**HEALTH ROOF**

**ALL IN ONE SOLUTION**

**A CAPSTONE PROJECT REPORT**

*Submitted in partial fulfillment of the*

*requirement for the award of the*

*Degree of*

**BACHELOR OF TECHNOLOGY**

**IN**

**COMPUTER SCIENCE AND ENGINEERING**

*by*

**S.LAKSHMI PRASANNA (19BEC7317)**

**Y.AVINASH MANI KIRAN (19BCE7466)**

**M.AMULYA (19BCE7752)**

*Under the Guidance of*

**DR. SUMATHI**



SCHOOL OF COMPUTER SCIENCE AND ENGINEERING

VIT-AP UNIVERSITY

AMARAVATI- 522237

*DECEMBER 2022*

**CERTIFICATE**

This is to certify that the Capstone Project work titled “**HEALTH ROOF – ALL IN ONE SOLUTION**” that is being submitted by **S.LAKSHMI PRASANNA (19BCE7317) , Y.AVINASH MANI KIRAN (19BCE7466) , M.AMULYA (19BCE7752)** is in partial fulfillment of the requirements for the award of Bachelor of Technology, is a record of Bonafede work done under my guidance. It is certified that neither the entirety nor any portion of the information contained in this project work has ever been obtained from another source or submitted to a different institute or university for the purpose of receiving a degree or diploma.

Dr. SUMATHI

Guide

**The thesis is satisfactory / unsatisfactory**

**Internal Examiner External Examiner**

**Approved by**

**PROGRAM CHAIR DEAN**

B. Tech. ECE School Of Electronics Engineering

**ACKNOWLEDGEMENTS**

Any project's success, aside from our own efforts, greatly depends on the support and directions of many people. We would want to take this chance to express our gratitude to everyone who helped us complete this project successfully.

We would like to use this occasion to show our sincere appreciation and respect for our mentor, Assistant Professor Dr. Sumathi, for her outstanding direction, close supervision, and unwavering encouragement during the project effort. The blessings, assistance, and direction she provided will go a long way toward assisting us on the upcoming journey through life.

We would also want to take this occasion to thank Dr. G. Viswanathan, Vice Chancellor Dr. S.V. Kota Reddy, and Dr. Madhusudhana Rao.N,for enabling us to work there and for serving as our motivation during the course, School of Computer Engineering.

We also want to express our gratitude to the instructors and staff for their encouragement and assistance in getting this project done successfully. This encouraged us to learn the necessary material so that we could effectively finish our course work.

It gives us great pleasure to thank our friends for their support and persuasiveness in getting us to start and finish this project.

Finally, we would like to take this occasion to thank our family for their unwavering support, without which we could not have finished this task. We truly appreciate and thank everyone who contributed, both directly and indirectly, to the success of this initiative.

Place: Amaravati

Date: December 2022

**ABSTRACT**

Due to their lifestyle choices and the state of the environment today, individuals are susceptible to many diseases. To prevent such diseases from becoming severe, early detection and prediction of these disorders are crucial. Most of the time, it is challenging for doctors to appropriately diagnose ailments by hand. This study aims to detect and forecast patients with more typical illnesses. This might be done by making sure that this categorization accurately identifies people with diseases using a cutting-edge machine learning technique. Another difficult task is illness forecasting. Therefore, in this study, we created the disease detection website "Health Roof - All in One Solution" to address this issue. In predicting diseases, this is crucial. The suggested method uses machine learning algorithms like convolutional neural network (CNN) for automatic feature extraction and disease prediction to provide a wide disease prediction based on patient complaints. The data set was prepared by gathering disease symptoms, a person's lifestyle, and information on doctor visits. These factors are all taken into account in this general disease prediction.

An all-inclusive medical solution is Health Roof. Using the power of AI, this project unifies the diagnosis of 7 diseases onto a single platform. Based on the reports and data supplied by users, we are using this project to identify numerous ailments. Covid 19 Detection, Brain Tumor Detection, Breast Cancer Detection, Alzheimer's Detection, Diabetes Detection, Pneumonia Detection, and Heart Disease Detection are some of the diseases we are able to identify utilizing this initiative. The key benefit of this project is that we can quickly and easily receive the test results at home.

Our model had an accuracy of 73.54% in detecting Alzheimer's, around 100% in detecting brain tumours, 91.81% in detecting breast cancer, 93% in detecting Covid-19, 66.8% in detecting diabetes, 86.96% in detecting heart disease, and 83.17% in detecting pneumonia.

**TABLE OF CONTENTS**

|  |  |  |  |
| --- | --- | --- | --- |
| **S.No.** | **Chapter** | **Title** | **Page Number** |
| **1.** |  | **Acknowledgement** | **2** |
| **2.** |  | **Abstract** | **3** |
| **3.** |  | **List of Tables** | **6** |
| **4.** |  | **List of Figures** | **7** |
| **5.** |  | **List of Abbreviations** | **9** |
| **4.** | **1**  **1.1**  **1.2** | **Introduction**  **Objectives**  **Background and Literature Survey** | **10**  **11**  **12** |
| **5.** | **2**  **2.1**  **2.2**  **2.3**  **2.4** | **Health Roof – All in One Solution**  **Flowchart**  **Working Methodology**  **Method – 1 - Image Pre-Processing**  **Method – 2 - Data Pre-Processing** | **14**  **14**  **15**  **16**  **20** |
| **6.** | **3**  **3.1**  **3.2**  **3.3**  **3.4**  **3.5** | **Algorithms**  **Convolutional Neural Networks(CNN)**  **XGBoost**  **Machine Learning Bagging & Random Forest**  **Random Forest**  **VGG-16** | **24**  **24**  **32**  **36**  **28**  **40** |
| **7.** | **4**  **4.1**  **4.2**  **4.3**  **4.4**  **4.5**  **4.6**  **4.7** | **Disease**  **Alzheimer**  **Brain Tumor**  **Breast Cancer**  **Covid-19**  **Diabetes**  **Heart Disease**  **Pneumonia** | **42**  **42**  **42**  **43**  **44**  **44**  **45**  **46** |
| **8.** | **5**  **5.1**  **5.2** | **Requirements & Feasibility**  **Requirements**  **Feasibility** | **47**  **47**  **48** |
| **9.** | **6**  **6.1**  **6.2**  **6.3**  **6.4**  **6.5**  **6.6**  **6.7** | **Results & Discussions**  **Detected Results**  **Alzheimer Result**  **Brain Tumor Result**  **Breast Cancer Tumor Result**  **Covid-19 Result**  **Diabetes Result**  **Heart Disease Result**  **Pneumonia Result** | **49**  **50**  **50**  **51**  **52**  **53**  **53**  **55**  **56** |
| **10.** | **7** | **Conclusion & Future Work** | **57** |
| **11.** | **8** | **Appendix** | **58** |
| **12.** | **9** | **References** | **69** |

**List of Tables**

|  |  |  |
| --- | --- | --- |
| **Table No.** | **Title** | **Page no.** |
| **1** | **Diseases-Methodologies** | **15** |
| **2** | **Diseases-Accuracies** | **49** |

**List of Figures**

|  |  |  |
| --- | --- | --- |
| **Fig No.** | **Title** | **Page no.** |
| **1** | **Annual Number Of Deaths by cause** | **10** |
| **2** | **System Flow chart** | **14** |
| **3** | **Methodology 1 - Image Pre-Processing Architecture** | **16** |
| **4** | **Input Form for Alzheimer Detection** | **18** |
| **5** | **Input Form for Brain Tumor Detection** | **18** |
| **6** | **Input Form for Covid - 19 Detection** | **19** |
| **7** | **Input Form for Pneumonia Detection** | **19** |
| **8** | **Methodology 2 – Data Pre-Processing Architecture** | **20** |
| **9** | **Input Form for Breast Cancer Detection** | **22** |
| **10** | **Input Form for Diabetes Detection** | **22** |
| **11** | **Input Form for Heart Disease Detection** | **23** |
| **12** | **Home Page** | **23** |
| **13** | **CNN making prediction on a bird Image** | **24** |
| **14** | **CNN making prediction on Flower Images** | **25** |
| **15** | **Representing 8 in pixels** | **25** |
| **16** | **Showing image and filter** | **26** |
| **17** | **ReLU Activation Function** | **27** |
| **18** | **Max Pooling** | **27** |
| **19** | **Max Pooled Image** | **28** |
| **20** | **CNN Structure** | **28** |
| **21** | **Flattening the Pooled Image** | **29** |
| **22** | **Flattened Image fed to Fully Connected Layer** | **29** |
| **23** | **Full CNN Structure** | **30** |
| **24** | **Feeding Bird Image to the CNN** | **30** |
| **25** | **Two Different parts of a CNN Structure** | **31** |
| **26** | **Bagging and Boosting Techniques** | **32** |
| **27** | **XGBoost** | **34** |
| **28** | **16 layers of VGG – 16** | **41** |
| **29** | **VGG – 16 Architecture** | **41** |
| **30** | **Alzheimer** | **42** |
| **31** | **Brain Tumor** | **43** |
| **32** | **Breast Cancer** | **43** |
| **33** | **Covid – 19** | **44** |
| **34** | **Diabetes** | **45** |
| **35** | **Heart Disease** | **45** |
| **36** | **Pneumonia** | **46** |
| **37** | **Alzhiemer Result** | **50** |
| **38** | **Brain Tumor Result** | **51** |
| **39** | **Breast Cancer Result** | **52** |
| **40** | **Covid – 19 Result** | **53** |
| **41** | **Diabetes Result** | **54** |
| **42** | **Heart Disease Result** | **55** |
| **43** | **Pneumonia Result** | **56** |

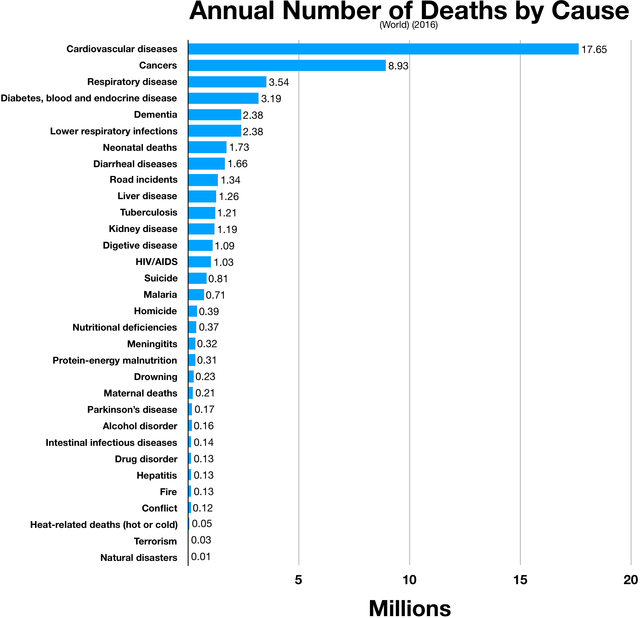
**List of Abbreviations**

|  |  |
| --- | --- |
| **convNet/CNN** | Convolutional Neural Network |
| **AI** | Artificial Intelligence |
| **CT** | Computed Tomography |
| **MRI** | Magnetic Resonance Imaging |
| **XGBoost** | Extreme Gradient Boosting |
| **VGG-16** | Visual Geometry Group |
| **Covid – 19** | Coronavirus disease |
| **BMI** | Body Mass Index |
| **ReLU** | Rectified Linear Activation Unit |
| **GBM** | Gradient Boosting Machine |
| **LASSO** | Least Absolute Shrinkage and Selection Operator |
| **ILSVRC** | ImageNet Large Scale Visual Recognition Challenge |
| **SARS-coV-2** | Coronavirus 2 |
| **GUI** | Graphical User Interface |
| **HTML** | Hypertext Markup Language |
| **CSS** | Cascading Style Sheets |
| **OpenCV** | Open Source Computer Vision Library |

**CHAPTER 1**

**INTRODUCTION**

Recent changes in people's lifestyles and socioeconomic circumstances have been influenced by swift technological advancements and the healthcare sector, increasing the likelihood that people will catch a variety of diseases. Major illnesses with a global impact include lung cancer, pneumonia, and brain tumors, among others.



**Fig 1. Annual Number Of Deaths by cause**

According to the World Health Organization, around 86000 persons had brain tumors in 2019 with an average survival rate of 35%, while lung cancer is a horrible illness that kills one in every five people worldwide, or 1.59 million people, accounting for 19.4% of all fatalities. The coronavirus pandemic, which has affected many nations and caused more than 37 million confirmed cases and more than 1 million fatalities worldwide, has raised awareness of diseases like pneumonia. These serious illnesses raise society demands and healthcare costs, which have an effect on the patient's general health. Determining a person's risk of getting one or more serious diseases is the main objective of disease detection. This calls for the careful analysis of a wide range of concerns, which consumes a substantial amount of time and money.

Numerous medical institutes across the world can now easily acquire medical datasets for health-related data. Examples of several types of medical data include image data, patient reports, and others. Medical data is particularly difficult to manage since it is very erratic, irregular, and contains unstructured data. Based on a variety of factors, including the patient's medical condition, the level of the doctor, and the variations in patient reports, among others, manual data entry is impossible and diagnosis is constrained. By including a machine learning-based disease detection module to help with disease prediction and diagnosis, these difficulties are solved.

Early disease identification may assist to reduce the number of fatalities. In recent years, some researchers have applied various machine learning-based methodologies to construct autonomous disease detection systems. The goal of the disease detection models is to integrate the domains of medicine and Artificial Intelligence (AI) so that people can see how effectively they can complement one another.

**1.1 Objectives**

The following are the objectives of this project:

* Integration of Artificial Intelligence with medical field
* Making disease detection a click away
* The idea of creating an online platform where we brought together all of these disease detections under the one roof.
* It contains both simple and advanced type of detection

**1.2 Background and Literature Survey**

More cures or longer survival times may result from early disease detection. This possibility has prompted public health initiatives that advise communities to undergo routine screening exams with the purpose of identifying particular chronic diseases, such as cancer, diabetes, cardiovascular disease, and so forth. One area that requires the assistance of human professionals is disease detection.

To address this, we came up with the idea of developing an online platform that would house all of these illness detections. It uses nearly all facets of machine learning, including deep learning, and includes both basic and sophisticated types of detections.

We all have been having a rough time lately, and it seems that everyone is afflicted with some sort of illness. For this reason, we wanted to gather everything together on one platform. The key benefit of this project is that we can quickly and easily receive the test results at home. Although we do not claim that these findings are flawless, the experiment we conducted may turn out to be a revolution in the years to come. We offer our approach for consolidating 7 illness detections onto one platform using the power of AI because we know that the future is all about AI.

We are all aware of the heroic role Artificial Intelligence (AI) is playing in modern medicine. Medical research has undergone a revolution thanks to AI, which is constantly surpassing our expectations. Medicine is changing as a result of AI. It aids medical professionals in making better treatment recommendations, future health projections, and more accurate patient diagnoses.AI has grown beyond our wildest dreams, detecting diseases like diabetes and breast cancer using basic machine learning models to diseases like the corona virus and brain tumors using segmentation and other sophisticated methods.

Medical diagnosis is one of the most often performed medical procedures in the area of medical diagnosis and therapy. A significant amount of analytical data is also provided by the identification of medical diseases, which creates the necessary conditions for machine learning in the field of medical diagnosis. By sorting and analyzing a significant amount of medical data, machine learning is able to produce specific data outcomes. Then, using machine learning techniques, a disease diagnosis model is created that can aid medical diagnosticians in making diagnoses.

There are numerous medical images since many diseases can no longer be identified by conventional clinical methods and must instead be evaluated by CT scan, MRI scan, and other methods as the work of medical institutions gets more and more challenging. However, many healthcare professionals are now limited to performing their own analyses of these medical images. As a result of medical staff having an excessive amount of subjective judgements about things like medical quality, knowledge level, and personal competence, it is simple to misinterpret medical imaging and come to incorrect diagnoses. Machine learning-based image processing technology gives us a way to understand visual content more accurately while effectively avoiding the influence of human variables.

We consulted various publications for various ailments, compared the benefits and drawbacks of various algorithms, and chose the most accurate algorithms possible. We used trained Convolutional Neural Networks (CNN) to detect Alzheimer's disease, custom Convolutional Neural Networks to detect pneumonia and Covid-19, Random Forest to detect breast cancer and diabetes, XGBoost to detect heart disease, and VGG-16 in conjunction with custom CNN to detect brain tumors.

**CHAPTER 2**

**HEALTH ROOF – ALL IN ONE SOLUTION**

**2.1 Proposed System**

Input

Feature  
Extraction

Preprocessing

Model  
Training

Model  
Evaluation

Disease  
Detection

Result

**Fig 2. System Flow Chart**

**2.2 Working Methodology**

|  |  |  |
| --- | --- | --- |
| **Disease** | **Methodology Used (1/2)** | **Training Algorithm Used** |
| Alzheimer Detection | Method 1- image as input | Trained CNN |
| Brain Tumor Detection | Method 1 - image as input | VGG – 16 and 3 levels of CNN |
| Breast Cancer Detection | Method 2 – csv file as input | Random Forest |
| Covid – 19 Detection | Method 1 - image as input | Custom-made CNN |
| Diabetes  Detection | Method 2 - csv file as input | Random Forest |
| Heart Disease Detection | Method 2 - csv file as input | XGBoost |
| Pneumonia Detection | Method 1 - image as input | Custom CNN |

**Table 1. Diseases - Methodologies**

**2.2.1 Method 1 – Image Pre-Processing**

(working with images as input)

MRI

Scan

Chest

X-Ray

CT

Scan

Image Acquisition

**Fig 3. Methodology 1 - Image Pre-Processing Architecture**

Pneumonia

Covid-19

Brain  
Tumour

Alzheimer

VGG-16

CNN

Feature

Extraction

Classification

Training

Pre-processing

1. **Image Acquisition Phase**

The first stage of the disease detection model is the collection of photos relevant to diseases. The CNN technique is utilized, therefore a large number of images must be trained on the model. In the context of this study, photographs offer crucial information for the identification of various disorders. Images like chest X-rays and CT scans can be used. The output of the first phase consists of photos that are used to train the model. Datasets were gathered from a variety of online resources. In Total we gathered 6,400 images for Alzheimer Detection,252 for Brain Tumor Detection,392 images for Covid-19 detection,5,858 images for Pneumonia Detection.

