

Review

Detection and Classification of COVID-19 by Radiological Imaging Modalities Using Deep Learning Techniques: A Literature Review

Albatoul S. Althenayan ^{1,2,*}, Shada A. AlSalamah ^{1,3}, Sherin Aly ⁴, Thamer Nouh ⁵
and Abdulrahman A. Mirza ¹

¹ Information Systems Department, College of Computer and Information Sciences, King Saud University, Riyadh 11451, Saudi Arabia

² Information Systems Department, College of Computer and Information Sciences, Imam Mohammed Bin Saud Islamic University, Riyadh 11432, Saudi Arabia

³ Digital Health and Innovation Department, Science Division, World Health Organization, 1211 Geneva, Switzerland

⁴ Institute of Graduate Studies and Research, Alexandria University, Alexandria 21526, Egypt

⁵ Trauma and Acute Care Surgery Unit, College of Medicine, King Saud University, Riyadh 11461, Saudi Arabia

* Correspondence: 438204280@student.ksu.edu.sa; Tel.: +009-665-0606-1614

Abstract: Coronavirus disease (COVID-19) is a viral pneumonia that originated in China and has rapidly spread around the world. Early diagnosis is important to provide effective and timely treatment. Thus, many studies have attempted to solve the COVID-19 classification problems of workload classification, disease detection, and differentiation from other types of pneumonia and healthy lungs using different radiological imaging modalities. To date, several researchers have investigated the problem of using deep learning methods to detect COVID-19, but there are still unsolved challenges in this field, which this review aims to identify. The existing research on the COVID-19 classification problem suffers from limitations due to the use of the binary or flat multiclass classification, and building classifiers based on only a few classes. Moreover, most prior studies have focused on a single feature modality and evaluated their systems using a small public dataset. These studies also show a reliance on diagnostic processes based on CT as the main imaging modality, ignoring chest X-rays, as explained below. Accordingly, the aim of this review is to examine existing methods and frameworks in the literature that have been used to detect and classify COVID-19, as well as to identify research gaps and highlight the limitations from a critical perspective. The paper concludes with a list of recommendations, which are expected to assist future researchers in improving the diagnostic process for COVID-19 in particular. This should help to develop effective radiological diagnostic data for clinical applications and to open future directions in this area in general.

Keywords: artificial intelligence; COVID-19; CXR; CT-scan; deep learning; diagnosis; image classification; multi-classes; pneumonia



Citation: Althenayan, A.S.; AlSalamah, S.A.; Aly, S.; Nouh, T.; Mirza, A.A. Detection and Classification of COVID-19 by Radiological Imaging Modalities Using Deep Learning Techniques: A Literature Review. *Appl. Sci.* **2022**, *12*, 10535. <https://doi.org/10.3390/app122010535>

Academic Editor: Josué Álvarez Borrego

Received: 24 August 2022

Accepted: 28 September 2022

Published: 19 October 2022

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Copyright: © 2022 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

COVID-19 is a form of viral pneumonia that has emphatically threatened the world's healthcare infrastructure. The virus affects the respiratory system (i.e., lungs) of the infected patient, which can lead to respiratory insufficiency. Since China reported its initial cases to the World Health Organization (WHO) in December 2019, there have been 588,757,628 confirmed cases of COVID-19 and 6,433,794 deaths—due to respiratory failure and injury to other major organs—reported to the WHO as of August 2022 [1]. The rapid spread of the disease has also increased the number of hospitalizations worldwide [2].

The standard way to diagnose COVID-19 involves the use of the reverse-transcription polymerase chain reaction (RT-PCR) method [3]. RT-PCR is “a highly sensitive technique for the detection and quantitation of messenger ribonucleic acid (RNA)”; it is used for

many types of viruses, including SARS-CoV-2. The technique involves taking a sample from a part of the body, such as the nose, to extract the virus' RNA [4]. Physicians are facing difficulties due to the limitations of this method, which include long waiting times for results, low availability of examination kits, and suboptimal sensitivity [5]. These limitations can compromise patients and further the spread of COVID-19.

In urgent cases, physicians can ask radiologists to complete the diagnostic process using radiological imaging modalities such as chest X-rays (CXRs) or computed tomography (CT) scans of the lungs. These images are then analyzed by a specialist to reach a final diagnosis. Accurate analysis of these medical images can help to overcome the limitations of RT-PCR [6].

However, the limited time available for radiologists to screen medical images and differentiate between COVID-19 infection and other lung diseases—especially under time constraints and high patient numbers—makes completing these processes difficult. In addition, radiologists with different skills and lower diagnostic accuracy may lead to an inaccurate diagnosis, producing incorrect results [7]. A recent trend in healthcare involves deploying artificial intelligence (AI) algorithms in image classification problems (e.g., identifying cardiovascular abnormalities, detecting fractures and other musculoskeletal injuries, and diagnosing neurological diseases) [8]. High-accuracy AI models have been developed through training on medical images, yielding promising results after learning complex problems in radiology. More broadly, machine learning and deep learning models have produced effective results in several key aspects of radiology [9].

Given the extent of the recent pandemic, many of the studies that are reviewed in Section 4 have used deep learning techniques to diagnose and detect COVID-19 pneumonia using medical imaging in a theoretical way that cannot be deployed clinically. Accordingly, there are many review articles in this field that highlight the efficiency of deep learning algorithms and imaging modalities for the diagnosis of COVID-19 [10,11].

The novelty of this review consists of the provision of a theoretical background on topics relating to these issues. In addition, the analysis developed in this review focuses on the methods and the main approaches that have been adopted in recent studies, seeking to identify limitations and gaps for further research and improvement. Finally, recommendations related to these limitations are presented, aiming to help researchers develop practical models that advance AI applications in radiology for the detection of pneumonia and produce reliable results for other future medical applications.

2. Survey Method

A literature review was performed in order to identify a broad range of deep learning approaches for addressing the detection and classification of COVID-19. We documented some details of the review's search process so that other researchers may more confidently use this literature review in future research. The Google Scholar [12], IEEE Xplore [13], and PubMed [14] databases were used to find candidate papers. Keyword searches included combinations of query terms such as "COVID-19", "pneumonia diagnosis", "multi-classes", "deep learning", "image classification", "hierarchical", and "CXR". The included publications on the COVID-19 pandemic were published from 2019 to 2022. A total of 184 articles were extracted, and the search results were reviewed and filtered, removing those that did not demonstrate learning by deep learning techniques along with those articles that achieved their results using simple binary classification methods. Additional exclusion criteria were applied to the papers that used the same approach that was followed by the selected papers after screening based on their titles and abstracts, and these papers were removed so that the same information was not repeated.

The final 47 studies meeting the inclusion criteria included journal articles and conference papers. In order to be included in the review, the papers needed to meet the following inclusion criteria:

- Articles that employ deep learning approaches.

- Articles that address the problem of detection, identification, and classification of COVID-19.
- Articles written in English (from any country).
- Articles published since the end of 2019.
- Peer-reviewed articles.

3. Background

3.1. Coronavirus Taxonomy

Pneumonia is an acute infection that attacks the respiratory system (i.e., lungs), caused by various pathogens. The virus that causes COVID-19 is referred to as severe acute respiratory syndrome coronavirus-2 (SARS-CoV-2) due to its high pathogenicity, comparable to SARS-CoV [15]. Figure 1 illustrates the most common types of pneumonia, in a hierarchical structure. The known types of pneumonia differ in terms of their characteristics, causes, symptoms, and diagnosis [16]. The International Classification of Diseases (ICD) classifies all diseases, including pneumonia, using a complex hierarchical structure, as shown in Figure 2 [17]. In the 10th revision of the ICD (ICD-10), a new type of viral pneumonia was introduced, belonging to the coronavirus family: the 2019 novel coronavirus, coded as “COVID-19”. In December 2019, COVID-19 originated in Wuhan (Hubei Province, China), after which it rapidly spread worldwide. Due to the high mortality rate of the disease and its global transmission, the WHO defined COVID-19 as a pandemic soon after.

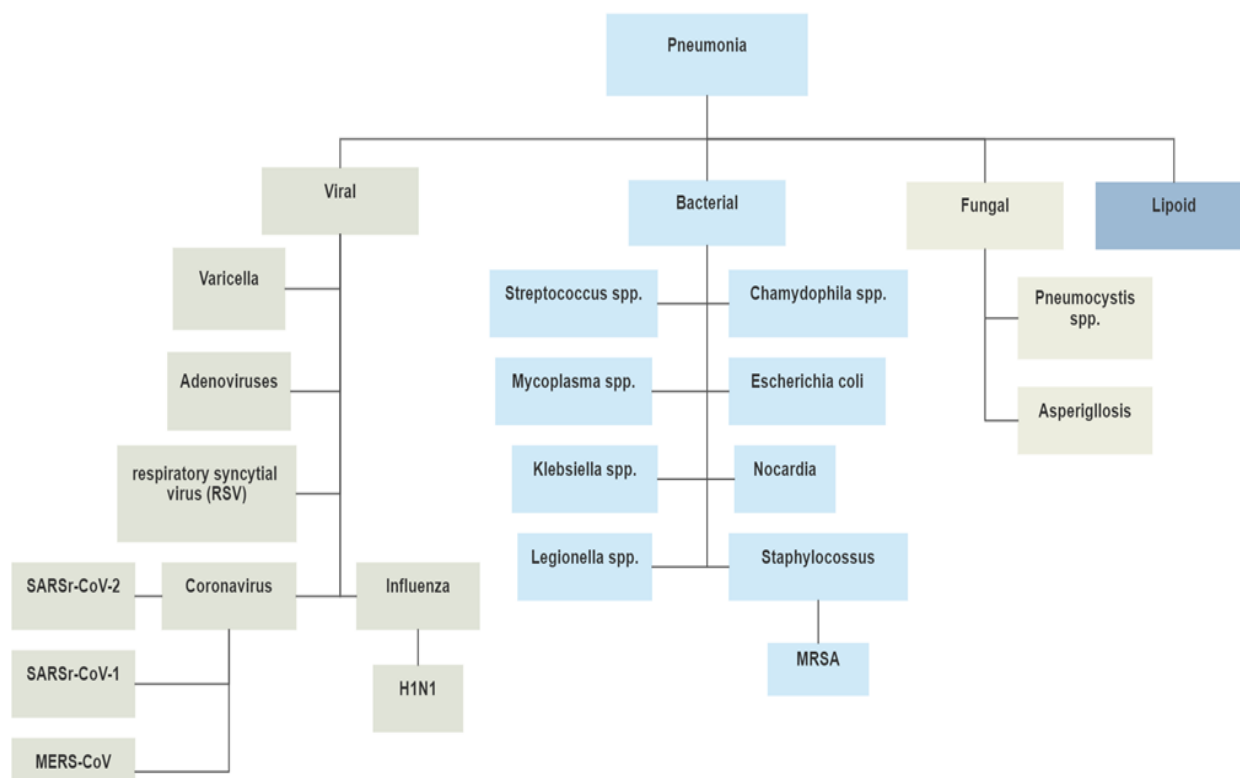


Figure 1. Hierarchical structure of the most prevalent pneumonia.

The first human coronaviruses began to emerge worldwide in the 1960s. Six human viruses belong to the coronavirus family, four of which are associated with mild symptoms, similar to those of the common cold and gastrointestinal system infections. The other two viruses are marked by differences, including their highly pathogenic nature, namely, Middle East respiratory syndrome coronavirus (MERS-CoV), which first originated in Saudi Arabia in 2012, and severe acute respiratory syndrome coronavirus (SARS-CoV-1), which emerged in Asia in February 2003 [18,19]. Furthermore, their impact on endemic countries’

health systems has been significant in raising mortality rates more than in other countries where the virus moved and did not originate, e.g., when MERS-CoV initially arose in the Middle East, most of the verified cases and deaths originated from Saudi Arabia [20]. At the beginning of the outbreak, COVID-19 was referred to as severe acute respiratory syndrome coronavirus-2 (SARS-CoV-2) due to its high pathogenicity, comparable to that of SARS-CoV [15].

ICD-10 Version:2019

Search [Advanced Search]

ICD-10 | Versions - Languages | Info

International Statistical Classification of Diseases and Related Health Problems 10th Revision (ICD-10)-WHO V expanded

Chapter X
Diseases of the respiratory system
(J00-J99)

Note: When a respiratory condition is described as occurring in more than one site and is not specifically indicated the lower anatomic site (e.g., tracheobronchitis to bronchitis in J40).

Excl.: certain conditions originating in the perinatal period (P00-P96)
certain infectious and parasitic diseases (A00-B99)
complications of pregnancy, childbirth and the puerperium (O00-O99)
congenital malformations, deformations and chromosomal abnormalities (Q00-Q99)
endocrine, nutritional and metabolic diseases (E00-E90)
injury, poisoning and certain other consequences of external causes (S00-T98)
neoplasms (C00-D48)
symptoms, signs and abnormal clinical and laboratory findings, not elsewhere classified (R00-R99)

This chapter contains the following blocks:

- J00-J06 Acute upper respiratory infections
- J09-J18 Influenza and pneumonia
- J20-J22 Other acute lower respiratory infections
- J30-J39 Other diseases of upper respiratory tract
- J40-J47 Chronic lower respiratory diseases
- J60-J70 Lung diseases due to external agents
- J80-J84 Other respiratory diseases principally affecting the interstitium
- J85-J86 Suppurative and necrotic conditions of lower respiratory tract
- J90-J94 Other diseases of pleura
- J95-J99 Other diseases of the respiratory system

Asterisk categories for this chapter are provided as follows:

- J17* Pneumonia in diseases classified elsewhere
- J91* Pleural effusion in conditions classified elsewhere
- J99* Respiratory disorders in diseases classified elsewhere

Figure 2. An example of ICD-10 hierarchical structure of the respiratory system diseases.

3.2. Radiological Imaging Modalities

Since the emergence of COVID-19, due to the high number of deaths, many organizations in different countries have sought to develop rapid and accurate diagnostic methods. Given that COVID-19 causes distinct spots on the lungs—as shown by the sample in Figure 3—a radiological examination of the chest is an important tool that can be used in the diagnostic process [19]. CT scans play an essential role in diagnosing pneumonia, especially given their efficiency and accuracy in detecting the features of pneumonia. Accordingly, many prior studies (in Section 4.3) have used the CT scan modality for the examination and diagnosis of COVID-19 despite the fact that CXR is more appropriate for the process, for the following reasons: firstly, it is difficult to control the spread of infectious diseases in CT suites and to decontaminate CT scanning machines; secondly, CT scans are costly, and are not available in all hospitals; thirdly, CT scanners are not portable [21]. In comparison, CXR is portable, cheap, simple to perform, and available in most hospitals. Therefore, CXR is the most commonly applied method for the radiological examination of the lungs [22]. The American College of Radiology (ACR) recommends that portable chest radiography tools should be used to limit the risk of disease transmission when scanning COVID-19 patients [23].



