4641_Project

CS4641: Brain Tumor Classification - Final Report

Introduction

Brain tumors are life-threatening conditions that require early and accurate diagnosis to improve patient outcomes. Medical imaging techniques and radiological expertise are essential in detecting brain tumors [3]. However, this is currently a very time consuming process [1]. The aim is to develop an automated classification system capable of automatically distinguishing between the two categories. This project uses a dataset of MRI images from Kaggle. The dataset contains patient brain images, with each image categorized as either containing a tumor or being tumor-free. There are 1311 brain images, 906 images having a tumor present (including Glioma, Meningioma, and Pituitary tumors) and 405 images having a tumor absent.

Current research has demonstrated the high effectiveness of using machine learning approaches to classify and achieve high accuracies. One approach involved applying transfer learning on a brain tumor dataset similar to this project and reporting high accuracies [1]. Another technique involved optimizing a CNN to further improve classification performance, with one study reporting an accuracy of 97% [2].

Problem

Between 2013-2017 in the United States, approximately 17,200 people died each year from a malignant brain tumor [4]. By training binary classifcation models to identify the presence of brain tumors, we provide a tool to aid in making more informed prognoses.

Methods

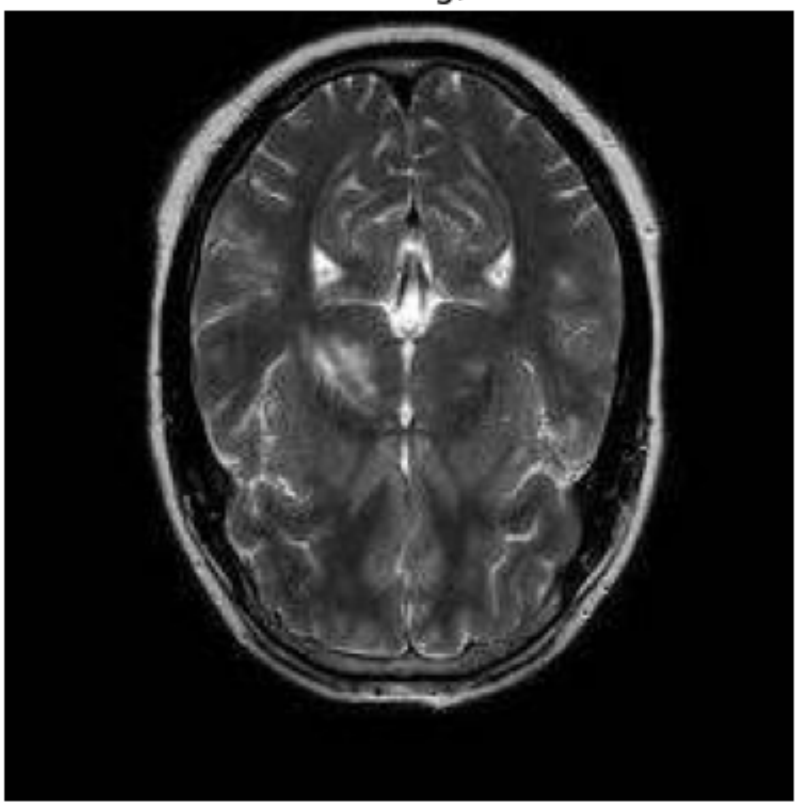
Data Preprocessing: Data Normalization

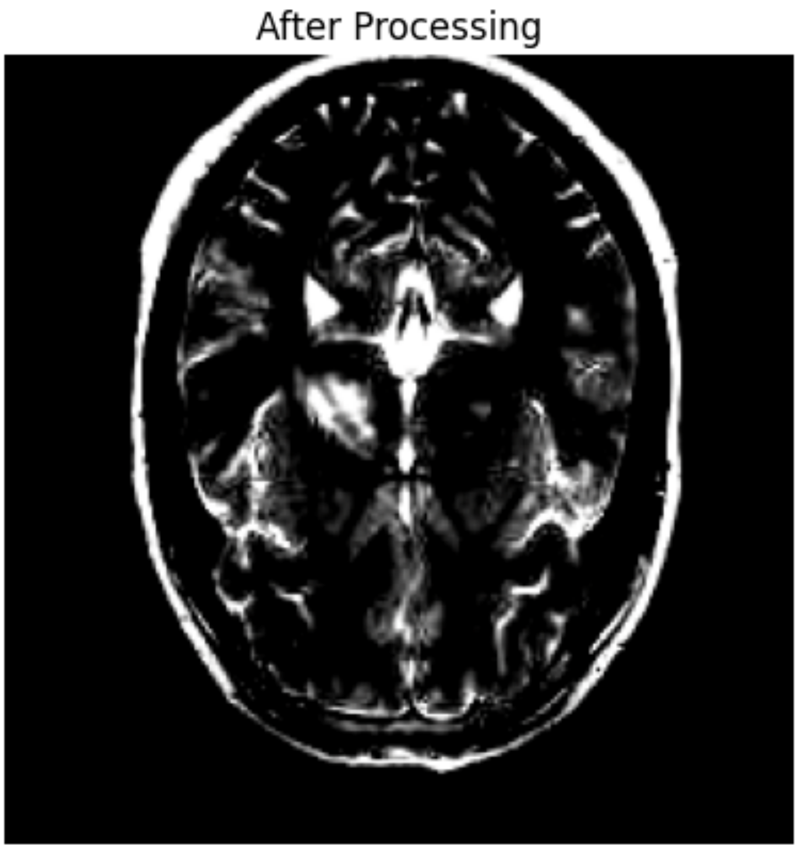
1. Data Normalization From our exploratory data analysis, we observed that not all images

had the same number of color channels. For images with only one color channel, i.e. grayscale images, we stacked the image three times to create three channels. As only one in a thousand images had more than three color channels, we dropped these images.The images were also normalized by the mean and standard deviation over the whole dataset. This standardizes contrast across images, and also speeds up convergence.

2. Reshaped Images We also reshaped all images to a width and height of 224 pixels to address different image sizes. This maintains the integrity of the image without cropping important information. The following figures show an example image from our dataset before and after preprocessing.

