Brain Hemorrhage Detection Using Hybrid Machine Learning Algorithm

by

Khondoker Nazia Iqbal
18101006
Istinub Azad
18101045
Md. Imdadul Haque Emon
18101049
Nibraj Safwan Amlan
18101596
Amena Akter Aporna
18301236

A thesis submitted to the Department of Computer Science and Engineering
in partial fulfillment of the requirements for the degree of
B.Sc. in Computer Science

Department of Computer Science and Engineering
Brac University
January 2022

Declaration

It is hereby declared that

1. The thesis submitted is my/our own original work while completing degree at Brac University.

2. The thesis does not contain material previously published or written by a third party, except where this is appropriately cited through full and accurate referencing.

3. The thesis does not contain material which has been accepted, or submitted, for any other degree or diploma at a university or other institution.

4. We have acknowledged all main sources of help.

Student's Full Name & Signature:

Khondoker Nazia Iqbal

Khondoker Nazia Iqbal
18101006

Istinub Azad

Istinub Azad
18101045

Md. Imdadul Haque Emon.

Md. Imdadul Haque Emon
18101049

Nibraj Safwan Amlan

Nibraj Safwan Amlan
18101596

Amena Akter Aporna

Amena Akter Aporna
18301236

Approval

The thesis/project titled "Brain Hemorrhage Detection Using Hybrid Machine Learning Algorithm" submitted by

1. Khondoker Nazia Iqbal (18101006)
2. Istinub Azad (18101045)
3. Md. Imdadul Haque Emon (18101049)
4. Nibraj Safwan Amlan (18101596)
5. Amena Akter Aporna (18301236)

Of Spring, 2022 has been accepted as satisfactory in partial fulfillment of the requirement for the degree of B.Sc. in Computer Science on January 16, 2022.

Examining Committee:

Supervisor:
(Member)

Md. Saiful Islam
Lecturer
Department of Computer Science and Engineering
Brac University

Co-Supervisor:
(Member)

Rafeed Rahman
Lecturer
Department of Computer Science and Engineering
Brac University

Head of Department:
(Chair)

Sadia Hamid Kazi, PhD
Chairperson and Associate Professor
Department of Computer Science and Engineering
Brac University

Abstract

Machine learning (ML) helps computers learn and program data without humans' help. According to data scientists, machine learning can extract 60% high-quality information, reduce the cost up to 46%, and increase operation speed by approximately 48% [1]. Recently, there has been successful implementation of machine learning in data analysis, computer vision, computer-aided diseases (CAD), and many more fields. Machine learning is broadly used in the medical industry because of its processing power for image data and pattern recognition quality. The image processing power of machine learning can be used in medical images to classify the brain images automatically. Segmentation and classification of brain image can provide valuable information and quantitative assessment of lesions which can be used for treatment strategies and predicting patient condition (Kamnitsas et al., 2017). According to research [2], an estimated 64-74 million people in the world are affected by traumatic brain injury every year. It affects the lives of nearly every one out of six persons. In our proposed system, we will use a hybrid approach of multiple machine learning algorithms together for the classification of CT brain images and diagnose brain disorders and diseases like brain hemorrhage. Some ML algorithms such as different 3D Convolutional Neural Networks (CNN) , AlexNet, DenseNet121, GoogleNet and some other models like Multilayer Perceptron Model (MLP), Support Vector Machine (SVM) and Random Forest (RF) have been applied successfully in this field in the past. Modifying previous methods, we want to build a hybrid machine learning algorithm by combining different CNN models like VGG-16, VGG-19, Random forest and Multilayer Perceptron (MLP) classifiers for detecting brain hemorrhage. We have used the VGG-16 and VGG-19 model to derive image features from the CT brain images and Random forest classifier and MLP classifier for testing the accuracy of our model. To test the efficiency of our system, we have used CT brain image datasets from Kaggle. The CT brain imaging data will be the input of our model and our model will detect brain hemorrhage and classify them into one of six classes: Epidural, Intraparenchymal, Intraventricular, Subarachnoid, Subdural and No Hemorrhage. Using our hybrid approach the best accuracy we achieved was around 97.24% using a combined approach of VGG-16 and Multilayer Perceptron classifier. Also we used Explainable AI to explain the prediction of the hemorrhagic classes.

Keywords: Hybrid Machine learning, Convolutional Neural Network (CNN), Multilayer Perceptron Model (MLP), Random Forest (RF), VGG-16, VGG-19, Brain Hemorrhage, Explainable AI

Acknowledgement

At first, all appreciation is owed to Allah who has helped us to accomplish our thesis work without any serious setbacks.
Secondly, our sincere appreciation to our supervisor, Md. Saiful Islam sir for his constant aids and directions to reach our goal. He was always there for us without any complain.
Thirdly, our eternal gratefulness to our co-supervisor, Rafeed Rahman sir for assisting and guiding us throughout the whole thesis work.
At the end, we express our unconditional love towards our parents. Without their throughout prayers and encouragement, it was impossible for us to pull up our graduation.

Table of Contents

List of Figures

List of Tables

Nomenclature

The next list describes several abbreviation that will be later used within the body of the document

AI Artificial Intelligence

CNN Convolutional Neural Network

CRF Conditional Random Field

CT Computed Tomography

MLP Multilayer Perceptron

MRI Magnetic Resonance Imaging

RF Random Forest

ROI Region of Interest

$SMOTE$ Synthetic Minority Oversampling Technique

SVM Support Vector Machine

Chapter 1

Introduction

The brain is considered to be the central organ of the sensory system. It is considered as the control unit of the human body as it directs most of the activities of a human body. Unlike all the body parts of a human body, the brain also faces some disorders when our brain is damaged or harmed. A brain hemorrhage is a condition when there is an internal bleeding between skull and brain area or inside the brain. Brain hemorrhages are sometimes life threatening and often lead to death. The ratio of death caused by brain hemorrhage is very high at present in the world. According to an article [3], brain hemorrhage occurs for 10–15 percent of all strokes and is associated with extremely high morbidity, and death rates have remained unchanged over the previous 40 years. So, the treatment for brain hemorrhage is becoming essential for us. To diagnose the patient's disease through the medical image, doctors usually suggest MRI or CT scan. CT scans are widely used compared to brain MRI. A Computed Tomography scan or CT scan is a medium used to detect brain abnormalities. It is a diagnosis where images of inside organs of the human body are taken using a series of X-rays. CT scanned images are more convenient than X-ray images because, in regular x-rays, there is only one tube fixed which can send rays in only one direction, but CT scan uses motorized x-rays which keep rolling over the human body, which helps to get a clearer image of the body. For our research, we have considered using CT brain images for detecting and classifying brain hemorrhage.

Usually, a radiologist or a physician searches for abnormality in CT images and prepares the report after the scan is done. Preparing the report for a scan takes a lot of time and effort. With the help of Computer-Aided Detection (CAD), it will be easier for radiologists to analyze data and generate reports. For the help of the radiologists on their task of interpreting medical images,Computer Aided Detection has been improved a lot in the last two decades. (Mohsen et al., 2012) [4].

Machine learning can be used for analyzing and detecting brain hemorrhage from CT scan data. In machine learning, CNN, SVM, RF, CRF are the common methods

that are used to detect brain hemorrhage from image data. Although there is already much work done in this field, there is still room for improvement. Still, brain segmentation using machine learning needs future research because till now it lacks accuracy, precision, and robustness (Balafar et al., 2010) [5].
For detecting and analyzing brain hemorrhage from CT scan data we applied four ML algorithms. We have used VGG-16, VGG-19, RF, and MLP algorithms. Here, the VGG-16, VGG-19 are used for feature extraction. After feature extraction, for testing, we have used the Random Forest and MLP algorithms. Also, we have tried to extract the Region of Interest using LIME which is an Explainable AI model.

1.1 Problem Statement

According to an article [6], neurological diseases were the most significant reason for disability-adjusted life years(95%) and the second biggest cause of mortality (88%–94%) worldwide in 2016. A massive portion of the world population is diagnosed on a daily basis for analyzing brain disorders and neurological diseases. These diagnoses can be done in many ways and imaging tests are mostly used from them. The two most popular brain imaging test techniques to determine brain abnormalities are MRI and CT scans. Usually, these scans are studied by a radiologist and other specialists manually, and they prepare the result, which is very time-consuming and takes too much effort. In a website it shows statistics that around 95.65 million MRI scans are annually performed in the US and Japan alone. The amount of medical data is adding up every day and getting more complex to analyze. This is very challenging to analyze this huge amount of data maintaining both accuracy and robustness. We can reduce the burden of workload of analyzing these data using machine learning or deep learning algorithms. Machine learning can be applied to process these medical scans and segment the brain tissue to separate the abnormalities in the brain (e.g. hemorrhage, various kinds of high-grade as well as low-grade gliomas, metastases, lesions and injury in brain tissue). Which can be a much more efficient and accurate way if it can be appropriately applied.

We want to build a robust way to determine the injuries and disorders like hemorrhage in the brain using different machine learning algorithms. We will use a hybrid machine learning model by combining different machine learning algorithms. We will use CNN architectures like VGG16, VGG19 for feature extraction and Random Forest (RF) classifier and Multilayer Perceptron Model (MLP) for calculating the accuracy of our model. This hybrid model will give more accurate results compared to the currently available methods. For testing our model and measuring the accuracy and robustness of it, we have used widely available CT brain imaging datasets from Kaggle so that we can compare our results with already applied methods.

