1. INTRODUCTION

COVID-19 (Coronavirus disease 2019) is a contagious virus that has been spreading all over the world since December 2019 in Wuhan (China). It is caused by a virus called by the International Committee on Taxonomy of Viruses (ICT) SARS-CoV-2 [1]. There are various methods to identify patients infected with COVID-19, the most effective test is Reverse Transcription Polymerase Chain Reaction (RT-PCR). In this method, a miniature target sequence of nucleic acids which consists of a fragment of DNA is copied many times to make it easier to detect. As a consequence, the test is positive when this amplification is detected (using a fluorophore-labeled probe). While RT-PCR tests have been widely used to detect COVID-19 and have played a crucial role in the pandemic response, they have certain limitations, particularly in terms of accuracy and scalability [2].

The accuracy of RT-PCR tests can vary depending on when the test is administered during the infection. They are generally more sensitive in the early stages of infection when viral loads are higher but may produce false negatives in later stages or if the sample quality is suboptimal. RT-PCR tests can sometimes yield false positives, especially if there is contamination during sample collection or processing. So RT-PCR test is not completely accurate and needs a second test for the diagnosis confirmation. Thus, RT-PCR test is time-consuming. It takes more than a day to get results. It is also costly. RT-PCR test is resource-intensive and can be relatively costly to administer, especially when large-scale testing is required. RT-PCR tests are not efficiently scalable because it requires specialized laboratory equipment, managed by professional technicians. Scaling up testing capacity can be challenging, especially during surges in cases, leading to delays in testing and reporting which makes it an expensive test and difficult to scale.

Other methods that are useful for detecting COVID-19 include clinical examination, pathological tests, and radiography. Due to their easy accessibility, many clinicians prefer to use chest imaging for the diagnosis of COVID-19. The role of chest imaging in assessing complications, disease progression, and prognosis of COVID-19 has been discussed by Inui et al. [3]. Also, the cochrane COVID-19 Diagnostic Test Accuracy Group [4] considered chest imaging as a good diagnostic tool for detecting and managing COVID-19 pneumonia. They have conducted different studies to evaluate the diagnostic accuracy of chest imaging in people of any age with suspected COVID-19. Their studies assessed chest CT, chest X-Ray, or ultrasound of the lungs for the diagnosis of COVID-19, using a reference standard that included RT-PCR.

Among medical imaging modalities, chest X-ray was...
generally preferred for different reasons. Chest X-Ray images are widely available and can be easily taken in most healthcare settings. They are less expensive compared to other imaging techniques like computed tomography (CT) scans. As COVID-19 is an emergency, X-Ray images are very beneficial because they provide quick results compared to RT-PCR tests, and they do not expose the patient to as much radiation as other imaging techniques like CT scans. X-Ray images can be useful in monitoring the progress of the disease and assessing the response to treatment. Cozzi et al. [5] investigated the principal radiological features of COVID-19 by describing the most important chest X-Ray findings in a selected cohort of patients. They also correlated the radiological appearance with the RT-PCR test and the outcome of the patients.

Accelerate diagnostic means using automatic tools that provide rapid results. The AI techniques have proven their effectiveness in different stages: the acquisition of medical images, then segmentation, and finally in the diagnosis of Covid-19. To illustrate the latest advances in medical imaging and radiology in the fight against COVID-19, Shi et al. [6] have illustrated the integration of AI with X-Rays, which are widely used in frontline hospitals.

Huang et al. [7] have provided a discussion of the challenges and perspectives of machine learning (ML) and deep learning (DL) in the detection of COVID-19. As a result, the researchers found that an AI model can be as accurate as experienced physicians in the diagnosis of COVID-19.

In general, Deep learning has shown massive potential for healthcare applications. Specifically, for medical data analysis and diagnosis through medical image processing [8]. Due to their ability to learn complex features and patterns from the images, deep learning models, especially convolutional neural networks (CNNs) have shown promising results in detecting COVID-19 from X-Ray images and have the potential to assist professionals in diagnosing the disease accurately [9].

However, some challenges need to be addressed, such as the lack of publicly available data for training and evaluating COVID-19 detection models which limit the ability to build accurate models. Another issue is the variation in X-Ray image quality, and not all images will be clear or well-defined, leading to issues with classification. Also, the generalization challenge, using X-ray images to detect COVID-19 may limit the generalization capability of the model when it comes to identifying the virus in other parts of the body.

Consequently, the main intention behind this survey is to provide an overview of CNNs that have contributed to tackling these challenges. We have organized the reviewed CNN Models into three main categories. The first one focuses on transfer learning from pre-trained CNN models. In the second, the model is developed from scratch and called the Custom CNN Model. The last category couples the two previous ones to constitute the Hybrid CNN Model. A summary for each study is presented by giving the dataset size, the number of classes, the model architecture, and the performance evaluation criteria (Accuracy, Sensitivity, and Specificity).

