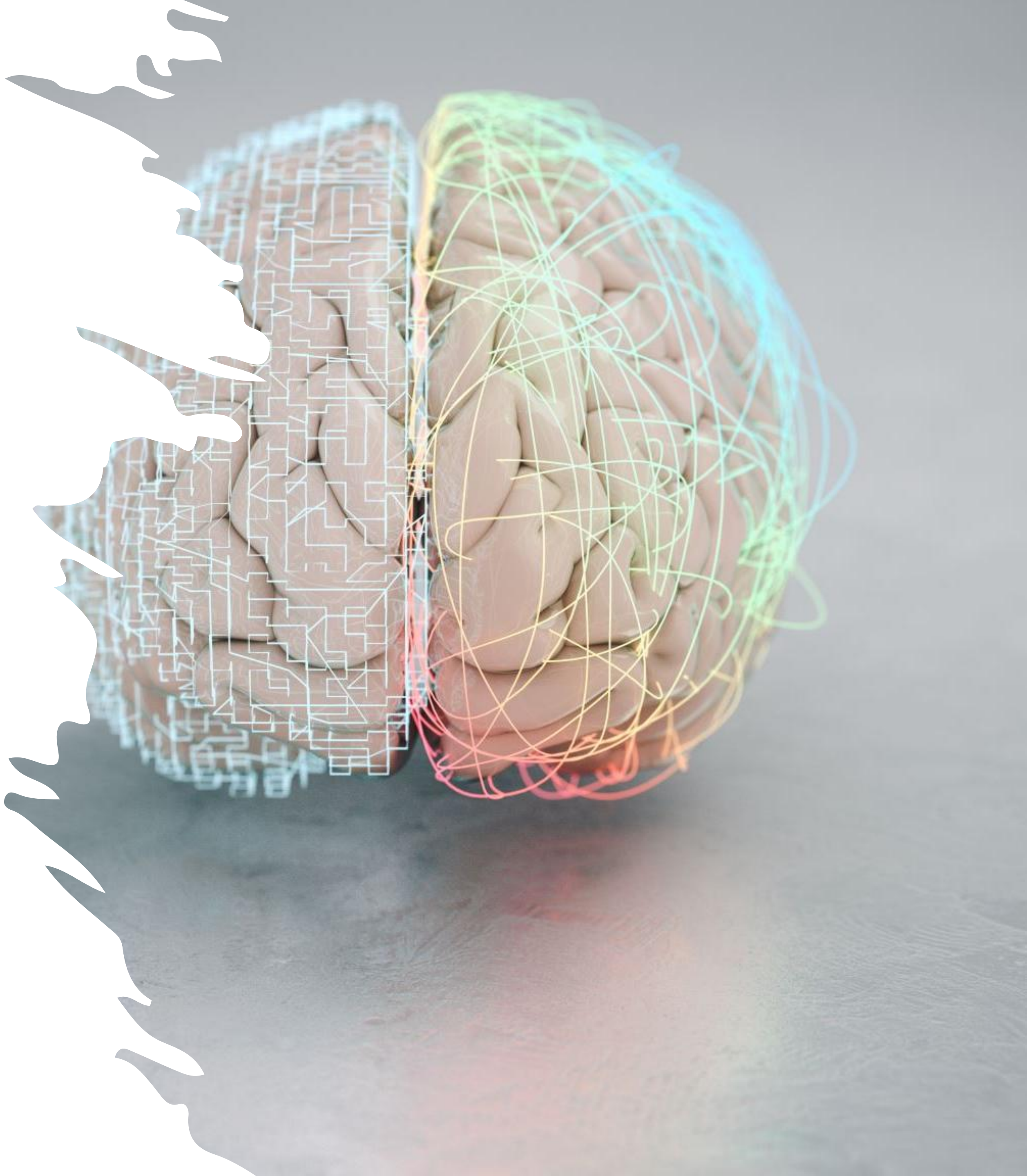
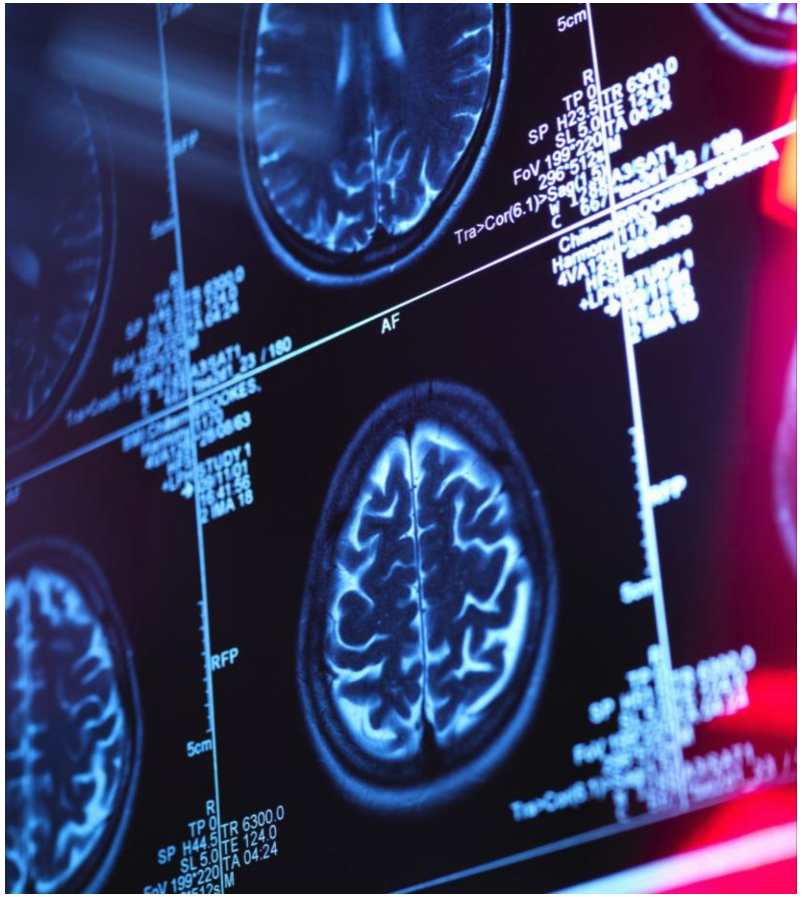
From Pixels to Prognosis: Deep Learning for Brain Tumor Detection



