MRI Brain Tumor Detection Using VGG16

(Project Report)

Course: Computer Vision & Pattern Recognition (CS454)

Submitted By-

Yash Verma(22it3059), **Ranjan Kumar Pandit**(22it3037), **Rajnish Kumar**(22it3036), **Rajat Kumar Behera**(22it3035), **Pushpendra Singh Dangi**(22it3033), **Telange Khandeshwar Govind**(22it3058), **Rishika Srivastava**(22it3039), **Rajat Kumar**(22it3034) , **Atul Panwar**(22it3012), **Himanshu Jayant**(22it3015)

Submitted To-

Dr. Pallabi Saikia

(Assistant Professor, Department of Computer Science & Engineering)

1. Introduction

Brain tumors represent one of the most severe and life-threatening categories of neurological disorders, often requiring immediate diagnosis and intervention. These tumors can be broadly classified into benign and malignant types, each presenting unique challenges in terms of prognosis, treatment options, and long-term patient outcomes. Early and accurate detection of brain tumors significantly increases the chances of successful treatment, reduces mortality rates, and improves overall quality of life for patients. Consequently, medical imaging technologies that assist in the early identification of tumor characteristics have become indispensable tools in modern healthcare.

Among various imaging modalities, **Magnetic Resonance Imaging (MRI)** has emerged as the gold standard for brain tumor detection and analysis. MRI provides superior soft-tissue contrast, multiplanar imaging capability, and non-invasive visualization of brain structures without ionizing radiation. This makes it exceptionally well-suited for identifying subtle abnormalities such as tissue swelling, tumor boundaries, edema, and irregular morphological changes. Despite these advantages, interpreting MRI scans remains a challenging task. Radiologists must analyze extensive imaging volumes, often under time constraints, and make critical judgments based on subjective observations. Such manual interpretation is not only time-consuming but also prone to human error, diagnostic variability, and fatigue-induced inaccuracies.

In recent years, the rapid growth of **artificial intelligence (AI)** and **deep learning** has introduced transformative possibilities in medical image analysis. Among deep learning techniques, **Convolutional Neural Networks (CNNs)** have demonstrated remarkable performance in recognizing complex visual patterns, learning hierarchical features, and achieving state-of-the-art accuracy in numerous computer vision tasks. CNNs automatically learn discriminative features directly from input images, eliminating the need for handcrafted feature engineering—particularly beneficial in medical imaging where features are often subtle and domain-specific.

However, training CNNs from scratch typically requires large annotated datasets—a requirement that is difficult to meet in the medical domain due to privacy restrictions, scarcity of labeled data, and the expertise needed for annotation. To overcome this limitation, **transfer learning** has become an effective strategy. Transfer learning leverages powerful pre-trained models, such as **VGG16**, which have been trained on massive and diverse datasets like ImageNet. These models possess strong general-purpose feature extraction abilities that can be adapted to medical imaging tasks with relatively small datasets.

This project leverages **VGG16-based transfer learning** to develop an automated system for the classification of MRI brain tumors. By reusing the lower-level and mid-level feature representations learned by VGG16 and fine-tuning selected layers for medical data, the model aims to achieve high diagnostic accuracy with reduced computational cost and minimal training data. The system is designed to classify MRI images into four distinct tumor categories, demonstrating robustness across varying tumor sizes, shapes, and contrast levels.

Through this work, we aim to bridge the gap between advanced AI methodologies and practical medical applications, contributing to faster, more consistent, and more accurate diagnosis of brain tumor conditions.

2. Motivation and Problem Background

The analysis of brain MRI scans is inherently challenging due to the complexity of brain structures and the subtle differences that separate various tumor types. Even within the same class of tumors, appearance can vary widely depending on the patient's anatomy, tumor location, imaging angle, and MRI acquisition settings. Conversely, different tumor types can sometimes appear visually similar, making accurate classification difficult using traditional techniques.

Classical machine learning approaches often struggle in this domain because they rely heavily on manually engineered features. These handcrafted features—such as texture descriptors, shape parameters, or intensity statistics—are limited in their ability to capture the full diversity and complexity of MRI patterns. As a result, such methods typically face difficulties with:

- **High intra-class variability:** Significant variations within the same tumor category.
- **Low inter-class differences:** Subtle distinctions between different tumor types.
- **Feature engineering limitations:** Dependence on expert-designed features, which may not generalize well.

