Brain Tumor Classification Using Convolutional Neural Networks

Morales Pérez Diego José[1], Ismael Eliezer Pérez Ruiz[1]

[1] Universidad Modelo, Mérida Yucatán, México

Abstract. Whether malignant or benign, brain tumors pose significant diagnostic challenges due to their impact on critical neural structures. This study evaluates the effectiveness of using Convolutional Neural Networks (CNNs) to assist in the detection and classification of brain tumors from images generated by Magnetic Resonance (MRI). The study used a publicly available dataset from Kaggle comprising 5,712 labeled images across four categories (glioma, meningioma, pituitary tumor, and no tumor) to train, validate, and test three CNN architectures: ResNet50, VGG16, and VGG19. Each model underwent the same preprocessing steps and training configuration to ensure a fair comparison. Evaluation metrics included accuracy, loss, confusion matrices, ROC curves, and precision-recall curves. VGG19 achieved the highest accuracy (98.95%) and exhibited the greatest stability among the models, followed by VGG16. ResNet50 demonstrated lower performance and greater variation in loss values. These results imply that incorporating CNN-based tools into medical diagnostics could improve the early detection and treatment planning of brain tumors, but should not replace medical professionals altogether. Further optimization and testing on larger datasets are recommended to validate generalizability and facilitate real-world clinical implementation.

Keywords: Brain Tumor, CNN, Weights, Nodes.

1 Introduction

Brain tumors are defined as groups of cells that are clustered within or in close proximity to the brain. These tumors can develop and form directly on the brain or in the vicinity of brain tissue. Consequently, they can be categorized based on their relative position within the patient's head. The categorization of brain tumors is typically based on their anatomical origin and biological characteristics. These include gliomas, which are tumors that form from glial cells and are often found in the spinal cord or brain; meningiomas, which form near or in the meninges; and pituitary tumors, which develop near the pituitary gland.[1][2]

According to the Preston Robert Tisch Brain Tumor Center, the global incidence rate of brain and other central nervous system (CNS) tumors is approximately 6.2 per 100,000 people per year. This number represents both malignant and benign tumors.[2]

A convolutional neural network (CNN) is a type of neural network trained to identify and recognize patterns in images for classification purposes. CNNs are composed of multiple layers, which are composed by filters that identify distinct characteristics of an image. The filtered images pass through, and the CNN can be adjusted by modifying the weights and biases for more significant results.[5]

Atika Akter et al. (2023) developed a customized deep convolutional neural network based on UNet capable of detecting, segmenting, and clasifying three types of brain tumors: glioma, meningioma and pituitary. The network was also able to indicate whether a tumor was malignant or benign. The images were first segmented using the UNet architecture, then passed through a preprocessing portion and finally, they were classified The classification process of this deep convolutional neural network (CNN) was executed by five pre-trained models. The pre-trained models employed in this study included VGG16, VGG19, EfficientNet B0, EfficientNet B7, and ResNet152V2. The results obtained from VGG16, ResNet152v2, and EfficientNet B0 were comparable, with all three models attaining a rate of 99 percent accuracy.[6]

T. Balamurugan et al. (2022) developed a customized deep convolutional neural network based on LuNet. This network is capable of detecting, segmenting, and classifying brain tumors as either malignant or benign. The classification process of this deep CNN, based on LuNet, was executed using the gray-level co-occurrence matrix (GLCM) and VGG16. A comparison was then made with AlexNet and ResNet.[7]

Muhammad Arif et al. (2022) developed a convolutional neural network (CNN) based on a genetic algorithm for classification and a UNet for segmentation. They compared the classification method using a genetic algorithm with multiple other CNNs, including VGG16, GoogLeNet, and AlexNet, to measure their performance against each other. As a result, the customized genetic algorithm-UNet hybrid far outperformed any other CNN it was compared against.[8]

2 Methodology

The methodology of this project consists of four stages. The first stage is dataset selection. The second stage is preprocessing the dataset images. The third stage is the architecture selection, training and evaluation. The fourth stage is comparing each architecture.

