Arash Mohammadi,
Yingxu Wang,
Nastaran Enshaei,
Parnian Afshar,
Farnoosh Naderkhani,
Anastasia Oikonomou,
Moezedin Javad Rafiee,
Helder C.R. Oliveira,
Svetlana Yanushkevich, and
Konstantinos N. Plataniotis

Challenges, opportunities, and applications

©SHUTTERSTOCK.COM/STEFANO GARAU

Diagnosis/Prognosis of COVID-19 Chest Images via Machine Learning and Hypersignal Processing

The novel coronavirus disease, COVID-19, has rapidly and abruptly changed the world as we knew it in 2020. It has become the most unprecedented challenge to analytic epidemiology (AE) in general and signal processing (SP) theories specifically. In this regard, medical imaging plays an important role for the management of COVID-19. SP and deep learning (DL) models can assist in the development of robust radiomics solutions for the diagnosis/prognosis, severity assessment, treatment response, and monitoring of COVID-19 patients.

We intend to present not only an overview of the current state, challenges, and opportunities of developing SP/DL-empowered models for the diagnosis/prognosis of COVID-19 but also the latest developments in the theoretical framework of AE and hypersignal processing (HP) for COVID-19 from the points of view of both SP and medical/pandemic control professionals. The imaging modalities and radiological characteristics of COVID-19 are then discussed. SL/DL-based radiomic models specific to the analysis of COVID-19 infection are described covering four domains, which encompass the segmentation of COVID-19 lesions, models for outcome prediction, severity assessment, and diagnosis/classification models. This work leads to the identification of significant open problems and opportunities

Digital Object Identifier 10.1109/MSP.2021.3090674
Date of current version: 27 August 2021

that challenge the state of the art of COVID-19 pandemic methodologies and solutions.

Introduction

We are facing an abruptly changed world because of the novel coronavirus outbreak identified as a pandemic by the World Health Organization (WHO). Given its high-contingency nature; relatively unknown behavior; systemic complications; and adverse effects, ranging from human fatalities to economic recessions across the world, it is of importance to develop efficient processing/learning models to help overcome this pandemic and be prepared for potential future ones.

The reverse-transcription polymerase chain reaction (RT-PCR) is the standard testing approach for the early diagnosis of suspected cases of COVID-19. The unavailability of enough RT-PCR testing kits, particularly in areas severely affected by the pandemic, and the test's relatively high and variable false-negative rate [i.e., highest during the first five days (up to 67%) and lowest on day 8 (21%) [1]] resulted in focusing on medical image radiomics [2] as a complementary source for diagnosis/prognosis.

Recent studies [3]–[5] have shown that chest computed tomography (CT) scans and chest radiography (CXR) reveal informative features of COVID-19 that can assist in the monitoring, severity assessment, and treatment of COVID-19 [6], [7]. According to the guidelines provided by the WHO, the use of chest imaging as a complementary source of data is recommended in different scenarios and stages of COVID-19 to help radiologists and physicians detect and evaluate the disease more accurately.

CT and CXR can decrease the false-negative rate at both the admission and discharge times. It is worth mentioning that chest CT has a key role for the diagnosis of COVID-19 in the very early stages of the infection and also to set up a prognosis. Comparisons between CT and RT-PCR at the early stages of COVID-19 infection show that CT abnormalities may appear before PCR positivity. In other words, CT has a greater sensitivity during the early stages of the infection. In addition, false negatives in RT-PCR results occur at both admission and discharge.

Finally, CT plays its role over the course of the disease for evaluating changes in severity and treatment adjustments. The key power of chest imaging is in its prognostic value to identify the severity of the disease as well as the likelihood of needing hospitalization and/or admission to the intensive care unit (ICU), especially in countries with limited human and economic resources. However, the interpretation of chest images for confirming suspected cases of COVID-19 and a severity assessment of the disease based on imaging findings are time-consuming and may be challenging.

The interpretation of CT and CXR images should be performed by expert thoracic radiologists, who may not be easily accessible, especially during an outbreak when the number of suspected cases of COVID-19 is growing exponentially. To address these issues, there has been a surge of interest in developing SP and DL techniques to extract informative features from chest images and help in the fast detection and risk assessment of the COVID-19 infection.

We would like to mention that, although research works on COVID-19-related topics have started very recently, the extensive number of research works that have been disseminated during this short period of time makes the topic mature. More specifically, extensive research works on the applications of signal/image processing and artificial intelligence (AI) for COVID-19 led to almost 1,200 publications by the end of November 2020, obtained from PubMed with the keyword "COVID-19" and either of the following keywords: "signal processing", "machine learning", "AI", or "deep learning". These publications cover several aspects/applications of SP/DL for COVID-19, including its diagnosis, classification, detection, segmentation, severity assessment, and survival analysis.

Review methodology

We defined the goal of our article as "identifying the gaps and developing a conceptual framework of the topic." Consequently, we followed the two phases of planning and conducting before reporting and dissemination [8], explained as follows:

- *Planning*: The diagnosis/prognosis of COVID-19 images was identified as the research question in the planning phase. The search strategy was based on the selected databases—Google Scholar and PubMed—with the selected keywords: "COVID-19" and ("Signal Processing" or "Machine Learning" or "AI" or "Deep Learning"). For any articles found, the reference list was also checked. To broaden the focus, not only the published articles but also preprints in Arxiv and Medrxiv were considered, along with both implementation and discussion articles, so as not to exclude useful insights. Since COVID-19 is an emerging field, we did not restrict the publication date of the included articles.
- *Conducting the review*: Following the defined review protocol, articles were retrieved. In this phase, articles whose methods were already included in other ones and that did not have added value for our research and/or had few citations were further excluded. Furthermore, we excluded articles without sufficient details and evidence. Data were, accordingly, extracted from the remaining articles, categorized, and summarized in a narrative manner.

In summary, in this feature article, we aim to present an overview of the current state of, challenges in, and opportunities for developing SP/DL-empowered models for the diagnosis and prognosis of the COVID-19 infection based on medical images. The article mainly focuses on problems and applications as well as how SP/DL models can be used to address those identified. In brief, we cover the following main topics:

- We focus on SP techniques specific to COVID-19 images and target specialized SP aspects of COVID-19 diagnoses, including AE and HP theory as advanced processing solutions of COVID-19.
- An investigation of the required medical background related to COVID-19 for the development of advanced

SP/DL models is included. An overview of different characteristics of COVID-19 that can be observed on chest images is presented.
- The potential applications of SP/DL-based models for the diagnosis and predictive prognosis of COVID-19 infections using medical images as the main source of data are introduced.
- DL radiomic directions specific to the analysis of COVID-19 infection are presented. We focus on DL-based solutions from the following four different aspects: the segmentation of COVID-19 lesions, models for outcome prediction in COVID-19 patients, DL/SP models for severity assessment, and the diagnosis and classification of COVID-19 cases.
- We introduce the challenges of, open problems with, and opportunities for developing intelligent and autonomous models for the diagnosis/prognosis of COVID-19.

Distinction from existing articles

As a final note, we would like to briefly elaborate on the differences between this article and a recent feature article [2] on radiomics and other surveys/tutorials [3]–[5] on COVID-19. In particular, [2] is focused on handcrafted and DL-based techniques for extracting features from cancer-related images. Cancer diagnosis is essentially a completely different task than that of COVID-19 analysis, which requires its own specific techniques and solutions. For instance, while nodules have solid shapes and defined locations, COVID-19 infection areas could be multifocal, ill defined, and diverse in pattern or (morphology). Therefore, the traditional radiomics filters are not applicable to the latter.

Furthermore, images from patients with pulmonary malignancies contain far fewer motion artifacts compared to COVID-19 ones, where patients suffer from dyspnea, calling for more advanced artifact-reduction techniques. DL models developed for cancer radiomics are also not transferable to COVID-19 without modifications. This is mainly due to the fact that, in a patient with COVID-19, a large number of slices may be affected, for which 3D analysis and more powerful resources are required.

With regard to differences with recent surveys/tutorials on COVID-19 research, [3] is limited to DL techniques. In this article, however, we also focus on SP modeling, applications, required medical background, and challenges/open problems. The work in [4] is more focused on the medical background and imaging modalities, and AI models are not discussed in a separate section, compared to this article, where different models and applications are separated and discussed. In [5], the authors review a subset of DL models without referring to SP methods and challenges.

AE and AI-based HP for COVID-19 image diagnoses

The worldwide outbreaks of COVID-19 and other contemporary contagious diseases have triggered a wide scope of transdisciplinary studies on epidemiology for their systematic treatment, control, prediction, prevention, management, and decision optimization. The transdisciplinary investigations into the COVID-19 pandemic have led to the emergence of AE underpinned not only by epidemiology, biology, and medical sciences, but also by computer, AI, big data, information, and signal and system sciences as well as mathematics, sociology, and economics.

Although the clinical features and imaging manifestations of COVID-19 commonly share certain characteristics, the mismatch between them is a challenging issue for its diagnosis. Furthermore, in some cases, there is a considerable mismatch between PCR and imaging manifestations. To move toward an effective detection/diagnosis and prevention plan for treatment, it is crucial to jointly incorporate epidemiology and imaging manifestations. In this section, therefore, we focus on AE and formal models of DL. This leads to the "DL Radiomics Specific to COVID-19" section, where DL applications will be explored and described.

AE models of COVID-19

AE is a transdisciplinary study of the cognitive, theoretical, and mathematical models of COVID-19 and other contagious diseases. It is recognized that AE may be better studied by big data explorations at the macro level rather than merely biological analyses at the micro level so as not to lose the forest for the trees [6]. The decision model of COVID-19 diagnoses may be described by a Cartesian product of the sets of symptoms [6] and test results.

Definition 1

Let the set of symptoms of COVID-19 be

$$\begin{aligned} S = \{&S_1(\text{fever}), S_2(\text{cough}), S_3(\text{breathing difficulty}), S_4(\text{chills}),\\ &S_5(\text{chill shaking}), S_6(\text{muscle pain}), S_7(\text{headache}),\\ &S_8(\text{sore throat}), S_9(\text{loss of taste/smell})\}, \end{aligned}$$

and the set of lab tests be

$$L = \{L_1(\text{nucleic acid}), L_2(\text{sore sample}), L_3(\text{lung image})\}.$$

The diagnosis outcomes E of COVID-19 infections are detected by the Cartesian product between the sets of logical values of detection symptoms E_S and lab confirmations E_L as follows [6]:

$$\begin{aligned} E \triangleq E_S \times E_L = \mathop{\boldsymbol{R}}_{i=1}^{9} S_i \Big| \mathrm{L} \times \mathop{\boldsymbol{R}}_{j=1}^{3} L_j \Big| \mathrm{L} \\ = \begin{cases} \left[\left(\mathop{\wedge}_{i=1}^{9} S_i \Big| \mathrm{L}\right) = T \,|\, \mathrm{L}\right] \wedge \left[\left(\mathop{\wedge}_{j=1}^{3} L_j \Big| \mathrm{L}\right) = T \,|\, \mathrm{L}\right] & /\!/ \text{ positive} \\ \left[\left(\mathop{\vee}_{i=1}^{9} S_i \Big| \mathrm{L}\right) = T \,|\, \mathrm{L}\right] \wedge \left[\left(\mathop{\vee}_{j=1}^{3} L_j \Big| \mathrm{L}\right) = T \,|\, \mathrm{L}\right] & /\!/ \text{ susceptibly positive} \\ \left[\left(\mathop{\vee}_{i=1}^{9} S_i \Big| \mathrm{L}\right) = T \,|\, \mathrm{L}\right] \wedge \left[\left(\mathop{\wedge}_{j=1}^{3} L_j \Big| \mathrm{L}\right) = F \,|\, \mathrm{L}\right] & /\!/ \text{ susceptibly negative} \\ \left[\left(\mathop{\wedge}_{i=1}^{9} S_i \Big| \mathrm{L}\right) = F \,|\, \mathrm{L}\right] \wedge \left[\left(\mathop{\wedge}_{j=1}^{3} L_j \Big| \mathrm{L}\right) = F \,|\, \mathrm{L}\right] & /\!/ \text{ negative}, \end{cases} \end{aligned} \tag{1}$$

where × represents a Cartesian product. The terms $T|L$ and $F|L$ denote a Boolean logical variable for true or false, respectively. The diagnosis results are classified in the categories of symptomatic positive, susceptibly positive, susceptibly negative, and negative. The big-R notation in (1) [9] is defined as follows.

Definition 2

The big-R calculus is a recursive operator for neatly manipulating a finite or infinite sequence of recurrent structures and a series of embedded functions, i.e.,

$$\begin{cases} \text{infinite sequence}: \mathrm{Q} \triangleq \left(\overset{\infty}{\underset{i=0}{R}} q_i\right) = (q_0, q_1, q_2, \ldots, q_k, \ldots) \\ \text{infinite set}: \mathrm{S} \triangleq \left\{\overset{\infty}{\underset{i=0}{R}} n_i\right\} = \{n_0, n_1, n_2, \ldots, n_k, \ldots\} \\ \text{infinite inductive functions}: \mathrm{F} \triangleq \overset{\infty}{\underset{k=1}{R}} f^k(f^{k-1}) = f^k(f^{k-1}(\ldots(f^0)\ldots)), \exists f^0. \end{cases} \tag{2}$$

The big-R notation may not only reduce the complexity of problem modeling but also enhance the efficiency in recursive inferences for hard problems. It is necessary for complex structure or behavior modeling, as demonstrated in the cases of (6)–(9), as well as DL in (12) and (13). Equation (1) reveals that many important symptoms and diagnoses of COVID-19 are in the domain of advanced SP as a foundation for COVID-19 diagnoses.

Cancer diagnosis is essentially a completely different task than that of COVID-19 analysis, which requires its own specific techniques and solutions.

In AE, the reproductive ratio R_0 of a contagious disease is modeled as an exponential transmission series $N_{\text{inf}}(t)$ on the $t_0 + k$th day, which is estimated by a product of initial infections $N_{\text{inf}}(t)$ and the average reproductive rate raised to the kth power:

$$N_{\text{inf}}(t_0 + k) \triangleq \overline{R}_0{}^k N_{\text{inf}}(t_0),\ \overline{R}_0 > 1.0, k \geq 0, N_{\text{inf}}(t_0) \neq 0. \tag{3}$$

Definition 3

The average reproductive rate of a pandemic transmission is reduced to the kth root of the average ratio between the number of infections $N_{\text{inf}}(t_0 + k)$ cumulatively infected at $t_0 + k$ by each initial infection $N_{\text{inf}}(t_0)$, i.e.,

$$\overline{R}_0 \triangleq \sqrt[k]{\frac{N_{\text{inf}}(t_0 + k)}{N_{\text{inf}}(t_0)}}, k \geq 0, N_{\text{inf}}(t_0) \neq 0. \tag{4}$$

For instance, the WHO has empirically estimated the $\overline{R}_0$ of COVID-19 to be in the range of 2.24–4.00, which is considerably higher than values obtained in rigorous analyses with real-world data according to (3) and (4) treating a long series of macropandemic signals.

The reproductive rate $\bar{R}_0$ in AE has been adopted as the key indicator θ for the congruous severity classified in two categories by the threshold $\bar{R}_0 = 1.0$, i.e.,

$$\theta = \begin{cases} \text{congruous}, & R_0(t) > 1.0 \\ \text{incongruous}, & 1.0 \geq R_0(t) \geq 0 \end{cases}. \tag{5}$$

However, when investigating the nature of pandemic dynamics to rigorously predict pandemic trends, we found that, to model more general and complex pandemic dynamics, the reproductive rate must be treated as a series of variables $R_0(t)$ over time. This finding has led to the formal model of the series of the dynamic reproductive rates of pandemics.

Definition 4

The series of the dynamic reproductive rates of COVID-19 is recursively determined by a long chain of causal probabilities over time:

$$\overset{n}{\underset{t=2}{R}} R_0(t-1) \triangleq \overset{n}{\underset{t=2}{R}} \frac{N_{\text{inf}}(t-1)}{N_{\text{inf}}(t-2)}, R_0(0) = 1, N_{\text{inf}}(t) \neq 0. \tag{6}$$

Simulations based on real-world data have provided highly accurate predications based on the mathematical model of the AE theory and its dynamic predictability for the pandemic hypersignal series [6].

HP for COVID-19 image diagnoses

Hypersignals are a general structure of abstract or real-world signals beyond 1D or its parallel compositions. HP intends to provide a unified mathematical model for advancing 1D signal (voice and time series) processing to 2D (images) and nD (generic hypersignals) processing [9]. The hypersignals may be embodied by sequences of images (videos), language expressions and semantics, knowledge structures, neural networks, and AI systems. Therefore, HP demands novel theories, mathematical means, and algorithms.

Definition 5

General nD hypersignals in the framework of HP theory are modeled as follows:

$$\begin{cases} \text{voice}|\mathrm{V}: \mathrm{V} \triangleq B \times T \\ \text{image}|\mathrm{I}: \mathrm{I} \triangleq B \times B \\ \text{video}|\mathrm{M}: \mathrm{M} \triangleq B \times B \times T \\ \text{hypersignal}|\mathrm{H}: \mathrm{H} \triangleq \overset{n_i}{\underset{i=0}{R}} \overset{n_j}{\underset{j=0}{R}} \ldots \overset{n_k}{\underset{k=0}{R}} \Theta(i, j, \ldots, k), \end{cases} \tag{7}$$

where B stands for a byte, T represents time, and Θ is an instance of a general abstract hypersignal.

Although classic 1D signals, such as a time series and its counterpart in the frequency domain, in general and a voice signal in particular may not be involved in COVID-19 diagnoses, they are fundamentally compatible with HS as one of its dimensions.

Definition 6

The hyperstructure model (hyper-SM) of complex images as hypersignals is formally described in the following color schemes:

$$\text{image}|\text{SM} \triangleq \begin{cases} \text{image}|\text{FG} = \underset{i=1}{\overset{|X|}{R}} \underset{j=1}{\overset{|Y|}{R}} \text{pixel}(i,j)|\text{PG} \\ \text{image}|\text{FB} = \underset{i=1}{\overset{|X|}{R}} \underset{j=1}{\overset{|Y|}{R}} \text{pixel}(i,j)|\text{PB} \\ \text{image}|\text{FR}^* = \underset{i=1}{\overset{|X|}{R}} \underset{j=1}{\overset{|Y|}{R}} \text{pixel}(i,j)|\text{PR}^* \\ \text{image}|\text{FG}^* = \underset{i=1}{\overset{|X|}{R}} \underset{j=1}{\overset{|Y|}{R}} \text{pixel}(i,j)|\text{PG}^* \\ \text{image}|\text{FB}^* = \underset{i=1}{\overset{|X|}{R}} \underset{j=1}{\overset{|Y|}{R}} \text{pixel}(i,j)|\text{PB}^* \\ \text{image}|\text{FC} = \underset{i=1}{\overset{|X|}{R}} \underset{j=1}{\overset{|Y|}{R}} [\text{pixel}(i,j)|\text{PR}^* \\ \quad \times \text{pixel}(i,j)|\text{PG}^* \times \text{pixel}(i,j)|\text{PB}^*] \\ \quad = \underset{k=1}{\overset{3}{R}} \underset{i=1}{\overset{|X|}{R}} \underset{j=1}{\overset{|Y|}{R}} \text{pixel}(i,j)|\text{PC}_k, \end{cases} \tag{8}$$

where the 2D frames are represented in six types, including FG (gray), FB (black/white), FR^* (red), FG^* (green), FB^* (blue); the composite color frame FC is represented by a Cartesian product of the three primitive color types (FR^*, FG^*, FB^*). Although gray-scale images are normally used in COVID-19 diagnoses, user interfaces of medical images widely tend to use color representation for doctors, technicians, and patients.