Before Processing, Tumor: No

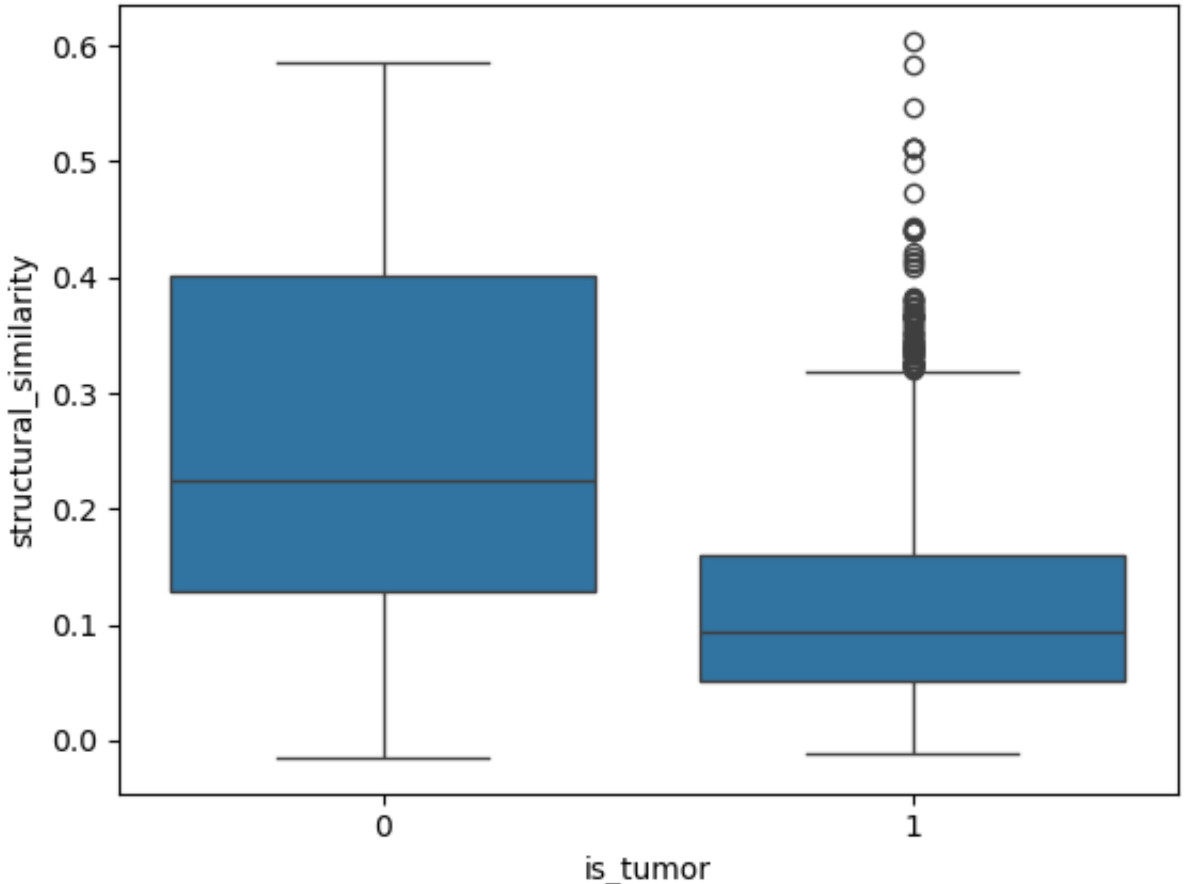




Data Preprocessing: Feature Engineering Symmetry

The key feature we engineered was a horizontal symmetry scoring using the structural similarity index measure. By a visual observation of the dataset, we found that healthy brains showed strong bilateral symmetry while tumors often created detectable asymmetries. From the figure below, we can clearly see a significant difference in structural similarity, where the healthy brains had a structural similarity mean of 0.23, while the brains with tumors exhibited a mean of 0.1.

However, there are several outliers in the tumor group (shown as individual circles) with unusually high symmetry scores up to about 0.6. These outliers might represent cases where tumors are positioned in ways that don’t significantly disrupt the brain’s bilateral symmetry. They could indicate limitations such as a sensitivity to image alignment as images might be skewed.



Machine Learning Methodologies

Convolutional Neural Networks

We have chosen CNNs for their effectiveness in image classification due to their ability to automatically learn feature representations, like spatial hierarchies, from raw image data through layers of convolutions and pooling.

Taking inspiration from AlexNet and VGG16, our CNN model utilizes a stripped-down version of these models with four main convolutional blocks followed by a sequence of linear layers. Within each block, each 2d convolutional layer maintains the spatial dimensions of their input while increasing the number of channels by utilizing padding in proportion to the kernel size. We note that this technique is commonly utilized in modern CNN models as border and edge information is better preserved. In the context of brain tumor classification, edge details of a tumor are potentially exhibited by contrast differences which could be critical for detection.

Next, a ReLU layer was applied to introduce non-linearity, followed by a max pooling layer to reduce the dimensionality to a quarter. This layer serves to reduce spatial dimensions while also selecting features by keeping the strongest activations in each window. In this way, each layer sees larger proportions of the original image, with a growing receptive field.

Moreover, batch normalization was also applied to normalize layer inputs during training which ensures the input distribution remains more stable. Moreover, as each sample is normalized using slightly different statistics, this added minor noise serves as an added regularizing effect. To that end, we also applied dropout layers between linear layers to reduce the effect of overfitting.

