Brain Tumor Classification Using Singular Value Decomposition (SVD)

Padmanabh Butala (pb8176@rit.edu)
Master of Science in Artificial Intelligence
Rochester Institute of Technology

Abstract

This research explores the use of Singular Value Decomposition (SVD) as a dimensionality reduction method in MRI brain tumor classification, coupled with multiple machine learning classifiers within a soft voting ensemble framework. The study highlights the effectiveness of SVD in extracting essential features from complex imaging data, thereby improving the accuracy of the classifiers. The findings demonstrate the technical advantage of this approach, positioning it as a computationally efficient alternative to traditional supervised learning methods in medical image analysis.

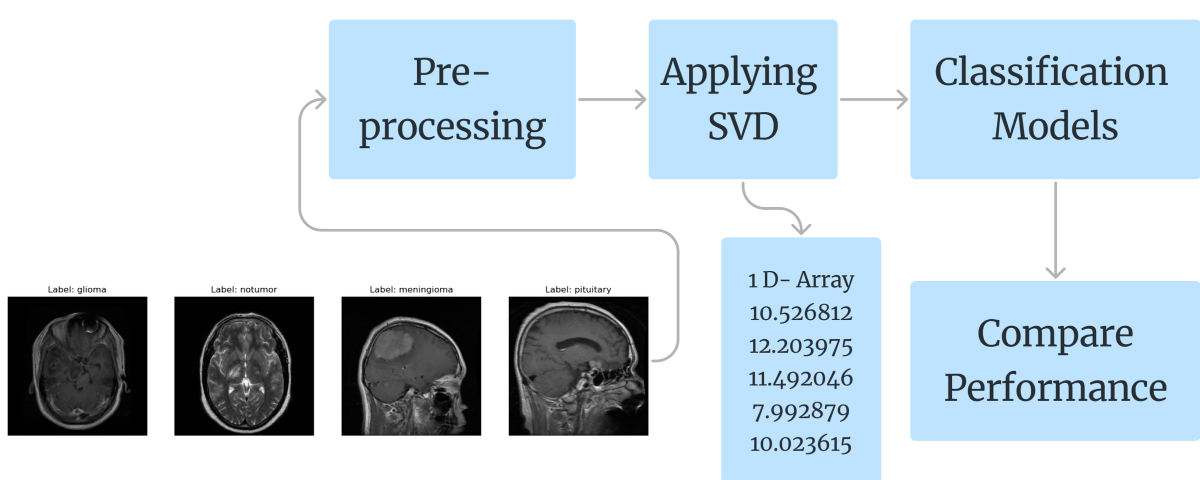


Figure 1: SVD Workflow

Introduction

Motivated by an academic exploration of Singular Value Decomposition (SVD) in Math 620, the project aims to evaluate the efficacy of SVD as a feature extraction technique in enhancing the accuracy of image classification models.

RQ 1] Can using SVD with different classification models help better identify brain tumors from MRI images?

RQ 2] How well does this SVD method work compared to the usual image analysis methods like CNN?

Method

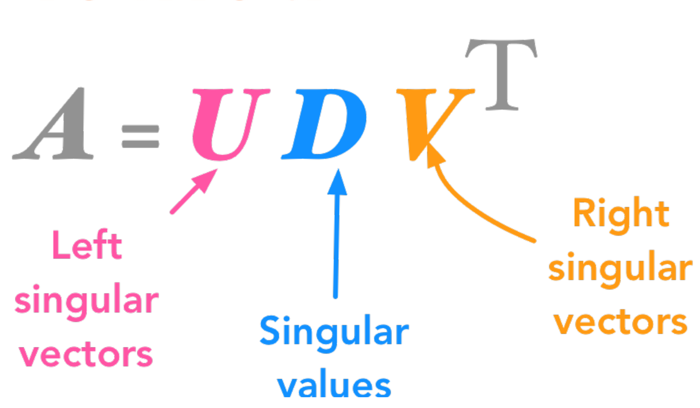


Figure 2: SVD Formula

A is the original matrix that we want to decompose.
U is an m×m orthogonal matrix whose columns are the left singular vectors of A.
V^T is the transpose of V, where V is an n×n orthogonal matrix whose columns are the right singular vectors of A

Σ (Sigma) is an m×n diagonal matrix with non-negative real numbers on the diagonal. These numbers are known as singular values and are typically arranged in descending order. The singular values are the square roots of the eigenvalues of A^T A or AA^T.
V and U are found such that they diagonalize A^T A and AA^T respectively.

In the context of image processing or machine learning, A might represent a set of flattened images stacked as rows in a matrix. The SVD is used to identify the principal components (singular vectors) that capture the most significant features of the data. The singular values indicate the importance or weight of each feature. Reducing dimensionality involves selecting a subset of the most significant singular values and their corresponding singular vectors.