A paradigm of using HP for COVID-19 is represented by image frame algebra (IFA) [9], which is a set of algebraic operators formally defined on the generic mathematical model of images and video frames modeled as SMs based on (8). IFA encompasses formal algebraic operators in the categories of logical, arithmetic, color manipulation, morphology, pattern recognition, and analytic operators. The novel analytic operators of IFA are image differentiation, similarity, classification, recognition, and image knowledge base manipulations.

AE is a transdisciplinary study of the cognitive, theoretical, and mathematical models of COVID-19 and other contagious diseases.

Definition 7

The mathematical model of image differentiation $\delta(I_1|\text{FG}, I_2|\text{FG})$ for COVID-19 image diagnosis according to IFA is as follows:

$\forall$ frame size X, Y, gray pixel $p(i,j)|\text{PG} \in [0, 255]$, and gray frame $I|\text{FG} = R_{i=1}^{|X|} R_{j=1}^{|Y|} p(i,j)|\text{PG}$,

$$\delta(I_1|\text{FG}, I_2|\text{FG}) \triangleq \frac{\sum_{i=1}^{|X|}\sum_{j=1}^{|Y|} |p_1(i,j)|\text{PG} - p_2(i,j)|\text{PG}|}{255|X|\bullet|Y|} \tag{9}$$
$$\delta(I_1|\text{FG}, I_2|\text{FG}) \in [0,1].$$

A COVID-19 image differential diagnosis (IDA) algorithm has been developed in our labs based on (9). The IDA algorithm is designed for diagnosing suspected COVID-19 patients using images of affected lung with two approaches: 1) time and b) spatial differentiation of the images:

$$\begin{cases} \text{time differentiation}: Diff_t \triangleq \delta_t(I_{t_1}, I_{t_1}) \\ \text{spatial differentiation}: Diff_s \triangleq \delta_s(I_L, I_R) \end{cases}, \tag{10}$$

where the first model expresses a sequential differentiation of image series with respect to time, and the second model denotes a differentiation between the left and right images of a symmetric structure, such as the lungs or two hemispheres of the brain.

The IDA algorithm is implemented in MATLAB, which processes and diagnoses suspected infected lung images by efficient and accurate hypersignal manipulations. The time series differential method may be applied to patients whose historical (heathy) lung image is available; however, the spatial differential method will be adopted for patients by using current left and right images, where one of them may be mirrored by 180° rotation horizontally. Both methods may also be implemented by comparing a patient's image against a standard reference image.

For instance, the results of COVID-19–affected lung images, as shown in Figure 1, may be diagnosed according to the IDA algorithm, where the highlighted spots are symptoms or diagnosed outcomes. The results may be further quantitatively analyzed against certain measures, such as the severity, affected ratio, dynamic development, and effect of treatments. Therefore, based on the IDA and IFA, fast screening, AI diagnosis, real-time monitoring, early warning, and rapid reaction can be enabled for COVID-19 diagnoses and treatments in hospitals, community centers, and homes. Further examples are discussed in the "Imaging Modalities and Radiological Characteristics of COVID-19" and "DL Radiomics Specific to COVID-19" sections.

DL in HP for COVID-19 image diagnosis

Learning is a cognitive process that cumulatively acquires knowledge or adaptively generates behaviors and skills. Machine learning (ML) mimicking the brain may be formally described as follows.

Definition 8

ML is classified into six categories: object identification, cluster classification, pattern recognition, functional regression, behavioral generation, and knowledge acquisition as follows:

FIGURE 1. A formal diagnosis of COVID-19–affected lung images by IFA.

$$\begin{cases} L_i(\mathbf{x}, \mathbf{P}|\mathbf{x} \subset \mathbf{X}) \triangleq \mathbf{x} = \mathbf{P}.\mathbf{x} & // \text{ object identification} \\ L_k(\mathbf{X}, \mathbf{P}) \triangleq \mathbf{X} \subset \mathbf{P} & // \text{ cluster classification} \\ L_r(\mathbf{X}\,\mathbf{P}) \triangleq \mathbf{X} = \mathbf{P} & // \text{ pattern recognition} \\ L_g(\mathbf{X}, \mathbf{P}) \triangleq \mathbf{X} \Rightarrow \mathbf{P}(\mathbf{X}) & // \text{ functional regression} \\ L_b(\mathbf{X}, \mathbf{P}) \triangleq \mathbf{X} \Rightarrow f(\mathbf{P}(\mathbf{X})) & // \text{ behavior generation} \\ L_k(\mathbf{X}, \mathbf{P}) \triangleq \mathbf{X} \Rightarrow c(\mathbf{X}) \uplus \mathbf{K} & // \text{ knowledge acquisition,} \end{cases} \tag{11}$$

where $\mathbf{X}$ is a given variable vector or matrix of characteristic attributes of a pattern $\mathbf{P}$, such as a frame of an image, a segment of voice, a stream of video, and a sequence of sentences; $f(\cdot)$ denotes a certain function on $\mathbf{X}$; $c(\mathbf{X})$ is a formal concept; and $\uplus$ is a composition of a concept $c(\cdot)$ with existing knowledge $\mathbf{K}$. The last category of machine knowledge learning was identified and developed by Yingxu Wang.

The recent technologies of deep neural networks (DNN) and recurrent neural networks (RNNs) for DL provide a promising approach to generic ML. A multilayer (k), multioutput (m) artificial neural network (ANN), ANN(n,k,m), is illustrated in Figure 2. The ANN(n,2,2) neural network can be recursively composed by {[3 · ANN(n, 1, 1)] ° [2 · ANN(3,1,1)]} in the given topological configuration, where ° represents a composition between two adjacent layers of the $\text{ANN}(n, 1, m_1)$ and $\text{ANN}(m_1, 1, m)$ networks.

Definition 9

Let $\mathbf{X} = R_{k=1}^{K} R_{i_k=0}^{n_k} x_{i_k}^k$, $\mathbf{W} = R_{k=1}^{K} R_{i_k=0}^{n_k} w_{i_k}^k$, and $\mathbf{Y}^k = R_{j=0}^{m_k} y_j^k$. The k-layer, m-output ANN(n,k,m), $\aleph_{\text{ANN}}^k(n, k, m)$ is a set of recursively configured single-layer, multioutput ANNs, which may be coordinately trained through all layers, i.e.,

$$\begin{aligned} \aleph_{\text{ANN}}^k(n, k, m) &\triangleq \mathop{R}_{k=1}^{K} \aleph_{\text{ANN}}^1(n_k, 1, m_k) \\ &= \mathop{R}_{k=1}^{K} \mathop{R}_{j=1}^{m_k} \aleph_{\text{ANN}}^1(n_{kj}, 1, 1) \\ &= \mathop{R}_{k=1}^{K} \mathop{R}_{j=1}^{m_k} \aleph_{\text{ANN}}^1(X_k, W_{kj}^*, Y_{kj}) \\ &= \mathop{R}_{k=1}^{K} \mathop{R}_{j=1}^{m_k} \left[Y_{i_k}^k(t) = f_{kj}\left(\sum_{i_k=0}^{n_k} W_{i_k}^{k*} X_{i_k}^k(t) \right) \right], \\ W^* &= \Gamma\left(\mathop{R}_{k=1}^{K} \mathop{R}_{j=1}^{m_k} \mathop{R}_{i=0}^{n_{kj}} \aleph_{\text{ANN}}^1(X_k, W_{kj}, Y_{kj}) \right), \end{aligned} \tag{12}$$

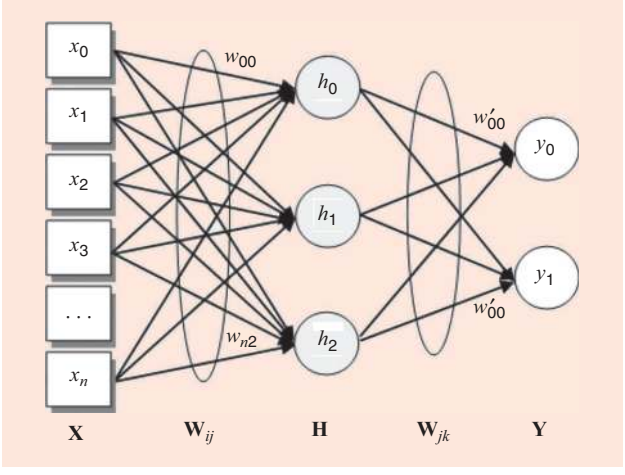

FIGURE 2. A general multilayer, multioutput deep ANN(n,k,m).

where $X^1 = R_{i_1=0}^{n_1} x_{i_1}^1(t)$ is the initial input vector, and $W^1 = R_{i_1=0}^{n_1} w_{i_1}^1(t)$ is the primitive weights corresponding to each input in X^1.

Many important and hard problems in ML are characterized as dynamic sequences of finite or infinite lengths. Such problems include video stream recognition, speech recognition, language translation, machine knowledge learning, and cognitive knowledge base manipulations as well as COVID-19 applications.

Definition 10

Let $X = R_{k=1}^{K} R_{i_k=0}^{n_k} x_{i_k}^k$, $W = R_{k=1}^{K} R_{i_k=0}^{n_k} w_{i_k}^k$, and $Y^k = R_{j=0}^{p_k} y_j^k$. The DNN($n$, k, m), $\aleph_{\text{DNN}}^k(n, k, m)$, is a k-layered recursive ANN where $k > 3$:

$$\begin{aligned} \aleph_{\text{DNN}}^k(n, k, m) &\triangleq \mathop{R}_{k=1}^{K} \aleph_{\text{ANN}}^k(n, k, m) \\ &= \mathop{R}_{k=1}^{K} \mathop{R}_{i_k=0}^{n_k} \left[Y_{i_k}^k = f\left(\sum_{i_k=0}^{n_k} W_{i_k}^k X_{i_k}^k \right) \right] = \mathop{R}_{k=1}^{K} \mathop{R}_{i_m=0}^{n_m} \left\{ y_{i_m}^m(t) \right. \\ &= \left. f\left(\sum_{i_m=0}^{n_m} \cdots \sum_{i_2=0}^{n_2} \sum_{i_1=0}^{n_1} w_{i_m}^m(t) \ldots w_{i_2}^2(t) w_{i_1}^1(t) x_i^1(t) \right) \right\}, \end{aligned} \tag{13}$$

where the activation function f is normally a continuously differentiable function, such as a sigmoid function for facilitating a least-square optimization during training.

DL methodologies underpinned by DNNs and RNNs have paved the way for solving a wide range of hard AI problems in COVID-19-related applications, as reviewed in the following sections of this article, particularly those of lung image diagnoses, chest image manifestations, related heart and brain image analyses, and pretrained neural networks for detecting typical patterns of COVID-19-related symptoms. However, it is noteworthy that DL technologies may not be the only solution for COVID-19 diagnoses and rapid analyses because DL theories and technologies for unsupervised, nondatadriven, and training-free neural networks are yet to be advanced and matured.

The fundamental theories of both AE and HP reveal the nature of COVID-19 and a rigorous approach to explain various practices in COVID-19 epidemiology and image diagnosis [6], [7]. They introduce mathematical rigor into COVID-19 pandemic analytics, which provides an indispensable framework for developing SP/ML models and algorithms for the manipulation of complex COVID-19 epidemiology and image processing. The applications of the formal models developed in the "AE and AI-Based HP for COVID-19 Image Diagnoses" section are demonstrated in the following sections throughout this article.

Based on the formal framework of COVID-19 epidemiology, the mathematical, methodological, and algorithmic insights on the COVID-19 pandemic have been explained for not only medical and pandemic control professionals but also for the SP community. The links between the overarching theories in this section and the state of the art of typical technologies and practice across the "Imaging Modalities and Radiological Characteristics of COVID-19," "DL Radiomics Specific to COVID-19," and "Challenges, Open Problems, and

Opportunities" sections are synergized, supported by a comprehensive literature tutorial.

The generic methodologies developed in the "AE and AI-Based HP for COVID-19 Image Diagnoses" section has led to the "Imaging Modalities and Radiological Characteristics of COVID-19" section for providing a set of COVID-19 imaging modalities and their radiological characteristics, which encompass a wide scope of advanced technologies, including CT, CXR, ultrasound, and MRI, as well as the multisystemic complications of COVID-19. It is recognized that all of these typical COVID-19 imaging technologies are highly dependent on the unified hyper-SM of generic image cognition and the mathematical model of IDAs for efficient COVID-19 image processing.

The links between the "AE and AI-Based HP for COVID-19 Image Diagnoses" and "DL Radiomics Specific to COVID-19" sections are demonstrated by DL technologies for radiomics specific to COVID-19 as medical applications of the general SP/ML methodologies, as modeled in the "DL in HP for COVID-19 Image Diagnosis" section. Particular examples are elaborated in the case studies of COVID-19 lesion segmentation, adverse outcome prediction, severity assessment, and diagnostic classifications.

The advantages of CXR over CT include its fast availability, ease of execution, and minimization of in-hospital transmissions.

The basic research in the "AE and AI-Based HP for COVID-19 Image Diagnoses" section has helped with the recognition and prediction of the trends in, challenges of, and opportunities for COVID-19 image diagnosis technologies discussed in the "Challenges, Open Problems, and Opportunities" section on a solid basis. A set of challenges and open problems will be discussed in depth for developing advanced COVID-19 diagnosis/prognosis models and technologies in the "Challenges, Open Problems, and Opportunities" section.

Imaging modalities and the radiological characteristics of COVID-19

The Fleischner Society and American College of Radiology, among others, recommend CT scans and CXR for COVID-19 patients with moderate to severe cases [10]. Among the chest imaging modalities, CXR is less sensitive and less specific compared to CT. The advantages of CXR over CT include its fast availability, ease of execution, and minimization of in-hospital transmissions. In addition, the CXR findings correlate well with CT findings.

In some situations where a fast assessment is necessary, the point-of-care lung ultrasound offers a radiation-free imaging modality with higher accuracy in patients without any previous cardiopulmonary disease [11]. Several recent studies have been conducted to investigate the specific characteristics of COVID-19 on the aforementioned chest images.

Different types of chest imaging patterns and distributions of lung involvement are related to the severity/stage of the COVID-19 infection and can help construct predictive SP/DL models to make decisions on hospital admission versus home isolation, non-ICU versus ICU hospital admission, monitoring of the treatment process, and the time of home discharge. In what follows, we present applications of different imaging modalities for COVID-19 diagnosis/prognosis. For each modality, we also present the required medical background related to COVID-19 for the development of SP/DL models.

CT scans

There has been considerable attention on CT imaging as the most useful imaging modality for representing COVID-19 infections. Figure 3 shows common CT patterns in COVID-19 patients, where the most prevalent are "ground glass opacities" (GGOs) and "consolidations." A GGO is a hazy transparent opacity that does not conceal lung vessels and bronchial areas. In a consolidation pattern, the air in the alveoli and peripheral bronchioles is replaced by a fluid, such as pus, water, blood, or an inflammatory material, obscuring the underlying distal airways and vascular margins.

In a research study on 645 confirmed COVID-19 patients, 88% of patients showed either pure GGOs, consolidation, or both. The appearance of pure GGO is more common in the early stage of the disease, while the appearance of GGOs with consolidations is more frequently seen in the progressive stage [12].

Another common CT pattern associated with COVID-19 is the so-called "crazy paving," referring to thickened interlobular septa and intralobular interstitium superimposed on GGOs. The crazy-paving pattern is more commonly seen

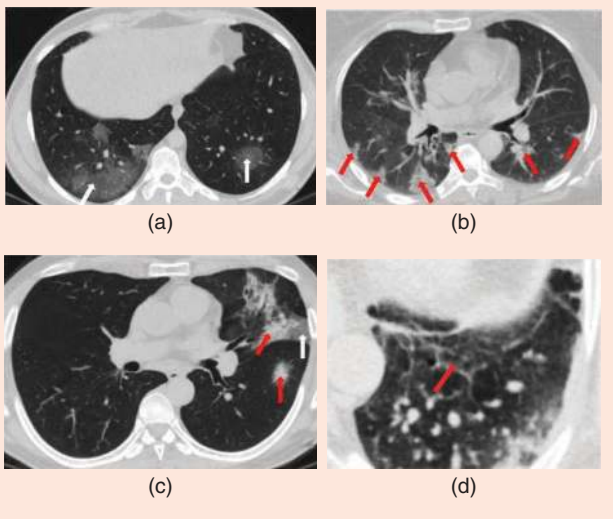

FIGURE 3. The most common CT patterns in COVID-19 patients. (a) An axial CT image of a 38-year-old man with bilateral ground glass opacities (GGOs) distributed in the posterior lung regions [13]. (b) An axial CT image of a 60-year-old woman showing scattered consolidation patterns with mainly peripheral distribution [13]. (c) An axial CT image of a 51-year-old man showing the appearance of GGOs (white arrow) and consolidations (red arrows) [13]. (d) A crazy-paving pattern in the axial CT image of a 43-year-old woman [13]. All CT images were obtained without contrast enhancement.

in the progressive stage of the disease [12]. The appearance of the crazy-paving/consolidation patterns as a sign of disease progression/severity can help radiologists evaluate the disease stage. Interlobular septal thickening, air bronchogram, and vascular enlargement are other CT findings in COVID-19 patients [12].

Distribution of lung involvement in COVID-19

CT findings of COVID-19 infections demonstrate that most of the COVID-19 patients have had "bilateral" and "multifocal" lung involvement. Bilateral involvement means that the lesions are distributed in both the right and left lungs, and multifocal involvement implies that more than one lobe (from five lobes) of the lung are affected by the disease. A systematic review of COVID-19 imaging findings declared that, in 17 out of 36 studies, bilateral lung involvement was much higher than unilateral lung involvement [12].

Researchers also showed that COVID-19 lesions, in most of the cases, are distributed in the lower lobes and have a "peripheral" instead of central distribution [12], [14]. Similarities between the CT features of COVID-19 and those of other viral pneumonias pose limitations in using CT images to diagnose COVID-19. However, in a study of 58 patients [15], six of seven radiologists could distinguish COVID-19 from other types of viral infections with an accuracy of 67%–93% and a specificity of 93%–100%. Peripheral distribution and GGOs were the most critical characteristics for distinguishing COVID-19 from non-COVID-19 pneumonia [15].

Correlation between CT findings and severity/stage of the disease

Since CT images provide high sensitivity for detecting COVID-19, they are reliable for developing SP/DL-based diagnosis models. Detecting the pattern of pulmonary involvement in COVID-19, including consolidation and/or crazy-paving patterns in chest CT images by SP/DL-powered networks, can help evaluate the disease severity. Quantifying the extent of lung involvement in COVID-19 patients is a deterministic criterion for assessing the disease's stage/severity.

Since CT images provide high sensitivity for detecting COVID-19, they are reliable for developing SP/DL-based diagnosis models.

Different CT severity measures have been introduced in the literature that can be mapped to disease severity. The manual calculation of these measures by radiologists is tedious and time-consuming. Severity measures that can be automatically quantified by SP/DL models are as follows:

Table 1. The correlation between the PO measure and COVID-19 stage [16].

COVID-19 Stage	PO
Moderate, median (interquartile range)	2.2 (0.4–7.1)
Severe, mean ± std	28.9 ± 19.2
Critical, mean ± std	49.6 ± 14.8

std = standard deviation.

Percentage of opacity

The percentage of opacity (PO) measures the volume of COVID-19 abnormalities related to the whole lung volume. It was reported in [16] that the POs for COVID-19 patients are divided into three different categories, as shown in Table 1.

