumor. From these more than 20000 images are selected. Among these, 80% are used for training and remaining 20% for testing purposes. Deep Learning toolbox in MATLAB 2018b is used for simulating transfer learning in the existing networks such as Alexnet and VGGnet. Among the 3 networks, VGG19 is most efficient.

Detection accuracy obtained for the above said networks are shown in table 2.

Table 2: Results

Detection					
Network	**No: of Layers**	**Training execution time (HH:MM:SS) (one time process)**	**Testing execution time for single image (secs)**	**Correctly detected out of 4689**	**Detection accuracy (%)**
Alexnet	8 stacks (25 layers)	11:38:24	7.59	4203	**89.64**
VGG16	16 stacks (41 layers)	21:32:56	19.97	4371	**93.22**
VGG19	19 stacks (47 layers)	24:45:15	20.23	4491	**95.78**

5. CONCLUSION

The proposed method can effectively detect brain tumours from MR images. Transfer learning is used for performing detection using a classification network. MR images are taken from the standard datasets available in the literature. We have done the work using 3 different networks as discussed earlier. From the obtained results, we can conclude that VGG19 can be used as the best detector among the three.

All the above said networks have input size limitation of 227x227 or 224x224. But most of the MRI data are of size 512x512. So resizing may cause data lose. In order to compensate for this, we can create a network which can process images of size 512x512.

REFERENCES

[1] Heba Mohsen a,*, El-Sayed A. El-Dahshan b,c El-Sayed M. El-Horbaty d, Abdel-Badeeh M. Salem d *"Classification using deep learning neural networks for brain tumors" Future Computing and Informatics Journal 3 (2018)*

[2] N. Varuna Shree, T. N. R. Kumar (2018) "Identification and classification of brain tumor MRI with feature extraction using DWT and probabilistic neural network*". Brain infomatic March 2018 ,volume 5 ,Issue 1, pp 23-30*.

[3] Zahra Sobhaninia, Safiyeh Rezaei, Alireza Noroozi, Mehdi Ahmadi, Hamidreza Zarrabi, Nader Karimi, Ali Emami,Shadrokh Samavi, "Brain Tumor Segmentation Using Deep Learning by Type Specific Sorting of Images"*, International conference on Computer Vision and Pattern Recognition (CVPR),* 20th Sep 2018.

[4] Mehdi Mohammadi, Ala Al-Fuqaha, Sameh Sorour and Mohsen Guizani, "Deep Learning for IoT Big Data and Streaming Analytics: A Survey", *IEEE Journal on Communication Surveys and Tutorials (COMS&T), pp 1-40, June 2018.*

[5] P. Chinmayi, L. Akilandeswari, M. Prabu Kumar, K. Murali Babu "An Efficient Deep Learning Neural Network Based Brain Tumor Detection System" *International*

Journal of Pure and Applied Mathematics Volume 117 No. 17, 151-160, 2017

[6] Bahadure NB, Ray AK, Thethi HP (2017) "*Image analysis for MRI based brain tumor detection and feature extraction using biologically inspired BWT and SVM*". *Int J Biomed Imaging 2017, Article ID 9749108,*

[7] Ali Işın, Cem Direkoglu, Melike Sah " *Review of MRI-based brain tumor image segmentation using deep learning methods"* Procedia Computer Science, volume 102, page 317-324

[8] Athency Antony, Ancy Brigit M.A , Fathima K.A. , Dilin Raju , Binish M.C "Brain Tumor Detection and Classification in MRI" *International Journal of Innovative Research in Science, Engineering and Technology 32,2007 Certified Organization Volume 6, Special Issue 5, March 2017*

[9] Lina Chato, Shahram Latifi Department of Electrical and Computer Engineering University of Nevada, Las Vegas (UNLV) "Machine Learning and Deep Learning Techniques to Predict Overall Survival of Brain Tumor Patients using MRI" *2017 IEEE 17th International Conference on Bioinformatics and Bioengineering*

[10] https://www.edureka.co/blog/deeplearning tutorial?utm_source=youtube&utm_campaign =deep-learning-180717-wr&utm_medium=description

[11]
https://www.mathworks.com/campaigns/offers/deelearning.html?s_eid=PSM_15

[12] Simonyan, Karen, and Andrew Zisserman. "Very deep convolutional networks for large-scale image recognition." arXiv preprint arXiv:1409.1556 (2014).

[13] Russakovsky, O., Deng, J., Su, H., et al. "ImageNet Large Scale Visual Recognition Challenge." *International Journal of Computer Vision (IJCV).* Vol 115, Issue 3, 2015, pp. 211–252

[14] Xiaobing Han, Yanfei Zhong, Liqin Cao 3,* and Liangpei Zhang, "Pre-Trained AlexNet Architecture with Pyramid Pooling and Supervision for High Spatial Resolution Remote Sensing Image Scene Classification".