Artificial Intelligence-Based Clinical Assistance Tool for COVID-19 Cases in Ethiopia

By

Mastewal Mathiwos Abebe

In partial fulfillment of the requirements for the Degree of Master of Science in Biomedical Engineering

(Bioinstrumentation and Imaging)

Center of Biomedical Engineering

Addis Ababa Institute of Technology

Advisor: Dawit Assefa Haile (PhD)

Co-advisor: Robel Kebede Gebre (PhD)

Addis Ababa, Ethiopia

December 2022

Declaration

By submitting this MSc thesis, I declare that the entirety of the work presented in this document is my original work. Where information has been derived from other sources, I confirm that this has been indicated in the thesis and that I have not previously in its entirety or part submitted it for obtaining any qualification.

Name: Mastewal Mathiwos Abebe

Signature: _____

Date: _____

This MSc thesis has been submitted for examination with my approval as an advisor.

_____

Certificate of Examination

Addis Ababa University

School of Graduate Studies

This is to certify that the thesis prepared by Mastewal Mathiwos Abebe entitled "*Artificial intelligence-based clinical assistance tool for COVID-19 cases in Ethiopia"* submitted in partial fulfillment of the requirements for the Degree of Master of Science in Biomedical Engineering (Bioinstrumentation and Imaging) complies with the regulations of the University and meets the accepted standards with respect to originality and quality.

Signed by the examining committee

Examiner___Dr. Melkamu Hunegnaw_____Signature _____ Date _____

Examiner ___Dr. Rahel Argaw_____Signature _____ Date _____

Advisor ____Dr. Dawit Assefa_____Signature _____ Date _____

Co-Advisor _____Signature _____ Date _____

_____

Chief of Department or Graduate program coordinator

Abstract

The COVID-19 crisis has had a devastating impact in terms of loss of human life and economic disruption. According to a world health organization report on March 2022, more than 6.1 million people have died worldwide as a result of this pandemic. Early and precise recognition of COVID-19 is key for treating patients and slowing the spread of the pandemic. Several artificial intelligence (AI) based solutions have been developed to facilitate the application of chest X-ray (CXR) imaging and anthropomorphic data for use as a COVID-19 screening tool in resource-limited settings. The current study aims to develop a COVID-19 diagnosis scheme based on a deep learning (DL) approach applied on X-ray image samples collected locally in Ethiopia as well as a severity prediction tool based on a machine learning (ML) approach applied on anthropomorphic data collected from the same patients. The study data for the DL approach consisted of 746 CXR labeled images collected from St. Peter's Specialized Hospital (SPSH) and Millennium COVID-19 care center (MCCC) in Ethiopia. The samples were composed of images acquired from patients treated for COVID-19 and normal controls. The DL model involved data preprocessing steps applied on the raw CXR images and was designed based on transfer learning approaches. The diagnostic prediction ability of the approach was measured in terms of probability score value and occlusion sensitivity map. The study data for the ML model development consisted of 308 anthropomorphic data (including demographic, comorbidity, and COVID-19 symptoms) collected from the MCCC to check the severity level of the COVID-19 cases and classify them into one of the three classes: moderate, sever or critical. A 5-fold cross-validation approach was used to train two popular ML approaches namely Support Vector Machine (SVM) and K-nearest Neighborhood (KNN) models. The DL method using ResNet-50 architecture achieved best classification performance with a validation accuracy of 94.20 % in accurately classifying COVID-19 and normal cases. The SVM model achieved a better prediction ability than K-NN with an overall accuracy of 89.9 % in predicting the severity status of the COVID-19 cases.

Keywords: COVID-19, Chest X-ray, Anthropomorphic Data, Deep Learning, Machine Learning

Acknowledgments

First of all, I would like to thank the Almighty God for his unconditional favor towards me and for giving me the strength, ability, and opportunity to undertake this thesis.

It is a great pleasure to give my sincere gratitude to my supervisor, Dr. Dawit Assefa Haile, and co-supervisor, Dr. Robel Kebede, for their continuous guidance, mentorship, and encouragement, starting from the initial phase of the thesis proposal and throughout the process of researching and writing this thesis.

I would like to thank Addis Ababa University for providing a scholarship opportunity for my study. I would also like to express my appreciation to my hosting company, Memagi Medical Imports PLC, and all the staff members for their unlimited support while pursuing my studies. I would also like to thank the radiology department at St. Peters Specialized Hospital and the Millennium COVID-19 Care Center for their help during data collection.

I would like to give a special thanks to my parents for the countless hours of support, prayer, and care they have given me. This accomplishment would not have been possible without you. And also many thanks to all my immediate family who have supported, prayed, and given me encouraging words throughout my journey.

I would also want to thank my loved ones, who have supported me throughout the entire process, both by keeping me harmonious and by helping me put pieces together. I will be grateful forever for your love. Finally, I want to thank my friends and classmates for all the care and support they have given me.

Table of Contents

List of Figures

List of Tables

Acronyms

2D	Two Dimensional
AI	Artificial Intelligence
AUC	Area Under the Curve
CDC	Centers for Disease Control and Prevention
COVID-19	Corona Virus Disease 2019
CNN	Convolution Neural network
CT	Computed Tomography
CXR	Chest X-Ray
DNN	Deep Neural Network
DM	Diabetes Mellitus
DT	Decision Tree
ET	Ethiopia
FC	Fully Connected
FNR	False Negative Rate
JPEG	Joint Photographic Experts Group
ICTV	International Committee on Taxonomy of Viruses
ICU	Intensive Care Unit
IRB	Institutional Review Board
KNN	K-Nearest Neighbor
LR	Logistic Regression
MATLAB	Matrix Laboratory
MCCC	Millennium COVID-19 Care Center
ML	Machine Learning
MLP	Multi-Layer Perception
MRN	Medical Record Number
OS	Open Source
ROC	Receiver Operating Characteristic
RT-PCR	Reverse Transcriptase Polymerase Chain Reaction
SARS	Severe Acute Respiratory Syndrome
SPHMMC	Saint Paul Hospital Millennium Medical College
SVM	Support Vector Machine
TPR	True Positive Rate
VGG	Visual Geometry Group
WHO	World Health Organization

Chapter 1 Introduction

1.1 Background

COVID-19 is an infectious disease caused by the SARS-CoV-2 virus (coronavirus) [1]. The coronavirus emerged in Wuhan, China at the end of 2019 and immensely spread around the world [2][3]. According to a Word Health Organization (WHO) report from March 26th, 2022, COVID-19 infected over 483,556,595 individuals with 6,132,461 deaths worldwide [4]. Ethiopia represents 469,679 cases and 7,492 deaths by the pandemic [4]. The virus that causes COVID-19 changes over time [5]. The SARS-CoV-2 variants include alpha, beta, gamma, delta, and omicron [5]. The changes affect the virus's properties, including how easily it spreads, the associated disease severity, and the performance of vaccines [5]. Figure 1.1 below shows several countries, territories, and areas reporting the COVID-19 variants. Person-to-person spread became the main mode of COVID-19 transmission [6]. It is mainly through close-range contact (i.e., two meters) via respiratory particles. Prevention of COVID-19 includes taking personal preventive measures and prevention by vaccination [7].

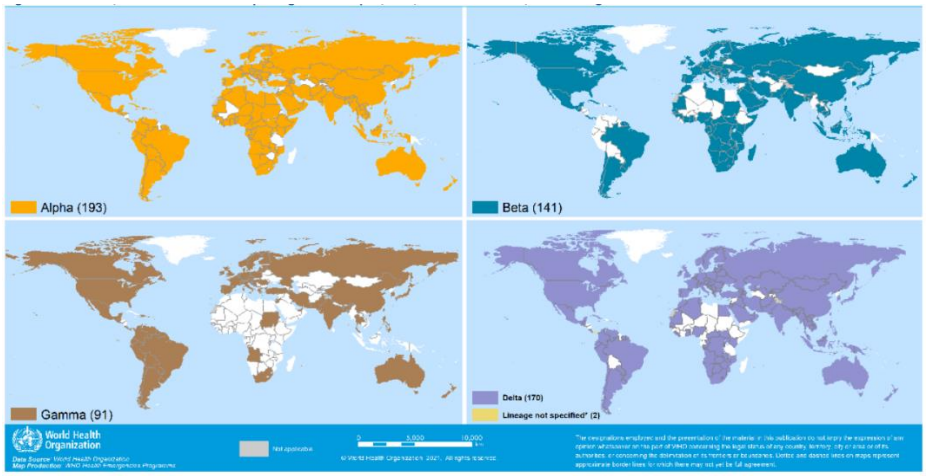


Figure 1.1 Countries, territories, and areas reporting SARS-CoV-2 variants alpha, beta, gamma, and delta, as of 31 August 2021 [8].

People with COVID-19 have had a wide range of symptoms. The spectrum of illness varies from mild to critical. Studies show that severe COVID-19 cases can be determined by socio-demographic characteristics and having a history of pre-existing comorbid illness [9][10]. The COVID-19 patterns in the infection presentation, progression, and outcome were different in Africa in comparison with the rest of the world [9]. Considering Ethiopia's population size (the second most populous country

in Africa), the ratio of COVID-19 cases per population is very low compared to most other African countries [11]. The study conducted in [9] reported the risk factors of disease severity among the Ethiopian population.

Early recognition of COVID-19 is key for treating infected patients and slowing the spread of the pandemic. COVID-19 is commonly diagnosed by a reverse transcriptase-polymerase chain reaction (RT-PCR), which can detect SARS-CoV-2 RNA from respiratory specimens (collected through a variety of means, such as nasopharyngeal or oropharyngeal swabs) [1][3] (see also Figure 1.2). However, RT-PCR is a time-consuming, laborious, and complicated manual process that is in short supply [12][13]. In a study done by [14], the sensitivity of a single RT-PCR test of upper respiratory tract samples was reported to be 82.2% and the sensitivity increases to 90.6% when patients are tested twice.

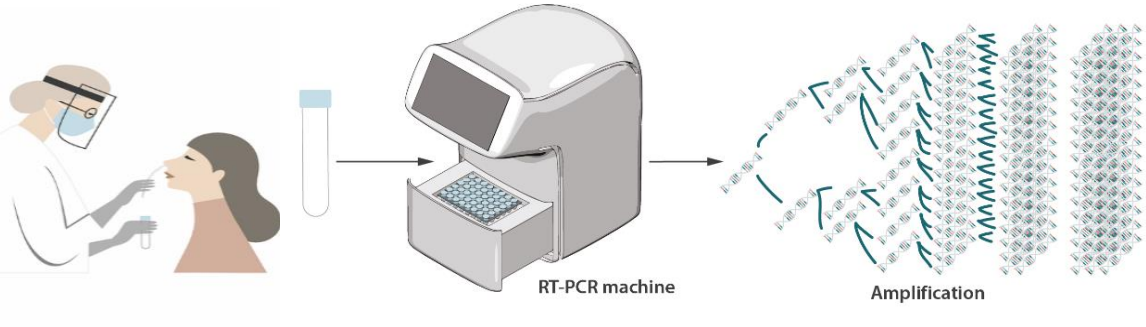


Figure 1.2 Illustration of RT-PCR-based COVID-19 diagnosis [15]

The other COVID-19 diagnostic method is lateral flow immunochromatography (Point-of-care-testing). The SARS-CoV-2 Immunoglobulin G (IgG)/ Immunoglobulin M (IgM) rapid test is a lateral flow chromatographic immunoassay capable of detecting antibodies against the SARS-CoV-2 virus [16] [17]. On average, IgM and IgG antibodies may take one to three weeks to develop after infection [18][19]. Some persons may not even develop detectable antibodies after COVID-19 infection, or antibody levels may decline over time to undetectable levels [19] [20]. The antibody tests are much affordable and faster than RT-PCT to use in resource-constrained areas, however, they are not as sensitive as the tests done using RT-PCR [21].

A radiography examination, chest X-ray (CXR), and computed tomography (CT) have been alternative screening methods utilized for COVID-19 screening [22]. CXR images are favored over CT images due to their lower radiation dose, widespread availability, and lower cost [21][23]. Other advantages of CXR include that it can be performed with portable equipment at the point of care,

which can minimize the risk of cross-infection related to patient transport [22]. However, the visual indicators of COVID-19 on CXR require experienced and expert radiologists to interpret the radiographic images. Figure 1.3 shows axial CT and CXR mages of patients with mild and severe COVID-19 cases.

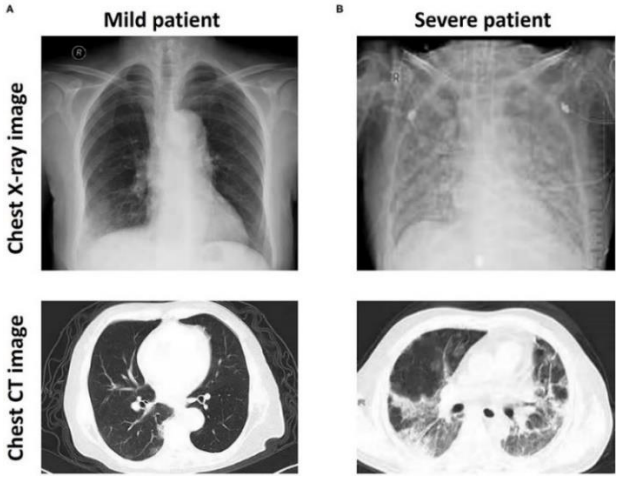


Figure 1.3 CXR and CT images with mild (A) and severe (B) COVID-19 [24].

Artificial intelligence (AI) approaches have made significant contributions to CXR image analysis and the results reported showed their high classification performance and less time-consumption [21] [25]. Deep learning (DL) is a subset of AI that uses deep neural networks (DNN) and that has been used previously to develop a COVID-19 diagnostic tool trained on CXR images [26]. Machine learning (ML) is the other AI technique used to predict the severity of COVID-19 cases where the training is made using patients' demographic data, previous medical history, and COVID-19 associated symptoms [27][28]. The field of AI can be represented by a Venn diagram that highlights the conceptual relationships between the different techniques (Figure 1.4).

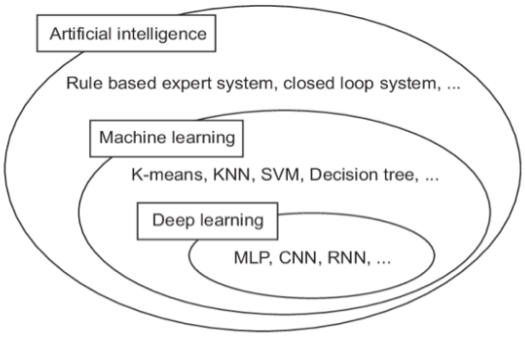


Figure 1.4 A Venn diagram of artificial intelligence [29].

