Project Title: Automated Detection of Covid-19 from CT Scans

ME 781 Course Project: Final report
Guide: Prof. Asim Tewari

Team members:

Naman Agarwal	170100049
Akhilesh Goyal	184100002
Manish	193020021
Bachalakuri Sowmya	180050018
Abhishek Narayan Chaudhury	19i190005

Problem Objective and Introduction

- We are a Mumbai based startup in the field of **AI-Healthcare** aiming to **redefine the healthcare services in India** and become a strong pillar in providing **reliable healthcare diagnostic services**

- We have developed a **robust** tool for early, fast, accurate and automated detection of COVID-19 infected patients using sophisticated Deep Learning algorithms

- We plan to enter the **B2B market** of the Indian healthcare industry with our healthcare solutions in medical image diagnostics like of CT scans, X-Rays, MRIs. Our startup aims to **serve people** in this time of need as **fast and cheap** as possible

- Our current aim is to **provide affordable and reliable diagnosis solutions to hospitals and doctors**, both in government as well as private institutions, for **Covid 19 detection** based on analysing CT scans of the patients. We plan to expand our products and services for other disease detection in near future.

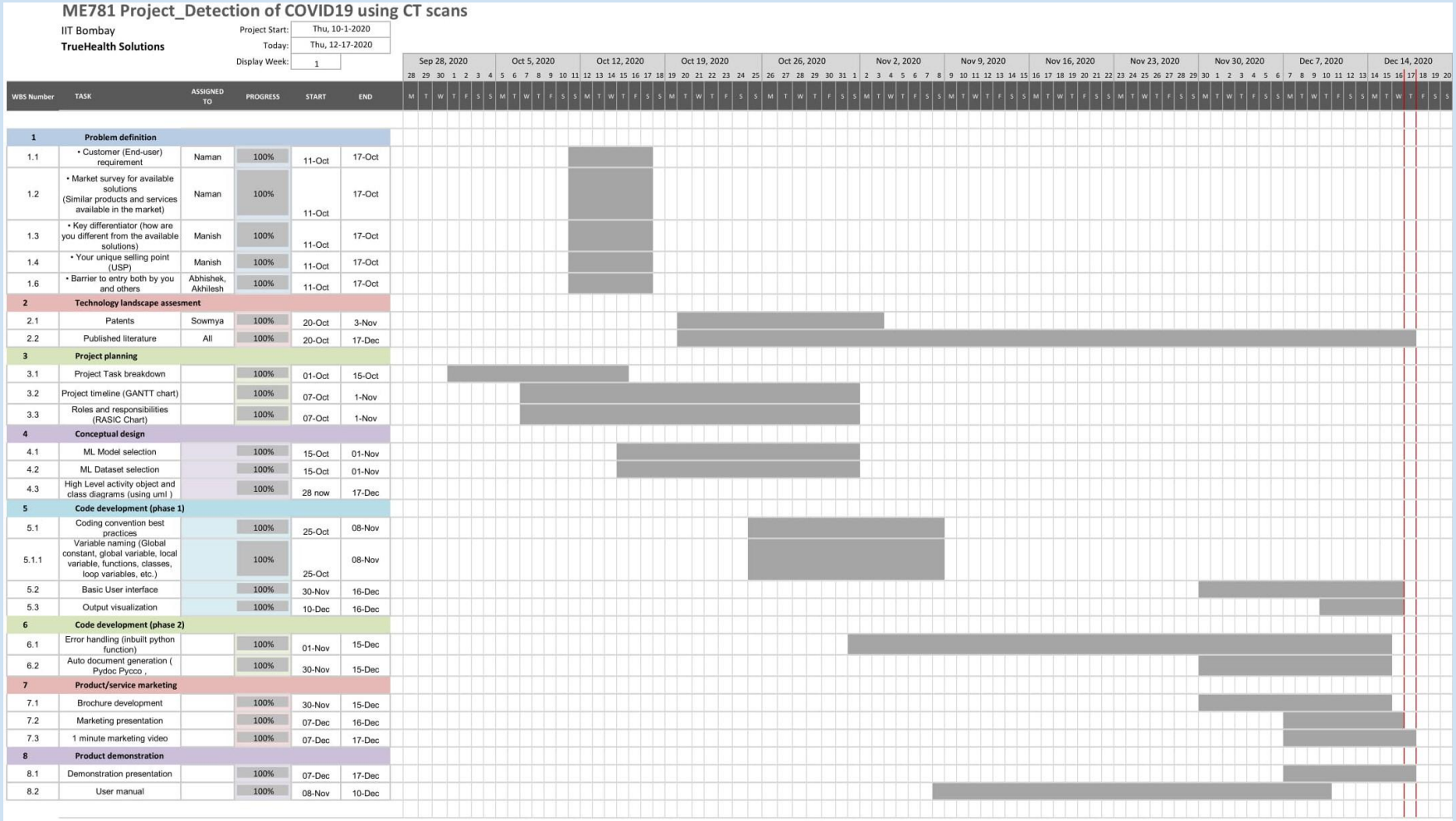
Problem Definition

Customer requirement	Fast, accurate, robust, and reliable diagnostic solutions to hospitals at an affordable price range. Doctors no longer have to wait long for medical image diagnosis results
Market Survey	**Qure.ai** (COVID-19 Response Solutions), **Niramai** (novel breast cancer screening solution), **Google's DeepMind** (retinal diagnosis), **Siemens Healthineers & Intel** (Cardiac MRI diagnosis)
