American Journal of Psychiatry and Neuroscience
2021; 9(4): 147-156
http://www.sciencepublishinggroup.com/j/ajpn
doi: 10.11648/j.ajpn.20210904.11
ISSN: 2330-4243 (Print); ISSN: 2330-426X (Online)

Human Visualization of Brain Tumor Classifications Using Deep CNN: Xception + BiGRU

Ashley Seong

Seoul International School, Gyunggi-do, South Korea

Email address:
seong.ashley.0308@gmail.com

To cite this article:
Ashley Seong. Human Visualization of Brain Tumor Classifications Using Deep CNN: Xception + BiGRU. *American Journal of Psychiatry and Neuroscience*. Vol. 9, No. 4, 2021, pp. 147-156. doi: 10.11648/j.ajpn.20210904.11

Received: September 27, 2021; **Accepted**: October 9, 2021; **Published**: October 19, 2021

Abstract: Throughout the world, brain tumors have become a medical priority as more people suffer from this malignant disease worldwide. In the field of computer science, researchers have been studying to utilize MRI scans to its fullest potential, in recognizing signs of tumors early on, and utilizing computers and convolutional neural networks to process massive amounts of patient data at once in hopes of saving lives. This investigation finds out the specifications of visualization of MRI scans and how filters and layers are used to identify lethal tumors in the brain. For one of our main methods, a pre-trained model to improve accuracy was used - the Xception model. This showed a contrast between previous existing models as those fully connected layers were added to the back of existing ones. Our main proposed model of Xception + Bidirectional GRU had the highest accuracy of 82% out of 7 different models. In our proposed model, Convolutional layers were used to extract specific features of an image and process other similar images in the same way. By using 3 layers of Convolution, Activation, and Max pooling, we saw the networks focus on the actual tumors in the brain by distinguishing patterns in images and focusing on that area to create visual representations. Principal components of this research were the ability to visualize abnormal features of brain scan images to filter out and layer regions to bring attention to tumors in the brain.

Keywords: Brain Tumor, Deep CNN, Xception, BiGRU

1. Introduction

1.1. Background

Cancer has become a medical phenomenon that contributes to the highest death rates around the world. Incidence rates remain consistently high in high-income countries (HIC), but the prevalence of risk factors including obesity, smoking, and physical inactivity has led to low and middle-income countries (LMIC) to have high rates of cancer as well [1]. As a result of this increase, cancer has become a known threat. However, brain cancer has been and is considered one of the most lethal and malignant cancers in people of all ages because the nervous system works directly with the brain to control the entirety of bodily function [2]. Out of the various types of brain cancer, pituitary tumor, meningioma tumor, and glioma tumors are focused throughout this article.

Approximately 238,000 new cases of brain and central nervous system cancer are diagnosed annually [3]. Although brain and nervous system cancers account for 3% of all cancers in the world, they have a mortality rate of 3.4 per 100,000 people [4]. Despite having a low overall mortality rate, these tumors are also one of the most common tumors in adolescence (21%) and have become the first leading cause of cancer deaths for males aged under 40 years and females aged under 20 years [5].

As brain cancer becomes a prevalent issue around the globe, methods of receiving faster and more accurate identification have come into question. Magnetic Resonance Imaging (MRI) scans have been long used to look at structures inside human bodies. However, new fields in medical science have built neural networks to train artificial intelligence [6]. Researchers at NYU Grossman School of Medicine in collaboration with Facebook AI were able to significantly analyze to what extent AI can accelerate MRI scanning and processing. They were able to remove roughly three-fourths of raw data and generate fast MRI scans that matched the standard, slower MRI process. As the AI MRI scans required up to four times less data than the standard,

patient imaging was much faster and as a result spent less time in the actual MRI machines. Through this study, researchers were able to underscore the clear benefits of investing in artificial intelligence in the medical field regarding MRI scans.

Although mortality rates of cancer originating from the brain are relatively low, mortality rates caused by metastasis (spread of cancer to other body parts) to the brain is high. This is an area of concern as the brain is one of the body's organs where metastasis occurs at a frequent rate. As metastasis has limited effective treatment and is difficult to identify and diagnose, patients' median survival rates are only a few months [7].

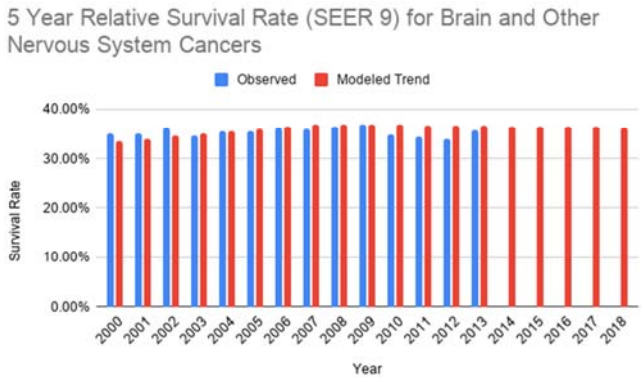


Figure 1. *5-year relatives survival rate for brain and other nervous system cancers.*

Figure 1 using a data set from (https://seer.cancer.gov/statfacts/html/brain.html), a graph was created comparing the 5-year survival rate (SEER 9) for brain and other nervous system cancers, with observed statistics versus the modeled trend. For years 2014-2018, observed data was not available, but the predicted model trends were given.

1.2. Objective

Our main objectives are to (i) adopt and incorporate deep learning techniques with pre trained Convolutional Neural Network models with fine tuning to fully quantify and classify brain tumor images, (ii) deliver them in a functioning high level of accuracy, (iii) bolster and underscore the efficiency of our method compared to traditional transfer learning and propagation neural network techniques, (iv) justify the use of 4 classifications: meningioma, glioma, pituitary tumors, and normal MRI scans compared to the typical 3 classifications, and (v) explore the usage of heatmaps in MRI scan images to portray anomaly sections of the brain. This paper will now explore contrasting related works, our material and methods used, results, discussions, and conclusions.