1.2 Problem statement

The COVID-19 crisis has had a devastating impact in terms of loss of human life and economic disruption. Recent research applied AI-based techniques to the fight against COVID-19. However, the datasets used in developing the AI models were derived from public repositories and do not represent the Ethiopian population dynamics. The datasets were primarily aggregated from various Western sources which miss out the Ethiopian demographic characteristics and other associated risk factors. Ethiopia, having its unique geographical and epidemiological profile in the world, justifies the need for its own dataset for use in automatic COVID-19 prediction. The current thesis work assimilated datasets collected from St. Peter's Specialized (SPSH) Hospital and Millennium COVID-19 care center (MCCC) in Ethiopia for developing an AI-based clinical assistance tool for COVID-19 diagnosis, leveraging the DL approach, and developing a severity prediction tool for the confirmed COVID-19 cases using ML based algorithm trained on local anthropomorphic data. The developed diagnostic and severity prediction models can be powerful tools for not only effectively identifying incoming patient cases but also assisting in identifying confirmed COVID-19 severity levels.

1.3 Research questions

- ✓ Can the DL model offer the best performance in the diagnostic prediction of COVID-19 using the Ethiopian population CXR image dataset?
- ✓ Can we distinguish patient cases with a viable option of COVID-19 diagnosis?
- ✓ Can the ML model offer the best performance in severity prediction of COVID-19 cases on data collected exclusively from the Ethiopian population?

1.4 Objectives of the study

1.4.1 General objective

The main objective of this thesis is to develop AI models for the diagnostic prediction of COVID-19 using radiographic images and severity prediction of COVID-19 cases using anthropomorphic data collected in Ethiopia.

1.4.2 Specific objectives

- ✓ To develop a DL-based COVID-19 diagnostic prediction model and evaluate its performance using CXR images collected from SPSH and MCCC in Ethiopia. Performance evaluation of the DL model will also be checked using open-source image data sets.
- ✓ To use probability scoring for diagnostically predicted classes and to use sensitivity map-based visualization of decision-making areas by the DL model.
- ✓ To develop ML-based models that allow best severity prediction of confirmed COVID-19 cases using anthropomorphic data collected in Ethiopia.

1.5 Organization of the thesis

The rest of the thesis has been organized into five chapters. Chapter 2 includes a literature review of previous works done on diagnostic prediction and severity prediction of COVID-19 using DL and ML methods. Chapter 3 presents the materials used and methods followed in the proposed COVID-19 diagnostic and severity prediction scheme while Chapter 4 presents the major findings. Discussions and limitations of the study are presented in Chapter 5. Lastly, Chapter 6 presents concluding remarks and future directions of the study.

Chapter 2 Literature review

2.1 Applications of DL in the diagnostic prediction of COVID-19 using radiographic images

A previous study reported in [30] used a method based on DL networks to classify COVID-19 based on X-ray images making use of the ResNet-50 pre-trained model. The authors used the Kaggle dataset of COVID-19, which contained 50 CXR images of the lungs, where 25 CXR images were for patients treated for COVID-19 and the rest 25 CXR images were acquired from healthy subjects. For their experiment, a 5-fold and 10-fold cross-validation were used to split the dataset. Their experimental results show that 5 folds offer more effective results than 10 folds with an accuracy of 97.28%. Another study reported in [31] developed a COVID-19 detection framework that used the notion of a pre-trained model to automatically classify CXR images into three COVID-19 severity classes: normal, moderate, and severe. The suggested approach combined transfer learning with three prominent pre-trained CNN models: Resnet-50, GoogleNet, and AlexNet. Their system considered 1491 CXR, including 1335 normal cases, 106 moderate cases, and 50 severe cases. The dataset was divided into three parts: 70% for training, 15% for validation, and 15% for testing. ResNet-50 outperformed the other architectures with an overall accuracy of 87.8%. In another study [32], an ensemble deep convolution neural network model ‘‘CoVNet-19’’ was proposed to diagnose COVID-19-infected patients using CXR images. They used 798 images taken from patients who tested positive for COVID-19, 2345 pneumonia infected patients, and 2341 CXR images of healthy cases. They used gradient-weighted class activation mapping for easy visualization of the results by radiologists and other medical experts (see also Figure 2.1)Their experimental results show that the overall classification accuracy obtained with their proposed approach for three-class classification among COVID-19, pneumonia, and normal was 98.28%, along with an average precision and Recall of 98.33% and 98.33%, respectively. The binary classification between non-COVID-19 and COVID-19 CXR images was reported with an overall accuracy of 99.71%.

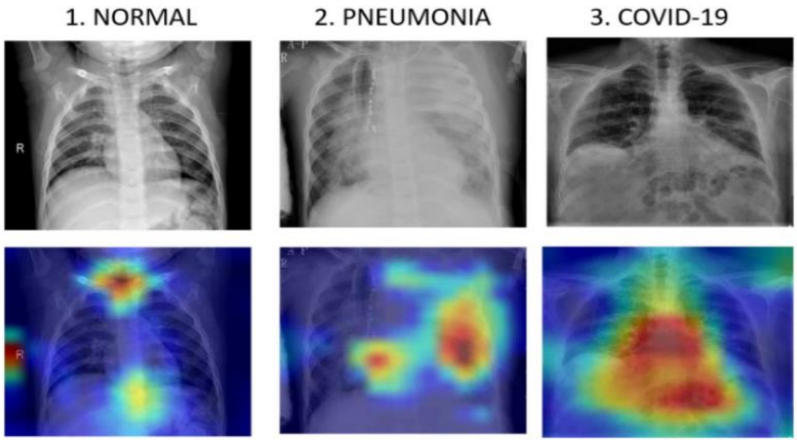


Figure 2.1 Gradient weighted class activation visualization [32].

In other research [33], DenseNet-121 was used to construct a DL-based strategy for detecting COVID-19 patients. The suggested system was trained and tested using the COVIDx dataset, which included 13,800 chest radiography pictures from 13,725 patients. Their dataset includes only 238 COVID-19-confirmed chest radiography images. The model was put to the test for two-class classification (COVID-19 and non-COVID-19) and three-class classification (COVID-19, pneumonia, and normal). Their network then achieved an accuracy of 96.49% for the two-class classification and 93.71% for the three-class classification.

In a study reported in [34], a method for detecting coronavirus (COVID-19 causing and other coronaviruses) based on deep transfer learning and several pre-trained models was proposed. VGG16, VGG19, DenseNet-201, Inception-ResNet-V2, Inception V3, Resnet-50, and MobileNet V2 are the seven pre-trained models the authors used for comparative analysis. The dataset for the experiments comprised 6087 CXR images from normal, pneumonia, and coronavirus cases, Out of the 6087 cases, 231 images were collected from confirmed COVID-19 cases. In their system, the training and validation data partitions were kept at 80:20 ratio. Based on their results, Inception-Resnet-V2 and the DensNet-201 models provided better accuracy than the other models; 92.18% accuracy for Inception-ResNetV2 and 88.09% accuracy for Densnet201 (see Figure 2.2 for the model the authors used).

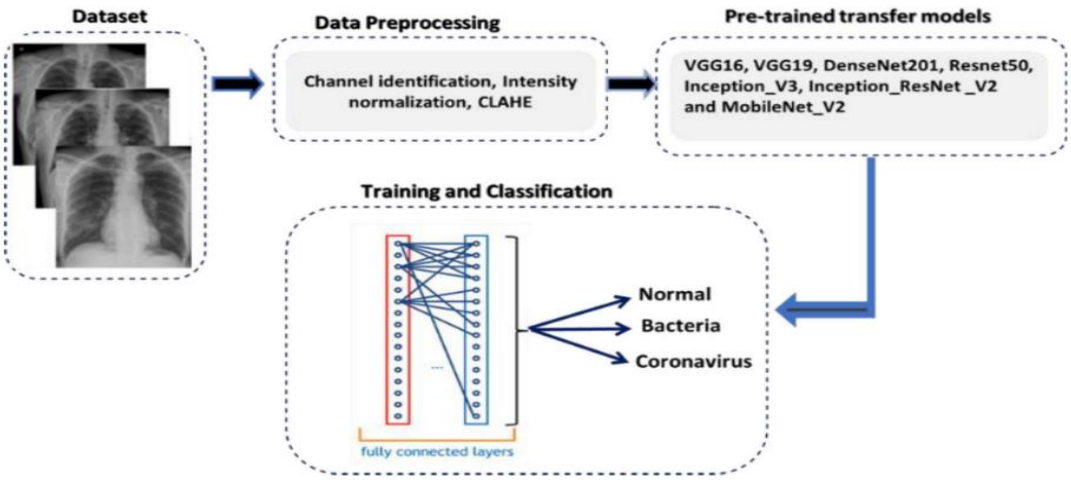


Figure 2.2 Block diagram of the methodology proposed in [34].

The study conducted in [35] proposed COVID-Net to detect distinctive abnormalities in CXR images among samples of 13,870 taken from COVID-19 and pneumonic cases as well as normal controls available at the COVIDx dataset. The dataset contained only 266 COVID-19 cases. The authors compared the performance of the COVID-Net model with that of ResNet-50 and VGG-19 models by using the same training dataset. The sensitivity of the models in identifying COVID-19 cases were 91%, 83.0%, and 58.7% using COVID-Net, ResNet-50, and VGG-19, respectively.

In another study conducted in [36], five pre-trained CNN-based models (ResNet50, ResNet101, ResNet152, InceptionV3, and Inception-ResNetV2) were proposed for the detection of COVID-19 cases based on analysis of CXR images. The authors implemented three different binary-class classifications with four classes comparing COVID-19 (n = 341) against normal/healthy samples (n = 2800), viral pneumonia (n = 1493), and bacterial pneumonia (n = 2772). Considering the results obtained, they have reported that the pre-trained ResNet50 model provided the highest classification performance (96.1% accuracy for Dataset-1, 99.5% accuracy for Dataset- 2 and 99.7% accuracy for Dataset-3) among the other four used models. According to studies performed in [33] [49] [52] and [53], the transfer learning approach was used for the diagnostic prediction of COVID-19 cases with small data sets, and they reported high performance accuracy. Some other DL-based studies proposed previously for use in the diagnosis of COVID-19 based on CXR images are shown in Table 2.1.

Table 2.1 Other studies proposed for DL-based COVID-19 diagnosis using CXR images.

Title	Architecture	Images source	Image distribution	Class	Performance
Recognition of coronavirus disease using a deep learning network, 2021 [30]	ResNet-50	CXR	25 COVID-19 and 25 normal	2 (COVID-19 and normal)	Accuracy 97.28 %
A deep learning model for the detection and analysis of COVID-19 patients, 2021 [32]	CoVNet-19	CXR	798 COVID-19, 2341 normal, 2345 pneumonia	3 (COVID-19, pneumonia, and normal) 2 (non-COVID-19 and COVID-19)	Accuracy 98.28 %, Accuracy 99.71 %
Explainable COVID-19 diagnosis based on chest X-ray images, 2020 [37]	Deep COVID Explainer: (VGG-19 + DenseNet-161)	CXR	1213 COVID-19, 8636 normal and 6038 pneumonia	3 (normal, pneumonia, and COVID-19)	PPV 96.12% and Recall 94.3%
A depth-wise separable deep neural network with white balance and CLAHE for detection of COVID-19, 2020 [38]	CovidLite	CXR	536 COVID-19, 619 viral pneumonia, 668 normal	2 (COVID-19, and normal), 3 (normal, COVID-19, and viral pneumonia)	Accuracy 99.58%, Accuracy 96.43%
Artificial intelligence applied to chest X-rays can aid in the diagnosis of COVID-19 infection, 2020 [39]	ResNet-50	CXR	324 COVID-19, 276 non-COVID-19	2 (COVID-19 and non-COVID-19)	Sensitivity 80% and Specificity 81%
COVID-19: automatic detection from X-ray images utilizing transfer learning with convolutional neural networks, 2020 [40]	VGG19	CXR	224 COVID-19 700 pneumonia, 504 healthy	3 (COVID-19, normal, and pneumonia)	Accuracy 96.78%, Sensitivity 98.66% and Specificity 96.46%
Automated detection of COVID-19 cases using deep neural networks with X-ray images, 2020 [41]	Dark COVID Net (Darknet-19, YOYO)	CXR	125 COVID-19, 500 pneumonia 500 no-findings	2 (COVID-19, and normal), 3 (COVID-19, normal, and pneumonia)	Accuracy 98.08%, Accuracy 87.02%
Classification of COVID-19 in chest Xray images using	DeTrace ResNet18	CXR	105 COVID-19, 11 SARS, 80 normal	3 (COVID-19, SARS, and normal)	Accuracy 95.12%, Sensitivity

DeTraC deep convolutional neural network, 2020 [42]	97.91%, and Specificity 91.87%

2.2 Application of ML in severity prediction of COVID-19 using anthropomorphic data

Research conducted in [28] aimed to develop a model that can predict whether a person is affected by COVID-19 or not by using the SVM model (Figure 2.3). Their study used 200 records performed on COVID-19 cases using 8 attributes (including COVID-19-related symptoms and medical conditions), and the infection status was predicted as severely infected, moderately infected, and not infected. The proposed model offered an overall accuracy of 87 %. Underlying medical conditions associated with a higher risk of severe COVID-19 were described in [60]. Cardiovascular diseases appear to be the most common, along with hypertension, diabetes, chronic respiratory diseases, and cancer. Certain demographic features have also been associated with more severe illnesses. Males have comprised a disproportionately higher number of critical cases and deaths in multiple cohorts worldwide [10] [61]. According to the Centers for Disease Control and Prevention (CDC) review [62], some baseline anthropomorphic characteristics are described as risk factors for COVID-19 spread and the risk of becoming ill from the pandemic. Table 2.2 shows increased risk factors for COVID-19 severity [63].

Table 2.2 Anthropometric risk for COVID-19 infection [63].

Age	Cardiovascular disease	Heart disease and failure
Asthma	Cerebrovascular diseases	Hypertension
Autoimmune disorders	Chronic digestive, kidney, liver, and respiratory disorders	Immune system disorders
COPD	Diabetes	Male gender
Cancer	Drinking alcohol	Overweight
Race	Respiratory system diseases	Smoking status

Another study proposed in [43] developed a COVID-19 severity prediction model based on SVM. The authors trained the model by using clinical and laboratory features that are associated with COVID-19 cases. In their study, they used 336 cases of COVID-19 patients, and they found that their model was able to predict patients' severity conditions with up to 77.5 % accuracy.

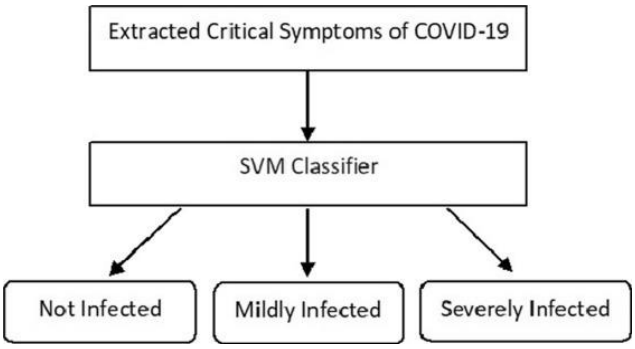


Figure 2.3 Block diagram of the suggested COVID-19 detection methodology in [28].