Figure 3. Pulmonary radiography of a person infected with COVID-19.

3.3. Hierarchical Image Classification

The data classification process aims to classify objects into related, predefined classes or categories. Image classification is an important topic in the research area of image processing, helping to classify images in different domains [24]. As shown in Figure 1, the nature of COVID-19 is organized in a hierarchical structure, which means that this is a hierarchical classification problem. Class hierarchy in hierarchical classification models supposes that an “IS-A” relationship exists between any node and its parent [25]. In hierarchical levels, the labels in high-level nodes are referred to as coarse-grained nodes that inherit all of the features from their parent node and have special features that will be passed down to their child nodes. In addition, the leaf node in the structure is referred to as fine-grained, which is the last level of nodes in the structure, and it inherits all of the features of its parent, and does not have children. Therefore, if the output of a classifier is class fine#2, it is natural to say that it also belongs to classes coarse#1.1 and coarse#1, as shown in Figure 4, where R is the root node.

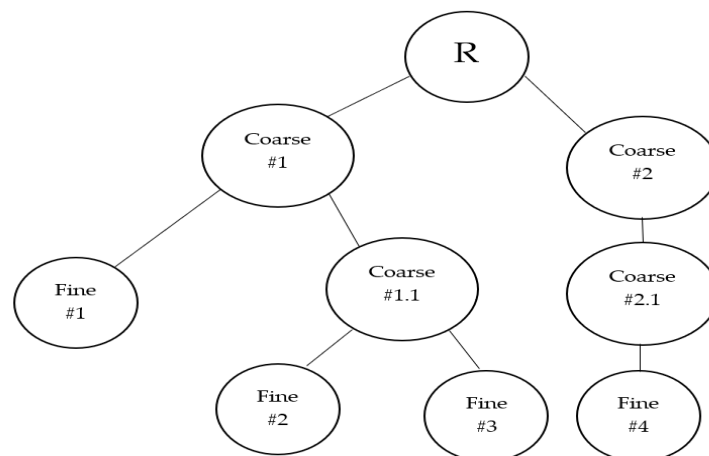


Figure 4. An example of the hierarchical tree structure.

Factors that improve performance when using classification models include—but are not limited to—the sharing of features between classes that fall on the same path in hierarchical classifications, and using the relationships between coarse- and fine-grained classes [26]. Despite this, most of the literature has focused on solving multi-class problems using flat classification (see examples in Sections 4.1, 4.2, 4.3, 4.5, and 4.6). Flat classification is a straightforward approach wherein there is no inherent hierarchy between the classes, which does not help the model to learn relevant features from the different

classes. Nevertheless, most classification problems—especially in the medical field—are hierarchical [27].

3.4. Deep Learning for Image Classification

Deep learning is an active subdomain of AI that has emerged recently. The term “deep” signifies that the neural network consists of a large number of hidden layers. Deep learning solves data-related problems effectively with minimal guidance from the developer. In addition, deep learning helps to reduce the time needed to solve problems involving big data [28].

In the field of medical imaging, deep learning has achieved remarkable successes in feature learning and image classification [29]. Consequently, this has encouraged researchers to explore deep learning techniques in greater depth. Compared to the existing diagnostic imaging-checking procedures that rely entirely on radiologists, deep learning algorithms perform excellently [30]. In particular, convolutional neural networks (CNNs) have yielded good results in previously impossible cases [31]. This is due to the fact that CNNs can detect and learn important features that radiologists cannot easily notice using the naked eye [32]. Therefore, deep learning offers novel models for image classification and medical image diagnostics, achieving excellent performance [33]. Furthermore, it is expected that deep learning techniques can help radiologists in the process of assessment and diagnosis.

3.5. Multimodal Data Fusion using Deep Learning

Multimodal data fusion is the process of learning features from heterogeneous data by integrating data from different sources (or types) into a single model to produce a unified result [34]. Since a single modality rarely provides complete knowledge of any real-world area of interest, multimodal data fusion is often essential. This is a form of multimodal big data that has a high volume and a high variety of data, and consists of several modalities. Diverse characteristics are extracted from multiple data sources and fed into the model, yielding a model with rich information—more so than is the case with a single modality, as mentioned above. This leads to significant performance improvements compared to the use of only a single feature modality [35] due to comprehensive characterization through mapping between different data types/sources.

Deep learning methods support the architecture of multimodal data fusion. Deep learning techniques have achieved substantial progress in multimodal structures in different domains, including medical assistant diagnosis [36]. Generally, the performance of multimodal deep learning models depends on the availability of a platform with a high computing capability to serve as the training device. The learning approach for multimodal deep learning models involves extracting the features of each modality separately. In turn, they are transformed into high-abstraction representations, which are concatenated into vectors as global representations of the multimodal data fusion that are used at the end by the deep learning model [37].

4. Deep Learning for COVID-19 Detection and Classification

COVID-19 medical image classification has recently gained significant attention as a research field due to the ongoing pandemic. Several researchers have developed deep learning classification and detection models to diagnose COVID-19 accurately and efficiently by classifying radiological images, as shown in this section. However, most existing approaches to the COVID-19 classification problem contain gaps and areas to extend, which we attempt to clarify in this paper. In the following subsections, a review is presented of studies that have adopted contrasting approaches to the problems in the literature. Table 1 summarizes most of the key approaches addressed in these studies.

4.1. Identification of COVID-19 from CXR Images using Ensemble of Deep Learning Models

Ensemble learning involves capturing the outputs produced by different classifiers, which helps to ensure a robust prediction and increases the accuracy of deep learning models [38]. Chouhan et al. [39] designed a model based on the performance of the following pre-trained deep learning models for the detection of pneumonia: AlexNet, DenseNet121, ResNet18, Inception-v3, and GoogLeNet. The dataset was obtained from the Guangzhou Women and Children's Medical Center (GWCMC); it consisted of a total of 5232 images labelled with the following classes: viral pneumonia, bacterial pneumonia, and normal. Random distribution of data into training (5232 images) and test (624 images) datasets was provided to avoid bias in performance. Some of the steps for data preprocessing include noise addition, random horizontal flip, and random-resized crop as augmentation techniques, with the images resized to 224×224 pixels. The classification results from the pre-trained CNN models were combined into a prediction vector, and majority voting was applied to generate a final prediction output. The proposed deep learning framework of the ensemble model achieved 96.4% testing accuracy.

Khan et al. [40] compared the performance of deep hybrid learning (COVID-RENet-1 and COVID-RENet-2) and deep boosted hybrid learning (COVID-RENet-1 and COVID-RENet-2) with the well-established CNNs (i.e., VGG-16/19, GoogLeNet, InceptionV3, ResNet-18/50, SqueezeNet, DenseNet-201, and Xception). Initially, the dataset images were resized to 224×224 pixels, followed by data augmentation with the help of rotation, reflection, and scaling operations. The balanced public CXR dataset from the GitHub and Kaggle repositories was divided into 3224 COVID-19-infected and 3224 healthy cases. In order to reduce bias, the dataset images were randomly distributed into training and validation datasets, with 80% of data for system training and 20% for system validation. Within the deep boosted hybrid learning framework, transfer-learning-based fine-tuned deep COVID-RENet-1 and 2 were used for feature extraction, followed by eigenvector-based transformation using principal component analysis (PCA) and SVM as the final classifier. The authors observed that the deep hybrid learning models outperformed the well-established CNN models, while the deep boosted hybrid learning models performed the most effectively.

Al-Waisy et al. [41] also proposed an ensemble COVID-19-DeepNet system based on the integration of a deep belief network and a convolutional deep belief network. The purpose of the system was to differentiate between healthy and COVID-19-infected CXR images. A public dataset consisting of 24,000 CXR images was obtained from GitHub, Radiopaedia, the Italian Society of Medical and Interventional Radiology, the Radiological Society of North America, and Kaggle. A balanced dataset with 400 images for COVID-19 and 400 images for normal cases was used to obtain an augmented dataset with around 24,000 images of 128×128 pixels in size. Some of the different operations used for data augmentation included random sampling of data into mutually exclusive training, contrast balancing using contrast-limited adaptive histogram equalization (CLAHE), noise removal with the help of a Butterworth bandpass filter, horizontal flipping, and five-degree rotation. To avoid bias in the results, the data were randomly distributed into training (75%), testing (25%), and validation (10%) sets. The results from two different deep learning approaches—namely, the deep belief network and the convolutional deep belief network—were fused together by computing predicted probability scores, and the final decision regarding classification was based on decision-level fusion. The proposed model was compared with the state-of-the-art approaches (e.g., COVID-Net, ResNet 50, COVID-ResNet, and EfficientNet-B3 model), and it outperformed other approaches, with a detection accuracy rate of 99.93%.

In another study, Mazaar et al. [42] proposed a hybrid model that exploits the potential of deep learning and transfer learning techniques to develop accurate and robust models for detecting COVID-19. The preprocessing function involved transforming the dimensions of the original dataset into 120×120 pixels. Three basic blocks were used for building seven different variants. These building blocks included (1) a CNN block, (2) a transfer learning

block using VGG16 or VGG19, and (3) a machine learning block. A publicly available dataset obtained from Kaggle was used with a private dataset from Asir Hospital, Abha, Saudi Arabia, which consisted of a total of 4103 CXR images labelled as COVID-19, viral, or normal. To avoid bias in the performance, a random distribution of data into training (3279 images) and validation (820 images) sets was performed. The hybrid model using VGG-16 with transfer learning was able to outperform all other variants, with an accuracy of 97.6%.

Bhowal et al. [38] applied the Choquet integral for ensemble deep learning models and evaluated the fuzzy measures for each classifier using coalition game theory (Shapley value), information theory, and lambda fuzzy approximation. Some of the preprocessing functions involved downscaling of original images to 512×512 pixels using bicubic interpolation and color-space translation to convert the original dataset images from RGB to grayscale. For deep feature extraction, three standard deep learning models with pre-trained weights on the ImageNet dataset were used—namely, VGG-16, Xception, and Inception-v3. The authors obtained a public dataset of CXR images from two repositories (namely, GitHub and Kaggle), which was used to create a single dataset called the Novel COVID-19 Chest X-ray Repository. The dataset consisted of 752 COVID-19-infected, 1584 viral pneumonia, and 1639 normal CXR images. Random sampling of images was used for the training, validation, and testing datasets, with the ratio of training, testing, and validation data as 77%:20%:3%, respectively. Finally, the results were combined to generate a final output as one of the following three classes: COVID-19, pneumonia, or normal. The ensemble model outperformed many recent methods, with an area under the curve (AUC) of 0.97 and a validation accuracy of approximately 98.99%.

Another method used ECG data for the classification of COVID-19 [43]. The proposed automated tool consisted of four steps: ECG trace image preprocessing, deep feature extraction and feature incorporation, hybrid feature selection, and classification. In the preprocessing stage, the dimensionality of the input image was reduced from the original image (ranging between 952×1232 and 2213×1572 pixels) to 224×224 pixels for ResNet-50, ShuffleNet, and MobileNet and 229×229 pixels for Inception-v3, Xception, and Inception-ResNet. The dataset consisted of 250 scans of cases with COVID-19, 848 for trace records of cases with a present or former myocardial infarction and irregular heartbeat, and 859 normal images. To avoid the classification bias arising from class imbalances, the number of images per class was kept equal (i.e., 250 images per class) for the binary and multi-class classification. The proposed method for hybrid feature extraction utilized 10 different DL-based approaches. These networks include Inception-ResNet, ResNet-18, ResNet-50, ShuffleNet, Inception-v3, MobileNet, Xception, Dark-Net-19, DarkNet-53, and DenseNet-201. The fully connected deep features obtained from the different networks were followed by hybrid feature selection utilizing a forward search with a random forest classifier. The performance of the proposed system was compared with ResNet-50, ShuffleNet, and MobileNet to reveal the superior performance of the proposed method.

Rajaraman et al. [44] utilized deep ensemble learning via a proposed network that consisted of two stages of pre-training followed by the actual deep learning architecture developed specifically for the recognition of COVID-19, which provided an input CXR of 256×256 pixels in size. The first stage of pre-training was composed of different deep learning models (i.e., ResNet18, VGG-16, VGG-19, Xception, Inception-V3, Dense-Net-121, MobileNet-V2, and NasNet-Mobile) with pre-trained weights on an ImageNet dataset, followed by successive layers for zero-padding, fully convolutional layers, pooling layers, dropout, and softmax activation. The second stage of pre-training was based on the first-stage pre-training model, followed by additional pooling, dropout, and softmax activation layers. The two stages of pre-training were concatenated with additional pooling and activation layers to provide the final recognition results. The ensemble learning leveraged the top three, top five, and top seven results from the proposed models with majority voting, simple averaging, and weighted averaging to obtain the final results. A total of 720 CXRs from the Montreal COVID-19, Twitter-COVID-19, RSNA, CheXpert, and NIH

collections were used in this study. These data were randomly split into 80% for training and 20% for validation to avoid bias in the performance of the proposed system. The proposed system with deep ensemble learning was able to achieve an accuracy of 90.97%.

The proposed deep stacked ensemble in [45] requires the preprocessing stage to reduce the size of the original images to 224×224 pixels, which is given as input to the feature extraction module leveraging different deep learning models (i.e., ResNet, MobileNet, Inception, DenseNet, and NasNet). The two models with the best performance are stacked to form a deep ensemble model for COVID-19 prediction. The dataset used in this study contained a total of 2905 images (219 COVID-19 images and 2686 normal class images), which were randomly divided into training (80%) and test sets (20%) by maintaining consistency in class labels at each partition and minimizing classification bias. However, the overall class imbalance was present throughout the training data. The proposed stacked ensemble was able to provide the best accuracy result of 95.1% on the given dataset. The development of COVID-Net and COVID-EnsembleNet was discussed in [46]. The COVID-Net architecture was based on four pairs of consecutive convolution and maximum-pooling layers, with filters ranging from 16 in the first pair to 128 in the fourth pair. These convolutional layer blocks were followed by the flattening operation and two fully connected layers with 128 and 2 filters, where each fully connected layer employed the softmax regression activation function. The COVID-EnsembleNet was constructed using the proposed COVID-Net architecture along with the existing VGG-16 architecture. The dataset used in this study contained 1281 positive cases and 3269 negative cases, with a random distribution of data into training (3641 images), testing (455 images), and validation (455 images) sets to avoid classification bias. Some of the preprocessing functions included image resizing to 224×224 pixels, image normalization for faster network convergence, and data augmentation techniques such as random horizontal and vertical image flipping, followed by a random 72-degree rotation. The proposed ensemble model was able to provide binary accuracy of 99.56% and multiclass accuracy of 97.56%.

