A Novel and Efficient Methodology Based on Blended Machine Learning Techniques for Brain MRI Classification

By

Jawad Haider

In partial fulfilment of the requirement for the degree

Bachelor of Science in Computer Science

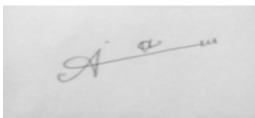
Department of Computer Science, School of Arts & Science

University of Central Asia, Naryn campus, Kyrgyz republic

(2022)

DECLARATION BY AUTHOR

I/we certify that this work has not been accepted in substance for any degree and is not concurrently being submitted for any degree other than that of Bachelor of Science in Computer Science being studied at the Department of Computer Science, School of Arts & Science, University of Central Asia, Kyrgyz Republic. I/we also declare that this work is the result of my/our own findings and investigations except where otherwise identified by references and that I/we have not plagiarized another's work.



Jawad Haider

DECLARATION BY SUPERVISOR

I, the undersigned hereby certify that I have read this project report and finally approve it with recommendation that this report may be submitted by the authors above to the final year project evaluation committee for final evaluation and presentation, in partial fulfillment of the requirements for the degree of Bachelor of Science in Computer Science at the Department of Computer Science, School of Arts & Sciences, University of Central Asia, Kyrgyz Republic.

Dr. Muhammad Fayaz

ABSTRACT

Nowadays, brain abnormalities have been proven as a life-threatening disease; treatment of which causes tremendous physical, emotional, and financial strains on families, communities, and health systems. The early brain tumor detection serves as the first step towards the cure of brain cancer. Several methods in literature for this purpose are either computationally too complex and time consuming or both. Our proposed work provides a better solution to this problem along with high accuracy and consistency for prediction of the Brain MRIs. Our proposed work consists of three-stage methodology namely – preprocessing, feature extraction and classification. In preprocessing filters and transforms were implemented to get quality data/image. Filters like median filter, prewitt and Laplacian filters were used, and log transform along with contrast stretching transforms were implemented on the grayscale MRI images to enhance the image quality and edges of the MRI. At the end of preprocessing stage, the obtained noise reduced, and quality enhanced images were converted to RGB images. Then during the feature extraction stage, the statistical entities were calculated for red, green, and blue channels. Mean, contrast, correlation, energy, entropy, homogeneity, kurtosis, skewness, and variance were the features considered and extracted for each channel of the RGB images and saved in separate files for each channel. This stage provided the final dataset in form of three files for red, green, and blue channels to be used in classification stage. Different machine learning models were used in classification stage, namely artificial neural networks, decision tree, naïve bayes, k-NN, random forests and support vector machines. Using the predictions of the model for each channel, the results were considered after majority voting. For each model the performance was evaluated using the metrics of accuracy, recall, precision, and f1-score. To better understanding the performance of the model for each channel, we also calculated the accuracy, recall, precision, and f-score for both training dataset and testing dataset. A great amount of experimentation was done on tuning the hyperparameters of the models, selecting those parameters which yielded the best results in terms of accuracy. Lastly, after the performance evaluation of each model on each channel and the comparison of the six models helped us to conclude the best model as Decision Tree having an accuracy of 97.04%. The results of decision tree indicated not only that our proposed model performed better using decision tree on our dataset, but the comparative analysis also proved that our final model performed better than the well-known state-of-art proposed models in literature for brain MRI classification.

Keywords: Image processing, machine learning, statistical features, Brain MRI classification.

ACKNOWLEDGEMENTS

Throughout the final year project, I have received huge support and assistance. I would first like to thank my supervisor, Dr. Muhammad Fayaz, whose expertise were invaluable from the step of formulating the research topic to the end of this great work. Your insightful and experienced feedback and support pushed me to sharpen my understanding as well as the perfect culmination of this research.

Other than that, I would also want to extend my utmost gratitude to all the Computer Science Department staff especially our department head Dr. Ayman Aljarbouh for helping us in every possible way to complete our work without any administrative hurdles.

Lastly, I would also like to take this opportunity to express my gratitude and love for my parents for their unwavering support during all the challenges I faced relevant to the given thesis.

CONTENTS

Table of Contents

List of Tables:

List of Figures:

CHAPTER 1: INTRODUCTION

Human brain is considered as the most complex organ in the body. Billions of neural and non-neural cells form the nerve tissue which makes the brain. Brain processes sensory information and directs the motor response. It controls and coordinates all the body functions i.e homeostasis, learning, memory, thought formation, and other cognitive activities. The networks of neurons continuously grow, build new synapsis, and die, but the uncontrolled growth of these cells leads to the formation of brain tumors and infection of adjoining organs. Depending on the location and cell type, the progression of the tumor may be rapid or slow.(Rogers, 2022). The brain composed of three parts a) Grey-matter – responsible for nervous signal processing, consists of neurons nucleus and dentaries. b) White matter – composed of fiber-like cell structures; the axons, and c) Cerebrospinal fluids – a colorless fluid around the brain tissue, protecting and secreting different hormones. Tumors affect these regions, causing abnormal body functions and destroying the healthy cells. (Ullah, Lee, Fayaz, & Sciences, 2019). Tumors are classified into benign and malignant. Benign tumors are non-progressive and non-cancerous, hence consider to be less harmful as they originate and grow slowly in the brain, not invading other body parts. While malignant tumors are much aggressive, grow much faster and spreading into other body parts; mostly leading to death. Primary malignant tumors originate with in the brain itself and secondary malignant tumors originate in other body parts (breast, lungs and skin) and spread into the brain. (Nicki, Brian, & Stuart, 2010; Reihl et al., 2021). Brain tumors are the leading cause of cancer-related death in the world. According to World Health Organization (WHO), cancer is the leading cause of death globally, nearly 10 million deaths in 2020 were attributed to cancer. WHO also concluded that many cancers can be cured if detected early and treated effectively (WHO, 2022). Hence, an important step towards the treatment of brain cancer is its early detection. Therefore, the aim of this paper is to help in providing accurate, fast, and robust machine learning model so that the brain tumors will be detected early, and proper screening and treatment will be started, saving lives, and reducing the number of deaths around the world. Additionally, it will also reduce the cancer burdens of tremendous physical, emotional, and financial strains on individuals, families, communities, and health systems.

Medical image analysis plays significant role in clinical studies and research (Othman, Abdullah, & Kamal, 2011). These techniques enable doctors and radiologist to process and examine the disease with more accurate information and then prescribe the diagnosis. Medical image analysis helps in understanding and visualizing the internal body structure, finding the abnormalities, and further study the disorder. Hence, the principal factor in the diagnosis of brain tumor is the medical images obtained from various imaging sources. Several medical imaging modalities are developed keeping in mind different factors and limitations either of the internal body structure or the mechanism of the equipment. Mammograms, Computed Tomography scans (CT-scans), X-rays, and Magnetic Resonance Imaging (MRI). Each process can be employed with its own benefits, limitations and side effects. (Othman et al., 2011). For the diagnosis of brain tumors MRI and CT scans are preferred by researchers over other image modalities. However, the capability of soft tissue discrimination, its non-invasiveness and high degree of spatial resolution makes the MRI option advantageous over CT scans (Shenton et al., 1992). Moreover, MRI is more safe choice as it does not use harmful radiations, it is painless and have insignificant known risks (da Silva, 2007; Kazmi, Qureshi, & Rashid, 2007; Mishra, 2010).

Classification of brain MRI images into normal and abnormal categories is the first steps towards the early detection of brain tumors. However, the large amount of data from MRI makes their manual classification tedious, error-prone, time-consuming and resource dependent as it needs an expert for the classification. Experts face great difficulty in assessing and interpreting the images to detect the tumor among these large amount of images (Ullah et al., 2019). This makes it necessary to develop and implement an automatic image analyzing system. The system should be easy to use, fast and accurate to categorize the images. Much research has been done in this regard and in literature one can find wide range of such medical diagnostic techniques obtained by applying various complex image processing methods coupled with different machine learning algorithms. These methods have the drawback of either being computationally too complex and time consuming or both.

For brain MRI classification, two types of machine learning systems are implemented – supervised and unsupervised. In supervised systems the algorithms like Artificial Neural Network (ANN), k- Nearest Neighbor (k-NN) and Support Vector Machine (SVM) are used. On the other hand, for unsupervised systems the methods of Self-Organizing Map (SOM) and Fuzzy c-means were explored. Subsequent researches showed that the supervised classification methods gave more accurate results as compare to the unsupervised ones.(Saritha, Joseph, & Mathew, 2013).

Perhaps, the information from brain MRI can be analyzed and processed using both supervised and unsupervised algorithms and is grouped into normal and abnormal categories, but the accuracy depends on how we extract the features form the images and how much relevant they are to determine the group. Few of such well known methods include the Independent Component Analysis (ICA), Fourier and wavelet transformed-based techniques. (Chaplot, Patnaik, Jagannathan, & control, 2006; Mallat, 2009), methods of statistical feature extraction like entropy, skewness, kurtosis, quartiles, mode, mean, median, and standard deviation. (Begg, Palaniswami, & Owen, 2005). As it is important to go for the most meaningful and relevant features, but it increases the computational burden of the classifier. Consequently, under these constraints, the feasible choice is to select a method of feature extraction, which can provide the minimum number of relevant features as possible to have the complete characteristic anatomy of the tumor. This way, the extra computational complexities for unnecessary feature extraction can be reduced. One of such method is using image transformations. In this work we implemented two of such transforms- log transform and contrast stretching. Log transform helped to restore the pixel intensities that are otherwise missed due to wide range of intensity values or lost at the expense of high intensity values. Contrast stretching is implemented to further increase the pixel value range by rescaling the pixel values in the input image. We structure our methodology as: Image pre-processing, then Feature extraction and lastly Classification.

Prior to embarking on working with the data, it is essential to preprocess it. During preprocessing, various types of noises are removed. Medical images are corrupted by Rician noise and salt-and-pepper noise. Accurate analysis of the medical images is only effective from a high-quality image being preprocessed, having the required brightness, details of edges and minimum noise. For brain MRI, the median filter helps better than any other filter to remove the salt-and-pepper noise (Ullah et al., 2019). Preprocessing is followed by feature extraction where the input image/data is transformed into a set of useful features, which can be used for the classification stage. Feature extraction serve as dimensionality reduction which gives only the quantitative measure of the important features and reducing the amount of unnecessary data. This helps the classification algorithms to process the reduced input instead of extremely large data – which consumes more time along with huge computational resources. Furthermore, the large data contains redundant information, that the classifier does not require to give optimal output. While the transformation of image data into reduced number of features helps to a great extent, but great consideration is to be put into the set of prominent features, which can more likely help in obtaining the desired information for the good performance of the algorithms, instead of giving the full image data to it.

Important features to consider for feature extraction from brain MRI are the dominant edges, position, shape, and texture of the image.

However, along with the importance, feature extraction also brings challenges to deal with. Different methods exist for this purpose, for instance, Gabor feature, Discrete Wavelet transform, Spectral Mixture Analysis, Principal Component Analysis, Minimum Noise fraction Transform. Implementing dimensionality reduction, we can focus on few key features. Some well-known algorithms for this purpose are Linear Discriminant analysis, Genetic algorithm, and Independent Component Analysis (Ullah et al., 2019).