Percentage of high opacity and lung high-opacity score

The percentage of high opacity (PHO) and lung high-opacity score (LHOS) introduced in [17] quantify the volume of consolidation regions in the whole lung and across the lobes, respectively.

CT severity score

The authors of [18] used a severity measure for COVID-19 patients, referred to as the *CT score*, that measures the extent of involvement based on a semiquantitative scoring for each of the five lobes. The score ranges from zero to five and is computed as shown in Table 2. The overall CT score would be between zero and 25, which is the sum of the lobar scores. (Some studies use a different scoring scale, which ends up with a CT score between zero and 20 [14]).

Francone et al. [18] conducted research on 130 COVID-19 patients and evaluated the correlation between the CT score and disease severity. They showed that the CT score is strongly correlated with the COVID-19 clinical stage and severity. For patients in the severe or critical categories, the CT score is significantly higher than for patients in the mild category [18]. A CT score greater than 18 (out of 25) can be used as a predictor of mortality in COVID-19 patients [18].

The CT score is highly correlated with a patient's age. In [18], the authors revealed that the CT score in patients older than 50 years was significantly higher than those in the age range of 26–50 years. For patients in the late stages of the disease, the CT score is higher than for those in the early stages. The CT score, together with the patient's age, can be used to predict a COVID-19 patient's death.

CXR

Some studies report that CXR images often show no lung infection in COVID-19 patients at early stages, resulting in a low sensitivity of 69% for the diagnosis of COVID-19 [19]. However, CXR is helpful for the prediction of the clinical

Table 2. The scoring system for measuring the CT severity score [18].

Lobe involvement rate	Score
No involvement	0
Involvement of less than 5%	1
Involvement from 5% to 25%	2
Involvement from 26% to 50%	3
Involvement from 51% to 75%	4
Involvement of higher than 75%	5

outcome and detection of COVID-19 in areas with limited access to reliable RT-PCR testing kits.

The most commonly observed patterns in CXR of COVID-19 patients are GGOs and consolidations with bilateral peripheral distribution [19]. The pre-existence of medical conditions, such as heart or other lung diseases, makes the interpretation of CXR images challenging. Therefore, the interpretation of CXRs in younger patients would be more reliable and predictive.

In [20], the authors developed a scoring approach for the severity assessment and outcome prediction of COVID-19 patients between the ages of 21 and 50 years based on their CXR images. In their scoring system, each lung is divided into three zones. A binary score is then given to each zone based on the appearance/absence of COVID-19 abnormalities, and the total score is in the range of zero to six. Their study on 338 patients demonstrates that there is a significant correlation between a CXR score greater than two and hospital admission. They also reported that a CXR score greater than three could predict the need for intubation.

Using a lung severity measure, referred to as the *radiographic assessment of lung edema* (*RALE*) *score*, the authors in [21] quantify the extent of lung involvement and compute correlations with the risk of ICU admission for COVID-19 patients. Recent research works have demonstrated the potentials of developing SP/DL-based models for grading the disease stage and performing outcome prediction using CXR images.

COVID-19 CT scans and CXR data sets

To ensure model generalization for clinical use, it is beneficial to train SP/DL models based on a diverse set of data sets acquired from different scanners and health centers covering a wide range of patients. CT images represent different resolutions and contrasts, depending on the type of scanner, image acquisition approach, and thickness of the slices. It is, therefore, necessary to make CT images consistent before feeding them into the processing and learning models. For a list of available CT and CXR imaging data sets along with their COVID-19-related information, please refer to [13] and [22], respectively.

It should be noted that most of the available data sets contain a specific imaging modality collected based on various equipment and protocols. This, combined with the diversity of the patient populations, makes the development of ML methods that generalize well a very challenging task.

To address this challenge, the Radiological Society of North America (RSNA) and the Society of Thoracic Radiology created the "RSNA International COVID-19 Open Radiology Database" (RICORD) [23]. RICORD is the first multi-institutional and multinational data set of thoracic CT scans and chest X-rays of COVID-19-confirmed cases. It comprises 240 thoracic CT scans and 1,000 chest radiographs annotated and manually segmented by thoracic specialists. Given the heterogeneity of the data and its free availability, this set creates a great opportunity for the research community to create better models.

Ultrasound and MRI

Besides the advantages of using CT or CXR combined with an RT-PCR test for a correct and precise diagnosis/prognosis of COVID-19, these imaging modalities have limitations, including diagnostic accuracy, logistic challenges, time-consuming assessment, and the use of ionizing radiation [11]. Despite the low sensitivity of ultrasound for the diagnosis of COVID-19 patients in the mild and moderate categories, lung ultrasound has shown high-sensitivity results in critical cases.

Due to its low cost, portability, ease of use, and being radiation free, lung ultrasound can play a crucial role in the follow-up and monitoring of patients in the ICU. Furthermore, ultrasound has been widely used for the diagnosis and monitoring of COVID-19 in pregnant women.

In Italy, health professionals used lung ultrasound as a screening tool and developed a lung ultrasound score for evaluating the severity of the disease in COVID-19 patients [24]. In another study with 93 patients, where 27 (29%) of them tested positive for COVID-19 by RT-PCR or CT, ultrasound imaging achieved a sensitivity of 89% and specificity of 59% [11]. Considering a subgroup of 37 patients without any cardiopulmonary disease, an assessment based on ultrasound revealed a sensitivity of 100% and specificity of 76% [11].

Thus, ultrasound represents a valuable imaging modality for the detection or assessment of COVID-19 severity, mainly in patients without any medical history of cardiopulmonary disease. Table 3 summarizes the

Table 3. The COVID-19 severity measures based on different imaging modalities.

Reference	Modality	Severity Measure	Lesion ROIs	Lung ROIs	Quantification Method	Range
[16]	CT	PO	All lesions	Whole lungs	$\frac{\text{lesion volume}}{\text{lung volume}}$	0–100
[14] and [18]	CT	CT severity score	All lesions	Lobewise	Scoring system	0–20 or 0–25
[17]	CT	PHO	Consolidations	Whole lungs	$\frac{\text{lesion volume}}{\text{lung volume}}$	0–100
[17]	CT	LHOS	Consolidations	Lobewise	Scoring system	0–20
[20]	CXR	CXR score	All lesions	Three specific zones for each lung (a total of six zones)	Scoring system	0–6
[21]	CXR	RALE score	Consolidations	Lung radiographic quadrant	Scoring system	0–48
[24]	Ultrasound	Lung ultrasound score	GGO and consolidations	Six specific regions for each lung (a total of 12 regions)	Scoring system	0–36

COVID-19 severity measures based on different imaging modalities.

MRI is a radiation-free medical imaging modality that can demonstrate common pulmonary image findings of COVID-19 pneumonia. However, the American College of Radiology recommends that health care practitioners minimize the use of MRI for patients with suspected or confirmed COVID-19 except for absolutely necessary cases [25].

MRI scanning has a longer acquisition time and increases the risk of disease transmission, which is critical in the case of a contagious disease, such as COVID-19. Although best known for affecting the lungs and causing respiratory infection, numerous research studies and case reports indicate that the COVID-19 disease may affect other body organs, such as the brain, heart, and kidneys. The complications and dysfunctions of other organs due to COVID-19 conditions, usually assessed by the MRI technique, are outside the scope of this article.

In summary, investigating the imaging manifestations of COVID-19 pneumonia and its radiological characteristics can help develop more powerful processing/learning models and identify new applications of SP/DL in this context. The specific imaging features, such as lesions' "peripheral" and "multifocal" distributions, help radiologists distinguish COVID-19 pneumonia from other types of community-acquired pneumonia (CAP). These imaging features can be extracted from medical images using SP techniques and incorporated in DL-based diagnostic models to enhance the model's overall performance.

The severity measures, summarized in Table 3, can be quantified using DL-based segmentation models by segmenting the target regions of interest (ROIs) of the lungs and lesions. Radiological research indicates that CXR and ultrasound imaging findings can assist in the severity assessment and treatment monitoring of COVID-19 patients admitted to the ICU, which calls for developing SP/DL-based prognosis models using these imaging modalities.

DL radiomics specific to COVID-19

In this section, we first present an overview of potential applications of SP/DL models for the diagnosis and predictive prognosis of COVID-19 infections. Then, we present different DL-based radiomic models specific to the analysis of the identified applications.

Applications

Chest imaging provides an important source of data for the diagnosis/prognosis of COVID-19 infection, assessment of the treatment response, and monitoring of COVID-19 patients. Generally speaking, applications of SP/DL models for COVID-19 diagnosis/prognosis can be classified into the following four categories:

- *Localizing COVID-19 lesions and identifying their types*: The pattern and extent of chest imaging findings is related to the stage and severity of COVID-19 and affects the treatment decision making. In the "Segmentation of COVID-19 Lesions" section, we further describe the applications of SP/DL models in localizing the involved areas and demonstrating imaging features on CT scans.
- *Outcome prediction (COVID-19 prognosis)*: To efficiently manage the limited medical resources during the pandemic, it is vital to accurately predict the risk of poor outcomes in COVID-19 patients. Some essential outcomes in COVID-19 patients are the 1) mortality risk, 2) progression to the severe/critical stage, 3) need for ICU admission/mechanical ventilation, and 4) length of the hospital stay. Predictive models are required to compute the probability of poor outcomes to help health-care professionals deliver appropriate services to high-risk patients. This application domain is presented in detail in the "Predictive Models for Adverse Outcome Prediction in COVID-19 Patients" section.
- *Severity assessment of COVID-19*: Chest imaging can be used to assess the lung infection severity in COVID-19 patients. The calculation of the percentage of parenchymal involvement and CT severity score can be achieved by segmenting the infected regions and lung areas in chest images. This is required to evaluate and quantify the severity and then predict the prognosis of the COVID-19 infection. In the "DL/SP Models for COVID-19 Severity Assessment" section, we present different SP/DL models developed for computing the lung infection rate and CT severity score metrics as two commonly used criteria for the severity assessment of COVID-19 infection.
- *Diagnosis of COVID-19 pneumonia from CAPs*: The most common COVID-19 symptoms (cough, shortness of breath, and fever) overlap with CAP symptoms. CAP is mainly caused by a bacterial infection but can also be caused by viruses. In most cases, microbiological tests for CAP, such as cultures of sputum and blood, are time-consuming, with poor sensitivity and specificity, and are not enough to identify the main pathogen. Thus, the PCR test via nasopharyngeal or oropharyngeal swab has been used for the correct identification of the source of the viral cause of CAP (like influenza). For COVID-19, a variation of PCR test, RT-PCR, has been used as the gold standard. Due to the test inaccuracy, the decision on treatment may be incorrect, i.e., antibacterial drug therapy can be administrated to all CAP patients who may not be confirmed as COVID-19 positive, but, for patients who test positive for COVID-19, this treatment is not required. Cases of negative results on RT-PCR with persistent COVID-19 symptoms are submitted to chest imaging evaluation. This application domain is further discussed in the "COVID-19 Classification Models" section, where different SP/DL models developed for COVID-19 diagnosis are presented.

Segmentation of COVID-19 lesions

In this section, we provide an overview of segmentation networks developed in the context of COVID-19. Segmentation allows physicians to identify the type and location of COVID-19 lesions, evaluate the extent of lung

involvement, and quantify the lung severity measures. Segmentation-based infection quantification models can be used to evaluate the effectiveness of different treatment solutions. Since segmentation is the first step of the handcrafted radiomics workflow, we also discuss handcrafted radiomics in this section.

Segmentation allows physicians to identify the type and location of COVID-19 lesions, evaluate the extent of lung involvement, and quantify the lung severity measures.

Imaging modalities used for the segmentation of COVID-19 lesions

Since CT images provide the most accurate COVID-19 manifestations for the grading and evaluation of infections, they have been widely used for COVID-19 lesion segmentation. In the literature, some studies [26]–[28] developed 2D models for the segmentation of lung infections in each CT slice. There have also been some 3D segmentation models that take the 3D CT volumes as input and segment the lung abnormalities on a patient-level basis.

It is worth mentioning that 3D segmentation models, which are able to segment areas of infection in the whole lung volume and assist in the patient-level assessment of the disease, would be more helpful in real practice. However, they need a large amount of pixel-level annotated lung volumes for training, which is not always practical. Furthermore, the development/training process of 3D segmentation models is more time-consuming as compared to 2D models.

Since the use of portable CXRs is more feasible for patients in the ICU, it is essential to develop segmentation models for the severity assessment of COVID-19 patients based on CXR images. A study conducted based on 2,951 COVID-19 CXR images, performed the lung infection segmentation as the first step of its COVID-19 diagnosis pipeline [29]. Using a human–machine collaboration, the researchers provided the first open access COVID-19 CXR data set with ground-truth infection masks.

ROIs

The ROIs would be different in COVID-19 segmentation models based on the research objective and can be classified into the following three main categories:

- *Localizing the infection regions without considering their types*: These studies perform a two-way segmentation approach, assigning each pixel of the CT images to either an infection or a background class.
- *Segmenting different types of COVID-19 lesions*: As mentioned previously, different types of COVID-19 infections can be correlated to the stage/severity of the disease. In this regard, some studies have segmented different types of COVID-19 lesions under different classes to further evaluate the severity of the disease [26], [30], [31]. For instance, a 3D DeepLabv3 segmentation model was proposed in [30] using CT images of 4,154 patients. The constructed model segments lung regions and different types of COVID-19 infections, including consolidations, GGOs, pulmonary fibrosis, interstitial thickening, and pleural effusion.
- *Segmentation of the lungs, lobes, and pulmonary regions*: Researchers who aim to quantify the extent of lung involvement and determine the COVID-19 severity measures consider segmenting lung/lobe/pulmonary regions besides the COVID-19 lesions. Advanced segmentation models can be trained to quantify different severity measures, such as PO, PHO, CT score, and LHOS, based on CT images. This will be further discussed in the "DL/SP Models for COVID-19 Severity Assessment" section.

Next, we investigate different DL architectures proposed for the segmentation of COVID-19 lung abnormalities.

DL architectures for the segmentation of COVID-19 lesions

Due to the distributed nature of COVID-19 infections (as discussed in the "Imaging Modalities and Radiological Characteristics of COVID-19" section), acquiring high-quality annotations is a challenging task. Consequently, on the one hand, the majority of recent studies leveraged well-established segmentation networks as a basis of their model to segment COVID-19 areas of infection. On the other hand, some research works proposed innovative encoder–decoder networks specifically designed for the segmentation of COVID-19 lesions. Here, we present these two categories.

COVID-19 lesion segmentation via well-established models

Since the U-Net model has shown superior performance for medical image segmentation, the majority of models used for the segmentation of COVID-19 lung abnormalities are developed based on it. Other segmentation architectures, such as fully convolutional networks, U-Net++, and VB-Net, have also been used to segment COVID-19 lesions in chest CT images [32], [33].

Segmentation models [developed based on convolutional neural networks (CNNs)] contain a contracting path (encoder) for extracting informative features from the input images and an expanding path (decoder) for reconstructing the mask representing the ROIs. Adopting pre-existing CNN—like densely connected convolutional networks (DenseNet), VGG, and residual neural network (ResNet)—blocks in a segmentation model's encoder path will result in extracting higher-resolution features from CT images [17], [34].

Since COVID-19 lung infections appear at varying sizes across the lungs, DL-based segmentation models containing only a single kernel size may fail to segment all lesions. In this regard, multiscale feature fusion, i.e., the integration of dilated convolutions with different dilation rates, can be added into segmentation models to help capture COVID-19 lung abnormalities in different scales [35].

Additionally, the integration of attention mechanisms in COVID-19 lesion segmentation studies [31] has shown promising results. Given the distributed nature of COVID-19

infections across the lungs, attention mechanisms can improve the segmentation performance by ignoring irrelevant features and focusing more on helpful regions for the target task.

Furthermore, to improve the prediction accuracy of COVID-19 segmentation models while preserving their computational efficiency, attention gates have been incorporated. For example, the authors of [26] proposed a new attention mechanism that enables the basic U-Net model to better understand the contextual information associated with COVID-19 abnormalities from CT slices. SCOAT-Net, proposed in [32], is developed based on U-Net++, incorporating spatialwise and channelwise attention modules and allowing the successful extraction of helpful features at both the pixel and channel levels. Their fivefold cross-validation results on a data set containing 1,117 annotated CT images from 19 COVID-19 patients achieved a Dice score of 89.48% and a sensitivity of 88.74%, yielding better results than other attention-based segmentation networks.

DL-based segmentation models need a large amount of pixel-level labeled data for efficient training. However, providing a high-quality annotated data set for training COVID-19 lesion segmentation models is challenging for several reasons. First, different types of COVID-19 lesions, such as GGOs and consolidations, have shown different shapes, scales, and contrasts. As shown in Figure 4, in comparison to lung cancer tumors, which have round shapes with clear boundaries from healthy lung tissues, COVID-19 areas of infection have irregular shapes, nonuniform contrasts, and boundaries that are hardly distinguishable from the surrounding regions.

Since COVID-19 lung infections appear at varying sizes across the lungs, DL-based segmentation models containing only a single kernel size may fail to segment all lesions.

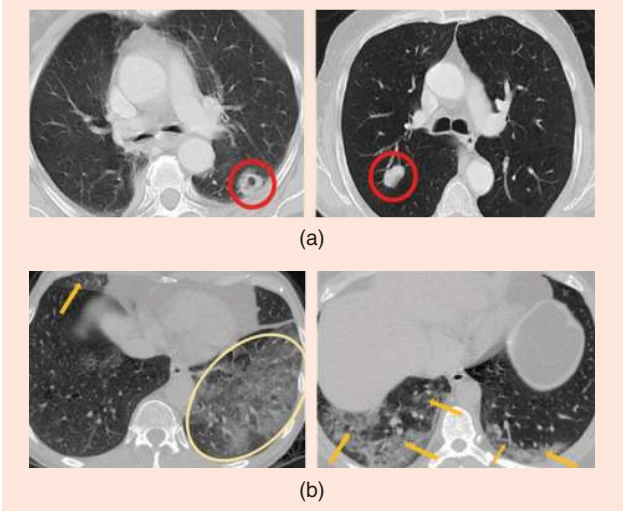

FIGURE 4. A comparison of lung abnormalities in (a) lung cancer [81] and (b) COVID-19 pneumonia disease from axial chest CT images [13]. As can be seen, lung cancer tumors have round shapes with clear boundaries from healthy lung tissue. In comparison, COVID-19 areas of infections have irregular shapes hardly distinguishable from the surrounding regions.

Second, experts may have different labeling methods, causing interobserver variabilities. Both inter- and intraobserver variabilities might be intensified due to the blurred boundaries and irregular shapes of COVID-19 lesions when experts annotate chest images. Such challenges result in a data set that is not accurate at the pixel level, which will, in turn, reduce the reliability and performance of the segmentation model. To address these issues, one approach [35] is to use a noise-robust loss function to efficiently train a COVID-19 segmentation network using low-quality training data sets.

Innovative COVID-19 specific encoder–decoder networks

Capitalizing on the fact that COVID-19 infections are distributed across the lung volume, appearing in varying shapes and sizes, some studies have developed innovative models for the segmentation of COVID-19 opacifications from scratch. These new modeling efforts target the following two key challenges of COVID-19 segmentation.

On the one hand, the training/fine-tuning process of massive segmentation networks with tens of millions of learning parameters is very expensive and time-consuming. This is not acceptable during the pandemic, when there is urgency to develop/train robust models in a timely fashion. It is necessary to design powerful but lightweight segmentation networks for the localization of COVID-19 lesions. For example, Qiu et al. [28] proposed a compact (lightweight) segmentation network based on 100 annotated CT slices from more than 40 COVID-19 patients.