```
==========================================================================================
Layer (type:depth-idx)                   Output Shape              Param #
==========================================================================================
SanityCheckBasicCNN                      [1]                       --
├─Sequential: 1-1                        [1, 64, 112, 112]         --
│    └─Conv2d: 2-1                       [1, 64, 224, 224]         4,800
│    └─BatchNorm2d: 2-2                  [1, 64, 224, 224]         128
│    └─ReLU: 2-3                         [1, 64, 224, 224]         --
│    └─MaxPool2d: 2-4                    [1, 64, 112, 112]         --
├─Sequential: 1-2                        [1, 128, 56, 56]          --
│    └─Conv2d: 2-5                       [1, 128, 112, 112]        204,800
│    └─BatchNorm2d: 2-6                  [1, 128, 112, 112]        256
│    └─ReLU: 2-7                         [1, 128, 112, 112]        --
│    └─MaxPool2d: 2-8                    [1, 128, 56, 56]          --
├─Sequential: 1-3                        [1, 256, 28, 28]          --
│    └─Conv2d: 2-9                       [1, 256, 56, 56]          294,912
│    └─BatchNorm2d: 2-10                 [1, 256, 56, 56]          512
│    └─ReLU: 2-11                        [1, 256, 56, 56]          --
│    └─MaxPool2d: 2-12                   [1, 256, 28, 28]          --
├─Sequential: 1-4                        [1, 512, 14, 14]          --
│    └─Conv2d: 2-13                      [1, 512, 28, 28]          1,179,648
│    └─BatchNorm2d: 2-14                 [1, 512, 28, 28]          1,024
│    └─ReLU: 2-15                        [1, 512, 28, 28]          --
│    └─MaxPool2d: 2-16                   [1, 512, 14, 14]          --
├─Sequential: 1-5                        [1, 1]                    --
│    └─Flatten: 2-17                     [1, 100352]               --
│    └─Linear: 2-18                      [1, 2048]                 205,522,944
│    └─ReLU: 2-19                        [1, 2048]                 --
│    └─Dropout: 2-20                     [1, 2048]                 --
│    └─Linear: 2-21                      [1, 512]                  1,049,088
│    └─ReLU: 2-22                        [1, 512]                  --
│    └─Dropout: 2-23                     [1, 512]                  --
│    └─Linear: 2-24                      [1, 64]                   32,832
│    └─ReLU: 2-25                        [1, 64]                   --
│    └─Linear: 2-26                      [1, 1]                    65
==========================================================================================
```