The remainder of this paper is organized as follows. Section 2 presents the description of related work, where we classify CNN models through three categories: 1- Transfer Learning from pre-trained models; 2- Custom model, and 3- Hybrid model. Section 3 is devoted to some preliminary notions dealing with image classification, Datasets, CNN model architecture. Finally, the paper is concluded in Section 4 with a discussion of the related works.

2. RELATED WORK

Deep learning has developed good solutions with high accuracy for medical image segmentation and classification in the health sector [10-13]. Many deep learning techniques based on medical images chest X-Ray have been developed to detect COVID-19 [14]. Jain et al. [15] achieved good results with their model in detecting COVID-19. They used a dataset containing 1215 images, including 250 COVID-19 ones. They got 98.93% of accuracy, 98.66% of specificity, 96.39% of precision, and 98.15% F-1 score.

Ucar and Korkmaz [16] obtained 98.26% accuracy employing a dataset of 5949 chest X-Ray images containing 76 COVID-19 images. Wang et al. [17] suggested a deep learning model drawing out visual characteristics from CT Scan images categorizing Covid19. They exploited a Dataset of 1065 computed tomography scans with 740 images of patients suffering from viral pneumonia, and 325 images are of COVID-19 patients. The model attains a good accuracy of 79.3%. The model is used to distinguish between COVID-19 and other typical viral pneumonia.

CNN models perform feature extraction of medical images (chest X-Rays) of patients suspected to be infected with the virus and classification. CNNs extract relevant features using convolutional layers that apply filters on the images, reduce the dimensionality of feature maps with pooling layers, and use fully connected layers to perform classification by generating class probabilities. The model's weights are optimized during training to minimize the loss and regularized to prevent overfitting. Transfer learning is another popular technique used in CNNs to improve feature extraction. It involves using pre-trained models that have already learned relevant features from large datasets and fine-tuning them on a specific task. This helps to speed up the training process and improve the accuracy of the model.

Different approaches for X-Ray image classification based on CNN are proposed. Some are focusing on transfer learning from pre-trained CNN models. Others are developed from scratch (custom CNN models). Other works couple these two approaches and constitute an approach called the « hybrid CNN Model ». In this section, we present an overview of related work, where we organize them through three categories: 1- Transfer Learning from pre-trained models; 2- Custom model, and 3- Hybrid model.

2.1 Transfer learning from pre-trained CNN models

A pre-trained model is a model that is trained on large datasets to extract high-level features from images such as VGG, Inception, Resnet, etc. Using a pre-trained model is justified for several motives. Firstly, expensive computation power is required if the training models are large and the datasets are sizeable. The second reason is that the process of training in large models can be slow. It can take several weeks. Finally, a pre-trained model can gain time in order to generalize the network and accelerate the convergence.

Adapting a pre-trained model and applying it to a new dataset or task by fine-tuning is the principle of the Transfer Learning technique.

In this section, we give an overview of related studies that focus on transfer learning (TL) from pre-trained CNN models and an X-Ray dataset.
Traditional, parallel convolutional layers, and residual connections are three ideas integrated in the developed model DCNN (deep convolutional neural network) [18]. The proposed model accomplished an 86.6% of F1-score when trained from scratch. It got 89.4% using the transfer learning from different domains of the targeted dataset, and it obtained 97.6% with transfer learning from the same domain of the targeted dataset. However, there is a crucial issue according to the source data type utilized by the TL compared to the target dataset. This in terms of data features, sizes, and characteristics. Also, it has been demonstrated that the transfer learning from different domains has a little effect on the performance, as lightweight models trained from scratch perform close to transferred standard models [19].

The Table 1 ranks recent research and exploits several metrics, such as Accuracy, Sensitivity, and Specificity, to compare the performance of models.

In each study, the first column of the table illustrates the best pre-trained model in terms of accuracy when compared to the others.

### 2.2 Custom CNN models

Custom CNN models are deep learning architectures that are designed from scratch to solve a specific problem or task. They are constructed using a series of layers and tailored to the specific requirements of the problem at hand. Compared to pre-trained models, custom CNN models can have a much higher level of control over the architecture’s structure, allowing for more flexibility and customization. There were several contributions in which authors presented their own CNN models without depending on pre-trained models. The Table 2 presents some of these studies aiming to classify chest X-Ray images as COVID-19 or other classes.