On the other hand, labeling a large amount of data for training DL-based segmentation models in a short period is impractical. To address this issue, Fan et al. [27] developed a novel COVID-19 lesion network where a parallel partial decoder aggregates high-resolution features, and attention modules are then implemented to improve the model representations. A semisupervised approach is proposed where the segmentation model can be trained using a few labeled CT images and a large number of unlabeled data.

In [36], researchers proposed an innovative method for training a COVID-19 lesion segmentation model with no labeled data. Using different operations, such as random shape/noise generation and image filtering, they synthesized COVID-19 lesions and inserted them into healthy chest CT scans to make training pairs. The experiments on two different public data sets indicated their model's effectiveness. Table 4 provides more details on the different COVID-19 lesion segmentation models.

Handcrafted radiomics

Handcrafted radiomics refers to the process of extracting several quantitative and semiquantitative features from the ROI with the ultimate goal of diagnosis/prediction. Compared to DL techniques, handcrafted radiomics is less common in the problem of COVID-19 analysis, as it requires a fine delineation of the infected regions and a prior knowledge of the types

Table 4. The COVID-19 lesion segmentation models.

Reference	Input Data	Data Set Size	Data Set Diversity	Model Description	Type of Supervision	ROI	Validation Type	Bias/Overfitting Prevention
[17]	CT scans	9,749 CT volumes	Multicenter from Canada, Europe, and the United States	3D U-Net with dense block in the encoder	Fully supervised	Multiclass lesions (GGOs and consolidations), lungs, lobes	Train/test split	Data augmentation
[26]	CT scans	471 slices: 100 slices from 40 patients and 371 slices from 9 patients	Two public data sets	U-Net based with an embedded new attention mechanism	Fully supervised	COVID-19 lesions [binary/multiclass (GGOs and consolidations)]	Threefold cross validation	Data augmentation
[27]	CT scans	100 annotated CT slices from >40 patients and 1,600 CT slices from 20 patients	Multicenter	Innovative model	Semisupervised	COVID-19 lesions [binary/multiclass (GGOs and consolidations)]	External validation	NA
[28]	CT scans	3,102 annotated CT slices	Four public data sets	Innovative model (lightweight architecture)	Fully supervised	COVID-19 lesions	Fivefold cross validation	Data augmentation
[29]	CXR	2,951 COVID-19 annotated CXR images	Multicenter	Three segmentation networks with four different backbones: DenseNet, Inception-v3, ResNet-50, and CheXNet	Fully supervised	Covid-19 lesions	Fivefold cross validation	Data augmentation, encoders (DenseNet, Inception-v3, and ResNet-50) pretrained on ImageNet, and CheXNet pretrained on a CXR data set
[30]	CT scans	617,775 CT images from 4,154 patients and 4,695 annotated slices	Multicenter	DeepLabv3 based	Fully supervised	Lung regions and multiclass COVID-19 lesions including consolidations, GGOs, pulmonary fibrosis, interstitial thickening, and pleural effusion	Fivefold cross validation and external validation	The AI-based lesion quantification could evaluate the effectiveness of different treatment solutions.
[31]	CT scans	2,506 slices with COVID-19-infected lesions and 2,274 noninfected slices from 18 COVID-19 patients and 18 non-COVID people	Single center	U-Net with ResNet-101 as the backbone encoder empowered by the attention mechanism and pyramid convolution block	Fully supervised	COVID-19 lesions (including GGOs, interstitial infiltrates, and consolidations)	Fivefold cross validation	Encoder (ResNet-101) pretrained on ImageNet
[32]	CT scans	1,117 annotated CT images from 19 COVID-19 patients For external validation: eight CT slices from two COVID-19 patients	Single center	Based on U-Net++ with attention mechanisms	Fully supervised	COVID-19 lesions	Fivefold cross validation	NA
[33]	CT scans	249 CT images for training 300 CT images for validation	Training: multicenter Validation: single center	VB-Net based	Fully supervised	COVID-19 lesions, lungs, and lobes	Train/test split	NA
[35]	CT scans	558 COVID-19 CT volumes containing 76,250 CT slices	10 different hospitals	U-Net based	Fully supervised	COVID-19 lesions	Train/test split	Random Gaussian noises for data augmentation
[36]	CT scans	For lesion segmentation: 453 healthy CT volumes and 18 COVID-19 CT volumes For lung segmentation: 515 lung CT volumes	Three data sets for lesion segmentation and three data sets for lung segmentation	3D U-Net	Weakly supervised (label free)	COVID-19 lesions under one class and lung segmentation as a preprocessing step	External validation	NA

NA = not applicable; CheXNet: chest X-ray network.

of features to extract. Nevertheless, it benefits from more interpretability, as the features are self-designed.

As shown in Figure 5, handcrafted radiomics, utilized in a few COVID-19 studies, follows a multistep process. First, the infected regions are annotated. Consequently, several features are extracted from the segmented regions and fed to a conventional model, such as support vector machine (SVM), logistic regression, and decision tree, for the final decision.

Handcrafted features cover a wide range of categories, including first-order (basic intensity and shape-based features); second-order (texture features extracted from various matrices); and more advanced features, including those calculated from Fourier and wavelet transforms. Intensity, shape-based, and/or texture-based features as well as other COVID-19-related features, such as CT quantification metrics, can be leveraged [34], [46].

The authors of [37] segmented COVID-19 lesions and lung fields from the CT images of 1,658 patients confirmed with COVID-19 disease and 1,027 CAP patients. Then, based on the specific characteristics of the COVID-19 manifestations in the CT scans, they extracted a set of handcrafted features to diagnose COVID-19 cases from other types of pneumonia.

First, they calculated the PO measure for each CT volume and observed that the lung infection rate in COVID-19 patients is generally higher than in other types of CAPs, saying that PO can be helpful in COVID-19 screening models. They also extracted other volume features, including the infection rate for each lobe and pulmonary segment. Intensity-related features were extracted using intensity histograms. Since COVID-19 lung involvement is usually bilateral and multifocal, they considered the total number of infection regions for each lung field as another discriminating feature. To take into account the peripheral distribution of COVID-19 lesions, they constructed infection and lung boundary surfaces and measured the distance of each infection surface from the nearest lung boundary surface. Their experimental results showed that their diagnosis model could yield the best results when using the handcrafted features specifically designed based on COVID-19 characteristics.

Saygili [38] used three different data sets of thoracic CT and CXR from different countries and applied classical steps, i.e., preprocessing, feature extraction, and classification. The feature extractors included histogram of oriented gradient, gray-level co-occurrence matrices, scale-invariant feature transform, and local binary pattern (LBP). The evaluated classifiers include the *k*-nearest neighbors (k-NNs), SVM, bag of trees, and kernel extreme learning machine.

The best results are reported for a binary classification case (COVID-19 versus pneumonia and no findings) where the LBP + SVM reaches an accuracy of 99.02%. For another data set, the binary classification (positive and negative cases for COVID-19) using the LBP + k-NN reaches an accuracy of 98.11%. These results show the potential of classical methods, which, unlike DL approaches, are not data hungry and take less time to process. Radiomics in COVID-19 studies are mostly used in adverse outcome prediction models, explained in the next section.

Segmentation guideline

The main building block of most of the discussed segmentation techniques is CNN, whose performance is greatly improved by adding skip connections, resulting in the popular U-Net architecture. Although U-Net leads to promising results, it is restricted to 2D images; therefore, it is more suitable for specific types of COVID-19 images, such as CXRs.

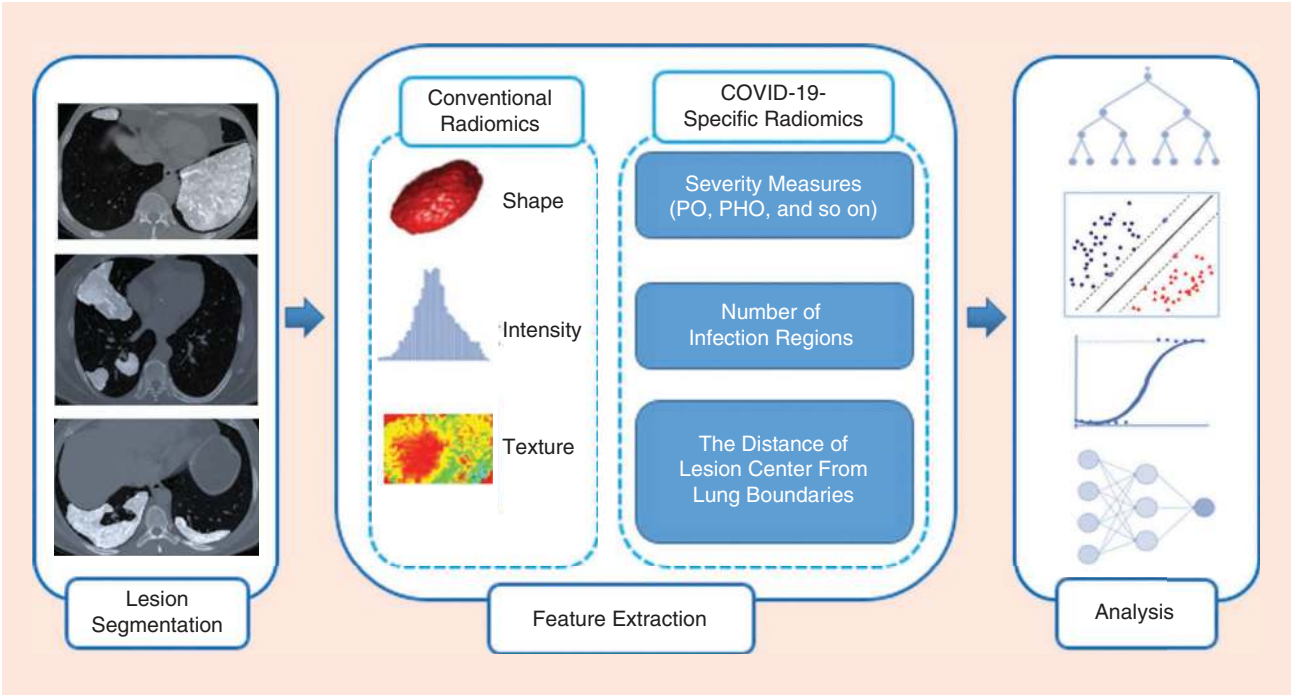

FIGURE 5. A handcrafted radiomics workflow [48].

CT scans can be processed individually, but useful information might be lost by not considering the relations between consecutive slices.

When dealing with 3D images, 3D U-Net is a more appropriate choice. It, however, suffers from memory limitations and, thus, cannot process the whole volume at once. V-Net architecture, on the other hand, benefits from improved convergence through residual connections and a better memory footprint.

All of the discussed segmentation techniques can be enhanced by introducing attention modules to obtain detailed region boundaries, which is a challenge in COVID-19 segmentation, as the boundaries of the ROI are vague compared to other problems, such as lung cancer segmentation. Furthermore, multitask training can be adopted when the data are scarce, but labels associated with other tasks are available. This approach leads to feature sharing and more generalizability. As a final note, it is worth mentioning that, although handcrafted radiomics is not a segmentation technique, it relies on accurate segmentation, as these features are commonly extracted from segmented regions.

Models for adverse outcome prediction in COVID-19 patients

As stated previously, for the efficient utilization of limited medical resources during the COVID-19 pandemic, it is critically important to accurately predict the risk of poor outcomes and help health-care professionals deliver appropriate services to high-risk patients. Although image-driven features have shown high correlation with COVID-19 outcomes, they are not the only influential factors. In other words, radiologists use image-driven features together with other clinical and risk factors to make the final decision.

Some of the clinical/laboratory pieces of information used for COVID-19 outcome prediction are patients' symptoms, laboratory test results, oxygen saturation, and comorbid diseases. Chronic lung diseases, obesity, hypertension, cardiovascular diseases, and diabetes are examples of comorbidities that increase the risk of adverse outcomes in COVID-19 pneumonia. In this section, we focus on predictive models that leverage chest medical images to estimate the risk of adverse outcomes in COVID-19 patients. In what follows, predictive models are categorized and described in terms of their model structure and target outcome (Table 5).

Structure of COVID-19 outcome prediction models

To effectively benefit from heterogeneous data resources in predicting poor outcomes in COVID-19 patients, conventional ML methods or hybrid models can be utilized, as explained here.

Conventional ML models

In some COVID-19 predictive studies, CT radiomics and other informative features specific to COVID-19 are extracted from chest images in a preprocessing step. Extracted features are then used together with clinical/laboratory data to train a shallow classifier, such as logistic regression or random forest (RF).

For instance, Chao et al. [39] extracted lobewise and whole-lung involvement metrics from CT images. These CT features, together with patients' clinical information, including age, sex, vital signs, and laboratory findings, were then fed into an RF classifier to predict the need for ICU admission in COVID-19 patients.

Homayounieh et al. [45] used a multiple logistic regression model to predict the risk of ICU admission or death in COVID-19 patients based on CT radiomics and clinical information.

Table 5. The predictive models for outcome prediction in COVID-19 patients.

Reference	Data Set Size	Data Set Diversity	Input Data	Model	Target Outcome
[39]	295 COVID-19 patients	Multicenter	CT scans and clinical/laboratory data	Random forest	Need for ICU admission
[40]	1,170 COVID-19 patients	Multicenter	CT scans and blood/urine test results	DL-based hybrid model	Risk of mortality
[41]	1,003 patients	Multicenter	Clinical, biological, and CT radiomics	DL pipeline for segmentation/DL pipeline for predicting severity evolution	Progression risk
[42]	236 patients	Single center	Clinical parameters and CT metrics	Logistic regression	ICU admission or death versus no ICU admission or death
[43]	60 COVID-19 patients	Single center	CT scans and clinical/laboratory data	DL for volume calculation/correlation with clinical factors	Need for ICU admission/mechanical ventilation
[44]	31 COVID-19 patients	Multicenter	CT radiomics	Random forest/logistic regression	Length of hospital stay
[45]	315 COVID-19 patients	Single center	CT quantification metrics and clinical/laboratory data	Multiple logistic regression	Risk of ICU admission or death
[46]	693 patients	Multicenter	Radiomic CT features and clinical/biological attributes	Ensemble consensus-driven learning	Severe versus nonsevere/short- versus long-term prognosis
[47]	133 patients in the mild stage	Single center	Temporal information of CT scans and clinical/laboratory data	A joint multilayer perceptron and LSTM network	Progression from the mild to sever/critical stage

Their study incorporated clinical data, including patients' symptoms, laboratory test results, oxygen saturation, and comorbid diseases. They indicated that CT radiomics, including pulmonary involvement and the type of infections in each lobe, could be superior to radiologists' visual assessments in predicting COVID-19 outcomes. Following these studies, one can conclude that adding clinical information to CT features can improve the overall outcome prediction performance.

Hybrid models

Hybrid models (such as multiple models, a mixture of experts, and ensemble models) are of high importance in the field of medical imaging, typically improving the initial results. While hybrid models can be developed in a variety of forms, they are mostly adopted in COVID-19 analysis in two main ways, i.e., a combination of handcrafted features and clinical/laboratory information or of DL-driven features and clinical/laboratory information:

- *Combinations of handcrafted features and clinical/laboratory information*: As explained previously, although handcrafted radiomics typically relies on fine annotations and prior knowledge of the features to be extracted, it benefits from domain knowledge. Combining the CT handcrafted features and clinical/laboratory information in a hybrid model, therefore, can help in achieving improved overall results. For instance, hybrid models can be developed to predict the probability that mild COVID-19 patients deteriorate into the severe/critical stage. In this regard, [47] developed a hybrid model where clinical/laboratory data are fed into a multilayer perceptron (MLP). The output then is integrated with extracted handcrafted features from serial CT scans and fed into a long short-term memory (LSTM) network followed by fully connected layers. The LSTM network could detect the temporal dependencies between the vector of features and achieved an area under the curve of 0.92 in distinguishing mild patients who are more likely to deteriorate into the severe/critical stage.
- *Combination of DL-driven features and clinical/laboratory information*: The superiority of hybrid models that take advantage of both image-driven DL features and clinical factors is investigated in [40]. In this study, a hybrid model is designed to predict the probability of mortality outcome in COVID-19 patients. The authors used 19,685 labeled CT slices and the blood and urine test results of 1,170 patients. The developed model consists of two successive CNN networks for analyzing CT images, a DNN network for the analysis of clinical features, and a penalized logistic regression to integrate image-driven DL features and the DL features extracted from clinical data to predict mortality outcomes. Improvements are reported when image-driven and clinical features are jointly used.

Next, the use of different model architectures and data sources in predicting COVID-19 adverse outcomes is further investigated in COVID-19 studies.

Target outcomes in COVID-19 patients

Generally speaking, in the context of COVID-19 prognosis, the target outcomes of interest include progression to the severe/critical stage, mortality risk, the need for ICU admission/ventilation, and the length of hospital stay. Here, we present different COVID-19 outcome prediction models based on these target outcomes:

- *Risk of progression to more severe stages*: Lassau et al. [41] developed a DL-based framework incorporating EfficientNet and ResNet50 as encoders to predict the severity evolution in COVID-19 patients using CT features and clinical data. Their results, obtained based on a multicenter data set containing 1,003 COVID-19 patients, demonstrated that CT features, including the GGO/crazy-paving/consolidation extents and total lung involvement, are significantly associated with the risk of disease progression to the more severe stages.
- *Mortality risk*: Chassagnon et al. [46] used an ensemble of classifiers, including nearest neighbor, SVM, decision trees, RFs, adaptive boosting, extreme gradient boosting, Gaussian naive Bayes, and MLP, to predict a short-term negative outcome (death in fewer than four days) or a long-term negative outcome (patients who did not recover after 31 days: either died after four days or were still intubated). They trained their model based on a multicenter cohort of 693 patients and obtained promising results on multiple external validation sets.
- *Need for ICU admission*: The authors of [42] implemented a logistic regression model to predict the risk of ICU admission/death based on the clinical parameters and lung involvement metrics of 236 COVID-19 patients from one health center. Lung involvement metrics can also be used to predict ICU admission or death. PO and PHO are calculated in [43] for potential correlation with clinical and laboratory factors. The obtained results show that patients with a high PO and PHO have greater need for ICU admission/mechanical ventilation.
- *Length of hospital stay*: The need for short- or long-term hospital stay for COVID-19 patients can be estimated using handcrafted features followed by an ML classifier. In [44], handcrafted features, such as first-order, second-order, and/or shape features, are extracted from CT images and fed into an RF classifier to predict the length of hospital stay of COVID-19 patients.

In summary, existing research works show that adding clinical information, including age, gender, vital signs, and blood and urine tests (laboratory findings), to image-driven features can improve the overall outcome prediction performance. In particular, clinical data (including age and gender) should be combined with image radiomic features to predict a short- or long-term negative outcome.

Furthermore, CT quantification metrics, i.e., CT features including the GGO/crazy-paving/consolidation extents, and the total lung involvement are strongly associated with the risk of disease progression to the more severe stages. It was shown that such CT quantification metrics could be superior to

radiologists' visual assessments in predicting COVID-19 outcomes. CT quantification metrics should be jointly used with clinical and laboratory parameters, i.e., patients' symptoms, laboratory test results, oxygen saturation, and comorbid diseases, for predicting the risk of ICU admission/death.

Predictive modeling guideline

Since predictive models deal with forecasting COVID-19 outcomes, they commonly incorporate clinical information as well as imaging data. When extracting features from images, one can choose to utilize a handcrafted workflow or a DL model, depending on the requirements. The former results in more interpretable features and requires fewer computational resources, while the latter is more computationally expensive and less interpretable. DL-based solutions, however, do not rely on annotated regions and a prior knowledge of the features to extract. In other words, DL techniques learn features that contribute the most to the problem at hand, where the features do not necessarily resemble radiological findings.