A novel ensemble deep learning model for the detection of COVID-19 from CT images was developed in [47], utilizing 2500 lung CT images from COVID-19 patients, along with 2500 CT images of lung tumors and 2500 CT images of normal lungs from a hospital. These images were randomly distributed into 6000 images for training and 1500 images for validation. Transfer learning was used for model parameter initialization, followed by deep feature extraction using three pre-trained deep convolutional neural network models, namely, AlexNet, GoogLeNet, and ResNet. The ensemble classifier EDL-COVID was obtained via relative majority voting of the aforementioned individual classifiers. With an average accuracy of 99.1%, precision of 99.1%, and recall of 99.6%, the ensemble classifier was able to outperform the results of the individual classifiers.

Taken together, the results of these previous studies strongly indicate that deep learning models built using multiple learning algorithms and, in particular, with ensemble learning classifiers, can benefit from improved performance. Nevertheless, it is worth mentioning that these studies have some limitations, not least because building ensemble models based on a single modality (in this case, mainly CT over CXR) leads to clinical challenges because, practically speaking, it is not preferable to expose the patient to CT radiation, and CXR imaging is not sufficient for the diagnosis of COVID-19 without any additional data [48].

4.2. Identification of COVID-19 in CXR Images Using Deep Learning Models

A substantial body of research has been published recently on models for the detection and classification of COVID-19 using binary/multi-class classifiers for CXR images with off-the-shelf networks. In this review, we chose not to address studies with solutions involving binary classification. This is because the ability of AI systems to differentiate between the different classes has considerably improved, as these systems are able to learn from diverse data belonging to different classes [49]. Abdelsamea et al. [50] proposed the use of a CNN called Decompose, Transfer, and Compose (DeTraC); this method helps to

deal with any irregularities and the limited availability of annotated CXR images. Their study used data from different sources—80 cases of normal CXR images with 4020×4892 pixels from the Japanese Society of Radiological Technology (JSRT), along with 105 and 11 cases of COVID-19 and SARS, respectively, with 4248×3480 pixels. Data augmentation techniques were used to generate 1764 samples from the original limited dataset. The different augmentation techniques included random up/down and right/left flipping, random translation, and rotation using five different angles. A histogram modification technique was applied to the augmented images to enhance the contrast of the images. The augmented dataset was randomly divided into 70% training and 30% validation sets to minimize classification bias. An AlexNet network based on shallow learning was used for the class decomposition layer, and different ImageNet pre-trained CNN networks were used for the transfer learning stage. The high-dimensional feature space was substantially reduced using PCA. The highest accuracy was achieved by VGG19 in DeTraC. The accuracy rate after applying the model was 93.1%.

Brunese et al. [51] implemented a deep learning model using a dataset containing 6523 CXR images collected from three different CXR image sources. The dataset was labelled with 250 COVID-19 images, 2753 images belonging to patients with other pulmonary diseases, and 3520 normal patients. The preprocessing stage allowed the reduction of the image dimensions to 224×224 pixels as well as random distribution into training (2000 images), testing (1100 images), and validation (803 images) sets. Data augmentation was performed via random clockwise and counterclockwise rotation by 15 degrees. The proposed approach is based on a threefold method: Firstly, a process for the detection of any type of pneumonia in the CXR image is conducted. Secondly, if the lungs are not normal, then the system tries to classify between COVID-19 and other pneumonia. Finally, in the event of COVID-19 classification, the images are used to identify the area in the CXR that indicates the presence of COVID-19. The researchers applied the VGG-16 CNN model with 16 layers, which yielded an accuracy of 98% for the detection of COVID 19.

The limited number of CXR images that exist for COVID-19 research constituted the focus of the study undertaken by Loey et al. [52]. The dataset used in this research was created by Dr. Joseph Cohen from the University of Montreal; it consists of total 307 CXR images, including 69 from COVID-19 patients, while the remainder belong to normal, bacterial, and viral pneumonia patients. The proposed model consists of two stages: In the first stage, a generative adversarial network (GAN) is used to generate additional images to increase the size of the existing limited dataset. In the second stage, deep transfer learning is used in the training, validation, and testing phases of the proposed model. For their investigation, the researchers selected the following deep transfer learning models: AlexNet, GoogLeNet, and ResNet18. Each of these networks can take an input image of 512×512 pixels in size. The choice of the models was due to their architectures, which contain a small number of layers, thereby reducing the processing time, the memory consumed, and the proposed model's complexity. The highest test accuracy (80.6%) for a scenario that included all four classes was achieved by the GoogLeNet framework.

Oh et al. [53] proposed a batch-based CNN approach with a probabilistic Grad-CAM saliency map that was compatible with a batch-based approach. This approach considered the limited availability of CXR images for the classification of COVID-19. The researchers used a public dataset containing 502 CXR images, including 180 COVID-19 images, 191 normal images and 113 belonging to three other classes: viral pneumonia, bacterial pneumonia, and tuberculosis. To reduce classification bias in the system, the data were randomly distributed into training (354 images), validation (49 images), and test (99 images) sets. The CXR images were first preprocessed for data normalization, and the images were resized to 224×224 pixels to obtain the preprocessed data. The preprocessed data were then fed into a segmentation network, from which lung areas could be extracted for network training and classification using patch-by-patch training and inference. The final decision regarding the network classification was based on majority voting. The experimental results were stable with a small dataset and achieved 88.9% accuracy using the proposed

batch-based CNN approach. The effects of patch size, different segmentation methods (e.g., U-Net, FC-DenseNet63, FC-DenseNet-103), and training dataset size were also evaluated in relation to the overall performance of the proposed system.

An attention-based VGG-16 model was used for the classification of COVID-19 in [54]. A total of 4901 image data were used in this study from three different datasets with their own sets of unique challenges and limitations. The original image size was reduced in dimensions to 224×224 pixels. To reduce the classification bias, data from each dataset were randomly divided into 70% training and 30% validation sets. The proposed attention-based method was based on four main building blocks: an attention module, a convolution module, FC layers, and a softmax classifier. The attention module was used to capture the spatial relationships of visual clues in the COVID-19 CXR images. The output from the attention module was given as an input to both maximum pooling and average pooling on the input tensor, which was the fourth pooling layer of the VGG-16 model in the proposed method. After that, these two resultant tensors (maximum-pooled 2D tensor and average-pooled 2D tensor) were concatenated to one another to perform a convolution, followed by the fully connected layers and the softmax classifier to give the final output. Based on the inherent characteristics and limitations of each dataset, the performance of the proposed approach varied, with the accuracy ranging between 80% and 87%.

A deep learning pipeline for the diagnosis and discrimination of viral, non-viral, and COVID-19 pneumonia CXR images was developed in [55]. The dataset used in this study included data from two public datasets: CheXpert and CC-CXRI. The total CXR images included 1571 COVID-19 images, 5656 viral pneumonia images, 11,591 other pneumonia images, and 10,477 normal images. The CXR images were resized to 512×512 pixels. The common thoracic disease detection module classified the standardized CXR images into 14 different classes. Multiple external validations were performed, with an average ratio of random training, validation, and testing data distribution amounting to 80%, 10%, and 10%, respectively. The following three modules provide the main functionality of the proposed deep learning pipeline: (1) a CXR standardization module, (2) a common thoracic disease detection module, and (3) a final pneumonia analysis module. The pneumonia analysis module consists of a lung-lesion segmentation model and a final classification model to estimate the subtype of pneumonia and the severity of COVID-19. The lung-lesion segmentation training was based on 1016 CXR images that were manually segmented into four categories and common lesions to develop a model that could differentiate between COVID-19 and other types of pneumonia. The final classification model was developed based on the DenseNet-121 architecture, which was able to perform lung-lesion segmentation and pneumonia diagnosis. The results showed that the proposed deep learning pipeline was able to predict COVID-19 pneumonia with an AUC of 86.8%, along with a recall of 80.65% and precision of 82.05%.

Fusion Module–Hand-Crafted Features–Deep Learning Features (FM–HCF–DLF) is another model for COVID-19 CXR classification given in [56]. The study made use of an imbalanced dataset containing 220 images for COVID-19, 27 for normal lungs, 11 for SARS, and 15 for pneumonia. In the preprocessing part of the system, a 1D Gaussian operator was used for noise removal and image smoothing for the input images, followed by resizing the original images down to 299×299 pixels. The FM model incorporates the fusion of hand-crafted features with the help of local binary patterns (LBPs) and deep learning. The deep learning (DL) features are computed using the convolutional neural network (CNN)-based Inception-v3 framework, followed by a multilayer perceptron (MLP) to provide the final output classification. The proposed method's performance was compared with that of traditional ML algorithms to highlight the superior performance of the proposed model, which achieved 94.08% accuracy.

CVDNet is a deep convolutional neural network (CNN) model to distinguish COVID-19 infection from normal lungs and other pneumonia cases using chest X-ray images, as presented in [57]. The proposed architecture is based on a residual neural network and is constructed by using two parallel levels with different kernel sizes to capture the local

and global features of the inputs. This model is trained on a publicly available dataset containing a combination of 219 COVID-19, 1341 normal, and 1345 viral pneumonia CXR images. The images are randomly distributed into 70% for training, 10% for validation, and 20% for testing to reduce classification bias. Some of the preprocessing functions include cropping and resizing the original images to provide input images of 256×256 pixels in size. Two streams with four parallel residual blocks are used for deep feature extraction, followed by feature concatenation leading to a final residual block, and ending with a fully connected layer and a softmax classifier. The proposed system can provide an accuracy of 96.69%, which is comparable to state-of-the-art methods for the classification of COVID-19.

Azemin et al. [58] used a ResNet-101-based deep learning model. A total of 10,359 images were used in this study, of which 154 were from COVID-19. Despite the over-whelming imbalance in the dataset used in this study, the authors failed to provide an adequate strategy for mitigating the effects of data imbalance on the overall system performance. The input data were randomly distributed into training (3063 images), validation (1313 images), and test (5828 images) sets to evaluate the performance of the proposed system. The best accuracy provided by the proposed system was 71.9%, which is considerably lower than that achieved in many of the studies discussed in this section.

Despite the findings of these studies, there are notable limitations in terms of small sample sizes, the use of too few pneumonia classes, and dependence on one modality. These limitations have implications for the potential use of these research findings in real-world healthcare applications.

4.3. Identification of COVID-19 in Chest CT Images Using Deep Learning Models

Multiple studies have used chest CT scans to detect and classify COVID-19 in images. CT images are chosen due to the ease of detecting abnormal regions in infected lungs, along with their accuracy in extracting features of pneumonia. However, in practice, CT scans are not the most suitable option for COVID-19 diagnosis due to the limited availability of the necessary equipment and the high cost associated with the process. In addition, both the American College of Radiology and the Italian Society of Radiology (SIRM) do not recommend chest CT scans as a screening tool for COVID-19 [59]. Based on these guidelines, many studies refrain from using CT images as the input modality for the classification of COVID-19. Consequently, this subsection of the paper considers studies that ultimately decided to rely on CT scans as the primary modality for data collection for the development of their deep-learning-based models for the classification of COVID-19.

Amyar et al. [60] designed a slice-level classification model for three learning tasks—segmentation, classification, and image reconstruction—for CT scan images. A CNN model was used, the architecture of which consists of a common encoder and two decoders based on U-Net. A common encoder module was used for the three tasks, taking a CT scan as its input, and its output was then used for image reconstruction through the first decoder module, followed by the segmentation task completed by the second decoder module and multi-class classification for COVID-19 (and other lung diseases) performed by the multilayer perceptron. Different state-of-the-art models were compared to this classification model; the authors used AlexNet, VGG-16, VGG-19, ResNet-50, 169-layer DenseNet, InceptionV3, Inception-ResNet v2, and EfficientNet. The dataset was collected from three hospitals and contained 1369 images divided into three classes: 425 for normal cases, 449 for COVID-19 cases, and 495 for other infections. In order to avoid classification bias, the original balanced dataset was randomly distributed into training (1069 images), validation (150 images), and test (150 images) sets for performance evaluation. The preprocessing stage involved the conversion of original images down to $256 \times 256 \times 3$ pixels and $512 \times 512 \times 3$ pixels, and these two sizes were used as inputs to the proposed models, with the smaller images providing a comparatively better performance. The proposed model outperformed state-of-the-art methods for the classification and segmentation of COVID-19, achieving an accuracy of 94.67% for the classification task and a dice coefficient of 88.0% for the segmentation task.

Li et al. [61] developed a 3D deep learning framework using the architecture of COVNet to identify COVID-19 from CT scan images. The initial input data of 3D CT scans were preprocessed to reduce the original dimensions to $512 \times 512 \times 3$ pixels, followed by the extraction of the region of interest (i.e., lungs) using a U-Net-based segmentation model. The preprocessed image was given as an input to COVNet to extract visual features from 2D local and 3D global images. The COVNet framework leverages ResNet-50 as the backbone model for deep feature extraction from CT slices, which are combined using a maximum pooling operation, followed by a fully connected layer and softmax activation function to generate a probability score for classification (COVID-19, CAP, and non-pneumonia). The study used a dataset obtained from six hospitals, amounting to 4356 CT scan images divided into three classes: 1296 for COVID-19, 1735 for CAP, and 1325 for non-pneumonia. To remove classification bias, the original balanced dataset was randomly distributed into 3918 CT images for model training and 434 CT images for model testing. The proposed framework achieved an AUC of 96% for the identification of COVID-19 and 95% for CAP on CT scan images.