The next stage after feature extraction is classification. Here, the classifier uses the preprocessed images with selected features and categorize them into normal and abnormal classes. Classification has two parts, training, and testing. Training part is where the classifier is supervised by providing the inputs and their corresponding output labels; it trains on these data and build the model to predict/classify the unseen data. During the testing part, the model is tested on data it has not seen before. The majority voting strategy is used to get the final results if there are several classifiers used for one dataset. Then performance is evaluated. Loss and accuracy of the classifier will determine if the model should be selected or rejected. After several experiments on classifiers using Artificial Neural Network (ANN), Support Vector Machines (SVM), Decision Trees (DT), Random forests (RF), Naïve Bayes (NB) and k-nearest neighbors (K-NN), we selected decision tree for its best performance among these classifiers as well among other well-known state-of-the-art techniques used in literature. Our approach provides accuracy close to expert level with high consistency.

The objectives of this work were two folds:

- The first objective was to extract a smaller number of informative features from brain MRI images because we can’t feed images directly to machine learning algorithms. So, our first objective was to not only extract the features but features of high level of information.
- The second objective was classifying the brain MRI images with high accuracy by proposing a novel and efficient brain MRI classification methodology.

Hence, we proposed two novel methods as given below to obtain the above two objectives.

- To achieve the first objective, we proposed a novel method of features extraction, we converted the grayscale brain MRI images to RGB images and calculated the nine statical features of each channel of RGB images, i.e., nine features for red channel, nine for green channel and nine for blue channel and stored them with labels in three separate files.

- To obtain the second objective first we applied several machine learning algorithms separately on the features obtain from red, green, and blue channels of RGB images, and considered the best algorithm with the highest accuracy. Next, we proposed a new method named as majority voting, in this technique the classification results are obtain from each channel by using a high accuracy machine learning algorithm and then the high mode class is considered as output class.

Hence, the above are both novel methods of features extraction and classification respectively and these are main contributions in this work.

The structure of this report is organized as below: section 2 presents the literature review, in section 3 the proposed methodology is explained in detail, section 4 is about implementation detail, results and discussion, with insightful performance evaluation of each algorithm, and the report is concluded in section 5 with future recommendations.

The abbreviations with their corresponding descriptions are listed in Table 1.1

Table 1.1 Abbreviation and their description

Notation	Description
MRI	Magnetic Resonance Imaging
RGB	Red, green, and Blue
ANN	Artificial Neural Network
DT	Decision Tree
KNN	k-nearest neighbor
RF	Random Forest
NB	Naïve Bayes
SVM	Support Vector Machine
GF	Grayscale Features

CHAPTER 2: LITERATURE REVIEW

Medical Image analysis has been extensively explored and great amount of research has been put into the building of automatic tools for it in recent years. These methods allow researchers and doctors in studying the internal working of soft tissues without the invasive surgery. Two of the most used and preferred techniques for this purpose in both research labs as well as in hospitals is Magnetic Resonance Imaging and Computed Tomography scans.

MRI is considered more suitable because it provides high resolution of brain anatomy along with sharp contrast. MRI is also not harmful as no such radiations are used and MRI is non-invasive. It provides accurate information parameters of spin-spin relaxation times, spin-lattices, flow velocity, proton density and chemical shifts. These parameters help in precise characterization of the brain tissue (Georgiadis et al., 2008).

Image segmentation plays crucial role in brain MRI classification. An approach for brain MRI segmentation was proposed by Zanaty and Aljahdali in 2011, (Zanaty & Aljahdali, 2011) using automatic fuzzy algorithms. These algorithms used spatial constraints for their object function as well as for validation of the procedure for clustering. Employing the intra-cluster distance measures, the working of their method was tested on synthetic images having different noise levels. The proposed method worked better on segmentation of brain MRI. Similarly, for extracting/segmenting Brain MRI, Somasundaram et al. (Somasundaram, Kalaiselvi, & medicine, 2010) proposed two other algorithms, BEA for T2-weighted scans. Here, lowpass filter was used to reduce the noise, then the image was diffused to improve the boundaries of the brain. By obtaining the intensity threshold value using Ridle's method they produced a binary image. Then by applying the morphological operations on this image, they extracted the brain with highest connected component analysis. The methodology was implemented on 2D and 3D information of slices, and the performance was better than brain surface extractor (BSF) and brain extraction tool (BET).

Further improved techniques for image segmentation were introduce over these methods by Vasuda et al. (Vasuda, Satheesh, & Engineering, 1713). To obtain an accurate image segmentation of MRI images, they used Fuzzy c-Means (FCM) for fuzzy clustering. To deal with the drawback of FCM, the cluster center and membership value updating criterion was modified which effectively reduced the computational time for convergence rate. Moreover, a two-phase method for image segmentation was also proposed by (Logeswari, Karnan, & Engineering, 2010). Noise from the image was removed during the first phase. The second phase consist of vector quantization of the images coupled with Hierarchical Self Organizing Map to get higher value of tumor pixel and computational speed. In 2014, Joseph et al. introduced a two stage methodology – image segmentation and tumor detection (Joseph, Singh, Manikandan, & Technology, 2014) implementing K-means clustering with morphological filtering for image segmentation and tumor location detection in brain MRI respectively. It was a simple method to detect the abnormal brain images. Morphological filtering avoided the misclustering of regions after the segmentation.

Another study by Chap lot et al. in 2006 (Chaplot et al., 2006), used wavelets as inputs and compared the performance of self-organizing maps (SOM) and support vector machine. They concluded that SVM were better at classifying the brain MRI into normal and abnormal images at the accuracy rate of 98 % while the accuracy of SOM was 96%. They implemented 2D discrete wavelet transforms and for the decomposition they used Daubechies filters. However, Maitra et al. (Maitra, Chatterjee, & Control, 2006) introduced an improved version of orthogonal discrete wavelet transform for feature extraction – the Slantlet transform, which proved a better time localized space information for non-stationary MRI images. The feature vector from this feature extraction method gave an improved performance on the training of the backpropagation neural network based binary classifier. The classifier achieved 100% accuracy on Alzheimer brain MRIs.

In 2010, El-Dahshan et al. (El-Dahshan, Hosny, & Salem, 2010) proposed a hybrid methodology with three stages – feature extraction, dimensionality reduction and classification. For feature extraction discrete wavelet transform was used, for dimensionality reduction principal component analysis was used to effectively select the important features without putting burden on the model. Then, feed forward backpropagation artificial neural network (FA-NN) and k-NN were implemented for the classification. K-NN overperformed FA-NN with accuracy of 98%. Another three-staged classification method was introduced by Zhang et al. (Zhang, Dong, Wu, & Wang, 2011) with same steps as El-Dahshan et al. but using Scaled Conjugate Gradient (SCG) during backpropagation neural network to get the optimal weights. Their model performed with 100% accuracy on 66 images with the computational time of 0.0451s each image. Another similar work was done by Fayaz et al. in (Fayaz, Shah, Wahid, Shah, & Recognition, 2016). The stages were pre-processing where median filter was used to remove the salt-and-pepper noise from the MRIs and images were converted into RGB, feature extraction; where red, green and blue channels were extracted from the RGB images and for each channel statistical features were calculated. Then the k-NN was implemented to classify the images. An accuracy of 98% was obtained on training data, a 95% on testing data for normal category while 100 % accuracy on training and 90% on test for abnormal images were the results.

Different variations of this three-stage methodology were proposed by several other works. For example, Wahid et al. suggested a method much similar to Fayaz et al. in (Wahid, Ghazali, Fayaz, Shah, & Bio-Technology, 2016). The changes were introduced in feature extraction stage, more focus was put on color moments and textures of the MRIs. For classification they implemented a logistic function based probabilistic model. As a common criterion for the performance evaluation of this model, they compared it with the well know existing classifiers like SVM, Naïve Bayes, ANN and normal densities based linear classifier. To improve their classifier, they used 10-fold cross validation. The resulting model gave an accuracy of 90.66%. Another similar work was done by Ullah et al. 2016,(Ullah, Fayaz, Iqbal, & Science, 2016). They employed data mining techniques on medical data to improve the classification of various MRIs like heart ailments, breast cancer and lung cancer. A k-NN algorithm was implemented by Ullah et al. (Ullah, Lee, Khan, Fayaz, & Iqbal, 2018) and classification accuracy was much improved. Then, Suhaimi et al. worked on designing CNNs for (functional MRI) fMRI classification and concluded that using a feature map size selection on fMRI and MRI datasets. Furthermore, a study by Saleh et al.(Saleh & Al-Bakry, 2017) implemented agency system classification for the tumors. Salt-and-pepper noise was removed during the preprocessing then multilevel threshold segmentation was used to detect tumors; opening and closing morphological operations were also used. The tumor segmentation was done using watershed transform and using binary object feature method, the essential features were extracted. Then finally the classification was done with 100% accuracy. Their system classified brain tumor images into Metastases and Gliomas, meningioma benign, meningioma tumor and normal tissue.

Another multilevel-stage approach was proposed by Keerthana et al. 2018 (Keerthana & Xavier, 2018), their system classified the tumor images along with providing healthy advice and disease description for the user. They used data mining techniques. Pre-processing, segmentation, feature extraction and classification. SVM along with genetic algorithm for enhancing the features and SVM parameters were used to identify the type of brain tumor. Mathew et al. 2017 (Mathew & Anto, 2017) also employed a similar four stage method. Pre-processing was done using anisotropic diffusion filter, DWT was implemented for feature extraction and then SVM was applied on the final data for the classification. Korolev et al. 2017 (Korolev, Safiullin, Belyaev, & Dodonova,

2017) implemented the residual and plain 3D convolutional neural network architecture while skipping the feature extraction step to classify MRI images for Alzheimer.

Indeed, the list of literary works on the medical image analysis and neuroimaging data is inexhaustible. However, it is necessary to end this section with some of the famous and streamlined works done recently which had far reaching applications and revolutionized the area. For further reading one can refer to the works by (Akkus, Galimzianova, Hoogi, Rubin, & Erickson, 2017; Lundervold & Lundervold, 2019; Mostapha & Styner, 2019; Yasaka, Akai, Kunimatsu, Kiryu, & Abe, 2018; Zhu et al., 2015) for comprehensive reviews of the papers from 2017-2020. Several classical machine learning approaches as well as deep learning-based models are discussed providing the overview of each, performance, speed, and the properties of each method. The terms, definitions, and algorithms are explained extensively in much easy to understandable way along with focus on the research area of MRI brain images.