It should be noted that the use of handcrafted features that have been specifically designed based on COVID-19 lung manifestations, such as the infection rate in each lung, lobe, or pulmonary segment, can help in achieving better results in predicting COVID-19 adverse outcomes. Hybrid models have been developed for COVID-19 outcome prediction research using medical images and clinical data. Hybrid modeling is a fruitful direction for the development of more advanced predictive models, as hybrid models are more robust in integrating heterogeneous data resources and yielding more accurate predictions.

DL/SP models for COVID-19 severity assessment

Severity essentially refers to how much the lungs are affected and involved in the disease. COVID-19 severity assessment is of high importance due to its unique role in risk management and resource allocation. In this section, we present the existing severity assessment methodologies from two perspectives, i.e., imaging modality and type of assessment (Table 6).

Imaging modality used for the severity assessment of COVID-19

The severity of the COVID-19 can be assessed using both CXR and CT scans, where the latter, due to its 3D nature, is capable of providing a more accurate estimate of the lung involvement. Here, we provide a few examples of recent works using CXR or CT for the severity assessment of COVID-19 infection.

CXR for severity assessment

To utilize CXR for severity assessment, irrelevant, low-quality, and negative COVID-19 images need to be excluded prior to the analysis [75].

CT for severity assessment

CT scans are utilized by Li et al. [34] and divided into severe and nonsevere groups. The nonsevere cases may progress into the severe class during the treatment. Since severe and nonsevere patients have different treatment regimens, the same grouping is performed in [49]. CT scans from multiple centers are utilized by Ghosh et al. [50] for a generalizable severity assessment.

Table 6. The COVID-19 severity assessment models.

Reference	Input Data	Data Set Size	Data Set Diversity	Objective	Segmentation Method	Type of Assessment
[50]	CT	509 CT images from 101 COVID-19 patients	Multicenter	Diagnosis	DL, traditional, and manual	Handcrafted radiomics
[73]	CT	51 patients	One hospital	Diagnosis	Automatic followed by manual adjustment	Volume calculation
[16]	CT	842 COVID-19 CT volumes for segmentation; 126 COVID-19 patients classified into four clinical stages: 6 mild, 94 moderate, 20 severe, and 6 critical cases	One hospital	Progression assessment	DL	Volume calculation
[34]	CT	531 thick-section CT scans from 204 COVID-19 patients	One hospital	Progression assessment	DL	Volume calculation
[75]	X-ray	131 portable CXR images from 84 COVID-19 patients	One public data set	Diagnosis	Not required	DL
[49]	CT	176 patients	Seven hospitals with different scanners	Diagnosis	DL	Volume calculation
[76]	CT	1,110 COVID-19 patients	One public data set	Diagnosis	Traditional	Handcrafted radiomics
[74]	CT	346 COVID-19 patients	Two hospitals	Progression assessment	DL	DL
[77]	CT	99 COVID-19 patients	Two institutions	Correlation with clinical factors	DL	Handcrafted radiomics
[78]	X-ray	468 COVID-19 patients	One hospital	Progression assessment	Not required	Siamese neural networks

As stated previously, segmentation models are, typically, used to quantify different severity measures, such as PO, PHO, CT score, and LHOS, based on CT images. More specifically, to quantify (PO, CT score) and (PHO, LHOS), the model learns to segment COVID-19 infections and COVID-19 high-opacity infections, respectively. The whole lung region and lobe regions are segmented to measure (PO, PHO) and (CT score, LHOS).

For instance, in [16], researchers segmented the lung regions and COVID-19 lesions using U-Net-based commercial software and determined the PO in COVID-19 patients. Based on the clinical data, they could map the PO measure to the severity of the disease. It was concluded in this study that the median PO for patients in the moderate category is 2.2 (0.4, 7.1), for those in the severe group is 28.9 ± 19.2, and for those in the critical category is 49.6 ±14.8.

Along a similar path, the authors of [33] developed a VB-Net-based segmentation model using 249 CT volumes of COVID-19 patients to segment the lung regions, lobes, and lung infections. Their model could quantify the PO measure with an error of 0.3% over a test set of 300 CT volumes collected at a single hospital. Commercial U-Net-based software has also been used in some research works for the quantification of COVID-19 abnormalities and determining the severity of the disease [16].

Severity Assessment Types

Generally speaking, two types of severity assessment can be defined within the COVID-19 literature. The first one is to consider a classification approach, where different discrete labels are defined to assess the severity. The second type, however, aims at calculating the degree/portion of lung involvement as a measure of severity. Although the second type, referred to as *quantification*, is often followed by a classification paradigm, the degree of lung involvement is essentially embedded in the feature vector. In the following sections, we further elaborate on these two severity assessment types.

COVID-19 severity classification

Similar to most of the classification problems, COVID-19 severity classification can be solved using either handcrafted or DL methods.

- *Handcrafted radiomics*: While different engineered features have the potential to distinguish between severe and nonsevere cases, Ghosh et al. [50] proposed a handcrafted feature, referred to as the L_{norm}. This feature is defined using the maximum bone reference (B), minimum air reference (A), and mean gray-scale intensity of the lesion (L), as follows

$$L_{\text{norm}} = 100 \times \frac{L-A}{B-A}, \quad 0 \leq L_{\text{norm}} \leq 100. \tag{14}$$

  The optimum cutoff value to distinguish between severe and nonsevere cases using the L_{norm} is then obtained based on a receiver operating characteristic curve analysis. Other traditional handcrafted features, such as first-order histogram features and/or texture-based ones, can be incorporated followed by a regression model to distinguish between severe and nonsevere patients. First-order histogram features are also used in [77] for severity classification.
- *DL*: To identify the discrete severity scores of COVID-19 patients, CNN-based models can alternatively be developed [75]. A two-stage DL framework is proposed in [74] for COVID-19 severity classification. In the first stage, CT scans are individually fed to a U-Net model, whose extracted features are stored for the second stage. Through the second stage, the feature vectors are fed to a bidirectional LSTM model for the final classification.

Severity assessment via quantification

Although quantification is performed by calculating the lung and infection volume in most of the studies, it is also possible to adopt a different approach, such as using a Siamese neural network. Here, we discuss recent works performed along these two directions:

- *Quantification via volume calculation*: The PO index and average infection Hounsfield unit (HU) are calculated in [34] to quantify the severity, followed by dividing the patients into two groups of severe and nonsevere. The infection and GGO ratios are calculated in [49]. These two measures, along with several other quantitative features, are further fed to an RF classifier to classify patients as severe and nonsevere.
- *Quantification via Siamese neural networks*: This approach consists of two identical models, in terms of weights and parameters, with the goal of finding the similarity between the two inputs. Beside having several applications, Siamese models, in particular, convolutional Siamese neural networks, can be adopted for COVID-19 severity assessment [78]. In such scenarios, the Euclidean distance between the two final layers is calculated as a measure of the difference between the inputs. Therefore, the distance between a COVID-19 and normal scan can show the degree of abnormality. Utilizing a pool of normal images, the median distance can represent the severity.

The taxonomy of the COVID-19 prognosis models is presented in Figure 6.

Severity assessment guideline

Quantifying the extent of lung involvement in COVID-19 patients is the main criterion for assessing the disease's stage/severity. Different severity measures have been introduced in the literature with the goal of providing a discrete or continues severity score. While the former is less accurate, it can be performed using conventional classification techniques. It is, therefore, faster and not necessarily dependent on annotation. The latter, although being more burdensome, results in a more accurate estimation of the severity by determining the degree of lung involvement.

However, the manual calculation of these measures by radiologists is tedious and time-consuming, therefore motivating the development of SP/DL models to automatically quantify severity measures. Different types of chest imaging

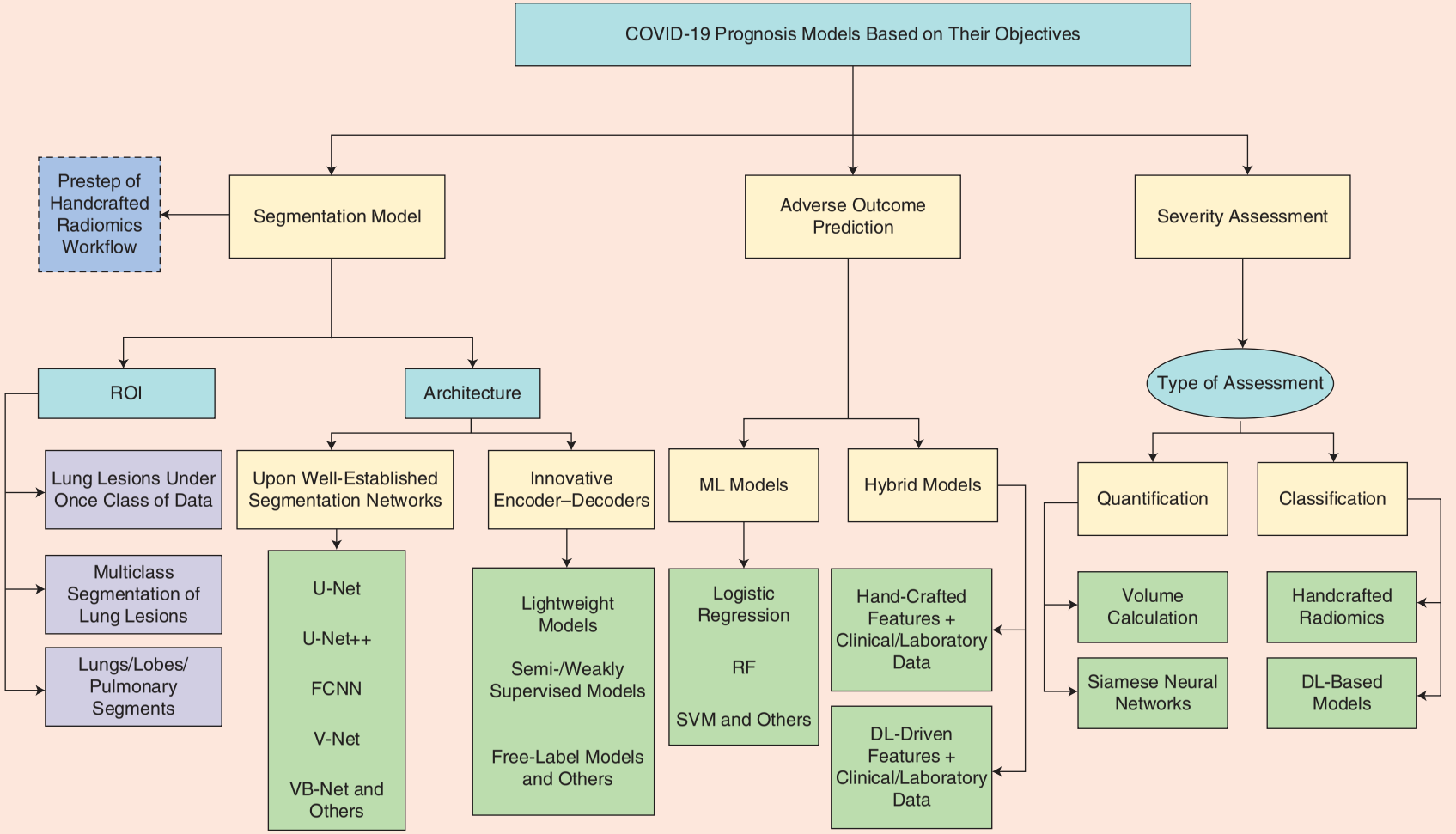

FIGURE 6. The taxonomy of the COVID-19 prognosis models based on their objectives. FCNN: fully convolutional NNs.

patterns and distributions of lung involvement are related to the severity/stage of COVID-19 infection and can help in constructing SP/DL models for autonomous severity assessment. In particular, the appearance of the crazy-paving/consolidation patterns is a sign of disease severity and can be used for model development.

Among different imaging modalities, CT images provide the high sensitivity required for evaluating changes in the severity of the disease. Particularly, CT images can be used for the detection of pulmonary involvement patterns, including consolidation and/or crazy-paving patterns, to help evaluate the disease severity.

COVID-19 classification models

The development of DL-based COVID-19 classification models can be approached from four main perspectives, as shown in Figure 7. The first aspect is the annotation dependency of the developed frameworks. The second one is whether the proposed methods consider a binary or multiclass classification, followed by the third perspective, which focuses on the imaging modality used for the classification, as, based on the modality, different solutions are admissible. The DL model architecture is another important aspect of the developed frameworks. Next, we discuss these four categories in detail. Table 7 summarizes how different studies approach the aforementioned categories.

Annotation dependency

Annotation dependency refers to whether the developed COVID-19 classification models rely on annotated images as inputs. Annotation can be related to either segmenting the whole lung region or the infected areas from the chest image. In this regard, we categorized studies into three groups of 1) no annotation required, 2) lung segmentation required, and 3) infection segmentation required. These three groups are discussed in the following sections.

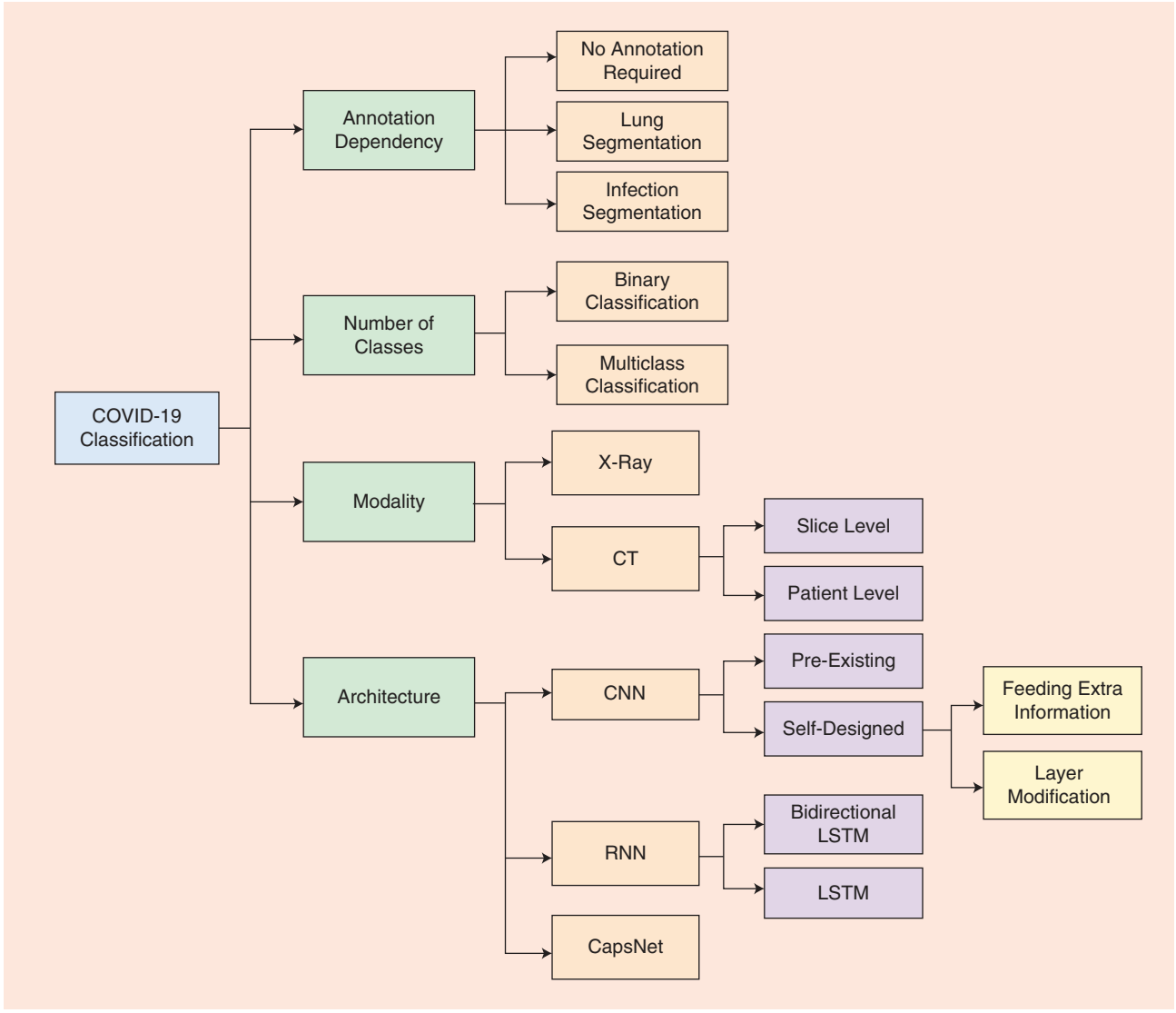

FIGURE 7. The taxonomy of the COVID-19 classification techniques using DL. CapsNet: capsule network.

Table 7. The COVID-19 classification models.

Reference	Input Data	Data Set Size	Data Set Diversity	Number of Classes	Architecture	Bias and Overfitting Prevention	Annotation Dependency
[51]	CT	419 COVID-19 486 non-COVID	18 centers	Binary	Inception and ResNet CNN	Data augmentation and transfer learning	Lung segmentation
[52]	CT	4,356 chest CT exams from 3,322 patients	Six hospitals	Multiclass	ResNet50 CNN	Data augmentation	Lung segmentation
[72]	X-ray	403 COVID-19 721 normal	Multicenter	Binary	VGG CNN	Generative adversarial network-based data augmentation and transfer learning	Not required
[53]	X-ray	266 COVID-19 5,538 non-COVID 8,066 normal	Multicenter	Binary	CapsNet	Loss modification and transfer learning	Not required
[54]	CT	44 COVID-19 55 non-COVID	Three hospitals	Binary	Inception CNN and ensemble of classifiers	Transfer learning	Infection segmentation
[22]	X-ray	266 COVID-19 5,538 non-COVID 8,066 normal	Multicenter	Multiclass	CNN	Data augmentation and transfer learning	Not required
[55]	X-ray	45 COVID-19 1,591 non-COVID 1,023 normal	Multicenter	Multiclass	ResNet CNN	Data augmentation and transfer learning	Not required
[56]	X-ray	50 COVID-19 50 normal	One public data set	Binary	ResNet CNN	Transfer learning	Not required
[57]	X-ray	127 COVID-19 500 non-COVID 500 normal	Two public data sets	Multiclass and binary	Darknet [you only look once (YOLO) base model]	NA	Not required
[58]	X-ray	266 COVID-19 5,538 non-COVID 8,066 normal	Three public data sets	Multiclass	Ensemble CNN	Data augmentation	Not required
[48]	CT	449 COVID-19 495 non-COVID 425 normal	Three hospitals	Multiclass	Encoder–decoder CNN	Multitask learning	Not required
[70]	CT	127 COVID-19 500 non-COVID 500 normal	Two public data sets	Multiclass	Xception CNN	Transfer learning	Not required
[63]	X-ray	1,525 COVID-19 1,525 non-COVID 1,525 normal	Several public data sets	Multiclass	LSTM	NA	Not required
[59]	X-ray	25 COVID-19 25 non-COVID	Two public data sets	Binary	CNN	NA	Not required
[61]	X-ray	599 COVID-19 24,622 non-COVID 18,881 normal	Multiclass	One-class anomaly detection	EfficientNet CNN	Data augmentation and transfer learning	Not required
[69]	CT	159 COVID-19 90 non-COVID	Multicenter	Binary	ResNet CNN	Data augmentation and transfer learning	Lung segmentation
[62]	CT	150 COVID-19 150 non-COVID 150 normal	Two hospitals	Multiclass	Multiscale CNN	Loss modification and data augmentation	Lung segmentation
[60]	CT	88 COVID-19 101 non-COVID 86 normal	Hospitals of two provinces in China	Multiclass	CNN	NA	Lung segmentation
[64]	CT	540 COVID-19 229 normal	One hospital	Binary	CNN	Data augmentation	Lung segmentation
[65]	CT	110 COVID-19 224 non-COVID 175 normal	Three hospitals	Multiclass	CNN	Changing sampling probability	Infection segmentation
[66]	CT	20 COVID-19 282 non-COVID	One public data set	Multiclass	Bidirectional LSTM and CNN	Changing sampling probability and data augmentation	Lung segmentation
[67]	CT	146 COVID-19 149 normal	One hospital	Binary	CNN	Data augmentation	Lung segmentation
[68]	CT	366 COVID-19	Four centers	Binary	3D CNN with integrated clinical data	Loss modification and data augmentation	Lung segmentation

COVID-19 classification without annotation

Studies that do not include any segmentation as a preprocessing step essentially feed the developed model with raw images. As CXR images are single slices and simpler to process compared to CT scans, they are utilized without annotation in most of the studies. The work in [57] is an example of such a study, where raw CXR images are fed to a DL model for binary and multiclass COVID-19 classification.

