An Empirical Assessment of Discriminative Deep Learning Models for Multiclassification of COVID-19 X-rays

Sunday Adeola Ajagbe[1,2,*], Pragasen Mudali[2] and Matthew Olusegun Adigun[2]

[1]Abiola Ajimobi Technical University, Ibadan, Nigeria
[2]University of Zululand, Kwadlangezwa, South Africa

Abstract

The current era of pandemic and infectious diseases demands contemporary technologies across many industries. Modern inventions and technology have advanced significantly, primarily in disease detection, control, and influence over intelligent healthcare instruments and facilities. Artificial intelligence (AI), a rapidly developing technology, was critical in the detection of COVID-19 based on the X-ray modality. Most existing studies focus on binary classification using DL models and COVID-19 X-rays and there is a limited assessment of the strengths and weaknesses of deep learning (DL) models in multiclassification of COVID-19 using X-ray images. Therefore, this study focuses on the empirical assessment of discriminative DL models for multiclassification of COVID-19 X-rays. We suggest four phases of approaches, namely, data acquisition, preprocessing, modeling, and training for four classes of diseases (normal, pneumonia, COVID-19, tuberculosis) as well as evaluation phases in this investigation. Convolution neural network, (CNN) recurrent neural network (RNN), and multilayer perceptron (MLP) were discriminative DL models implemented. The CNN model demonstrates an effective and valuable approach for the multiclassification of diseases as classification accuracies of 0.9066, 0.6278, and 0.7652 were obtained for CNN, RNN, and MLP respectively. Discriminative DL models demonstrate the feasibility of multiclassifying COVID-19 X-rays. Implementing this approach will alleviate the burden on radiologists and other medical professionals while enhancing the precision and effectiveness of COVID-19 diagnosis.

Keywords

Artificial intelligence (AI), Discriminative deep learning, COVID-19 X-ray images,, Multiclassification classification, Machine learning.

1. Introduction

Recently, a new virus known as COVID-19 began to infect the lungs as well as the upper respiratory tract. On the scale of a worldwide epidemic, the incidence and mortality rates have been increasing daily [1, 2]. The disease has been diagnosed using chest X-ray images, which have been shown to help monitor a variety of lung disorders. Chest X-ray pictures are recognized to be useful for the observation and assessment of several lung conditions, including pneumonia, hernias, atelectasis, infiltration, and tuberculosis. The Wuhan region of China conducted the initial investigation of COVID-19 in late 2019 [3]. The virus primarily affects the airway and subsequently the lungs of individuals infected. It manifests as an infection affecting the upper respiratory tract and lungs [4]. Research has demonstrated the value of chest X-rays in monitoring the damage that COVID-19 does to lung tissue. Thus, chest X-ray pictures could potentially be utilized to identify COVID-19.

Technological development is changing with human evolution. The current era is one in which demand for contemporary technologies is high across many industries. Modern inventions and technology have advanced significantly, primarily in disease detection, control, and influence over intelligent healthcare instruments and facilities. Artificial intelligence (AI), a rapidly developing technology, was critical in the diagnosis of COVID-19. Previously, AI was employed in the analysis of medical images, resulting in improved accuracy and directly or indirectly contributing to the reduction of time and labor required for COVID-19 detection, which are significant factors. AI approaches such as deep learning

ICAIW 2024: Workshops at the 7th International Conference on Applied Informatics 2024, October 24–26, 2024, Viña del Mar, Chile
*Corresponding author.
230015266@stu.unizulu.ac.za (S. Ajagbe); MudaliP@unizulu.ac.za (P. Mudali); adigunm@unizulu.ac.za (M. Adigun)
0000-0002-7010-5540 (S. Ajagbe); 0000-0003-4720-1726 (P. Mudali); 0000-0001-6256-5865 (M. Adigun)

(DL) and machine learning (ML) have been increasingly important in recent years for healthcare applications [5, 6]. Deep learning methods are an effective method for automatically analyzing COVID-19 detected in CT scans and X-ray pictures [7]. These are the two imaging modalities used to diagnose COVID-19 [6]. A few systems have been improved by taking into account pre-trained models, while some are using customized neural networks using DL methods and input images from CT and X-ray specimens. There have been a few DL-based methods for illness detection using chest X-ray pictures in the literature [8, 9].

The primary goal of [10] was to review the DL methods for the detection and prediction of pandemic and [11] entails the applications (reviews of current advancements), challenges of the internet of things (IoT) and convolution neural networks (CNN) which are forth industrial revolution (4IR) technologies that are prominent in pandemic prevention and control. In [11], the taxonomy of DL was presented, and it was broadly classified into three; discriminative, generative, and hybrid learning. Convolution neural network, (CNN) recurrent neural network (RNN), and multilayer perceptron (MLP) are categorized as discriminative DL models [12].

Each of these categories has a different structure and learning pattern. Researchers and experts select and use these methods even in COVID-19 classification and diagnosis without a thorough investigation of the strengths and weaknesses of these methods. This may be informed by the hyperparameter of the deep learning models or data dimensionality. In addition, most existing studies focus on binary classification using DL models and COVID-19 X-ray image datasets as radiologists' complaints of complex traits pose difficulties in interpreting COVID-19 X-ray images. Unlike many other DL-based methods that have been proposed in the literature, the DL-based approaches in this work are proposed for multiclassification of COVID-19 based on chest X-ray images that go beyond binary classification and, in a sense, this study is very promising to address other related infectious diseases and pandemic diagnosis thereby control the spread of pandemic or infectious diseases. This study also expands the evaluation methods beyond accuracy as it focuses on the empirical assessment of discriminative DL models for multiclassification of COVID-19 X-rays. Thus, the objective of this study is to:

- carries out a multiclassification of COVID-19 chest X-ray using discriminative DL models.
- evaluate the performance discriminative DL models for multiclassification of COVID-19 chest X-ray using accuracy (Acc), sensitivity (Sen), specificity (Spe), F1 measure, and confusion matrix.

This is how the rest of the paper is organized. In addition to providing a brief overview of the approach, Section 2 provides the background on the discriminative DL models for image processing, while Section 3 includes a review of the relevant works. Section 4 describes in full the methodology employed in the study. The outcomes and analysis are covered in Section 5, and the study's conclusions and future suggestions are presented in Section 6.