KNN algorithm was also used in another study reported in [44] to predict COVID-19 survivals (recovery and death) by using two parameters (age and gender) of patients containing 730 records. To evaluate the performance of the ML model, 10-fold cross-validation was used as a resampling procedure to validate the model. KNN algorithm-based outcome prediction results were compared to other algorithms including logistic regression, decision tree, SVM, and multi-layer perception. Whereas the KNN predicts the outcome of the COVID-19 cases (based on age and gender) with up to 3.3% improved accuracy than other algorithms.

Another study conducted in [45] demonstrated the capability of ML algorithms in determining whether a given patient (COVID-19 infected or suspected) is more likely to survive than die or vice versa. Their algorithm was trained on a database extracted from confirmed and suspected COVID-19 infections, which was compiled and made publicly available by the Mexican federal government. The database includes 28 features for each patient with a medical history, demographic data, and information related to the COVID-19 episode. Their proposed method aimed to detect high-risk patients with high accuracy and to improve hospital capacity planning and time management. Their proposed ML algorithm was based on a multilayer feed-forward network with two sigmoid neurons in a single hidden layer and two softmax neurons in the output layer. Their proposed algorithm offered an accuracy, specificity, and sensitivity value of 93.5%, 90.9%, and 96.1%, respectively.

In another study, the classification accuracy of a dataset with some ML models was compared [46]. In their study, naive Bayes, KNN, SVM, and decision tree algorithms were trained on a COVID-19 database collected by the Mexican federal government with a sample record of 96,839. The best result was obtained from the SVM algorithm in classifying COVID-19 positive and negative cases with an accuracy of 100 %.

Prediction of COVID-19 patients using a supervised ML algorithm was conducted in [47]. In their research, in order to evaluate the collected datasets of COVID-19 patients, five ML based classifiers

were used: Nave Bayes, SVM, logistic regression, KNN, and decision tree model, as shown in Figure 2.4. COVID-19 datasets from the repository were combined and re-examined to remove incomplete entries, and a total of 2500 cases were finally utilized. After performing classification, it was uncovered that the decision tree classifier gives better outcomes. Increasing the data size in a model is suggested as it will increase its performance. They have reported that the accuracy, specificity , and sensitivity reach values up to 93.5%, 90.9%, and 96.1%, respectively.

In another study, ML based model to predict the disease severity and outcome in COVID-19 patients was presented [48]. The study was performed using 287 COVID-19 patient samples acquired from the King Fahad University Hospital, Saudi Arabia. The data were analyzed using three classification algorithms, namely logistic regression, random forest (RF), and extreme gradient boosting. A 10-K cross-validation was applied for data partitioning. The authors reported that RF performed best with an overall classification accuracy of 95%.

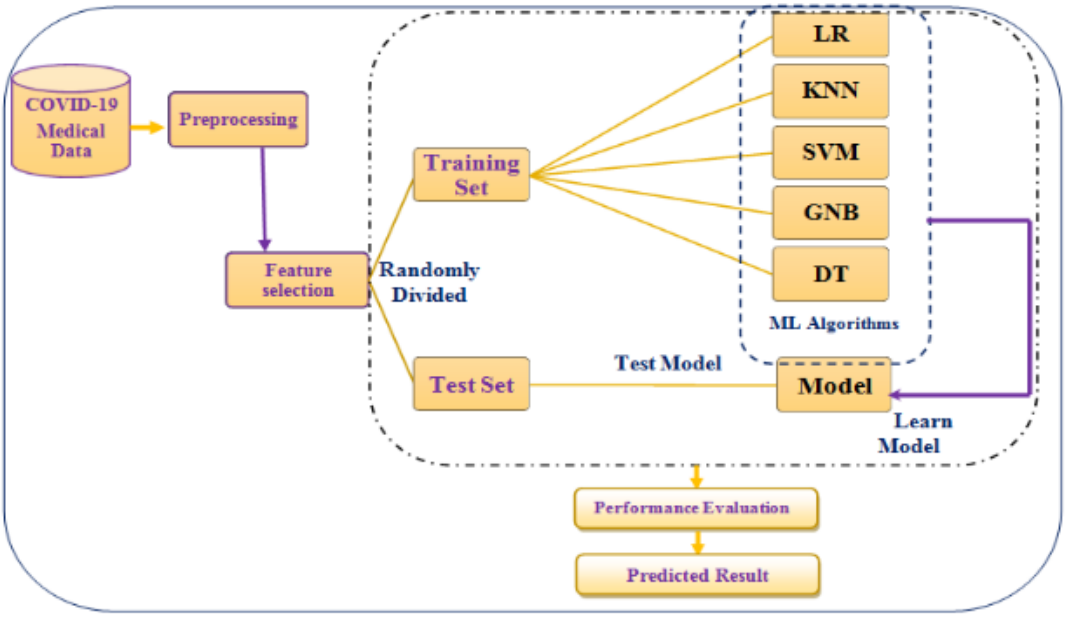


Figure 2.4 Structure of six supervised ML models for predicting COVID-19 disease [47].

A summary of the studies based on ML method is presented in Table 2.3 including the model/s used, the number of features, as well as the classification type, considered.

Table 2.3 A summary of methods used for COVID-19 severity prediction using anthropomorphic data.

Title	**Model**	**Number of data**	**Features**	**Classes**	**Performance**
A novel approach to predict COVID-19 using a support vector machine, 2021 [28]	SVM	200	8	3 (not-infected, infected, severely infected)	Accuracy 77.5%
Prediction of COVID-19 Possibilities using KNN classification algorithm, 2020 [44]	KNN	730	2	2 (recovered, deceased)	Accuracy 80.4%
Classification of COVID-19 dataset with some machine learning methods, 2020 [46]	SVM	96,839	19	2 (infected, not infected)	Accuracy 100%
Prediction of a COVID-19 patient using a supervised machine learning algorithm, 2021 [47]	Ensembled supervised	2500	12	2 (infected, not infected)	Accuracy 93.5%
A combination of four clinical indicators predicts the severe-critical symptom of patients infected with COVID-19, 2020 [43]	SVM	336	25	2 (severe, critical)	ROC-AUC 99.96%

Chapter 3 Materials and methods

This chapter discusses the proposed DL-based COVID-19 diagnostic prediction tool as well as the ML-based COVID-19 severity status prediction scheme. The general methodological pipeline for both schemes is shown in Figure 3.1.

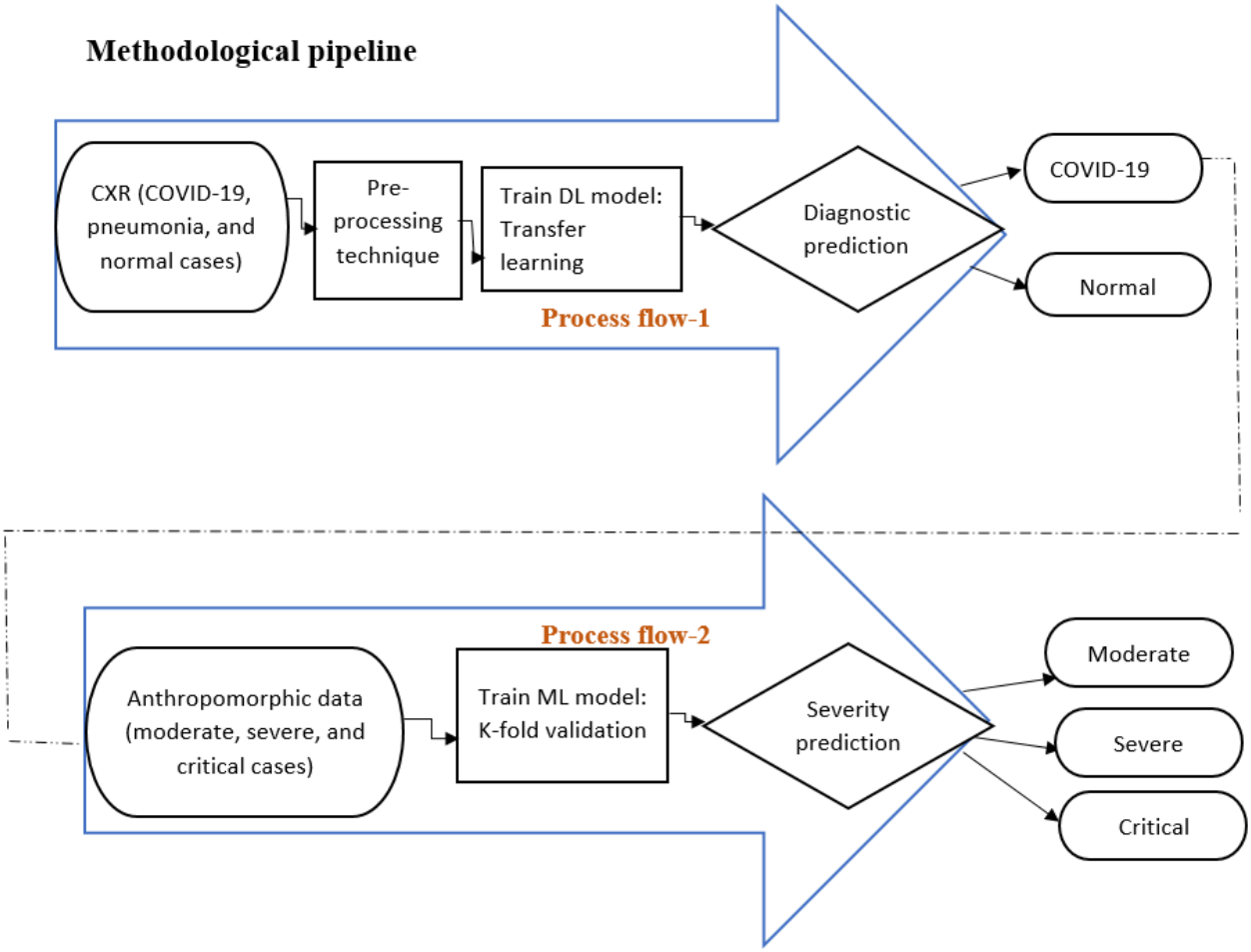


Figure 3.1 Process flow-1 (top) illustrates a methodological pipeline for the deep learning (DL) based diagnostic prediction model development using chest x-ray (CXR) images; process flow-2 (bottom) illustrates a methodological pipeline for the machine learning (ML) based COVID-19 severity prediction model development using anthropomorphic data.

3.1 Diagnostic prediction of COVID-19

The CXR images were used in the diagnostic prediction model development and preprocessing techniques were performed on the CXR images before further processing as illustrated in Figure 3.2.

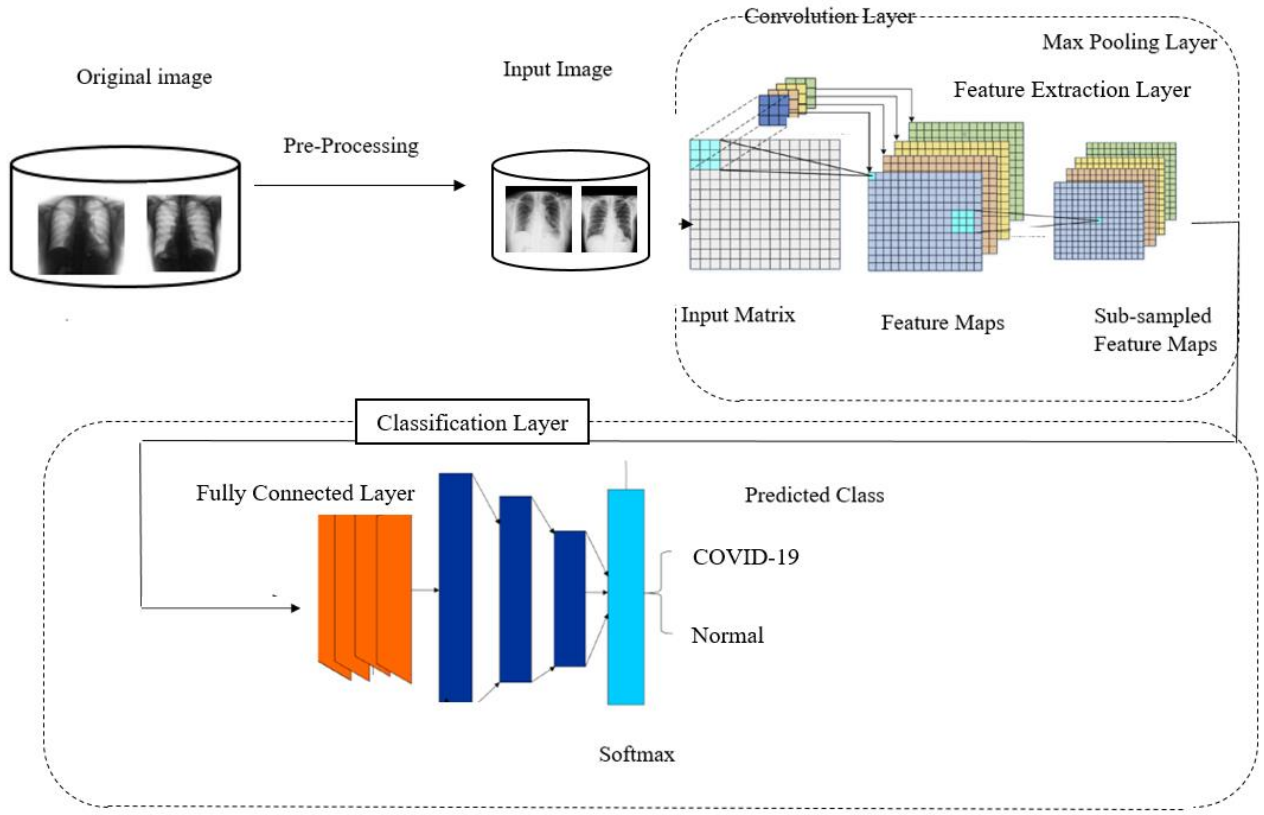


Figure 3.2 A schematic representation of the proposed COVID-19 diagnostic prediction model.

3.1.1 Study dataset

Data source

The CXR image dataset used to develop the proposed DL model was collected from St. Peter's Specialized Hospital (SPSH) and Millennium COVID-19 care center (MCCC) in Ethiopia. MCCC was at the time the largest COVID-19 center in the country, with a capacity of 1000 beds in the intensive care unit (ICU).

Study period and data sample

The CXR images from SPSH were collected between December 2, 2020, and May 11, 2021, whereas the MCCC data was collected between July 17, 2021, and October 16, 2021. The CXR dataset consisted of the following patient cases (see also Table 3.1):

- ✓ COVID-19 cases: 65 from SPSH and 308 from MCCC;
- ✓ Normal cases: 373 from SPSH;

Inclusion and exclusion criteria

The inclusion criteria for the COVID-19 cases in MCCC and SPSH was all patients who were confirmed to have COVID-19 using both RT-PCR and CXR imaging during the study period. The inclusion criteria for the normal cases in SPSH was all cases with SARS-CoV-2 negative test results and healthy CXR diagnoses. The other patients whose triage status was unknown were excluded from this study.

Sample size

All cases meeting the inclusion criteria during the study period were included.