2. Related Works

Swato et al., have used a public CE-MRI data set (Cheng, 2017) to train convolutional neural networks for specific types of brain cancers. This article used a pre-trained deep CNN model and a block-wise fine-tuning strategy to evaluate CE-MRI datasets. They were able to achieve an average accuracy of 94.82% under five-fold cross-validation and used traditional machine learning incorporated with deep learning methods using CNNs. They classified three types of brain tumors: meningioma, glioma, and pituitary tumors [8].

Deepak et al., used a pre-trained GoogLeNet to identify and analyze MRI images of the brain. This experiment used a similar five-fold cross-validation process from an MRI dataset on figshare, outputting an accuracy of 98%. This paper specifically evaluated the system with fewer training samples and implied transfer learning as a useful technique in limited medical imaging. They classified three types of brain tumors: meningioma, glioma, and pituitary tumors [9].

Sumitra et al., suggested Neural Network techniques for classification of MRI of the human brain. The PCA and Neural Network technique utilized dimensionality reduction, feature extraction, and classification. The Back Propagation Neural Network classifier classified subjects as normal, benign, and malignant images. The accuracy for this method was ranging from 100% to 73%. BPN was used to train, test, and classify tumors for its fast-training speed [10].

Seetha et al., used Fuzzy C Means based segmentation, texture, and shape feature extractions. They further used SVM and DNN based classifications to result in tumor or normal brain images. CNN used a deep learning method, using image net database pre-trained models. This method showed the training accuracy to be 97.5% [11].

Afshar et al., found CNNs to require large amounts of data, therefore switching to capsule networking that proposed to revolutionize deep learning. Capsule networks were found to be robust to rotation and affine transformation and required less training data, specifically targeting CNN's flaws. The accuracy found for CapsNet imaging was 78%, while CNN's accuracy imaging was at 61.97%. Therefore, the Capsule networks efficiently overcame the shortcomings of CNN [12].

3. Materials & Methods

3.1. Data Description

Data set used involved two sections: testing and training. With 4 classifications, there were glioma tumor, meningioma tumor, pituitary tumor, and no tumor. The testing files contained 100 files for glioma tumor, 115 files for meningioma tumor, 74 files for pituitary tumors, and 105 files for no tumors. For the training set, glioma tumor carried 826 files, meningioma tumor carried 822 files, pituitary tumor had 827 files, and no tumor carried 395 files. Overall, 3264 files were used in our data set. This data set can be analyzed and credited in this link: https://www.kaggle.com/sartajbhuvaji/brain-tumor-classification-mri?select=Training [13].

