

Brain Tumor Detection Using Convolutional Neural Network with Dense Connections

Para Rajesh

Research Scholar

Department of Computer Science and Engg.,
Faculty of Engineering and Technology,
Annamalai University, Chidambaram,
Technology, Annamalai Nagar - 608002,
Tamil Nadu, India.

Dr. A.Punitha

Department of Computer Science and Engg.
Faculty of Engineering and Technology,
Annamalai University, Chidambaram,
Annamalai Nagar - 608002,
Tamil Nadu, India.

P.Chandra Sekhar Reddy

Computer Science and Engg. Dept.,
Gokaraju Rangaraju Institute of
Engineering and Technology
Hyderabad-500090,
Telangana, India

Abstract- Radiologists examine various Magnetic Resonant Image (MRI) sequences produced using a multimodal imaging approach when treating a brain tumor. Research in neuro-oncology has recently concentrated on cutting-edge MRI methods that attempt to link histological characteristics to radiological phenotypes like cellularity or vascularity. The fundamental imaging procedure used in clinical practice today consists of T1-weighted sequences, which show anatomy, and T2-weighted sequences, which show oedema and can evaluate cellularity. Radiologists continue to attempt to diagnose a suspected glioma and a brain lesion on an MRI using the histological subtype. However, a significant problem with this diagnostic is that it is quite subjective. They only base the determination on qualitative, subjective characteristics such "low-to-moderate oedema could be reflecting the features of tumour," as accurate, quantitative thresholds have not yet been established. The diagnostic accuracy for finding the tumor is still a challenge; it is now moderate to low. A few years ago, AI made its way into neuro-oncology. AI algorithms can be utilized to segment and identify the subtypes of tumors in MR images in addition to detecting their presence. The goal of this study is to identify brain tumors using the Convolutional Neural Network (CNN) variation DenseNet. Under various architectural patterns, the performance of the pre-trained DenseNet and ResNet models were compared.

Keyword: glioma, Magnetic Resonant Image, Convolutional Neural Network, DenseNet, ResNet, brain tumor.

Introduction

The tumor is an uncontrolled proliferation of undesirable tissue that has formed a fibrous mesh inside of our brain. Radiologists frequently employ magnetic resonance imaging (MRI) to examine brain malignancies. Nowadays, manually examining the enormous volume of MRI (magnetic resonance imaging) images and identifying a brain tumor is an extremely time-consuming and

imprecise task. It might have an impact on the patient's appropriate medical care. Again, because there are so many image datasets involved, it can take a very long time. The appearance of brain tumor cells and normal tissue is quite similar, making it challenging to identify a tumor. A very accurate automatic tumor detection system is therefore necessary. By utilizing deep learning techniques [3, 4], the experiments conducted in

this paper determines if the brain is healthy or diseased. In this study, a pre-trained convolutional neural network method for detecting brain tumors using 2D MRIs of the brain was developed. This method might be followed by segmentation and classification models for tumor localization and subtype categorization. To properly train the model, this study uses a variety of MR images with variable tumor sizes, locations, forms, and image intensities. The performance of the pre-trained model is also assessed by contrasting it with models having residual connections.

A recently proposed CNN architecture called DenseNet [1, 2] has an intriguing connectivity pattern where each layer is connected to every other layer inside a dense block. All layers in this scenario have access to the feature maps from the layers that came before them, which promotes extensive feature reuse. Because of this, the model is smaller and less prone to overfitting. In addition, the shortcut paths give the loss function direct supervision over each individual layer, resulting in implicit deep supervision. DenseNet is a perfect fit for per-pixel prediction analysis with all these favorable characteristics. Performance is improved when dense blocks are used in place of convolutions in the architecture. When traditional CNNs delve deeper, issues start to appear. This is because the distance between the input and output layers gets so great that information can vanish before it gets there. The interaction pattern between layers proposed in conventional architecture is made simpler by DenseNets. A huge dataset of T1-weighted contrast-enhanced MRI images that have undergone preprocessing is used to train the DenseNet models. The DenseNet model's internal architecture is modified in terms of the number of classification layers, the way the layers are arranged next to one another, the hyperparameters, and the values are optimized to fine-tune the model's performance and improve its accuracy. The number of MRI slices is increased using a data augmentation approach in this study, and the overfitting problem is resolved by adding dropout layers to the models.

Related Work

MRI is the most popular because it offers radiation-free, higher resolution images. Radiologists can diagnose brain disorders using medical imaging data obtained through the non-invasive MRI method [8], [9]. The Computer-Aided Diagnosis (CAD) approach, on the other hand, was created to find brain cancers in their early stages without the need for human participation. Based on MRI pictures, CAD systems can generate diagnostic reports and give the radiologist advice [10].