USP	Easy-to-use **"plug and play" user-friendly interface** to assist doctors identify potential patients at early stages. **High accuracy (>90%), sensitivity** and **specificity** as compared to normal radiologists and other available solutions
Protection of USP	Code encryption and secure and reliable cloud and continuous updation of data for the model Patenting and copyrighting of our startup solutions, and codebase **OR** keep it as industry secret
Barrier to Entry	Healthcare data protection laws and extent of medical data Automated algorithm to discover a biomarker and build a test Patents and copyright issues from other startups/companies working in the field
Product & Operational Structure	CT scans upload to our server - > Model extracts important lung features from the image -> Model predicts the probability that the patient is infected with Covid-19

Technology landscape assessment

Patents	https://patents.justia.com/assignee/qure-ai-technologies-private-limited
Published literature	https://pubs.rsna.org/doi/10.1148/radiol.2020200905 https://www.sciencedirect.com/science/article/pii/S0010482520301645?via%3Dihub https://www.sciencedirect.com/science/article/pii/S0169260720314140 https://doi.org/10.1038/s41467-020-17971-2 https://www.nature.com/articles/s41591-020-0931-3 https://www.medrxiv.org/content/10.1101/2020.03.24.20042317v1 https://pubs.rsna.org/doi/10.1148/radiol.2020201237
Open libraries	Numpy, opencv, matplotlib, keras, Pandas, tensorflow, glob, nibabel
Proprietary libraries	https://github.com/wangshuocas/COVID-19 https://github.com/ChenWWWeixiang/diagnosis_covid19

Project timeline (GANTT chart)



