Brain Tumor Classification using CCN

*Marcelo Delgado, †Isaac Madrigal, ‡José Antonio Mora and §Mariela Valerio
School of Computer Science
Universidad de Costa Rica, San Jose, Costa Rica
*marcelo.delgadomora@ucr.ac.cr
†isaac.madrigalsilva@ucr.ac.cr
‡jose.moramonge21@ucr.ac.cr
§ mariela.valerio@ucr.ac.cr

***Abstract*—This paper investigates the use of Convolutional Neural Networks (CNN) for classifying brain tumor images to support the precision of medical diagnosis. The developed classification model differentiates among four categories: non-tumor, glioma, meningioma, and pituitary tumors. Using a Kaggle dataset containing 2,870 training images and 394 test images, the CNN model was trained and evaluated to optimize the prediction accuracy. This model seeks to reduce diagnostic errors in medical imaging and assist healthcare professionals by improving image analysis reliability. The findings show the potential of CNN to improve tumor classification, which could have a positive impact on diagnostic precision and treatment planning. However, there are still limitations that need to be addressed and worked on, such as the ability of the model to generalize on new, unseen data, a common problem with machine learning models. This suggests that while these types of models have the potential to aid in the medical field, they should complement rather than replace trained professionals, who can use these tools to provide a more accurate and final diagnosis, due to the critical nature of the practice in the medical field.**

I. Introduction

Artificial Intelligence or AI has recently risen in the computer science field, constantly demonstrating its potential and the benefits it brings to modern computing and daily life. This discipline gives computers human-like capabilities to solve complex problems that traditionally computers were not able to. One of the current challenges that benefit the most from the constant improvement in AI is the identification and classification of images, a task that is usually easy for the human eye and mind, but complex for computers until the implementation of AI [1].

Classification algorithms and models have gained increasing interest due to the range of applications and fields where they can serve a purpose: medicine, vehicle navigation, and robotics are just some of the areas benefited from them. The process of Classification involves obtaining a great quantity of data, called a dataset, to train an AI model to find or classify patterns. Then, it involves a process of preprocessing, where images are prepared and converted to data that can be interpreted and compared by computers. Then, a technique or algorithm is used to make decisions and classify the images. Within these techniques there are the traditional ones, using machine learning, that tend to have acceptable results, but they also tend to be outperformed by a great margin by more recent techniques that use Deep Learning, with the disadvantage that these techniques need a lot more information to train the model successfully. Within the Deep Learning techniques, there is the Neural Networks, that is the most successful technique for image classification by a great margin. Lastly, the model must be tested, using test data that was excluded from the training dataset to verify the precision and accuracy of the model for the identification and classification of images [2].

The principal objective of this project is the classification of brain tumor images by the use of Convolutional Neural Networks (CNN's) to help reduce and prevent errors from doctors on data analysis in medical examinations resulting in misdiagnosis.

II. Theoretical Background

A. AI and Machine Learning

To understand the significance of Convolutional Neural Networks in medical image classification, it is important to first explore the fundamental concepts of artificial intelligence and machine learning, including their definitions and functionalities. Both Artificial Intelligence and Machine Learning have a lot of theoretical information worth talking about. In summary, AI refers to the capability of a machine to imitate intelligent human-like behavior, and it consists of various fields, one of them being Machine Learning. Machine learning enables a computer to use an algorithm to learn from and make predictions based on data. In terms of image recognition, Machine Learning revolutionized the field with its appearance by providing different methods to analyze visual data. The capacity to adapt and improve over time as data is provided to them enables complex image interpretation. There are various types of machine learning that vary in the way they learn; one of the standouts of this type is supervised learning, which involves training a model on labeled datasets and the model learns the relationship between the data and their label, facilitating the prediction or classification it makes. In contrast, there is unsupervised learning that does not include any labels and success is determined by the reduction or increase in a predefined cost function [3].

B. Convolutional Neural Networks

As a subfield of Machine Learning comes Deep Learning. This introduces the concept of artificial neural networks heavily inspired by the way biological nervous systems operate, they consist of multiple layers and are also known as deep networks. The basic concept behind Deep Learning and Neural Networks is the idea of neurons where the data is trained and modified, the combination of these neurons takes an input, and an output is predicted, generated or classified.

Neural networks have seen different implementations and modifications over the years, but a critical innovation came with the Convolutional Neural Network (CNN), which is designed for processing grid-like data such as images. Its hierarchical structure is inspired by the way the human brain processes visual information, making it capable of identifying patterns in a group of images. [4].