CHAPTER 3: PROPOSED METHODOLOGY

Our proposed work comprised of three stages namely the preprocessing, then feature extraction, and finally the classification. The abstract diagram of the methodology is given in Figure 3.1

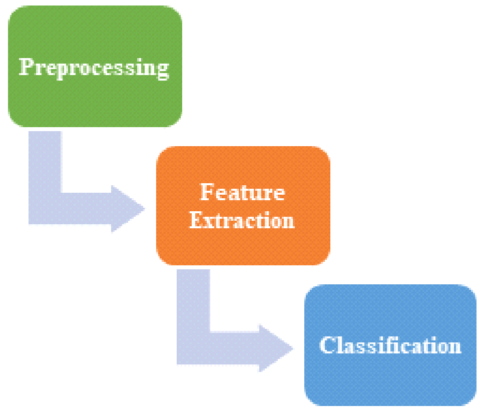


Figure 3.1 Abstract model of the proposed methodology

The preprocessing step essentially focuses on the noise removal, for the purpose of which different filters and transforms were applied. After filtering and transforming the greyscale images, the resulting smoothened and enhanced images were converted into RGB scale. Conversion to RGB images concludes the preprocessing stage. The next step is feature extraction. As RGB images contains the three channels – red, green and blue, during feature extraction stage, for each channel the statistical features were extracted. This provided a two-fold advantage. Firstly, it reduced the size of data to be processed by the algorithms in later stages and secondly, it also helped us to focus on the important features instead of feeding the algorithm with redundant data. Hence, the computational time as well as computational memory was reduced to minimum without effecting the necessary details. The statistical features extracted were contrast, correlation, energy, entropy, homogeneity, kurtosis, mean, skewness, and variance. The data for each channel was saved in separate files.

After feature extraction, the classification stage included experimentation with different classical machine learning algorithms to find which provides better performance. Different algorithms like Artificial Neural Networks, Decision Tree, k-NN, Naïve Bayes, Support Vector Machine, and Random Forest and were implemented to classify the images into normal and abnormal based on the statistical features. The algorithms were implemented for each channel separately and then the final predictions were found using majority votes. Lastly, the performance of each of the algorithm was evaluated on both testing and training data. The results gave a great insight into the algorithm to be selected as well as the usage of filters, transforms and features extracted. The detailed structure diagram in Figure 3.2 illustrates the overall steps performed during each step. The next modules will provide a detailed discussion on each stage and actions performed during each stage.

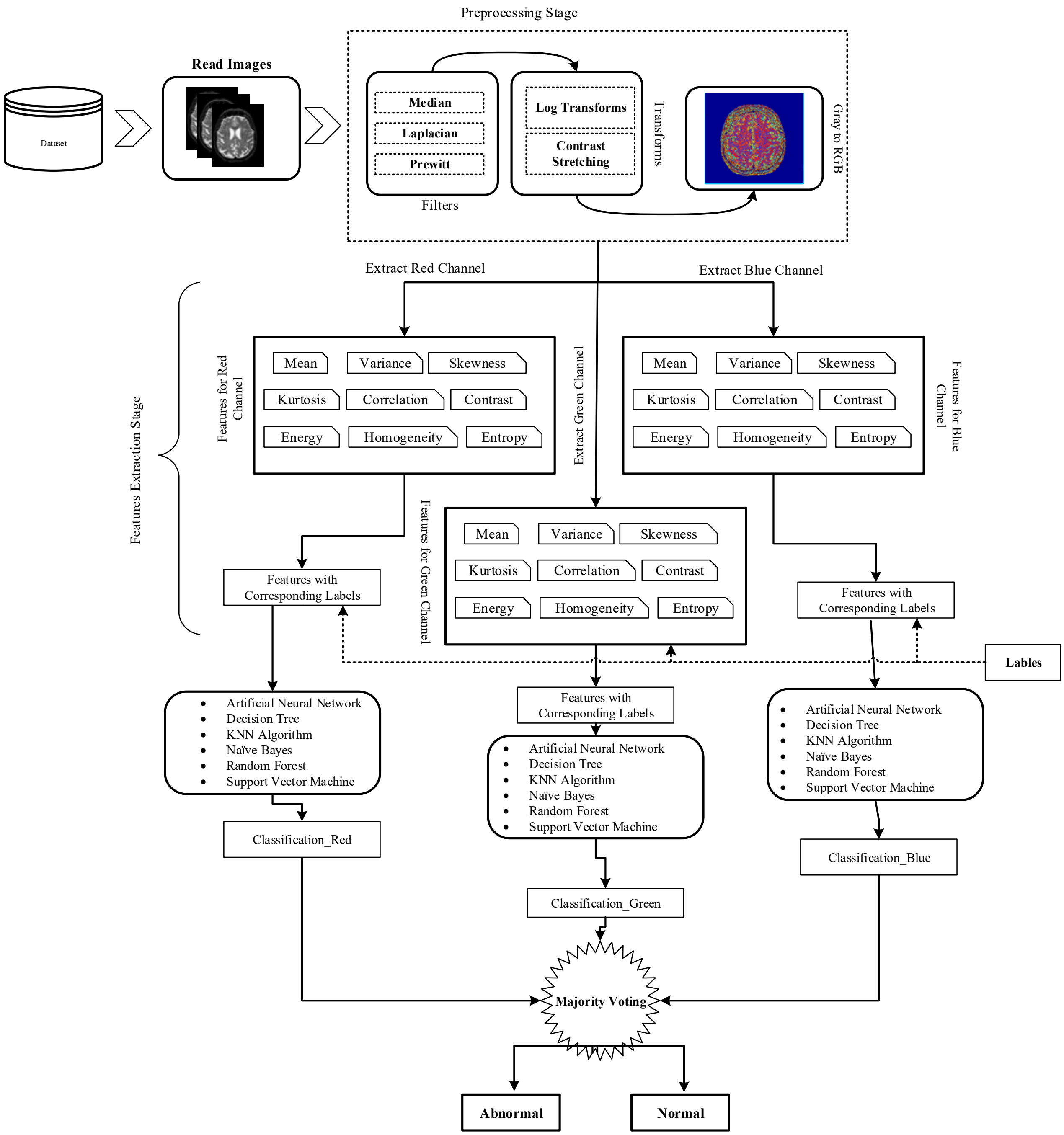


Figure 3.2 Detail Conceptual diagram of the proposed work.

3.1 PREPROCESSING

The first step toward building an accurate and robust model for classification is to preprocess the data before feeding it to the classifier. During medical image analysis the preprocessing stage is focused on removing the noise from the medical images. MRI are contaminated with different types of noises such as spackle noise, salt-and-pepper noise and Gaussian noise. For each noise removal we use specific type of filters like wiener filter, mean filter, Laplacian filter etc. Brain MRIs are particularly affected by salt-and-pepper noise which highly degrade the details in the image. To enhance and restore these images the required filters and transforms were implemented. A detailed diagram of this stage is illustrated in Figure 3.3 below.

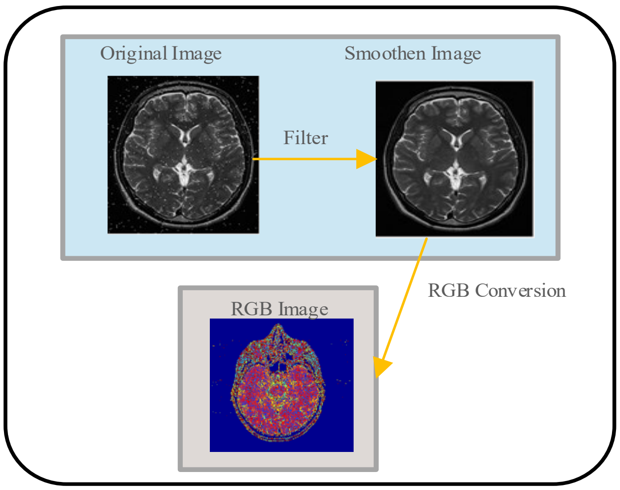


Figure 3.3 Detailed diagram for preprocessing stage

3.1.1 Filters:

Firstly, median filter is used to remove salt-and-pepper noise from the MRI. As it is the most common filter for such noise removal (Nazir, Wahid, Ali Khan, & Systems, 2015). Median filter is a non-linear filter that removes Impulse and Salt-and-pepper noise which is characterize by the black and white spots randomly distributed over the image. Figure 3.7 shows the original image and the corresponding median filtered image. Apart from removing the noise from the MRI, it is sufficiently important to also detect the edges. To do so we must implement both first derivative filter – Prewitt filter, and second derivative filter – Laplacian filter. As we transit from one boundary to the next, there is a great change of intensity, but it remains constant in the dark and white regions.

Figure 3.4 a) and b) represents the intensity profile of the region around the point of transition. First derivative has an increasing positive value while being zero in the dark and white regions. First derivative has maximum or peak at the edge. Since the first derivative is increasing before the edge, the value for second derivative is positive and as first derivative is decreasing after the edge, the corresponding value for second derivative is negative. At the edge the second derivative is zero a phenomenon known as zero-crossing. (Chityala & Pudipeddi, 2020).

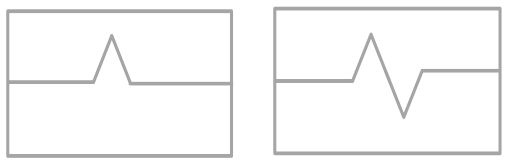


Figure 3.4 a) First derivative profile b) Second derivative profile

An acquired MRI image will have noise that may affect the detection of zero-crossing. If the change of intensity is too rapid in the profile, some spurious edges will be detected by the zero-crossing. To avoid issues due to noise or rapid intensity change, the image is pre-processed before application of second derivative filter.

A popular first derivative filter is Prewitt (Prewitt & Psychopictorics, 1970) The masks for the filter are as show in Figure 3.5 below.

-1	-1	-1
0	0	0
1	1	1

-1	0	1
-1	0	1
-1	0	1

Figure 3.5 Prewitt masks a) horizontal and b) vertical edges.

As the sum of the coefficients is 0, this filter does not affect pixels with constant grayscale. Hence, the filter does not reduce noise as can be seen in the values of the coefficient. See the Figure 3.7. After the first derivative, the second derivative is used to determine the edges. One of the popularly used second derivative filter is Laplacian. A Laplacian of continuous function is given by equation 1 below:

$$L(x,y) = \nabla^2 f(x,y)$$

$$\frac{\partial^2 f(x,y)}{\partial x^2} + \frac{\partial^2 f(x,y)}{\partial y^2} \qquad (1)$$

Most widely used masks for Laplacian filters are given below in Figure 3.6. Laplacian filter obtains the edges without introducing any noise. The zero introduction of noise is because of the sum of the values in the mask is equal to zero.

0	-1	0
-1	4	-1
0	-1	0

-1	-1	-1
-1	8	-1
1-	-1	-1

Figure 3.6 Laplacian Masks

The Figure 3.7 below gives the images for all the filters.

Original Image	Median Filtered	Prewitt Filtered	Laplacian Filtered

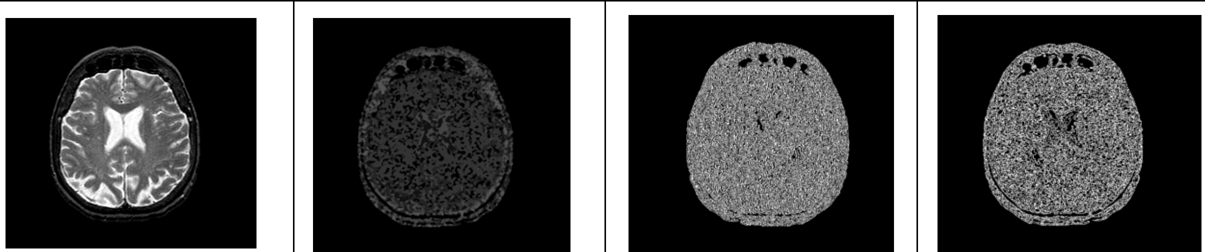


Figure 3.7 The original image with the corresponding filtered images

For more comprehensive details the reader can refer to (Marr & Hildreth, 1980; Martelli & processing, 1972; Nazir et al., 2015; Robinson, 1977).