Applications of machine learning (ML) and deep learning (DL) in the field of medical imaging have significantly enhanced the CAD process [11 - 13]. These methods improve the CAD system's ability to accurately detect brain cancers. The principles of extraction of features, selection of features, and classification form the foundation of machine learning techniques. The tumor region of the human skull is segmented using a variety of feature extraction techniques, such as thresholding-based, clustering-based, contour-based, and texture-based [14]. These methods extract features from MRI images, choosing just the most significant features through a feature selection process. High accuracy can be attained by extracting features with considerable discriminatory information [15]. However, it is easy to delete essential details from the original image via features extraction [16].

However, DL techniques deal with this problem by taking the original image as input [17]. They don't need handcrafted traits, in other words, to be classified. Convolutional Neural Network (CNN) is one of the DL models that offer a variety of convolution layers [18] that are capable of extracting features from the images [19]. When given a large dataset, which is sometimes difficult to come by in the field of medical imaging, CNN fared well [20]. The use of transfer learning is one strategy to handle this problem. A model that has previously been trained using a sizable dataset relevant to a different domain is utilized for classification in transfer learning [21, 22]. Such information aids the model in producing accurate results on a modest dataset [23].

CNN offers a segmentation-free approach that does away with the necessity for manually created

feature extractor methods. Numerous academics have suggested various CNN architectures as a result. A significant amount of picture data, including multiclass brain tumor detection, was reported by the majority of CNN models. The work proposed in [24] used a CNN model with 16 layers [24]. On two datasets that are openly available, the CNN model was tested. The tumor area was extracted from the MRI images using the Fuzzy C-Means clustering technique proposed in [25]. They contrasted it with six existing machine learning models and presented a new CNN-based approach. The reported prediction accuracy of 97.9% performs better than previous models.

For a huge image collection, CNN seems to be performing well. However, it also has the following two major drawbacks:

- For training, the CNN model needed a sizable amount of images, which is frequently challenging in the medical imaging area.
- Convolutional Neural Networks (CNN) can classify images that are very similar to the dataset very effectively. CNNs, on the other hand, have trouble classifying images that are slightly rotated or tilted. As new image variants were added during training in this study through data augmentation, this issue was addressed.

This method uses a pre-trained model with modified weights for classification tasks. Another advantage is that since only the fully linked layers of the model need to be trained, it does not necessitate a significant amount of computer resources. Different transfer learning models have been employed to diagnose brain cancers because of these benefits. For instance, Talo et al. identified normal and abnormal brain MRI images using a pre-trained ResNet34 model. To achieve high prediction accuracy, a significant amount of data augmentation is also conducted [26].

Methodology

The overall process flow adopted in this research for detecting the presence of cancer in the MR images is presented in Fig. 1. Initially MR images are preprocessed to remove the unwanted noise and enhance the quality of the images. Wiener filter minimizes the overall mean square error by the process of inverse filtering and noise smoothing. Later the denoised image contrast is improved using histogram equalization. The contrast enhanced MRI images are then fed to a trained binary classifier for detecting the presence of tumor.

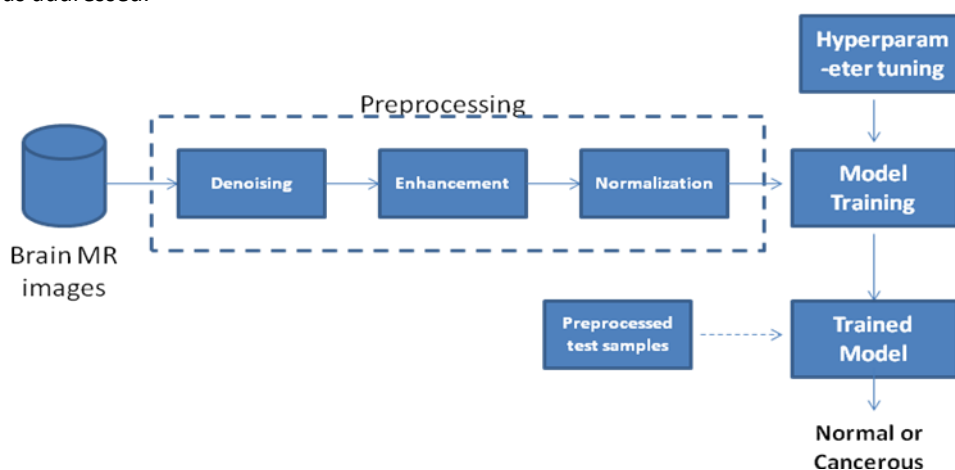


Fig. 1 Block Diagram of Brain Tumor Detection using Pre-trained Model

The MR image samples were extracted from three different sources; figshare, SARTAJ dataset, Br35H. The dataset contains 9523 images of human brain MRI images belonging to any one of the following classes: glioma, meningioma, no tumor, and pituitary. No tumor class images were taken from the Br35H dataset. For training the

binary classifier no tumor images are labeled as '0' and other categories are labeled as '1' with a train test split ratio was 80:20.

