PREDICTION OF COVID-19 RISK OF A PERSON BY ANALYZING COMPUTED TOMOGRAPHY IMAGES USING CONVOLUTIONAL NEURAL NETWORKS

A THESIS SUBMITTED TO
THE GRADUATE SCHOOL OF INFORMATICS
OF
MIDDLE EAST TECHNICAL UNIVERSITY

BY

KAAN TOPÇU

IN PARTIAL FULFILLMENT OF THE REQUIREMENTS
FOR
THE DEGREE OF MASTER OF SCIENCE
IN
INFORMATION SYSTEMS

JANUARY 2024

Approval of the thesis:

PREDICTION OF COVID-19 RISK OF A PERSON BY ANALYZING COMPUTED TOMOGRAPHY IMAGES USING CONVOLUTIONAL NEURAL NETWORKS

submitted by **KAAN TOPÇU** in partial fulfillment of the requirements for the degree of **Master of Science in Information Systems Department, Middle East Technical University** by,

Prof. Dr. Banu Günel Kılıç
Dean, Graduate School of **Informatics** ____________

Prof. Dr. Altan Koçyiğit
Head of Department, **Information Systems** ____________

Assist. Prof. Aybar Can Acar
Supervisor, **Medical Informatics, METU** ____________

Examining Committee Members:

Prof. Dr. Tuğba Taşkaya Temizel
Data Informatics, METU ____________

Assist. Prof. Aybar Can Acar
Medical Informatics, METU ____________

Prof. Dr. Tunca Doğan
Computer Engineering, Hacettepe University ____________

Date:

I hereby declare that all information in this document has been obtained and presented in accordance with academic rules and ethical conduct. I also declare that, as required by these rules and conduct, I have fully cited and referenced all material and results that are not original to this work.

Name, Surname: Kaan Topçu

Signature :

ABSTRACT

PREDICTION OF COVID-19 RISK OF A PERSON BY ANALYZING COMPUTED TOMOGRAPHY IMAGES USING CONVOLUTIONAL NEURAL NETWORKS

Topçu, Kaan
M.S., Department of Information Systems
Supervisor: Assist. Prof. Aybar Can Acar

January 2024, 56 pages

In this thesis, 4 main research questions are answered to evaluate the performance of convolutional neural networks (CNN) in predicting Covid-19 risk by using computed tomography (CT) images. The CT images by Yang et al., 2020[1] are utilized for this study. It contains 349 CT images labeled as being positive for Covid-19 from 216 patient and 397 CT images that are negative. Different CNNs like VGG-16, ResNet-18, ResNet-50, DenseNet-121, DenseNet-169, EfficientNet-B0, and EfficientNet-B1 are experimented and evaluated accordingly.

The first research question investigates the performance of the CNNs without pre-training them. The second one evaluates the effect of transfer learning for each CNN. The third research question studies the impact of source dataset's domain used for transfer learning. Finally, whether the performance of the networks can be maintained or improved by training the networks partially is analyzed.

All four research questions are evaluated by comparing the accuracy, F1-score and AUC values.

Keywords: convolutional neural network, Covid-19, image classification, computed tomography

ÖZ

KİŞİLERİN COVİD-19 RİSKİNİN, BİLGİSAYARLI TOMOGRAFİ GÖRÜNTÜLERİNİN EVRİŞİMLİ SİNİR AĞLARI KULLANILARAK TAHMİNLENMESİ

Topçu, Kaan
Yüksek Lisans, Bilişim Sistemleri Bölümü
Tez Yöneticisi: Dr. Öğr. Üyesi. Aybar Can Acar

Ocak 2024 , 56 sayfa

Bu tezde, evrişimsel sinir ağlarının Covid-19 riskini tahmin etmedeki performansını değerlendirmek için 4 ana araştırma sorusu yanıtlanmaktadır. Bu çalışma için Yang ve diğerleri[1] tarafından elde edilen bilgisayarlı tomografi görüntüleri kullanılmıştır. Çalışma, Covid-19 pozitif olarak etiketlenmiş 216 hastadan gelen 349 CT görüntüsü ve negatif olan 397 CT görüntüsünü içermektedir. Çalışmada VGG-16, ResNet-18, ResNet-50, DenseNet-121, DenseNet-169, EfficientNet-B0 ve EfficientNet-B1 gibi farklı ağlar kullanılmış ve bunların performansları değerlendirilmiştir.

İlk araştırma sorusu, evrişimsel sinir ağlarının ön eğitim olmadan performansını incelemektedir. İkinci soru, her evrişimsel sinir ağı için öğrenme transferinin etkisini değerlendirir. Üçüncü araştırma sorusu, öğrenme transferi için kullanılan kaynak veri kümesinin etkisini incelemektedir. Son olarak, ağların performansının kısmi eğitimle sürdürülüp sürdürülemediğini veya iyileştirilip iyileştirilemediğini analiz edilmektedir.

Dört araştırma sorusuna cevap verebilmek için geliştirilen modeller geliştirildikten sonra doğruluk değeri, F1 değeri ve AUC değerlerini karşılaştırılarak değerlendirilmiştir.

Anahtar Kelimeler: evrişimsel sinir ağları, Covid-19, görüntü sınıflandırma, bilgisayarlı tomografi

to my family

ACKNOWLEDGMENTS

I would like to express my sincere gratitude to my thesis supervisor Asst. Prof. Aybar Can Acar for his guidance and comments throughout this study.

I am also grateful to my fiance, Bilge Yalınkılıçlı who is always with me when I need her and who always makes me happy.

I would like to thank my mother Gülsün Topçu, my father Mustafa Topçu and my brother Fatih Topçu for their encouragement and belief in me throughout my life.

TABLE OF CONTENTS

LIST OF TABLES

TABLES

LIST OF FIGURES

FIGURES

LIST OF ABBREVIATIONS

CT	Computed Tomography
CNN	Convolutional Neural Network
ReLU	Rectified Linear Unit

CHAPTER 1

INTRODUCTION

1.1 Background and Motivation

In late 2019, the the novel coronavirus, SARS-CoV-2, emerged and led to a global health crisis spreading rapidly across the borders and affecting millions of lives. With the evolution of the pandemic, the need for accurate and efficient methods for early detection and risk assessment of Covid-19 remained a paramount issue. As a diagnostic tool in fighting with Covid-19, computed tomography (CT) imaging has emerged since it provides detailed visualizations of lung abnormalities associated with the disease. Convolutional Neural Networks (CNNs) have shown great success in medical image analysis tasks leveraging the power of artificial intelligence and deep learning. In this context, this masters' thesis aims to develop and validate a state-of-the-art CNN-based framework for the prediction of Covid-19 risk in individuals by analyzing Computed Tomography images.

One of the crucial points of effective disease management and resource allocation is the rapid and accurate identification of individuals at high risk of Covid-19. Traditional methods like polymerase chain reaction (PCR) tests played a crucial role in detecting active infections but they haven't always provided timely and definite results. Computed tomography imaging has become a complementary tool to such PCR tests, enabling visualizations of the abnormalities associated with Covid-19. CT images have aided in the early detection of the disease and provided valuable insights into the progression of the Covid-19 in affected individuals thanks to rich information that they have contained. [3, 4, 5].

The field of medical imaging analysis has taken a huge step and has been revolutionized by Convolutional Neural Networks, which demonstrated exceptional capabilities in image classification, segmentation, and object detection etc. The nature of CNNs allows them to learn hierarchical and discriminative features from the input images and makes them useful tools for complex medical image analysis. This has resulted in an increase in the use of CNN-based approaches in the automatic diagnosis of various medical conditions such as lung diseases and infectious diseases. It is possible to develop a robust and accurate predictive tool for assessing an individual's risk of Covid-19 by training a CNN model on a large dataset of CT images from Covid-19 patients.

1.2 Objectives and Research Questions

This masters' thesis seeks to address the need for reliable COVID-19 risk prediction by proposing an innovative deep learning-based approach. The primary objectives of this research are as follows:

- To investigate different Convolutional Neural Network (CNN) architectures optimized for COVID-19 risk prediction using Computed Tomography (CT) images, and compare their performance in terms of accuracy, sensitivity, specificity, and computational efficiency.

There are different CNN architectures used in the medical image analysis area, and these will be explained in detail in the following sections. This research question focuses on the exploration of multiple CNN architectures for Covid-19 risk prediction based on CT images. The study will focus on designing a novel architecture or modifying the existing ones to enhance their performance in detection of infectious individuals and identifying the ones at high risk of Covid-19. The CNN models will be optimized to extract relevant features from CT scans and increase the performance of predictions of Covid-19 infection and severity of the disease. The research will evaluate the performance of each CNN architecture using a CT image dataset and employing different metrics such as accuracy, sensitivity, specificity, and other relevant measures. In addition to those, the computational efficiency and run time of the proposed models ensure their applicability in the real-world settings. This research aims to identify the most effective and efficient model for Covid-19 risk prediction by developing and comparing different CNN architectures.

- To test whether transfer learning influences the performance of convolutional neural networks in the context of Covid-19 risk prediction using Computed Tomography (CT) images.

This research question aims to investigate the impact of transfer learning on the performance of convolutional neural networks (CNNs) when applied to the task of Covid-19 risk prediction using CT images. Transfer learning involves the use of knowledge gained from the pre-trained models on a different dataset to improve the performance on the target task. The research will explore whether pre-training CNN models on a large dataset, such as ImageNet, and fine-tuning them on the Covid-19 CT image dataset leads to a better performance compared to training the CNNs from scratch. The research will analyze different CNN architectures such as VGG16, ResNet50, DenseNet121, EfficientNet-B1, and assess the influence of transfer learning on their ability to predict Covid-19 risk. The findings from the research will provide insights about the effectiveness of transfer learning as a tool to increase the performance of CNNs for Covid-19 risk prediction and its possible implications for other medical imaging tasks.

- To investigate the impact of using a source dataset in the similar domain as the target dataset (e.g., using LUNA16 (Lung Nodule Analysis 2016) dataset for Covid-19 CT dataset) for transfer learning in Covid-19 risk prediction using computed tomography (CT) images and assess whether this approach improves the performance of the convolutional neural network models.

This research question explores the potential benefits of using a source dataset, such as LUNA16 (Lung Nodule Analysis 2016) dataset, which is in the same domain as the target Covid-19 risk prediction dataset (CT images). The effects of transfer learning on ImageNet dataset are analyzed with the previous research questions. In addition to that, the study will investigate the transfer learning process, where pre-trained

CNN models from the source dataset will be fine-tuned and modified accordingly for Covid-19 risk prediction. The outcomes of CNN models will be compared with the ones using transfer learning with different domain. This analysis is expected to provide insights for the potential advantages of having prior knowledge in the medical domain and its potential implications for practical implementation in the medical imaging domain.

- Can the performance of deep networks be maintained or improved by partially training only the last layers, while freezing the first layers, for various convolutional neural network architectures in the context of Covid-19 risk prediction using Computed Tomography (CT) images?

This research question aims to explore the effects of partial training in deep neural networks. To eliminate the computation costs from the high parameter counts in the deep architectures of DenseNet121, DenseNet169, ResNet50, EfficientNet-B0, and EfficientNet-B1, partial training can be a good alternative. The investigation will examine whether training only last layers while freezing the earlier layers would can lead to increased performance at the cost of reduced efficiency in the model. The selection of layers to be trained in different models will be based on the unique topology of each network. The research question would evaluate the impact of this training methodology on the performance of the models and provide insights about the partial training for the medical imaging task.

1.3 Scope and Significance of the Study

This masters' thesis focuses on developing and validating a state-of-the-art Convolutional Neural Network (CNN) framework to predict Covid-19 risk in individuals by analyzing computed tomography (CT) images. The study will leverage the power of deep learning and medical imaging to accurately assess the Covid-19 risk of individuals by using their CT scans. In the research, different CNN architectures like VGG-16, ResNet-18, ResNet-50, DenseNet-121, DenseNet-169, EfficientNet-B0, and EfficientNet-B1 will be explored, analyzed, and optimized for Covid-19 risk prediction. The dataset used for training and evaluation includes a diverse set of

CT images from Covid-19 patients, as well as non-Covid-19 patients. The Covid-19 dataset has 349 CT images from 216 patients. The utility of the dataset has been confirmed by a senior radiologist in Tongji Hospital, Wuhan, China [1, 5]. The study involves data preprocessing, CNN model development, model comparison, fine-tuning, and evaluation of the models. Performance metrics such as accuracy, sensitivity, specificity, F1 score, and computational efficiency will be considered to assess the effectiveness of the developed CNN-based approach in a more comprehensive way.

The development of an accurate and efficient CNN-based approach would contribute to Covid-19 risk prediction using CT images and it has potential implications:

1. Early Detection and Risk Assessment: The proposed CNN-based approach has the potential to detect Covid-19 in individuals in early phases of the disease even before the symptoms appear. By having such an early detection system, the healthcare professionals can prioritize high-risk patients for further testing, isolation, early intervention etc. This would help prevent the spread of the virus and increase patient management capabilities.

