Empirical Evaluation of Brain Tumor Classification using benchmark Deep Learning Models

Raza Imam
MBZUAI
Abu Dhabi, United Arab Emirates
raza.imam@mbzuai.ac.ae

Mohammed El-Amine Azz
MBZUAI
Abu Dhabi, United Arab Emirates
mohammed.azz@mbzuai.ac.ae

Mohammed Talha Alam
MBZUAI
Abu Dhabi, United Arab Emirates
mohammed.alam@mbzuai.ac.ae

1 Introduction

Impactful solutions are being offered by the context-aware deployment of deep learning methodologies to enhance medical diagnostics. The World Health Organization (WHO) states that a correct diagnosis of a brain tumor entails its discovery, localization, and classification based on its degree, kind, and severity. This research comprises finding the tumor, grading it according to type and location, and classifying it according to grade in the diagnosis of brain tumors using magnetic resonance imaging (MRI). This approach has experimented with using several models rather than a single model for classification task in order to categorize brain MRI data.

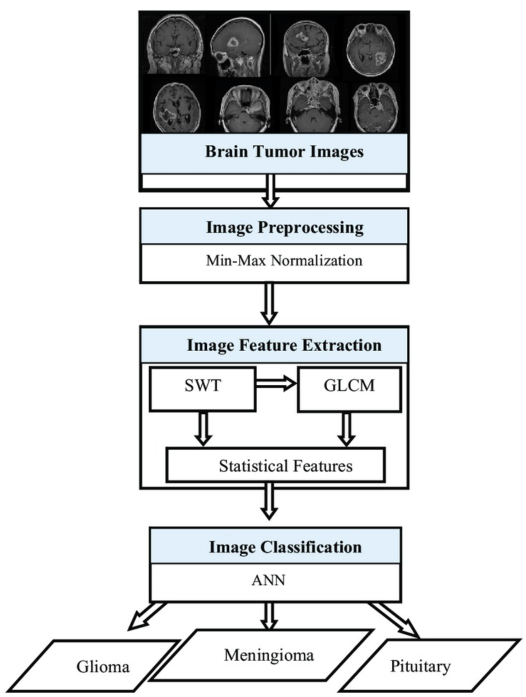

Figure 1: Architecture of the proposed study

Using three pathogenic forms of brain tumor (glioma, meningioma, and pituitary tumor), we aim to propose an accurate and automated classification scheme in this work. Towards this classification task, our goal is to empirically assess how well the most common benchmark models perform. Our goal of offering the finest classification model will open up new research directions in terms of choosing the right model for real-world brain tumor classification deployment. Utilizing tested models, the acquired characteristics are classified. The approach of our experiments that we will be acquiring is also represented in Figure 1 [1]. Subsequently, a thorough assessment of the suggested system is performed utilizing several effective evaluation metrics of classification tasks along with comparing several analytical factors, such as how well each model performs with less training samples from practicality aspect and how overfitting with lower training samples affects performance of the classifier.

1.1 Literature Review

The classification of brain tumors using MRI data has been the subject of numerous studies based on convolutional neural networks in recent years. Many of these methods make use of hybrid approaches, and many also offer technical variations on widely used deep learning models. In [2], Deepak and Ameer describe a classification method for the 3-class classification issue that combines transfer learning with GoogleNets. They used a number of evaluation metrics, with a mean classification accuracy of 98%, including area under the curve (AUC), precision, recall, F-score, and specificity. Kader et al. [3] have made use of the benefit of differential CNN by generating extra differential feature maps. With the capacity to categorize a sizable database of pictures with high accuracy of about 98%, their approach demonstrated a considerable improvement for the brain MRI classification problem. In [4], the authors proposed a hybrid architecture by adopting GoogleNet as a based CNN model while tweaking the last few layers for the specific Brain Tumor Classification. Their proposal attained the classification accuracy of 99.67%. Moreover, [5] conducted a multi-class study of Brain Tumor MRI Images as they propose a CNN model for early diagnoses purposes with fully optimized framework. Compared to the conventional CNN models, their solution attained an accuracy of 98.14%.

2 Motivation

Early brain tumor identification and classification represent a significant area of research in the field of medical imaging. It helps in choosing the most appropriate line of action for treatments to save patients' lives. In both children and adults, brain tumors are regarded as one of the most severe disorders. Brain tumors account for 85% to 90% of all major malignancies of the Central Nervous System (CNS). An estimated 11,700 people receive a brain tumor diagnosis each year. For those with a malignant brain or CNS tumor, the 5-year survival rate is around 34% for males and 36% for women [6]. It is difficult to treat brain tumors as we know that our brain has a very complex structure having tissues that are linked to each other in a complex manner. Often, producing MRI results is very difficult and time-consuming in underdeveloped nations due to a shortage of skilled medical professionals and a lack of understanding of malignancies.