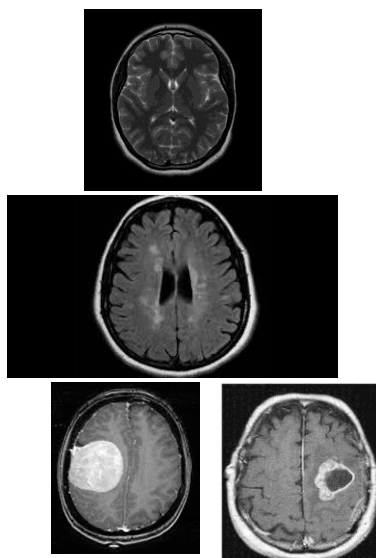


Fig. 2 Sample Images a) & b) Normal c) & d) Tumorous images

The Wiener filter is an optimal linear filter which seeks the linear time-invariant filter. Its output comes as close as possible to the original signal. In other words, the goal is to minimize the mean square error (MSE) between the expected noise-free signal and the actual output signal. The Wiener filter assumes that the input is the sum of valuable signals and noise. Both of which are generalized stationary processes. Their second-order statistical characteristics are known and hence it is not adaptive and always implemented

in the frequency domain. The weiner filter treats images and noises as a random processes and the aim is to estimate f' of the raw image in a way that the MSE is minimized which can be solved as optimization problem. Denoised image and the respective raw image was shown in Fig. 2. It is mathematically expressed as follows;

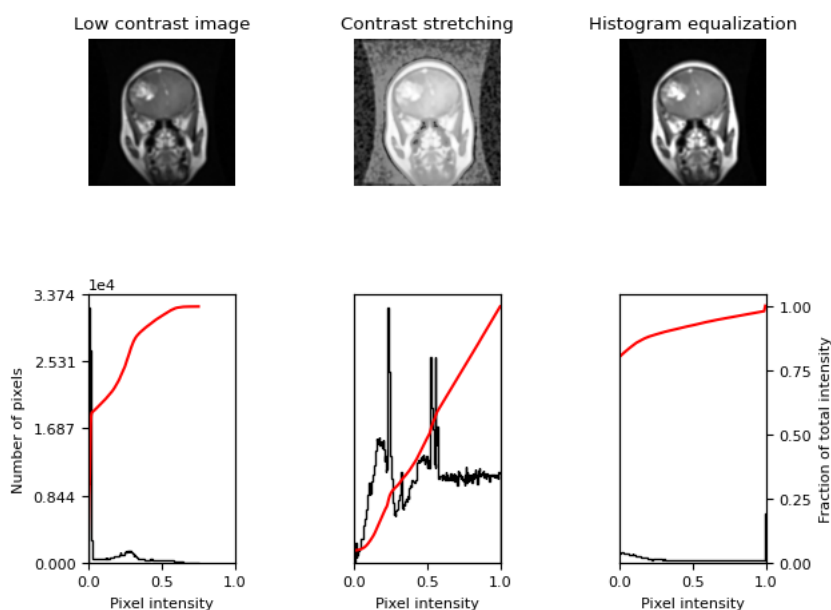
$$\min e^2 = E\{(f - f')^2\} \quad \text{Eq. 1}$$


Fig. 3 Contrast enhancement using Histogram Equalization

A crucial pre-processing step for effective computer aided diagnosis of MR images is contrast

enhancement. The main goal of image enhancement is to improve the quality of the

source image or, more specifically, to draw attention to a specific features of the image. A pre-processing method for enhancing low contrast photos is histogram equalization. To create a flattened histogram with a high contrast image, it increases the dynamic range of pixel values and equalizes all pixels [7].The probability distribution of the image's gray level values is the foundation of the theory supporting histogram equalization [8].The pixels of an image I are represented as a discrete function with a range of 0 to $L-1$. For an 8-bit image, L has a value of 256. The probability of appearance of gray scale r_k is

$$Pr(r_k) = \frac{n_k}{n} \quad k = 0,1,2 \dots \dots, L - 1 \quad \text{Eq. 2}$$

where r_k represents the k^{th} gray level, n_k represents the count of pixels with r_k as gray level, and n is total number of pixels in the image.

