ISYE 6740 – Fall 2024
Final Project
Brain Tumor MRI Dataset Analysis

Team Members

Deng Kun, Liang Yawen, Ng Wee Ding

Problem Statement

Brain tumors are critical medical conditions requiring precise and timely diagnosis for effective treatment. Magnetic Resonance Imaging (MRI) is widely used for diagnosing brain tumors; however, manual interpretation is expensive, time-consuming, error-prone, and relying heavily on the radiologist's expertise. With growing volumes of MRI data, automated diagnostic methods are needed to support healthcare professionals by improving accuracy and reducing human error, and most importantly, saving cost and faster turnaround.

Our project aims to build a robust pipeline for automated analysis of brain tumor MRI images. Using both unsupervised and supervised machine learning techniques, as well as the advanced deep learning model, we strive to classify and detect tumors effectively. Our approaches include:

- Unsupervised methods: Clustering algorithms, specifically PCA, K-Means, and GMM.
- Supervised methods: Traditional classifiers, specifically Bayes, KNN, Logistic Regression, and SVM to predict tumor classes.
- Deep learning: Multilayer Perceptron model and CNN-based architecture, specifically ResNet model, to leverage advanced feature extraction for improved diagnostic accuracy.

Data Source

[1] BRAIN TUMOR MRI DATASET:
HTTPS://WWW.KAGGLE.COM/DATASETS/MASOUDNICKPARVAR/BRAIN-TUMOR-MRI-DATASET/DATA

Data Description

The Brain Tumor MRI Dataset [1] contains 7023 MRI images that are divided into four categories, namely Glioma, Meningioma, No Tumor and Pituitary. The downloaded datasets are organized into two directories, Training and Testing, and 4 subfolders under each directory to hold the above-mentioned four categories of images. The downloaded datasets have the data already pre-partitioned into training and testing. To have comparable results across, we decided to follow the originally pre-partitioned datasets as it is, where we have 82% of training data and 18% of testing data. This is important for our study such that results are consistent and comparable. Since the dataset has the 4 classes distributed almost evenly on both training and testing dataset [FIGURE 1]., no further work needed for dataset reorganization or separation.
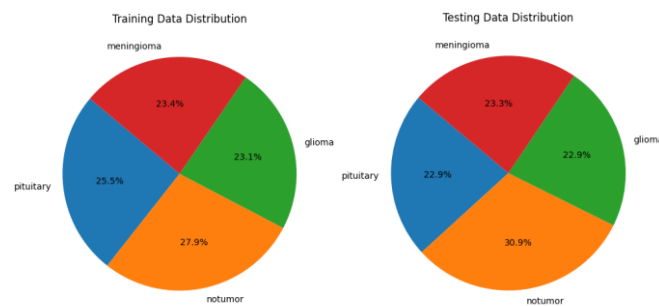

FIGURE 1: DISTRIBUTION OF THE 4 CLASSES OF TUMORS IDENTIFICATION IN BOTH TRAINING AND TESTING DATASET.