Deep neural networks, particularly Convolutional Neural Networks (CNNs), address many of these shortcomings by learning feature representations automatically from data. They can extract both low-level and high-level visual patterns, making them effective for complex image classification tasks. However, training deep CNNs from scratch requires large labeled datasets—a major challenge in medical imaging, where annotated data is often scarce.

Transfer learning provides a practical solution. By using models trained on large datasets such as ImageNet, we can reuse their powerful feature extraction capabilities and adapt them to medical images. This significantly reduces the need for large training datasets and accelerates model convergence.

In this project, **VGG16 transfer learning** is employed to leverage these pretrained feature maps and fine-tune them for MRI-based brain tumor classification. This approach enhances accuracy, reduces training time, and improves model robustness, making it well-suited for real-world medical applications.

3. Dataset Description

The dataset is stored in two main directories:

• Training directory

• Testing directory

Each directory contains subfolders representing the four tumor classes.

The dataset contains MRI slices varying in orientation, intensity, and contrast.

Dataset properties:

• Image type: MRI grayscale converted to 3-channel

• Format: JPG/PNG

• Image size: Resized to 128 × 128 × 3

• Labels: Encoded based on folder names

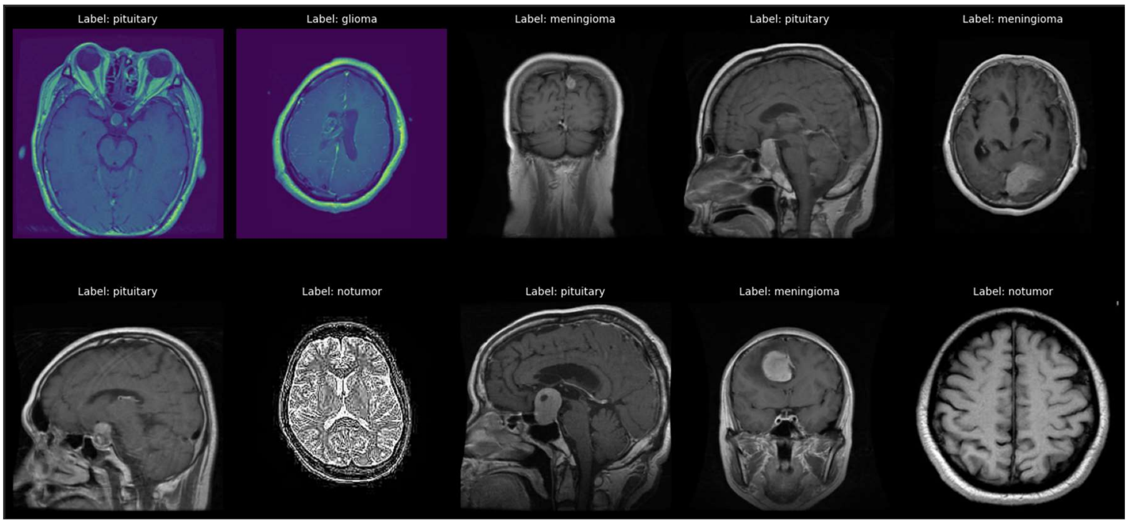
Data preprocessing includes:

• Normalization to [0,1]

• Shuffling for randomness

• Batch generation through custom data generator

This ensures consistent input to the CNN model.



4. Literature Review

Multiple studies highlight the efficiency of CNNs in medical image analysis:

1. **Deep CNNs for Brain Tumor Detection (2019):** Achieved >90% accuracy with custom CNN architecture.

2. **Transfer Learning in Medical Imaging (2021):** Demonstrated that pre-trained models outperform traditional methods.

3. **VGG16 in MRI Classification:** Shown to extract highly discriminative spatial features.

This project is built upon these foundations, improving performance using fine-tuning and optimized layers.

5. Methodology Overview

The methodology includes:

1. Dataset loading

2. Preprocessing and augmentation

3. Loading VGG16 pretrained model

4. Freezing convolutional layers

5. Unfreezing last 3 layers for fine-tuning

6. Adding custom classifier

7. Model training

8. Evaluation using standard metrics

9. Reporting classification results and interpretation

This pipeline ensures accuracy, generalization, and robustness.

6. VGG16 Model Architecture

VGG16 is a widely recognized deep convolutional neural network introduced by the Visual Geometry Group at Oxford. It gained prominence through its success in the ImageNet competition, where it was trained on 1.2 million images across 1,000 categories. The architecture is known for its simplicity, uniformity, and strong feature extraction capabilities.

