

Diagnosing Patients Through Magnetic Resonance Imaging

Team Name: TumorTrace

Team Members: Jimmy Vu, William Akins, Chen Zhang

February 10th, 2025

1 Problem Introduction

Magnetic Resonance Imaging (MRI) is a critical diagnostic tool in modern medicine. It provides high-resolution images of internal body structures [3]. However, analyzing MRI scans for accurate diagnosis remains to be a challenging task that requires extensive expertise and time. In particular, detecting and classifying abnormalities such as tumors, infections, or degenerative diseases from MRI data can be complicated due to the variations in image quality, patient anatomy, and disease presentation [10]. Misdiagnoses or delays in interpretation can lead to severe consequences, such as ineffective treatments and worsened patient outcomes. This project aims to develop a deep learning-based approach to assist in diagnosing medical conditions from MRI scans. The goal is to improve the efficiency and accuracy of medical image analysis, potentially reducing diagnostic errors and assisting radiologists in clinical decision-making. Our focus will be on classifying MRI scans to detect brain tumors. By utilizing publicly available MRI datasets, we seek to develop a model that generalizes well across different patient cases.

2 Motivation

Accurate and timely diagnosis is essential in medical imaging, particularly in MRI analysis, where early detection of abnormalities can significantly impact treatment success. However, radiologists frequently face high workloads, analyzing hundreds of scans daily, which can lead to fatigue and increased diagnostic variability. In addition, MRI scans contain extensive details that require careful examination, often across multiple 2D slices, to reconstruct a comprehensive 3D understanding of a patient's anatomy [3]. This complex and time-intensive process can result in inconsistencies in diagnoses due to differences in radiologists' expertise, experience, and mental fatigue levels. Another challenge in MRI interpretation is the inherent variability in image quality [11], including the differences in resolution, noise levels, and contrast due to varying machine specifications and scanning protocols, which make manual analysis even more challenging and require specialized expertise to ensure diagnostic accuracy.

Deep learning models provide an opportunity to address these issues by automating and standardizing image analysis. Studies have shown that deep learning approaches can significantly improve diagnostic efficiency by rapidly processing large volumes of MRI scans with high accuracy, reducing interpretation time while maintaining reliability [12]. By leveraging deep learning, we aim to streamline the MRI diagnostic workflow and enhance both efficiency and diagnostic confidence, highlighting potential areas of concern, reducing the chances of oversight, and offering consistent assessments.

3 Methodology

We plan to develop a model which intakes a dataset of MRI scan images, and following many preprocessing techniques produces a classifier for images if they show signs of being infectious or cancerous. Many other scientists have sought to develop similar models [4], but we will attempt to differentiate ourselves via employing new combinations of techniques and validation metrics.

There are many model variants that can be employed for this challenge [5], but we have chosen to focus our development on three primary ones to determine which combination works most effectively. The first model is expected to be the standard metric: 3D Convolutional Neural Networks (CNN). It is an obvious choice, given that the dataset makes use of higher dimensional values that standard CNNs are not natively equipped to train [5]. This technique does necessitate additional preprocessing to allow for 3D scaling, but should provide a good baseline. The second model to be deployed will be the Inception model, which is primarily known for its performance efficiencies. It is expected that this model will utilize its resources more effectively than the other intensive models, and given the potential scale and dimensionality of the dataset would be very much appreciated when running complex computations. The third model would be a ResNet. Given the complexity of the data correlating to large-layered neural networks, a ResNet accounts for possible vanishing gradient problems via skip connections [9].

For analysis, we will employ two main metrics of quantitative model performance: F1-score and accuracy. Accuracy is an intuitive metric of model performance, but may be inadequate unless our dataset has an even distribution of class labels which is not necessarily true. F1-score is less intuitive of a result, but it would not be influenced by dataset issues like uneven distribution of classes. For visualizing our results, a confusion matrix allows for a better understanding of our model's true positive (TP), false positive (FP), true negative (TN), and false negative (FN) results.

To further validate model robustness, we will conduct ablation studies to compare the impact of different preprocessing techniques, architecture variations, and hyperparameter tuning. Cross-validation will ensure that our results are not biased by specific data splits. We will also analyze misclassified cases to understand failure modes, such as whether certain tumor types are more prone to misclassification. In addition, we will compare our models against existing benchmarks in the literature review to determine how well our approach improves upon prior methods.

4 Potential Issues

An issue lies in the current datasets available, specifically how they are assembled in terms of what MRIs record. MRIs operate via capturing body structure in 2D segments, with each slice providing a 2D "slice" of that body part [3]. A complete MRI scan is, therefore, composed of a collection of these images at their varying points of recording, which ideally provides a complete 3D perspective on the internal structure of the body [3]. However, the majority of datasets of MRI scans are solely composed of these 2D images of different recordings of the same scan, losing an entire dimensionality of information. This may result in some scans having a false assertion of disease due to its visual appearance not being fully represented at that level. Data linking these images will have to be performed to properly have a complete data entry that contains all the information. This process involves collecting the batch of images that correspond to each

other, but also ensuring any subsequent image preprocessing techniques won't deform the continuity between these sliced images.

Furthermore, traditional image processing networks are not natively designed to account for an additional dimensionality, so more specific variations of these algorithms would have to be implemented. This issue is addressed in the chosen models we have decided to proceed with stated in the prior sections.