1.2 Research Objective

On the basis of a statistic from the WHO (World Health Organization), neurological illnesses claim the lives of an estimated 6.8 million individuals each year and the expense of neurological diseases in Europe was projected to be at 139 billion euros[7]. Till now, many scholars did research and found different ways to predict and diagnose brain disorders. The goal of this study is to see how well hybrid models can predict and categorize different types of brain hemorrhage from CT scans of Brain. We have chosen to conduct research in this field because we wanted to build an accurate model to determine brain hemorrhage and reduce the computational cost of detecting neurological diseases. Also we are interested to examine how a hybrid approach performs against the previously applied approaches in this field.

Chapter 2

Related Work

Sage and Badura, (2020) presented an approach [8] of detecting intracranial hemorrhage (ICH) in computed tomography head images. They have used a ResNet architecture of dual branch convolutional neural networks. They have classified the CT images into five different subtypes of intracranial hemorrhage. In their research, they have collected in total 372,556 images which form 11,454 computed tomography series belonging to 9997 patients. In this research work, a CNN model is used which is a double branched model. They used this model to get the features. And to conduct classification task, they have used a RF classifier and SVM classifier. Their model has achieved 96.7% accuracy in intraventricular hemorrhage, and in intraparenchymal hemorrhage, they have gained 93.3% accuracy.

A deep learning system that is based on a Restricted Boltzmann Machine (RBM) as representational translation and a Random Forest (RF) as task-specific was proposed in a 2017 article [9]. This study used a publicly accessible BRATS database from the MICCAI BRATS Challenge. They provided an approach for data comprehension and feature derivation from MRI image data that combined RF and RMB. They have used methods like-Triple Cascaded Framework, Anisotropic Convolutional Neural Network,Multi-View Fusion, Augmentation for training and testing ,Uncertainty Estimation of Segmentation Result. The memory requirements for using a 2D CNN slice-by-slice are relatively minimal. Here they have used datasets BraTS 2017 and BraTS 2018 for validation of their used methods . Both datasets use the identical collection of training photos from 285 people, with 75 222 LGG instances and 210 HGG cases. Images 223 from 46 and 66 patients with traumatic brain injury are included in the BraTS 2017 and 2018 validation sets, respectively. Images from patients with traumatic brain injury are included in the BraTS 2017 and 2018 testing sets, respectively. In the verification and validation sets, the grades of brain tumors 225 remain unknown. FLAIR, T1, and T2 were used to scan each patient. The original photos were captured from various angles with an isotropic resolution. As a result, the suggested Cascaded approach is highly suited for segmenting tumor regions.

Mitra et al., (2016) showed that Traumatic Brain Injury (TBI) can be identified using the Random Forest regression model [10]. To classify severe brain injury and healthy control subjects, they developed machine learning techniques. By using their proposed method, they claimed to analyze and categorize the whole-brain network. The proposed method has two main phases: (a) Feature extraction by network connectivity analysis; and (b) Discriminative extraction of features via classifier. At first, they begin by utilizing a probabilistic tractography approach to achieve entire-brain tractography. Using Network-Based Statistic (NBS), they discovered statistical significance and discriminative structural relations between these two data sets. Finally, they use Random Forest classification to classify the patients and extract the features, which helps differentiate them from healthy participants. They compared their implemented method with the previous ones and found a mean accuracy close to 68.16%±1.81% and a mean sensitivity around 80.0% ±2.36% to classify the TBI patients accurately.

A paper proposed by Biomedical Image Analysis Group of Imperial College London, UK [11], presented a double pathway, 11-layer deep, 3D CNN for division of brain lesions. They examined the advantages of utilizing small convolutional portions in 3D CNNs. Furthermore, they asserted to apply 3D completely associated CRF in clinical information without increasing the computational expense. Additionally, they utilize similar convolutional pathways for multi-scale transforming, which effectively consolidates neighborhood and logical data, altogether further developing division results. Right away, they applied a 3D CNN model to section the head injuries with a precision level. Then, at that point, a completely 3D CRF model is utilized to play out a total division planning on the recently extricated division highlights from the CNN model. Their proposed framework helps segment brain lesions of 3 kinds of cerebrum wounds like- TBI, brain tumors, and ischemic stroke. Moreover, in view of the little convolutional bits of 3D CNN, their proposed technique assists with fragmenting cerebrum injuries at a lower cost without expanding the teachable boundaries. Around 66.6% mean DSC on the validation fold is accomplished by the 3D adaptation of profound media with CRF. Some best, regulated division techniques for cerebrum injuries depend on voxel-wise classifiers, Convolutional Neural Network, Conditional Random Field, Random Forests. A 3D brain examine division is accomplished by preparing each 2D cut freely, which is seemingly a non-ideal utilization of the volumetric clinical image information.

A novel system for MRI data segmentation is developed by a research work using 2.5D CNNs [12]. In this paper, they have used 2.5D CNN because its intra-receptive field is much bigger than the slice receptive field. To segment the property of tumor structure, they use three networks. At first, using Wnet they segregated the entire tumor. After that, TNet extracts the tumor center from the cropped image domain. Finally, ENet separates the developing tumor core from the next cropped zone. The author presented a 2.5D CNN with a large domain and a smaller slice receptive field to balance memory consumption and model complexity. Their method was top scorers for BraTS 2017 dataset. Also, they proposed a CRF method for the after-processing of the MRI data. Their method scored 3.28±3.88 (Enhancing tumor core), 3.89±2.79 (Whole tumor), and 6.48±8.26 (Tumor core), which is better than the other methods.

To detect Alzheimer's and MCI disease, the authors of a paper [13] demonstrated the probability of using functional MRI scans. They have combined two different methods in order to develop a system which can provide the best performance. In order to do this, they combined two feature selection techniques such as univariate t-test and Recursive Feature Elimination. The proposed method extracts the huge amount of information changes because of Alzheimer or MCI disease. In contrast to all other previous research, the proposed method calculates the highest accuracy of classification. To increase the amount of discriminative patterns a support vector machine was employed using a recursive feature elimination (SVM-RFE) and LASSO in concurrence with the univariate t-test. In the ADNI2 cohort, the technique got maximum accuracies of 98.86% and 98.57% for Alziemer disease and mild cognitive impairment vs. CN, respectively. Some same kinds of accuracies were 98.70% and 94.16% found in the in-house cohort.

Zacharaki et al., (2009) presented a classification scheme to differentiate adult brain tumors in their research work [14]. In order to do this, they have used a support vector machine classifier which helps to classify texture patterns. This study employs a variety of methodologies and materials. A multiparametric framework is used to classify brain tumors and predict their degree of malignancy. A multiparametric imaging profile also uses the Pattern classification algorithm. In order to identify different glioma grades and metastasis, they used a cross validation method. They have combined different classification techniques such as Linear Discriminant Analysis, K-Nearest Neighbour and nonlinear support vector machines. They have also used t-test and constrained linear discriminant analysis as a feature ranking approach. Overall, the average number of chosen features in classification tasks has decreased by 33%.

Classification of CT brain hemorrhage using a 3D CNN network is done in the paper [15] of Jnawali, Kamal, et al. In their proposed model, they first extract the features from the CT images. After that, they used a logistic function in order to detect brain hemorrhage. They have collected their dataset from a health system named Geisinger Health System. This dataset contains 40,357 images of the CT brain, where 30,001 images indicate no hemorrhage, and 10,356 images indicate that there are hemorrhages. In order to decrease the imbalance of data, they have increased the data set to 276,237 pictures by translation, rotation, and mirroring. They examined AlexNet and GoogleNet for two distinct 2D CNN architectures. Its network consisted of four CNN layers. In this network, there are two distinct max pool layers. There are other two layers which are completely connected with each other and finally there is one output layer. Their proposed model led to an accuracy of 87%. Their model was the first example of applying a 3D CNN model in large head CT images to detect acute brain hemorrhage.

Chapter 3

Methodology

3.1 Background Study

Several types of deep learning approaches are used for segmentation and classification of medical images to determine diseases and abnormalities. In our research work, we are interested in deep neural networks like Convolutional Neural Network (CNN), some transfer learning models (VGG16, VGG19) and some machine learning classifiers like Random Forest (RF) and Multi-Layer Perceptron (MLP). Also we've tried to build an explainer model using CNN and Explainable AI.

3.1.1 Convolutional Neural Network (CNN)

The deep learning method specially constructed to process 2D images for deep analysis is known as Convolutional Neural Network (CNN). It can also be used for one and three dimensions. In deep learning architecture, it is referred as the most popular artificial neural network. It is used in data analysis, language processing, and other classification problems. But it is widely used in analyzing and recognizing images because of its high accuracy feature. It analyzes images by detecting patterns using layers that are not visible. These hidden layers of the model are called the convolutional layers. CNN is formed with one input layer, multiple convolutional layers, and non-convolutional layers, and one output layer. The neurons of these layers are connected with the nearby same weighted neurons. The two primary operations of CNN are convolution and pooling. Convolution is the process of filtering the inputs and results in a feature map. These feature maps for different inputs are stacked together, and it provides the output. The pooling operation mainly reduces the number of parameters. This operation is performed in each feature map. In 2012, in the ImageNet competition, a dataset of around a million pictures with 1000 distinct classes was used in this model of CNN; CNN's nearly minimized the error

rates in half compared with the previously best computing approaches [16].
CNN generates assumptions by categorizing each voxel in a picture and assigning voxel-wise segmentation labels based on the local and contextual visual information. At the network's cascaded layers, sequential convolutions of the input with numerous filters are needed to achieve this process. A 3D Deep CNN with multiple scales has been shown to yield excellent performance on brain hemorrhage detection. Each layer has Cl FMs, known as channels. A particular pattern is detected as a group of neurons which is represented by the weights of the kernel which are joined with the FM.

