Medical Imaging Data Science Project: Brain Tumor Detection

This report details the execution, analysis, and results of a deep learning project for brain tumor detection using MRI images, as conducted in the provided Jupyter Notebook.

1. Project Objectives :

The primary objective of this project was to gain hands-on experience in medical imaging data science by performing a binary classification task (Tumor vs. No Tumor) on MRI scans. The work emphasizes technical proficiency in deep learning and a clear discussion of clinical relevance.

Core Objectives:

- **Data Analysis & Preprocessing:** Understand and prepare the image dataset.
- **Baseline Modeling**: Establish a performance floor for deep learning.
- **Deep Learning & Transfer Learning**: Build and train both a custom CNN and an EfficientNetB0 model.
- **Evaluation & Comparison**: Assess models using clinically relevant metrics (Precision, Recall, F1-score).
- **Clinical Discussion**: Outline how the model could support radiologists.

2. Exploratory Data Analysis (EDA)

The EDA focused on inspecting the raw data, image characteristics, and the distribution of the target variable.

2.1. Dataset Size and Class Distribution

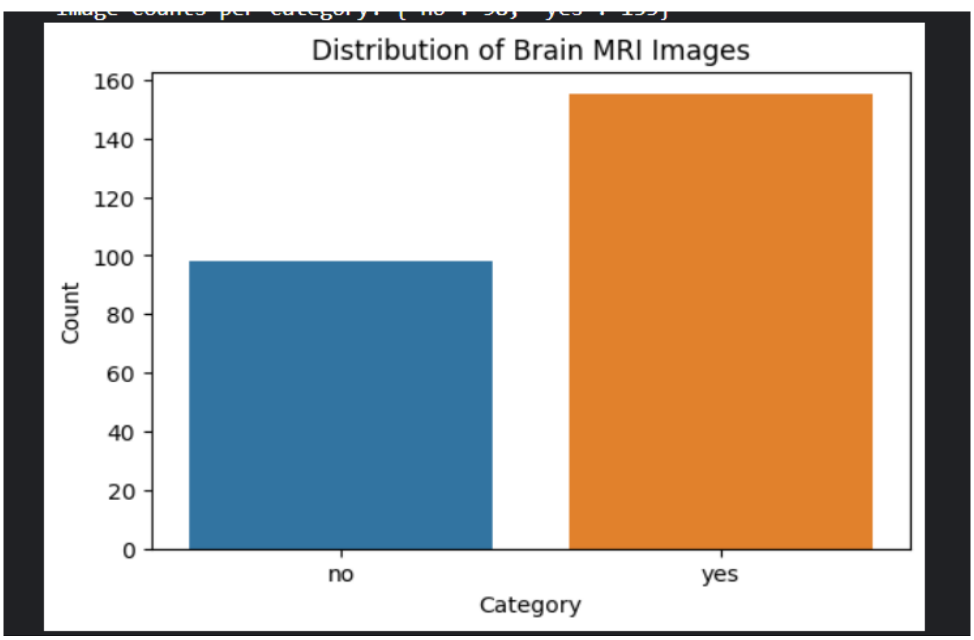
The dataset contains a total of **253** MRI scans, split into two categories. The analysis revealed a critical **class imbalance**.