3.1.2 Transforms:

Filters are implemented for the sake of enhancing the quality of images so that important details can be quantified and visualized. To further improve the quality of image we use transform. Transforming the pixel values to a new output value using a mapping function helps to improve images quality.

Log transform helps to restore the pixel intensities that are otherwise missed due to wide range of intensity values or lost at the expense of high intensity values. Log transform at (i, j) for intensities ranging from $[0, \mathrm{L}-1]$ is given by the equation 2 below:

$$t(i,j) = k\,log\big(1 + I(i,j)\big) \qquad (2)$$

Where k = $\frac{L-1}{\log(1+|I_{max}|}$ and I_{max} is maximum magnitude value and $I(i,j)$ is the intensity value of the pixel in the input image at (i,j). Log transform makes the image brighter. Figure 3.8 shows the log transformed image.

Contrast stretching is implemented to further increase the pixel value range by rescaling the pixel values in the input image. Contrast stretching is given by the equation 3 below:

$$t(i,j) = 255 * \frac{I(i,j) - a}{b - a} \tag{3}$$

Where $I(i,j)$ is intensity at (i,j) a and b are the minimum and maximum pixel values in the input images. The resulted image can be seen in Figure 3.8.

Hence, transforms will improve the feature extraction stage by providing accurate values for statistical profiles of the images.

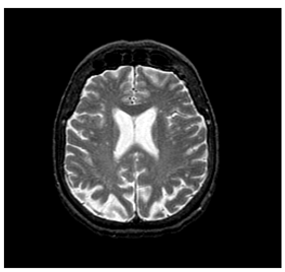
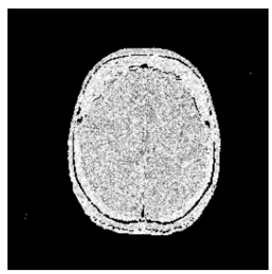
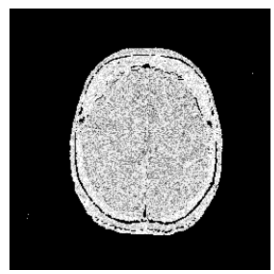
Original Image	**Log Transformed**	**Contrast Stretching**
		

Figure 3.8 Log transformed and Contrast Stretched Images

For further readings on image enhancement the reader can use (Hong, Wan, Jain, & intelligence, 1998; Osher & Rudin, 1990; Pal, King, & CYBER., 1981). Lastly, the smoothen grayscale images were converted into RGB for further processing in feature extraction stage.

3.2 FEATUER EXTRACTION

Feature extraction plays an important role in image processing and machine learning models. Before the data is given to the model for classification or recognition of images, the images are filtered and transformed (Kumar & Bhatia, 2014). Feature extraction refers to the process of transforming the raw data into numerical features that can be processed which preserve the information from the original dataset; followed by the preprocessing methods (MathWorks, 2022). Figure 3.9 provides the detailed structural diagram for feature extraction stage.

We obtained the statistical features for all the preprocessed images. Statistical moments like mean, variance, kurtosis, skewness as well as co-occurrence matrix features, namely energy, correlation, entropy, homogeneity, and contrast are extracted from each channel separately into their corresponding files. These statistical modes provided the best option for training our models during classification stage and are hence the most vital part of our proposed work, on which all the performance of our machine learning models depended. As Figure 3.9 shows, we took the RGB image output of the preprocessing stage and then found the above statistical features for each channel and saved them in their respective files.

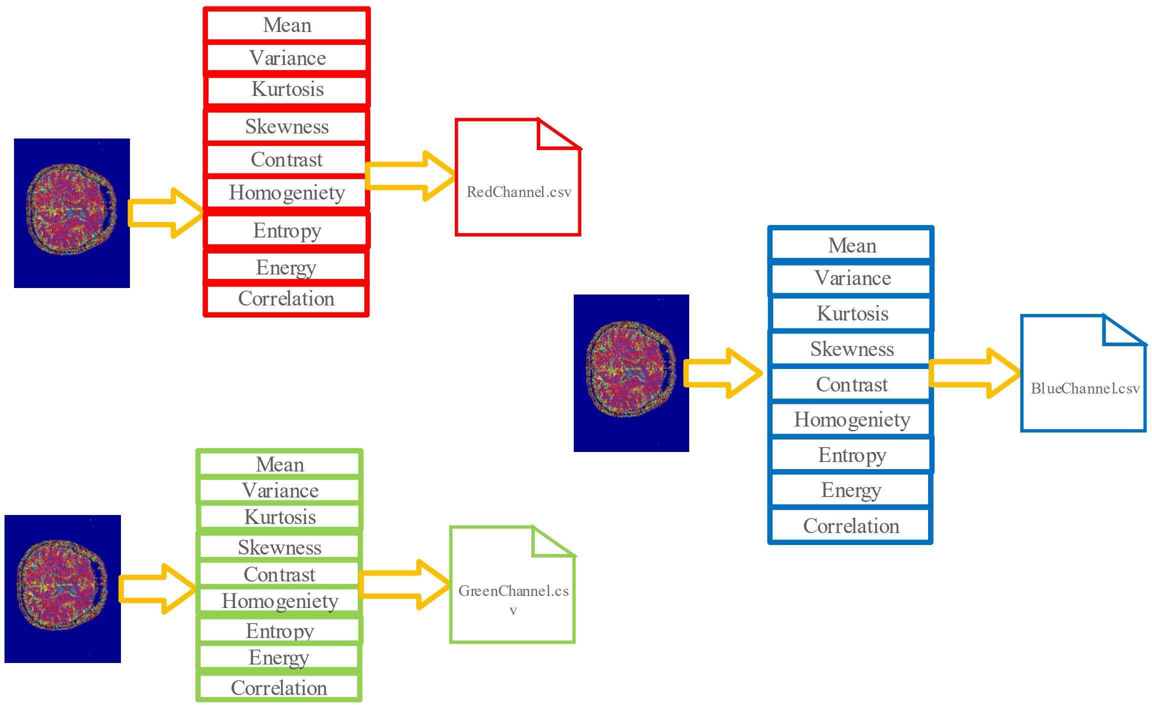


Figure 3.9 Structural Diagram for Feature Extraction Stage

The Equation (4-7) gives the mean, variance, skewness, and kurtosis of the image.

$$Mean = \frac{\sum_{i=1}^{n} x}{n} \tag{4}$$

$$variance = \frac{\sum_{i=1}^{n} (x_i - x)^2}{n} \tag{5}$$

$$Skewness\ (X) = \frac{\sum_{i=1}^{n} (x_i - \bar{x})^3}{n} \tag{6}$$

$$Kurtosis\ (X) = \frac{\sum_{i=1}^{n} (x_i - \bar{x})^4}{n} \tag{7}$$

Here, n denotes the total number of pixels in the image, $\bar{x}$ represents the mean value of the image pixel values, and x is the image pixel value.

Mean is the first-order statistical analysis calculation, and its value represents the intensity of the image from the entire pixels. The bright and dark values effect mean of the image.

Variance is the second-order and its values indicates the contrast of the image. (Sudiro, Kardian, Madenda, Hermanto, & Engineering, 2021). Skewness of the image is a measure of the symmetry, while kurtosis measures the peak and flatness of the pixel values in relation to their normal distribution. (Fayaz et al., 2021).

In equation (8-12) represents the mathematical forms of entropy, correlation, contrast, homogeneity, and energy respectively.

$$Entropy = -\sum_{x}\sum_{y}\lambda_{i,j}\, log_2\lambda_{i,j} \tag{8}$$

$$Correlation = \frac{1}{\sigma_\alpha\sigma_\beta}\sum_{x}\sum_{y}(x-\bar{x}_\alpha)((x-\bar{x}_\beta)\,\lambda_{i,j} \tag{9}$$

$$Contrast = \sum_{i}\sum_{j}(\lambda_{i,j})^2\ (i-j)^2 \tag{10}$$

$$Homogeneity = \sum_{i}\sum_{j}\frac{\lambda_{i,j}}{|i-j|} \tag{11}$$

$$Energy = \sum_{i}\sum_{j}(\lambda_{i,j})^2 \tag{12}$$

Entropy is the measure of the total content of the image. Correlation provides pixel similarity at pixel distance. Contrast gives the difference between the highest and lowest intensity values of the image. Homogeneity describes the similarity of the pixels in different location. And finally, energy is for the image is the localized change of the image, it is the rate of change in the color/brightness/magnitude of the pixels over local areas.

3.3 CLASSIFICATION

In the proposed study, we have implemented six state-of-art machine learning algorithms for classification. Decision Tree, Random Forests, k-NN, SVM, ANN and Naïve bayes. Decision tree performed better than other algorithms. In our work, we used three DTs, one each for red, green, and blue channels. After training the model on each channel, we implemented majority voting method to finalize the classification category of the input data. The next step after majority voting was to evaluate the performance evaluation of each algorithm's results. Figure 3.10 shows the detailed diagram for this stage, where we have implemented each of the six models for each channel and then applied majority voting after which the performance metrics for the model were calculated to find the best performing model.