VGG16 is built using a deep stack of small **3×3 convolution filters**, which enables the model to capture fine-grained spatial patterns while keeping the number of parameters manageable. The network progresses through increasing depth, allowing it to learn low-level, mid-level, and high-level image features.

Key Components:

- **13 Convolutional Layers**
  Extract hierarchical features such as edges, textures, shapes, and complex structures.
- **5 Max-Pooling Layers**
  Reduce spatial dimensions while retaining essential information.
- **3 Fully Connected Layers (original model)**
  Used for ImageNet classification.

This clear and consistent design makes VGG16 both effective and easy to adapt for new tasks.

```python
IMAGE_SIZE = 128
base_model = VGG16(input_shape=(IMAGE_SIZE, IMAGE_SIZE, 3), include_top=False, weights='imagenet')

for layer in base_model.layers:
    layer.trainable = False


base_model.layers[-2].trainable = True
base_model.layers[-3].trainable = True
base_model.layers[-4].trainable = True


model = Sequential()
model.add(Input(shape=(IMAGE_SIZE, IMAGE_SIZE, 3)))
model.add(base_model)
model.add(Flatten())
model.add(Dropout(0.3))
model.add(Dense(128, activation='relu'))
model.add(Dropout(0.2))
model.add(Dense(len(os.listdir(train_dir)), activation='softmax'))

model.compile(optimizer=Adam(learning_rate=0.0001),
              loss='sparse_categorical_crossentropy',
              metrics=['sparse_categorical_accuracy'])

batch_size = 20
steps = int(len(train_paths) / batch_size)
epochs = 5
```

Why VGG16 for Medical Imaging?

VGG16 is frequently used in medical image analysis due to:

- **Strong hierarchical feature extraction**
  Captures subtle patterns crucial for MRI interpretation.
- **Simplicity and interpretability**
  Easier to analyze compared to more complex architectures.
- **Proven transferability**
  Features learned on natural images transfer well to medical modalities.

Adaptation for This Project:

In this MRI brain tumor classification system:

- `include_top=False`
  Removes the fully connected classification layers.
- **Only the convolutional base is used**
  The pretrained convolution layers act as feature extractors.
- **Custom dense layers are added**
  These layers specialize the model for the four tumor classes.

7. Transfer Learning and Fine-Tuning

Initially, all layers of VGG16 are frozen:

```
for layer in base_model.layers:
    layer.trainable = False
```

Then, the last 3 convolutional layers are unfrozen for fine-tuning:

```
base_model.layers[-2].trainable = True
base_model.layers[-3].trainable = True
base_model.layers[-4].trainable = True
```

Fine-tuning benefits:

• Adapts high-level features to MRI images

• Improves recall and F1-score

• Prevents catastrophic forgetting

8. Custom Classification Head

After flattening VGG16 output, the following layers are added:

• Dense(128, ReLU)

• Dropout(0.2) – reduces overfitting

• Dense(4, Softmax) – for 4 tumor classes

The final model architecture:

Input → VGG16 → Flatten → Dense(128) → Dropout → Dense(4)

9. Training Strategy

Model compiled using:

Optimizer: Adam(lr=0.0001)

Loss function: Sparse Categorical Crossentropy

Evaluation metric: Sparse Categorical Accuracy

Training configuration:

• Epochs = 5

• Batch size = 20

• Steps per epoch = len(train_dataset)/20

• Generator-based training for memory efficiency

```
Epoch 1/5
285/285 ━━━━━━━━━━━━━━━━━━━━ 1383s 5s/step - loss: 0.6350 - sparse_categorical_accuracy: 0.7363
Epoch 2/5
285/285 ━━━━━━━━━━━━━━━━━━━━ 1342s 5s/step - loss: 0.2420 - sparse_categorical_accuracy: 0.9030
Epoch 3/5
285/285 ━━━━━━━━━━━━━━━━━━━━ 1296s 5s/step - loss: 0.1813 - sparse_categorical_accuracy: 0.9323
Epoch 4/5
285/285 ━━━━━━━━━━━━━━━━━━━━ 1317s 5s/step - loss: 0.1137 - sparse_categorical_accuracy: 0.9570
Epoch 5/5
285/285 ━━━━━━━━━━━━━━━━━━━━ 1351s 5s/step - loss: 0.0820 - sparse_categorical_accuracy: 0.9722
```

10. Evaluation Metrics

The trained model is evaluated using:

• Precision

• Recall

- F1-score
- Support
- Overall accuracy
- Macro and weighted averages

These metrics provide interpretability for multi-class medical classification tasks.