Within a neuron's receptive field are the input ls of the kernel which are joined affect its activity. Its size φl, adds up at each layer l, where Kl, Tl and N3 are vectors are given by the 3D vector. CNN has usually minimized the Cross-Entropy through the cost function by using the following mathematical linear formula:

$$J\left(\Theta; I^i, c^i\right) = -\frac{1}{B}\sum_{i=1}^{B} \log\left(P\left(Y = c^i \mid I^i, \Theta\right)\right) = -\frac{1}{B}\sum_{i=1}^{B} \log\left(pc^i\right) \tag{3.1}$$

After doing the analysis, CNN models provide results. The accuracy of these results depends upon the variety of CNN models that are being used. In the proposed system, three CNN models have been used, and the accuracy of the results is provided according to the comparison that has been made among them. The basic CNN architecture comprises five layers of CNN. For example, Pooling Layer, Convolutional Layer, Dropout, Activation Functions, and Fully Connected Layer. A CNN architecture is formed by stacking these layers.

Furthermore, a CNN has typically been divided into three phases: the first one is convolutional layers with activation, such as ReLU. Then, the second one is a pooling layer for size reduction, typically max-pooling. At last, to classify the features maps, a flatten layer is applied with a last fully connected layer.

3.1.2 VGG -16

VGG-16 is a deep CNN with a total layer count of 16. These 16 layers are combined all together on top of one another to form a deep neural network. Two famous researchers, K. Simonyan and A. Zisserman proposed this model in 2014 [17]. The renowned ILSVRC(Imagenet) competition was won using this model. The number of parameters of this model is very high which is approximately 138 million. Within this architecture, there are different layers like :

- Convolution Layers
- Pooling Layers
- Flat Layer
- Softmax Layer

Though there are a vast number of parameters in different architectures, in this VGG-16 model, there are 3x3 filter convolution layers that shift only one pixel over the input matrix. Moreover, it utilizes same filler and max pool layer of the 2x2 filter which shifted only two pixels over the input matrix [18]. The order of these convolutional and max-pooling layers is maintained throughout the 16 layers of this whole architecture. There are three layers which are fully connected. These fully connected layers are combined together to form a pile of convolutional layers of different depths. The sixteenth layer is the soft-max layer which is used to show the output. The architecture of VGG-16 network is shown Figure 3.1

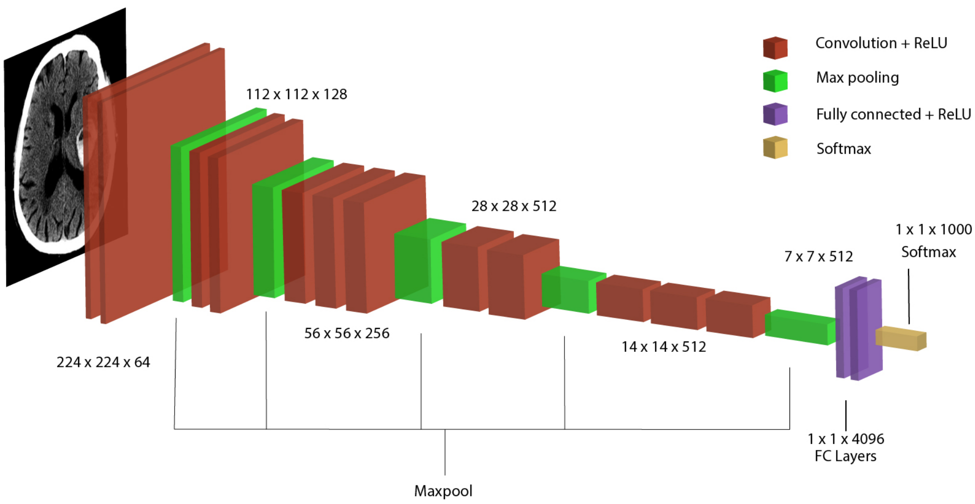

Figure 3.1: VGG-16 Network Architecture

In VGG-16, there is an activation function for the neural network, which is known as Rectified Linear Unit (ReLU). The following equation defines the ReLU function:

$$y = \max(0, x) \tag{3.2}$$

In this equation, the output is y, and the input is x. If the input is positive, it will directly pass through the output; otherwise, the output will be 0. The softmax formula is as follows:

$$\sigma(\vec{z})_i = \frac{e^{z_i}}{\sum_{j=1}^{K} e^{z_j}} \tag{3.3}$$

.

Here,
σ= softmax
z = vector input
K = class count
e^{z_i} = vector input exp. function
e^{z_j}= vector output exp. function

VGG-16 is used for different image classification problems. The only disadvantage is that because of the depth and multiple fully connected nodes, it requires enormous memory space. For this reason, this model is a bit slow to train the dataset, but it is very easy to implement in image classification.

3.1.3 VGG-19

VGG19 is a variation of VGG which has 16 convolution layers, 3 fully connected layers, maxPool layers, and only one SoftMax layer. VGG19 is significantly better than VGG16, but it uses more RAM. The fixed VGG19 is a feature extraction network that has been pre-trained with ImageNet and is utilized efficiently in the VGG19-based fusion architecture. Figure 3.2 shows a diagram of the architecture.

The image module is loaded to preprocess the picture object, and the preprocess input module is imported to appropriately scale pixel values for the VGG19 model. Arrays are processed using the NumPy module. The VGG19 model module is also used to generate a new model that is a subset of the layers in the main VGG19

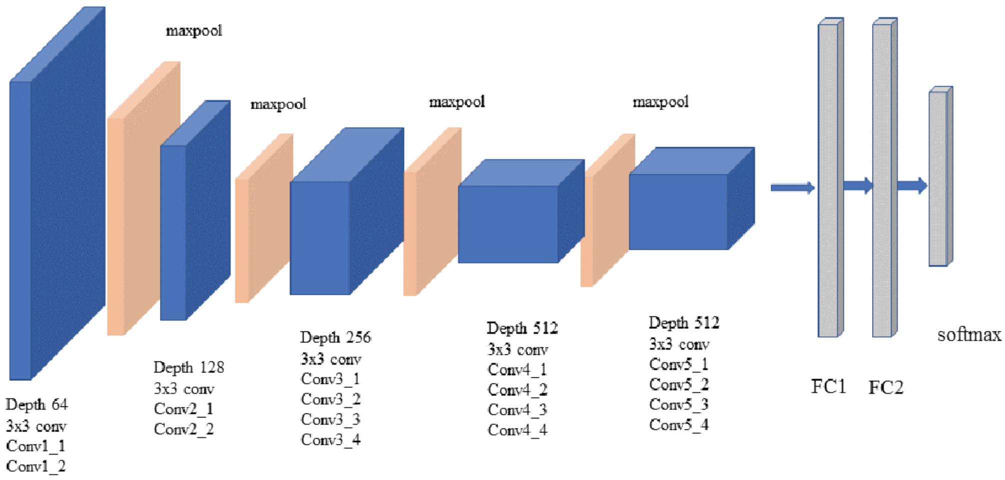

Figure 3.2: VGG-19 Network Architecture [19]

model. The activation of the layer or the feature map would be the output of a specific convolutional layer, as we all know. The pre-trained weights for the ImageNet dataset are then fed into the VGG19 model. In the VGG19 model, one or more dense layers follow a series of convolutional layers. Including the top allows you to choose whether or not you want the final dense layers. The input layer to the last max-pooling layer labeled (7 * 7 * 512) is regarded as the feature extraction element of the model. In contrast, the rest of the network is considered the classification aspect of the model. We must load the input image with the model's expected size after defining the model. This network was given a fixed size (224 * 224) RGB image as input, indicating that the matrix was shaped (224, 224, 3). The image PIL object must then be converted to a NumPy pixel data array. In a NumPy array with three layers of colors (RGB), the image pixels are represented as integers in the range of 0–255. Moreover, we can train our model using the raw pixel range; however, it reduces model accuracy and requires higher training time and data set.

Using two-scale decomposition, the VGG19 network uses the base and intricate portion of the input images (Hui Li et al., 2018). Maximum and average fuse rules are employed to fuse the basis parts. On the other hand, detailed sections are linked by a map derived from the feature maps of the k^{th} layer's j^{th} detailed content. The softmax is used to create the starting weight map (W_j^k). The Softmax formula is as follows:

$$W_j^k = \frac{C_j^k(s,t)}{\sum_{n=1}^{J} C_n^k(s,t)} \tag{3.4}$$

the formula of k layer in VGG19 is determined as follows:

$$\varphi_j^{k,p} = \phi_k \left(I_j^d\right) \tag{3.5}$$

Here,
W_j^k = softmax operator
j = number of activity levels
ϕ_k = layer of VGG19 network where p $\varepsilon\{1,2,3,...,P = 64*2^{k-1}\}$

As a result, the VGG19 is being evaluated in this study for further improvement to improve the proposed methodology's accuracy.