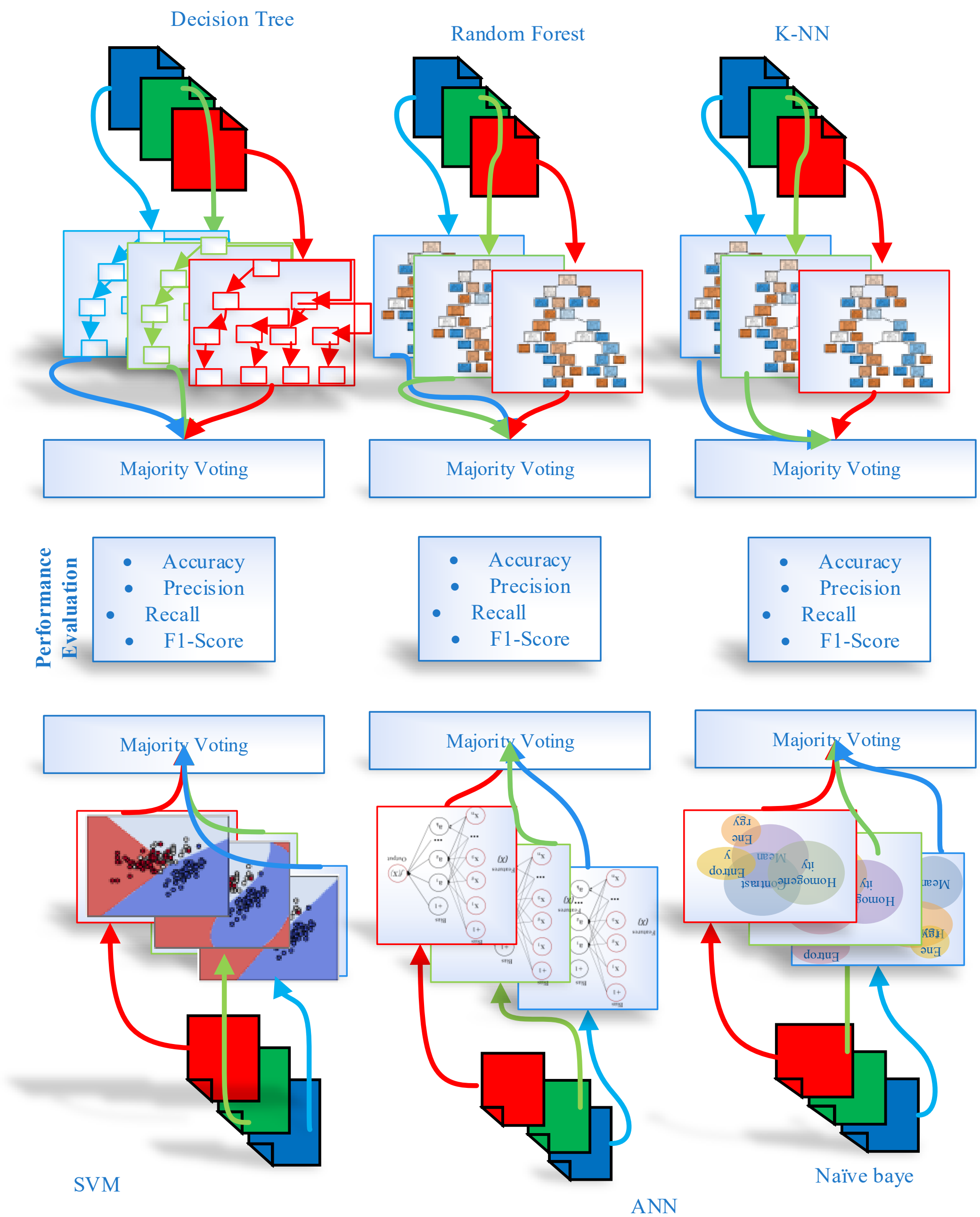


Figure 3.10 Structural Diagram for Classification Stage.

The remaining part of this section provides a concise description of each algorithm, their hyperparameters and tuning done for the experimentation. For comprehensive study and mathematical details, the reader can refer to the provided references of review papers and library links. For code details see GitHub.

3.3.1 Decision Tree:

Decision tree is a predictive modeling method widely used in machine learning, statistics, and data mining. The tree model is used to go from observations/features about an item to conclusions about the target variable. The target values are discrete (class labels) for classification and continues real number values for regression.

In the tree model, the branches represent the conjunctions of features that lead to the class labels. The leaves represent the final output of the model – labels or real number values. (Wu et al., 2008). The model implemented for this work used entropy as criterion for measuring the impurity. Other options include Gini Index, and Misclassification. All other parameters for the classifier were experimented with different values as well as default values. (F. a. V. Pedregosa, G. and Gramfort, A. and Michel, V., and Thirion, and Weiss, & Cournapeau, 2011).

3.3.2 Random Forest:

A perturb-and-combine technique algorithm for trees, random forest is a meta estimator that fits a number of decision tree classifiers on various sub-samples of the dataset and uses averaging to improve the predictive accuracy and control over-fitting. (Breiman, 2001). Hence, several trees which affect the accuracy are obtained and the final results depend on the voting results of the trees formed.

After experimenting with different number of estimators; 10, 100, 200; maximum depth 5,8,10, and criterion 'gini' for Gini impurity and 'entropy' for information gain, the best parameters were a maximum depth of 10 with number of estimators 100. The scikit-learn library is used for code implementation.(F. a. V. Pedregosa, G. and Gramfort, A. and Michel, V. et al., 2011).

3.3.3 k-NN:

k-NN is used for both supervised and unsupervised learning. For supervised learning it is implemented for classification and regression. K-NN is based on the principal to find a predefined number of training samples closest in distance to the new point, and hence, predicts a label using these. The distance can be any distance metric measure but the standard Euclidean distance is mostly preferred. The number of samples is user-defined (k-nearest neighbors) but depends on the local density of the points. k-NN are considered as non-generalized classifiers as they simply remember all the training data. In our implementation we set parameter of k =3. See Table 4.1.

3.3.4 SVM:

SVM are supervised learning models used for both classification and regression problems. SVM maps training examples to points in space so as to maximize the width of the gap between the two categories. They form the hyperplane or a set of hyperplanes in a n- dimensional space, which can be used for classification, regression, or tasks of outlier detection. (F. Pedregosa, 2011). After testing different parameter combinations, our model produced better results for linear kernel.

3.3.5 ANN:

ANNs are inspired by biological neural networks. ANN is a collection of connected nodes; each neuron sums up the output of previous units and apply an activation function to decide whether to fire the signal to the next layer of neurons or not. Some commonly used activation functions are Rectified Linear Unit (ReLU), softmax, sigmoid, and hyperbolic tangent. The number of output units and its activation function are important for classification and regression.

Here, we assume the reader has some intermediate level understanding of the working of neural networks along with its architecture. Read more at (Lundervold & Lundervold, 2019; F. a. V. Pedregosa, G. and Gramfort, A. and Michel, V. et al., 2011; Shen, Wu, & Suk, 2017).

The architecture of our implementation was a three hidden layer with a dropout of 30% along with Rectified learn Unit, we compiled the model with adam optimizer and binary_crossentropy for loss. Finally, the training was done with a batch size of 20 for 100 epochs. See Table 4.1 for more details on the parameters and hyperparameters for each model.

3.3.6 Naïve bayes:

Naïve bayes algorithm is a supervised learning algorithm based on the implementation of Bayes' theorem with the naïve assumption of conditional independence between every pair of features given the value of the class variable. The advantage of this method is that, given a small training data, a good estimate of the necessary parameters is accomplished. (F. a. V. Pedregosa, G. and Gramfort, A. and Michel, V. et al., 2011).

In our implementation we used BernoulliNB as we have to accomplish binary classification. Table 4.1 shows the hyperparameters for each classifier.

CHAPTER 4: IMPLEMENTATION/EXPERIMENTAL RESULTS/DISCUSSION

4.1 IMPLEMENTAION SETUP

Implementation of the work was done in Python 3.8.10 install on Intel® Core i5 having 8 GB RAM and some of the part for preprocessing and feature extraction was completed in MATLAB version 9.11.0.1847648 (R2021b) online.

The MRI dataset used for our work is achieved from(Keith A. Johnson, 2021). We selected 140 images – 90 normal and 50 abnormal images: for our work. The abnormal images were of glioma, meningioma, Alzheimer, Alzheimer plus visual agnosia pic's disease, sarcoma, and Huntington's.

Preprocessing was done on all 140 grayscale MRI images. Here, filtering was applied for removal of noise and transforms were applied to enhance the image details and quality, then the grayscale images were converted into RGB for next stage. Next, in feature extraction the statistical features were extracted from each channel and saved in separate file. Finally, during classification six machine learning models were used to find which model performs the best. Giving each channel's features to models and using majority voting to get the results. Accuracy, Precision, Recall and F1-Score were calculated to evaluate the performance measure of each of the six model for both testing and training. A more detailed discussion on each stage was given before in the modules for preprocessing, feature extraction and classification under methodology section. Table 4.1 below shows implementation details of the machine learning algorithms we used.

Table 4.1 Details of Parameters for each Classifier

Algorithm		Parameters			Testing/Training Split	
1	Artificial Neural Network	**Layers**	**Units/Neurons**	**Activation**	**Batch Size**	10
		Input layer (1)	9	ReLU	**Epochs**	100
		Hidden layer (3)	32, 64, 128	ReLU	**Test set**	30%
		Drop out Layer	-	rate = 0.3	**Train set**	70%
		Output layer (1)	1	Sigmoid	**Learning rate**	0.5
2	Decision Tree	Criterion = 'entropy'			R, G, B (30,70) %,	
3	k-NN	n_neighbors = 3,			R, G, B (30,70) %	
4	Naïve Bayes	BernoulliNB,			R, G, B (30,70) %	
5	Random Forest	criterion = 'gini', n_estimator = 10, max_depth=10			R, G, B (30,70) %	
6	SVM	kernel = 'linear'			R, G, B (30,70) %	

During the implementation several parameters were experimented with and were adjusted for optimal model performance. The above table shows the final best performing parameters. As there are few parameters for each model, it is because the default values are not necessary felt to be mentioned.

4.2 RESULTS AND DISCUSSION

To get the results, we decided to select decision tree as the best model based on its accuracy of 97.04%. As we obtained three files (red, green, and blue) for each channel in feature extraction stage, in our work we used three same machine learning models with different parameters. Each prediction from the classifier is evaluated using performance metrics of accuracy, precision, recall and f1-score. After getting the results from each classifier we used majority voting strategy to finalize the prediction and then again, we calculated the performance metrics both for training set and testing set. In majority voting strategy we compared all the three predictions for one input and based on majority votes the output was chosen. For instance, if two of the channels' prediction is 1 for an input, but the third channel model predicted 0 then the input is classified as 1 using majority voting. Similarly, if all three models predicted same prediction, then the majority voting gave the same output. Table 4.2 shows the values for the performance metrics, for each model on the result after majority voting on testing and training.

Table 4.2 Performance Metrics of each Classifier on Testing and Training Dataset

Algorithms	**Decision Tree**		**SVM**		**Random Forest**		**Artificial Neural Network**		**k-NN**		**Naïve Bayes**	
	Training Score	Testing Score	Training Score	Testing Score	Training Score	Testing Score	Training Score	Testing Score	Training Score	Testing Score	Training Score	Testing Score
Accuracy	100%	97.04%	65.30%	45.23%	100%	80.95%	96.93%	73.80%	63.26%	69.04%	*70%*	*52.38%*
Precision	100%	100%	54.28%	14.28%	100%	85.71%	97.22%	60%	53.34%	42.28%	*100%*	*0%*
Recall	100%	88.90%	51.35%	15.38%	100%	46.15%	94.95%	46.15%	42.10%	25%	*3.3%*	*0%*
F1-Score	100%	94.11%	52.77%	14.81%	100%	60%	96.07%	43.47%	39.95%	3.19%	-	-

Here, for we will only discuss the confusion matrix, accuracy, precision, recall and f1-score for the whole model i.e the results after majority voting. To avoid too much redundancy here we did not discuss the results of the model on each channel. Below is the concise definition of terms in confusion metrics and the performance metrics – accuracy, precision, recall and f1-score.

True Positives: The outcome where the model correctly predicts the positive class.

False Positives: The outcome where the model incorrectly predicts the positive class.

False Negatives: The outcome where the model incorrectly predicts the negative class.

True Negatives: the outcome where the model correctly predicts the negative class.

Accuracy is the simplest metric as it is just the number of correct predictions divided by the total number of predictions, it gives the proportion of correct predictions. Precision is the fraction of true positive instances among all true instances. Recall is also known as sensitivity of the model and is always reported alone with precision. Recall is the fraction true positives and the sum of true positives and false negatives. A perfect model has recall and precision equals to 1. Usually, the precision and recall are combined to avoid giving more importance to precision over recall in another metric called f1-score, it's the harmonic mean of the precision and recall.

