BRAIN TUMOR CLASSIFICATION USING MRI SCANS

MOHMMAD AYAAN
(2022302)
DEPARTMENT: *CSAI*
INDRAPRASTHA INSTITUTE OF INFORMATION TECHNOLOGY, DELHI

NAMAN BIRLA
(2022310)
DEPARTMENT: CSE
INDRAPRASTHA INSTITUTE OF INFORMATION TECHNOLOGY, DELHI

***Abstract*— Detecting and classifying brain tumors accurately is crucial for effective treatment planning. This study proposes a novel approach using machine learning techniques, including deep learning algorithms and ensemble methods, to enhance brain tumor detection and classification. The results demonstrate the potential of this approach to improve patient care through timely and accurate diagnoses.**

I. INTRODUCTION

A brain tumour refers to a discrete accumulation of cells in the brain, classified as benign or malignant, that pose a serious health risk due to their ability to raise intracranial pressure. Importance cannot be. Stated that the importance of early detection and accurate classification of brain tumours is not excessive. This is learned as a process of this compulsive responsibility for the interconnectedness of a cell to be seen by the brain's responsibility, classification, and locality. -taking advantage of the copy, the research process is well structured. Strives to do so, facilitating faster and more accurate clinical interventions to improve patient outcomes.

II. PROBLEM STATEMENT AND MOTIVATION

Brain cancer is relentless, impacting both the young and the old with its devastating effects. For patients and doctors alike, the journey of battling brain tumours is fraught with challenges. The intricacies and complexities of these tumours make accurate identification and classification a daunting task. Despite strides in medical technology, the manual interpretation of MRI scans remains susceptible to errors, leaving room for misdiagnosis and delayed treatment. Moreover, the sheer volume of data generated from these scans overwhelms human capacity, increasing the likelihood of critical details being overlooked.

Combining machine learning and artificial intelligence offers a promising solution in response to these challenges. Deep learning algorithms, such as convolutional neural networks (CNNs) and artificial neural networks (ANNs), have demonstrated a remarkable ability to identify and classify brain tumours from MRI images. By automating this process, these technologies improve the precision and speed of tumour detection and alleviate the burden on healthcare professionals. A robust and reliable automated system for tumour classification could revolutionize patient care, providing timely and accurate diagnoses worldwide and ultimately enhancing treatment outcomes and quality of life.

III. LITERATURE REVIEW

We reviewed different papers for Brain Tumor disease detection papers. Most papers were focused on data preprocessing and data augmentation while working with machine learning techniques segmentation was considered as an important parameter. We used model training and evaluation, techniques like random forests and decision trees have become popular due to their ability to combine the strengths of multiple classifiers. Additionally, principal component analysis (PCA) has been used to simplify feature extraction and reduce the complexity of input data, making classification models more efficient and easier to understand. By combining these approaches, we aim to improve brain tumor detection and classification, ultimately enhancing patient care and outcomes.

We used various Machine Learning techniques in this research like:

1. Principal Component Analysis (PCA), a dimensionality reduction technique that transforms high-dimensional data into a lower-dimensional space while preserving its essential features.
2. Decision Trees, Classifier, a predictive modelling technique that partitions data into branches based on feature values to make sequential decisions
3. Random Forests Classifier, learning method that combines multiple decision trees to improve classification accuracy and generalization to new data.
4. KNN Classifier, a non-parametric algorithm that assigns a data point to the majority class among its k nearest neighbours in the feature space.
5. SVM (Support Vector Machine): SVM is a supervised machine learning algorithm that classifies data by finding an optimal line or hyperplane that maximizes the distance between each class in an N-dimensional space

In SVMs there are various kernel functions that take data as input and convert it into the required shape.

   a) SVM with linear kernel, creates a linear decision boundary to separate classes in the data.
   b) SVM with Radial Base Function (RBF) kernel, to achieve high accuracy by effectively capturing complex patterns in the data through non-linear decision boundaries.
   c) SVM with polynomial kernel, a machine learning model that separates data by fitting a polynomial decision boundary to achieve classification

6. Neural Networks: A machine learning process that works in a layered structure and uses a dense structure of interconnected nodes to learn the data. It mimics the human brain while making decisions.
   a) Convolutional Neural Networks (CNNs): Deep learning models specifically designed for analyzing grid data like images with the help of a layered structure that learns the data patterns and features.

The model's reliability is tested using various parameters like accuracy, recall, precision and f1-score and the Confusion Matrix.

We plot the precision-recall curve. High precision implies low false positive rates. High recall implies low false negative rates. High precision and high recall are desired, i.e. the higher the area under precision-recall curve, the more reliable the model is. F1 Score is the harmonic mean of precision and recall and contains both the parameters equally, hence the better the F1 score, the better the model.

The ROC-AUC curve is also a good parameter for checking the model's reliability, but it is used for Binary Classification hence, not included in this research.

