Brain tumor classification using machine learning models

Hager Sherif
Dept. of Systems & biomedical engineering of Cairo University
Cairo, Egypt
hager.sherif2001@hotmail.com

Bassant Medhat
Dept. of Systems & biomedical engineering of Cairo University
Cairo, Egypt
bassantmedhat49@gmail.com

Mariam Ahmed
Dept. of Systems & biomedical engineering of Cairo University
Cairo, Egypt
mariamturkey2010@gmail.com

Yousef Adham
Dept. of Systems & biomedical engineering of Cairo University
Cairo, Egypt
yousefdhm@gmail.com

***Abstract*—Early detection of brain tumors is crucial for successful treatment. Manual analysis of MRI images is time-consuming and error-prone. This paper compares models for classifying glioma, meningioma, pituitary tumors, and healthy brains. Preprocessed MRI images underwent feature extraction and dimensionality reduction. Various machine learning models were trained and evaluated: K-Nearest Neighbors, Random Forest, Support Vector Machines, SGD,and Ensemble. KNN achieved 93% accuracy, showing potential to assist radiologists in early and accurate brain tumor classification.**

Keywords—Brain tumors, Machine learning

I. Introduction

Brain tumors are a severe disease that can impair brain function and lead to death if left untreated. The number of patients dying from brain tumors has risen dramatically in recent decades, posing major challenges for healthcare providers. Early detection and timely treatment are critical for improving patient outcomes.

The development of machine learning models to classify different brain tumor types from MRI data has the potential to assist radiologists and improve diagnosis. Accurately identifying tumor types can guide treatment decisions as some tumors require more aggressive therapies than others.

This paper aims to develop and evaluate machine learning models for classifying MRI scans into four brain tumor classes: glioma, meningioma, pituitary tumor, and healthy brains. The ability to accurately classify different tumor types can help physicians determine the appropriate treatment options and monitor patients' responses. Several machine learning models are compared, including random forest (RF), K-nearest neighbors (KNN), support vector machines (SVM), SGD,and Ensemble. The MRI images were first preprocessed by intensity normalization, smoothing, and resizing to reduce noise and variation. Dimensionality reduction is then performed using principal component analysis to reduce the feature, and then fed as input to the models and the output is the tumor class.

The models were trained on a subset of the data and evaluated on the remaining data. The results show that the accuracy of the KNN model was better than RF, SVM, SGD,and Ensemble 93, 89, 91,56 and 89 %, respectively.

II. Literature Review

A. 1st paper (Feature extraction from MRI ADC images for brain tumor classification using machine learning techniques)

[1] They used 1599 labeled MRI brain ADC (Apparent Diffusion Coefficient) image slices, 995 malignant and 604 benign, from 195 patients who were radiologically diagnosed as brain tumor patients to extract features.

The features extracted included mean pixel values, skewness, kurtosis, features of Grey Level Co-occurrence Matrix (GLCM), mean, variance, energy, entropy, contrast, homogeneity, correlation, prominence, and shade from MRI ADC images of each patient.

The validation of the extracted features was measured using ANOVA f-test.

The extracted features were used as input to several machine learning models, which are listed in the following table.

TABLE I. Results f the cross validation experiment

The table visualize the performance of each machine learning algorithm received at the cross-validation experiment over the training data set and the standard deviations for each result

Algorithm	Mean accuracy	Accuracy as percentage (%)	Standard deviation (SD)
Logistic regression	0.753378	75.33	0.034451
Linear discriminant analysis	0.748898	74.89	0.036810
k-Nearest neighbors classifier	0.828459	82.84	0.030710
Decision tree classifier	0.800764	80.07	0.045553
GaussianNB	0.748922	74.89	0.052582
SVC	0.815082	81.50	0.043396
Random forest classifier	0.843629	84.36	0.042054

The Random Forest classifier was chosen as the final model due to its highest accuracy, which was measured at 90.41% after the hyperparameter tuning process. The final model was able to predict malignant and benign neoplasms with 90.41% accuracy.