2. Resource Allocation: By detecting the high-risk patients, it is possible to direct the limited medical resources to them. This would help optimize allocation of hospital beds, medical staff, medical supplies in Covid-19 cases with an accurate risk prediction using CT images.

3. Complementary Diagnostic Tool: Computed Tomography imaging is known to provide valuable insight into lung abnormalities related with Covid-19. Integration of this CNN-based risk prediction approach to existing diagnostic methods like PCR tests would increase the capability to fight with Covid-19 and enhance the accuracy of Covid-19 diagnosis.

4. Generalizability: The research will evaluate the performance of different CNN architectures to assess the generalizability across different datasets and in medical domain. The findings of the research would provide guidance for selecting the most effective CNN model for Covid-19 risk prediction.

5. Potential for Future Pandemics: The developed CNN-based approach would serve as a base point for predicting risk in future pandemics. The transferability of the approach to other medical imaging problems would help to apply it in broader domains beyond Covid-19.

Overall, a successful and validated CNN-based approach for Covid-19 risk prediction using CT images can have significant benefits for public health and clinical practices. By providing reliable and timely Covid-19 risk assessments, the research would assist better patient management, medical resource allocation, and pandemic strategies. Moreover, the findings can contribute to advancing the field of medical imaging, deep learning, and its implications in medical domain, building new approaches by using artificial intelligence in healthcare for disease diagnosis and risk assessment.

1.4 Overview of the Thesis Structure

The thesis is structured into several chapters, each contributes to the overall understanding and accomplishment of the objectives. Chapter 2 presents a comprehensive literature review on Covid-19 diagnosis, risk assessment, medical imaging modalities, convolutional neural networks. Chapter 3 dives deep into the methodology, data preprocessing, evaluation metrics, training process and hyperparameter settings. Chapter 4 presents and gives details about the experimental results and discussions, keynoting the performance of different CNN models with their limits and strengths. Chapter 5 finalizes the thesis with the summary of the findings, contributions, future research and improvement points.

In conclusion, the thesis aims to contribute to the literature covering the methods and studies used to fight with Covid-19 by providing an accurate and efficient CNN-based tool for risk prediction. By leveraging the power of Convolutional Neural Networks and medical imaging, the research aims to make accurate and timely predictions for individuals at high risk of Covid-19.

CHAPTER 2

LITERATURE REVIEW

2.1 Overview of COVID-19 and Its Impact

The coronavirus disease, Covid-19 has started in the late 2019 and rapidly evolved into a global pandemic and affected millions of lives and transformed the societies worldwide[6]. This section provides an overview of Covid-19, its expressions and effects, transmission, and the impact it created on society in different aspects.

As of 2023, October 12, there have been 771.191.203 confirmed cases of Covid-19, including 6.961.014 deaths, reported to World Health Organization (WHO)[7].

Covid-19 primarily affects the respiratory system and can cause mild to severe illness in individuals. Fever, couch, fatigue, breath difficulty, and loss in taste and smell senses are the common symptoms people experience[8]. Some infected individuals may remain asymptomatic or experience non-typical symptoms, which makes it difficult to detect and diagnosis earlier[9, 10]. The virus mainly spreads through respiratory droplets when an infectious person coughs, sneezes, or talks[11]. It can also be transmitted through contaminated surfaces in case of close contact with the surface[12].

The unpredicted scale and severity of the Covid-19 pandemic affected healthcare systems globally. Hospitals, healthcare facilities and professionals have suffered from the huge volume of patient admissions, shortage of medical supplies, equipment etc[13]. Medical imaging, especially computed tomography (CT) scans, surged as an important diagnostic tool for detecting Covid-19.

While CT scans have been valuable tools in Covid-19 diagnosis, they also have some limitations and challenges. CT scans expose patients to ionizing radiation, and the availability of CT scanners are limited in some hospitals and even cities.

Covid-19 has had huge consequences on public health and socioeconomic aspects. Governments and health organizations applied strict precautions such as lockdowns[14], travel restrictions[15], social distancing to decrease the rate of the spread of the virus. While these precautions were necessary, they have led to significant disruptions in economies, education systems, mental health, and overall well-being of society. The pandemic emphasizes the importance of global collaboration, scientific research, and rapid development strategies in medical domain like effective vaccine development and treatment[16].

The world has faced an unpredicted global health crisis and impacted individuals and nations worldwide. According to (Miller, 6 *of the worst pandemics in history*)[17] Covid-19 pandemic is one of the 6 worst pandemics in history. The integration of medical imaging, particularly CT scans, has played a vital role in fighting the pandemic with the detection and management of Covid-19. However, it is essential to continue developing new methods and optimizing the current imaging techniques while considering the challenges and limitations to make timely and accurate predictions. The ongoing effort in research, public health, advancements in medicine and technology will help for effective strategies to prevent or decrease the effects of future pandemics.

2.2 Medical Imaging

Medical imaging comprises a range of techniques which are used to visualize the internal structures of the human body for diagnosis and research purposes. As explained before, CT scans, which are a measure of medical imaging, were widely used during Covid-19 pandemic. This section provides an overview of medical imaging, including its methods, advantages, disadvantages, and practical applications.

2.2.1 Methods of Medical Imaging

There are different methods used for medical imaging based on their purpose and the context in which they are used.

1. X-ray Imaging: X-ray imaging uses electromagnetic radiation to create two-dimensional images of the body. It is commonly used to examine bones and detect conditions such as fractures, tumors, and lung diseases.

2. Computed Tomography (CT): CT imaging utilizes a series of X-ray images taken from different angles to generate detailed cross-sectional images of the body. CT scans are particularly useful for examining the brain, chest, abdomen, and bones, providing enhanced visualization of anatomical structures.

3. Magnetic Resonance Imaging (MRI): Magnetic Resonance Imaging employs a powerful magnetic field and radio waves to generate detailed images of the body's organs and tissues. It is especially effective for visualizing soft tissues, such as the brain, spinal cord, muscles, and joints.

4. Ultrasound Imaging: Ultrasound uses high-frequency sound waves to create real-time images of organs, blood vessels, and developing fetuses. It is safe, non-invasive, and widely used in obstetrics, cardiology, and various other medical fields.

5. Nuclear Imaging: Nuclear medicine techniques involve the administration of small amounts of radioactive substances, called radiotracers, which emit gamma rays. These rays are detected by specialized cameras to create images that provide functional and metabolic information about organs and tissues.

2.2.2 Pros and Cons of Medical Imaging

There are some advantages that healthcare professionals apply medical imaging while they have some limitations and disadvantages.

1. Pros:

   - Early Detection and Diagnosis: Medical imaging allows for the early detection and diagnosis of various conditions, enabling timely interventions and improved patient outcomes. This helped a lot to people who are asymptomatic or experience non-typical symptoms for Covid-19.

   - Non-Invasive or Minimally Invasive: Many imaging techniques are non-invasive, meaning they do not require surgical procedures, reducing patient discomfort and risks.

   - Visualizing Internal Structures: Medical imaging provides visualizations of internal structures that may not be accessible through physical examination alone. For some Covid-19 cases, the exact diagnosis can only be made by medical imaging modalities.

2. Cons:

   - Ionizing Radiation: Techniques like X-ray and CT imaging expose patients to ionizing radiation, which carries potential health risks, particularly with repeated or excessive exposure.

   - Limited Accessbility and Cost: Certain imaging modalities, such as MRI and advanced nuclear medicine techniques, may have limited availability due to their high cost and specialized requirements. Advanced imaging technologies can be expensive to acquire, operate, and maintain, limiting their accessibility in certain healthcare settings. For the times like pandemic, in which the medical resources are hardly available, the cost of the medical imaging is more than regular times.

2.2.3 Applications of Medical Imaging

There are many practical applications of medical imaging used for different purposes. One of the most used areas is screening and preventive medicine. Certain medical imaging techniques like mammography for breast cancer screening or CT scans or lung cancer screening are utilized for early detection and even preventive healthcare. Medical imaging techniques can be used for diagnostic of a disease. They play crucial roles in diagnosing various conditions, fractures, tumors, cardiovascular diseases, neurological disorders, and respiratory conditions. They can also be used to help in planning surgeries, monitoring treatment responses and assessing post-treatment outcomes. Apart from the cases in which medical imaging techniques are applied directly to individuals, they are highly valuable for conducting research studies and investigating the nature of the diseases. By so, new drugs and treatment modalities are developed and medical knowledge advances in overall.

2.2.4 Conclusion

Medical imaging includes various methods that have transformed healthcare. These methods allow doctors to see inside the body without surgery, helping with diagnosis, treatment planning, and research. While each method has strengths and weaknesses, together, they've greatly improved patient care and medical progress. Ongoing research and better technology will keep pushing medical imaging forward, making patient care and disease management even better.

2.3 What is Computed Tomography (CT)

Computed tomography is one of the medical imaging techniques which uses X-ray technology and computer processing to generate cross-sectional images of the body. CT scanning provides three dimensional visualizations of the internal body parts which allows them to be used for detection, diagnosis, and monitoring of a wide range of medical conditions. To analyze lungs, CT screening provides more detailed information than conventional X-rays and offers better care for patients.

2.3.1 Principle and Process of CT Imaging

CT scans use X-rays to detect and record the amount of radiation absorbed by different tissues from different angles around the body. The CT scanner takes a series of X-ray projections from different angles as it moves around the body. This rotation can be either a 360-degree or partial based on the CT scanning technique. X-rays pass through the body and the detector records the X-ray during each rotation and is used to reconstruct detailed cross-sectional images of the body. Creating cross-sectional images from projection data is called image reconstruction and is one of the fundamental processes of CT imaging. It commonly uses Radon transform as a mathematical technique which uses different algorithms like filtered back projection or iterative reconstruction methods. The resulting images provide a comprehensive view to allow the detection and diagnosis of the abnormalities in the lungs.

2.3.2 Advantages of CT Imaging for Lung Evaluation

There are numerous reasons why CT imaging is preferred over other medical imaging methods which can be listed as follows:

- High-resolution Imaging: CT scanners can provide images with high-resolution and detailed visualizations of lung structures. This is very helpful in cases where lung abnormalities are subtle.

- Multiplanar Imaging: CT images provide images from multiple angles and planes which allow professionals to get a more comprehensive view of the internal structure of the body. It facilitates accurate diagnosis and treatment planning.

- Contrast Enhancement: Contrast enhancing agents can be applied to CT imaging. This way, highlighting the tissues that otherwise would be difficult to distinguish from their surroundings is possible.

- Minimally Invasive Procedures: CT scans help professionals apply minimally invasive procedures like lung biopsies and needle aspirations more accurately. This reduces the chances of problems during the procedure and makes it easier to detect and diagnose the medical conditions.

2.3.3 Conclusion

Computed Tomography (CT) is a powerful medical imaging technique that offers detailed and comprehensive assessment of lungs. The advantages mentioned above make it an excellent choice for lung imaging. CT plays a vital role in diagnosing and monitoring lung diseases, guiding interventions and assisting in treatment planning in many cases. Despite its numerous advantages, it is important to consider other medical imaging modalities based on the requirement and context of the medial issue.

2.4 Role of Medical Imaging in COVID-19 Diagnosis and Risk Assessment

Medical imaging, particularly chest imaging techniques like X-ray and computed tomography (CT) has played a crucial role in fighting with Covid-19. These medical imaging modalities provide valuable understanding of the pulmonary effect of the disease by aiding in early detection, monitoring the progression of the disease, evaluating the adverse effects of the disease on the lungs. This section explains the role of medical imaging in Covid-19 diagnosis and risk assessment.

2.4.1 Early Detection and Diagnosis

Different methods are used to detect and diagnose Covid-19 through pandemic and medical imaging is one of those modalities. Chest X-ray and CT scans reveal the abnormalities associated with Covid-19 like ground-glass opacities (Health.com, n.d.)[18], consolidations (Bhatt, Ganatra, Kotecha, 2021)[19], and bilateral lung involvement (Sun et al., 2020)[20]. These imaging findings, combined with the clinical test results, and stages of exposure to the disease helps accurate and timely diagnosis of Covid-19. Based on this accurate and timely diagnosis, proper isolation techniques can be applied, and the spread of the disease can be prevented.

2.4.2 Differential Diagnosis

It is very vital to differentiate Covid-19 from other respiratory infections with similar clinical findings. The distinctive radiographic patterns observed in Covid-19 by applying medical imaging techniques such as peripheral and multifocal ground-glass opacities, helps healthcare professionals to distinguish it from other viral or bacterial pneumonias. This differentiation between Covid-19 and other ones is critical for appropriate steps to triage, treatment, and infection control purposes.

