Multimodal Deep Learning for Predicting Mechanical Ventilation Duration from Chest X-ray and Clinical Data

by

Abrar Shahrier Arnab
21201451
Yeasin Fiyaz
21201087
S M Mahidur Rahman
24241319
Umma Souda Alfa
21201008

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

Department of Computer Science and Engineering
BRAC University
January 2026.

Declaration

It is hereby declared that

1. The thesis submitted is my/our own original work while completing degree at Brac University.
2. The thesis does not contain material previously published or written by a third party, except where this is appropriately cited through full and accurate referencing.
3. The thesis does not contain material which has been accepted, or submitted, for any other degree or diploma at a university or other institution.
4. We have acknowledged all main sources of help.

Student's Full Name & Signature:

Abrar Shahrier Arnab
(21201451)

S M Mahidur Rahman
(24241319)

Umma Souda Alfa
(21201008)

Yeasin Fiyaz
(21201087)

Approval

The thesis/project titled "Multimodal Deep Learning for Predicting Mechanical Ventilation Duration from Chest X-ray and Clinical Data" submitted by

1. Abrar Shahrier Arnab (21201451)
2. S M Mahidur Rahman (24241319)
3. Umma Souda Alfa (21201008)
4. Yeasin Fiyaz (21201087)

of Spring 2025 has been accepted as satisfactory in partial fulfillment of the requirement for the degree of B.Sc. in Computer Science in Spring 2026.

Examining Committee:

Supervisor:
(Member)

Dr.Swakkhar Shatabda

Professor
Department of Computer Science and Engineering
Brac University

Thesis Coordinator:
(Member)

Mr.Rafeed Rahman

Senior Lecturer
Department of Computer Science and Engineering
Brac University

Head of Department:
(Chair)

Dr.Sadia Hamid Kazi

Chairperson and Associate Professor
Department of Computer Science and Engineering
Brac University

Abstract

In critical care facilities, invasive mechanical ventilation is now a key criterion for determining the severity of illness caused by the COVID-19 pandemic. Both clinical decision processes and the efficient use of critical care facilities can be improved by using the predicted duration of invasive ventilation. In order to predict the invasive ventilator days, the study proposes the use of a multimodal learning architecture that utilizes both clinical data and chest X-ray images.The Cancer Imaging Archive (TCIA) dataset included chest X-ray images and clinical data for patients. A total of 213 full instances were retained after rigorous data cleansing and mapping using patient identifiers. Z-score normalization was used to normalize the clinical data, and intensity normalization and scaling were used to normalize the chest X-ray DICOM image data. The study used a continuous regression model to predict the ventilation day outcome.The study used a variety of deep learning and machine learning models, including ResNet-18, DenseNet-121, Random Forest, Linear Regression, and Gradient Boosting. A multimodal attention-based fusion model was used to combine the image and clinical data. The study used MAE, RMSE, R^2, Pearson correlation, and Spearman correlation to evaluate the model's performance.The results of the experiments show that multimodal deep-learning models have moderate performance, and the ensemble-based clinical models perform better than the linear ones. Besides, the results highlight the significant role of clinical data, and image data provides marginal performance improvements.

Keywords: COVID-19, machine learning, Multimodal , Regression analysis, DenseNet, ResNet

Table of Contents

List of Figures

List of Tables

Chapter 1

Introduction

COVID-19 was one of the deadliest viruses with a high fatality rate. Across the world 704,753,890 people are reported to have COVID-19, according to Worldometer. Physicians and other healthcare professionals frequently struggle to administer treatments because there are no data-driven tools to identify and predict risk. The fatality rate may rise sharply as a result. Identifying and predicting the patients who need ventilation support and the duration of ventilation support is our main goal.

1.1 Background

The pandemic caused by the COVID-19 virus highlighted the problems associated with the treatment of serious respiratory infections. The epidemic is a significant barrier to the global healthcare system. A significant complication of SARS-CoV-2 infection is the occurrence of acute respiratory failure. While mechanical ventilation is an important therapeutic strategy, there are significant individual variations between patients in the need for the ventilator.From a therapeutic point of view, it is imperative to forecast the duration of mechanical ventilation. Prolonged use of mechanical ventilation has a strong association with ventilator-associated pneumonia, lung injury, muscle weakness, prolonged intensive care unit stay, and increased healthcare costs. On the other hand, weaning patients from mechanical ventilation can lead to a high rate of respiratory relapse, causing increased mortality rates due to re-intubation.The early prediction of the duration of breathing can help in providing better care for patients.However, it is still difficult to predict the period of mechanical ventilation because of the complexity of the COVID-19 infection.
A number of factors, including the extent of lung involvement, intensity of inflammatory responses, presence of any comorbid conditions, and physiological changes, affect the progression of COVID-19 infection. In view of the poor predictive accuracy of clinical scoring systems, it is difficult to predict these factors using statistical models.During the epidemic, the use of chest X-ray imaging has largely been applied in the frontline diagnosis and monitoring of the disease due to its accessibility, cost-effectiveness, and short imaging time. The radiographic features that provide vital information regarding the severity of the disease are the presence of bilateral opacities, consolidation, and the presence of lung involvement. Additional vital information regarding the physiological state of the body can be obtained from the vital signs, oxygen requirements, tests, and the patient's history. The analysis of

the information from each imaging modality in isolation is limited.
However, the recent developments in deep learning (DL) techniques have shown promising results in deriving high-level representations from medical images, while machine learning techniques have already proven their efficiency in analyzing structured clinical data. Nevertheless, using a single source of information is not adequate for achieving the desired performance; besides that, the complex nature of critical illness is not reflected by using a single source. Multimodal learning is proposed to achieve better results.
Multimodal deep learning can uncover hidden correlations between radiological features and clinical indicators that are hard to spot using conventional analytics. By combining the features of chest X-rays and the clinical profile of the patient, the model can better capture the complexity of the disease. This is especially important in the prediction of the duration of the ventilator since the damage is occurring simultaneously.
There is a scarcity of research on the application of AI-based CDSs to make predictions on the duration of mechanical ventilation using multimodal deep learning techniques. Most existing research is biased toward mortality prediction or disease classification, while the duration of mechanical ventilation is still an understudied topic. In addition, the lack of complete patient information, class imbalance, and the heterogeneous presentation of chest radiographs pose difficulties in developing the model.
Keeping the above limitations in mind, the main contribution of the proposed research is the development of a multimodal deep learning model that uses chest X-ray imaging and clinical information to make a prediction on the duration of mechanical ventilation among COVID-19 patients. The proposed model can provide a better representation of the duration of mechanical ventilation among patients.
The proposed model can be useful in the advancement of intelligent healthcare systems. The proposed model can be applied to other forms of acute respiratory failure, not just COVID-19.

1.2 Rational of the Study or Motivation

The COVID-19 pandemic had a huge impact on public health sectors. Mechanical ventilation is a life-saving part that is used in intensive care units (ICUs) for saving the life of a patient. It is used in severe respiratory failure, lung diseases, pneumonia, and other critical conditions. Despite mechanical ventilation being important for saving lives, prolonged dependence on ventilatory support is associated with serious complications such as ventilator-associated pneumonia, lung injury, muscle weakness, increased mortality risk, and extended hospital stays. So predicting the mechanical ventilation is our main purpose.
In the current situation, the duration of mechanical ventilation days is widely dependent on physicians' and doctors' experience, intermittent clinical measurements, and radiographic interpretations. This may be subjective and may vary significantly across physicians and doctors. Early intervention of ventilation support days could significantly support clinical decision-making, improve ICU resource allocation, ventilation support, and treatment planning, and help the patient and doctors to identify high-risk patients who may require long vent days.

For patients on ventilation support, chest X-ray imaging is frequently carried out and gives useful perceptions into lung pathology, disease progression, and treatment plans. Simultaneously, clinical variables, including vital signs, laboratory results, and demographic factors, encompass additional physiological information. The maximum prediction models use only image or clinical data but don't use both. As a result, they were unable to find the full potential from their prediction model. Single-modality approaches frequently fail to achieve good performance in ICU or vent days outcome prediction tasks but recent activity has shown remarkable success in medical image analysis. To learn more precisely, multimodal deep learning produces a promising solution. Furthermore, attention-based fusion mechanisms allow the model to dynamically prioritize the most relevant data from each modality, potentially improving prediction accuracy.
Another major limitation of many models is interpretability. Black-box can make predictions without any explanation. This creates a lack of trust and has huge real-world adoption. Explainable artificial intelligence (XAI) techniques such as Grad-CAM for imaging data and feature-level analysis can provide visual and quantitative explanations, helping physicians understand why a patient is predicted to require prolonged ventilation and ICU support.
According to [17], of 29,141 cancer patients, 7791 patients were sent for a COVID-19 test. Among them, 1359 (17%) were tested positive. Of them, 84% were hospitalized, and 217 patients were taken to the ICU. The fatality rate is about 33%. According to [1], severe complications that have been reported to occur in 33% of patients with COVID-19 include acute respiratory distress syndrome, acute renal failure, acute respiratory injury, septic shock, and severe pneumonia. Many of them needed to be admitted to the ICU or receive ventilation support. Due to a lack of ICU facilities and ventilation, many patients died due to not getting proper treatments.
So, after getting motivated by these challenge, this study proposes an explainable deep learning framework that integrates the X-ray images and clinical data to predict the duration of mechanical ventilation. The suggested method seeks to produce precise, comprehensible, and clinically significant predictions by fusing convolutional neural networks (CNN) for image feature extraction, attention-based multimodal fusion, and explainability techniques. Early risk assessment, improved clinical decision-making, and more individualized and effective critical care management are all possible with such a system.

1.3 Problem Statement

Mechanical ventilation is one of the most basic life-support services provided in the intensive care unit for patients who have severe respiratory problems. While the need for assisted ventilation may be necessary for sufficient gas exchange, the number of days a patient needs assisted breathing remains a very uncertain process. The number of ventilation days can vary considerably among patients.
Prolonged ventilatory support leads to unfavorable clinical outcomes, such as ventilator-associated infections, lung injury, poor recovery, prolonged ICU stay, and increased healthcare load. On the other hand, early discontinuation of ventilatory support can cause respiratory failure and subsequent intubations. Prediction of ventilatory support duration is therefore vital for ensuring better patient care.

Current methods for determining the number of days it will take to ventilate patients involve an amalgamation of subjectivity, difficulty in replication, and a degree of nonlinearity. Such methods rely on the judgment of doctors and the interpretation of clinical lab values that yield unreliable information and the observation of the patient's chest X-ray. The weaknesses associated with standard statistical modeling and rule-based scoring methods, such as poor predictability and inadequate estimates, also apply.

Despite the fact that chest X-ray images contain valuable information regarding pulmonary involvement, imaging information cannot completely represent systemic physiological instability. Clinical factors also do not provide visualization of pulmonary pathology directly. Previous models based only on a single modality cannot make use of complementary information and therefore provide less accurate results. Furthermore, the approaches to the problem using deep learning methods tend to be black-box solutions, providing little insight into the solution being provided. Lack of interpretability means that the use of these approaches in an ICU setting could be rendered impossible.

Consequently, it can be deduced that the need for a reliable, accurate, and interpretable prediction model that can predict the duration of mechanical ventilation through the analysis of chest X-ray images, along with clinical data, is a prerequisite. This constitutes the main aim of the current investigation.

1.4 Objectives

The study's goals are to create and assess a data-driven multimodal prediction model that uses clinical information and chest radiographs to forecast how long hospitalized COVID-19-positive patients would need invasive mechanical ventilation. The proposed study's goals are:

- Instead of making categorical forecasts, the duration of mechanical ventilation should be precisely predicted in terms of the number of days by formulating the prediction as a regression issue.
- To create and apply machine learning models that use clinical data, including vital signs, test results, comorbid conditions, and demographic information, to forecast the length of mechanical ventilation.
- To create deep learning-based feature extractor models using chest X-ray pictures and assess the models' capacity to identify patterns linked to the deterioration of respiratory diseases.
- To assess the effectiveness of the suggested multimodal prediction model, which uses the multimodal learning framework to combine clinical and imaging data, and to compare the multimodal model's performance with that of the unimodal models, which use clinical and imaging data independently to predict the length of mechanical ventilation.
- Using a variety of regression performance metrics, including MAE, RMSE, R^2, Pearson correlation, and Spearman correlation, compare the performance of the conventional machine learning model and the suggested deep learning-based multimodal model.

- Using a variety of explainable AI techniques, such as Grad-CAM for imaging data and feature importance for clinical data, to enhance the model's interpretability and boost the clinical community's confidence in the suggested prediction model.

- To assess the suggested prediction model's viability in an intensive care unit setting.

1.5 Methodology in Brief

To attain the goals of this work, we propose a multimodal, AI-based research methodology that unites both medical data (structured patient data) and medical imaging data (chest X-ray images). This dual approach makes it possible to have a broader and more in-depth study of finding the duration of ventilation support days because both quantitative and visual evidence of the disease spread are attained. Overall workflow includes the following key steps:

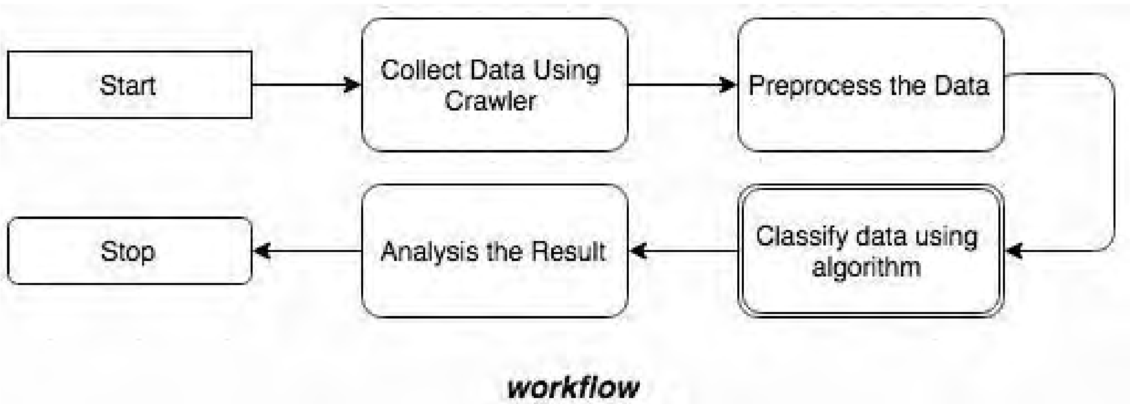

Figure 1.1: Workflow

- **Data Collection:** Public datasets de-identified were downloaded from reputable medical databases, for instance, The Cancer Imaging Archive (TCIA). Every dataset contains paired chest X-ray images alongside corresponding clinical data (comorbidities, therapy, age), which share a common patient ID that associates with the former. This ensures consistency, ethical use, reproducibility.

- **Data Pre-processing :** Clinical information was pre-processed by removing missing or inconsistent information, followed by encoding categorical variables into binary or numerical variables. Imaging information was standardized through resizing, normalization, and format change (from DICOM to JPEG) to ensure its compatibility with deep-learning models.

- **Feature Extraction and Modeling:**
  - **Clinical Informatio::** Labelled using standard machine-learning methods such as Support Vector Machine (SVM), K-Nearest Neighbors (KNN), linear regression, random forest, and Gradient Boosting.

  - **Imaging Data:**Handed over deep learning networks like DenseNet and ResNet for high-level computer-automated extraction from the radiography for regression analysis risk and pulmonary disease. Also, CLIP embedding was done using MedCLIP and a multimodal LLM with MedGemma.

- **Multimodal Integratio::** Features thus obtained from the imaging and clinical models are subsequently combined at a terminal level with a method that entails concatenating the two feature vectors prior to subjecting them to a meta-classifier for simultaneous analysis. For this particular situation or scenario, we considered MedGemma as the first option. It supports reasoning, summary generation, reports, and responses. This lets the algorithm access both radiological and physiological information, which is predictive even if life is simple.

- **Evaluation and Interpretation:** With respect to all these models, we use standard quantitative metrics like MAE, RMSE, R^2 , PCC and SCC etc. Owing to feature importance analysis tools like Grad-CAM and others, we can even understand the contribution of varying attributes or image regions to the final prediction. The methodical approach that combines the efficacy of multiclass AI with clinical expertise not only helps in the prediction of finding ventilation days but also serves as an essential link in the chain that had been formed by the research studies that preceded this work.