ME781 Project_Detection of COVID19 using CT scans

IIT Bombay

TrueHealth Solutions

Project Start: Thu, 10-1-2020

Today: Thu, 12-17-2020

Display Week: 1

WBS Number	TASK	ASSIGNED TO	PROGRESS	START	END
1	**Problem definition**				
1.1	• Customer (End-user) requirement	Naman	100%	11-Oct	17-Oct
1.2	• Market survey for available solutions (Similar products and services available in the market)	Naman	100%	11-Oct	17-Oct
1.3	• Key differentiator (how are you different from the available solutions)	Manish	100%	11-Oct	17-Oct
1.4	• Your unique selling point (USP)	Manish	100%	11-Oct	17-Oct
1.6	• Barrier to entry both by you and others	Abhishek, Akhilesh	100%	11-Oct	17-Oct
2	**Technology landscape assesment**				
2.1	Patents	Sowmya	100%	20-Oct	3-Nov
2.2	Published literature	All	100%	20-Oct	17-Dec
3	**Project planning**				
3.1	Project Task breakdown		100%	01-Oct	15-Oct
3.2	Project timeline (GANTT chart)		100%	07-Oct	1-Nov
3.3	Roles and responsibilities (RASIC Chart)		100%	07-Oct	1-Nov
4	**Conceptual design**				
4.1	ML Model selection		100%	15-Oct	01-Nov
4.2	ML Dataset selection		100%	15-Oct	01-Nov
4.3	High Level activity object and class diagrams (using uml)		100%	28 now	17-Dec
5	**Code development (phase 1)**				
5.1	Coding convention best practices		100%	25-Oct	08-Nov
5.1.1	Variable naming (Global constant, global variable, local variable, functions, classes, loop variables, etc.)		100%	25-Oct	08-Nov
5.2	Basic User interface		100%	30-Nov	16-Dec
5.3	Output visualization		100%	10-Dec	16-Dec
6	**Code development (phase 2)**				
6.1	Error handling (inbuilt python function)		100%	01-Nov	15-Dec
6.2	Auto document generation (Pydoc Pycco ,		100%	30-Nov	15-Dec
7	**Product/service marketing**				
7.1	Brochure development		100%	30-Nov	15-Dec
7.2	Marketing presentation		100%	07-Dec	16-Dec
7.3	1 minute marketing video		100%	07-Dec	17-Dec
8	**Product demonstration**				
8.1	Demonstration presentation		100%	07-Dec	17-Dec
8.2	User manual		100%	08-Nov	10-Dec

Roles and responsibilities (RASIC Chart)