<table>
<thead>
<tr>
<th>Best Model &amp; Year</th>
<th>Number of Pre-Trained Models</th>
<th>Dataset</th>
<th>Classification</th>
<th>Accuracy (%)</th>
<th>Sensitivity (%)</th>
<th>Specificity (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Googlenet [22] 2020</td>
<td>Three (3): Alexnet, Googlenet, and Restnet18</td>
<td>307 CHEST X-RAY images including 69 images of COVID-19</td>
<td>4 classes: (Covid19, normal, pneumonia bacterial, pneumonia virus)</td>
<td>80.56</td>
<td>80.56</td>
<td>/</td>
</tr>
<tr>
<td>VGG16 [23] 2020</td>
<td>Nine (9)</td>
<td>/</td>
<td>3 classes: (Covid19, Normal, pneumonia bacteria)</td>
<td>95.88</td>
<td>/</td>
<td>/</td>
</tr>
<tr>
<td>VGG16 [24] 2020</td>
<td>/</td>
<td>8474 CHEST X-RAY images</td>
<td>3 classes: (Covid19, infected pneumonia, no-COVID-19 infected pneumonia, normal (non-pneumonia))</td>
<td>94.50</td>
<td>98.40</td>
<td>98.00</td>
</tr>
<tr>
<td>VGG16 [25] 2020</td>
<td>/</td>
<td>396 CHEST X-RAY images including 132 of COVID-19</td>
<td>3 classes: (Covid-19, healthy, normal)</td>
<td>86.00</td>
<td>86.00</td>
<td>93.00</td>
</tr>
<tr>
<td>ResNet50V2 [27] 2020</td>
<td>5 (VGG16 ResNet50V2 XceptionMobileNet DenseNet121)</td>
<td>678 CHEST X-RAY Images</td>
<td>4 classes: (normal, bacterial, COVID-19, viral)</td>
<td>98.15</td>
<td>98.26</td>
<td>98.89</td>
</tr>
<tr>
<td>VGG16 [28] 2021</td>
<td>VGG16, VGG19, ResNet, DenseNet, and InceptionV3</td>
<td>120 CHEST X-RAY Images</td>
<td>2 classes: (Negative; Positive)</td>
<td>80.00</td>
<td>80.00</td>
<td>/</td>
</tr>
<tr>
<td>Densenet161 [29] 2021</td>
<td>3 (VGG16 Densenet161 ResNet50)</td>
<td>1431 CHEST X-RAY Images</td>
<td>2 classes: (Covid-19, pneumonia)</td>
<td>97.10</td>
<td>100</td>
<td>/</td>
</tr>
<tr>
<td>vgg19 [30] 2022</td>
<td>One (vgg19)</td>
<td>3797 CHEST X-RAY Images</td>
<td>3 classes: (Covid; pneumonia, healthy)</td>
<td>97.11</td>
<td>97.00</td>
<td>/</td>
</tr>
</tbody>
</table>
### Table 2. Custom CNN models

<table>
<thead>
<tr>
<th>Model &amp; Year</th>
<th>Dataset</th>
<th>Classification</th>
<th>Accuracy (%)</th>
<th>Sensitivity (Recall) (%)</th>
<th>Specificity (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>CovXNet [31] 2020</td>
<td>5856 CHEST X-RAY images including 305 images of Covid-19</td>
<td>4 classes: (Covid-19, Normal, viral, bacterial pneumonia)</td>
<td>90.00</td>
<td>89.00</td>
<td>89.00</td>
</tr>
<tr>
<td>CoroNet [33] 2020</td>
<td>1300 CHEST X-RAY images including 290 images of Covid-19</td>
<td>4 classes: (Covid, Pneumonia bacterial, pneumonia viral, normal)</td>
<td>89.60</td>
<td>89.92</td>
<td>96.40</td>
</tr>
<tr>
<td>multi-CNN [34] 2020</td>
<td>950 CHEST X-RAY images including 453 images of Covid-19</td>
<td>2 classes: (Covid, Normal)</td>
<td>91.16</td>
<td>98.50</td>
<td>/</td>
</tr>
<tr>
<td>COVID-NET [35] 2020</td>
<td>13,975 CHEST X-RAY images</td>
<td>3 classes: (Normal, Pneumonia, Covid)</td>
<td>93.30</td>
<td>91.00</td>
<td>/</td>
</tr>
<tr>
<td>Covid-xnet [36] 2020</td>
<td>/</td>
<td>2 classes: (normal, Covid)</td>
<td>94.43</td>
<td>92.53</td>
<td>96.30</td>
</tr>
<tr>
<td>DarkCovidNet [37] 2020</td>
<td>/</td>
<td>2 classes: (Covid, no-findings)</td>
<td>98.08</td>
<td>95.13</td>
<td>95.30</td>
</tr>
<tr>
<td>[38] 2020</td>
<td>/</td>
<td>2 classes: (Covid; normal)</td>
<td>93.00</td>
<td>/</td>
<td>/</td>
</tr>
<tr>
<td>CNN [39] 2021</td>
<td>1184 CHEST X-RAY images</td>
<td>2 classes: (Covid-19; normal)</td>
<td>98.65</td>
<td>98.49</td>
<td>98.82</td>
</tr>
<tr>
<td>OptCoNet [40] 2021</td>
<td>900 CHEST X-RAY images</td>
<td>3 classes: (Covid, normal, pneumonia)</td>
<td>97.78</td>
<td>97.75</td>
<td>96.25</td>
</tr>
<tr>
<td>2019NCOV [41] 2021</td>
<td>9544 Chest X-rays Images</td>
<td>2 classes: (Covid, normal)</td>
<td>93.80</td>
<td>99.50</td>
<td>/</td>
</tr>
<tr>
<td>SARS-Net [42] 2022</td>
<td>13975 CHEST X-RAY images</td>
<td>3 classes: (Normal, Pneumonia, Covid)</td>
<td>97.60</td>
<td>92.90</td>
<td>/</td>
</tr>
<tr>
<td>2dCNN-BiCuDNNLSTM [43] 2022</td>
<td>6863 CHEST X-RAY images (1000 COVID-19 patients, 3863 normal cases, and 2000 pneumonia patients)</td>
<td>3 classes: (Covid, normal, pneumonia)</td>
<td>93.00</td>
<td>100</td>
<td>/</td>
</tr>
<tr>
<td>RT2022 [44] 2022</td>
<td>5255 CHEST X-RAY images</td>
<td>2 classes: (Covid, normal)</td>
<td>90.60</td>
<td>/</td>
<td>/</td>
</tr>
</tbody>
</table>