Data Pre-processing

The image sizes are non-uniform, a preprocessing step is required to standardize the sizes for all the images. We resized the images into 256x256 and removed the extra margins. To give you an idea how the images looked like after the preprocessing, 10 images are picked randomly in training set, for each category to be demonstrated in [FIGURE 2].
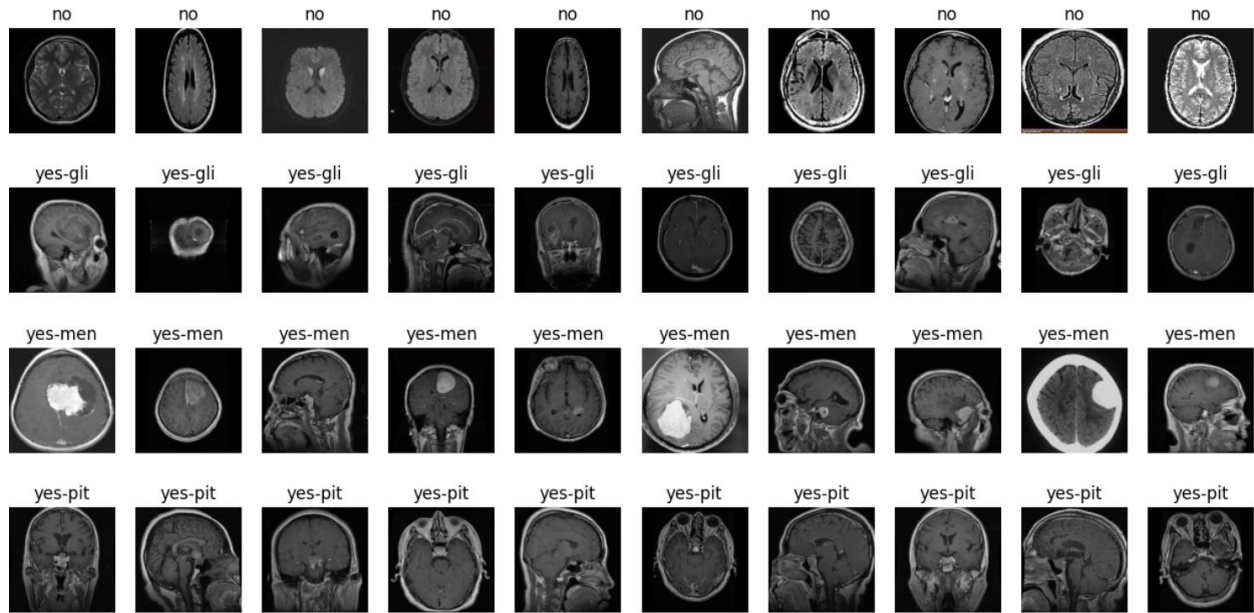
FIGURE 2: 10 MRI IMAGES OF NO TUMOR, AND 10 IMAGES OF EACH OF THE 3 TYPES OF BRAIN TUMORS (GLIOMA, MENINGIOMA, PITUITARY) [1]

As you can see, it wouldn't be straight forward to classify the tumors manually without specialized experts' opinions. Hence, an automatic systematic approach is needed.

Methodology

This is an image classification problem. We study the application of unsupervised models and supervised models on the datasets and evaluate their performance. For unsupervised learning, the models considered are PCA and Kernel PCA. For supervised learning, the models considered are k-nearest neighbors (KNN), support vector machine (SVM), logistic regression. We further develop 2 types of neural-network models, multi-layer perceptron (MLP)[6] model and convolutional neural network (CNN) model - Residual Network (ResNet). We are interested in learning the models' performance on images classification, evaluated with metrics, such as accuracy, precision, recall, F1-score and ROC curves.

Our code is written in Python and utilizes libraries from the *sklearn* and *PyTorch* packages.

Results Analysis

Unsupervised Learning – PCA

Principal component analysis (PCA) is a technique for dimensional reduction and features extractions. We fit the PCA on the training dataset, the projections of the training and test data's top two principal components, are plotted as the following [FIGURE 3]. This is implemented with *scikit's PCA* package. The results shown that the MRI images datasets from each class cannot be linearly separated, there is no straight line that can separate the images of each of the four classes. Moreover, the classes are highly overlapping one another, which makes it impossible to separate them effectively purely with PCA technique.
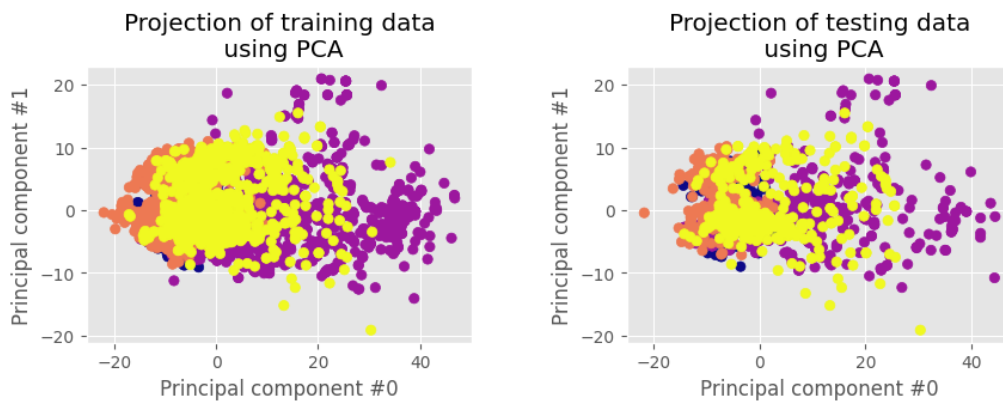

