International Journal of Engineering Technology and Management Sciences
Website: ijetms.in Issue: 2 Volume No.7 March - April – 2023
DOI:10.46647/ijetms.2023.v07i02.011 ISSN: 2581-4621

A Novel Brain Tumor Classification Model Using Machine Learning Techniques

Saikat Sundar Pal[1], Prithwish Raymahapatra[2], Soumyadeep Paul[3], Sajal Dolui[4], Dr. Avijit Kumar Chaudhuri[5], Sulekha Das[6]
[1,2,3,4] UG - Computer Science and Engineering, Techno Engineering College Banipur, Habra, Kolkata
[5,6]Assistant Professor, Computer Science and Engineering, Techno Engineering College Banipur, Habra, Kolkata
[1]0000-0002-1042-622X, [2]0000-0001-7147-9491, [3]0000-0003-2188-4469, [4]0000-0002-1005-9765, [5]0000-0002-5310-3180,[6]0000-0002-6641-3268

Abstract
The objective of this research work is to classify brain tumor images into 4 different classes by using Convolutional Neural Network (CNN) algorithm i.e. a deep learning method with VGG16 architecture. The four classes are pituitary, glioma, meningioma, and no tumor. The dataset used for this research is a publicly available MRI Image dataset of brain tumor with 7023 images. The methodology followed in this project includes data pre-processing, model building, and evaluation. The dataset is pre-processed by resizing the images to 64x64 and normalizing the pixel values. The VGG16 architecture is used to build the CNN model, and it is trained on the pre-processed data for 10 epochs with a batch size of 64. The model is evaluated using the area under the operating characteristic curve (AUC) metric of the receiver. The results of this project show that the CNN model with VGG16 architecture achieves an AUC of 0.92 for classifying brain tumor images into four different classes. The model performs best for classifying meningioma with AUC of 0.90, followed by pituitary with AUC of 0.91, glioma with AUC of 0.93, and no tumor with AUC of 0.89. In conclusion, the CNN model with VGG16 architecture is an effective approach for classifying brain tumor images into multiple classes. The model achieves high accuracy in identifying different types of brain tumors, which could potentially aid in early diagnosis and treatment of brain tumors.
Keywords - CNN, VGG16, AUC, Brain Tumor

Introduction:

Brain tumor image classification using convolutional neural network (CNN) is a challenging task in medical image analysis. In this project, we will use the VGG16 model, which is a widely used deep learning architecture for image classification tasks. The goal of this project is to classify brain tumor images into four different classes: pituitary, glioma, meningioma, and no tumor. We will use the Area Under the Operating Characteristic Curve (AUC) of the receiver as the evaluation metric for our model. AUC, or Area Under the Curve, is an important metric used in machine learning(ML) to evaluate the performances of binary classification models. It measures the performance of the given model by distinguishing between positive and negative samples.

In a problem that is binary classified, the model predicts a probability score for each sample, and the AUC represents the probability that the model will rank a positive sample that is randomly chosen higher than a negative sample that is also randomly chosen. The AUC ranges from 0 to 1, where an AUC of 0.5 indicates that the model is not better than random guessing, while an AUC whose value is 1, indicates a perfect model. AUC is a popular metric because it is more robust than accuracy when dealing with imbalanced datasets. It also provides a comprehensive evaluation of the model's performance, taking into account both false positive and false negative rates.

Furthermore, AUC is useful for comparing the performance of different models, as it is independent of the classification threshold used to make predictions. This allows for a fair comparison of models even if they have different thresholds.

International Journal of Engineering Technology and Management Sciences
Website: ijetms.in Issue: 2 Volume No.7 March - April – 2023
DOI:10.46647/ijetms.2023.v07i02.011 ISSN: 2581-4621

Overall, AUC is a valuable metric in machine learning that provides insight into the performance of models that are binary classified and can aid in selecting the best model for a given problem.
To accomplish this, we will first pre-process the dataset, which consists of brain tumor images. We will then train the VGG16 model using transfer learning with pre-trained weights on ImageNet. Next, we will fine-tune the model on our dataset, followed by evaluating the model's performance on the test set using the AUC metric. The final output of our project will be a model that can accurately classify brain tumor images into one of the four classes with high accuracy and AUC. This model can be used as an effective tool for early diagnosis and treatment of brain tumors.

Literature Review:

Author	Feature/methods	Performance
Machhale et al.[3]	SVM-KNN	Sensitivity: 100% Specificity: 93.75% Accuracy: 98%
Zacharaki et al.[4]	Cross-validation using different classifiers (LDA,k-NN,SVM)	Sensitivity: 75% Specificity: 100% Accuracy: 96.4%
Pan et al.[5]	Segmentation results	Sensitivity: 85% Specificity: 88% Accuracy: 80%
Afshar et al.[6]	Capsule network method	Accuracy: 86.56%
Zia et al.[7]	Window based image cropping	Sensitivity: 86.26% Specificity: 90.90% Accuracy: 85.69%
Sajjad et al.[8]	CNN with data augmentation	Sensitivity: 88.41% Specificity: 96.12% Accuracy: 94.58%
Badža and Barjaktarovic[9]	CNN	Accuracy: 95.40%
Cheng et al.[10]	Feature extraction: Intensity, histogram, GLCM, BOW, classification Methods: SVM, SRC, KNN	Accuracy: 91.28%
Paul et al.[11]	CNN	Accuracy: 84.19%
Huang et al.[12]	convolutional neural network based on complex networks (CNNBCN)	Accuracy: 95.49%

International Journal of Engineering Technology and Management Sciences
Website: ijetms.in Issue: 2 Volume No.7 March - April – 2023
DOI:10.46647/ijetms.2023.v07i02.011 ISSN: 2581-4621

Methodology:

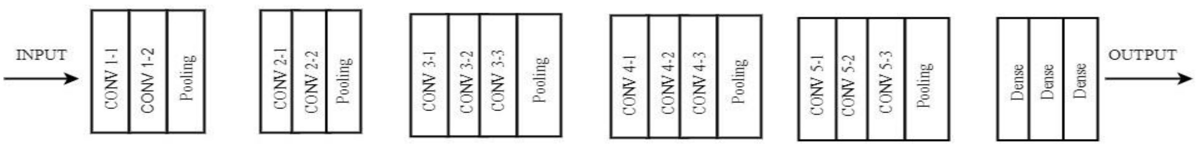
- Data: The first step of this whole research work was selecting the dataset. In this case, we have chosen a dataset which is a mixture of figshare, SARTAJ and Br35H dataset. This dataset contains 7023 MRI images of the human brain, classified into 4 classes: glioma - meningioma - no tumor and pituitary. The particular reason behind working with this dataset is that this dataset consists of a lot of sample images and a lot of research works has already been done with this dataset and got remarkable results. The dataset was divided in two parts: training set and testing set. The training set contains a total of 5712 images and the testing set consists of 1311 images. Each of the two consists of all the 4 classes i.e. glioma, meningioma, no tumor and pituitary.

Table 1

	Training Set	Testing Set
glioma	1321	300
meningioma	1339	306
no tumor	1595	405
pituitary	1457	300

Research Method:

- Convolutional Neural Network: CNN[23] stands for Convolutional Neural Network, which is a type of deep neural network i.e. deep learning commonly used in image and video recognition and also to process any tasks. The key characteristic of CNN is that it has ability to automatically learn and extract features from the raw data, in this case, images or videos. These features are learned through a process of convolution, where the network applies a complete set of filters or kernels to the input image to identify patterns and structures in the data. The output of the convolutional layers is then passed through a series of pooling layers, which reduce the spatial size of the features and help to increase the network's ability to generalize to new images. After the pooling layers, the resulting features are flattened into a vector and fed into fully connected layer, where the network can make predictions based on the learned features. CNNs have proven to be highly effective in a range of computer vision tasks, including classification of the image, the detection of objects, and segmentation, and have achieved state-of-the-art performance on many benchmark datasets.
- VGG16: VGG16 [22] is a convolutional neural network (CNN) architecture used to win the 2014 ILSVR (ImageNet) competition. Today it is considered one of the excellent machine vision model architectures. The great feature of VGG16 is that it avoids lots of hyper parameters, we focused on 3x3 filter convolution layers in step 1, and always used the same padding and max pool layers of 2x2 filters in step 2. Maximum pool layers consistently throughout the architecture. In the end, there are 2 Fully Connected Layers (FCs) and a softmax for the output layer. The 16 in VGG16 means there are 16 layers with weights. This network is quite large and has about 138 million parameters.



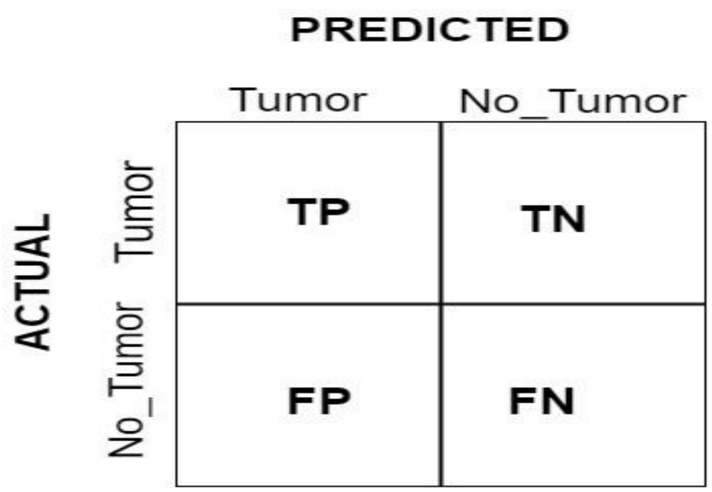
VGG-16

DOI:10.46647/ijetms.2023.v07i02.011 ISSN: 2581-4621

From the given images, it's clear that in between input and output layering part, there are total 16 numbers of layers. This layers generate the desired output of a given input.

- ImageNet: The ImageNet [24] weights for VGG16 are pre-trained weights that have been learned on the large-scale ImageNet dataset. These weights are often used as starting point for transfer learning in computer vision tasks. The size of the ImageNet weights for VGG16 is approximately 528 MB. This includes the weights for all the layers in the network, as well as the biases for the fully connected layers.
- ImageDataGenerator: In Keras, the ImageDataGenerator[25] class is used for image generation and data augmentation. This class provides set of functions for pre-processing and data augmentation on the input images. It generates batches of tensor image data using real-time data augmentation. This allows you to train deep learning models on a large dataset without having to load all the images into memory at once. Instead, the ImageDataGenerator loads the images in batches and applies various image transformations on the fly.

PRIMARY WORK: The first step of this whole research work was selecting the dataset. In this case, we have chosen a dataset which is a combined form of figshare, SARTAJ and Br35H dataset. This dataset contains 7023 MRI images of the human brain, classified into 4 classes: glioma - meningioma - no tumor and pituitary. The particular reason behind working with this dataset is that this dataset has a lot of sample images and a lot of research works has already been done with this dataset and got remarkable results.
After selection of the dataset, we have used the VGG16 model that came out in 2014 which is one of the best CNN models available right now and is used in many classification models over other models like AlexNet which are less discriminative.
Post training the model over the dataset, we tested it over the testing set and got remarkable results with the classifications. The various parameters of measuring the performance i.e. accuracy, recall, precision, specificity, F1-score and AUC of this research are depicted later.
Confusion Matrix: Confusion matrix [17] i.e. also called as error matrix, is one type of matrix or a table where we put the results of the MLR model i.e. the test data. Confusion matrix is the shortest way to see and understand the result of the model. In confusion matrix there are total four variables as – TP, TN, FP, FN. TP stands for 'true positive' that shows the total number of positive data classified accurately. TN stands for 'true negative' that shows the total number of negative data classified accurately. FP stands for 'false positive' which indicates the real value is negative but predicted as positive. FP is called as TYPE 1 ERROR. FN stands for 'false negative' which indicates the real value is positive but predicted as negative. FN is also called as TYPE 2 ERROR.