### Table 3. Hybrid CNN models

<table>
<thead>
<tr>
<th>Model &amp; Year</th>
<th>Dataset</th>
<th>Classification</th>
<th>Used Pre-Trained Models</th>
<th>Accuracy (%)</th>
<th>Sensitivity (Recall) (%)</th>
<th>Specificity (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Covid</td>
<td>/</td>
<td>/</td>
<td>VGG19, InceptionV3, MobileNet</td>
<td>97.88</td>
<td>98.67</td>
<td>98.00</td>
</tr>
<tr>
<td>ConvLSTM</td>
<td>/</td>
<td>/</td>
<td>Vgg16</td>
<td>96.91</td>
<td>98.33</td>
<td>98.68</td>
</tr>
<tr>
<td>[47] 2020</td>
<td>6432 CHEST X-RAY images</td>
<td>3 classes:COVID-19 positive, pneumonia positive, Normal</td>
<td>/</td>
<td>96.74</td>
<td>/</td>
<td>/</td>
</tr>
<tr>
<td>[49] 2021</td>
<td>/</td>
<td>/</td>
<td>VGG16+ SVM, VGG-16, ResNet50</td>
<td>98.82</td>
<td>99.82</td>
<td>/</td>
</tr>
<tr>
<td>[50] 2022</td>
<td>6432 X-ray images</td>
<td>3 classes: - COVID-19(positive), non-COVID-19 (negative), normal (negative))</td>
<td>MobileNetV2</td>
<td>96.00</td>
<td>95.00</td>
<td>91.00</td>
</tr>
<tr>
<td>[51] 2022</td>
<td>912 CHEST X-RAY images for regular people, and 912 CHEST X-RAY images for COVID-19 infected people</td>
<td>InceptionResnet-v2</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td></td>
</tr>
<tr>
<td>[52] 2022</td>
<td>2613 CHEST X-RAY images</td>
<td>2 classes:(Covid; normal)</td>
<td>Inception-v3, ResNet50, ResNet101</td>
<td>99.00</td>
<td>98.00</td>
<td>100</td>
</tr>
<tr>
<td>[53] 2022</td>
<td>6000 CHEST X-RAY Images</td>
<td>3 classes: (Covid, pneumonia, normal)</td>
<td>EfficientNetB0</td>
<td>98.00</td>
<td>100</td>
<td>99.00</td>
</tr>
</tbody>
</table>
2.3 Hybrid CNN models

Hybrid CNNs are models which combine predesigned CNNs with other machine learning techniques such as Support Vector Machine (SVM). The idea is simply to add new blocks to the pre-trained CNN model. For example, in the study [45], the process of Covid-19 detection using chest X-rays operated on three popular CNN models: VGG19, MobileNet, and InceptionV3 by adding two mechanisms: ConvLSTM and SE block. ConvLSTM means Convolutional Long Short-Term Memory which is a layer used to encode the spatial dependency among the feature maps received from the CNN's last layer of convolution and also to improve the model's image representational capability. The other block is squeeze-and-excitation block. It is then added and used to assign weights to significant local features. These two mechanisms are employed to improve the classification strength of CNN models through VGG19 + ConvLSTM + SE block, Inception V3 + ConvLSTM + SE block, and MobileNet + ConvLSTM + SE block. Accuracies using only VGG16, Inception V3, and MobileNet are respectively: 96.01%, 95.22%, and 96.98%. The accuracies after adding blocks are respectively: 97.88%, 97.23%, and 97.80%.