IV. DATASET DETAILS

Our dataset used for training and testing was taken from Kaggle. "It contains 7023 human brain MRI pictures. They were sorted into four groups: glioma, meningioma, no tumour, and pituitary tumours. Images labelled "no tumour" are all from Br35H category. An important step was leaving out some dataset images. Previous research and our own model training showed the glioma class was mislabelled. To avoid errors, we swapped in fig-share images instead of using those potentially wrong glioma cases. This upholds the dataset's accuracy." as mentioned on Kaggle.

One key aspect is the variety in image sizes present in the dataset. To reduce potential issues during training and testing the model, a specific process was followed. This involved resizing all images to the same size after removing extra margins. Carefully preparing the images this way aimed to make the model more accurate and consistent for all images. Around 5712 images were used to develop and evaluate the model through training. Another 1311 different images were used as a test set, which was employed to assess how well the trained model could classify brain tumour images accurately. Splitting the data like this allowed for proper testing of the model's performance and ability to work reliably across various image samples. We also used mixed dataset to check out model for better accuracy.

The training dataset has the following class-wise distribution.

Out of 5712 images, we had 1321 glioma files, 1339 meningioma files, 1595 notumor and 1457 pituitary images as our training data. [Figure 1]

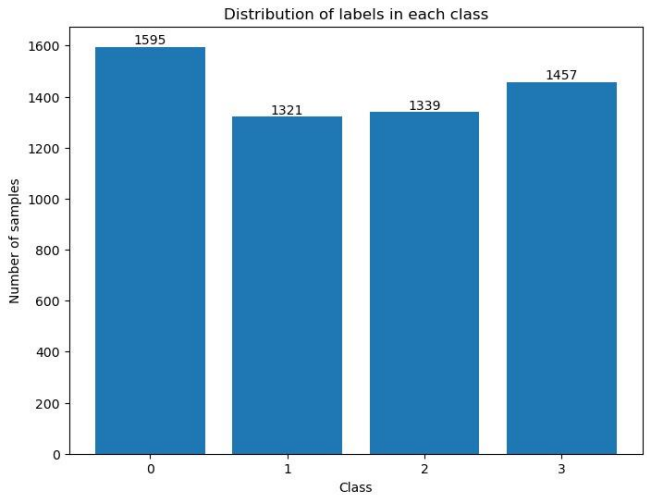


Figure 1: Distribution of training dataset

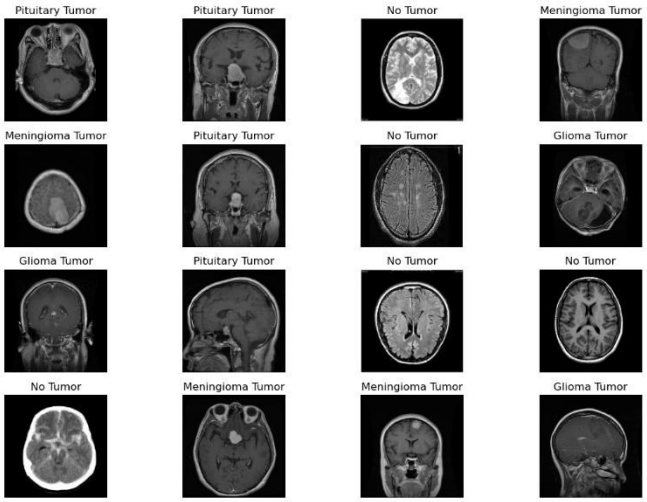


Figure 2: Glimpse of the training dataset

To get an idea of the dataset, the inter-class dependency, we plot a scatterplot of the dataset based on the first two features. [Figure 3]

The plot clearly shows that the data is non- linear i.e. not linearly separable.

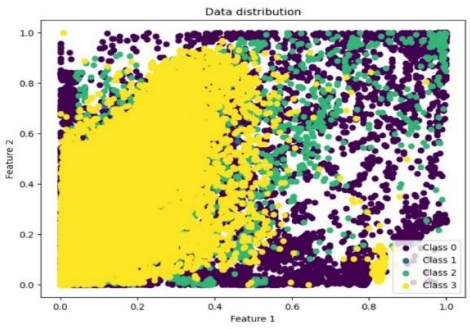


Figure 3: Dataset distribution based on first two features

The test set consists of 1311 test samples for the model to be tested upon.

V. VISUALISATIONS

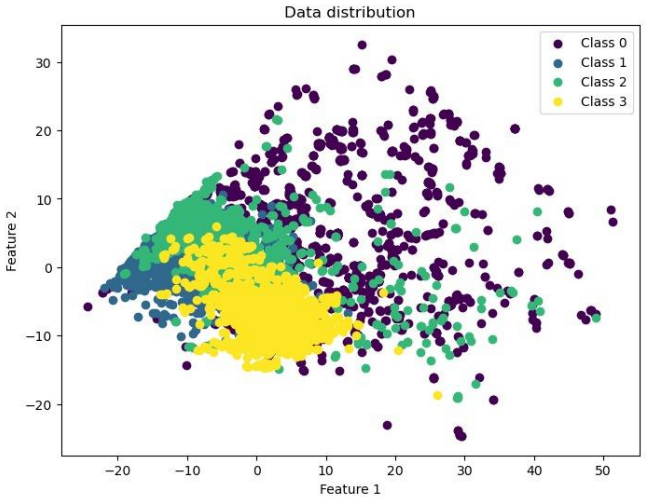


Figure 4: Dataset distribution after applying PCA.

