International Journal on Engineering Technology (InJET)

www.kec.edu.np/journal
Volume 1, issue no. 1, Nov 2023, Page 11-26
Received Date: July 13, 2023
Accepted Date: Sept 28, 2023
ISSN: 3021-940X (print)

Brain Tumor Detection Using Convolutional Neural Networks: A Comparative Study

Bibhusha Ojha[1], Ruman Maharjan[2*], Tirtha Acharya[3 *]

[1]*Department of Computer and Electronics Engineering, Kantipur Engineering College, Dhapakhel, Lalitpur, Nepal, kan075bct021@kec.edu.np*
[2]*Department of Computer and Electronics Engineering, Kantipur Engineering College, Dhapakhel, Lalitpur, Nepal, kan075bct050@kec.edu.np*
[3]*Department of Computer and Electronics Engineering, Kantipur Engineering College, Dhapakhel, Lalitpur, Nepal, achtirtha@gmail.com*

Abstract

Using Magnetic Resonance Imaging (MRI) images to detect brain tumors by medical practitioners is mundane and prone to errors. Misdiagnosis of brain tumors can be life-threatening, so to lessen misdiagnosis, computational techniques can be used in concert with medical professionals. Deep learning approaches have been gaining popularity in modeling and developing systems for medical image processing that can detect abnormalities quickly. The methods proposed herein are based on Convolutional Neural Networks (CNN) trained on the 'BR35H::Brain Tumor Detection 2020' dataset. A custom CNN architecture was designed, followed by the utilization of transfer learning with four pre-trained models: InceptionV3, ResNet101, VGG19, and DenseNet169 and a comparative analysis of these architectures has been presented in this paper. The experimental results show that the DenseNet169 model outperformed other models with a training accuracy of 99.83 %, test accuracy of 99.66%, precision of 99.67%, and recall of 99.67%. Additionally, ResNet101 has a 95.92% test accuracy, VGG19 has a test accuracy of 97.83%, the custom architecture has a test accuracy of 98.16%, and InceptionV3 has the lowest test accuracy of 91.66%. It has been concluded that DenseNet169 provides better results for the classification of brain tumors than other models.

Keywords: Brain Tumor, Deep Learning, Convolutional Neural Network, Magnetic Resonance Imaging, Transfer Learning, Image Classification

1. Introduction

Brain tumors are characterized by the aberrant growth of tissues within the brain (Tandel, et al., 2019)[(Gore, 2020) (DeAngelis, 2001). There are different types of tumors identified by their texture, location, and shape (Mabray, 2015) (Cha, 2006) (Ranjbarzadeh, et al., 2021). Analyzing these tumors presents certain challenges due to their diverse morphology, size, location, and appearance within the brain. Magnetic Resonance Imaging (MRI) is mainly used for detecting brain tumors by medical professionals. However, the existing diagnostic methods rely heavily on subjective human expertise, increasing the likelihood of erroneous tumor identification. Various deep learning methodologies have been used for the purpose of brain tumor detection and classification (Mamun, et al., 2022) (Mamun, et al., 2022) (Mahmud, et al., 2022). CNN, a powerful image processing and computing method, uses deep learning to perform each generative and descriptive task and is seen typically exploiting machine vision: image and video recognition, recommender systems and linguistic communication processes (Chattopadhyay, 2022). Deep learning models set an exciting trend in machine learning owing to their

**Corresponding Author*

ability to effectively capture intricate relationships without necessitating an extensive number of nodes, like in shallow architectures such as Support Vector Machines (SVM) and K-nearest neighbor (KNN) (S. Sarkar, 2020). Brain Tumor detection poses challenges in selecting appropriate features and needs an effective model to extract the relevant features from these images. To overcome these challenges, this paper presents a comprehensive study of a custom CNN architecture and four pre-trained CNN models. The primary objective of this study is to evaluate the performance of these models and determine the model with the highest accuracy in detecting brain tumors from the MR images.

The custom architecture has been proposed in this study since it can be tailored to the intricacies of brain tumor detection allowing the network structure to better address the unique characteristics of tumor images. By customizing the architecture, specific convolutional layers, activation functions, and pooling strategies can be implemented for detecting features that might be overlooked by more generic pre-trained models. It also enables us to incorporate domain-specific knowledge into the architecture. Proposing a custom CNN architecture reflects our commitment and contribution to the medical image classification domain that can potentially offer insights into other image classification and analysis tasks.

The proposed custom CNN architecture consists of five convolutional layers, each followed by a max-pooling layer, and concludes with two fully connected layers. The Rectified Linear Unit (ReLU) activation function is employed after each convolutional layer, while the Softmax activation function is utilized in the final layer for accurate probability predictions. To address the issue of overfitting, dropout regularization is incorporated into the model architecture (H. Ide, 2017) (H. H. Tan, 2019).

In the transfer learning paradigm, the well-established InceptionV3, ResNet101, VGG19, and DenseNet169 models were employed as starting points due to their exceptional performance on diverse image classification tasks. These models have been trained on vast and diverse datasets and have learned valuable features that can expedite the detection of brain tumors. However, instead of employing these models in their entirety, only the initial layers were utilized while keeping their weights frozen. This strategy allowed the models to benefit from the comprehensive knowledge gained during their extensive training on large-scale datasets. By fine-tuning the classification layer to align with the unique requirements of brain tumor detection, the transfer learning models were effectively customized to enhance their suitability for this specific application domain.

Metrics such as accuracy, precision, recall, specificity, and F1 score were utilized to comprehensively assess the best-performing models' capabilities in accurately identifying brain tumors. The results revealed notable distinctions in the performance of the custom architecture and the transfer learning models, highlighting the advantages and limitations of each approach.

2. Related Works

In one of the research works, the Convolutional Neural Network CNN (S. Sarkar, 2020) was implemented, which drives an overall accuracy of 91.3% and a recall of 88%, 81%, and 99% in the detection of meningioma, glioma, and pituitary tumor respectively. Deep learning architecture leveraged 2D Convolutional Neural Networks for classification of different types of brain tumors from MRI scans. In this paper, techniques like data acquisition, data preprocessing, pre–model, model optimization, and hyperparameter tuning were applied. Moreover, the 10-fold cross-validation was performed on the complete dataset to check for the model's generalizability.

In (H. Mohsen, 2018), Fuzzy C-Means (FCM) segmentation was applied to separate the tumor and non-tumor regions of the brain. Also, wavelet features were extracted by using multilevel Discrete Wavelet Transform (DWT). Finally, Deep Neural Network (DNN) was incorporated for brain tumor classification with high accuracy. This technique was compared with KNN, Linear Discriminant Analysis (LDA), and Sequential Minimal Optimization (SMO) classification methods. An accuracy of 96.97% using DNN-based brain tumor classification. However, the complexity was very high, and performance was abysmal (H. Mohsen, 2018).