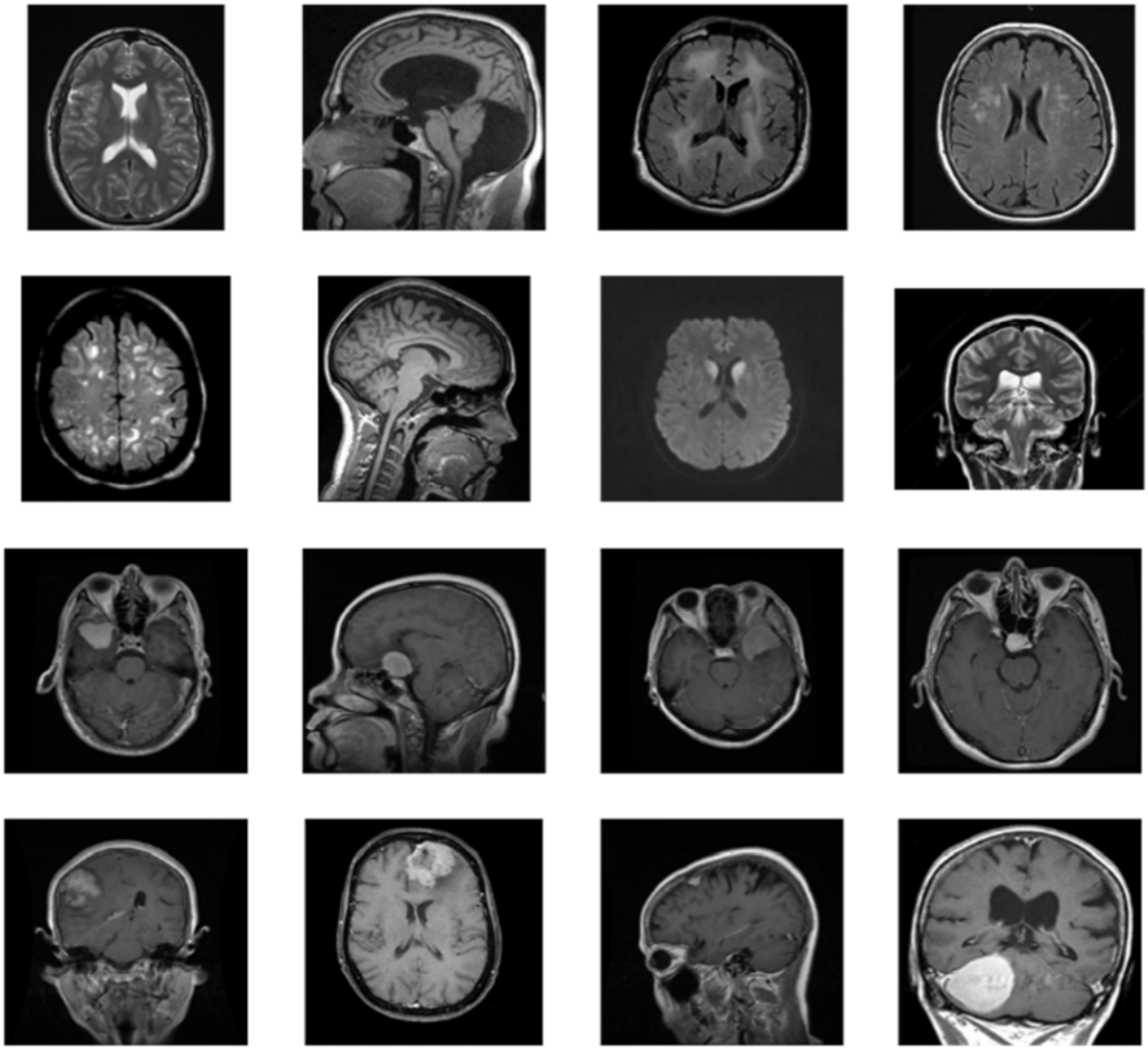


Figure 2. *Data visualization from the given dataset, downloaded from Kaggle website.*

3.2. Data Preprocessing

Since the size of the data is relatively insufficient to train a deep learning model, we had to multiply the data before putting it into the model and running it. For efficient data augmentation, we used Keras' ImageDataGenerator function. Through ImageDataGenerator, shear range, zoom range, horizontal flip, vertical flip, rotation range, width shift range, height shift range, etc. can be adjusted. All images were divided by 255 for normalization, and 30% of the training set was used as the validation set.

3.3. Convolutional Neural Network (CNN)

A Convolutional Neural Network (CNN) consists of a convolution layer, a pooling layer, and a fully connected layer. When CNN gets an input image, it first creates a convolution layer through a filter and produces a feature map, which is called a kernel. After that, the pooling layer reduces the size of the feature map by calculating the average or maximum value of the feature map. These are called max pooling and average pooling, respectively. A fully connected layer is the same as a deep neural network; the main purpose of this layer is to classify objects with activation functions. For multi-class classification, the softmax function is used as the activation function; for binary classification, the sigmoid function is mainly used [14].

3.4. Pre-trained CNN

To extract features from images, we used pre-trained CNN models such as VGG16, VGG19, MobileNet, Inception-Resnet_v2, and Inception_v3. These pre-trained models can be downloaded from Keras and were pre-trained on a dataset named ImageNet. Because CNN layers such as pooling and conv layers are properly arranged and pre-trained with a large image set in advance, the accuracy is relatively higher than that of training through a general CNN model. In particular, these models extract features of images to be employed by the user, and the layers to be classified after they are defined by the user [15].