1. **Image Pre-processing Phase**

At this stage, the image is altered to improve image quality. Online photos boost the diversity of information, but they also have the potential to have noise, artefacts, spurious pixels, and surface roughness. Additionally, they could be twisted or confusing. The collected photos were recorded with appropriate average resolutions in order to avoid these issues and prevent image intensity loss. Since they are of different sizes, the photos in the dataset are altered to have a shape of (176,176) = (image width, image height) for Alzheimer,(224,224) for Brain Tumor,(224,224) for Covid-19,(150,150) for Pneumonia. Normalization is used to scale pixel values to lie between 0 and 1. The purpose of feature extraction is to enable the CNN model to discover pertinent features that may be applied to categorize a certain class. The end result is a collection of photos with improved quality or unwanted parts deleted.

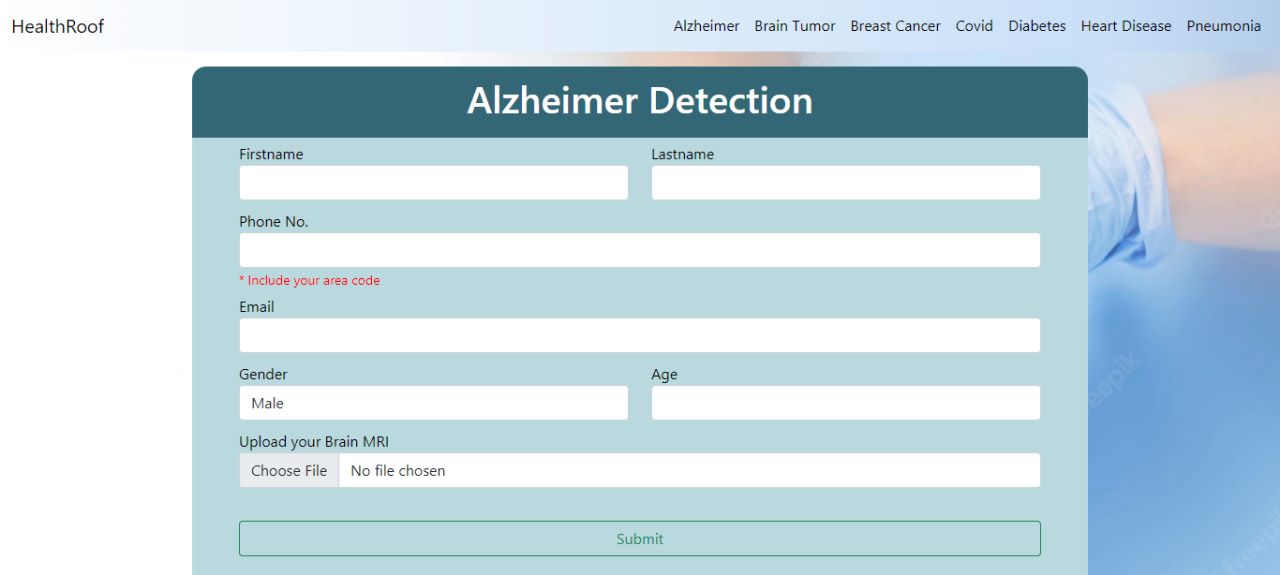
1. **Training Phase**

The third phase, training, is where a deep learning algorithm is chosen. A deep learning algorithm like CNN is one that has already been mentioned. There are several ways that algorithms can learn. Certain data types and certain algorithms complement one another well. CNN is skilled at employing visuals. Which deep learning technique is employed should be determined by the type of data. This stage results in the models built using the data learned. After comparing different algorithms and their accuracies for different diseases, we used customized CNN for Alzheimer Detection,Covid-19 Detection and pneumonia Detection, and VGG-16 and 3 layers of CNN for Brain Tumor Detection.

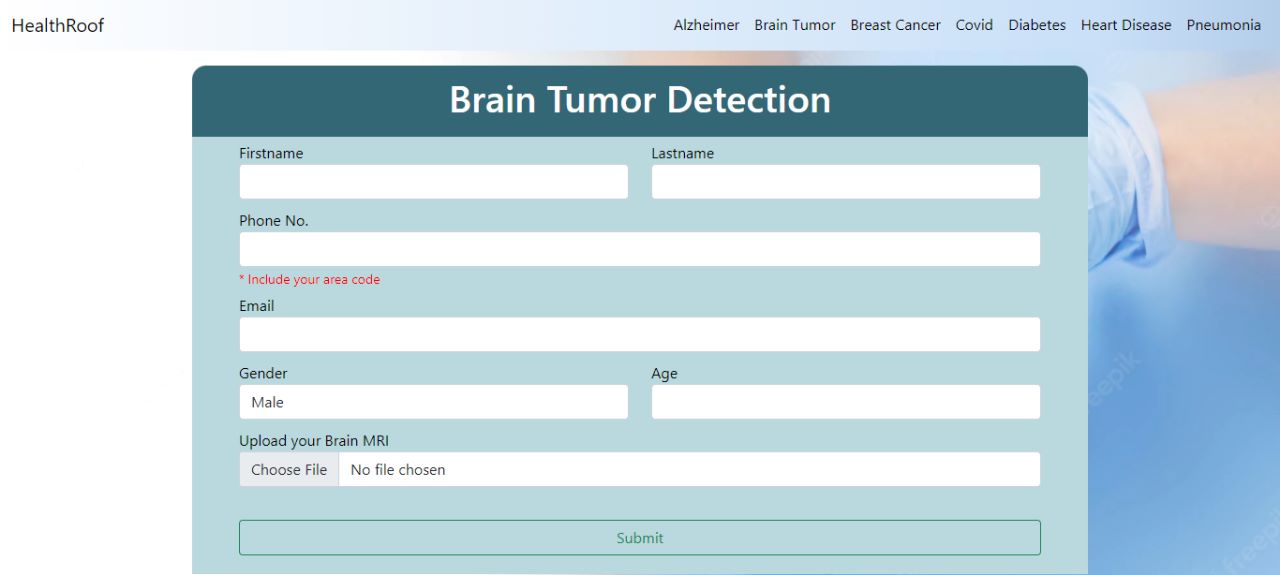
1. **Classification Phase**

The trained model predicts which class an image belongs to in the final phase of classification. For Alzheimer, a model trained to differentiate Mild Demented, Moderate Demented, Non Demented, Very Mild Demented in images and classifies images according to that. For Brain Tumor, a model trained to discriminate between a tumorous brain and a normal brain in MRI pictures and should classify images accordingly. For Covid-19,model is trained to classify as covid and normal. For Pneumonia, model is trained to differentiate images as normal person and a person suffering with Pneumonia. Each image is assigned a probability score by the model, indicating the likelihood that it belongs to a particular class.

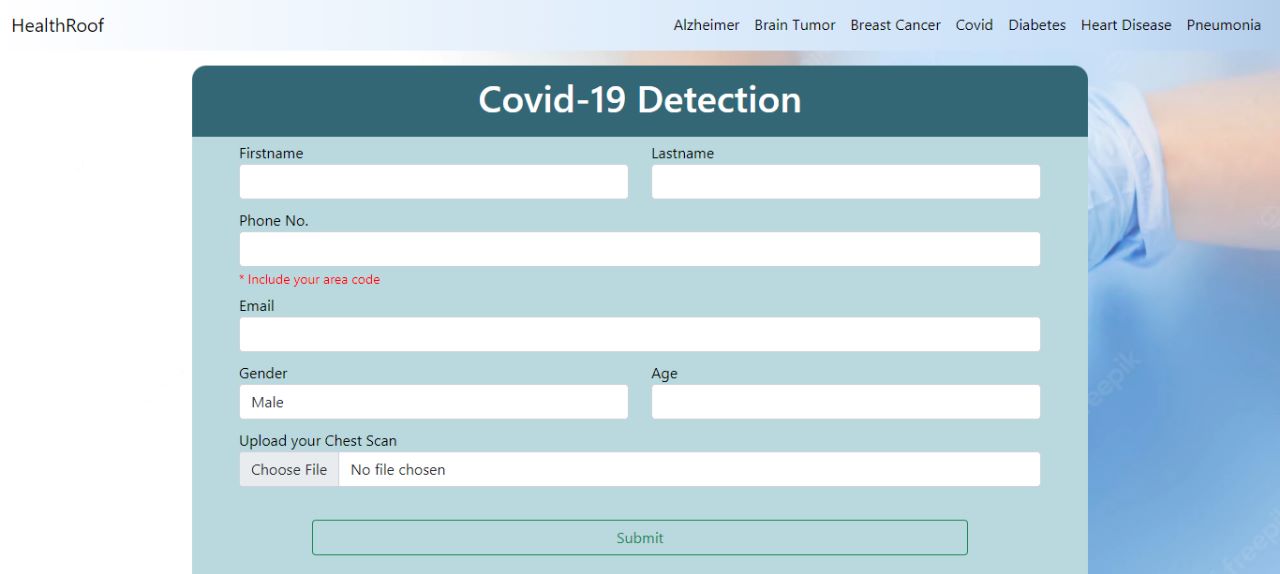
**INPUT FORMS**



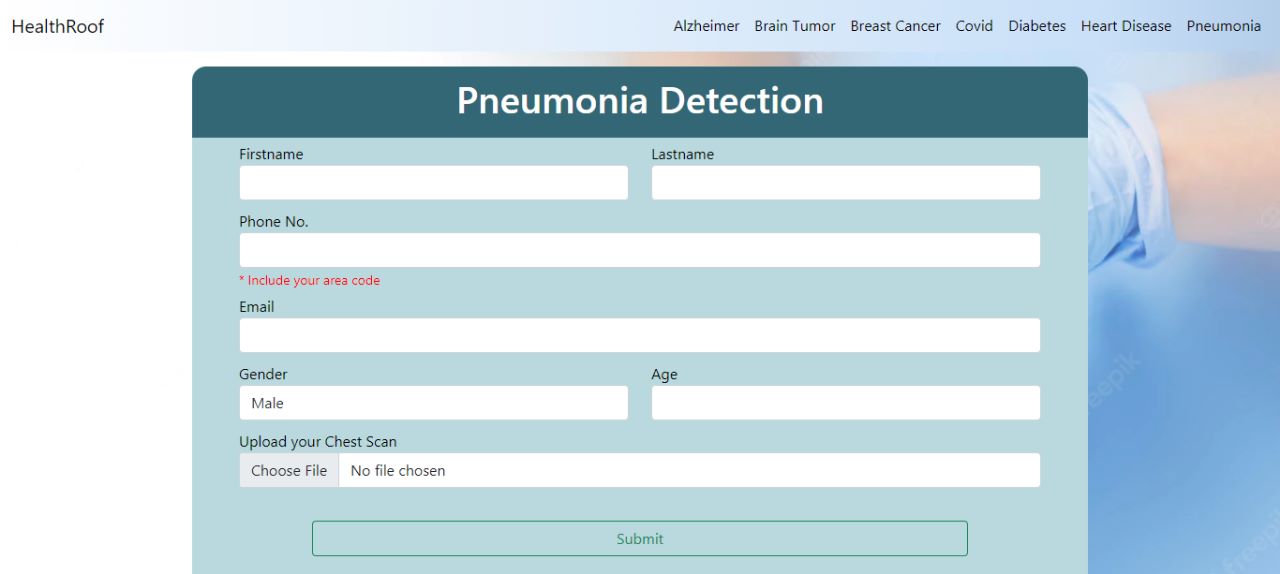
**Fig 4. Input Form for Alzheimer Detection**

****

**Fig 5. Input Form for Brain Tumor Detection**

****

**Fig 6. Input Form for Covid – 19 Detection**

****

**Fig 7. Input Form for Pneumonia Detection**

**2.2 Method 2 – Data Pre-Processing**

In this system, we apply various machine learning approaches, such as feature scaling, model building, and data pre-processing, to produce results, fast coding To visualize the data, we will use a variety of Python modules, including pandas and NumPy. To get our final result, we combine the best model with the highest accuracy out of five separate models.

**ADVANTAGES**

1. Increased Accuracy

2. Low time and cost-friendly

**DISADVANTAGES**

1. Interpretation of Result.

2. Data Acquisition

Dataset

Data Pre-Processing

Feature Selection

Prediction

Performance  
Evaluation

Results

**Fig 8. Methodology 2 – Data Pre-Processing Architecture**

1. **Data Pre-Processing**

The data set's data will undergo pre-processing. There are 32 records and 570 fields in the Breast Cancer data collection. There are a total of 9 records and 769 fields in the Diabetes data collection. There are 304 fields and 14 records in the Heart Disease data set. Certain records have some values that are missing. The other records will go through pre-processing while those records are replaced or eliminated.

1. **Feature Selection and Reduction Phase**

Age and Gender are common inputs ,hence they are taken from the patient for every disease. In Breast Cancer Detection, among all records Concave points Mean, Area Mean, Radius Mean, Perimeter Mean, Concavity Mean are important to detect breast cancer so these are collected from the user ,rest are ignored. In Diabetes Detection, all records are important ,hence we collected data like number of pregnancies, glucose concentration, blood pressure, skin thickness, insulin, BMI, Diabetes pedigree are taken from the user. In Heart Disease Detection, among all records ,data like Old Peak, Maximum Heart Rate Achieved, Exercise induced angina, number of major vessels, types of chest pain, Thal are obtained from user and rest are ignored.

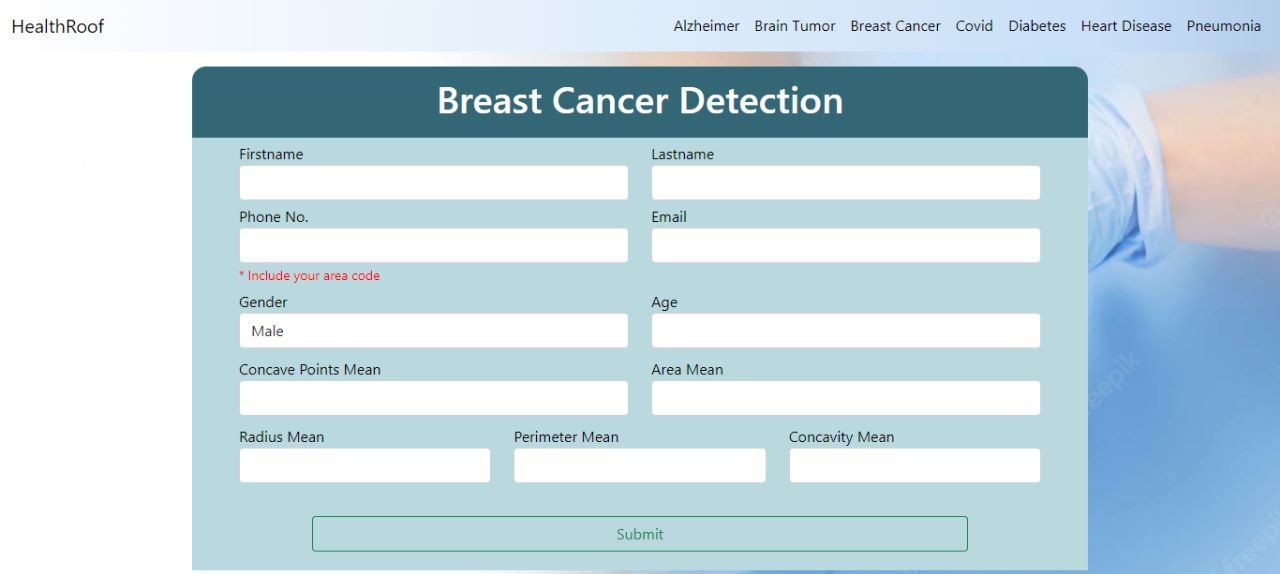
1. **Prediction Phase**

In this phase ,output is predicted using suitable algorithms. Comparing accuracies of different algorithms for different disease detections suitable algorithms are used. Random Forest is used to detect Breast Cancer and Diabetes and XGBoost is used to detect Heart Disease.

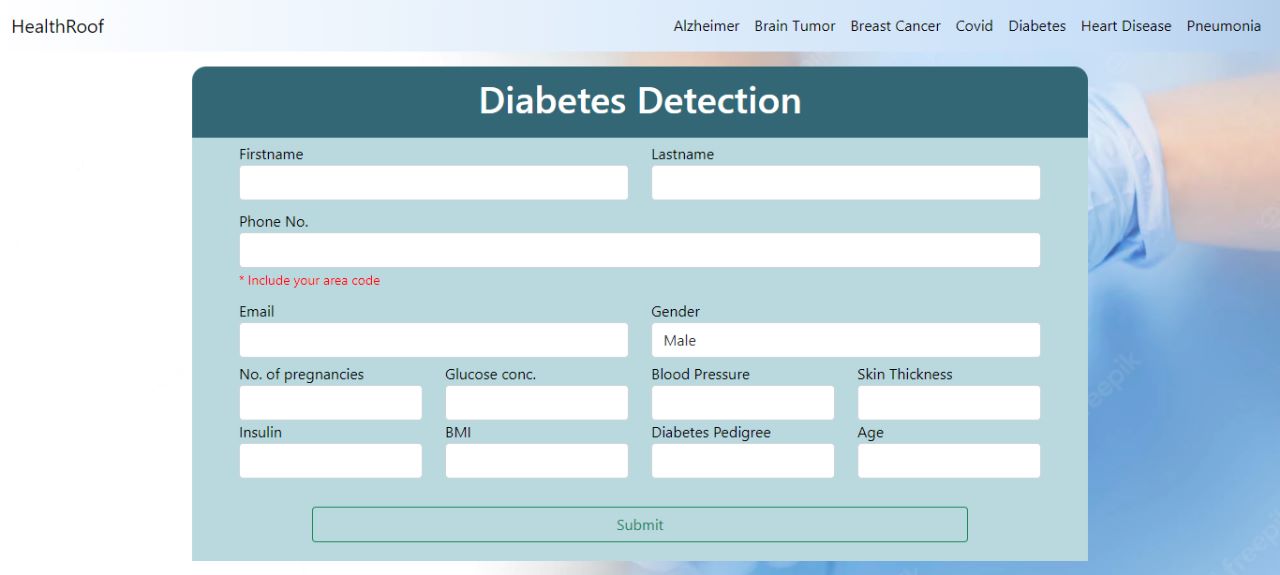
1. **Classification Phase**

Clustering and decision tree feature criteria were applied based on variables. Each clustered dataset also has the classifiers applied to measure their performance. The findings are connected to the top models based on their low rate of error.