In order to further check the efficacy of the DL model, open source CXR data sets were also acquired from additional four publicly available repositories, namely: COVID-19 radiography database [16] from Kaggle; Montgomery County CXR set [17] from the National Library of Medicine, USA; CXR images [18] from Kaggle, and Shenzhen CXR image set [19] from the National Library of Medicine, China. The CXR images collected from open source consisted of (see also Table 3.1):

- ✓ COVID-19 cases: 1200 from the COVID-19 radiography database [16];
- ✓ Normal cases: 394 from the COVID-19 radiography database [16], 80 from Montgomery County [17], 400 from Kaggle [18], and 326 from Shenzhen [19];

Ethical clearance for use of the CXR images was obtained from SPSH's institutional review board (IRB) with reference number V264/07/04/2021. St. Paul's Hospital Millennium Medical College's (SPHMMC) IRB issued the ethical clearance for the data collected from MCCC with reference number Pm2367. All CXR images were accessed by using the Medical Record Number (MRN) and patient name. Protected identifiable information such as patient name, patient sex, institution name, station name (radiology), patient id, patient birth date, and operator's name present in the metadata (header) was de-identified (anonymized). The anonymized CXR dataset was coded as follows. For example, in COVID-ET-PX, "COVID" represents COVID-19 confirmed case. For normal cases, "COVID" would be replaced by "NORMAL". "ET" represents data that has been acquired in Ethiopia. "P" is used for images collected from SPSH and would be replaced by "SP" for images collected at MCCC while "X" represents a specific patient referring number. The images were structured in a folder to be used for training, testing, and validation and were labeled by a radiologist (original sample images are presented in Figure 3.3).

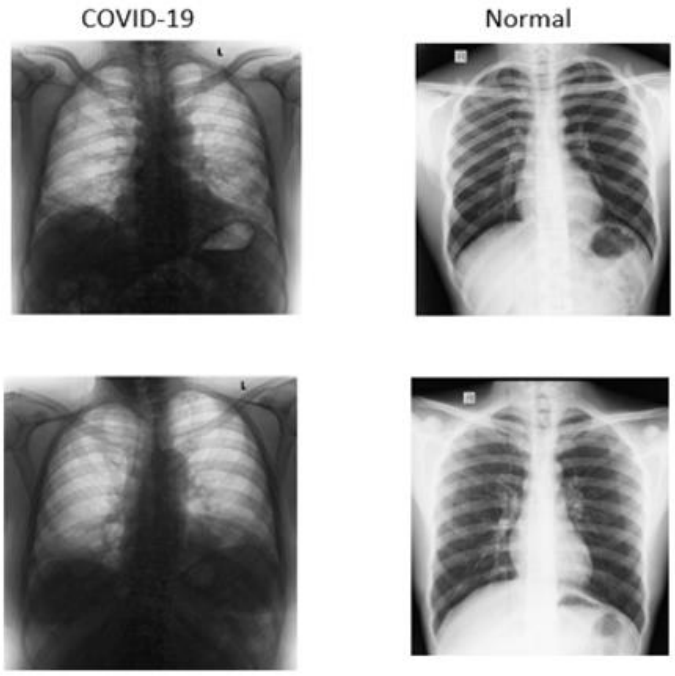


Figure 3.3 Randomly selected sample images from the two classes: COVID-19 (left) and normal (right).

Table 3.1 CXR images collected from different sources.

Source	Case	Data type	# of Images
Ethiopia (SPSH and MCCC)	COVID-19	CXR image	373
Ethiopia (SPSH)	Normal	CXR image	373
Open source [49]	COVID-19	CXR image	1200
Open source [78][79][51][52]	Normal	CXR image	1200

3.1.2 CXR image preprocessing

Two preprocessing techniques were applied on the CXR images before further analysis namely image complement and augmentation, as explained below.

Image complement

The CXR images collected from MCCC were all in a digital negative format (dark and bright pixel intensity values switched) compared to those acquired from SPSH and the open-source data repositories. The images were specifically acquired for clinical diagnosis purposes (not directly for research). Therefore, an image complement was applied to the CXR images collected from the MCCC. In the output image, dark areas become light, and light areas become dark. The output image

has the same size and class as the input image. Sample CXR images collected from the MCCC and their complements and respective histograms are shown in Figure 3.4.

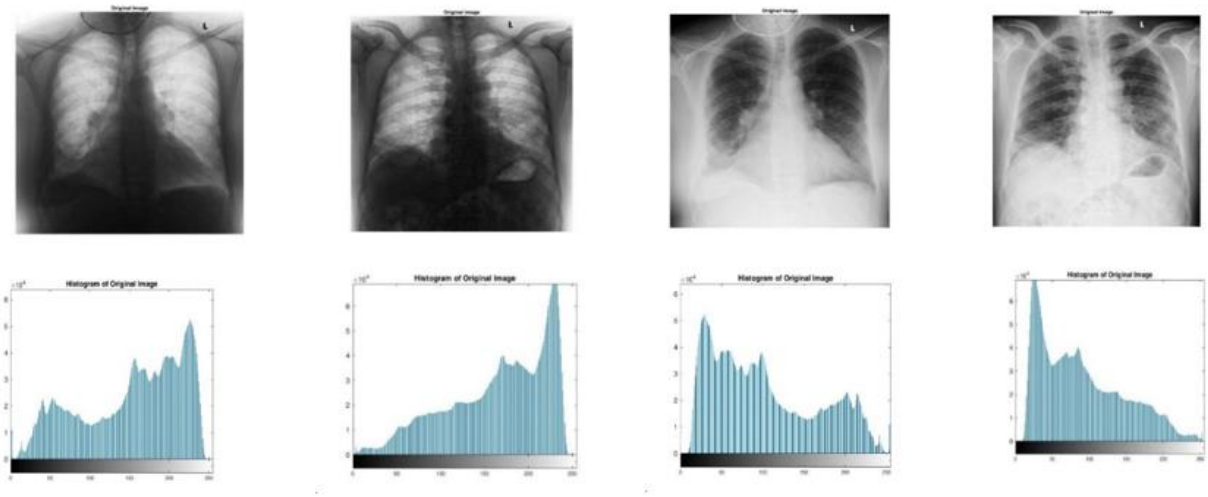


Figure 3.4 Top: Original CXR sample images collected from MCCC (1st and 2nd column) and their complements (3rd and 4th column); Bottom: corresponding histograms.

Image augmentation

A randomized augmentation including rotation, scaling, flip, reflection, and shear was applied to the training dataset (Figure 3.5). Augmentation enhances network training to be invariant to distortions and helps in reducing network overfitting [53][87][55].

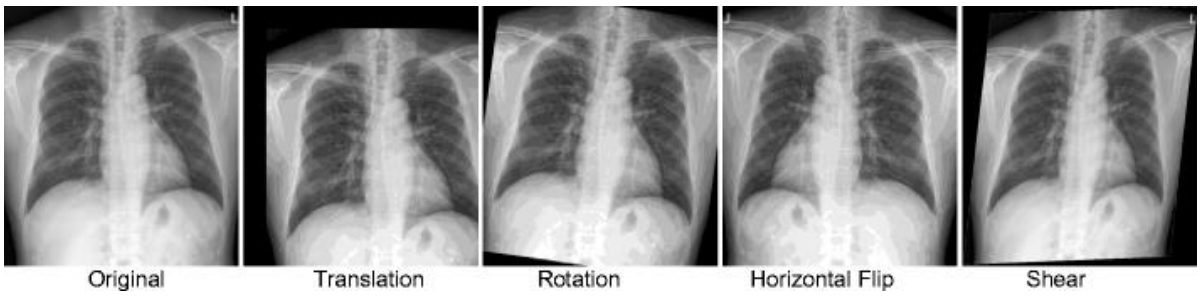


Figure 3.5 Illustration of CXR image data augmentation.

3.1.3 Network training

As explained in Chapter 2, there are different deep transfer learning approaches used in the literature for various applications. Two of the most popular ones, namely ResNet-50 and DenseNet-201, have been adopted in the current study during training. These two DL architectures are in favor given

image classification results reported in previous literatures (as also listed in Chapter 2). The binary classification was used to discriminate COVID-19 samples from that of the normal controls.

Data partitioning

A random image split was performed to use 60% of the image samples as a training dataset, 30% for validation, and the remaining 10% for testing the model. The validation dataset is used to provide an unbiased evaluation of a model fit on the training dataset while tuning model hyperparameters. The test dataset is used to describe the evaluation of a final tuned model on an unseen dataset. Three different scenarios were considered during partitioning:

- ✓ Only the Ethiopian dataset was used for training (373 COVID-19, 373 normal);
- ✓ The Ethiopian and open-source datasets were mixed for training (1200 + 373 COVID-19 and 1200 + 373 normal), and
- ✓ Only the open-source dataset was used for training (1200 COVID-19 and 1200 normal).

Considering the less prevalence of COVID-19 in the Ethiopian population [56], the DL model's performance was also checked by using only four images as a test dataset for COVID-19 cases (which is 1% of normal cases), whereas the split ratio of the other dataset was kept the same. In order to reduce biases in selecting the experiment data, a random sampling approach was used.

Transfer learning

The transfer learning techniques changed the last three layers of the neural network's architecture: the fully-connected layer, softmax layer, and classification output layer, and the 1000 classes used in ResNet-50 and DenseNet-201 were changed to three (in case of multi-class classification) and two (in case of binary classification) as also illustrated in Figure 3.6 for the case of ResNet-50.

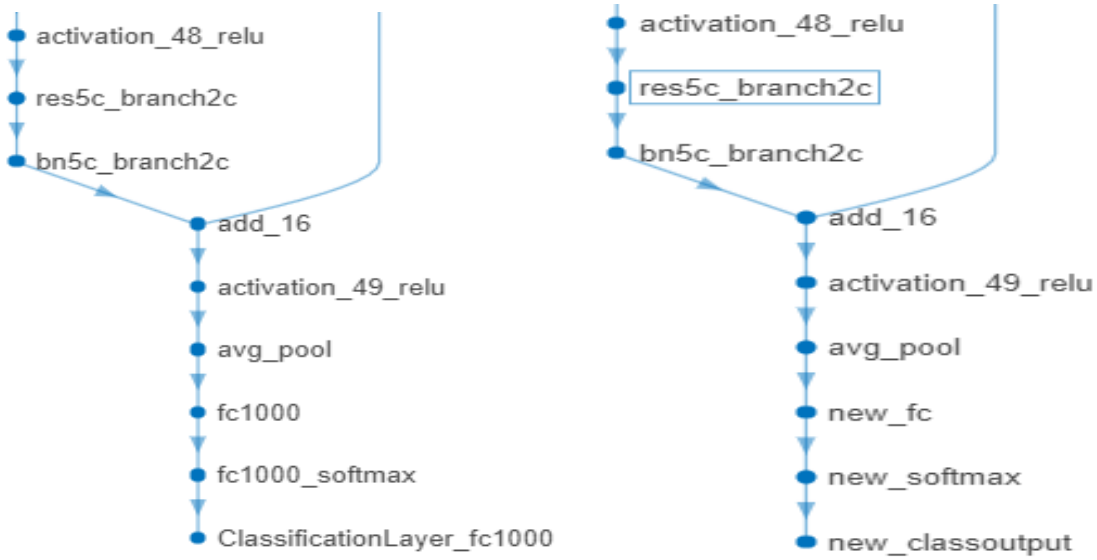


Figure 3.6 The original ResNet-50 last layers (left), and the transferred ResNet-50 last layers (right).

Training parameters of the network

Three parameters are important in setting up the DL scheme during training: epoch size, batch size as well as the learning rate. These parameters are picked to give best training accuracies at the end. A full pass through the entire data set is called an epoch. Maximum number of epochs to train the model can be specified. It is recommended to select a smaller number of epochs for small networks or for fine-tuning and transfer learning, where most of the learning is already done.

There are different options used to optimize the network hyperparameters including ‘rmsprop’ (root mean square propagation) and ‘sgdm’ (stochastic gradient descent with momentum). One other popular optimizer is Adams, derived from adaptive moment estimation, which has been used successfully in different studies. In the current study, Adam optimizer was used to network update parameters. Accordingly, the maximum number of epochs, the minimum-batch size, and the learning rate were set. To increase data randomness, data shuffling was enabled. While training the network, the progress was checked in terms of accuracy and loss at different iteration points and epochs. Figure 3.7 presents a sample training progress for multi-class classification using ResNet-50.

Experimental settings

A personal computer with a Microsoft Windows operating system, Intel® Core™ i5-72000U CPU @2.50GHz-2.71GHz, 8.00 GB RAM was used for implementation of the DL scheme on a MatLab (R2020a) environment.

Probability scoring and prediction visualization

The developed DL model has interpretability and visualization advantages to assist clinicians in providing a final diagnosis. It displays the probability score value for the diagnostically predicted class for a given input image. The visual explanations of the prediction were achieved using an occlusion sensitivity map to facilitate the investigation of the region of interest. The sensitivity map dictates the DL model decision-making areas on the CXR images that provide more contribution for the predicted class.

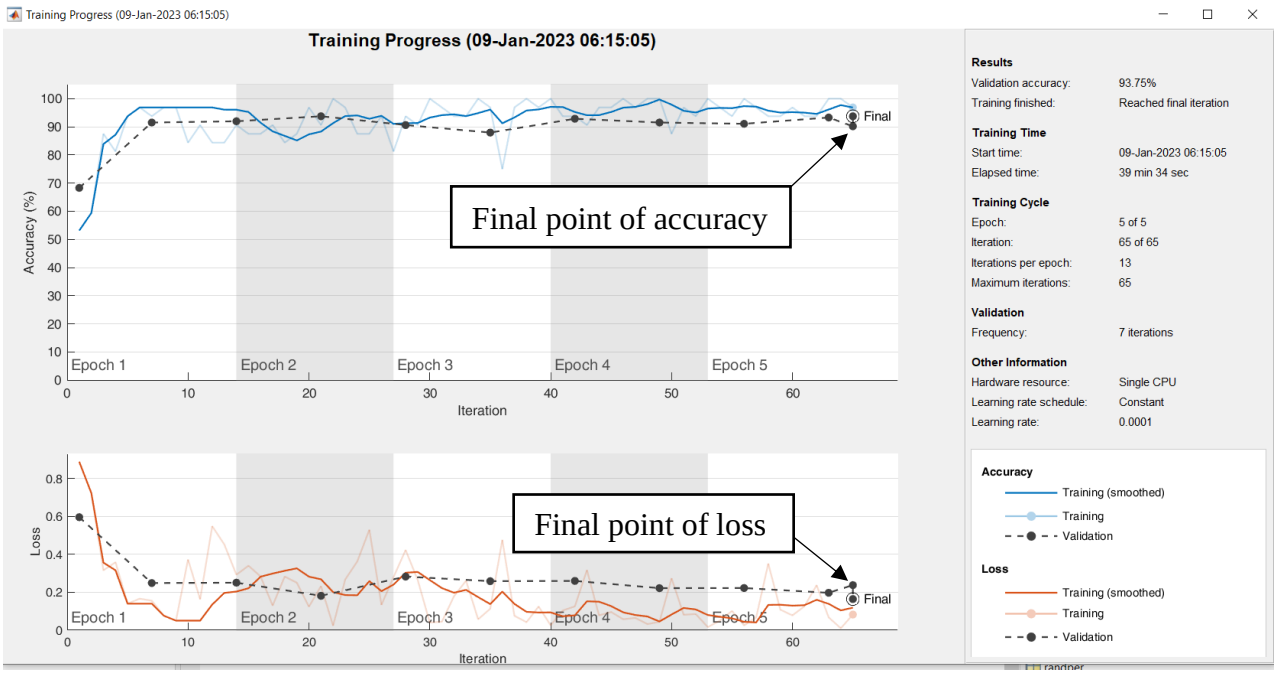


Figure 3.7 A sample training progress for the binary class classification using ResNet-50.