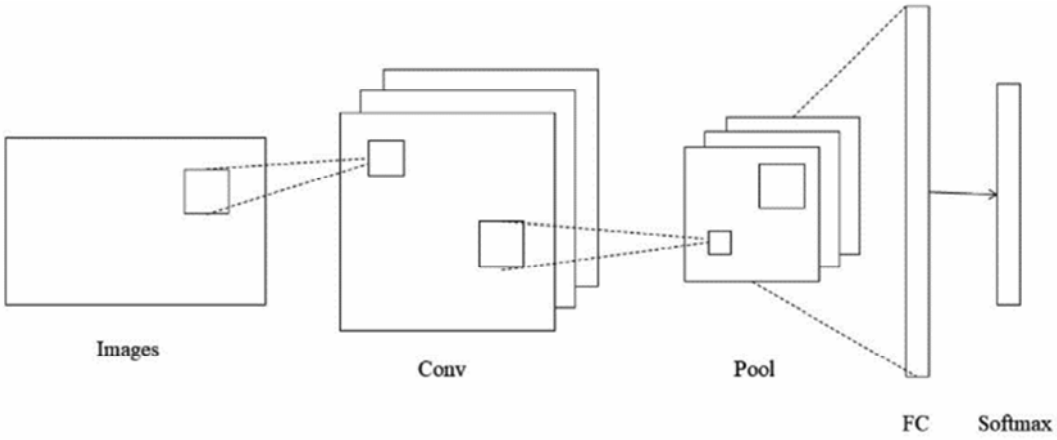


Figure 3. *Overall architecture of the CNN.*

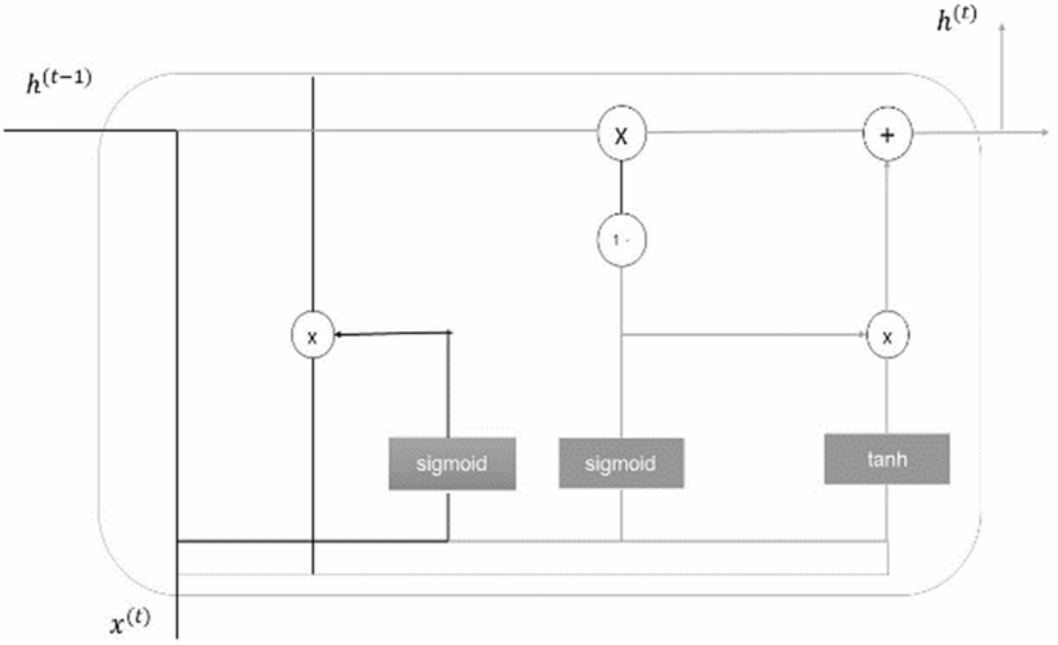


Figure 4. *Overall architecture of the GRU.*

3.5. Gated Recurrent Units (GRU)

$$z_t = \sigma(W_z[h_{t-1}, x_t] + b_z) \quad (1)$$

$$r_t = \sigma(W_r[h_{t-1}, x_t] + b_r) \quad (2)$$

$$\tilde{h}_t = \tanh(W_h[r_t \odot h_{t-1}, x_t] + b_h) \quad (3)$$

$$h_t = (1 - z_t) \odot h_{t-1} + z_t \odot \tilde{h}_t \quad (4)$$

In the case of LSTM, there were three gates: a forget gate, an input gate, and an output gate, but in the GRU, only two gates are used: a reset gate and an update gate. In addition, the cell state and hidden state are combined to express a single hidden state. The formula to find the reset gate corresponds to Equation (2) in the formula above. This is a method used to obtain the hidden state of the previous time and the x of the current time by applying the activation function sigmoid. The result will have a value between 0 and 1, which can be interpreted as information about how much to use the value of the previous hidden state. The value from the reset gate is not used as it is but is reused by expression (3). In equation (3), it is calculated by multiplying the hidden state of the previous time by the reset gate. The update gate plays a similar role to the input and forget gates of LSTM, and the key is to obtain the ratio of how much past and present information will be reflected. As a result of Equation (1), z reflects how much current information will be used. And (1-z) reflects how much to use for past information. So, each role can be viewed as an input and forget gate of the LSTM, and finally, the hidden state of the output value at the present time can be obtained through Equation (4) [16].

3.6. Bidirectional GRU

A sequence processing model called a Bidirectional GRU (BiGRU) consists of two GRUs. One takes the input from a

forward direction while the other takes the input in a backward direction. It utilizes the bidirectional recurrent neural networks that only use input and forget gates in the entire process [17].

3.7. Grad CAM

After classifying the pre-trained CNN models to obtain accuracy, Gradient-Weighted Class Activation Map (Grad-CAM) was used to check which part had abnormalities. Grad-CAM provides the cause for the classification result, and uses Global Average Pooling (GAP), instead of the fully connected layer used before final classification in the existing CNN model. It is shown through the heat map; the purple part indicates normal while the red indicates abnormal parts [18].

3.8. Proposed Model

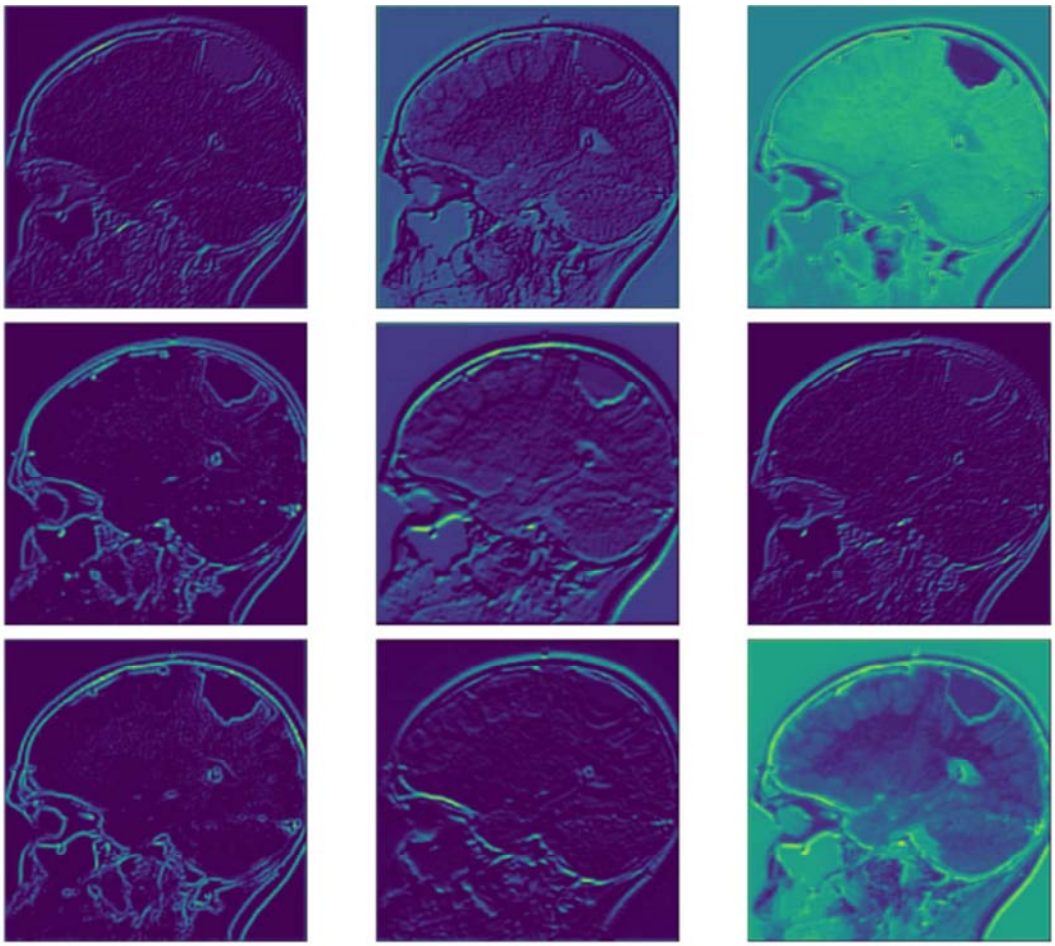
To improve classification accuracy, the Xception model - one of the pre-trained models - was used. In addition, bidirectional GRU was used for more accurate classification, contrasting with the existing models (in which a fully connected layer is added at the back of the model). Furthermore, a dense layer with 512 nodes was added as well as a dropout layer to prevent overfitting. The optimizer used nadam, the learning rate set at 0.001, and early stopping was set to stop training when the validation loss fell.