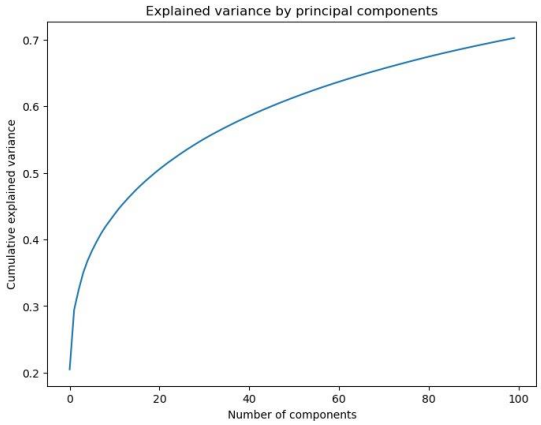


Figure 5: Cumulative Explained Variance by principal components

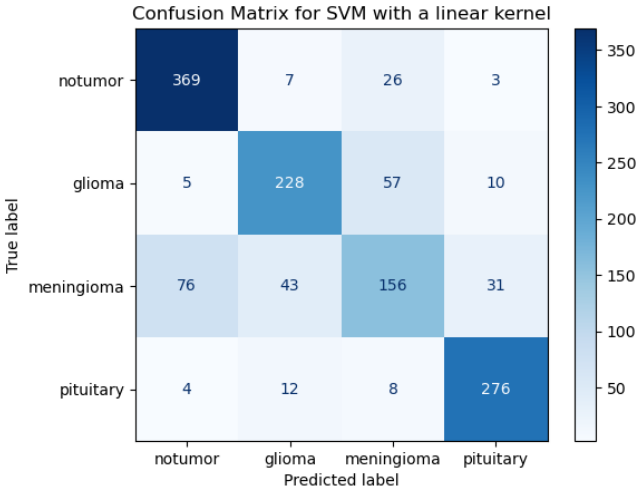


Figure 6: Confusion Matrix for SVM with a linear kernel

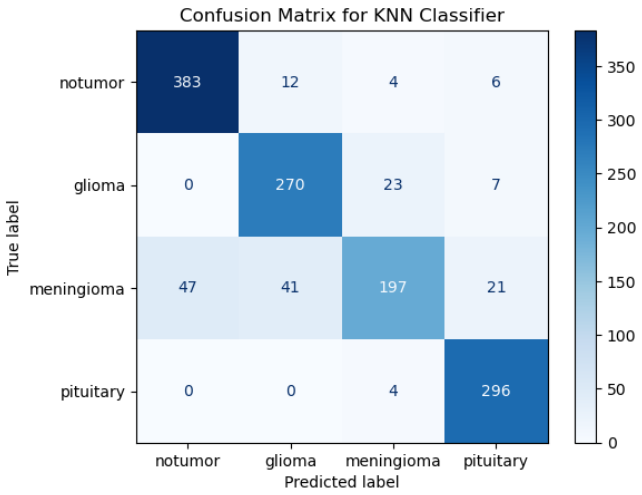


Figure 7: Confusion Matrix for KNN classifier

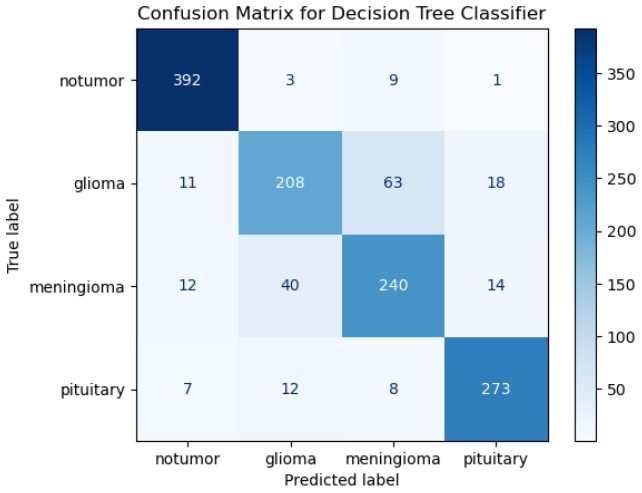


Figure 8: Confusion Matrix for Decision Tree Classifier

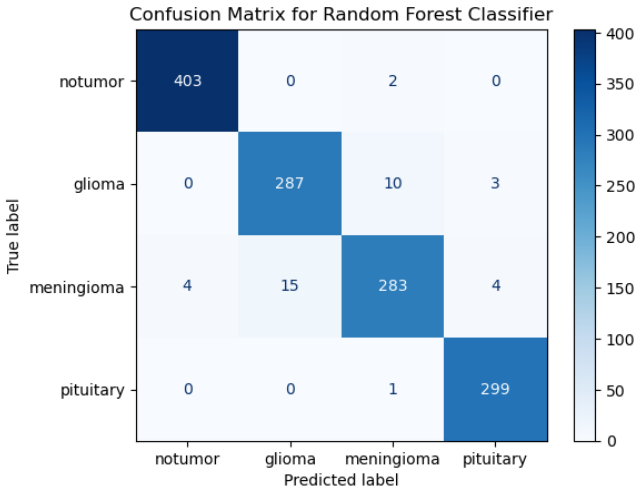


Figure 9: Confusion Matrix for Random Forests Classifier

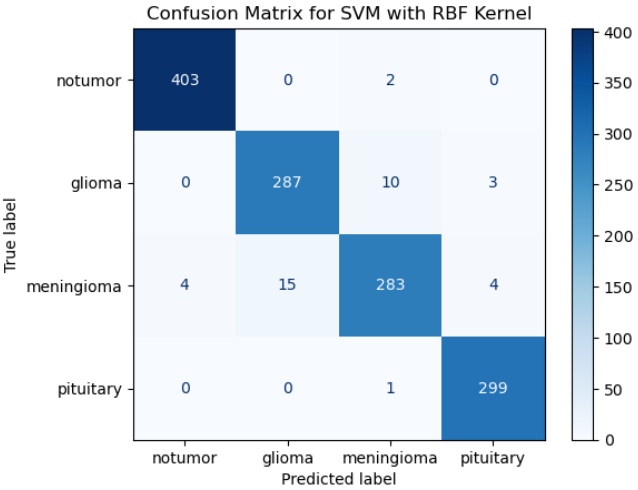
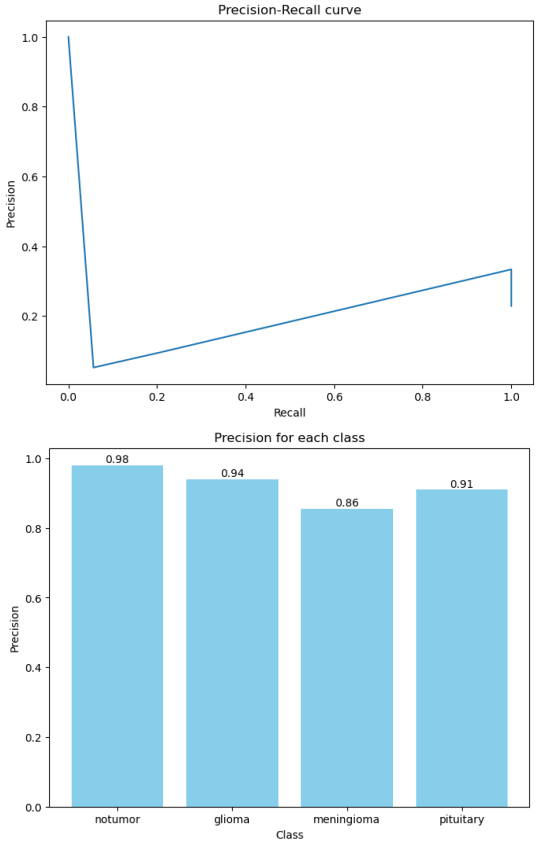