1.6 Scopes and Challenges

In our study, there are some constraints and challenges. There are also some scopes in our study.

- **Scopes:**
  - **Advanced Machine Learning Use:** Several techniques from deep learning and machine learning models such as Random Forest, Support Vector Machines (SVM), Neural Networks, and K-Nearest Neighbors (KNN) have been used, each of them tailored to suit medical data.
  - **Multimodal Data Integration:** Investigating the creation of comprehensive, multidimensional predictive models through the integration of clinical, imaging, laboratory, and demographic data is termed multimodal data fusion using longitudinal modeling to generate a comprehensive diagnosis for ventilated patients.
  - **Real-world Applicability:** Spanning the gaps between model-theoretic development and clinical practice implementation by developing machine learning models that are clinically interpretable and deployable in actual healthcare settings.

- **Challenges:**
  - **Data Set Limitations:** Some of the major issues are the completeness, availability, and quality of high-quality data sets. Some of the major issues that can limit the generalizability of the method are retrospective biases, missing data, and geographically limited data sets.

- **Interpretability of the Model**: Clinical interpretation of machine learning techniques, particularly deep learning techniques, is becoming increasingly complex because they are sophisticated. In a bid to enable clinical practitioners to be able to depend on and respond accordingly to prediction data, clinically transparent and interpretable models must be developed.
- **Challenges of Real-Time Deployment:** Infrastructure, integration with current clinical practice, confidentiality of the information, and consistent and reliable operation independent of changing healthcare conditions are the main challenges to the real-time deployment of ML models in actual clinical practice.

Chapter 2

Literature Review

The severity of COVID-19 in various patient types has been the subject of numerous research studies. Almost every person has similar comorbidities, like fever, cough, hypertension, and diabetes. Scientists and researchers used different ML models to predict that these patients are in LCT (long-term COVID-19). We found different study gaps from different studies, such as data limitation, less use of ML and DL methods, etc. We aim to work on these gaps to reach our desired objectives and goals.

The literature on COVID-19, post-COVID conditions, and their consequences for patients is extensively reviewed in Chapter 2, with special emphasis on the application of AI and machine learning for clinical risk prediction. The first section of this chapter discusses the long-term health problems caused by COVID-19. It also highlights the weakness of traditional statistical approaches in dealing with complex multimodal medical data. To predict the severity of COVID-19, ICU admission, mortality, pulmonary problems, and long-term health consequences, it examines a broad spectrum of research on deep learning and machine learning techniques, including ensemble methods, logistic regression, random forests, support vector machines, and convolutional neural networks. It also focuses on studies that apply explainable AI, multimodal data fusion, and image-based approaches. Some of the most important data on symptom patterns, comorbidities, and post-COVID interstitial lung disease are also discussed in this chapter. Some of the important gaps in the literature that have also been discussed in this chapter include real-time prediction issues, data set limitations, lack of longitudinal modeling, inadequate multimodal data fusion, data imbalance, and interpretability issues. The proposed multimodal and interpretable AI framework developed in this thesis is highly motivated by these issues.

2.1 Preliminaries

Apart from the clinical emergency that resulted from the COVID-19 outbreak, one of the biggest challenges that face modern healthcare systems today involves patients who require mechanical ventilation support. This support has been identified as a life-sustaining intervention for patients who experience acute respiratory failure. The duration of time for which patients require mechanical ventilation support has been identified as varying significantly and plays an important role in the survival of patients who require this support. The duration of time for which patients require mechanical ventilation support has been identified as technologically significant and clinically significant. For example, the prolonged duration for which patients require mechanical support has been identified as contributing to an increase in complications for patients who require this support. Similarly, premature extubation for patients who require mechanical support has been identified as leading to respiratory failure and has also been identified as leading to re-intubation for patients who may die as a result of this complication. Therefore, the accurate estimation of the expected duration of mechanical ventilation support has been identified as important for patients who require this support during their treatment. During the COVID-19 outbreak, which resulted in a respiratory illness for patients across the world, hospitals across the world were forced to deal with an unprecedented increase in patients who required mechanical support for survival. The clinical state of ventilated patients depends on the complex interaction of pulmonary involvement, physiological instability, laboratory biomarkers, comorbid conditions, and patient demographics. The chest X-ray image provides crucial insights into lung pathology, such as consolidation, opacities, and disease progression, while clinical variables provide insights into the patient's systemic condition and health status. When each of the data sources is considered individually, the predictive potential of the model is limited. When combined, the data provide an overall view of the severity of the patient's respiratory failure.

Traditional statistical techniques were initially used to model the relationship between the data sources and the patient condition but were found to be ineffective, especially considering the linearity and feature engineering constraints of the traditional approach. The traditional approach becomes even more limited when applied to the clinical environment, where the data are heterogeneous, incomplete, and constantly changing. As such, the conventional approach to predictive modeling has been found to be inadequate for the generalization of patient populations and environments.

However, a promising method that can efficiently learn from large-scale data sets, even with non-linear correlations, can be achieved with the recent developments in machine learning (ML) and deep learning (DL) methods. Discriminative characteristics from large-scale chest X-ray images have been efficiently learned using deep learning methods, such as convolutional neural networks. Machine learning methods have also shown encouraging results in learning organized clinical data to recognize physiological patterns that are unique to each patient.

Recently, multimodal deep learning techniques have shown promising results in integrating both imaging and clinical information into a single predictive framework. This can be considered a major improvement over conventional approaches to using single modalities of information, where the interaction between modalities can be difficult to understand and interpret. This is particularly true when estimating the length of breathing because the patient's health as well as the level of lung damage might vary at the same time.

Even though the prediction of mechanical ventilation duration is important, it has not been given much attention in comparison to other important outcomes, such as mortality rates and disease classification. The literature that is currently being used is mostly based on single modality data, which is mostly based on a few features. However, a significant gap exists in the standardization of these methods, which are mostly based on single modalities of information.

This study proposes a unique deep learning method for predicting the duration of mechanical ventilation using multimodal data, such as chest X-rays and other data. This is anticipated to increase forecast accuracy, and as a result, the outcomes are anticipated to make a substantial contribution to the field of pandemic management, which is currently of utmost concern. The proposed method is also expected to contribute significantly to the domain of intelligent respiratory care management. The prediction of mechanical ventilation duration is important since it is expected to improve critical care planning, reduce the adverse effects of mechanical ventilation, and hence improve patient outcomes.

2.2 Review of Existing Research

The study titled "A convolutional attention mapping deep neural network for classification and localization of cardiomegaly on chest X-rays" proposes an automated system that utilizes a weakly supervised learning method for the detection of cardiomegaly from chest X-ray images and the localization of the region of the heart that is affected. In this study, the proposed system is implemented using a weakly supervised learning method because the process of diagnosing the disease from the X-ray image is time-consuming and also because it is difficult to obtain bounding box and segmentation annotations. In this study, the proposed system is implemented using the NIH ChestXRay14 dataset, which contains more than 100,000 images from 14 different diseases, including cardiomegaly. In this study, the proposed system is implemented using convolutional neural networks such as DenseNet-121, DenseNet-201, ResNet-50, and Inception-ResNet-V2. Among these networks, DenseNet-121 performed better than the other networks. In this study, the novelty of the proposed system is the Attention Mapping Mechanism (AMM) that assigns different weights to the feature maps instead of the use of all the features equally as in the conventional Global Average Pooling (GAP) method. In this study, the results obtained from the experiments were compared using different metrics such as precision, recall, F1-score, and AUC. From the results obtained from the experiments, it is clear that the proposed system using DenseNet-121 with the Attention Mapping Mechanism CXA_Dense121 performed better than the other approaches that use the GAP method in all the metrics. In this study, the proposed system obtained a result of 0.87 for precision, 0.85 for recall, 0.86 for F1-score, and 0.89 for AUC. In this study, the attention heatmaps also clearly highlighted the region of the heart.[15]

The research entitled "Post-COVID-19 Condition Prediction in Hospitalized Cancer Patients: A Machine Learning-Based Approach" was executed by Mohammadi et al.[23] and the Sechenov Stop COVID Research Team, and it appeared in the journal. The purpose of this research is to forecast how probable post-COVID infection complications are to develop patients. This means that these researchers attempted to apply their understanding of various machine learning classifiers to discern these critical factors that forecast complications from COVID infection. They used logistic regression, random forest, SVM, KNN, and neural networks to forecast PCC in patients. Data were gathered during the acute period of the disease and subsequently followed up at 6 and 12 months post-discharge. The model's performance was assessed using sensitivity, specificity, and the area under the receiver operating characteristic curve (AUC). KNN demonstrated the best predictive performance, with an AUC of 0.80, sensitivity of 0.73, and specificity of 0.69.[23]. Pre-existing comorbidities and severe COVID-19 were significant predictors of post-acute sequelae of SARS-CoV-2 infection (PCC). The long-term consequences of COVID-19 on patients are better understood because of this study.

Post-COVID-19 interstitial lung disease is a considerable clinical issue, as a substantial proportion of survivors exhibit enduring fibrotic alterations in lung tissue. A study titled "Post-COVID-19 interstitial lung disease: Insights from a machine learning radiographic model" by Karampitsakos et al. (2023) [16]. Machine learning-based radiography provides valuable insights into post-COVID interstitial lung disease (ILD). The focus of this study is to increase diagnostic accuracy, as well as guide specific treatment protocols, by quantifying various radiographic as well as functional impairments in post-COVID-19 interstitial lung disease. The study employed a quantitative technique, analyzing high-resolution computerized tomography (HRCT) images of 232 participants through Imbio Lung Texture Analysis, version 2.1, from five ILD referral centers in Greece, specifically three months post-infection, as measured through lung function testing as well as HRCT images analyzed by the Imbio machine learning model. The study also emphasizes the importance of screening for potential long-term pulmonary issues in post-COVID-19 cases.

Wells, Devaraj, and Desai (2021) [32] discuss the uncertainties about interstitial lung disease, which may appear as a result of the COVID-19 infection. They identify some things about it that we currently don't understand in terms of the healthy lungs. These researchers examine the long-lasting health impact on an individual's lungs as a result of living through a serious case of COVID-19. They highlight the case of interstitial lung disease and how it presents in the CT scan performed on the patients. They examine it further at the six-month mark to understand what remains as a residual impact. At the six-month mark, 62% showed some abnormality in the CT scan. Meanwhile, 35% manifested some kind of fibrotic changes in the lungs.. The research indicates that mechanical ventilation and ARDS are prominent predictors of lifelong pulmonary impairment.

Cordelli et al. (2024) [10]performed a study on pulmonary Long Covid sequelae utilizing machine learning methodologies to forecast long-term pulmonary difficulties in COVID-19 patients based on clinical data titled "Machine learning predicts pulmonary Long Covid sequelae using clinical data." The study included 152 COVID-19 patients from three hospitals in Northern Italy (2020–2021), collecting clinical, laboratory, and demographic data during hospitalization, and labeling patients based on the presence or absence of pulmonary sequelae at CT scan 12 months post-discharge. The research utilized shallow learners (kNN, SVM, and decision tree), ensemble methods (random forest, AdaBoost, stacking, and majority voting), and a multimodal classifier selection that integrates classifiers trained on various data modalities (e.g., demographics, hospitalization data, ventilation) with a meta-classifier (e.g., SVM). The study demonstrated multiple limitations, such as the exclusive focus on hospitalized patients (excluding mild or home-treated cases), the absence of data regarding vaccination status and previous infections, a limited sample size (n=152), and the omission of radiological imaging due to exposure risks and lack of standardized follow-up. This indicates the necessity for future research utilizing larger, more diverse datasets and the investigation of deep learning methodologies.

The editorial titled "COVID-19, cancer post-pandemic risk, and the radiation oncology physicist" by Mary Beth Allen (2022) [3] addresses how the COVID-19 pandemic has affected cancer care, especially how it has changed screening, treatment, long-term outcomes, and the role of medical physicists in a changed healthcare system. The paper states that fresh information about extended COVID continues to come out and is not yet clear, especially when it comes to its link to cancer progression or the chance of getting cancer again. This editorial contributes to the discussion by defining COVID-19 as a long-term risk factor in oncology, talking about how fewer people are getting screened for cancer (for example, breast screening dropped by 90.8%), problems from extended COVID, and the necessity for treatment regimens to keep becoming better.

David J. Pinato et al. (2021) [26] conducted a study titled "Prevalence and impact of COVID-19 sequelae on treatment and survival of patients with cancer who recovered from SARS-CoV-2 infection." The purpose of the study was to assess the prevalence of post-COVID-19 sequelae in cancer patients. It is a study based on data from the OnCovid registry, NCT 04393974, a large European effort that amasses real-world clinical data on cancer patients diagnosed with COVID-19. It found that higher mortality was independently associated with these outcomes. The study has several significant limitations, including its retrospective design, reliance on symptom-based reporting of sequelae without imaging or laboratory confirmation, and relatively short follow-up, median of approximately 4 months. The authors require more prospective studies to understand PCC and how to take care of cancer patients in this setting.

Monroy-Iglesias et al. (2022) [24] performed "Long-term effects of COVID-19 in Cancer Patients — The Experience from Guy's Cancer Centre" and was a study that was observational in nature. The main aim of the study was, in essence, evaluating the prevalence of long-term symptoms of COVID-19 in Cancer Patients, understanding the descriptions of long-term symptoms of COVID-19 in Cancer Patients, as well as understanding a potential link between long-term symptoms of COVID-19 in Cancer Patients and demographic factors, Cancer, and Cancer Treatment Strategy, in essence, understanding how long-term symptoms of COVID-19 uniquely affect Cancer Patients. The study integrated a dataset from an internally ethically approved Guy's Cancer Cohort (reference number: 18/NW/0297). Clinical and demographic data were matched to patient data. The percentage of participants who reported having long COVID-related symptoms in this research was 51.3%, while fatigue (78%) was the most common, followed by breathlessness (53.7%), cognitive impairment (46.3%), sleep disturbances, inability to taste food, sadness etc's [24]. The research reported in this article appears to be the first examining long COVID in cancer patients. Those who receive neoadjuvant treatment, palliative treatment, require oxygen treatment, or were hospitalized during their COVID illness have higher rates of long COVID. The research has significant implications: It suggests that over half of cancer patients who have an episode of COVID-19 may be impaired in ways that will limit successful rehabilitation as well as negatively impact their quality of life. The research underscores the pressing need to design better COVID rehabilitation approaches tailored to cancer treatment. Estiri et al. (2021) conducted a study in a European multicenter patient data and developed an ML model that can predict COVID-19 mortality about 60 days ahead. [20].