Das et al. [70] and Islam et al. [63] also utilized CXR images without annotation for a three-way COVID-19 classification. Although using CT scans, [48] is independent from segmented inputs. It exploits annotation labels in the output layer to develop a multitask training framework. In other words, both classification and segmentation are aimed at in this work.

Lung segmentation for COVID-19 classification

Lung segmentation is the first step in many COVID-19 classifications studies, as it eliminates unessential information. Gozes et al. [69], for instance, used a pretrained U-Net model for this task. Since the segmentation model should be able to annotate the lungs even in the presence of COVID-19 opacities, the U-Net model is fine-tuned on a data set of interstitial lung disease cases. More advanced lung segmentation models are also used in COVID-19 studies. The authors of [62], for instance, proposed a multiwindow U-Net that incorporates several windows instead of the standard HU window. Furthermore, this study uses a sequential information attention module to integrate all CT slices.

Infection segmentation for COVID-19 classification

Besides lung segmentation, some COVID-19 classification studies rely on segmenting the pulmonary regions of infection. Xu [65], for instance, used a 3D CNN model trained on pulmonary tuberculosis for infection segmentation. Although this model is not trained on a COVID-19 data set, it can still extract candidate patches. The annotation results are consequently used to form cubic patches around the regions of infection, which are then fed to the classification model. Based on the COVID-19 characteristics, such as GGO, Wang et al. [54] manually delineated the CT scans to extract all of the ROIs, from which two or three patches are randomly selected as the input to the CNN model for classification purposes.

COVID-19 classification types

Binary or *multiclass COVID-19 classification* refers to whether the problem is considered as COVID-19 versus all other possible categories as one class or if all of the classes are treated separately. These two approaches are investigated in the following sections.

Binary COVID-19 classification problems

The work in [57] is an example of binary classification, where the goal is to distinguish between COVID-19 and non-COVID cases. Non-COVID cases include both normal and pneumonia patients. The authors of [56] explore three different COVID-19-related binary classification problems, in each of which COVID-19 is classified against a different class, including viral pneumonia, bacterial pneumonia, and normal. The obtained results show that COVID-19 is best distinguishable from bacterial pneumonia. Besides positive and negative COVID-19, patients can be classified based on other clinical outcomes. Meng et al. [68], for instance, consider high and low risk as the binary classification labels.

Multiclass COVID-19 classification problems

The authors of [57], besides considering a binary classification problem, try to solve a multiclass classification consisting of three classes of COVID-19, pneumonia, and normal. The obtained accuracy, however, is lower than the binary scenario. The same categorization is followed in [63]. The authors of [48] also followed a three-way classification, with the difference that all diseases other than COVID-19 are considered as the "others" class to be classified against COVID-19 and normal subjects. COVID-19, pneumonia, and other diseases are considered as three separate classes in [70]. Since [65] uses annotated infection patches as inputs to a CNN model, it also considers an irrelevant-to-infection class to exclude incorrectly segmented areas.

It is worth mentioning that, unlike binary and multiclass approaches, COVID-19 classification is considered as a one-class anomaly detection in [61], where the model's output is the anomaly score of the input, along with a confidence score that determines the model's confidence in its prediction. Consequently, subjects with a high anomaly score or low confidence score are considered as positive for COVID-19.

Imaging modality used for COVID-19 classification

CXR and CT are two common imaging modalities considered in COVID-19 classification studies. These two modalities, however, require different processing strategies, as described in the following sections.

COVID-19 classification via CXR images

CXR images are 2D, and, as such, processing techniques to incorporate the relation between images are not required. CXR images can be independent inputs to a DL model. The works in [57], [63], and [70] are examples of using CXR images for classification tasks.

COVID-19 classification via CT scans

Unlike CXR images, CT scans are 3D in the sense that each patient is associated with several 2D slices. As a result, analyzing CT scans requires specific strategies, the first of which is a slice-level classification, where slices are treated independently with the goal of assigning labels to separate slices. Patient-level classification, on the other hand, tries to make the final decision using all of the available slices.

- *Slice-level classification*: The work in [48], as an example of a slice-level classification algorithm, uses separate

slices as inputs to a DL model, where slices are gathered from three different data sources and preprocessed to have consistent size, resolution, and contrast. In [62], patient-level labels are assigned to all of the slices, and a 2D CNN model is leveraged. This strategy, however, can cause inconsistency when a slice without any visible manifestation is assigned with a COVID-19 or pneumonia label. The work in [67] is another example of a slice-level classification model, where target slices are manually selected to train the CNN model. At the test time, however, this study averaged over all of the probabilities to form the patient-level classification. Therefore, the underlying study can be considered as cross sections of the slice- and patient-level classifications, bringing us to the discussion in the next part, i.e., patient-level classification.

- *Patient-level classification*: Patient-level classification using CT scans requires a voting strategy to combine the slice-level outcomes. The voting mechanism is of particular importance, as the whole CT volume typically cannot be processed at once. Different voting mechanisms have been developed in the literature, including the following items:
  - *Volumetric scoring*: In [69], 2D slices are first processed to form the slice-level outcomes. Summing over the activation maps of the detected positive slices, consequently, results in the aforementioned volumetric score. It is worth mentioning that only activations above a predefined threshold are considered in the summation. The obtained COVID-19 score can also be considered as the extent of the disease in a patient's lungs.
  - *Pooling operations*: For patient-level classification, one approach is to combine different models (e.g., parallel CNNs) in a parallel architecture. The results from individual slices can then be aggregated through pooling operations [60]. Similarly, Li et al. [52] incorporate parallel CNNs, the results of which are aggregated through a max-pooling operation.
  - *Whole CT volume*: To leverage the information from all of the CT scans and capture their relations, Wang et al. [64] feed their developed CNN model with the whole CT volume, which is concatenated with the segmented lung mask. The same strategy of feeding the whole CT volume is also used in [68].
  - *Bayesian merging*: A noisy-or Bayesian function is adopted in [65] to combine the outcomes of several infection patches.
  - *RNN-based merging*: Using RNNs is another strategy to combine the slice-level information and consider the spatial relations. This group of models is discussed in the "DL Architectures for COVID-19 Classification" section.
  - *Multistage frameworks*: Designing a multistage framework is also a common patient-level classification approach. Mei et al. [51], for instance, designed a two-stage workflow, where, in the first stage, abnormal slices are detected using a previous pretrained pulmonary tuberculosis detection model. The top 10 candidate slices are then fed to another CNN, in stage two, to identify slices with positive COVID-19. The final outcome is ultimately set as the average of the slice-level prediction of a patient's 10 most abnormal candidates.

DL architectures for COVID-19 classification

Although different DL architectures are applicable to the task of image classification, in the COVID-19 scenario, discriminative models, including CNNs, RNNs, and capsule networks (CapsNets), are the most commonly used ones. These networks and how they are incorporated in COVID-19 classification studies are explained here.

CNN-based COVID-19 classification models

CNNs are a stack of convolutional and pooling layers, often followed by fully connected ones. Since the trainable filters share weight across the whole image, these networks are computationally effective and can extract local features from the input. CNNs have shown promising results in the field of image processing, including COVID-19 classification. Although it is possible to design a CNN from scratch, most of the studies have built their models upon pre-existing successful CNN models:

- *Pre-existing CNN models*: Since the start of the outbreak, the following pre-existing CNN models for COVID-19 classification have been used:
  - *Darknet-19 model*: The DarkCovidNet model proposed in [57] is a modification of the Darknet-19 model, which is the basis of the you only look once object detection system. The proposed DarkCovidNet consists of 17 convolutional layers, which are followed by pooling layers and, eventually, one fully connected layer for the final classification.
  - *Inception model*: This is another CNN model commonly utilized in COVID-19 studies. In [70], for instance, an extreme version of the Inception model, referred to as *Xception*, is used. Two other variations of Inception, referred to as *InceptionV3* and *Inception-ResNetV2*, are exploited in [56], along with three variations of the popular ResNet architecture, namely, ResNet50, ResNet101, and ResNet152. The obtained results show superior performance for ResNet50 model.
  - *ResNet models*: ResNet50 is the basis of the model proposed in [55], referred to as the *COVID-ResNet*. This model is trained in three stages, where, in each stage, the image size is increased gradually. The ResNet50 model of [69] is followed by a gradient-weighted class activation mapping localization to verify the pathological areas focused through the training process. The resulting map can provide insights to the radiologist.
  - *VGG models*: Besides the variations of ResNet [69] and Inception, VGG models, such as VGG19 [59], are common choices.

- *EfficientNet*: Utilizing compound coefficients to scale up CNNs, this is another architecture used for COVID-19 classification in [61].
- *Ensemble models*: Besides adopting the pre-existing CNN architectures, it is also possible to develop ensemble frameworks to leverage the potentials of different CNN models. Such a strategy is used in [58], where several models, including VGG and ResNet, are combined.

- *Self-designed CNN models*: Based on the identified requirements, some studies have designed their own specific CNN models for COVID-19 classification. A multiscale CNN, for instance, is proposed in [62], where intermediate CNN representations are aggregated through a global max-pooling operation to make the final decision. The self-designed CNN model proposed by Wang et al. [64] consists of three subsequent blocks, the first of which is a vanilla 3D CNN, followed by a residual block. The last part is a progressive classifier, containing convolutional and fully connected layers.

  Besides focusing on designing the layers of a CNN, another strategy is to feed the model with information other than the raw image. Such a strategy is leveraged in [65], where the distance between the center of infection and pleura is concatenated with a fully connected layer. This distance can contribute to a more accurate classification, as COVID-19 infection has a pleural distribution, partly distinguishing it from other diseases. Meng et al. [68] utilized patients' clinical factors, such as gender, age, and chronic disease history, as the additional information to be concatenated with the CNN's fully connected layer. More heterogeneous factors, including travel and exposure history as well as symptomatology, are incorporated in the model designed by Mei et al. [51].

RNN-based COVID-19 classification models

RNNs are especially useful in medical imaging when the goal is to process the whole volume or analyze follow-up studies. Since RNNs are subject to the problem of vanishing gradients, LSTM networks are commonly used as an effective alternative. The vanilla LSTM is not designed for extracting local features from images, and, as such, this network is often combined with a CNN to make use of its weight-sharing advantages.

Such a model is utilized in [63] for COVID-19 classification, resulting in a CNN–LSTM design. In the underlying study, 12 convolutional layers are first incorporated to extract features from CXR images. The output of the CNN is then fed to an LSTM, the result of which determines the probability of COVID-19, pneumonia, and normal classes. While a conventional LSTM considers only forward relations, bidirectional LSTMs additionally take the backward relations into account. Such models are incorporated in [66] for COVID-19 classification.

CapsNet-based COVID-19 classification models

CapsNets are relatively new DL architectures, proposed to solve the incapability of CNNs to recognize spatial information. Each capsule in a CapsNet consists of several neurons to represent an object's instantiation parameters as well as its existence probability.

The main feature of the CapsNet is its routing-by-agreement process, through which capsules in a lower layer predict the outcome of capsules in the next layer. The parent capsules take these predictions into account based on the similarity (agreement) between the prediction and actual outcome. Using the routing by agreement, CapsNet is capable of recognizing spatial relations between image instances and, therefore, handling much smaller data sets compared to CNNs. The authors of [53] recently exploited CapsNets for the problem of COVID-19 classification using CXR, showing improvements over the CNN counterparts. The proposed architecture, referred to as *COVID-CAPS*, consists of several convolutional, pooling, and capsule layers, the output of which determines the probability of positive COVID-19. A similar architecture can be utilized with CT scans by combining slice-level final capsules to form a patient-level decision.

The ultimate goal of developing COVID-19 diagnosis/prognosis models is to use them in clinical applications and reduce the health-care system's workload during pandemic conditions.

Classification guideline

COVID-19 classification can be performed in a slice- or patient-level manner. While the former is more simple, requiring less-complicated models and computational resources, it does not utilize the available information in a whole volume or study. While CNNs are the architecture of choice within the medical imaging domain, they might not be able to capture the spatial information between infections, which is of high importance in COVID-19 classification.

CapsNets may result in a more promising performance, especially when a large data set is not available. When processing a whole volume or even if follow-up studies are considered, analyzing all of the images at once may not be possible using conventional models. RNNs are more proper choices for processing serial images. CNNs and CapsNets, however, can be utilized by designing a voting strategy between slices.

Challenges, open problems, and opportunities

In this section, first, we focus on the limitations and challenges of developing COVID-19 diagnosis/prognosis models. Then, we discuss open problems and potential opportunities for SP research by highlighting the problems and challenges of developing SP/DL models for COVID-19 management.

Challenges in developing COVID-19 diagnosis/prognosis models

The ultimate goal of developing COVID-19 diagnosis/prognosis models is to use them in clinical applications and reduce the health-care system's workload during pandemic conditions.

Some models proposed for the diagnosis and prognosis of COVID-19 have shown successful results in real applications and enhanced the performance of junior radiologists to senior level [30]. In what follows, we summarize the limitations and challenges specific to the development of COVID-19 diagnosis/prognosis models:

- Different types of respiratory infections have shown similarities with COVID-19 pneumonia in lung manifestations, which makes it challenging for both radiologists and DL-based models to distinguish COVID-19 disease from other types of CAPs based on medical images. One characteristic that helps differentiate COVID-19 images from other CAPs is the multifocal and peripheral distribution of infection regions across the lung. In this regard, extracting handcrafted radiomic features based on COVID-19 imaging characteristics can help enhance DL models' performance in diagnosing COVID-19 cases.
- The ambiguous boundaries and irregular shapes of COVID-19 lung infections exacerbate inter- and intraobserver variabilities when annotating infection regions by expert radiologists. This makes it necessary to work with a team of expert radiologists to annotate lung lesions accurately and take the average of annotations, which is impractical in pandemic conditions. This issue results in generating inconsistent and inaccurate annotations that cause bias in DL models. Noise-robust loss functions and adaptive self-ensembling training frameworks can help achieve better segmentation results while being trained on inaccurate annotated data sets.
- The lack of annotated labels in developing COVID-19 lesion segmentation models is another challenge in this context. Inspired by the intensity range and specific characteristics of COVID-19 lesions in chest images, it is possible to synthesize lesions inside healthy lung images and generate pairs of chest images and their associated masks. This synthesized data set can then be used for training segmentation models [36] to tackle the lack of annotated labels. Semisupervised learning, incorporated in [27], is another method that can be used to address this issue.
- To accelerate the annotation process, human–machine collaboration methods can be developed. It is shown in [29] and [33] that collaborative interactions between radiologists and DL models can dramatically reduce the annotation time.
- To improve the accuracy of COVID-19 outcome prediction and severity assessment methods, CT severity measures can be quantified by DL-based segmentation models and then incorporated within prediction and assessment models. Furthermore, there is no consensus on which severity scoring system to use, which is a challenge for the adoption of automatic assessment models. This requires further research on the evaluation, comparison, and design of a unified severity measure.
- Due to restrictions imposed on preserving patients' privacy and also health centers' strict data-sharing protocols, collecting sufficient data is a common problem in AI-based research in the medical domain. This issue, however, gets more challenging during a pandemic when an unknown disease is threatening people's lives, and researchers need to develop robust models in a short period of time. Conventional data-augmentation strategies are the most commonly used methods to overcome this problem by applying conventional transformations on original images [61], [69].

  Generative adversarial networks have been used in [72] to augment the training data set by generating fake instances from original images. Pretraining DL models on both natural-image data sets [28], [56] and similar medical data sets [53] have shown promising results in improving COVID-19 diagnosis/prognosis models. Developing DL models using multitask learning to perform COVID-19 classification and lesion segmentation at the same time can provide improved results [48].

  Incorporating CapsNets, which are less data demanding compared to CNNs and can be trained using smaller data sets, is another approach that has shown potential for the development of COVID-19 diagnosis models [53]. A class-imbalanced data set is another common challenge in the medical imaging domain, including COVID-19 models. Modified loss functions can tackle this issue by assigning more penalties to the misclassified instances/pixels of the minority class [26], [62], [68]. Data resampling is another strategy that can address this problem. Li et al. [71] introduced a new offline sampling strategy that ranks the non-COVID-19 samples based on their diversity and difficulty levels. The most informative samples can then be used to decrease the training time while maintaining the overall performance of the model.

Open problems

In this section, we focus on open problems and potential opportunities for SP research by highlighting the problems and challenges of developing SP/DL models for COVID-19 management.

- As stated previously, to have an effective detection/diagnosis and prevention treatment plan, it is crucial to jointly incorporate epidemiology and imaging manifestations. Currently, works on AE and medical imaging are mainly performed separately. A focus on the development of diagnosis/prognosis DL models for COVID-19 based on a simultaneous epidemiology study has not yet been explored. This opens a fruitful direction for future research for the development of a comprehensive diagnosis/prognosis DL modeling framework.
- COVID-19 patients suffer from dyspnea; as such, there are inevitable motion artifacts in the acquired images. This is in contrast to most other medical images, where motion artifacts are rarely present. The artifacts in the COVID-19 images sometimes overlap with the main areas of infection, making the diagnosis/prognosis challenging even for experienced radiologists.

To eliminate the effect of artifacts, most of the studies simply remove the noisy data from the data set. This, however, reduces the generalizability and applicability of the model in clinical practice. An alternative solution is to the approach of using advanced artifact-reduction techniques, among which the adaptive techniques are of higher capability, as they can adjust and track the signal under noisy conditions.

- COVID-19 involves a large volume of the lung and is sparsely distributed around the lung volume. This is, in particular, in contrast to medical images in which the ROI is found in a specific location of the organ. Analyzing and extracting patterns from the COVID-19 images require sparse filtering techniques within the SP domain.
- As COVID-19 infection is distributed in the whole lung volume, the relation between the image slices is of high diagnostic and prognostic importance, calling for specific 3D filtering and pattern-recognition approaches.
- A key issue with chest CT scans is exposing patients to harmful radiation. In this regard, using low- or ultralow-dose scanning is of high interest. More importantly, low-dose examinations are associated with less cancer risk, especially in young women.

Figure 8 shows the obtained CT scans for three different patients at three different dose levels, i.e., standard, low, and ultralow. According to this figure, although low- and ultralow-dose images have more visible artifacts, they can still reveal the presence of COVID-19 infection. The artifacts, however, can hamper the effective training of the model. Furthermore, collecting a data set of low-dose scans may not resolve this issue, as, besides dose, other factors, such as the patient's weight, can influence the quality of the image, leading to a wide variety of possible artifacts. This calls for advanced SP/DL models that can cope with images at different resolutions while providing the same level of diagnosis/prognosis performance.