FIGURE 3: PCA AND KERNEL PCA (RADIAL BASIS FUNCTION, RBF) ON TESTING AND TRAINING DATASETS

Unsupervised Learning – Clustering Algorithms: K-means and EM Clustering

K-means clustering is a type of hard clustering algorithm that partitions data into K-distinct clusters, whereas Expectation-maximization (EM) clustering is a type of soft clustering that assigns probabilities to cluster memberships.
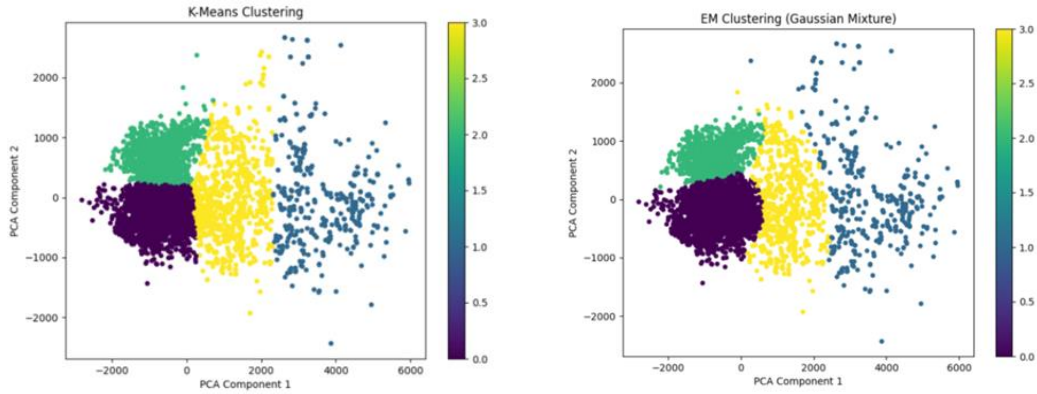
FIGURE 4: K-MEAN CLUSTERING AND EM CLUSTERING (GAUSSIAN MIXTURE) ON TRAINING DATASETS

Our results [TABLE 1] run on the brain tumor dataset indicate that both methods achieve moderate accuracy, with EM performing slightly better, but both methods struggled with the data's inherent complexity. Due to the nature of the PCA projected data complexity and non-linearity, limited clustering performance is achieved.

	K-mean clustering	EM clustering
Accuracy	47.44%	49.12%
Adjusted Rand Index (ARI)	0.1711	0.1763

TABLE 1: K-MEAN CLUSTERING AND EM CLUSTERING (GAUSSIAN MIXTURE) ON TRAINING DATASETS

Our study indicates that MRI images have complex structures that are hard to separate using PCA - unsupervised methods. Dimensionality reduction with PCA likely removed important tumor-specific features, impacting clustering results. However, PCA with combination of supervised learning technique, such as SVM, might work better [6]. Our report will further investigate this area subsequently.

Supervised Learning – Bayes' Theorem

The Naive Bayes classifier achieved an overall accuracy of 60.45% [FIGURE 5], with strong performance for No Tumor and Pituitary, but struggled significantly with Meningioma. While the simplicity of Naive Bayes makes it a quick baseline, more sophisticated models and feature engineering are recommended for improved performance.
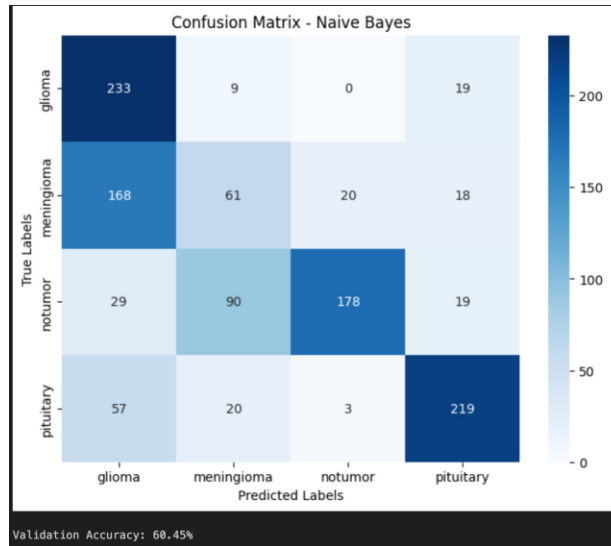
FIGURE 5: NAIVE BAYES CONFUSION MATRIX BASED ON VALIDATION DATASET