TABLE II. Classification report: performance of Random Forest after hyperparameter optimization

Tumor type	Precision (%)	Recall (%)	F1-score (%)	Support
Malignant	91	93	92	299
Benign	88	85	86	181
Accuracy			90	480
Macro average	87	85	86	480
Weighted average	87	87	87	480

B. *2nd paper (MRI-based brain tumor detection using convolutional deep learning methods and chosen machine learning techniques)*

[2] In this study They used a dataset containing 3264 MRI brain images, similar to the dataset we used in building our model. They proposed two deep learning methods and several machine learning approaches for diagnosing three types of brain tumor - glioma, meningioma, and pituitary gland tumors - as well as healthy brains without tumors.

To prepare the data for analysis, the authors applied preprocessing and augmentation algorithms to the MRI brain images. They then developed a new 2D Convolutional Neural Network (CNN) and applied six machine learning techniques to classify brain tumors. The 2D CNN achieved exemplary performance and optimal execution time without latency, making it a promising approach for diagnosing brain tumors.

They also used feature extraction based on shape and intensity features, and applied PCA to improve performance. These techniques were used to classify brain tumors and healthy brains without tumors, achieving comparable results to the first study.

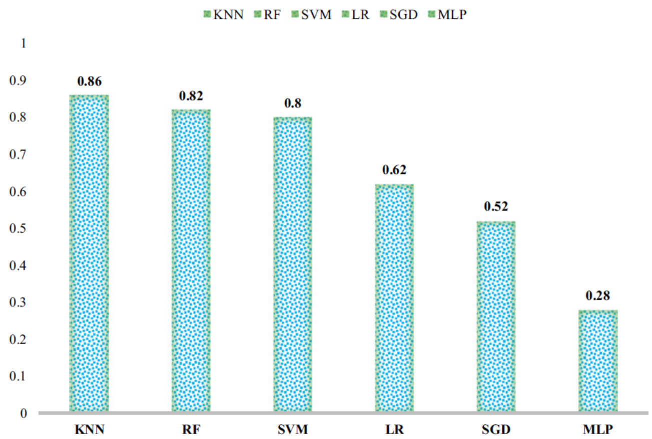

Fig. 1. Comparison of classification accuracy rates of machine learning classifiers

C. *3rd paper (Computer-assisted brain tumor type discrimination using magnetic resonance imaging features)*

[3]In this study, a comprehensive review of recent research on brain tumor multiclass classification using MRI is provided (from sources like conferences, journals, research groups and other published material).

The paper discusses the pipeline of brain tumor classification then discusses findings and results.

- Brain tumor imaging and dataset acquisition: There are different medical imaging modalities including (SPECT), (CT), (MRI).
- Preprocessing: Preprocessing in medical image analysis is an optional step. These steps include skull stripping, noise removal ,etc.
- Segmentation: Brain tumor classification schemes using MRI images may or may not perform segmentation. We discuss Non-segmentation based multi-classification techniques.
- Feature extraction: Basic quantitative features may include average, variance, correlation, contrast, entropy and inertia. In addition to spatial domain features like Gray Level Co-occurrence Matrix (GLCM), Laplacian of Gaussian (LoG), Color co-occurrence histograms (CCH),etc.
- Feature reduction: Popular feature reduction methods include Principal Component Analysis, Linear Discriminant Analysis.

TABLE III. List of some of the Non-segmentation based multi-classification work review

Study	Findings					
	Imaging	***Pre Processing***	***Dataset***	***Feature set***	***Classification***	***Evaluation***
Zulpe et al. [4]	T2 weighted PD	Noise suppression equalization outlier elimination	Patients=4 Images-80 AS=20 MEN=20	GLCM=444 textural features=16	FFNN	Accuracy=97.5%
Rajini et al [5]	T1 weighted	Noise removal	Images=110	GLCM based features=6	Decision tree	Average sensitivity=96 Specificity=100 Accuracy=99

a.