Sethi et al. (2024) [29] published a study and got some key findings. In their research, they employed ML methods to distinguish the severity of the patients. They performed logistic regression, XGBoost, Naive Bayes, and SVM. Among them, the

XG model was the superior, with an accuracy of 95% and an AUC of 0.99. A research paper named "Machine Learning-Based Prediction of COVID-19 Severity and Progression to Critical Illness Using CT Imaging and Clinical Data" by Purkayastha et al. (2024) [27] conducted their work using radiomics data to predict the severity of COVID-19. They performed the test with a population of 981 patients. Of them, 274 patients developed severe disease. The strongest performance was achieved by radiomics features and clinical data, with an ROC AUC of 0.76 compared to 0.70 for visual CT variables score and clinical variables. The research entitled "Prediction of COVID-19 Mortality in Patients with Cancer" by Mantha et al. (2020) [21] looked to create a machine learning predictive model to assess the risk of mortality within 45 days following a COVID-19 diagnosis in cancer patients, who are particularly vulnerable due to immunosuppression and comorbidities. The dataset comprised 820 COVID-19-positive cancer patients from Memorial Sloan Kettering Cancer Center, with clinical, laboratory, and demographic information gathered at diagnosis, and records were meticulously scrutinized to guarantee precise inclusion of essential predictors.[21] The researchers employed Random Survival Forests (RSF), an ensemble learning technique for time-to-event data, utilizing R packages randomForestSRC, ROCR, and cvAUC, with model efficacy assessed via repeated 10-fold cross-validation. While its average AUC value of 0.81 makes it part of the first machine learning models for COVID-19 positive individuals with cancer diagnoses, the study reflects the influence various general factors like age, health conditions before the illness, and even hypoxia may contribute to the risk of death, as well as factors specific to the individuals with active or past cancers like the types of cancer or metastasis development. Some of the study limitations include the need for external verification before a clinical use may be made for the results and the low representation of some lab values like the inclusion of D-dimers within only 7% of the data. "Long COVID in cancer patients: predominance of symptoms in the majority of patients over an extended duration" study conducted by Dagher et al. (2023)[11]. The study focused on understanding the duration of symptoms among cancer patients. In an attempt to monitor the long-term effects of COVID for up to 14 months, researchers used electronic patient-reported outcome surveys and examined electronic medical records for remote symptom tracking. Overall, 312 provided longitudinal data; 188 cancer patients showed post-COVID symptoms. Longitudinal clinical, therapeutic, demographic, and symptom-duration information was collected for up to 14 months. The results indicate that, among those who did so, 60% continued with symptoms of fatigue, sleep disturbances, and gastrointestinal problems, offering new, more granular insights into long COVID in cancer patients. However, there are notable limitations to the study: it is single-center, which limits how broadly the results apply; its retrospective nature makes it hard to draw causal conclusions from it; it relied on subjective reports, without standardized quality-of-life measures; and there is a risk that symptoms from cancer itself, or its treatment, might confound the results because of the possible overlapping with COVID-related symptoms. The COVID-19 pandemic has had a major impact on healthcare delivery, particularly in the area of cancer treatment, as observed by Allen (2022)[3]. The work presents a range of public health information, emphasizing discrepancies in healthcare and a reduction in cancer screening. It focuses on the fact that COVID-19 has not gone away, impacting cancer care, and the need for urgent resumption of pre-COVID cancer screening and preventive measures. Manea and colleagues used clinical analysis, radiological data, and demographic information in a 2025 study to evaluate the severity of COVID-19 in those suffering from cancer. Their investigation included 137 individuals who had various kinds of cancerous tumors. Manea and colleagues evaluated severity through oxygen level. Their models included Naive Bayes, KNN, SVM, and decisions trees. It was highlighted that the top-scoring models were those of KNN, which were evaluated to 100% and 98.3%, and the models managed to attain an exceptionally high AUC. [9] Navlakha et all.(2021) [25] relied on effectively developed machine learning models to assess how serious COVID-19 might become in cancer patients. This is a reflection of how effectively machine learning models can be applied to effectively foresee and forecast illnesses. To this effect, this particular study has relied on machine learning models to effectively classify patients depending on the clinical characteristics. More so, it has relied on a random forest classifier to foresee the severity of COVID-19 infection in cancer patients. To this effect, there have been 267 clinical parameters applied in this particular case. Though there is a possibility of applying time-series data and clinician input to effectively apply this model in real-time, this has been a limiting factor within this particular instance of a very small dataset. This aspect also calls for better algorithms in effectively forecasting the outcomes beyond the early stages. The research conducted by Aktar et al. (2021) [2] uses machine learning techniques for discovering comorbidities and symptoms that elevate the mortality risk in COVID-19 patients. The authors underscore the viewpoint that chronic health concerns like diabetes, heart disease, and chronic obstructive pulmonary disease have a significant impact in deciding the outcome in the case of the pandemic (Aktar et al., 2021). The authors have made use of machine learning in conducting a meta-analysis on the data that could shed some light in deciding the important comorbid conditions along with their respective symptoms responsible for death caused by the COVID-19 virus. The data that has been collectively analyzed shows that the total number of infected members with COVID-19 across 141 countries is 481,289. It has been underscored that some of the limitations were the potential threats that could be faced based on the data that could be freely obtained, along with the threat of not knowing the severity of the comorbid conditions. Furthermore, future research in this direction could be made with the use of data that has been collectively obtained along with the use of cross-validation in obtaining better accuracy in meeting the research goals. Diabetes, hypertension, along with heart conditions, could be some of the significant comorbid conditions that could be responsible in the case of death caused by the COVID-19 virus. The study was designed with the aim of comparing how cancer patients and healthy individuals would react in terms of developing antibodies following three COVID-19 vaccine brands: BNT162b2, AZ1222, and Coronavac. They considered cancer patients as a demographic that would react negatively to various diseases and possibly vaccine medications based on their condition and/or medication. Therefore, they wondered what vaccine would yield a better response from cancer patients. For them to understand and identify critical factors—such as how well a vaccine works—considering factors like how well a vaccine works with respect to breakthrough infection risk, they analyzed a dataset with 246 healthy individuals and 389 cancer patients, including factors like IgG antibody results in a lab, possible data from a CT scan or X-ray configured on a scale to judge disease severity. In both patients and healthy individuals, Pfizer-BioNTech COVID-19, referred to as BNT162b2, showed an increase in IgG Antibodies compared to CoronaVac

produced by Sinovac. CoronaVac had 7.34 times more breakthrough infections in healthy individuals compared to patients receiving BNT162b2. Gender did not have an effect on post-vaccination side effects observed in patients and healthy individuals. This study has provided new perspectives on understanding the participation of the immune system in cancer patients receiving various COVID-19 vaccines and CoronaVac produced by Sinovac. Cancer patients should be provided with effective and appropriate vaccines to ensure maximum protection from COVID-19 in the future. Responses were identified according to sex.
The study, titled "Identification of Risk Factors of Long COVID and Predictive Modeling in the RECOVER EHR Cohorts," was authored by Chengxi Zang, Yu Hou, Edward J. Schenck, Zhenxing Xu, Yongkang Zhang, and others from Weill Cornell Medicine and the University of Florida (Zang et al., 2024) [33].The project aimed to discover parameters linked to the onset of long COVID (post-acute sequelae of SARS-CoV-2 infection, PASC) and to create predictive models utilizing electronic health records (EHR) from extensive clinical research networks. The research employed machine learning approaches, such as regularized Cox proportional hazards models, gradient boosting machines, and deep neural networks. The research utilized two extensive EHR datasets from the INSIGHT and OneFlorida+ clinical research networks, comprising de-identified health information from millions of patients in New York and Florida. The predictive models demonstrated high accuracy for illnesses such as dementia, COPD, and heart failure (C-index ¿ 0.8), intermediate performance for conditions including pulmonary fibrosis and diabetes (C-index 0.7–0.8), and diminished predictive capability for conditions including fatigue and depression (C-index ¡ 0.7). The AUROC results validated the diverse prediction efficacy under varying settings.
Kwon et al. (2022) [18] explored COVID-19 outcomes in cancer patients utilizing a substantial cohort from the University of California Health system database. This retrospective cohort study mainly tried to assess the impact of cancer-related variables on the severity of COVID-19 outcomes. Since this study aimed at identifying cancer-related risk factors that might affect COVID-19 outcomes within 30 days of infection, the paper managed to collect data from a pool of patients within the University of California Health system's electronic health records of patients within the last 17 hospitals. Eventually, this pool contained over 409,000 patients who were subjected to testing for SARS-CoV-2 infection in 2020. Among this pool of patients, there were 1,781 patients who tested positive for COVID-19. Furthermore, this pool contained patients who were diagnosed with cancer, which equalled 49,918 patients. A multivariable logistic regression model was employed to examine data concerning cancer types and treatments. Patients with BCR/ABL-negative myeloproliferative neoplasms exhibited a hospitalization risk exceeding twice. Individuals on immunosuppressive medicines (e.g., venetoclax and methotrexate) have almost a threefold increased risk of hospitalization due to COVID-19. Cancer patients had a lower likelihood of testing positive for COVID-19 in comparison to non-cancer patients. The machine learning study facilitates a comprehensive knowledge of the relationship between cancer therapy and COVID-19 severity, emphasizing at-risk cancer populations, such as individuals with myeloproliferative neoplasms or those undergoing specific treatments such as methotrexate and venetoclax.
Lin and Fang (2025)[**?**] investigated transfer learning approaches with chest X-rays in estimating ICU admission for COVID-19 patients.This solves the urgent problem

of early triage with COVID-19 by estimating ICU admitting patients from their initial chest X-rays so that better resource allocation may occur under low bed and resource availability. It uses two publicly held data sets: COVID-19-NY-SBU (1,384 patients; 899 AP/PA first-encounter CXRs after filtering: 269 ICU, 630 non-ICU, defined by ICU charges or in-hospital mortality) and The MIDRC-RICORD-1c dataset contains 417 chest X-rays of 361 patients. Every image was labeled based on the amount of lung involvement by pathology: if the opacities were more than four lung zones, it was labeled as an ICU case; otherwise, it was non-ICU. Pre-processing involved doing the images before training, where the images were first processed by a U-Net model to separate the lung area, then cropped to just that area. The images were enhanced with CLAHE to make the details clearer, resized to 224×224 pixels, and normalized according to the input specifications of the model. Since the number of ICU cases was less than that of non-ICU cases, the classes were balanced by generating more ICU samples by subjecting the existing ones to small random rotations (between 15° and +15°). This oversampling was applied only to the training set in order not to bias test or validation data.The models were evaluated with different versions of the DenseNet121 model with transfer learning: one trained for ImageNet (ImageN-SBU), one trained for chest X-rays (TorchX-SBU), and two extended models combining SBU and RICORD datasets, one with the ELIXR encoder (TorchX-SBU-RSNA and ELIXR-SBU-RSNA). They only adjusted the final block of convolution and added two fully connected layers with dropout and batch normalization. To improve explanation of predictions, they used Grad-CAM to show what lung areas the model utilized. Data were split with 5-fold cross-validation; for ImageN-SBU/TorchX-SBU, the train sets were 172 ICU and 403 non-ICU images, while the extended models were trained on 266 ICU and 642 non-ICU cases (RICORD images being reserved for training purposes only). The performances were uniform in their improvements: ImageN-SBU had a balanced accuracy of 61.7% (AUC 0.657), TorchX-SBU 65.1% (AUC 0.711), TorchX-SBU-RSNA 68.9% (AUC 0.772), and ELIXR-SBU-RSNA 69.8% (AUC 0.777), with specificity up to 80%. Overall, the domain-specific pretraining and dataset expansion models made improved early risk predictions with more reliable models, allowing physicians to prioritize COVID-19 patient monitoring and ICU readiness better even with limited data.

According to a research paper, Deep Learning–Based Time-to-Death Prediction Model for COVID-19 Patients Using Clinical Data and Chest Radiographs, Matsumoto et al. (2023)[22] developed a prognostic model integrating chest radiographs and clinical data to predict survival in COVID-19 patients.This research addressed the problem of predicting time-to-death among hospitalized COVID-19 patients to enable early risk stratification and resource allocation by clinicians.The researchers utilized the Stony Brook University COVID-19 dataset (TCIA), consisting of 1,356 patients with chest X-rays (AP view upon admission) and clinical data (demographics, comorbidities, vital signs, and laboratory tests). A single X-ray per patient closest to admission was contributed, and the data were separated into training (1,082 patients) and testing (274 patients) randomly, while using fivefold cross-validation on the training set. Images were rescaled (256×256, 320×320, 512×512) and augmented (rotation, shift, brightness, horizontal flip), and the clinical characteristics were age, gender, oxygen saturation, comorbidities, and blood markers. The model suggested was a DeepSurv with CNN structure, in which the CNN extracted fea-

tures from X-rays and combined them with clinical information through a Multi-Layer Perceptron for prediction of time-to-event. This approach outperformed both traditional and single-modal models: DeepSurv-CNN achieved c-index of 0.82 and Brier score of 0.20, which beat 0.77 for DeepSurv-only (clinical), 0.70 for CNNSurv (images alone), and 0.71 for Cox proportional hazards. The model also demonstrated stable time-dependent AUC (¿0.8 during the first week), and saliency maps confirmed infiltrated lung regions were driving predictions. Overall, the study solved the problem of integrating imaging and clinical data into a single framework to provide more accurate and interpretable prognosis, giving doctors a tool to prioritize care and plan ICU resources.
Lung shape differences on baseline CT scans have been shown to help predict COVID-19 severity with AI algorithms that utilize shape and infiltrates (Hiremath et al., 2024)[14].The aim of this study was to tackle the challenge of estimating early and clinically understandable predictors of COVID-19 severity from baseline chest CT scans since clinicians needed robust tools to predict which patients would develop severe disease and be exposed to long-term effects like fibrosis or long COVID.
The study used a multi-site retrospective data set of 3,230 COVID-19 patients with 213 healthy controls (NLST), with patients classified as healthy, mild (not ventilator), and severe (ventilator). Among these, D1 (Renmin Hospital, Wuhan) was split into training (416) and internal testing (419), and three external data sets (D2: 113, D3: 2000, D4: 282) and D5 (213 healthy controls) for validation. Human annotators annotated lung regions and infiltrates within the training dataset by expert radiologists, and a Difference Atlas (DA/DEA) was created to quantify CT-registered differences in lung shape by employing signed distance functions. Methodologically, CovSafeNet, an AI model with two 3D CNN streams - one addressing lung shape change analysis (M1) and the other addressing infiltrate analysis (M2), with joint output for severity prediction - was created. Automatic segmentation of lungs and infiltrates was done with a 3D U-Net, and explainability was integrated with Grad-CAM maps. Results showed that severe patients exhibited impressive lung shape deforms (especially at mediastinal and basal surfaces), decreased lung volumes and elevated CT intensities compared to slight ones, and CovSafeNet better than clinical-only and radiomics models with an internal AUC of 0.890 and external AUCs of 0.654–0.769. In summary, the research illustrated that incorporating lung morphology and infiltrates into an interpretable AI framework is able to present early severity prediction and long-term risk determination, allowing physicians equipment for triage, ventilator planning, and monitoring patients who are likely to experience future lung damage.
The research study called "Semi-supervised classification of disease prognosis using CR images with clinical data structured graph" published by Bai, Li, and Nabavi (2022)[6], which had proposed a semi-supervised graph neural network approach that integrates chest radiographs and clinical data to improve COVID-19 prognosis prediction.This research addressed the problem of proper prediction of ICU admission in COVID-19 patients by integrating chest radiographs (CR images) with clinical data, especially when labeled data is scarce. Researchers used the COVID-19-NY-SBU dataset (TCIA) containing 1,342 patients (258 ICU, 1,085 non-ICU) with CR images and clinical variables such as protein levels, respiratory rate, age, BMI, smoking status, and lab test. They proposed a Graph Markov Neural Network (CGMNN) model combining image features (learned with the help of an autoen-

coder) and clinical data in a graph framework, with each patient as a node and edges formed with KNN-based similarity. It used semi-supervised learning, where both labeled and unlabeled patients were both used in feature learning with the assistance of the Expectation-Maximization (EM) algorithm. 70% of patients were utilized for training and the other 30% was utilized as validation and testing sets. The evaluation revealed that the CGMNN model performed much better compared to single-modality models such as ResNet (images alone) or Random Forest (clinical alone), with accuracy of 0.82, a precision of 0.81, a sensitivity of 0.82, and F1 score of 0.76. Overall, this research resolved the issue of limited labeled data by demonstrating that semi-supervised, graph-based fusion of imaging and clinical features produces a more stable and more accurate prognosis prediction model to assist clinicians with ICU triage and patient treatment in the COVID-19 pandemic.

2.3 Summary of Key Findings

Particularly, among high-risk populations, an abundance of research has studied the implementation of machine learning ML and AI in COVID-19 prediction, diagnosis, and outcome analysis.
In the early age, the research mainly focused on diagnostic purposes by analyzing the radiographic image and clinical data. Over time, studies began to incorporate more complex tasks for predictions. Researchers tried different methods of statistical analysis, like logistic regression, decision trees, and machine learning algorithms like random forest, gradient boosting, deep neural networks, etc.

2.3.1 Symptom Patterns and Comorbidities

According to [2], a meta-analysis that includes 26 studies, reflecting that over 13,400 patients, who are mainly from China, encountered the following symptoms:
Fever (88.26%) Cough (63.68%) Fatigue (40.48%) Dyspnea (26.49%)
The most commonly seen comorbidities were hypertension (23.41%), diabetes (11.84%), and cardiovascular disease (10.00%). Chronic conditions like COPD, cerebrovascular disease, and kidney disease are also strong indicators of mortality.
On the contrary, about 1143 patients indicated a reduced total comorbidity load (merely 13.39% exhibited any comorbidity), yet still recognized hypertension and diabetes as significant predictors of mortality. The model confirmed it by its high predictive accuracy of about 0.80 [2]. A study was conducted in Malaysia. Their core study was analyzing the antibody (BNT162b2, AZ1222, Coronavac) in cancer patients versus healthy individuals. There were 246 normal people with no other disease, and there were 389 cancer patients. In that test, BNT162b2 consistently produced higher IgG antibody levels than CoronaVac in both groups. Coronavac was also linked to a significantly higher risk of breakthrough infections, especially in healthy individuals (7.34 times higher odds than BNT162b2). Sex doesn't affect the vaccine. [30]

2.3.2 Machine learning in Risk Prediction for Cancer Patients

Researchers used ML to identify the outcomes of COVID-19 among cancer patients. A study from Memorial Sloan Kettering involving 348 cancer patients used ML models to stratify severity based on their clinical and imaging data. Another longitudinal study from 4 different hospitals in Moscow, Russia, used ML models like KNN, SVM, and neural networks to predict the risk of PCC [25]. Among these ML models, KNN scored the highest accuracy of 0.8. Higher PCC risk is related to several indicators, like infection, low BMI, and liver disease. Dementia and heart failure achieved high predictive accuracy with a C-index of greater than 0.80.