Another study proposed a DL-based pipeline for CT images called CoviWavNet for the automatic diagnosis of COVID-19 [62]. In the preprocessing phase of the proposed system, data from two different datasets are concatenated and augmented to increase the size of each of the datasets and reduce overfitting. Some of the augmentation processes performed include scaling, shearing, rotation, flipping on the x- and y-axes, and random translation in the x- and y-directions. Finally, the augmented images are resized to $227 \times 227 \times 3$ pixels. The proposed CoviWavNet uses multilevel discrete wavelet transform (DWT) and heatmaps of the approximation levels to train three different ResNet models for classification. To examine the effect of the combination of spatial and spectral-temporal information on diagnostic accuracy, deep spectral-temporal features are generated from ResNet using transfer learning and integrated with deep spatial features extracted from ResNet models trained with the original CT slices. To reduce the dimensionality, the most valuable feature is selected using the minimum-redundancy-maximum-relevance (mRMR) technique and used as inputs to three support-vector machine (SVM) classifiers. The performance of the proposed system achieves accuracy of 98.62%, precision of 99.54%, an F1-score of 99.62%, and recall of 99.69%.

Alshazly et al. [63] applied two separate CT datasets for developing a deep-learning-based system for the classification of COVID-19 using the most advanced networks, such as SqueezeNet, Inception, ResNet, ResNeXt, Xception, ShuffleNet, and DenseNet. The original images from the two datasets were randomly distributed into 60% for training and 40% for testing, and were resized to $253 \times 349 \times 3$ pixels. Data augmentation methods were implemented to effectively increase the number of training samples for improved generalization. Some of these methods included random horizontal flipping, normalization, cropping, blurring, Gaussian noise addition, and brightness and contrast improvement. To assess the performance of the proposed models, stratified K-fold ($K = 5$) cross-validation was used to ensure class-level consistency in each of the five folds. The proposed model achieved high accuracy, with 99.4%, and an F1-score of 99.4%.

Another study proposed a combination of radiomics and artificial intelligence for the analysis of medical images using a CAD framework with four phases [64]: image preprocessing, feature extraction, feature fusion, and classification. In the first phase, the images are analyzed using two texture-based radiomic approaches: gray-level co-variance matrix (GLCM), and discrete wavelet transform (DWT). The radiomics and original CT images are resized to $227 \times 227 \times 3$ pixels before they are given as inputs for the feature generation and extraction phase. Different data augmentation techniques are performed on the radiomics and original CT data, including shearing, scaling, random translation, and rotation. In the second phase of the proposed framework, three residual networks (ResNets) are used for deep feature extraction. In the third phase, these features are fused together using discrete cosine transform (DCT). In the fourth phase, three machine learning classifiers are used to perform the classification procedure. The dataset used in this study

was a benchmark 2D CT dataset containing a total of 2482 CT images, with 1252 for COVID-19, while the remaining 1230 were non-COVID-19. The proposed framework was able to provide accuracy of 99.6% and an F1-score of 99.6%.

The uAI Intelligent Assistant Analysis System (IAAS) is a deep-learning-based software platform developed by United Imaging Medical Technology Company Limited (Shanghai, China) for the classification of COVID-19 [65]. The IAAS software has an underlying deep learning architecture consisting of a modified 3D convolutional neural network and a combined V-Net with bottleneck structures. A total of 2460 images were used in this study (2215 for COVID-19, while the rest were normal). The authors failed to provide details regarding the data imbalance issues with respect to the distribution of the training and validation sets. At the same time, detailed information regarding preprocessing procedures and final performance was also not provided. One potential reason could be that the main purpose of the study was to assess the feasibility of utilizing the uAI IAAS as a diagnostic tool for COVID-19 from CT images.

Shah et al. [66] utilized different deep learning frameworks to differentiate between the CT images of COVID-19 and non-COVID-19 patients. CTnet-10 was designed for the diagnosis of COVID-19, developed using six successive layers of convolution and maximum pooling, followed by flattening and dropout layers, and ending in a softmax classification layer. Some of the other models that were tested included DenseNet-169, VGG-16, ResNet-50, InceptionV3, and VGG 19. The performance of the different deep learning models was assessed to highlight the most suitable option for the classification of COVID-19. For the CT-net model, the input image size was 128×128 pixels. For the VGG-19 model, the image dimensions used were $224 \times 224 \times 3$ pixels. The images were randomly distributed into 80% for training, 10% for validation, and 10% for testing. CT-net provided an accuracy of 82.1%, while the VGG-19 model was able to provide an accuracy of 94.5%.

Wang et al. [67] developed a deep-learning-based framework for CT scans of COVID-19 cases. The dataset used in this study was based on 1065 CT images of COVID-19 patients as well as patients who had a prior history of typical viral pneumonia. Unfortunately, the authors did not mention the class-wise distribution of data for the different lung diseases and COVID-19, limiting the ability to assess class balance for this study. The dataset was randomly divided into one training subset (320 CT images), one internal validation subset (455 CT images), and one external validation cohort (290 CT images). The proposed architecture consisted of the preprocessing module, deep feature extraction module, and final classification module. The preprocessing module involved the conversion of the original images into grayscale, followed by grayscale binarization, background area filling, color reversal, and ROI selection. The preprocessed images were rescaled to $299 \times 299 \times 3$ pixels before being used for deep feature extraction. The transfer learning process involved training with a predefined model, which in this study was the GoogLeNet-Inception-v3 deep learning architecture. After feature extraction, the final step was to provide multi-class classification using an ensemble of classifiers to improve performance. The proposed model was able to provide an accuracy of 89.5% on internal validation and 79.3% on external validation, which is significantly lower than the performance provided by prior studies discussed in this section.

As this review's findings indicate, many studies have been published that rely on CT scans to diagnose COVID-19. However, these studies not only suffer from limitations, but also produce models that cannot be applied practically and economically in routine clinical practice. This is especially due to the cost and limited availability of CT scans, as well as the circumstances of the ongoing pandemic. In comparison, CXR has better availability and ease of execution, and minimizes in-hospital transmission; it is neither time-consuming with lengthy waiting times, such as the CT scan procedure, nor does it require wearing a special suit. In addition, CT scans are often optimally applied in critical cases where the infected lungs are very clear for radiologists.

4.4. Identification of COVID-19 from CXR Images using Deep Learning Models in a Hierarchical Classification Scenario

One of the earliest studies that applied hierarchical image classification using deep learning was carried out in 2015 [68]. A hierarchical deep CNN (HD-CNNs) model was proposed by embedding deep CNNs into a two-level category hierarchy, where the easily distinguishable classes were classified using a coarse-category classifier and difficult classes were classified using fine-category classifiers. Hierarchical classification models have a noticeable impact on reducing classification errors.

Despite their advantages in classifying COVID-19 due to the similarity in symptoms and physical features shared with many other diseases (e.g., viral pneumonia, bacterial pneumonia, tuberculosis), only one study by Pereira et al. [69] has used a deep hierarchical model for the classification of COVID-19. The study leveraged deep learning with a pre-trained CNN model to classify 1144 CXR images for 7 flat labels and 14 hierarchical labels of multi-class pneumonia classification, including COVID-19 and healthy lungs. To reduce classification bias, the dataset was randomly distributed into training and validation sets of 70% and 30%, respectively. To reduce the effect of imbalanced data on the final performance of the classifier, different resampling algorithms were employed (i.e., ADASYN, SMOTE, SMOTE-B1, SMOTE-B2, AI-KNN, ENN, RENN, TomekLinks (TL), and SMOTE + TL). In the feature extraction phase, different hand-crafted features were used, including oriented basic image features (oBIFs), locally encoded transform feature histograms (LETRISTs), and local directional numbers (LDNs), using local phase quantization (LPQ) and deep learning (Inception-v3) methods. Both early and late fusion techniques were employed, followed by data resampling and multi-class and hierarchical classifiers. For multi-class classification, different traditional ML methods were employed, namely, support-vector machines, multilayer perceptrons, random forests, and decision trees. For the hierarchical classification, Clus-HMC—an unsupervised predictive clustering algorithm—was selected as the hierarchical classifier. The experiment proved that the hierarchical structure for the classification of COVID-19, which was tested on the RYDLS-20 dataset, achieved a higher F1-score (89%) with early fusion and BSIF, EQP, and LPQ features, compared to a flat structure for classification (83%). However, these findings are limited due to the way in which the authors classified pneumonia, as well as the extraction of features from a single modality, which were applied on a small sample size.

4.5. Identification of COVID-19 in Chest CT and CXR Images Using Multimodal Deep Learning Models

Models based on data from diverse radiological imaging modalities can achieve a higher accuracy rate compared to models that use only one imaging modality. However, analyzing CXR and CT images for the same patient in the context of the COVID-19 pandemic is impractical in clinical practice and is considered a significant gap in prior studies.

Horry et al. [70] proposed multimodal imaging data to detect COVID-19 by combining CXR, ultrasound, and CT scans using a VGG-19 model. A publicly accessible dataset gathered from different data sources, containing 1118 COVID-19 images, 996 pneumonia images, and 60,533 images with no findings, was used in the experiment. To prevent bias in classification, the original dataset was randomly distributed into 20% for validation and 80% for training. To enhance different features in the original images, the contrastive limited adaptive histogram equalization (CLAHE) method was used. Since the different deep learning models have different limitations in terms of image size, the original image dataset had to be modified due to the utilization of different deep learning models, and the input image sized varied from 224×224 pixels for VGG variants to 299×299 pixels for Inception-v3. The performance of the proposed approach varied based on the type of data modality being used for system training, such that the highest performance was provided by the proposed models trained on ultrasound data and the lowest performance was provided for the proposed models trained on CT scan data. The results indicate that

ultrasound images were more accurate compared to CXR and CT scans, with a precision of 100% compared to 86% and 84%, respectively.

Vinod et al. [71] demonstrated the possibility of integrating CXR images with CT scans in deep learning models to diagnose COVID-19 patients. A deep COVIX-Net model was used to classify CXR and CT images into one of the following three classes: normal, COVID-19, and pneumonia. The dataset was obtained from the Kaggle and GitHub repositories, consisting of 9000 CXR images (3000 pneumonia, 3000 COVID-19, and 3000 normal) and 6000 CT scans (3000 pneumonia and 3000 COVID-19). The original dataset was randomly distributed between 80% training and 20% validation subsets; this ensured that the developed system was able to minimize the classification bias. The images from the original dataset were converted to 224×224 pixels, and these images were subjected to different techniques for feature extraction. The features were extracted from the medical images using the following techniques: texture, gray-level co-occurrence matrix (GLCM), gray-level difference method (GLDM), fast Fourier transform (FFT), and discrete wavelet transform (DWT). From the different feature extraction techniques used, some of the different statistical features used for model training included average, skewness, kurtosis, energy, average deviation, dimension, RMS, consistency, average gradient, minimum, and median. Unfortunately, there is no further information regarding the structure of deep COVIX-Net shown in the paper. The outputs from feature extraction from different techniques were given as inputs to a random forest classifier to provide the final classification output. The proposed model achieved promising outcomes, with 96.8% accuracy for the CXR images and 97% for CT scans.

Yadav et al. [72] proposed a deep unsupervised framework (Lung-GANs) to classify lung diseases based on chest CT scans and CXR images using unlabeled data. The proposed method involved the use of six large, publicly available datasets consisting of 38,155 CXR and CT scan images from healthy, sick, tuberculosis (TB), viral pneumonia, COVID-19, and non-COVID-19 classes. A number of different image preprocessing procedures were leveraged for modifying the original dataset images, such as color mode conversion, image resizing (original images were converted to 512×512 pixels), and image normalization. After that, the preprocessed images were randomly distributed into training (70%) and validation (30%) sets to reduce classification bias. The generator module in Lung-GAN takes in a 100-dimensional vector as its input and outputs a single 512×512 image. The discriminator module in Lung-GAN is a CNN architecture that can differentiate between real and synthesized images of 512×512 pixels in size as inputs, and its output is given as a probability value specifying whether the image is real or not. The authors developed an ensemble of classifiers, with linear support-vector classification (SVC) and random forests serving as the base classifiers, which were combined with predictions from each classifier to produce the final result. The CNN architecture was used for both models. The performance of a GAN-based single framework for all binary classifications on all datasets achieved higher accuracy compared to other state-of-the-art unsupervised models in this area. The breakdown of the performance of Lung-GAN for the different datasets showed diverse results, which varied considerably from one dataset to another. The accuracy values ranged between 94% and 99.5%, with an average accuracy of 97.6%. This shows that the proposed system is sensitive to variations in dataset characteristics (i.e., noise, image dimensions, quality and quantity of data, metadata).

Kalaiselvi et al. [73] designed three artificial intelligence models: sANN ML, using machine learning; pVGG TL, using transfer learning; and pCNN DL, using deep learning. The purpose of each model was to detect COVID-19 from CXR and CT scan images. Each model used the ReLU and E-Tanh activation functions. The public dataset was collected from different research and medical centers, and contained a total of 650 CXR images and 746 CT images divided into positive and negative COVID-19 cases. The training and validation sets contained a total of 625 images for training and 10 images for validation. Despite the considerable data imbalance, the study failed to shed light on the techniques to be used for mitigating the underlying issues. For the different models highlighted above,

a VGG-16 network, ANNs with different activation functions, and CNNs with different activation functions were used to provide the backbone for classification. The sANN ML and pVGG TL models achieved high accuracy (100%) in the detection of COVID-19 from CXR images. However, pCNN DL did not perform well, which could be attributed to the small size of the dataset. Notably, each model had a low detection performance for CT scan images compared to CXR images. In addition, the E-Tanh activation function yielded positive results for CXR images.

Another study made use of 8879 CXR and 3724 CT images for the training and development of the proposed deep learning model [74]. The authors did not specify the distribution of the original data between training, test, and validation sets. In the preprocessing stage of the proposed model, contrast-limited adaptive histogram equalization (CLAHE) was used for contrast and feature enhancement. The use of different data augmentation strategies, such as random transformations (i.e., rotation, horizontal and vertical translations, zooming, and shearing), ensured that the system could be generalized well to unknown data. All of the data were resized to 512×512 pixels before being used in the system. In the first stage of the proposed model, an Inception-v3 deep model was trained for the recognition of COVID-19 using multimodal learning by leveraging data from both modalities, i.e., X-ray and CT scan. The second stage was based on a convolutional neural network architecture for recognizing three types of lung disease. The third stage was based on transfer learning from pulmonary nodule segmentation in CT scans to produce binary masks for segmenting similar regions in the given data. Ultimately, this method showed an accuracy of 99.4%, precision of 99.5%, recall of 99.1%, and F1-score of 99.3%.