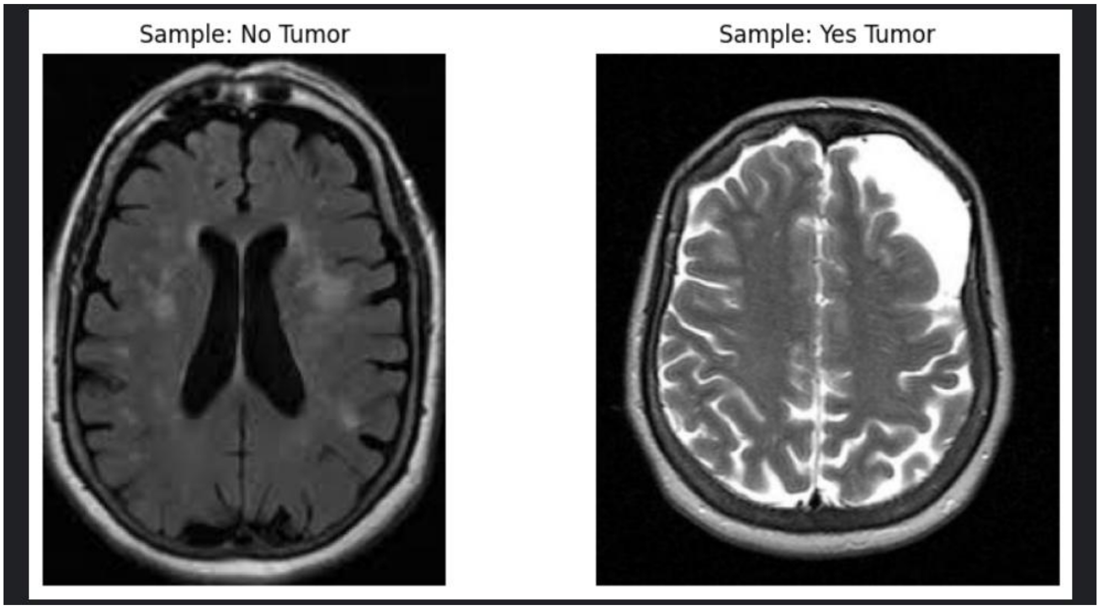
Category	Label	Count	Percentage
No Tumor	no	98	**38.7%**
Tumor Present	yes	155	**61.3%**
Total Images		**253**	**100%**

EDA Finding: The imbalance (61.3% majority class) necessitates mitigation techniques to ensure the model does not become biased toward predicting 'yes' for all inputs.

2.2. Sample Visualization and Image Characteristics

Sample images were visualized to confirm the presence and nature of the data. All images were fixed to a target size of **(224, 224)** for consistency across all models.





3. Data Preprocessing and Augmentation

The data was prepared for deep learning, ensuring consistency, proper scaling, and an expansion of the training data's diversity.

3.1. General Preprocessing

- **Image Resizing:** All images were set to **(224, 224)**.
- **Normalization:** Pixel values were scaled to the **[0, 1]** range by setting rescale=1./255.
- **Data Split:** The data was loaded into generators with an 80/20 split:
    - **Training Set:** 203 images
    - **Validation Set:** 50 images

3.2. Data Augmentation Strategy

The ImageDataGenerator applied robust augmentation to the **training set** to improve generalization:

- **Rotation:** rotation_range=15
- **Shifts:** width_shift_range=0.1, height_shift_range=0.1
- **Shearing/Zoom:** shear_range=0.1, zoom_range=0.1
- **Flipping:** horizontal_flip=True

3.3. Class Weight Calculation

To combat the identified class imbalance, class_weights were calculated and applied to the loss function during model training.

Class Index	Category	Weight
0	no (Minority)	**1.2848**
1	yes (Majority)	**0.8185**

Rationale: The higher weight on the minority class (1.28) forces the models to put more emphasis on correctly classifying 'No Tumor' cases.

4. Simple Baseline Model

A Logistic Regression model was trained on the raw, flattened pixel data to set a critical performance benchmark.

4.1. Method

Raw images were loaded, flattened into a vector of 150,528 features, and used to train a LogisticRegression model with max_iter=1000.

Baseline Result

Model	Validation Accuracy
Logistic Regression (Baseline)	**0.7647 (76.47%)**

Conclusion: The deep learning models must achieve an accuracy significantly higher than 76.47% to demonstrate superior feature learning capabilities.

5. Deep Learning Model Training and Evaluation

Two primary deep learning models were trained. Both utilized the same data generators, class weights, and callbacks (EarlyStopping, ReduceLROnPlateau).

5.1. Model A: Custom Convolutional Neural Network (CNN)

This model was built from scratch and trained using the Adam optimizer and binary_crossentropy loss.

Architecture Summary

The architecture featured 3 convolutional blocks (32, 64, 128 filters) followed by a 512-unit dense layer and a 0.5 dropout layer before the final sigmoid output.

- **Total Trainable Parameters:** 44,396,609.

Training Results

Metric	**Train (Best Epoch 4)**	**Validation (Best Epoch 4)**
Accuracy	0.8104	**0.7600**
Loss	0.4916	**0.5018**

Custom CNN Evaluation (Validation Set)

Metric	**Precision**	**Recall (Sensitivity)**	**F1-Score**	**Support**
no (No Tumor)	0.69	0.47	0.56	19
yes (Tumor)	0.73	**0.87**	0.79	31
Overall Accuracy			**0.72**	50

Evaluation: The model's primary strength is **'Yes Tumor' Recall (87%)**. This means it is highly effective at identifying actual tumor cases (minimizing False Negatives), which is a critical priority in medical screening.

