A novel Gray-Scale spatial exploitation learning Net for COVID-19 by crawling Internet resources

Mohamed E. ElAraby [a], Omar M. Elzeki [b,f,*], Mahmoud Y. Shams [c], Amena Mahmoud [d], Hanaa Salem [e]

[a] Faculty of Computers and Artificial Intelligence, Beni-Suef University, Beni-Suef 62511, Egypt
[b] Faculty of Computers and Information, Mansoura University, Mansoura 35516, Egypt
[c] Faculty of Artificial Intelligence, Kafrelsheikh University, Kafr El Sheikh 33511, Egypt
[d] Faculty of Computers and Information, Kafrelsheikh University, Kafr El Sheikh 33511, Egypt
[e] Faculty of Engineering, Delta University for Science and Technology, Gamasa, Egypt
[f] Faculty of Computer Science, New Mansoura University, Gamasa, Egypt

A B S T R A C T

Today, the earth planet suffers from the decay of active pandemic COVID-19 which motivates scientists and researchers to detect and diagnose the infected people. Chest X-ray (CXR) image is a common utility tool for detection. Even the CXR suffers from low informative details about COVID-19 patches; the computer vision helps to overcome it through grayscale spatial exploitation analysis. In turn, it is highly recommended to acquire more CXR images to increase the capacity and ability to learn for mining the grayscale spatial exploitation. In this paper, an efficient Gray-scale Spatial Exploitation Net (GSEN) is designed by employing web pages crawling across cloud computing environments. The motivation of this work are i) utilizing a framework methodology for constructing consistent dataset by web crawling to update the dataset continuously per crawling iteration; ii) designing lightweight, fast learning, comparable accuracy, and fine-tuned parameters gray-scale spatial exploitation deep neural net; iii) comprehensive evaluation of the designed gray-scale spatial exploitation net for different collected dataset(s) based on web COVID-19 crawling verse the transfer learning of the pre-trained nets. Different experiments have been performed for benchmarking both the proposed web crawling framework methodology and the designed gray-scale spatial exploitation net. Due to the accuracy metric, the proposed net achieves 95.60% for two-class labels, and 92.67% for three-class labels, respectively compared with the most recent transfer learning Google-Net, VGG-19, Res-Net 50, and Alex-Net approaches. Furthermore, web crawling utilizes the accuracy rates improvement in a positive relationship to the cardinality of crawled CXR dataset.

1. Introduction

The first confirmed unknown cause of pneumonia was found in Wuhan; the disease was named COVID-19 declared by WHO. The WHO confirmed that the Chinese epidemic of COVID-19 on the last of 2019 to be a community health crisis of global concern posturing a great threat to nations with the lack of healthcare organizations. The WHO reported that COVID-19 could be erased by primary recognition, lockdown, suitable caution, and carrying out a consistent tracking system by defining the COVID-19 characteristics [1]. The main step in fighting COVID-19 spread is observing the infected cases, letting those

contaminated search for suitable analysis and instant diagnosis care, besides separating them to decrease the virus spread. Generally, Reverse Transcription-enzyme Polymerase Chain Reaction (RT-PCR) is the most important selection tool used to recognize COVID-19 cases [2,3].

Artificial Intelligence (AI) tools are currently utilized to assist with the study of the virus to identify the correct antivirus and appropriate treatment with the least side effects. A model based on the forecast to deal with COVID-19, which was discovered in China, at the end of 2019, is being investigated [4–6].

Currently, the major challenge is not only classifying the CXR images, but also collecting the X-ray image datasets that are randomly

Table 1

A comparison between previously well-known articles and the current proposed approach in terms of methods, architecture, fine-tuning, hyperparameter optimization, number of classes, and accuracy.

Study	Methods	Architecture	Fine Tuning	Hyperparameter Optimization	Number of Classes	Accuracy (%)
Das et al. [19]	Transfer Learning	Xception	√	�	3	97.40
Singh et al. [20]	Deep Learning	CNN	√	√	2	92.55
Afshar et al. [21]	Deep Learning	Capsule Network CNN	√	√	4	95.70
Ucar and Korkmaz [22]	Transfer Learning	Deep Bayes-SqueezeNet	√	√	3	98.3
Khan et al. [23]	Transfer Learning	Xception	√	√	4	89.60
Elzeki et al. [24]	Deep Learning	CXRVN	√	√	2	96.70
Zhang et al. [25]	Deep Learning	CAAD	√	√	2	95.18
Elzeki et al. [26]	Transfer Learning	VGG-19	√	√	2	99.00
Narin et al. [27]	Transfer Learning	InceptionV3, ResNet50, ResNet101, ResNet152, Inception-ResNetV2	√	√	2	99.70
The Proposed Approach	Deep Learning + Crawler	GSEN	√ √	√ √	2 3	95.60 92.76

distributed in different sites such as GitHub, Mendeley, Kaggle, Google drive, and the most common websites that host the CXR or any dataset related to COVID-19. With the crawling procedure, the web pages of the World Wide Web are being methodically looks and gathered information using search engines. This is performed by indexing the web pages and updating their contents. Furthermore, the web crawlers can copy the webpages for the processing stages by a most common search engine by simultaneously indexing and updating the downloaded webpages to facilitate the searching strategies by the users [7,8].