Figure 10: Confusion Matrix for SVM with RBF Kernel



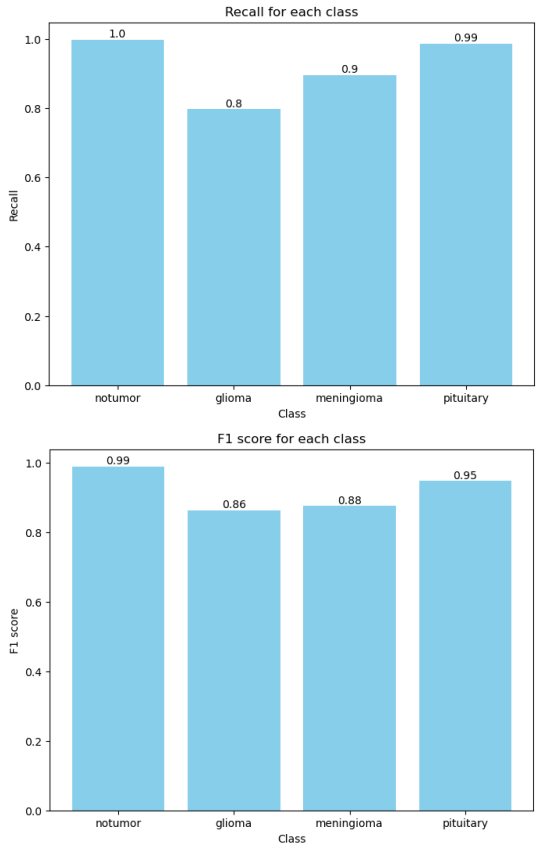


Figure 11: Classification Report for Random Forests Classifier

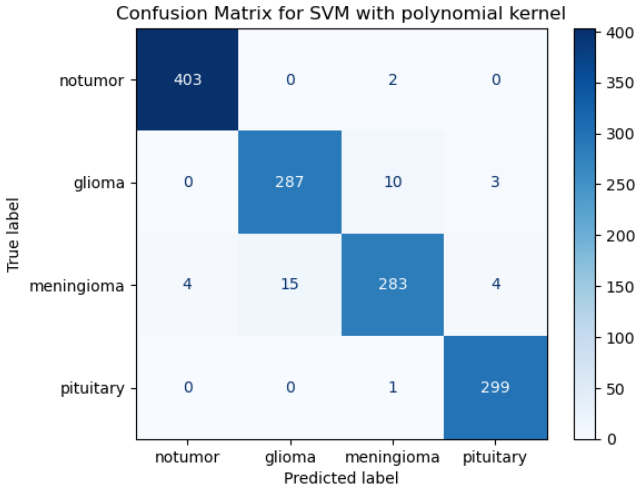