Naive Bayes Classifiers

The second algorithm we chose to implement is a naive bayes classifier. The symmetry score serves as a key feature for the classifier by providing a continuous numerical feature that captures global structural symmetry. This enables a probabilistic estimate of tumor presence based on symmetry deviation, offering an interpretable metric that aligns with medical domain knowledge.

The classifier assumes independence between features and calculates the probability of each class given the feature values using Bayes' theorem: P(tumor	symmetry). While the naive assumption of feature independence might seem limiting, this simplicity allows for clear probabilistic interpretations of the relationship between brain symmetry and the presence of tumors. This potentially aligns with how medical professionals reason about diagnostic probability.

Random Forests

Our third approach is a Random Forest classifier, which builds upon symmetry analysis in a different manner. This ensemble method constructs multiple decision trees, each learning optimal threshold values for the symmetry scores to learn decision rules based, thus differentiating between healthy and tumor-present brains.

The strength of the Random Forest classifier lies in its ability to capture complex, non-linear relationships in the symmetry data while maintaining robustness through its ensemble nature.

Both approaches complement our primary CNN model by offering more interpretable results and efficient computational requirements. The Naive Bayes classifier provides clear probability estimates that medical professionals can easily interpret, while the Random Forest offers insights into critical symmetry thresholds that distinguish healthy from pathological cases. Together, these classical machine learning approaches provide valuable perspectives on tumor detection while maintaining computational efficiency and interpretability.