2.1 Dataset Selection

Research for the dataset was conducted using the Kaggle website, which has a wide variety of projects, datasets, and documents related to medical scenarios and conditions.[9] The Brain Tumor MRI Dataset, created by Masoud Nickparvar, was used to train, validate, and test different architectures. It contains 7,023 images, labeled by professionals, of human brains classified into four categories: glioma, meningioma, no tumor, and pituitary tumor. However, the quantity of this data had to be reduced to prevent overtraining of the CNNs, resulting in a total of 5,712 images.

2.2 Preprocessing of the Dataset's Images

Starting with the preprocessing, the image size was modified to 225x225 to prevent overloading the CNNs while maintaining image quality and data integrity. Then, the data was organized into batches of 30 images each. Lastly, the balanced weights function was used to obtain the weights of each class in the dataset.

Using the data.take function from the numpy library[14], we specified the percentages of data, they were 75% for training, 10% for validation and 15% for testing

2.3 Architecture Selection, Training, Validation and Testing

The architectures to be evaluated in this project were selected based on a review of the literature. The selected architectures are:

ResNet50: ResNet50 is a 50-layer convolutional neural network (CNN) that uses residual blocks to simplify deep CNN training. This architecture has proven to be highly effective for tasks such as image classification, consisting of 50 layers of convolutional, grouping, and an exit layer with a softmax trigger for image classification; the convolutional layers work by analyzing the characteristics of images, such as texture, shapes, and borders. The grouping layers combine the outputs of the convolutional layers, thus reducing the computational power required to train the CNN.[10]

VGG16: The VGG16 consists of 16 layers: 13 convolutional layers with ReLU functions and three fully connected layers. These fully connected layers process the data obtained from the convolutional layers to produce an output that uses a softmax function to classify images.[11]

VGG19: The VGG-19 is a 19-layer convolutional neural network and a direct upgrade to the VGG16 architecture. The VGG19 CNN has 13 convolutional layers with ReLU functions at the end of each layer and three fully connected layers. These layers process data obtained by the convolutional layers for the output layer, which uses a softmax function to classify images.[12]

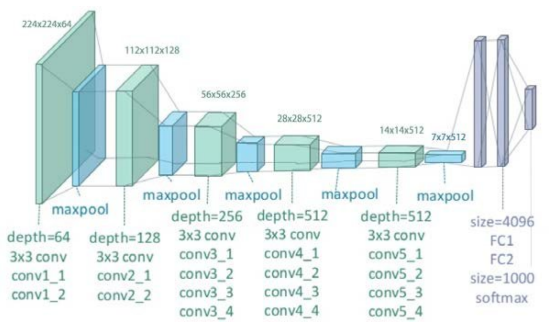

Fig. 1. Architecture of the VGG19 model, 16 convolutional layers and 3 fully connected layers

After loading the models, the top layer is replaced with a layer built from scratch. The input size was specified using the shape of our image and the last layers of the model were made trainable, in that way, the values of the CNNs can be modified directly. Then, a Global Average Pooling 2D layer was added for categorization, along with a normal dense layer, a dropout layer, and an exit layer [13]. All of our classes were loaded into the exit layer and a softmax activation function was used to detect said classes.

After compiling our model, the training and validating is done within 10 epochs. Then, the data from the test and validation portions was uploaded upon the model, as well as the weights for each class. With this data, the CNNs can execute the training and validation stages of the project.

After finishing the training and validation sections, the test portion of our dataset was uploaded into the model. With its weight values, tests of the data were made on images that the model has never seen before. This makes the results more reliable and valid.

2.4 Comparison of the Convolutional Neural Networks

The performance of the models were evaluated via graphs using the data obtained from the training, validation, and testing sections. In this case accuracy, loss, the confusion matrix, the ROC curve, and the precision-recall curve will be used to evaluate the models.

After obtaining the data and graphs for each convolutional neural network (CNN), the results will be compared to select the best performer.

3 Results

3.1 Accuracy and Loss

In the Accuracy and Loss section of the results, the ResNet50 model achieved an accuracy score of 88.60% in the testing portion and 81.93% in the validation portion. Meanwhile, it obtained loss values of 0.3091 and 0.6472 in the training and validation portions, respectively. However, the loss values of the model varied and regressed significantly during the compilation phase, making the model less reliable.

The VGG16 model achieved an accuracy score of 98.60% in the testing phase and 95.79% in the validation phase. It obtained loss values of 0.0395 and 0.1162 in the training and validation phases, respectively. The loss of the VGG16 model varied much less significantly than that of the ResNet50 model, meaning that the VGG16 model is more stable.