The Table 3 presents a brief overview of significant hybrid CNN Models.

3. CNNS MATERIALS AND TECHNIQUES

Deep learning techniques have the potential for modeling large sets of complex data. Building and training accurate and efficient deep learning models are based on key concepts. First, the neural network architecture is the foundation of a deep learning model. It is composed of layers of interconnected artificial neurons that process the input data and generate output predictions. The design of the architecture is crucial to the model's performance. Other concepts must be considered carefully such as the selection of appropriate activation functions used to introduce non-linearity into the output of a neural network. Also, the regularization techniques are used to reduce variance in the model by introducing restrictions that prevent it from overfitting to the training data. The accuracy and efficiency of the deep learning model can significantly differ depending on the hyperparameters that are selected. Hyperparameter tuning is a critical step to fine-tune the performance of the model but is also a challenging and time-consuming task.

CNN is one of the most widely used models in the domain of deep learning, because the appropriate features of input data are automatically extracted without any intervention of humans. CNN covers a wide range of topics: we are automatically extracted without any intervention of deep learning, because the appropriate features of input data.

3.2 Dataset

To combat the pandemic of COVID-19, open source data and methods are needed to enable the global scientific community to collaborate on research that is verifiable and transparent [59, 60]. As a result, a set of electronic medical repositories are developed to help researchers locate the appropriate resource. In our context, we are interested in public chest X-ray datasets. Several datasets are developed such as PACDCHES1 [61]. It is reported at San Juan hospital in Spain by radiologists, the project (COVID chest X-ray) [62] has the aim of collecting X-ray images showing COVID-19 from online and others sources [63]. In general, most of the available X-ray datasets for COVID-19 classification are collected from Kaggle and Github platforms, such as: Covid-19 radiography database [64], COVID-19 image-dataset [65], Covid-chest X-ray-dataset [66]. However, these datasets are limited by the number of lung images with COVID-19 infection. And this is not efficient for training a deep learning model. The model may overfit the data. Therefore, a hybrid dataset combining different repositories is used to train the existing works [67].
The dataset is split into different sets, once it is determined. The purpose of splitting is to avoid overfitting. In deep learning, overfitting occurs when the capacity to learn is so large that the network is learning wrong features instead of meaningful patterns. The original dataset is usually divided into three sets: training, validation, and test sets. The model is trained to employ the training set. The validation set contains examples utilized to vary parameters in the learning process. The final model evaluates the testing set and compares it to the previous datasets. Generally, there are no rules that specify how to partition the data. This may depend on the predictor's number in a predictive model or the original data pool size. Each dataset split should contain subfolders. The number of subfolders depends on the number of output classifications (binary, ternary, ...), for example, in binary classification, we find for each dataset two subfolders: Covid and normal.

### 3.3 Preprocessing Techniques

Preprocessing techniques are methods aiming to transform data before it is delivered to the machine learning or deep learning algorithm. Training a CNN on unprocessed images can lead to poor classification results, which can be interpreted in terms of overfitting issues. Overfitting is considered to be a major problem in deep learning. To reduce this problem, an important step must be added as a preprocessing step. It is called data augmentation, which is the process of artificially increasing the amount of data by generating new data points from existing data [68]. Data augmentation techniques include:

- **Resizing**: To standardize the dataset, all the images are resized into a fixed size such as 224×224 or 299×299.
- **Flipping or Rotating**: To increase the sample size of the datasets, horizontal and vertical flipping are mainly used.
- **Scaling or Cropping**: To reduce the redundancy, because not all parts of the images need to be used.
- **Brightness or Intensity**: To increase or decrease the images brightness.

Traditional image preprocessing techniques such as Contrast Stretching, Histogram Equalization, Smoothing, Sharpening, Adaptive Winner Filter, Histogram Enhancement, and Color Space Conversion, are used in some works [69-72]. As a consequence, the preprocessing step is essential to improve data quality and accurate AI Models.

### 3.4 Basic CNN Architecture for Classification

The structure of CNNs is inspired by neurons in the brains of humans or animals. It works in a similar way to a conventional neural network [73]. The CNN composition consists of a series of connected layers of three types: convolutional layers, pooling layers, and fully connected layers. The CNN architecture is based on a combination of these layers [74]. The role of the convolutional layer is to subjugate the input data (image) to the convolution operation to produce a series of output features.