The objective of the histogram equalization is to obtain a uniform distribution of the gray levels in the image. $Ps(s)$ is a pdf of the image after equalization has been done. An equalized histogram, or $Ps(s)$, has a constant distribution across all conceivable values. The following transfer function changes $Pr(r)$ into $Ps(s)$.

$$s = T(r) = (L - 1) \int_0^r Pr(w) dw \quad \text{Eq. 3}$$

By differentiating Eq. 3 with respect to r

$$\begin{aligned} \frac{ds}{dr} &= \frac{d}{dr} T(r) \\ &= \frac{d}{dr} (L - 1) \int_0^r Pr(w) dw \\ &= (L - 1)Pr(r) \end{aligned} \quad \text{Eq. 4}$$

Relation between the probability distribution of raw and contrast enhanced image can be expressed as;

$$\begin{aligned} Ps(s) &= Pr(r) \left| \frac{ds}{dr} \right| \\ &= Pr(r) \left| \frac{1}{(L - 1)Pr(r)} \right| \\ &= \frac{1}{(L - 1)}; 0 \leq L - 1 \end{aligned} \quad \text{Eq. 5}$$

Therefore, $Ps(s)$ is a normalized distribution, and the transfer function given in Eq.3 can be used to equalize the histogram. The result of histogram equalization is shown in Fig. 3.

DenseNet Architecture

In DenseNet architecture every layer is connected with other directly which ensure maximum flow of information across layers. This helps to solve the problem of vanishing gradient in dense layered models. DenseNets utilize the capability of the network through feature reuse rather than obtaining representational power through very deep or broad architectures. As there is no requirement to learn redundant feature maps through this sort of connection, DenseNets require

a smaller number of parameters than a comparable traditional CNN. Additionally, several ResNets variations have shown that many layers can be deleted because they are seldom contributing. ResNets actually have a large number of parameters because each layer has weights that must be learned. Instead, DenseNets layers only add a tiny number of new feature-maps and are extremely constrained. Due to the previously noted information flow and gradients, very deep networks also had difficulty being trained. As every single layer has immediate exposure to the gradients from the error function and the input image, DenseNets are able to resolve this problem.

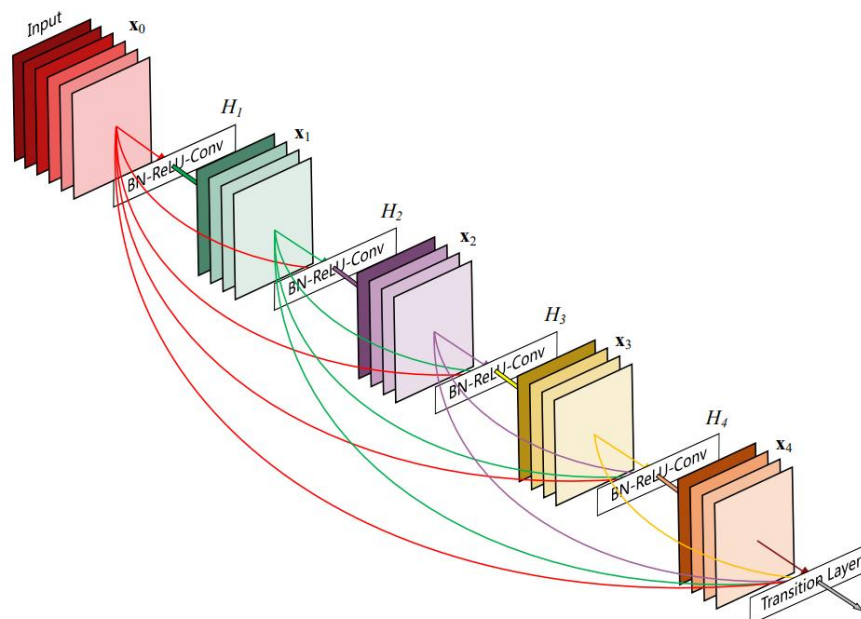


Fig. 4 Schematic view of connection pattern in DenseNet [5]

Following the application of a composite of operations, conventional feed-forward neural networks transfer the output of the layer to the following layer. Typically, this composite operation

$$x_l = H_l(x_{l-1}) \quad \text{Eq. 6}$$

Expanding this behavior by including the skip connection, ResNets reformulated this operation as follows;

$$x_l = H_l(x_{l-1}) + x_{l-1} \quad \text{Eq. 7}$$

DenseNets concatenate rather than sum the layer's output feature maps and incoming feature maps. As a result, the equation changes once again to;

$$x_l = H_l(x_0, x_1, \dots, x_{l-1}) \quad \text{Eq. 8}$$

DenseNets are divided into DenseBlocks, where the number of filters varies from block to block but

consists of an activation function, a batch normalization, and a convolution operation or pooling layers. This operation can be mathematically expressed as;

the dimensions of the feature maps remains constant inside each block. Transition Layers are the layers in between them that handle the downsampling by performing batch normalization, 1x1 convolution, and 2x2 pooling layers. Every layer of the feature maps' concatenation results in an increase in the channel dimension. If the H_l layer consistently generates k feature maps, it is possible to generalize this to the l -th layer as follows;

$$k_l = k_0 + k * (l - 1) \quad \text{Eq. 9}$$

Every layer has access to the collective knowledge and its previous feature maps. Then, as concrete k feature maps of information, each layer is adding fresh information to this body of knowledge. A very basic diagram of the DenseNet-121's architecture, which was used in this work, is shown in Fig. 5.