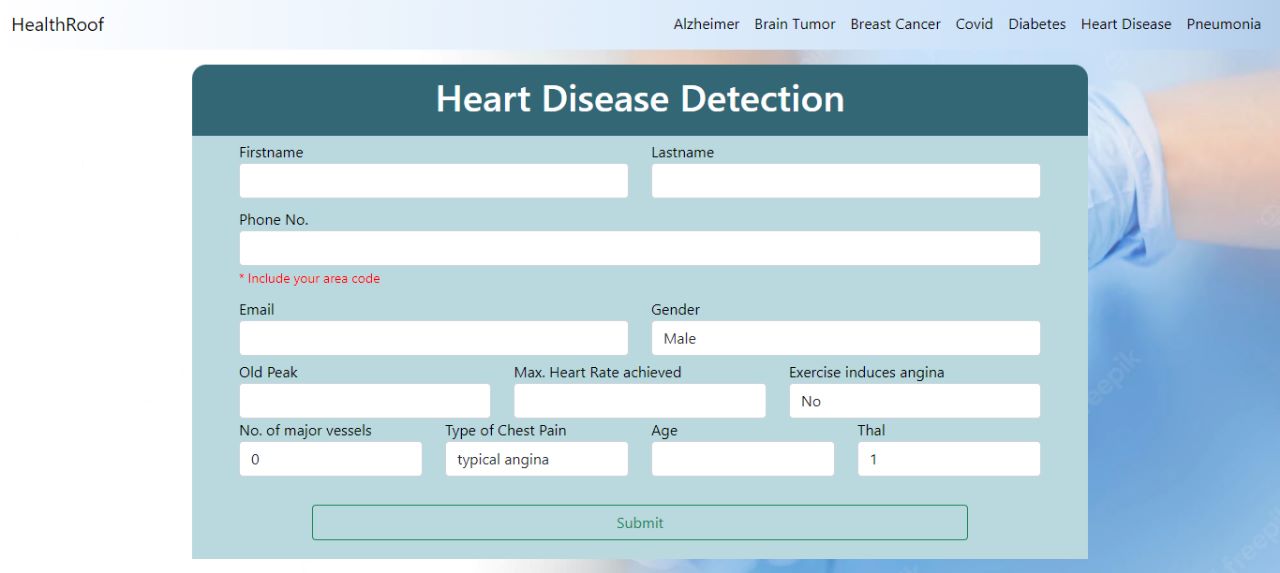
**INPUT FORMS**



**Fig 9. Input Forms for Breast Cancer Detection**

****

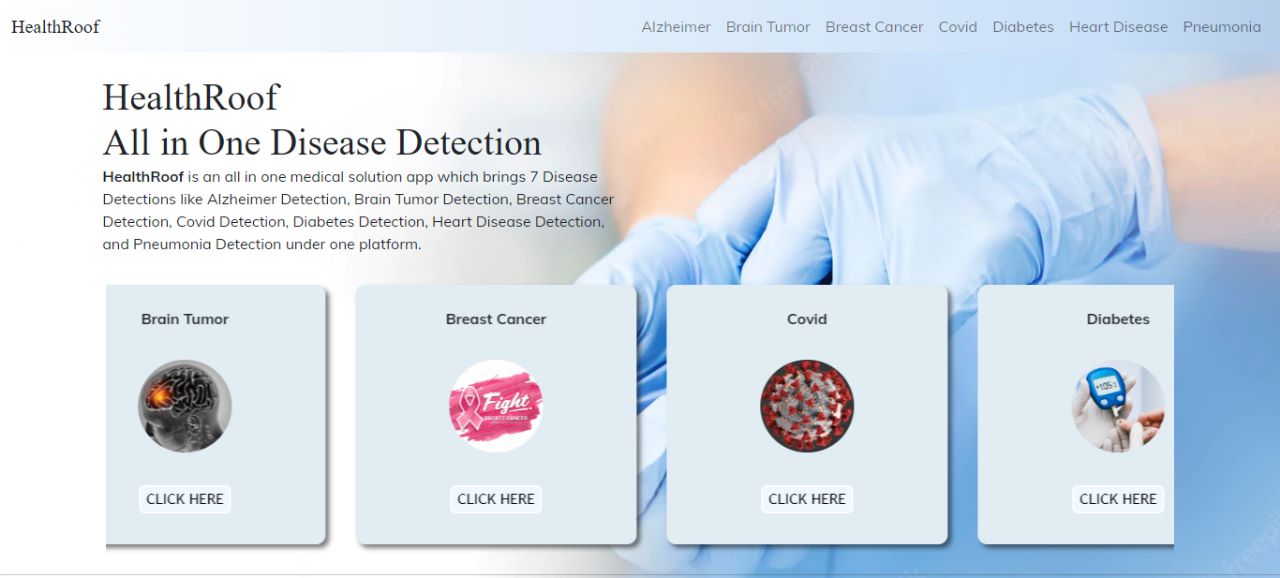
**Fig 10. Input Form for Diabetes Detection**

****

**Fig 11. Input Form for Heart Disease Detection**

**HOME PAGE**

We can navigate to the respective disease from navbar or carousels and input the required fields to get respective results.

****

**Fig 12. Home Page**

**CHAPTER 3**

**ALGORITHMS**

**3.1 CONVOLUTIONAL NEURAL NETWORKS(CNN)**

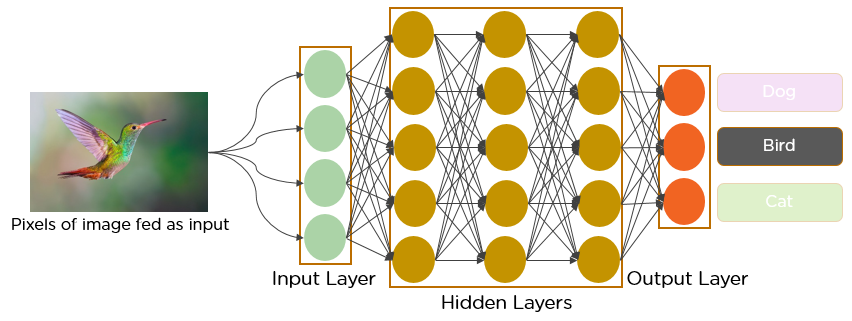
The potential gap between people and machines has been seamlessly closed because to advancements in artificial intelligence. And data enthusiasts from all over the world work on a variety of AI-related topics to make visions a reality. One such fantastic area is the field of computer vision. In order to utilise the knowledge for various tasks and processes, this field tries to enable and configure machines to view the world like humans do (such as Image Recognition, Image Analysis and Classification, and so on). And the developments in computer vision made possible by deep learning—particularly those involving the Convolutional Neural Network algorithm—have been a great success.

**3.1.1 CNN introduction**

Director of Facebook's Al Research Group, Yann LeCun was a convolutional neural network pioneer. In 1988, he created the first convolutional neural network, known as LeNet. For character recognition tasks like reading zip codes and numbers, LeNet was employed.

Have you ever wondered how object detection helps to construct self-driving cars, how facial recognition works on social media, or how disease identification is done using visual imaging in healthcare? All of this is made possible by convolutional neural networks (CNN). Here is a convolutional neural network example that demonstrates how it functions:

Consider trying to determine whether a photograph of a bird actually depicts a bird or another object. The first thing you do is provide the input layer with the image's pixels arrays as a neural network (multi-layer networks used to classify things). The feature extraction process is carried out by the hidden layers through various calculations and manipulations. The extraction of features from the image is done by a number of hidden layers, including the convolution layer, the ReLU layer, and the pooling layer. A fully connected layer that identifies the object in the image is present at the end.



**Fig 13. CNN making prediction on a bird Image**

**3.1.2 Convolutional Neural Network: What Is It?**

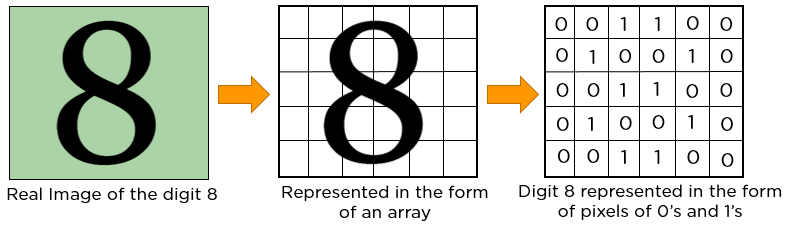
A feed-forward neural network called a convolutional network analyses visual images by processing data in a grid-like architecture. It is also know as a ConvNet. To find and categorize items in an image, a convolutional neural network is employed.

The neural network shown below can distinguish between the rose and orchid flower types.



**Fig 14. CNN making prediction on Flower Images**

Every image on CNN is represented as an array of pixel values.



**Fig 15. Representing 8 in pixels**

**3.1.3 Convolutional neural network layers**

Multiple hidden layers in a convolution neural network aid in information extraction from a picture.

CNN's four key layers are as follows:

Layers:

1. Convolution,

2. ReLU

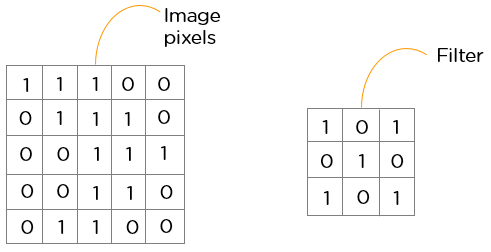
3. Pooling

4. A layer with all connections

**3.1.3.1 Convolution Layer**

The process of removing useful elements from an image begins with this. Multiple filters work together to perform the convolution action in a convolution layer. Each image can be thought of as a matrix of pixel values.

Take into account the following 5x5 image, where each pixel's value is either 0 or 1. A filter matrix with a 3x3 dimension is also present. To obtain the convolved feature matrix, move the filter matrix over the image and compute the dot product.

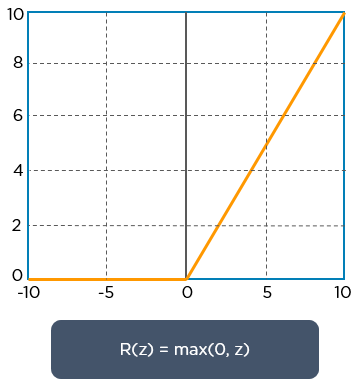


**Fig 16. Showing image and filter**

**3.1.3.2 Layer ReLU**

The rectified linear unit is referred to as ReLU. The next step is to transfer the feature maps to a ReLU layer after they have been retrieved.

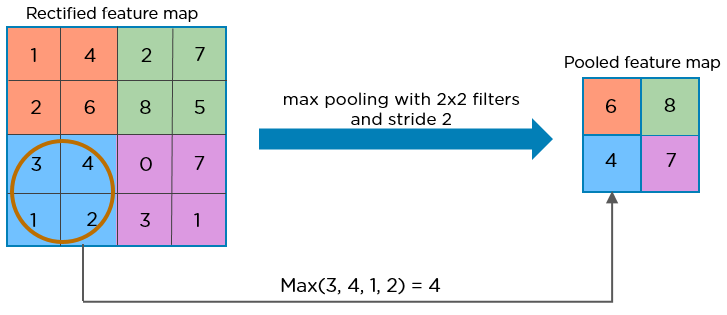
ReLU executes an operation element-by-element, setting all the negative pixels to 0. The result is a corrected feature map, and it gives the network non-linearity. The graph of a ReLU function is shown below:



**Fig 17. ReLU Activation Function**

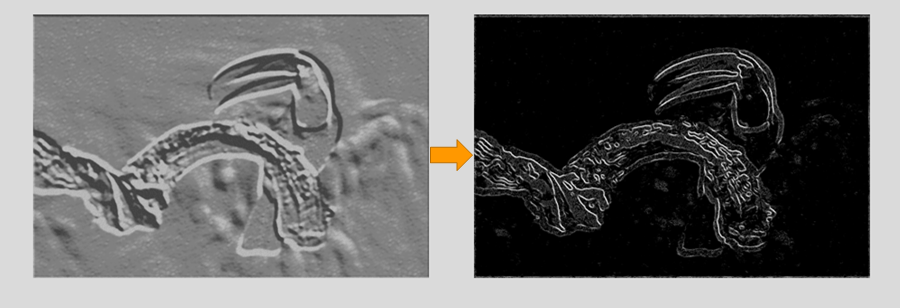
**3.1.3.3 A pooling layer**

The down sampling process of pooling lowers the feature map's dimensionality. To create a pooled feature map, the rectified feature map is now passed through a layer of pooling.



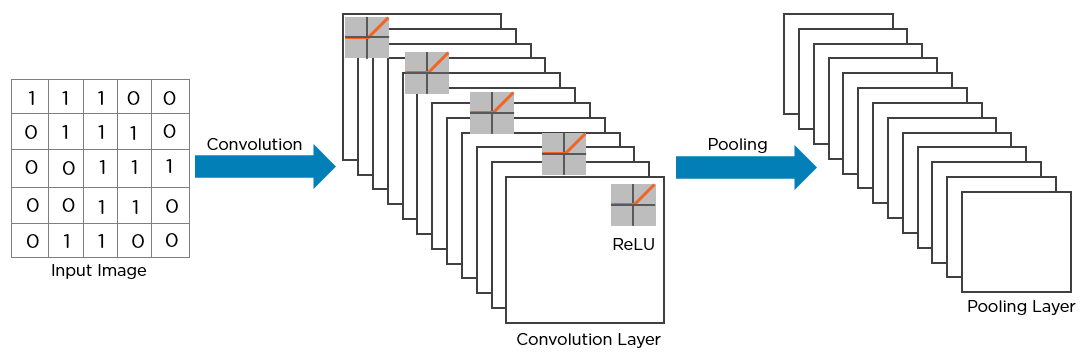
**Fig 18. Max Pooling**

To distinguish distinct portions of the image, such as edges, corners, bodies, feathers, eyes, and beak, the pooling layer employs a variety of filters.



**Fig 19. Max Pooled Image**

Here is how the convolution neural network is currently structured:



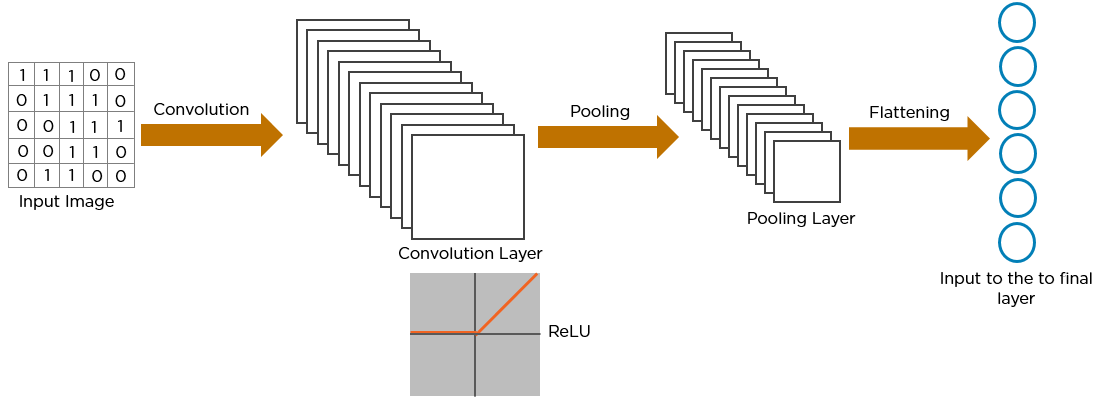
**Fig 20. CNN Structure**

Flattening is the procedure's following phase. The generated 2-Dimensional arrays from pooled feature maps are all flattened into a single, lengthy continuous linear vector.

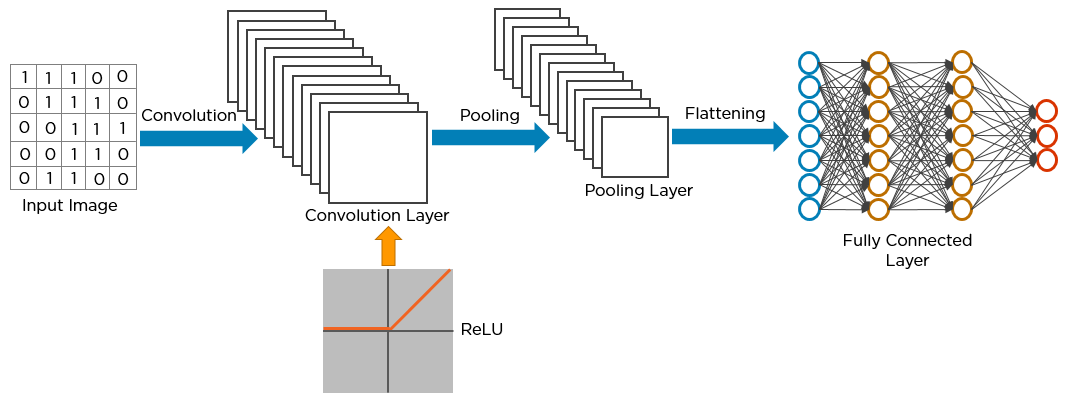


**Fig 21. Flattening the Pooled Image**

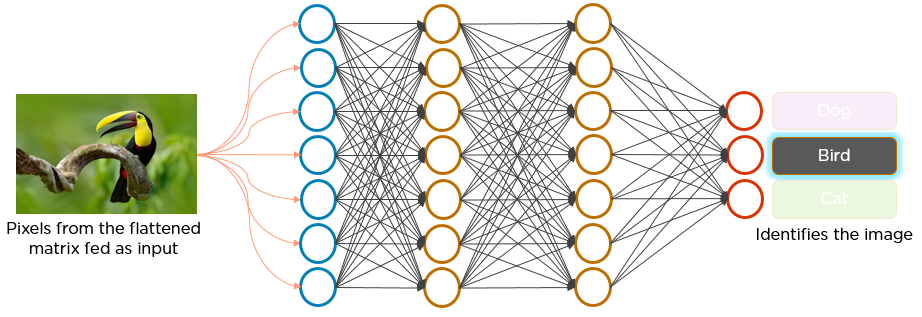
To categorize the image, the flattened matrix is provided as input to the fully linked layer.



**Fig 22. Flattened Image fed to Fully Connected Layer**



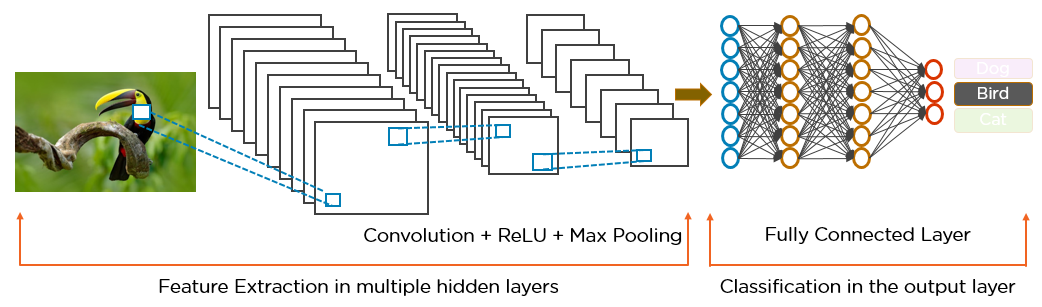
**Fig 23. Full CNN Structure**



**Fig 24. Feeding Bird Image to the CNN**

**Exactly how does CNN identify a bird?**

* The convolutional layer that executes the convolution operation receives input from the image's pixels, creating a convolved map.
* The image is processed with numerous convolutions and ReLU layers for locating the features.
* Different pooling layers with varying filters are used to identify certain sections of the image.
* The pooled feature map is flattened and sent to a fully connected layer to get the final result.