2.4.3 Disease Monitoring and Progression

Not only detection of Covid-19 but also monitoring the progression of it is very important for the sake of individuals' health. Vital signs, such as temperature, respiratory rate, heart rate, blood pressure, hematological and biochemistry parameters, and signs and symptoms of venous or arterial thromboembolism, should be monitored (World Health Organization, 2023). Repeating chest CT scans enables clinicians to evaluate the progress and evolution of the disease, identify complications, and take precautions based on those. Continuous monitoring provides valuable information about the success of the therapy applied and guides for escalating or de-escalating of the interventions.

2.4.4 Assessment of Disease Severity

Medical imaging provides essential information regarding the severity of Covid-19 and its short- and long-term effects on the lungs. CT images allow professionals to interpret lung abnormalities, such as lung lobes, consolidations, or ground-glass opacities, and evaluate their extent and distribution within the lungs. Additionally, this imaging plays a crucial role in assessing the progression of the disease over time and the effectiveness of treatment strategies, enabling healthcare professionals to make informed decisions about patient care and resource allocation. Furthermore, by monitoring changes in lung conditions through medical imaging, appropriate interventions and rehabilitation programs can be addressed to Covid-19 patients.

2.4.5 Prognostic Value and Risk Assessment

Imaging findings obtained through CT scans have demonstrated prognostic value in Covid-19[21, 22]. After assessing the severity of the disease, it is easier to determine the level of risk, prescribe appropriate care, and manage resources effectively, including decisions regarding the need for intensive care unit admission or hospitalization. Furthermore, mortality rates can be reduced, and overall patient outcomes can be improved with the ability to predict the course of the disease based on these imaging findings. This targeted approach in patient management can also help in allocating limited healthcare resources more efficiently, especially during times of increased demand or healthcare crises.

2.4.6 Follow-up and Recovery Assessment

Medical imaging plays a crucial role in the follow-up and assessment of patients recovering from Covid-19. It enables the tracking of the resolution of lung abnormalities and the assessment of recovery progress. These findings are invaluable in making decisions about when to discontinue isolation, when it's safe to return to work, and monitoring for potential complications. Additionally, the data from these scans contribute to ongoing research efforts aimed at improving our understanding of the virus and refining treatment protocols.

2.4.7 Conclusion

Medical imaging, particularly chest CT, is valuable in many aspects of the fight against Covid-19. These images provide valuable insights into the disease, enabling early detection, diagnosis, and risk evaluation. However, it is still important not to rely solely on CT scans and to consider traditional clinical methods. All these factors significantly contribute to increasing humanity's ability to combat Covid-19.

2.5 Existing Approaches and Techniques for COVID-19 Risk Prediction

Considering the advantages and applications of medical imaging, particularly CT scans, they are widely used during the pandemic to predict Covid-19 risk of the individuals. Both as an area of research and practical use, predicting the risk of Covid-19 infection and disease severity has been crucial. In addition to medical imaging modalities, different approaches and techniques have been employed to develop models and algorithms for risk prediction. This section explains the existing approaches and related works in Covid-19 prediction.

2.5.1 Machine Learning-Based Risk Prediction

Different machine learning techniques have been used for Covid-19 risk prediction. Those approaches train models on large datasets containing clinical and demographic information to detect patterns and predict risk. They frequently focused on the recourse allocation since the demand for all types of medical resources was extremely high and appropriate management was necessary during the pandemic. Some notable projects include: Moulaei et al. (2020) developed different machine learning algorithms to predict Covid-19 mortality using the patient's data at the first time of admission and choose the best algorithm as a predictive tool. In another project, Yan et al., (2020)[23] developed a machine learning model to support decision making and logistical planning in healthcare systems. For this study a database of blood samples from 485 infected patients in the region of Wuhan, China used to identify crucial predictive biomarkers of disease mortality. In their article they suggest a simple and operable decision rule to quickly predict patients at the highest risk, allowing them to be prioritized and potentially reducing the mortality rate. Gupta et al. (2020)[24] examined the performance of 22 prognostic models for Covid-19 using data of patients in the intensive care unit. The models were compared by their ability to predict clinical deterioration and mortality using information available at the time of admission. By this study they showed two different models had the highest prediction performance for predicting deterioration over 24 hours and within 14 days from admission. Kamran et al. (2020) [25] created and validated a simple and transferable machine learning model for patients with Covid-19 by using their electronic health record data to predict their clinical deterioration. They used a US hospital for model training and validation and tested their model on patients' data who are admitted to another 12 US hospitals. With this study they identified low-risk patients and high-risk patients and calculated how many bed days hospitals could save per patient if low-risk patients were discharged early.

2.5.2 Radiomics-Based Risk Prediction

Radiomics-based risk prediction involves extracting a large number of quantitative features from medical images, such as CT scans, and MRI images, X-rays. These features is utilized to achieve details, patterns, and textures within the images which can not be detected by the human eye. Radiomics-based risk prediction models for Covid-19 assessment are constructed, combining these features with machine-learning models and clinical and demographic data. Notable works in this field include:" models are built for Covid-19 assessment. Notable works include: Li et al. (2020)[4] aimed to develop an automated system to detect Covid-19 using chest CT scans and assess its performance. They created a deep learning model which extracts visual features from chest CT scans to identify Covid-19. They used 3322 patients' CT scans collected between August 2016 and February 2020. They also measured how well their model is in distinguishing Covid-19 and other lung conditions. Similar to Li et al.[4], Zhang et al. (2020)[3] developed an AI system using data from 3777 patients to diagnose the complications of Covid-19. With the AI system it is desired to offer accurate clinical prognosis combining with clinical data. Its purpose is to provide rapid diagnosis and assist healthcare professionals in making timely treatment decisions and allocating resources effectively. The AI system has been made available globally to support clinicians in their efforts to fight Covid-19.

2.5.3 Clinical Scoring Systems

Clinical scoring systems are the third method used for Covid-19 predictions. These are the tools used to assess, quantify, and predict the clinical aspects of patients' conditions. By evaluating a set of criteria or variables, numerical scores are assigned based on the results. The total score is then used to determine the severity of the disease, predict outcome, or identify risk factors. Notable works include:

Knight et al. (2020)[26], created and validated a practical Covid-19 risk assessment tool, the 4C Mortality Score to predict mortality rate of patients admitted to hospital. They trained the tool with the data of patients admitted to UK hospitals between February and May 2020. Then validated it the ones between May and June 2020. The 4C Mortality Score includes eight variables with high scores mean higher mortality rate. Even if its applicability is ambiguous, it might help patient categorization with some improvements.

In another paper, Sun et al. (2021) [27], developed an automated scoring system, known as Covid-19 Acuity Score (CoVA). The score states the likelihood of hospitalization, illness level, and likelihood of death within 7 days for patients. The study included 9381 patients' data and 30 predictors. The top predictors are age, diastolic blood pressure, blood oxygen saturation, COVID-19 testing status, and respiratory rate. This scoring system helps triaging patients and optimizing patient care. All 3 modalities are exemplified with just a few examples from the existing literature of Covid-19 risk prediction. The field is rapidly evolving with new publications and new studies.

2.6 Related Studies on CT Image Analysis for COVID-19 Risk Prediction

As mentioned above, radiomics-based risk prediction methods have been highly used for Covid-19 risk prediction. Since this thesis focuses on the risk prediction of Covid-19 by using CT images, some research in the literature is mentioned in this section. Researchers have applied different approaches to extract the features from the CT images and build predictive models. Some notable studies that used CT image analysis for predicting Covid-19 is as follows:

Song et al. (2020)[28] developed a deep learning-based model for Covid-19 risk prediction using CT images from 88 Covid-19 patients, 100 bacterial pneumonia patients, and 86 healthy individuals. Their model highlighted key features like ground-glass opacities which help medical professionals for diagnosis and successfully distinguished patients with Covid-19 and bacterial pneumonia.

Zheng et al. (2020) [29] focused on rapid and accurate diagnosis of Covid-19 using chest CT scans. They developed a weakly-supervised deep learning model which does not need lesion annotations. For each patient the lung region was segmented, and deep neural network model predicted the Covid-19 infection probability with a threshold of 0.5. Their model processes a single patient's CT volume in nearly 2 seconds which provides a fast and accurate way for predicting Covid-19.

Ardakani et al. (2020)[30], studied convolutional neural networks in distinguishing between Covid-19 and other pneumonia cases to increase the level of diagnostic accuracy. They used 1020 CT slices from 108 Covid-19 patients and 86 patients with non-Covid pneumonia. They employed 10 CNN models, and ResNet-101 and Xception had the best performance. ResNet-101 achieved an accuracy over 99.5% in distinguishing Covid-19 from other non-Covid cases. Similarly, Xception performed an accuracy around 99%. They also compared these accuracy levels with the performance of radiologists in diagnosing Covid-19 via CT scans which have an accuracy of 86%.

Since considering only CT scans or clinical findings might be misleading especially in the early phases of the disease, Mei et al. (2020)[31] combined the clinical findings, laboratory tests, and exposure history with the CT scans to diagnose patients who are Covid-19 positive. Their proposed AI system has correctly identified 68% of the Covid-19 positive patients whose T scans are normal and classified as Covid-19 negative by the radiologists. By having both Y scans and relevant clinical information, their solution highly aids in diagnosis Covid-19 patients and it increases the accuracy level even in cases that CT scans might be misleading for medical professionals. By doing so, higher false negative in detection of Covid-19 can be decreased.

Cui et al. (2021) studied to create a prognostic tool for predicting poor outcomes in Covid-19 patients using CT images. They used images from the early days which are 0 to 6 days of the symptoms and late days which are more than 7 days from 492 patients. They applied LASSO approach to two separate groups and created distinct scores to evaluate predictions of poor outcomes. By interpreting those created scores, their findings suggested that using CT images to detect Covid-19 is more effective for the late-phase Covid-19 patients. For early-phase patients, considering clinical findings along with CT images resulted better.

These studies highlight the potential of CT imaging in diagnosing Covid-19 and predicting the severity and risk level of the disease. In some studies, it is shown that clinical findings would help a lot to CT imaging results. By using radiomics features and advanced machine learning and deep learning techniques, these studies offer valuable insight for resource allocation, clinical decision-making, and risk evaluation as well as disease diagnosis.

2.7 Convolutional Neural Network Models

As mentioned above, in this thesis different convolutional neural networks are utilized for answering the related research questions. To have a better understanding of the convolutional neural networks, in this section an overview of the convolutional neural networks, the motives to use them in imaging tasks and medical imaging and the explanations of the architectures of the selected networks is included.

Convolutional neural networks are widely used subsets of deep learning models. They are one of the various types of neural networks which are used for different purposes and data types. They are extremely useful for processing and analyzing visual data, including images and videos. Their ability to automatically learn hierarchical representations directly from pixel data helped them to revolutionize image processing and computer vision. They have been extremely popular in medical imaging tasks in recent years, due to their performance and ability to extract meaningful features from complex images like CT images.

2.7.1 Understanding Convolutional Neural Networks

Convolutional neural networks consist of multiple interconnected layers, each serving a specific purpose in the process. These layers are the mathematical functions used to extract features from the input and pass that output to the following layer. There are mainly two components of the convolutional neural networks. The first one is the feature extraction, which has the relevant layers for feature extraction. The second one is the fully connected layer which gives the output, predicting the image class for this thesis based on the features retrieved in the previous layers.

Convolutional neural networks have three types of layers: convolutional layers, pooling layers, and fully connected layers:

Convolutional layers are the first layers in the network responsible for extracting the different features of the input image. Most of the computations happen in the convolutional layers which can be assumed as the core building blocks of CNNs. The kernel or filter moves across the image and checks whether a feature is present in the image. Numerical values are obtained from the image in this layer, and it allows CNN to interpret the image and obtain relevant patterns from it such as edges, textures, and shapes.

Pooling layers reduce the dimension and parameters in the input and results in a loss in the information as well. By so, it reduces the complexity and computational expenses and improves the efficiency of the CNN. They extract the most prominent features while reducing the complexity. Different operations can be applied for pooling purposes depending on the need and mechanism utilized.

Fully connected layers follow the convolutional and pooling layers and are utilized to make predictions based on the features learned in the previous layers. As the name suggests all the inputs or nodes from one layer to node of the next layer. Since they increase the density of the network, all the layers in the network are not fully connected layers.

2.7.2 Advantages of CNNs in Image Processing and Medical Imaging

Convolutional neural networks are highly utilized for image processing tasks including in medical imaging. They provide several advantages in image processing tasks.

CNNs are capable of automatically learning and extracting hierarchical features from images. They can easily identify patterns, edges, shapes, textures etc. in the image and understand complex visual information through convolutional layers.

CNNs have the translation invariance which is the capability of learning translationally invariant properties of the image. This allows them to recognize patterns regardless of their location in the image. The translation invariance capability is very crucial in medical imaging since the position of an anomaly might vary in the image.

One of the key strengths of convolutional neural networks is adaptability and transfer learning. Available CNN models trained on large datasets like ImageNet can be reused for specific imaging tasks. Using these pre-trained models significantly increases the performance of the networks and reduces the time required to develop models. This is deeply analyzed to answer one of the research questions of the thesis and will be explained in detail in the following sections.