2. Background on the Discriminative DL in Image Processing

This section discusses a summary of the recent literature that has chosen the discriminative DL-based methods to carry out to carry out this work. The primary goal here is to produce a study that reviews current advancements in DL-based COVID-19 diagnosing systems. A class of deep learning models that is employed to deliver a discriminative function in classification tasks is called discriminative DL. Discriminative DL models are commonly developed to enhance the ability to classify patterns by characterizing the distributions of classes based on accessible data [13]. Discriminative models primarily consist of MLP, CNN, RNN, and their respective variations [13].

1. Convolutional Neural Network (CNN): CNN is the predominant model utilized in DL perhaps because of the ability to incorporate many layers in the learning process. The approach has gained widespread usage, particularly in the field of image processing, in recent years. CNN is composed of several layers, including a convolutional layer, an activation function, pooling, and a fully connected layer [13, 14]. Convolutional layers are typically arranged sequentially to extract

feature patterns from the basic characteristics of images and progress towards more complex features [15, 16]. Activation functions in CNN architecture are mathematical functions that map incoming inputs to a specific range or selectively accept and discard certain input values. Pooling layers, however, allow for the reduction of feature matrix size by sampling. The fully connected layer is responsible for performing the classification process based on the features gathered via convolution, activation function, and pooling. This layer functions similarly to a traditional artificial neural network. Before the classification phase, the feature matrices are converted into feature vectors by a process known as flattening. The output of CNN is obtained according to equation 1. Where $\sum X_i^{l-1}$ represent the feature obtained from the previous layers, and k_{ij}^l and b_i^{l-1} represent the adjustable Kanels and training bias respectively. Figure 1 shows the general framework of the CNN classifier in this case.

$$X_j^l = f\left(\sum_{i \in Mj} X_i^{l-1} * k_{ij}^l + b_j^l\right) \tag{1}$$

2. Multi-layer perceptron (MLP): Advancements in DL and transformer particularly, MLP models have introduced novel network architectures for computer vision challenges. While these models have demonstrated effectiveness in various vision tasks, such as image identification, there are still difficulties in applying them to low-level vision [17]. The Perceptron Learning algorithm is derived from the previously discussed back-propagation rule. This algorithm can be implemented using any programming language, including Python which is specifically in the context of this study [18].
3. Recurrent Neural Networks (RNN): The progress in the field of image compression networks utilizing RNN is very limited in comparison to CNN and auto-encoders. The suggested solutions [19] utilize a comprehensive image resolution network that incorporates residual scaling, RNN, and entropy coding based on DL [20].

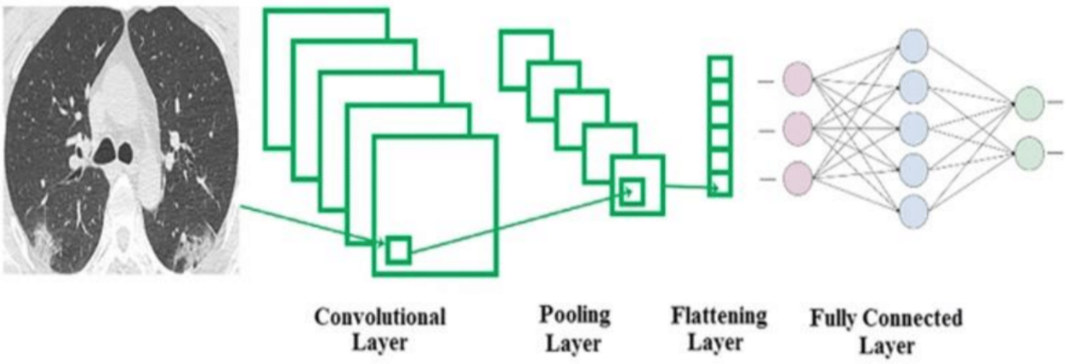

Figure 1: General framework of the CNN classifier [13].

3. Review of the Related Works on DL Models for COVID-19 Diagnosis.

This section provides a summary of related works that depicts the recent studies of DL models for COVID-19 diagnosis. It is presented in Table 1 that identified the models used, the contributions as well as the limitations of recent studies in this area.

COVID-19 datasets may exhibit an unequal distribution of samples across different classes, leading to a potential bias in the model towards the class with the highest number of samples. There is a scarcity of high-quality labeled datasets, particularly during the initial phases of a pandemic. Scarcity

Table 1
Review of the related works

Ref	Models	Contributions	Limitations
[21]	DenseNet 121	The model possesses unique characteristics due to its pre-existing training for the identification of different lung ailments. To address the issue of class imbalance, the weighted loss function was employed.	The model was trained using a limited dataset. The dataset's class imbalance requires attention. It focuses on binary classification.
[22]	CNN	The suggested CNN model, developed from scratch, achieved higher accuracy when trained on a specific medical dataset compared to pre-trained models based on a more general ImageNet dataset. The model was trained using limited resources and within a short timeframe.	The model underwent training using a restricted dataset. The study was limited to two classes.
[23]	COVID-Net	The COVID-Net architecture was made publicly available for open access. Utilized lightweight design patterns. Implemented selective long-range connectivity in strategic areas to enhance representational capacity and streamline training, while ensuring optimal computational complexity and memory efficiency.	Enhancing sensitivity is crucial to minimize the number of COVID-19 cases that go undetected. It focuses on binary classification.
[24]	Deep CNN based on COVID-Net	Regularization techniques were employed to address the issue of data imbalance. Preprocessing the data enhanced the performance of the model.	The model suffers the issue of class imbalance. It focuses on binary classification.
[25]	ResNet 50 & SVM	The utilization of pre-trained networks has simplified the process of classification tasks. The calculation of pandemic uncertainty was performed to ensure the accurate identification of classes.	Exclusively concentrated on the posterior-anterior (PA) perspective of the X-rays, hence unable of distinguishing alternative viewpoints of the X-ray images. The X-ray images containing multiple disease symptoms were not efficiently classified. It focuses on binary classification.
[26]	CNN	Utilizing preprocessing procedures (histogram equalization algorithm, and a bilateral low-pass filter) effectively to improve the performance of CNNs in the detection of COVID-19 from chest X-ray.	Possible constraints on the capacity to apply findings to a wider context, resulting from particular methods used to prepare the data and the unique attributes of the dataset.
[27]	CVDNet	The utilization of batch normalization approach enhances the rate at which convergence occurs during the training process. The vanishing gradient problem was resolved through the utilization of a residual network.	Enhancing sensitivity is necessary to minimize the number of undetected COVID-19 instances. The study was limited to binary classification.

can result in overfitting, a situation where the model exhibits good performance on training data but performs badly on unseen data. Most importantly, limited studies focus on the performance of the discriminative DL models on multi-classification of X-ray images using the COVID-19 dataset. Rather, the focus has been on binary classifications of the COVID-19 X-ray. In addition, there are limited studies on multiclassification of pandemics and infectious as most of the existing studies focus on binary classification.