- Accuracy**:** In any model, it represents the ratio of number of times the model is able to make the correct prediction with the total number of predictions.

International Journal of Engineering Technology and Management Sciences
Website: ijetms.in Issue: 2 Volume No.7 March - April – 2023
DOI:10.46647/ijetms.2023.v07i02.011 ISSN: 2581-4621

- Sensitivity**:** We defined it as the ratio of number of times a model is able to make the positive prediction to the total number of correct predictions.
- Specificity**:** We defined it as the ratio of number of times a model is able predict that the result will be negative to the total number of times it has made the correct prediction.
- Precision**:** Precision is the method in which way one can say how correctly predicted cases actually turned positive.
- Recall: Recall is calculated as the ratio of the number of positive samples correctly classified as positive to the total number of positive samples. Recall measures the ability of a model to detect positive samples. The higher the recall, the more positive samples are found.
- F1_Score: F1 score is the measurement of accuracy and it is the harmonic mean of precision and recall. Its maximum value can be 1 and minimum value can be 0.
- AUC & ROC: AUC [26] stands for Area Under the ROC Curve, which is a popular evaluation metric in machine learning for binary classification problems. The ROC (Receiver Operating Characteristic) curve is a graphical representation of the performance of a binary classifier, and the AOC measures the area under this curve. In a problem that is binary classified, the classifier tries to predict whether an input belongs to a positive or negative class. The ROC curve plots the true positive rate (TPR) against the false positive rate (FPR) for different classification thresholds. The TPR is the ratio of correctly predicted positive samples to the total number of actual positive samples, and the FPR is the ratio of incorrectly predicted positive samples with the total number of actual negative samples. The AOC ranges from 0 to 1, with higher values indicating better performance. A perfect classifier would have an AOC of 1, while a completely random classifier would have an AOC of 0.5. The AOC is a useful evaluation metric because it takes into account all possible classification thresholds and provides a single number to compare the performance of different classifiers. However, it should be noted that the AOC only measures the overall performance of a classifier, and other metrics such as precision, F1 score, and recall may be more appropriate depending on the specific problem and application.

DEVELOPING EQUATION OF CONFUSION MATRIX:
Let's take-
TP= TRUE POSITIVE
TN= TRUE NEGATIVE
FP= FALSE POSITIVE
FN= FALSE NEGATIVE
FPR= FALSE POSITIVE RATE
Now,

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

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

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

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

International Journal of Engineering Technology and Management Sciences
Website: ijetms.in Issue: 2 Volume No.7 March - April – 2023
DOI:10.46647/ijetms.2023.v07i02.011 ISSN: 2581-4621

$$\text{Recall} = \frac{\text{TP}}{\text{TP+FN}}$$

$$FPR = \frac{FP}{TN + FP}$$

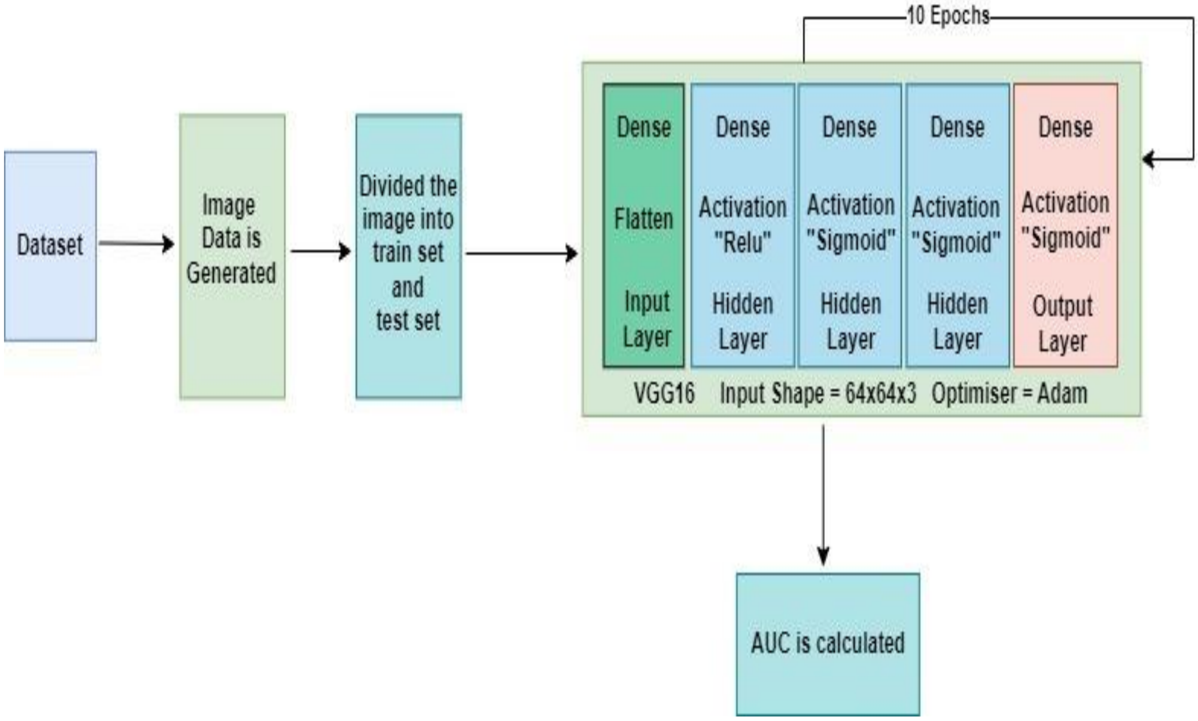
$$\text{F1\_Score} = \frac{2 * \text{Recall} * \text{Precision}}{\text{Recall + Precision}}$$

$$\text{auc}=\frac{1}{2}-\frac{\text{FPR}}{2}+\frac{\text{recall}}{2}$$

Procedure:

- Define the model architecture using the VGG16 pre-trained model as a base and add new classifier layers on top.
- Load the pre-trained weights for the VGG16 model.
- Freeze all the layers of the VGG16 model to prevent them from being updated during training.
- Add new fully connected classifier layers with appropriate activation functions and kernel initializers.
- Compile the model with appropriate optimizer and loss function, and evaluate using relevant metrics like accuracy, precision, recall, AUC, and F1 score.
- Augment the data using ImageDataGenerator to increase the size of the training dataset.
- Fit the model to the augmented data and evaluate the model on the test data.
- Calculate and print relevant metrics like accuracy, precision, recall, specificity, and F1 score for the test dataset.
- Calculate and print the AUC (Area under the Curve) score.
- Plot the diagnostic learning curves (loss and accuracy) for both training and validation data.

FLOWCHART:



International Journal of Engineering Technology and Management Sciences
Website: ijetms.in Issue: 2 Volume No.7 March - April – 2023
DOI:10.46647/ijetms.2023.v07i02.011 ISSN: 2581-4621

RESULTS AND DISCUSSION: After analysing this model we get the results that are given below.

Table 2-For EPOCHS-2

ATTRIBUTES	VALUE RANGE(%)
Accuracy	87.41
Recall	92.98
Specificity	70.64
Precision	90.50
F1_Score	91.0
AUC	81.66

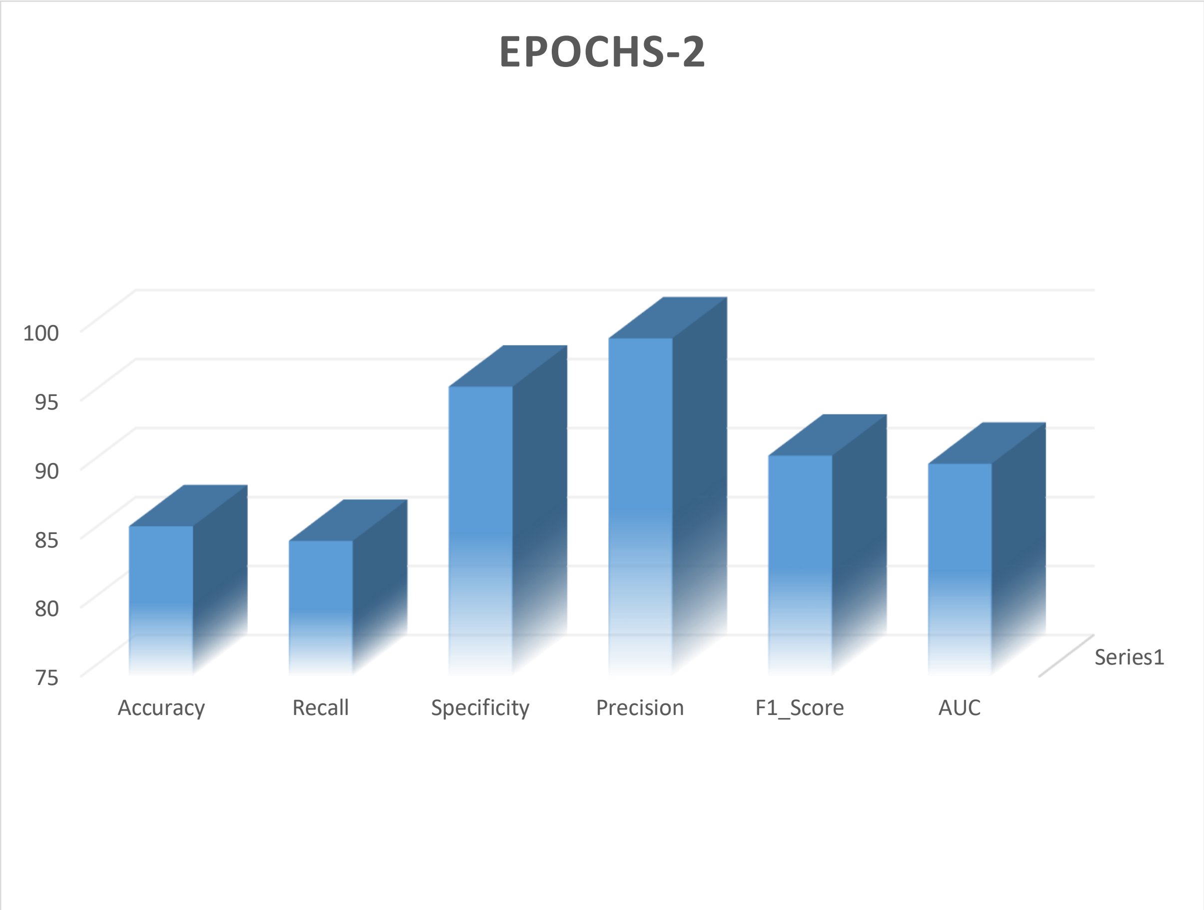


Table 3-For EPOCHS-5

ATTRIBUTES	VALUE RANGE(%)
Accuracy	91.60
Recall	91.98
Specificity	89.91
Precision	97.62
F1_Score	94.72
AUC	90.95