3.1.4 The DL method performance evaluation

Performance evaluation of the DL based classification scheme was carried out using five different matrices: sensitivity, specificity, precision, accuracy as well as fl-score. These matrices are computed using the number of true positives (TP), true negatives (TN), false positives (FP), and false negatives (FN) as shown in the equations below.

$$\text{Sensitivity} = \frac{\text{TP}}{\text{TP} + \text{FN}} = \text{Recall}$$

$$\text{Specificity} = \frac{\text{TN}}{\text{TN} + \text{FP}}$$

$$\text{Precision} = \frac{\text{TP}}{\text{TP+FP}} = \text{Postive Predictive Value (PPV)}$$

$$\text{Accuracy} = \frac{\text{TP} + \text{TN}}{\text{TP} + \text{TN} + \text{FP} + \text{FN}}$$

$$\text{F1} - \text{score} = 2\text{x}\frac{\text{Precision x Recall}}{\text{Precision} + \text{ Recall}}$$

3.2 Severity prediction of COVID-19 cases

The anthropomorphic data collected from confirmed COVID-19 cases was used in the severity prediction model development. The schematics of the proposed COVID-19 severity prediction model based on ML is illustrated in Figure 3.8.

3.2.1 Study dataset

Data source

The study participants were all laboratory-confirmed COVID-19 positive cases who were admitted to the care center during the study period.

Study period and data sample

The anthropomorphic data from the MCCC were collected between July 17, 2021, and October 16, 2021. The anthropomorphic data includes the following COVID-19 patient cases: moderate cases (n = 84), severe cases (n = 197), and critical cases (n = 27) (see Table 4.2 for the summary). The study participants were all laboratory-confirmed COVID-19 positive cases who were admitted to the care center during the study period.

Inclusion and exclusion criteria

The inclusion criteria for the COVID-19 cases in the care center was all confirmed COVID-19 patients admitted with filled-out triage sheets and also having CXR diagnostic results. The other patients whose triage status is unknown and whose CXR images were not available were excluded from this study.

Sample size

All cases meeting the inclusion criteria during the study period were included.

The St. Paul's Hospital Millennium Medical College's (SPHMMC) IRB issued the ethical clearance for the anthropomorphic data collected from MCCC with reference number Pm2367. The anthropomorphic data was accessed from the patient's triage sheet by using the MRN and the protected identifiable information was de-identified.

A total of 15 features acquired from the COVID-19 patient triage sheets have been used for the ML model development. The 15 features contain demographic characteristics, medical history, and COVID-19 associated symptoms. The list of the 15 features is presented in Table 3.3.

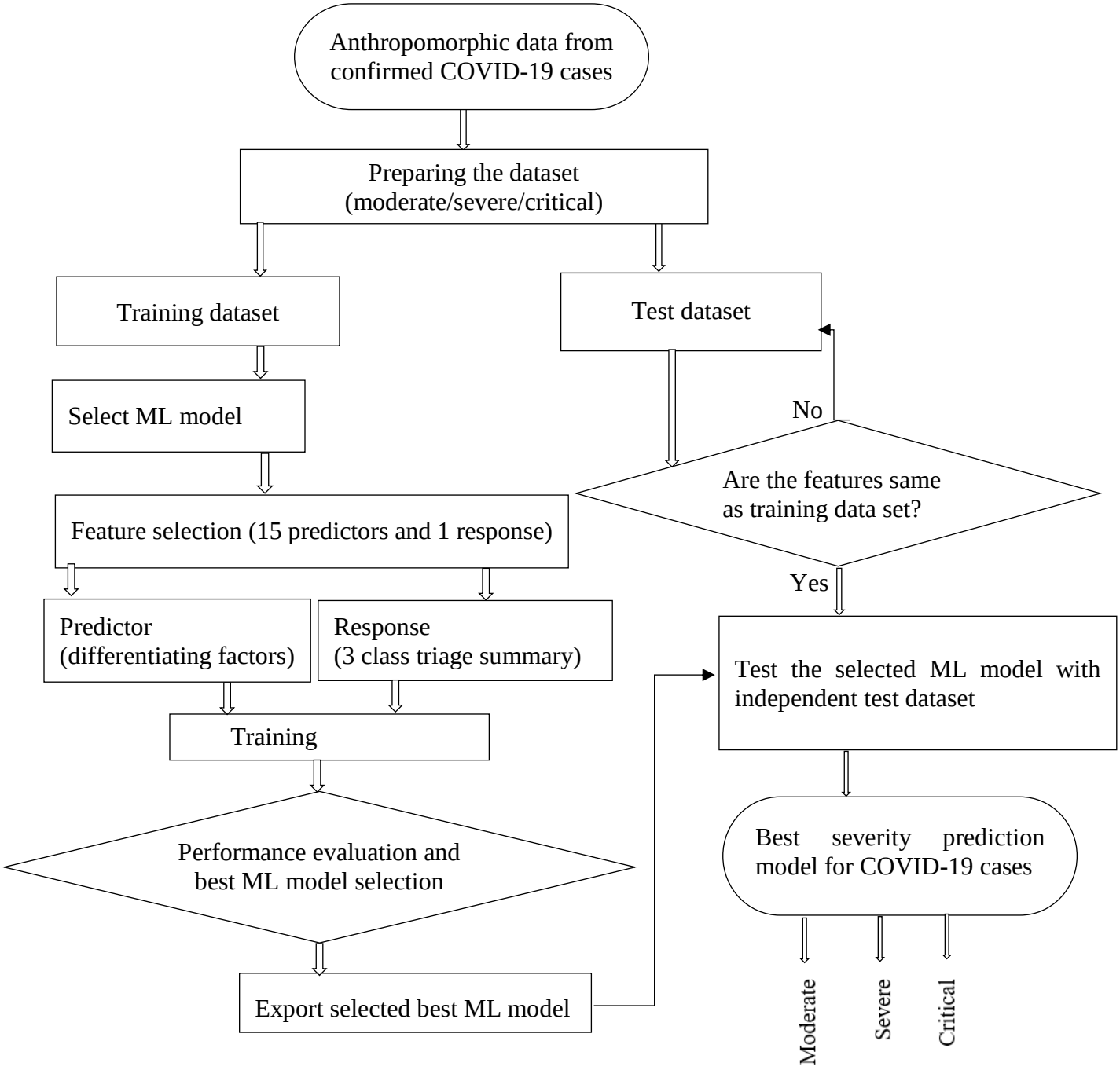


Figure 3.8 Schematic representation of the proposed COVID-19 severity prediction model based on machine learning (ML).

Table 3.2 The total number of collected anthropomorphic data from confirmed COVID-19 cases.

Source	Cases	Data type	Number of records
Ethiopia, MCCC	Moderate	Anthropomorphic	84
Ethiopia, MCCC	Severe	Anthropomorphic	197
Ethiopia, MCCC	Critical	Anthropomorphic	27

3.2.2 Anthropomorphic data preparation

The differentiating features were used to predict the severity status (response) of the patients into either moderate, severe, or critical cases. Patients with SpO_2<=65% and/or respiratory distress with or without additional symptoms/comorbidities are considered critical. Cases with 65%<=SpO_2<=93%, acute chest pain, and/or coughing up blood with or without additional symptoms/comorbidities are considered severe. While patients with normal SpO_2 level (i.e. >= 93%) but with shortness of breath (SOB), cough, active vomiting, acute abdominal pain, and/or stable comorbidities such as asthma, cardiac illness, Diabetic Mellitus (DM), and hypertension are considered moderate. The above is used as the ground truth information while developing the ML based severity prediction scheme. The two additional patient information used during the algorithm development are age and gender.

Each of these 15 features are inserted as input to the ML algorithm in a weighted sense. In the current study, the percentage of cases with that specific symptom/comorbidity has been used as a weighting factor. For example, if a patient is Asthmatic and 10 of the total population (i.e. 308) are confirmed Asthma cases, then this patient case will assume a weight factor of 3.2. If the patient is not Asthmatic, then he/she will get a zero weight. All weights are calculated based on this principle. Two of the features (out of 15) deviate from this rule, namely age, and gender. If a patient is between ages 15 and 29, he/she will assume a weighting factor of 6.5. If the age range is between 30 and 44, then a weighting factor of 23.5 is used, and so on. Similarly, if the gender happens to be male, then he will assume a weighting factor of 57.5 and if it is a female, then she will assume a weighting factor of 42.5. Table 3.3 summarizes the characteristics and differentiating factors presented on the studied patients for severity prediction.

Two of the most popular ML approaches are considered namely a support vector machine (SVM) and a K-nearest neighbor (KNN) and the efficacy of both has been checked in terms of predicting the severity level of COVID-19 based on the available anthropomorphic dataset.

Table 3.3 Differentiating features used for predicting COVID-19 severity levels.

No	Feature	Number of records	Percentage (%)
1	Age Range: 15 to 92	308	100
	15 to 29	19	6.5
	30 to 44	69	23.5
	45 to 59	92	31.3
	>= 60	114	38.7
2	Gender: Male	177	57.5
	Female	131	42.5
3	SOB: Yes	188	61
	NO	120	39
4	Cough: Yes:	247	80.2
	No:	61	19.8
5	65<=SpO_2<=93% on room air: Yes	192	62.3
	No	116	37.7
6	Respiratory distress: Yes	11	3.6
	No	297	96.4
7	Acute chest pain: Yes	59	19.2
	No	249	80.8
8	SpO_2 <= 65%: Yes	17	5.5
	No	291	94.5
9	Coughing up blood: Yes	6	2
	No	302	98
10	Actively vomiting: Yes	13	4.2
	No	295	95.8
11	Acute abdominal pain: Yes	10	3.2
	No	298	96.8
12	Asthma: Yes	10	3.2
	No	298	96.8
13	Cardiac illness: Yes	11	3.6
	No	267	96.4
14	DM (diabetes mellitus): Yes	57	18.5
	No	251	81.5
15	Hypertension: Yes	73	23.7
	No	235	76.3
	Triage summary: Moderate	84	27.3
	Severe	197	64
	Critical	27	8.7

There were other features included in the datasheet as differentiating features but not used in the ML model development due to the unavailability of data on the triage checklist during the study period.

These include active seizure, Systolic Blood Pressure (SBP) <= 90 mmHg, capillary refill > 2 sec, Glasgow Coma Scale (GCS) <= 13, change in mentation/confusion, diabetic + Random Blood Sugar (RBS) test > 350mg/dl, SBP >= 180 or Diastolic Blood Pressure (DBP) >= 110, pregnancy (1st and 2nd trimester), history of cancer diagnosis, and previously bedridden patients.

Note that a personal computer with a Microsoft Windows operating system, Intel® Core™ i5-72000U CPU @2.50GHz-2.71GHz, 8.00 GB RAM was used for implementation of the ML scheme on a MatLab (R2020a) environment.

3.2.3 The K-fold cross-validation

A 5-fold cross-validation approach has been used. In the first iteration, the first fold was used to test the model, and the rest were used to train the model. In the second iteration, the second fold was used as the testing set while the rest served as the training set. This process was repeated until each of the five folds had been used as the testing set (see also
Figure 3.9).

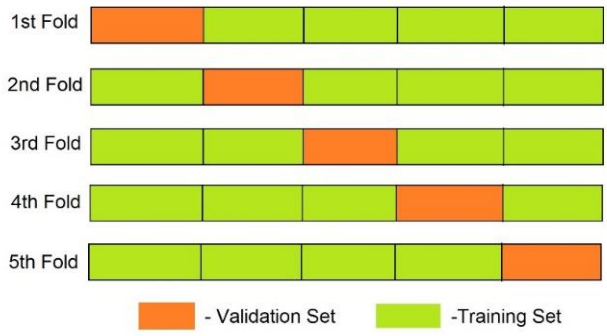


Figure 3.9 Schematic representation of the 5-fold cross-validation.

3.2.4 The ML model performance evaluation

The performance of the ML model in predicting the severity status of COVID-19 cases was evaluated using accuracy and the area under the receiver operating characteristic curve (ROC-AUC). The ROC curve is a probability curve plotted with a false positive rate (FPR) on the x-axis and a true positive rate (TPR) on the y-axis. A confusion matrix is used to describe and visualize the performance of the ML model and also provides insight into what the model misclassified. The confusion matrix consists of true positives (TP), true negatives (TN), false positives (FP), and false negatives (FN) (**Figure 3.10**). In addition, the model build time was also compared for both the DL and the ML models.

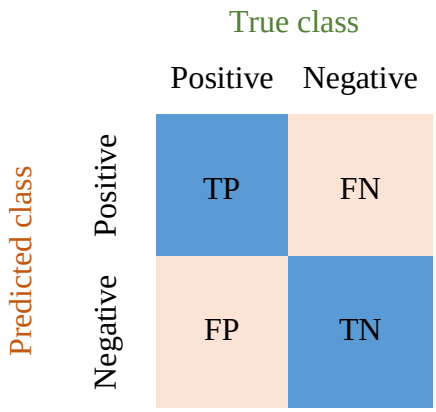


Figure 3.10 Theoretical structure of the confusion matrix.

Chapter 4 Results

4.1 Results of diagnostic prediction

4.1.1 DL model performance (using the Ethiopian CXR dataset)

The proposed DL model using the Ethiopian CXR images achieved a validation (using 30% of the total dataset) precision, sensitivity, accuracy, specificity, and F1-score of 98.06%, 90.18%, 94.20%, 98.21%, and 93.95%, respectively using the ResNet-50 model and 98.98%, 86.61%, 92.86%, 99.11% and 92.11, respectively using the DenseNet-201 model. The confusion matrix of the balanced test dataset (using 10% of the total dataset) using ResNet-50 show that out of 38 COVID-19 test data samples, 36 are correctly predicted as COVID-19 and 2 of COVID-19 cases were predicted as false negatives (see also Figure 4.1). In the case of the DL model performance evaluation using only 4 COVID-19 CXR images as a test dataset, the DL model classified 3 images as true positives and 1 image as false negative (see also Figure 4.1). The ResNet-50 model offered sensitivity and accuracy of 94.74% and 96.05% respectively using the balanced test dataset and 75% and 95.24% respectively using the imbalanced test dataset.

During the training of the DL models, ResNet-50 took around 39.34 minutes to complete the training whereas DenseNet-201 took around 183.59 minutes. A comparison of the computational time requirement of the two models, clearly indicates that the DL model building time of DenseNet-201 was significantly higher than that of the ResNet-50 model.