Acknowledgement

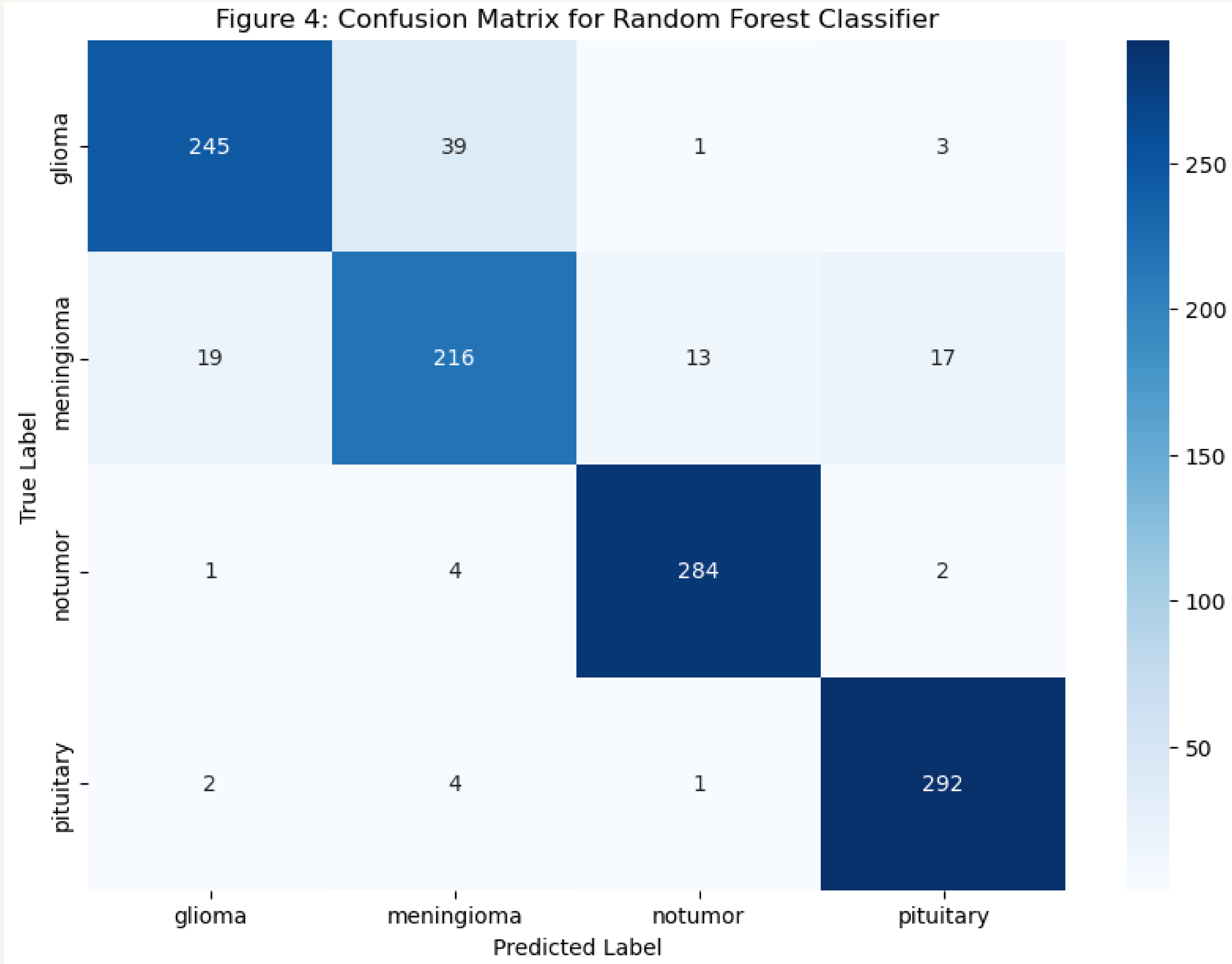
The dataset used for this project is from kaggle.
Brain Tumor MRI Dataset (kaggle.com)

Results

Classifier	Accuracy	F1 Score
Logistic Regression	0.79	0.78
Gradient Boosting	0.86	0.86
Random Forest	0.91	0.91
SVM	0.81	0.81
XGBoost	0.91	0.9
Voting Classifier	0.893	0.892

Figure 3: Classification Model Performance Comparison Table

Figure 4: Confusion Matrix for Random Forest Classifier



Classifier	Accuracy
SVD + Random Forest	0.91
CNN	0.9618

Figure 4: Comparison Between my method vs preferred method

Conclusion

In conclusion, while SVD combined with classification models like Random Forest shows promise for identifying brain tumors from MRI images, it currently does not outperform CNNs. CNNs remain the preferred method due to their higher accuracy in image analysis. Future work could explore Mutual Information accelrated SVD (MI-SVD) to potentially achieve better results than CNNs.