Ibrahim et al. [75] examined the effects of four deep learning models for the classification of COVID-19 using multimodal data consisting of CXR and CT images. The first stage of the proposed system is responsible for performing different image preprocessing functions, such as resizing, image augmentation (i.e., flipping, rotation, and skewing), and random data distribution of the total dataset (75,000 images) into training (70%) and validation (30%) subsets. Unfortunately, the authors did not provide the distribution of the dataset, limiting our ability to assess the level of balance in the dataset between the individual classes. The input images were resized to 224×224 pixels, which is a standard size that is suitable for input to the different deep learning models used in this study. The second and third stages are responsible for deep feature extraction and image classification using the following four networks: VGG19-CNN, ResNet152-v2, ResNet152-v2 + gated recurrent unit (GRU), and ResNet152-v2 + bidirectional GRU (Bi-GRU). Of the four networks mentioned, the best results were provided by ResNet152-v2 + bidirectional GRU, with an accuracy of 98%, precision of 99.5%, recall of 98%, and F1-score of 98.24%.

Sharma et al. [76] used the VGG-16 deep learning model for the classification of COVID-19. Before identifying the different lung infections, different preprocessing functions and data augmentation operations were performed to minimize the classification bias, enhance the system's generalizability, and improve the quantity and diversity of the data. Open-source data were acquired for this study, namely, the COVIDx CT-2A dataset, which includes 194,922 images from 3745 patients. From the open-source data, original data were used and randomly distributed into 80% for training and 20% for testing. The authors of this paper failed to highlight the class-wise distribution of the dataset used in their study, limiting our ability to assess level of class imbalance. Augmented images were converted into 512×512 pixels before being used for model training and validation. For the classification of COVID-19, the proposed model was able to provide the best performance, with 99.2% accuracy, 99.6% precision, and 99.8% recall.

A deep transfer learning algorithm was proposed in another study [77] to provide a rapid-response-based system for the classification of COVID-19 using multimodal data for CXR and CT images. Data from different publicly available sources were consolidated and utilized in this study. For example, a total of 6111 CXR images were acquired from two separate sources. Similarly, another data source was used for acquiring a total of 1252 CT scans. In the preprocessing stage, the original images were converted into 512×512 pixels.

The data were randomly split into training, validation, and testing subsets of 64%, 20%, and 16%, respectively. VGG-19 was used for transfer learning in this study; this deep learning model is based on 16 convolution layers and 3 fully connected layers, followed by 5 maximum-pooling layers and a softmax layer. The VGG-19 model is followed by the Grad-CAM model, which takes an image input to provide improved visualization output for detecting regions of interest. Once the predicted label has been calculated using the VGG-19 model, Grad-CAM is applied to the last convolutional layer of the VGG-19 model. Based on the different experiments performed, the best results obtained by the proposed algorithm provided an accuracy of 95.61%, precision of 88%, recall of 97%, and F1-score of 92%.

In another study [78], two separate deep learning models were used for the classification of COVID-19. The major difference between the two models was in terms of the data modalities used: the first model used CNN with CT and X-ray images separately, whereas the second model used CNN with both CT and X-ray images simultaneously. The dataset used in this study contained CXR and CT images divided into three separate classes: COVID-19, normal lungs, and pneumonia. The data were randomly distributed into training (3135 images) and testing (900 images) subsets. In the preprocessing stage, the sizes of the CXR and CT images were reduced to 299×299 pixels and 512×512 pixels, respectively. In the first model, the architecture consisted of different convolutional layers, pooling layers, dropout layers, and a softmax classification layer. In the second model architecture, the data from two modalities were provided to two separate architectures (where each architecture was based on the first model), followed by the concatenation of data from two models after the flattening layer and softmax classification layer. The deep learning model trained on multimodal data was able to outperform models trained on single data modalities, with accuracy, precision, and recall of 99% each.

4.6. Identification of COVID-19 in Chest CT or CXR Images with Clinical/lab Test Features Using Multimodal Deep Learning Models

The use of deep learning for multimodal data fusion has a substantial effect on medical applications. The EMIXER model is the only model that has been proposed in the literature to utilize CXR images along with radiologists' reports to classify CXR images and generate diagnostic reports [79]. The study was based on end-to-end multimodal data fusion that combined CXR images and corresponding text reports. CNN and recurrent neural network (RNN) models were used with the MIMIC-CXR dataset, which contains CXR images with associated reports. EMIXER is composed of five different modules, namely, the image generator (used to synthesize X-ray images from a prior noise distribution conditioned on label information), image-to-report decoder (used to provide an output text report when an X-ray image is given as an input), image discriminator (used to differentiate between real and synthetic X-ray images), text discriminator (used to differentiate between real and synthetic X-ray reports), and joint discriminator (used to combine X-ray images and text to discriminate between real and synthetic data). The COVID-19 classification task dataset contained 14 different classes that were resized to 128×128 pixels. Overall, EMIXER used 100,000 real images as well as 300,000 synthetic images; the addition of synthetic images led to a considerable improvement in results, with a maximum accuracy value of 92.4%. The researchers observed that EMIXER improved the classification of COVID-19 from CXR images. Unfortunately, the authors did not specify the class-level information or the training/validation split for the dataset used, limiting our ability to assess class-level imbalance and associated biases in performance.

Mei et al. [80] suggested that developing a model based only on CT scans may lead to limited negative predictive power. Therefore, the authors proposed the use of chest CT images along with clinical symptoms (e.g., fever, cough, and cough with sputum), laboratory testing (e.g., white blood cells, neutrophils, percentage neutrophils, lymphocytes, and percentage lymphocytes), and exposure history. The dataset was collected from 905 patients across 18 healthcare centers in China, where 419 cases were positive and

486 were negative, to provide a balanced dataset, with an image size of $512 \times 512 \times 3$ pixels. The original data were randomly distributed into 60% training, 10% validation, and 30% testing sets. The Inception–ResNet-v2 model was used to process the CT images, while support-vector machine (SVM), random forest, and MLP classifiers were used for clinical information. The joint model achieved better performance (AUC = 0.92) compared to the CNN model trained only on CT scans (AUC = 0.86); furthermore, it outperformed the MLP model trained on clinical information (AUC = 0.80).

Chen et al. [81] also found that the diagnostic model composed of radiological semantic features with clinical features was significantly different. The dataset used in the experiment (CT scans and clinical information) was collected from five independent hospitals in China. A total of 136 cases were labelled as COVID-19 or non-COVID-19. The researchers identified 18 radiological semantic features and 17 clinical features, including demographic information, daily body temperature, blood pressure, heart rate, clinical symptoms, history of exposure to epidemic centers, total white blood cell (WBC) counts, lymphocyte counts, lymphocyte ratios, neutrophil count, neutrophil ratios, procalcitonin (PCT), C-reactive protein (CRP) levels, and erythrocyte sedimentation rates (ESR). To compare diagnostic performance, the authors developed three models: one for clinical features, a second for CT scan features, and a third for the proposed model that combined them. The proposed model outperformed the others, with an AUC of 0.986, while the AUC values for the clinical and CT scan feature models were 0.952 and 0.969, respectively.

The combined use of CT scan images and clinical findings also occurred in another study [82]. Notably, this research also achieved promising diagnostic results for COVID-19. The study's dataset consisted of 168 patients, including 88 COVID-19-positive and 80 COVID-19-negative patients. The latter category included patients with bacterial infection, Mycobacterium tuberculosis complex, influenza virus A, influenza virus B, influenza virus B and mycoplasma, and mycoplasma. The total dataset was randomly distributed into a training subset (118 patients) and a testing subset (50 patients). The data from continuous variables within the clinical information were categorized into different classes using means, medians, and standard deviations. The categorical data were analyzed using chi-squared or Fisher's tests. Using 10 variables, logistic regression analysis was performed with ROC curves for analyzing the performance. Data from CT scans allowed the separation of visual features (i.e., the number of affected lobes and segments, segments with peripheral GGO, consolidation, air bronchograms, crazy-paving patterns, subpleural curvilinear lines, bronchiectasis, and patchy lesions), which were used to categorize between COVID-19 and non-COVID-19 patients. The predictive model utilizing clinical information with CT scan features was able to achieve good results, with an AUC of approximately 0.91.

In addition, Chen et al. [83] noted that there is a lack of awareness of the importance of biomedical features and choosing the right technical approach to diagnose COVID-19. Therefore, they proposed a late fusion deep learning–machine learning multimodal diagnostic approach to classify 214 patients with non-severe COVID-19, 148 patients with severe COVID-19, 129 patients with other viral infections, and 198 uninfected individuals. The data for 689 patients were collected from different hospitals in China, and the healthy cases (control group) were selected from patients who made up a regular annual physical examination cohort. For the preprocessing stage, the original image dataset was reduced to 512×512 pixels. The data were randomly distributed into 80% training and 20% validation subsets. The features used in the model were clinical (23), lab testing (10), and CT scan features. A customized ResNet CNN model was used for CT scan images and applied with three different ML models—random forests, SVM, and k-nearest neighbors (kNN)—for clinical findings and lab testing. The best performance with regard to all metrics was achieved when integrating SVM with ResNet, with an overall multimodal classification accuracy of 99.8%. This was a higher result compared to the use of a single modality (unimodality), which achieved 75.5% accuracy on clinical features alone, 67.7% accuracy on lab test features alone, and 90.8% accuracy on CT scan data alone.

A number of different deep learning models were used in another study, and their performance was compared to highlight the most suitable alternative for the classification of COVID-19 [84]. The private dataset was collected from King Fahad University Hospital, Dammam, KSA, and consisted of 270 cases. The preprocessing step was able to reduce the image size to 224×224 pixels (for case 2) and 300×300 pixels (for case 3), followed by random data distribution into 80% training and 20% validation sets. Different data augmentation techniques were employed, such as flipping (both horizontal and vertical), rotation, shifting, cropping, blurring, zooming, rescaling, and shearing. Three different cases were used in this study for selecting the most suitable deep learning model for the classification of COVID-19. Case 1 used clinical data for the training and validation of a 13-layer deep learning model. Case 2 used CXR data for training and validation on a diverse range of CNN architectures (i.e., ResNet, DenseNet, VGG19, EfficientNet). Case 3 used multimodal data (i.e., clinical and CXR images) with transfer learning to learn weights for EfficientNet, which was used as a backbone architecture for the classification of COVID-19. For case 3, which used multimodal data, the proposed system was able to provide the best performance (accuracy of 97%, recall of 98.6%, precision of 97.8%, and F1-score of 98.2%).

Similarly, Attaullah et al. [85] used multimodal data of symptoms and CXR images for the development of a deep-learning-based model for the classification of COVID-19. The public dataset contained a total of five classes: bacterial, COVID-19, non-COVID-19 viral, initial-stage COVID-19, and normal. The total dataset was subjected to random splitting into an 80% training set and 20% testing set. The images in the publicly available dataset were converted into 150×150 pixels as one of the main preprocessing steps. For the symptom data, the preprocessing step involved the removal of duplicate rows and null values and the application of resampling techniques to address the class imbalance issues. This step was followed by the training and validation of the preprocessed data using logistic regression and CNN models. For image data, different data augmentation techniques—zooming, rotation, and translation—were performed during CNN training. The CNN model was trained on the transformed images, and the decision tree model was trained on the labeled results of the previous two trained models to provide the final output, with an accuracy of 78.88%.

A fully automated hybrid framework based on capsule networks (CT-CAPS) and random forest classifiers was used in another study for the classification of COVID-19 using chest CT images and clinical/demographic data [86]. Private CT scans and the associated clinical/demographic data were collected for a total of 312 patients (176 COVID-19 patients, 60 pneumonia patients, and 76 normal cases) in this study. The dataset was randomly split into training (60%), validation (10%), and testing (30%) subsets. A capsule-network-based framework—namely, CT-CAPS—was used in this study, consisting of a stack of convolutional, pooling, batch normalization, and capsule network layers to extract slice-level feature maps from CT images in the first stage of the proposed model. The second stage of the proposed model leveraged the maximum pooling output of the first stage, followed by a conventional multilayer perceptron for final classification. The proposed model was able to provide 90.8% accuracy, 94.5% precision, 86% recall, and an AUC of 0.92.

Jiao et al. [87] leveraged deep learning models for the classification and severity prediction of COVID-19; a total of 1834 patients' CXR images and clinical data were used. The data were randomly distributed into 70% for training, 10% for validation, and 20% for testing. The deep learning features extracted from the model and the clinical data were used to predict the risk of COVID-19 progression. All images and masks were resized to 512×512 pixels in size and were normalized before being given as inputs to the segmentation model (i.e., U-Net). For the severity prediction model, the CXR image data were segmented using a pre-trained U-Net architecture, followed by feature extraction using a VGG-13 model with five encoder and five decoder blocks to learn the transformation from input images and binary masks. The final results for disease progression were computed based on the combined weighted sum of the individual image and clinical data

scores. The model using combined data was able to provide better prediction performance on internal and external testing.

For the development of a deep-learning-based system for the classification of COVID-19 in [88], data from 654 patients with a total of 5645 CXR images were acquired. Imaging and clinical data were used to train five longitudinal transformer-based networks, applying fivefold cross-validation. In the preprocessing stage, some of the different functions used to modify the images included inversion, padding, resizing (512×512 pixels), pixel value normalization, and scaling. The data were randomly divided into 80% for the training subset and 20% for the validation subset. The extracted features from CXRs were combined using global average and global maximum pooling operations, followed by two fully connected layers and a softmax layer to provide risk probability. The deep learning model with the combined data modalities was able to provide the best performance, with 73.2% accuracy and a 70.7% F1-score.

Despite the superior results that these studies have achieved by using multimodal models, the existing systems still suffer from some gaps. These include the fact that the radiologist's report in the EMIXER model is generated from the CXR image only, without using any clinical information that may support the report's findings. Moreover, relying on a small sample size of CT images to diagnose COVID-19 patients using multimodal deep learning is another limitation.

Table 1. Summary of the data extracted for each paper included in our review.