- COVID-19 is relatively new, and, as such, large data sets are not easily accessible. Therefore, the developed SP/DL models should be capable of handling small data sets and yet capturing informative features.
- COVID-19 lesion segmentation models are mainly developed using the CNN architecture. Since convolutional kernels are local operators, CNN models cannot fully capture long-range dependencies in the underlying image. In this context, transformers have recently revolutionized natural language processing (NLP) research due to their capability of learning long-range dependencies in a sequence of tokens. Following their promising results in different NLP tasks, transformers have found their way to the image processing domain.

One important challenge associated with COVID-19 analysis is the disease manifestation in patients with complications other than COVID-19.

The potentials of applying transformers for medical image segmentation have recently been explored in different studies [79], [80], where it was shown that transformers can outperform CNN-based attention networks. For example, TransUNet [79] adds transformers to the encoder path of the segmentation network (on image patches extracted by CNN blocks) to better learn the global context. Using a pure transformer as the encoder, [80] proposed a segmentation network for 3D medical image segmentation that can learn global information at different

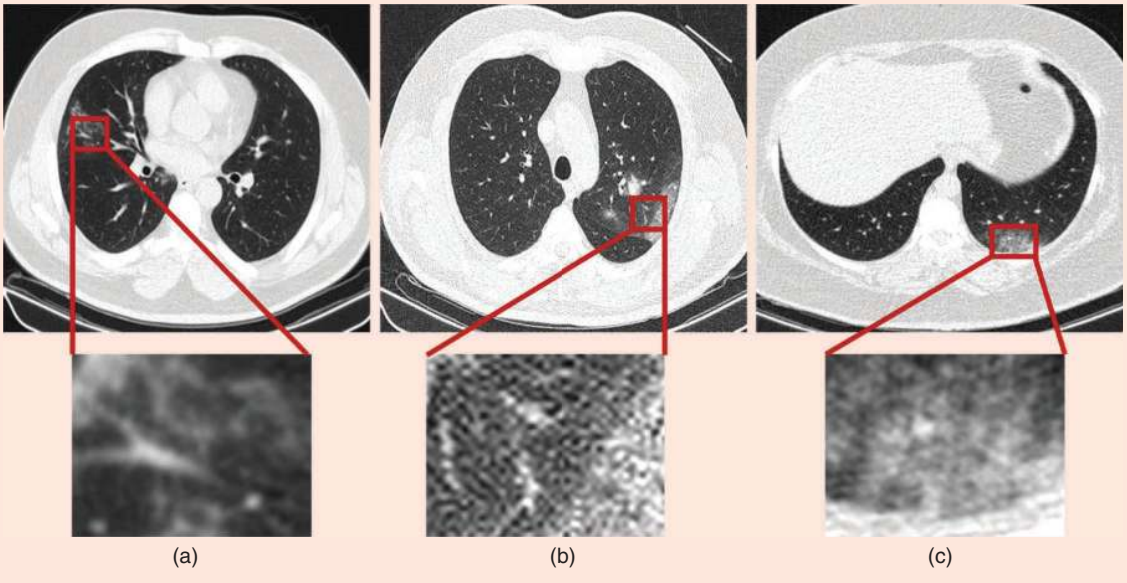

FIGURE 8. Acquired CT images with three different dose levels for three COVID-19 patients: (a) standard, (b) low, and (c) ultralow doses [82].

scales. Capitalizing on the superior performance achieved across various medical segmentation tasks, including multiorgan segmentation from CT scans and cardiac segmentation from MRI, one can conclude that transformers are solid encoders for COVID-19 segmentation tasks. Given the distributed nature of COVID-19 infection regions across the lung, learning long-range dependencies in chest images is of high importance in the context of COVID-19 lesion segmentation. Consequently, and on the basis of the promising results achieved by applying transformers for medical image segmentation, the development of transformer-based segmentation networks is of significant potential for detecting COVID-19 areas of infection at different scales.

- To encourage physicians and health professionals to confidently utilize DL models, it is important to provide explanations and interpretations on the internal behavior of the DL models and achieved results and, therefore, eliminate the "black-box perception." Regarding the black-box nature of DL models, communicating explainable outcomes to physicians is essential for the clinical adoption of implemented DL models. Several explainability techniques are leveraged in COVID-19 studies, the simplest of which is to verify the outcomes with a radiologist. This approach is, however, time-consuming and burdensome.

  Techniques providing heat maps of the most important regions of the input image are also popular within COVID-19 studies. One of the commonly used heat-map techniques is class activation mapping (CAM), utilized in [62] at different feature levels. Gradient-weighted CAM (Grad-CAM), visually depicting the deep model's decision, is also a CAM approach with the advantage of not requiring retraining. The Grad-CAM outcome shows how the developed model pays more attention to the regions of infection of the chest radiographs in [57] and [63]. A saliency map has also shown interpretable outcomes within the COVID-19 studies [62].

  Despite the advances in improving the explainability of the models, there are still examples for which the models fail to provide a clear explanation. Furthermore, heat maps do not provide enough explanation of the unique features they use to distinguish between COVID-19 and CAP cases.
- Due to the policy of protecting people's privacy and also immediate quarantine of mild cases without further examinations, scans with nonsevere symptoms are missing from most of the public data sets, and models are mostly developed based on patients with severe lung lesions who are at the late/advanced stages of the disease. The models, therefore, are biased toward severe cases and cannot be easily generalized.
- Evaluating the developed SP/DL models' performance in an unseen domain results in a decrease in the sensitivity of COVID-19 diagnosis. Most of the developed models, however, incorporate data coming from a single hospital, without a cross-center validation. In other words, the impact of equipment differences is not fully considered yet, and data from different sources are required to verify the generalizability of the models. It is worth noting that integrating data from multiple centers has specific challenges. Although institutions commonly follow general CT acquisition guidelines and protocols, there may still be inconsistencies in parameters such as slice thickness and the reconstruction algorithm. In this regard, we would like to suggest the following harmonization techniques:

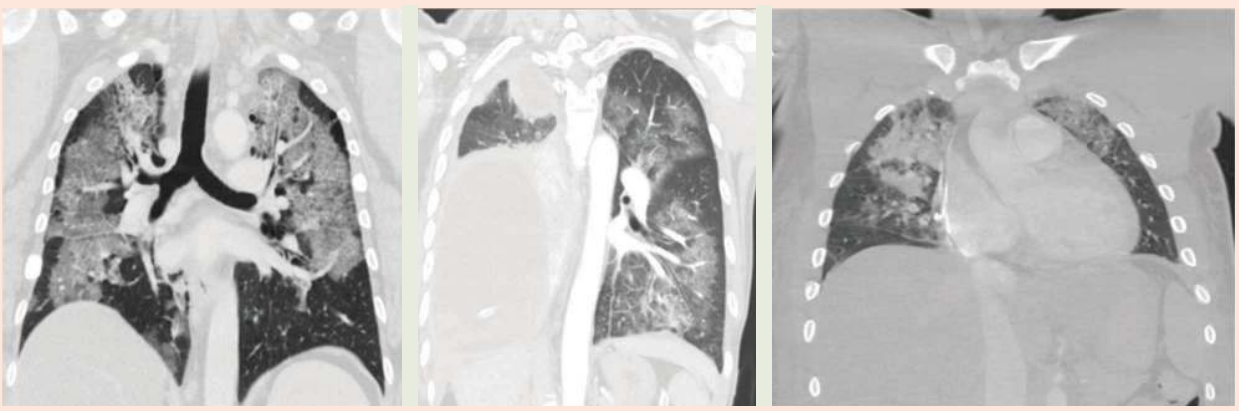

FIGURE 9. Acquired images for three patients suspected of having COVID-19. These patients have pre-existing conditions interfering with the diagnosis of COVID-19: (a) a 47-year-old male with fatty embolism and pulmonary edema, (b) a 51-year-old female with a history of right lung cancer and right lower lobectomy, and (c) a 27-year-old male with a gunshot injury in the left hemithorax with pulmonary contusion as well as left hemothorax and pneumothorax.

1) One important strategy to be considered is selecting radiomics features based on their stability, referring to the degree of dependency between the features and data source. Stable features are more reliable and more independent from the acquisition parameters. Test–retest is one of the most common techniques of assessing the stability of the radiomics.
2) It is of high importance to include data from several centers in both the training and test sets.
3) It is also suggested that different SP methods be followed to harmonize the image intensities through normalization, noise estimation, and denoising, e.g., through wavelets, anisotropic diffusion, and bilateral filtering.

- One limitation associated with many COVID-19 studies is that they try to distinguish COVID-19 cases from normal ones or categorize normal and non-COVID pneumonia cases as one class. Studies that consider a separate CAP class also report a relatively poor performance in distinguishing the COVID-19 and CAP classes. This calls for developing models with stronger backbone architectures and higher capacities. Furthermore, pneumonia incidence samples are older compared to the COVID-19 ones, and images from pneumonia patients with COVID-19 symptoms are not included in the data sets.
- Although hybrid models, which combine images and other relevant clinical information, can play an important role in COVID-19 analysis, few data sets are accompanied by demographic and clinical risk factors.
- One important challenge associated with COVID-19 analysis is the disease manifestation in patients with complications other than COVID-19. Several diseases can impact the lung tissue and interfere or change the appearance of COVID-19. Interstitial lung diseases as well as pleural or cardiac diseases may have imaging manifestations that can mask superadded COVID-19 and make it challenging for the interpreting radiologist. As shown in Figure 9, it is not clear if the abnormalities are related to COVID-19. This calls for developing more advanced SP solutions and unique features to facilitate COVID-19.

Conclusions

Medical imaging plays an important role in the diagnosis and management of COVID-19 infection. SP methods coupled with DL models can help to develop robust autonomous solutions for the diagnosis/prognosis of COVID-19 based on chest images. In this article, an integrated sketch is presented for designing and developing intelligent models for the COVID-19 infection diagnosis/prognosis. The latest developments on the theoretical framework of AE and HP for COVID-19 have been formally elaborated. Advanced SP methodologies and DL models for the diagnosis and prognosis of COVID-19 are presented, taking into consideration major challenges and opportunities.

This article provides the SP community with a comprehensive introduction to various solutions to COVID-19 radiomics. In addition, the article provides the required radiological background, available resources, and challenges/opportunities for extensive future SP research in this multidisciplinary domain to serve our diligent role in combating the COVID-19 pandemic and possible future similar ones.

SP methods coupled with DL models can help to develop robust autonomous solutions for the diagnosis/prognosis of COVID-19 based on chest images.

Acknowledgments

This project was partially supported by the Department of National Defence's Innovation for Defense Excellence and Security program, Canada. We would like to thank the consulting committee and the editor-in-chief of *IEEE Signal Processing Magazine* for their two-round reviews and encouraging comments.

Authors

Arash Mohammadi (arash.mohammadi@concordia.ca) is with the Concordia Institute for Information System Engineering, Concordia University, Montréal, Quebec, H3G-1M8, Canada. He is a Senior Member of IEEE.

Yingxu Wang (yingxu@ucalgary.ca) is with the Department of Electrical and Computer Engineering, University of Calgary, Calgary, Alberta, T2N 1N4, Canada. He is a Fellow of IEEE.

Nastaran Enshaei (n_enshae@encs.concordia.ca) is with the Concordia Institute for Information Systems Engineering, Concordia University, Montréal, Quebec, H3G-1M8, Canada. She is a Graduate Student Member of IEEE.

Parnian Afshar (p_afs@encs.concordia.ca) is with the Concordia Institute for Information Systems Engineering, Concordia University, Montréal, Quebec, H3G-1M8, Canada. She is a Graduate Student Member of IEEE.

Farnoosh Naderkhani (farnoosh.naderkhani@concordia.ca) is with the Concordia Institute for Information Systems Engineering, Concordia University, Montréal, Quebec, H3G-1M8, Canada. She is a Member of IEEE.

Anastasia Oikonomou (anastasia.oikonomou@sunnybrook.ca) is with the Department of Medical Imaging, Sunnybrook Health Sciences Centre, University of Toronto, Toronto, M4N 3M5, Canada.

Moezedin Javad Rafiee (moezedinjavad.rafiee@muhc.mcgill.ca) is with the Department of Medicine and Diagnostic Radiology, McGill University Health Center–Research Institute, Montréal, Quebec, H3H 2E9, Canada.

Helder C.R. Oliveira (helder@schulich.ucalgary.ca) is with the Department of Electrical and Computer Engineering, University of Calgary, Calgary, Alberta, T2N 1N4, Canada. He is a Member of IEEE.

Svetlana Yanushkevich (syanshk@ucalgary.ca) is with the Department of Electrical and Computer Engineering,

University of Calgary, Calgary, AB, T2N 1N4, Canada. She is a Senior Member of IEEE.

Konstantinos N. Plataniotis (kostas@ece.utoronto.ca) is with the Department of Electrical and Computer Engineering, University of Toronto, Toronto, M5S 3G4, Canada. He is a Fellow of IEEE.

References

[1] L. M. Kucirka, S. A. Lauer, O. Laeyendecker, D. Boon, and J. Lessler, "Variation in false-negative rate of reverse transcriptase polymerase chain reaction-based SARS-CoV-2 tests by time since exposure," *Ann. Intern. Med.*, vol. 173, no. 4, pp. 262–267, 2020.

[2] P. Afshar, A. Mohammadi, K. N. Plataniotis, A. Oikonomou, and H. Benali, "From handcrafted to deep-learning-based cancer radiomics: Challenges and opportunities," *IEEE Signal Process. Mag.*, vol. 36, no. 4, pp. 132–160, July 2019. doi: 10.1109/MSP.2019.2900993.

[3] F. Shi, J. Wang, J. Shi, Z. Wu, Q. Wang, Z. Tang, K. He, Y. Shi et al., "Review of artificial intelligence techniques in imaging data acquisition, segmentation and diagnosis for COVID-19," *IEEE Rev. Biomed. Eng.*, vol. 14, pp. 4–15, 2021. doi: 10.1109/RBME.2020.2987975.

[4] D. Dong, Z. Tang, S. Wang, H. Hui, L. Gong, Y. Lu, Z. Xue, H. Liao, et al., "The role of imaging in the detection and management of COVID-19: A review," *IEEE Rev. Biomed. Eng.*, vol. 14, pp. 16–29, Jan. 2020.

[5] M. Jamshidi, A. Lalbakhsh, J. Talla, Z. Peroutka, F. Hadjilooei, P. Lalbakhsh, M. Jamshidi, L. La Spada, et al., "Artificial intelligence and COVID-19: Deep learning approaches for diagnosis and treatment," *IEEE Access*, vol. 8, pp. 109,581–109,595, June 2020. doi: 10.1109/ACCESS.2020.3001973.

[6] Y. Wang, K. N. Plataniotis, J. Z. Wang, M. Hou, M. Zhou, N. Howard, J. Peng, R. Huang et al., "The cognitive and mathematical foundations of analytic epidemiology," in *Proc. IEEE 19th Int. Conf. Cogn. Informatics Cogn. Comput. (ICCI*CC'20)*, Tsinghua Univ., Beijing, Sept. 2020.

[7] Y. Wang, F. Karray, S. Kwong, K. N. Plataniotis, H. Leung, M. Hou, E. Tunstel, I. J. Rudas, Ljiljana Trajkovic et al., "Perspectives on the philosophical, cognitive and mathematical foundations of symbiotic autonomous systems (SAS)," in *Philosophical Transactions of Royal Society (A).* Oxford, U.K., 379(x), pp. 1–26, 2021. doi: 10.1098/rsta.2020.0362.

[8] J. Ramey and P. G. Rao, "The systematic literature review as a research genre," in *Proc. IEEE Int. Professional Commun. Conf.*, 2011, pp. 1–7.

[9] Y. Wang, "On visual semantic algebra (VSA): A denotational mathematical structure for modeling and manipulating visual objects and patterns," *Int. J. Softw. Sci. Comput. Intell.*, vol. 1, no. 4, pp. 1–16, 2009. doi: 10.4018/jssci.2009062501.

[10] G. D. Rubin, C. J. Ryerson, L. B. Haramati, N. Sverzellati, J.P. Kanne, S. Raoof, N. W. Schluger, A. Volpi et al., "The role of chest imaging in patient management during the COVID-19 pandemic: A multinational consensus statement from the Fleischner society," *Chest*, vol. 158, no. 1, pp. 106–116, 2020. doi: 10.1016/j.chest.2020.04.003.

[11] S .L. Haak, I. J. Renken, L. C. Jager, H. Lameijer, and B. B. Y. van der Kolk, "Diagnostic accuracy of point-of-care lung ultrasound in COVID-19," *Emerg. Med. J.*, vol. 38, no. 2, pp. 94–99, 2020. doi: 10.1136/emermed-2020-210125.

[12] Z. Sun, N. Zhang, Y. Li, and X. Xu, "A systematic review of chest imaging findings in COVID-19," *Quantit. Imag. Med. Surg.*, vol. 10, no. 5, p. 1058, 2020. doi: 10.21037/qims-20-564.

[13] P. Afshar, S. Heidarian, N. Enshaei, F. Naderkhani, M. J. Rafiee, A. Oikonomou, F. B. Fard, K. Samimi et al., "COVID-CT-MD: COVID-19 computed tomography (CT) scan dataset applicable in machine learning and deep learning," *Nature Sci. Data*, vol. 8, no. 1, pp. 1–8, 2021. doi: 10.1038/s41597-021-00900-3.

[14] A. Bernheim, X. Mei, M. Huang, Y. Yang, Z. A. Fayad, N. Zhang, K. Diao, B. Lin et al., "Chest CT findings in coronavirus disease-19 (COVID-19): Relationship to duration of infection," *Radiology*, vol. 295, no. 3, p. 200463, 2020. doi: 10.1148/radiol.2020200463.

[15] H. X. Bai, B. Hsieh, Z. Xiong, K. Halsey, J. W. Choi, T. M. L. Tran, I. Pan, L. B. Shi et al. "Performance of radiologists in differentiating COVID-19 from viral pneumonia on chest CT," *Radiology*, vol. 296, no. 2, p. 200823, 2020. doi: 10.1148/radiol.2020200823.

[16] L. Huang, R. Han, T. Ai, P. Yu, H. Kang, Q. Tao, and L. Xia, "Serial quantitative chest CT assessment of COVID-19: Deep-learning approach," *Radiol.: Cardiothorac. Imag.*, vol. 2, no. 2, p. e200075, 2020.

[17] S. Chaganti, P. Grenier, A. Balachandran, G. Chabin, S. Cohen, T. Flohr, B. Georgescu, S. Grbic et al., "Automated quantification of CT patterns associated with COVID-19 from chest CT," *Radiol.: Artif. Intell.*, vol. 2, no. 4, p. e200048, 2020. doi: 10.1148/ryai.2020200048.

[18] M. Francone, F. Iafrate, G. M. Masci, S. Coco, F. Cilia, L. Manganaro, V. Panebianco, C. Andreoli et al., "Chest CT score in COVID-19 patients: Correlation with disease severity and short-term prognosis," *European Radiol.*, vol. 30, no. 12, pp. 6808–6817, 2020. doi: 10.1007/s00330-020-07033-y.