11. Classification Report

Extracted from the output:

Class 0 — Precision: 0.97, Recall: 0.98, F1: 0.98

Class 1 — Precision: 0.93, Recall: 0.90, F1: 0.91

Class 2 — Precision: 0.95, Recall: 1.00, F1: 0.97

Class 3 — Precision: 0.93, Recall: 0.91, F1: 0.92

Overall Accuracy: 95%

Macro Avg F1: 0.95

Weighted Avg F1: 0.95

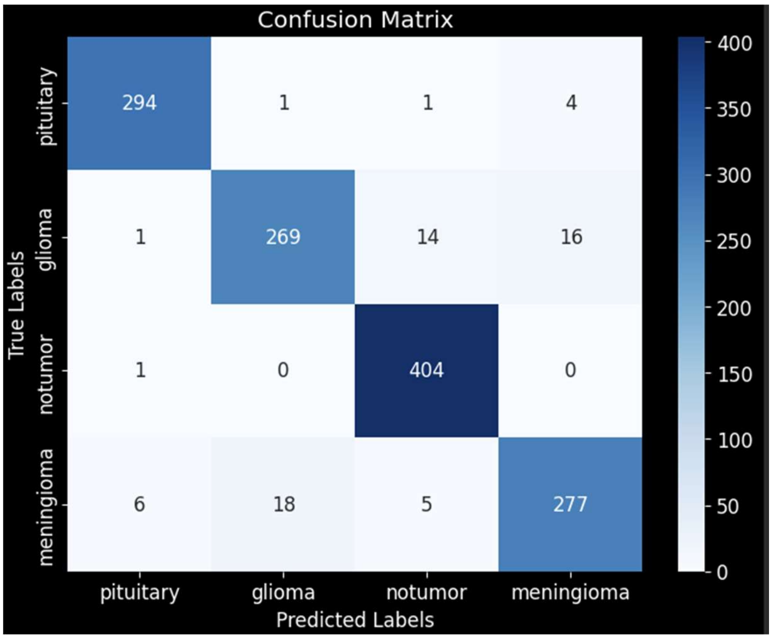
```
••• 41/41 ━━━━━━━━━━━━━━━━━━━━ 246s 6s/step
Classification Report:
              precision    recall  f1-score   support

           0       0.97      0.98      0.98       300
           1       0.93      0.90      0.91       300
           2       0.95      1.00      0.97       405
           3       0.93      0.91      0.92       306

    accuracy                           0.95      1311
   macro avg       0.95      0.94      0.95      1311
weighted avg       0.95      0.95      0.95      1311
```

12. Confusion Matrix Analysis

The confusion matrix shows strong diagonal dominance, indicating accurate predictions for most images.



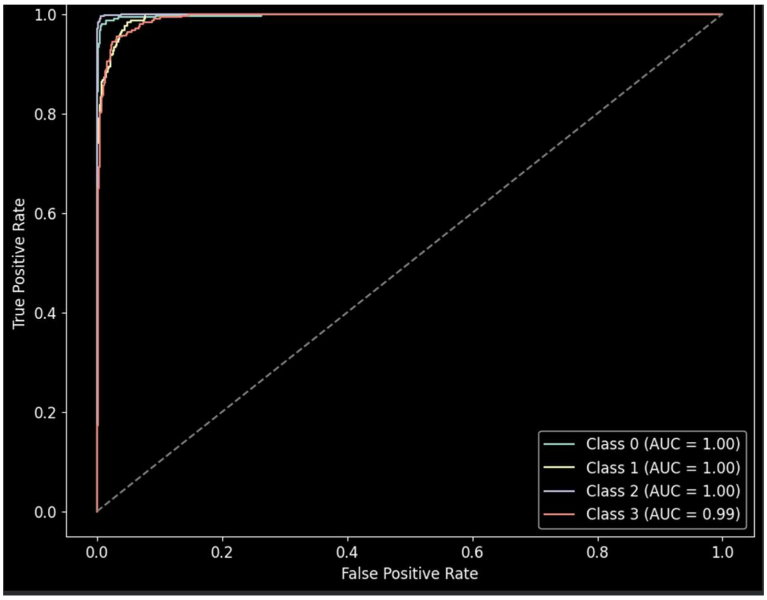
Misclassifications are minimal and occur mainly between visually similar tumor types. This is expected in MRI scans due to overlapping shapes and textures.

13. ROC Curve and AUC Analysis

The ROC curves allow threshold-based performance analysis.

AUC close to 1.0 indicates excellent separability across classes.