Pooling layers can reduce the size of the input image which comes from the convolution results without losing important features. This reduces the number of parameters and computations in the network. This leads to identifying patterns in the data more quickly and accurately in the network.

Finally, the last type of layers is Fully Connected (FC) layers. They are usually positioned before the output layer and form the final layers of a CNN structure. After feature extraction, comes the role of fully connected layers to split the data into different classes, so as to classify the input (see Figure 2 [75]).

![Figure 2. CNN architecture](image)

Usually, a complementing layer is added in the CNN architecture: The dropout layer. It is a technique that drops some neurons toward the next layer and doesn't modify all others. The dropped neuron does not participate in the back-propagation or the forward-propagation during the training process, but it is a part of the full-scale network to perform predictions during the testing process. In the prediction process, a technique called the Activation Function is adopted. It selects a neuron that should be activated or not. This means that it decides whether the neuron’s input to the network is significant or not. ReLU (Rectified Linear Unit), Softmax, Sigmoid, and tanH (Hyperbolic tangent) are the most used activation functions.

- **ReLU**: It converts all values of the input to positive numbers. The central advantage of ReLU compared to others is the lower computational load, because it does not activate, simultaneously, all the neurons.
- **Softmax**: It is usually applied as a final activation function in a neural network. It aims at normalizing a network's output to a probability distribution over a set of classes (predicted output classes).
- **Sigmoid**: It is used for binary classification in the CNN model. The input of the sigmoid function is real numbers, and the output value is limited between 0 and 1.
- **Tanh**: It is equivalent to the sigmoid function. Its input is real numbers, but the output is limited to between -1 and 1.

### 3.5 Popular CNN Models Categories

CNN architectures have undergone numerous recent developments. Bhatt et al. [76] summarized 8 categories of CNN architectures: (1) Spatial exploitation, (2) multi-path, (3) depth, (4) breadth, (5) dimension, (6) channel boosting, (7) feature-map exploitation, and (8) attention-based CNN.

**Spatial Exploitation (SE)**: In order to improve performance, spatial filters have been used. Spatial filters typically consist of a small matrix of weights (usually 3x3 or 5x5) applied to each pixel in the image. The strength of the feature detected in the image is determined by the weights in the filter. Convolutional operations are affected by the size of the filter. Small filters are usually used to extract fine-grained information, while large filters are used to extract coarse-grained information.

**Multi-path (MP)**: Shortcut connections or numerous pathways allow the flow of information between layers by evading some in-between levels [77]. Through cross-layer connectivity, the network is divided into sections. These
pathways extend the gradient to lower layers, which resolve the vanishing gradient issue.

**Depth (Dp):** Network depth is a key factor in determining the performance of a network and its ability to learn [78-81]. Deep networks are more effective at representing specific classes of functionality than shallow systems.

**Breadth (B):** Besides depth, width is a key factor in network learning process. It is confirmed that using ReLU activation functions in neural networks needs to be wide enough to maintain a universal approximation property, while at the same time increasing in depth [82]. However, the main issue is that a set of layers may fail to learn features even if the depth is raised. Furthermore, if the maximum width of a network is not greater than the input dimension, then it is not possible for any deep network to approximate a class of continuous functions on a compact set [83]. Consequently, to address this issue, researchers focused on wide and thin designs than deep and narrow ones.

**Dimension (Dm):** separable convolutions (called also depth-wise separable) are introduced to improve the efficiency of normal convolutions. The separable convolutions use point-wise to encode the spatial one, and depth-wise to encode the channel-wise.

**Channel boosting (CB):** The channel boosting is called also as the input channel dimension. It adds supplementary learners in the CNN for the enhancement of the network’s representation [84]. The input representation is an influential parameter in network learning. Network performance can be hampered by poor diversity and information about the class in the input.

**Feature-map exploitation (FE):** The selection of feature maps can have a significant impact on the improvement of network generalization. Different feature extraction steps are achieved, allowing for diverse types of features. However, excessive feature sets may provide a noise effect, leading to the over-fitting of the network [85].

**Attention (A):** In CNN, improving representation and overcoming computational limitations are the aims of using the concept of attention. It also gives intelligence to a CNN in order to differentiate elements even in complex scenarios and busy backdrops.

The Table 4 lists some CNN models basing on each cited category [33, 41].