2.3.3 Cancer-Related Risk Factors in COVID-19

Researchers have used EHR and sought to identify cancer-specific risk factors of 49,918 patients from 17 hospitals in California, influencing COVID-19 severity. The cancer patients who were suspected to have COVID-19 were less likely to test positive. If anyone is infected with COVID-19, they face higher risks. Hospitalization risk can be doubled in those who have BCR/ABL-negative myeloproliferative neoplasms. On the other hand, the risk can be multiplied by three times for the patients who are on venetoclax and methotrexate.

2.3.4 Post-COVID Interstitial Lung Disease and Post-Pandemic Trends in Thyroid Cancer

High-resolution CT scans and Imbio Lung Texture of 232 patients showed that most of the patients have good lung conditions after COVID-19, with the early indicators of fibrosis. Though a small percentage of the group developed ILD, they responded positively to the antifibrotic therapy. The study of thyroid cancer shows how COVID-19 treatment delays impacted thyroid cancer. The study also reflects that the tumor size was stable during delays, and post-lockdown patients were found with more multifocal tumors and lymph node metastases.

2.3.5 Research Gaps

Though there is notable progress, a significant number of gaps have been identified in the current study.

- **Dataset Limitations**: Public data sets are not available due to privacy issues. There is a lack of standardized labels, such as symptoms and complications.
- **Bias**: Training datasets are biased to specific populations like China and Russia. No robust cross-institutional validation or domain adaptation techniques are employed to test model performance on external cohorts
- **Temporal Dynamics and Longitudinal Modeling**: These models are built upon the foundation of this static data. RNN, KNN, and other deep learning and machine learning models are not frequently employed.

- **Data Imbalance**: Missingness is often ignored, potentially impacting robustness and model fairness. Class imbalance (e.g., rare comorbidities, pediatric or elderly patients) remains a significant issue, with minimal exploration of solutions such as SMOTE, class weighting, or focal loss.

- **Lack of Interpretability**: Permutation importance alone is insufficient for model transparency at the patient level.

- **Limited Use of ML**: They rely mostly on basic models like logistic regression or statistical analysis. No advanced ML method has been used

- **Real-Time Deployment**: Only a few models are designed for real-time symptom analysis. No specific mention of real-time streaming data in emergency or ICU conditions.

- **Lack of Personalized Models**: Most models are generalized and lack patient-level customization.

- **Post-COVID ILD and Imaging Gaps**: Limited imaging data and studies on ILD progression.

- **Miscellaneous Gaps**: Gaps in multi-modal integration and underutilization of genomic data. Lack of clustering techniques for unsupervised patient stratification. There are also no investigation records of the time of tumour progression related to COVID-19

Although numerous research have examined the COVID-19 results in patients using machine learning techniques like KNN, SVM, and Random Forest, most of the previous studies used single-modality datasets, e.g., imaging or clinical data. This renders them less capable of describing the intricate relationship between physiological variables and radiograph abnormalities. Moreover, the majority of those studies were interested in the short-term severity of COVID-19 infection compared to the long-term post-COVID consequences for ICU survivors.

This study, by contrast, presents a multimodal AI system that incorporates information from de-identified, real-world sources, such as chest X-ray imaging, alongside clinical records. This method of fusion focuses especially on a deeper view of the post-covids through numerical as well as vision feature fusions. It also focuses especially on interpretability, utilizing correlation maps, confusion matrices, and visualization techniques like Grad-CAM to promote clinical interpretability, a feature that previous studies did not focus on. Our research addresses the gap in how to integrate information from various modalities, how to explain models, and how to enable specific-patient predictions: Our aim is to develop an integrated approach to enable more personalized and actionable health decisions for the patients who are in ventilation or need ventilation support.

Chapter 3

Requirements, Impacts, and Constraints

3.1 Final Specifications and Requirements

In this section, the technological and methodological prerequisites for the practical use of predictive models for the duration of ventilator use in the ICU are presented.

- **Technical Requirements:** A regression model can be trained with clinical data sets with the support of high-performance computer resources.
- **Software Requirements:** Libraries in a Python programming environment are data analysis and model development libraries like `scikit-learn`, `pandas`, `numpy`, `matplotlib`, and `seaborn`.
- **Data Requirements:** organized health care files contain extensive information about patients, including test results, ventilation status, and vital signs. To perform any form of machine learning activity, there is a need to clean and format the data appropriately.
- **Methodological Requirements:** Techniques for ensemble regression, including hyperparameter tuning, cross-validation, and evaluation metrics such as MAE, RMSE, R^2, PCC, SCC, etc., also include Random Forest and Gradient Boosting.

3.2 Societal Impact

Predictive ventilator duration modeling has some possible social significance.

- **Health and Safety:** Not only can waiting times be reduced, but the quality of patient care and even mortality rates be decreased.
- **Legal Considerations:** Proper anonymization of the data and compliance with various patient privacy regulations (e.g., HIPAA or GDPR) are also important considerations.
- **Cultural and Social Impacts:** Medical staff might also need to undergo education or training to build their levels of confidence in using AI-assisted decision-making in the ICU.

3.3 Environmental Impact

While mainly computational, the effects of training models on the environment can be considered:

- This is done through the reduction of carbon emissions, which is made possible through the usage of energy-efficient computer gear.
- Environmental expenditures can be reduced by cloud-based training that involves efficient resource utilization.
- Resource conservation and sustainability are indirectly affected through the overall positive social benefit of more effective ICU care.

3.4 Ethical Issues

Some of the key ethical commitments at the heart of this project include:

- Ensuring patient information privacy and misusing sensitive information are prevented.
- Avoiding algorithmic bias that might result in misclassification or low representation of several patient groups.
- Professional accountability, clinically validating model predictions before any clinical application.

3.5 Standards

- Standards for data processing and privacy (HIPAA, GDPR) were followed.
- Reproducibility procedures, code documentation, and version control are examples of software development standards.
- Academic criteria for repeatability and transparency in AI healthcare research were followed.

3.6 Project Management Plan

- **Schedule:** The phases followed in the project were data collection/cleansing, developing models, evaluation, as well as documentation. Time-allocated milestones existed for each level.
- **Budget:** Computer resources, software licenses (if necessary), and access to data are also part of the estimated budget.
- **Resource Management:** Tasks in statistical analysis, model training, data pretreatment, and documentation that are assigned to various members of a team.

3.7 Risk Management

- **Data Quality Risks:** There are various ways to reduce missing values in patient data through the application of preprocessing techniques.
- **Model Performance Risks:** Other methods of correcting overfitting or poor generalization include cross-validation and adjusting hyperparameters.
- **Technical Risks:** There are backup systems in place to mitigate the effects of software or even hardware failure.
- **Ethical/Compliance Risks:**Stricter adherence to the anonymization and consent procedures ensures effective management of non-compliance with the regulations.

3.8 Economic Analysis

- **Cost Estimation:** Data collection, staff time, and computer resources are all expenses.
- **Benefits:** Better use of ICU resources, shorter ventilation times for patients, and even monetary savings through efficient hospital operations.
- **Cost-Benefit Analysis:** The benefits of clinical outcomes and increased business efficiencies far outweigh the investment costs of predictive modeling.

Chapter 4

Proposed Methodology

This chapter discusses a technique for predicting the invasive ventilator duration of COVID-19 patients using clinical and imaging information. The chapter discusses a systematic, design-oriented technique for multimodal learning, model selection, data preprocessing, and performance evaluation. The problem is described as a regression problem for estimating the duration of the ventilator, which is a critical clinical indicator of the severity of the disease. To ensure consistency, a series of extensive preprocessing techniques are discussed for clinical data and chest radiograph data. The chapter also discusses the selection and usage of a series of machine learning and deep learning models, including Gradient Boosting, Random Forest, Linear Regression, ResNet-18, and DenseNet-121. To improve the performance of the model, a multimodal integration technique is discussed for integrating deep image representations with structured clinical data. The proposed technique also discusses the challenges of a small dataset, patient variability, and computational feasibility. Finally, the chapter discusses an assessment technique for improving the interpretability of the model predictions using regression and correlation metrics and also discusses the usage of Grad-CAM and SHAP.

4.1 Design Process or Methodology Overview

The major aim of the design method was to produce a robust predictive model on the results of patients with COVID-19. The following crucial steps were used as a guide in developing a methodology based on a design approach:

- **Problem Definition:** The main challenge that this work tries to address is predicting the length of mechanical breathing, in terms of days, of COVID-19-infected patients. The key challenge of this work, in order to address this clinically relevant metric of disease severity, is a continuous regression task.

- **Data Preprocessing:** To fulfill the input requirements of deep learning models, image data was scaled up, improved, and normalized. To ensure data quality, clinical data were processed through data cleaning, standardization, encoding of categorical data if necessary, and elimination of data with incomplete or inconsistent information in terms of ventilation days.

- **Model Selection:** The chest X-ray image dataset was used with deep learning models ResNet18 and attention-enhanced DenseNet121. These models aimed to detect high-level radiological features with respect to respiratory deterioration. To aid in regression on the ventilation day, clinical factors as well as imaging characteristics were incorporated with multimodal learning.

- **Integration Strategy:** The multimodal (MM-CXR-CLIN) model was used to integrate the imaging and clinical information by combining the structured clinical information with the deep features from the convolutional neural network. Using this approach, the evaluation of the contribution of each data source to the prediction of the ventilation day was made possible through the use of unimodal and multimodal methods.

- **Performance Evaluation:** Regression-based evaluation metrics like Mean Absolute Error (MAE), Root Mean Square Error (RMSE), coefficient of determination R^2, Pearson Correlation Coefficient (PCC), and Spearman Correlation Coefficient (SCC) are used as evaluation metrics for each of these models. Together, these metrics assess correlation with actual ventilation duration values, prediction accuracy, and error rates.

- **Constraint Handling:** The limited size of the dataset ($n = 213$ ventilation instances), the heterogeneity in the patient clinical profile, as well as the limited computing resources, have to be kept in mind. These have contributed to the selection of more expressive attention models in combination with effective models such as ResNet18. Grad-CAM has been employed to address the issue of the effectiveness of the models.

Thus, this methodical approach was able to address the practical constraints of the data's variability, interpretability concerns, and computer feasibility while at the same time being consistent with the intended purposes of the research.

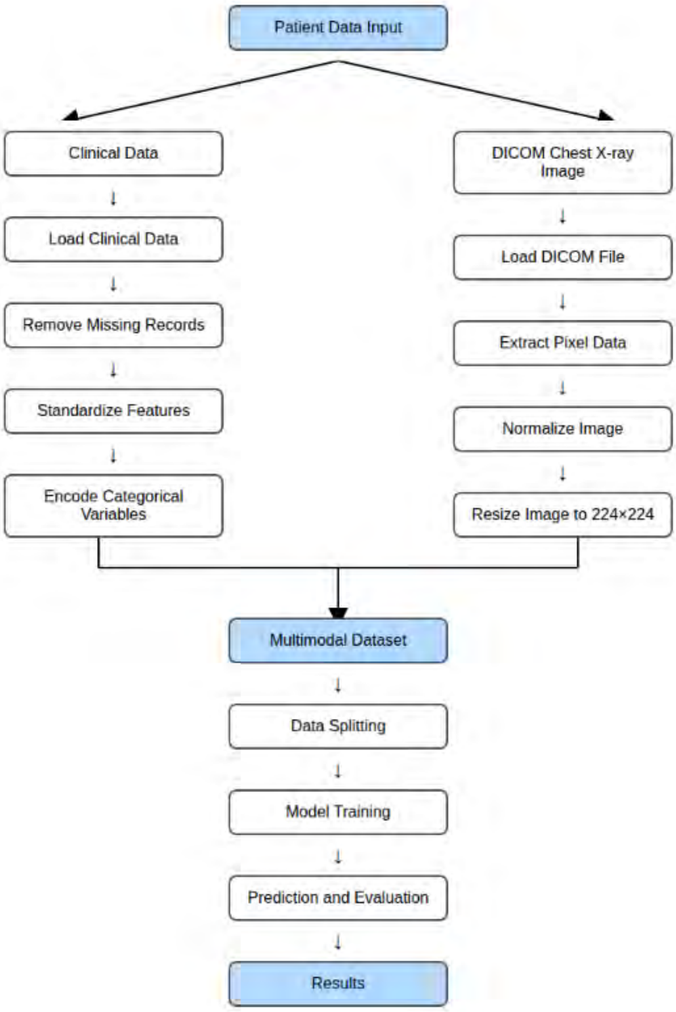

Figure 4.1: Methodology
[13]

4.2 Preliminary Design or Design (Model) Specification

Several models that were equally promising were constructed and tested concurrently to attain the goals that were stipulated:

- **Image-Based Models:** Four state-of-the-art Convolutional Neural Networks (CNNs) were put into practice:
  - *DenseNet*: Use dense connections to prevent vanishing gradients and reuse of features.
  - *ResNet*: To allow for extremely deep networks while maintaining stable convergence, residual connections are used.
- **Clinical Data Models:** Several machine learning algorithms were explored to capture relationships in structured clinical data:
  - *Linear Regression*: For binary outcome prediction and baseline linear modeling.

- *Gradient Boosting:* An ensemble learning technique that builds multiple weak models sequentially, correcting previous errors to improve accuracy and capture non-linear relationships between features.
- *Random Forest*: A strategy based on ensemble learning to prevent overfitting and to find non-linear interactions between features.

- **Simulation and Verification:** The respective dataset (images/medical records) was used for training and evaluating each model. Multiple simulations were conducted for each model to evaluate the performance using the respective dataset and the ground truth labels assigned to the data. The best model for each modality was identified by evaluating the outcome of each model compared to the others for each modality, providing insights on the integration of the models using the hybrid approach.

This multi-model approach guaranteed that the final answer is correct while at the same time being generalizable through providing a methodical way of comparing deep learning approaches with conventional machine learning approaches.
.

4.3 Data Collection

For this research, we tried to find the dataset that covers who needs ventilation support. But there was 1384 data or rows, and ventilation was applicable for 213. We looked at the TCIA, We found chest X-ray images of patients and clinical data from the TCIA. The patient's X-ray image and clinical data are linked by patient ID.

4.3.1 Data Cleaning

The dataset used in this study was obtained from a publicly available COVID-19 clinical and imaging repository [28]. Initially, the dataset contained records for a large number of patients with heterogeneous clinical characteristics. As the objective of this work is to predict the duration of mechanical ventilation, only patients with valid and reliable ventilation-day information were retained. Records with missing or undefined ventilation duration values were removed. In addition, patients without corresponding chest X-ray DICOM files were excluded to ensure consistency in the multimodal learning framework. After these filtering steps, a total of 213 patient samples with complete clinical data, ventilation-day labels, and associated chest X-ray images were included for further analysis.

4.3.2 Data Transformation

Prior to model training, both imaging and clinical data were transformed to ensure compatibility with deep learning models. The pictures of the chest X-rays were then intensity-normalized, scaled to a fixed resolution of 224×224, and transformed into three-channel representations. In order to improve the model's generalization performance, various data augmentation techniques such as rotation and random horizontal flipping were used during the training process.

All the variables in the clinical data were converted to a numerical format with zeros indicating the absence of a value. In order to ensure the consistency of the features' scale values, z-normalization is utilized as a method for standardization. Prior to training the model, the variable 'ventilation-day' is transformed using a $\log(1+x)$ method.

4.3.3 Data Reduction

The original clinical dataset contained a large number of variables, including patient identifiers, demographic information, comorbidities, and laboratory measurements. Several features were not directly relevant to the prediction of ventilation duration or contained a high proportion of missing or inconsistent values. Such features were removed to reduce noise and improve model robustness. After excluding non-informative and incomplete variables, a reduced set of clinically meaningful features was retained. These selected clinical features were then integrated with imaging-derived features within a multimodal learning framework, with ventilation duration (in days) defined as the final regression output.