In the paper, a Brain Cancer Detection and Classification System (D. Joshi, 2010) has been developed using ANN. Image processing techniques such as histogram equalization, image segmentation, image enhancement, and feature extraction have been used. The proposed approach using ANN as a classifier for the classification of brain images provides a good classification efficiency compared to other classifiers. The sensitivity, specificity, and accuracy are also improved. The proposed approach was computationally effective and yielded (D. Joshi, 2010).

The study, (Cinar, 2022) is aimed to compare the models used for tumor detection from brain MRI images within the scope of deep learning methods. For this reason, the five most used convolutional neural networks for brain tumor classification are discussed. VGG19, DenseNet169, AlexNet, InceptionV3, and ResNet101 models, which are Convolutional Neural Network (CNN) architectures, were used. MR images, which underwent the same dataset and preprocessing processes, were trained with these models with the same hyperparameters. As a result of the study, the ResNet101 model obtained the highest accuracy rate with an accuracy value of 98. 6%. In addition, the VGG19 model showed a very high accuracy rate of 97.2%. Other models have accuracy values of InceptionV3 94.3%, DenseNet169 92.8%, and AlexNet 89.5%, respectively (Cinar, 2022).

This research paper (A. Naseer, 2021) introduces two deep learning models aimed at binary (normal vs. abnormal) and multiclass (meningioma, glioma, and pituitary) brain tumor classification, utilizing two publicly available datasets comprising 3064 and 152 MRI images, respectively. While a 23-layer convolution neural network (CNN) effectively handles the first dataset's substantial data, the challenge of overfitting arises with the smaller second dataset. To mitigate this, transfer learning is employed, combining the VGG16 architecture with the researchers' "23 layers CNN" design. Notably, the study stands out by comparing its models with existing literature, yielding promising results. The proposed models achieve impressive classification accuracy rates, up to 97.8% and 100% (A. Naseer, 2021).

Brain tumors, both cancerous and non-cancerous, pose significant health risks, making their early and accurate detection crucial. This study (Heidari & Rafatirad, 2020) conducts a comprehensive review of brain tumor detection methods, spanning two decades from 2000 to 2020, involving 20 research papers. The focus is on MRI image analysis, where two key issues, Image Restoration and Image Enhancement, are identified and effectively addressed. The research introduces a novel approach employing Convolutional Neural Networks (CNNs) to enhance classification accuracy while overcoming dataset image algorithm errors. Implemented using Python and TensorFlow, this method showcases promising results. The study delves into quantitative aspects of brain tumor analysis, encompassing shape, texture, and signal intensity, aiming for high accuracy with minimal errors, thus laying the foundation for future advancements in this critical field (Heidari & Rafatirad, 2020).

3. Methodology

The dataset used in this study consisted of labeled Brain MRI scans. To enhance the model's capacity for accurate prediction for any given input, there must be ample data. To address this requirement, data augmentation techniques were employed, resulting in an augmented dataset with an increased number of images. These pre-processed images, enriched through data augmentation, were subsequently utilized to train the CNN models for brain tumor detection.

The dataset was split into three sub-categories: 80% in the train set, 10% in the validation set, and the remaining 10% in the test set. Before model training and testing, the images were resized to ensure uniformity, with a standardized resolution of 256x256 pixels. However, for compatibility with the InceptionV3 architecture, which required a specific image size of 299x299 pixels, the images were resized accordingly.

Finally, five different models were trained on the training set, and the performance of each model was considered to find the best-performing model.

3.1. Dataset

'Br35H::Brain Tumor Detection 2020' is publicly available in the Kaggle dataset containing MRI images of brain tumors (A. Naseer, 2021). The dataset was categorized into two portions: those with brain tumors and those without brain tumors. The dataset contains 1,500 images for each category, making a total of 3,000 images.

3.2. Dataset augmentation

To amplify the dataset's size and diversity, a data augmentation technique was employed, randomly augmenting each original image by a factor of five- each image was flipped horizontally and vertically, randomly rotated, and random noise (Gaussian noise) was added and subjected to random adjustment of brightness.

The augmented images were subsequently stored in separate folders named 'yes augmented' and 'no augmented.' This results in a total of 7,500 augmented images for each category. The original images in the 'yes' and 'no' folders are merged with the augmented images to create a final dataset of 18,000 images, with 9,000 images for

each category. By augmenting the dataset into 18,000 images, the models have been exposed to a diverse range of samples which reduces the risk of overfitting and helps build a resilient model. The number of augmented images maintains a balance between the available resources and meaningful experimentation.

3.3. Dataset division

For model training, validation, and evaluation, the dataset was thoughtfully partitioned into three distinct subsets, adhering to an 8:1:1 ratio. The training set contains 14400 images, with 7180 images in the 'yes' category and 7220 images in the 'no' category. The validation set contains 1800 images, with 904 images in the 'yes' category and 896 images in the 'no' category. Finally, the test set contains 1800 images, with 916 images in the 'yes' category and 884 images in the 'no' category.

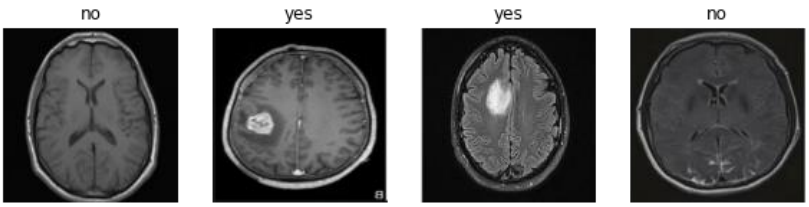

Fig 1.1: Dataset

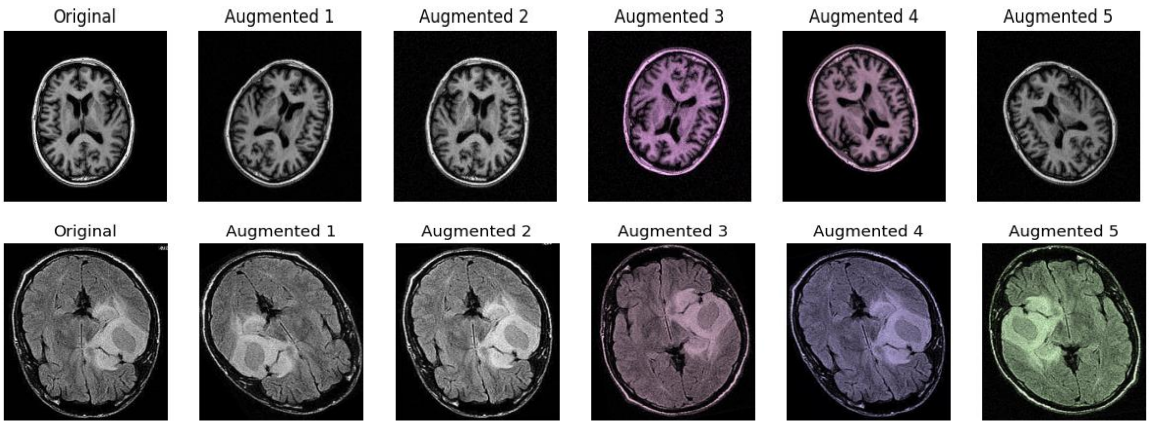

Fig 1.2: Augmented Dataset

3.4. Model Selection

The performance of five different approaches for brain tumor detection was compared in this study:

1. Custom CNN architecture:
   A custom Convolutional Neural Network (CNN) architecture was developed, comprising five convolutional layers, each followed by a max-pooling layer, and two fully connected layers. The Rectified Linear Unit (ReLU) activation function was applied after each convolutional layer, while the Softmax activation function was used in the final layer. A dropout regularization of 20% was implemented to prevent overfitting. The model was trained for 30 epochs using the Adam optimizer and the sparse categorical cross-entropy loss function.
   This CNN architecture has five convolutional layers. They are
   - Conv2D layer with 32 filters and a kernel size of (3,3)
   - Conv2D layer with 64 filters and a kernel size of (3,3)
   - Conv2D layer with 64 filters and a kernel size of (3,3)
   - Conv2D layer with 64 filters and a kernel size of (3,3)
   - Conv2D layer with 128 filters and a kernel size of (3,3)

Convolutional layers are the ones that perform the feature extraction from the input image, so they are critical components of the CNN architecture. The proposed CNN architecture is as follows:

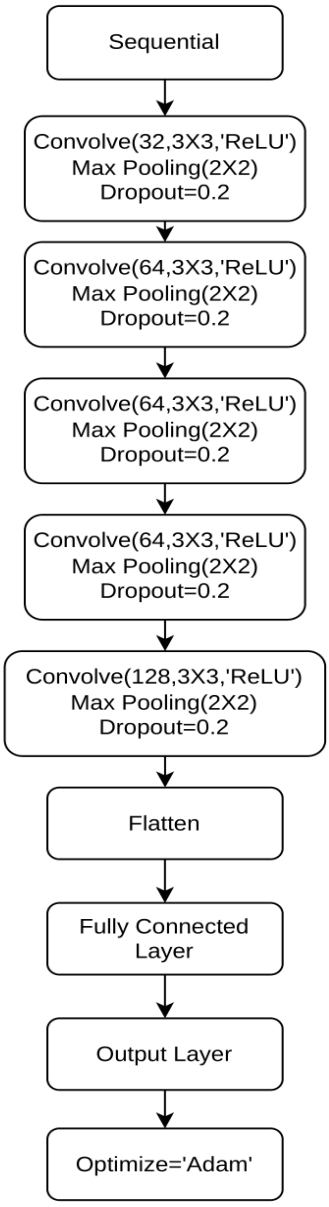

Fig 2: Proposed CNN architecture

2. Transfer Learning Models

Transfer learning provides the opportunity to leverage the pre-trained weights of the models, learned through many hours of training on powerful GPUs (Heidari & Rafatirad, 2020) (Rahman, et al., 2020). They are frequently used with pre-trained neural networks (Cao, et al., 2010) (Zhuang, et al., 2020) (Loey, et al., 2021) since they are time and cost-effective (Thenmozhi & Reddy, 2019).

2.1. Inception V3:

The Inception V3 model is an upgrade to the Inception V1 model widely employed for image detection. (Liu, et al., 2021) (Jignesh Chowdary, et al., 2020) (Xia, et al., 2017) (Tio, 2019) (Mujahid, et al., 2022). Inception V3 was published in 2015, with 42 layers and decreased error rates than its predecessors. Convolution, pooling, dropout, fully connected, and softmax are the steps in the Inception process (Andrew & Santoso, 2022) (Al Husaini, et al., 2022).

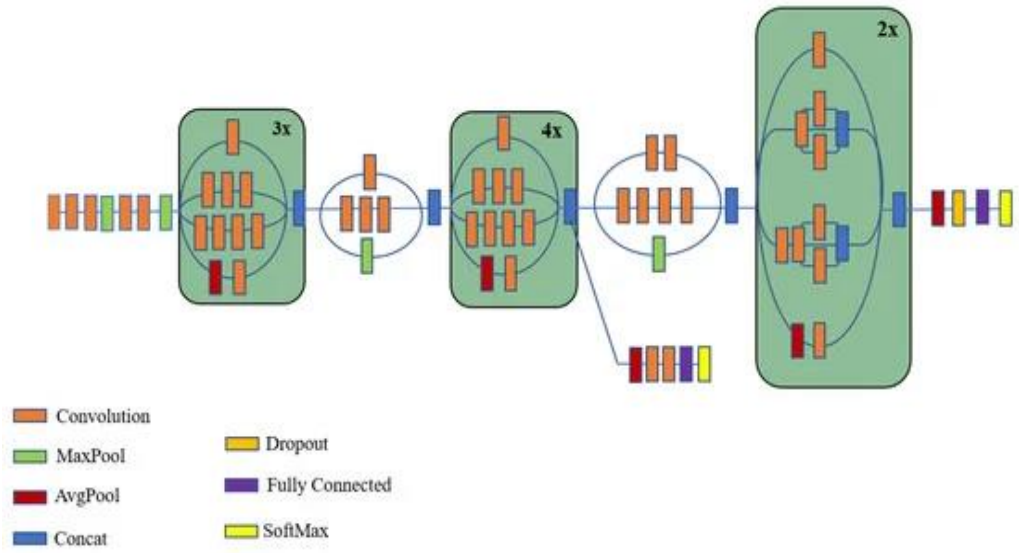

Fig 3: Inception V3 architecture [32]

The Inception v3 model was employed as a pre-trained feature extractor, with an additional CNN classifier incorporated on top. The weights of the Inception model were frozen, ensuring they remained unchanged throughout the training process. The classification layer was customized to align with the requirements of brain tumor detection. The classification layer comprised two distinct classes for prediction: non-tumorous and tumorous. This layer was trained on the dataset.

For the training phase, the model was optimized using the Adam optimizer and the categorical cross-entropy loss function. The training process was iterated for 30 epochs, allowing the model to refine its classification capabilities.

2.2. ResNet-101:

ResNet-101 is a 101-layered CNN architecture trained on the ImageNet dataset comprising over a million images (Akiba, et al., 2017). This architecture, based on the Residual neural network learning approach, is one of the most in-depth proposed architectures for ImageNet (He K, 2016). The main difference between Resnet-101 and other architectures is that it optimizes the residues between input and desired convolution properties. Unlike alternative architectures, desired functionalities are obtained more readily and efficiently. As a result, residual optimization can reduce the number of parameters in a more complex network (Chung YM, 2021).