5.2. Model B: EfficientNetB0 (Transfer Learning)

This model leveraged pre-trained weights to utilize robust features learned from the massive ImageNet dataset.

Training Strategy

1. **Stage 1 (Feature Extraction):** EfficientNetB0 base was frozen; only the custom classification head was trained. (Best Val Acc: 0.62)
2. **Stage 2 (Fine-Tuning):** The top 30 layers of the base were unfrozen and trained with a very low learning rate (1e-5).

Training Results

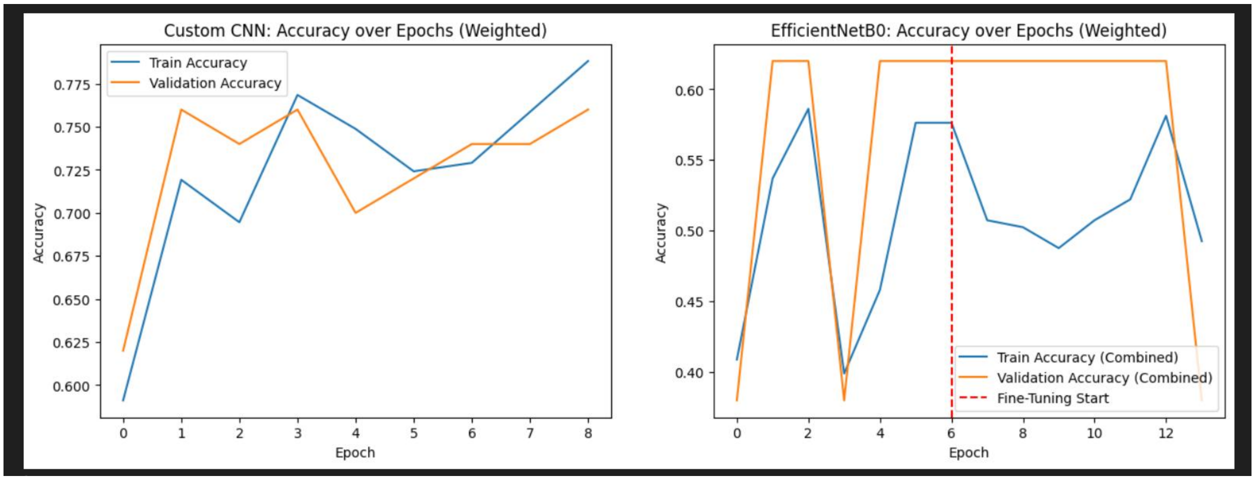
The model consistently stabilized at a low accuracy, even after fine-tuning.