International Journal of Engineering Technology and Management Sciences
Website: ijetms.in Issue: 2 Volume No.7 March - April – 2023
DOI:10.46647/ijetms.2023.v07i02.011 ISSN: 2581-4621

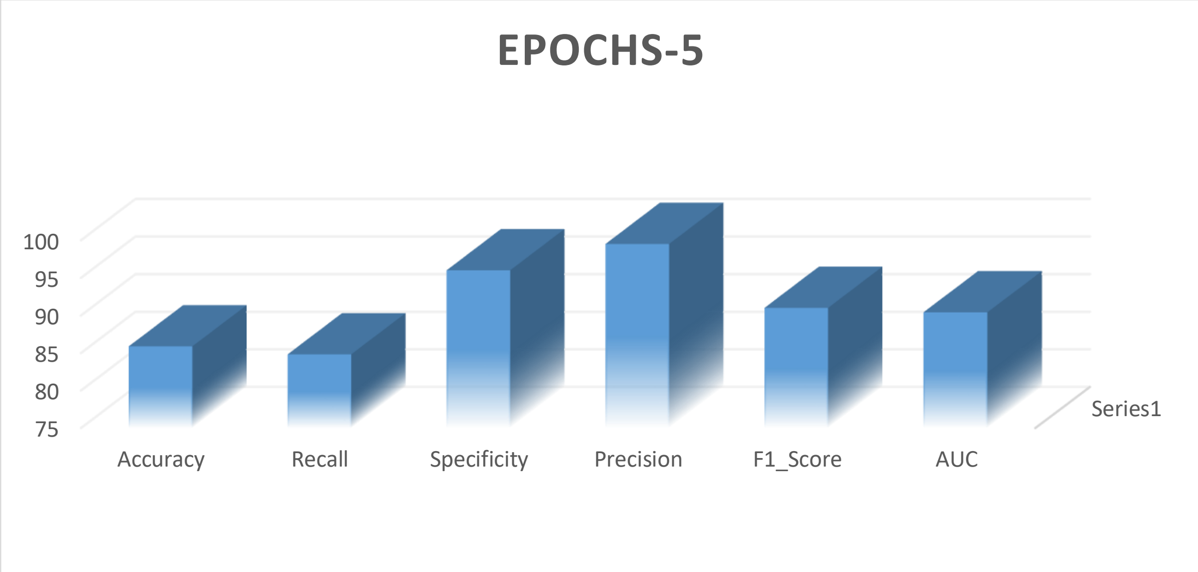


Table 4-For EPOCHS-9

ATTRIBUTES	VALUE RANGE(%)
Accuracy	91.60
Recall	91.98
Specificity	89.91
Precision	97.62
F1_Score	94.72
AUC	90.95

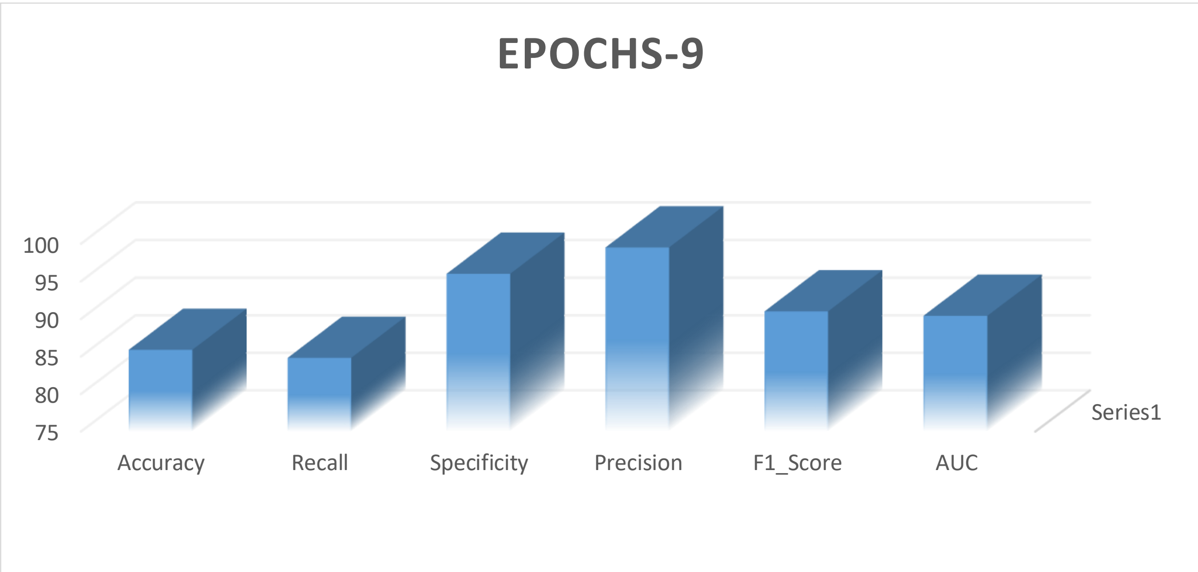


Table 5-For EPOCHS-10

ATTRIBUTES	VALUE RANGE(%)
Accuracy	92.75
Recall	92.96
Specificity	91.83
Precision	98.02
F1_Score	95.42

International Journal of Engineering Technology and Management Sciences
Website: ijetms.in Issue: 2 Volume No.7 March - April – 2023
DOI:10.46647/ijetms.2023.v07i02.011 ISSN: 2581-4621