3.1.4 Random Forest (RF)

Random Forest (RF) is a machine learning algorithm that helps classify data using the concept of Decision Trees (DTs). It can perform regression works, as well as classification tasks. It is also widely used in medical image classification. The idea of the Random Forest algorithm is the same as an actual forest. A random forest consists of decision trees that are independent and constructed using random values. The accuracy of prediction is based on the number of decision trees. If there are more decision trees in the forest, then the accuracy of prophecy will be higher. The forests that are created can be preserved and used on other data in the future. A random forest algorithm is used to impute the missing values from a dataset. It can also handle a large number of datasets. Because of its easy and flexible feature, it is used in classification and regression problems. It determines the importance of several variables in the categorization. According to a journal article [20], testing a dataset of 2400 medical images, Random Forest showed an average precision of 93.10% in classifying those images, which is 4% higher than the MSVM method. The architecture diagram is presented in Figure 3.3

The random forest training method uses the standard approach of bootstrap aggregation. Suppose there is a dataset D and multiple decision trees like DT1, DT2, DT3,. . . .,DTN. From the dataset D, different samples from the dataset like S1, S2, S3,. . . .,SN will be given to DT1, DT2, DT3,. . . .,DTN decision trees. During the training period, different decision trees will result in different accuracy levels. After that, the mean of the accuracy will be calculated, which will be the accuracy level of the random forest. Depending on the accuracy level, random forest predicts the

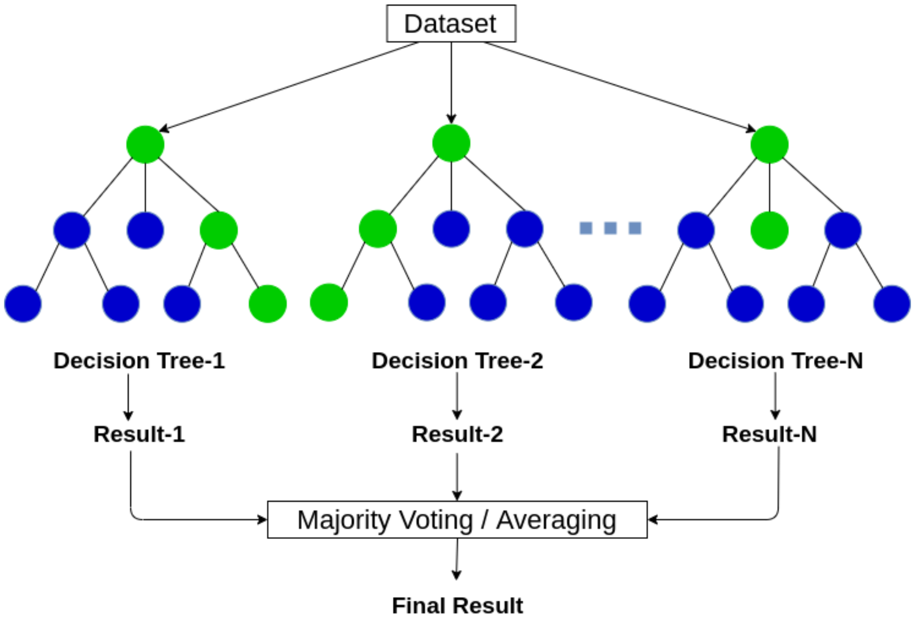

Figure 3.3: Random Forest [21]

best result from the decision trees.
After training, summing the predictions from all the various regression trees on x':

$$f = \frac{1}{N} \sum_{n=1}^{N} f\left(x'\right) \tag{3.6}$$

Here,
f = Final prediction
x' = Unseen sample
N = Number of bootstrap pattern repetition (n = 1, 2, 3 . . . . , N)

3.1.5 Multi-layer Perceptron (MLP)

MLP is a type of ANN model which belongs to the feedforward network category. Frank Rosenblatt developed the perceptron algorithm at the Cornell Aeronautical Laboratory in 1958 [22]. It is a supervised learning model which generates a set of outputs from the collection of different inputs. It uses a backpropagation method to train the network. This backpropagation method helps calculate a loss function's gradient concerning all of the network's weights. Multi-layer perceptron (MLP) is applied in stock analysis, image identification, spam detection, and election vote predictions.

In the Multilayer perceptron method, there are three layers which are given below -

- Input Layer
- Hidden Layer
- Output Layer

By using these layers, the interconnected nodes transfer information to each other. The input layer takes an input that will produce output and pass through the output layer. This network must have at least one hidden layer to perform computation and different operations on the input data. Multilayer Perceptron process architecture is shown in the below Figure 3.4

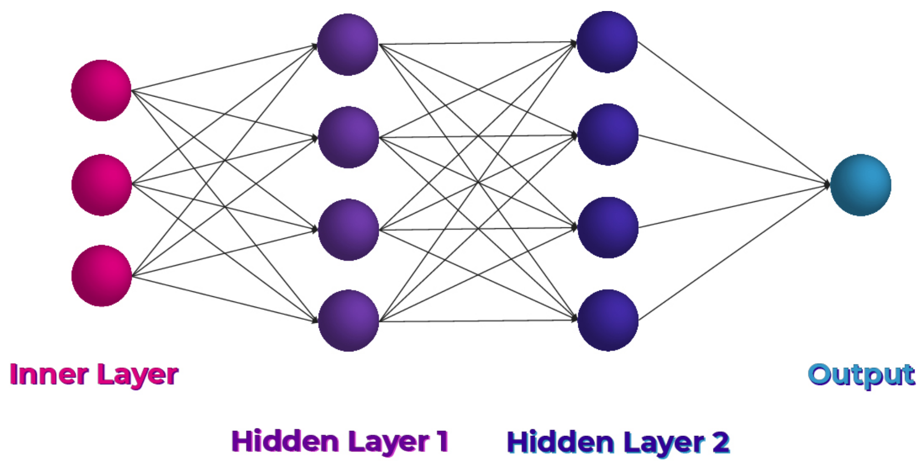

Figure 3.4: Multilayer Perceptron Architecture

Different weighted values are assigned to each layer, which is the central concept of the Multi-layer perceptron process. The equation which represents each layer is given below:

$$y = f(WxT + b) \tag{3.7}$$

Here input vector is x, the set of assigned weights to each layer is W, b is the biased vector and f is the activation function similar to the above mentioned ReLU function.

The output of the MLP network can be described by the following equation -

$$\begin{aligned} y_{\mathrm{k}}^{0} &= f_{\mathrm{k}}^{0}\left(b^{0}{}_{\mathrm{k}} + \sum_{i=1}^{S} w_{\mathrm{ik}}^{0} y_{\mathrm{i}}^{\mathrm{h}}\right) \\ &= f_{\mathrm{k}}^{0}\left(b_{\mathrm{k}}^{0} + \sum_{i=1}^{S} w_{\mathrm{ik}}^{0} f_{\mathrm{i}}^{\mathrm{h}}\left(b_{\mathrm{i}}^{\mathrm{h}} + \sum_{j=1}^{N} w_{\mathrm{ji}}^{\mathrm{h}} x_{\mathrm{i}}\right)\right) \end{aligned} \tag{3.8}$$

Here k = 1 , ... L

In the above equation [23], h indicates the hidden layer elements and o signifies the output layer's elements. Moreover, $b_{\mathrm{i}}^{\mathrm{h}}$ indicates the bias of nodes i in the hidden layer and $b_{\mathrm{k}}^{\mathrm{o}}$ is the bias of the nodes k of the output layer. Here, w indicates the weight connected to the nodes and f represents the transfer function. The Multilayer Perceptron process works well with large input data and can achieve the same accuracy ratio even with small input datasets. With the MNIST dataset, a lightweight MLP (2–3 layers) may easily achieve high accuracy. The only disadvantage of this model is the huge number of parameters.

3.1.6 Explainable AI (Lime)

Explainable AI was created to explain any "Blackbox" ML model in which data points are supplied and a prediction is formed.Usually we have no idea what happens inside the black box until explainable AI was a thing. Lime is one of the most famous Explainable Ai techniques proposed by Ribeiro et al., which is a novel explanation approach that can explain the reason of prediction of any classifier by learning the used model locally around the prediction [24]. The author claimed that lime can demonstrate the flexibility of any machine learning or deep learning model related to text data or image classification using deep neural networks.
When it comes to understanding the basis of prediction for any machine learning model, LIME can be a very good solution. LIME examines the prediction of a black box machine learning model and generates new data which contains perturbed samples and the prediction itself for that particular model. LIME works for tabular, text and image data. For image data, LIME generates super-pixels with positive weights for the predicted classes which is the region of the image for which the decision was predicted by the model. In the paper [24], the author used a pretrained Inception model and showed the superpixel explanation of the top three classes for an image. While generating explanations with LIME, they only kept the superpixel for the prediction and the rest parts were grayed out.

Figure 3.5: Original Image

(a) Electric Guitar (b) Acoustic Guitar (c) Labrador

Figure 3.6: Explanation of the top three classes using LIME [24]

3.2 Dataset

For our research, we have collected a brain CT hemorrhage dataset [25] from kaggle which is a set of 82 patient's brain CT scans and they are collected from the Al Hilla Teaching Hospital, Iraq. This dataset contains a total of 5002 computed tomography head images in an equal ratio of brain and bone images. We are interested in the brain region so we have taken only the brain portion of the dataset which contains 2501 images. Among them, 2183 images are normal CT scan images (no hemorrhage) and 318 images are hemorrhagic CT scan images. The hemorrhagic images are of five classes: Intraventricular, intraparenchymal, subarachnoid, epidural and subdural. Some of the images had multiple classes of hemorrhage while most of them had only one class. To reduce complexity of classification, we discarded the images with multiple hemorrhage classes and at last ended with 2476 images where 293 images were hemorrhagic. The collected CT scan images have a dimension of 650 x 650. All the CT head images are in .jpg format. The class count in the dataset and sample image from dataset are given below in Figure 3.7 and Figure 3.8.