References	Data Used	Sample Size	Dataset Balance	Balance Strategy	Model Type	Classification Method	Performance
Chauhan et al. [39]	CXR	5232 images	No	Augmentation techniques	Ensemble	Multi-class	Accuracy = 96%
Khan et al. [40]	CXR	3224 positive, 3224 negative	Yes	-	Ensemble	Binary	Accuracy = 98% F-score = 98%
Al-Waisy et al. [41]	CXR	400 negative, 400 positive	Yes	-	Ensemble	Binary	Accuracy = 99%
Bhowal et al. [38]	CXR	752 COVID-19 1584 viral, 1639 normal	No	Augmentation techniques	Ensemble	Multi-class	AUC = 97% Accuracy = 99%
Mazaar et al. [42]	CXR	219 COVID-19, 1345 viral pneumonia, 1341 normal	No	Augmentation techniques	Ensemble	Multi-class	Accuracy = 97.8%
Attallah et al. [43]	ECG records and images	250 COVID-19, 859 normal 848 others	Yes	-	Ensemble	Multi-class	Accuracy = 91.6% Precision = 91.8% Recall = 91.6%
Rajaraman et al. [44]	CXR	360 normal 360 COVID-19	Yes	-	Ensemble	Binary	Accuracy = 90.97% AUC = 95.08 Precision = 93.94 F1 = 90.91
Bharadwaj et al. [45]	CXR	219 COVID-19, 2686 normal	No	Not mentioned	Ensemble	Binary	Accuracy = 95.1% Precision = 100% Recall = 97%
Al-Mansur et al. [46]	CXR	1281 COVID-19, 3269 healthy lungs	No	Augmentation techniques	Ensemble	Binary	Accuracy = 97.56%
Tao et al. [47]	CT-scan	2500 COVID-19, 2500 normal, 2500 lung tumors	Yes	-	Ensemble	Multi-class	Accuracy = 99.1% Precision = 99.1% Recall = 99.6%
Abdelsamea et al. [50]	CXR	105 COVID-19, 80 normal, 11 SARS	No	Augmentation techniques	Single modality	Multi-class	Accuracy = 93%

Table 1. Cont.

References	Data Used	Sample Size	Dataset Balance	Balance Strategy	Model Type	Classification Method	Performance
Brunese et al. [51]	CXR	250 COVID-19, 6273 other pulmonary diseases and normal	No	Augmentation techniques	Single modality	Multi-class	Accuracy = 98%
Loey et al. [52]	CXR	69 COVID-19, 79 normal, 79 bacterial, 79 viral	Yes	-	Single modality	Multi-class	Accuracy = 81%
Oh et al. [53]	CXR	180 COVID-19, 191 normal, 113 others	Yes	-	Single modality	Multi-class	Accuracy = 89%
Sitaula et al. [54]	CXR	4901 total images	No	Not mentioned	Single modality	Multi-class	Accuracy = 80–87% Precision = 91–96% Recall = 77–95% F1-score = 83–93%
Wang et al. [55]	CXR	1571 COVID-19, 5656 viral pneumonia, 11,591 other pneumonia, 10,477 normal	No	Augmentation techniques	Single modality	Multi-class	AUC = 86.8%, Recall = 80.65%, Precision = 82.05%
Shankar et al. [56]	CXR	220 COVID-19, 27 normal, 11 SARS, 15 pneumonia	No	Not mentioned	Single modality	Multi-class	Accuracy = 94.08%, Precision = 94.85%, F1-score = 93.2%
Ouchica et al. [57]	CXR	219 COVID-19, 1341 normal, 1345 viral	No	Not mentioned	Single modality	Multi-class	Accuracy = 96.69% Precision = 96.72% Recall = 96.84% F1-score = 96.68%
Azemin et al. [58]	CXR	154 COVID-19, 5828 no findings, 2166 opacity, 2210 no opacity	No	Not mentioned	Single modality	Binary	Accuracy = 71.9% Precision = 77.3% Recall = 71.8%
Amyar et al. [60]	CT-Scan	449 COVID-19, 425 normal, 495 others	Yes	-	Single modality	Multi-class	Accuracy = 95%
Li et al. [61]	CT-Scan	1296 COVID-19, 1735 CAP, 1325 non-CAP	Yes	-	Single modality	Multi-class	Accuracy = 96%
Attallah et al. [62]	CT-scan	7264 positive, 6382 normal	Yes	Augmentation techniques	Single modality	Binary	Accuracy = 98.62% Precision = 99.54% F1-score = 99.62% Recall = 99.69%
Alshazly et al. [63]	CT-scan	2517 COVID-19, 758 normal, 1644 others	No	Augmentation techniques	Single modality	Multi-class	Accuracy = 99.4% Precision = 99.6% Recall = 99.1% F1-score = 99.4%
Attallah et al. [64]	CT-scan	1252 COVID-19, 1230 non-COVID-19	Yes	Augmentation techniques	Single modality	Binary	Accuracy = 99.6% Precision = 99.72% Recall = 99.47% F1-score = 99.6%
Zhang et al. [65]	CT-scan	2215 COVID-19, 245 normal	No	Not mentioned	Single modality	Binary	Not mentioned

Table 1. Cont.

References	Data Used	Sample Size	Dataset Balance	Balance Strategy	Model Type	Classification Method	Performance
Shah et al. [66]	CT-scan	349 positive, 463 negative	No	Augmentation techniques	Single modality	Binary	Accuracy = 94.5%
Wang et al. [67]	CT-scan	Total 1065 images including COVID-19	N/A	Not mentioned	Single modality	Binary	Accuracy = 89.5%
Periera et al. [69]	CXR	90 COVID-19, 1000 normal, 54 others	No	Resampling algorithms	Single modality	Multi-class	F-score = 89%
Horry et al. [70]	CXR/CT-Scan/Ultrasound	CXR: 115 COVID-19, 322 pneumonia, 60,361 no finding, CT-scan: 349 COVID-19, 397 non-COVID-19 Ultrasound: 654 COVID-19, 277 Pneumonia, 172 no finding.	No	-	Multi-modal	Multi-class	Precision = 100% for Ultrasound Precision = 86% for CXR Precision = 84% for CT-Scan
Vinod et al. [71]	CXR/CT-Scan	3000 CT-scan, COVID-19, 3000 CT-scan pneumonia, 3000 CXR COVID-19, 3000 CXR pneumonia, 3000 CXR normal	Yes	-	Multi-modal	Multi-class	Accuracy = 96% for CXR Accuracy = 97% for CT-Scan
Yadav et al. [72]	CXR/CT-Scan	38,155 CXR and CT-scan	No	Not mentioned	Multi-modal	Multi-class	Accuracy = 96% for CXR, 97% for CT-scan.
Kalaiselvi et al. [73]	CXR/CT-Scan	650 CXR, 746 CT-Scan	No	Not mentioned	Multi-modal	Binary	Accuracy = 100% for CXR
El-Banaa et al. [74]	CXR/CT-scan	5719 COVID-19, 2485 normal, 2122 bacterial, 2277 viral	No	Augmentation techniques	Multi-modal	Multi-class	Accuracy = 99.4% Precision = 99.5% Recall = 99.1% F1-score = 99.3%
Ibrahim et al. [75]	CXR/CT-scan	75,000 for COVID-19, normal, pneumonia, lung cancer	No	Augmentation techniques	Multi-modal	Multi-class	Accuracy = 98% Recall = 98% Precision = 99.5% F1-score = 98.24%
Sharma et al. [76]	CXR/CT-scan	Total images = 194,922	No	Augmentation techniques	Multi-modal	Multi-class	Precision = 99% Recall = 91% F1-score = 89%
Panwar et al. [77]	CXR/CT-scan	526 CXR and CT-scan for COVID-19, 1252 CT-scan for COVID-19, 1230 CT-scan for other, 5856 CXR for normal and pneumonia	No	Not mentioned	Multi-modal	Multi-class	Accuracy = 95.61%

Table 1. Cont.

References	Data Used	Sample Size	Dataset Balance	Balance Strategy	Model Type	Classification Method	Performance
Ouahab et al. [78]	CXR/CT-scan	1345 CXR COVID-19, 1345 CXR normal, 1345 CXR pneumonia, 1345 CT COVID-19, 1345 CT normal, 1345 CT pneumonia	Yes	-	Multi-modal	Multi-class	Accuracy = 99% Recall = 99% Precision = 99%
Biswal et al. [79]	CXR/Radiologist report	377,110 CXR	N/A	Not mentioned	Multi-modal	Multi-class	Accuracy = 92% AUC = 90%
Mei et al. [80]	CT-Scan/Symptoms/Lab Tests/Exposure history to COVID-19	415 COVID-19, 486 negative	Yes	-	Multi-modal	Binary	AUC = 92%
Chen et al. [81]	CT-Scan/Clinical information	70 COVID-19, 66 non-COVID-19,	Yes	-	Multi-modal	Binary	AUC = 98.6%
Yang et al. [82]	CT-Scan/Clinical information/Lab Tests	88 COVID-19, 80 other pneumonias	Yes	-	Multi-modal	Multi-class	AUC = 91%
Xu et al. [83]	CT-Scan/Clinical information/Lab Tests	689 cases	Yes	-	Multi-modal	Multi-class	Accuracy = 99%
Khan et al. [84]	CXR/Clinical data	222 COVID-19, 48 normal	No	Augmentation techniques	Multi-modal	Binary	Accuracy = 97% Recall = 98.6% Precision = 97.8% F-Score = 98.2%
Attah Ullah et al. [85]	CXR/Symptoms	200 bacterial, 290 COVID-19, 180 viral, 130 normal	No	Resampling techniques/ Augmentation techniques	Multi-modal	Multi-class	Accuracy = 78.88%
Afshar et al. [86]	CT-scan/Clinical data	176 COVID-19, 76 normal, 60 CAP.	No	Augmentation techniques	Multi-modal	Multi-class	Accuracy = 90.8%
Jiao et al. [87]	CXR/Clinical data	Total data = 1834 patients	N/A	Not mentioned	Multi-modal	Binary	C-index = 0.805
Cheng et al. [88]	CXR/Clinical data	Total data = 5645 cases	N/A	Not mentioned	Multi-modal	Binary	Accuracy = 73.2% Recall = 70.7% Precision = 71.4% F1-score = 74.6%

5. Recommendations to Bridge the Gap

Based on the limitations of prior studies and the gaps identified by this literature review, this section provides recommendations covering different aspects that can help researchers in this area. These recommendations are as follows:

- For problems based on radiological imaging modalities, deep learning has achieved remarkable successes in feature learning and image classification. The architecture of

CNNs has resulted in their status as pioneers in the field of image classification and detection [89]. These systems can detect and learn what radiologists cannot notice using the naked eye, and they significantly outperform traditional techniques, even in previously impossible cases [32].

- Referencing the superior results achieved by hierarchical classification compared to flat classification in previous studies—as shown in the literature—and due to the natural hierarchical structure of diseases developed by the ICD-10 [69], hierarchical classification can improve performance compared to flat classification [27].
- Multimodal deep learning models that combine different types of data in a process of data fusion are more accurate compared to single-modality models, especially in medical research. In contrast, radiologists have mentioned the difficulties associated with relying on portable CXRs alone in facilitating the accurate diagnosis of COVID-19.
- Multi-class classifiers (with a large number of classes) are more efficient and lead to more reliable results compared to binary classifiers.
- Based on the various limitations related to CT imaging modalities—especially in the context of the ongoing pandemic—CXRs are the most recommended method for the radiological examination of the lungs.
- Deep learning models perform better with larger datasets compared to smaller datasets. Accordingly, large amounts of training data from all included classes play a critical role in the success of the model. Therefore, deep learning models should be verified on larger datasets.
- Class imbalance scenarios are mostly found in the domain of health [90], especially in diagnostics and disease detection. Data deficiency resulting from class imbalance has a significant impact on the performance of deep learning models, increasing the difficulty of the learning process and reducing the accuracy [91]. Although there are different ways to solve this problem, it is better to avoid having a large difference in the numbers of images in different classes.
- Most studies in this field have been tested and evaluated using public datasets that are available online, which contain COVID-19 image samples. There is no guarantee that the included cases are COVID-19 cases, and there is also a possibility of duplicating images across these repositories. This makes it difficult to guarantee the performance of the tested models, especially on large datasets. In contrast, private datasets tend to be more reliable and authenticated.
- Model performance should not be evaluated based on accuracy alone, despite the importance of accuracy as a basic evaluation metric. As is well known, each classification scheme has a different number of evaluation metrics (e.g., a multi-class classification model is based on eight criteria, a multi-label classification has four criteria, and a hierarchical classification involves six criteria) [92]. In summary, the assessment measures for COVID-19 classification models require consideration of all related criteria [2,49].

6. Conclusions and Future Perspectives

Despite substantial improvements in AI models and the emergence of extensive research in healthcare applications, there is still a shortage in applying and utilizing AI models in healthcare. Many studies have explored the problem of the classification and detection of COVID-19, especially in light of the influence of the ongoing pandemic. This review shows that most prior studies of COVID-19 image classification are misleading because they used multi-class classifiers (i.e., with only three or four classes) with larger samples for the COVID-19 class compared to the samples for other classes; this gives the algorithm an incorrectly high percentage of sensitivity. In addition, most studies addressing the detection and classification of pneumonia in medical images have focused on differentiating between two or three classes, e.g., viral, bacterial, and COVID-19 infections. Despite this, the ability of AI systems to differentiate between various classes is increasing as they learn from a greater number of classes. Moreover, all approaches that have been used in the literature are based on classifying the dataset using a flat structure; only one study has

addressed the efficiency of hierarchical classification compared to flat classification. In the literature, CNN-based classifiers have also been applied on the hierarchical ETHEC dataset. The results demonstrated that the hierarchical structure incorporated the loss function and enhanced generalization across classes. This can be attributed to the exploitation of shared features between classes at different levels, which assists in overcoming the data scarcity problem [93]. In addition, in multiple domains, classifiers in hierarchical models have been shown to reduce classification errors and break down the problem; this leads to a better performance compared to flat models [94].

Surprisingly, several techniques have been used for the detection and classification of COVID-19, but limited research has addressed multimodal deep learning models for heterogeneous data types. It is also notable that most existing models focus on a single feature modality (i.e., medical images), while multimodal features (i.e., those combining more than one aspect of COVID-19 health information, such as medical images, diagnostic data, medication data, and laboratory data) contribute to a superior performance in disease diagnostic processes [35,95]. Furthermore, prior studies have primarily considered the CT scan as the main radiological imaging modality for all infected cases during the ongoing pandemic. Finally, regarding the approaches mentioned earlier in this review, and taking into account its limitations, there is clearly still room to enhance and improve research in this field. By considering the gaps and exploiting the remarkable successes of each prior approach, it is expected that these can be combined into one model that may see widespread clinical adoption in the future.

Author Contributions: Conceptualization, S.A.A., A.A.M. and A.S.A.; methodology, S.A.A., S.A. and A.A.; software, A.A.; validation, S.A.A., S.A., T.N. and A.M.; formal analysis, S.A.A., S.A. and A.A.; investigation, A.S.A.; resources, S.A.A., S.A. and A.A.; data curation, A.S.A.; writing—original draft preparation, A.S.A.; writing—review and editing, S.A.A., S.A., T.N. and A.M.; visualization, S.A.A., S.A., T.N., A.A.M. and A.S.A.; supervision, S.A.A. and A.A.M.; project administration, S.A.A., A.A.M. and A.S.A.; funding acquisition, S.A.A. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Data Availability Statement: Not applicable.