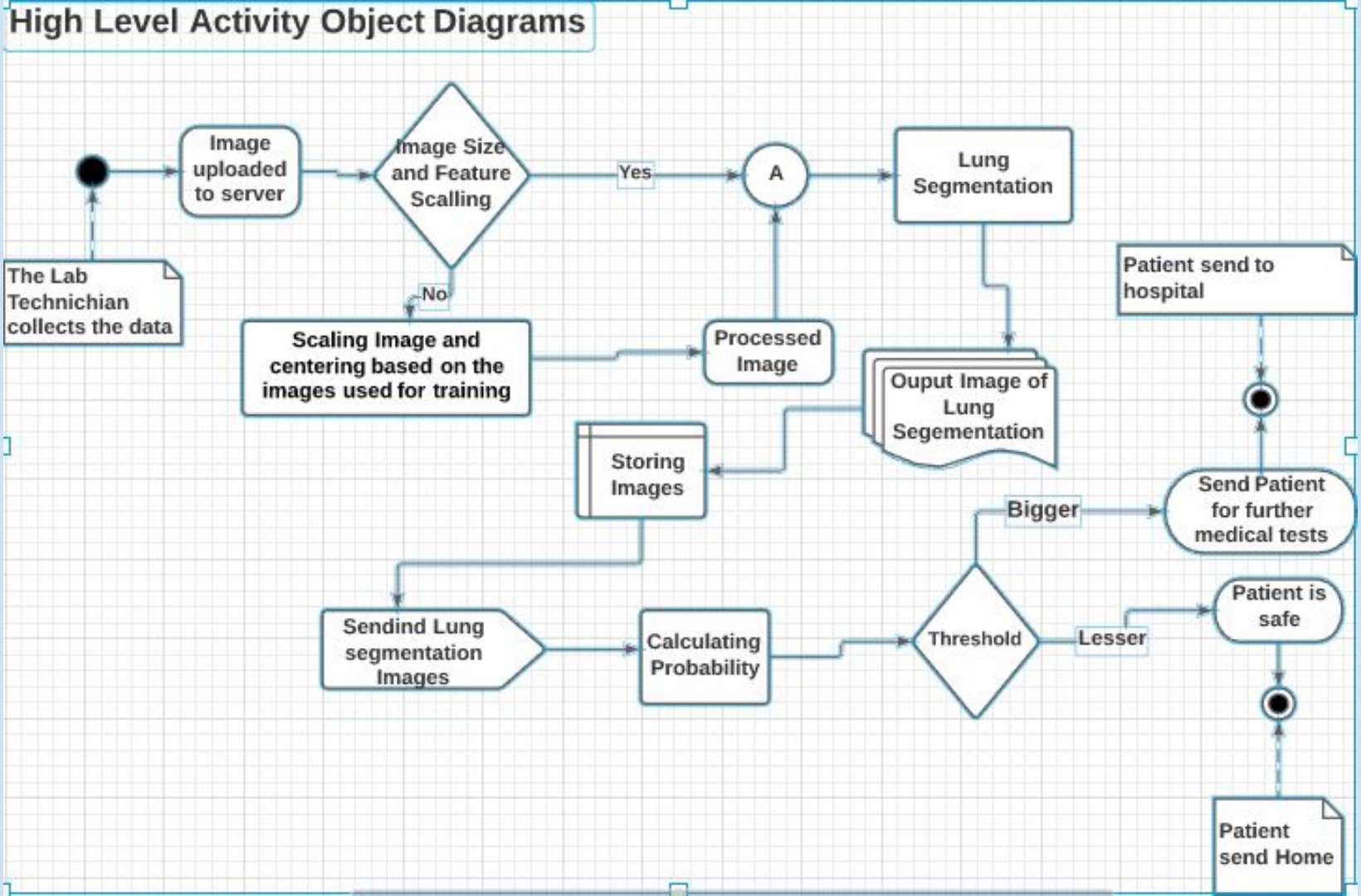
RASIC Chart

Tasks	WBS	Abhishek	Akhilesh	Manish	Naman	Sowmya
Customer (End-user) requirement	1.1	I	I	I	R	I
Market survey for available solutions (Similar products and services available in the market)	1.2	I	I	I	R	I
Key differentiator (how are you different from the available solutions)	1.3	C	I	R	A	I
Your unique selling point (USP)	1.4	I	I	R	A	I
How can you protect your USP	1.5	I	I	R	A	I
Barrier to entry both by you and others	1.6	R	S	I	I	I
Patents	2.1	I	I	I	I	R
Published literature	2.2	S	S	S	R	S
Project Task breakdown	3.1	I	I	I	R	I
Project timeline (GANTT chart)	3.2	I	R	I	C	I
Roles and responsibilities (RASIC Chart)	3.3	I	R	I	C	I
ML Model selection	4.1	I	I	I	R	I
ML Dataset selection	4.2	S	S	R	S	S
High Level activity object and class diagrams (using uml)	4.3	I	R	I	I	I
Choice of language, operating system, data set, algorithms and dependencies on external libraries	4.4	R	I	I	I	I
Coding convention best practices	5.1					
Variable naming (Global constant, global variable, local variable, functions, classes, loop variables, etc.)	5.1.1	I	I	I	I	R
Coding style	5.1.2	I	I	I	I	R
Indentation style	5.1.3	I	I	I	I	R
Error handling (inbuilt python function)	6.1	R	I	I	I	I
Auto document generation (Pydoc Pycco ,	6.2	I	I	R	I	I
Version control using GitHub	6.3	I	I	I	I	R
Brochure development	7.1	A	R	S	C	I
Marketing presentation	7.2	S	R	C	S	I
Demonstration presentation	8.1	S	S	R	I	I
User manual	8.2	I	R	I	A	I

R: Responsible: the person responsible for this task

A: Approve: the person giving the approval
S: Supporting: the people giving support for the completion of the task
I: Informed: the people to inform about the task
C: Consulted: people who can act as expert in regard to the task

Rule: Only one A and possible one R per Row


High Level Activity Object Diagrams
Image uploaded to server
Image Size and Feature Scalling
Yes
A
Lung Segmentation
The Lab Technichian collects the data
No
Scaling Image and centering based on the images used for training
Processed Image
Patient send to hospital
Ouput Image of Lung Segementation
Storing Images
Send Patient for further medical tests
Bigger
Sendind Lung segmentation Images
Calculating Probability
Threshold
Lesser
Patient is safe
Patient send Home

USER INTERFACE

We're committed to a clean and safe facility. Submit your health declaration

English

IIT Bombay, Powai, Mumbai 400076, Maharashtra, India

Tel: 000-000-0000

TrueHealth solutions

Image Diagnosis made simple!