4.3.4 Summary of Preprocessed Data

The dataset used in this study consists of 213 patient records and originally included 131 attributes. After preprocessing, three redundant columns containing only missing values were removed, leaving 19 analytically meaningful variables. The dataset provides a comprehensive overview of COVID-19 patients, integrating demographic, clinical, and symptomatic information. Demographic variables include patient identifiers and gender. Clinical outcomes capture COVID-19 status, survival outcome, ICU admission, and use of ventilation. Comorbidities and treatment history such as malignancies, hypertension, smoking status, and prior use of antibiotics or NSAIDs are also included. Symptom-related variables are represented in binary format, indicating the presence or absence of features such as cough, dyspnea, nausea, vomiting, diarrhea, abdominal pain, and fever. Contextual information is given through the type of visit (inpatient or emergency room). Notably, this dataset is reliable and prepared for statistical and predictive modeling because there are no missing values for its major features.

Dataset Overview

- **Size:** 213 records $\times$ 131 columns
- **Valid Columns:** 131 (output label: invasive_vent_days)
- **Key Attributes:**
  - **Patient Info:** to_patient_id, gender
  - **Clinical Outcomes:** covid19_status, last.status (0/1 outcome), is_icu, was_ventilated
  - **Medical History:** malignancies_v, htn_v (hypertension), antibiotics_use_v, nsaid_use_v, smoking_status

- **Symptoms:** cough_v, dyspenea_admission, nausea_v, vomting_v, diarrhea_v, abdominal_pain_v, fever_v
- **Contextual Info:** visit_concept (e.g., Inpatient Visit, Emergency Room Visit)

- **Data Types:**
  - **Categorical:** covid19_status, gender, visit_concept
  - **Binary/Numeric:** Most clinical and symptom fields (0 = absent/negative, 1 = present/positive)

- **Missing Data:**
  - None in major variables
  - Only the placeholder Unnamed columns contained missing values and were dropped

- **Preprocessing Applied:**
  1. Retained structured binary/categorical format for analysis
  2. Ensured no null values in relevant features
  3. Treated patient IDs as identifiers, not analysis features

4.4 Implementation of Selected Design

We have worked on two types of datasets. One is clinical data, which are in CSV form, and the other is imaging data. SO when we downloaded the dataset, It was a huge datasets of 500 GB. One patient has many x-ray images. But there was no specific timeline mentioned in the metadata. The date was set randomly, and there is no chance of understanding the sequence of the collection of x-ray images. Some patients have only one X-ray; some have more X-ray images. They are classified as CHEST AP PORT and CHEST AP VIEW ONLY. So we choose CHEST AP PORT for our dataset, and where there are many X-ray images, we consider the latest CHEST AP PORT from their random date. The date is random because all the dates were from 1900 to 1901, and covid-19 came in 2019-2020. So keeping that in mind, we selected our population. We double-checked with our clinical data that all the patients are selected properly or not. After finalizing datasets and preprocessing our clinical data, we proceeded to test with pretrained models like ResNet and DenseNet and embedding methods like Grad-CAM, MedGemma, and MedCLIP. Our imaging data was in DICOM (Digital Imaging and Communications in Medicine) or (.dcm) format. The imaging dataset was split into 3 parts: test, train, and val. We applied the 70-15-15 method for splitting the dataset.

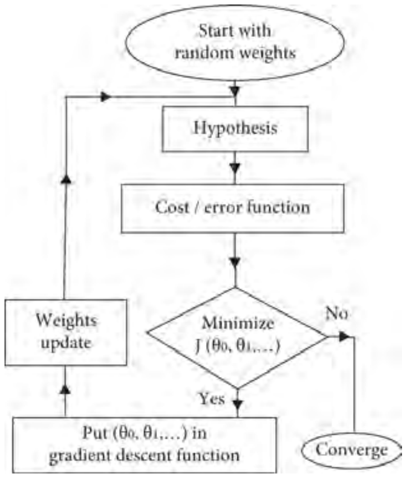

Figure 4.2: linear Regression
[7]

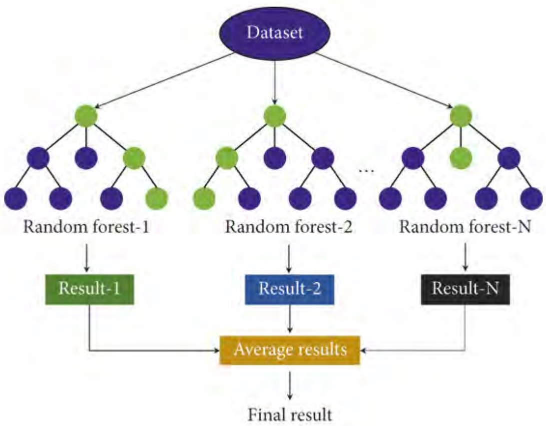

Figure 4.3: Random Forest
[8]

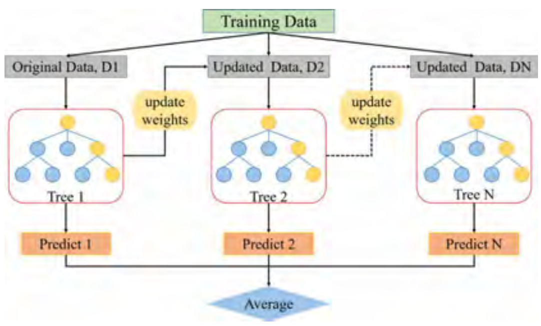

Figure 4.4: Gradient Boosting
[13]

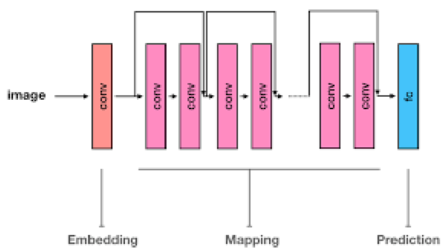

Figure 4.5: Resnet18
[19]

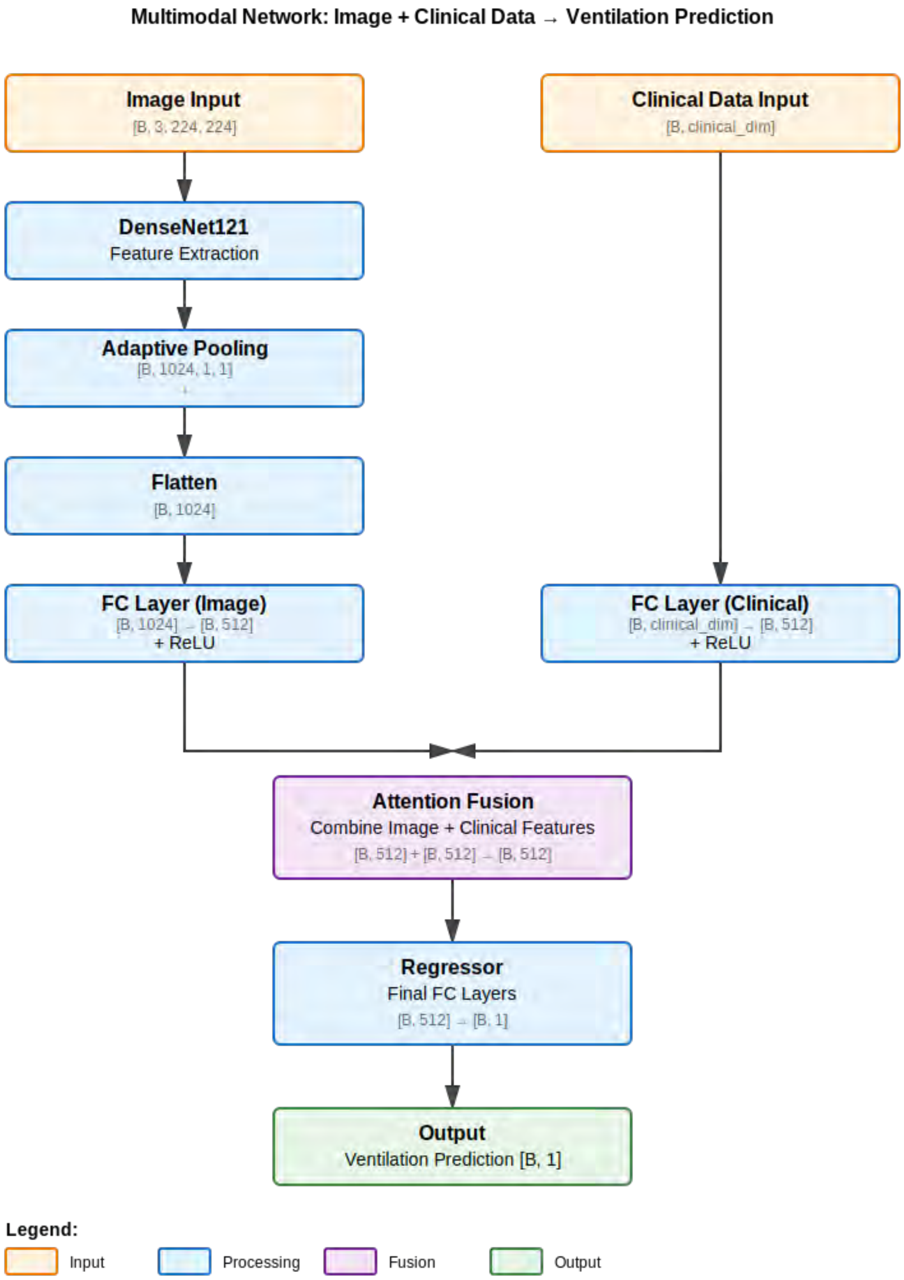

Figure 4.6: MM-CXR-CLIN-Attention Network
[12]

Chapter 5

Result Analysis

This chapter offers a comprehensive analysis of the experimental outcomes from both deep learning and conventional machine learning models. For the assessment of the prediction accuracy and reliability, the performance of the models is checked using various regression-based parameters, such as Mean Absolute Error (MAE), Root Mean Squared Error (RMSE), coefficient of determination (R^2), Pearson Correlation Coefficient (PCC), Spearman Correlation Coefficient (SCC) and ROC(graph). ROC analysis was performed after thresholding the continuous prediction to evaluates model's ability to discriminate the prolonged ventilation cases. The application of ensemble-based models, such as Gradient Boosting Regression and Random Forest Regression, reveals a considerable outperformance of Linear Regression, as the models are able to capture the non-linear relationships present in the ICU clinical data, as observed from the experimental outcomes of models that used only clinical information. The application of multimodal deep learning-based models that incorporate both chest X-ray images and clinical information reveals moderate performance, with DenseNet-121 performing slightly better than ResNet-18. For the purpose of highlighting the important factors and regions that affect the predictions, the study incorporates SHAP and Grad-CAM-based visual and feature-based explainability techniques. The comparative analysis reveals that the most important modality for the prediction of ventilator duration is the structured clinical information, while the application of the imaging modality provides limited advantages.

5.1 Performance Evaluation

We used both deep learning models like DenseNet and ResNet and classic machine learning models like linear and logistic regression to test how well the suggested approach worked. In order to ensure a full validation of the performance of the regression, various indicators were used to measure the performance. Larger variances were used to calculate the weights to obtain the mean amount of prediction errors using the Root Mean Squared Error (RMSE). The measure used to obtain the strength of the linear association between the two values was the Pearson Correlation Coefficient (PCC). The Spearman's Correlation Coefficient (SCC) was also evaluated for its ability to withstand non-linear trends and outliers by assessing the monotonic relationship between predictions and true outcomes. ROC curve also shows true positive which indicates true predictions in this case.

Results on Clinical Data (CSV file)

For the clinical dataset, We performed random forest regression, gradient boosting, and linear regression. The dataset was first divided into three subsets: 70% for training, 15% for validation, and 15% for testing. This split ensured that model performance could be evaluated objectively on unseen data while allowing enough samples for both learning and hyperparameter tuning. The primary evaluation metric was **MAE**,**RMSE**,R^2, **SCC,PCC** and ROC representing the accuracy among all predictions.

Accuracy Matrix	**value**
MAE	5.47
RMSE	7.22
R^2	0.62
PCC	0.79
SCC	0.80

Table 5.1: Performance of Gradient Boosting on clinical data

Accuracy Matrix	**value**
MAE	5.86
RMSE	7.69
R^2	0.57
PCC	0.76
SCC	0.76

Table 5.2: Performance of Random Forest on clinical data

Accuracy Matrix	**value**
MAE	11.44
RMSE	14.07
R^2	-0.42
PCC	0.38
SCC	0.28

Table 5.3: Performance of Linear Regression on clinical data

Results on Image + Clinical Data (Deep Learning Models)

For image data, three deep learning architectures were tested:

Accuracy Matrix Name	Value
MAE	6.89
RMSE	9.30
R^2	0.31
PCC	0.61
SCC	0.62

Table 5.4: Performance of ResNet-18 on imaging multimodal data

Accuracy Matrix Name	Value
MAE	6.594
RMSE	8.568
R^2	0.453
PCC	0.716
SCC	0.678
ROC	0.821

Table 5.5: Performance of DenseNet-121 on imaging multimodal data

Among deep learning models, DenseNet achieved the highest accuracy, while ResNet performed almost similarly to DenseNet but lagged behind.

5.1.1 Evaluation Metrics and Interpretation

To comprehensively evaluate the predictive performance of the proposed machine learning and deep learning models, multiple regression-based evaluation metrics were employed. Considering the fact that ventilator time is a continuous variable, which is subject to a wide degree of variation, an exclusive dependence upon a single performance criterion can lead to results which may be either invalid or incomplete. Therefore, in order to ensure a valid and complete analysis, a combination of error-based, variance-based, and correlation-based metrics was utilized.

5.1.2 Evaluation Metrics

Mean Absolute Error (MAE)

Mean Absolute Error (MAE) measures the average absolute difference between the predicted ventilator days and the true observed values:

$$\text{MAE} = \frac{1}{n}\sum_{i=1}^{n} |y_i - \hat{y}_i| \tag{5.1}$$

The metric appears to be very natural to clinical practitioners, as it can be interpreted in terms of days. The ventilated days that are predicted are likely to vary less from the real result when the MAE value is low.

Root Mean Squared Error (RMSE)

Root Mean Squared Error (RMSE) is defined as:

$$\text{RMSE} = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(y_i - \hat{y}_i)^2} \tag{5.2}$$

Because the squaring factor is included, the RMSE penalizes bigger prediction mistakes more severely than the MAE. This aspect is especially important in the intensive care unit, where considerable mistakes in ventilator duration can occur.

Coefficient of Determination (R^2)

The coefficient of determination quantifies the proportion of variance in ventilator days explained by the predictive model:

$$R^2 = 1 - \frac{\sum_{i=1}^{n}(y_i - \hat{y}_i)^2}{\sum_{i=1}^{n}(y_i - \bar{y})^2} \tag{5.3}$$

A strong explanatory model will result when the R^2 value is closer to 1. On the other hand, values closer to or below 0 indicate limited predictive value for the model with poor performance relative to a mean-based model.

Pearson Correlation Coefficient (PCC)

The Pearson Correlation Coefficient (PCC) evaluates the linear association between predicted and actual ventilator days:

$$\text{PCC} = \frac{\text{cov}(y, \hat{y})}{\sigma_y \sigma_{\hat{y}}} \tag{5.4}$$

A higher PCC reflects stronger linear agreement and indicates that the predicted trends closely follow the true clinical patterns.

Spearman's Rank Correlation Coefficient (SCC)

Spearman's Rank Correlation Coefficient (SCC) measures monotonic relationships between variables and is defined as:

$$\text{SCC} = 1 - \frac{6\sum d_i^2}{n(n^2 - 1)} \tag{5.5}$$

Clinical datasets, which usually display non-linear correlation, non-normal distributions, and outliers, are most appropriate when using this metric. Thus, SCC, which measures rank-based consistency instead of rigid linear dependence, is an improvement over PCC.

Receiver Operating Characteristics (ROC) Curve

A graphical representation of how a binary classifier system's diagnostic ability is affected when a change is made to the system's discrimination threshold is given by the Receiver Operating Characteristic Curve. The True Positive Rate and False Positive Rate are shown on this curve.

$$\text{TPR} = \frac{TP}{TP + FN} \tag{5.6}$$

$$\text{FPR} = \frac{FP}{FP + TN} \tag{5.7}$$

$$\text{ROC} = \{(\text{FPR}(t), \text{TPR}(t)) \mid t \in R\} \tag{5.8}$$

$$\hat{y}_{\text{bin}} = \begin{cases} 1, & \hat{y} > \tau \\ 0, & \hat{y} \leq \tau \end{cases} \tag{5.9}$$

5.1.3 Interpretation of Clinical Machine Learning Results

The regression models trained exclusively on structured clinical features demonstrated notable performance variation across different algorithms.