AUC	92.40

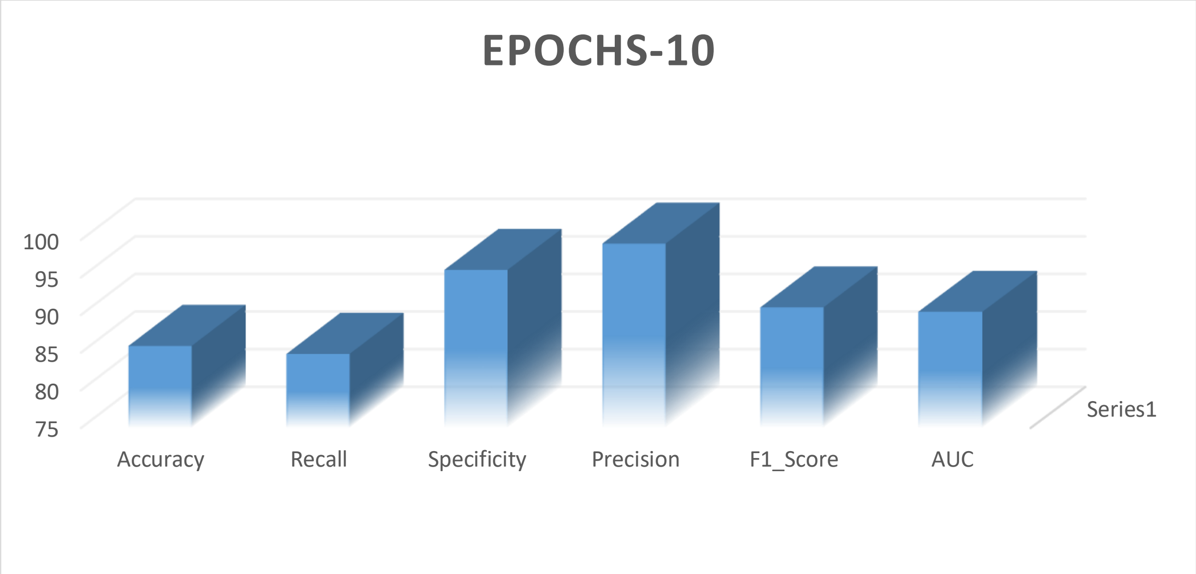
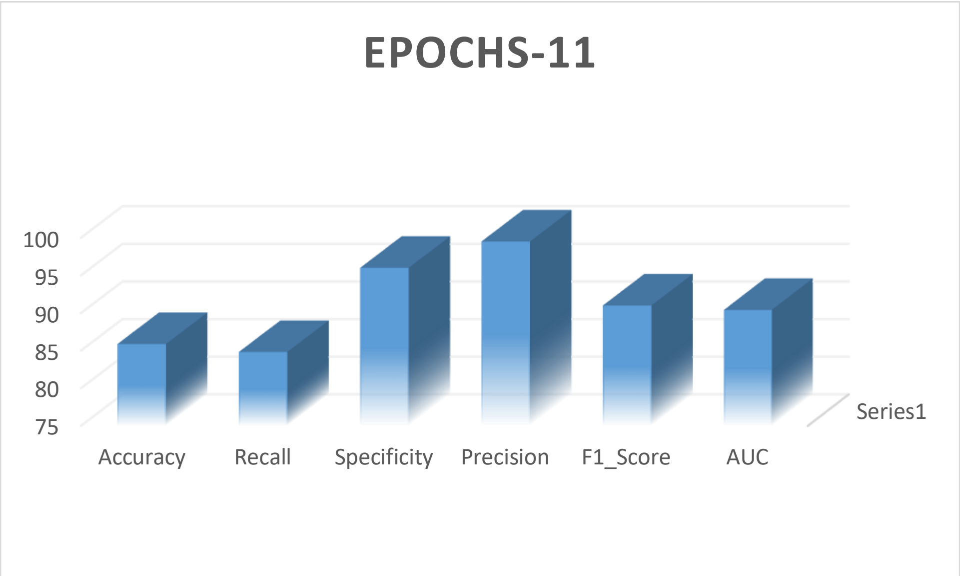


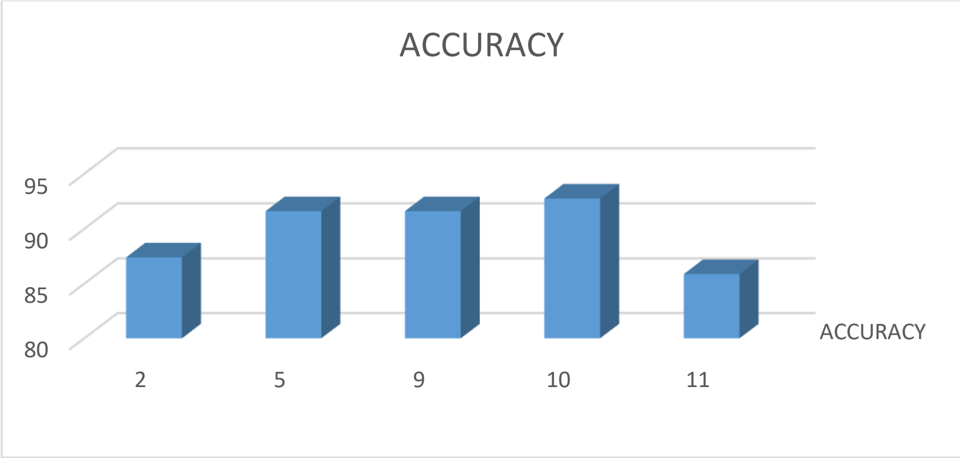
Table 6-For EPOCHS-11

ATTRIBUTES	VALUE RANGE(%)
Accuracy	85.88
Recall	84.82
Specificity	96.0
Precision	99.50
F1_Score	91.0
AUC	90.41