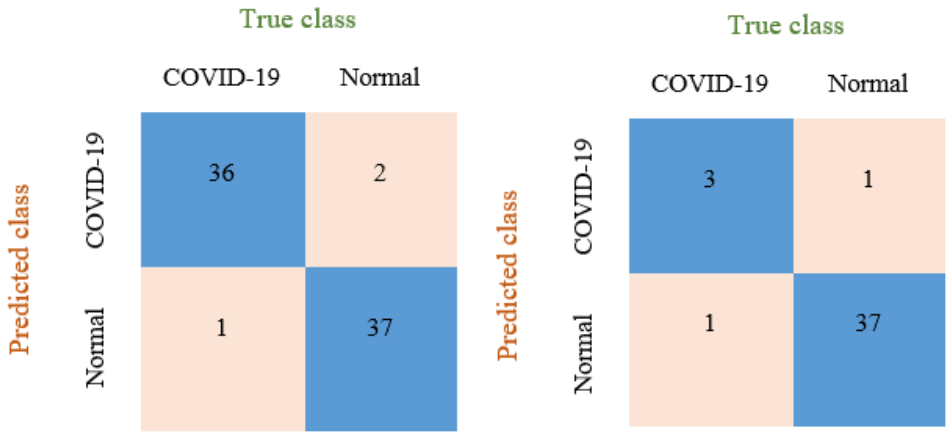


Figure 4.1 Confusion matrix with balanced test dataset (left) and with imbalanced test dataset (right) using Ethiopian CXR dataset.

Table 4.1 Comparison of the classification performance of ResNet-50 and DenseNet-201 DL models using CXR images taken from the Ethiopian dataset.

Ethiopian CXR dataset	COVID-19 vs Normal (Validation dataset)		COVID-19 vs Normal (Balanced test dataset)	COVID-19 vs Normal (Imbalanced test dataset)
Model	ResNet-50	DenseNet-201	ResNet-50	ResNet-50
Precision	98.06 %	98.98 %	97.30 %	75 %
Sensitivity (Recall)	90.18 %	86.61 %	94.74 %	75 %
Accuracy	94.20 %	92.86 %	96.05 %	95.24 %
Specificity	98.21 %	99.11 %	97.37 %	97.37 %
F1-score	93.95 %	92.11 %	96 %	75 %

4.1.2 DL model performance (using combined Ethiopian and open-source CXR datasets)

The classification of CXR images on the combined Ethiopian and open-source datasets achieved an accuracy of 97.35 % using the ResNet-50 model (Table 4.2). The elapsed time taken by the ResNet-50 model to build the binary- classification model on the combined dataset was 157.7 min. A summary of the results is presented in Table 4.2 below. Testing of the other DL model, DenseNet 201, was found computationally infeasible to be applied on the combined dataset.

Table 4.2 Binary classification performance of the ResNet-50 model applied on the combined CXR images acquired from both the Ethiopian and open-source databases.

Combined Ethiopian and open-source CXR datasets	Precision	Sensitivity (Recall)	Accuracy	Specificity	F1-score	Elapsed time
COVID-19 vs Normal	96.86 %	97.77 %	97.35 %	96.82 %	97.37 %	157.7 min

4.1.3 Deep learning model performance (using the open-source CXR dataset)

The binary classification scheme applied on CXR images acquired from the open-source database achieved an overall accuracy of 98.75 % using the ResNet-50 model (Table 4.3). The elapsed time when building the ResNet-50 model in the binary-class classification applied on the open-source dataset was around 137 min. Once again, testing of the other DL model, DenseNet 201, was found computationally infeasible to be applied on the open-source dataset.

Table 4.3 Binary classification performance of the ResNet-50 model when applied on the CXR images acquired from the open-source database.

Open-source dataset	Precision	Sensitivity (Recall)	Accuracy	Specificity	F1-score	Elapsed time
COVID-19 vs Normal	100 %	97.50 %	98.75 %	100 %	98.73 %	137 min

4.2 Visualization of predicted images

Figure 4.2 presents the sensitivity map generated after evaluating the performance of the DL-based schemes in predicting the different classes. In general, the models achieved maximum probability scores in predicting true positives. For example, the probability scores of the ResNet-50 model in predicting the two true COVID-19 cases presented in Figure 4.2 were 0.99999 and 0.98082, respectively. The same model resulted in probability scores of 0.0000049 and 0.0083337, respectively, in predicting the two confirmed COVID-19 cases as normal. Similar results were obtained when considering the normal cases. A summary of the probability scores computed for two randomly selected COVID-19 and normal cases is presented in Table 4.4 while the respective sensitivity maps are depicted in Figure 4.2. Note that red signatures on the sensitivity map indicate areas of maximum values on the occlusion map.

Table 4.4 Computed class prediction probability scores for selected CXR image samples taken from COVID-19 and normal cases.

No.	**Class**	**Predicted class**	**Probability score**	**Top predicted class**

1	COVID-19	COVID19	0.99999	COVID-19
		Normal	0.0000049	
2	COVID-19	COVID19	0.9808200	COVID-19
		Normal	0.0083337	
3	Normal	Normal	0.9960200	Normal
		COVID19	0.0008083	
4	Normal	Normal	0.9980400	Normal
		COVID19	0.0000020	

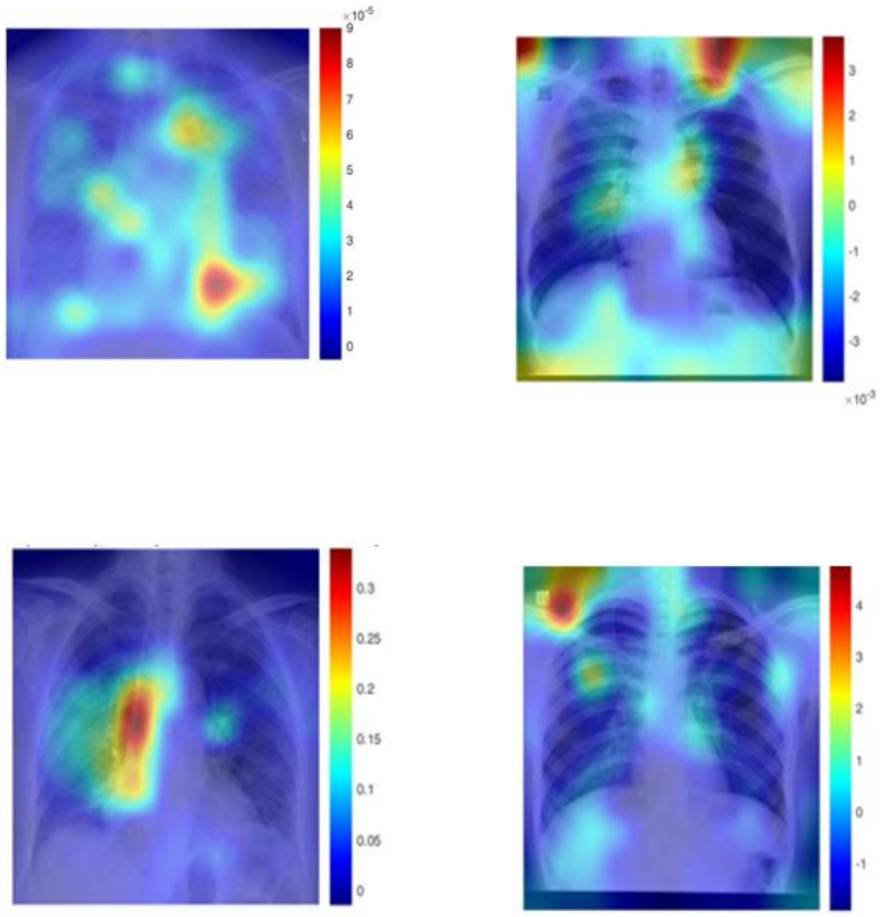


Figure 4.2 Randomly selected sample images with occlusion sensitivity mapping from the two classes namely: COVID-19 (1st column) and normal (2nd column).

4.3 Results of severity prediction

As explained in the previous chapter, two models are compared for their efficacy in predicting the severity status of the confirmed COVID-19 cases: SVM and KNN. When predicting the severity status, the SVM model offered an overall accuracy of 89.9% while it took the model just over 1.52 seconds to complete prediction of all samples into their respective classes: moderate, severe, or critical. While the KNN model performed with an accuracy of 89.6% with an elapsed time of 3.67 seconds. The true positive rates for classifying critical, moderate, and severe cases by using the SVM model were 100%, 61.7%, and 95.2%, and by that of KNN were 95.8%, 66.0%, and 94.2%, respectively. A summary has been presented in terms of the confusion matrix presented in Figure 4.3.

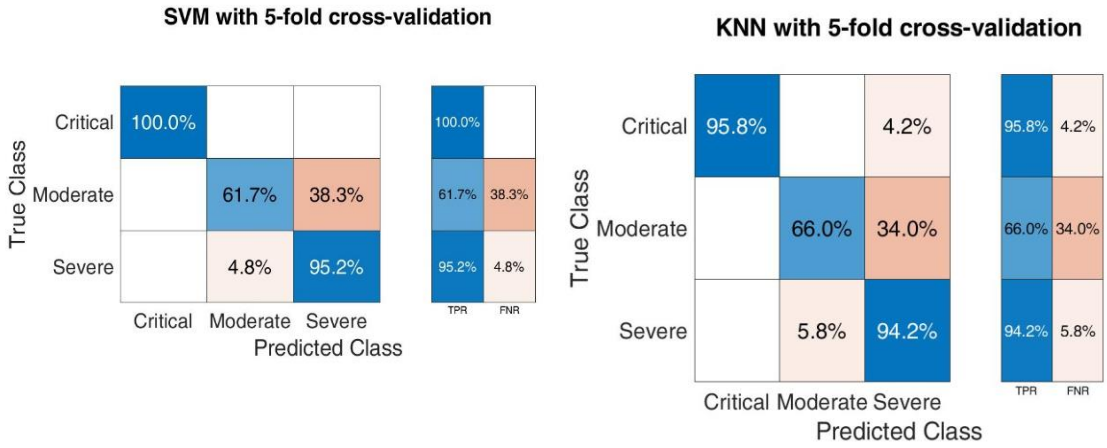


Figure 4.3 Computed confusion matrix including the true positive rate (TPR) and false negative rate (FNR) obtained using the SVM model (left) and the KNN model (right) with a 5-fold cross-validation.

The prediction accuracy of the proposed models was further validated using ROC curves and the respective AUC values, as also presented in Figure 4.4 and Figure 4.5 using SVM and KNN, respectively. The AUC values computed when using SVM were 1.0, 0.95, and 0.95 when predicting critical, severe, and moderate COVID-19 cases, respectively. The AUCs computed when using the KNN model were 0.98, 0.85, and 0.80 when predicting critical, severe, and moderate cases, respectively.

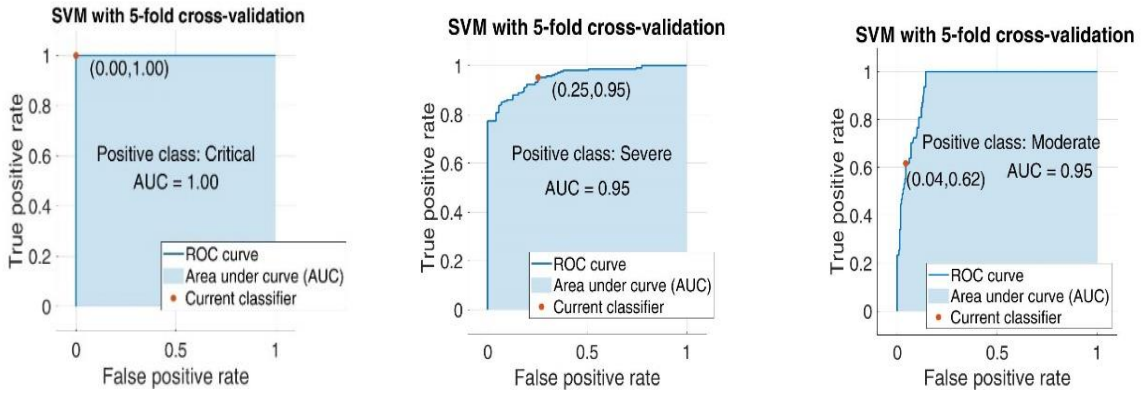


Figure 4.4 ROC curves and respective AUC values computed when using the SVM model in predicting critical (left), severe (middle), and moderate (right) COVID-19 cases.

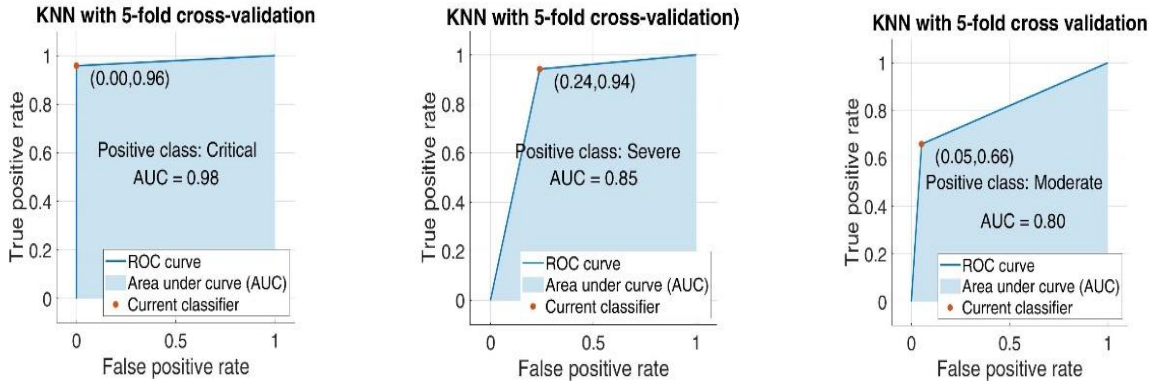


Figure 4.5 ROC curves and respective AUC values computed when using the KNN model in predicting critical (left), severe (middle), and moderate (right) COVID-19 cases.

Chapter 5 Discussion

5.1 Overview of the studies and the major findings

This work mainly focused on developing and studying automatic models for COVID-19 diagnostic prediction based on DL as well as severity prediction based on ML method using Ethiopian datasets. The DL model was trained on CXR images collected from SPSH and MCCC in Ethiopia. In addition, the performance of the DL model was validated on data acquired from an open-source database. The CXR image data distribution was performed by splitting the data into 60% for training, 30% for validation, and the remaining 10% for testing the DL model. The performance of the DL model was also checked on an imbalanced test dataset using only 4 COVID-19 CXR images (which is 1 % of the normal cases). The ML model was trained on anthropomorphic data collected from confirmed COVID-19 cases in MCCC.

The study exploited ResNet-50 and DenseNet-201 neural network architectures for developing the DL model. Useful image preprocessing techniques on the CXR images and transfer learning techniques on the DL models were included while developing the diagnostic prediction model. Every COVID-19 predicted image was identified using probability scoring and a decision visualization highlighted by the DL model. The two ML methods, namely SVM and KNN, were trained with a 5-fold cross-validation approach to select the best-performing COVID-19 severity case prediction tool.