CNNs have the capability to automate the image analysis process for medical professionals in detection and diagnosis of the diseases, identifying hidden patterns which are hard to detect by the human eye. Automated and objective analysis of medical images can assist radiologists and clinicians.

Parameter sharing is a fundamental technique in CNNs which utilize the same weights and biases in the entire image. Sharing parameters allow CNNs to apply the same learned features across different regions of the image. The reduction in the number of parameters serves to a memory and computational cost saving, which makes the CNNs applicable for real-world problems. Thanks to parameter sharing, the network learns to recognize the essential features regardless of their location in the image.

CNNs have revolutionized image processing, especially in medical imaging tasks. In medical imaging tasks, CNNs have demonstrated high performance in disease detection and diagnosis, risk prediction, etc. which helps a lot to medical professionals. The use of CNNs would increase the overall capability of medical services and increase patient care by accurate and efficient analysis of medical images, which enables timely interventions and personalized treatments.

2.7.3 Detailed Explanation of the Selected CNN Architectures (e.g., VGG-16, ResNet-18, ResNet-50, DenseNet-121, DenseNet-169, EfficientNet-B0, and EfficientNet-B1)

In this thesis, different networks are utilized, applied to CT images, and evaluated. To have a better understanding of what methodology is followed, which steps are taken, detailed explanation of the networks and their architecture is explained in this section.

2.7.3.1 VGG-16

The VGG-16 Net was proposed by Karen Simonyan and Andrew Zisserman of the Visual Geometry Group Lab of Oxford University in 2014 in the paper "Very Deep Convolutional Networks for Large-Scale Image Recognition"[32]. With this model Simonyan & Zisserman, 2014 won 1st and 2nd place in object detection and classification in the 2014 ILSVRC challenge. It is one of the most popular networks for image classification and is easy to use with transfer learning.

The architecture of the network can be seen in Figure 2.1:

The VGG-16 Net consists of 16 layers, mostly convolutional layers. The network uses relatively small 3x3 filters in the convolutional layers compared to 11x11 filters in AlexNet or 7x7 filters ZFNet. The network architecture can be divided into five groups based on the number of layers and the specific configurations of convolutional

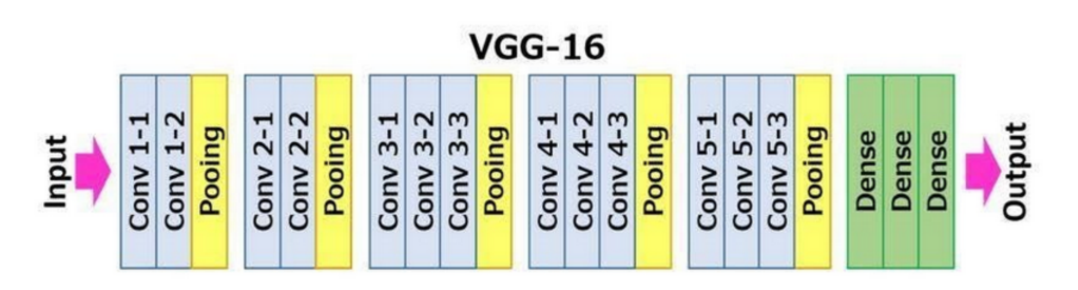

Figure 2.1: VGG-16 Architecture

and fully connected layers.

The input to the VGG-16 Net is typically a 224x224 RGB image. The network consists of thirteen convolution layers denoted as Conv1-1 to Conv5-3. These layers execute spatial convolution to the input image and extract relevant features and patterns in different scales. After each set of convolutional layers in the five groups in, max pooling is applied as Pool1 to Pool5. These layers reduce the dimensionality of the features and retain the most prominent features. The last three layers of the VGG-16 Net are the fully connected layers which are denoted as FC6 to FC8. These layers help to make predictions using the features learned in the previous layers. While FC6 and FC7 have 4096 neurons, the last fully connected layer's neuron number depends on the number of target classes.

Throughout the network, Rectified Linear Unit (ReLU) is used as activation function after each convolutional layer and fully connected layer. By this way, non-linearity is obtained which enables the network to learn complex representations from the images. Moreover, the network introduces dropout regularization in two of the fully connected layers.

With the 16 layers, the network has approximately 138 million trainable parameters, which makes it a relatively large model. Among the networks used for this thesis, it has the highest number of parameters.

2.7.3.2 ResNet-18

ResNet18 is a convolutional neural network introduced in the paper "Deep Residual Learning for Image Recognition" by He et al. in 2016[33]. An ensemble of the residual nets won the 1st place in the ILSVRC 2015 classification task which achieves 3.57% error on ImageNet. Deeper than VGG nets but they have lower complexity, and they are effective in especially in image classification tasks. One of the main problems that the ResNet models solve is the vanishing gradient. With the increase in the depth of the network, gradients diminish to very small values by the chain rule and fail to make significant progress or improvement in the learning process. Thanks to residual connections introduced in the network, training of very deep networks is enabled and learning more complex representations is allowed.

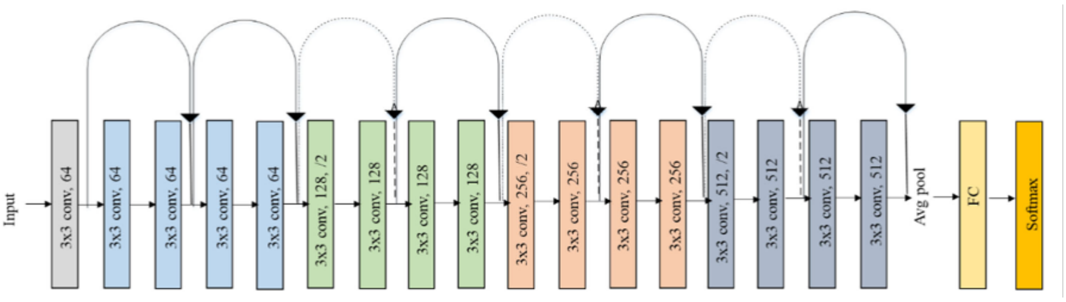

Figure 2.2: ResNet-18 Architecture

The architecture can be seen in Figure 3.1. The input to ResNet18 is typically a 224x224 RGB image. Compared to other residual nets, the other variants in the ResNet family, ResNet18 is relatively shallow and consists mostly convolutional layers. The network contains a series of convolutional layers divided into residual blocks followed by batch normalization and ReLU activation functions. Each residual block contains two convolutional layers with 3x3 kernels. The network applies average pooling and global average pooling for downsampling the dimensions. In the end of the network, there is a fully connected layer which is followed by a softmax function and used for classification purposes.

2.7.3.3 ResNet-50

ResNet50 is a convolutional neural network architecture from the ResNet (Residual Network) family like ResNet-18. Similar to ResNet-18, ResNet-50 is highly used in image classification since its effectiveness.

The network consists of 50 layers, most of them are convolutional layers. Thanks to the residual connections explained above, vanishing gradient problem is eliminated by He et al, 2016, ResnNet-50[33]. So, training such a deep network is enabled and complex representation learning is allowed.

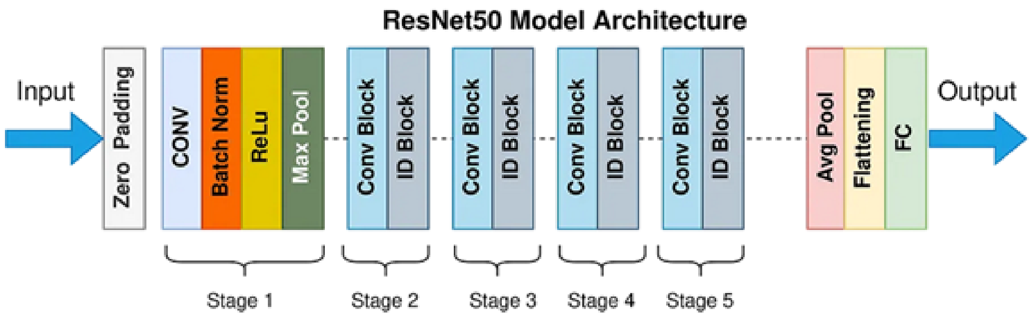

Figure 2.3: ResNet-50 Architecture

The input to ResNet50 is typically a 224x224 RGB image as in ResNet18. The architecture comprises various elements in its nature and starts with a 7x7 kernel convolutional layer along with 64 kernels with a stride of 2 which is followed by a 2-sized stride max pooling. After the max pooling, there are a series of convolutional layers which are made of a sequence of three layers. The sequence patterns include a 3x3 convolution with 64 kernels, 1x1 convolution with 64 kernel and 1x1 convolution with 256 kernels. After these 9 layers, another sequence of 12 convolutional layers follows the previous layers. 1x1 convolution with 128 kernels, 3x3 convolutions with 128 kernels and 1x1 convolutions with 512 kernels are iterated 4 times. Following that, there are 18 layers which are iterated 6 times. The iterated sequence contains 1x1 convolutions with 256 kernels, 3x3 convolutions with 256 kernels, and 1x1 convolutions with 1024 kernels. The network's last convolutional layer sequence includes 9 more convolutional layers which contains a sequence of 3 layers which are 1x1 convolutions with 512 kernels, 3x3 convolutions with 512 kernels, and 1x1 convolutions with 2048 kernels repeated 3 times. In the end, average pooling and a fully connected layer concludes the network in which the fully connected layer utilizes the softmax activation function.

2.7.3.4 DenseNet-121

DenseNet-121 is a convolutional neural network architecture introduced in the paper "Densely Connected Convolutional Networks" by Huang et al. in 2017[34]. It is part of the DenseNet family of models. These models are known for their dense connectivity patterns and efficient use of parameters. DenseNet models are popular for image classification and object detection tasks. Different than traditional convolutional neural networks, in a DenseNet architecture, each layer is connected directly with every other layer rather than passing feature maps sequentially.

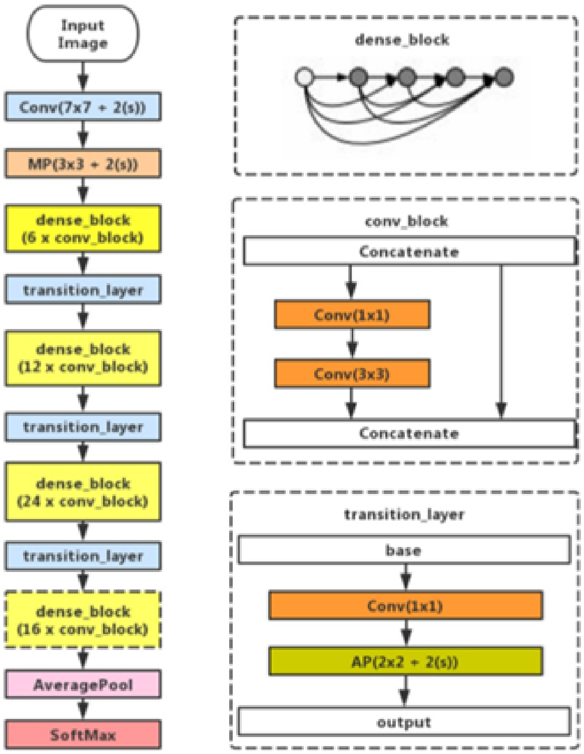

Figure 2.4: DenseNet-121 Architecture

Similar to the previously mentioned networks above the input to DenseNet-121 is typically a 224x224 RGB image. The overall architecture can be seen in Figure 2.4. The network starts with a convolutional layer utilizes 64 filters of size 7x7 with a stride of 2. This layer is followed by a basic pooling layer which applies 3x3 max pooling with a stride of 2.

Following the initial layers, Dense Block 1 exists which consists of 2 consecutive convolutional layers repeated 6 times. Transition Layer 1 contains one convolutional layer and applies average pooling follows Dense Block 1.

After those layers, Dense Block 2 contains two convolutional layers and repeats 12 times and Transition Layer 2 contains one convolutional layer along with average pooling.

Afterwards, Dense Block 3 which employs 2 two convolutions iterated 24 times. Transition Layer 3 follows, composed of one convolutional layer and average pooling.

The network continues with Dense Block 4 incorporates two convolutions repeated 16 times. Finally, the Global Average Pooling layer accepts all the feature maps of the network for the classification process, leading to the output layer.

DenseNet-121 utilizes batch normalization after every convolution layer and uses rectified linear unit (ReLU) activation function applied after the bath normalization.

2.7.3.5 DenseNet-169

Similar to DenseNet-121, DenseNet-169 is part of the DenseNet family of models. It follows the same principles as DenseNet-121 but has a deeper architecture by Huang et al., 2017[34].