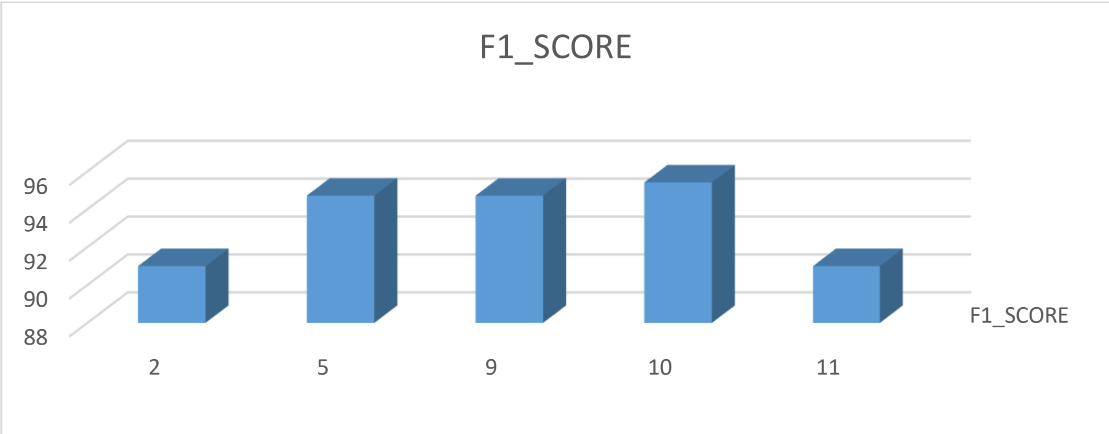


COMPARISON:
ACCURACY:

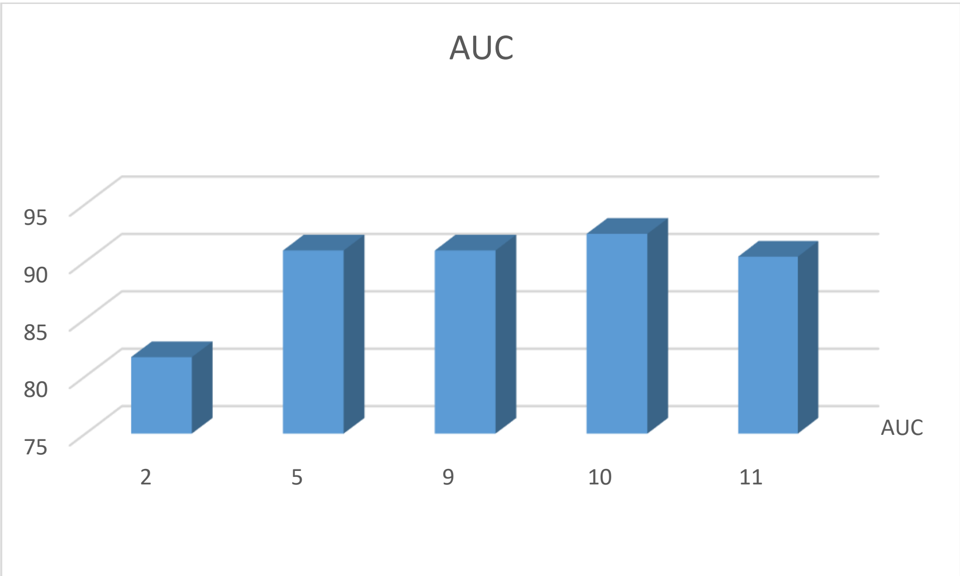
International Journal of Engineering Technology and Management Sciences
Website: ijetms.in Issue: 2 Volume No.7 March - April – 2023
DOI:10.46647/ijetms.2023.v07i02.011 ISSN: 2581-4621



F1_SCORE:



AUC:



CONCLUSION:

This article focuses on the identification and classification of different MRI images of brain tumor into its respective classes i.e. Meningioma, Glioma, Pituitary and No tumor by transfer learning approach using Convolutional Neural Network (CNN) as the working model with VGG16 architecture with sigmoid and relu activation function, and calculating the AUC of the model which depicts how efficiently the model is working and how accurately it is classifying those images. The AUC of this model is 0.92 which depicts that this model is highly efficient to classify those images.

DOI:10.46647/ijetms.2023.v07i02.011 ISSN: 2581-4621

The model achieves high accuracy in identifying different types of brain tumors, which could potentially aid in early diagnosis and treatment of brain tumors.

FUTURE SCOPE:

As the AUC of this model is very high so this model can be used in future for other disease dataset and also other dataset. In future we will collect data from various nursing homes and hospitals and will train this model on the same.

REFERENCES:

[1]https://www.kaggle.com/datasets/masoudnickparvar/brain-tumor-mri-dataset?select=Training
[2]https://www.researchgate.net/publication/346391164_Study_on_image_processing_using_deep_learning_techniques
[3] Machhale K, Nandpuru HB, Kapur V, Kosta L. MRI brain cancer classification using hybrid classifier (SVM-KNN). In: 2015 International Conference on Industrial Instrumentation and Control (ICIC). IEEE; 2015, pp. 60–65
[4] Zacharaki EI, Wang S, Chawla S, et al. Classification of brain tumor type and grade using MRI texture and shape in a machine learning scheme. Magn Reson Med. 2009;62:1609- 1618. https://doi.org/10.1002/mrm.22147
[5] Zacharaki EI, Wang S, Chawla S, et al. Classification of brain tumor type and grade using MRI texture and shape in a machine learning scheme. Magn Reson Med. 2009;62:1609- 1618. https://doi.org/10.1002/mrm.22147
[6] Afshar P, Mohammadi A, Plataniotis KN. Brain tumor type classification via capsule networks. In: 2018 25th IEEE International Conference on Image Processing (ICIP). IEEE; 2018, pp. 3129–3133.
[7] Zia R, Akhtar P, Aziz A. A new rectangular window based image cropping method for generalization of brain neoplasm classification systems. Int J Imaging Syst Technol. 2018;28:153-162. https://doi.org/10.1002/ima.22266
[8] Sajjad M, Khan S, Muhammad K, et al. Multi-grade brain tumor classification using deep CNN with extensive data augmentation. J Comput Sci. 2019;30:174-182. https://doi.org/10.1016/j.jocs.2018.12.003
[9] Badža MM, Barjaktarovic MČ. Classification of brain tumors from MRI images using a convolutional neural network. Appl Sci. 2020;10:1999. https://doi.org/10.3390/app10061999
[10] Cheng J, Huang W, Cao S, et al. Enhanced performance of brain tumor classification via tumor region augmentation and partition. PLoS One. 2015;10:e0140381. https://doi.org/10.1371/journal.pone.0140381
[11] Paul JS, Plassard AJ, Landman BA, Fabbri D. Deep learning for brain tumor classification. Proc SPIE. 2017;10137: 1013710
[12] Huang Z, Du X, Chen L, et al. Convolutional neural network based on complex networks for brain tumor image classification with a modified activation function. IEEE Access. 2020;8:89281-89290. https://doi.org/10.1109/ACCE SS.2020.2993618
[13] Afshar P, Plataniotis KN, Mohammadi A. Capsule networks for brain tumor classification based on MRI images and coarse tumor boundaries. In: ICASSP 2019 – 2019 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP). IEEE; 2019, pp. 1368–1372.
[14] Hussain L, Awan IA, Aziz W, et al. Detecting congestive heart failure by extracting multimodal features and employing machine learning techniques. Biomed Res Int. 2020;2020:1-19. https://doi.org/10.1155/2020/4281243
[15] Hussain L, Saeed S, Awan IA, et al. Detecting brain tumor using machine learning techniques based on different features extracting strategies. Curr Med Imaging Rev. 2018;14:595-606. https://doi.org/10.2174/1573405614666180718123533
[16] Fazli S, Pour HM, Bouzari H. Particle filter based object tracking with sift and color feature. 2009 2nd Int Conf Mach Vision, ICMV. 2009;2009:89-93. https://doi.org/10.1109/ICMV. 2009.47

[17]Pal, S. S., Paul, S., Dey, R., Das, S., & Chaudhuri, A. K. Determining the probability of poverty levels of the Indigenous Americans and Black Americans in US using Multiple Regression.
[18]Bialynicky-Birula I, Mycielski J. Relations for information entropy in wave mechanics. Commun Math Phys. 1975;44: 129-132
[19] Chen T, Xu R, He Y, Wang X. Improving sentiment analysis via sentence type classification using BiLSTM-CRF and CNN. Expert Syst Appl. 2017;72:221-230. https://doi.org/10.1016/j.eswa.2016.10.065
[20] He K, Zhang X, Ren S, Sun J. Deep residual learning for image recognition. In: IEEE Conference on Computer Vision and Pattern Recognition; 2016.
[21]Ray, A., & Chaudhuri, A. K. (2021). Smart healthcare disease diagnosis and patient management: Innovation, improvement and skill development. Machine Learning with Applications, 3, 100011.
[22] Qassim, H., Verma, A., & Feinzimer, D. (2018, January). Compressed residual-VGG16 CNN model for big data places image recognition. In *2018 IEEE 8th annual computing and communication workshop and conference (CCWC)* (pp. 169-175). IEEE.
[23] Girshick, R. (2015). Fast r-cnn. In *Proceedings of the IEEE international conference on computer vision* (pp. 1440-1448).
[24] Krizhevsky, A., Sutskever, I., & Hinton, G. E. (2017). Imagenet classification with deep convolutional neural networks. *Communications of the ACM*, *60*(6), 84-90.
[25] Shorten, C., & Khoshgoftaar, T. M. (2019). A survey on image data augmentation for deep learning. *Journal of big data*, *6*(1), 1-48.
[26] Lobo, J. M., Jiménez-Valverde, A., & Real, R. (2008). AUC: a misleading measure of the performance of predictive distribution models. *Global ecology and Biogeography*, *17*(2), 145-151.
[27] Saha, S., Mondal, J., Arnam Ghosh, M., Das, S., & Chaudhuri, A. K. Prediction on the Combine Effect of Population, Education and Unemployment on Criminal Activity Using Machine Learning.
[28] Jalal deen K; Karthigai Priya G; Magesh B; Kubendran R. "Identification of CT Lung Tumor Using Fuzzy Clustering Algorithm". International Research Journal on Advanced Science Hub, 3, Special Issue ICEST 1S, 2021, 30-33. doi: 10.47392/irjash.2021.016
[29] Hari Prasada Raju Kunadharaju; Sandhya N.; Raghav Mehra. "Detection of Brain Tumor Using Unsupervised Enhanced K-Means, PCA and Supervised SVM Machine Learning Algorithms". International Research Journal on Advanced Science Hub, 2, Special Issue ICSTM 12S, 2020, 62-67. doi: 10.47392/irjash.2020.262
[30] Naveenkumar C; Praveenkumar M; Praveenkumar S; Karthik V; Parthasarathy .. "Study on the Effect different materials in Traffic signal pole using numerical simulation approach". International Research Journal on Advanced Science Hub, 1, 1, 2019, 33-41. doi: 10.47392/irjash.2019.06
[31] Gladia Angeline P; Jagadesh S; Jeba Christina D; Nevetha G; Poornima L. "Efficient and Enhanced Data Encryption in Cloud Computing". International Research Journal on Advanced Science Hub, 2, 3, 2020, 1-4. doi: 10.47392/irjash.2020.16