### Table 4. Summary of popular CNN Models in each category

<table>
<thead>
<tr>
<th>CNN Model &amp; Year</th>
<th>Category</th>
<th>Dataset</th>
<th>Main Findings</th>
<th>Depth</th>
<th>Parameter</th>
<th>Error Rate</th>
<th>Input Size</th>
</tr>
</thead>
<tbody>
<tr>
<td>LeNet1998</td>
<td>SE</td>
<td>MNIST</td>
<td>1st CNN architecture</td>
<td>0.060M</td>
<td>0.95</td>
<td>227×227×3</td>
<td>/</td>
</tr>
<tr>
<td>AlexNet2012</td>
<td>SE</td>
<td>ImageNet</td>
<td>Dropout + ReLU</td>
<td>8</td>
<td>60M</td>
<td>16.4</td>
<td>224×224×3</td>
</tr>
<tr>
<td>ZNet2014</td>
<td>SE</td>
<td>ImageNet</td>
<td>Intermediate layers visualization</td>
<td>8</td>
<td>60M</td>
<td>11.7</td>
<td>224×224×3</td>
</tr>
<tr>
<td>VGG2014</td>
<td>SE</td>
<td>ImageNet</td>
<td>Small filter size + increase depth</td>
<td>16−19</td>
<td>138 M</td>
<td>7.3</td>
<td>224×224×3</td>
</tr>
<tr>
<td>GoogleNet2015</td>
<td>SE</td>
<td>ImageNet</td>
<td>different filter sizes + increase depth +block concept</td>
<td>22</td>
<td>4 M</td>
<td>6.7</td>
<td>224×224×3</td>
</tr>
<tr>
<td>Inception-V3</td>
<td>Dp+B</td>
<td>ImageNet</td>
<td>Small filter size+ better feature representation</td>
<td>48</td>
<td>23.6 M</td>
<td>3.5</td>
<td>229×229×3</td>
</tr>
<tr>
<td>High way 2015</td>
<td>Dp+MP</td>
<td>CIFAR-10</td>
<td>1st idea multipath</td>
<td>19-32</td>
<td>2.3 M</td>
<td>7.76</td>
<td>32×32×3</td>
</tr>
<tr>
<td>Inception ResNet</td>
<td>Dp+B+MP</td>
<td>CIFAR-10CIFAR-100</td>
<td>Residual links</td>
<td>164</td>
<td>55.8 M</td>
<td>3.52</td>
<td>229×229×3</td>
</tr>
<tr>
<td>Residual Attention2017</td>
<td>A</td>
<td>CIFAR-100</td>
<td>1st model to introduce attention mechanism</td>
<td>452</td>
<td>8.6 M</td>
<td>3.90</td>
<td>40×40×3</td>
</tr>
<tr>
<td>Squeeze-and-excitation Networks2017</td>
<td>FE</td>
<td>ImageNet</td>
<td>Modeled interdependencies between channels</td>
<td>152</td>
<td>27.5</td>
<td>2.3</td>
<td>229×229×3</td>
</tr>
<tr>
<td>Channel Boosted CNN 2018</td>
<td>CB</td>
<td>/</td>
<td>Boosted original channels with additional information by artificial channels</td>
<td>/</td>
<td>/</td>
<td>/</td>
<td>224×224×3</td>
</tr>
<tr>
<td>EdgeNet2019</td>
<td>Dm</td>
<td>/</td>
<td>Introduced concept of visual intelligence at the edge</td>
<td>/</td>
<td>/</td>
<td>/</td>
<td>320×320×3</td>
</tr>
<tr>
<td>DiceNET 2021</td>
<td>Dm</td>
<td>/</td>
<td>Introduced dimension-based CNN, including height, width, and depth</td>
<td>/</td>
<td>/</td>
<td>/</td>
<td>/</td>
</tr>
</tbody>
</table>

### 3.6 CNN model performance

The evaluation of a CNN algorithm’s classification performance is important. Generally, Accuracy, Recall/Sensitivity, Precision, Specificity, F1 Score, confusion matrix, Area Under Curve (AUC) are the most used metrics for the evaluation. To test a CNN model, a variety of these evaluation metrics are mixed.

Metrics formulas base on the terms TP, TN, FP, FN, where:

- **TP** means True positive, the number of cases that were correctly predicted as positive;
- **TN** denotes True Negative, the number of cases that were correctly predicted as negative;
- **FP** signifies False Positive, the number of cases that were incorrectly predicted as positive;
- **FN** means False Negative, the number of cases that were incorrectly predicted as negative.

The accuracy is a metric used to measure the percentage of the true and negative cases that are correctly predicted. It is calculated by the formula:

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

The sensitivity (called also Recall) is a measure of a model's ability to correctly identify true positives. It is given
by the formula:

\[
Recall = \frac{TP}{TP + FN}
\] (2)

The specificity is the percentage of negative cases correctly predicted. It is calculated by the formula:

\[
Specificity = \frac{TN}{TN + FP}
\] (3)

The precision measures the proportion of positive prediction results that are correct.

\[
Precision = \frac{TP}{TP + FP}
\] (4)

There are additional metrics, such as the confusion matrix and the Area under the Curve (AUC), that the present paper does not mention.