Supervised Learning – KNN

The KNN model performed well with an accuracy of 85.48% [FIGURE 6]., achieving strong classification results for most classes. However, the confusion between Glioma and Meningioma highlights areas for improvement in feature representation and model optimization.
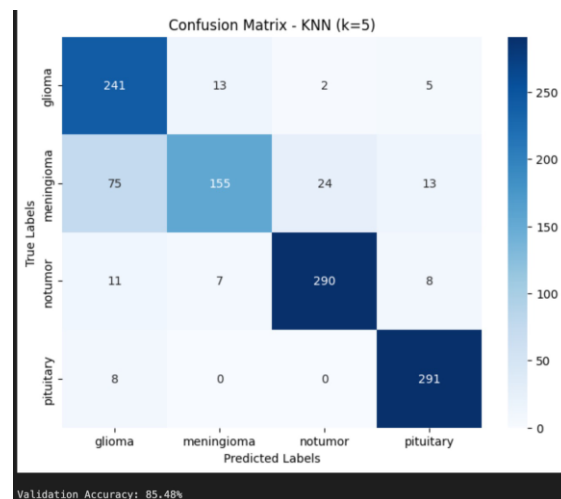

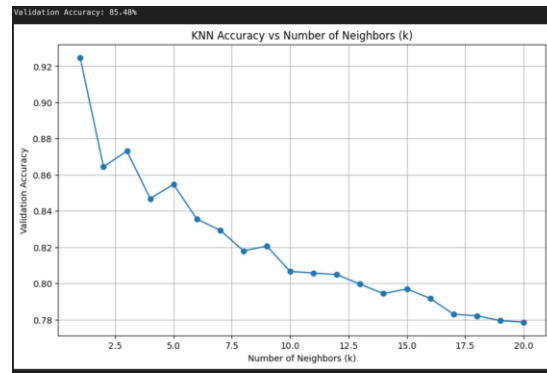FIGURE 6: KNN CONFUSION MATRIX BASED ON VALIDATION DATASET

FIGURE 7: KNN ACCURACY VS NUMBER OF NEIGHBOR

Supervised Learning – Logistics Algorithm

Logistic Regression performed well with an overall accuracy of 86% [FIGURE 8]., particularly excelling in the No Tumor and Pituitary categories. However, the model shows room for improvement in distinguishing between Glioma and Meningioma, suggesting the need for enhanced feature representation or alternative classification techniques.
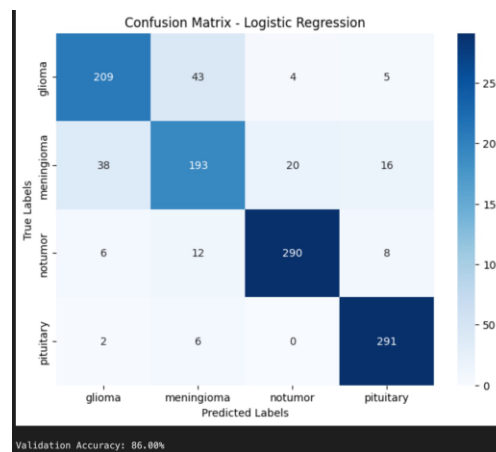

FIGURE 8: LOGISTIC REGRESSION CONFUSION MATRIX BASED ON VALIDATION DATASET

Supervised Learning – SVM and PCA + SVM

SVM with an RBF kernel performed exceptionally well, achieving an accuracy of 88.45% [FIGURE 9]. It excels in distinguishing between most tumor types, particularly No Tumor and Pituitary, but still shows some confusion between Glioma and Meningioma. With appropriate feature engineering and further optimization, SVM could achieve even better performance.
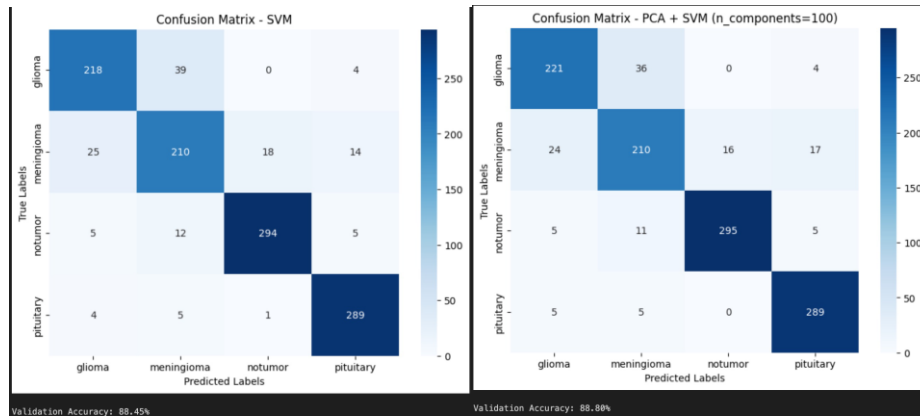
FIGURE 9: SVM AND PCA-SVM CONFUSION MATRIX BASED ON VALIDATION DATASET