Gradient Boosting Regression

In terms of predicting the length of an intrusive ventilator, the best model was the Gradient Boosting Regression model. The model had an RMSE of 7.22 days and an MAE of 5.47 days. The predictions made by the model were, on average, within 5 to 7 days compared to actual values. The model had a respectable level of accuracy in its predictions, as it was able to explain almost 62% of the variability in the ventilator time with a R^2 value of 0.62. The model had a remarkable level of correlation performance, as it had a Pearson Correlation Coefficient (PCC) value of 0.79 and a Spearman Correlation Coefficient (SCC) value of 0.80. The model was successful in picking up both linear and non-linear relationships, as these values were high and showed a strong agreement between the predicted and actual values. The Gradient Boosting model, which is excellent at picking up intricate relationships between different parameters, which are often non-linear and complex in nature, is what made this model perform exceptionally well. The model is highly reliable and successful at forecasting patient outcomes in critical situations.

Random Forest Regression

In terms of forecasting the duration of ventilation for patients infected with COVID-19, the Random Forest Regression model offered competitive and stable results. The model was able to achieve an RMSE of 7.69 days and an MAE of 5.86 days, showing slightly less precise results compared to the Gradient Boosting model but still relatively similar to observed values. The model offered moderate predictive accuracy, accounting for 57% of the variance in the duration of ventilation as shown by the R2 value of 0.57. The model's ability to handle linear and non-linear relationships between variables is further shown by the high correlation values of 0.76 and 0.76, respectively, using the Pearson and Spearman correlation coefficients. The high correlation values indicate that there is a strong concurrence between expected and actual outcomes. The high correlation values are further evidence of the model's robustness and proficiency in accurately predicting the duration of ventilation based on the complex relationships between variables. The model's strength and proficiency in dealing with complex relationships between variables made it particularly useful in forecasting the duration of ventilation. The model's strength and proficiency in dealing with complex relationships between variables made it particularly useful in forecasting the duration of ventilation.

Linear Regression

In comparison to more advanced algorithms like the Gradient Boosting and the Random Forest, the Linear Regression model underperformed when it came to the prediction of the duration of the ventilation. The estimates were off by a considerable margin, as indicated by the MAE of 11.44 days and the RMSE of 14.07 days. This shows that the model was not making accurate predictions when it came to the duration of the ventilation. It even underperformed when compared to a simple model that predicted the mean for every data point. This is indicated by the negative R2 of -0.42. This shows that the linear regression model was not able to grasp the complexity of the problem of the duration of the ventilation in the ICU. This is further indicated by the weak relationship between the expected and the actual values, as indicated by the PCC of 0.38 and the SCC of 0.28. This shows that the model was not able to grasp the relationship between the features and the target variable. This shows that the linear regression model was not able to grasp the complicated and complex relationship between the features and the target variable. This demonstrates that, in comparison to the linear regression model, more sophisticated models, such as Gradient Boosting and Random Forest, are better able to understand the nuanced and complex dynamics of the problem.

5.1.4 Interpretation of Multimodal Deep Learning Results

Deep convolutional neural networks were used to integrate clinical information with chest X-ray images for multimodal learning.

ResNet-18

The following performance was obtained with the ResNet-18 architecture. We got MAE of 6.89 days. Root Mean Square Error (RMSE) of 9.30 days and $R^2$0.31 and corelation matrix PCC of 0.61 and SCC of 0.62 These findings reveal that radiographic characteristics offer significant but insufficient information about ventilator time, indicating a moderate predictive potential.

DenseNet-121

DenseNet-121 demonstrated improved performance. The model resulted in a mean absolute error of 6.59 days and a root mean squared error of 9.26 days for the prediction of the duration of mechanical ventilation. The coefficient of determination, R^2, was 0.453, indicating a fair proportion of the variance, while the Pearson correlation coefficient (PCC) and the Spearman correlation coefficient (SCC) were 0.66 and 0.64, respectively, indicating strong agreement between the predicted and actual values. The dense connectivity structure played a significant role in this regard by improving the propagation of features and the reuse of learned patterns.

5.1.5 Training Curve:

From the analysis of the learning curve, it is evident that our suggested model converges fast and consistently across all folds in cross-validation. There is a consistent decrease in both training and validation losses, and there is no sign of unstable behavior, which is a sure indicator of a well-optimized model and a level of complexity that is under control. Moreover, the small margin between training and validation losses is an indicator that there is virtually no overfitting and that early stopping can be used to reduce training time without compromising performance.