Acknowledgments: The authors would like to thank the editor and reviewers for spending their valuable time reviewing and polishing this article.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. World Health Organization. WHO Coronavirus (COVID-19) Dashboard. 2022. Available online: <https://covid19.who.int/> (accessed on 17 August 2022).
2. Chowdhury, M.E.H.; Rahman, T.; Khandakar, A.; Mazhar, R.; Kadir, M.A.; Bin Mahbub, Z.; Islam, K.R.; Khan, M.S.; Iqbal, A.; Al Emadi, N.; et al. Can AI Help in Screening Viral and COVID-19 Pneumonia? *IEEE Access* **2020**, *8*, 132665–132676. [CrossRef]
3. Maharjan, N.; Thapa, N.; Magar, B.P.; Maharjan, M.; Tu, J. COVID-19 Diagnosed by Real-Time Reverse Transcriptase-Polymerase Chain Reaction in Nasopharyngeal Specimens of Suspected Cases in a Tertiary Care Center: A Descriptive Cross-sectional Study. *J. Nepal. Med. Assoc.* **2021**, *59*, 464. [CrossRef] [PubMed]
4. MedicineNet. Definition of RT-PCR. 2021. Available online: <https://www.medicinenet.com/rtpcr/definition.htm> (accessed on 26 July 2022).
5. Swapnarekha, H.; Behera, H.S.; Nayak, J.; Naik, B. Role of intelligent computing in COVID-19 prognosis: A state-of-the-art review. *Chaos Solitons Fractals* **2020**, *138*, 109947. [CrossRef] [PubMed]
6. Khorasani, A.; Chegini, A.; Mirzaei, A. New Insight into Laboratory Tests and Imaging Modalities for Fast and Accurate Diagnosis of COVID-19: Alternative Suggestions for Routine RT-PCR and CT—A Literature Review. *Can. Respir. J.* **2020**, *2020*, 4648307. [CrossRef] [PubMed]
7. Yang, T.; Wang, Y.-C.; Shen, C.-F.; Cheng, C.-M. Point-of-Care RNA-Based Diagnostic Device for COVID-19. *Diagnostics* **2020**, *10*, 165. [CrossRef] [PubMed]
8. HealthITAnalytics. Top 5 Use Cases for Artificial Intelligence in Medical Imaging. 2021. Available online: <https://healthitanalytics.com/news/top-5use-cases-for-artificial-intelligence-in-medicalimaging> (accessed on 15 July 2022).

9. Roberts, M.; Covnet, A.; Driggs, D.; Thorpe, M.; Gilbey, J.; Yeung, M.; Ursprung, S.; Aviles-Rivero, A.I.; Etmann, C.; McCague, C.; et al. Common pitfalls and recommendations for using machine learning to detect and prognosticate for COVID-19 using chest radiographs and CT scans. *Nat. Mach. Intell.* **2021**, *3*, 199–217. [CrossRef]
10. Ghaderzadeh, M.; Asadi, F. Deep Learning in the Detection and Diagnosis of COVID-19 Using Radiology Modalities: A Systematic Review. *J. Health Eng.* **2021**, *2021*, 6677314.
11. Aljondi, R.; Alghamdi, S. Diagnostic Value of Imaging Modalities for COVID-19: Scoping Review. *J. Med. Internet Res.* **2020**, *22*, e19673. [CrossRef]
12. Google Scholar. 2022. Available online: <https://scholar.google.com/> (accessed on 23 July 2022).
13. IEEE Xplore. Ieeexplore.ieee.org. 2022. Available online: <https://ieeexplore.ieee.org/Xplore/guesthome.jsp> (accessed on 15 July 2022).
14. PubMed. 2022. Available online: <https://pubmed.ncbi.nlm.nih.gov/> (accessed on 26 September 2022).
15. Pal, M.; Berhanu, G.; Desalegn, C.; Kandi, V. Severe Acute Respiratory Syndrome Coronavirus-2 (SARS-CoV-2): An Update. *Cureus* **2020**, *12*, e7423. [CrossRef]
16. World Health Organization (WHO). Pneumonia. 2021. Available online: <https://www.who.int/news-room/fact-sheets/detail/pneumonia> (accessed on 27 July 2022).
17. ICD-10 Version:2019, Icd.who.int. 2021. Available online: <https://icd.who.int/browse10/2019/en/#/> (accessed on 22 September 2022).
18. Centers for Disease Control and Prevention. Middle East Respiratory Syndrome (MERS). 2022. Available online: <https://www.cdc.gov/coronavirus/mers/index.html> (accessed on 15 August 2022).
19. Helmy, Y.A.; Fawzy, M.; Elasad, A.; Sobieh, A.; Kenney, S.P.; Shehata, A.A. The COVID-19 Pandemic: A Comprehensive Review of Taxonomy, Genetics, Epidemiology, Diagnosis, Treatment, and Control. *J. Clin. Med.* **2020**, *9*, 1225. [CrossRef]
20. Desforgues, M.; le Coupanec, A.; Dubeau, P.; Bourgouin, A.; Lajoie, L.; Dubé, M.; Talbot, P.J. Human Coronaviruses and Other Respiratory Viruses: Underestimated Opportunistic Pathogens of the Central Nervous System? *Viruses* **2019**, *12*, 14. [CrossRef] [PubMed]
21. Mohammadi, A.; Wang, Y.; Enshaei, N.; Afshar, P.; Naderkhani, F.; Oikonomou, A.; Rafiee, J.; de Oliveira, H.R.; Yanushkevich, S.; Plataniotis, K.N. Diagnosis/Prognosis of COVID-19 Chest Images via Machine Learning and Hypersignal Processing: Challenges, opportunities, and applications. *IEEE Signal Process. Mag.* **2021**, *38*, 37–66. [CrossRef]
22. Candemir, S.; Antani, S. A review on lung boundary detection in chest X-rays. *Int. J. Comput. Assist. Radiol. Surg.* **2019**, *14*, 563–576. [CrossRef]
23. Jacobi, A.; Chung, M.; Bernheim, A.; Eber, C. Portable chest X-ray in coronavirus disease-19 (COVID-19): A pictorial review. *Clin. Imaging* **2020**, *64*, 35–42. [CrossRef]
24. Ponnusamy, R.; Sathyamoorthy, S.; Manikandan, K. A Review of Image Classification Approaches and Techniques. *Int. J. Recent Trends Eng. Res.* **2017**, *3*, 2455–1457. [CrossRef]
25. Silla, C.; Freitas, A. A survey of hierarchical classification across different application domains. *Data Min. Knowl. Discov.* **2010**, *22*, 31–72. [CrossRef]
26. Guo, Y.; Liu, Y.; Bakker, E.M.; Guo, Y.; Lew, M.S. CNN-RNN: A large-scale hierarchical image classification framework. *Multimed. Tools Appl.* **2017**, *77*, 10251–10271. [CrossRef]
27. Zimek, A.; Buchwald, F.; Frank, E.; Kramer, S. A Study of Hierarchical and Flat Classification of Proteins. *IEEE/ACM Trans. Comput. Biol. Bioinform.* **2010**, *7*, 563–571. [CrossRef]
28. Huang, K.; Hussain, A.; Wang, Q.; Zhang, R. *Deep Learning: Fundamentals, Theory and Applications*; Cognitive Computation Trends; Springer: Berlin/Heidelberg, Germany, 2019.
29. Lee, J.-G.; Jun, S.; Cho, Y.-W.; Lee, H.; Kim, G.B.; Seo, J.B.; Kim, N. Deep Learning in Medical Imaging: General Overview. *Korean J. Radiol.* **2017**, *18*, 570–584. [CrossRef]
30. Cozzi, A.; Schiaffino, S.; Arpaia, F.; Della Pepa, G.; Tritella, S.; Bertolotti, P.; Menicagli, L.; Monaco, C.G.; Carbonaro, L.A.; Spairani, R.; et al. Chest x-ray in the COVID-19 pandemic: Radiologists’ real-world reader performance. *Eur. J. Radiol.* **2020**, *132*, 109272. [CrossRef]
31. Zhou, S.; Rueckert, D.; Fichtinger, G. *Handbook of Medical Image Computing and Computer Assisted Intervention*; Academic Press: Cambridge, MA, USA, 2019.
32. Kiryu, S.; Yasaka, K.; Akai, H.; Nakata, Y.; Sugomori, Y.; Hara, S.; Seo, M.; Abe, O.; Ohtomo, K. Deep learning to differentiate parkinsonian disorders separately using single midsagittal MR imaging: A proof of concept study. *Eur. Radiol.* **2019**, *29*, 6891–6899. [CrossRef] [PubMed]
33. Razzak, M.; Naz, S.; Zaib, A. Deep Learning for Medical Image Processing: Overview, Challenges and the Future. In *Lecture Notes in Computational Vision and Biomechanics*; Springer: Berlin/Heidelberg, Germany, 2017; pp. 323–350.
34. Lahat, D.; Adali, T.; Jutten, C. Multimodal Data Fusion: An Overview of Methods, Challenges, and Prospects. *Proc. IEEE* **2015**, *103*, 1449–1477. [CrossRef]
35. Xu, M.; Ouyang, L.; Gao, Y.; Chen, Y.; Yu, T.; Li, Q.; Sun, K.; Bao, F.S.; Safarnejad, L.; Wen, J.; et al. Accurately Differentiating COVID-19, Other Viral Infection, and Healthy Individuals: A Multimodal Late Fusion Learning Approach. *J. Med. Internet Res. Med. Inform.* **2020**, *23*, e25535. [CrossRef] [PubMed]

36. Gao, J.; Li, P.; Chen, Z.; Zhang, J. A Survey on Deep Learning for Multimodal Data Fusion. *Neural Comput.* **2020**, *32*, 829–864. [[CrossRef](#)] [[PubMed](#)]
37. Yang, L.; Tao, L.; Chen, X.; Gu, X. Multi-scale semantic feature fusion and data augmentation for acoustic scene classification. *Appl. Acoust.* **2020**, *163*, 107238. [[CrossRef](#)]
38. Bhowal, P.; Sen, S.; Yoon, J.H.; Geem, Z.W.; Sarkar, R. Choquet Integral and Coalition Game-Based Ensemble of Deep Learning Models for COVID-19 Screening from Chest X-ray Images. *IEEE J. Biomed. Health Inform.* **2021**, *25*, 4328–4339. [[CrossRef](#)]
39. Chouhan, V.; Singh, S.K.; Khamparia, A.; Gupta, D.; Tiwari, P.; Moreira, C.; Damaševičius, R.; de Albuquerque, V.H.C. A Novel Transfer Learning Based Approach for Pneumonia Detection in Chest X-ray Images. *Appl. Sci.* **2020**, *10*, 559. [[CrossRef](#)]
40. Khan, S.; Sohail, A.; Khan, A.; Hassan, M.; Lee, Y.; Alam, J.; Basit, A.; Zubair, S. COVID-19 detection in chest X-ray images using deep boosted hybrid learning. *Comput. Biol. Med.* **2021**, *137*, 104816. [[CrossRef](#)]
41. Al-Waisy, A.S.; Mohammed, M.A.; Al-Fahdawi, S.; Maashi, M.S.; Garcia-Zapirain, B.; Abdulkareem, K.H.; Mostafa, S.A.; Kumar, N.M.; Le, D.-N. COVID-DeepNet: Hybrid Multimodal Deep Learning System for Improving COVID-19 Pneumonia Detection in Chest X-ray Images. *Comput. Mater. Contin.* **2021**, *67*, 2409–2429. [[CrossRef](#)]
42. Qaid, T.S.; Mazaar, H.; Al-Shamri, M.Y.H.; Alqahtani, M.S.; Raweh, A.A.; Alakwaa, W. Hybrid Deep-Learning and Machine-Learning Models for Predicting COVID-19. *Comput. Intell. Neurosci.* **2021**, *2021*, 9996737. [[CrossRef](#)]
43. Attallah, O. An Intelligent ECG-Based Tool for Diagnosing COVID-19 via Ensemble Deep Learning Techniques. *Biosensors* **2022**, *12*, 299. [[CrossRef](#)] [[PubMed](#)]
44. Rajaraman, S.; Sornapudi, S.; Alderson, P.O.; Folio, L.R.; Antani, S.K. Analyzing inter-reader variability affecting deep ensemble learning for COVID-19 detection in chest radiographs. *PLoS ONE* **2020**, *15*, e0242301. [[CrossRef](#)] [[PubMed](#)]
45. Bharadwaj, L.; Boddada, R.R.; Gajapaka, M.; Vardhan, S. COVID-19 Classification Using Staked Ensembles: A Comprehensive Analysis. *arXiv* **2021**, arXiv:2010.05690.
46. Al-Monsur, A.; Kabir, R.; Ar-Rafi, A.M.; Nishat, M.M.; Faisal, F. Covid-EnsembleNet: An Ensemble Based Approach for Detecting COVID-19 by utilising Chest X-ray Images. In Proceedings of the 2022 IEEE World AI IoT Congress (AIIoT), Seattle, WA, USA, 6–9 June 2022; pp. 351–356. [[CrossRef](#)]
47. Zhou, T.; Lu, H.; Yang, Z.; Qiu, S.; Huo, B.; Dong, Y. The ensemble deep learning model for novel COVID-19 on CT images. *Appl. Soft Comput.* **2021**, *98*, 106885. [[CrossRef](#)] [[PubMed](#)]
48. Sarma, A.; Heilbrun, M.; Conner, K.; Stevens, S.; Woller, S.; Elliott, C. Radiation and Chest CT Scan Examinations. *Chest* **2012**, *142*, 750–760. [[CrossRef](#)]
49. Albahri, O.; Zaidan, A.; Zaidan, B.; Abdulkareem, K.H.; Al-Qaysi, Z.; Alamoodi, A.; Aleesa, A.; Chyad, M.; Alesa, R.; Kem, L.; et al. Systematic review of artificial intelligence techniques in the detection and classification of COVID-19 medical images in terms of evaluation and benchmarking: Taxonomy analysis, challenges, future solutions and methodological aspects. *J. Infect. Public Health* **2020**, *13*, 1381–1396. [[CrossRef](#)]
50. Abbas, A.; Abdelsamea, M.; Gaber, M. Classification of COVID-19 in chest X-ray images using DeTraC deep convolutional neural network. *Appl. Intell.* **2020**, *51*, 854864. [[CrossRef](#)]
51. Brunese, L.; Mercaldo, F.; Reginelli, A.; Santone, A. Explainable Deep Learning for Pulmonary Disease and Coronavirus COVID-19 Detection from X-rays. *Comput. Methods Programs Biomed.* **2020**, *196*, 105608. [[CrossRef](#)]
52. Loey, M.; Smarandache, F.; Khalifa, N.E.M. Within the Lack of Chest COVID-19 X-ray Dataset: A Novel Detection Model Based on GAN and Deep Transfer Learning. *Symmetry* **2020**, *12*, 651. [[CrossRef](#)]
53. Oh, Y.; Park, S.; Ye, J.C. Deep Learning COVID-19 Features on CXR Using Limited Training Data Sets. *IEEE Trans. Med. Imaging* **2020**, *39*, 2688–2700. [[CrossRef](#)]
54. Sitaula, C.; Hossain, M.B. Attention-based VGG-16 model for COVID-19 chest X-ray image classification. *Appl. Intell.* **2020**, *51*, 2850–2863. [[CrossRef](#)] [[PubMed](#)]
55. Wang, G.; Liu, X.; Shen, J.; Wang, C.; Li, Z.; Ye, L.; Wu, X.; Chen, T.; Wang, K.; Zhang, X.; et al. A deep-learning pipeline for the diagnosis and discrimination of viral, non-viral and COVID-19 pneumonia from chest X-ray images. *Nat. Biomed. Eng.* **2021**, *5*, 509–521. [[CrossRef](#)]
56. Shankar, K.; Perumal, E. A novel hand-crafted with deep learning features based fusion model for COVID-19 diagnosis and classification using chest X-ray images. *Complex Intell. Syst.* **2020**, *7*, 1277–1293. [[CrossRef](#)] [[PubMed](#)]
57. Ouchicha, C.; Ammor, O.; Meknassi, M. CVDNet: A novel deep learning architecture for detection of coronavirus (COVID-19) from chest x-ray images. *Chaos Solitons Fractals* **2020**, *140*, 110245. [[CrossRef](#)]
58. Azemin, M.Z.C.; Hassan, R.; Tamrin, M.I.M.; Ali, M.A. COVID-19 Deep Learning Prediction Model Using Publicly Available Radiologist-Adjudicated Chest X-Ray Images as Training Data: Preliminary Findings. *Int. J. Biomed. Imaging* **2020**, *2020*, 8828855. [[CrossRef](#)]
59. Fusco, R.; Grassi, R.; Granata, V.; Setola, S.V.; Grassi, F.; Cozzi, D.; Pecori, B.; Izzo, F.; Petrillo, A. Artificial Intelligence and COVID-19 Using Chest CT Scan and Chest X-ray Images: Machine Learning and Deep Learning Approaches for Diagnosis and Treatment. *J. Pers. Med.* **2021**, *11*, 993. [[CrossRef](#)] [[PubMed](#)]
60. Amyar, A.; Modzelewski, R.; Li, H.; Ruan, S. Multi-task deep learning based CT imaging analysis for COVID-19 pneumonia: Classification and segmentation. *Comput. Biol. Med.* **2020**, *126*, 104037. [[CrossRef](#)] [[PubMed](#)]