III. Dataset & features

The dataset used throughout the project is publicly provided by a Kaggle user: “MASOUD NICKPARVAR”. It consists of, 7023 MRI images with a resolution of 512*512 pixels, and split into 4 classes: glioma - meningioma - no tumor and pituitary.

The dataset contains a folder for testing, and a folder for training. We split the training folder into training and validation sets, using an 80%, & 20% split respectively. So in total we used 4570 training images, 1142 validation images, and 1311 testing images.

Pre-processing the data was a crucial part of the project, as some of the data did have some noise. Firstly, we normalized the images, then applied a Gaussian blur filter to smooth the images.

I. Examples:

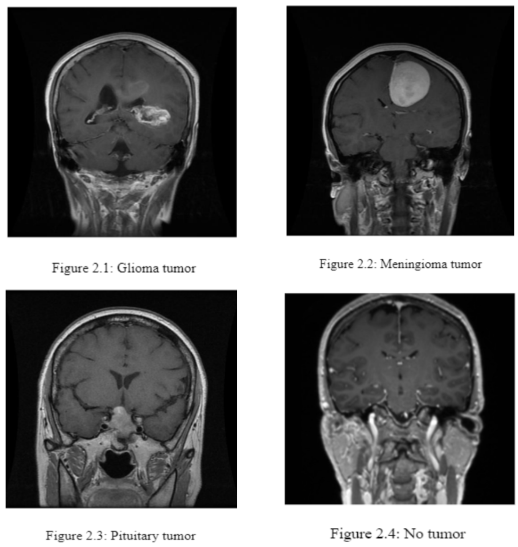

Fig. 2. Dataset samples

IV. Methods

After many careful considerations, we decided to train multiple machine learning models to compare their performances. But firstly, we would need to preprocess the images, by normalization and denoising, then extract relevant features and finally train the model.

But unfortunately training the models on the features extracted from the dataset yielded worse results, than training the models on the images directly after pre-processing, so we decided to skip feature extraction. Instead, we used PCA for dimensionality reduction of the (256*256) images. Only the first 100 principal components were used.

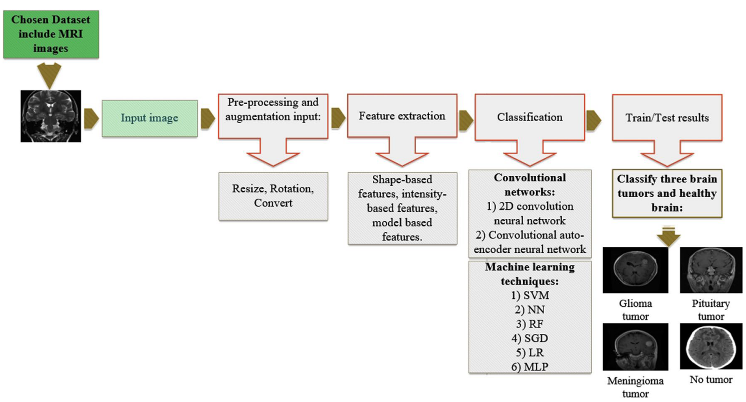

Fig. 3. Algorithm methodology

Algorithms used:

SVM (Support Vector Machine): A supervised learning model that aims to find an optimal hyperplane to separate data points in a high-dimensional space using a margin-maximizing approach, given by: maximize 2/||w|| subject to y_i(w^T x_i + b) >= 1 for all i, where w is the weight vector, b is the bias term, x_i is a data point, and y_i is its corresponding class label.

KNN (k-Nearest Neighbors): A non-parametric classification algorithm that assigns a class label to a data point based on the majority class of its k nearest neighbors in the feature space, where the class assignment is determined by: y = mode({y_1, y_2, ..., y_k}), where y is the predicted class label, y_i is the class label of the i-th nearest neighbor, and mode() returns the most frequent class label.