Figure 12: Confusion Matrix for SVM with Polynomial Kernel

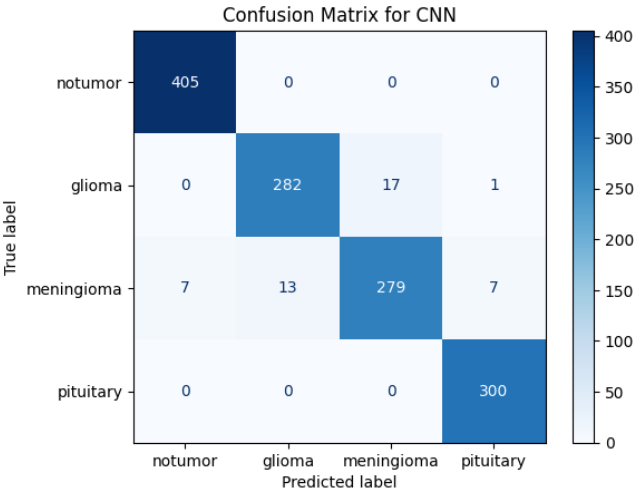


Figure 13: Confusion Matrix for CNN

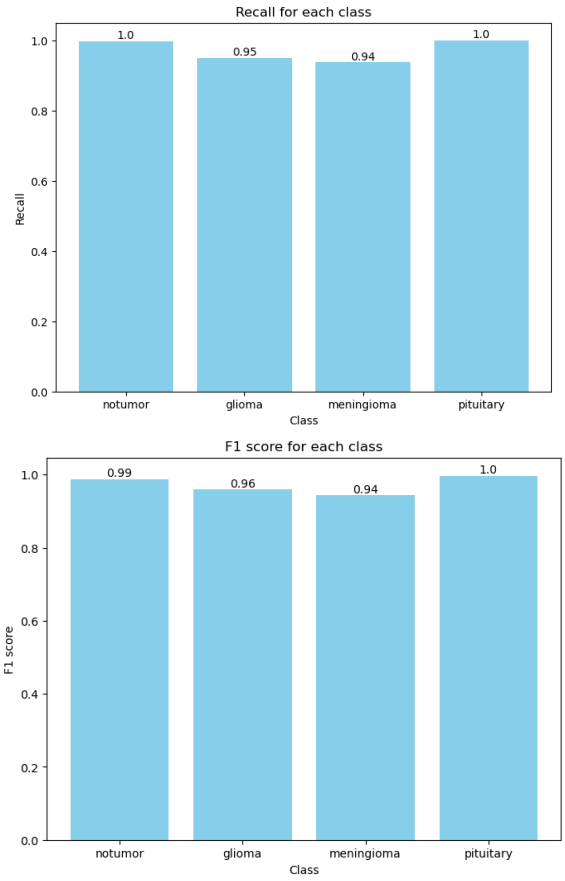
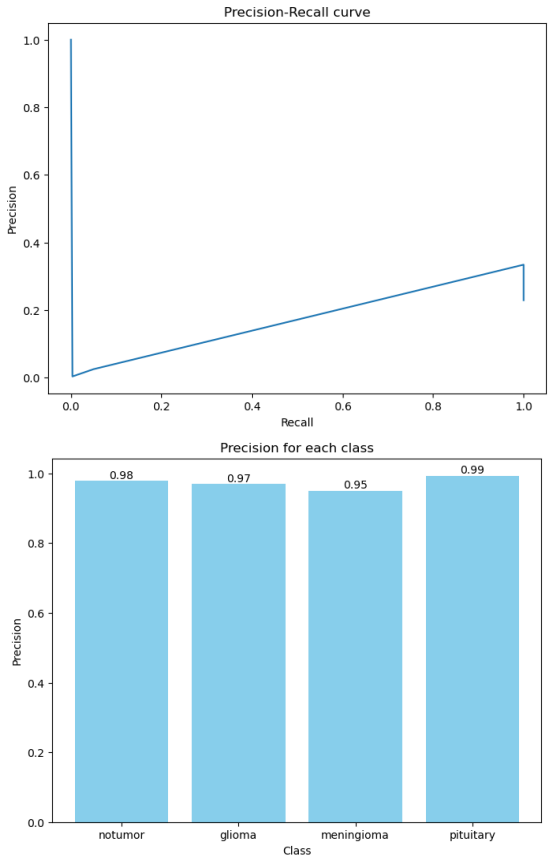
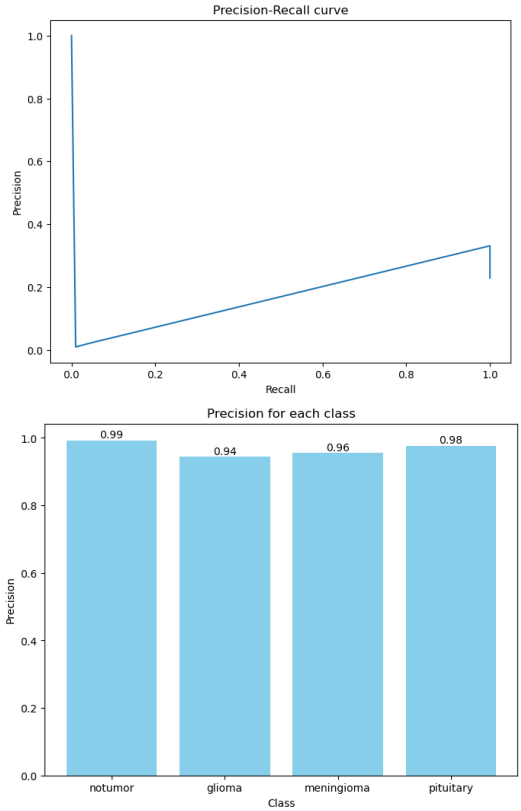