Using PCA before SVM slightly improves accuracy (+0.35%) [FIGURE 9]. This suggests that reducing dimensionality with PCA helps the SVM model focus on more informative features while ignoring noise or redundant data.
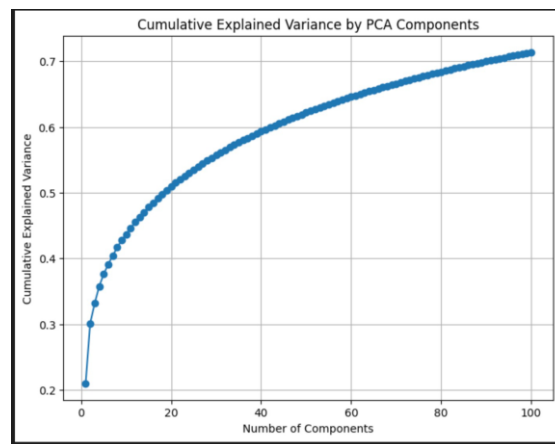

FIGURE 10: CUMULATIVE EXPLAINED VARIANCE BASED ON NUMBER OF PCA COMPONENTS

In comparison to Naive Bayes, KNN, Logistic regression, and SVM models; SVM and PCA + SVM offer the best performance for this dataset, achieving accuracy above 88%. PCA marginally improves SVM performance while reducing computational complexity, making it a strong choice for preprocessing. Naive Bayes, while efficient, is unsuitable for this dataset due to its simplistic assumptions. Finally, both KNN and Logistic Regression provide strong baseline models but fall short of SVM's performance, especially for overlapping classes. Below are some additional confusion matrix insights.

Glioma and Meningioma Misclassification:
- Across all models, there is significant confusion between Glioma and Meningioma due to overlapping feature spaces.

- SVM and PCA + SVM reduce this misclassification compared to simpler models like Naive Bayes and KNN.

No Tumor and Pituitary:
- All models classify No Tumor and Pituitary accurately, with minimal misclassification.
- SVM and PCA + SVM achieve near-perfect classification for these categories.

Deep Learning – Multilayer Perceptron (MLP) and Convolutional Neural Network (CNN) - ResNet

Multilayer Perceptron (MLP) is a class of artificial neural networks to solve classification problems, which are known to handle non-linear characteristics in complex datasets. In a nutshell, MLP networks contain multiple layers of connected nodes, as known as neurons, where their weights can be trained iteratively until the ideal until the model converges. The setup of our MLP model contains: 2 hidden layers, each having 64 and 32 neurons respectively.

Our MLP Architecture:
INPUT LAYER (128 NEURONS) --> HIDDEN LAYER 1 (64 NEURONS) --> HIDDEN LAYER 2 (32 NEURONS) --> OUTPUT LAYER (4 NEURONS)