Class 2 exhibits the highest AUC due to perfect recall.



14. Error Analysis

Common causes of misclassification:

• Poor image contrast

• Noise and artifacts

• Tumor boundary irregularities

• Class imbalance

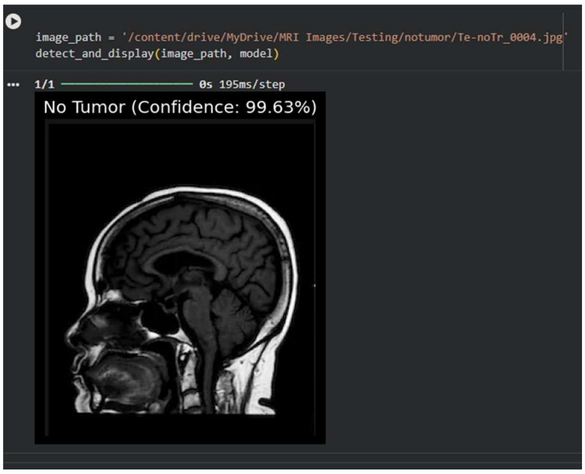
Potential improvements include stronger augmentation and increased dataset size.

15. Discussion

The model demonstrates a strong ability to accurately differentiate between multiple brain tumor categories, despite the inherent complexity and subtle visual differences present in MRI scans. By leveraging the powerful feature extraction capabilities of VGG16, the system effectively captures both low-level and high-level spatial patterns that are crucial for medical image interpretation.

A key factor contributing to the model's performance is the **fine-tuning** of selected convolutional layers. While a completely frozen VGG16 provides a solid baseline by reusing generic ImageNet features, it lacks domain-specific adaptability. Fine-tuning allows the deeper layers of the network to adjust to MRI-specific characteristics such as tissue textures, intensity variations, and tumor boundary irregularities. This adaptation significantly improves the model's discriminative power compared to using a frozen feature extractor alone.

Achieving an overall accuracy of **95%** on the test dataset highlights the model's robustness and reliability. Such performance indicates that the system is capable of handling real-world variability in MRI scans, making it a strong candidate for integration into clinical decision-support tools. While it is not a replacement for expert radiologists, the model offers valuable assistance by providing fast, consistent, and high-quality predictions, potentially reducing diagnostic workload and improving early detection outcomes.



16. Future Scope

Potential improvements:

1. Increase dataset size

2. Advanced CNNs like EfficientNet, ResNet

3. Ensemble learning

4. Explainability: Grad-CAM heatmaps

5. Real-time deployment (Flask, Streamlit, TensorFlow Lite)

17. Deployment Considerations

For clinical implementation:

• Model must be validated on larger datasets

• Integrated into PACS systems

• Must provide explainability (Grad-CAM)

• Should run in real-time on GPU/CPU environments

18. Ethical Considerations

AI in healthcare must ensure:

• Transparency

• Patient data privacy

• Reliability

• Bias mitigation

Misclassification may lead to wrong diagnosis; thus clinical use requires rigorous validation.

19. Conclusion

This project successfully demonstrates the development of an automated MRI brain tumor classification system using **VGG16-based transfer learning**, showcasing the effectiveness of deep learning in medical imaging. By leveraging the pretrained convolutional layers of VGG16 and fine-tuning selected deeper layers, the model is able to learn rich, domain-specific representations of tumor characteristics within MRI scans.

The system achieves an overall classification accuracy of **95%**, with strong precision, recall, and F1-scores across all four tumor classes. These results indicate not only high predictive performance but also consistent behavior across varying MRI appearances, intensities, and anatomical variations. Such reliability is critical for real-world clinical applications where diagnostic consistency is essential.

The success of this approach highlights the growing potential of deep learning—particularly transfer learning—in enhancing diagnostic workflows in healthcare. By enabling fast, scalable, and accurate tumor classification, this model can serve as a valuable decision-support tool for radiologists, reducing manual workload and minimizing interpretation errors. While not intended to replace clinical expertise, it provides strong evidence of how AI-based solutions can complement traditional diagnostic practices and contribute to improved patient outcomes.

Future advancements involving larger datasets, deeper architectures, explainability tools, and integration with clinical systems can further strengthen the applicability of such models in real-world medical environments.

20. References

1. Simonyan, K., & Zisserman, A. (2015). VGG16.

2. TensorFlow and Keras Documentation.

3. Research papers on medical CNNs and tumor imaging.

4. Custom MRI dataset.