Moreover, the variability and updated version of the dataset inside the host websites required simultaneous updates and adapted to handle the new instances and cases. Therefore, this work is extended to use a cloud computing environment with elastic web crawling to manage and crawl the X-ray images from different sites and resources to provide enough X-ray images in diagnosis and classification issues as well as provide the updated versions of the images in the websites. The crawling process is a continuous process that focuses on enriching a system with valid input based on seeking a set of certain criteria. COVID-19 is a trend changeable and mutated virus, and its effects are always changed over time and in a continuous change rather than the other diseases. Also, COVID-19 mutation leads to lockdown and threats to human lives which requires healthcare organizations to discover new strains of the corona virus as a response to the multiple coronavirus strains variants. Our attention is on designing a system that handles and adapts to these continuous variants. Thus, the proposed classifier has to be adapted and enhanced continuously by updating its dataset and retraining the classifier, from this point was the intention for the crawler. The crawler is proposed to be responsible for fetching and continuously updating the dataset from not only one source but also multiple sources that can be added to make our classifier is listening and triggered by these sources continuously. Moreover, the cloud is one of the computation paradigms and environments characterized by the resilience in the computation resource allocation that is needed for crawling multiple sources of datasets, this requires varied processing and storage, and coverage or availability geographically. Ultimately, we can itemize the necessity and advantages like the following:

- Adapting the classifier with the scans of the newly discovered cases,
- Continuous mutation of the virus thus discovering new or changed effects that might appear in new shapes in the scan,
- Adapting the dataset automatically without manual action interferences,
- Covering more than one source of datasets by minimal efforts, just adding the source seeds in a configurable way,
- The dataset of the classifier is a key factor in its design, so an adaptive dataset leads to the adaptive classifier,
- Leveraging the cloud computing feature in process automation,

- Designing the crawler over the cloud to erase infrastructure hassles such as settlement, preparation, scaling, operating, and monitoring.

Classification as a utility tool for many healthcare applications especially the COVID-19 pandemic is widely presented to diagnosis the patients based on image processing of CT and/or CXR images. The major motivation of this work is answering the question "How can crawling and classification help COVID-19 defense? To answer this question there are two interconnected axes, the first one the gathering and collecting the data that are commonly imbalanced and updated constantly and this required a crawling procedure. The second axis is the classification of the collected images as a utility to diagnosis the infected COVID-19 patients.

The main contribution of the proposed paper is shown as in the following points:

1. Adapting the crawled updated CXR COVID-19 images datasets using a web crawler-based cloud environment.
2. Designing a novel Gray-Scale Spatial Exploitation Net (GSEN) to detect infected COVID-19 cases easily.
3. Optimizing the hyperparameters of GSEN by using Stochastic Gradient Descent (SGD) Optimizer.

The organization of the paper is the literature survey investigated in Section (2). While in Section (3) the proposed method includes a web COVID-19 crawler and the utilized deep learning architecture. The results, comparative evaluation, and discussion are introduced in Section (4). Finally, the conclusion and future directions are clarified in Section (5).

2. Related work

Collecting different datasets with different locations is essential to track the variability and availability of the characteristics of this COVID-19 dataset. There are many different approaches based on AI to categorize and diagnose the infected cases that carry the virus. These approaches used CXR images to monitor and investigate the change of lung structure and chest depending on the analysis of X-ray images. The results prove the ability of X-rays to investigate and identify the infected persons. Furthermore, CT scans of the chest are utilized to clear more details inside the chest and the number and volumes of the batches. PCR tests are also utilized to determine the infected cases, but still, different problems are realized due to the instability of the results obtained [9–12].

COVID-19 data are being distributed across many data resources that contain low-quality data during inconsistent data structures [13,14]. Rao et al. [15] recommended a framework for collecting data and the possibility to identify COVID-19 cases. The data collected can be

employed to help in diagnosis and discover COVID-19 infected cases. While the progression of occurrences suggests that the virus might be transferred by a symptomless carrier, Hassanien et al. [16] present an algorithm for classifying COVID-19 X-Ray images that contain 40 contrast-enhanced lung X-rays with a fixed size 512 × 512 given that fifteen normal and the remaining 25 images are positive COVID-19 lungs using CXR images. Shinde et al. [17] presented a review for Predicting COVID-19 cases and the recent attempts of these prototypes for monitoring this pandemic.

Augmentation of datasets may be an effective solution for a limited number of collected datasets and to be adept with the applied DNN approach [3]. Pneumonia recognition for the applied CXR images with fine-tuned transfer learning is presented by Loey et al. [18]. They utilized Generative Adversarial Networks (GAN) to augment a limited number of COVID-19 CXR images. Furthermore, they classify the images using the most common transfer learning Alexnet, VGG19, Googlenet, and Resnet50.

Traditionally, different studies have been introduced to detect COVID-19 cases employing CXR images since the initial public release of COVID-19 datasets, several of which have leveraged variants of the COVID-19 to conduct such studies. It is important to note that as new patient cases are continuously added and made publicly available regularly, the COVID-19 dataset continues to evolve. To review the previously established research, a comparison is conducted in Table 1 regarding the used classifiers and obtained accuracy. Das et al. [19] presented an extreme version of inception called Xception that can be validated based on the applied fine-tuning of the weights of networks. In Singh et al.[20] DCNN with hyperparameter values are applied to ensure that adaptive multi-objective differential evolution is fine-tuning to classify COVID-19 and normal cases. The COVID-CAPS architecture was developed by Afshar et al. [21] using a capsule CNN network to identify bacterial, normal, viral, and COVID-19 cases, with the architecture's accuracy reaching 95.70 percent. SqueezeNet architecture is designed and fine-tuned for diagnosing COVID-19 cases using additive neural networks. Ucar and Korkmaz [22] present Bayesian optimization hyperparameters for classifying normal, pneumonia, and covid-19 cases. Khan et al. present a transfer learning depending on the Xception architecture that is used to classify four classes with 89.60 percent accuracy [23].

Moreover, Elzeki et al. [24] proposed a lightweight covid-19 chest x-ray network called CXRVN to classify two and three classes using a generative adversarial network for imbalanced applied COVID-19 datasets. The confidence-aware anomaly detection (CAAD) model is presented by Zhang et al. [25]. Moreover, a transfer learning model based on VGG-19 to classify CXR imbalanced dataset images for specified COVID-19 patients is presented by Elzeki et al. [26]. Narin et al. [27] presented transfer learning based on InceptionV3, ResNet50, ResNet101, ResNet152, Inception-ResNetV2 with an average accuracy reached to 99.70% to classify two classes of CXR images. As investigated in Table 1 the related current work is based on CXR images concerning the applied study, the methods used, the architecture, the fine-tuning, the hyperparameters, the number of classes, and the accuracy.