Figure 14: Classification report for SVM with RBF kernel





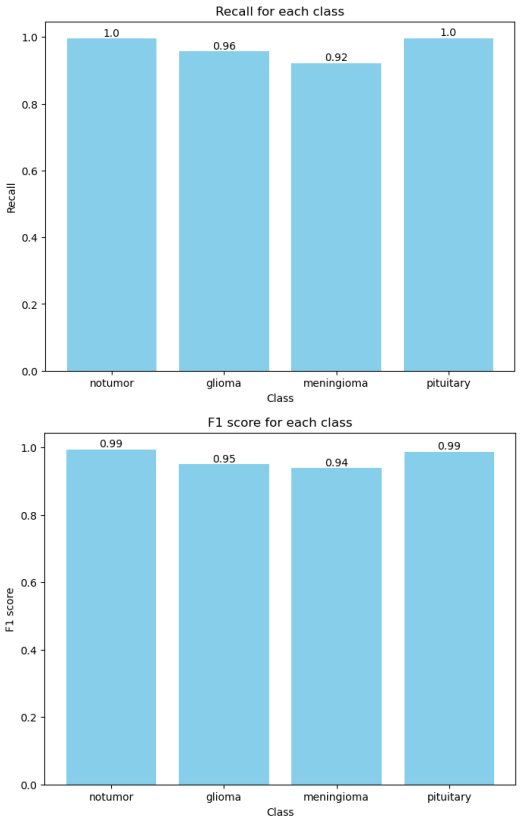


Figure 15: Classification report for SVM with polynomial kernel

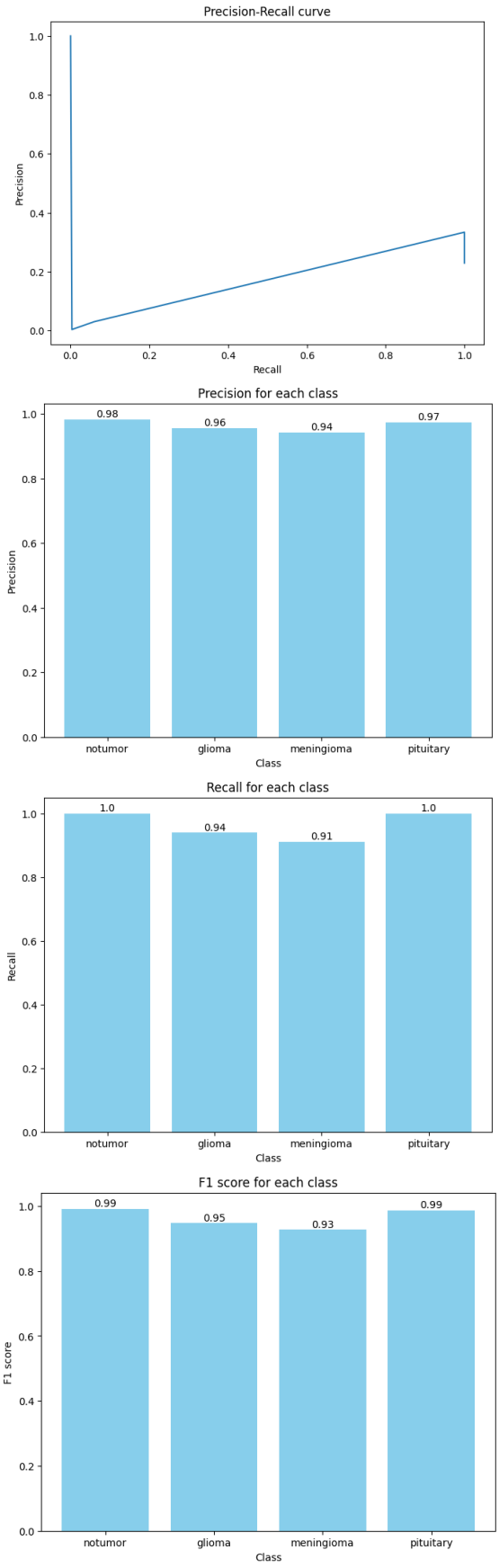


Figure 16: Classification report for classifier using CNN

VI. WORK DONE

We utilized a dataset containing approximately 5712 training MRI images to identify and differentiate brain tumors. The images are initially converted to the same size (of 150x150) and are converted to grayscale to maintain consistency in the data [Figure 1]. This converted dataset is then used for training.

A 150x150 size image consists of 22500 features (each feature being a pixel). Principal Component Analysis was applied to reduce the feature space to 100, while keeping the essential features (and patterns) of the dataset intact. PCA helped in preprocessing and normalizing the data while retaining crucial information necessary for accurate tumor classification. PCA helped in making the data less scattered as compared to dataset without having PCA applied [Figure 4]. This reduced-dimensionality dataset is further used for training the models. The plot of cumulative explained variance with respect to each principal component was captured [Figure 5]. The explained variance increases with the number of principal components signifying that PCA sorts the data in decreasing order of feature importance.

Initially, we tried a linear classifier to classify the data. Since we already observed that the dataset is nonlinear [Figure 2], this was not expected to work. We use Support Vector Machine with a linear kernel to classify the samples. SVM works well with higher dimensional data, hence we use that. We get an accuracy of 78.49% which is not good at all. The confusion matrix for the SVM with linear kernel is shown in [Figure 6]

We shift to non-linear models since the dataset is non-linear. We use the K-nearest Neighbours classifier initially. K nearest neighbors classifies samples based on the labels of its K-nearest neighbors. This is effective in situations where similar classified points are located nearby. In the dataset, KNN Classifier gave an accuracy of 87.41% which is much better than a linear model. The corresponding Confusion Matrix is shown in [Figure 7].
It shows that, while datapoints near to each other have same label, that is not always the case. We can further improve the accuracy by employing techniques that focus more on feature extraction of the sample than using the nearby points to classify it.

We use the ensembling methods: the Decision Tree Classifier and the Random Forests Classifier since they focus more on feature extraction of high dimensional data and mitigating overfitting.