4. Methodology

The current study expands the investigation on COVID-19 detection beyond normal (healthy) vs infected (unhealthy) using chest X-ray images, the author explores a suitable X-ray dataset that has four different cases (tuberculosis, pneumonia, COVID-19 and normal) to investigate the robustness of discriminative DL model in the multiclassification of COVID-19 X-ray dataset. The study methodology phases align with some elements of CRoss Industry Standard Process for Data Mining (CRISP-DM) that proposed a systematic approach that categorizes operations into four levels of abstraction: phases, generic tasks, particular tasks, and processes integrated into the life cycle of data mining projects [28]. In this study also, we suggest a systematic approach that is categorized into four-phase approaches. The suggested framework for the empirical assessment of discriminative DL methods of COVID-19 detection using X-ray images is demonstrated in Figure 2. They suggest a four-phases approaches, namely, data acquisition, preprocessing, modelling, and training as well as evaluation phases. Where the suggested COVID-19 detection method uses chest X-ray images as input. Three DL-based models were taken into consideration, as was previously mentioned: CNN, RNN, and MLP.

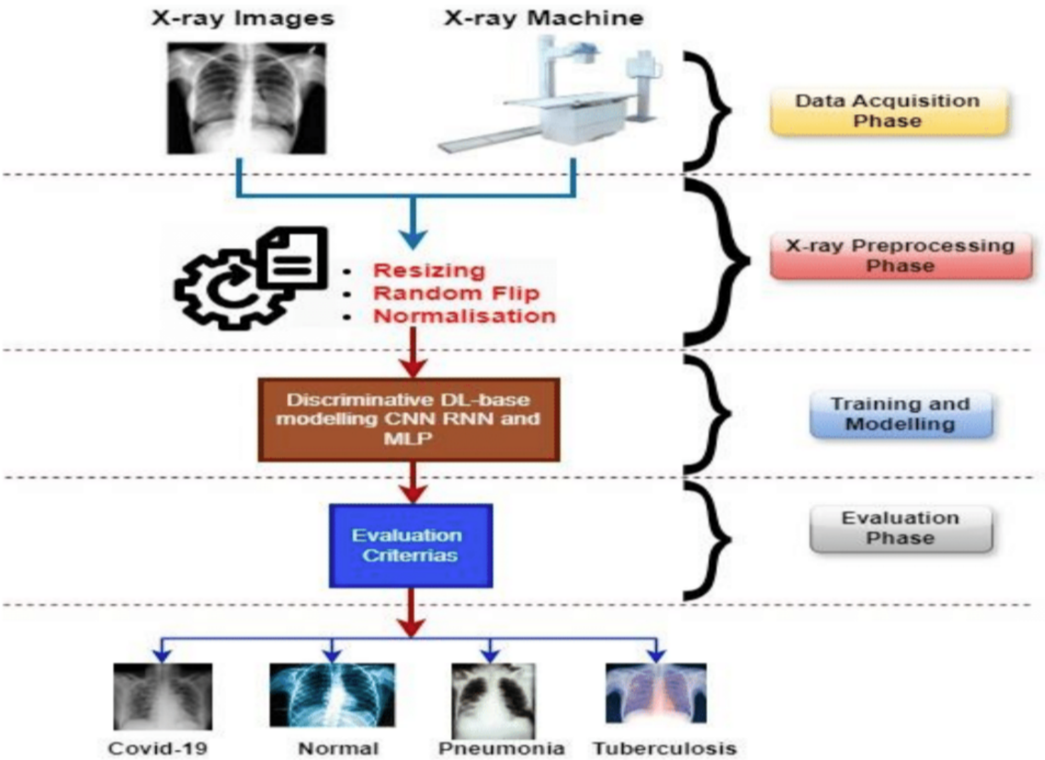

Figure 2: Framework for the empirical assessment discriminative DL methods COVID-19 detection using X-ray images.

4.1. Acquired Dataset and Description

The data for this study is a secondary X-ray image dataset from a popular repository named, Kaggle, https://www.kaggle.com/datasets/jtiptj/chest-xray-pneumoniacovid19tuberculosis. It was titled Chest X-Ray (Normal/Pneumonia/Covid-19/Tuberculosis). The content of the dataset is structured into three (3) folders (train, test, val) and contains subfolders for each image category or class (Normal, Pneumonia, COVID-19, and Tuberculosis), this makes it useful for multi-classification studies. The creation of the dataset was inspired by AI applications, specifically to detect and classify these diseases from X-ray images. A total of 7135 x-ray images are present. Figure 3 presents eight (8) samples of the acquired dataset. Images 1, 4, 5, and 8 show pneumonia infection, image 2 shows tuberculosis infection, images 3 and 7 are normal chest X-ray images while image 6 shows the COVID-19 infection dataset. The websites from where the X-ray images were taken asserted that they had addressed the ethical issues related to obtaining and utilizing the images.

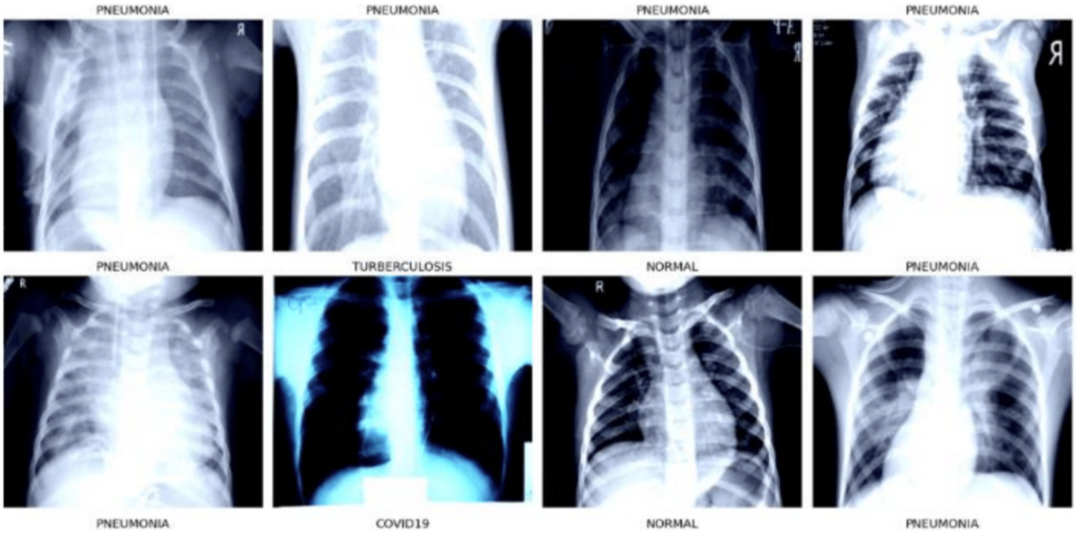

Figure 3: Sample X-ray dataset acquired.