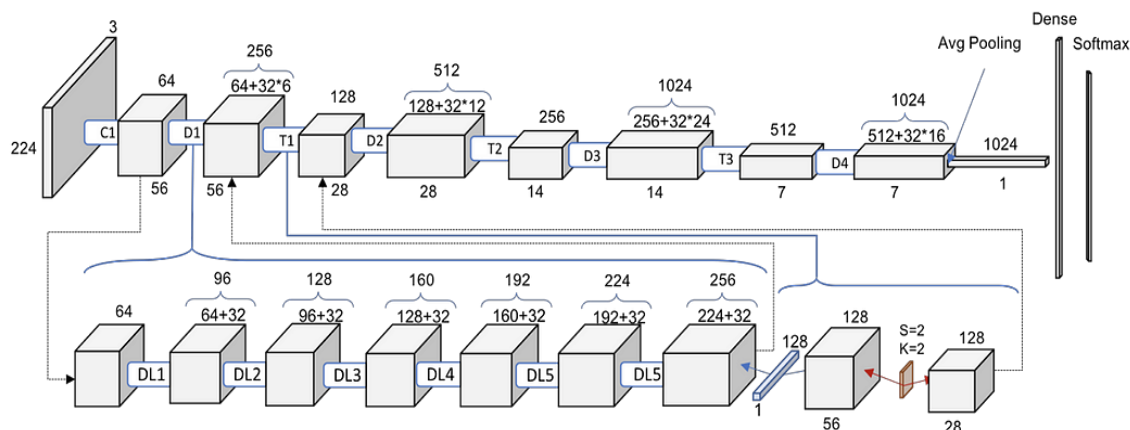


Fig. 5 Schematic view of Dense and Transition Blocks in a DenseNet-121 architecture [6]

With 128 filters, Transition Block functions as a 1x1 convolution. Following a 2x2 pooling with a stride of 2, the volume's size and the number of feature maps are divided in half.

Experiments And Results

A random search was utilized in this study's studies to determine the values of batch size and learning rate. The parameter values were then chosen as candidates from the results of the random search. The ideal range of values from the chosen candidates is determined by evaluating each of these values from the chosen candidates. The hyperparameter batch size and learning rate that will be employed during the model's training will presumably be determined based on the results of the random search. The datasets are trained with a growth rate of 12 utilizing the bottleneck layers and compression of the DenseNet model. The range of learning rates ranges from 0.001 to 0.1, while the range of batch sizes is from 16 to 128. The tests employ momentum of 0.9 and weight decay of 0.0001 and the model was trained for 50 epochs.

A reliable deep learning technique called transfer learning uses model parameters that were previously trained on a bigger collection of data. Additionally, pre-trained weights are used for transfer learning in order to learn more quickly and improve our model's accuracy. Only the fully connected layers are trained during the transfer learning process, with the CNN layers being frozen. As a result of the need to cope with brain MRI images, CNN layers were trained to improve model

performance. The weights were adjusted from the features of the pre-trained dataset to the specific features of the brain MRI image dataset. A rectified linear unit (ReLU) activation layer, batch normalization (BN), and convolution (convs) are all components of the dense block found in the DenseNet design. The final dense block is followed by a global average pooling layer, which is then fed into a Softmax classifier. Given that DenseNet contains L layers, there will be direct connections to $L(L + 1)/2$ levels.

A method for standardizing network inputs called batch normalization was used with either the activations of a preceding layer or the inputs itself. DenseNet splits into Dense blocks with the same size but different sets of filters. Batch normalization is applied via Transition Layer using downsampling. Finding the average for each segment of the feature map is part of average pooling. As a result, the average value of each square on the function map is down-sampled. In order to fine-tune the weights, the Cross Entropy loss function and an Adamax optimizer was used in the experiments. The Adam version of gradient descent is expanded with AdaMax, which generalizes the method to the infinite norm (max) and produces a more successful optimization.

Discussions

After training the model was able to classify the images it had not previously noticed. It perfectly differentiated between images with brain tumors and those with normal conditions. The trained model is capable of classifying the test images

since it is trained to provide the least amount of validation loss. The denoising step included in the pre-processing phase helped the model to achieve better results. The result of the denoising using weiner filter is shown in Fig. 6. Recall, F1-score, accuracy, and precision are used to assess the model. Each statistic for our proposed model's two

classes—brain tumor and normal-healthy—is identified. The model's performance was also contrasted with other optimization algorithms. Additionally, it is evaluated by comparison with other CNN architecture including ResNet and Inception.