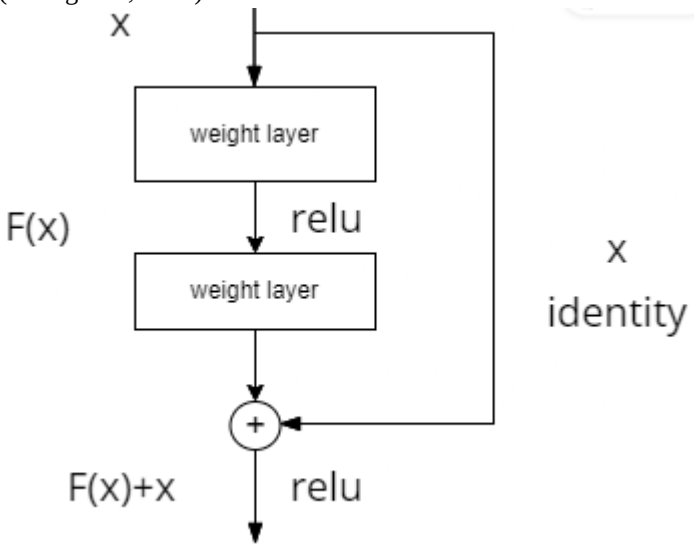

Fig 4: ResNet-101 architecture (Koç M, 2019)

A pre-trained ResNet-101 model was utilized and augmented with an additional fully connected layer. The weights of the ResNet-101 model were frozen to prevent retraining and the classification layer was customized to suit the use case. The classification layer was then trained on the dataset. The model was trained for 30 epochs using the Adam optimizer and the categorical cross-entropy loss function. The softmax activation function was implemented during the training process.

2.3. VGG-19:

Simonyan and Zisserman of the University of Oxford developed the VGG-19 model, a 19-layer (16 conv., 3 fully connected) CNN that strictly utilized 3x3 filters with stride and pad of 1, as well as 2x2 max-pooling layers with stride 2 (Simonyan, 2014).

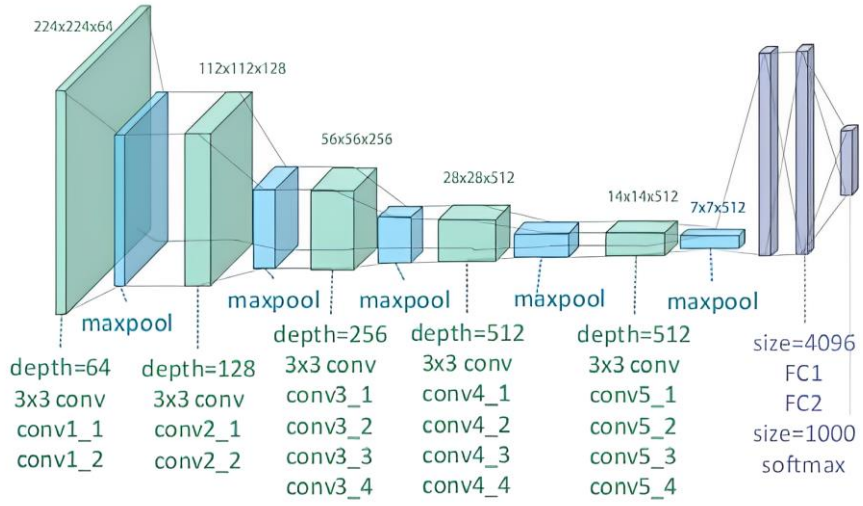

Fig 5: VGG-19 architecture (Zheng, 2018)

A pre-trained VGG-19 model was used to leverage the power of transfer learning. The classification layer of the model was customized as per the requirement of the task of brain tumor detection. The model was iterated for 30 epochs using Adam optimizer and sparse categorical cross-entropy, which fine-tuned the model and enhanced its performance over the consecutive epochs.

2.. DenseNet-169:

DenseNet-169 has 169 layers, which include convolutional layers, dense blocks, and batch normalization. There is a 7x7 convolution layer with stride 2 in the first part of the architecture. The layer that follows is a 3x3 stride 2 max pooling layer. There are four dense blocks, each followed by a transition layer. The number of layers in each dense block varies. For DenseNet169 there are (6,12,32,32) layers in the dense block. The layers between dense blocks are the transition layers which are responsible for convolution and pooling. The transition layer has a 1x1 convolution layer with a 2x2 average pooling and a stride of 2. The convolution operations in the architecture are the bottleneck layers. A 1x1 convolution reduces the number of channels in the input and a 3x3 convolution performs the convolution operation. The additional input acquired from all previous layers is sent to the feature maps in all subsequent layers in Densenet. Each layer gathers data from the layers above it. Because each layer receives feature maps from all preceding layers, the network may be narrower and tighter, resulting in fewer channels (Nair, 2022).

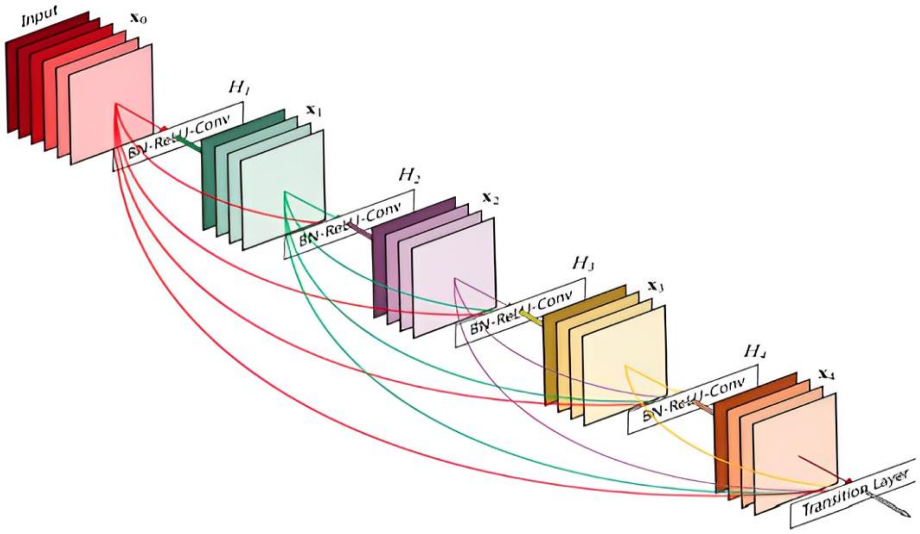

Fig 6.1: DenseNet-169 architecture (Nair, 2022)

Layers	Output Size	DenseNet-121	DenseNet-169	DenseNet-201	DenseNet-264
Convolution	112 × 112	7 × 7 conv, stride 2			
Pooling	56 × 56	3 × 3 max pool, stride 2			
Dense Block (1)	56 × 56	[1 × 1 conv; 3 × 3 conv] × 6	[1 × 1 conv; 3 × 3 conv] × 6	[1 × 1 conv; 3 × 3 conv] × 6	[1 × 1 conv; 3 × 3 conv] × 6
Transition Layer (1)	56 × 56	1 × 1 conv			
	28 × 28	2 × 2 average pool, stride 2			
Dense Block (2)	28 × 28	[1 × 1 conv; 3 × 3 conv] × 12	[1 × 1 conv; 3 × 3 conv] × 12	[1 × 1 conv; 3 × 3 conv] × 12	[1 × 1 conv; 3 × 3 conv] × 12
Transition Layer (2)	28 × 28	1 × 1 conv			
	14 × 14	2 × 2 average pool, stride 2			
Dense Block (3)	14 × 14	[1 × 1 conv; 3 × 3 conv] × 24	[1 × 1 conv; 3 × 3 conv] × 32	[1 × 1 conv; 3 × 3 conv] × 48	[1 × 1 conv; 3 × 3 conv] × 64
Transition Layer (3)	14 × 14	1 × 1 conv			
	7 × 7	2 × 2 average pool, stride 2			
Dense Block (4)	7 × 7	[1 × 1 conv; 3 × 3 conv] × 16	[1 × 1 conv; 3 × 3 conv] × 32	[1 × 1 conv; 3 × 3 conv] × 32	[1 × 1 conv; 3 × 3 conv] × 48
Classification Layer	1 × 1	7 × 7 global average pool			
		1000D fully-connected, softmax			