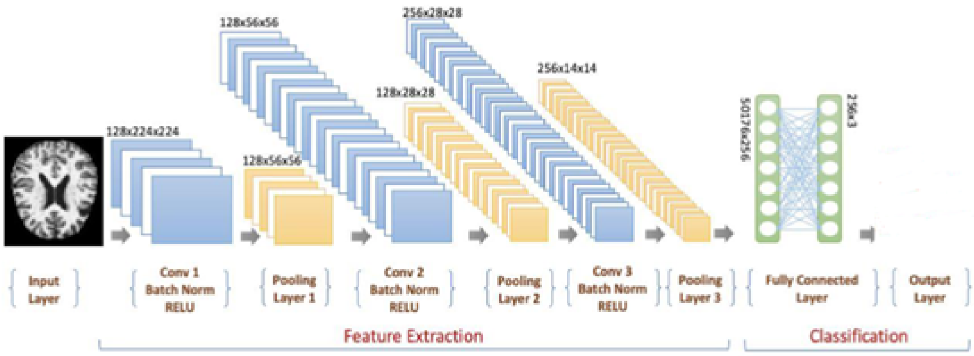

Figure 2.5: DenseNet-169 Architecture

DenseNet-169 architecture can be seen in Figure 2.5. The input to DenseNet-169 is typically a 224x224 RGB image. As in DenseNet-121, DenseNet-169 consists of four dense blocks, each having multiple convolutional layers. Each dense block starts with a convolutional layer and contains a series of convolutional layers in which batch normalization is applied and ReLU activation function is used. Each dense block is followed by a transition layer which contains a 1x1 convolution layer and pooling to reduce the number of channels to handle the model's complexity.

The network starts with a 7x7 convolutional layer with a stride of 2 and utilizes 64 filters. To downsample the feature maps, max pooling follows the convolutional layer. Following the initial convolutional layer, the four dense blocks and transition blocks are repeated. The first dense block comprises 6 layers, the second dense block is made up of 12 layers, the third and fourth dense blocks consist of 32 layers. Each layer in all four dense blocks comprises 4 units. Except the fourth dense block, each dense block is followed by a transition block which involves a 1x1 convolution and 2x2 average pooling. In the end of the network, for classification purposes, the network employs a global average pooling and uses a fully connected layer containing 1000 nodes and softmax activation function.

2.7.3.6 EfficientNet-B0

EfficientNet-B0 is a convolutional neural network architecture introduced in the paper "EfficientNet: Rethinking Model Scaling for Convolutional Neural Networks" by Tan et al. in 2019[35]. It is one of the networks of EfficientNet family. These networks are more compact compared to the previous ones explained above yet powerful. The paper emphasizes balancing accuracy and computational resources. The paper by Tan & Le, 2019[35] showed a better scaling of all the dimensions of the network which are width, depth, and resolution scaling. They are trying to find the best coefficients for width, depth and resolution that maximizes the accuracy of the network considering the available resources. The model was inspired by the Mnas-Net by Tan et al (2019)[35].

Typically, a 224x224 RGB image is given as an input to the model. The building block of this network is the mobile inverted bottleneck MBConv called inverted residual block with an additional squeeze and excitation block. The residual blocks used in the flow of deep neural networks follow the pattern of starting wide in the beginning of the convolution block and get narrower along the depth of the network till the end. In the end they become wide again with the added information. Thus, mostly a wide-narrow-wide pattern is followed. In the inverted residual block, the story changes and narrow-wide-narrow sequence is followed to use computational resources efficiently and increase non-linearity and learning capacity. The explained architecture of the network can be found in Figure 2.6.

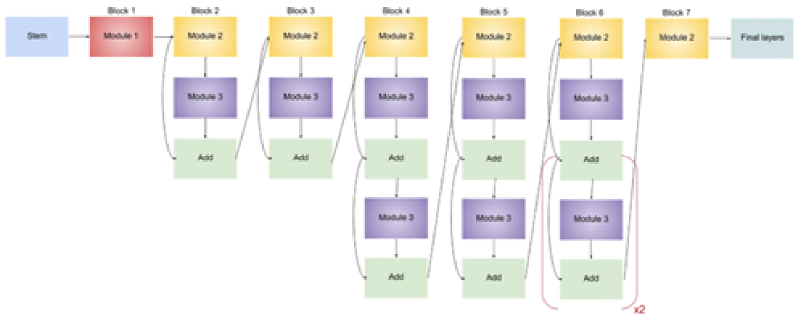

Figure 2.6: Architecture for EfficientNet-B0 (x2 means that modules inside the bracket are repeated twice)

2.7.3.7 EfficientNet-B1

Similar to EfficientNet-B0, EfficientNet-B1 is a convolutional neural network architecture introduced in the paper "EfficientNet: Rethinking Model Scaling for Convolutional Neural Networks" by Tan et al. in 2019[35] and is one of the networks of EfficientNet family. The network architecture can be seen in Figure 2.7.

The flow in the network is very similar to EfficientNet-B0 to use computational resources efficiently and increase non-linearity and learning capacity. The only difference is the EfficientNet-B1 has more sub blocks than EfficientNet-B0.

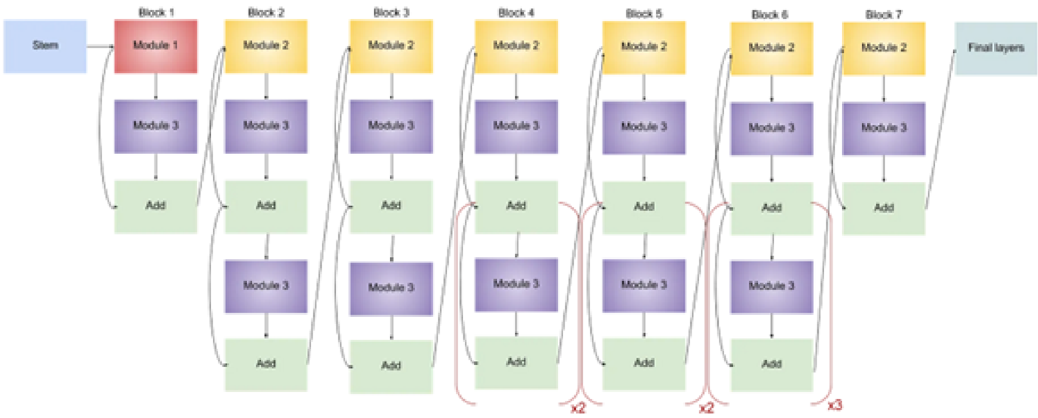

Figure 2.7: Architecture for EfficientNet-B1 (x2 and x3 means that modules inside the bracket are repeated twice and third times)

CHAPTER 3

METHODOLOGY

3.1 Description of the CT Image Dataset Used

The Covid-19 CT dataset used in this study which is collected by Yang et al., 2020[1] contains 349 CT images labeled as being positive for Covid-19 from 216 patient and 397 CT images that are negative. Examples of CT images from patients diagnosed with Covid-19 can be seen in Figure 3.1.

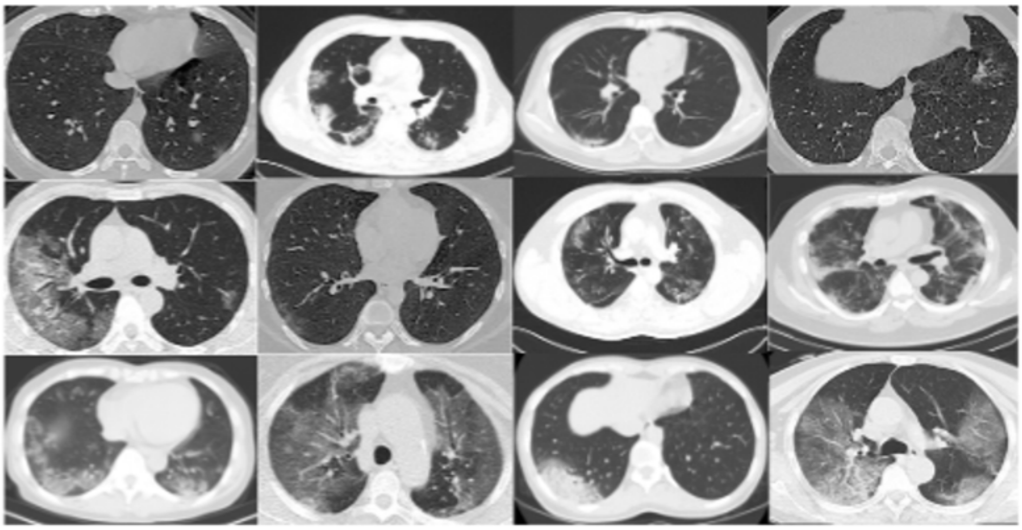

Figure 3.1: Examples of Positively Labeled CT-Scans for Covid-19

Table 3.1: Sizes of CT Images

	Height	Width
Minimum Pixel Value	153	124
Average Pixel Value	491	383
Maximum Pixel Value	1853	1485

These CT images have different sizes and the minimum, average, and maximum values can be seen in Table 3.1.

For the positively labeled patients 169 of them have age information and 137 of them have gender information and their distribution can be seen in Figure 3.2 and Figure 3.3. Gender and age features are not included to evaluate the model performance in this study.

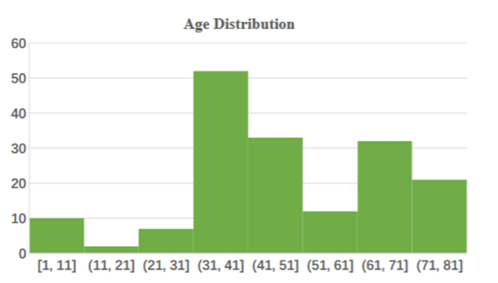

Figure 3.2: Age distribution of Covid-19 patients

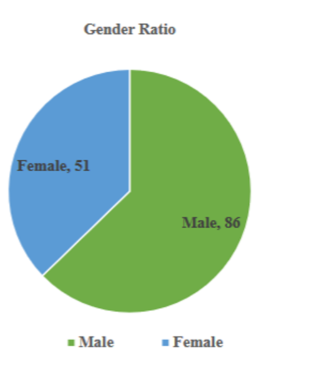

Figure 3.3: The gender ratio of Covid-19 patients

3.2 Data Preprocessing

In medical imaging, like in many other data problems, data preprocessing is one of the most crucial steps to ensure the quality and the reliability of the analysis. This section explains the data preprocessing steps taken on the Covid-19 T image dataset before utilizing it for training and evaluation.

The required Python packages which are essential for handling medical images and utilizing the networks explained in Section X. Since GPUs are preferred over CPUs for image processing tasks as mentioned by Woodward et al.[36], 2012 GPU acceleration is leveraged for this thesis.

The Covid-19 CT image dataset utilized in this study consists of two classes based on their Covid-19 status: CT images from the patients diagnosed with Covid-19 and CT images from non-Covid-19 patients.

The relevant data transformation steps are taken for both datasets. To prepare for training the deep learning model, the following transformations are applied. Images are resized to 256x256 to make the input images in the same size for neural networks. This standardization is crucial for compatibility with the model architecture that demands fixed-size inputs. Randomized cropping is made during training to augment the dataset and to increase model robustness. Normalization of the pixels reduces the impact of the variations across images and helps the model to learn features more effectively. To achieve this, image pixel values are normalized by using the values seen in Table 3.2 for each channel. These values are widely used for normalization and found to work well in practice as Krizhevsky et al. (2012)[37] explained in their paper. The normalization is done by subtracting the mean and dividing by the standard deviation for each channel.

Table 3.2: Mean and Standard Deviation Values for Each Channel

Channel	Mean	Standard Deviation
Red	0.485	0.229
Green	0.456	0.224
Blue	0.406	0.225

The dataset is split into three subsets as training set, validation set and test set. In order to ensure a balanced representation of both classes in the training, validation, and testing phases of the study, the dataset was split into three subsets following a 60-20-20 ratio. This maintains the overall distribution of data and guarantees that both classes are proportionally present in each subset. Each subset is loaded by using custom data loaders to facilitate efficient processing of batches during training and evaluation.

All these preprocessing operations serve essential purposes in the deep learning models. These steps ensure data quality, model compatibility and robustness which increase the effectiveness and overall capability of the trained neural networks.

3.3 Performance Metrics for Risk Prediction

The evaluation of the models is vital to assess their performance in predicting Covid-19 risk from CT images. Various metrics are used to comprehensively evaluate models' performance and provide insights about their capabilities and limitations.

Confusion matrix is used to derive several performance metrics. The mainly contains four components. True positive (TP), true negative (TN), false positive (FP), and false negative (FN). These metrics are the tools to assess models' ability. As their name suggests, true positive and true negative are straightforward. True positive is the number of Covid-19 positive cases where the model classified as positive. True negative is the number of Covid-19 negative patients that model predicts as negative. False positive on the other hand, is the number of instances that the model falsely indicates Covid-19 positive when the patient is Covid-19 negative. Lastly, false negative is the number of instances where the model incorrectly predicts Covid-19 negative while the patient is Covid-19 positive. By using these values, different interpretations can be made, and other metrics can be derived.