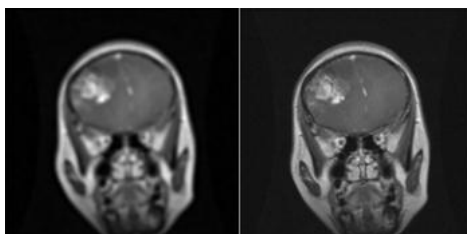
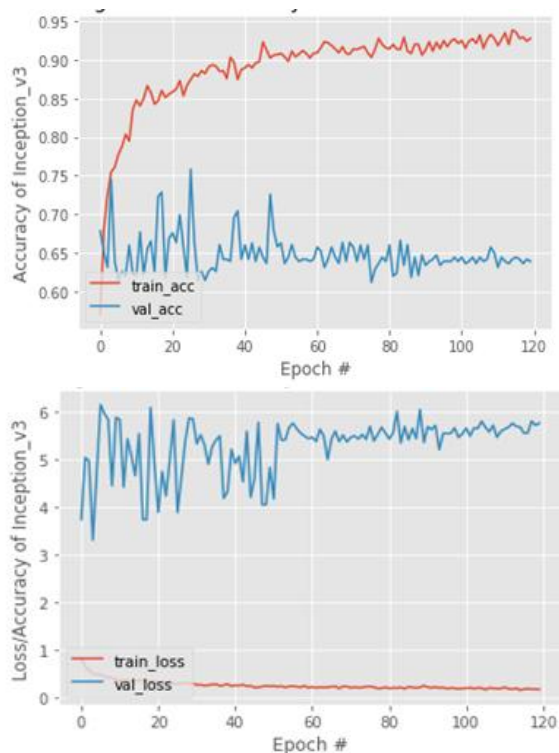


Fig. 6 Result after Denoising using weiner filter a. Raw image and b) Denoised image

The sample images that a model was trained on determine its training accuracy. Fig. 7 presents the performance of the models for each epoch. Here, the accuracy for each epoch was tabulated, and a final plot was then created. As a result, the Plot informs us that the model's accuracy rises together with the number of epochs. DenseNet model achieves the highest training accuracy (91.94%), which is superior to ResNet-50 and Inception V3 model accuracy levels of 86.3% and 76.2%, respectively. Inception model v3's

validation accuracy and training accuracy diverge significantly, and it may therefore be claimed that the model is overfitting regardless of the use of dropout, batch normalization, or regularization techniques.



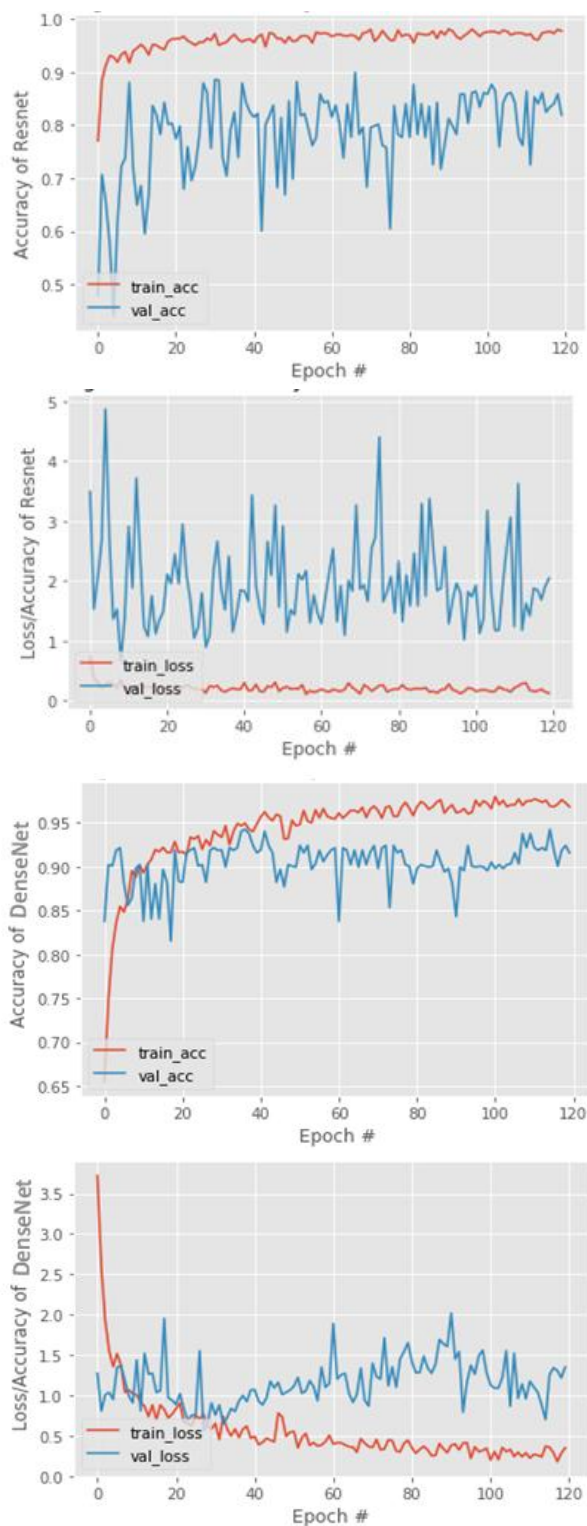


Fig. 7 Training and Validation performance of CNN Architectures

Training loss here reflects how effectively the model is actually learning for each iteration and which attributes to incorporate so that for the following iteration model will make less mistakes and can differentiate the brain tumor and normal

cases in a more accurate manner. Less loss value indicates that the DenseNet model is very accurate and that there are less mistakes made while classifying cases of brain tumors and normal cases.