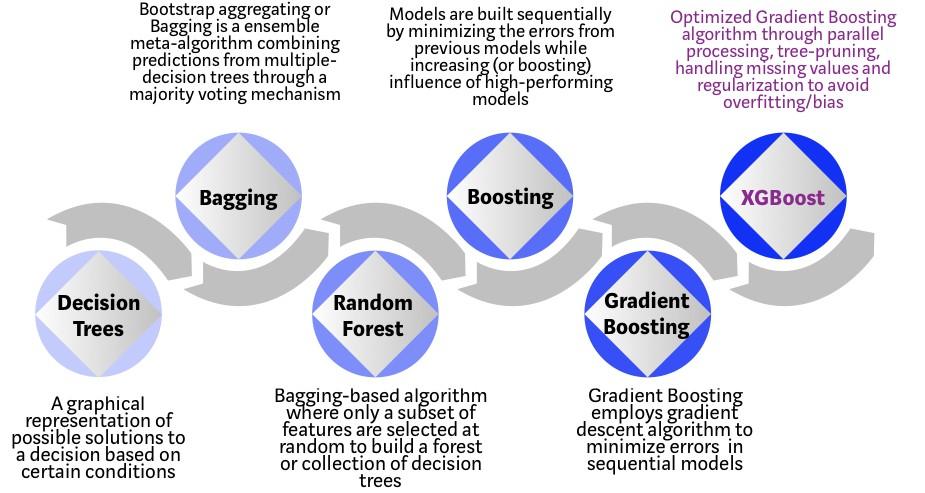


**Fig 25. Two Different parts of a CNN Structure**

**3.2 XGBoost Technique**

**3.2.1 Explain XGBoost**

The decision-tree based ensemble machine learning algorithm XGBoost uses a gradient boosting architecture. In prediction problems involving unstructured data, artificial neural networks usually outperform all existing algorithms or frameworks (pictures, text, etc.). However, it is currently believed that decision tree-based algorithms are best in class for small to medium-sized structured/tabular data. Please see the graph below for an explanation of how tree-based algorithms have evolved throughout time.



**Fig 26. Bagging and Boosting Techniques**

At the University of Washington, a research endeavor led to the creation of the XGBoost algorithm. At the SIGKDD Conference in 2016, Tianqi Chen and Carlos Guestrin gave a paper presentation that lit up the machine learning community. Since its launch, this algorithm has received praise for not just taking first place in a number of Kaggle competitions but also for powering a number of innovative business applications. As a result, the XGBoost open source projects have a large community of data scientists participating, with about 350 contributors and 3,600 GitHub commits.

The following characteristics of the algorithm set it apart from others:

* A variety of applications: can be used to resolve challenges involving regression, classification, ranking, and custom prediction.
* Portability: Easily uses Linux, Windows, and OS X.
* Languages: Julia, C++, Python, R, Java, Scala, and all other popular programming languages are supported.
* Cloud Integration: Compatible with Flink, Spark, and other ecosystems, and supports AWS, Azure, and Yarn clusters.

**3.2.2 How can I develop my understanding of XGBoost?**

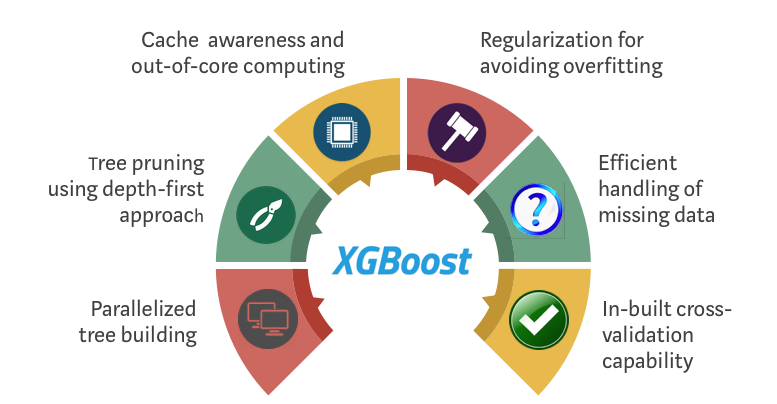
While decision trees are very understandable and straightforward to comprehend in their most basic forms, developing intuition for the next generation of tree-based algorithms can be challenging. To further see how tree-based algorithms have developed, see the straightforward illustration below.

Consider yourself a recruiting manager who is interviewing a number of individuals with outstanding credentials. It is possible to think of each stage of the development of tree-based algorithms as a particular interviewing procedure.

1. Decision Tree : Every recruiting manager has a list of requirements, including education level, years of experience, and interview performance. A decision tree is comparable to a hiring manager conducting interviews with candidates according to their own standards.
2. Bagging : Alternatively Instead of just one interviewer, there is now a panel of interviewers who each cast a vote. Using a democratic voting procedure, bagging or bootstrap aggregating combines the information from each interviewer to determine the final outcome.
3. Random Forest : This bagging-based algorithm's main distinction is that just a portion of the characteristics are chosen at random. In other words, each interviewer will only evaluate the candidate according to a set of randomly chosen criteria (e.g. a technical interview for testing programming skills and a behavioral interview for evaluating non-technical skills).
4. Boosting : Boosting is an alternative strategy in which each interviewer modifies the evaluation standards in light of comments made by the previous interviewer. This "boosts" the interview process' efficiency by utilizing a more dynamic evaluation procedure.
5. Gradient Boosting : A special case of boosting in which errors are reduced using the gradient descent algorithm is used, for example, by strategy consulting firms to select out less competent candidates through case interviews.
6. XGBoost : Think of XGBoost as gradient boosting on "steroids" (Extreme Gradient Boosting is its official name, after all!). In order to produce better results with fewer computer resources and in the shortest amount of time, it is the ideal mix of software and hardware optimization techniques.

**3.2.3 Why is XGBoost so effective?**

Both XGBoost and Gradient Boosting Machines (GBMs), ensemble tree approaches, use the gradient descent architecture to boost weak learners (CARTs in general). But XGBoost enhances the fundamental GBM architecture with system optimization and algorithmic improvements.



**Fig 27. XGBoost**

**3.2.4 How XGBoost improves the common GBM method**

**3.2.4.1 System Optimization:**

1. **Parallelization**

The sequential tree construction method is approached by XGBoost using a parallelized implementation. This is made possible by the interchangeability of the two inner loops that calculate the features and the outer loop that counts the leaf nodes of a tree when creating base learners. This nesting of loops restricts parallelization because the outer loop cannot be initiated until the inner loop, which is the more computationally intensive of the two, has been finished. As a result, the order of the loops is adjusted utilising initialization, a global scan of all instances, and sorting using parallel threads to save run time. By balancing any parallelization overheads in computation, this choice enhances algorithmic performance.

1. **Tree Pruning**

Within the GBM framework, the halting criterion for tree splitting is greedy in nature and is dependent on the negative loss criterion at the split point. Instead of starting with the first criterion, XGBoost uses the supplied "max depth" option and prunes the trees backward. Significantly better computational performance is achieved with this "depth-first" strategy.

1. **Hardware Optimization**

This algorithm has been created to utilise hardware resources as effectively as possible. By allocating internal buffers in each thread to hold gradient statistics, cache awareness does this. Additional improvements, such "out-of-core" computation, maximise disc space while managing large data frames that don't fit in memory.

* + - 1. **Improvements to the algorithm:**

1. **Regularization**

To avoid overfitting, it penalizes more complex models using both LASSO (L1) and Ridge (L2) regularization.

1. **Sparsity Awareness**

XGBoost handles various forms of sparsity patterns in the data more effectively by automatically "learning" the best missing value based on training loss and naturally admits sparse features for inputs.

1. **Weighted Quantile Sketch**

To efficiently identify the best split points across weighted datasets, XGBoost uses the distributed weighted Quantile Sketch algorithm.

1. **Cross-validation**

The algorithm includes a cross-validation approach that is incorporated into each iteration, eliminating the need to explicitly program this search and to specify the precise number of boosting iterations necessary in a single run.

**3.3 Machine Learning Bagging and Random Forest Ensemble Algorithms**

One of the most well-liked and effective machine learning algorithms is Random Forest. It is a form of ensemble machine learning algorithm known as bagging or Bootstrap Aggregation.  
The Bagging ensemble technique and the Random Forest algorithm for predictive modelling are introduced in this post.

We will be familiar with:  
1. The bootstrap approach, which uses samples to estimate statistical quantities.  
2. Using a single training dataset, the Bootstrap Aggregation technique can generate numerous distinct models.  
3. The Random Forest algorithm, which modifies Bagging slightly and produces an extremely potent classifier.

**3.3.1 Bootstrap Method**

Let's quickly review the bootstrap, a crucial foundational approach, before moving on to bagging. A strong statistical technique for estimating a quantity from a sample of data is the bootstrap.

If the quantity is a descriptive statistic like a mean or a standard deviation, this is the easiest to comprehend.  
Assume we want to estimate the sample mean and that we have a sample of 100 values (x).

The mean can be determined directly from the sample using the formula

mean(x) = 1/100 \* sum (x)

We are aware of the tiny sample size and error in our mean. Using the bootstrap method, we may refine the estimate of our mean:

1. Generate a large number (like 1000) of random subsamples with replacement (meaning we can select the same value multiple times).  
2. Determine each subsample's mean.  
3. Determine the average of all the means we've gathered, then use that figure as our estimated mean for the data.  
As an illustration, suppose we obtained the mean values 2.3, 4.5, and 3 from 3 resamples.  
3. By averaging these, we can determine that the estimated mean for the data is 3.367.  
Other values, such as the standard deviation and even those utilized in machine learning techniques, such as learned coefficients, can be estimated using this method.

**3.3.2 Bootstrap Aggregation (Bagging)**

The ensemble method known as "Bagging" is straightforward and extremely effective.

An ensemble method is a strategy that combines predictions from various machine learning algorithms to provide predictions that are more accurate than those from any one model.

For algorithms with large volatility, Bootstrap Aggregation is a general method that can be used to lower the variance. Decision trees, such as classification and regression trees, are an approach with a high variance (CART).

Trees used for decision making are sensitive to the particular data used for training. The resulting decision tree, and consequently the predictions, can differ significantly if the training data is altered (for example, if a tree is trained on a fraction of the training data).

Applying the Bootstrap process to a high-variance machine learning system, usually a decision tree, is known as bagging.

Assume we are using the CART algorithm and a sample dataset of 1000 instances (x). The CART algorithm would function as follows while bagging.

1. Produce numerous (say, 100) arbitrary replacement subsamples from our dataset.  
2. Use each sample to train a CART model.  
3. Determine the average forecast from each model given a fresh dataset.

In general, ensemble learning is a model that generates predictions using a variety of distinct models.

The ensemble model is typically more flexible (less biased) and less data-sensitive because it combines individual models (less variance).

The two most common ensemble techniques are boosting and bagging.

**1. Bagging:** The concurrent training of many different models. The data used to train each model is chosen at random.

**2. Boosting:** The successive training of several different models Each individual model gains knowledge from the errors of the preceding model.

**3.4 Random Forest**

Bagging serves as the ensemble method and decision trees serve as the individual models in the ensemble model known as random forest.  
Let's examine the magic of randomness in more detail:

**Step 1:** choose n (say, 1000) randomly chosen subsets from the training set.

**Step 2:** Train n decision trees, such as 1000.

1. Only one decision tree is trained from a single random subset.
2. Each decision tree's best splits are determined by a random collection of features

**Step 3:** Each tree independently predicts the data points in the test set.

**Step 4:** Make your final forecast.

Random Forest determines the final forecast for each candidate in the test set by taking the class (such as cat or dog) that received the most votes.

**3.4.1 AdaBoost (Adaptive Boosting) (Adaptive Boosting)**

The boosting ensemble model AdaBoost performs very well with the decision tree. The secret to improving models is to draw lessons from past errors, such as incorrectly classified data points. AdaBoost corrects its errors by giving misclassified data points more weight.  
Let's show you how AdaBoost adjusts.  
Initialize the weights of the data points in step 0.

**Step 0 :** The initial weight for each point should be 1/100 = 0.01 if the training set has 100 data points.

**Step 1 :**Train a decision tree first.

**Step 2:** Determine the decision tree's weighted error rate (e). The number of incorrect guesses out of all forecasts is the weighted error rate (e), and you handle incorrect predictions differently depending on the weight of the corresponding data point. When calculating the weight, the related inaccuracy will be weighted more heavily the higher the weight (e).

**Step 3:** Determine the ensemble weight for this decision tree.  
This tree's weight is determined by learning rate times log ((1 — e)/e)

1. A tree will be awarded less voting power in the subsequent voting if its weighted error rate is higher.
2. A tree will be awarded more decision power during subsequent votes if its weighted error rate is lower.

**Step 4:** Adjust point weights that were incorrectly assigned.

Every data point has a weight of 1,

1. The weight remains the same if the model correctly predicted this data point.
2. If the model miscalculated this data point, the new weight of this point is the previous weight multiplied by np.exp (weight of this tree)  
     
   Note :The misclassified data point by this tree will receive more of a boost (importance) the higher the weight of the tree (the more accurately this tree performs). After all the incorrectly classified points have been rectified, the weights of the data points are normalized.

**Step 5 :** Repeat Step 1 in Step 5 (until the number of trees we set to train is reached)

**Step 6 :** Make your final prediction in step six.

By multiplying the forecast by the weight (of each tree), the AdaBoost creates a new prediction (of each tree). Naturally, the tree with the most weight will have the greatest ability to sway the outcome.

**3.4.2 Gradient boosting**

Another boosting model is gradient boosting. Remember that the secret to improving your model is to learn from your past errors. Instead of changing the weights of the data points, Gradient Boosting directly learns from the error—the residual error.  
Here is an example of how Gradient Boost learns.

**Step 1 :** Train a decision tree first.

**Step 2 :** Use the decision tree you have trained to make predictions.

**Step 3:** Determine the decision tree's residual and save the residual errors as the new y.

**Step 4 :** Repeat Step 1 in Step 4 (until the number of trees we set to train is reached)

**Step 5 :** Make your final prediction in step five.

By simply adding the predictions, the Gradient Boosting generates a new forecast (of all trees).

**3.5 VGG – 16**

VGG-16 is a convolutional neural network model trained on the ImageNet dataset. It was developed by researchers at the University of Oxford and was used to win the ILSVRC-2014 competition. The VGG-16 model consists of 16 layers and was created by modifying the architecture of the VGG-19 model. The model is widely used as a benchmark for image classification and has been used in a number of research papers. It is also often used as a base model for training custom image classification models.

**3.5.1 About VGG-16 Architecture**

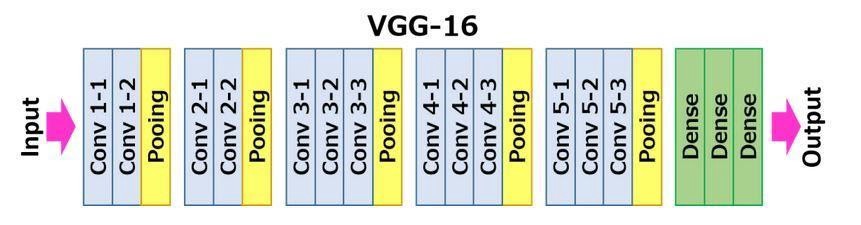
A ConvNet is another name for a convolutional neural network, which is a type of artificial neural network. The input layer, output layer, and convolutional layer all output layer is one of the hidden layers. The CNN (Convolutional Neural Network) variation known as VGG16 is one of the best computer vision models available today. In order to analyze the networks and improve the depth of the model, the creators used an architecture with exceptionally small (3 3) convolution filters, which showed a significant improvement over previous methods. With the depth extended to 16–19 weight layers, around 138 trainable parameters were produced.

**3.5.2 What's the use of VGG16?**

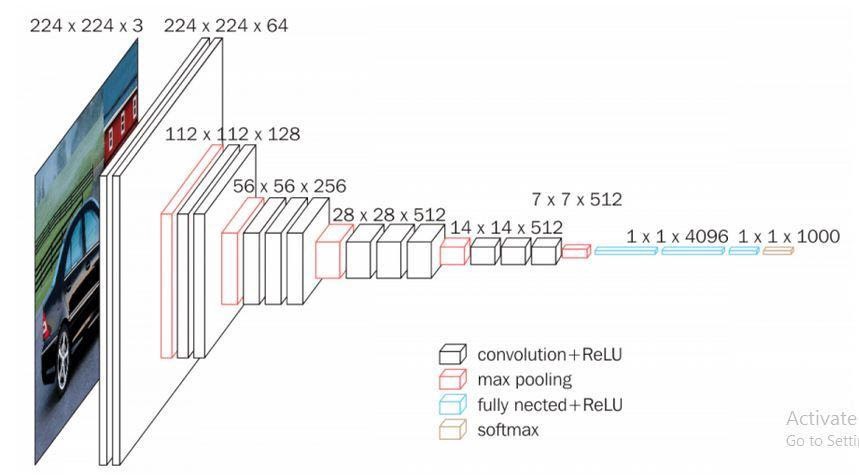
With 1000 images divided into 1000 separate categories, the object identification and classification algorithm VGG16 has a classification accuracy of 92.7%. It is a popular method for categorizing photographs and is easy to use with transfer learning.

**3.5.3 VGG – 16 Architectures**

* VGG16's 16th digit denotes its 16 weighted layers. Thirteen convolutional layers, five Max Pooling layers, three Dense layers, and a total of 21 layers make up VGG16, but only sixteen of them are weight layers, also known as learnable parameters layers.
* VGG16 requires an input tensor with three RGB channels and a size of 224, 244.
* The most distinctive feature of VGG16 is that it prioritized convolution layers of a 3x3 filter with stride 1 rather than a large number of hyper-parameters and consistently employed the same maxpool layer and padding as a 2x2 filter with stride 2.
* Throughout the whole architecture, the convolution and max pool layers are uniformly ordered.
* There are 64 filters in the Conv-1 Layer, 128 filters in Conv-2, 256 filters in Conv-3, and 512 filters in Conv-4 and Conv-5.
* A stack of convolutional layers is followed by three Fully-Connected (FC) layers, the third of which conducts 1000-way ILSVRC classification and has 1000 channels. The first two FC layers have 4096 channels each (one for each class). The soft-max layer is the last one.