HOME ABOUT US BLOG RESULTS COLLABORATE FAQs HELP

N

SUBSCRIBE

SEARCH

PATIENT INFO. | SYMPTOMS | STATUS | IMAGING | RESULTS | TREATMENT

TrueHealth solutions

Your one-stop solution!

Enter patient details below

Patient's Name: <First> <Last>

Age: _____

Gender: _____

Weight: _____

Registration no.: _____

Upload medical history: Upload

PREVIOUS

NEXT

Book Now

Registration no.: _____

Upload medical history: Upload

PREVIOUS NEXT

TrueHealth Solutions: Product demonstration

Reach out to customercare@ths.com if you are facing any issue in software. You can also refer to the product demonstration video and FAQ section for more details.

Reports and Diagrams

Class Diagrams: We have imported functions and classes and other subroutines from different well known python libraries (like keras, scikit-learn, etc.) to make our code time efficient and optimized. Those Subroutines are explicitly described in the documentation of those well known python libraries. We also didn't require to build some class structure, hence, class diagrams are not required.

Unit Test report: Our model expects 2-D CT Scan of Lung Section as input as also mentioned in the User Manual details. After uploading image to the server:

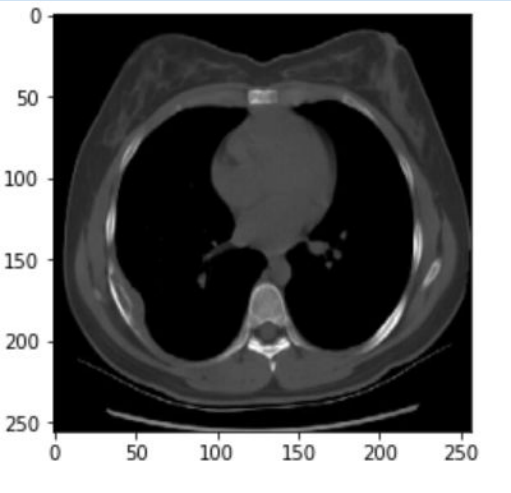
1. Image is pre-processed and centralized as per our model
2. Image is Feature Normalized as per our model parameters requirements

After performing the above two steps, image is sent to further processing for Lung Segmentation and Image Analysis. We perform data pre-processing with a good enough time on our end so that no discrepancies can come in the time of prediction which then successfully passes all testing parameters required in the code testing part. Hence, there isn't requirement for unit testing in our model as we ONLY input 2-D images in model which can be handled smoothly by the model.

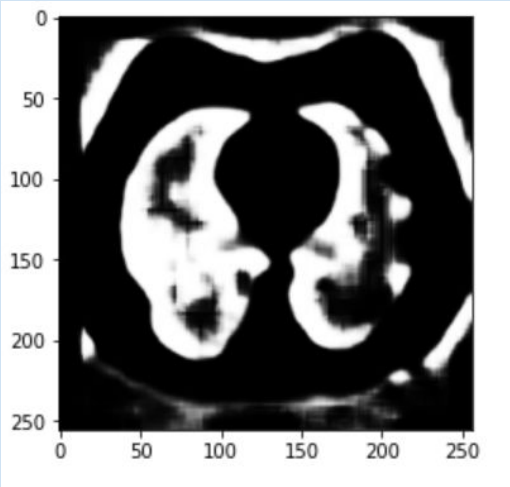
Model Output Visualization

Our model accepts 2-D image of Lung Section and outputs a probability . Below we show a output of our Lung Segmentation Model based on which predictions are done .

Original Lung CT Scan



Processed Lung Segmentation Image



Dataset details	Training data	Validation data
No. of covid images	337	12
No. of non-covid images	385	12

Chest CT scan dataset with 2 classes

Model details:

DenseNet121 model pretrained on ImageNet dataset was used. Fully connected layers were modified to classify CT images as 'COVID'/ 'NON-COVID'. SGD is used as optimizer and learning rate is reduced over time using ReduceLROnPlateau keras callback method.

Results	Cross-entropy Loss	Accuracy	MSE loss
Training data	0.0094	0.9958	0.0020
Validation data	0.1921	0.9167	0.0644

Results obtained after training the model for 100 epochs. A training log file is provided separately.

Thank you!