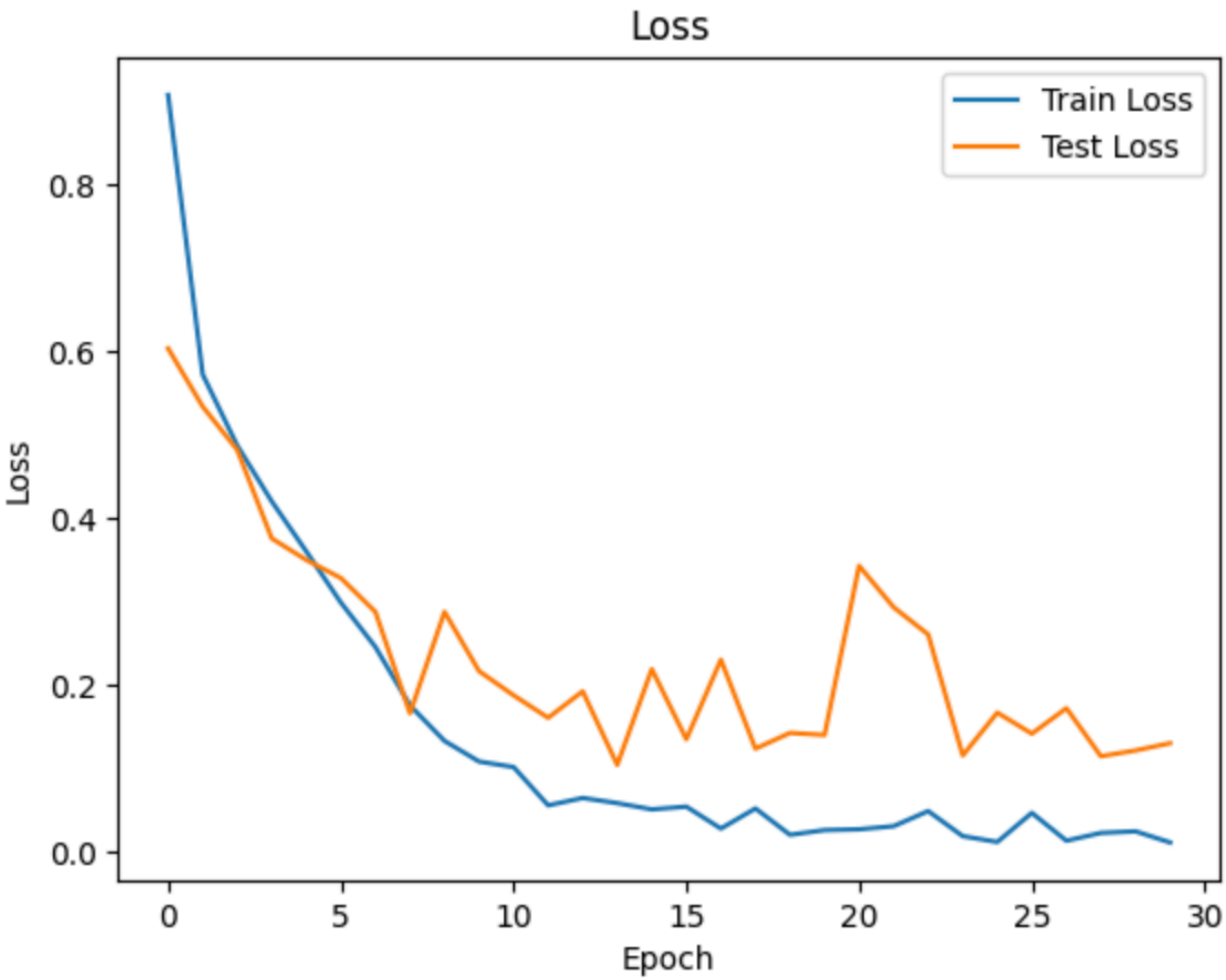
Results

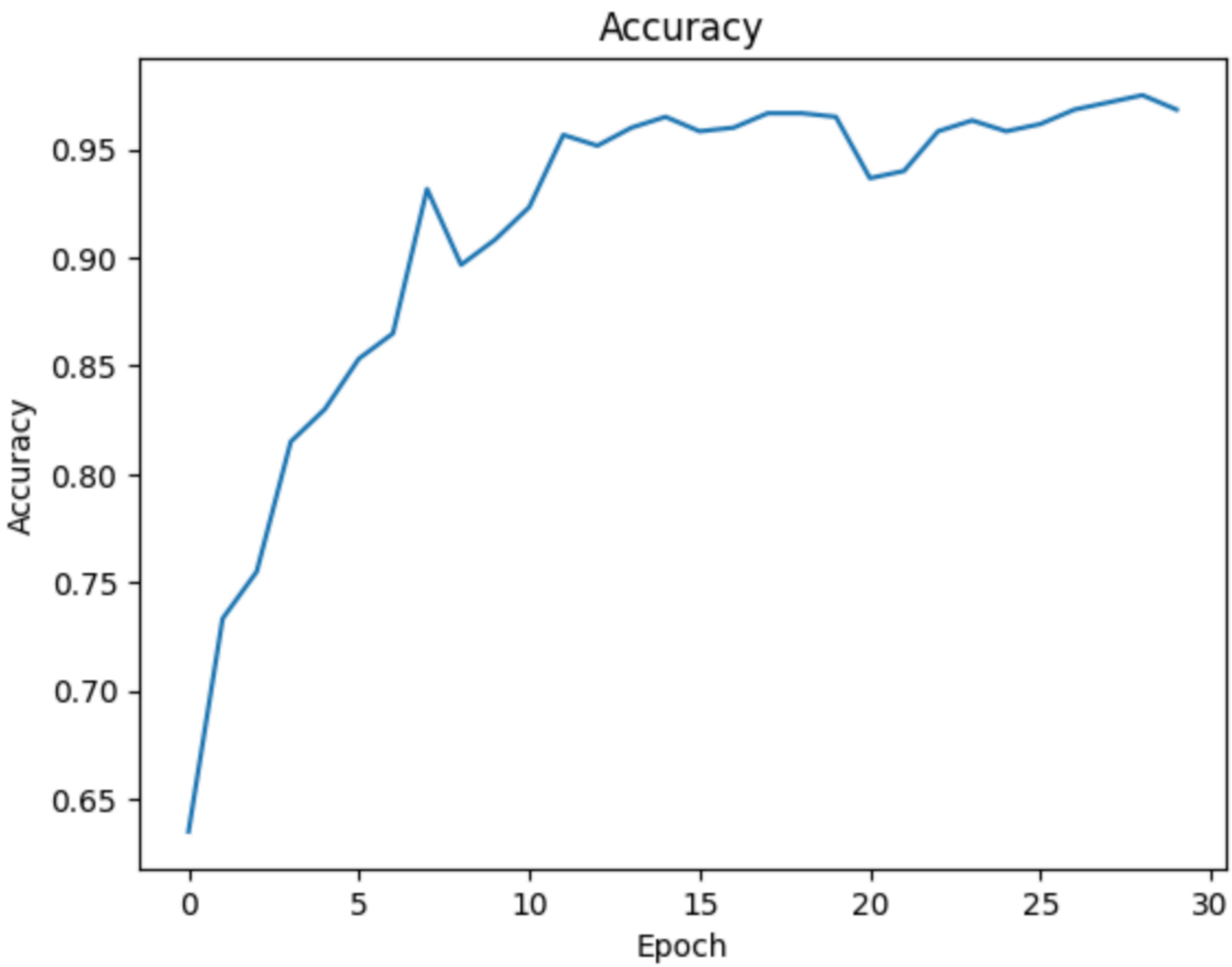
The data normalization, resizing, and channel adjustments likely improved feature consistency across images, allowing the model to focus on key visual differences between tumor and non-tumor images.