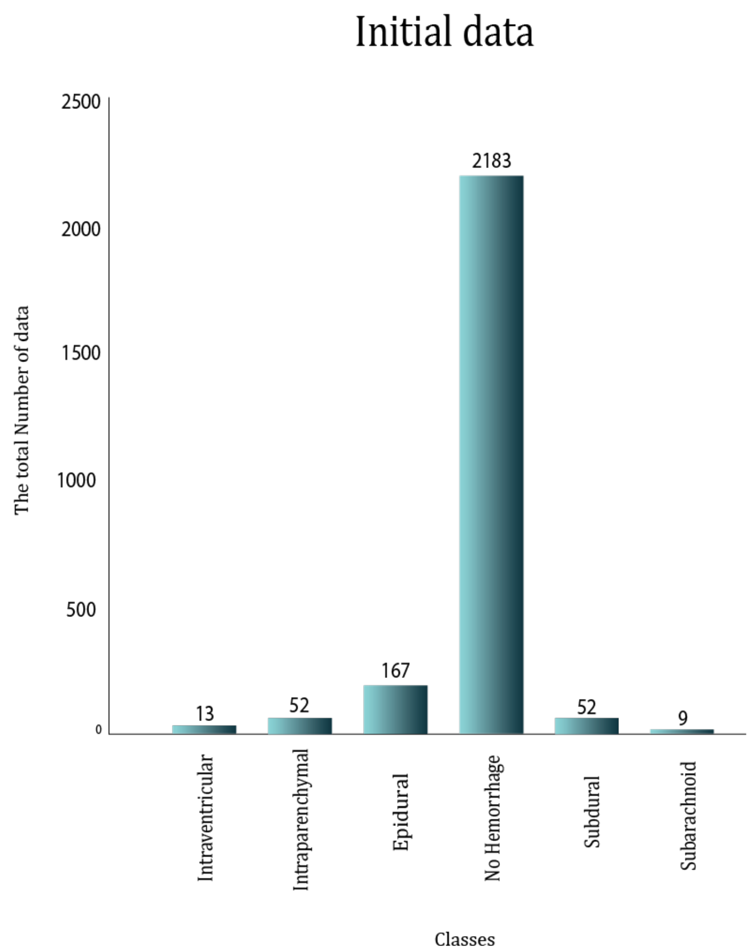

Figure 3.7: Dataset Class Count

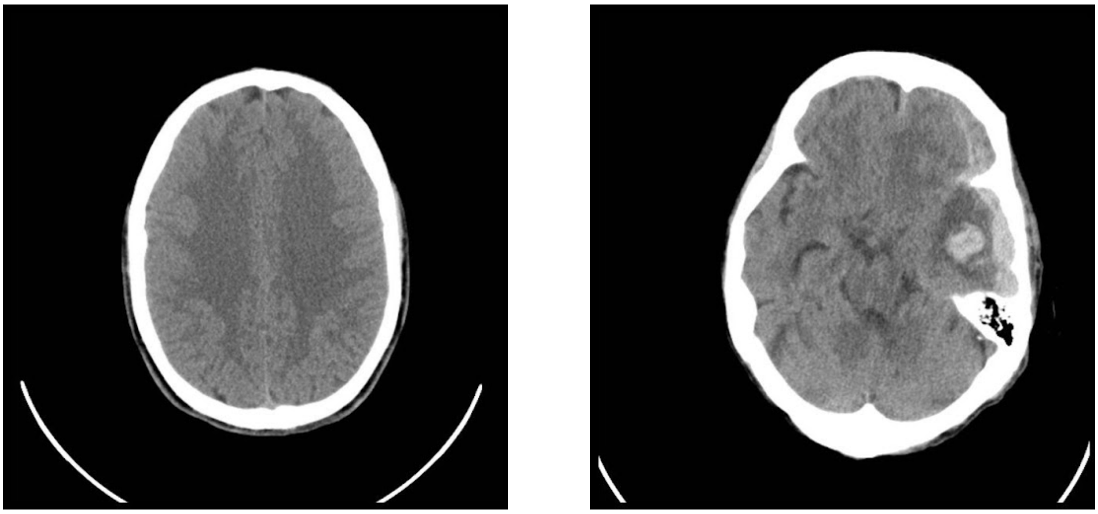

Figure 3.8: Sample of Dataset (a) Non-hemorrhagic brain (b) Hemorrhagic brain

3.3 Research Workflow

In our proposed method, we will use different machine learning models to determine and classify neurological disorders (brain hemorrhage). CNN architectures, random forest, Multi-layer Perceptron (MLP) algorithms will be used in our research. To design the best possible model, choosing the correct model space given the training data is essential. We have collected CT brain image datasets from Kaggle. We will at first pre-process the data and balance the data classes using augmentation and “Smote”. Features will be extracted from the images using pre-trained transfer learning models (VGG16/VGG19) and Random forest and MLP classifiers will be used to test the data and calculate the accuracy of the model. If our model gives less accuracy, we will modify our model until we receive our desired accuracy. After we reach a satisfactory result, we will compare our model accuracy with different existing algorithms to find the accuracy level of our model.
In the below Figure 3.9, our research's workflow diagram is shown:

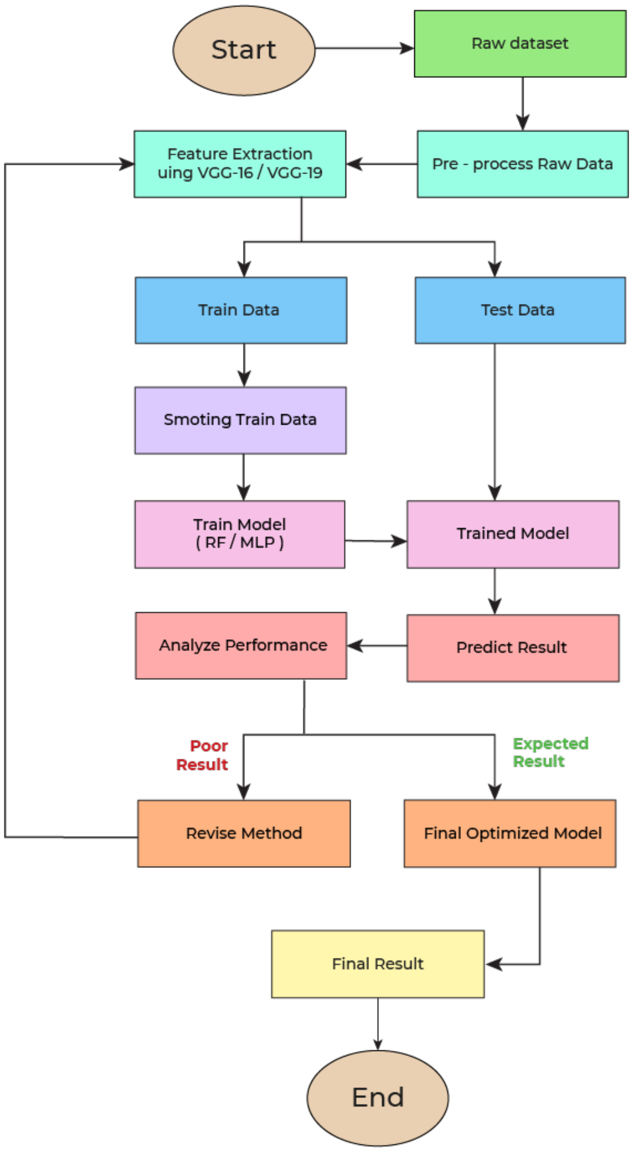

Figure 3.9: Workflow Diagram

3.4 Data Preprocessing

3.4.1 Resizing Image

Our collected images have a dimensions of 650 x 650. For VGG models, the recommended image size is 228x228 as the VGG models are originally trained on 228x228 images. Another problem with using high resolution images is, it makes the overall process very slow and resource hungry. So at the time of reading the images, we have resized them to 228 x 228.

3.4.2 Augmentation

If we see the class count in the dataset, we can see that the positive image percentage is 11.83% and the negative image percentage is 88.17%. This is a huge imbalance in the data. Also if we see the hemorrhagic portion in the image dataset, the "Subarachnoid" class only has 9 images and the "Intraventricular" class has only 13 images. This amount of data is not sufficient to conduct a good training for any machine learning classifier. So we used augmentation to increase the count of data for the positive classes in the dataset. We have used two parameters for aumanting per image, they are: random 45% rotation and a random zoom level in range of 80% to 120%. Using these parameters, we generated 7 images for every positive image in the dataset.
The overall process steps are shown in the below figure.