Random Forest: An ensemble learning method that combines multiple decision trees to make predictions by averaging or voting on the outputs of individual trees, employing bagging and random feature selection techniques. The prediction in a random forest is determined by: y = mode({f_1(x), f_2(x), ..., f_n(x)}), where y is the predicted class label, f_i(x) is the class assigned by the i-th decision tree, and n is the total number of trees in the forest.

We applied the Stochastic Gradient Descent (SGD) algorithm to optimize our linear model with logistic loss and L2 regularization. SGD updates model parameters incrementally by computing the gradient on mini batches of the training data, making it particularly useful for large datasets where it can converge faster than batch gradient descent. We set the learning rate to 0.001 and the maximum number of iterations to 1000 to ensure convergence.

We have also applied ensemble learning, which is a machine learning technique that combines multiple models to improve predictive performance. Furthermore, we have used logistic regression, random forest and support vector in the ensemble learning.

V. Experiments/Results/Discussion

To achieve accurate results for the models, we need to fine tune each model's hyperparameters. Therefore, we have decided to use the grid search algorithm, to determine the appropriate hyperparameters for each model.

Hyperparameters:

- SVM: linear kernel and C= 1.
- Random forest: 100 n estimators.
- KNN: 3 nearest neighbors.

Primary metrics:

The primary metrics used for evaluating the models performance are: accuracy, precision and recall.