In this research, we are motivated to tackle some important concerns that are encountered by researchers in detecting CXR images. We crawl different datasets that updated the enrolled chest X-ray image simultaneously. The proposed architecture is based on a web CXR image crawler over a cloud computing environment then GSEN are applied to classify CXR images of the COVID-19 cases is. The proposed GSEN architecture not only classifies the CXR images but also extracts the most significant features of the images. Moreover, we visualize CXR images to investigate the characteristics of the enrolled images that might be helpful to correctly classify the normal and COVID-19 cases.

3. Proposed method

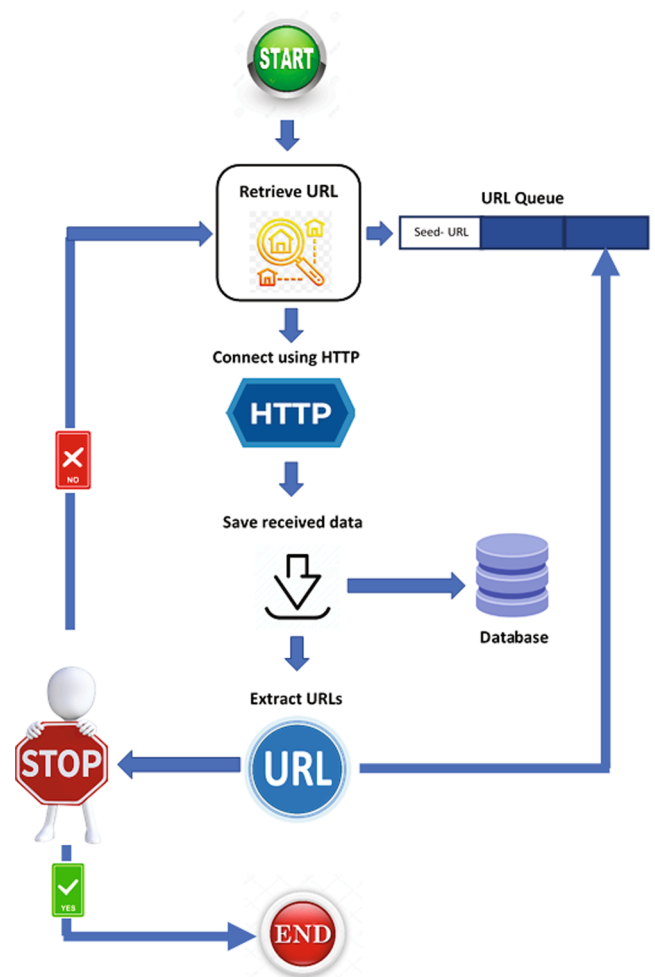The basic conceptions of web COVID-19 crawler and deep learning

Fig. 1. The procedures of the designed crawler.

based on GSEN are being discussed in the following section.

3.1. Web COVID-19 crawler

The crawler is the backbone and the stand of many systems such as general search engines, specific search engines, anti-virus, knowledge banks, and many other systems. On the Internet, Crawler looks for information that it assigns to certain groups, then lists and catalogs it so that the crawled data can be collected and checked. Until a crawl is started, the activities of certain computer applications need to be identified. Then the crawler executes these instructions instantly. An index is created with the results of the crawler, which can be accessed via output tools. The competence a crawler can learn from the Internet depends on simple commands. Fig. 1 is a flowchart that describes the procedures of the designed crawler process [28–30].

As the main component, it is responsible for gathering the base of data for the others component in the system. Crawler's performance is the main factor that the main research challenge to enhance and optimize. The performance is measured by execution time, scalability, availability, and flexibility. These factors are highly related to the computation paradigm that is used for running. To build a web crawler, there are two required phases. The first phase is finding the seed page. The seed page is the page that we will provide to the crawler. Ideally, it is only needed to provide the crawler with the link for the seed page to start the crawler off on its merry way. But depending on the approach needs, a list may be provided as a dictionary, data frame, etc. In such cases, all the same ideas and concepts are applied, and the approach is
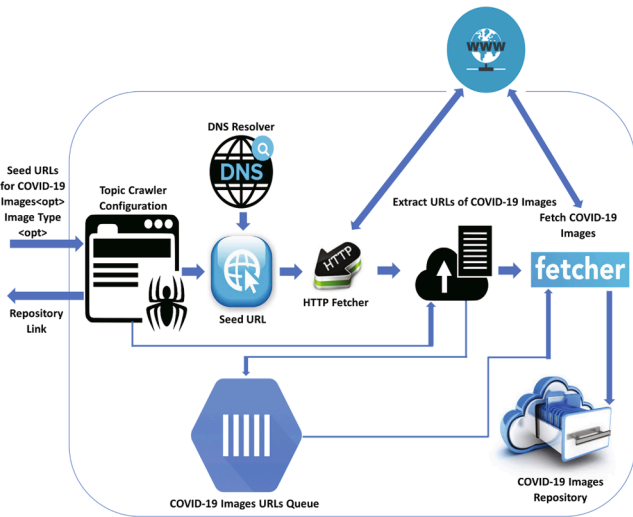
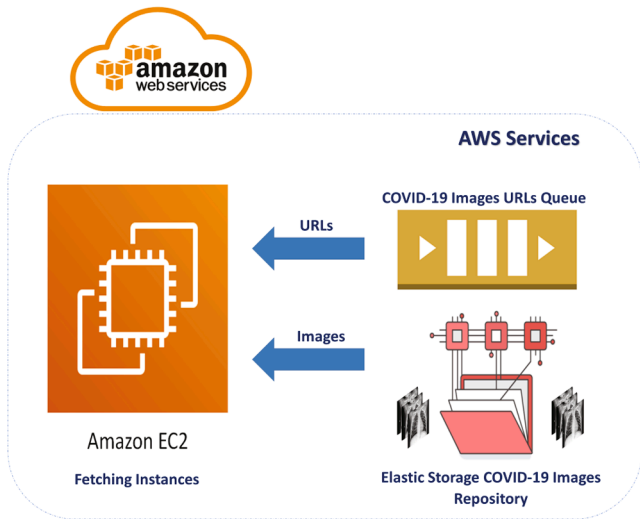**Fig. 2.** Steps of crawling CXR COVID-19 images.