Precision which measures the accuracy of the positive predictions made by the model. It is computed as the ratio of true positive to the sum of true positive and false positive.

$$Precision = \frac{TP}{TP + FP} \tag{3.1}$$

Recall, which is also named as sensitivity, measures the model's ability to evaluate all positive instances. It is calculated as the ratio of TP to the sum of TP and FN.

$$Recall = \frac{TP}{TP + FN} \tag{3.2}$$

F1 score is another metric which combines precision and recall which provides a balanced metric between two. It is especially useful when there is an imbalance between the classes.

$$F1Score = \frac{2 \times Precision \times Recall}{Precision + Recall} \tag{3.3}$$

F1 Score=(2 x Precision x Recall)/(Precision+Recall)

Accuracy is a fundamental metric that quantifies the overall model predictions' correctness. It is computed as the ratio of the sum of total correctly predictions to the total number of cases. It is defined as follows:

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \tag{3.4}$$

The Receiver operating characteristic (ROC) curve is a graphical representation of the trade-off between true positive rate and false positive rate at various thresholds. It plots the True Positive Rate against the False Positive Rate. The Aera under the curve (AUC) is the area under the ROC curve and serve as a comprehensive metric for the models' ability.

A higher AUC indicates better model performance. For example, when AUC equals to 1, it means that the model makes perfect predictions. When it is 0.5, it is equivalent to random guessing which makes the model without any discriminatory power. The illustrations can be seen below3.4[2].

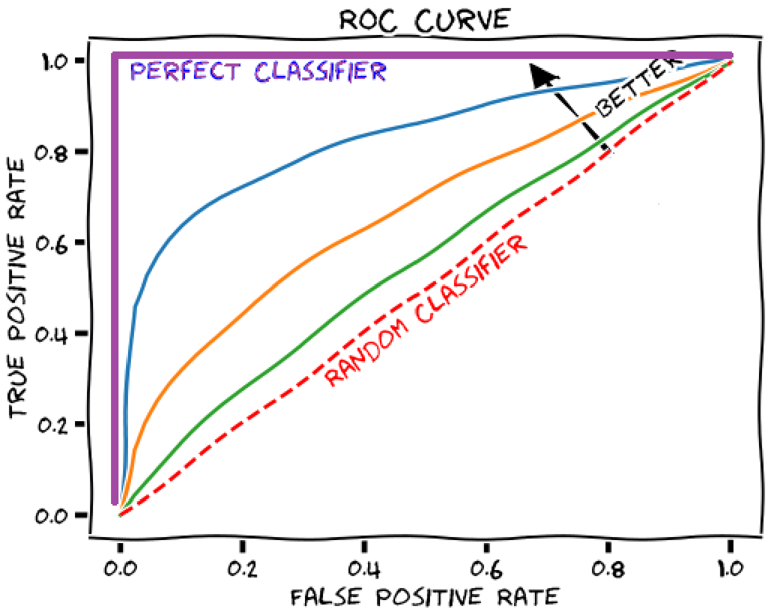

Figure 3.4: The ROC Curve[2]

3.4 Training Process and Hyperparameters Settings

Throughout this study, Adam optimizers are applied and different results for them are compared that will be quantified in the following section. There are some advantages of it compared to the other optimizers. It is straightforward to implement, it requires low memory, and less tuning than any other optimization algorithm. Kingma et al. (2014) [38], and He et al. (2016) [33] utilized the Adam optimizer in image classification tasks and evaluated the performance of it. The default values of Adam optimizer are 0.001, 0.9, 0.999 and 1e-8 for learning rate, Beta1, Beta2, and epsilon, respectively.

As a loss function, cross-entropy loss was utilized for this study, since it is a commonly used loss metric in image classification tasks especially in convolutional neural networks. Simonyan et al. (2014)[32] and Krizhevsky et al. (2012)[37] utilized the cross-entropy loss function and discussed the network performance in image classification tasks. It measures the performance of a classification task when the output is a value between 0 and 1. It measures the distance between the predicted and true probability distribution and the formula is as follows:

$$L(y, \hat{y}) = ylog(\hat{y}) + (1 - y)log(1 - \hat{y}) \tag{3.5}$$

The number of epochs selected for all the runs in this study was 70. While deciding the number of epochs that the model would run, it is important to observe the training loss and validation loss over epochs. As it can be seen in Figure 3.5, the training loss stabilizes before the 70th epoch, while the validation loss starts to increase around the 70th epoch. To prevent overfitting, stopping the model training at the 70th epoch is a good choice.

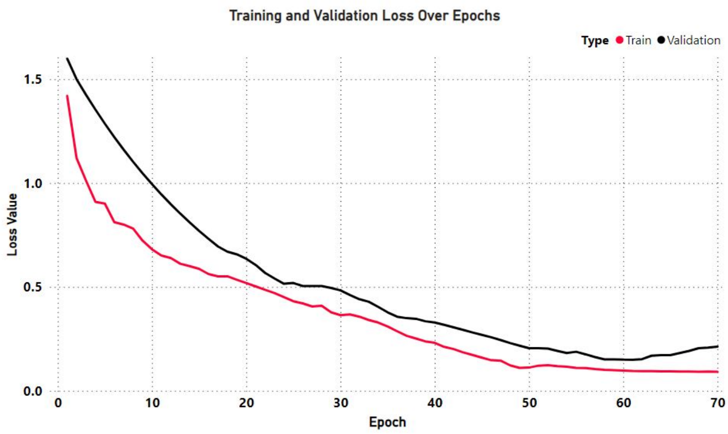

Figure 3.5: Training and Validation Loss Over 70 Epochs

Every model was run 30 times with different seeds. After obtaining the values for each model, the standard deviation and mean values are calculated for the 30 runs. After that, the upper confidence level is calculated by adding one standard deviation to the mean. Similarly, the lower confidence level is calculated by subtracting one standard deviation from the mean. While answering the effect of partial training of the networks, there are more than one upper-level values. These are obtained by running the network five times with the same seed, resulting in slightly different values for the same seed. It means that the models are run 150 times to answer that research questions.

CHAPTER 4

RESULTS AND DISCUSSION

4.1 Presentation and Analysis of the Results

To answer the first research question that the thesis focuses, which is to develop different convolutional neural network architectures optimized for COVID-19 risk prediction using computed tomography (CT) images, and to compare their performance in terms of accuracy, sensitivity, specificity, and computational efficiency; different networks are trained from scratch by using the parameters explained above and the results shown in Table 4.1 were obtained.

Table 4.1: Performance of Randomly Initialized Networks

	Accuracy	F1 Score	ROC - AUC
Network	mean$\pm std$	mean$\pm std$	mean$\pm std$
VGG-16	0.60$\pm$0.06	0.61$\pm$0.07	0.64$\pm$0.06
ResNet-18	0.66$\pm$0.05	0.69$\pm$0.06	0.76$\pm$0.02
ResNet-50	0.59$\pm$0.05	0.63$\pm$0.09	0.66$\pm$0.05
DenseNet-121	0.69$\pm$0.05	0.68$\pm$0.06	0.79$\pm$0.05
DenseNet-169	0.71$\pm$0.05	0.71$\pm$0.06	0.81$\pm$0.05
EfficientNet-B0	0.63$\pm$0.05	0.67$\pm$0.05	0.70$\pm$0.03
EfficientNet-B1	0.71$\pm$0.05	0.70$\pm$0.05	0.81$\pm$0.05

The mean of accuracy for VGG-16, ResNet-50, and EfficientNet-B0 networks are below 0.65. Moreover, ResNet-18 and DenseNet-121 have the upper level of accuracy below 0.70. DenseNet-169 and EfficientNet-B0 have values above 0.70. It is seen that DenseNet-169 and EfficientNet-B1 gives the best results. A similar trend is also valid for F1 score and AUC between the networks in which DenseNet-169 and

EfficientNet-B1 have the highest values among all.

Even if these two networks seem the best performing ones, a statistical comparison should be made. Hypothesis testing is applied to compare the models and whether a statistically significant difference between the performance of the models. The following hypothesis is constructed and tested accordingly:

H_0 : There is no significant difference between the two models

H_1 : There is a significant difference between the two models

To evaluate the hypothesis tests, Wilcoxon Signed-Rank Test is constructed and the p-values comparing different models performance are obtained. A lower p-value indicates that there is strong evidence to reject the null hypothesis in favor of the alternative hypothesis. Lower p-value concludes the observed difference between the two groups is unlikely to have occurred by random chance there is a statistically significant difference between the models. The related p-values while comparing the randomly initialized networks performance are obtained and can be seen in 4.1. DenseNet-169 and EfficientNet-B1 has lower p-values than 0.05 for every network comparison. By examining these values, it can be asserted that there is strong evidence suggesting a significant difference in the performance of these two models compared to other models. Therefore, the null hypothesis is rejected, concluding that there is a statistically significant difference between DenseNet-169 and other networks in favor of DenseNet-169. Similarly, there is a statistically significant difference between EfficientNet-B1 and other networks in favor of EfficientNet-B1. Only two comparisons have p-values higher than 0.05: VGG-16 vs. ResNet-50 and DenseNet-169 vs. EfficientNet-B1, making it difficult to determine which one is statistically better than the other.

For the sake of simplicity, these two networks were selected to answer the remaining research questions in this thesis.

After understanding that different networks perform differently in predicting Covid-19 risk by using computed tomography images, models are run with parameters pre-trained on ImageNet to understand whether applying transfer learning influences the performance of convolutional neural networks in the context of Covid-19 risk predic-

Model Comparison	p-value
ResNet-50 vs. DenseNet-169	1.243450e-14
VGG-16 vs. DenseNet-169	3.375078e-14
ResNet-50 vs. EfficientNet-B1	2.433609e-13
VGG-16 vs. EfficientNet-B1	3.002043e-13
ResNet-50 vs. DenseNet-121	3.403677e-11
VGG-16 vs. DenseNet-121	2.023217e-10
EfficientNet-B0 vs. EfficientNet-B1	2.260165e-10
DenseNet-169 vs. EfficientNet-B0	3.881979e-10
DenseNet-121 vs. EfficientNet-B0	2.305757e-07
ResNet-18 vs. ResNet-50	5.327708e-07
ResNet-18 vs. EfficientNet-B1	1.859967e-06
ResNet-18 vs. DenseNet-169	1.296514e-05
VGG-16 vs. ResNet-18	7.046067e-05
ResNet-18 vs. DenseNet-121	2.409992e-04
ResNet-50 vs. EfficientNet-B0	5.734774e-04
VGG-16 vs. EfficientNet-B0	2.420950e-02
ResNet-18 vs. EfficientNet-B0	2.825698e-02
DenseNet-121 vs. DenseNet-169	4.830954e-02
DenseNet-121 vs. EfficientNet-B1	4.970954e-02

P-value = 0.05

VGG-16 vs. ResNet-50	2.804057e-01
DenseNet-169 vs. EfficientNet-B1	3.325669e-01

Figure 4.1: Wilcoxon Signed-Rank Test for Randomly Initialized Networks Performances

tion using computed tomography images. The following results in Table 4.2 are obtained. By comparing the values in Table 4.1 to Table 4.2, it is seen that every network investigated in this thesis uses the advantage of transfer learning. For each network, transfer learning improves network performance by around 15%. This improvement was observed by many in literature: Kandel et al. (2020)[39] experiences similar improvement by applying transfer learning to detect diabetic retinopathy (DR) using VGG-16 and VGG-19 networks. With their paper, they explore the use of pre-trained CNN models for feature extraction and fine-tuning on a specific medical imaging task, showcasing the potential of transfer learning in healthcare applications. At this point, it is important to emphasize that even if the target dataset's domain, which is medical domain for this thesis, is different from the one which the pretraining is done, e.g. ImageNet, a daily object dataset, transfer learning has a positive impact on improving the performance of models. This can be interpreted as the learned features in ImageNet dataset are transferable and effectively reused for CT image classification task. It means that capturing general features such as edges, textures and shapes from the

ImageNet dataset would help the network to learn better. By already having useful information after pretraining, to adapt weights to the medical domain would be faster compared to the model start from random initialization. Thanks to these, the model focuses on learning specific details for Covid-19 prediction rather than starting from scratch.

Even if there are improvements in every network, similar to the results obtained with random initialized networks, DenseNet-169 and EfficientNet-B1 have the best performance compared to others. In terms of accuracy, these networks' mean values are above 0.80 while the others have less than 0.80 values. Similarly, these two networks have the highest AUC, and F1 score values compared to other networks. However, it is also necessary to compare them by using Wilcoxon Signed-Rank Test. After comparing these models performance the p-values are obtained in 4.2.