The VGG19 model achieved an accuracy score of 98.95% in the testing phase and 96.32% in the validation phase while obtaining loss values of 0.0365 and 0.1386, respectively. The VGG19 model had the least loss variation and regression, making it the most stable.

Of all the models, the VGG19 model showed the most accuracy and the least loss regression, making it the best model in this regard.

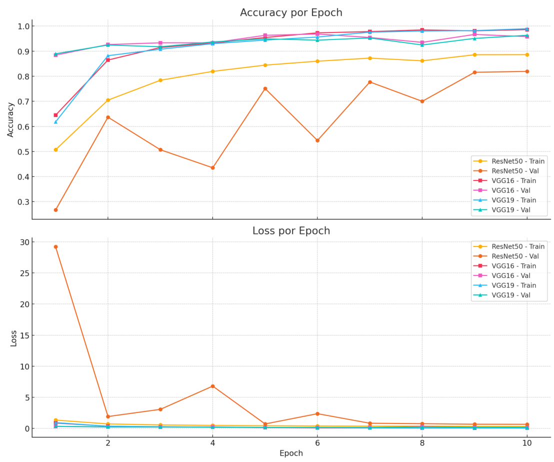

Fig. 2. Accuracy and Loss comparison graphs of the ResNet50, VGG16 and VGG19 models. In the accuracy graph, we can see the development respecting the values of accuracy of all CNNs. In the loss graph, we can see the values of loss of each CNN through the epochs

3.2 ROC And PR Curves

In the ROC curve portion of the results, the ResNet50 model acquired an area under the curve[15] (AUC) value of 91%, while the VGG16 model acquired an AUC value of 99%. And for the PR curve portion of the results, the ResNet50 model acquired an area under the curve (AUC) value of 91%, while the VGG16 model acquired an AUC value of 99%. While both the VGG16 and VGG19 models acquired similar, valid values, the ResNet50 model deviated considerably, eliminating it as a candidate for this criterion.

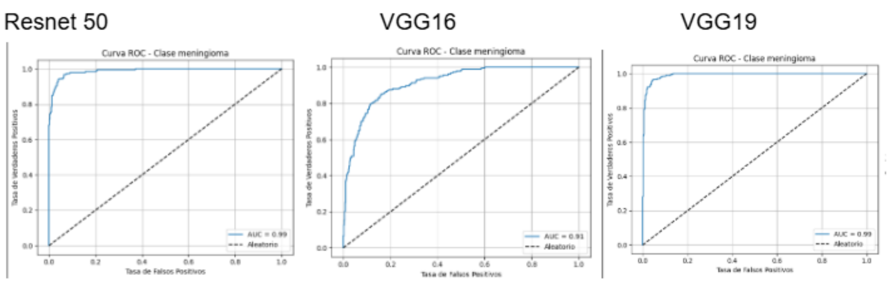

Fig. 3. ROC curve graphs of the ResNet50, VGG16 and VGG19 models. Here, we can see that the VGG16 and the VGG19 curves have a high value of True positives at the beginning, this means that the rate of positive detection is high without introducing false positives. In contrast, the Resnet 50 Curve has a relatively low value of true positives at the start.

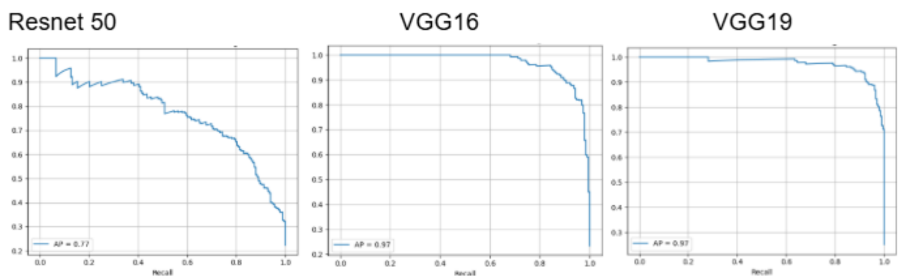

Fig. 4. PR curve graphs of the ResNet50, VGG16 and VGG19 models. Here, we can see that the VGG16 and VGG19 curves have a similar AUC, but the Resnet 50 Curve has a low AUC.