4.2. Data Pre-processing

The images acquired exhibit inconsistencies due to the dataset being obtained from multiple origins employing diverse equipment and attributes. Therefore, it is necessary to do pre-processing on the data before feeding it to the model for further processing. To make the input chest X-ray images compatible with the three DL-based models considered for empirical assessment in this study, images were first scaled to 224 × 224 pixels. However, to enhance the algorithm's ability to generalize, this research refrains from using substantial pre-processing processes. Therefore, three pre-processing steps were used for the training process:

4.2.1. Random flip

Random flip is an example of augmentation, it was used to generate variants of x-ray image dataset for this study. This is medical data so not much augmentation should be done to avoid misrepresenting the data and having false results. The random flip seems to be something easy that can be done without much effect.

4.2.2. Resizing

A consistent 224 x 224-pixel resizing is needed and applied to every image.

4.2.3. Normalization

Min-max normalization was employed to ensure consistent intensity across all photos. The intensity value of all images, in the range ([0.485, 0.456, 0.406]), was normalized to the intensity range of 224 × 224 following equation 2.

$$F_i^{new} = \frac{f_i^{old} - (F)}{max(F) - min(F)} \tag{2}$$

4.3. Training and Modelling

Various kernel functions were utilized in CNN, RNN, and MLP models to classify the acquired dataset for COVID-19 classification, employing a distinct approach. The models for COVID-19 detection were subsequently trained (fine-tuned) using chest X-ray images. The model was trained for 28 epochs. The initial optimizer used was the Adam Optimiser and the CrossEntropy loss was the loss function that was used for the training process. The architecture including a total of 23 layers was built for the investigation. Python programming was employed in the software development process, where the names of functions and parameters for the layers were directly specified in the respective sections for layer names and parameters. The presence of four dimensions in the first layer, known as the picture input layer, is a result of conducting trials involving various input photos of varying sizes as part of the study. Figures 4 - 6 show the parameters setting for the CNN, RNN, and MLP models' architecture respectively.

Layer (type)	Output Shape	Param #
Conv2d-1	[-1, 32, 224, 224]	896
MaxPool2d-2	[-1, 32, 112, 112]	0
BatchNorm2d-3	[-1, 32, 112, 112]	64
Conv2d-4	[-1, 64, 112, 112]	18,496
MaxPool2d-5	[-1, 64, 56, 56]	0
BatchNorm2d-6	[-1, 64, 56, 56]	128
Conv2d-7	[-1, 128, 56, 56]	73,856
MaxPool2d-8	[-1, 128, 28, 28]	0
BatchNorm2d-9	[-1, 128, 28, 28]	256
Conv2d-10	[-1, 256, 28, 28]	295,168
MaxPool2d-11	[-1, 256, 14, 14]	0
BatchNorm2d-12	[-1, 256, 14, 14]	512
Dropout2d-13	[-1, 256, 14, 14]	0
Linear-14	[-1, 512]	25,690,624
Linear-15	[-1, 4]	2,052

Total params: 26,082,052
Trainable params: 26,082,052
Non-trainable params: 0

Input size (MB): 0.57
Forward/backward pass size (MB): 34.84
Params size (MB): 99.50
Estimated Total Size (MB): 134.91

Figure 4: Parameters setting for the CNN model architecture.

4.4. Evaluation Criteria

The evaluation criteria proposed for the empirical assessment of the CNN, RNN, and MLP in the multiclassification of COVID-19 X-rays are Acc, Sen, Spe, F1 measure, and confusion matrix. This study utilizes various parameters, including false positive (FP), true positive (TP), true negative (TN), and false negative (FN), as well as dependent variables such as Acc, Sen, Spe, F-1 measure and confusion matrix. These variables were obtained through mathematical analysis of the aforementioned parameters and were used to evaluate the results. In this case, TP represents the frequency with which the real patient data is correctly identified as patients by classification. False positive (FP), on the other hand, refers to the instances where non-patient data is incorrectly classified as patient data. TN represents the frequency at which data that is not related to patients is incorrectly identified as not being related to patients due to the classification process. FN, however, involves categorizing the patient data as non-patient data in a similar manner. The mathematical definitions of Acc, Pre, Sen, Spe, F-1 measure and confusion matrix, values obtained using these parameters are given in equations 3 – 7. Spe, F-1 score, and Acc were computed using a threshold (cut-off) value of 0.5, this is according to [29, 30, 31].

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \tag{3}$$

$$Precision = \frac{TP}{TP + FP} \tag{4}$$

$$Sensitivity = \frac{TP}{TP + FN} \tag{5}$$

$$Sensitivity = \frac{FP}{TN + FP} \tag{6}$$

$$F1 - Score = 2 * \frac{Precision * Recall}{Precision + Recall} \tag{7}$$

```
==========================================================================================
================
Layer (type (var_name))                 Kernel Shape      Output Shape      Param #
Mult-Adds
==========================================================================================
================
COVIDImageRNN (COVIDImageRNN)           --                [32, 4]           --
--
├─RNN (basic_rnn)                       --                [32, 224, 128]    45,312
324,796,416
│    └─weight_ih_l0                     [128, 224]                          ├─28,672
│    └─weight_hh_l0                     [128, 128]                          ├─16,384
│    └─bias_ih_l0                       [128]                               ├─128
│    └─bias_hh_l0                       [128]                               └─128
├─Linear (FC)                           --                [1, 32, 4]        516
516
│    └─weight                           [128, 4]                            ├─512
│    └─bias                             [4]                                 └─4
==========================================================================================
================
Total params: 45,828
Trainable params: 45,828
Non-trainable params: 0
Total mult-adds (M): 324.80
==========================================================================================
================
Input size (MB): 6.42
Forward/backward pass size (MB): 7.34
Params size (MB): 0.18
Estimated Total Size (MB): 13.95
==========================================================================================
```

Figure 5: Parameters setting for the RNN model architecture.