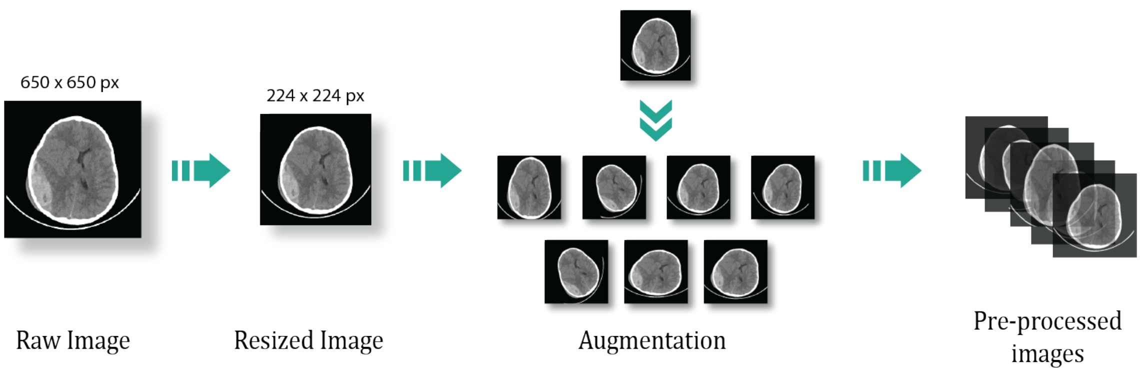

Figure 3.10: Image Pre-processing

3.5 Feature Extraction

VGG16 and VGG19 were used for the feature extraction part. To extract the features from the dataset, we applied these algorithms one at a time. After that we checked accuracy using different machine learning algorithms.
While extracting the features using VGG16 and VGG19 we have used pre-trained weights from imagenet and we have also turned off trainable parameters for every layer in the models. Because these two models will be used for feature extraction instead of training and prediction. While initializing the models the input shape was set to (224,224,3) for both models and 14,714,688 parameters were generated by VGG16 whereas 20,024,384 parameters were generated by VGG19. After running the models we extracted "25,088" features separately using VGG16 and VGG19 for every single image. These features will be used for training and testing for Multi-layer Perceptron and Random forest classifiers.

3.6 Train - Test Split

The features we extracted from the VGG16 or VGG19 models both were divided into 80:20 ratio for train and test. We utilized a built in function of sklearn, named as train-test-split for dividing the extracted features. We used 80% of the extracted features to train the machine learning models and for the prediction task, we used the rest 20%. The data count after train-test-split is displayed in Figure 3.11.

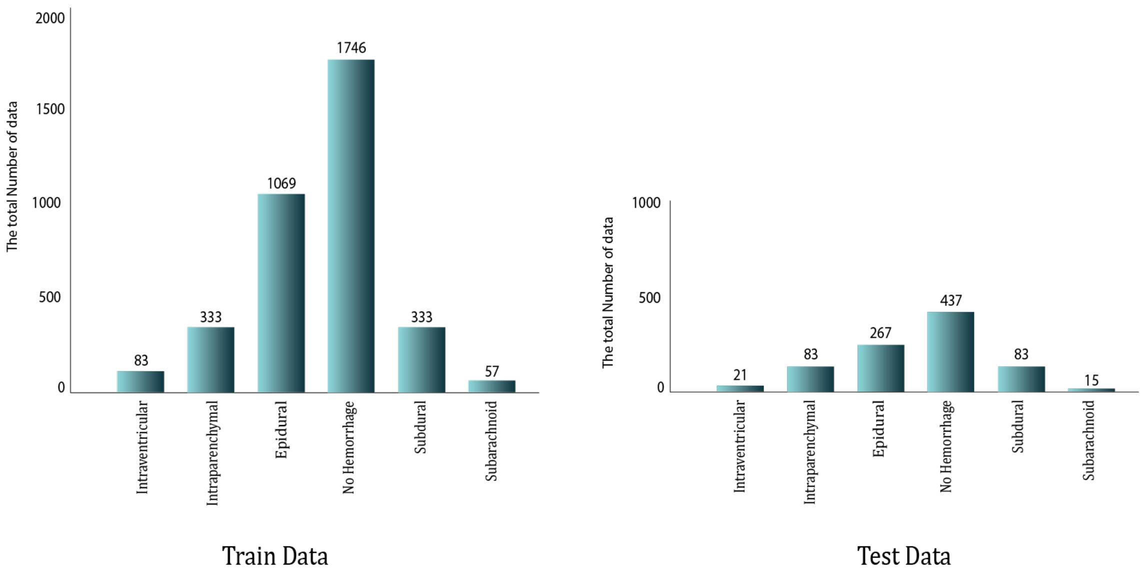

Figure 3.11: Train and Test Data Amount

3.6.1 SMOTE Train Data

After the split, the training data is still imbalanced. We can see that The non-hemorrhagic class data count is still very high compared to the classes like intraventricular, subdural or subarachnoid. So to make the training data completely balanced, we conducted data SMOTE (Synthetic Minority Oversampling Technique) on the training dataset. SMOTE will generate more data from the currently available data from each class and make the data count the same for all the classes.

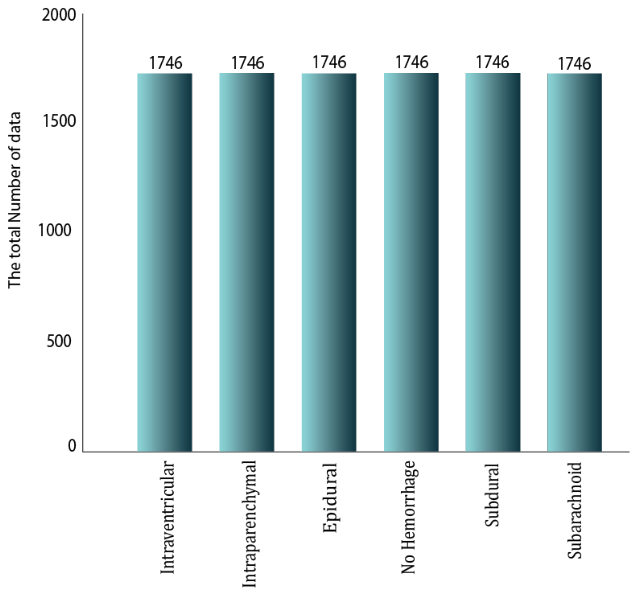

Figure 3.12: Training Data After Applying SMOTE

Chapter 4

Model Implementation

The implementation of the proposed model for brain hemorrhage detection is defined in this section. The model's implementation consists of several phases like processing input data, deriving features, classification, and checking. Processing input data is a stage for qualifying the brain hemorrhage dataset to be used as an input to the detection.

For the model implementation, in order to extract the features we used two previously trained models: VGG16 and VGG19. The features and weights of images are extracted and they are transferred to the ML models. Moreover, we used Random Forest (RF) and Multi-layer Perceptron (MLP) models to fit the features and derive the accuracy rate or prediction. Thus a hybrid approach was made by implementing the above mentioned models.

This chapter also includes the outcomes of the proposed model's implementation for detecting brain hemorrhage.

4.1 Workflow Overview

We have followed some steps to build the best model for predicting brain hemorrhage and calculate the accuracy correctly. The overview of our research is shown in this section:

- At first, we resized the CT brain images from (650 x 650) to (228 x 228) and performed minor preprocessing on our dataset.
- The the hemorrhagic images were augmented with random 45° rotation and random (80-120)% zoom range. For every hemorrhagic image, 7 more augmented images were generated.

- Feature selection and feature extraction have been performed by using VGG-16 and VGG-19 models. We used them one by one to derive the features from the images, these features are then used for training the Multilayer Perceptron and Random Forest classifier models.
- We have split the extracted features into 80:20, where we have used 80% of our features to train the models and 20% test data were used to test the accuracy.
- We applied SMOTE on train data to balance the classes in the data.
- The ML models then were trained by the training data.
- At last, we calculated the accuracy of our system using the testing data.

4.2 Results

After training the machine learning models, we predicted the result for test data using each classifier. While extracting the features using VGG16, the MLP classifier showed 97.24% accuracy and Random Forest showed 92.72% accuracy. And when we extracted the features using the VGG19 model, we got 97.02% accuracy from the MLP classifier and 94.59% accuracy from Random Forest. The accuracy scores are given in the below table and also accuracy comparison is shown in the table and in the bar chart 4.1.

	MLP Classifier	**Random Forest**
VGG-16 Feature Extraction	97.24%	92.72%
VGG-19 Feature Extraction	97.02%	94.59%

Table 4.1: Accuracy Scores

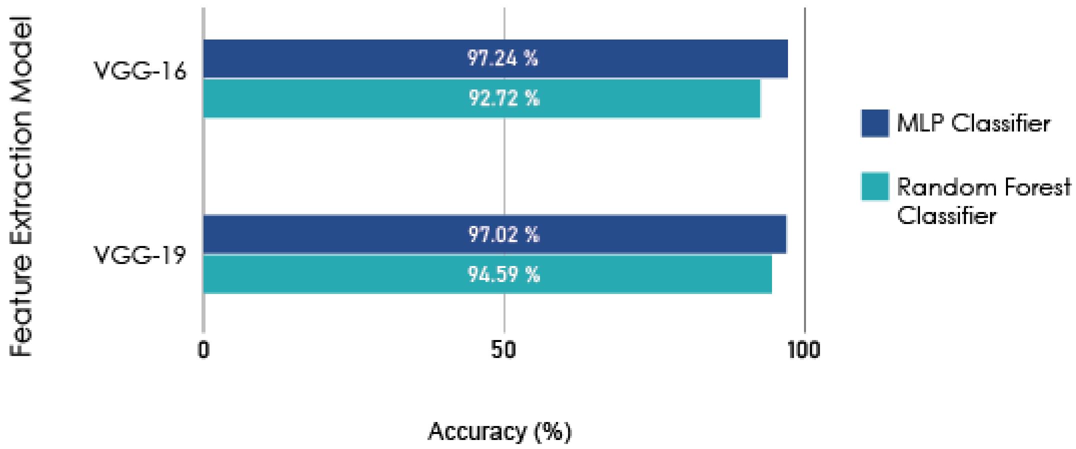

Figure 4.1: Accuracy Scores

The following figures are the confusion matrices of all the approaches to have a better idea about the predictions.