ResNet stands for Residual Network, it is a class of convolutional neural network (CNN). ResNet50 refers to deep neural network architecture that consists of 50 weight layers (48 convolution layers, 1 maxpool and 1 average pool layer. It is a variant of ResNet model. It is a popular choice for image classification problems, on huge datasets like ImageNet, due to its proven performance. [5][8]. We implemented the MLP and CNN models using *sklearn.neural_network* and *pytorch* packages.
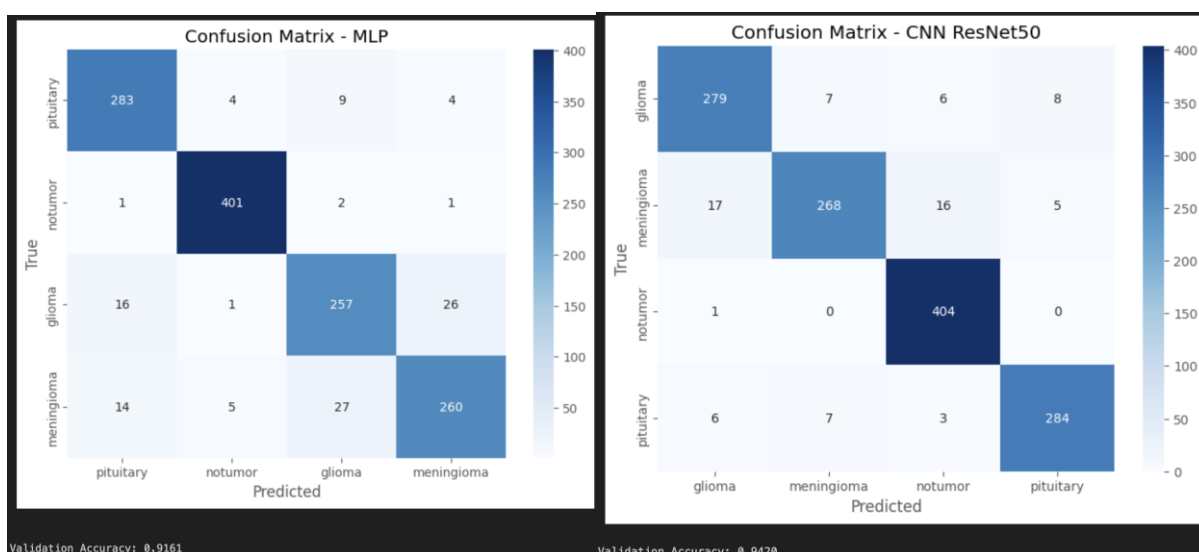

FIGURE 11: CONFUSION MATRIX FOR MLP AND RESNET50 MODELS ON TESTING DATASETS

Our study [Figure 11] shows that MLP and ResNet50's validation accuracy are 91.61% and 94.20% respectively. Neural Network models have outperformed all traditional unsupervised and supervised models which we have studied earlier.
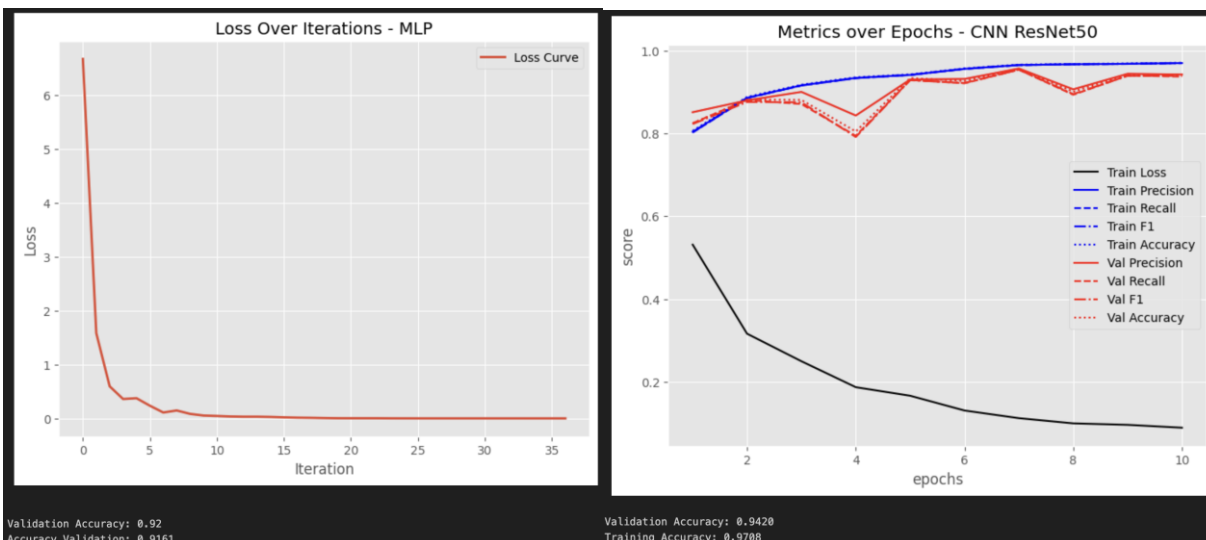