The performance evaluators used are given by the equations (13-16) below:

$$Accuracy = \frac{True\ Positives + Ture\ Negatives}{All\ Samples} \tag{13}$$

$$Precision = \frac{True\ Positives}{True\ Positives + True\ Negatives} \tag{14}$$

$$Recall = \frac{True\ Positives}{True\ Positives + False\ Negatives} \tag{15}$$

$$F1\text{-}Score = \frac{2*(Precision*Recall)}{Precision + Recall} \tag{16}$$

4.2.1 Results and Discussion for ANN

The training and testing confusion matrices for ANN is given in Figure 4.1 and Figure 4.2 From the Table 4.1 in the implementation section, we used (98, 42) 30/70% test train split. So, using the confusion matrix from Figure 4.1 for ANN shows that, from 62 abnormal images, 60 were accurately classified, but two abnormal images were classified as normal on training dataset. Similarly, out of 36 normal images 35 were accurately classified as normal, while 1 normal image was wrongly classified as abnormal, which gives the accuracy of 96.03 % for training set.

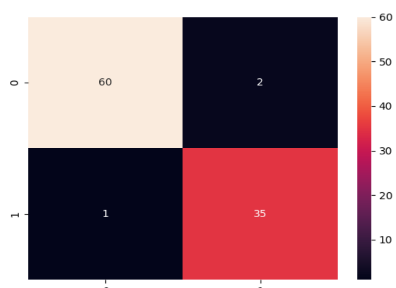


Figure 4.1 Confusion Matrix of ANN for Training Dataset

On the testing dataset the confusion matrix shows that out of 32 abnormal images, 24 were correctly classified while 8 abnormal images were inaccurately classified as normal. See Figure 4.2.

While out of 10 normal images half were misclassified as abnormal. The overall accuracy for testing dataset was 64.42%.

For all the performance metrics see Table 4.2 and Figure 4.3 below.

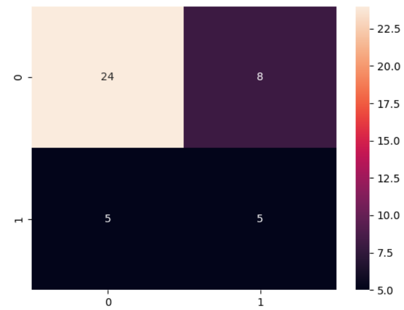


Figure 4.2 Confusion Matrix of ANN for Testing Dataset

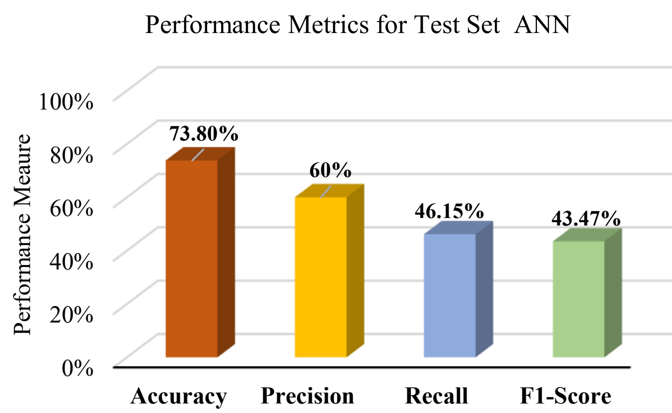


Figure 4.3. Graphical Representation of Performance Metrics for ANN

4.2.2 Results and Discussion for DT

The training and testing set confusion matrices for DT were calculated. From the dataset of 140 images, 105 were given for training (70%) and 35 (30%) were used for testing see the Table 4.1 from previous section. As the confusion matrix in Figure 4.4 illustrates, out of 64 abnormal images it accurately classified all of 64 as abnormal. For 41 normal images, it also classified all of them correctly as normal. Hence, an accuracy of 100% on training data.

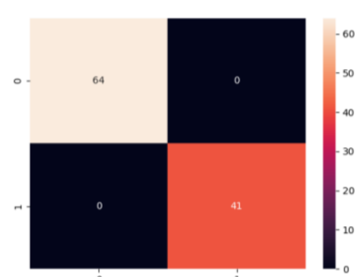


Figure 4.4 Confusion Matrix of DT for Training Dataset

However, for testing dataset the performance of DT out of 27 abnormal images, it classified 26 as abnormal, while it accurately classified all 8 normal images as normal.Figure 4.5 The accuracy for testing data was 97.46% due to one inaccurate classification of abnormal image.

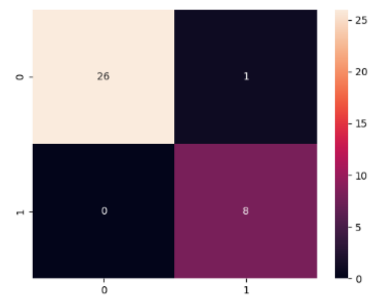


Figure 4.5 Confusion Matrix of DT for Testing Dataset

Refer to Table 4.2 for all the performance metrics values of DT. Figure 4.6 shows the visualization of the performance metrics.

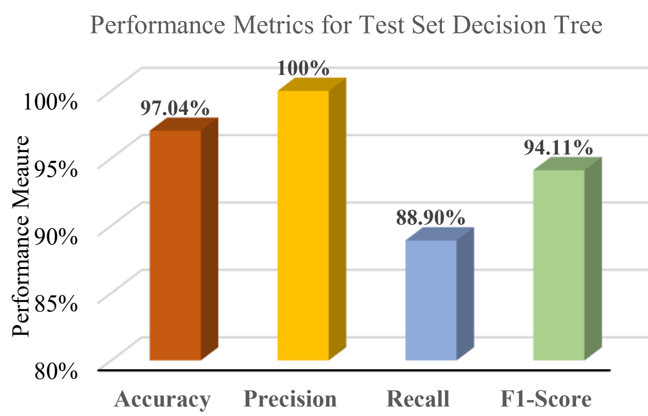


Figure 4.6 Graphical Representation of Performance Metrics for DT

4.2.3 Results and Discussion for k-NN

The training and testing set confusion matrices for k-NN were calculated to see the performance of the classifier. Table 4.1 from the implementation section. We split the dataset into 98 training, 42 testing dataset out of 140 images.

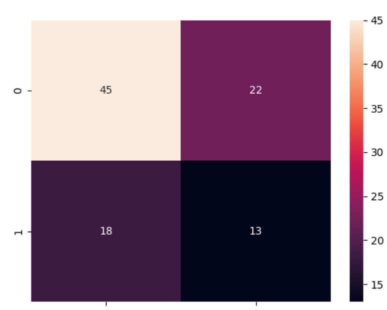


Figure 4.7 Confusion Matrix of k-NN for Training Dataset

The confusion matrix in Figure 4.7 gives the values for training as; out of 67 abnormal images, 45 were accurately classified and out of 31 normal images only 18 were correctly classified. The accuracy of k-NN for training was 59.18% and testing as 69.04%. See testing set graphical representation of performance metrics for k-NN from Table 4.2 in Figure 4.8 above.

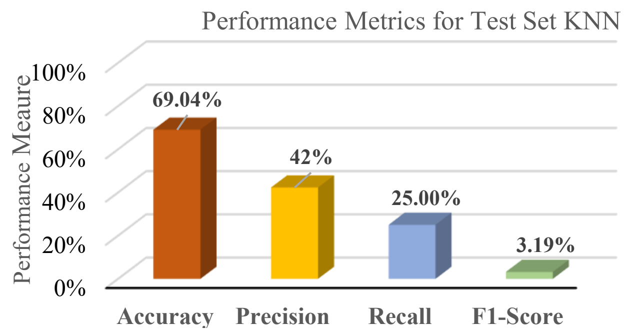


Figure 4.8 Graphical Representation of Performance Metrics for k-NN

On testing data, the performance was even worst, as indicated by Figure 4.9 Out of 36 abnormal images only 23 were correctly classified, and out of 6 normal images none were classified correctly. The graphical representation of the testing accuracy is given in Figure 4.8 above.

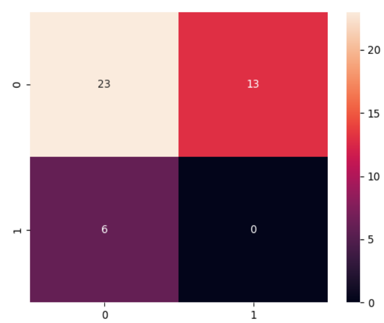


Figure 4.9 Confusion Matrix of k-NN for Testing Dataset

4.2.4 Results and Discussion for NB

The training and testing set confusion matrices for NB is given in Figure 4.10 and Figure 4.11 respectively. NB classifier was given the same train-test split dataset. A 98 train, 42 testing images.

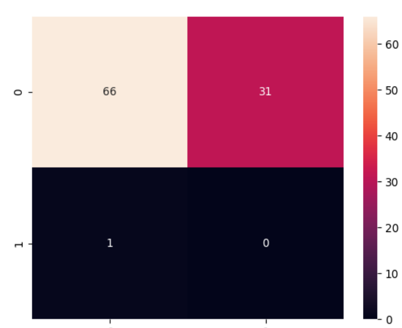


Figure 4.10 Confusion Matrix of NB for Training Dataset

NB classified 66 abnormal images correctly out of 91, while it was not able to correctly classify 1 normal image. The accuracy on training dataset was 67.34%.

For the testing dataset the NB was able to classify 27 abnormal images out of 42 abnormal images. As illustrated by Figure 4.11 0 out of 0 normal images were accurately classified for the testing dataset.

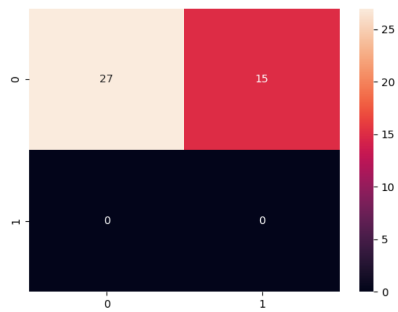


Figure 4.11 Confusion Matrix of NB for Testing Dataset

However, only 27 abnormal images out of 42 were accurately classified by NB. The testing accuracy was 52.38%. To see more detail metrics value, refer to Table 4.2. Figure 4.12 shows the graphical visualization of the performance metrics.

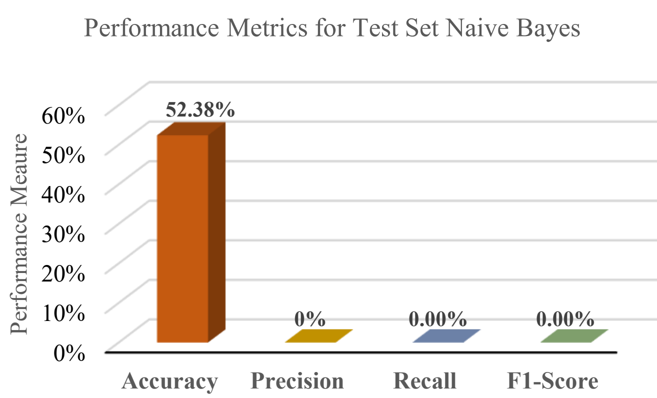


Figure 4.12 Graphical Representation of Performance Metrics for NB

4.2.5 Results and Discussion for RF

The confusion matrices for RF are shown in Figure 4.13 and Figure 4.14. A similar test-train split as for other models was done for Random Forest as well. 98-42. The confusion matrix obtained in Figure 4.13 presents that Random Forest was able to classify all 61 abnormal images as well as all 32 normal images correctly, scoring a 100% accuracy on training dataset.