The ResNet-50 model was able to classify the Ethiopian CXR images with a validation precision of 98.06%, a sensitivity of 90.18%, an accuracy of 94.20%, a recall of 98.21%, and an F1-score of 93.95. The F1-score offered by the ResNet-50 with the balanced and imbalanced test dataset was 96% and 75% respectively. By combining the Ethiopian dataset with the open-source dataset, an accuracy of 97.35% was obtained in the classification of COVID-19 and normal cases. On the severity prediction of the COVID-19 cases, the SVM model achieved an accuracy of 89.9% while the KNN achieved an accuracy of 89.6% using a 5-fold cross-validation approach.

The transfer learning approach has been predominant in literatures related to DL-based COVID-19 researches [30][34][37][57]. In the current work, a transfer learning approach was adopted by changing the last three layers of the neural network, and the training of the DL model was

performed with the hyper parameters of minimum batch size 32, epoch 5, an initial learning rate of 0.0001, and Adam optimizer.

Following the studies conducted by the authors in [37] and [57], image preprocessing is one of the major steps before training a DL model. In that regard, an image complement technique was applied on the CXR images collected from MCCC. Unlike the CXR images collected from SPSH and the open-source repositories, the MCCC's CXR images were in the form of image negatives (see Figure 3.4, for example). Data augmentation was the other useful method applied on the original CXR image datasets before further processing.

Two DL models, ResNet-50 and DeseNet-201, were used in the current study as other researchers has already demonstrated the promising performance of the two models when applied in COVID-19 diagnostic prediction [30][58]. It was found in the current study that the DenseNet-201 model requires an expressively higher training time than the ResNet-50 model (see also Figure 5.1), indicating that it has a higher computational cost [34].

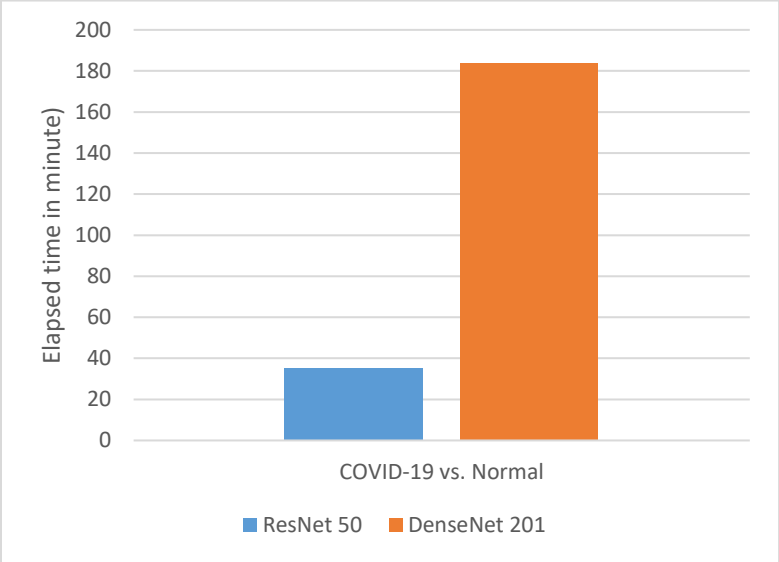


Figure 5.1 A graphical summary of the elapsed times taken by the DL models when applied on the Ethiopian local dataset.

It is known that in a given test performed at a given COVID-19 Center, most would test negative while a smaller number would test positive for the disease. This is the actual scenario. In order to

test the efficacy of the proposed deep learning model in such scenarios, an imbalanced test dataset-based experiment was performed. On selecting the experimental dataset in this case, a random sampling approach was attempted to reduce bias. Simple random sampling is one of the most successful methods several researchers use to mitigate sampling bias in deep learning applications. It ensures that everyone in the population has an equal chance of being selected. It has been seen in the result that, using the imbalanced test dataset on the developed DL model, there is a slight decrease in the accuracy and specificity of the model (see also Figure 4.1). The sensitivity and precision decreased significantly, mainly due to the fact that with the small number of given test data (in our case, 4 COVID-19 images), the single image misclassification cost was high (25%). The sensitivity, precision, and F1 score increased when the number of test datasets increased, as shown in Table 4.1.

Several studies have been carried out previously to make DL more interpretive in COVID-19 diagnostic prediction. Research done by the authors of [59], [60], and [61] suggested mappings (Grad-CAM, occlusion sensitivity) to provide an explainable view of the DL models focusing areas. Grad-CAM and occlusion sensitivity usually return qualitatively similar results [62]. However, the Grad-CAM map can miss fine details more than the occlusion sensitivity map [62]. In the current study, a probability scoring and occlusion sensitivity mapping have been used to make the proposed DL model more interpretable and explainable (see also Table 4.4 and Figure 4.2).

Different researchers reported that in most types of pneumonia cases, the bacteria or the virus reproduces itself and spreads throughout the lung areas quickly [63] [64]. On the other hand, COVID-19 pneumonia infects small areas and settles in, and then the COVID-19 virus uses the immune system to start spreading into other parts of the lung over time [63] [64]. Similar to the study done by the authors of [60] and [65], the current study was able to show that the DL model visualization for COVID-19 cases focused around the lung region. Whereas in normal cases, the DL model appeared to be more interested in the edges. The results shown in the probability scoring and the visualization option can support clinicians and enhance their screening skills to make faster and more accurate diagnoses.

Different studies observed the achievement of ML techniques in classifying COVID-19 severity status. A study conducted by the author of [28] found the performance of ML in classifying

COVID-19 cases using the SVM model and reported an accuracy of 77.5%. In another study reported by the authors in [44] found the performance of KNN with a cross-validation approach and achieved an accuracy of 80.4%. The authors of [44] and [66] also evaluated their ML model using a cross-validation approach. In the current work, both SVM and KNN were employed with a 5-fold cross-validation approach. Figure 5.3 shows the comparison of the confusion matrix generated by the trained SVM and KNN models. It was observed that KNN has a higher elapsed time than SVM to predict the COVID-19 severity cases (Figure 5.2). Some studies have shown the development of ML methods trained on data from laboratory results, demographic information, and clinical data to predict the severity status of COVID-19 patients. The research conducted by [67] used laboratory data to identify patients who require special attention and forecast their length of stay in specialized care. Another study used demographic and clinical parameters to use ML methods to stratify the need for ICU admissions, and their results showed that SVM can perform with an accuracy of 88.1% [68]. In the current work, focus was given to using anthropomorphic features containing demographic information, symptoms, and general medical conditions of COVID-19 patients to establish the ML based approach that can predict severity levels (Table 3.3). Results have shown that the SVM model can predict the severity status of COVID-19 patient cases with higher accuracy than KNN exploiting 15 differentiating simple features without the need for laboratory testing.

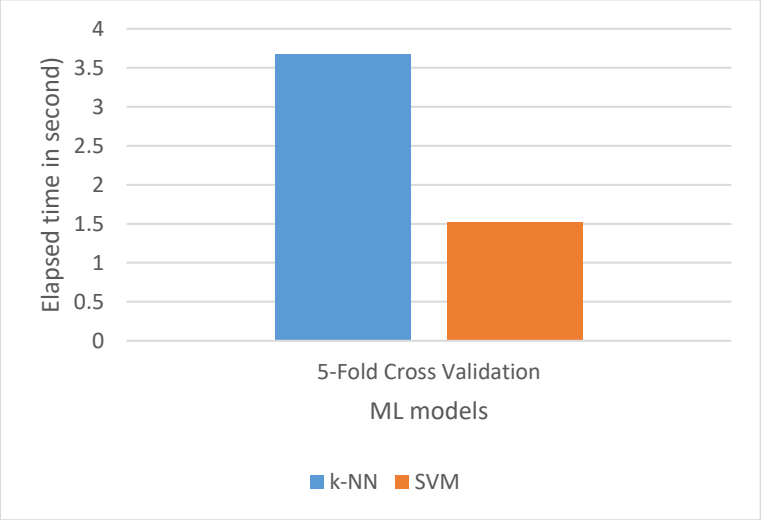
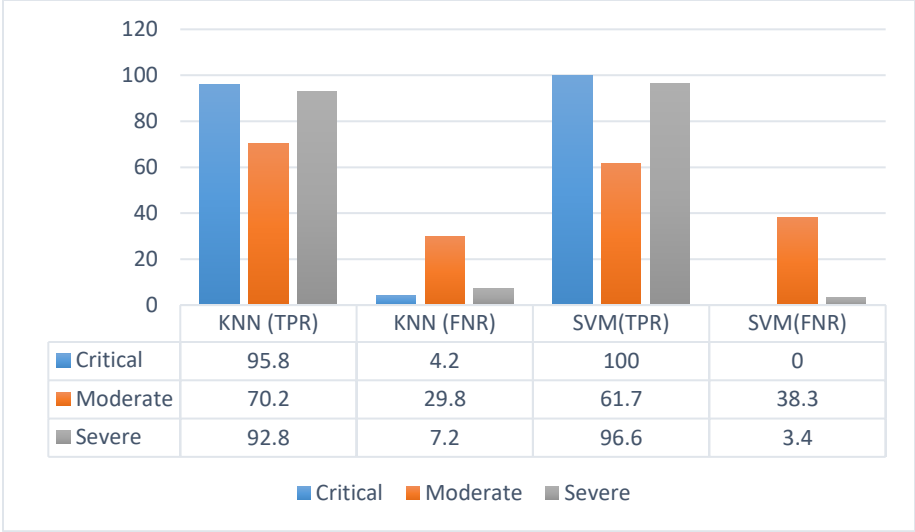


Figure 5.2 A graphical summary of the elapsed times taken by the ML models when applied on the Ethiopian local dataset.



.

Figure 5.3 Comparison of the confusion matrix for the trained SVM and KNN models.

5.2 Significance and limitations of the study

Early screening and severity status prediction of COVID-19 are key both in terms of slowing the speed of the spread and viably treating incoming patients. The AI detection method could deliver many potential benefits. From a methodical perspective, this study illustrates a step toward the potential of an AI-driven system for COVID-19 diagnostic and severity prediction, based on DL and ML models. The proposed method used a wide variety of metrics, including accuracy, sensitivity, specificity, precision, F1-score, confusion matrix, and ROC-AUC to evaluate the performance of models. From a data perspective, this study provides new datasets to augment Ethiopian datasets with existing publicly available datasets, which were used by the research community to develop AI-driven COVID-19 detection models. As far as the literature review carried out in the current study, collecting and utilizing the Ethiopian dataset for DL model development and anthropomorphic data for severity prediction based on ML has never been done before. This can minimize the drawbacks and issues of data inclusiveness for AI applications. From a clinical standpoint, AI models could be used as a screening tool to assist radiologists screen suspected COVID-19 patients, thereby shortening the waiting time for clinical decisions. It is noted that the development of AI-driven detection of COVID-19 infection does not intend to replace the

RT-PCR test, as it is the gold standard in diagnosing COVID-19. Rather, AI-based detection aims to augment the inaccessibility of RT-PCR machines in many countries.

In the proposed study, there was a limitation in the number of datasets acquired, especially in the Ethiopian dataset while establishing both diagnostic prediction and severity prediction models. Usually, for the ML method, the larger the dataset, the better the result will be. Preparing and collecting data from different institutes might improve the results and make them more general. The other limitation is that only frontal CXR was used, but it has been shown that lateral view CXR is also helpful in the diagnosis of abnormalities in CXR [38]. Such and similar other issues were beyond the scope of the current study.

Chapter 6 Conclusions and future works

6.1 Conclusions

This thesis demonstrated the potential of an artificial intelligence-based approach for the diagnostic and severity prediction of COVID-19 using data collected from Ethiopia. The proposed diagnostic prediction model was developed using the DL method for discriminating COVID-19 and normal cases using CXR images and the overall results were well acceptable. Experiments were performed on both balanced and imbalanced test datasets. The imbalanced test dataset experiment was performed to consider the lower prevalence of COVID-19 cases compared to normal cases.

The proposed severity prediction model was developed using the ML method for predicting the severity status of COVID-19 cases as moderate, severe, or critical. The ML method also demonstrated commendable results using demographic data, medical history, and COVID-19-related symptom data collected from confirmed COVID-19 patients. Based on the specific objective, the following conclusions can be made:

- ✓ The ResNet-50 model with the Adam optimizer, initial learning rate of 0.0001, minimum batch size of 32, and Epoch of 5, achieved a test accuracy and F1-score of 96.05% and 96% using the balanced test dataset, and 95.24% and 75% using the imbalanced test dataset;
- ✓ The availability of class-predicting scores and visual explanations using an occlusion sensitivity mapping offers a rapid second opinion to assist clinicians in the final interpretation of CXR images, and
- ✓ The ML model (using SVM) can effectively identify the severity status of patients among those exposed to COVID-19 with an accuracy of 89.9%. This allows healthcare front-liners to isolate and give timely treatment to incoming patients. Furthermore, this will aid in resource allocation based on the severity status of the patients.

6.2 Future works

There are still rooms to enhance the performance of the proposed DL and ML based approaches so that they be clinically applicable.

- ✓ To apply multi-class classification based on CXR images of different pneumonia types (bacterial, viral, and fungal infections).
- ✓ To have plenty of data to train the DL models from scratch instead of adopting transfer learning techniques.
- ✓ To try combination of other DL architectures for building an improved accuracy in COVID-19 detection system using the CXR images.
- ✓ To consider other differentiating features that are available on the patient's triage sheet, such as clinical manifestations, for developing an improved ML model.

List of references

[1] WHO, “Coronavirus disease (COVID-19),” 2020.

[2] M. A. Shereen, S. Khan, A. Kazmi, N. Bashir, and R. Siddique, “COVID-19 infection: Origin, transmission, and characteristics of human coronaviruses,” *J. Adv. Res.*, vol. 24, pp. 91–98, 2020, doi: 10.1016/j.jare.2020.03.005.

[3] V. Nikolaou, S. Massaro, M. Fakhimi, L. Stergioulas, and W. Garn, “COVID-19 diagnosis from chest x-rays: developing a simple, fast, and accurate neural network,” *Health Information Science and Systems*, vol. 9, no. 1. 2021, doi: 10.1007/s13755-021-00166-4.

[4] WHO, “WHO Coronavirus (COVID-19) Dashboard,” 2022, [Online]. Available: https://covid19.who.int/.

[5] WHO, “Tracking SARS-CoV-2 variants,” no. March 22, 2022, 2022, [Online]. Available: https://www.who.int/en/activities/tracking-SARS-CoV-2-variants/.

[6] K. McIntosh, M. S. Hirsch, and A. Bloom, “Coronavirus disease 2019 (COVID-19),” *UpToDate Hirsch MS Bloom*, vol. 5, no. 1, 2020.

[7] C.-19 V. Tracker, “VACCINES CANDIDATES IN CLINICAL TRIALS,” [Online]. Available: https://www.who.int/publications/m/item/draft-landscape-of-covid-19-candidate-vaccines (Accessed on October 20, 2020).

[8] “COVID-19 Weekly Epidemiological Update,” no. August, 2021, [Online]. Available: https://reliefweb.int/sites/reliefweb.int/files/resources/20210831_Weekly_Epi_Update_55.pdf.