[19] H. Y. F. Wong, H. Y. S. Lam, A. H. T. Fong, S. T. Leung, T. W. Y. Chin, C. S. Y. Lo, M. M. S. Lui, J. C. Y. Lee, K. W. H. Chiu, T. W. H. Chung, and E. Y. P. Lee et al., "Frequency and distribution of chest radiographic findings in COVID-19 positive patients," *Radiology*, vol. 296, no. 2, pp. E72–E78, 2020.

[20] D. Toussie, N. Voutsinas, M. Finkelstein, M. A. Cedillo, S. Manna, S. Z. Maron, A. Jacobi, M. Chung et al., "Clinical and chest radiography features determine patient outcomes in young and middle age adults with COVID-19," *Radiology*, vol. 297, no. 1, pp. E197–206, 2020. doi: 10.1148/radiol.2020201754.

[21] D. Cozzi, M. Albanesi, E. Cavigli, C. Moroni, A. Bindi, S. Luvarà, S. Lucarini, S. Busoni, L. N. Mazzoni, and V. Miele, "Chest X-ray in new coronavirus disease 2019 (COVID-19) infection: Findings and correlation with clinical outcome," *La Radiologia Medica*, vol. 125, no. 8, pp. 730–737, 2020. doi: 10.1007/s11547-020-01232-9.

[22] L. Wang, Z. Q. Lin, and A. Wong, "COVID-NET: A tailored deep convolutional neural network design for detection of COVID-19 cases from chest X-ray images," *Sci. Rep.*, vol. 10, no. 1, p. 19549, 2020. doi: 10.1038/s41598-020-76550-z.

[23] E. B. Tsai, S. Simpson, M. P. Lungren, M. Hershman, L. Roshkovan, E. Colak, B. J. Erickson, G. Shih, A. Stein, J. Kalpathy-Cramer, and J. Shen, "The RSNA international COVID-19 open radiology database (RICORD)," *Radiology*, vol. 299, no. 1, pp. E204–E213, 2021. doi: 10.1148/radiol.2021203957.

[24] L. Vetrugno, T. Bove, D. Orso, F. Barbariol, F. Bassi, E. Boero, G. Ferrari, and R. Kong, "Our Italian experience using lung ultrasound for identification, grading and serial follow-up of severity of lung involvement for management of patients with COVID-19," *Echocardiography*, vol. 37, no. 4, pp. 625–627, 2020. doi: 10.1111/echo.14664.

[25] "Guidance on COVID-19 and MR use," ACR. Accessed: May 31, 2020. [Online]. Available: https://www.acr.org/Clinical-Resources/Radiology-Safety/MR-Safety/COVID-19-and-MR-Use

[26] K. T. Rajamani, H. Siebert, and M. Heinrich, "Dynamic deformable attention (DDANET) for semantic segmentation," medRxiv, 2020.

[27] D. P. Fan, T. Zhou, G. P. Ji, Y. Zhou, G. Chen, H. Fu, J. Shen, and L. Shao, "INf-NET: Automatic COVID-19 lung infection segmentation from CT images," *IEEE Trans. Med. Imag.*, vol. 39, no. 8, pp. 2626–2637, 2020.

[28] Y. Qiu, Y. Liu, and J. Xu, "Miniseg: An extremely minimum network for efficient COVID-19 segmentation," 2020, arXiv:2004.09750.

[29] A. Degerli, M. Ahishali, M. Yamac, S. Kiranyaz, M. E. Chowdhury, K. Hameed, T. Hamid, R. Mazhar et al., "COVID-19 infection map generation and detection from chest X-ray images," *Health Inf. Sci. Syst.*, vol. 9, no. 1, pp. 1–16, 2021.

[30] K. Zhang, X. Liu, J. Shen, Z. Li, Y. Sang, X. Wu, Y. Zha, W. Liang et al., "Clinically applicable AI system for accurate diagnosis, quantitative measurements, and prognosis of COVID-19 pneumonia using computed tomography," *Cell*, vol. 181, no. 6, pp. 1423–1433, 2020. doi: 10.1016/j.cell.2020.08.029.

[31] B. Zheng, Y. Liu, Y. Zhu, F. Yu, T. Jiang, D. Yang, and T. Xu, "MSD-Net: Multi-scale discriminative network for COVID-19 lung infection segmentation on CT," *IEEE Access*, vol. 8, pp. 185,786–185,795, 2020. doi: 10.1109/ACCESS.2020.3027738.

[32] S. Zhao, Z. Li, Y. Chen, W. Zhao, X. Xie, J. Liu, D. Zhao, and Y. Li, "SCOAT-Net: A novel network for segmenting COVID-19 lung opacification from CT images," *Pattern Recognit.*, p. 108109, 2021.

[33] F. Shan, Y. Gao, J. Wang, W. Shi, N. Shi, M. Han, Z. Xue, D. Shen, and Y. Shi, "Lung infection quantification of COVID-19 in CT images with deep learning," 2020, arXiv:2003.04655.

[34] Z. Li, Z. Zhong, Y. Li, T. Zhang, L. Gao, D. Jin, Y. Sun, X. Ye et al., "From community-acquired pneumonia to COVID-19: A deep learning-based method for quantitative analysis of COVID-19 on thick-section CT scans," *European Radiol.*, vol. 30, no. 12, pp. 6828–6837, 2020. doi: 10.1007/s00330-020-07042-x.

[35] G. Wang, X. Liu, C. Li, Z. Xu, J. Ruan, H. Zhu, T. Meng, K. Li et al., "A noise-robust framework for automatic segmentation of COVID-19 pneumonia lesions from CT images," *IEEE Trans. Med. Imag.*, vol. 39, no. 8, pp. 2653–2663, 2020. doi: 10.1109/TMI.2020.3000314.

[36] Q. Yao, L. Xiao, P. Liu, and S.K. Zhou, "Label-free segmentation of COVID-19 lesions in Lung CT," 2020, arXiv:2009.06456.

[37] F. Shi, L. Xia, F. Shan, B. Song, D. Wu, Y. Wei, H. Yuan, H. Jiang et al., "Large-scale screening of COVID-19 from community acquired pneumonia using infection size-aware classification," vol. 66, no. 6, p. 065031, 2020.

[38] A. Saygili, "A new approach for computer-aided detection of coronavirus (COVID-19) from CT and X-ray images using machine learning methods," *Appl. Soft Comput.*, vol. 23, p. 107323, 2021.

[39] H. Chao, X. Fang, J. Zhang, F. Homayounieh, C. D. Arru, S. R. Digumarthy, R. Babaei, H. K. Mobin et al., "Integrative analysis for COVID-19 patient outcome prediction," *Med. Image Anal.*, vol. 67, p. 101,844, 2020. doi: 10.1016/j.media.2020.101844.

[40] W. Ning, S. Lei, J. Yang, Y. Cao, P. Jiang, Q. Yang, J. Zhang, X. Wang et al., "Open resource of clinical data from patients with pneumonia for the prediction of

COVID-19 outcomes via deep learning," *Nature Biomed. Eng.*, vol. 4, no. 12, pp. 1197–1207, 2020. doi: 10.1038/s41551-020-00633-5.

[41] N. Lassau, S. Ammari, E. Chouzenoux, H. Gortais, P. Herent, M. Devilder, S. Soliman, O. Meyrignac et al., "AI-based multi-modal integration (ScanCov scores) of clinical characteristics, lab tests and chest CTs improves COVID-19 outcome prediction of hospitalized patients," Inria Saclay Ile de France. Res. Rep. hal-02586111v3, 2020.

[42] D. Colombi, F. C. Bodini, M. Petrini, , G. Maffi, N. Morelli, G. Milanese, M. Silva, N. Sverzellati et al., "Well-aerated lung on admitting chest CT to predict adverse outcome in COVID-19 pneumonia," *Radiology*, vol. 296, no. 2, pp. E86–E96, 2020. doi: 10.1148/radiol.2020201433.

[43] V. Mergen, A. Kobe, C. Blüthgen, A. Euler, T. Flohr, T. Frauenfelder, H. Alkadhi, and M. Eberhard "Deep learning for automatic quantification of lung abnormalities in COVID-19 patients: First experience and correlation with clinical parameters," *Eur. J. Radiol. Open*, vol. 7, p. 100,272, Jan. 2020. doi: 10.1016/j.ejro.2020.100272.

[44] H. Yue, Q. Yu, C. Liu, Y. Huang, Z. Jiang, C. Shao, H. Zhang, B. Ma et al., "Machine learning-based CT radiomics method for predicting hospital stay in patients with with SARS-COV-2 infection: Multicenter study," *Ann. Transl. Med.*, vol. 8, no. 14, 2020. doi: 10.21037/atm-20-3026.

[45] F. Homayounieh, S. Ebrahimian, R. Babaei, H. K. Mobin, E. Zhang, B. C. Bizzo, I. Mohseni, S. R. Digumarthy et al., "CT radiomics, radiologists and clinical information in predicting outcome of patients with COVID-19 pneumonia," *Radiology: Cardiothoracic Imaging*, vol. 2, no. 4, p. e200322, 2020. doi: 10.1148/ryct.2020200322.

[46] G. Chassagnon, M. Vakalopoulou, E. Battistella, S. Christodoulidis, T. N. Hoang-Thi, S. Dangeard, E. Deutsch, F. Andre et al., "AI-driven quantification, staging and outcome prediction of COVID-19 pneumonia," *Med. Image Anal.*, vol. 67, p. 101,860, 2020. doi: 10.1016/j.media.2020.101860.

[47] X. Bai, C. Fang, Y. Zhou, S. Bai, Z. Liu, L. Xia, Q. Chen, Y. Xu et al., "Predicting COVID-19 malignant progression with AI techniques," 2020, medRxiv 2020.03.20.20037325.

[48] A. Amyar, R. Modzelewski, H. Li, and S. Ruan, "Multi-task deep learning based CT imaging analysis for COVID-19 pneumonia: Classification and segmentation," *Comput. Biol. Med.*, vol. 126, p. 104,037, 2020. doi: 10.1016/j.compbiomed.2020.104037.

[49] Z. Tang, W. Zhao, X. Xie, Z. Zhong, F. Shi, T. Ma, J. Liu, and D. Shen, "Severity assessment of COOVID-19 using CT image features and laboratory indices," *Phys. Med. Biol.*, 2021.

[50] B. Ghosh, N. Kumar, N. Singh, A. K. Sadhu, N. Ghosh, P. Mitra, and J. Chatterjee, "A quantitative lung computed tomography image feature for multi-center severity assessment of COVID-19," medRxiv, 2020.

[51] X. Mei, H. C. Lee, K. Y. Diao, M. Huang, B. Lin, C. Liu, Z. Xie, Y. Ma et al., "Artificial intelligence-enabled rapid diagnosis of patients with COVID-19," *Nat. Med.*, vol. 26, no. 8, pp. 1224–1228, 2020. doi: 10.1038/s41591-020-0931-3.

[52] L. Li, L. Qin, Z. Xu, Y. Yin, X. Wang, B. Kong, J. Bai, Y. Lu et al., "Using artificial intelligence to detect COVID-19 and community-acquired pneumonia based on pulmonary CT: Evaluation of the diagnostic accuracy," *Radiology*, vol. 296, no. 2, pp. E65–E71, 2020. doi: 10.1148/radiol.2020200905.

[53] P. Afshar, S. Heidarian, F. Naderkhani, A. Oikonomou, K. N. Plataniotis, and A. Mohammadi, "COVID-CAPS: A capsule network-based framework for identification of COVID-19 cases from X-ray images," *Pattern Recog. Lett.*, vol. 138, pp. 638–643, Oct. 2020. doi: 10.1016/j.patrec.2020.09.010.

[54] S. Wang et al., "A deep learning algorithm using CT images to screen for corona virus disease (COVID-19)," medRxiv, 2020.

[55] M. Farooq and A. Hafeez, "COVID-Resnet: A deep learning framework for screening of COVID-19 from radiographs," 2020, arXiv:2003.14395.

[56] A. Narin, C. Kaya, and Z. Pamuk, "Automatic detection of coronavirus disease (COVID-19) using X-ray images and deep convolutional neural networks," 2020, arXiv:2003.10849.

[57] T. Ozturk, M. Talo, E. A. Yildirim, U. B. Baloglu, O. Yildirim, and U. R. Acharya, "Automated detection of COVID-19 cases using deep neural networks with X-ray images," *Comput. Biol. Med.*, vol. 121, p. 103,792, June 2020. doi: 10.1016/j.compbiomed.2020.103792.

[58] M. Karim et al., "Deep COVID explainer: Explainable COVID-19 predictions based on chest X-ray images," 2020, arXiv:2004.04582.

[59] E. E. Hemdan, M. A. Shouman, and M. E. Karar, "COVIDx-Net: A framework of deep learning classifiers to diagnose COVID-19 in X-ray images," 2020, arXiv:2003.11055.

[60] S. Ying et al., "Deep learning enables accurate diagnosis of novel coronavirus (COVID-19) with CT images," medRxiv, 2020.

[61] J. Zhang et al., "Viral pneumonia screening on chest X-ray images using confidence-aware anomaly detection," 2020, arXiv:2003.12338.

[62] S. Hu, Y. Gao, Z. Niu, Y. Jiang, L. Li, X. Xiao, M. Wang, E. F. Fang, W. Menpes-Smith, J. Xia, and H. Ye, "Weakly supervised deep learning for COVID-19 infection detection and classification from CT images," *IEEE Access*, vol. 8, pp. 118,869–118,883, June 2020. doi: 10.1109/ACCESS.2020.3005510.

[63] M. Z. Islam, M. M. Islam, and A. Asraf, "A combined deep CNN-LSTM network for the detection of novel coronavirus (COVID-19) using X-ray images," *Inf. Med. Unlocked*, vol. 20, p. 100,412, Jan. 2020. doi: 10.1016/j.imu.2020.100412.

[64] X. Wang, X. Deng, Q. Fu, Q. Zhou, J. Feng, H. Ma, W. Liu, and C. Zheng, "A weakly-supervised framework for COVID-19 classification and lesion localization from chest CT," *IEEE Trans. Med. Imag.*, vol. 39, no. 8, pp. 2615–2625, 2020. doi: 10.1109/TMI.2020.2995965.

[65] X. Xu, X. Jiang, C. Ma, P. Du, X. Li, S. Lv, L. Yu, Q. Ni et al., "A deep learning system to screen novel coronavirus disease 2019 pneumonia," *Engineering*, vol. 6, no. 10, pp. 1122–1129, 2020. doi: 10.1016/j.eng.2020.04.010.

[66] A. Mohammed, C. Wang, M. Zhao, M. Ullah, R. Naseem, H. Wang, M. Pedersen, and F. A. Cheikh, "Weakly-supervised network for detection of COVID-19 in chest CT Scans," *IEEE Access*, vol. 8, pp. 155,987–156,000, Aug. 2020. doi: 10.1109/ACCESS.2020.3018498.

[67] S. Yang, L. Jiang, Z. Cao, L. Wang, J. Cao, R. Feng, Z. Zhang, X. Xue, Y. Shi, and F. Shan, "Deep learning for detecting corona virus disease 2019 (COVID-19) on high-resolution computed tomography: A pilot study," *Ann. Transl. Med.*, vol. 8, no. 7, p. 450, 2020. doi: 10.21037/atm.2020.03.132.

[68] L. Meng, D. Dong, L. Li, M. Niu, Y. Bai, M. Wang, X. Qiu, Y. Zha et al., "A deep learning prognosis model help alert for COVID-19 patients at high-risk of death: A multi-center study," *IEEE J. Biomed. Health Inform.*, vol. 24, no. 12, pp. 3576–3584, 2020.

[69] O. Gozes, M. Frid-Adar, N. Sagie, H. Zhang, W. Ji, and H. Greenspan, "Coronavirus detection and analysis on chest CT with deep learning," 2020, arXiv:2004.02640.

[70] N. N. Das, N. Kumar, M. Kaur, V. Kumar, and D. Singh, "Automated deep transfer learning-based approach for detection of COVID-19 infection in chest X-rays," *IRBM*, July 2020.

[71] Y. Li, D. Wei, J. Chen, S. Cao, H. Zhou, Y. Zhu, J. Wu, L. Lan et al., "Efficient and effective training of COVID-19 classification networks with self-supervised dual-track learning to rank," *IEEE J. Biomed. Health Inf.*, vol. 24, no. 10, pp. 2787–2797, 2020. doi: 10.1109/JBHI.2020.3018181.

[72] A. Waheed, M. Goyal, D. Gupta, A. Khanna, F. Al-Turjman, and P. R. Pinheiro, "COVIDGAN: Data augmentation using auxiliary classifier GAN for improved COVID-19 detection," *IEEE Access*, vol. 8, pp. 91,916–91,923, May 2020. doi: 10.1109/ACCESS.2020.2994762.

[73] P. Lyu, X. Liu, R. Zhang, L. Shi, and J. Gao, "The performance of chest CT in evaluating the clinical severity of COVID-19 pneumonia: Identifying critical cases based on CT characteristics," *Investigative Radiol.*, vol. 55, no. 7, pp. 412–421, 2020. doi: 10.1097/RLI.0000000000000689.

[74] Y. Feng, S. Liu, Z. Cheng, J. Quiroz, P. Chen, Q. Lin, L. Qian, X. Liu et al., "Severity assessment and progression prediction of COVID-19 patients based on the lesion encoder framework and chest CT," medRxiv, 2020.

[75] J. Zhu, B. Shen, A. Abbasi, M. Hoshmand-Kochi, H. Li, and T. Q. Duong, "Deep transfer learning artificial intelligence accurately stages COVID-19 lung disease severity on portable chest radiographs," *PLOS One*, vol. 15, no. 7, pp. 1–11, 2020. doi: 10.1371/journal.pone.0236621.

[76] S. S. Yip, Z. Klanecek, S. Naganawa, J. Kim, A. Studen, L. Rivetti, and R. Jeraj, "Performance and robustness of machine learning-based radiomic COVID-19 severity prediction," medRxiv, 2020.

[77] W. Cai, T. Liu, X. Xue, G. Luo, X. Wang, Y. Shen, Q. Fang, J. Sheng et al., "CT quantification and machine-learning models for assessment of disease severity and prognosis of COVID-19 patients," *Acad. Radiol.*, vol. 27, no. 12, pp. 1665–1678, 2020. doi: 10.1016/j.acra.2020.09.004.

[78] M. D. Li, N. T. Arun, M. Gidwani, K. Chang, F. Deng, B. P. Little, D. P. Mendoza, M. Lang et al., "Automated assessment and tracking of COVID-19 pulmonary disease severity on chest radiographs using convolutional Siamese neural networks," *Radiol.: Artif. Intell.*, vol. 2, no. 4, p. e200079, 2020. doi: 10.1148/ryai.2020200079.

[79] J. Chen, Y. Lu, Q. Yu, X. Luo, E. Adeli, Y. Wang, L. Lu, A. L. Yuille et al., "Transunet: Transformers make strong encoders for medical image segmentation," 2021, arXiv:2102.04306.

[80] A. Hatamizadeh, D. Yang, H. Roth, and D. Xu, "UNETR: Transformers for 3D medical image segmentation," 2021, arXiv:2103.10504.

[81] P. Afshar, A. Mohammadi, P. N. Tyrrell, P. Cheung, A. Sigiuk, K. N. Plataniotis, E. T. Nguyen, and A. Oikonomou, "DRTOP: Deep learning-based radiomics for the time-to-event outcome prediction in lung cancer, *Sci. Rep.*, vol. 10, no. 1, pp. 1–15.

[82] P. Afshar et al., "COVID-19 low-dose and Ultra-low-dose CT scans," *IEEE Dataport*. [Online]. Available: https://ieee-dataport.org/open-access/covid-19-low-dose-and-ultra-low-dose-ct-scans. doi: 10.21227/sed8-6r15.

SP