Travis Bates

Capstone 3 Presentation



Problem Statement

What opportunities exist for oncologists to automate and decrease tumor diagnosis time by 50% while maintaining an accuracy of 95% or higher with the use of machine learning and MRI images to detect the presence of a tumor?

Background

- Brain tumors are among the most serious neurological disorders, often requiring timely and precise diagnosis to improve patient outcomes.
- MRI imaging is one of the most reliable imaging techniques for diagnosing brain abnormalities, but manual interpretation is time-consuming and subject to human error.
- Recent advancements in deep learning, particularly convolutional neural networks (CNNs), have shown great promise in medical imaging tasks. By training a model on labeled MRI scans, we can automate the tumor detection process, reduce diagnostic workload, and potentially enhance diagnostic accuracy

Dataset Overview

The dataset contained 7,023 images of human brain MRI scans which are classified into 4 classes:

- Glioma
- Meningioma
- Pituitary
- No Tumor

Source: Kaggle- https://www.kaggle.com/datasets/masoudnickparvar/brain-tumor-mri-dataset

Data Wrangling And EDA

Meta Data	Raw Images	Cleaned Images
Image Sizes	There were 388 different images sizes in the data set	All images converted to (224, 224)
Image Modes	There were 4 different modes in the data set.	All images converted to RGB