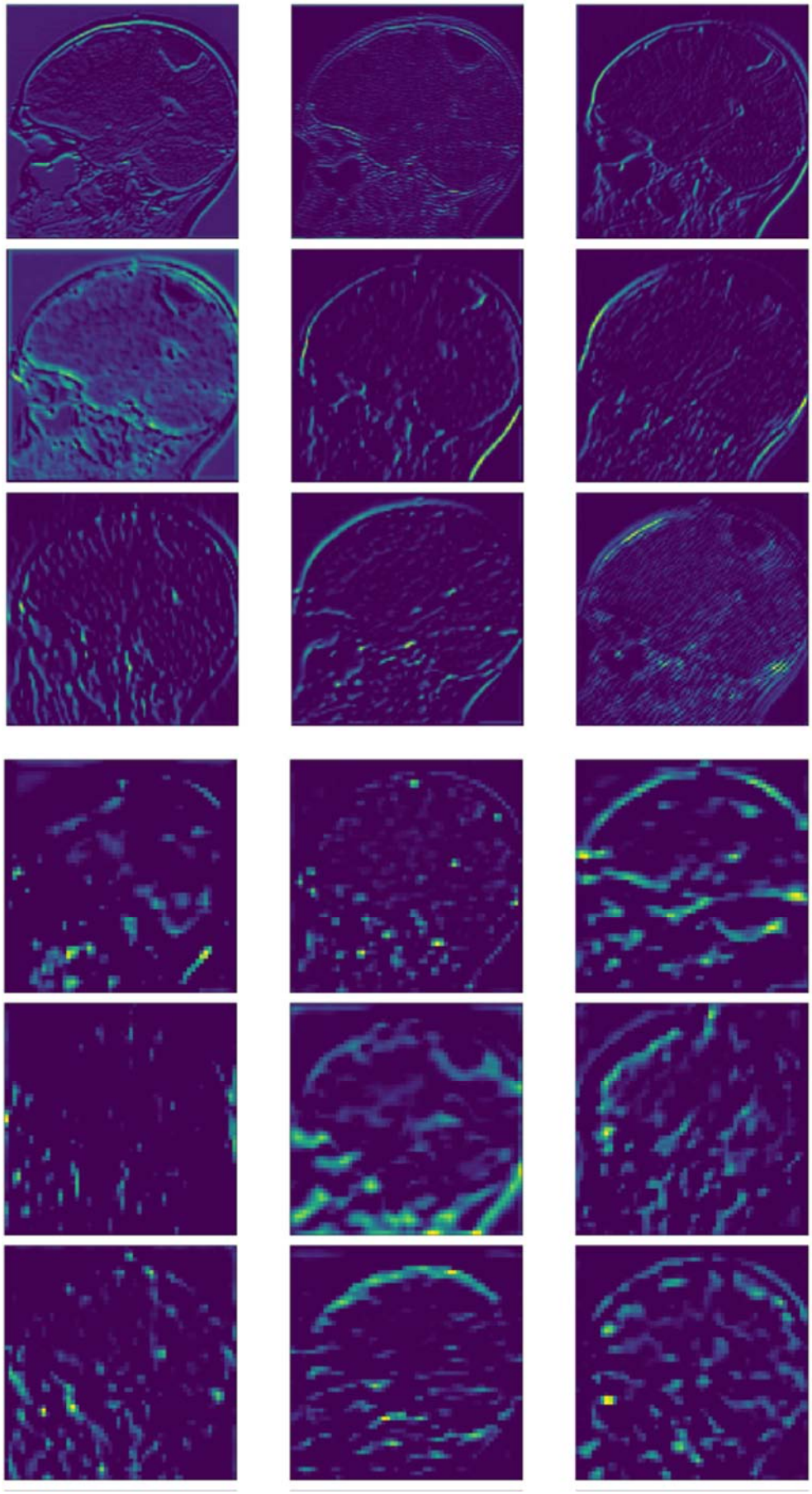
4. Results

4.1. Visualization of Convolution Layers

The first layer of CNN is to gather a collection of different types of edge detectors. Almost all of the information in the initial photo is preserved during this stage of activation. As the layers move up, activation becomes more abstract and visually difficult to understand. The representation of the upper layers shows less information about the visual content of the image but more information about the class of the image. In the first layer, all filters are active on the input image, but as layers move up, the filters become inactive. This means that the pattern encoded in the filter did not appear in the input image.

This shows some of the important features that deep neural networks typically exhibit in learned representations. The features extracted from the layer become more and more abstract along each depth of the layer. The activation of higher floors results in less and less visual information about a particular input, and more of a list of targets. Deep neural networks behave like a pipeline of information cleansing over the source data being inputted. Repetitive transformation filters out irrelevant information, and useful information is highlighted and improved.