5. Result and Discussions

This section presents the results attained by three types of Neural Network architectures (discriminative DL models) on the joint diagnosis of COVID-19, tuberculosis, and pneumonia. The task is conceptualized as a multiclass classification problem on chest X-ray images. We also discussed and analyzed the outcome of this investigation in this section. Three discriminative DL models are investigated in the COVID-19 X-ray image dataset. Section 5.1 presents the outcome of the investigation while Section 5.2 presents the analysis of the outcome.

5.1. Outcome of the discriminative DL model for multi-classification of COVID-19 X-ray

The investigation results are based on the multi-classification discussed in Section 5.1. Table 2 shows the performance of the three discriminative DL over the COVID-19 X-ray dataset earlier discussed in

```
==========================================================================================
==
Layer (type:depth-idx)                   Output Shape              Param #
==========================================================================================
==
COVIDImageMLP                            [32, 4]                   --
├─Linear: 1-1                            [32, 512]                 77,070,848
├─BatchNorm1d: 1-2                       [32, 512]                 1,024
├─LeakyReLU: 1-3                         [32, 512]                 --
├─Dropout: 1-4                           [32, 512]                 --
├─Linear: 1-5                            [32, 512]                 262,656
├─BatchNorm1d: 1-6                       [32, 512]                 1,024
├─LeakyReLU: 1-7                         [32, 512]                 --
├─Dropout: 1-8                           [32, 512]                 --
├─Linear: 1-9                            [32, 512]                 262,656
├─BatchNorm1d: 1-10                      [32, 512]                 1,024
├─LeakyReLU: 1-11                        [32, 512]                 --
├─Dropout: 1-12                          [32, 512]                 --
├─Linear: 1-13                           [32, 4]                   2,052
==========================================================================================
==
Total params: 77,601,284
Trainable params: 77,601,284
Non-trainable params: 0
Total mult-adds (G): 2.48
==========================================================================================
==
Input size (MB): 19.27
Forward/backward pass size (MB): 0.79
Params size (MB): 310.41
Estimated Total Size (MB): 330.46
==========================================================================================
```

Figure 6: Parameters setting for the MLP model architecture.

Section 5.2 in the empirical assessment study. Figure 7 depicted the Loss vs Epoch curve for training and validation.

Table 2
Performance of discriminative DL models in a multiclassification of COVID-19 X-ray

Discriminative DL models	Classes of infection	Accuracies	Loss	Pre	Sen	F1-measure	Specificity
CNN	COVID-19	0.9066	0.009944	1.0000	0.7736	0.8723	1.0000
	Normal			0.8978	0.8632	0.8802	0.9572
	Pneumonia			0.9214	0.9615	0.9410	0.9160
	Tuberculosis			0.7018	0.9756	0.8163	0.9767
RNN	COVID-19	0.6278	0.029787	0.9706	0.3113	0.4714	0.9985
	Normal			0.6218	0.3162	0.4193	0.9162
	Pneumonia			0.6099	0.9179	0.7329	0.3990
	Tuberculosis			0.6129	0.4634	0.5278	0.9836
MLP	COVID-19	0.7860	0.015943	0.8837	0.7170	0.7917	0.9850
	Normal			0.7853	0.5940	0.6764	0.9292
	Pneumonia			0.8036	0.9128	0.8547	0.7717
	Tuberculosis			0.5385	0.8537	0.6604	0.9589

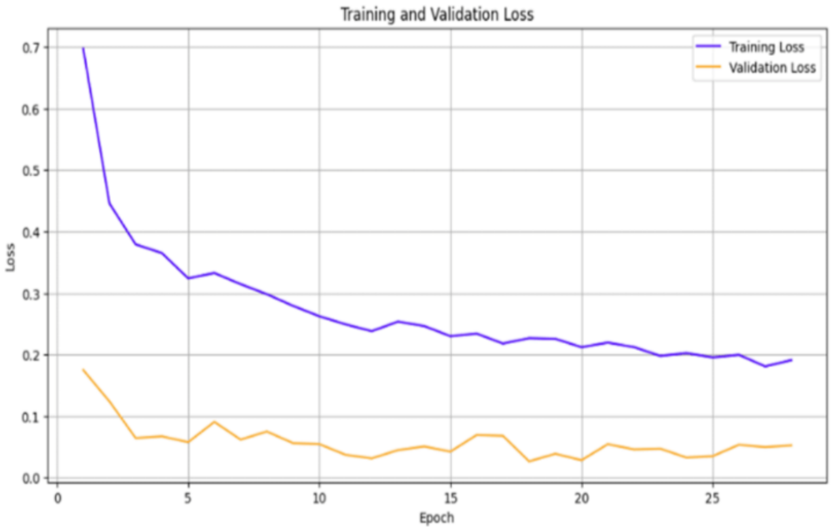

Figure 7: Loss vs Epoch curve for COVID-19 X-ray CNN training and validation.

5.2. Analysis of the outcome of discriminative DL model for multiclassification

These findings emphasize the capabilities of discriminative DL models in enhancing X-ray diagnosis and clinical environments. Upon analyzing the outcomes of the empirical assessment of the discriminative DL model for the multiclassification of COVID-19 X-ray presented in Table 2, it becomes evident that each of these discriminative DL models has areas of strength as well as weaknesses, depicting the average performance of the models, and it shows that CNN significantly outperformed the two other discriminative DL models. Figure 7 displays the curve depicting the relationship between loss and epoch for the training and validation of a COVID-19 X-ray CNN. The training loss exhibits a consistent decline with each subsequent epoch, suggesting that the model is progressively acquiring knowledge and effectively adapting to the training data.