Convolutional Neural Networks are similar to traditional Neural Networks in the sense of neurons that receive inputs, perform operations and produce an output. It also still consists of one score function for the entirety of the network and a loss function. The main difference between the implementations is the ability to encode image-specific features into the architecture making CNN a more suitable option for image-related tasks. [4]

The architecture of a CNN consists of three main types of layers: convolutional layers, pooling layers, and fully connected layers.

A **Convolutional Layer** is responsible for detecting local patterns of the input data, edges in images for example, by performing a mathematical operation called a convolution, this operation involves ”sliding” filters known as kernels, which are matrices of size smaller than the full data, and computing the dot product between the kernel and the area of the input it overlaps and repeats this across the entire input. There are some hyper parameters important for the convolution operation like the **stride** which defines how far the kernel moves across the input and the **Padding** which is a process of adding extra pixels around the border of the input data to make its size more consistent. After the convolution operation, an activation function is applied to the feature that maps the convolution operation generated. This function is usually a non-linear function like **ReLu** which replaces all negative values in the maps with zeroes. [4]

The **Pooling Layers**, using a max-pooling technique, are in charge of reducing the complexity of the data and the number of parameters without sacrificing significant features with the goal of mitigating overfitting, which will be explained in detail later. [4]

Finally, there are **Fully-connected Layers** at the end which consist of neurons that like their name implies, are fully connected to the other layers without being connected to any layers within them. With this, they add high-level features that were extracted by the previous layers into predictions. [4]

One of the main challenges of Convolutional Neural Networks is the need for a lot of computational resources, especially when the input data is on the larger side of things, one of the solutions to this challenge is to use smaller filters divided across different convolutional layers or the use of *Parameter Sharing* which bases itself on the idea that if one region in the convolutional layer is useful to compute one bigger region, it might be useful in another region.

C. Artificial Intelligence in the Medical Field

All of the concepts mentioned above come together in modern artificial intelligence, which has experienced tremendous growth in recent years and has been integrated into various disciplines, greatly benefiting areas like machine learning. One field that has notably benefited from AI is medicine, where AI inclusion has led to a reduction, and in some cases even elimination, of certain human errors. Many of these errors, previously arising from data analysis involved in medical exam evaluations, often led to incorrect diagnoses. AI has significantly helped reduce such mistakes, improving accuracy and reliability in the medical field. [5]

On the other hand, machine learning applications extend beyond merely diagnosing conditions. As mentioned in [6], machine learning can be valuable for disease prognosis, enabling predictions about a patient’s possibility to develop certain conditions or relapse to previous diagnosis, and, eventually, it may even be used to forecast a person’s life expectancy.

III. Selected Task

Based on the previous theoretical background, this project aims towards the use of machine learning for the classification of brain tumor images. The accurate classification of these images impacts patient diagnosis and treatment decisions directly. Developing a reliable classification method would help the medical field in more informed decisions and cost reduction.

To accomplish this purpose, the goal is to train a Convolutional Neural Network model so that, given an image, it is able to identify the type of brain tumor that presents the brain shown. For this, the images will be divided into 4 groups:

- Non-tumor
- Glioma tumor
- Meningioma tumor
- Pituitary tumor

Included in the task there’s several important sub-tasks such as, the selection of a dataset containing images for training the model, image preparation to ensure consistency, training the model itself and evaluating the performance of the model.

IV. Implementation

First, the data to be used for the program was selected. This data was obtained from the Kaggle page **https://www.kaggle.com/datasets/sartajbhuvaji/brain-tumor-classification-mri**, from a sample of brain images with different pathologies. In total, the data consists of 2,870 images for model training (826 glioma, 822 meningioma, 395 without tumor, and 827 pituitary) and 394 images for the testing stage (100 glioma, 115 meningioma, 105 without tumor, and 74 pituitary).

As the first step in implementation, the libraries needed for program development were imported:

- TensorFlow for building and training the model.
- Numpy for numerical operations.
- Matplotlib and Seaborn for data visualization.
- Sklearn for model evaluation (confusion matrix).

A. Data Preparation

In this section, the paths are set for the images used in both training and testing. The data is then loaded using **tf.keras.preprocessing.image_dataset_from_directory**. With this function, the images are organized into batches of 32, resized to 256x256 pixels, and automatically labeled according to the folder they are located in.

B. Model Selection

In this part of the code, a sequential convolutional neural network (CNN) is built with the following characteristics:

- A **Conv2D** layer with 32 filters that analyzes the images to identify patterns, along with a **MaxPooling2D** layer to reduce the size of the features, focusing on selecting the most important information.
- A **Flatten** layer that converts the two-dimensional features into a one-dimensional array.
- A fully connected **Dense** layer with 256 neurons and **Leaky ReLU** activation with an alpha of 0.1.
- An output layer with 4 neurons (one per class) and **softmax** activation to generate classification probabilities.