Fig. 3. CXR COVID-19 image crawler architecture based on cloud computing.

modified to meet the problem.

The second phase is examining the speed page. Examining the maker page, the list of all brands is arranged alphabetically on a single web-page. In most cases, when a page has several items (search results), they are split and distributed across multiple web pages and stored as separate links, which can be found on a nag page tab. The crawler will have to be able to find all such nav page links. In the case of the maker page, all the information is listed on a single page. Algorithm 1 summarizes the steps of the proposed COVID-19 crawler as follows.

Cloud computing is one of the advanced computation paradigms that have tremendous capabilities and features. Its features proved that it is the most proper computation paradigm for this research topic. Cloud computing is characterized by elasticity which means scalability and expandability, flexibility, reliability, region covering, and pay-as-use or on-demand.

The proposed architecture for the crawler focused on two phases; the crawling steps, and the mounting on the cloud computing resources. In the first phase of crawling steps as shown in Fig. 2, the initial inputs of crawling are entered as configurations such as seed URLs for the sites that contain images of covid19 or referral links to pages that contain images for any level of depth. Moreover, the additional configurations like the number of images and types that are targeted of crawling. These configurations are the guidance of the crawling steps. The seed URLs are saved as initial values in the URLs queue. The HTTP fetcher retrieves the content of the referenced page, the parser extracts the included URLs from that page to start adding them to the queue, and the image fetcher retrieves the image URL from the queue to retrieve its image and retailer it in the COVID-19 image repository for each URL in the queue. The second phase of mounting the crawler on cloud computing is shown in Fig. 3 The extracted URLs are stored in a queue service which is structured as a first-in-first-out pattern. The queue size is configured and scaled on-demand in the runtime without affecting the operations. The extracted images from the crawled pages are stored in an elastic storage service. Elastic storage is storage that can be scaled up or down on-demand in runtime. Moreover, elastic storage can be shared or replicated over different geographical regions. This storage can be archived or backed up on demand.

The EC2 is an elastic cloud computing service. It is a virtual machine that is configured firstly, prepared with crawl APIs, and made as a virtual machine image to be instantiated many times on demand in the runtime. EC2 is the link between the queue and elastic storage, which reads the next URL from the queue and marks this URL with a lock to block other instances from working on it. The newly extracted URLs are inserted at the end of the queue and stored in the extracted images in the elastic storage.
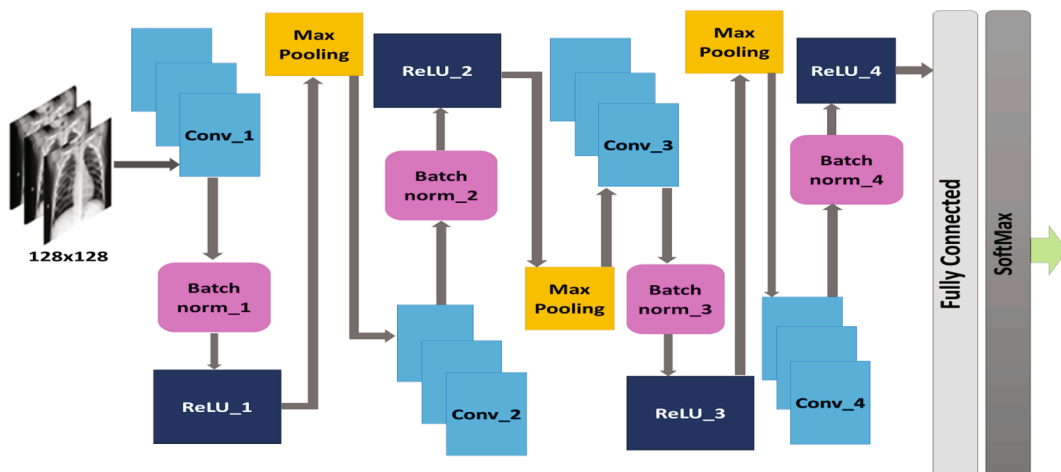

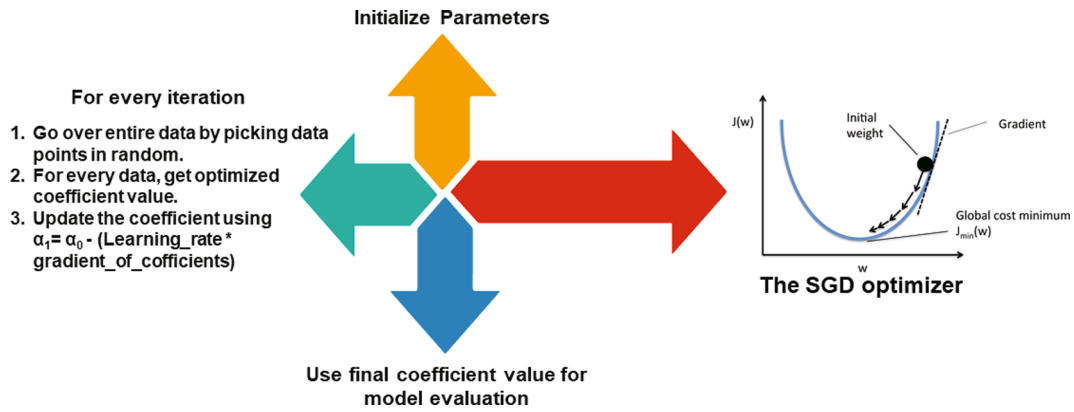**Fig. 4.** The proposed GSEN structure for CXR COVID-19 images.

Fig. 5. The architecture of the SGD optimizer.