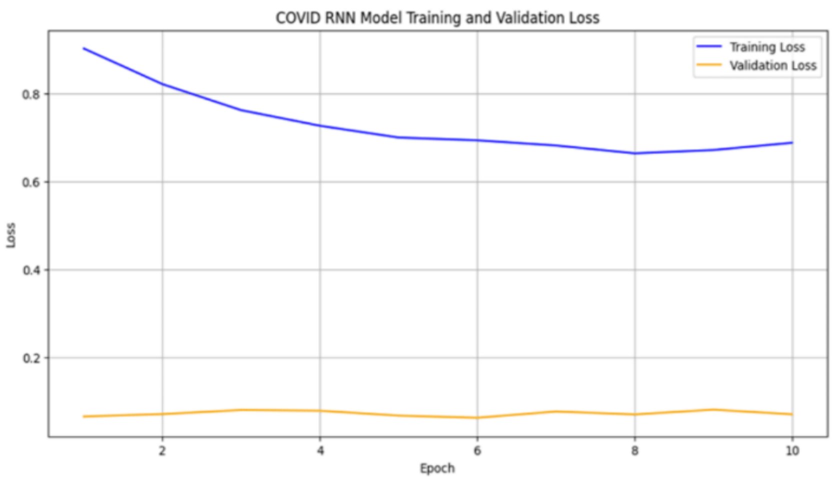

Figure 8: Loss vs Epoch curve for COVID-19 X-ray RNN training and validation.

There is a significant and rapid decrease in the early stages, followed by a steady slowdown, indicating that the model immediately grasps the fundamental patterns and then refines them. Similarly, the validation loss first lowers but then levels off after a few epochs. The validation loss remains very steady with slight fluctuations after the initial decline, suggesting that the model is effectively adapting to the unseen validation data. The persistent decrease in training loss demonstrates the successful acquisition of knowledge from the training data. The consistent validation loss indicates that the model is not

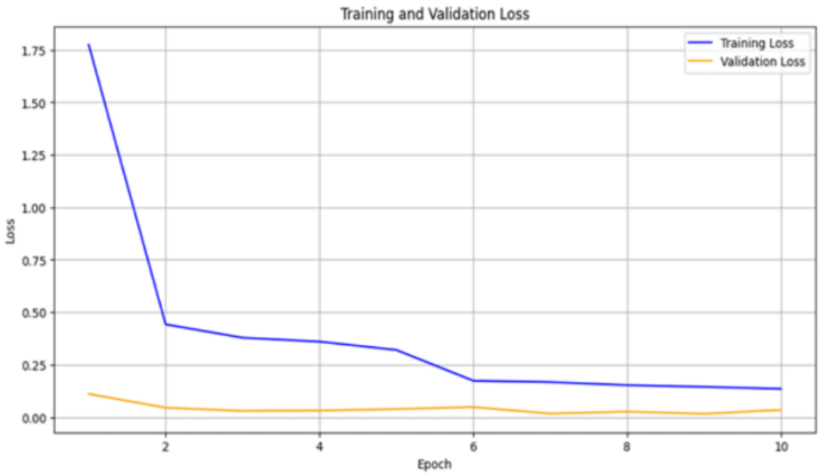

Figure 9: Loss vs Epoch curve for COVID-19 X-ray MLP training and validation.

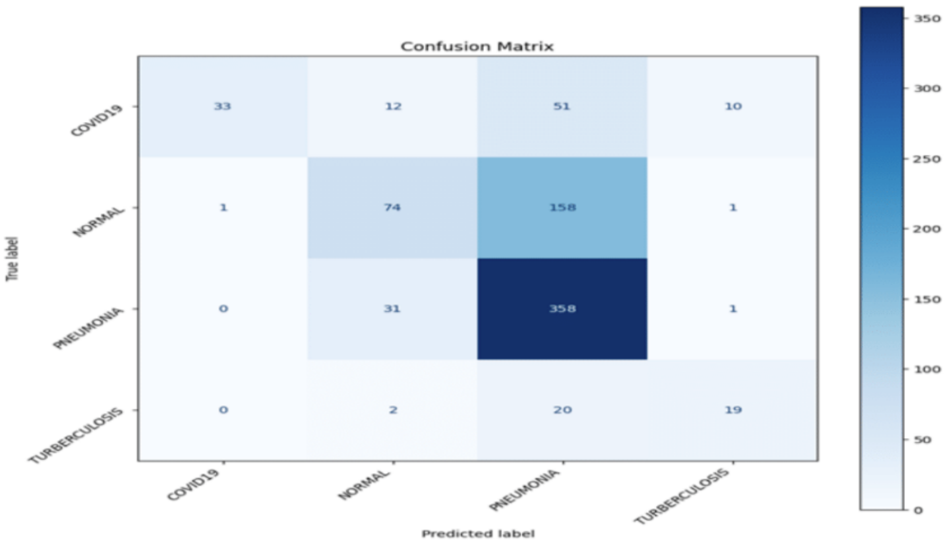

Figure 10: CNN confusion matrix.

suffering from overfitting and is effectively adapting to new data.

The curve illustrating the correlation between loss and epoch for the training and validation of the COVID-19 X-ray RNN is shown in Figure 6. The training loss consistently declines from one epoch to the next, indicating that the model is gaining knowledge and improving its performance on the training data. The fall in the second graph is less pronounced than in the first graph, indicating a positive trend in learning progress. The validation loss has a reasonably low value in comparison to the training loss, indicating that the model is not suffering from overfitting and is effectively generalizing to the validation data. The validation loss exhibits slight changes but maintains a generally constant trend. The training loss exhibits a decreasing trend, but towards the end, it shows a little increase, maybe indicating the onset of overfitting. It is something that should be observed in future periods. The low and stable validation loss signifies a commendable level of generalization performance.

Figure 8 displays the curve depicting the relationship between loss and epoch for the training and validation of the COVID-19 X-ray MLP. The first training loss exhibits a significant value, but swiftly diminishes within the initial epochs, suggesting that the model promptly grasps the essential patterns inside the training data. Following the first decline, the training loss consistently decreases, indicating that the model is gradually enhancing its alignment with the training data. The validation loss has an initial low value and maintains a consistently low level throughout the epochs, with minimal variations.