Depending on the tumor's severity—that is, its location, size, and type—different treatment methods are possible. The most common technique for treating brain tumors at the moment is surgery since it has no adverse implications on the brain. There are different medical imaging techniques that are used to view the internal structures of a human body in order to discover any abnormalities. The most often used of them to identify brain tumors is Magnetic Resonance Imaging (MRI) since it can show abnormalities that may not be seen or just dimly visible on computed tomography (CT) [7]. However, the rush of patients makes it difficult, chaotic, and perhaps error-prone to manually review these images. Automated classification methods based on machine learning and artificial intelligence have regularly outperformed manual classification in terms of accuracy in order to solve this issue. Therefore, recommending a system that does detection and classification utilizing deep learning algorithms employing the above-mentioned benchmark models would be helpful for radiologists and other medical professionals.

3 Evaluation

For the typical assessment of a classifier, numerous performance metrics are specified. The most frequently cited quality metric is classification accuracy. The number of samples that were successfully classified in a classification is measured as a percentage of all the samples of data, and this percentage is known as accuracy. For the projected models that we will use, which are as follows, we seek to get individual classification accuracy results from our tests.

1. CNN
2. ANN
3. Transfer Learning
4. ResNets
5. MobileNets
6. SSD
7. Vision Transformer
8. GoogleNet

Along with the aforementioned models, we also intend to use the most recent benchmark studies and other hybrid techniques as a comparison, such as Convolutional XGBOOST, or transfer learning to our classification task using a pre-trained Deep Network and GoogleNet features.

When the test dataset has an identical amount of samples from each category, classification accuracy is a useful metric to assess performance. Nevertheless, the dataset we aim to use for this categorization problem under discussion is uneven. This calls for a more thorough assessment of the suggested system using more evaluation metrics. To evaluate the effectiveness of our tumor categorization method, we employed confusion matrices. A confusion matrix presents a graphical summary of correct and incorrect categorization. The effectiveness of the relevant implemented models that we will employ in our categorization will be empirically evaluated by the confusion matrix. In our dataset, the different categorical classes of brain tumor types include meningioma, glioma, and pituitary tumor.

$$Precision = \frac{TruePositive}{TruePositive+FalsePositive}$$
$$Recall = \frac{TruePositive}{TruePositive+FalseNegative}$$
$$Specificity = \frac{TrueNegative}{TrueNegative+FalsePositive}$$

The performance of the classifier for each type of tumor can be measured using several metrics obtained from a confusion matrix. The three most important indicators are precision, recall (or sensitivity), and specificity. Precision, recall (or sensitivity), and specificity are crucial indicators, and they are determined using the relations described above. If we obtain high specificity values across all classes, this suggests that we have successfully identified samples free of any class of brain tumor. For each class, the harmonic mean of recall and precision yields the F-score, another significant statistical classification metric. We also seek to use the average F-Score due to the existence of imbalance among the 3 classes. We also intend to compare several analytical factors, such as how well each model performs with less training samples from practicality aspect and how overfitting with lower training samples affects performance of the classifier.

4 Resources

We will need resources, including programming tools, computing resources, datasets, and reference materials, to perform the comparative study as stated above. The classifiers mentioned above including CNN, ANN, Transfer Learning, ResNets, MobileNets, SSD, Vision Transformer, and GoogleNet, were created with the use of programming tools. They will mostly be implemented in the Python programming language, utilizing frameworks and libraries including are Pytorch, Pandas, Numpy and Scikit-Learn. Computing power will be needed to enable the training of robust deep learning-based models. GPUs that are accessible through services like Google Colaboratory and Kaggle can assist us in getting strong training outcomes under time-intensive conditions. Other activities, such data preparation, analysis, and inference, will be carried out using the Jupyter and Anaconda environments on local workstations.

Our dataset is a combination of three datasets taken from Kaggle namely figshare, SARTAJ and Br35H. This dataset contains 7022 images of human brain MRI images which are classified into 4 classes: glioma, meningioma, no tumor and pituitary. No tumor class images were taken from the Br35H dataset. In addition, 1311 pictures of brain tumors are included for the model's validation.

5 References

1. Sathi, Khaleda & Islam, Md. (2021). Multi-classification of Brain Tumor Images Based on Hybrid Feature Extraction Method. 10.1007/978-3-030-68154-8_83.
2. Deepak, S., & Ameer, P. M. (2019). Brain tumor classification using deep CNN features via transfer learning. Computers in biology and medicine, 111, 103345.
3. Abd El Kader, I., Xu, G., Shuai, Z., Saminu, S., Javaid, I., & Salim Ahmad, I. (2021). Differential deep convolutional neural network model for brain tumor classification. Brain Sciences, 11(3), 352.
4. Raza, A., Ayub, H., Khan, J. A., Ahmad, I., S. Salama, A., Daradkeh, Y. I., ... & Hamam, H. (2022). A hybrid deep learning-based approach for brain tumor classification. Electronics, 11(7), 1146.
5. Irmak, E. (2021). Multi-classification of brain tumor MRI images using deep convolutional neural network with fully optimized framework. Iranian Journal of Science and Technology, Transactions of Electrical Engineering, 45(3), 1015-1036.
6. Kaggle. Brain Tumor Classification (MRI). https://www.kaggle.com/datasets/sartajbhuvaji/brain-tumor-classification-mri [Accessed on: 18 October 2022]
7. Maravilla KR, Sory WC. Magnetic resonance imaging of brain tumors. Semin Neurol. 1986 Mar;6(1):33-42. doi: 10.1055/s-2008-1041445. PMID: 3332409.