This standardization and preprocessing contributed significantly to achieving high accuracy quickly.

Convolutional Neural Networks

The CNN model achieved a high accuracy of 96.83% after 30 epochs. Both training and validation accuracies are closely aligned, indicating strong generalization and minimal overfitting. Precision is 97.11% , recall is 96.79%, and our F1 score is 96.95%. Precision, recall, and F1-scores for both classes are high. This suggests the model has balanced classification capabilities and it has the potential for reliable tumor classification.





The figures above show the training and testing loss, as well as accuracy over 30 epochs. Based on these figures, we can see that the training and validation performance metrics show the CNN model is learning effectively. It is achieving a high accuracy of over 95% in distinguishing tumors from non-tumor images. The training loss steadily decreases over the epochs and converges to a low value, indicating the model is becoming more confident in its predictions of the training data. The test loss initially decreases and then begins to show some fluctuations, suggesting some overfitting. These fluctuations are minor and the high training accuracy and stable performance show a high potential for generalization. Overall, the CNN shows great potential for reliable and accurate brain tumor classification.

Naive Bayes Classifiers

The bayes classifier exhibits the following performance metrics:

	precision	recall	f1-score	support
0	0.78	0.51	0.62	311
1	0.62	0.84	0.71	289
accuracy			0.67	600
macro avg	0.70	0.68	0.67	600
weighted avg	0.70	0.67	0.67	600

These results show that our model has an accuracy of around 70%, which is lower than our original goal of 90%. This can be attributed to the fact that the symmetry feature we engineered is inaccurate for some of the images where the brainscan is not centered in the image. When the image is not centered, the symmetry score is skewed, and this affects the classification. To ensure that our model was working correctly, we inserted noisy features into the bayes classifier, and received a much lower accuracy, which tells us that our model works correctly with the symmetry feature. This model has potential for higher performance metric values in the future with better engineering of the symmetry feature, such as performing transforms on the misaligned images, or using different features, such as clustering, to identify the tumors.