3.2. Deep learning-based GSEN architecture

Deep learning is one of the commonly utilized techniques that offer state-of-the-art precision. It has an effective role in image processing especially medical applications that have not been accomplished before [31]. Deep learning in health care addresses a wide variety of topics, stretching on or after cancer screening and tracking disease to tailored treatment recommendations. Today, different sources of data radiological imaging including CT, X-ray, MRI scans, pathology as well as genomic sequences recently, have provided large quantities of data to physicians [32–35]. DCNN is considered as one of the most crucial techniques that are used to extract the features of the enrolled images as well as a promising classifier of that feature. Generally, DCNN consists of three layers that are convolution, pooling, and full connection. The convolution layer is used as a feature extractor layer to classify COVID-19 X-ray input images. The pooling layer is a downsample layer that controls the over-fitting problem. The fully connected layer is used for the classification of normal/abnormal cases.

In this paper, a system based on GSEN was built to be used in COVID-19 x-ray image classification to normal and abnormal categories. The percentage of COVID-19 virus existence can be visualized and tracked based on the expectation–maximization algorithm and SoftMax activation function. The architecture of the proposed system is shown in Fig. 4. Initially, the main difference described as "Any Gray-Scale Spatial Exploitation Net (GSEN) is a specific CNN architecture" that have major pros for Chest images can be itemized as:

a. The GSEN is a *simple* architecture that consists only of four convolutional blocks each is composed of a single convolutional, Batch Normalization, ReLU activation function, and Max Pooling. On other hand, the traditional CNN consists of Multiple Complex Blocks like Convolutional, Residual, and much more.
b. The GSEN is a *fast* net, which is a lightweight architecture in memory space, initialization stage, number of parameters, number of blocks, and the structure of each block. Hence, the GSEN are designed specifically for grayscale images.
c. The GSEN architecture is an *efficient* net, that deals more efficiently in processing the grayscale images without any transformation, scaling, or floating-point operations.
d. The GSEN architecture is the cross *computable* net that can execute grayscale processing on CPU/GPU while the traditional CNN must execute only on GPU.

The proposed system is executed using four convolutional layers, by which each layer covers a batch normalization, and Rectified Linear Unit (ReLU) activation function. SGD optimizer is applied to the proposed GSEN architecture to learn the parameters of the features that are convoluted such that SGD conveys to zero means and the variance are normalized [36,37].

The Rectified Linear Unit (ReLU) is the function that is activated based on the hidden layer such that the convoluted input feature "x" of the ReLU is mathematically introduced as in Eq. (1).

$$ReLU(x) = \begin{cases} x, & x \geq 0 \\ 0, & x < 0 \end{cases} \tag{1}$$

The normalization of the batch has been determined based on the mean and variance for the enrolled features "x" as shown in Eqs. (2), and (3). The mean is the expected value of $\times$ such that:

$$\mu = E[x_i] \tag{2}$$

The variance is determined by: -

$$\sigma^2 = E(x_i - \mu] \tag{3}$$

The normalized value of $\times$ is given by Eq. (4) as follows:

$$\dot{x} = \frac{x_i - \mu}{\sqrt{\sigma^2 + \in}} \tag{4}$$

The score vectors of the enrolled CXR COVID-19 images symbolized by S^{w_x}, such that the probability of the scored values is given in Eq. 5.

$$P_i = \frac{e^{S^{w_x}}}{\sum_{i=1}^{n} e^{S^{w_n}}} \tag{5}$$

3.3. SGD optimizer

The redundancy of data is very valuable to the routine utilization of optimizers like SGD. Meanwhile, the learning rate varies from comparatively major to minor that well known as the scheduling process. Consequently, it is necessitating to estimate the parameters to be convergence by futzing the parameters ultimately [38]. Practically, the main usage of the SGD is to reduce the cost function given a great number of training sets that lead to very expensive procedures [39]. The minimization of the SGD cost is investigated in Equations (1), and (2) given the simple linear in Equation (6).

$$Y = a_0 + a_1 X + x \tag{6}$$

Where X, Y represents the independent, dependent variables, respectively, x is the error. a_0, a_1 are the interception, and the slope of SGD optimizer, respectively.

For a given parameter a_0, a_1 the cost function (s) can be determined as in Equation (7) as follows.

$$s(a_0, a_1) = \frac{1}{2n} \sum_{i=0}^{n} \left(h_a(x^i) - y^i \right)^2 \tag{7}$$

As shown in Fig. 5, the main purposes of the SGD optimizer are to minimize the cost function such that $minimize_{a_0, a_1} s(a_0, a_1)$ [40].

Table 2

Five instances average running time in seconds for multiple VMs based on the Amazon Cloud Computing environment.

Number of Images	50	100	150	200	250	300
1-Instance	102.176	102.7766	307.882	407.488	493.718	587.12
2-Instances	53.424	106.774	167.226	220.944	270.002	321.488
3-Instances	34.998	73.844	113.006	150.532	183.17	214.632
4-Instances	30.626	63.892	91.91	124.92	146.094	176.756
5-Instances	22.954	49.828	78.106	103.008	124.896	124.848

The parameters are initially utilized, at that time altered iteration of the gradient descent yields the parameters to the global minima over the data path called the trajectory. In various technologies, Batch Gradient Decent is accumulated by summing up, which results in time consumption and, as a result, higher complexity of the chosen platform. The SGD looks at one sample or a small subset or mini-batch example and make a progress in the space of parameters to the global minima. Instead of the SGD receipts the batch simply and directly to achieve the global–local minimum, the SGD cost is adapted by randomly shuffling the dataset of the trained CXR COVID-19 images such that the number of trained instances is calculated. As a result, the features are selected depending on the summation accumulated by the first iteration in that direction and decrease gradually towards the global minima. Therefore, to optimize the accuracy of the system, it is preferred to use as many numbers as possible of iterations to get optimized results of the learning process [40–42].

4. Experimental results and discussion

In this paper, we designed the proposed architecture to perform two scheduling processes that are sequentially carried out. The first process is the crawling process which includes the collection of the CXR image from different datasets. Afterward, the crawled CXR COVID-19 images are utilized to process both feature extraction and classification stages based on using GSEN architecture.