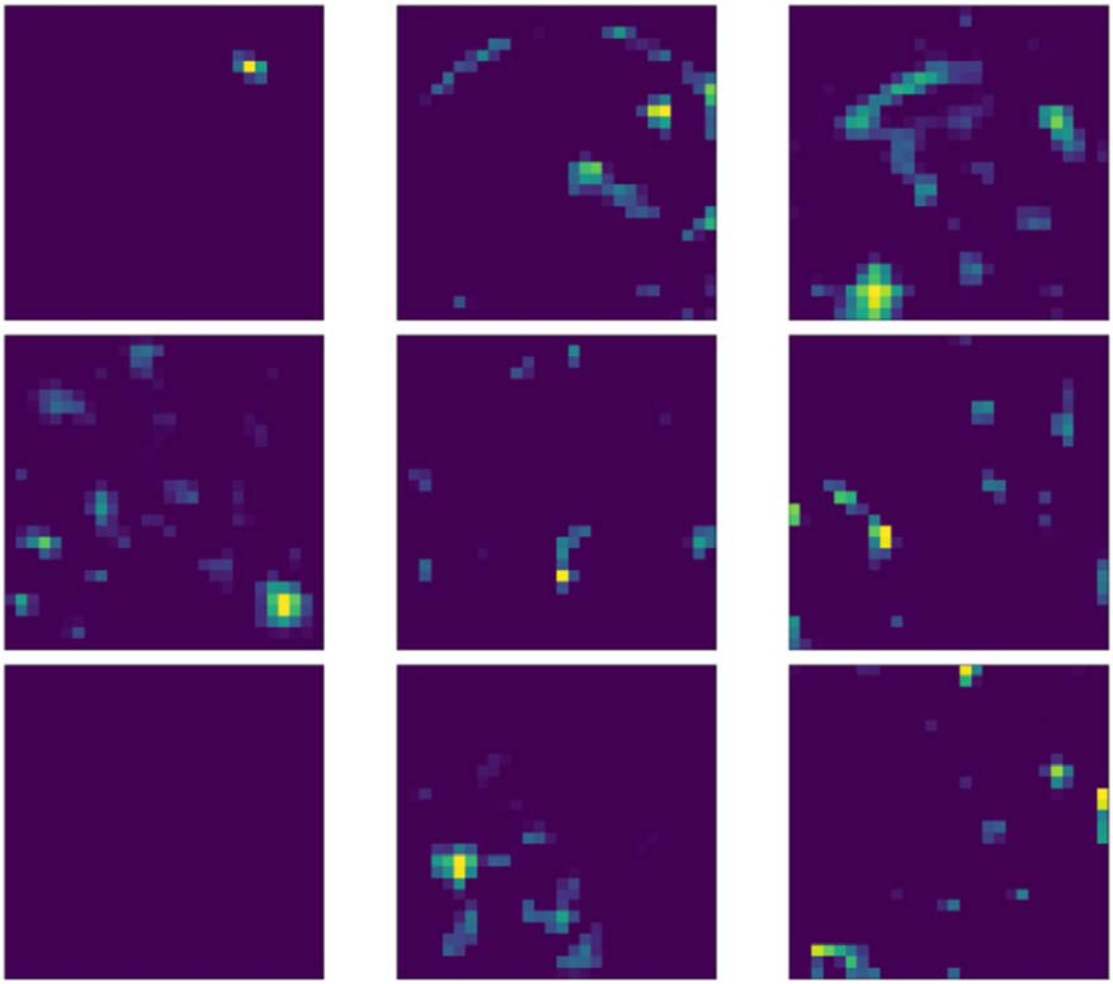


Figure 5. *Visualization of the convolution layers; shallow layer to deeper layer.*

4.2. Visualization of Pooling and Activation Layers

Convolutional layers were used to extract core features of an image and further use these distinct features to identify images that contained features of the same sort. In this run, we used 3 layers of CNN including Convolution, Activation, and Max Pooling. With these three layers, we started to see the network focus on regions such as the meningioma tumor in the actual brain. These types of features would allow the CNN through deep learning to differentiate between meningioma, pituitary, and glioma tumors. These neural networks are able to distinguish patterns in images that become akin to what the human eye can do, in focusing on one area and region to create a visual representation.

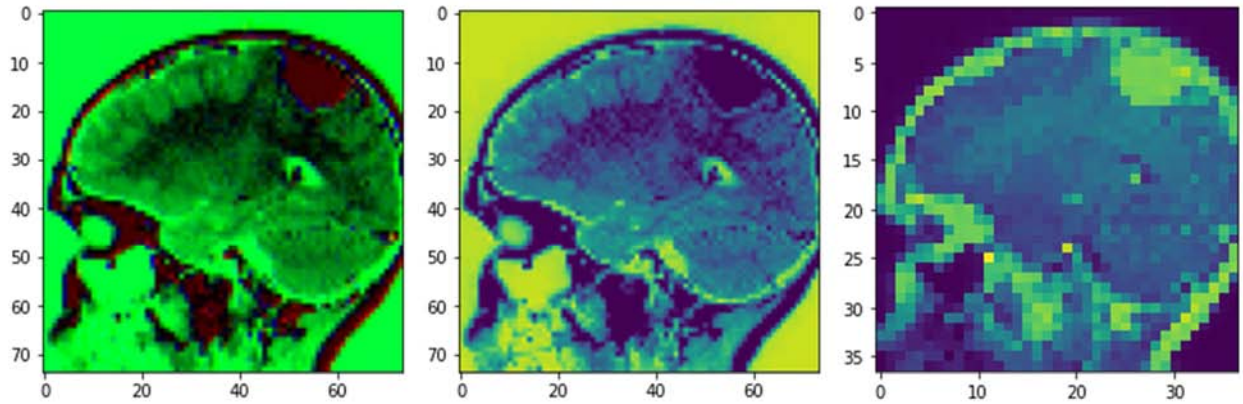


Figure 6. *Visualization of the convolution layers; shallow layer to deeper layer.*

4.3. Accuracy of Proposed Model

Figure 7. *Graph for accuracy comparison; proposed model and other deep learning models.*

In our main proposed model of Xception + Bidirectional GRU, the accuracy was the highest with an accuracy of 82.00%. The model with the second highest accuracy was Xception with a 78.00% accuracy. In comparison, VGG 16 and VGG 19 were around the 71.00% accuracy range. The lowest model accuracy was 44.70%, coming from MobileNet. Inception_V3 and DenseNet had similar accuracy percentages with 74.00% and 73.50% respectively.