Random Forests

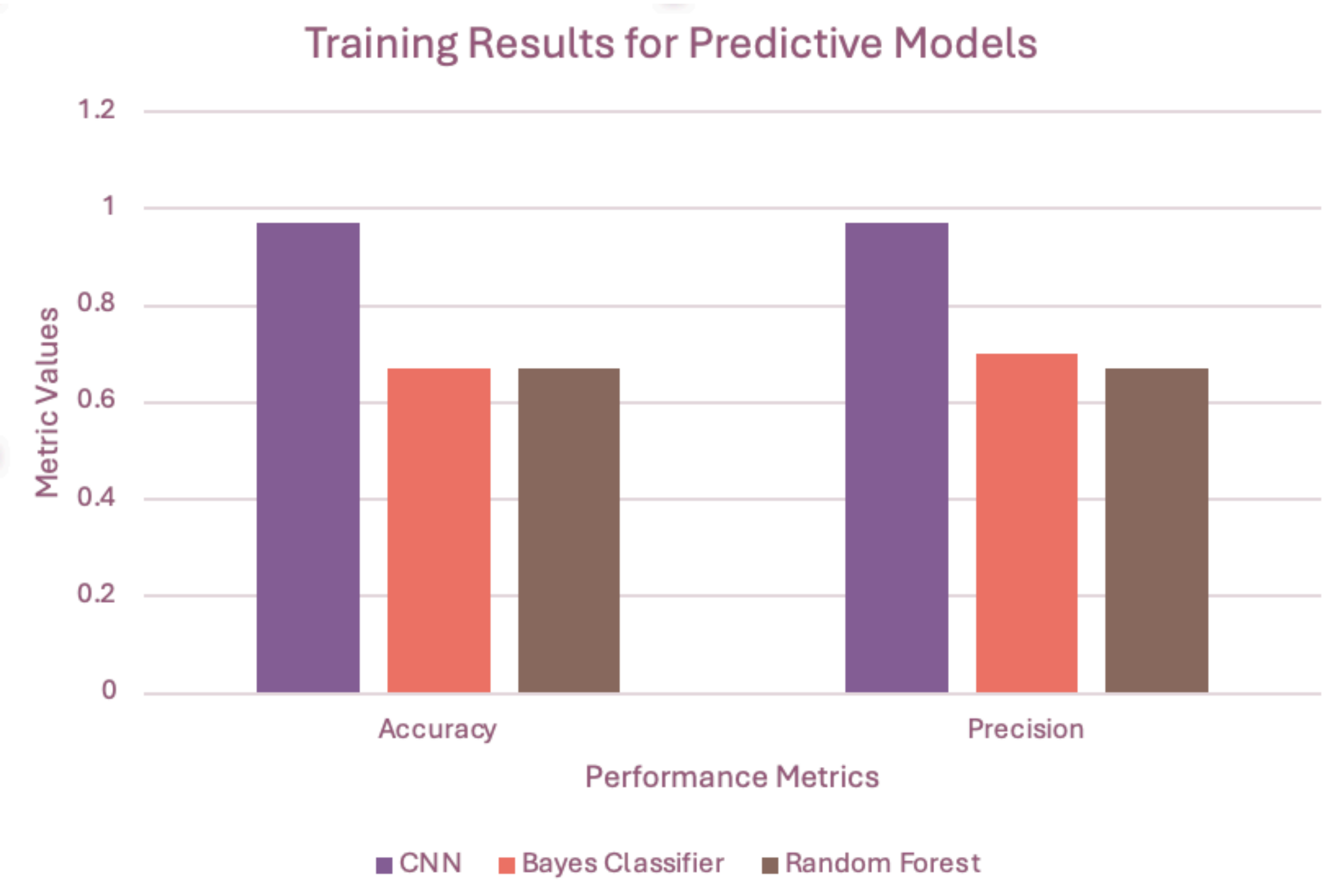
The random forest has the following performance metrics:

	precision	recall	f1-score	support
0	0.70	0.64	0.67	311
1	0.64	0.70	0.67	289
accuracy			0.67	600
macro avg	0.67	0.67	0.67	600
weighted avg	0.67	0.67	0.67	600

The random forest has an accuracy of around 67%, which is again lower than our original goal of 90%. Similar to the bayes classifier, this can be attributed to inaccuracies within the symmetry feature. This model also has potential for higher performance metric values in the future with better engineering of the symmetry feature, such as performing transforms on the misaligned images, or using different features, such as clustering, to identify the tumors.

Discussion

Our analysis of three distinct approaches - Convolutional Neural Networks, Naive Bayes, and Random Forest - for brain tumor classification reveals important insights about the strengths and limitations of each method. The CNN demonstrated superior performance with 96.83% accuracy, significantly outperforming both the Naive Bayes (70%) and Random Forest (67%) classifiers.