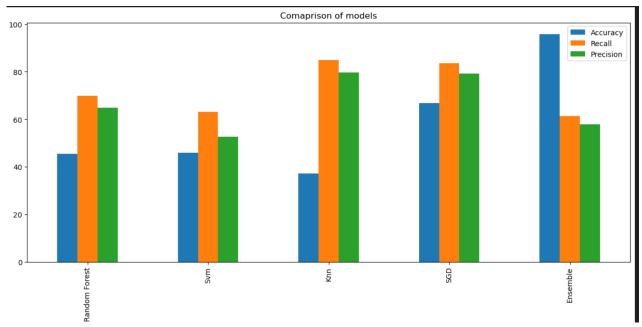

Fig 4. Models performance

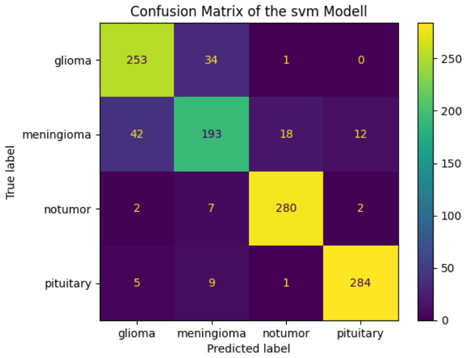

Fig. 4. SVM model's confusion matrix

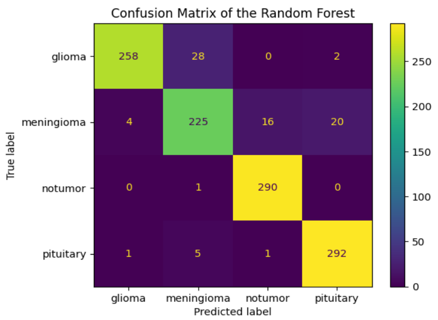

Fig. 5. Random forest model's confusion matrix

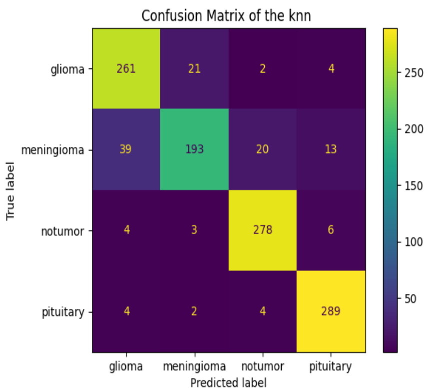

Fig. 6. KNN model's confusion matrix

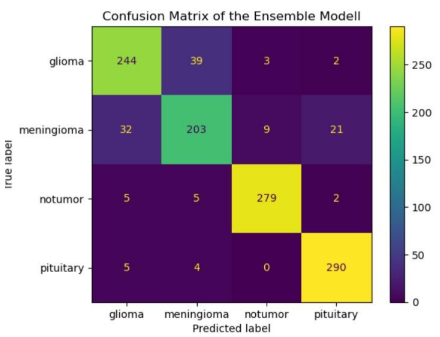

Fig. 7. Ensemble model's confusion matrix

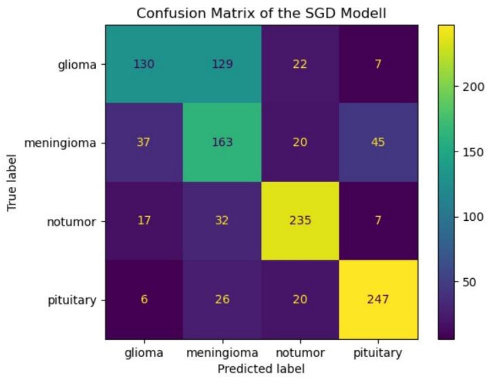

Fig. 8. SGD model's confusion matrix

VI. Conclusion/Future Work

Conclusion:

In conclusion, pre-processing the images using normalization and smoothing algorithms trains more accurate models. The random forest model is the most accurate out of the 3 models we've trained.
The Random Forest model outperformed SVM and KNN in classifying brain tumor MRI images due to its ability to capture complex relationships, handle noisy data, provide feature importance analysis, handle imbalanced data, and leverage ensemble learning for improved accuracy.

Future work:

Having access to a larger dataset, more computational resources, and more team members, would allow us to train models on larger training data, fine tune the models to a very fine degree, and try other computationally expensive models such as Hierarchical clustering. Overall, having access to the pre-mentioned resources would enable us to create more accurate and real-world applicable machine learning models.

References

[1] Sahan M. Vijithananda, Mohan L. Jayatilake, Badra Hewavithana, Teresa Gonçalves, Luis M. Rato, Bimali S. Weerakoon,Tharindu D. Kalupahana, Anil D. Silva &Karuna D. Dissanayake 21, Article number: 52 (2022).Feature extraction from MRI ADC images for brain tumor classification using machine learning techniques.

[2] Soheila Saeedi, Sorayya Rezayi, Hamidreza Keshavarz & Sharareh R. Niakan Kalhori BMC Medical Informatics and Decision Making 23, Article number: 16 (2023) . MRI-based brain tumor detection using convolutional deep learning methods and chosen machine learning techniques

[3] Iqbal S, Khan MUG, Saba T, Rehman A. Computer-assisted brain tumor type discrimination using magnetic resonance imaging features. Biomed Eng Lett. 2017 Oct 4;8(1):5-28. doi: 10.1007/s13534-017-0050-3. PMID: 30603187; PMCID: PMC6208555.[PubMed]

[4] Zulpe N, Pawar V. GLCM textural features for brain tumor classification. IJCSI International Journal of Computer Science Issues. 2012;9(3):354–359. [Google Scholar]

[5] Rajini NH, Narmatha T, Bhavani R. Automatic classification of MR brain tumor images using decision tree. In: Proceedings of international conference on electronics, vol. 31; 2012.

Used libraries:

Numpy, pandas, os, cv2, seaborn, matplotlib, sklearn, skimage

Contributions

Yousef Adham: half of the report's content, and report organization.

Bassant Medhat: literature Review, pre processing and building some models.

Hager Sherif : literature Review , feature extraction and grid search in models.

Mariam Ahmed: Abstract and introduction in paper, PCA,ROC curve and test data in models.