Data had class imbalance ratio of 2.51

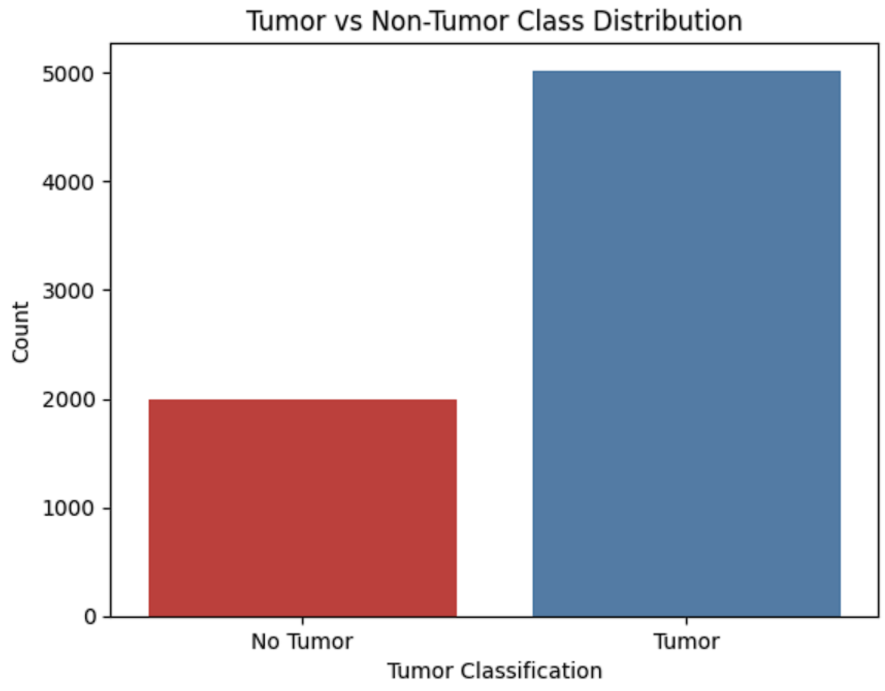


Figure 1: Class distribution between tumor and non-tumor images

Models Tested

Feature	MobileNetV2	EfficientNetB0	Custom CNN
Type	Pretrained (Transfer Learning)	Pretrained (Transfer Learning)	From Scratch
Input Size	224×224×3	224×224×3	224×224×3
Parameters	~2.2 million	~5.3 million	~12.9 million
Pretrained Weights	ImageNet	ImageNet	None
Model Size	Lightweight	Medium	Lightweight
Train Time	Fast (~minutes)	Moderate (~minutes)	Fast

Training Strategy

Transfer learning:

- Frozen base
- Fine-tuned top layers

Loss function:

Binary Crossentropy / Focal Loss

Optimizer:

Adam

Other techniques use:

Early Stopping, Reduce LR on plateau

Evaluation Metrics

Selected Metrics	Importance
Accuracy	Reflects how often the model correctly identifies both tumor and non-tumor cases. However, can be misleading with imbalanced data.
Precision	In healthcare scenarios missing a tumor (false negative) could delay diagnosis and treatment.
Recall (Sensitivity)	Helps reduce false alarms, avoiding unnecessary stress, further testing, or treatments for healthy individuals.
F1-Score	Ensures a balanced view, especially when classes are imbalanced (e.g., more "Tumor" than “No Tumor").
Confusion Matrix	Gives a transparent view into model behavior across both classes.