Fig 6.2: Detailed DenseNet architectures (Simonyan, 2014)

The DenseNet model, a pre-trained convolutional neural network, was employed in this study to harness the benefits of transfer learning. Specifically, the classification layer of the DenseNet model was tailored to suit the unique requirements of brain tumor detection. The classification layer had two output neurons, detecting the presence or absence of brain tumors. Through an iterative training process spanning 30 epochs, the model was refined using the Adam optimizer and sparse categorical cross-entropy loss function. This iterative training approach effectively fine-tuned the model, leading to significant performance improvements throughout the training iterations.

3.5. Performance Evaluation Metrics

The following metrics were considered to evaluate the machine learning model and to analyze its performance.

3.5.1. Precision:

Precision is the proportion of Predicted Positive cases that are Real Positives (Powers, 2011). It is represented as

$$Precision = TP/(TP + FP) \quad \text{(Equation 1)}$$

where TP = True Positive, the number of positive samples correctly classified

FP = False Positive, the number of samples misclassified as positive.

3.5.2. Recall:

The proportion of Real Positive cases that are accurately predicted Positive is called recall or sensitivity (Powers, 2011). The recall can be represented by the following formula

$$Recall = TP/(TP + FN) \quad \text{(Equation 2)}$$

where FP = False Negative, the number of samples misclassified as negative.

3.5.3. Specificity:

Specificity is the proportion of real negative cases accurately predicted as Negative (Powers, 2011). It can be calculated using the following formula

$$Specificity = TN/(TN + FP) \quad \text{(Equation 3)}$$

W here TN = True Negative, the number of negative samples correctly classified

3.5.4. F1-score:

The F1-score is calculated from precision and recall (Powers, 2011). It represents precision and recall symmetrically in the same metric. The F1-score can be mathematically represented as

$$F1 - score = 2 * (precision * recall)/(precision + recall) \quad \text{(Equation 4)}$$

4. Experimental Analysis

The results of the five different models— custom CNN architecture, InceptionV3, ResNet101, VGG-19, and DenseNet169 architectures on the augmented brain tumor dataset regarding the training, test, and validation accuracy are mentioned in the table below.

Table I. Comparison of different approaches

Methodology	Training Accuracy	Validation Accuracy	Test Accuracy
Custom CNN architecture	0.9951	0.9816	0.9816
Inception-V3	0.9547	0.9261	0.9166
ResNet-101	0.9747	0.9609	0.9592
VGG-19	0.9911	0.9767	0.9783
DenseNet-169	0.9983	0.9978	0.9966

After analyzing the models, it was observed that amongst CNN architecture and transfer learning models (InceptionV3, ResNet-101, VGG-19, and DenseNet-169), DenseNet-169 outperformed other models with 99.83

% training accuracy and 99.66% test accuracy. The Inception V3 model gave the lowest training accuracy of 95.47% and the lowest test accuracy of 91.66%. Accuracy and loss graphs were plotted for all the different models. The orange lines represent training accuracy and loss for respective graphs, while the blue lines represent validation accuracy and loss. The training accuracy concerning validation accuracy graphs are given below:

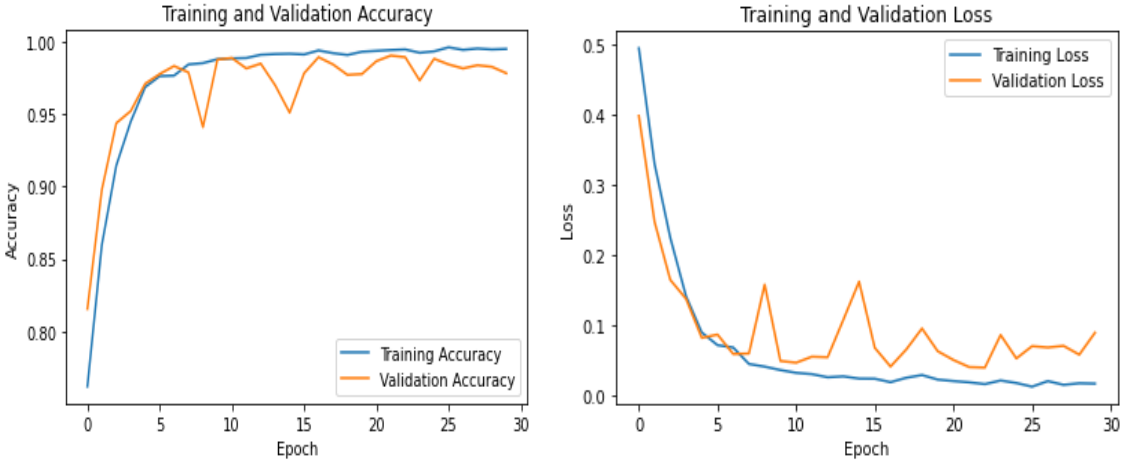

Fig 7.1: Model accuracy and loss graph for custom CNN model

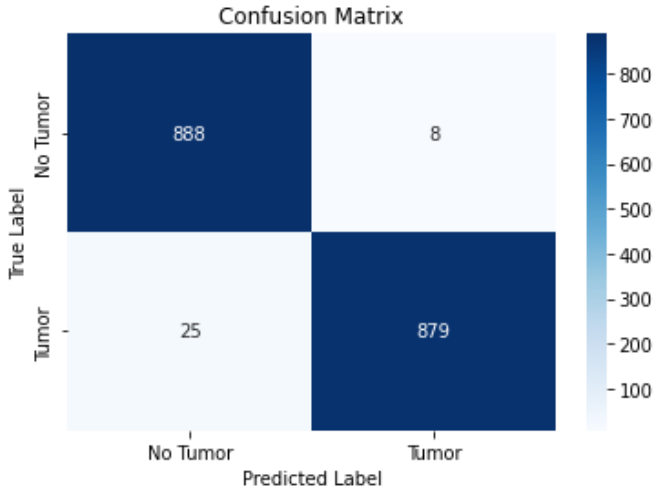

Fig 7.2: confusion matrix for custom CNN model

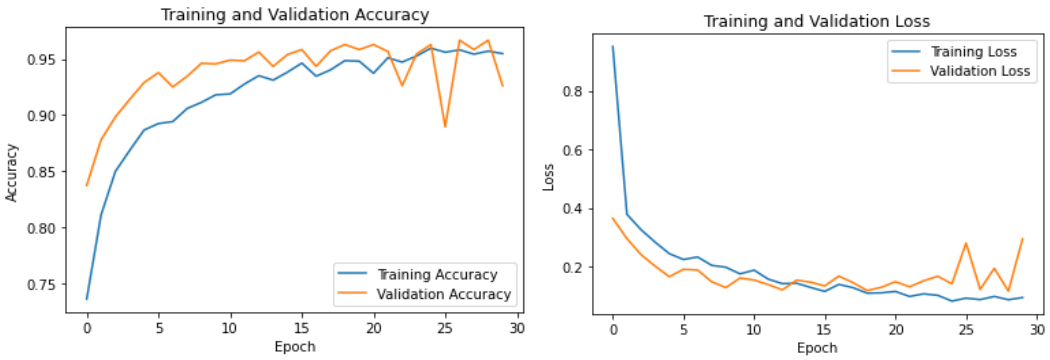

Fig 7.3: Model accuracy and loss graph for InceptionV3 model

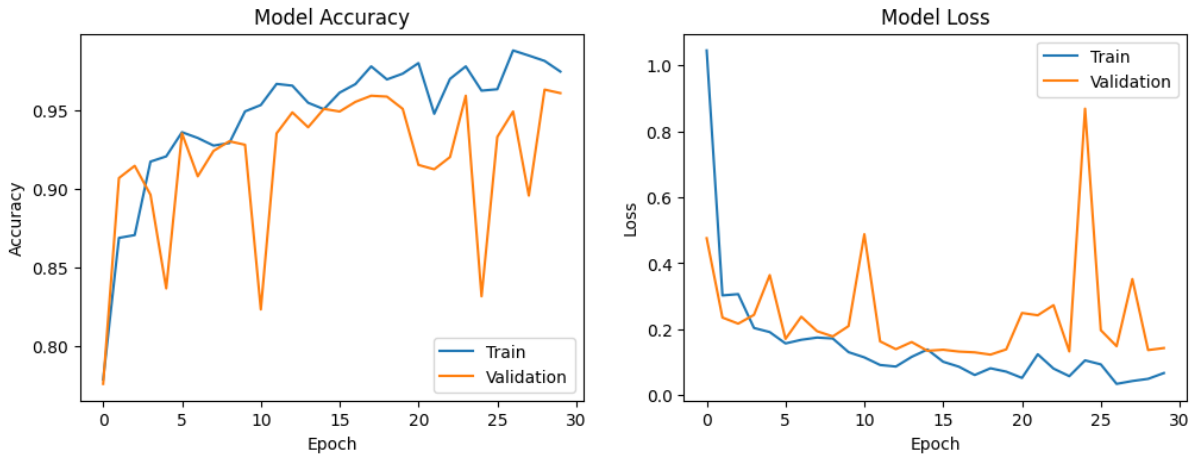

Fig 7.4 : Model accuracy and loss graph for ResNet-101 model

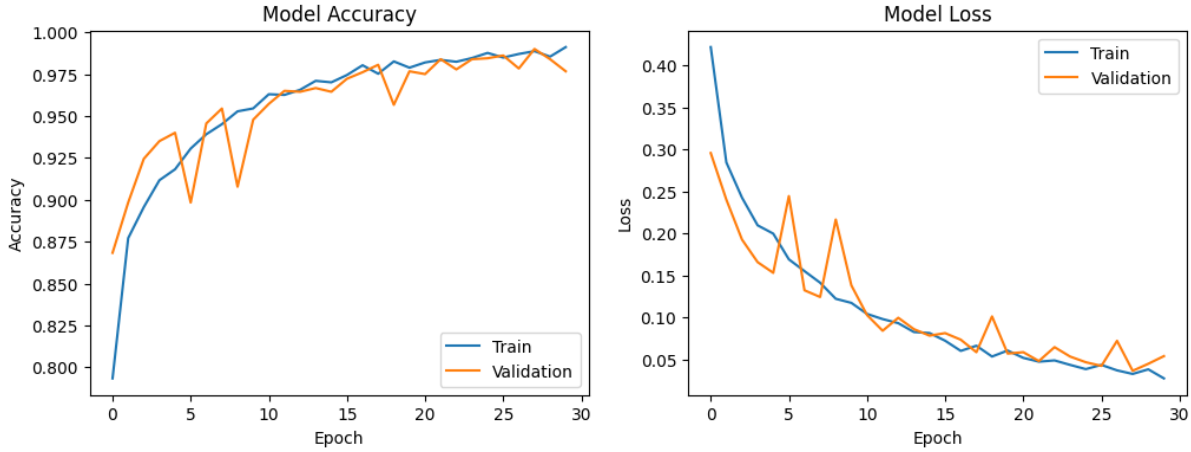

Fig 7.5 : Model accuracy and loss graph for VGG-19 model

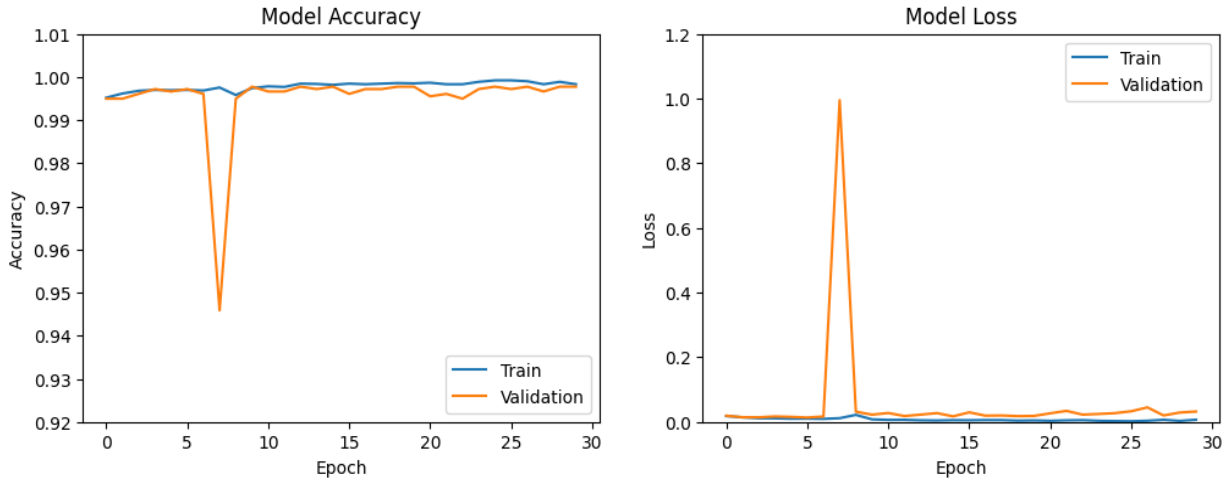

Fig 7.6 : Model accuracy and loss graph for DenseNet-169 model

```
Model: "sequential"
_________________________________________________________________
 Layer (type)                Output Shape              Param #
=================================================================
 densenet169 (Functional)    (None, 8, 8, 1664)        12642880

 flatten (Flatten)           (None, 106496)            0

 dense (Dense)               (None, 256)               27263232

 dense_1 (Dense)             (None, 2)                 514

=================================================================
Total params: 39906626 (152.23 MB)
Trainable params: 39748226 (151.63 MB)
Non-trainable params: 158400 (618.75 KB)
_________________________________________________________________
```