Next, the model is compiled using the **SparseCategoricalCrossentropy** loss function, along with the **Adam** optimizer. Additionally, **accuracy** is used as a metric to monitor performance during the model training process.

It is important to note that various techniques were tested to identify the best model for the problem, including hyperparameter tuning. The techniques tested included:

- Changing the **number of neurons**: 128 and 256 neurons.
- Using different **activation functions** for hidden layers, which determine the output of each neuron: ReLU, Leaky ReLU with an alpha of 0.01, Leaky ReLU with an alpha of 0.1, and Parametric ReLU.
- Methods to reduce overfitting: **dropout**, which randomly deactivates a percentage of neurons in each layer during training. Also **regularization**, which adds a penalty term to the loss function.
- **Data augmentation**, to expand the training dataset by subtly modifying existing images, such as rotating them or zooming in on them.

However, dropout, regularization, and data augmentation were not used in the final model, as applying any of these techniques considerably worsened the model's performance.

C. Model Training

The training is performed with 50 epochs. Early stopping is also used to stop the training process if validation accuracy does not improve for 5 consecutive epochs, restoring the best weights.

D. Model Deployment

Once the model has been trained, it can be saved if the corresponding code line is enabled (**model.save("models/trained_model_4")**). Similarly, the model can be loaded if there is already a trained version.

E. Evaluations and Results

After training, the model makes predictions on the test image set using the **model.predict()** function. This testing method generates a list of probabilities for each class, where the class with the highest probability is selected for the final prediction.

Finally, a confusion matrix is used to evaluate the model's performance. This matrix helps verify the number of correct and incorrect predictions and provides information on where the model made errors. For this, the confusion matrix from **sklearn** was used.

V. Results

For the evaluation of the results of the model, some of the most important metrics provided in [7] will be used and explained.

In order to validate a machine learning model, its prediction performance is evaluated by providing an unbiased estimate of the model performance using never before seen data that was excluded from the training phase, called the testing data. To evaluate the model, different sorts of metrics are used to quantify the performance of the machine learning model. It is important to provide and evaluate different metrics, as there is no one single metric that can accurately define the quality of a model. For classification tasks, the most common metrics utilized are Accuracy, Precision, Recall and Loss.

The results for the metrics on the trained model can be summarized as the following:

- **Accuracy:** It denotes the fraction of the samples that were correctly classified out of all predictions made. The model achieved an accuracy of 0.76.
- **Precision:** Also called Positive Predictive Value, it denotes the fraction of the positively classified samples that are actually positive. The model achieved a precision of 0.76.
- **Recall:** Also called Sensitivity, it denotes the fraction of positive samples that were correctly retrieved. The model achieved a recall of 0.76.
- **Loss:** It denotes how wrong a model's predictions are. The loss of the model is 96.16.

Even though machine learning models are evaluated using the testing data, it can be useful to include basic metrics of how the model performs on the training data, in order to evaluate the presence of **overfitting**, where the model performs well on the training data but poorly on the testing data, or **underfitting**, where the model performs poorly on both [8]. Metrics such as Accuracy and Loss for the training data can be summarized as the following:

- **Accuracy:** The model achieved an accuracy of 0.99 for the training data.

- **Loss:** The model achieved a loss of 0.1170 for the training data.

For classification tasks, the results of the model can be summarized on a confusion matrix, that in the case of the present problem, would be a multi-class confusion matrix presented in Figure 1:

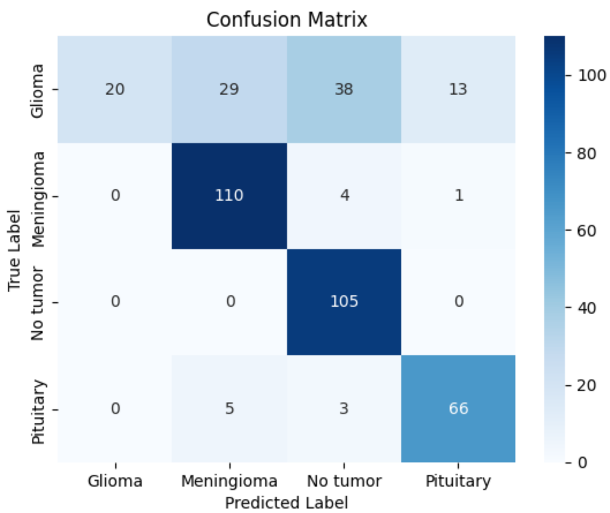

Fig. 1. Confusion Matrix of the model.