****

**Fig 28. 16 layers of VGG – 16**

****

**Fig 29. VGG – 16 Architecture**

**CHAPTER 4**

**DISEASES**

* 1. **ALZHEIMER DETECTION**

Alzheimer's disease is a progressive brain disorder that affects memory, thinking, and behavior. It is the most common cause of dementia in older adults. The disease is characterized by the abnormal accumulation of amyloid protein in the brain, which leads to the death of brain cells and the loss of cognitive function. The symptoms of Alzheimer's disease typically develop slowly over a period of years and may include memory loss, difficulty communicating, and changes in behavior. There is currently no cure for Alzheimer's disease, but there are treatments that can help manage the symptoms.



**Fig 30. Alzheimer**

* 1. **Brain Tumor**

The most frequent solid tumor in children is a brain tumor. Primary brain tumors are identified in about 4,000 children and teenagers in the US each year. Primary brain tumors originate in the brain and often do not spread to other parts of the body. Brain tumors are the primary malignancy of the central nervous system. Tumors that begin in the brain's cells are known as benign or malignant brain tumors. An unnatural development of tissue is a tumor.

A benign tumor does not contain cancerous cells and typically does not come back after removal. The majority of benign brain tumors do not infect adjacent tissue because they have distinct borders. However, due to their size and placement in the brain, these tumors can produce symptoms that are comparable to those of malignant tumors. cancerous brain tumors.



**Fig 31. Brain Tumor**

* 1. **Breast Cancer**

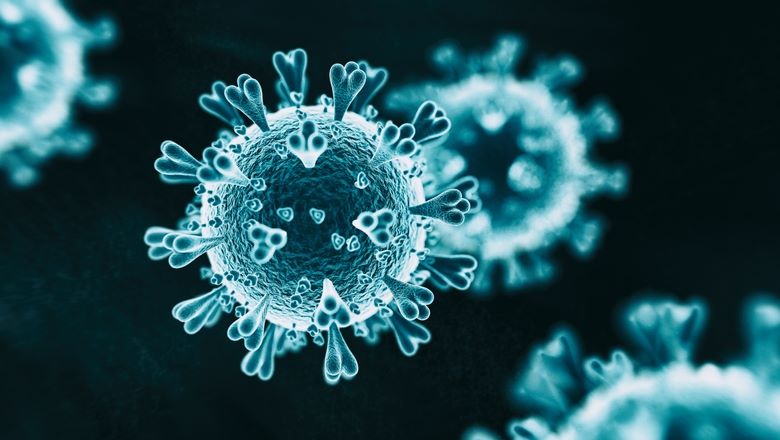
Breast cancer is a type of cancer that affects the cells in the breast. It is the most common cancer in women and the second most common cause of cancer death in women, after lung cancer. There are several different types of breast cancer, but the most common is ductal carcinoma, which begins in the cells of the milk ducts. Risk factors for breast cancer include being female, older age, a family history of breast cancer, and certain genetic mutations. Symptoms of breast cancer may include a lump or thickening in the breast or under the arm, changes in the size or shape of the breast, and discharge from the nipple. Treatment for breast cancer may include surgery, chemotherapy, radiation therapy, and hormonal therapy.



**Fig 32. Breast Cancer**

* 1. **Covid – 19**

COVID-19 is a infectious disease caused by the severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2). It was first identified in Wuhan, China in December 2019 and has since become a global pandemic. The symptoms of COVID-19 can range from mild to severe and may include fever, cough, difficulty breathing, and body aches. The virus is primarily spread through respiratory droplets produced when an infected person speaks, coughs, or sneezes. There is currently no specific treatment for COVID-19, and most people who contract the virus recover on their own. However, some people, particularly older adults and those with underlying health conditions, are at higher risk of severe illness or death from COVID-19.



**Fig 33. Covid - 19**

* 1. **Diabetes**

Diabetes is a long-term disease that alters how the body uses sugar (glucose). Type 1 diabetes and type 2 diabetes are the two basic kinds of the disease. The immune system of the body targets and kills the cells in the pancreas that make insulin in type 1 diabetes, an autoimmune illness. A hormone called insulin controls how much sugar is in the blood. Type 1 diabetics must inject themselves with insulin or use an insulin pump to manage their blood sugar levels.

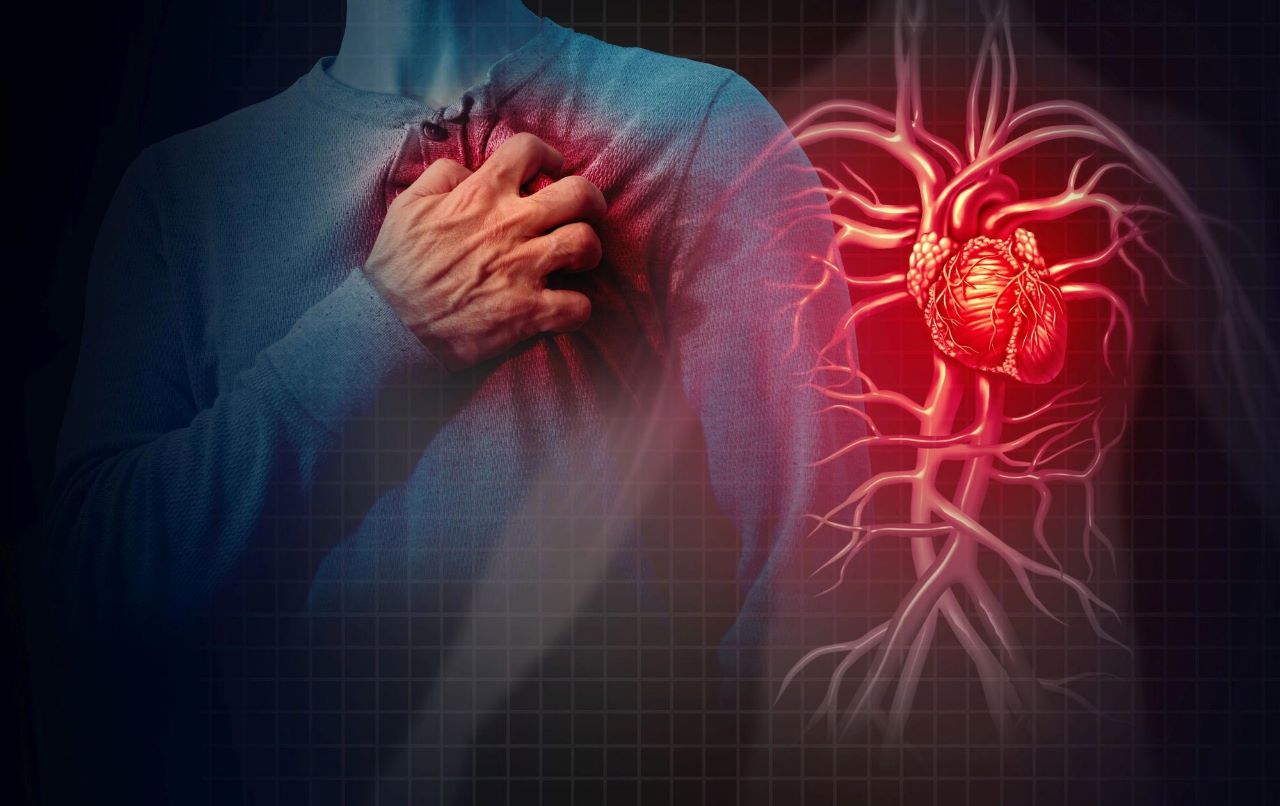
Insulin resistance and/or insufficient insulin production are two metabolic disorders that can lead to type 2 diabetes. Blood sugar levels may rise as a result of this. A typical treatment for type 2 diabetes is a medicine, nutrition, and exercise together. If left untreated, both types of diabetes can result in major health issues such heart disease, nerve damage, blindness, kidney failure, and amputations.



**Fig 34. Diabetes**

* 1. **Heart Disease**

Heart disease is a term used to describe a range of conditions that affect the heart. Some common forms of heart disease include coronary artery disease, heart attack, heart failure, and arrhythmias. Risk factors for heart disease include high blood pressure, high cholesterol, diabetes, smoking, obesity, and a family history of heart disease. Symptoms of heart disease may include chest pain, shortness of breath, and palpitations. Treatment for heart disease may include lifestyle changes, medication, and medical procedures.

****

**Fig 35. Heart Disease**

* 1. **Pneumonia**

Pneumonia is an infection of the lungs that can be caused by a variety of microorganisms, including bacteria, viruses, fungi, and parasites. It is characterized by inflammation of the lung tissue and the filling of the alveoli (small air sacs) with fluid or pus. Pneumonia can range in severity from mild to severe, and it can be life-threatening, particularly in older adults, young children, and people with weakened immune systems. Symptoms of pneumonia may include fever, cough, difficulty breathing, chest pain, and fatigue. Pneumonia is usually treated with antibiotics, although viral pneumonia may not respond to these medications. Other treatments may include oxygen therapy, bronchodilators to open the airways, and supportive care to manage symptoms.



**Fig 36. Pneumonia**

**CHAPTER 5**

**5.1 REQUIREMENTS**

**5.1.1 Functional Requirements**

* GUI(Graphical User Interface) with User

**5.1.2 Software Requirements**

* Python (version 3.6 or higher)
* Anaconda (for various python libraries)
* TensorFlow
* Flask(backend)
* Html(front end)
* CSS
* Visual Studio Code(IDE)
* OpenCV

**5.1.3 Hardware Requirements**

* Processor: Pentium IV or higher
* RAM: 256 MB
* Space on Hard Disk: minimum 512MB

**5.1.4 Debugger and Emulator**

* Any Browser(Particularly Chrome)

**5.2 FEASIBILITY**

**5.2.1 Economical**

* Since our concept is entirely software-based and deals with image processing and identification, its implementation is cost-effective.

**5.2.2. Time**

* Overall Project Timeline
* Time allotted for planning and strategy
* Number of project Modules
* Deadline of Each Module
* Short term goals
* Long Term goals
* Effective model so that implementation and testing goes hand in hand

**5.2.3. Scope**

* Social
* Time Saving
* economical
* reliability
* Health

**CHAPTER 6**

**RESULTS AND DISCUSSIONS**

|  |  |  |
| --- | --- | --- |
| **Disease** | **Output Format** | **Accuracy Rate(%)** |
| Alzheimer Detection | Mild Demented  Moderate Demented  Non Demented  Very Mild Demented | 73.54% |
| Brain Tumor Detection | Tumor  No Tumor | 100% |
| Breast Cancer Detection | Benign(not cancer)  Malignant(Cancer) | 91.81% |
| Covid-19 Detection | Covid Positive  Covid Negative | 93% |
| Diabetes Detection | Positive  Negative | 66.8% |
| Heart Disease Detection | Positive  Negative | 86.96% |
| Pneumonia Detection | Positive Negative | 83.17% |

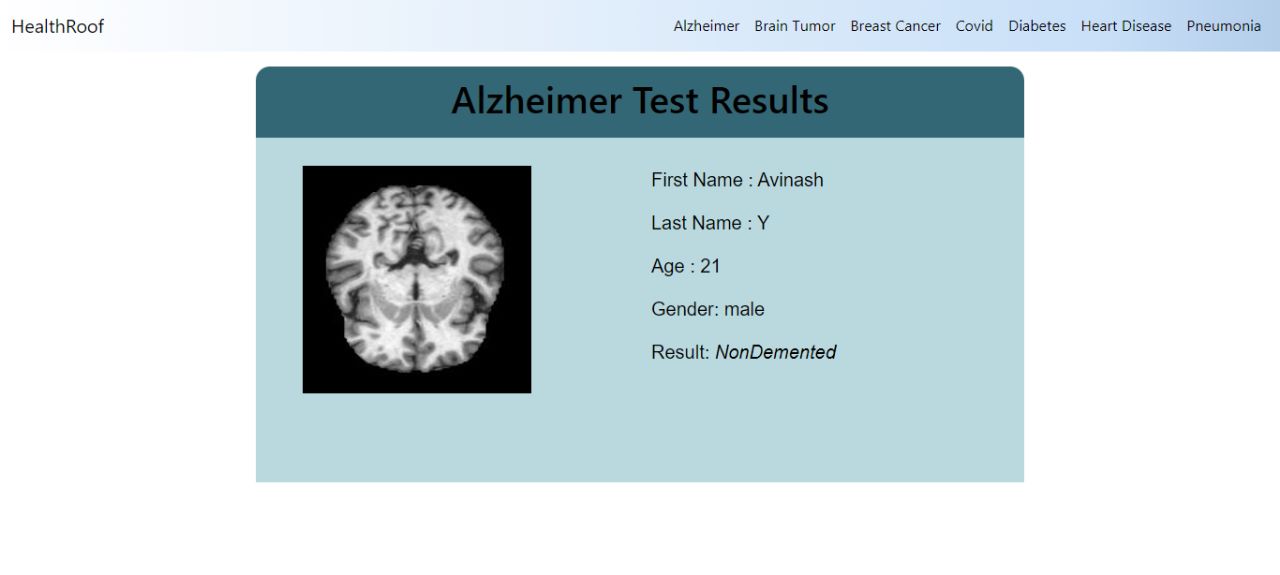
**Table 2 : Diseases - Accuracies**

**DETECTED RESULTS**

**6.1.Alzheimer Detection**

**6.1.1 Risk Factors**

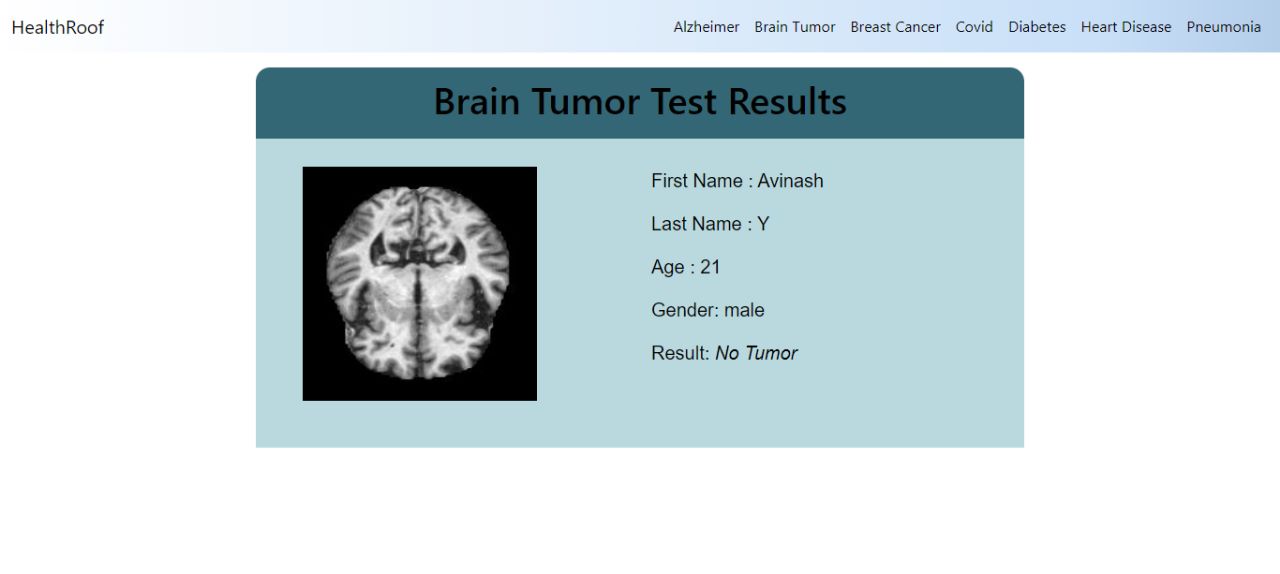
* Increasing Age
* Family history and genetics
* Down syndrome results in Alzheimer’s
* Mild Cognitive Impairment
* Head trauma
* Air Pollution
* Excessive alcohol consumption
* Poor sleep patterns
* Lifestyle and heart health
* Lifelong learning and social engagement

**Fig 37. Alzheimer Result**

**6.2.Brain Tumor Detection**

**6.2.1 Risk Factors**

* **Radioactive contamination:** Ionizing radiation is a category of radiation that increases the risk of brain tumors in those who have been exposed to it. Ionizing radiation examples include atomic bomb radiation exposure and radiation therapy used to cure cancer.
* **Family history:** A tiny percentage of brain tumors are found in patients who have a family history of the disease or genetic abnormalities that enhance the chance of developing brain tumors.

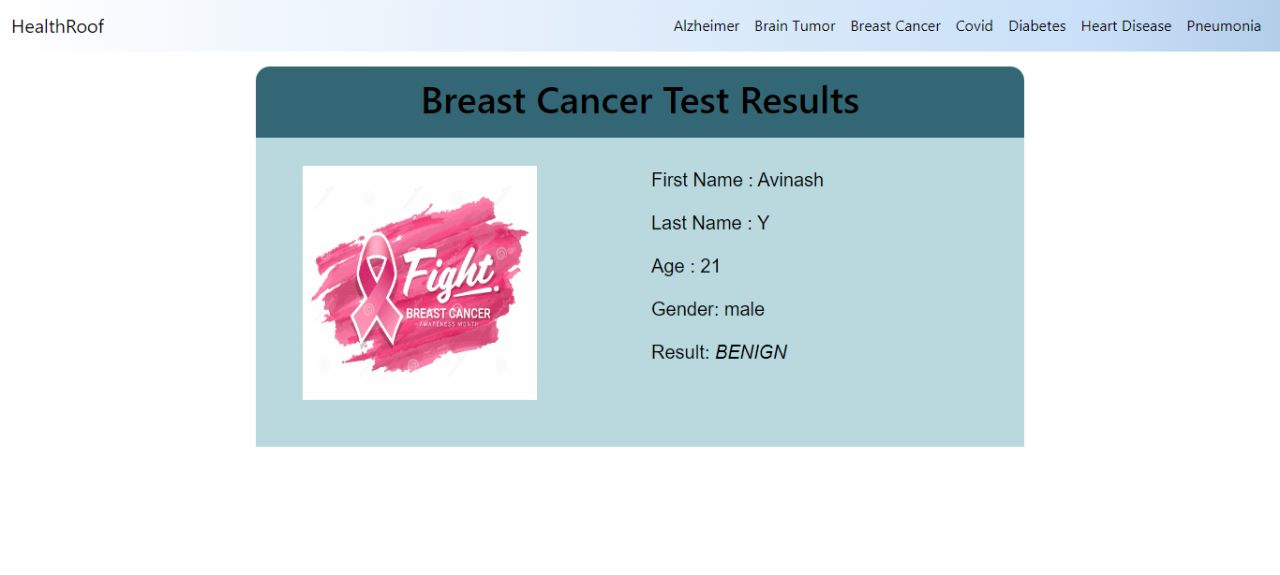
****

**Fig 38. Brain Tumor Result**

**6.3.Breast Cancer Detection**

**6.3.1 Risk Factors**

* Being female
* Increasing age
* A personal history of breast condition
* A personal history of breast cancer
* A family history of breast cancer
* Inherited genes that increase cancer risk
* Radiation exposure
* Obesity
* Beginning your period at a younger age
* Beginning menopause at an older age.
* Having your first child at an older age.
* Having never been pregnant.
* Postmenopausal hormone therapy.
* Drinking alcohol.