Model Comparison	p-value
DenseNet-169 vs. EfficientNet-B0	5.329071e-15
EfficientNet-B0 vs. EfficientNet-B1	9.769963e-14
VGG-16 vs. DenseNet-169	3.677059e-13
VGG-16 vs. EfficientNet-B1	4.218847e-12
ResNet-50 vs. EfficientNet-B0	3.134559e-10
DenseNet-121 vs. DenseNet-169	4.440531e-09
ResNet-18 vs. DenseNet-169	1.370374e-08
DenseNet-121 vs. EfficientNet-B0	1.692581e-08
ResNet-18 vs. EfficientNet-B0	2.317554e-08
VGG-16 vs. ResNet-18	5.402057e-08
ResNet-50 vs. DenseNet-169	3.486416e-07
VGG-16 vs. ResNet-50	5.367255e-06
VGG-16 vs. DenseNet-121	5.680464e-06
ResNet-50 vs. EfficientNet-B1	8.872968e-06
DenseNet-121 vs. EfficientNet-B1	1.045182e-05
ResNet-18 vs. EfficientNet-B1	4.646973e-05
VGG-16 vs. EfficientNet-B0	8.226752e-02
DenseNet-169 vs. EfficientNet-B1	1.909597e-01
ResNet-18 vs. ResNet-50	6.253678e-01
ResNet-18 vs. DenseNet-121	8.112492e-01
ResNet-50 vs. DenseNet-121	8.632650e-01

P-value = 0.05

Figure 4.2: Wilcoxon Signed-Rank Test for ImageNet Pretrained Networks Performances

Table 4.2: Performance of ImageNet Pretrained Networks

	Accuracy	F1 Score	ROC - AUC
Network	mean$\pm std$	mean$\pm std$	mean$\pm std$
VGG-16	0.73$\pm$0.04	0.75$\pm$0.04	0.84$\pm$0.02
ResNet-18	0.77$\pm$0.03	0.78$\pm$0.04	0.86$\pm$0.02
ResNet-50	0.77$\pm$0.03	0.79$\pm$0.05	0.87$\pm$0.02
DenseNet-121	0.77$\pm$0.03	0.78$\pm$0.04	0.86$\pm$0.02
DenseNet-169	0.81$\pm$0.04	0.82$\pm$0.03	0.90$\pm$0.02
EfficientNet-B0	0.72$\pm$0.03	0.71$\pm$0.05	0.81$\pm$0.02
EfficientNet-B1	0.80$\pm$0.03	0.80$\pm$0.03	0.88$\pm$0.02

By examining the p-values, it is evident that DenseNet-169 and EfficientNet-B1 have p-values lower than 0.05 for every network comparison. Based on these values, it can be asserted that there is strong evidence suggesting a significant difference in the performance of these two models compared to other models. Therefore, the null hypothesis is rejected, and it is concluded that there is a statistically significant difference between DenseNet-169 and other networks in favor of DenseNet-169. Similarly, there is a statistically significant difference between EfficientNet-B1 and other networks in favor of EfficientNet-B1. The comparisons with p-values higher than 0.05 are the outcomes of other network comparisons.

After seeing that applying transfer learning even with a dataset that is in the different domain has a positive impact on the performance of predicting Covid-19 risk by using CT images it is also inquired to investigate the impact of using a source dataset in the similar domain for the target dataset (e.g., using LUNA16 (Lung Nodule Analysis 2016) dataset for Covid-19 CT dataset) for transfer learning in Covid-19 risk prediction using computed tomography (CT) images and assess whether this approach improves the performance of the convolutional neural network models. The motive behind this is to understand how different pretraining methodologies affect the performance of convolutional neural networks in predicting Covid-19 risk. The choice of initialization and pretraining method can considerably influence a model's ability to capture features and provide information from them.

Table 4.3: Performance of DenseNet-169 and EfficientNet-B1 with Different Weights Initialization Mechanisms

Network	Initilization Method	Accuracy	F1 Score	ROC - AUC
		mean$\pm std$	mean$\pm std$	mean$\pm std$
DenseNet-169	Random Initialization	0.71 ± 0.05	0.71 ± 0.07	0.81 ± 0.06
	Pretrained on ImageNet	0.82 ± 0.03	0.82 ± 0.03	0.90 ± 0.02
	Pretrained on LUNA16	0.81 ± 0.03	0.82 ± 0.03	0.88 ± 0.02
	Pretrained on ImageNet, Finetuning on LUNA16	0.82 ± 0.03	0.81 ± 0.03	0.89 ± 0.02
EfficientNet-B1	Random Initialization	0.70 ± 0.05	0.69 ± 0.05	0.80 ± 0.05
	Pretrained on ImageNet	0.81 ± 0.03	0.80 ± 0.03	0.89 ± 0.02
	Pretrained on LUNA16	0.79 ± 0.03	0.80 ± 0.03	0.88 ± 0.02
	Pretrained on ImageNet, Finetuning on LUNA16	0.78 ± 0.03	0.80 ± 0.03	0.87 ± 0.02

At this point, it is important to explain the LUNA16 (Lung Nodule Analysis 2016) dataset in more detail. The LUNA16 dataset is a widely used dataset in the field of medical imaging, especially for detection of lung nodules in CT scans. It was created for the LUNA16 Grand Challenge. The dataset is annotated, including details such as the nodule's coordinates, diameter, and whether it is benign or malignant. The data consists of 6,691 patients in which 2,526 ones are diagnosed with positive 4,165 cases are diagnosed with negative.

To compare the effects of initialization and pretraining methods, DenseNet-169 and EfficientNet-B1 were utilized since they have the highest performance in random initialization and pretrained on ImageNet data. They both are pretrained on LUNA16 (Lung Nodule Analysis 2016) dataset and firstly pretrained on ImageNet and fine-tuned on LUNA16 dataset and the following performance metrics are obtained in Table 4.3. Intuitively, it is expected that pretraining on LUNA16 dataset would perform better than pretraining on ImageNet. However, as seen in the table domain specific pretraining can be as beneficial as pretraining on ImageNet but do not better perform. As mentioned in the preprocessing steps, normalization is applied to all images. This might caused a possible information loss in the images. The images might have converged to daily images and the complexity required from the model might have decreased and pretraining on LUNA16 dataset does not improve more than pretraining on ImageNet dataset.

Tschandl et al. (2019)[40] applied the same approach in skin lesion segmentation and compared the results of random initialized network performance, networks pretrained on ImageNet dataset, and pretrained on skin lesion dataset. They have similar results with this thesis' findings such that while random initialized network had the poorest performance, the other methods have better but similar performances with each other.

After observing that even using ImageNet for pretraining can increase the performance of the networks as much as using a medical-domain-dataset, it is further inquired whether the performance of deep networks can be maintained or improved by partially training only the last layers, while freezing the first layers, for various convolutional neural network architectures in the context of COVID-19 risk prediction using CT images. To answer this for the two networks, DenseNet-169 and EfficientNet-b1 perform better than the others are selected, and models are run and evaluated accordingly.

DenseNet-169 and EfficientNet-b1 have 14,149,480 and 7,794,184 trainable parameters consecutively. To investigate the partial training performance, the layers starting from the end of the network which have the 12%, 25%, 50%, and 75% of the total trainable parameters that network has are determined.

As seen in Figure 4.3, for the randomly initialized network, the performance does not show a significant trend with different layer share percentages. On the other hand, the model achieved a higher accuracy when pretrained on ImageNet compared to random initialization which ranges 77% to 86%. It is seen that it is possible to obtain higher performance when less layers are frozen, and more layers are trained. With an increase in the number of trained parameters there is a noticeable improvement in accuracy. The difference between the two initialization methods is getting smaller when less layers are trained. It makes sense since as the number of layers pretrained on ImageNet is getting smaller, the model converges to random initialization. This convergence occurs around the first 25% of the network. Thus, even with the ImageNet dataset, the first quarter of the layers have the same performance as the random initialization. Thus, there is no need to pretrain the whole network, it is enough to pretrain only the last 75% of the network and freeze the first quarter to improve runtime performance.

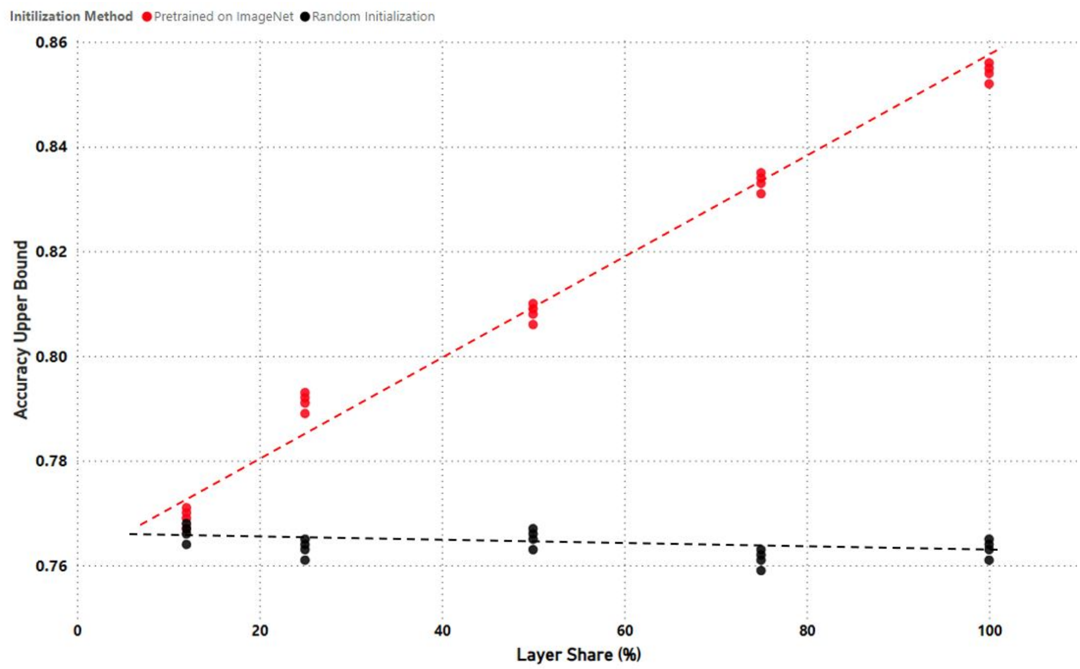

Figure 4.3: Randomly Initialized and Pretrained on ImageNet DenseNet-169 Accuracy for Training Different Percentages of the Network's Layers

By looking at Figure 4.3, it is seen that there is an increase in the performance of the model with an increase in the number of layers trained in case of pretraining on ImageNet. It means that the learning process is going on throughout the whole network. There are features throughout the whole network, and in all layers that the model learns.

By looking at Figure 4.3, it is seen that the performance of the model does not change with the change in the number of parameters trained in randomly initialized network. The performance of a case where the last 25% of the network is trained and a case where the last 75% of the network is trained is very similar. This can be interpreted in a way that parameters not trained but frozen are not critical for Covid-19 prediction. These parameters can be considered redundant or less influential others since they don't significantly impact model's ability to learn. The task of Covid-19 risk prediction by using CT images might be comparatively simple and does not contain complexity captured by all the parameters. (Goodfellow, Bengio, Courville, 2017)[41] discusses how simpler tasks may not require full complexity of the network to perform better. Similarly, it can be interpreted as the network has more capacity than the requirement of the task, but it has enough flexibility to adapt and learn the relevant features for the task by using the unfrozen parameters.

4.2 Strengths and limitations of the proposed approach

This study has some limitations that can be eliminated and investigated further in future research. One of the main limitations is the limited dataset of CT images for training and evaluation. Since the performance of deep learning models directly relies on the properties of the dataset, the Covid-19 CT dataset contains 349 CT images labeled as being positive for Covid-19 from 216 patient and 397 CT images that are negative might limit the performance of the study. The CT images may represent a specific population which might negatively affect the generalizability of the study to different populations with different characteristics. In this study, the images are resized to 256x256 to make the input images in the same size for neural networks as it is crucial for compatibility with the model architecture used. However, when the images are resized to a smaller size, it can result in some loss of detail since it involves averaging or subsampling pixel values. While splitting the dataset into training, validation and test subsets presence of the same individuals' CT images in multiple subsets are not eliminated. This is also a limitation of the study. Certain individuals' CT images may be present in both the training and validation sets, and possibly in the test set as well. This overlap introduces the risk of data leakage, potentially leading to an overestimation of the model's performance due to its exposure to familiar cases during training and validation.

The study provides a comprehensive analysis of different popular convolutional neural networks like VGG-16, ResNet-18, ResNet-50, DenseNet-121, DenseNet-169, EfficientNet-B0, and EfficientNet-B1 and allows an understanding of the performance of different architectures. Evaluation of both randomly initialized and ImageNet pretrained versions of the networks provides insights about the effect of transfer learning in prediction of Covid-19. Moreover, applying different datasets for pretraining helps to understand the dataset domain influence in the performance of the model. Finally, the effects of partial training are well evaluated to have a comprehensive understanding of different layers of the networks and their impact on the learning process.

CHAPTER 5

CONCLUSION

As explained in the previous chapters this thesis mainly focuses on answering four research questions about the performance and behaviors of convolutional neural networks in predicting Covid-19 risk by using CT images. We seek to answer these through comparison of different networks' performance, investigating the effects of transfer learning on the performance, the influence of the dataset domain which is used in transfer learning, and performance changes in case of partial training of the networks.