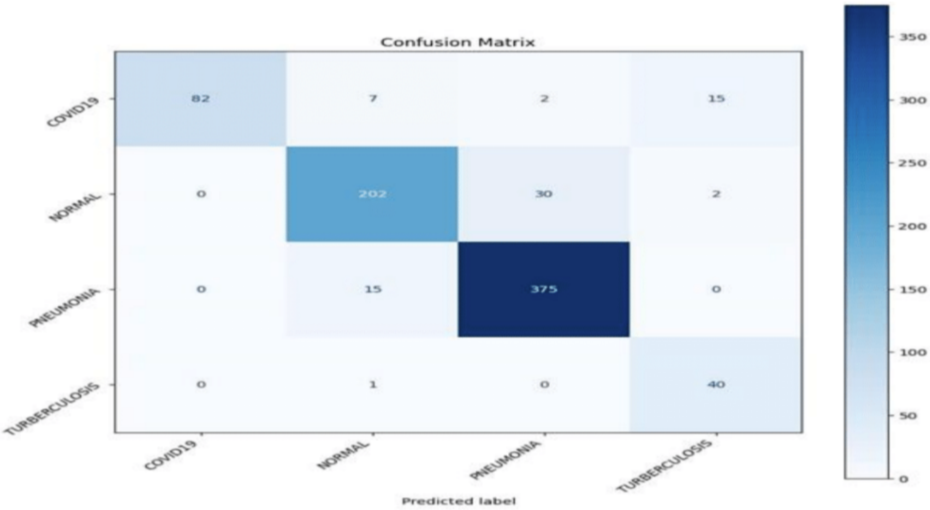

Figure 11: RNN confusion matrix.

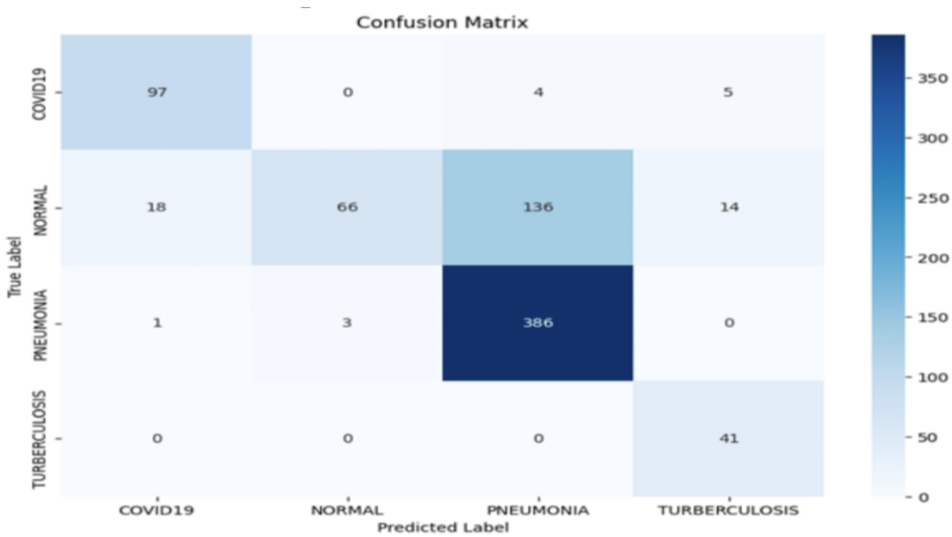

Figure 12: MLP confusion matrix

The model's validation loss demonstrates stability and a low value, indicating that it is effectively generalizing to the validation data and avoiding overfitting. The sharp and rapid decrease in training loss indicates that the model is quickly acquiring the fundamental characteristics of the data. The consistent and little validation loss seen during the training phase suggests successful generalization and the absence of overfitting.

Figures 10 - 12 are the confusion matrixes for CNN, RNN, and MLP respectively in this multiclassification study. The models demonstrate rapid acquisition of features and accurate X-ray image classification after only a few epochs of training. In addition, this slightly parallels the CNN model. This means CNN and the MLP are the best at the classification problem. But in all, CNN proved to be an outstanding model among the three discriminative DL explored in this study.

6. Conclusion and Future Scope

The goal of the research was to empirically evaluate the strengths and weaknesses of discriminative DL models in a multiclassification task using a dataset of COVID-19 X-rays, with the aim of achieving accurate results for disease diagnosis. This study expands the investigation of discriminative DL models over the COVID-19 X-ray dataset beyond binary classification tasks to multiclassification tasks due to the evolution of related infectious diseases. The detection models built using CNN, RNN, and MLP models were able to establish that the discriminative DL models can perform multiclassification tasks using COVID-19 X-ray datasets (tuberculosis, pneumonia, COVID-19, and normal) which are complex traits and pose difficulties for radiologists to interpret. Different preprocessing techniques such as data normalization, resizing, and random flip flop were used. The high accuracies achieved indicate that the discriminative DL models can identify unique features in the COVID-19 X-rays, enabling the deep networks to accurately differentiate between the images. These trained models may effectively reduce the workload of radiologists and other medical practitioners and increase the accuracy and efficiency of COVID-19 diagnosis. The CNN model demonstrates an effective and valuable approach for the multiclassification of diseases. This approach for identifying individuals with COVID-19 using chest X-ray pictures can be employed for preliminary screening purposes. To enhance the performance of the developed model, it is advisable to train the models on different datasets other than chest X-ray as this may be the subject of comparison to improve these models in the future, in a sense, these models in this research can be trained on an alternative dataset (CT scan, audio etc) to acquire a more profound understanding of the performance. Researchers can also explore the performance of several CNN architectures, such as AlexNet, ResNet50, and InceptionV3, and utilize metaheuristic optimization algorithms may be explored to set hyperparameter turning of DL network and improve the data exploration for such experiments in the future. These models and architecture can also be trained on a multimodal COVID-19 dataset to get a reliable diagnosis of the infection. A real-time scalable with an IoT framework COVID-19 diagnostic program can be created based on existing research findings to facilitate its implementation in medical practice.

Acknowledgments: Authors acknowledge the Centre of Excellence, University of Zululand for the support received for this work.

References

[1] A. AlMohimeed, H. Saleh, N. El-Rashidy, R. M. Saad, S. El-Sappagh, S. Mostafa, Diagnosis of covid-19 using chest x-ray images and disease symptoms based on stacking ensemble deep learning, Diagnostics 13 (2023) 1968.

[2] H. Florez, S. Singh, Online dashboard and data analysis approach for assessing covid-19 case and death data, F1000Research 9 (2020).

[3] W. W. H. Organization), Coronavirus disease (covid-19) pandemic, World Health Organization (2020).

[4] A. M. Ismael, A. Şengür, Deep learning approaches for covid-19 detection based on chest x-ray images, Expert Systems with Applications 164 (2021) 114054.