61. Li, L.; Qin, L.; Xu, Z.; Yin, Y.; Wang, X.; Kong, B.; Bai, J.; Lu, Y.; Fang, Z.; Song, Q.; et al. Using artificial intelligence to detect COVID-19 and community-acquired pneumonia based on pulmonary CT: Evaluation of the diagnostic accuracy. *Radiology* **2020**, *296*, E65–E71. [[CrossRef](#)] [[PubMed](#)]
62. Attallah, O.; Samir, A. A wavelet-based deep learning pipeline for efficient COVID-19 diagnosis via CT slices. *Appl. Soft Comput.* **2022**, *128*, 109401. [[CrossRef](#)] [[PubMed](#)]
63. Alshazly, H.; Linse, C.; Abdalla, M.; Barth, E.; Martinetz, T. COVID-Nets: Deep CNN architectures for detecting COVID-19 using chest CT scans. *PeerJ Comput. Sci.* **2021**, *7*, e655. [[CrossRef](#)]
64. Attallah, O. A computer-aided diagnostic framework for coronavirus diagnosis using texture-based radiomics images. *Digit. Health* **2022**, *8*, 205520762210925. [[CrossRef](#)] [[PubMed](#)]
65. Zhang, H.; Zhang, J.; Zhang, H.; Nan, Y.; Zhao, Y.; Fu, E.; Xie, Y.; Liu, W.; Li, W.; Zhang, H.; et al. Automated detection and quantification of COVID-19 pneumonia: CT imaging analysis by a deep learning-based software. *Eur. J. Nucl. Med. Mol. Imaging* **2020**, *47*, 2525–2532. [[CrossRef](#)]
66. Shah, V.; Keniya, R.; Shridharani, A.; Punjabi, M.; Shah, J.; Mehendale, N. Diagnosis of COVID-19 using CT scan images and deep learning techniques. *Emerg. Radiol.* **2021**, *28*, 497–505. [[CrossRef](#)] [[PubMed](#)]
67. Wang, S.; Kang, B.; Ma, J.; Zeng, X.; Xiao, M.; Guo, J.; Cai, M.; Yang, J.; Li, Y.; Meng, X.; et al. A deep learning algorithm using CT images to screen for Corona Virus Disease (COVID-19). *Eur. Radiol.* **2020**, *31*, 6096–6104. [[CrossRef](#)]
68. Yan, Z.; Zhang, H.; Piramuthu, R.; Jagadeesh, V.; DeCoste, D.; Di, W.; Yu, Y. HD-CNN: Hierarchical Deep Convolutional Neural Networks for Large Scale Visual Recognition. In Proceedings of the IEEE International Conference on Computer Vision (ICCV), Santiago, Chile, 7–13 December 2015; pp. 2740–2748.
69. Pereira, R.M.; Bertolini, D.; Teixeira, L.O.; Silla, C.N.; Costa, Y.M. COVID-19 identification in chest X-ray images on flat and hierarchical classification scenarios. *Comput. Methods Programs Biomed.* **2020**, *194*, 105532. [[CrossRef](#)] [[PubMed](#)]
70. Horry, M.J.; Chakraborty, S.; Paul, M.; Ulhaq, A.; Pradhan, B.; Saha, M.; Shukla, N. COVID-19 Detection through Transfer Learning Using Multimodal Imaging Data. *IEEE Access* **2020**, *8*, 149808–149824. [[CrossRef](#)]
71. Vinod, D.N.; Jeyavadhanam, B.R.; Zungeru, A.M.; Prabakaran, S. Fully automated unified prognosis of COVID-19 chest X-ray/CT scan images using Deep COVIX-Net model. *Comput. Biol. Med.* **2021**, *136*, 104729. [[CrossRef](#)]
72. Yadav, P.; Menon, N.; Ravi, V.; Vishvanathan, S. Lung-GANs: Unsupervised Representation Learning for Lung Disease Classification Using Chest CT and X-ray Images. *IEEE Trans. Eng. Manag.* **2021**, 1–13. [[CrossRef](#)]
73. Padmapriya, S.T.; Kalaiselvi, T.; Somasundaram, K.; Kumar, C.N.; Priyadharshini, V. Novel Artificial Intelligence Learning Models for COVID-19 Detection from X-ray and CT Chest Images. *Int. J. Comput. Intell. Control.* **2021**, *13*, 9–17.
74. El-Bana, S.; Al-Kabbany, A.; Sharkas, M. A multi-task pipeline with specialized streams for classification and segmentation of infection manifestations in COVID-19 scans. *PeerJ Comput. Sci.* **2020**, *6*, e303. [[CrossRef](#)]
75. Ibrahim, D.M.; Elshennawy, N.M.; Sarhan, A.M. Deep-chest: Multi-classification deep learning model for diagnosing COVID-19, pneumonia, and lung cancer chest diseases. *Comput. Biol. Med.* **2021**, *132*, 104348. [[CrossRef](#)] [[PubMed](#)]
76. Sharma, Y.; Furqan, A. Combination of computed tomography images with a chest X-ray diagnostic system using deep learning. *Int. J. Adv. Multidiscip. Sci. Res.* **2022**, *5*, 24–54. [[CrossRef](#)]
77. Panwar, H.; Gupta, P.; Siddiqui, M.K.; Morales-Menendez, R.; Bhardwaj, P.; Singh, V. A deep learning and grad-CAM based color visualization approach for fast detection of COVID-19 cases using chest X-ray and CT-Scan images. *Chaos Solitons Fractals* **2020**, *140*, 110190. [[CrossRef](#)]
78. Ouahab, A. Multimodal Convolutional Neural Networks for Detection of COVID-19 Using Chest X-Ray and CT Images. *Opt. Mem. Neural Netw.* **2021**, *30*, 276–283. [[CrossRef](#)]
79. Biswal, S.; Zhuang, P.; Pyrros, A.; Siddiqui, N.; Koyejo, S.; Sun, J. EMIXER: End-to-end Multimodal X-ray Generation via Self-supervision. *arXiv* **2020**, arXiv:2007.05597.
80. Mei, X.; Lee, H.-C.; Diao, K.-Y.; Huang, M.; Lin, B.; Liu, C.; Xie, Z.; Ma, Y.; Robson, P.M.; Chung, M.; et al. Artificial intelligence-enabled rapid diagnosis of patients with COVID-19. *Nat. Med.* **2020**, *26*, 1224–1228. [[CrossRef](#)] [[PubMed](#)]
81. Chen, X.; Tang, Y.; Mo, Y.; Li, S.; Lin, D.; Yang, Z.; Yang, Z.; Sun, H.; Qiu, J.; Liao, Y.; et al. A diagnostic model for coronavirus disease 2019 (COVID-19) based on radiological semantic and clinical features: A multi-center study. *Eur. Radiol.* **2020**, *30*, 4893–4902. [[CrossRef](#)]
82. Qin, L.; Yang, Y.; Cao, Q.; Cheng, Z.; Wang, X.; Sun, Q.; Yan, F.; Qu, J.; Yang, W. A predictive model and scoring system combining clinical and CT characteristics for the diagnosis of COVID-19. *Eur. Radiol.* **2020**, *30*, 6797–6807. [[CrossRef](#)]
83. Xu, M.; Ouyang, L.; Han, L.; Sun, K.; Yu, T.; Li, Q.; Tian, H.; Safarnejad, L.; Zhang, H.; Gao, Y.; et al. Accurately Differentiating Between Patients with COVID-19, Patients with Other Viral Infections, and Healthy Individuals: Multimodal Late Fusion Learning Approach. *J. Med. Internet Res.* **2021**, *23*, e25535. [[CrossRef](#)]
84. Khan, I.U.; Aslam, N.; Anwar, T.; Alsaif, H.S.; Chrouf, S.M.B.; Alzahrani, N.A.; Alamoudi, F.A.; Kamaleldin, M.M.A.; Awary, K.B. Using a Deep Learning Model to Explore the Impact of Clinical Data on COVID-19 Diagnosis Using Chest X-ray. *Sensors* **2022**, *22*, 669. [[CrossRef](#)]
85. Attaullah, M.; Ali, M.; Almuftareh, M.F.; Ahmad, M.; Hussain, L.; Jhanjhi, N.; Humayun, M. Initial Stage COVID-19 Detection System Based on Patients' Symptoms and Chest X-Ray Images. *Appl. Artif. Intell.* **2022**, *36*, 2055398. [[CrossRef](#)]

86. Afshar, P.; Heidarian, S.; Naderkhani, F.; Rafiee, M.J.; Oikonomou, A.; Plataniotis, K.N.; Mohammadi, A. Hybrid Deep Learning Model For Diagnosis of COVID-19 Using Ct Scans And Clinical/Demographic Data. In Proceedings of the 2021 IEEE International Conference on Image Processing (ICIP), Anchorage, AK, USA, 19–22 September 2021; pp. 180–184. [[CrossRef](#)]
87. Jiao, Z.; Choi, J.W.; Halsey, K.; Tran, T.M.L.; Hsieh, B.; Wang, D.; Eweje, F.; Wang, R.; Chang, K.; Wu, J.; et al. Prognostication of patients with COVID-19 using artificial intelligence based on chest x-rays and clinical data: A retrospective study. *Lancet Digit. Health* **2021**, *3*, e286–e294. [[CrossRef](#)]
88. Cheng, J.; Sollee, J.; Hsieh, C.; Yue, H.; Vandal, N.; Shanahan, J.; Choi, J.W.; Tran, T.M.L.; Halsey, K.; Iheanacho, F.; et al. COVID-19 mortality prediction in the intensive care unit with deep learning based on longitudinal chest X-rays and clinical data. *Eur. Radiol.* **2022**, *32*, 4446–4456. [[CrossRef](#)] [[PubMed](#)]
89. Rawat, W.; Wang, Z. Deep Convolutional Neural Networks for Image Classification: A Comprehensive Review. *Neural Comput.* **2017**, *29*, 2352–2449. [[CrossRef](#)]
90. Li, D.-C.; Liu, C.-W.; Hu, S.C. A learning method for the class imbalance problem with medical data sets. *Comput. Biol. Med.* **2010**, *40*, 509–518. [[CrossRef](#)]
91. Johnson, J.M.; Khoshgoftaar, T.M. Survey on deep learning with class imbalance. *J. Big Data* **2019**, *6*, 27. [[CrossRef](#)]
92. Sokolova, M.; Lapalme, G. A systematic analysis of performance measures for classification tasks. *Inf. Process. Manag.* **2009**, *45*, 427–437. [[CrossRef](#)]
93. Dhall, A.; Makarova, A.; Ganea, O.; Pavllo, D.; Greeff, M.; Krause, A. Hierarchical Image Classification using Entailment Cone Embeddings. 2020, 3649–3658. In Proceedings of the 2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops (CVPRW), Seattle, WA, USA, 14–19 June 2020. [[CrossRef](#)]
94. Kowsari, K.; Sali, R.; Ehsan, L.; Adorno, W.; Ali, A.; Moore, S.; Amadi, B.; Kelly, P.; Syed, S.; Brown, D. HMIC: Hierarchical Medical Image Classification, A Deep Learning Approach. *Information* **2020**, *11*, 318. [[CrossRef](#)]
95. Vaishya, R.; Javaid, M.; Khan, I.; Haleem, A. Artificial Intelligence (AI) applications for COVID-19 pandemic. *Diabetes Metabolic Syndrome: Clin. Res. Rev.* **2020**, *14*, 337–339. [[CrossRef](#)]