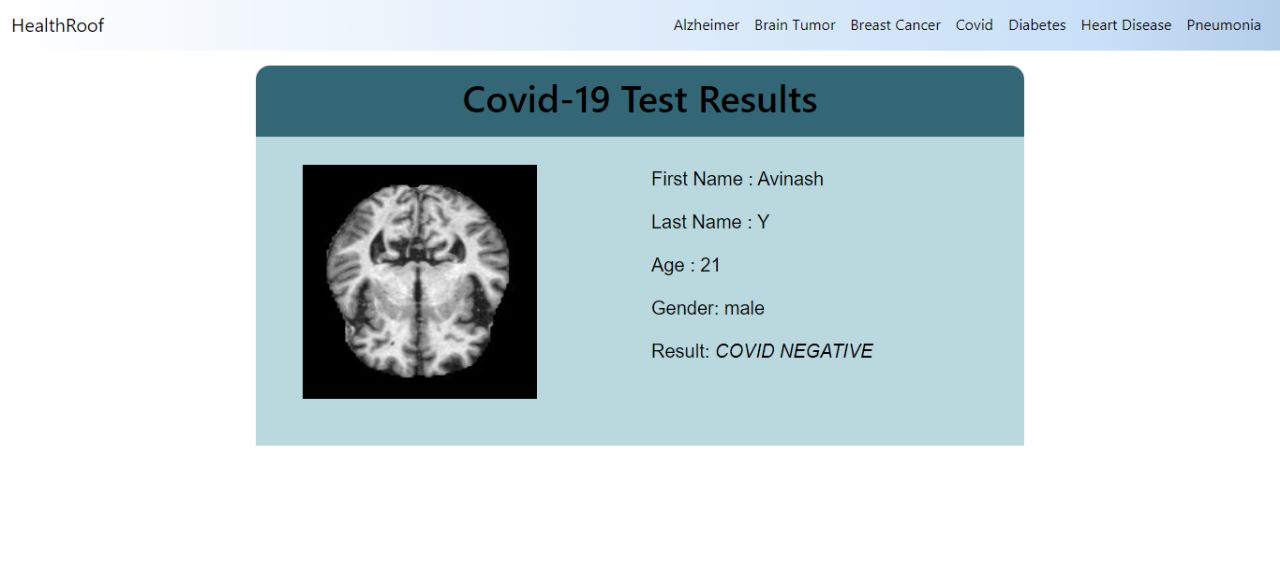
****

**Fig 39. Breast Cancer Result**

**6.4.Covid – 19 Detection**

**6.4.1 Risk Factors**

* Close contact (within 6 feet, or 2 meters) with someone who has COVID-19
* Being coughed or sneezed on by an infected person

****

**Fig 40. Covid – 19 Result**

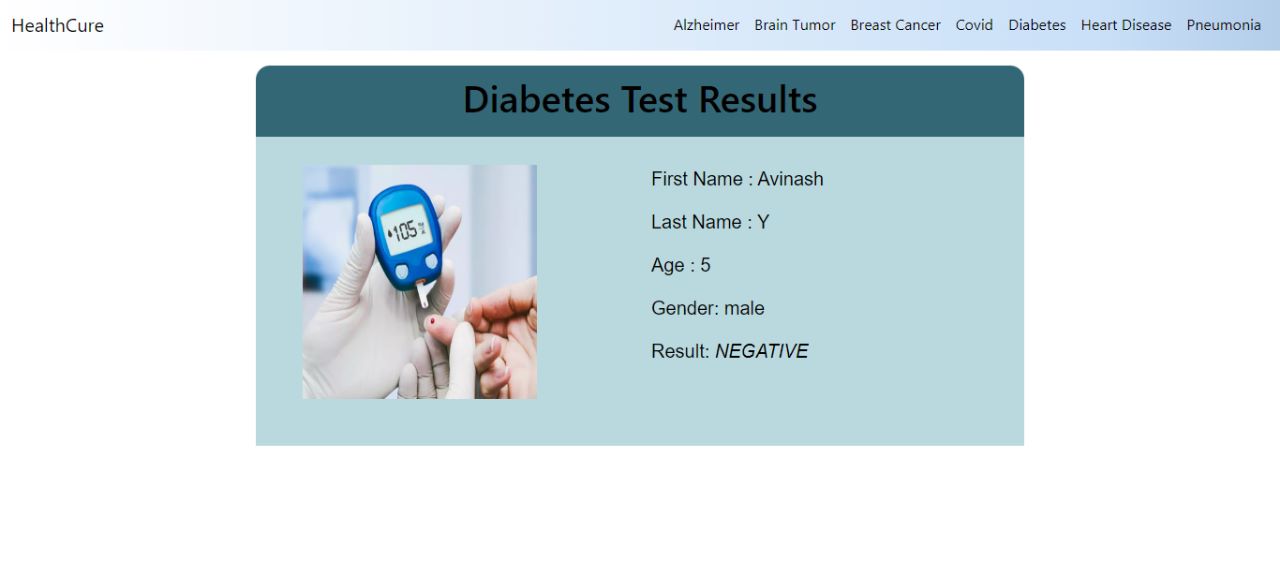
**6.5 Diabetes Detection**

**6.5.1 Risk Factors for type 1 diabetes**

* Family history.
* Environmental factors.
* The presence of damaging immune system cells (autoantibodies)
* Geography

**6.5.2 Risk Factors for type 2 diabetes**

* Weight
* Inactivity
* Family History
* Race or ethnicity
* Age
* Gestational diabetes
* Polycystic ovary syndrome
* High blood pressure
* Abnormal cholesterol and triglyceride levels

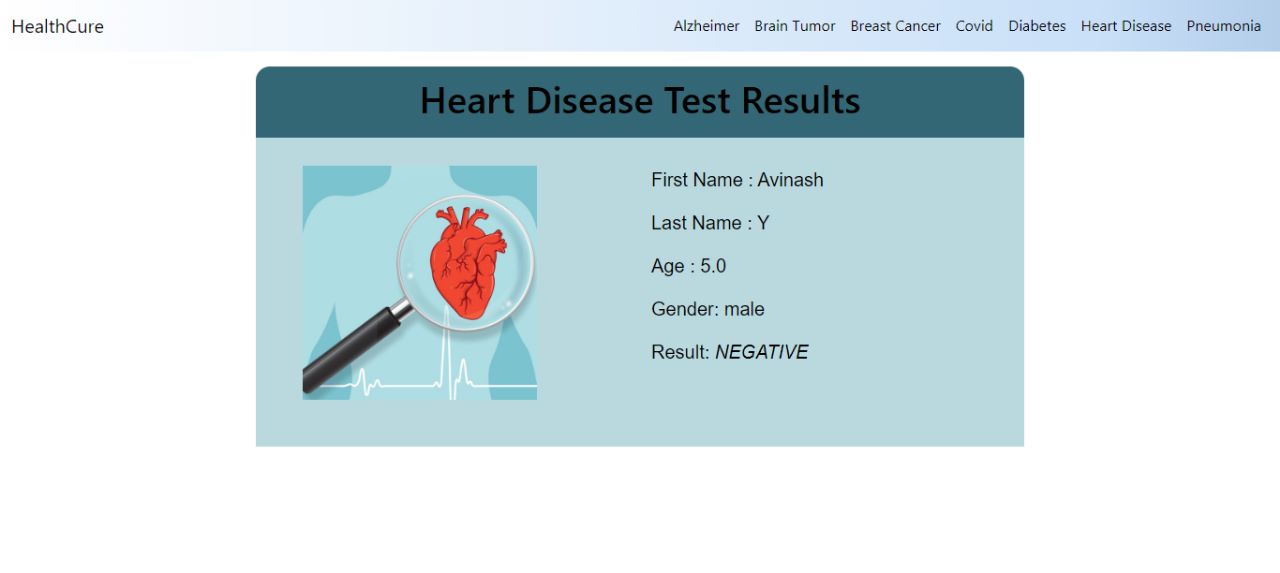
****

**Fig 41. Diabetes Result**

* 1. **Heart Disease Detection**

**6.6.1 Risk Factors**

* Age
* Sex
* Family history
* Smoking
* Poor diet
* High blood pressure
* High blood cholesterol levels
* Diabetes
* Obesity
* Physical inactivity
* Stress
* Poor dental health

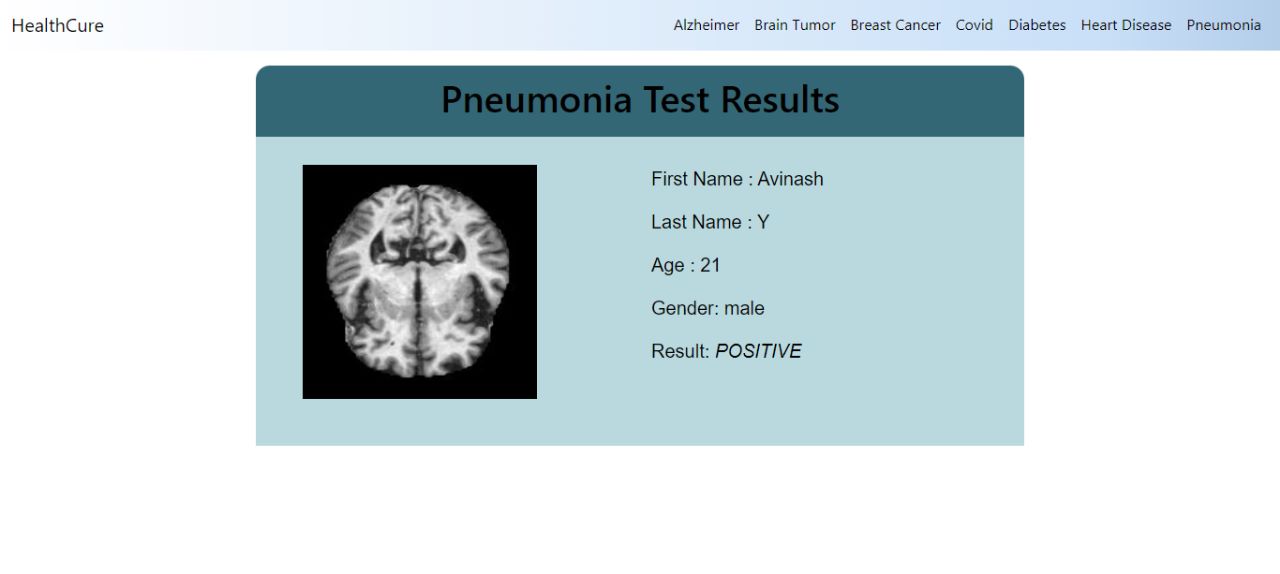


**Fig 42. Heart Disease Result**

* 1. **Pneumonia Detection**

**6.7.1 Risk Factors**

* Being hospitalized
* Chronic disease
* Smoking
* Weakened or suppressed immune system



**Fig 43. Pneumonia Result**

**CHAPTER 7**

**CONCLUSION AND FUTURE WORK**

Currently, there isn't a single test that can identify every condition. This is due to the fact that various diseases can impact the body in various ways and have various origins. As a result, numerous tests are required to correctly diagnose various disorders. One area that requires the assistance of human professionals is disease identification; this process is streamlined by Health Roof.  
  
 In general, a comprehensive medical history and physical examination are the first steps in the diagnosis of a condition. A healthcare professional may request more tests, such as blood tests, imaging scans, or biopsies, in order to confirm or exclude a diagnosis based on the findings of these first tests.

It is crucial to remember that early disease detection and diagnosis can enhance treatment outcomes and raise the likelihood of a full recovery. As a result, it's critical to be knowledgeable about the warning signs and symptoms of common illnesses and to schedule routine checkups and screenings with a healthcare professional.

It will be more productive to increase the number of diseases that can be identified using X-ray scans or simply by entering simple data. We can add further features, such as the ability to display warnings if a person is discovered to be positive. Even can be changed by giving users the option to schedule appointments online based on their severity.

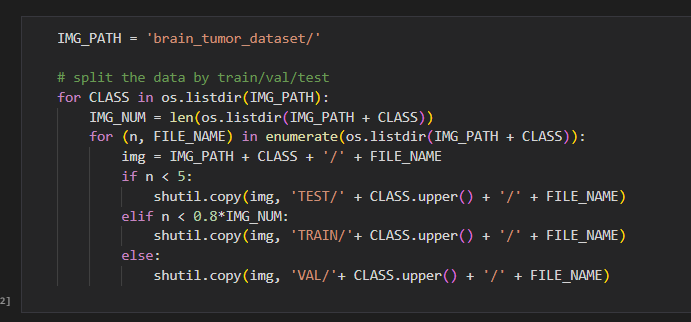
An all-inclusive medical solution is **“Health Roof – All in One Solution”**. Using the power of AI, this project unifies the diagnosis of 7 diseases onto a single platform. The key benefit of this project is that we can quickly and easily receive the test results at home.