4 Conclusion

This project built, compiled and compared the results between the ResNet50, VGG16 and VGG19 CNNs for detection and classification of the types of brain tumors, based on the location relative to the brain and its characteristics, using a dataset located via Kaggle. Between all of them, the VGG19 CNN obtained the highest degree of accuracy, and a similar AUC in comparison with the VGG16 in both the ROC and PR curve, but way higher than the ResNet50 CNN. The results of the evaluation show that using CNN for detection and classification of brain tumors is feasible. An improvement upon the current accuracy, ROC and PR curve values is possible with the enabling and/or disabling of layers for training, validated by the various tests made upon the models, in which different numbers of layers were enabled or disabled, but a more in-depth analysis would be required. The implementation of CNN onto the medical workspace could significantly improve the detection rates of brain tumors, this for the early treatment of said tumors, but should not be used instead of the diagnosis given by a professional of healthcare. CNNs can be used to assist in the work of medical professionals, but they cannot substitute the experience and criteria used by the personnel, because, even though it is minimal, the CNNs by themselves can still make errors.

References

1. American Association of Neurological Surgeons. (2024, June 25). Brain Tumors - AANS. AANS. https://www.aans.org/patients/conditions-treatments/brain-tumors/
2. “How common is a brain tumor,” Tisch Brain Tumor Center, Jun. 26, 2024. https://tischbraintumorcenter-duke-edu.translate.goog/blog/how-common-brain-tumor?_x_tr_sl=en&_x_tr_tl=es&_x_tr_hl=es&_x_tr_pto=sge#:~:text=Los%20tumores%20cerebrales%20son%20bastante,tanto%20tumores%20malignos%20como%20benignos
3. Ibm. (2025b, April 16). Convolutional Neural Networks. What are convolutional neural networks? https://www.ibm.com/think/topics/convolutional-neural-networks
4. ¿Qué son las redes neuronales convolucionales? (n.d.). MATLAB & Simulink. https://la.mathworks.com/discovery/convolutional-neural-network.html
5. Mishra, M. (2025, January 27). Convolutional neural networks, explained. Towards Data Science. https://towardsdatascience.com/convolutional-neural-networks-explained-9cc5188c4939/
6. Akter, A., Nosheen, N., Ahmed, S., Hossain, M., Yousuf, M. A., Almoyad, M. a. A., Hasan, K. F., & Moni, M. A. (2023). Robust clinical applicable CNN and U-Net based algorithm for MRI classification and segmentation for brain tumor. Expert Systems With Applications, 238, 122347. https://doi.org/10.1016/j.eswa.2023.122347
7. Balamurugan, T., & Gnanamanoharan, E. (2022). Brain Tumor Segmentation and Classification using hybrid Deep CNN with LuNet Classifier. Research Square (Research Square). https://doi.org/10.21203/rs.3.rs-1599383/v1
8. Arif, M., Jims, A., F, A., Geman, O., Craciun, M., & Leuciuc, F. (2022). Application of genetic Algorithm and U-Net in brain tumor segmentation and Classification: a deep learning approach. Computational Intelligence and Neuroscience, 2022, 1–11. https://doi.org/10.1155/2022/5625757

9. Brain Tumor MRI dataset. (2021, September 24). Kaggle. https://www.kaggle.com/datasets/masoudnickparvar/brain-tumor-mri-dataset?select=Testing
10. Kundu, N. (2023, January 23). Exploring ReSNet50: An In-Depth look at the model architecture and code implementation. Medium. https://medium.com/@nitishkundu1993/exploring-resnet50-an-in-depth-look-at-the-model-architecture-and-code-implementation-d8d8fa67e46f
11. Hassan, M. U. (2023, May 9). VGG16 – Convolutional Network for Classification and Detection. Neurohive / Neural Networks. https://neurohive.io/en/popular-networks/vgg16/
12. https://www.geeksforgeeks.org/vgg-net-architecture-explained/
13. Team, K. (n.d.). Keras: Deep Learning for humans. https://keras.io/
14. https://www.geeksforgeeks.org/numpy-take-python/
15. Classification: ROC and AUC. (n.d.). Google for Developers. https://developers.google.com/machine-learning/crash-course/classification/roc-and-auc