4.1. Crawling experiment

In this step, the crawling experiment is concentrating on the performance assessment with the crawling architecture overcloud environment. Accordingly, the execution time is determined in seconds to ensure the duration executed to crawl the images with minimum elapsed time. The experiment was performed five times indicates that the elapsed time varies related to the applied Virtual Machines (VM) that are crawled, such that the 1-VM, 2- GSEN VMs, 3-VMs, 4-VMs, and 5-VMs are utilized. On the other hand, the total number of the crawling pages and the extracted CXR COVID-19 images are fixed. In this work, the proposed crawler architecture is implemented using Java language concluded IDE development based on Oxygen and JDK-8. Furthermore, we utilized a central database that is designed to crawl the extracted Uniform Resource Locators URLs from different distributed web pages. To signboard the URL status, each URL record has a flag and an identifier that might be executed or not.

Over the cloud environment, the experimental results were performed such that each crawler instance is operating in a distinct VM based on Amazon Elastic Cloud Computing EC2. At this time, the Amazon EC2 provides various instance selection types; nevertheless, we have instantiated type t2.micro. The T2.micro category includes Giga-Byte memory, Intel Xeon processors, clocked at up to 3.3 GHz, and EBS memory. In addition, the performance of the t2.micro network varies from low to medium and the internet speed during the current test ranges from 140 to 210 Mbps. We created instances in the Northern Virginia (US East) region. Additionally, the database is designed to centralize all robot instances and is built on top of the Amazon Relational Database Service (RDS). Typically, RDS services are simple to extend relational databases in the cloud and allow you to focus on your
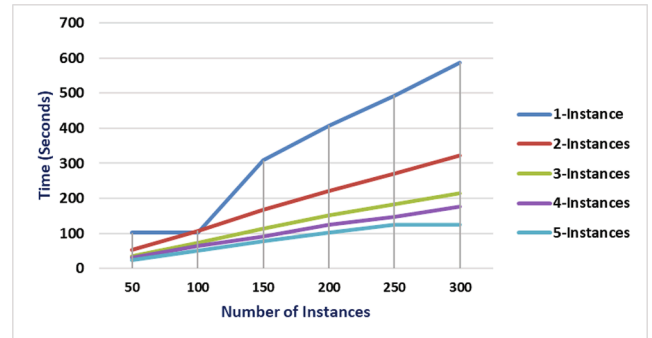

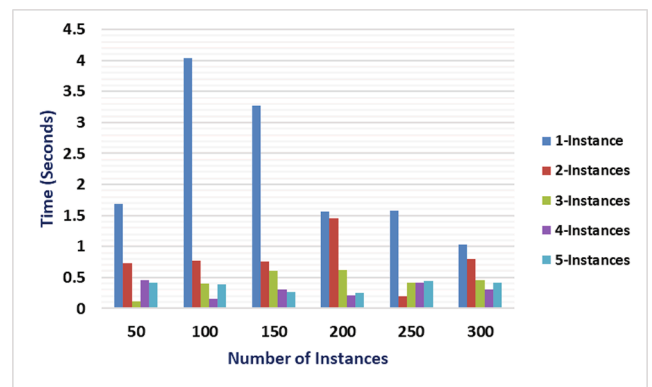**Fig. 6.** The running time operates on numerous VMs on Amazon Cloud.

Fig. 7. Five instances STD running time in seconds for multiple VMs based on Amazon Cloud Computing environment.

business and applications. Therefore, we used the MySQL DB instance, whose type is DB.t2.micro. Additionally, DB.t2.micro has a vCPU, 1 GiB of RAM, and general-purpose database (SSD) storage. Therefore, the origination zone is the same region and region as the robot versions in Northern Virginia (US East).

4.1.1. Statistical analysis

In this work, the experimental results consist of two important variables, the first one is the instances numbers and the second is the number of crawled CXR images. Consequently, we investigated the proposed crawler architecture using five instances. The instances are increasing from one instance to five instances. In the proposed, we determine statistical parameters, these parameters include the average, and standard deviation, min, and max, of the crawling time (seconds) obtained in the training phase. Accordingly, the experimental results of this experiment are shown in Table 2 and Fig. 6 as follows.

Here we noticed that the average running time is increased concerning the number of instances. For example, in the case of five instances, the running time reached 124.848 s. Also, the standard deviation (STD) for running time is decreased concerning the number of instances. For example, in the case of five instances, the STD running time reached 0.417008393 s as shown in Fig. 7. Also, maximum and
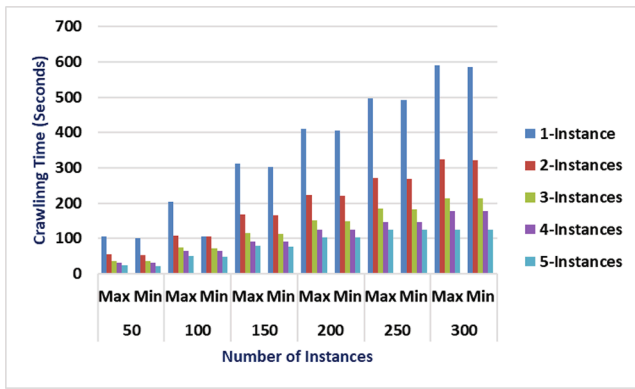
Fig. 8. Five instances maximum and minimum running time in seconds for multiple VMs based on Amazon Cloud Computing environment.

minimum running time in seconds for multiple VMs based on the Amazon Cloud Computing environment is represented in Fig. 8. This means that the proposed crawler architecture is incrementally varied with the number of CXR COVID-19 images as well as the number of instances.

5. Classification of the crawled dataset

5.1. PC properties

The experimental results for classification-based GSEN were performed using a package of software containing MATLAB-2020a established on a workstation Core i7, 16-RAM, NVIDIA 4G- GT 740 m GPU environment.