5 Related Work

Previous works have looked into utilizing deep learning as a means to identify disease in patients via MRI imaging, and thus we have taken inspiration from some to see how we could further elevate their processes and methodologies. Other works will often go through with more mining and search for MRI scans and data collection [6], and while this does have the benefit of widening the available set, it also results in drawbacks such as additional and nonstandardized preprocessing techniques having to be performed. Instead we choose to go with a synergized dataset from an online dataset repository, which standardizes the preprocessing step.

6 Datasets

For this project, we use a brain MRI image dataset [7]. This dataset contains 2D slices of an individual's MRI, each with a label for three types of brain tumors: pituitary, Meningioma and Glioma.

We preprocess this data by turning groups of 2D slices into a 3D datapoint showing an individual's complete MRI. We split the data using 5-fold cross-validation whereby we partition the data into 80% training and 20% testing, iteratively training the model 5 total times, each with a unique training/testing partitioning. We benchmark our model using the work of Abdusalomov et. al [1] which achieved an accuracy of 99.5% for binary classification of brain tumor or no brain tumor. By combining our multi-class classification results for the three types of brain tumors we detect, we can compare our results to this benchmark. To measure our model's performance for specifically multi-class classification, we use accuracy and F1-score, though there is no related work that could provide a benchmark for the specific case of multi-class classification.

7 Ethical Implications

Ethics are of primary concern when developing a medical model. In the domain of medicine, hallucinations from models are flatly unacceptable. Furthermore, automation bias—the tendency to trust and rely on what a machine tells you [8]—could cause doctors to make incorrect diagnoses, causing patients mental distress and inefficient use of medical resources. Though it is impossible to completely prevent hallucinations from our model, we can mitigate the harmful effects by displaying our results in probabilities rather than facts. This means rather than telling a doctor “this person has a tumor”, our model would say “there is a x% chance this person has a tumor.” This would counteract automation bias as doctors would feel encouraged to make the final decision (being yes or no there is a tumor).

References

- [1] Abdusalomov, A. B., Mukhiddinov, M., & Whangbo, T. K. (2023, August 18). Brain tumor detection based on deep learning approaches and Magnetic Resonance Imaging. *Cancers*.
<https://pmc.ncbi.nlm.nih.gov/articles/PMC10453020/>
- [2] Ahmad, I., & Alqurashi, F. (2024, October 15). *Early cancer detection using deep learning and medical imaging: A survey*. *Critical Reviews in Oncology/Hematology*.
<https://www.sciencedirect.com/science/article/pii/S1040842824002713#sec0020>
- [3] Butman, J. (n.d.). *Neuroradiologist offers insight on MRI scans*. NCI.
<https://www.cancer.gov/rare-brain-spine-tumor/blog/2022/neuroradiology>
- [4] Datta, P., & Rohilla, R. (2019). An introduction to deep learning applications in MRI images. *2019 2nd International Conference on Power Energy, Environment and Intelligent Control (PEEIC)*, 458–465. <https://doi.org/10.1109/peeic47157.2019.8976727>
- [5] Dorfner, F. J., Patel, J. B., Kalpathy-Cramer, J., Gerstner, E. R., & Bridge, C. P. (2025, January 3). *A review of deep learning for brain tumor analysis in MRI*. *Nature News*.
<https://www.nature.com/articles/s41698-024-00789-2>
- [6] Horasan, A., & Güneş, A. (2024, August 27). *Advancing Prostate Cancer Diagnosis: A deep learning approach for enhanced detection in MRI images*. MDPI.
<https://www.mdpi.com/2075-4418/14/17/1871>
- [7] Kumar, H. (2021, July 4). Brain Tumor Classification MRI images. Kaggle.
<https://www.kaggle.com/datasets/jarvisgroot/brain-tumor-classification-mri-images>
- [8] Kahn, L., & Probasco, E. (2024, November 20). *AI safety and automation bias*. Center for Security and Emerging Technology. <https://cset.georgetown.edu/publication/ai-safety-and-automation-bias/>
- [9] Lee D, Lee J, Ko J, Yoon J, Ryu K, Nam Y. Deep Learning in MR Image Processing. *Investig Magn Reson Imaging*. 2019 Jun;23(2):81-99. <https://doi.org/10.13104/imri.2019.23.2.81>
- [10] Pinto-Coelho, L. (2023, December 18). *How artificial intelligence is Shaping Medical Imaging Technology: A survey of innovations and applications*. Bioengineering (Basel, Switzerland).
<https://pmc.ncbi.nlm.nih.gov/articles/PMC10740686/>
- [11] Ma, Y., Lou, J., Tanguy, J.-Y., Corcoran, P., & Liu, H. (2024). Rad-IQMRI: A benchmark for MRI image quality assessment. *Neurocomputing*, 602, 128292.
<https://doi.org/10.1016/j.neucom.2024.128292>
- [12] Foti, G., & Longo, C. (2024, September 13). *Deep learning and AI in reducing magnetic resonance imaging scanning time: Advantages and pitfalls in clinical practice*. *Polish journal of radiology*.
<https://pmc.ncbi.nlm.nih.gov/articles/PMC11497590/>