FIGURE 12: TRAINING LOSS OVER ITERATIONS FOR MLP AND RESNET50 ON TRAINING DATASETS

The training loss is decreasing fast on both MLP and ResNet50 models [Figure 12] as it stabilizes in less than 10 iterations.
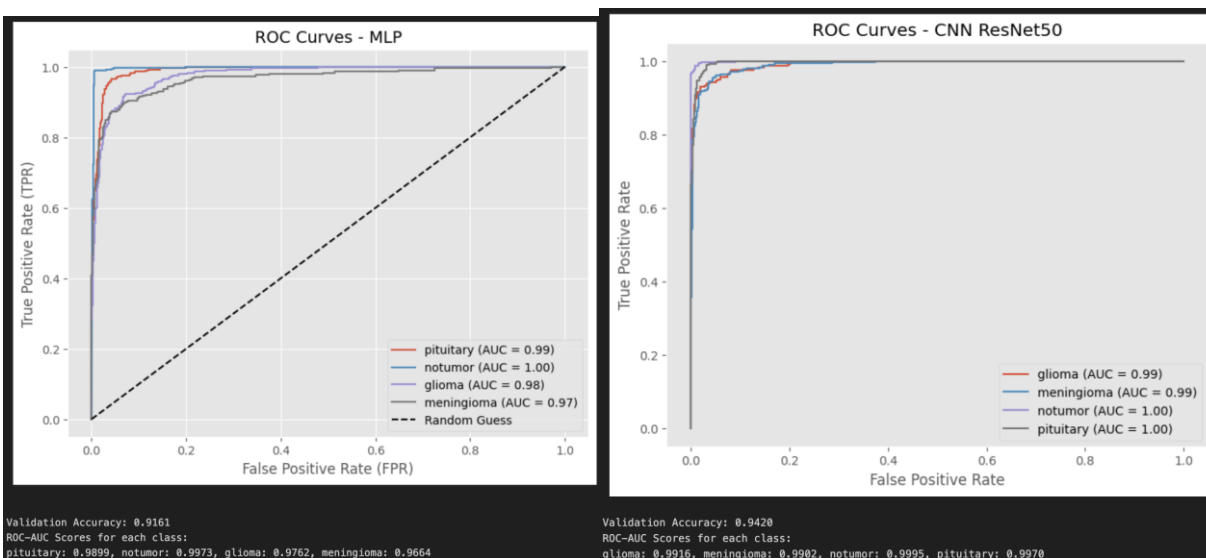

FIGURE 13: ROC CURVES FOR MLP AND RESNET50 ON TESTING DATASETS

Both models achieve great classification results on the brain tumor MRI images dataset. One interesting note is that when MLP detects *notumor*, it is more precise than ResNet50 [FIGURE 11].

In short, although ResNet50 has higher accuracy rates than MLP model in our case, practically we might adopt MLP model as cost of false negative for *notumor* is huge and we'd rather lower the accuracy but having the model to reduce the false negative for *notumor* class.

Images which are misclassified are plotted [FIGURE 14] in comparison with images which are classified correctly.
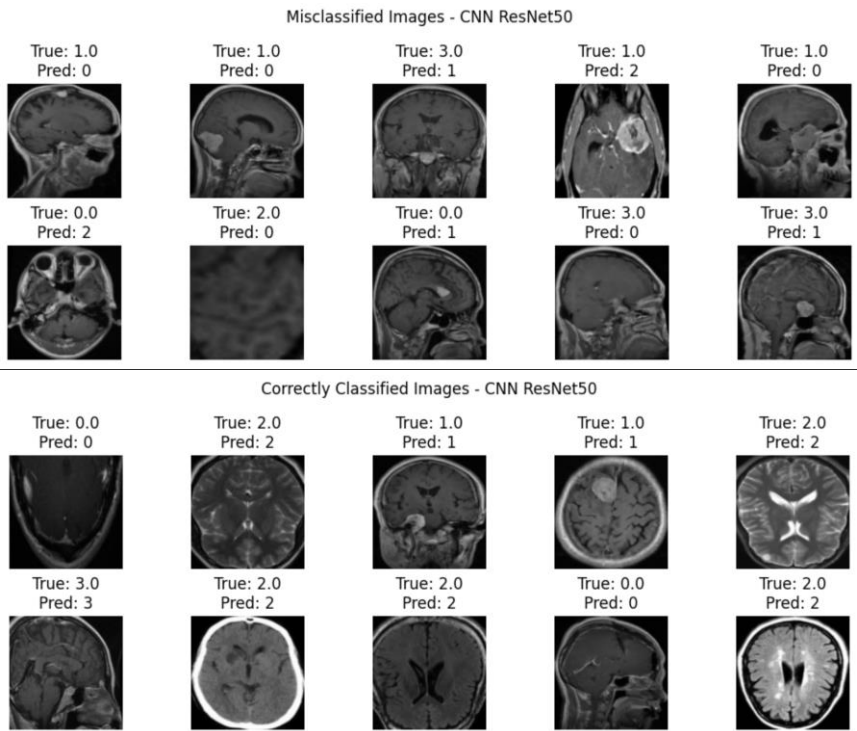