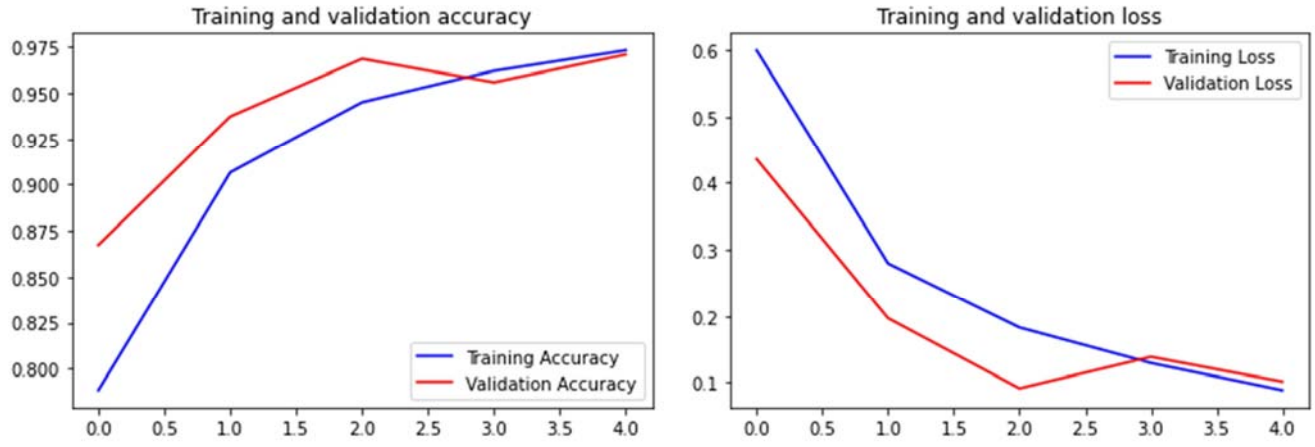


Figure 8. *Graph for loss and accuracy from training and validation sets.*

Looking at the two graphs above, it can be seen that during model training, the accuracy of the training set and the accuracy of the validation set increased almost continuously, reaching about 97.5%. It is shown that the loss of the training set and the loss of the validation set also decreased continuously during model training. However, it was found that overfitting occurred because the accuracy in the actual test set was about 82%.

As shown by the red heat mapping in this brain scan, it accurately aligns with the part of the meningioma tumor, the white oval shaped tumor. This shows the validity in this proposed model used in utilizing heat maps.

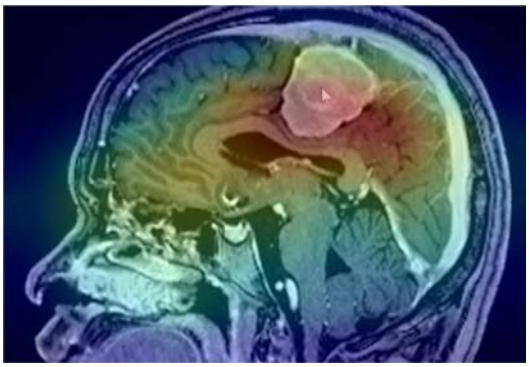


Figure 9. *Result of grad-CAM: meningioma tumor.*

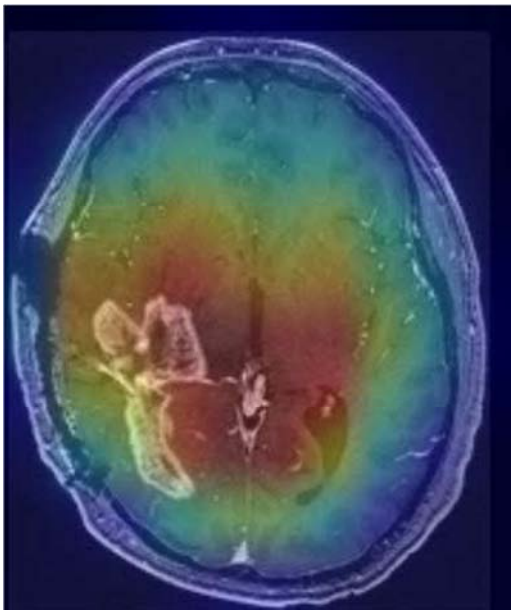


Figure 10. *Result of grad-CAM: glioma tumor.*

In this figure, the glioma tumor can be seen as the white area, which the heat map was in the vicinity of. As this figure does not show the complete strength of the heat maps, this was one example of a weakness in this proposed model.

Figure 11. *Result of grad-CAM: pituitary tumor.*

In this figure, the heat map centered in the area towards the pituitary gland, showing the efficacy of this proposed model. Although the gland itself is not opaquely visible, the heat map shows the suggested region.

5. Discussion

5.1. Principal Finding

A principal component to this research was the visualization of abnormal and malignant aspects in MRI Brain scans that was not seen in previous related works. By using heat maps to visualize the specific areas of tumors in the brain, comparisons can be drawn to the related works "Brain tumor classification using deep CNN features via transfer learning" and "Brain tumor classification using back propagation neural network". Although these related works used similar techniques and applications to show results of neural networking, they lacked the element of heat mapping and visualization that brings this research to another layer. This research was able to imply and underscore the similar steps of Convolutional Neural Networking to that of the human eye; As the features extracted from layers improved and became more abstract, the activation of higher floors became less about visual information and more about a list of targets.