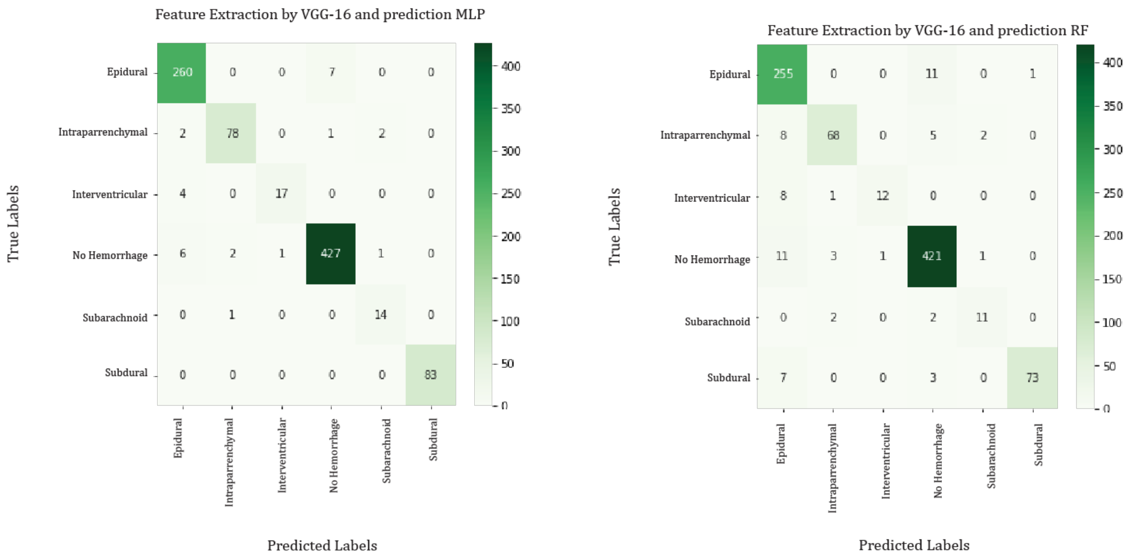

Figure 4.2: Feature Extraction by VGG-16 and Prediction by MLP Random Forest (Multiclass)

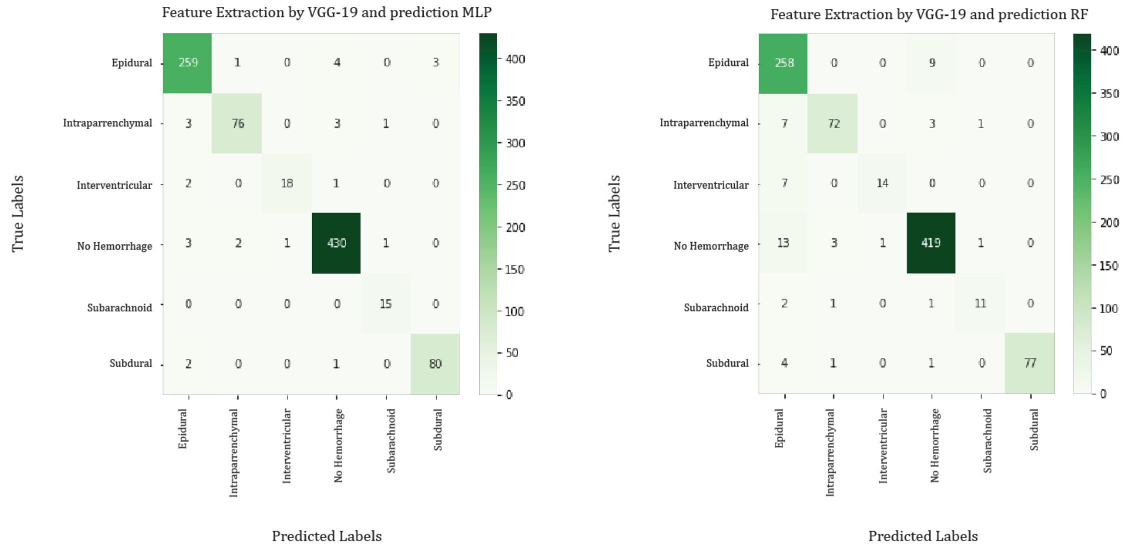

Figure 4.3: Feature Extraction by VGG-19 and Prediction by MLP Random Forest (Multiclass)

The best accuracy is obtained from the VGG16-MLP model. This model scored a 97.24% accuracy.

To evaluate the performance of our approach, we have compared our result with a previous paper [15]. In this paper they used three 3D CNN architectures to predict the brain hemorrhage. In the table below, we have shown the accuracy of 3D CNN approach vs our hybrid approach in the same field.

Model	**Precision**	**Recall**	**F1-Score**
3D CNN (Previous Approach)	0.80	0.77	0.78
VGG16-MLP	0.95	0.96	0.97
VGG16-RF	0.93	0.84	0.92
VGG19-MLP	0.95	0.95	0.97
VGG19-RF	0.94	0.88	0.94

Table 4.2: Accuracy Comparison

From the table, we can see that our every hybrid approach outperforms the previously applied 3D CNN approach

4.2.1 Lime Explanation

We have used a CNN model using three convolutional layers, two max pooling layers and two dense layers and trained the model to explain with LIME. The intention behind this was to see if we can extract the region for which a particular type of hemorrhage was predicted. We were successfully able to generate the correct superpixels of the region of interest and generated an output image with the area which includes the region of interest (ROI) for an image.
In the explanation below, we showed three images for every explanation. The first one is the original image, the second image has the marked hemorrhagic region and the third one is generated using the LIME image explainer.

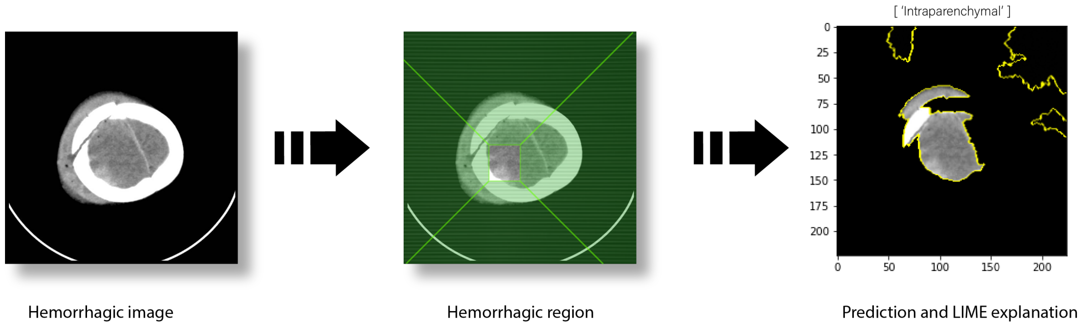

Figure 4.4: Explaining Intraparenchymal Class with LIME

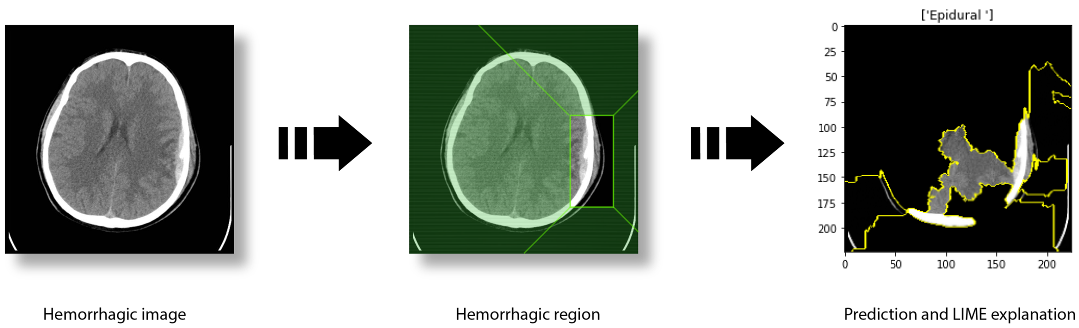

Figure 4.5: Explaining Epidural Class with LIME

Using LIME, the correct region of the image can be extracted. However, the explanation is not always perfect. For example, often extra regions are generated alongside the positive region. Also sometimes, it is completely outside of the positive region. This is shown in the examples below.