Fig 7.7: Model summary for DenseNet-169 model

Finally, to assess the validity of the best-performing model, various performance metrics were analyzed. The DenseNet-169 model had a precision of 99.67% which means that out of all the positive predictions made by the model, approximately 99.67% of them are accurate or true positives, a recall of 99.67% which means that out of all the actual positive cases of brain tumors, approximately 99.67% of them were correctly identified and classified as positive, a specificity of 99.67% which means that out of all the true negative cases in the dataset, approximately 99.67% of them are correctly identified as true negatives and an F1-score of 99.61% which means that the model has achieved a high level of accuracy in both precision and recall. The model's precision, recall, and F1-score metrics demonstrate that it can adapt to new data and can accurately predict the presence of tumors or no presence of tumors on unseen images.

The generalizability of the model is evaluated using the confusion matrix (fig.8) and K-fold cross-validation (fig.9). There were 904 tumorous images and 896 non-tumorous images in the test images. The model predicted 3 non-tumorous images as tumorous and 3 tumorous images as non-tumorous. The model was also evaluated using the K-fold cross-validation technique. In the process of model validation, each of the 10-fold underwent a validation procedure comprising 10 epochs. Notably, several epochs yielded a remarkable accuracy of 100%. The highest accuracy of 100%, coupled with a negligible loss of 0.0000152 was achieved. Conversely, the lowest accuracy was 99.82%, accompanied by a loss of 0.0055. These results underscore the model's ability to consistently attain high accuracy levels, with occasional peaks at 100% accuracy. The average accuracy across all folds was 99.92%, reflecting the model's consistently high performance.

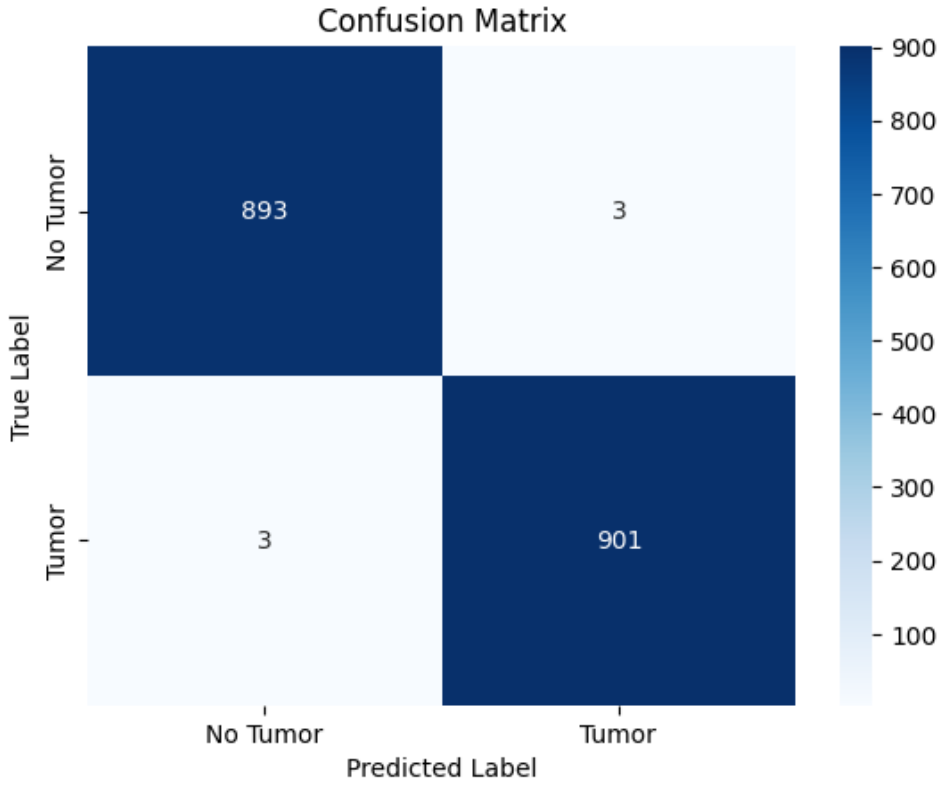

Fig 8: Confusion matrix for DenseNet-169

Fold	Loss	Accuracy
1.0	0.0029	99.98
2.0	0.0006	99.83
3.0	0.001	99.97
4.0	0.0032	99.86
5.0	0.0037	99.91
6.0	0.00011686	100.0
7.0	0.0013	99.97
8.0	0.0055	99.82
9.0	1.5258e-05	100.0
10.0	0.0094	99.93

Fig 9: K-fold Cross Validation

5. Conclusion and Future Enhancements

The paper discussed and implemented various deep-learning models that can be utilized for brain tumor detection in MRI images. The paper also compared five different convolutional neural networks (CNN) architectures that were trained on 14400 images, validated, and tested on 1800 images each. The results from the experimental analysis showed that the DenseNet-169 architecture exhibited a higher training accuracy of 99.83%, 99.66% test accuracy, a precision of 99.67%, a recall of 99.67%, a specificity of 99.67%, and an F1-score of 99.61%. These results highlight the effectiveness of using deep neural networks in detecting brain tumors in MRI images. The DenseNet-169 architecture can potentially be used in medical image classification by medical professionals for diagnosing brain tumors.

Due to the limitation of powerful GPUs, the paper focused on binary classification. However, with access to powerful GPUs and a variety of brain MRI scans, a more robust model can be built to suit the use case by

classifying brain tumors based on their location and size. Such advancements will ultimately benefit patients and medical professionals.

Acknowledgment

The authors would like to extend their gratitude to the Department of Computer and Electronics Engineering at Kantipur Engineering College for their invaluable guidance and unwavering support throughout the research. The insights shared by the faculty members were of utmost importance for enhancing the study's relevance. The Department created an encouraging environment that fostered research and innovation. The authors are particularly thankful to the college for nurturing a research culture that has led to the authors' intellectual growth. Finally, the authors are grateful to the college for their direct and indirect assistance during the research.

References

A. Naseer, T. Y. A. A. T. S. K. Z., 2021. Computer-aided brain tumor diagnosis: performance evaluation of deep learner cnn using augmented brain MRI. *International Journal of Biomedical Imaging*, Volume 20.
Akiba, T., Suzuki, S. & Fukuda, K., 2017. *Training resnet-50 on imagenet.* s.l.:arXiv.

Al Husaini, M. et al., 2022. Thermal-based early breast cancer detection using inception V3, inception V4 and modified inception MV4.. *Neural Computing and Applications*, 34(1), pp. 333-348.