[5] M. Sevi, İ. Aydin, Covid-19 detection using deep learning methods, in: 2020 International conference on data analytics for business and industry: way towards a sustainable economy (ICDABI), IEEE, 2020, pp. 1–6.

[6] S. Kumari, E. Ranjith, A. Gujjar, S. Narasimman, H. A. S. Zeelani, Comparative analysis of deep learning models for covid-19 detection, Global Transitions Proceedings 2 (2021) 559–565.

[7] T. Anwar, S. Zakir, Deep learning based diagnosis of covid-19 using chest ct-scan images, in: 2020 IEEE 23rd international multitopic conference (INMIC), IEEE, 2020, pp. 1–5.

[8] J. Oguntoye, O. Awodoye, J. Oladunjoye, B. Faluyi, S. Ajagbe, E. Omidiora, Predicting covid-19 from chest x-ray images using optimized convolution neural network, LAUTECH Journal of Engineering and Technology 17 (2023) 28–39.

[9] A. Hayat, P. Baglat, F. Mendonça, S. S. Mostafa, F. Morgado-Dias, Novel comparative study for the detection of covid-19 using ct scan and chest x-ray images, International Journal of Environmental Research and Public Health 20 (2023) 1268.

[10] S. A. Ajagbe, M. O. Adigun, Deep learning techniques for detection and prediction of pandemic diseases: a systematic literature review, Multimedia Tools and Applications 83 (2024) 5893–5927.

[11] S. A. Ajagbe, M. O. Adigun, J. B. Awotunde, J. B. Oladosu, Y. J. Oguns, Internet of things enabled convolutional neural networks: applications, techniques, challenges, and prospects, IoT-enabled Convolutional Neural Networks: Techniques and Applications (2023) 27–63.
[12] I. H. Sarker, Deep learning: a comprehensive overview on techniques, taxonomy, applications and research directions, SN computer science 2 (2021) 420.
[13] H. Yasar, M. Ceylan, A novel comparative study for detection of covid-19 on ct lung images using texture analysis, machine learning, and deep learning methods, Multimedia Tools and Applications 80 (2021) 5423–5447.
[14] S. A. Ajagbe, J. B. Awotunde, H. Florez, Ensuring intrusion detection for iot services through an improved cnn, SN Computer Science 5 (2023) 49.
[15] F. Hardalac, H. Yaşar, A. Akyel, U. Kutbay, A novel comparative study using multi-resolution transforms and convolutional neural network (cnn) for contactless palm print verification and identification, Multimedia Tools and Applications 79 (2020) 22929–22963.
[16] J. Hernandez, H. Florez, An experimental comparison of algorithms for nodes clustering in a neural network of caenorhabditis elegans, in: Computational Science and Its Applications–ICCSA 2021: 21st International Conference, Cagliari, Italy, September 13–16, 2021, Proceedings, Part IX 21, Springer, 2021, pp. 327–339.
[17] Z. Tu, H. Talebi, H. Zhang, F. Yang, P. Milanfar, A. Bovik, Y. Li, Maxim: Multi-axis mlp for image processing, in: Proceedings of the IEEE/CVF conference on computer vision and pattern recognition, 2022, pp. 5769–5780.
[18] T. P.S, K. P, K. P, Image compression using mlp neural network, 2017.
[19] M. Weber, C. Renggli, H. Grabner, C. Zhang, Lossy image compression with recurrent neural networks: from human perceived visual quality to classification accuracy, arXiv preprint arXiv:1910.03472 (2019).
[20] K. Islam, L. M. Dang, S. Lee, H. Moon, Image compression with recurrent neural network and generalized divisive normalization, in: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2021, pp. 1875–1879.
[21] M. C. Arellano, O. E. Ramos, Deep learning model to identify covid-19 cases from chest radiographs, in: 2020 IEEE XXVII International Conference on Electronics, Electrical Engineering and Computing (INTERCON), IEEE, 2020, pp. 1–4.
[22] K. F. Haque, F. F. Haque, L. Gandy, A. Abdelgawad, Automatic detection of covid-19 from chest x-ray images with convolutional neural networks, in: 2020 international conference on computing, electronics & communications engineering (iCCECE), IEEE, 2020, pp. 125–130.
[23] L. Wang, Z. Q. Lin, A. Wong, Covid-net: A tailored deep convolutional neural network design for detection of covid-19 cases from chest x-ray images, Scientific reports 10 (2020) 19549.
[24] J. D. Arias-Londono, J. A. Gomez-Garcia, L. Moro-Velazquez, J. I. Godino-Llorente, Artificial intelligence applied to chest x-ray images for the automatic detection of covid-19. a thoughtful evaluation approach, Ieee Access 8 (2020) 226811–226827.
[25] H. Asgharnezhad, A. Shamsi, R. Alizadehsani, A. Khosravi, S. Nahavandi, Z. A. Sani, D. Srinivasan, S. M. S. Islam, Objective evaluation of deep uncertainty predictions for covid-19 detection, Scientific Reports 12 (2022) 815.
[26] M. Heidari, S. Mirniaharikandehei, A. Z. Khuzani, G. Danala, Y. Qiu, B. Zheng, Improving the performance of cnn to predict the likelihood of covid-19 using chest x-ray images with preprocessing algorithms, International journal of medical informatics 144 (2020) 104284.
[27] P. K. Chaudhary, R. B. Pachori, Automatic diagnosis of covid-19 and pneumonia using fbd method, in: 2020 IEEE International Conference on Bioinformatics and Biomedicine (BIBM), IEEE, 2020, pp. 2257–2263.
[28] C. E. Durango Vanegas, J. C. Giraldo Mejía, F. A. Vargas Agudelo, D. E. Soto Duran, A representation based on essence for the crisp-dm methodology, Computación y Sistemas 27 (2023) 675–689.
[29] S. A. Ajagbe, A. A. Adegun, P. Mudali, M. O. Adigun, Performance of machine learning models for pandemic detection using covid-19 dataset, in: 2023 IEEE AFRICON, IEEE, 2023, pp. 1–6.
[30] S. Akter, F. J. M. Shamrat, S. Chakraborty, A. Karim, S. Azam, Covid-19 detection using deep

learning algorithm on chest x-ray images, Biology 10 (2021) 1174.
[31] R. Rajagopal, Comparative analysis of covid-19 x-ray images classification using convolutional neural network, transfer learning, and machine learning classifiers using deep features, Pattern Recognition and Image Analysis 31 (2021) 313–322.