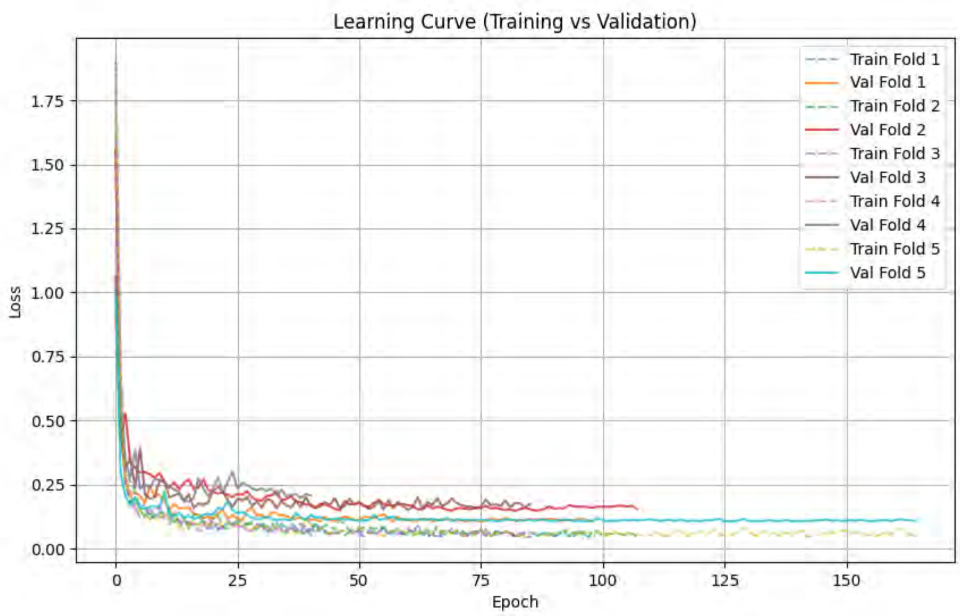

Figure 5.1: Training Curve

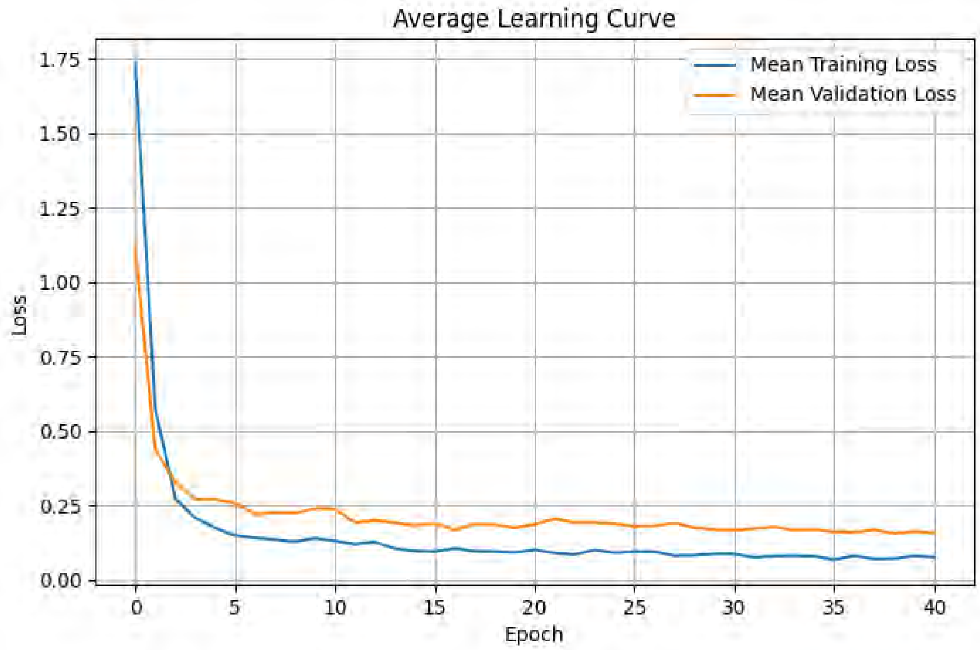

Figure 5.2: Average Mean Curve

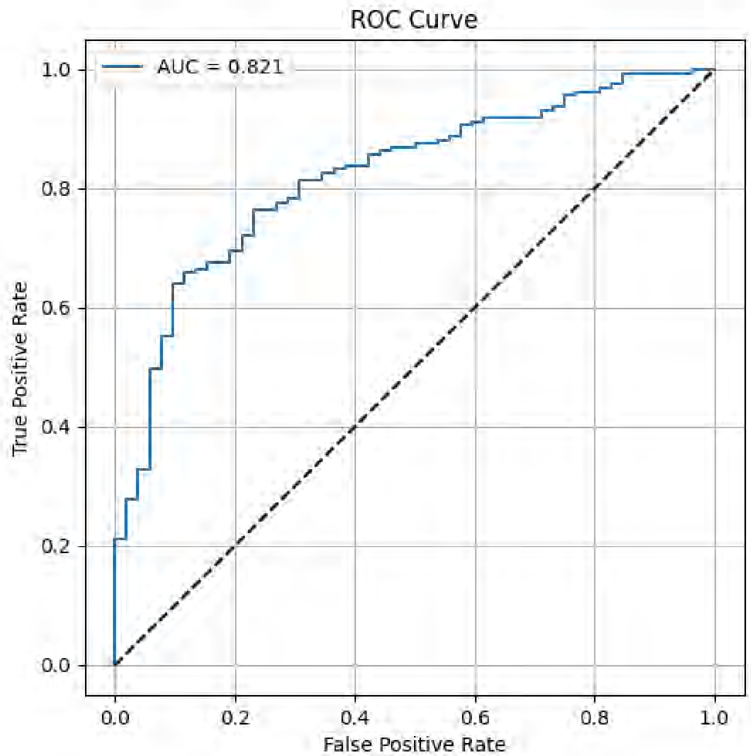

Figure 5.3: ROC Curve

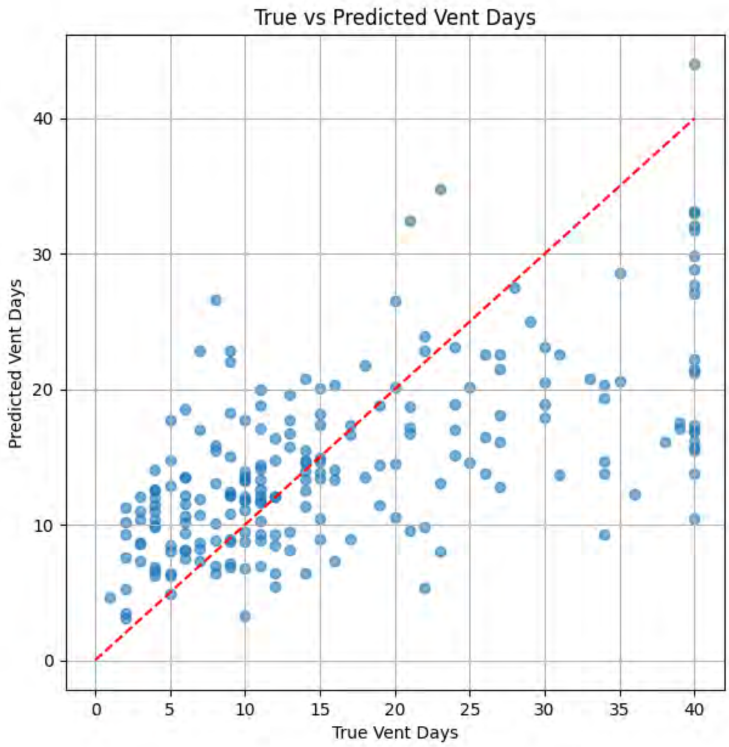

Figure 5.4: True vs Predicted

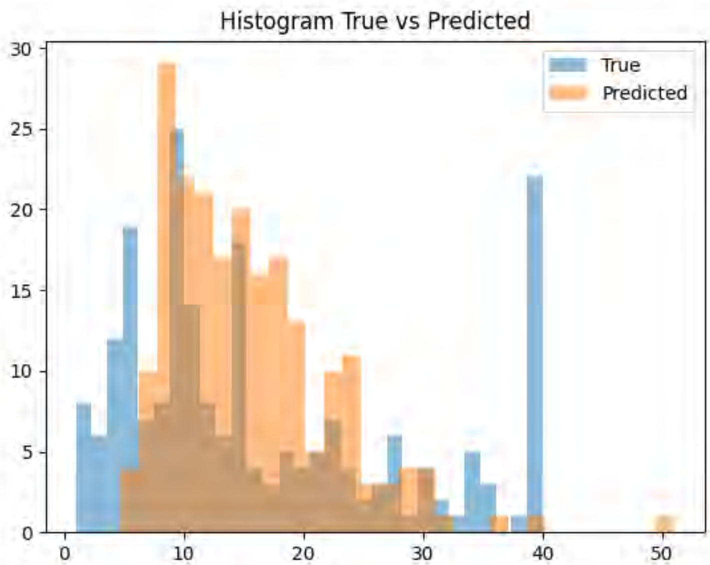

Figure 5.5: Histogram

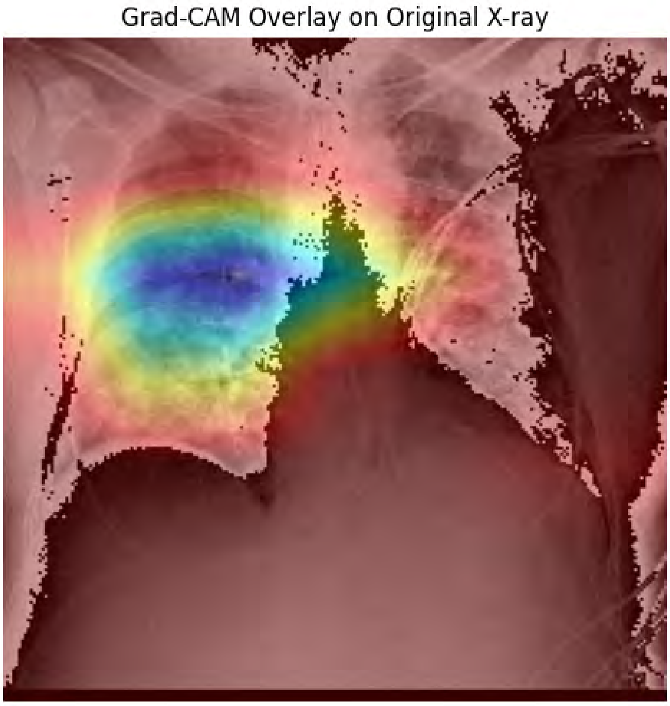

Figure 5.6: Grad CAM Heat map

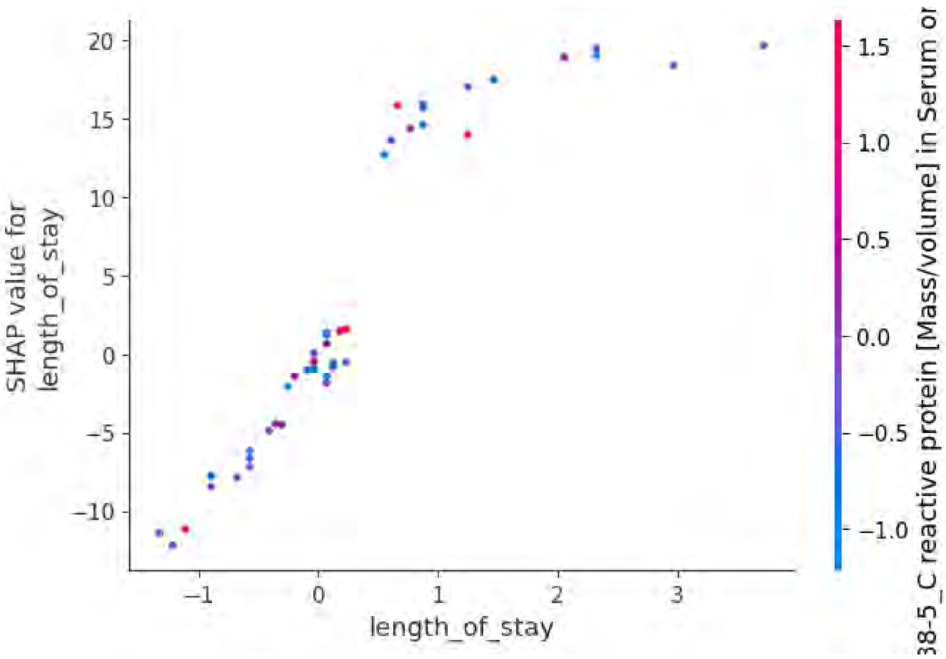

Figure 5.7: SHAP Value for length of Stay

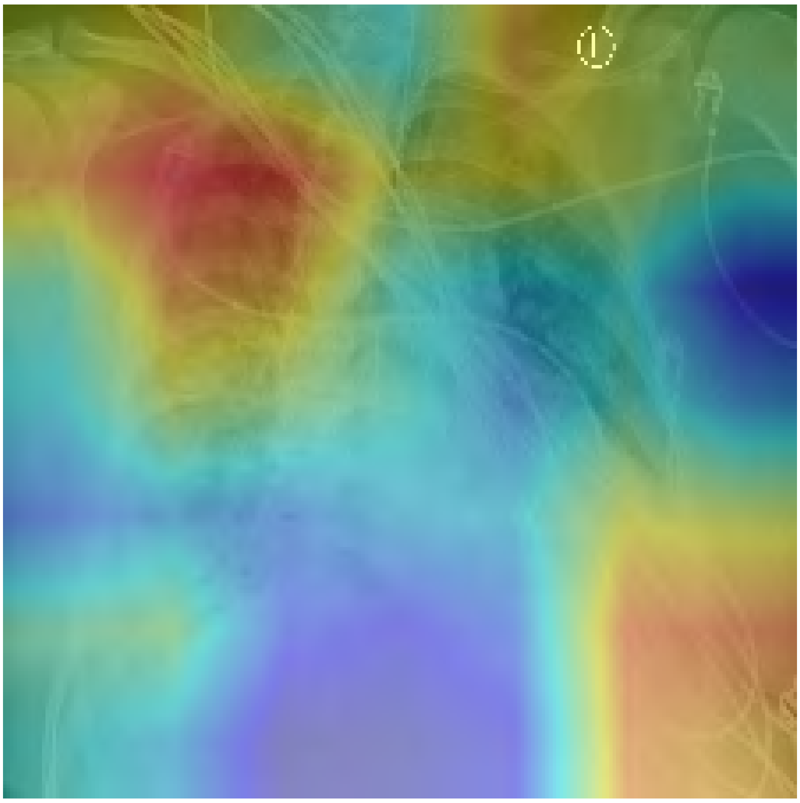

Figure 5.8: Attention Map

5.1.6 Comparative Analysis

Based on the multimodal deep learning performance comparison, the following significant results have been observed:

- The advantage of the multimodal approach, combining clinical and image modalities, can be seen from the better predictive capabilities observed in the multimodal deep learning models compared to the image-based deep learning models.

- DenseNet-based models have better performance than ResNet-based models in terms of MAE and RMSE, reflecting the better capabilities of the former in the reuse and learning of features, based on the multimodal deep learning architectures considered.

- Item correlation values show significant improvement with the incorporation of clinical modalities, with the values of PCC and SCC increasing in the multimodal models, reflecting better agreement between expected and actual values for the duration of the ventilator.

- In terms of overall accuracy and variance, the ensemble machine learning models, based solely on the use of structured clinical data, have better performance than the multimodal deep learning models, even with the improved capabilities observed in the multimodal models.

- Based on the results, the limitations in the use of the multimodal approach can be seen from the relatively low R2 values observed for the multimodal

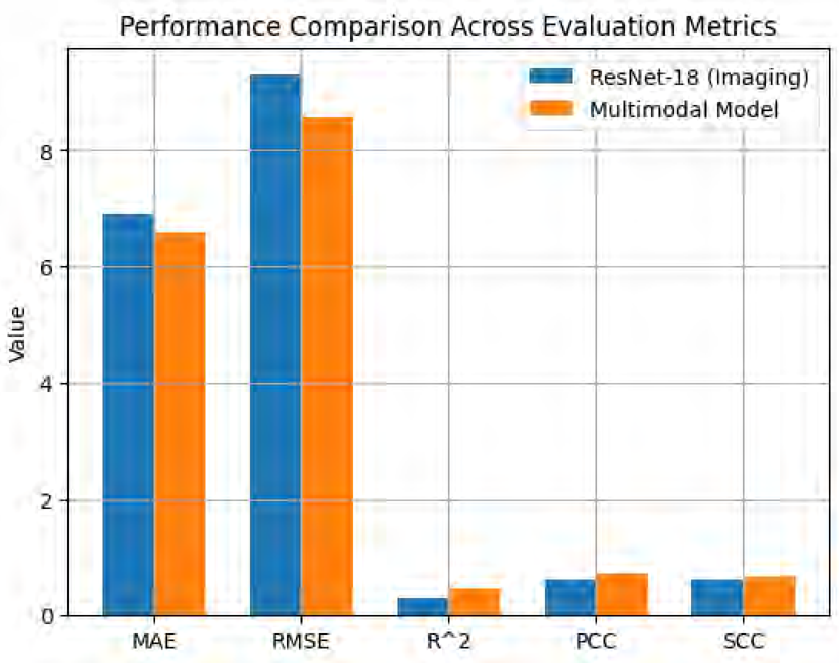

Figure 5.9: Performance Comparison

deep learning models, reflecting the need to use larger datasets to maximize the potential of the deep learning models in the prediction of patient outcome in the Ventilation.

5.1.7 Summary

In brief, the results of this trial only reinforce that, when it comes to knowing how long you'll need a ventilator, organized clinical data still represents the most effective means of assessment. Of all the models assessed, Gradient Boosting Regression emerged as the most reliable and accurate means of assessment. Although multimodal deep learning, using chest X-ray images, provided additional avenues of assessment, it did not prove to be superior to advanced ensemble machine learning approaches, at least not with the data provided.
A good, reliable, and clinically interpretable approach to ventilator duration prediction can be achieved by combining error-based metrics (MAE, RMSE), variance-based metrics (R^2), and correlation-based metrics (PCC, SCC).

5.2 Analysis of Design Solutions

In this section, the design options will be evaluated using a set of predefined criteria: accuracy, efficiency, feasibility, cost, and strengths and weaknesses. This will help in determining the best option for forecasting the duration of invasive ventilation required by ICU patients. The ensemble machine learning algorithms, such as Random Forest and Gradient Boosting, were found to have better predictive accuracy. These algorithms resulted in lower error rates and higher correlation compared to the linear algorithms. On the other hand, the Linear Regression algorithm demonstrated low accuracy, as shown by the negative coefficient of determination, and could not represent the actual distribution of the data.

Due to the relative simplicity of linear regression to mathematically formalize, it requires very little time to train. It also requires very little computational power. The trade-off is that you pay with less predictive accuracy. Gradient Boosting and Random Forests require more effort to train, but once they are trained, they require very little computational power to run in real-time. To determine the feasibility, we considered their interpretability, ease of integration with existing hospital information systems, and their robustness to missing values. Tree-based ensemble methods have been shown to be useful in ICU environments due to their robustness to varying factors.

Although deep learning techniques provide good representations, their complexity, along with decreased interpretability, may impede their rapid adoption, especially in therapeutic settings, without additional tools to facilitate explainability. From a cost perspective, licensing costs are low because all methods are based on open-source platforms. However, increased computational requirements may also result in indirect costs, such as system maintenance and computational resource usage. In terms of ensemble methods, these techniques provide good accuracy with relatively low computational requirements, thus providing a good balance between cost and accuracy. Gradient Boosting and Random Forest methods appear to provide a good balance between predicted accuracy, robustness, and feasibility, as derived from overall performance evaluation. These techniques have good overall performance, as they are capable of good generalization and robustness to non-linear data patterns, although their major drawback is decreased interpretability, as seen with other linear methods. Although Linear Regression is simple, easy to implement, and requires minimal computational resources, it fails to provide adequate representations to address the complexity of ICU data. In conclusion, ensemble-based learning systems are seen as the most appropriate design, considering overall performance, while at the same time being feasible and cost-effective.

5.3 Final Design Adjustments

There are a lot of modifications and ideas that can be implemented to improve model interpretability and reliability, as inspired by the performance of regression models in predicting the duration of invasive ventilation.

- **Model Selection:** Across all metrics, Gradient Boosting Regression performed best, followed closely by Random Forest. This underscores the need to use models that can effectively handle non-linear relationships, as linear regression did poorly. Thus, it is recommended that the basic prediction framework be built around ensemble tree-based models.

- **Hyperparameter Tuning:** We could even improve performance by tweaking other important parameters, such as learning rate, number of estimators, depth, and min samples per leaf. At first we kept the learning rate at 0.0001 and the epoch at 50. We got an R^2 value of 0.28. But later after hyperparameter tuning, we make the learning rate 0.00001, batch size= 8 and the epoch

300. Then our R^2 value improved to 0.45. The dataset maintains a strong correlation with actual values but methods like grid search and Bayesian optimization could improve MAE and RMSE.

- **Feature Engineering:** Supplying tree-based models with more detailed clinical information, such as ventilation ratios, patient severity scores, or aggregated vital signs over time, may improve the models' prediction accuracy.
- **Explainability:** Clinician trust in the model outputs may increase if we are able to demonstrate the effect of certain features on the predicted ventilator duration using tools such as SHAP values or Partial Dependence Plots.
- **Data Quality and Augmentation:** In addition to the augmentation techniques, maintaining detailed and accurate clinical records may help stabilize model training and improve model generalization.

5.4 Statistical Analysis

The evaluation of the prediction models was carried out using various statistical metrics and methods, which are as follows:

- **Mean Absolute Error (MAE)** – It determines the average of the absolute difference between the number of days observed and the number of days anticipated on the ventilators.
- **Root Mean Squared Error (RMSE)** – It shows the level of weight given to the greater deviations in order to account for the overall prediction error.
- **Coefficient of Determination (R^2)** – It shows the percentage of the variation in the length of time on the invasive ventilators that the model is able to explain.
- **Pearson Correlation Coefficient (PCC)** – It shows if there is a linear relationship between the expected days on the ventilators and the actual days observed.
- **Spearman Correlation Coefficient (SCC)** – It shows the monotonic, rank-based correlation in order to identify the linear as well as non-linear relationships.

Model	**MAE (days)**	**RMSE (days)**	**R^2**	**PCC**	**SCC**
Gradient Boosting	5.47	7.22	0.62	0.79	0.80
Random Forest	5.86	7.69	0.57	0.76	0.76
Linear Regression	11.44	14.07	-0.42	0.38	0.28
ResNet-18 (Image)	6.89	9.30	0.31	0.61	0.57
DenseNet-121	6.594	8.563	0.453	0.716	0.678

Table 5.6: Performance of different models on clinical and imaging data

While Linear Regression scored poorly ($R^2 = -0.42$), Gradient Boosting and Random Forest demonstrated significant predictive accuracy and high correlation across both linear and non-linear metrics, indicating the existence of non-linear correlations inherent in ICU clinical data.

5.5 Comparisons and Relationships

From the above, some key observations can be made:

- **Gradient Boosting vs. Random Forest:** When compared to the performance of the Random Forest model, where MAE is 5.86 days and RMSE is 7.69 days, Gradient Boosting outperforms it slightly with MAE at 5.47 days and RMSE at 7.22 days. This indicates that gradient boosting is better at predicting the values. Moreover, the correlation between the actual and predicted values is also high, as indicated by PCC values close to 0.79/0.76 and SCC values close to 0.80/0.76.

- **Linear Regression:** On the other hand, the performance of the Linear Regression model is poor, as indicated by MAE at 11.44 and RMSE at 14.07. Moreover, the $\mathbf{R^2}$ is negative, which indicates that the prediction of the duration of ICU ventilator use cannot be performed using linear regression.

- **Clinical Data Insights:** It is quite obvious that the duration of ventilator use is dependent on various patient vitals, severity scores, and intervention history, which interact with each other in a non-linear fashion.

- **Comparison with Literature:** The paper titled "Predicting Mechanical Ventilation and Mortality in COVID-19 Using Radiomics and Deep Learning on Chest Radiographs: A Multi-Institutional Study" they predicted wheather they required ventilation or not and mortality prediction. Their mAUC was 0.75. We got our ROC 0.821. They used Resnet-18 and didnot multimodal it combibinig with clinical and imging data.[5]. Another paper named "Predicting, Analyzing and Communicating Outcomes of COVID-19 Hospitalizations with Medical Images and Clinical Data" by (Stritzel Raidou, Eurographics VCBM 2022) showed ther results Training: 0.80 and Validation: 0.78. Whereas we performed ROC better than their baseline model Support Vector Machine (SVM).[31]

5.6 Discussions

Key results were observed in predicting the length of stay of the ICU patients on invasive ventilation support. Tree-based ensembles, particularly the gradient boosting regression model, were found to be the most effective among the several approaches. The clinical data's complex non-linear correlations and interactions were best captured by gradient boosting, obtaining a mean absolute error of 5.47 days, R^2 of 0.62, and high Pearson and Spearman correlation values of 0.79 and 0.80, respectively. Although Random Forest Regression performed well on various datasets used in the study, its performance was slightly less optimal compared to Gradient Boosting. On the other hand, the performance of Linear Regression was found to be poor, with low correlation values and negative R^2 values, proving the limitation of linear models in predicting intricate outcomes in the ICU setting. The ROC curve usually used for classification so, we made a threshold of 7 days. If vent_days¿ 7 will be regarded as prolonged ventilated days.[4]. The average duration of stay of ICU patients on invasive ventilation assistance was found to be moderately predicted by the multimodal deep learning model employing both clinical data and chest X-ray images. Although DenseNet-121 performed slightly better than the ResNet-18 model, thanks to better feature propagation, the lower R^2 values of the latter model indicated that the information in the chest X-ray images was limited in predicting the length of stay of the patients on the ventilators, especially when the dataset size is limited.

Structured clinical information was found to be the most informative modality in predicting the length of stay of the ICU patients on the ventilators, as observed in the comparison of the three different models. Ensemble learning methods were found to perform the best in obtaining the most precise and clinically interpretable results, with Gradient Boosting being the best model among the three. The multimodal deep learning model's performance is constrained by the amount of the dataset and the model's complexity, but the model's high PCC and SCC values demonstrated its dependability in producing accurate results that match the real values.

The results of the study are clinically important in that the precise prediction of the length of stay of the ICU patients on the ventilators could be of great value in the future.

Chapter 6

Conclusion

6.1 Transition from Results

The performance of regression models in predicting invasive ventilator duration using clinical data in a structured format. Gradient Boosting Regression and Random Forest Regression also outperformed Linear Regression with a better fit of complex nonlinear relationships of interactions and maintaining a strong correlation with the observation of outcomes. These Findings present a clear basis on which the study can be concluded by emphasizing ensemble methods and providing a background for actionable insights and Recommendations

6.2 Summary of Findings

This study aimed to examine predictive modeling in regards to invasive ventilator duration in ICU patients based on structured clinical features. Three different machine learning models were utilized: Gradient Boosting Regression (GBR), Random Forest Regression (RFR), and Linear Regression were implemented and evaluated. The key findings are as follows:

- The best performance was observed using Gradient Boosting Regression with MAE = 5.47 days, RMSE = 7.22 days, $R^2 = 0.62$, PCC = 0.79, and SCC= 0.80 indicating that it successfully captured the complexity of the interaction of the clinical data. After multimodal (Dense Net+ single-layer MLP) MAE= 6.594 days RMSE= 8.56 days $R^2 = 0.453$, PCC=0.716 and SCC=0-.678

- Random Forest Regression was competitive for the given MAE of 5.86 days: 7.69 days, RMSE: 7.69 days, $R^2 = 0.57$, PCC: 0.76, and SCC: 0.76.