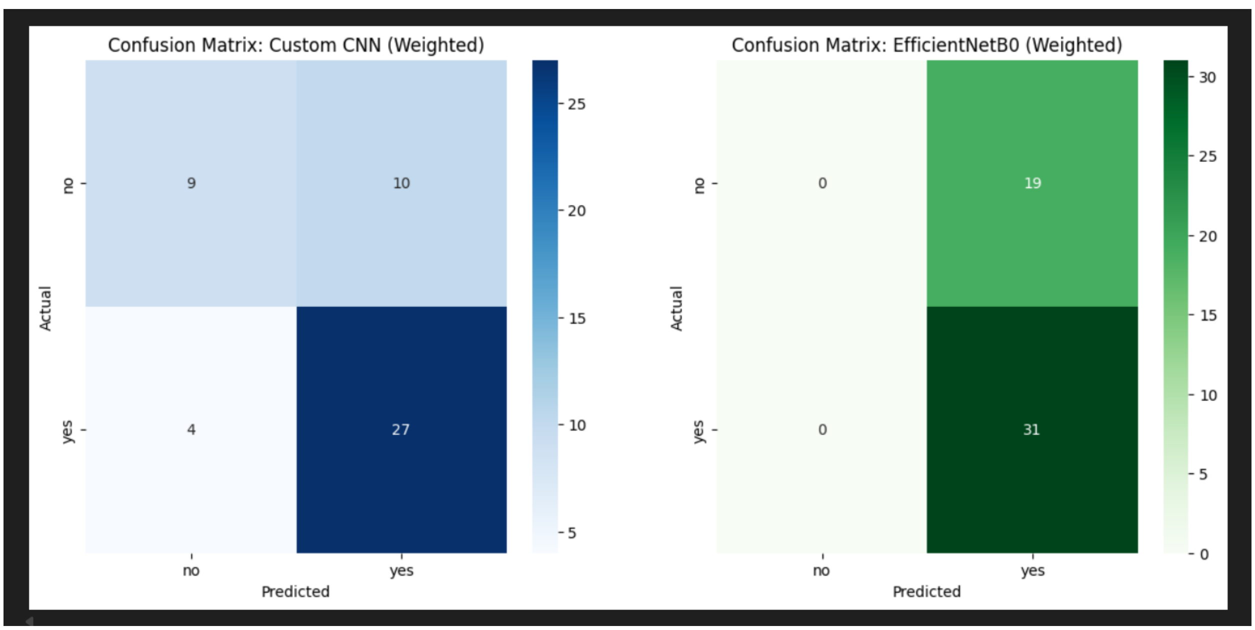
Metric	Train (Best Epoch 8)	Validation (Best Epoch 8)
Accuracy	0.5193	**0.6200**
Loss	0.7053	**0.6692**

EfficientNetB0 Evaluation (Validation Set)

Metric	Precision	Recall (Sensitivity)	F1-Score	Support
no (No Tumor)	**0.00**	**0.00**	0.00	19
yes (Tumor)	0.62	1.00	0.77	31
Overall Accuracy			**0.62**	50

Critical Evaluation: The model exhibited a classification failure, predicting **'yes' (tumor) for every single sample** (0 True Negatives, 0 False Positives). Its 62% accuracy is merely a reflection of the majority class distribution and its inability to learn the features relevant to the minority 'no tumor' class.


Custom CNN: Accuracy over Epochs (Weighted)
Train Accuracy
Validation Accuracy
Accuracy
Epoch
EfficientNetB0: Accuracy over Epochs (Weighted)
Train Accuracy (Combined)
Validation Accuracy (Combined)
Fine-Tuning Start
Accuracy
Epoch


Confusion Matrix: Custom CNN (Weighted)
no
yes
Actual
Predicted
9
10
4
27
Confusion Matrix: EfficientNetB0 (Weighted)
0
19
0
31

6. Conclusion and Comparative Analysis

The table below provides a final performance comparison across all models.

Model	Overall Accuracy	'Yes Tumor' Recall (Sensitivity)	Failure Mode
Logistic Regression	**0.7647**	N/A	High-dimensional data risk.
Custom CNN (Weighted)	0.72	**0.87**	High False Positive rate (low 'no' Recall).
EfficientNetB0 (Weighted)	0.62	1.00	**Total classification collapse (predicts 'yes' every time).**

Final Model Selection

The **Custom CNN (Weighted)** is selected as the best performing model based on its **clinically relevant Recall of 87%** for the 'Tumor Present' class. Although its overall accuracy is slightly below the baseline, its ability to successfully identify most actual tumor cases is the priority in screening.

7. Clinical Implications

The development of this model demonstrates a high potential to support radiologists in brain tumor screening and diagnosis.

1. Prioritize Urgent Cases

The model quickly scans every MRI and automatically flags the patients who most likely have a tumor. This helps the medical team:

- Triage the workload: Doctors see the most critical scans first.

- Speed up treatment: Faster review means a quicker diagnosis and treatment plan for urgent cases.

2. Reduce Diagnostic Errors

The model acts as an independent safety check—a "second pair of eyes"—for the radiologist. This helps to:

- Catch subtle tumors: The AI can spot tiny or complex abnormalities that might be missed during a rapid human review.
- Increase confidence: When both the doctor and the AI agree on a diagnosis, the result is more reliable.

3. Improve Workflow Efficiency

The AI takes over the routine, repetitive tasks of screening, managing the department's heavy workload by:

- Filtering normal scans: The AI quickly confirms the large volume of non-critical cases.
- Saving time: This frees up the radiologist's valuable time, allowing them to focus their expertise on the most complex or challenging patient films.