5.2. Dataset characteristics

Several publicly available datasets, especially in medical applications, recently exist on different websites, such as GitHub, Mendeley, Google Drive, and so on. Most of these websites' pages host and contain the latest COVID-19 images, whether X-ray and/or CT images that are constantly updated. In this work, we crawled different CXR COVID-19 images from different websites. We implemented the experimental

results using two different datasets. The first dataset is called MOMA dataset from Mendeley [43] consists of 603 chest X-ray images. The 603 chest X-ray images were composed of 221 COVID-19, 382 non-COVID images, samples are shown in Fig. 9. 382 non-COVID images consist of 148 pneumonia and 234 normal cases. The second dataset is collected using our proposed crawler from Kaggle and Github which consists of COVID and non-COVID images as well. We randomly crawled different updated versions from the mentioned sites and the collected datasets were trained, tested, and validated based on fixed-parameter sets illustrated as follows.

5.3. Parameter's setting

In this section, all experiments were performed based on fixed hyperparameter values. The applied values of the proposed GSEN architecture are investigated in Table 3.

5.4. Measures of performance

To assess the performance of the proposed GSEN architecture, the confusion matrix containing sensitivity, precision, specificity, accuracy, and F1 score are utilized to ensure the reliability of the proposed GSEN architecture as shown in Eqs. ((8)–(12)) as follows.

$$Recall = Sensitivity = TP/(TP + FN) \tag{8}$$

$$Specificity = TN/(TN + FP) \tag{9}$$

$$Precision = TP/(TP + FP) \tag{10}$$

$$Accuracy = (TP + TN)/(TP + TN + FP + FN) \tag{11}$$

Table 3
List of hyperparameter values of the proposed GSEN architecture.

Hyper-Parameter	Value
Learning-Rate	0.01
Batch-Size	32
Momentum	0.8
Weight-Decay	0.001
Maximum no. of iterations (2-classes)	200
Maximum no. of iterations (3-classes)	600
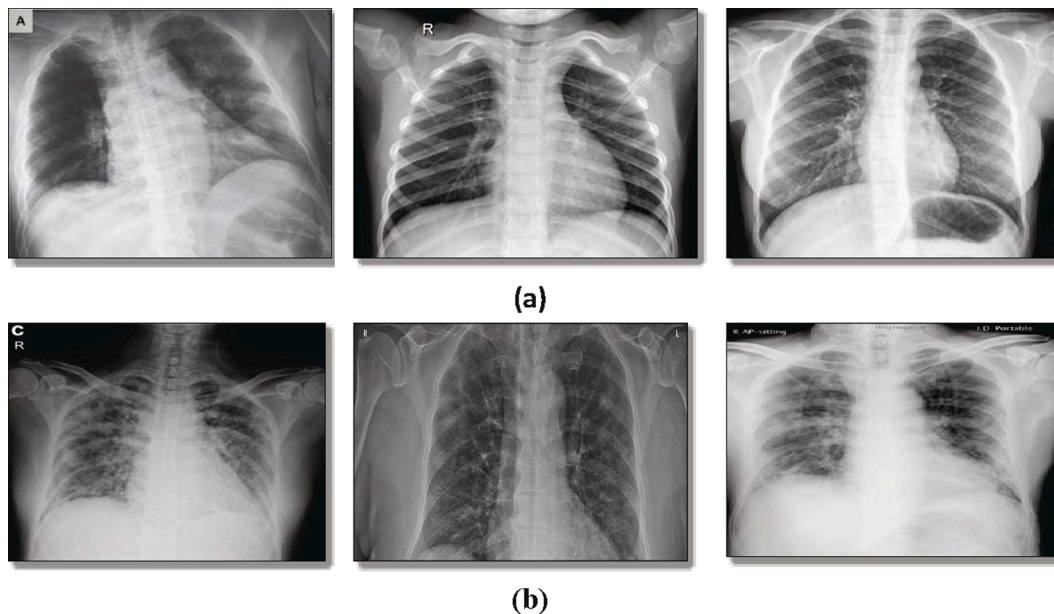
Fig. 9. Samples of MOMA dataset [43] (a) Normal X-Ray images, (b) infected X-Ray COVID-19 images.

	Normal	COVID-19	
Normal	**57** 50.40%	**4** 3.50%	**93.40%** 6.60%
COVID-19	**1** 0.9%	**51** 45.10%	**98.10%** 1.90%
	98.30% 1.70%	**92.70%** 7.30%	**95.60%** 4.40%

Fig. 10. The confusion matrix of the tested 113 CXR COVID-19 images.

$$F1 - score = 2TP/(2TP + FP + FN) \tag{12}$$

Where TP: True Positive, TN: True Negative, FP: False Positive, FN: False Negative.

5.5. First experiment: Crawling of two classes (COVID-19 or normal)

In this scenario of experiments, we used 113 tested X-ray images crawled from different two sources of webpages. The first source collected by Moony was founded on the Kaggle.com website as in the following link "https://www.kaggle.com/paultimothymooney/chest-xray-pneumonia". They collected 5,863 CXR COVID-19 images with two labels "Pneumonia" and "Normal", we utilized only 234 out of 1341 normal cases. The second source collected from both Faizan (GitHub-website) founded in the following link "https://github.com/smfai200/Detecting-COVID-19-in-X-ray-images", and Bachir (Kaggle-website) as in the following link "https://www.kaggle.com/bachrr/covid-chestxray", and we utilized only 221 COVID-19 cases. Therefore 455 COVID-19/Normal are successfully collected, and we tested 113 out of 455 images (~25%) of the total collected CXR COVID-19 images were tested. The confusion matrix of the first scenario of an experiment is shown in Fig. 10.

As investigated in the above Fig. 10, the accuracy of the proposed GSEN architecture of the crawled 113 CXR COVID-19 images reached 95.60% which indicates the reliability of the proposed GSEN architecture based on the crawled images over the cloud computing environment. The accuracy and loss of this experiment are shown in Fig. 11 by which we found that after 200 iterations the accuracy is stabled to be 95.60% with minimum loss value. Table 4 investigate briefly the final results of the proposed GSEN for the 113 crawled 113 tested CXR images. We found that the validation accuracy was 95.58% with a minimum loss of 0.5797 after 200 iterations and 20 epochs. The elapsed time was 22 min and 56 s indicates that the proposed architecture is more speed to recognize and classify CXR COVID-19 images with higher accuracy and minimum loss. The histogram of the normal and COVID-19 cases of the 113 tested images was shown in Fig. 12 such that 58 normal and 55 COVID-19 cases were successfully classified.