Andrew, A. & Santoso, H., 2022. Compare VGG19, ResNet-50, Inception-V3 for review food rating.. *SinkrOn*, 7(2), pp. 845-494.

Cao, B. et al., 2010. *Adaptive transfer learning.* Atlanta, GA, USA, AAAI Conference on Artificial Intelligence.

Cha, S., 2006. Update on Brain tumor imaging: From anatomy to physiology. *Am. J. Neuroradiol*, pp. 475-487.

Chattopadhyay, M. M., 2022. MRI-based brain tumour image detection using CNN based Deep Learning Method. *Neuroscience Informatics*, 2(4), p. 100060.

Chung YM, H. C. L. A. S. C. T., 2021. A hybrid deep learning architecture and its application to skin lesion classification. *Mathematics*, 9(22), p. 2924.

Cinar, N. K. B. ,. K. M., 2022. *Comparison of deep learning models for brain tumor classification using mri images.* s.l., International Conference on Decision Aid Sciences and Applications (DASA).

D. Joshi, N. R. V. M., 2010. *Classification of brain cancer using artificial neural network.* s.l., 2nd International Conference on Electronic Computer Technology.

DeAngelis, L., 2001. Brain tumors. *New England Journal of Medicine*, 344(2), pp. 114-123.

Gore, D. a. D. V., 2020. Comparative study of various techniques using deep learning for Brain tumor detection. *International Conference for Emerging Technology (INCET)*, pp. 1-4.

H. H. Tan, K. H. L., 2019. *Vanishing Gradient Mitigation with Deep Learning Neural Network Optimization.* Sarawak, Malaysia, 2019 7th International Conference on Smart Computing & Communications (ICSCC).

H. Ide, T. K., 2017. *Improvement of learning for CNN with ReLU activation by sparse regularization.* Anchorage, AK, USA, 2017 International Joint Conference on Neural Networks (IJCNN).

H. Mohsen, E.-S. A. E.-D. E.-S. M. E.-H. A.-B. M. S., 2018. Classification using deep learning neural networks for brain tumors. *Future Computing and Informatics Journal*, 3(1), pp. 68-71.

He K, Z. X. R. S. S. J., 2016. *Deep residual learning for image recognition.* USA, Las Vegas, IEEE conference on computer vision and pattern recognition.

Heidari, M. & Rafatirad, S., 2020. *Using transfer learning approach to implement convolutional neural network model to recommend airline tickets by using online reviews.* Zakynthos, Greece, 15th IEEE International Workshop on Semantic and Social Media Adaptation and Personalization (SMA).

Jignesh Chowdary, G., Punn, N., Sonbhadra, S. & Agarwal, S., 2020. *Face mask detection using transfer learning of inceptionv3.*. Sonepat, India, In Big Data Analytics, Proceedings of the 8th International Conference.

Koç M, Ö. R., 2019. Enhancing Facial Expression Recognition in the Wild with Deep Learning Methods Using a New Dataset: RidNet. *Bilecik Seyh Edebali University Journal of Science,* 6(2), pp. 384-396.

Liu, Z. et al., 2021. Deep learning framework based on integration of S-Mask R-CNN and Inception-v3 for ultrasound image-aided diagnosis of prostate cancer. *Future Generation Computer Systems,* Volume 114, pp. 358-367.

Loey, M., Manogaran, G., Taha, M. & Khalifa, N., 2021. A hybrid deep transfer learning model with machine learning methods for face mask detection in the era of the COVID-19 pandemic. *Measurement,* p. 108288.

Mabray, M. B. R. a. C. S., 2015. Modern Brain Tumor Imaging. *Brain Tumor Research and Treatment,* 3(1).

Mahmud, M., Mamun, M. & Abdelgawad, A., 2022. *A Deep Analysis of Textual Features Based Cyberbullying Detection Using Machine Learning.*. Maidu, Egypt, IEEE Global Conference on Artificial Intelligence and Internet of Things (GCAIoT).

Mamun, M. et al., 2022. *Vocal FeatureGuided Detection of Parkinson's Disease Using Machine Learning Algorithms.* New York, NY, USA,, IEEE 13th Annual Ubiquitous Computing, Electronics, and Mobile Communication Conference (UEMCON), .

Mamun, M. et al., 2022. *Can Machine Learning Techniques Enable to Predict Heart Diseases?.* New York, USA, IEEE 13th Annual Ubiquitous Computing, Electronics, and Mobile Communication Conference (UEMCON).

Mujahid, M. et al., 2022. Pneumonia Classification from X-ray Images with Inception-V3 and Convolutional Neural Network. *Diagnostics,* 12(5), p. 1280.

N. A. K. S. R. El Atrache, A. A. M. K., n.d. *Deep learning-based approach for diagnosing covid-19 using chest x-ray images.* s.l., AIP Conference Proceeding.

Nair, K. a. D. A. a. G. R. a. P. A., 2022. *Analysing X-Ray Images to Detect Lung Diseases Using DenseNet-169 technique.* [Online]
Available at: http://dx.doi.org/10.2139/ssrn.4111864

Powers, D. &. A., 2011. Evaluation: From precision, recall and F-measure to ROC, informedness, markedness & correlation. *J. Mach. Learn. Technol.,* p. 2229–3981.

Rahman, T. et al., 2020. Transfer learning with deep convolutional neural network (CNN) for pneumonia detection using chest X-ray.. *Applied Sciences,* 10(9), p. 3233.

Ranjbarzadeh, R. et al., 2021. *Brain tumor segmentation based on deep learning and an attention mechanism using MRI multi-modalities brain images.,* s.l.: s.n.

S. Sarkar, A. K. S. C. S. A. J.-S. S. H.-C. K., 2020. A CNN based approach for the detection of brain tumor using MRI scans. *Test Engineering and Management,* Volume 83, p. 16580 – 16586.

Simonyan, K. Z. A., 2014. *Very Deep Convolutional Networks for Large-Scale Image Recognition,* s.l.: arXiv technical report.

Simonyan, K. Z. A., 2014. *Very Deep Convolutional Networks for Large-Scale Image Recognition,* s.l.: arXiv technical report.

Tandel, G. et al., 2019. A review on a deep learning perspective in brain cancer classification.. *Cancers,* 11(1), p. 111.

Thenmozhi, K. & Reddy, 2019. U.S. Crop pest classification based on deep convolutional neural network and transfer learning. *Computers and Electronics in Agriculture,* p. 104906.

Tio, A., 2019. *Face shape classification using inception v3.* s.l.:arXiv.

Xia, X., Xu, C. & Nan, B., 2017. *Inception-v3 for flower classification.* Chengdu, China, 2nd IEEE International Conference on Image, Vision and Computing (ICIVC).

Zheng, Y. &. Y. C. &. M. A., 2018. Breast cancer screening using convolutional neural network and follow-up digital mammography.

Zhuang, F. et al., 2020. *A comprehensive survey on Transfer Learning.* s.l., IEEE.