While for testing dataset 6 out of 8 normal images were correctly classified and 27 out of 33 abnormal images were accurately classified as show in Figure 4.14.

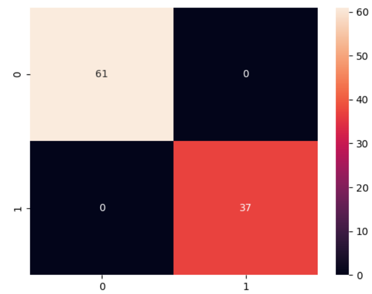


Figure 4.13 Confusion Matrix of RF for Training Dataset

The testing accuracy remained at 80.95%, see Table 4.2 for values of other metrics while Figure 4.15 shows the graphical representation of the accuracy, precision, recall and F1-Score for Random Forest.

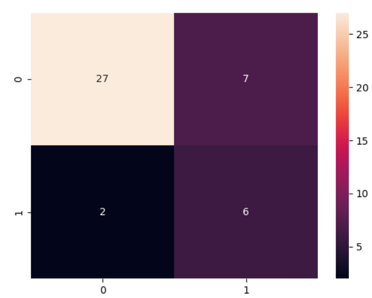


Figure 4.14 Confusion Matrix of RF for Testing Dataset

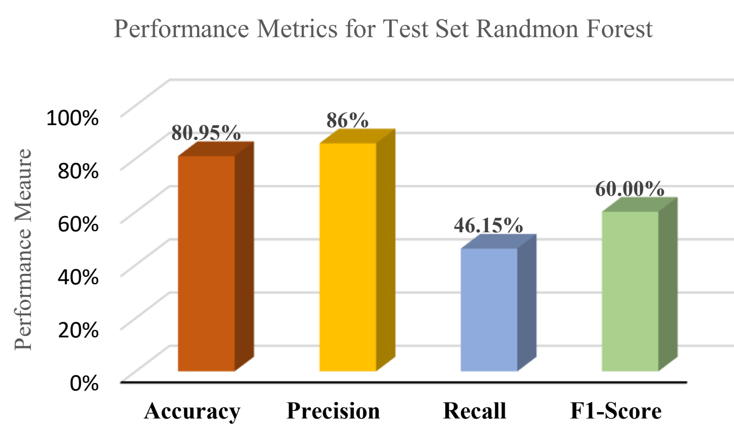


Figure 4.15 Graphical Representation of Performance Metrics for RF

4.2.6 Results and Discussion for SVM

The training and testing set confusion matrices for SVM is given in Figure 4.16 and Figure 4.17. For SVM the train test split remains as 98 – training and 42 – testing. The classification accuracy for training dataset was at lowest among all the models, a 65.04% after k-NN 59.18% and naïve bayes 67.04%. SVM on training classified 45 abnormal images correctly out of 63. And 16 normal images were accurately classified out of 35. See the Figure 4.16.

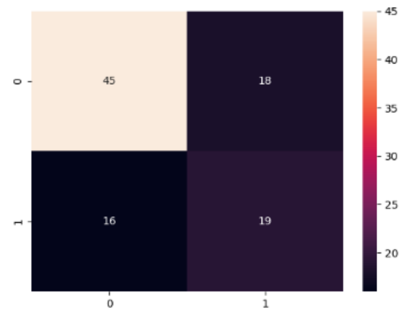


Figure 4.16 Confusion Matrix of SVM for Training Dataset

On testing dataset, as Figure 4.17 displays, SVM was able to correctly classify 17 abnormal images out of 28 and 2 normal images out of 14. See Table 4.2 and for its graphical representation see Figure 4.18 below.

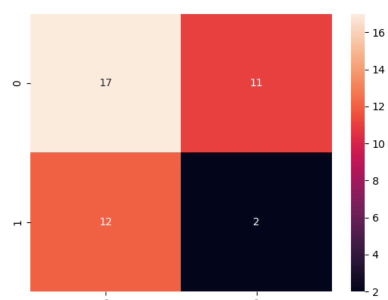


Figure 4.17 Confusion Matrix of SVM for Testing Dataset

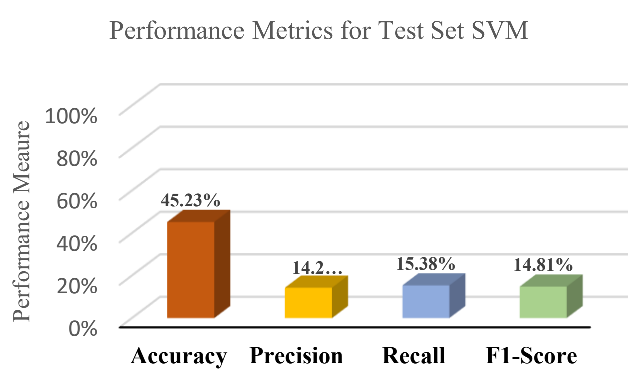


Figure 4.18 Graphical Representation of Performance Metrics for SVM

An internal model level comparison is provided below with the ranking of the six classifiers used based on their testing Table 4.3 and training Table 4.4 accuracies. The best performing model after all the preliminary testing on training set and testing set is found to be Decision Tree with an accuracy of 97.04 % on testing and unseen data sets.

Table 4.3 Ranking of Classifiers based on Accuracy measure for Testing Dataset

Rank	Algorithm	Testing Accuracy
1	Decision Tree	97.04%
2	Random Forest	80.95%
3	Artificial Neural Network	73.80%
4	k-NN	69.04%
5	Naïve Bayes	52.38%
4	Support Vector Machine	45.23%

Table 4.4 Ranking of Classifiers based on Accuracy measure for Training Dataset

Rank	Algorithm	Training Accuracy
1	Decision Tree	100%
1	Random Forest	100%
2	Artificial Neural Network	96.93%
3	Naïve Bayes	70.40%
4	Support Vector Machine	65.30%
5	k-NN	63.28%

Given below is the ranking of our models for the testing datasets. Figure 4.19 shows the visualization of the ranking.

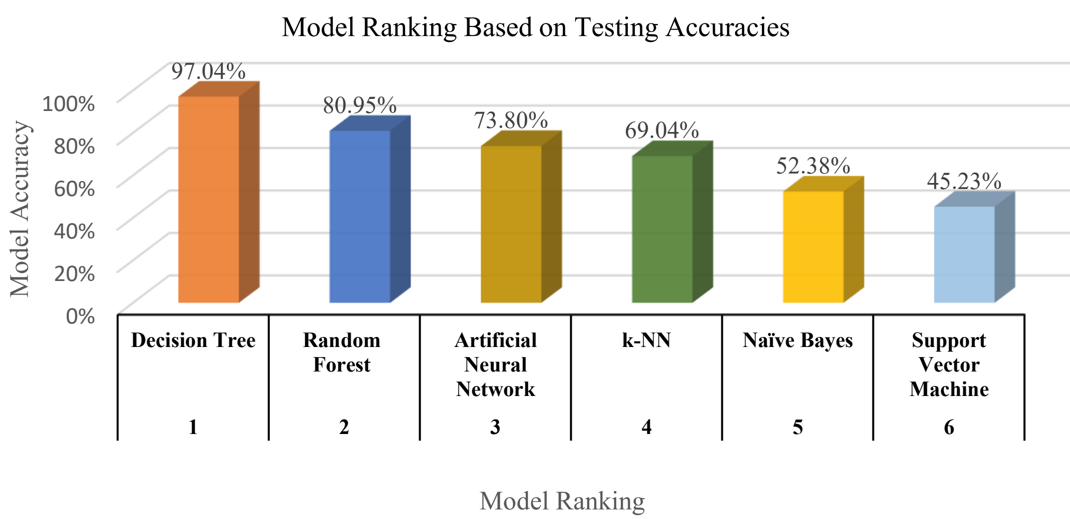


Figure 4.19 Graphical Representation of Ranking of the Models

A more robust way to check the selected model's performance is to see how the model performs in comparison to the well-known state-of-the-art proposed models in literature. For this purpose, the proposed method is compared with the counterpart algorithms in Table 4.5 and the results indicate that the performance of the proposed model is better as compared to the other methods for classification of brain MRI images. The proposed method also takes less time for computation.

Table 4.5 Classification accuracy of proposed method along with other well-known methods

S.No.	Methodology	Classification Accuracy
1	CM + ANN (Nazir et al., 2015)	92%
2	DWT+ ANN (Rajini & Bhavani, 2011)	90%
3	Zahid *et al.* (Ullah et al., 2019)	95%
4	Fayaz *et al.* (Fayaz et al., 2016)	96%
5	**Proposed Methodology**	**97.04%**

CHAPTER 5: CONCLUSION AND FUTURE WORK

In this work we undertook the study of binary classification for brain MRI images into normal and abnormal categories with a three-stage methodology. The methodology consisted of preprocessing, feature extraction and classification. In preprocessing stage, filters and transforms were implemented to get quality data/image. Filters like median filter, prewitt and Laplacian filters were used, and log transform along with contrast stretching transforms were implemented on the grayscale MRI images. At the end of preprocessing stage, the obtained noise reduced, and quality enhanced images were converted to RGB images. Then during the feature extraction stage, the statistical entities were calculated for red, green, and blue channels. Mean, contrast, correlation, energy, entropy, homogeneity, kurtosis, skewness, and variance were the features considered and extracted for each channel of the RGB images and saved in files for each channel. This stage provided the final dataset in form of three files for red, green, and blue channels to be used in classification stage. Different machine learning models were used in classification stage, namely artificial neural networks, decision tree, naïve bayes, k-NN, random forests and support vector machines. Using the predictions of the model for each channel, the results were considered after majority voting. For each model the performance was evaluated using the metrics of accuracy, recall, precision and f1-score. For the better understanding of the performance of each model, we also calculated the accuracy, recall, precision, and f-score for both training dataset and testing dataset. A great amount of experimentation was done on tuning the hyperparameters of the models, selecting those parameters which yielded the best results in terms of accuracy. Lastly, after the performance evaluation of each model on each channel and the comparison of the six models helped us to conclude the best model as Decision Tree having an accuracy of 97.04%. The results of decision tree indicated that our proposed model performed better using decision tree on our dataset.

Future recommendations to be considered are to increase the dataset and test the proposed model. However, the advances in deep learning and the new emerging models for MRI images required huge amount of data. Unfortunately, getting medical data is not easy so we have to rely on small datasets; and on image processing techniques to select prominent features and apply classical machine learning algorithms. It is strongly recommended to use deep learning models instead of classical machine learning algorithms. Furthermore, deep learning models also eliminate the steps of image processing like filtering, transforming and the steps of feature extraction.

Apart from using deep learning methods with large dataset, another recommendation to be made here is the deployment of the best performing model as web application so that any user can upload their brain MRI and check the results in a user-friendly graphical web interface.