The CNN's exceptional performance can be attributed to several factors. First, its hierarchical feature learning capability allows it to automatically discover and learn complex patterns in the brain MRI images. The architecture's design, with four convolutional blocks followed by linear layers, enables the model to capture both local features (like tumor edges and texture) and global patterns such as overall brain structure. The high precision (97.11%) and recall (96.79%)

suggest the model is equally capable of identifying both tumor and non-tumor cases, making it reliable for clinical applications.

However, the Naive Bayes and Random Forest classifiers, while showing promise in their interpretability, faced significant limitations. Both models relied heavily on the engineered symmetry feature, which proved to be sensitive to image alignment issues. When brain scans were not perfectly centered, the symmetry scores became unreliable, leading to decreased performance.

Despite their lower accuracy and precision, these classical machine learning approaches offer valuable complementary benefits. The Naive Bayes classifier provides interpretable probability estimates that medical professionals can easily understand, while the Random Forest offers insights into critical symmetry thresholds. These interpretability advantages could be particularly valuable in clinical settings where understanding the model's decision-making process is crucial.

Implications

The results of this study have several important implications for medical image analysis and clinical practice.

The CNN's high accuracy suggests it could serve as a reliable preliminary screening tool, potentially reducing medical professionals' workload and accelerating diagnosis timelines. However, the model should be viewed as an assistive tool rather than a replacement for medical expertise.

The performance gap between deep learning and traditional machine learning approaches emphasizes the importance of robust feature engineering. While engineered features like symmetry can be intuitive and interpretable, they may not capture the full complexity of medical imaging data.

The choice between high-performing but complex models (CNN) versus more interpretable but less accurate models (Naive Bayes, Random Forest) represents a crucial consideration for healthcare applications. A hybrid approach might be optimal, using the CNN for initial screening while leveraging the interpretable models for explanation and verification.

Conclusion \& Next Steps

Our study demonstrates that deep learning approaches, specifically CNNs, can achieve high accuracy in brain tumor classification from MRI images. The CNN model's 96.83% accuracy represents a significant achievement that could potentially impact clinical practice. While traditional machine learning approaches showed lower accuracy, they provide valuable complementary benefits through their interpretability and efficiency.

Future work could focus on improving the robustness of feature engineering for traditional models, particularly addressing image alignment sensitivity. Furthermore, developing hybrid approaches would combine the high accuracy of CNNs with the interpretability of traditional methods, perhaps through transfer learning methods to leverage existing performant models.

Moreover, future studies could look at expanding the dataset to include more diverse cases and tumor types, particularly smaller tumors that are harder to detect by the human eye. Lastly, implementing attention mechanisms in CNNs would allow for a better understanding of which image regions influence the classifier's decisions.

This research contributes to the growing body of evidence supporting the potential of machine learning in medical image analysis while highlighting important considerations for practical implementation in clinical settings.

References

1. Vinod Kumar Dhakshnamurthy, M. Govindan, Kannan Sreerangan, Manikanda Devarajan Nagarajan, and A. Thomas, "Brain Tumor Detection and Classification Using Transfer Learning Models," Feb. 2024, doi: https://doi.org/10.3390/engproc2024062001.
2. Muhammad Aamir et al., "Brain Tumor Detection and Classification Using an Optimized Convolutional Neural Network," Diagnostics, vol. 14, no. 16, pp. 1714–1714, Aug. 2024, doi: https://doi.org/10.3390/diagnostics14161714.
3. I. Ilić and M. Ilić, "International patterns and trends in the brain cancer incidence and mortality: An observational study based on the global burden of disease," Heliyon, vol. 9, no. 7, pp. e18222–e18222, Jul. 2023, doi: https://doi.org/10.1016/j.heliyon.2023.e18222.
4. Q. T. Ostrom, N. Patil, G. Cioffi, K. Waite, C. Kruchko, and J. S. Barnholtz-Sloan, "CBTRUS Statistical Report: Primary Brain and Other Central Nervous System Tumors Diagnosed in the United States in 2013–2017," Neuro-Oncology, vol. 22, no. Supplement_1, pp. iv1–iv96, Oct. 2020, doi: https://doi.org/10.1093/neuonc/noaa200.

Gantt Chart

Gantt Chart

Contribution Table

Name	Contribution
Sruthi	Data preprocessing
Vidhya	Visualizations
Sukshma	Results and Discussion
Rei	Data preprocessing, CNN
Keane	Data preprocessing, CNN