Top of Form

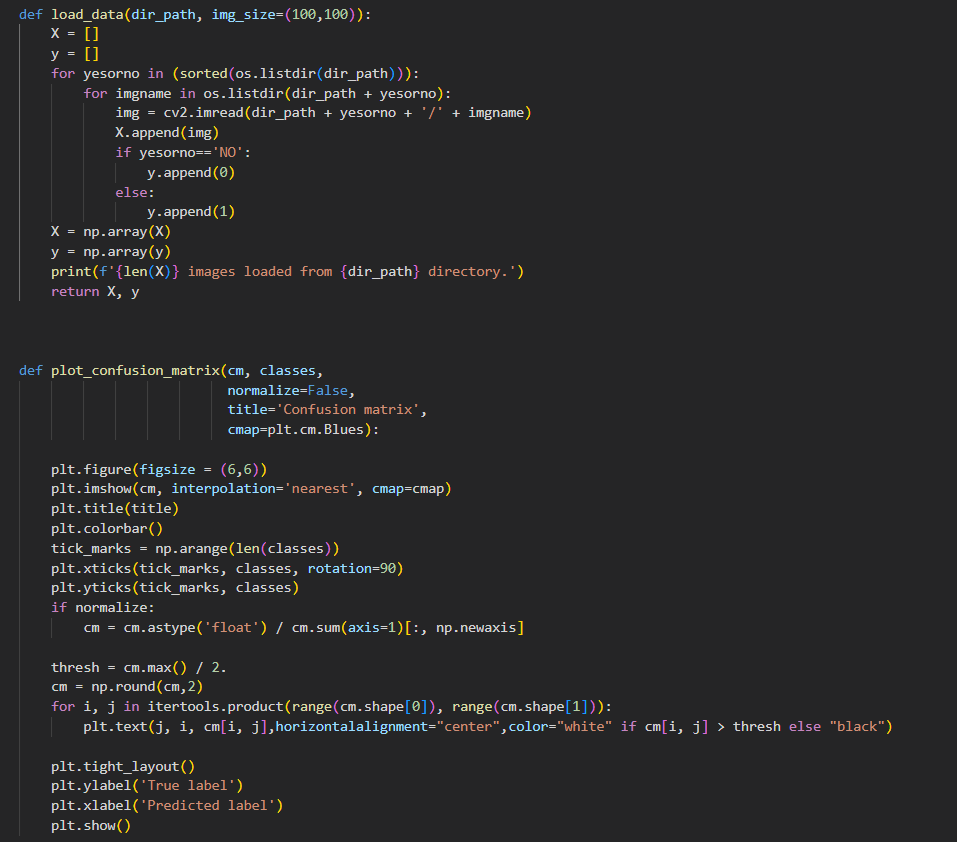
**CHAPTER 8**

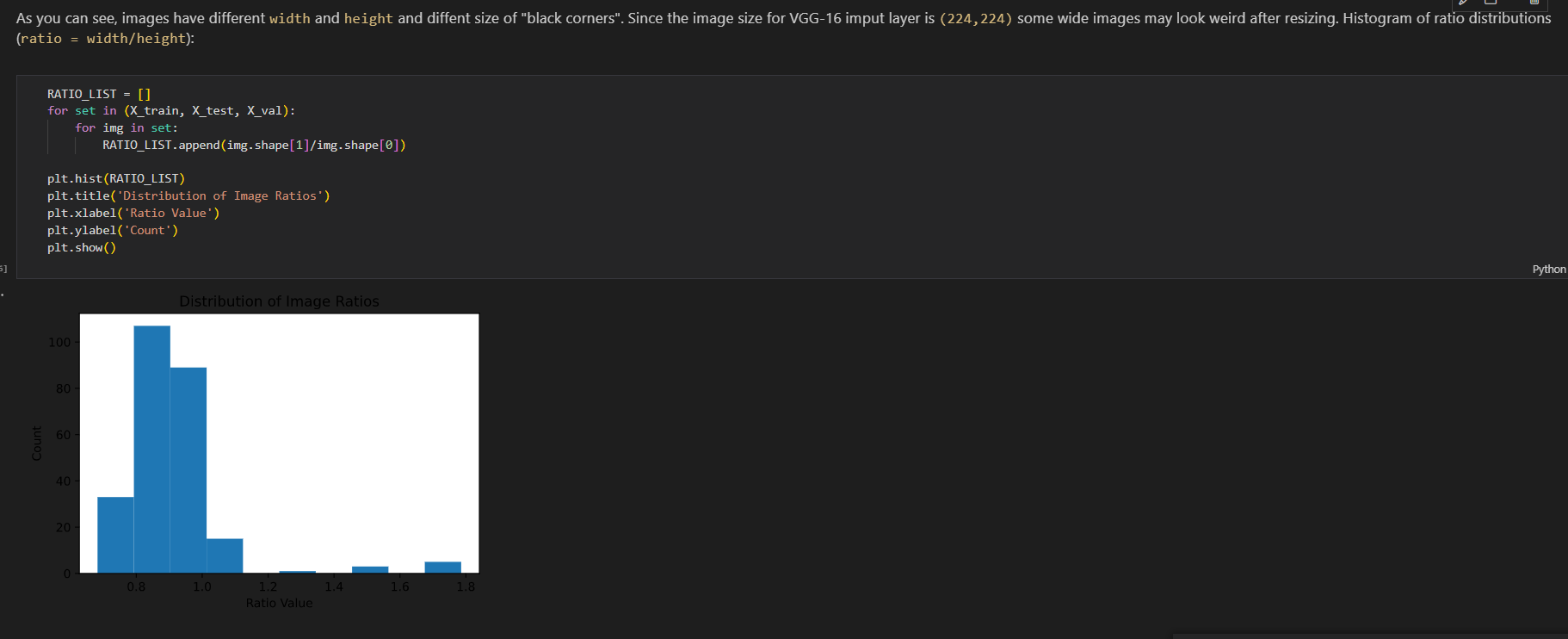
**APPENDIX**

**Splitting data into test and train**

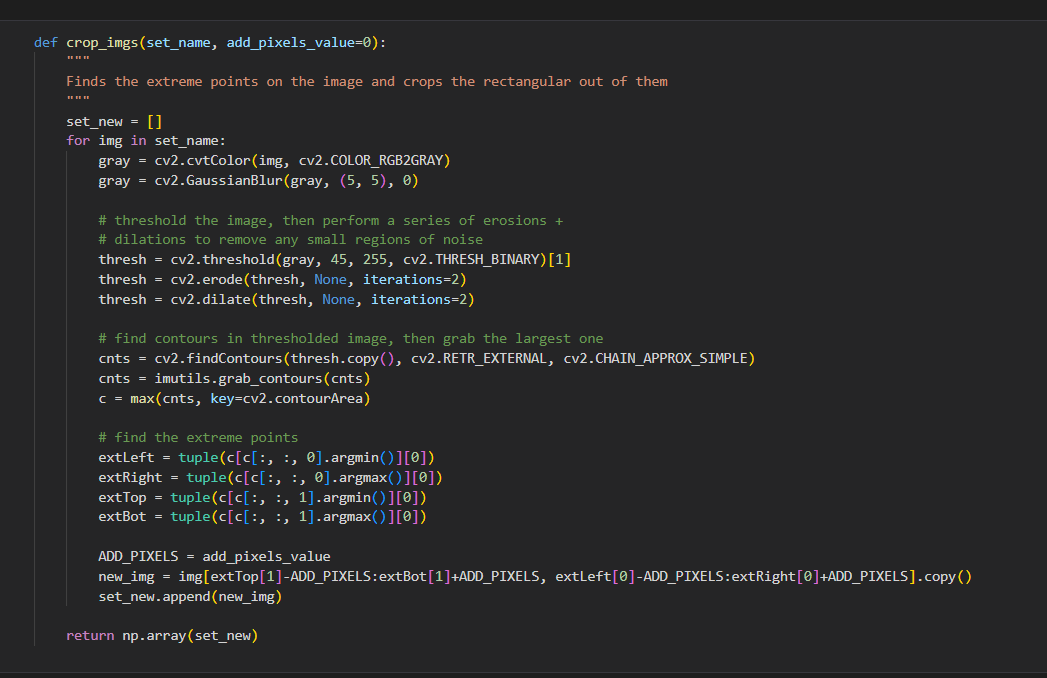
****

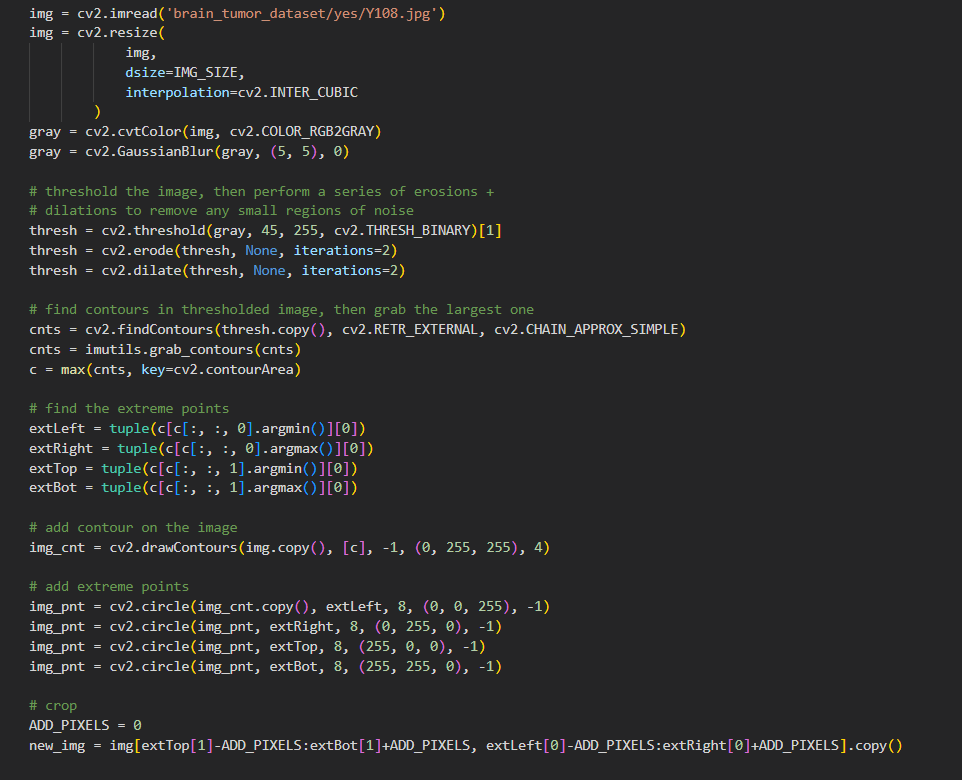
**Data Import and Pre-Processing**

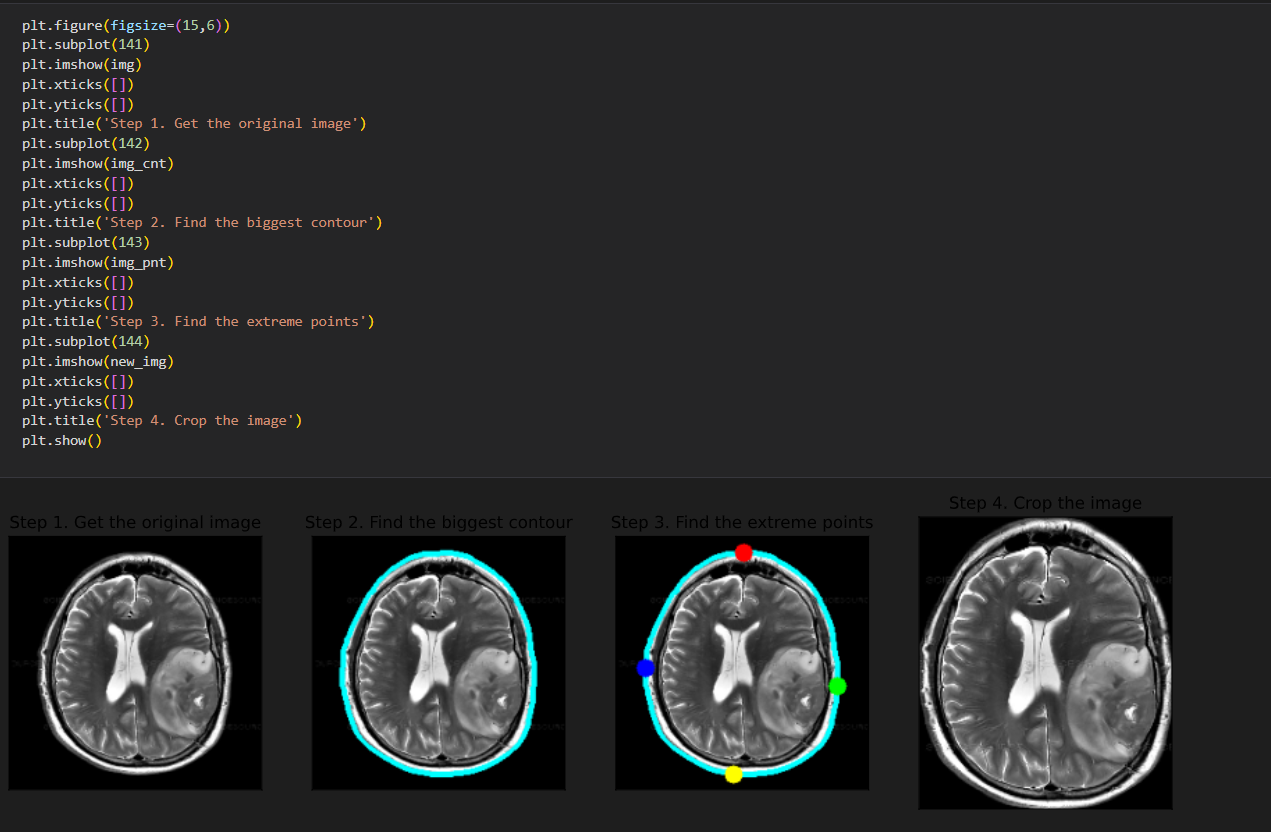
****

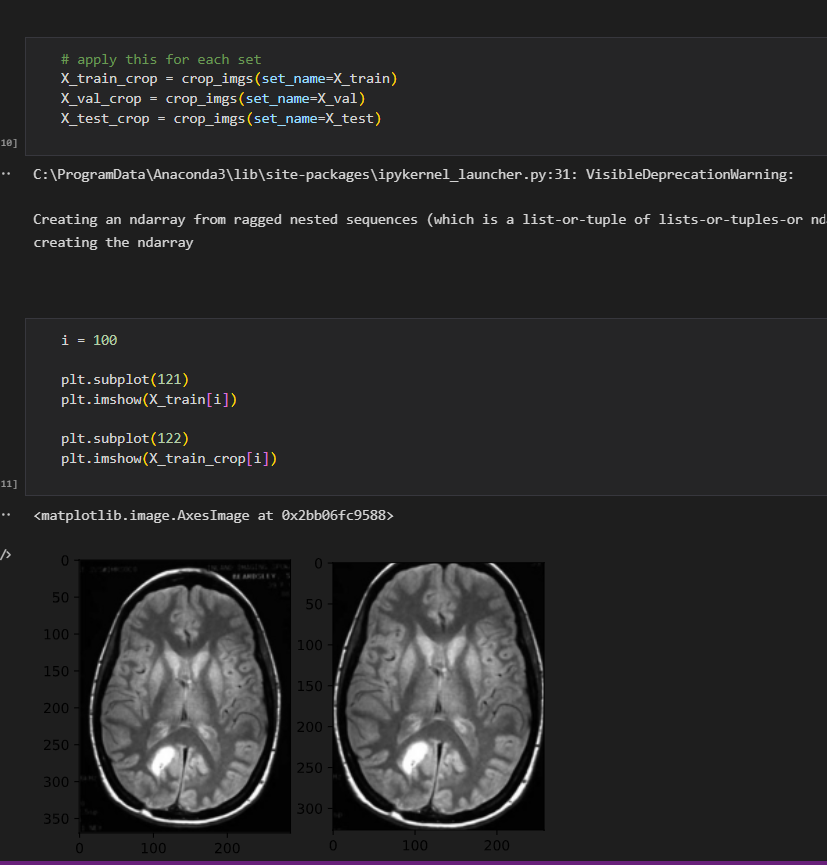
****

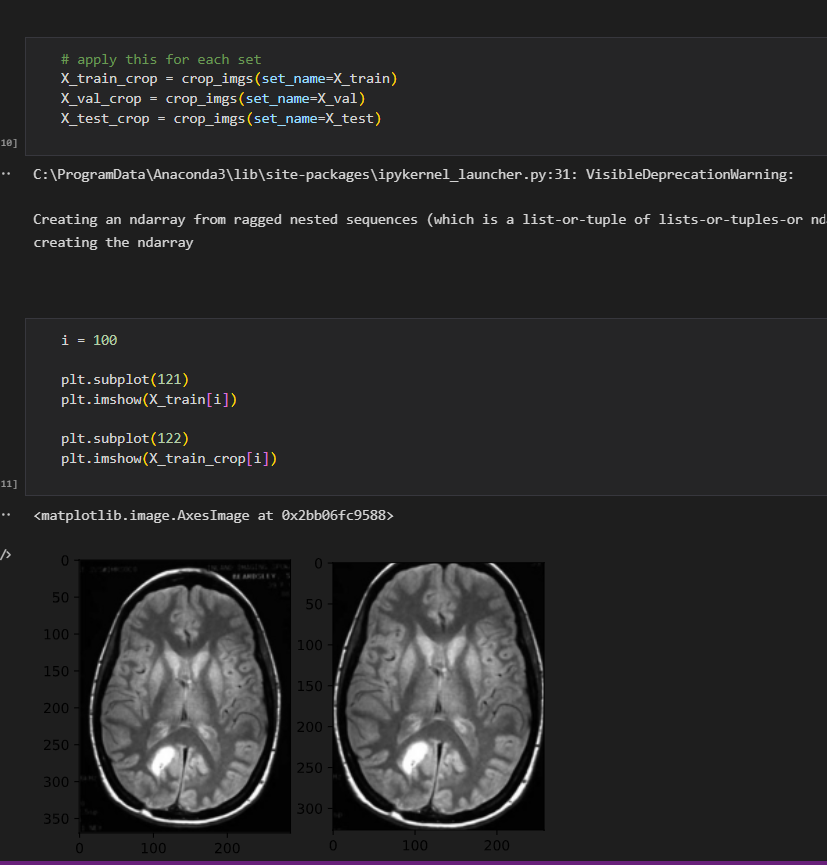
**Crop main part of image**

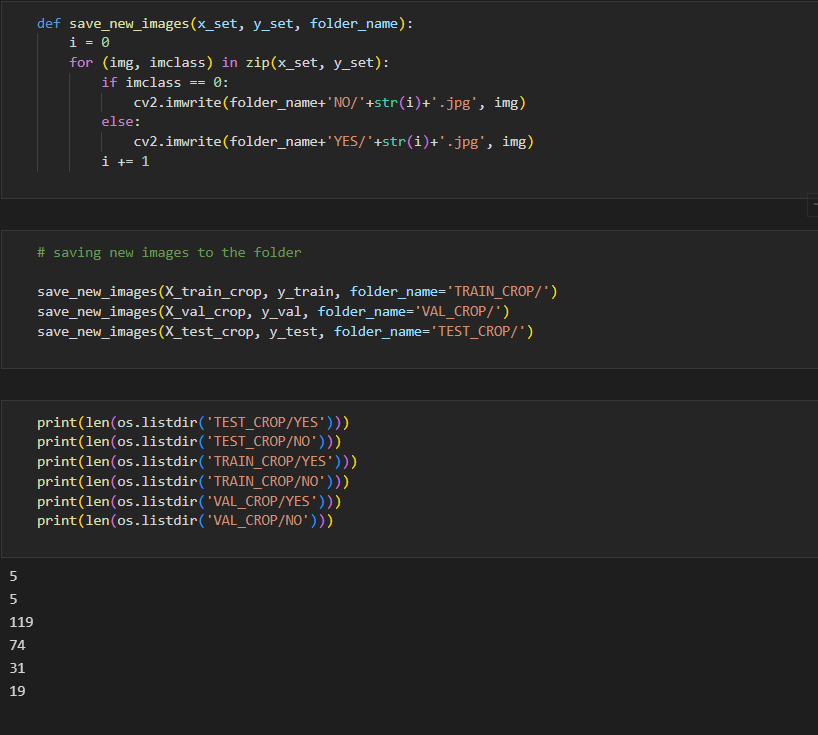
****

****

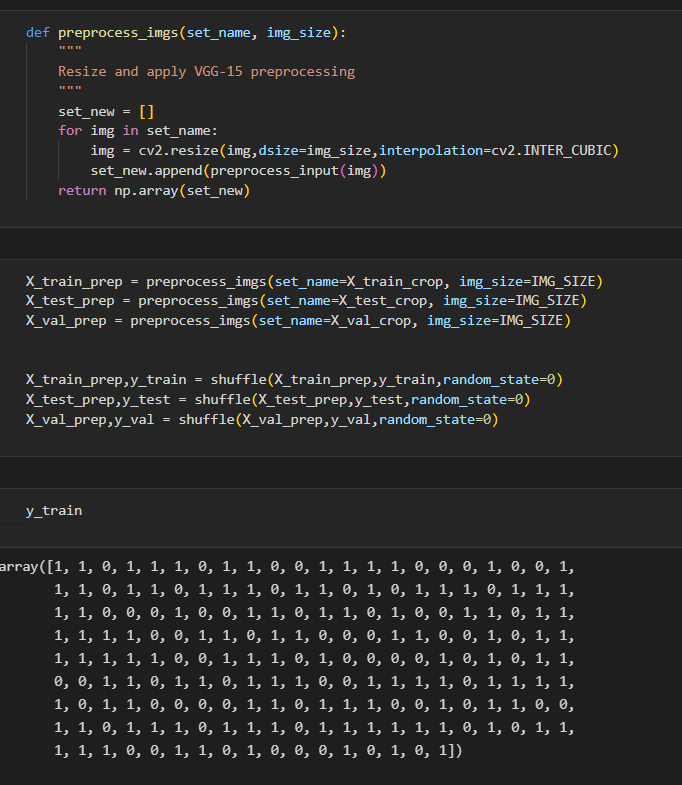
****

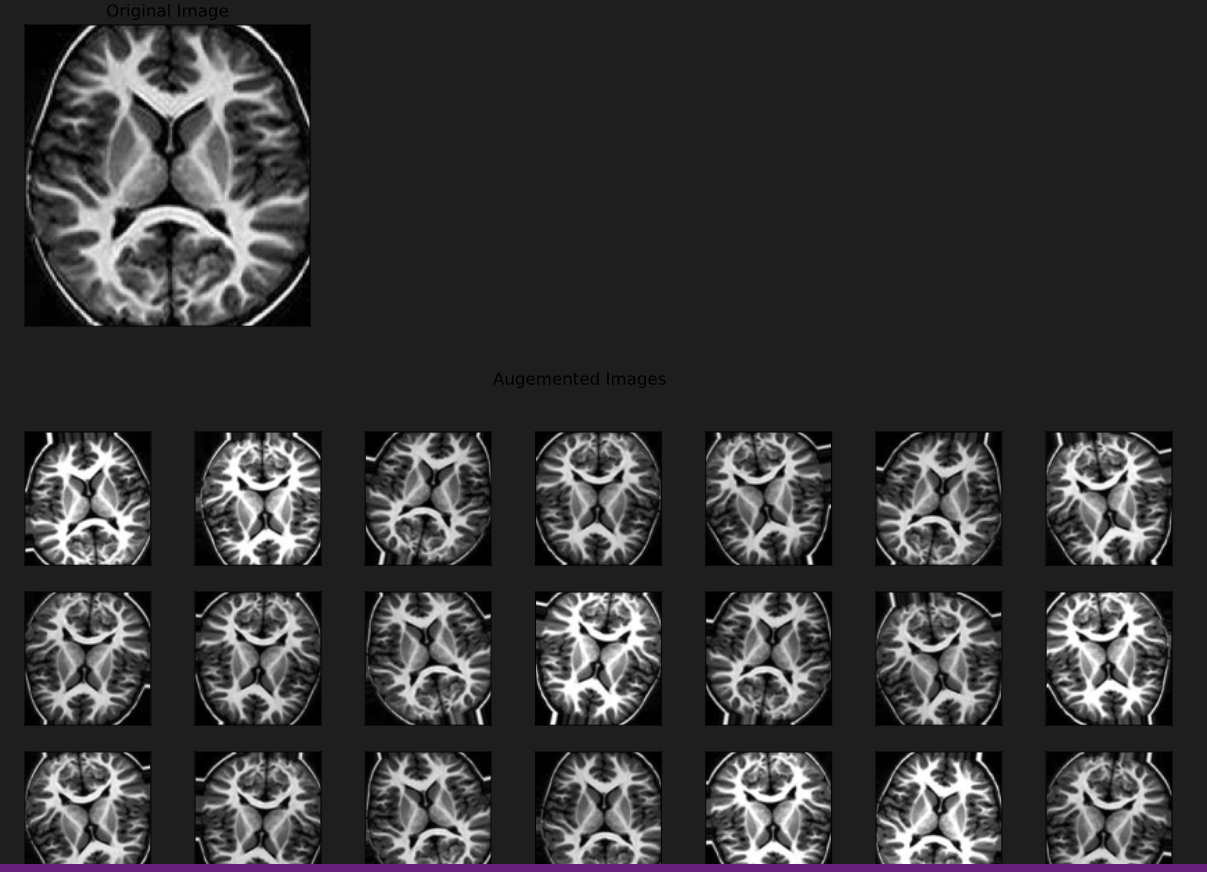
****

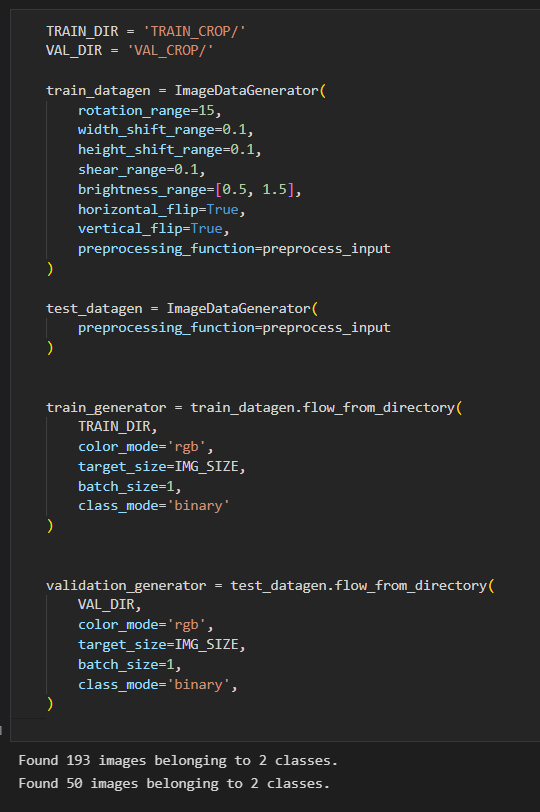
****

****

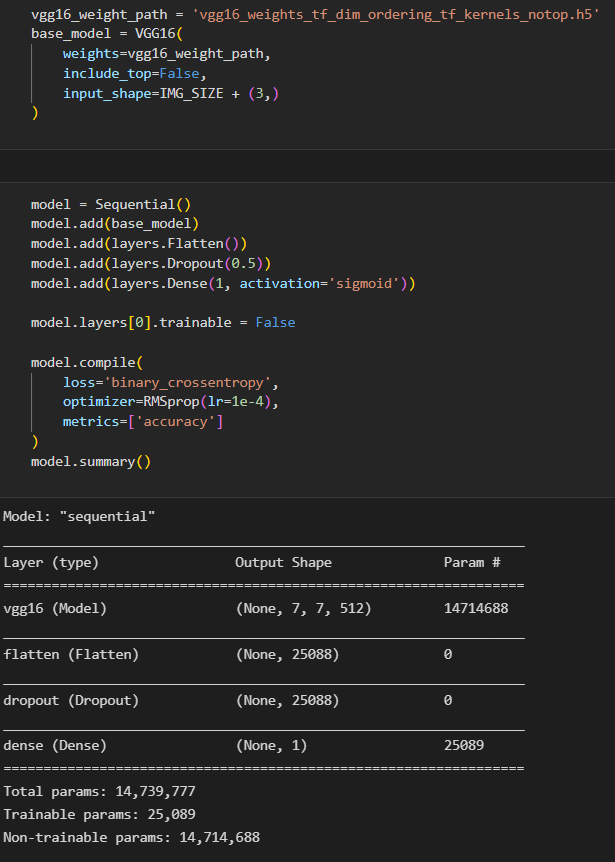
**Resizing and applying pre-processing required for respective algorithm**

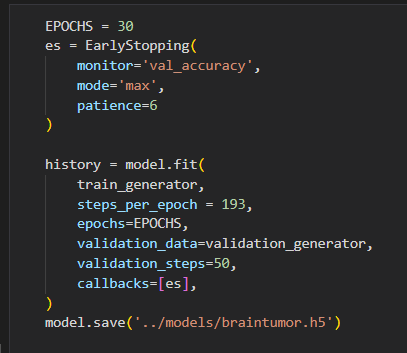
****

****

****

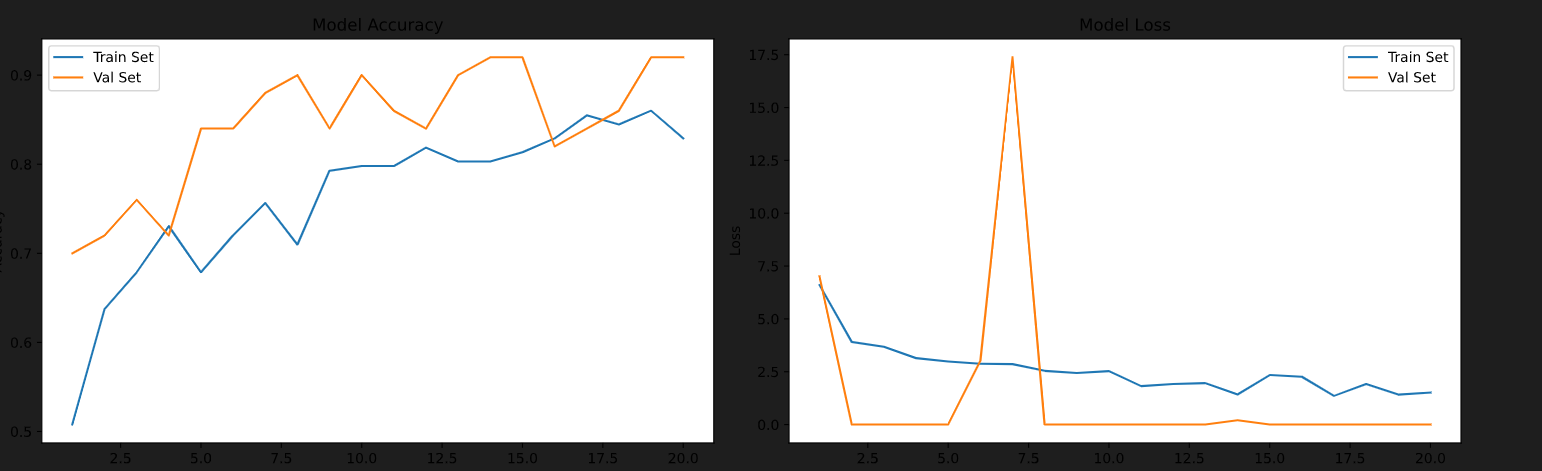
**Model Building**

****

****

**Plot Model Performance**

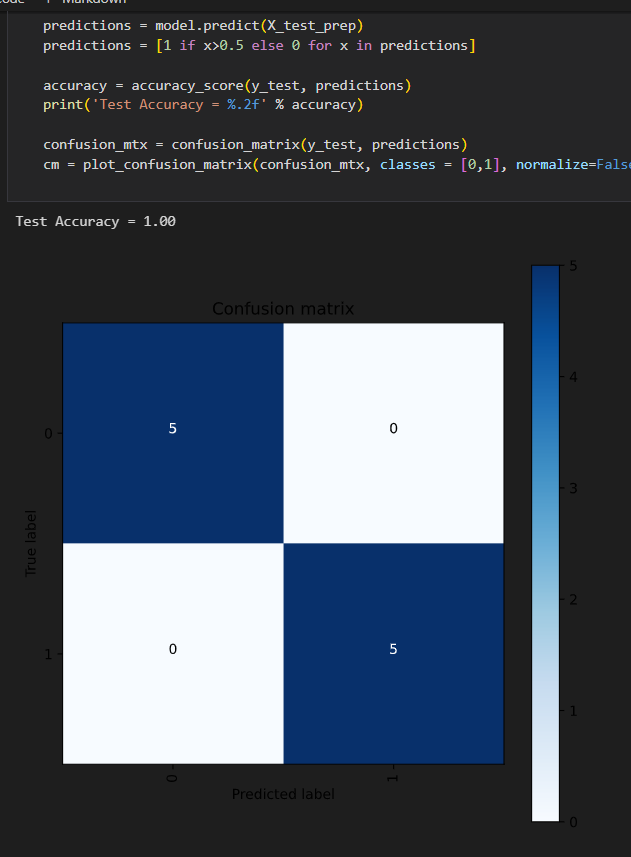
****

****

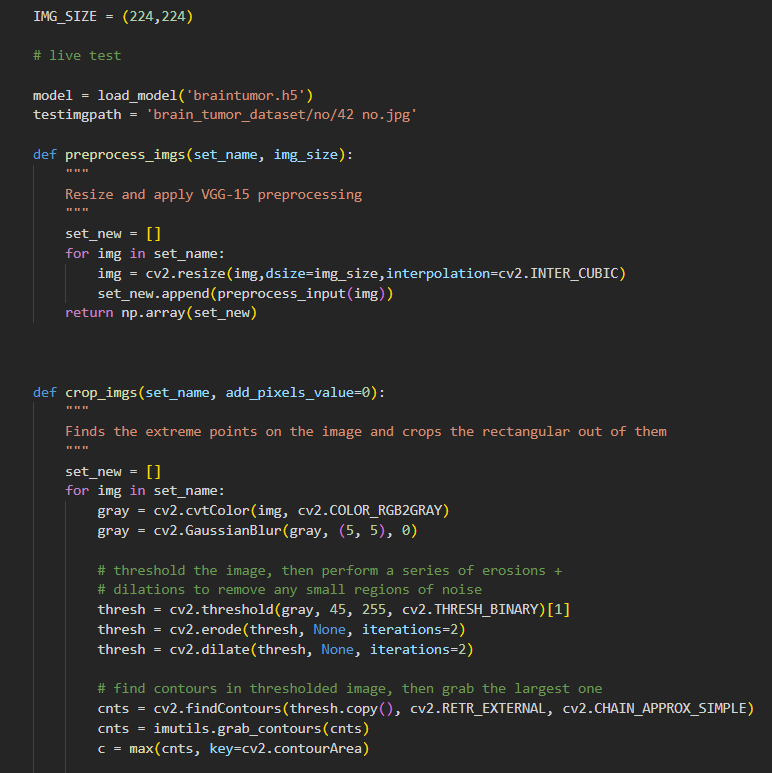
**Validate on validation set**

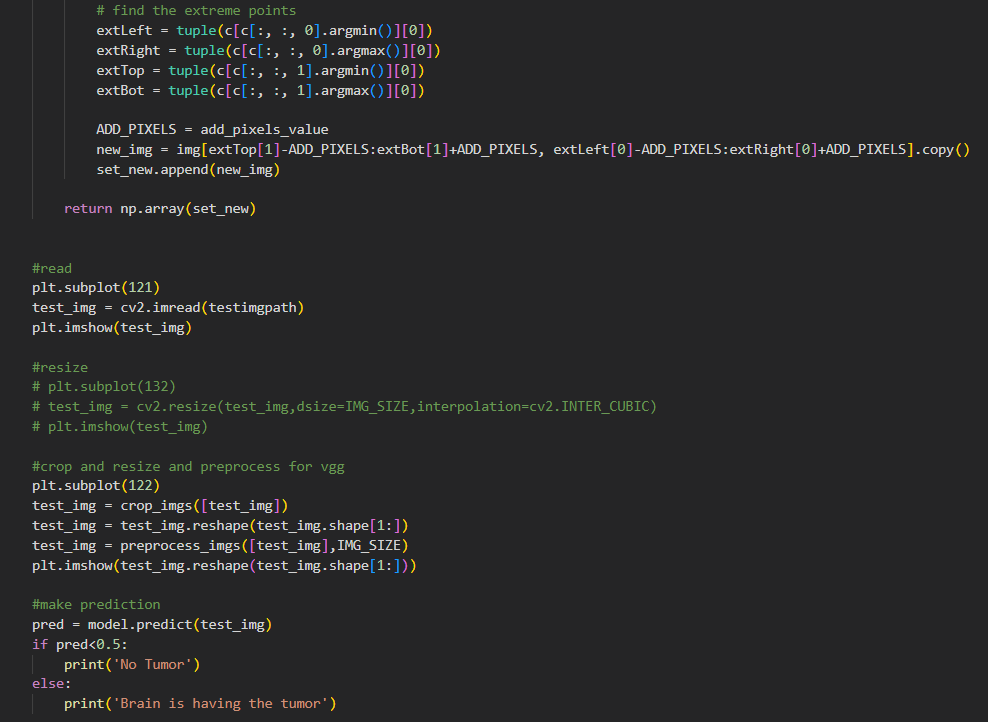
****

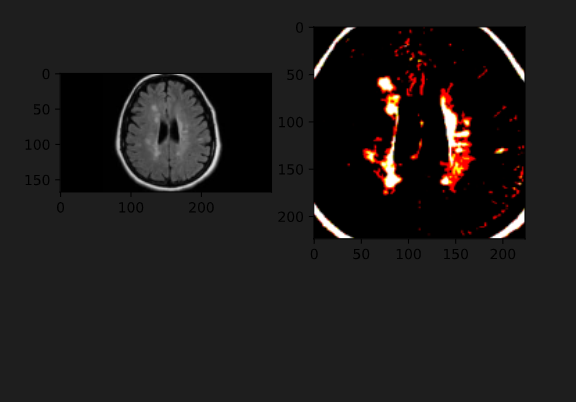
**Validate on test set**

****

**Conclusions**

****

****

****

**REFERENCES**

[1] B. Reyers and E. R. Selig, ‘‘Global targets that reveal the socialecological interdependencies of sustainable development,’’ Nature Ecol. Evol., vol. 4,no. 8, pp. 1011–1019, Jul. 2020.

[2] F. He and S. P. Hubbell, ‘‘Species—Area relationships always overestimate extinction rates from habitat loss,’’ Nature, vol. 473, no. 7347, pp. 368–371, May 2011.

[3] A. Perrin, A. Khimoun, B. Faivre, A. Ollivier, N. de Pracontal, F. Théron,M. Loubon, G. Leblond, O. Duron, and S. Garnier, ‘‘Habitat fragmentation differentially shape neutral and immune gene variation in a tropical bird species,’’ Heredity, vol. 126, no. 1, pp. 148–162, Sep. 2020.