Reference:

Akkus, Z., Galimzianova, A., Hoogi, A., Rubin, D. L., & Erickson, B. J. J. J. o. d. i. (2017). Deep learning for brain MRI segmentation: state of the art and future directions. *30*(4), 449-459.

Begg, R. K., Palaniswami, M., & Owen, B. J. I. t. o. B. E. (2005). Support vector machines for automated gait classification. *52*(5), 828-838.

Breiman, L. J. M. l. (2001). Random forests. *45*(1), 5-32.

Chaplot, S., Patnaik, L. M., Jagannathan, N. J. B. s. p., & control. (2006). Classification of magnetic resonance brain images using wavelets as input to support vector machine and neural network. *1*(1), 86-92.

Chityala, R., & Pudipeddi, S. (2020). *Image processing and acquisition using Python*: Chapman and Hall/CRC.

da Silva, A. R. F. J. M. i. a. (2007). A Dirichlet process mixture model for brain MRI tissue classification. *11*(2), 169-182.

El-Dahshan, E.-S. A., Hosny, T., & Salem, A.-B. M. J. D. s. p. (2010). Hybrid intelligent techniques for MRI brain images classification. *20*(2), 433-441.

Fayaz, M., Haider, J., Qureshi, M. B., Qureshi, M. S., Habib, S., & Gwak, J. J. I. A. (2021). An Effective Classification Methodology for Brain MRI Classification Based on Statistical Features, DWT and Blended ANN. *9*, 159146-159159.

Fayaz, M., Shah, A. S., Wahid, F., Shah, A. J. I. J. o. S. P., Image Processing, & Recognition, P. (2016). A robust technique of brain MRI classification using color features and K-nearest neighbors algorithm. *9*(10), 11-20.

Georgiadis, P., Cavouras, D., Kalatzis, I., Daskalakis, A., Kagadis, G. C., Sifaki, K., . . . biomedicine, p. i. (2008). Improving brain tumor characterization on MRI by probabilistic neural networks and non-linear transformation of textural features. *89*(1), 24-32.

Hong, L., Wan, Y., Jain, A. J. I. t. o. p. a., & intelligence, m. (1998). Fingerprint image enhancement: algorithm and performance evaluation. *20*(8), 777-789.

Joseph, R. P., Singh, C. S., Manikandan, M. J. I. J. o. R. i. E., & Technology. (2014). Brain tumor MRI image segmentation and detection in image processing. *3*(1), 1-5.

Kazmi, J. H., Qureshi, K., & Rashid, H. J. M. J. o. C. S. (2007). Enhanced MRA Images Quality Using Structure Adaptive Noise Filter And Edge Sharpening Methods. *20*(2), 99-114.

Keerthana, T., & Xavier, S. (2018). *An intelligent system for early assessment and classification of brain tumor.* Paper presented at the 2018 Second International Conference on Inventive Communication and Computational Technologies (ICICCT).

Keith A. Johnson, J. A. B. (Producer). (2021, 12 15). The Whole Brain Atlas. *The Whole Brain Atlas*. Retrieved from http://www.med.harvard.edu/AANLIB/home.html

Korolev, S., Safiullin, A., Belyaev, M., & Dodonova, Y. (2017). *Residual and plain convolutional neural networks for 3D brain MRI classification.* Paper presented at the 2017 IEEE 14th International Symposium on Biomedical Imaging (ISBI 2017).

Kumar, G., & Bhatia, P. K. (2014). *A detailed review of feature extraction in image processing systems.* Paper presented at the 2014 Fourth international conference on advanced computing & communication technologies.

Logeswari, T., Karnan, M. J. I. J. o. C. T., & Engineering. (2010). An improved implementation of brain tumor detection using segmentation based on hierarchical self organizing map. *2*(4), 591.

Lundervold, A. S., & Lundervold, A. J. Z. f. M. P. (2019). An overview of deep learning in medical imaging focusing on MRI. *29*(2), 102-127.

Maitra, M., Chatterjee, A. J. B. S. P., & Control. (2006). A Slantlet transform based intelligent system for magnetic resonance brain image classification. *1*(4), 299-306.

Mallat, S. G. (2009). A theory for multiresolution signal decomposition: the wavelet representation. In *Fundamental Papers in Wavelet Theory* (pp. 494-513): Princeton University Press.

Marr, D., & Hildreth, E. J. P. o. t. R. S. o. L. S. B. B. S. (1980). Theory of edge detection. *207*(1167), 187-217.

Martelli, A. J. C. g., & processing, i. (1972). Edge detection using heuristic search methods. *1*(2), 169-182.

Mathew, A. R., & Anto, P. B. (2017). *Tumor detection and classification of MRI brain image using wavelet transform and SVM.* Paper presented at the 2017 International Conference on Signal Processing and Communication (ICSPC).

MathWorks (Producer). (2022, 03 03). Feature Extraction. Retrieved from https://www.mathworks.com/discovery/feature-extraction.html

Mishra, R. (2010). *MRI based brain tumor detection using wavelet packet feature and artificial neural networks.* Paper presented at the Proceedings of the International Conference and Workshop on Emerging Trends in Technology.

Mostapha, M., & Styner, M. J. M. r. i. (2019). Role of deep learning in infant brain MRI analysis. *64*, 171-189.

Nazir, M., Wahid, F., Ali Khan, S. J. J. o. I., & Systems, F. (2015). A simple and intelligent approach for brain MRI classification. *28*(3), 1127-1135.

Nicki, R. C., Brian, R. W., & Stuart, H. R. J. E. (2010). Davidson's principles and practice of medicine. *21*, 744.

Osher, S., & Rudin, L. I. J. S. J. o. n. a. (1990). Feature-oriented image enhancement using shock filters. *27*(4), 919-940.

Othman, M. F. B., Abdullah, N. B., & Kamal, N. F. B. (2011). *MRI brain classification using support vector machine.* Paper presented at the 2011 fourth international conference on modeling, simulation and applied optimization.

Pal, S. K., King, R. J. I. T. S., MAN,, & CYBER. (1981). Image enhancement using smoothing with fuzzy sets. *11*(7), 494-500.

Pedregosa, F. (2011). Scikit-learn: Machine Learning in Python. *Journal of Machine Learning Research*, 2825--2830.

Pedregosa, F. a. V., G. and Gramfort, A. and Michel, V., and Thirion, B. a. G., O. and Blondel, M. and Prettenhofer, P., and Weiss, R. a. D., V. and Vanderplas, J. and Passos, A. and, & Cournapeau, D. a. B., M. and Perrot, M. and Duchesnay, E. (2011). Scikit-learn: Machine Learning in Python *Journal of Machine Learning Research, 12*, 2825-2830.

Prewitt, J. M. J. P. p., & Psychopictorics. (1970). Object enhancement and extraction. *10*(1), 15-19.

Rajini, N. H., & Bhavani, R. (2011). *Classification of MRI brain images using k-nearest neighbor and artificial neural network.* Paper presented at the 2011 International conference on recent trends in information technology (ICRTIT).

Reihl, S. J., Patil, N., Morshed, R. A., Mehari, M., Aabedi, A., Chukwueke, U. N., . . . Waite, K. J. m. (2021). A population study of clinical trial accrual for women and minorities in neuro-oncology Following the NIH Revitalization Act.

Robinson, G. S. J. O. E. (1977). Detection and coding of edges using directional masks. *16*(6), 166580.

Rogers, K. (Producer). (2022, 2 9). Britannica. *Britannica.com*. Retrieved from https://www.britannica.com/science/brain

Saleh, S. R., & Al-Bakry, A. M. (2017). *MRI images classification based on software agent.* Paper presented at the 2017 Annual Conference on New Trends in Information & Communications Technology Applications (NTICT).

Saritha, M., Joseph, K. P., & Mathew, A. T. J. P. R. L. (2013). Classification of MRI brain images using combined wavelet entropy based spider web plots and probabilistic neural network. *34*(16), 2151-2156.

Shen, D., Wu, G., & Suk, H.-I. J. A. r. o. b. e. (2017). Deep learning in medical image analysis. *19*, 221-248.

Shenton, M., Kikinis, R., Jolesz, F., Pollak, S., LeMay, M., Martin, J., . . . McCarley, R. J. N. E. J. M. (1992). Left-lateralized temporal lobe abnormalities in schizophrenia and their relationship to thought disorder: A computerized, quantitative MRI study. *327*(9), 604-612.

Somasundaram, K., Kalaiselvi, T. J. C. i. b., & medicine. (2010). Fully automatic brain extraction algorithm for axial T2-weighted magnetic resonance images. *40*(10), 811-822.

Sudiro, S. A., Kardian, A. R., Madenda, S., Hermanto, L. J. I. J. o. A. S., & Engineering. (2021). Mean and variance statistic for image processing on FPGA. *18*(1), 1-6.

Ullah, Z., Fayaz, M., Iqbal, A. J. I. J. o. M. E., & Science, C. (2016). Critical analysis of data mining techniques on medical data. *8*(2), 42.

Ullah, Z., Lee, S.-H., Fayaz, M. J. I. J. o. A., & Sciences, A. (2019). Enhanced feature extraction technique for brain MRI classification based on Haar wavelet and statistical moments. *6*(7), 89-98.

Ullah, Z., Lee, S.-H., Khan, M. N. A., Fayaz, M., & Iqbal, M. M. J. T. J. (2018). Features Reductions Using Color Moments and Classification of Brain MRI Using K-NN. *23*(04), 77-83.

Vasuda, P., Satheesh, S. J. I. J. o. C. S., & Engineering. (1713). Improved fuzzy C-means algorithm for MR brain image segmentation. *2*(5), 2010.

Wahid, F., Ghazali, R., Fayaz, M., Shah, A. S. J. I. J. o. B.-S., & Bio-Technology. (2016). Using Probabilistic Classification Technique and Statistical Features for Brain Magnetic Resonance Imaging (MRI) Classification: An Application of AI Technique in Bio-Science. *8*(6), 93-106.

WHO (Producer). (2022, 02 03). WHO. *who.int*. Retrieved from www.who.int/news-room/fact-sheets/cancer

Wu, X., Kumar, V., Ross Quinlan, J., Ghosh, J., Yang, Q., Motoda, H., . . . systems, i. (2008). Top 10 algorithms in data mining. *14*(1), 1-37.

Yasaka, K., Akai, H., Kunimatsu, A., Kiryu, S., & Abe, O. J. J. j. o. r. (2018). Deep learning with convolutional neural network in radiology. *36*(4), 257-272.

Zanaty, E., & Aljahdali, S. (2011). *Improving fuzzy algorithms for automatic image segmentation.* Paper presented at the 2011 International Conference on Multimedia Computing and Systems.

Zhang, Y., Dong, Z., Wu, L., & Wang, S. J. E. S. w. A. (2011). A hybrid method for MRI brain image classification. *38*(8), 10049-10053.

Zhu, X., Suk, H.-I., Zhu, Y., Thung, K.-H., Wu, G., & Shen, D. (2015). *Multi-view classification for identification of Alzheimer's disease.* Paper presented at the International Workshop on Machine Learning in Medical Imaging.