Table. 1 Performance Comparison of DenseNet model

Model	Accuracy	Precision	Recall	F1 Score	AUC RoC
DenseNet	0.91	0.85	0.98	0.91	0.92
Inception V3	0.76	0.74	0.82	0.78	0.71
ResNet	0.86	0.81	0.91	0.85	0.81

ResNet's representational capacity is constrained by the identity shortcut that stabilizes training, but DenseNet has a greater capacity due to multi-layer feature concatenation. Table 1's findings demonstrate that our proposed model performs

significantly better than ResNet. The robustness of the DenseNet design is further assessed using various optimization algorithms and the results are tabulated in Table 2

Table. 2 Performance Comparison of DenseNet model with various optimization algorithm

Model	Accuracy	Precision	Recall	F1 Score	AUC RoC
ADAMAX	0.91	0.85	0.98	0.89	0.92
ADAM	0.88	0.85	0.92	0.86	0.88
SGD	0.85	0.84	0.81	0.83	0.86

The main advantage of AdaMax over SGD is that it is far less sensitive to the choice of hyper-parameters. The second momentum component of the ADAM estimate method is used by AdaMax. This provides a more reliable answer to the weight optimization issue. By updating parameters with just one record, Stochastic Gradient Decent (SGD) found a solution to the Gradient Descent problem. SGD takes a while to converge, though, because

each record necessitates both forward and backward propagation. Adam experiences a weight decay problem and does not always converge to the best option. Thus it could be seen in the results that the model trained with ADAMAX optimizer yielded better results in terms of accuracy, precision, and other metrics.

Conclusion

Convolutional neural networks (CNNs) and the DenseNet architecture have been presented as a method for diagnosing brain tumors using MR images. There is a need to evaluate publications linked to the use of deep learning based diagnosis in a logical and efficient way given the ongoing growth of brain diseases and their respective research. Researchers and educators will benefit from this study's quantitative analysis of the

relevant literature. To automatically choose the optimal optimizer, a comparative study has been offered. Among the optimizers Adamax, Adam, and SGD, Adamax performs better than the others. The results of our studies demonstrated that the random search method was used to choose the learning rate and batch size. According to our tests, 64 is the ideal batch size to obtain enough training time for the DenseNet model, while 0.01 is the ideal learning rate. This range of values is the

most effective one to employ for training to get an average accuracy of about 91%.