[4] D. A. Wilkinson, J. C. Marshall, N. P. French, and D. T. S. Hayman, ‘‘Habitat fragmentation, biodiversity loss and the risk of novel infectious disease emergence,’’ J. Roy. Soc. Interface, vol. 15, pp. 1– 10, Dec. 2018.

[5] G. Strona and K. D. Lafferty, ‘‘Environmental change makes robust ecological networks fragile,’’ Nature Commun., vol. 7, no. 1, pp. 1–7, Aug. 2016.

[6] M. G. Betts, C. Wolf, M. Pfeifer, C. Banks-Leite, V. ArroyoRodríguez, D. B. Ribeiro, J. Barlow, F. Eigenbrod, D. Faria, R. J. Fletcher, and A. S. Hadley, ‘‘Extinction filters mediate the global effects of habitat fragmentation on animals,’’ Science, vol. 366, pp. 1236–1239, Dec. 2019.

[7] T. Telenský, P. Klvaňa, M. Jelínek, J. Cepák, and J. Reif, ‘‘The influence of climate variability on demographic rates of avian afroPalearctic migrants,’’ Sci. Rep., vol. 10, no. 1, pp. 1–11, Oct. 2020.

[8] J. T. Ibarra and K. Martin, ‘‘Biotic homogenization: Loss of avian functional richness and habitat specialists in disturbed Andean temperate forests,’’ Biol. Conservation, vol. 192, pp. 418–427, Dec. 2015.

[9] J. L. Blanchard, ‘‘A rewired food web,’’ Nature, vol. 527, no. 7577, pp. 173–174, Nov. 2015.

[10] R. Inger, E. Per, D. T. C. Cox, and K. J. Gaston, ‘‘Key role in ecosystem functioning of scavengers reliant on a single common species,’’ Sci. Rep., vol. 6, no. 1, pp. 1–5, Jul. 2016.

[11] R. S. C. Cooke, F. Eigenbrod, and A. E. Bates, ‘‘Projected losses of global mammal and bird ecological strategies,’’ Nature Commun., vol. 10, no. 1, pp. 1–8, May 2019.

[12] A. A. A. Bovo, K. M. P. M. B. Ferraz, M. Magioli, E. R. Alexandrino, É. Hasui, M. C. Ribeiro, and J. A. Tobias, ‘‘Habitat fragmentation narrows the distribution of avian functional traits associated with seed dispersal in tropical forest,’’ Perspect. Ecol. Conservation, vol. 16, no. 2, pp. 90–96,Apr. 2018.

[13] (2017). BirdLife International. IUCN Red List for Birds. Version 2017.1.BirdLife International. [Online].

[14] Data Zone. State of the World’s Birds: 2020 Annual Update.

**THANK YOU**