[9] T. W. Leulseged *et al.*, “COVID-19 disease severity and associated factors among Ethiopian patients: A study of the millennium COVID-19 care center,” *PLoS One*, vol. 17, no. 1 January, pp. 1–18, 2022, doi: 10.1371/journal.pone.0262896.

[10] Z. Wu and J. M. McGoogan, “Characteristics of and Important Lessons from the Coronavirus Disease 2019 (COVID-19) Outbreak in China: Summary of a Report of 72314 Cases from the Chinese Center for Disease Control and Prevention,” *JAMA - Journal of the American Medical Association*, vol. 323, no. 13. pp. 1239–1242, 2020, doi: 10.1001/jama.2020.2648.

[11] I. Consultant and A. Ababa, “Ethiopia ’ s Response to the Covid-19 Pandemic : Measures , Impacts and Lessons,” no. April, 2022.

[12] S. Basu, S. Mitra, and N. Saha, “Deep Learning for Screening COVID-19 using Chest X-Ray Images,” *2020 IEEE Symp. Ser. Comput. Intell. SSCI 2020*, no. Ml, pp. 2521–2527, 2020, doi: 10.1109/SSCI47803.2020.9308571.

[13] N. Narayan Das, N. Kumar, M. Kaur, V. Kumar, and D. Singh, “Automated Deep Transfer Learning-Based Approach for Detection of COVID-19 Infection in Chest X-rays,” *IRBM*, Jul. 2020, doi: 10.1016/J.IRBM.2020.07.001.

[14] T. C. Williams *et al.*, “Sensitivity of RT-PCR testing of upper respiratory tract samples for SARS-CoV-2 in hospitalised patients : a retrospective cohort study [version 2 ; peer review : 2 approved],” pp. 1–13, 2022.

[15] N. Jawerth, “How is the COVID-19 virus detected using real time RT–PCR?,” *IAEA Bull.*, vol. 61, no. 2, pp. 8–11, 2020.

[16] Y. Pan *et al.*, “Serological immunochromatographic approach in diagnosis with SARS-CoV-2 infected COVID-19 patients,” *J. Infect.*, vol. 81, no. 1, pp. e28–e32, 2020, doi: 10.1016/j.jinf.2020.03.051.

[17] C. Sciences, “SARS-CoV-2 (Covid-19): Diagnosis by IgG/IgM Rapid Test,” [Online]. Available: https://www.clinisciences.com/en/read/newsletter-26/sars-cov-2-covid-19-diagnosis-by-2264.html.

[18] K. Green *et al.*, “What tests could potentially be used for the screening , diagnosis and monitoring of COVID-19 and what are their advantages and disadvantages ? 2 NIHR Newcastle In Vitro Diagnostics Co-operative Newcastle upon Tyne NHS Hospitals Foundation Trust , Newcast,” *Cent. Evidence-Based Med. Dev. Promot. disseminates better Evid. Healthc.*, 2019.

[19] C. for D. C. and P. (CDC), “COVID-19: Interim Guidelines for COVID-19 Antibody Testing,” [Online]. Available: https://www.cdc.gov/coronavirus/2019-ncov/lab/resources/antibody-tests-guidelines.html.

[20] F. E. et Al., “COVID-19 diagnostic testing,” *Am. Nurse Journal*, 2020, [Online]. Available: https://www.myamericannurse.com/covid-19-diagnostic-testing/.

[21] R. Maia *et al.*, “Diagnosis Methods for COVID-19: A Systematic Review,” *Micromachines*, vol. 13, no. 8, pp. 1–17, 2022, doi: 10.3390/mi13081349.

[22] S. Inui *et al.*, “The role of chest imaging in the diagnosis, management, and monitoring of coronavirus disease 2019 (COVID-19),” *Insights Imaging*, vol. 12, no. 1, pp. 1–14, 2021,

doi: 10.1186/s13244-021-01096-1.

[23] E. S. Mtsweni *et al.*, “No 主観的健康感を中心とした在宅高齢者における 健康関連指標に関する共分散構造分析Title,” *Eng. Constr. Archit. Manag.*, vol. 25, no. 1, pp. 1–9, 2020, [Online]. Available: http://dx.doi.org/10.1016/j.jss.2014.12.010%0Ahttp://dx.doi.org/10.1016/j.sbspro.2013.03.034%0Ahttps://www.iiste.org/Journals/index.php/JPID/article/viewFile/19288/19711%0Ahttp://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.678.6911&rep=rep1&type=pdf.

[24] X. Qi *et al.*, “Abnormal Coagulation Function of Patients With COVID-19 Is Significantly Related to Hypocalcemia and Severe Inflammation,” vol. 8, no. June, pp. 1–10, 2021, doi: 10.3389/fmed.2021.638194.

[25] W. A. Hatamleh, H. Tarazi, C. Subbalakshmi, and B. Tiwari, “Analysis of Chest X-Ray Images for the Recognition of COVID-19 Symptoms Using CNN,” *Wirel. Commun. Mob. Comput.*, vol. 2022, 2022, doi: 10.1155/2022/3237361.

[26] S. Minaee, R. Kafieh, M. Sonka, S. Yazdani, and G. Jamalipour Soufi, “Deep-COVID: Predicting COVID-19 from chest X-ray images using deep transfer learning,” *Med. Image Anal.*, vol. 65, Oct. 2020, doi: 10.1016/j.media.2020.101794.

[27] Q. Liu *et al.*, “Machine learning models for predicting critical illness risk in hospitalized patients with COVID-19 pneumonia,” *J. Thorac. Dis.*, vol. 13, no. 2, pp. 1215–1219, 2021, doi: 10.21037/JTD-20-2580.

[28] S. Guhathakurata, S. Kundu, A. Chakraborty, and J. S. Banerjee, “A novel approach to predict COVID-19 using support vector machine,” *Data Sci. COVID-19*, no. 2021 May 21, pp. 351–364, 2021, doi: 10.1016/b978-0-12-824536-1.00014-9.

[29] Pmf-research, “ARTIFICIAL INTELLIGENCE IN MEDICAL IMAGING,” [Online]. Available: https://www.pmf-research.eu/en/artificial-intelligence-in-medical-imaging/.

[30] A. A. Abdulmunem, Z. A. Abutiheen, and H. J. Aleqabie, “Recognition of Corona virus disease (COVID-19) using deep learning network,” *Int. J. Electr. Comput. Eng.*, vol. 11, no. 1, pp. 365–374, 2021, doi: 10.11591/ijece.v11i1.pp365-374.

[31] S. Bhatti, A. Aziz, N. Nadeem, I. Usmani, and P. Muhammad, “Automatic Classification of the Severity of COVID-19 Patients Based on CT Scans and X-rays Using Deep Learning,” *Eur. J. Mol. Clin. Med.*, vol. 07, no. 10, pp. 1436–1455, 2020.

[32] P. Kedia, Anjum, and R. Katarya, “CoVNet-19: A Deep Learning model for the detection and analysis of COVID-19 patients,” *Appl. Soft Comput.*, vol. 104, p. 107184, Jun. 2021, doi: 10.1016/j.asoc.2021.107184.

[33] L. Sarker, M. Islam, T. Hannan, A. Zakaria, Z. Ahmed, and A. Zakaria, “COVID-DenseNet: A Deep Learning Architecture to Detect COVID-19 from Chest Radiology Images,” *Preprints*, no. May, 2020, doi: 10.20944/preprints202005.0151.v1.

[34] K. El Asnaoui and Y. Chawki, “Using X-ray images and deep learning for automated detection of coronavirus disease,” *J. Biomol. Struct. Dyn.*, vol. 39, no. 10, pp. 1–12, 2020, doi: 10.1080/07391102.2020.1767212.

[35] W. LINDA, “A tailored deep convolutional neural network design for detection of covid-19 cases from chest radiography images,” *J. Netw. Comput. Appl.*, 2020.

[36] Z. P. Ali Narin, Ceren Kaya, “Automatic Detection of Coronavirus Disease (COVID-19) Using X-ray Images and Deep Convolutional Neural Networks.”

[37] M. R. Karim, T. Döhmen, D. Rebholz-Schuhmann, S. Decker, M. Cochez, and O. Beyan, “DeepCOVIDExplainer: Explainable COVID-19 Diagnosis Based on Chest X-ray Images,” *arXiv Prepr. arxiv2004.04582*, 2020.

[38] M. Siddhartha and A. Santra, “COVIDLite: A depth-wise separable deep neural network with white balance and CLAHE for detection of COVID-19,” *arXiv Prepr. arXiv2006.13873*, 2020.

[39] I. Castiglioni *et al.*, “Artificial intelligence applied on chest X-ray can aid in the diagnosis of COVID-19 infection: a first experience from Lombardy, Italy,” *medRxiv*, 2020.

[40] I. D. Apostolopoulos and T. A. Mpesiana, “Covid-19: automatic detection from x-ray images utilizing transfer learning with convolutional neural networks,” *Phys. Eng. Sci. Med.*, p. 1, 2020.

[41] T. Ozturk, M. Talo, E. A. Yildirim, U. B. Baloglu, O. Yildirim, and U. R. Acharya, “Automated detection of COVID-19 cases using deep neural networks with X-ray images,” *Comput. Biol. Med.*, p. 103792, 2020.

[42] A. Abbas, M. M. Abdelsamea, and M. M. Gaber, “Classification of COVID-19 in chest X-ray images using DeTraC deep convolutional neural network,” *arXiv Prepr. arXiv2003.13815*, 2020.

[43] “Combination of four clinical indicators predicts the severe_critical symptom of patients

infected COVID-19 _ Elsevier Enhanced Reader.pdf.” 2020.

[44] R. Square, “Prediction of COVID-19 Possibilities using KNN Classification Algorithm,” 2020, [Online]. Available: https://assets.researchsquare.com/files/rs-70985/v2_stamped.pdf.

[45] “Identification of high-risk COVID-19 patients using machine learning _ Enhanced Reader.pdf.” .

[46] R. Article, “Classification of Covid-19 Dataset with Some Machine Learning Methods,” pp. 36–44, 2020.

[47] S. Malaysiana, R. Pesakit, A. Pembelajaran, and M. Diselia, “Prediction of COVID -19 Patient using Supervised Machine Learning Algorithm,” vol. 50, no. 8, pp. 2479–2497, 2021.

[48] S. S. Aljameel, I. U. Khan, N. Aslam, M. Aljabri, and E. S. Alsulmi, “Machine Learning-Based Model to Predict the Disease Severity and Outcome in COVID-19 Patients,” vol. 2021, 2021.

[49] Kaggle, “COVID-19 Radiography Database.” https://www.kaggle.com/datasets/tawsifurrahman/covid19-radiography-database.

[50] P. Mooney, “Kaggle.” https://www.kaggle.com/paultimothymooney/chest-xray-pneumonia.

[51] U. Montgomery County, MD, “Montgomery County chest X-ray set,” *Natl. Libr. Med. Natl. Institutes Heal. Bethesda, MD, USA*.

[52] C. S.-T. S. set-C. X-ray, “Shenzhen chest X-ray set,” *Natl. Libr. Med. Natl. Institutes Heal. Shenzhen, China*.

[53] V. Kumar, A. Zarrad, R. Gupta, and O. Cheikhrouhou, “COV-DLS: Prediction of COVID-19 from X-Rays Using Enhanced Deep Transfer Learning Techniques,” *J. Healthc. Eng.*, vol. 2022, 2022, doi: 10.1155/2022/6216273.

[54] C. Shorten and T. M. Khoshgoftaar, “A survey on Image Data Augmentation for Deep Learning,” *J. Big Data*, vol. 6, no. 1, 2019, doi: 10.1186/s40537-019-0197-0.

[55] A. A. Reshi *et al.*, “An Efficient CNN Model for COVID-19 Disease Detection Based on X-Ray Image Classification,” *Complexity*, vol. 2021, 2021, doi: 10.1155/2021/6621607.

[56] D. K. Deribe, “COVID-19 in Ethiopia: status and responses,” *The Royal Society of Tropical Medicine and Hygiene*, Addis Ababa, Jun. 20, 2020.

[57] M. Faisal, F. Albogamy, H. Elgibreen, M. Algabri, S. A. M. Alvi, and M. Alsulaiman, "COVID-19 diagnosis using transfer-learning techniques," *Intell. Autom. Soft Comput.*, vol. 29, no. 3, pp. 649–667, 2021, doi: 10.32604/iasc.2021.017898.

[58] A. Jaiswal, N. Gianchandani, D. Singh, V. Kumar, and M. Kaur, "Classification of the COVID-19 infected patients using DenseNet201 based deep transfer learning," *J. Biomol. Struct. Dyn.*, pp. 1–8, 2020.

[59] L. R. Baltazar *et al.*, "Artificial intelligence on COVID-19 pneumonia detection using chest xray images," *PLoS One*, vol. 16, no. 10 October, Oct. 2021, doi: 10.1371/JOURNAL.PONE.0257884.

[60] M. Nishio *et al.*, "Deep learning model for the automatic classification of COVID-19 pneumonia, non-COVID-19 pneumonia, and the healthy: a multi-center retrospective study," *Sci. Rep.*, vol. 12, no. 1, pp. 1–10, 2022, doi: 10.1038/s41598-022-11990-3.

[61] J. B. M. Salazar W. Espeche el.Ennis, "Covid-19 detection via deep neural network and occlusion sensitivity maps," *Ann Oncol*, no. January, pp. 2–5, 2020, [Online]. Available: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC7254017/pdf/main.pdf.

[62] I. The MathWorks, "Understand Network Predictions Using Occlusion," 2022, [Online]. Available: https://www.mathworks.com/help/deeplearning/ug/understand-network-predictions-using-occlusion.html.

[63] C. C. medical Professional, "COVID Pneumonia," 2022. https://my.clevelandclinic.org/health/diseases/24002-covid-pneumonia (accessed Oct. 08, 2022).

[64] M. Paul, "Why COVID-19 pneumonia lasts longer, causes more damage than typical pneumonia," *Nature*, Jan. 11, 2021.

[65] W. Zhao, W. Jiang, and X. Qiu, "Fine-tuning convolutional neural networks for covid-19 detection from chest x-ray images," *Diagnostics*, vol. 11, no. 10. 2021, doi: 10.3390/diagnostics11101887.

[66] C. N. Villavicencio, J. Jerison, E. Macrohon, X. A. Inbaraj, J. Jeng, and J. Hsieh, "COVID-19 Prediction Applying Supervised Machine Learning Algorithms with Comparative Analysis Using WEKA," 2021.

[67] M. Pourhomayoun and M. Shakibi, "Predicting mortality risk in patients with COVID-19 using machine learning to help medical decision-making," *Smart Heal.*, vol. 20, no.

November 2020, p. 100178, 2021, doi: 10.1016/j.smhl.2020.100178.

[68] H. M. Alshanbari, T. Mehmood, W. Sami, W. Alturaiki, M. A. Hamza, and B. Alosaimi, “Prediction and Classification of COVID-19 Admissions to Intensive Care Units (ICU) Using Weighted Radial Kernel SVM Coupled with Recursive Feature Elimination (RFE),” vol. 2, no. December 2019, 2022.