References

- [1] C. Zhang *et al.*, "ResNet or DenseNet? Introducing Dense Shortcuts to ResNet," *2021 IEEE Winter Conference on Applications of Computer Vision (WACV)*, Waikoloa, HI, USA, 2021, pp. 3549-3558,
- [2] Zhu, Yi, and Shawn Newsam. "Densenet for dense flow." *2017 IEEE international conference on image processing (ICIP)*. IEEE, 2017.
- [3] Nazir, Maria, Sadia Shakil, and Khurram Khurshid. "Role of deep learning in brain tumor detection and classification (2015 to 2020): A review." *Computerized medical imaging and graphics* 91 (2021): 101940.
- [4] Amin, Javaria, et al. "Brain tumor detection and classification using machine learning: a comprehensive survey." *Complex & intelligent systems* (2021): 1-23.
- [5] G. Huang, Z. Liu and L. van der Maaten, "Densely Connected Convolutional Networks," 2018.
- [6] <https://towardsdatascience.com/understanding-and-visualizing-densenets-7f688092391a>
- [7] Harpreet Kaur, Neelofar Sohi, "A Study for Applications of Histogram in Image Enhancement", *The International Journal of Engineering and Science (IJES)*, Vol.6, Issue.6, ISSN (e): 2319 – 1813 ISSN (p): 2319 – 1805 , pp: 59-63, 2017.
- [8] Rajulath Banu A.K, Dr. A. Ranjith Ram, "Contrast Enhancement of MRI Images: A Review", *International Journal of Emerging Technology and Advanced Engineering(IJETAE)*, ISSN 2250-2459, ISO 9001:2008 Certified Journal, Volume 5, Issue 6, June 2015
- [9] Gudigar A, Raghavendra U, San T, Ciaccio E, Acharya U. Application of multiresolution analysis for automated detection of brain abnormality using mr images: A comparative study. *Future Gener Comput Syst* 2019; 90:359–67.
- [10] Chen Y., Shao Y., Yan J., Yuan T.-F., Qu Y., Lee E., Wang S. A feature-free 30-disease pathological brain detection system by linear regression classifier. *CNS & Neurol Disorders-Drug Targets (Formerly Current Drug Targets-CNS & Neurological Disorders)* 2017;16(1):5–10.
- [11] Chen Y., Yang M., Chen X., Liu B., Wang H., Wang S. Sensorineural hearing loss detection via discrete wavelet transform and principal component analysis combined with generalized eigenvalue proximal support vector machine and tikhonov regularization. *Multimedia Tools Appl.* 2018;77(3):3775–3793.
- [12] Wang S.-H., Zhan T.-M., Chen Y., Zhang Y., Yang M., Lu H.-M., Wang H.-N., Liu B., Phillips P. Multiple sclerosis detection based on biorthogonal wavelet transform, rbf kernel principal component analysis, and logistic regression. *IEEE Access.* 2016;4:7567–7576.
- [13] Saiful Islam, Umme Sara, Abu Kawsar, Anichur Rahman, Dipanjali Kundu, Diganta Das Dipta, A.N.M. Rezaul Karim, and Mahedi Hasan. Sgbb: An efficient method for prediction system in machine learning using imbalance dataset. *Int J Adv Comput Sci Appl* 2021;12(3).
- [14] Shah FM, Hossain T, Ashraf M, Shishir FS, Al Nasim MA, Kabir MH. Brain tumor segmentation techniques on medical images-a review.
- [15] Komura D., Ishikawa S. Machine learning methods for histopathological image analysis. *Comput Struct Biotechnol J.* 2018;16:34–42.
- [16] Leng W.Y., Shamsuddin S.M. Writer identification for chinese handwriting. *Int J Adv Soft Comput Appl.* 2010;2(2):142–173.
- [17] Muhammad Arif F., Ajesh Shermin Shamsudheen, Geman Oana, Izdrui Diana, Vicoveanu Dragos. Brain tumor detection and classification by mri using biologically inspired orthogonal wavelet transform and deep learning techniques. *J Healthcare Eng.* 2022
- [18] Haitham Alsaif, Ramzi Guesmi, Badr M Alshammari, Tarek Hamrouni, Tawfik Guesmi, Ahmed Alzamil, and Lamia Belguesmi. A novel data augmentation-based

- brain tumor detection using convolutional neural network. *Appl Sci* 12(8):3773, 2022.
- [19] Hamza Rafiq Almadhoun and Samy S. Abu-Naser. Detection of brain tumor using deep learning. *Int J Acad Eng Res* 2022;6(3).
- [20] R. Sa, W. Owens, R. Wiegand, M. Studin, D. Capoferri, K. Barooha, A. Greaux, R. Rattray, A. Hutton, J. Cintineo, et al., Intervertebral disc detection in x-ray images using faster r-cnn, in: 2017 39th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC), IEEE, 2017, pp. 564–567.
- [21] Sadia Anjum, Lal Hussain, Mushtaq Ali, Monagi H. Alkinani, Wajid Aziz, Sabrina Gheller, Adeel Ahmed Abbasi, Ali Raza Marchal, Harshini Suresh, and Tim Q. Duong. Detecting brain tumors using deep learning convolutional neural network with transfer learning approach. *Int J Imag Syst Technol* 2022;32(1):307–323.
- [22] Muhannad Faleh Alanazi, Muhammad Umair Ali, Shaik Javeed Hussain, Amad Zafar, Mohammed Mohatram, Muhammad Irfan, Raed AlRuwaili, Mubarak Alruwaili, Naif H. Ali, and Anas Mohammad Albarrak. Brain tumor/mass classification framework using magnetic-resonance-imaging-based isolated and developed transfer deep-learning model. *Sensors* 2022;22(1):372.
- [23] Shin H.-C., Roth H.R., Gao M., Lu L., Xu Z., Nogues I., Yao J., Mollura D., Summers R.M. Deep convolutional neural networks for computer-aided detection: Cnn architectures, dataset characteristics and transfer learning. *IEEE Trans Med Imaging*. 2016;35(5):1285–1298.
- [24] Sultan H.H., Salem N.M., Al-Atabany W. Multi-classification of brain tumor images using deep neural network. *IEEE Access*. 2019;7:69215–69225.
- [25] T. Hossain, F. Shishir, M. Ashraf, M. Al Nasim, F. Shah, Brain tumor detection using convolutional neural network, in: (pp. 1–6). IEEE., 2019 May 3.
- [26] Talo M., Baloglu U.B., Yildirim Ö., Acharya U.R. Application of deep transfer learning for automated brain abnormality classification using mr images. *Cogn Syst Res*. 2019;54:176–188.