Random Forests grown using multiple decision trees yielded exceptional results yielding an accuracy of 92.52% in classifying the tumors whereas employing a single-decision tree for classification yielded an accuracy of just 84.9%. The corresponding confusion matrices for Decision Tree Classifier and the Random Forests Classifier are shown in [Figure 8] and [Figure 9] respectively.
The random nature of random forests while using multiple decision trees enhanced correct classification and generalization of new data. Application of PCA before training made it more efficient.

Owing to the high dimensionality and probably complex relationships between the data points, we try using SVMs again, with a non-linear kernel like RBF- Radial base function or Polynomial Kernel. Both polynomial and rbf generate a non-linear boundary. RBF is used when data is unevenly distributed whereas polynomial kernel is used when data cannot be separated using single line/hyperplane. Both outperformed the other models yielding very high accuracy.
We set the parameters C and "gamma" for the RBF kernel as 5 and 0.01 respectively. The greater the C, the lower the misclassification error and the smaller the "gamma", the farther the distance between data points to be considered while classification (this captures the uneven distribution of data). Setting these parameters too high and too low results in overfitting and hence we obtain the optimum value. SVM with RBF kernel gave an accuracy of 97.33% which is excellent.
The SVM with polynomial kernel considers relationships in higher dimensions and can generate non-linear decision boundaries. We consider the parameters degree and "gamma" where degree represents the degree of the polynomial to be used. We set the degree as 3 and "gamma" as 0.01. Again, using a very high degree or a very low gamma will result in overfitting and give poor results. This yields an accuracy of 96.95%.
Corresponding confusion matrices for SVM with RBF and polynomial kernels can be found in [Figure 10] and [Figure 12] respectively.

SVM with rbf kernel outperformed other models because it effectively captured complex relationships in the data. RBF kernel created non-linear boundaries in the data, enabling it to recognize intricate patterns in the dataset more accurately than other models. Additionally, SVMs, in general, are less prone to overfitting with small and medium sized datasets making it suitable for our research.

And finally, we use Neural networks which are highly appreciated for their capability to recognize complex patterns in dataset. A sequential 3-layered CNN was employed for classification with "relu" activation at each layer and softmax activation for the final layer. Here the input images are converted to one-hot encoding first and then fed to the network. We set the number of epochs to 10 and we obtained an accuracy of 96.58%. Corresponding confusion matrix can be found in [Figure 13].