### 3.7 Process of image classification using CNN models

The image classification technique is used for classifying or predicting the category of a particular object in an image. The main goal of this technique is identifying accurately the features in an image. In this study, we summarize the process of image classification based on the CNN model in 7 main steps (see Figure 3). The learning stage includes step 1 to step 5, while the inference one is determined by step 6 and step 7.

![Figure 3. Process of Image Classification using CNN](image)

**Step 1:** Choosing the Dataset, a large training dataset is essential;

**Step 2:** Preprocessing the dataset using different techniques;

**Step 3:** Construction of the CNN model;

**Step 4:** Training and testing the model on the chosen dataset (in step 1);

**Step 5:** The classification performance of a CNN algorithm is evaluated using metrics such as accuracy, specificity, F1 score, and sensitivity.

**Step 6:** Preprocessing the X-Ray image to be classified to enhance it and reduce noise or unwanted details;

**Step 7:** CNN classification of the introduced image.

### 4. DISCUSSION AND CONCLUSION

In literature, a lot of works demonstrate that the CNN model is a powerful method to classify Covid-19 using X-Ray images. We have reviewed significant research papers since the emergence of the virus. Based on the literature and after analyzing studies in this field, we have classified CNN Models into three categories. The first one focuses on transfer learning from pre-trained CNN models. In the second, the CNN model is developed from scratch and called the Custom CNN Model. The last category couples the two previous ones to constitute the Hybrid CNN Model. We have presented a summary of each model in each category by giving the dataset size, the number of classes, the model architecture, and the performance evaluation criteria (Accuracy, Sensitivity, and Specificity). As deep learning models need a large amount of data, dataset size is considered a critical factor in determining the classification performance of the CNN model. The majority of CNN models have achieved high performance for Covid-19 classification. Without ignoring that most research applies different techniques for preprocessing data to enhance their quality.

We further observed that models based on transfer learning from pre-trained models are the most used category in terms of performance and this is for multiple reasons. A pre-trained model that has learned general features can be used to improve the accuracy of the model on a new dataset. By transferring the knowledge acquired from a pre-trained model, transfer learning decreases the training time as well. Since the model has already learned the features, it requires less time to learn new features and patterns in the new domain. Fewer computational resources are needed to train a new model and achieve comparable results. Finally, Transfer learning allows models to be more easily scaled to new applications, platforms, and tasks, allowing for more efficient use of resources and increased flexibility. Without ignoring the great efforts of the developed models in the two other categories.

Custom CNN models can provide high accuracy if they are designed correctly and trained on large datasets. Still, they tend to require more computational power and time to train and optimize compared to pre-trained models and transfer learning techniques.

Hybrid CNN models can be challenging to design and optimize, and finding the optimal combination and tune hyperparameters can be time-consuming. However, these models can lead to superior performance and generalization ability compared to both transfer learning and custom CNN models.

The proposed models in the three categories have the same goal which is the contribution to reduce the ongoing pandemic’s impact by improving diagnosis accuracy and assisting research efforts.

We remarked, in our survey, that CNN Models for COVID-19 detection acquire data from a wide range of sources (labs, hospitals, universities, research centers, etc.). So developing a health data collection platform is a necessity that enhances the possibilities for storing, sharing, and analyzing health data among researchers in several fields. Through real-time data acquisition and AI machine learning, researchers have access to huge, diverse, and high-quality datasets. The platform can help them uncover new patterns, create novel therapies, interventions, and prevention methods, and make informed judgments, all of which can reduce
illness burden and improve patient outcomes through improved medical research.

We noticed, also, that the proposed models are not applicable or beneficial in hospitals or clinics, while most of these efforts and research aim to help the radiologist to make better decisions in tackling Covid-19. CNNs have the potential for early detection and diagnosis. CNNs have been used to monitor areas to ensure that people maintain social distancing. CNNs can accelerate the vaccine development process by using computational models to predict which parts of the virus are most prone to inducing an immune response. This helps researchers design and optimize vaccine candidates that can potentially protect against COVID-19.

Finally, we aim with this paper to provide important insights for CNN models based on X-Ray images to help researchers develop a standardized, reliable, and comprehensive system for COVID-19 diagnosis and control that can scale up globally and help mitigate the impact of the pandemic on public health and the economy.

REFERENCES


model for COVID-19 detection from chest X-rays.

https://doi.org/10.1016/j.eswa.2022.117812


https://doi.org/10.36227/techrxiv.12839204.v1