5.2. Limitation

Although this research was able to find key principal components of CNNs, there were also some limitations. One of these came from an accuracy that was not able to surpass 90%. This is seen as a limitation as it does not show reliable data throughout this neural network. A second limitation came from the fact that we used the simplest classification to network; in the computer vision field, classification, object detection, and segmentation are often 3 key factors, but we were only able to use classification as our data factor. Segmentation methods such as Unet and FCN (Fully Convolutional Network) were not utilized because we were unable to use a mask.

6. Conclusion

Through continued analysis and research, we concluded that neural networking did follow a pattern that showed visualization of the human eye. As more layers of CNN were added, filters became inactive, and patterns encoded thus did not appear in input images. This showed that important features of deep neural networking exhibited learned representations. By using 3 layers of CNN, Convolution, Activation, and Max Pooling, these layers were able to focus on the actual regions of tumors in the brain. The proposed model of Xception + Bidirectional GRU thus had the highest accuracy of 82%.

Features and layers were an important step in this study as we found out activation of higher floors focused on making visual information a target-based system. However, further

study is necessary to prove this data to be valid and reliable, as limitations were that accuracy was less than 90% and we were only able to utilize classification as our data factor. In further studies, it is important to use a mask that would allow us to use segmentation and increase the accuracy of our proposed models to over 90%.

References

[1] Torre, Lindsey A., Rebecca L. Siegel, Elizabeth M. Ward, and Ahmedin Jemal. "Global Cancer Incidence and Mortality Rates and Trends—An Update." Cancer Epidemiology Biomarkers & Prevention 25, no. 1 (2015): 16–27. https://doi.org/10.1158/1055-9965.epi-15-0578.

[2] Abou-Antoun, Tamara J., James S. Hale, Justin D. Lathia, and Stephen M. Dombrowski. "Brain Cancer Stem Cells in Adults and Children: Cell Biology and Therapeutic Implications." Neurotherapeutics 14, no. 2 (2017): 372–84. https://doi.org/10.1007/s13311-017-0524-0.

[3] L. Saenz del Burgo, R. M. Hernández, G. Orive, J. L. Pedraz, Nanotherapeutic approaches for brain cancer management. Nanomedicine. 10, 905–919 (2014).

[4] Saenz del Burgo, Laura, Rosa María Hernández, Gorka Orive, and Jose Luis Pedraz. "Nanotherapeutic Approaches for Brain Cancer Management." Nanomedicine: Nanotechnology, Biology and Medicine 10, no. 5 (2014). https://doi.org/10.1016/j.nano.2013.10.001.

[5] Siegel, Rebecca L., Kimberly D. Miller, and Ahmedin Jemal. "Cancer Statistics, 2020." CA: A Cancer Journal for Clinicians 70, no. 1 (2020): 7–30. https://doi.org/10.3322/caac.21590.

[6] "New Research Finds FastMRI Scans Generated with Artificial Intelligence Are as Accurate as Traditional MRI." NYU Langone News. Accessed June 30, 2021. https://nyulangone.org/news/new-research-finds-fastmri-scans-generated-artificial-intelligence-are-accurate-traditional-mri.

[7] Obenauf, Anna C., and Joan Massagué. "Surviving at a Distance: Organ-Specific Metastasis." Trends in Cancer 1, no. 1 (2015): 76–91. https://doi.org/10.1016/j.trecan.2015.07.009.

[8] Swati, Zar Nawab, Qinghua Zhao, Muhammad Kabir, Farman Ali, Zakir Ali, Saeed Ahmed, and Jianfeng Lu. "Brain Tumor Classification for MR Images Using Transfer Learning and Fine-Tuning." Computerized Medical Imaging and Graphics 75 (2019): 34–46. https://doi.org/10.1016/j.compmedimag.2019.05.001.

[9] Deepak, S., and P. M. Ameer. "Brain Tumor Classification Using Deep CNN Features via Transfer Learning." Computers in Biology and Medicine 111 (2019): 103345. https://doi.org/10.1016/j.compbiomed.2019.103345.

[10] Sumitra, N., and Rakesh Kumar Saxena. "Brain Tumor Classification Using Back Propagation Neural Network." International Journal of Image, Graphics and Signal Processing 5, no. 2 (2013): 45–50. https://doi.org/10.5815/ijigsp.2013.02.07.

[11] Seetha, J., and S. Selvakumar Raja. "Brain Tumor Classification Using Convolutional Neural Networks." Biomedical and Pharmacology Journal 11, no. 3 (2018): 1457–61. https://doi.org/10.13005/bpj/1511.

[12] Afshar, Parnian., Arash Mohammadi, and Konstantinos N. Plataniotis. "Brain Tumor Type Classification via Capsule Networks." 2018 25th IEEE International Conference on Image Processing (ICIP), 2018. https://doi.org/10.1109/icip.2018.8451379.

[13] Sartaj. "Brain Tumor Classification (MRI)." Kaggle, May 24, 2020. https://www.kaggle.com/sartajbhuvaji/brain-tumor-classification-mri?select=Training.

[14] Yamashita, Rikiya, Mizuho Nishio, Richard Kinh Do, and Kaori Togashi. "Convolutional Neural Networks: An Overview and Application in Radiology." Insights into Imaging 9, no. 4 (2018): 611–29. https://doi.org/10.1007/s13244-018-0639-9.

[15] Marmanis, Dimitrios, Mihai Datcu, Thomas Esch, and Uwe Stilla. "Deep Learning Earth Observation Classification Using ImageNet Pretrained Networks." IEEE Geoscience and Remote Sensing Letters 13, no. 1 (2016): 105–9. https://doi.org/10.1109/lgrs.2015.2499239.

[16] Chung, Junyoung., Gulcehre, Caglar, Cho, KyungHyun, & Bengio, Yoshua (2014). Empirical evaluation of gated recurrent neural networks on sequence modeling. arXiv preprint arXiv: 1412.3555.