**3.1 Research Design**

The research is designed to address the classification of brain tumors using deep learning models applied to MRI images. The study involves the collection, preprocessing, and augmentation of MRI datasets, followed by the development, training, and evaluation of four deep learning models: VGG16, VGG19, CNN, and InceptionV3. These models are trained and validated on a labeled dataset to classify brain tumors into four categories: glioma, meningioma, no tumor, and pituitary tumor. The overall goal is to determine the most accurate model for brain tumor classification.

**3.2 Data Collection**

The dataset used in this study is the "Brain Tumor MRI Dataset," sourced from Kaggle (masoudnickparvar/brain-tumor-mri-dataset). The dataset is divided into training and testing directories, with images categorized into four classes: glioma, meningioma, no tumor, and pituitary. The dataset is split into training and testing sets, with further splitting of the training data for validation purposes.

**3.3 Data Preprocessing**

Preprocessing is a crucial step to ensure that the MRI images are properly formatted and ready for model input. The images were resized to 224x224 pixels for consistency across all models. Data augmentation techniques were applied using the ImageDataGenerator class to enhance the training dataset by generating variations of the images. The augmentation process included random rotations, width and height shifts, shearing, zooming, and horizontal flipping. These techniques help prevent overfitting and improve the generalization capability of the models. Additionally, images were normalized using preprocessing functions specific to each model, such as vgg\_preprocess and inception\_preprocess.

**3.4 Feature Engineering**

Feature engineering in this study involved the use of pre-trained models, such as VGG16, VGG19, and InceptionV3, to leverage their learned features for brain tumor classification. The pre-trained models were used as feature extractors, with their convolutional layers frozen, while new fully connected layers were added to fine-tune the models for the specific task of brain tumor classification.

**3.5 Model Selection**

The study employs four deep learning models including VGG16, VGG19, InceptionV3, and a custom-built Convolutional Neural Network (CNN). The selection of these models is based on their established performance in image classification tasks, particularly in the medical imaging domain. Each model brings unique strengths, and their architectures are tailored to capture intricate features in MRI images of brain tumors.

3.5.1 VGG16 and VGG19

VGG16 and VGG19 are part of the Visual Geometry Group (VGG) models, which are characterized by their deep architectures consisting of multiple convolutional layers followed by fully connected layers. The key innovation of VGG models lies in the use of very small (3x3) convolution filters throughout the network, which allows them to capture fine details in the images while maintaining computational efficiency.

The general structure of a VGG model can be described by the following equations:

Convolutional Layers: Each convolutional layer applies a filter

Max Pooling: After a set of convolutional layers, max pooling is applied to reduce the spatial dimensions:

​

represents the local region in the feature map.

Fully Connected Layers: The output of the final convolutional layer is flattened and passed through one or more fully connected layers:

​

are the weights and biases of the fully connected layer.

Softmax Output: The final output layer applies a softmax function to produce class probabilities:

VGG16 consists of 16 layers (13 convolutional layers and 3 fully connected layers), whereas VGG19 extends this to 19 layers (16 convolutional layers and 3 fully connected layers). Both models are pre-trained on the ImageNet dataset and fine-tuned on the brain tumor MRI dataset to leverage their strong feature extraction capabilities.

3.5.2 InceptionV3

InceptionV3 is a deep convolutional network that is part of the Inception family. The model is designed to efficiently capture multi-scale features within the image by using the Inception module, which applies different sizes of convolution filters in parallel and concatenates their outputs.

The Inception module can be mathematically described as follows:

Parallel Convolutions: Multiple convolution operations are applied in parallel on the same input

​

​

are the corresponding biases.

Concatenation: The feature maps from the parallel convolutions are concatenated:

Reduction Layer: The concatenated output may be followed by a pooling or dimensionality reduction layer to control the computational cost.

InceptionV3 introduces several optimizations, including factorized convolutions, which further reduce the computational burden while maintaining high accuracy. The model's architecture is more complex than VGG, but it excels in capturing diverse features from the MRI images.

3.5.3 Custom CNN

The custom CNN is designed to serve as a baseline for comparison with the more complex architectures like VGG and Inception. It consists of a series of convolutional layers, followed by pooling layers, fully connected layers, and a softmax output layer.

The custom CNN can be represented by the following equations:

Convolutional Layer: As with VGG, the convolution operation is performed as:

Pooling Layer: A max pooling layer reduces the dimensions of the feature map:

Dense Layer: After flattening the pooled feature map, it is passed through dense layers:

Output Layer: The final dense layer applies the softmax function to produce the predicted probabilities for the four tumor classes:

This model, though simpler, is effective in capturing essential features and serves as a benchmark for evaluating the performance of more sophisticated models.

**3.6 Model Implementation**

Each of the selected models was implemented using TensorFlow and Keras. The VGG16 and VGG19 models were built by loading pre-trained weights and freezing the convolutional layers to retain the learned features. A custom fully connected layer with 256 units, followed by a dropout layer, was added to prevent overfitting. The CNN model was built from scratch, featuring three convolutional layers followed by max-pooling layers, and ending with a dense layer for classification. The InceptionV3 model was also implemented using a pre-trained base, with the top layers replaced by a global average pooling layer, a dense layer, and a dropout layer.

Each model was compiled using the Adam optimizer, with a learning rate of 0.0001, and trained using the categorical cross-entropy loss function. The models were trained for 50 epochs, with batch sizes of 32, and their performance was validated using a separate validation set.

**3.7 Training and Evaluation**

The models were trained on the augmented training dataset, and their performance was evaluated on the test dataset. Training was monitored using the accuracy and loss metrics for both training and validation sets. After training, the models were evaluated using precision, recall, F1-score, and confusion matrix metrics to assess their classification performance.

The evaluation involved generating predictions on the test dataset and comparing them with the true labels. The confusion matrix and classification report were generated to provide a detailed analysis of the models' performance across the different classes. The best-performing models were identified based on their accuracy, precision, recall, and F1-score.

**3.8 Experimental Setup**

The experiments were conducted using a system with sufficient computational resources, including a GPU for accelerated training. The models were implemented in Python using the TensorFlow and Keras libraries. Data visualization and analysis were performed using Matplotlib and Seaborn.

**3.9 Ethical Considerations**

Given that this research involved the use of publicly available medical data, care was taken to ensure that the data was used responsibly, with proper anonymization and no identifiable patient information. The research complies with ethical standards for data usage in medical research, focusing solely on the technical development of models for potential future application in clinical settings.