Results Comparison

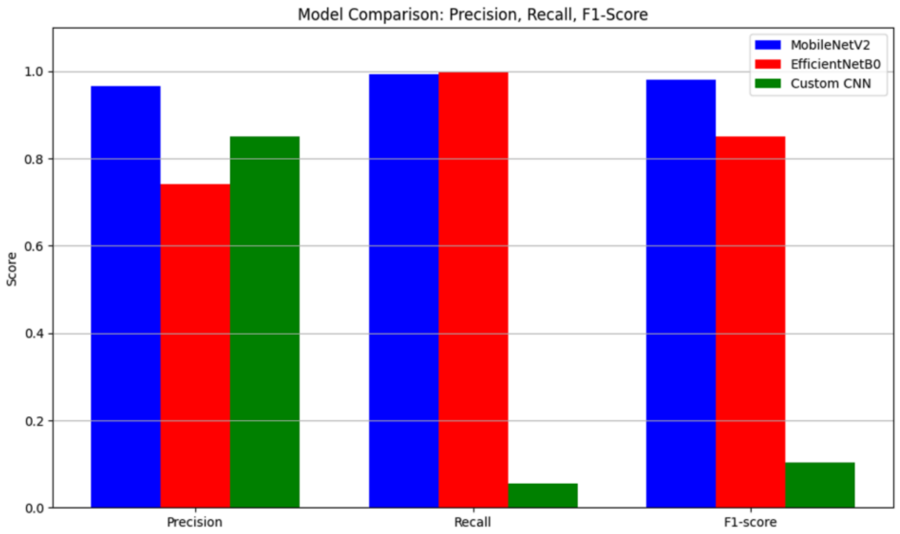


Figure 2: Model Comparison: Precision, Recall, F1-Score

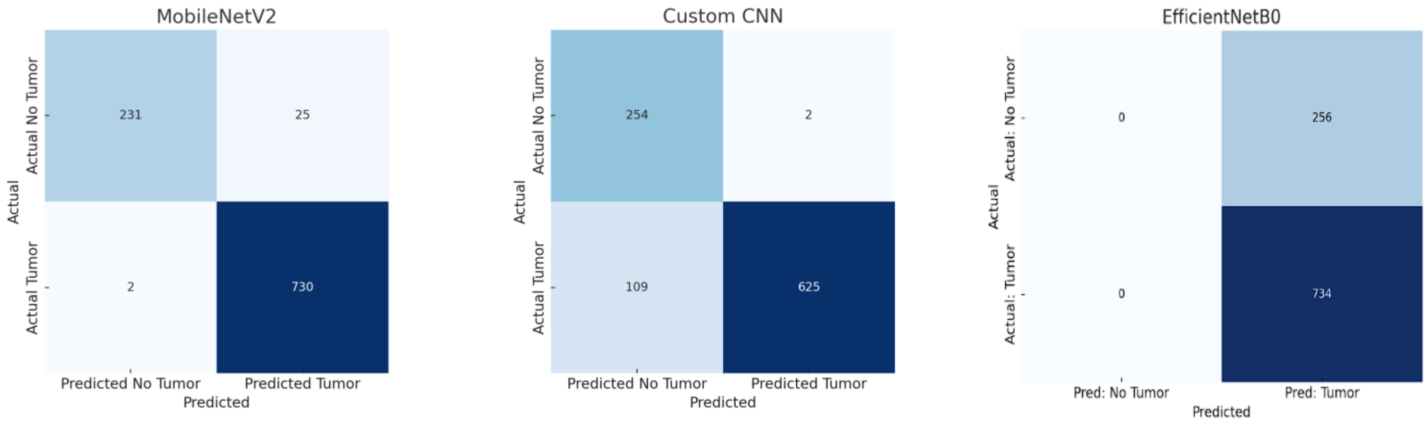


Figure 3: Confusion Matrix Comparison: MobileNet, Custom CNN, EfficientNet

Key Takeaways & Challenges

Key Takeaways

MobileNetV2 outperformed all models	• Achieved ~95% accuracy with very high recall, making it suitable for medical diagnosis where false negatives must be minimized.
Preprocessing was critical	• Using appropriate normalization (mobilenet_preprocess_input) and data augmentation improved generalization.
Balanced data improved model reliability	• Addressing class imbalance with augmentation and class weights helped avoid biased predictions.
Model interpretability & efficiency matters	• MobileNetV2 offered a good trade-off between performance and computational efficiency.

Challenges Encountered

Class Imbalance	• Led to poor generalization and biased predictions
EfficientNetB0 failed to generalize	• Despite fine-tuning, it consistently predicted one class due to likely preprocessing mismatches and sensitivity to hyperparameters.
Debugging	• Required careful inspection of generators, labels, and output shapes

Future Work

Future Work

- **Interpretability**: Implement explainability tools like Grad-CAM to visualize which regions of the MRI image the model uses to make its predictions. This can increase clinician trust and identify potential biases.
- **Multiclass Classification**: Expand the model from binary classification (tumor vs. no tumor) to classify tumor types (e.g., glioma, meningioma, pituitary tumor) to support more granular diagnostics.
- **Model Ensemble**: Combine predictions from multiple architectures (e.g., MobileNetV2, EfficientNet, and a custom CNN) through ensembling to potentially boost performance and robustness.
- **Real-Time Web or Mobile App Integration**: Develop a lightweight interface for doctors to upload MRI images and receive predictions, either via a web app or a mobile device with proper security protocols.
- **Clinical Trial or Validation Study**: real-world deployment in a medical context should include further validation on larger, diverse, and clinically annotated datasets. A human-in-the-loop review system is also recommended.
- **Include more metadata**: Adding patient metadata (age, sex, scanner type, etc.) as auxiliary features may improve model performance and interpretability.