FIGURE 14: RANDOM MISCLASSIFIED IMAGES AND CORRECTLY CLASSIFIED IMAGES FROM RESNET50 MODEL

Conclusion

We presented our results summary for all the machine learning methods, which covers accuracy, precision and sensitivity [8] on detecting, on particularly, benign tumor or no tumor, which is what patients and healthcare professionals are concerned about, as the cost of false negative might mean delayed treatments for patients.

	Accuracy	Precision *notumor*	Sensitivity *notumor*
K-means clustering	47.44%	17.89%	20.99%
EM (GMM)	49.12%	15.18%	18.02%
Naive Bayes	60.45%	88.56%	56.33%
KNN	85.48%	91.18%	91.77%
Logistic Regression	86%	92.36%	91.77%
SVM (RBF kernel)	88.45%	93.93%	93.04%
PCA + SVM	88.80%	94.86%	93.35%
ANN - MLP	91.61%	97.57%	99.01%
CNN - ResNet50	94.20%	94.17%	99.75%

TABLE 2: MODELS PERFORMANCE ON TESTING DATASET

In this report, we study unsupervised, supervised and neural networks machine learning models to solve the Brain tumor MRI images classification problem. Overall neural networks exhibit highest accuracy, precision and sensitivity, outperforming all traditional supervised and unsupervised models. Among all traditional models, SVM and Logistic Regression show strong performance. SVM with PCA achieves slightly better performance, taking advantage of dimensionality reduction of possibly noise and irrelevant features. Due to the nature of the dataset, which lacks clear features separability, unsupervised learning performs poorly as the accuracy is less than 50%.

Future Work

We suggest additional research, investigation and validation to be conducted on varieties of approaches [10], such as images segmentation, images augmentation, preprocessing with adaptive filtering, or exploring different types of networks. As the continuous advancement and effort of brain tumor detection systems there is always room for further exploration and improvement.

References

[1] *Brain Tumor MRI Dataset* https://www.kaggle.com/datasets/masoudnickparvar/brain-tumor-mri-dataset/data

[2] *Machine Learning Approach to Brain Tumor Detection and Classification*, By Alice Oh, Inyoung Noh, Jian Choo, Jihoo Lee, Justin Park, Kate Hwang, Sanghyeon Kim, Soo Min Oh https://arxiv.org/abs/2410.12692

[3] *Brain Tumor Types Classification using K-means Clustering and ANN Approach,* Angona Biswas; Md. Saiful Islam, and more, IEEE https://ieeexplore.ieee.org/document/9331115

[4] *Pre-trained deep learning models for brain MRI image classification*, Srigiri Krishnapriya, Yepuganti Karuna https://pmc.ncbi.nlm.nih.gov/articles/PMC10157370/

[5] *Deep Residual Learning for Image Recognition* Srigiri Krishnapriya, Kaiming He, Xiangyu Zhang, Shaoqing Ren, Jian Sun https://arxiv.org/pdf/1512.03385

[6] *Multilayer Perceptrons (MLP)* https://d2l.ai/chapter_multilayer-perceptrons/index.html

[7] *Brain MRI Image Segmentation and Classification using PCA-SVM and CNN,* S. P. S, V. S, Z. A. L, S. V, K. T and S. N. P, 2023 International Conference on Recent Advances in Science and Engineering Technology (ICRASET), B G NAGARA, India, 2023

[8] *ResNet-50 Introduction*, https://srsapireddy.medium.com/resnet-50-introduction-b5435fdba66f

[9] *Confusion Matrix*, https://scaryscientist.blogspot.com/2016/03/confusion-matrix.html

[10] Brain tumor MRI images identification and classification based on the recurrent convolutional neural network, Ramdas Vankdothu, Mohd Abdul Hameed, https://doi.org/10.1016/j.measen.2022.100412