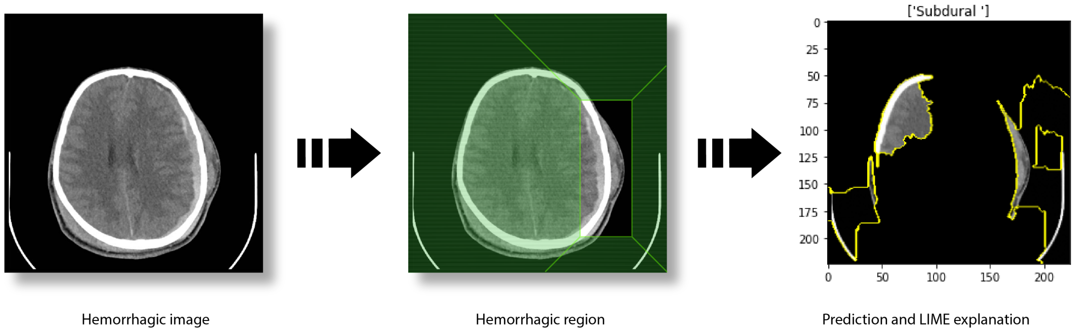

Figure 4.6: Explaining Subdural Class with LIME

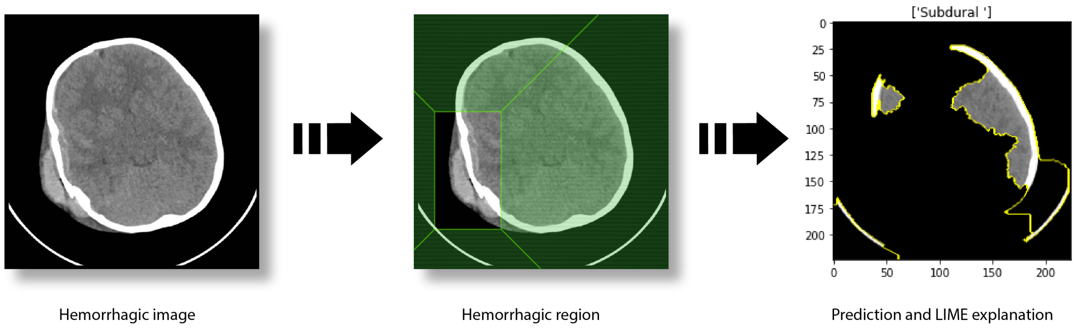

Figure 4.7: Example of Prediction from Completely Wrong Region

So, as we can see the region is not always perfect and precise so the result is not good enough to be used as a positive region extractor from any image prediction. However, this can be a good tool to understand and visualize the region behind a particular image classification or class prediction for any image classifier.

Chapter 5

Conclusion

In this paper, we showed that instead of applying any single model if we go to a hybrid approach we can determine brain disorders like hemorrhage more accurately. Our proposed model first derives features from the data using one model and then training and detection is conducted using other models. Also we tried to build a system which can extract the region of interest(ROI) from an image using LIME. Brain injury prediction is still a very complicated and lengthy process. The rate of brain injury will rise with time and computational burden will increase with it. In a report from Hossain et al. (2018) [26], head injury occurrence in Bangladesh is around 814 per 100000 people annually, having a mortality rate of approx 24 per 100000 people. TBI has a fatality rate of 30 per 100,000, resulting in an estimated 50,000 fatalities in the United States per year. The number of brain diseases is huge throughout the world and the complexity will increase with time. Different models of ML, and deep learning are used to predict brain disorders. Our proposed model can predict and classify hemorrhage at a 97.24% (VGG16-MLP) accuracy rate. Also using LIME we showed the region of image for which a particular prediction was triggered. These approaches can be a pathway to future studies for building more hybrid machine learning models in detecting other brain disorders and abnormalities. In future, we want to research more on the medical image classification domain and intend to build a robust system which can predict and classify disease more accurately and quickly and also can extract the exact ROI from the image. Also, we want to build a hybrid classifier for brain images which can predict any brain abnormalities from a given image.

Bibliography

[1] C. Dilmeganil, *41 statistics, facts forecasts on machine learning [2021]*, https://research.aimultiple.com/ml-stats/, Oct. 2020.

[2] M. C. Dewan, A. Rattani, S. Gupta, *et al.*, "Estimating the global incidence of traumatic brain injury," *Journal of neurosurgery*, vol. 130, no. 4, pp. 1080–1097, 2018.

[3] T. Steiner and E. Jüttler, *American guidelines for the management of spontaneous intracerebral hemorrhage in adults: European perspective*, 2008.

[4] H. Mohsen, E.-S. A. El-Dahshan, and A.-B. M. Salem, "A machine learning technique for mri brain images," in *2012 8th International Conference on Informatics and Systems (INFOS)*, IEEE, 2012, BIO–161.

[5] M. A. Balafar, A. R. Ramli, M. I. Saripan, and S. Mashohor, "Review of brain mri image segmentation methods," *Artificial Intelligence Review*, vol. 33, no. 3, pp. 261–274, 2010.

[6] V. L. Feigin, E. Nichols, T. Alam, *et al.*, "Global, regional, and national burden of neurological disorders, 1990–2016: A systematic analysis for the global burden of disease study 2016," *The Lancet Neurology*, vol. 18, no. 5, pp. 459–480, 2019.

[7] WHO, *Neurological disorders affect millions globally: Who report*, https://apps.who.int/mediacentre/news/releases/2007/pr04/en/index.html, Feb. 2007.

[8] A. Sage and P. Badura, "Intracranial hemorrhage detection in head ct using double-branch convolutional neural network, support vector machine, and random forest," *Applied Sciences*, vol. 10, no. 21, p. 7577, 2020.

[9] S. Pereira, R. Meier, R. McKinley, *et al.*, "Enhancing interpretability of automatically extracted machine learning features: Application to a rbm-random forest system on brain lesion segmentation," *Medical image analysis*, vol. 44, pp. 228–244, 2018.

[10] J. Mitra, K.-k. Shen, S. Ghose, *et al.*, "Statistical machine learning to identify traumatic brain injury (tbi) from structural disconnections of white matter networks," *NeuroImage*, vol. 129, pp. 247–259, 2016.

[11] K. Kamnitsas, C. Ledig, V. F. Newcombe, *et al.*, "Efficient multi-scale 3d cnn with fully connected crf for accurate brain lesion segmentation," *Medical image analysis*, vol. 36, pp. 61–78, 2017.

[12] G. Wang, W. Li, S. Ourselin, and T. Vercauteren, "Automatic brain tumor segmentation based on cascaded convolutional neural networks with uncertainty estimation," *Frontiers in computational neuroscience*, vol. 13, p. 56, 2019.

[13] D. T. Nguyen, S. Ryu, M. N. I. Qureshi, M. Choi, K. H. Lee, and B. Lee, "Hybrid multivariate pattern analysis combined with extreme learning machine for alzheimer's dementia diagnosis using multi-measure rs-fmri spatial patterns," *PloS one*, vol. 14, no. 2, e0212582, 2019.

[14] E. I. Zacharaki, S. Wang, S. Chawla, *et al.*, "Classification of brain tumor type and grade using mri texture and shape in a machine learning scheme," *Magnetic Resonance in Medicine: An Official Journal of the International Society for Magnetic Resonance in Medicine*, vol. 62, no. 6, pp. 1609–1618, 2009.

[15] K. Jnawali, M. R. Arbabshirani, N. Rao, and A. A. Patel, "Deep 3d convolution neural network for ct brain hemorrhage classification," in *Medical Imaging 2018: Computer-Aided Diagnosis*, International Society for Optics and Photonics, vol. 10575, 2018, p. 105751C.

[16] K. Kamnitsas, C. Ledig, V. F. Newcombe, *et al.*, "Efficient multi-scale 3d cnn with fully connected crf for accurate brain lesion segmentation," *Medical image analysis*, vol. 36, pp. 61–78, 2017.

[17] M. ul Hassan, *Vgg16 – convolutional network for classification and detection*, https://neurohive.io/en/popular-networks/vgg16/, Nov. 2018.

[18] R. Thakur, *Step by step vgg16 implementation in keras for beginners*, https://towardsdatascience.com/step-by-step-vgg16-implementation-in-keras-for-beginners-a833c686ae6c, Aug. 2019.

[19] R. Hewage, *Extract features, visualize filters and feature maps in vgg16 and vgg19 cnn models*, May 2020. [Online]. Available: https://towardsdatascience.com/extract-features-visualize-filters-and-feature-maps-in-vgg16-and-vgg19-cnn-models-d2da6333edd0 (visited on 07/16/2021).

[20] B. C. Ko, S. H. Kim, and J.-Y. Nam, "X-ray image classification using random forests with local wavelet-based cs-local binary patterns," *Journal of digital imaging*, vol. 24, no. 6, pp. 1141–1151, 2011.

[21] A. Sharma, *Decision tree vs. random forest - which algorithm should you use?* May 2020. [Online]. Available: https://www.analyticsvidhya.com/blog/2020/05/decision-tree-vs-random-forest-algorithm/?__cf_chl_f_tk=Y.0Qc0O.RjFETGjGTdxzE4kXEpg6ARwvc0xgBNUM690-1642345633-0-gaNycGzNCL0 (visited on 07/27/2020).

[22] Wikipedia, *Perceptron*, https://en.wikipedia.org/wiki/Perceptron, Aug. 2021.

[23] E. Agirre-Basurko, G. Ibarra-Berastegi, and I. Madariaga, "Regression and multilayer perceptron-based models to forecast hourly o3 and no2 levels in the bilbao area," *Environmental Modelling & Software*, vol. 21, no. 4, pp. 430–446, 2006.

[24] M. T. Ribeiro, S. Singh, and C. Guestrin, ""why should I trust you?": Explaining the predictions of any classifier," in *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, San Francisco, CA, USA, August 13-17, 2016*, 2016, pp. 1135–1144.

[25] M. Hssayeni, *Computed tomography images for intracranial hemorrhage detection and segmentation*, Mar. 2020. [Online]. Available: https://doi.org/10.13026/4nae-zg36.

[26] M. J. Hossain, J. Ferdoush, R. A. Shawon, S. R. Mashreky, and A. F. Rahman, "Pw 2842 epidemiology of traumatic head injury-nationally conducted survey findings from bangladesh," *Injury Prevention*, vol. 24, no. Suppl 2, A195, 2018.