After analyzing all these results obtained from the model runs by using different convolutional neural networks, it can be said that for this study, DenseNet-169 and EfficientNet-B1 had the best performance metrics among all networks when the networks are randomly initialized. Only these two networks accuracy upper level are above 0.76 while the others have even below 0.70.

After comparing the networks performance when they are randomly initialized, the networks are pretrained with ImageNet to see the effects of transfer learning. With the pretrained weights, all networks increased their performance around 10%-15% which results an accuracy upper level of 85% for EfficientNet-b1 and DenseNet-169.

To inquire the effects of the domain of the pretraining dataset, rather than using just ImageNet, the LUNA16 dataset which is in the same domain with the target dataset, CT images is utilized. It is seen that there is no such an improvement in the case where LUNA16 dataset is used meaning that even with the ImageNet dataset, same performance can be obtained with the LUNA16 dataset.

And finally, both randomly initialized networks and networks that are pretrained with ImageNet are partially run to evaluate their performance and to understand their complexity. By freezing the first layers and only training the last layers, it is seen that there is no improvement with the increase in number of trained parameters in the randomly initialized network. This can be interpreted as the task this study focuses on which is Covid-19 risk prediction using CT images might be comparatively simple and does not contain complexity captured by all the parameters of the network. On the other hand, in the case of pretrained on ImageNet the performance of the network increases with the increase in the number of trained parameters. This can be concluded as, to have the highest performance, the whole network pretrained on ImageNet should be trained.

There are some points that are not included in this study but can be further improved in the future. As mentioned in the limitations part, different datasets with different characteristics might be included for both training and evaluation. In this way, an increase in the number of CT images ss obtained and generalizability of the model is improved.

The dataset used in this study contains CT images from the early pandemic. Throughout the pandemic, the virus has evolved, and different variants of the Covid-19 have emerged. To include all these variants, it is necessary to use different CT images taken from patients which are diagnosed with different Covid-19 variants. With this improvement, this study might be applicable in the practical fields for the fight against Covid-19 in 2024, as well as other possible novel viruses affecting the respiratory system.

In the context of this, only CT images of the patients are utilized, and related evaluation methods are applied to the results. As an improvement point in the future, the categorical parameters of the patients like age, weight, height, blood test results, medical records, patient habits, etc. can be included in the model in addition to use of CT images.

REFERENCES

[1] X. Yang, X. He, J. Zhao, Y. Zhang, S. Zhang, and P. Xie, "Covid-ct-dataset: a ct scan dataset about covid-19," *arXiv preprint arXiv:2003.13865*, 2020.

[2] XKCD, "Roc-draft-xkcd-style.svg," 2023.

[3] K. Zhang, X. Liu, J. Shen, Z. Li, Y. Sang, X. Wu, Y. Zha, W. Liang, C. Wang, K. Wang, *et al.*, "Clinically applicable ai system for accurate diagnosis, quantitative measurements, and prognosis of covid-19 pneumonia using computed tomography," *Cell*, vol. 181, no. 6, pp. 1423–1433, 2020.

[4] L. Li, L. Qin, Z. Xu, Y. Yin, X. Wang, B. Kong, J. Bai, Y. Lu, Z. Fang, Q. Song, *et al.*, "Artificial intelligence distinguishes covid-19 from community acquired pneumonia on chest ct," *Radiology*, 2020.

[5] X. He, X. Yang, S. Zhang, J. Zhao, Y. Zhang, E. Xing, and P. Xie, "Sample-efficient deep learning for covid-19 diagnosis based on ct scans," *medrxiv*, pp. 2020–04, 2020.

[6] P. V. Markov, M. Ghafari, M. Beer, K. Lythgoe, P. Simmonds, N. I. Stilianakis, and A. Katzourakis, "The evolution of sars-cov-2," *Nature Reviews Microbiology*, vol. 21, no. 6, pp. 361–379, 2023.

[7] World Health Organization, "COVID-19 Data Dashboard." `https://covid19.who.int/`. Accessed: February 19, 2024.

[8] J. Mullol, I. Alobid, F. Mariño-Sánchez, A. Izquierdo-Domínguez, C. Marin, L. Klimek, D.-Y. Wang, and Z. Liu, "The loss of smell and taste in the covid-19 outbreak: a tale of many countries," *Current allergy and asthma reports*, vol. 20, pp. 1–5, 2020.

[9] X. Yu, D. Ran, J. Wang, Y. Qin, R. Liu, X. Shi, Y. Wang, C. Xie, J. Jiang, and J. Zhou, "Unclear but present danger: An asymptomatic sars-cov-2 carrier," *Genes & Diseases*, vol. 7, no. 4, pp. 558–566, 2020.

[10] V. Parcha, K. S. Booker, R. Kalra, S. Kuranz, L. Berra, G. Arora, and P. Arora, "A retrospective cohort study of 12,306 pediatric covid-19 patients in the united states," *Scientific reports*, vol. 11, no. 1, p. 10231, 2021.

[11] S. S. Diwan, S. Ravichandran, R. Govindarajan, and R. Narasimha, "Understanding transmission dynamics of covid-19-type infections by direct numerical simulations of cough/sneeze flows," *Transactions of the Indian National Academy of Engineering*, vol. 5, pp. 255–261, 2020.

[12] S.-E. Lee, D.-Y. Lee, W.-G. Lee, B. Kang, Y. S. Jang, B. Ryu, S. Lee, H. Bahk, and E. Lee, "Detection of novel coronavirus on the surface of environmental materials contaminated by covid-19 patients in the republic of korea," *Osong public health and research perspectives*, vol. 11, no. 3, p. 128, 2020.

[13] A. D. Kaye, C. N. Okeagu, A. D. Pham, R. A. Silva, J. J. Hurley, B. L. Arron, N. Sarfraz, H. N. Lee, G. E. Ghali, J. W. Gamble, *et al.*, "Economic impact of covid-19 pandemic on healthcare facilities and systems: International perspectives," *Best Practice & Research Clinical Anaesthesiology*, vol. 35, no. 3, pp. 293–306, 2021.

[14] D. Koh, “Covid-19 lockdowns throughout the world,” *Occupational Medicine*, vol. 70, no. 5, pp. 322–322, 2020.

[15] M. Chinazzi, J. T. Davis, M. Ajelli, C. Gioannini, M. Litvinova, S. Merler, A. Pastore y Piontti, K. Mu, L. Rossi, K. Sun, *et al.*, “The effect of travel restrictions on the spread of the 2019 novel coronavirus (covid-19) outbreak,” *Science*, vol. 368, no. 6489, pp. 395–400, 2020.

[16] B. S. Graham, “Rapid covid-19 vaccine development,” *Science*, vol. 368, no. 6494, pp. 945–946, 2020.

[17] K. Miller, “6 of the worst pandemics in history,” n.d. Accessed: February 19, 2024.

[18] Health.com, “Ground glass opacities in covid-19,” n.d. Accessed: [2023-12-12].

[19] A. Bhatt, A. Ganatra, and K. Kotecha, “Covid-19 pulmonary consolidations detection in chest x-ray using progressive resizing and transfer learning techniques,” *Heliyon*, vol. 7, no. 6, 2021.

[20] Z. Sun, N. Zhang, Y. Li, and X. Xu, “A systematic review of chest imaging findings in covid-19,” *Quantitative imaging in medicine and surgery*, vol. 10, no. 5, p. 1058, 2020.

[21] J. Lei, J. Li, X. Li, and X. Qi, “Ct imaging of the 2019 novel coronavirus (2019-ncov) pneumonia,” *Radiology*, vol. 295, no. 1, pp. 18–18, 2020.

[22] A. Bernheim, X. Mei, M. Huang, Y. Yang, Z. A. Fayad, N. Zhang, K. Diao, B. Lin, X. Zhu, K. Li, *et al.*, “Chest ct findings in coronavirus disease-19 (covid-

19): relationship to duration of infection," *Radiology*, vol. 295, no. 3, pp. 685–691, 2020.

[23] L. Yan, H.-T. Zhang, J. Goncalves, Y. Xiao, M. Wang, Y. Guo, C. Sun, X. Tang, L. Jing, M. Zhang, *et al.*, "An interpretable mortality prediction model for covid-19 patients," *Nature machine intelligence*, vol. 2, no. 5, pp. 283–288, 2020.

[24] R. K. Gupta, M. Marks, T. H. Samuels, A. Luintel, T. Rampling, H. Chowdhury, M. Quartagno, A. Nair, M. Lipman, I. Abubakar, *et al.*, "Systematic evaluation and external validation of 22 prognostic models among hospitalised adults with covid-19: an observational cohort study," *European Respiratory Journal*, vol. 56, no. 6, 2020.

[25] F. Kamran, S. Tang, E. Otles, D. S. McEvoy, S. N. Saleh, J. Gong, B. Y. Li, S. Dutta, X. Liu, R. J. Medford, *et al.*, "Early identification of patients admitted to hospital for covid-19 at risk of clinical deterioration: model development and multisite external validation study," *bmj*, vol. 376, 2022.

[26] S. R. Knight, A. Ho, R. Pius, I. Buchan, G. Carson, T. M. Drake, J. Dunning, C. J. Fairfield, C. Gamble, C. A. Green, *et al.*, "Risk stratification of patients admitted to hospital with covid-19 using the isaric who clinical characterisation protocol: development and validation of the 4c mortality score," *bmj*, vol. 370, 2020.

[27] H. Sun, A. Jain, M. J. Leone, H. S. Alabsi, L. N. Brenner, E. Ye, W. Ge, Y.-P. Shao, C. L. Boutros, R. Wang, *et al.*, "Cova: an acuity score for outpatient screening that predicts coronavirus disease 2019 prognosis," *The Journal of infectious diseases*, vol. 223, no. 1, pp. 38–46, 2021.

[28] Y. Song, S. Zheng, L. Li, X. Zhang, X. Zhang, Z. Huang, J. Chen, R. Wang, H. Zhao, Y. Chong, *et al.*, "Deep learning enables accurate diagnosis of novel coronavirus (covid-19) with ct images," *IEEE/ACM transactions on computational biology and bioinformatics*, vol. 18, no. 6, pp. 2775–2780, 2021.

[29] C. Zheng, X. Deng, Q. Fu, Q. Zhou, J. Feng, H. Ma, W. Liu, and X. Wang, "Deep learning-based detection for covid-19 from chest ct using weak label," *MedRxiv*, pp. 2020–03, 2020.

[30] A. A. Ardakani, A. R. Kanafi, U. R. Acharya, N. Khadem, and A. Mohammadi, "Application of deep learning technique to manage covid-19 in routine clinical practice using ct images: Results of 10 convolutional neural networks," *Computers in biology and medicine*, vol. 121, p. 103795, 2020.

[31] X. Mei, H.-C. Lee, K.-y. Diao, M. Huang, B. Lin, C. Liu, Z. Xie, Y. Ma, P. M. Robson, M. Chung, *et al.*, "Artificial intelligence–enabled rapid diagnosis of patients with covid-19," *Nature medicine*, vol. 26, no. 8, pp. 1224–1228, 2020.

[32] K. Simonyan and A. Zisserman, "Very deep convolutional networks for large-scale image recognition," *arXiv preprint arXiv:1409.1556*, 2014.

[33] K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, pp. 770–778, 2016.

[34] G. Huang, Z. Liu, L. Van Der Maaten, and K. Q. Weinberger, "Densely connected convolutional networks," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, pp. 4700–4708, 2017.

[35] M. Tan and Q. Le, "Efficientnet: Rethinking model scaling for convolutional

neural networks," in *International conference on machine learning*, pp. 6105–6114, PMLR, 2019.

[36] A. Woodward, P. Delmas, and T. Ikegami, "An optimal parameter analysis and gpu acceleration of the image receptive fields neural network approach," in *Proceedings of the 27th Conference on Image and Vision Computing New Zealand*, pp. 222–226, 2012.

[37] A. Krizhevsky, I. Sutskever, and G. E. Hinton, "Imagenet classification with deep convolutional neural networks," *Advances in neural information processing systems*, vol. 25, 2012.

[38] D. P. Kingma and J. Ba, "Adam: A method for stochastic optimization," *arXiv preprint arXiv:1412.6980*, 2014.

[39] I. Kandel and M. Castelli, "Transfer learning with convolutional neural networks for diabetic retinopathy image classification. a review," *Applied Sciences*, vol. 10, no. 6, p. 2021, 2020.

[40] P. Tschandl, C. Sinz, and H. Kittler, "Domain-specific classification-pretrained fully convolutional network encoders for skin lesion segmentation," *Computers in biology and medicine*, vol. 104, pp. 111–116, 2019.

[41] Y. Bengio, I. Goodfellow, and A. Courville, *Deep learning*, vol. 1. MIT press Cambridge, MA, USA, 2017.