VII. RESULTS AND ITS ANALYSIS

- Linear Model:
  Total number of tested samples: 1311
  Number of correctly classified samples: 1029
  Accuracy is 79.25% with a linear model which is not reliable. Hence, we switch to non- linear models to try and get better accuracy.
  [As shown in Figure 6]

- KNN Classifier:

Total number of tested samples: 1311
Number of correctly classified samples: 1146
Accuracy is 87.41% with KNN Classifier
Such accuracy suggest that the data of a particular class is not as closely distributed which can be seen from the plot as well, hence we need to switch our models to those that can extract the features from the data i.e. the decision trees and random forests. [As shown in Figure 7]

- Decision Trees:
  Total number of tested samples: 1311
  Number of correctly classified samples: 1113
  Accuracy is 84.90%
  We use random forests to grow multiple trees which will increase the accuracy [As shown in Figure 8]

- Random Forest:
  Total number of tested samples: 1311
  Number of correctly classified samples: 1206
  Accuracy is 91.99%
  [As shown in Figure 9]

  Classification Report [Figure 11]

	precision	recall	f1-score	support
notumor	0.98	1.00	0.99	405
glioma	0.94	0.80	0.86	300
meningioma	0.86	0.90	0.88	306
pituitary	0.91	0.99	0.95	300
accuracy			0.93	1311
macro avg.	0.92	0.92	0.92	1311
weighted avg.	0.93	0.93	0.92	1311

We can further try to enhance the accuracy of model by considering higher dimension relationships between features like using a polynomial or a radial basis function kernel with a support vector machine.

- SVM with RBF kernel:
  Total number of tested samples: 1311
  Number of correctly classified samples: 1276
  Accuracy is 97.33% when we use SVM with RBF kernel.
  [As shown in Figure 10]

  Classification Report [Figure 14]

	precision	recall	f1-score	support
notumor	0.98	1.00	0.99	405
glioma	0.97	0.95	0.96	300
meningioma	0.95	0.94	0.94	306
pituitary	0.99	1.00	1.00	300
accuracy			0.97	1311
macro avg.	0.97	0.97	0.97	1311
weighted avg.	0.97	0.97	0.97	1311

- SVM with polynomial Kernel:
  Using SVM with a non – linear kernel gives an accuracy of 97.02% which is a little less than SVM with RBF kernel.

  Classification Report [Figure 15]

	precision	recall	f1-score	support
notumor	0.99	1.00	0.99	405
glioma	0.94	0.96	0.95	300
meningioma	0.96	0.92	0.94	306
pituitary	0.98	1.00	0.99	300
accuracy			0.97	1311
macro avg.	0.97	0.97	0.97	1311
weighted avg.	0.97	0.97	0.97	1311

- Convolutional Neural Network (CNN)
  Total number of tested samples: 1311
  Number of correctly classified samples: 1266
  The CNN model gave an accuracy of 96.95% on the test set. The number of epochs was set to 10 and accuracy after each epoch was calculated. [Figure 13]

  Classification Report [Figure 16]

	precision	recall	f1-score	support
notumor	0.98	1.00	0.99	405

glioma	0.94	0.94	0.95	300
meningioma	0.94	0.91	0.93	306
pituitary	0.97	1.00	0.99	300
accuracy			0.97	1311
macro avg.	0.96	0.96	0.96	1311
weighted avg.	0.97	0.97	0.97	1311

CONCLUSION FROM RESULTS

From the results obtained, it is evident that the Support Vector Machine (SVM) classifier, particularly with a radial basis function (RBF) kernel, outperforms other models such as KNN, decision trees, and random forests. These findings suggest that SVM classifiers are well-suited for brain tumor classification, particularly with non-linear kernels like RBF and polynomial.CNNs are widely recognized for their outstanding performance in image classification tasks. In this study, it can be inferred that SVM classifiers and CNNs are the best models for brain tumor classification, as they consistently demonstrate superior accuracy compared to other traditional machine learning algorithms. Their ability to capture complex relationships in the data and handle high-dimensional feature spaces makes them superior to other models tested.

CONTRIBUTIONS

We both have equal contribution in the whole project.

Naman Birla : Provided relevant resources, final model is made by him and he has trained the final model.Helped in making the report and presentation.

Mohmmad Ayaan – Provided relevant resources solutions to the problem. Helped in making the report and presentation.

REFERENCES

[1] International Journal of Engineering Research and Technology. ISSN 0974-3154, Volume 12, Number 4 (2019), pp. 563-569 © International Research Publication House.

[2] An introduction to image classification from Deigned Models to End-to-End Learning by Klaus D. Toennies

[3] https://www.researchgate.net/publication/325116934_Image_classification_using_Deep_learning

[4] Enhanced Image Classification With a Fast-Learning Shallow Convolutional Neural Network Mark D. McDonnell and Tony Vladusich.

[5] Image denoising using deep CNN with batch renormalization

[6] https://doc-pak.undip.ac.id/id/eprint/12130/1/Paper_1570766875.pdf

[7] https://www.kaggle.com/datasets/masoudnickparvar/brain-tumor-mri-dataset