- Linear Regression indicated poor forecasting ability compared to other models with a substantially weaker MAE value = 11.44 days, RMSE = 14.07 days, $R^2 = 0.42$, PCC = 0.38, - highlighting the inadequacy of purely linear assumptions for ICU clinical data. Overall, ensemble-based models were superior for simulating ventilator duration due to their capacity to capture non-linear and complicated feature interactions. The reliability of Random Forest and Gradient Boosting in ICU predictive analytics is confirmed by the high correlation measures, which show a substantial agreement between expected and observed outcomes.

6.3 Contributions to the Field

The study provides several meaningful contributions:

- To show the effectiveness and superiority of ensemble regression methods (Gradient Boosting and Random Forest) in predicting ICU ventilator duration.
- It highlights feature-driven interpretability and robustness, which allows clinicians to understand and trust model predictions
- Provides a comparative assessment of classical and modern techniques of regression. illustrating the restrictions of linear models with respect to complex clinical data.
- It highlights feature-driven interpretability and robustness, which allows clinicians to understand and trust model predictions.

6.4 Recommendations for Future Work

Building upon this research, several directions can be taken to enhance model performance. and clinical utility

- **Multimodal Data Incorporation:** Imaging, laboratory, and Clinical notes can further improve predictive accuracy and interpretability.
- **Temporal Modelling:** Longitudinal patient information might improve the accuracy of predictions made by sequential models like LSTM or Transformer models.
- **External Validation:** Testing models on larger and more diverse ICU datasets This will ensure generalizability across a variety of hospitals with different patient populations.
- **Explainable AI Integration:** Provide advanced interpretability techniques such as Most feature attribution approaches, including model-agnostic ones, such as SHAP and LIME, are able to provide clinicians with actionable insights into feature importance and decision rationale.
- **Real-time Deployment:** The research on deployment in ICUs with real-time monitoring and automated alerts might improve patient care and operational efficiency

It has been demonstrated in this work that ensemble-based regression models are valuable, interpretable, and relevant techniques in the forecasting of ventilator duration in people with serious illness. Future work can extend this foundation towards multimodal, temporal, and real-time predictive systems, eventually culminating in better-informed clinical decision making, and resource management.

Bibliography

[1] The impact of covid-19 pandemic among cancer patients. *Int. J. Radiat. Oncol. Biol. Phys.*, 114(3):e91, October 2022.

[2] Sakifa Aktar, Ashis Talukder, Md. Martuza Ahamad, A. H. M. Kamal, Jahidur Rahman Khan, Md. Protikuzzaman, Nasif Hossain, A. K. M. Azad, Julian M. W. Quinn, Mathew A. Summers, Teng Liaw, Valsamma Eapen, and Mohammad Ali Moni. Machine learning approaches to identify patient comorbidities and symptoms that increased risk of mortality in COVID-19. *Diagnostics*, 11(8):1383, 2021.

[3] M. B. Allen. Covid-19, cancer post-pandemic risk, and the radiation oncology physicist. *Journal of Applied Clinical Medical Physics*, 23(5):e13628, 2022.

[4] J. M. Añón et al. Modelo de probabilidad de ventilación mecánica prolongada. *Medicina Intensiva*, 36:488–495, 2012.

[5] Joseph Bae, Saarthak Kapse, Gagandeep Singh, Rishabh Gattu, Syed Ali, Neal Shah, Colin Marshall, Jonathan Pierce, Tej Phatak, Amit Gupta, Jeremy Green, Nikhil Madan, and Prateek Prasanna. Predicting mechanical ventilation and mortality in covid-19 using radiomics and deep learning on chest radiographs: A multi-institutional study. *Diagnostics*, 11(10), 2021.

[6] Jun Bai, Bingjun Li, and Sheida Nabavi. Semi-supervised classification of disease prognosis using cr images with clinical data structured graph. In *Proceedings of the 13th ACM International Conference on Bioinformatics, Computational Biology and Health Informatics*, pages 1–9, 2022.

[7] S. Boyd and L. Vandenberghe. *Convex Optimization.* Cambridge Univ. Press, Cambridge, U.K., 2004.

[8] L. Breiman. Random forests. *Machine Learning*, 45(1):5–32, 2001.

[9] Elena-Victoria Manea (Carneluti), Carmen Duta, Mircea Sebastian Serbanescu, Ilona Georgescu, Stefan-Alexandru Artene, Andreea Denisa Hodorog, Irina Anna-Maria Stoian, Cristina Pană, Stefana Oana Popescu, Ligia Gabriela Tataranu, and Anica Dricu. Risk stratification of covid-19 severity in cancer patients using machine learning algorithms. *Preprints*, March 2025.

[10] Ermanno Cordelli, Paolo Soda, Sara Citter, Elia Schiavon, Christian Salvatore, Deborah Fazzini, Greta Clementi, Michaela Cellina, Andrea Cozzi, Chandra Bortolotto, Luisa Francini, Matteo Tortora, Isabella Castiglioni, Sergio Papa, Diego Sona, and Marco Alì. Machine learning predicts pulmonary long covid

sequelae using clinical data. *BMC Medical Informatics and Decision Making*, 24, 11 2024.

[11] H. Dagher, A. M. Chaftari, I. M. Subbiah, A. E. Malek, Y. Jiang, P. Lamie, B. Granwehr, T. John, E. Yepez, J. Borjan, C. Reyes-Gibby, M. Flores, F. Khawaja, M. Pande, N. Ali, R. Rojo, D. D. Karp, P. Chaftari, R. Hachem, I. I. Raad, and The University of Texas MD Anderson Cancer Center. Long covid in cancer patients: Preponderance of symptoms in majority of patients over long time period. *eLife*, 12:e81182, 2023.

[12] H. Duanmu, T. Ren, H. Li, N. Mehta, A. J. Singer, J. M. Levsky, M. L. Lipton, and T. Q. Duong. Deep learning of longitudinal chest x-ray and clinical variables predicts duration on ventilator and mortality in covid-19 patients. *BioMedical Engineering OnLine*, 21(77), Oct. 2022.

[13] J. H. Friedman. Greedy function approximation: A gradient boosting machine. *Annals of Statistics*, 29(5):1189–1232, 2001.

[14] Amogh Hiremath, Vidya Sankar Viswanathan, Kaustav Bera, Rakesh Shiradkar, Lei Yuan, Keith Armitage, Robert Gilkeson, Mengyao Ji, Pingfu Fu, Amit Gupta, Cheng Lu, and Anant Madabhushi. Deep learning reveals lung shape differences on baseline chest ct between mild and severe covid-19: A multi-site retrospective study. *Computers in Biology and Medicine*, 177:108643, 2024.

[15] Mohammed Innat, Md. Faruque Hossain, Kevin Mader, and Abbas Z. Kouzani. A convolutional attention mapping deep neural network for classification and localization of cardiomegaly on chest x-rays. *Scientific Reports*, 13(1):6247, 2023.

[16] T. Karampitsakos, V. Sotiropoulou, M. Katsaras, P. Tsiri, V. E. Georgakopoulou, I. C. Papanikolaou, E. Bibaki, I. Tomos, I. Lambiri, O. Papaioannou, E. Zarkadi, E. Antonakis, A. Pandi, E. Malakounidou, F. Sampsonas, S. Makrodimitri, S. Chrysikos, G. Hillas, K. Dimakou, N. Tzanakis, N. V. Sipsas, K. Antoniou, and A. Tzouvelekis. Post-covid-19 interstitial lung disease: Insights from a machine learning radiographic model. *Frontiers in Medicine*, 9:1083264, 2023.

[17] M. Kuderer, D. Choueiri, C. Shah, et al. Clinical characteristics, care trajectories, and mortality rate of sars-cov-2 infected cancer patients: A multicenter cohort study. *ResearchGate*, 2021.

[18] D. H. Kwon, J. Cadena, S. Nguyen, K. H. R. Chan, B. Soper, A. L. Gryshuk, J. C. Hong, P. Ray, and F. W. Huang. Covid-19 outcomes in patients with cancer: Findings from the university of california health system database. *Cancer Medicine*, 11(6):1855–1863, 2022.

[19] Y. LeCun, Y. Bengio, and G. Hinton. Deep learning. *Nature*, 521(7553):436–444, May 2015.

[20] T. Linden, F. Hanses, D. Domingo-Fernández, L. N. DeLong, A. T. Kodamullil, J. Schneider, M. J. G. T. Vehreschild, J. Lanznaster, M. M. Ruethrich, S. Borgmann, M. Hower, K. Wille, T. Feldt, S. Rieg, B. Hertenstein, C. Wyen,

C. Roemmele, J. J. Vehreschild, C. E. M. Jakob, M. Stecher, M. Kuzikov, A. Zaliani, H. Fröhlich, and LEOSS study group. Machine learning based prediction of covid-19 mortality suggests repositioning of anticancer drug for treating severe cases. *Artificial Intelligence in the Life Sciences*, 1:100020, December 2021. Erratum in: Artif Intell Life Sci. 2022 Dec;2:100032. doi: 10.1016/j.ailsci.2022.100032.

[21] Swapna Mantha, Jae Jee, Gerald A. Soff, and Michael S. Pessin. Prediction of covid-19 mortality in patients with cancer. *Blood*, 136:29–30, 2020. Epub 2021 Aug 3.

[22] Toshimasa Matsumoto, Shannon Leigh Walston, Michael Walston, Daijiro Kabata, Yukio Miki, Masatsugu Shiba, and Daiju Ueda. Deep learning–based time-to-death prediction model for covid-19 patients using clinical data and chest radiographs. *Journal of Digital Imaging*, 36:178–188, 2023.

[23] S. Mahvash Mohammadi, M. Rumyantsev, E. Abdeeva, D. Baimukhambetova, P. Bobkova, Y. El-Taravi, M. Pikuza, A. Trefilova, A. Zolotarev, M. Andreeva, E. Iakovleva, N. Bulanov, S. Avdeev, E. Pazukhina, A. Zaikin, V. Kapustina, V. Fomin, A. A. Svistunov, P. Timashev, and N. Avdeenko. Post-covid-19 condition prediction in hospitalized cancer patients: A machine learning-based approach. *Cancers*, 17(4):687, 2025.

[24] M. J. Monroy-Iglesias, K. Tremble, B. Russell, C. Moss, S. Dolly, A. Sita-Lumsden, and M. Van Hemelrijck. Long-term effects of covid-19 on cancer patients: the experience from guy's cancer centre. *Future Oncology*, 18(32):3585–3594, 2022.

[25] Saket Navlakha, Shaunak Morjaria, Ruben Perez-Johnston, Andy Zhang, and Ying Taur. Projecting covid-19 disease severity in cancer patients using purposefully-designed machine learning. *BMC Infectious Diseases*, 21(1):391, May 2021.

[26] David J Pinato, Josep Tabernero, Mark Bower, Lorenza Scotti, Meera Patel, Emeline Colomba, Saoirse Dolly, Angela Loizidou, John Chester, Uma Mukherjee, Alberto Zambelli, Alessia Dalla Pria, Juan Aguilar-Company, Diego Ottaviani, Amani Chowdhury, Eve Merry, Ramon Salazar, Alexia Bertuzzi, Joan Brunet, Matteo Lambertini, Marco Tagliamento, Anna Pous, Ailsa Sita-Lumsden, Krishnie Srikandarajah, Johann Colomba, Fanny Pommeret, Elia Seguí, Daniele Generali, Salvatore Grisanti, Paolo Pedrazzoli, Gianpiero Rizzo, Michela Libertini, Charlotte Moss, Joanne S Evans, Beth Russell, Nadia Harbeck, Bruno Vincenzi, Federica Biello, Rossella Bertulli, Raquel Liñan, Sabrina Rossi, Maria Carmen Carmona-García, Carlo Tondini, Laura Fox, Alice Baggi, Vittoria Fotia, Alessandro Parisi, Giampero Porzio, Maristella Saponara, Claudia Andrea Cruz, David García-Illescas, Eudald Felip, Ariadna Roqué Lloveras, Rachel Sharkey, Elisa Roldán, Roxana Reyes, Irina Earnshaw, Daniela Ferrante, Javier Marco-Hernández, Isabel Ruiz-Camps, Gianluca Gaidano, Andrea Patriarca, Riccardo Bruna, Anna Sureda, Clara Martinez-Vila, Ana Sanchez de Torre, Luca Cantini, Marco Filetti, Lorenza Rimassa, Lorenzo Chiudinelli, Michela Franchi, Marco Krengli, Armando Santoro, Aleix Prat, Mieke

Van Hemelrijck, Nikolaos Diamantis, Thomas Newsom-Davis, Alessandra Gennari, Alessio Cortellini, Judith Swallow, Chris Chung, Gino Dettorre, Neha Chopra, Alvin JX Lee, Christopher CT Sng, Yien Ning Sophia Wong, Myria Galazi, Sarah Benafif, Palma Dileo, Grisma Patel, Anjui Wu, Alasdair Sinclair, Gehan Soosaipillai, Eleanor Jones, Amanda Jackson, Martine Piccart, Oriol Mirallas, Anna Carbó, Isabel Garcia, Rachel Wuerstlein, Ricard Mesia, Clara Maluquer, Francesca D'Avanzo, Giuseppe Tonini, Salvatore Provenzano, Valeria Tovazzi, Corrado Ficorella, Paola Queirolo, Raffaele Giusti, Francesca Mazzoni, Federica Zoratto, Marco Tucci, Rossana Berardi, Annalisa Guida, Sergio Bracarda, and Maria Iglesias. Prevalence and impact of covid-19 sequelae on treatment and survival of patients with cancer who recovered from sars-cov-2 infection: evidence from the oncovid retrospective, multicentre registry study. *The Lancet Oncology*, 22(12):1669–1680, 2021.

[27] S. Purkayastha, Y. Xiao, Z. Jiao, R. Thepumnoeysuk, K. Halsey, J. Wu, T. M. L. Tran, B. Hsieh, J. W. Choi, D. Wang, M. Vallières, R. Wang, S. Collins, X. Feng, M. Feldman, P. J. Zhang, M. Atalay, R. Sebro, L. Yang, Y. Fan, and H. X. Bai. Machine learning-based prediction of covid-19 severity and progression to critical illness using ct imaging and clinical data. *Korean Journal of Radiology*, 22(7):1213–1224, 2021.

[28] J. Saltz, M. Saltz, P. Prasanna, R. Moffitt, J. Hajagos, E. Bremer, J. Balsamo, and T. Kurc. Stony brook university covid-19 positive cases [data set]. *The Cancer Imaging Archive (TCIA)*, 2021. Available at: https://doi.org/10.7937/TCIA.BBAG-2923.

[29] Sahil Sethi, Sushil Shakyawar, Athreya S. Reddy, Jai Chand Patel, and Chittibabu Guda. A machine learning model for the prediction of covid-19 severity using rna-seq, clinical, and co-morbidity data. *Diagnostics*, 14(12), 2024.

[30] Chin Song, Ros Bustamam, Gin Gan, Marniza Saad, Nur Satar, Alagu Ramasamy, I-Ching Sam, Yek-Ching Kong, Harenthri Rajah, Yoke-Fun Chan, Jolene Fu, Cheng Siang Tan, Mahmoud Danaee, Cheng Yip, Carla Gils, and Nirmala Bhoo-Pathy. Antibody response following covid-19 vaccination in malaysian cancer patients and healthy individuals. *Cureus*, 16, 11 2024.

[31] Oliver Stritzel and Renata Georgia Raidou. Predicting, analyzing and communicating outcomes of covid-19 hospitalizations with medical images and clinical data. In Thomas Schultz and H.-Y. Wu, editors, *Proceedings of the Eurographics Workshop on Visual Computing for Biology and Medicine (VCBM)*, Vienna, Austria, 2022. The Eurographics Association. Short Paper.

[32] A. U. Wells, A. Devaraj, and S. R. Desai. Interstitial lung disease after covid-19 infection: A catalog of uncertainties. *Radiology*, 298(1):1–3, 2021.

[33] C. Zang, Y. Hou, E. J. Schenck, Z. Xu, Y. Zhang, J. Xu, J. Bian, D. Morozyuk, D. Khullar, A. S. Nordvig, E. A. Shenkman, R. L. Rothman, J. P. Block, K. Lyman, Y. Zhang, J. Varma, M. G. Weiner, T. W. Carton, F. Wang, and R. Kaushal. Identification of risk factors of long covid and predictive modeling in the recovered ehr cohorts. *Communications Medicine*, 4(1), 2024.