VI. Conclusion

After exploring the capabilities of AI and Machine Learning models, specifically Convolutional Neural Networks, for the classification of images of brain tumors and its applicability in the medical field, several conclusions can be drawn.

After training the model using the dataset that included the images of different types of brain tumors, to distinguish between four categories, the model achieved an accuracy, precision and recall of around 76%, with a high loss of 96.16. This implies that the model results are moderately reliable, and there is a significant room for improvement, even more in clinical use, where a higher reliability is needed to assist on medical scenarios where patients are concerned.

The confusion matrix summarizes the problem, where it shows that the model struggles when classifying the Glioma class, which has a lot of wrongful classifications compared to the rest of the classes, that the model differentiates between very accurately.

When observing the training data metric results, the accuracy achieves a result of 99% and a low loss of 0.1170, that when compared to the testing data metric results, it shows that while it learned the training data well, it's ability to generalize new, unseen data is moderate, a sign of overfitting on the model. Even though various techniques and tactics were implemented to reduce this phenomenon on the model results, like data augmentation, dropout and regularization, these techniques did not improve the model results and even worsened them, suggesting that standard techniques are not always effective. So, it can be drawn that overfitting is a problem that is still present, as it is a common problem of Machine Learning models, and can only be reduced, but not completely discarded.

This results show that there is potential for the use of Machine Learning models that use Convolutional Neural Networks for the aid of medical professionals and their tasks, to reduce errors and mistakes, improve diagnosis accuracy and aid treatment, as they can aid in image classification in a moderately correct way, but it cannot replace the labor of the medical staff, only serve as a second opinion that needs to be checked by a trained professional, as the model can commit serious mistakes that are not acceptable for such delicate medical matters and cannot be trusted to make a final judgement decision.

As a reference of future work, there is still a lot of room for improvement in the model's prediction ability and the reduction of limitations of the model, that can be aided by research in trying alternative techniques, more complex reduction techniques, different and more advanced architecture implementations, more extensive hyper parameter tuning, using advanced preprocessing techniques and utilizing cross-validation metrics for improving the model's precision and performance.

References

[1] Pooja Kamavisdar et. Al ”A Survey on Image Classification Approaches and Techniques”, International Journal of Advanced Research in Computer and Communication Engineering, Vol. 2, Issue 1, January 2013.

[2] S. Y. Chaganti, I. Nanda, K. R. Pandi, T. G. N. R. S. N. Prudhvith, y N. Kumar, ”Image Classification using SVM and CNN”, 2020 International Conference On Computer Science, Engineering And Applications (ICCSEA), mar. 2020, doi: 10.1109/iccsea49143.2020.9132851.

[3] S. A. Bini, “Artificial intelligence, machine learning, deep learning, and cognitive computing: what do these terms mean and how will they impact health care?,” The Journal of Arthroplasty, vol. 33, no. 8, pp. 2358–2361, Feb. 2018, doi: 10.1016/j.arth.2018.02.067.

[4] K. O'Shea and R. Nash, ”An Introduction to Convolutional Neural Networks”. Cornell University. https://arxiv.org/abs/1511.08458.

[5] K. L. Castellano Sánchez, ”Modelo Machine Learning para la detección de tumores cerebrales mediante un conjunto de imágenes de resonancias magnéticas,” Universidad Católica de Santiago de Guayaquil, Guayaquil, Ecuador, 2022. http://repositorio.ucsg.edu.ec/bitstream/3317/20047/1/T-UCSG-PRE-ING-CIS-305.pdf

[6] M. Álvarez Vega, L. M. Quirós Mora, y M. V. Cortés Badilla, ”Inteligencia artificial y aprendizaje automático en medicina,” Revista Médica Sinergia, vol. 5, no. 8, pp. 557-567, 2020. https://www.medigraphic.com/cgi-bin/new/resumen.cgi?IDARTICULO=95310.

[7] G. Varoquaux y O. Colliot, Evaluating Machine Learning Models and Their Diagnostic Value, Neuromethods, pp. 601-630, 2023, doi: 10.1007/978-1-0716-3195-9_20.

[8] S. Pothuganti, “Review on over-fitting and under-fitting problems in Machine Learning and solutions,” International Journal of Advanced Research in Electrical, Electronics and Instrumentation Engineering, vol. 7, no. 9, Sep. 2018, doi: 10.15662/IJAREEIE.2018.0709015.