Table 4
The validation results of the crawled 113 CXR using GSEN architecture.

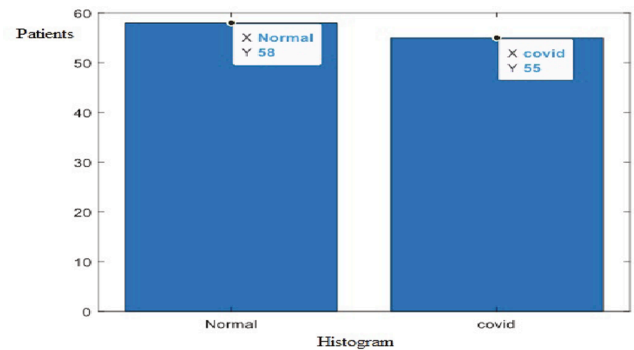| Epochs	Iteration	Time Elapsed	Validation Accuracy	Validation Loss
20	200	22 min and 56 sec	95.58%	0.5797

Fig. 12. The histogram of the classified CXR includes 58 normal and 55 COVID-19 cases.
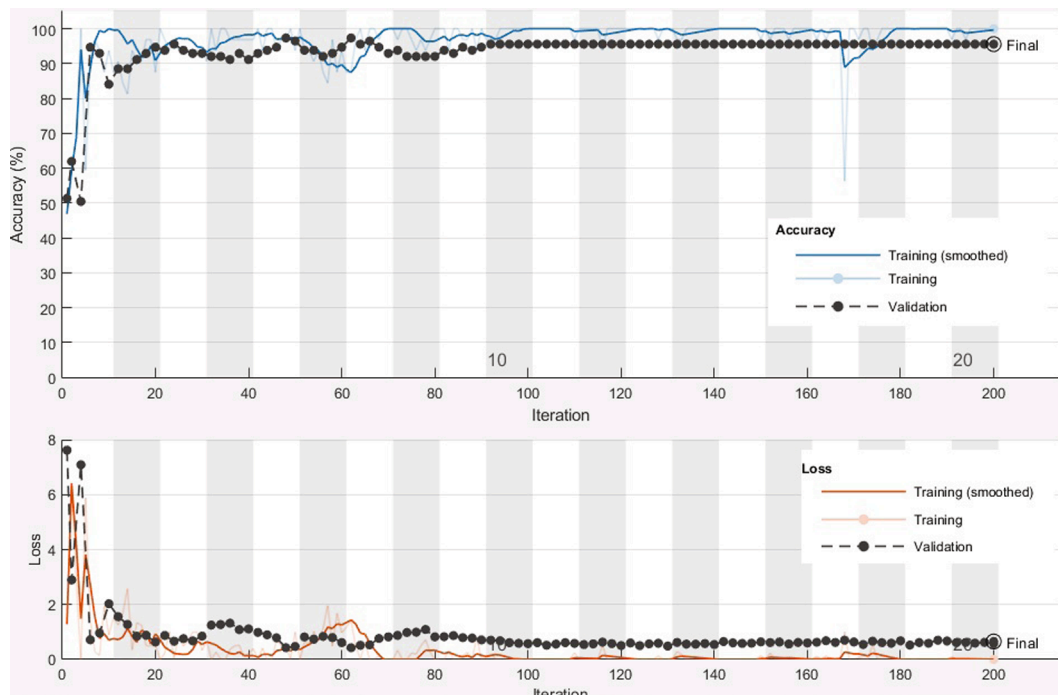

Fig. 11. The resulting accuracy and loss of the proposed GSEN after 200 iterations.

Table 5
The confusion matrix results were obtained measuring sensitivity, specificity, accuracy, precision, and F1 score in the first experiment.

GSEN -architecture	Dataset-1/Crawler-1
Sensitivity	93.40
Specificity	98.10
Accuracy	95.60
Precision	98.30
F1 score	95.80

The numerical results of the confusion matrix are shown in detail in Table 5 such that the proposed architecture accuracy was 95.60 % of the classified Normal/COVID-19 cases which proves the system robustness and reliability for the crawled CXR images.

5.6. Visualization

Deep dream evidence is a feature visualization technique in deep learning that synthesizes images that strongly activate network layers. The visualized features are essential to understand and diagnosis the proposed GSEN. This helps to visualize CXR COVID-19 images using the features represented by each convoluted layer.

The deep dream shows different insights at different levels and blocks whereas using different pyramid scales. The pyramid scale can help generate more informative images for layers at the beginning of the network. Fig. 13 The visualization of the classified CXR COVID-19 images including normal, and COVID-19 cases, (a) Activation channel for 2 CXR COVID-19 images, (b) CXR COVID-19 images at different pyramid

scales. Fig. 13 is a complex diagram representing the investigation of deep dream in the different layers in a different block. We notice the informative details increase at a higher level of pyramid scale and last layers that are interested in the contents of images.

5.7. Experiment 2: (Crawling of three classes (COVID-19, normal, and Pneumonia))

To boost the architecture's ability to classify more than two classes such as COVID-19, Normal, and Pneumonia, we expand the experimental results to classify three class labels. In this scenario of an experiment, we used 603 crawled CXR images found in [43], and the samples of the collected CXR COVID-19 images were shown in Fig. 14. We tested 25% of the overall 603 CXR COVID-19 images such that 150 CXR COVID-19 images were crawled and tested. The confusion matrix of the proposed GSEN based on the three-class labels are shown in Fig. 15. The classification accuracy of the three-class labels is 92.67%. Table 6 determines the validation accuracy, Number of epochs, Number of iterations, elapsed time, and validation loss of the tested 150 CXR COVID-19 images based on GSEN architecture.

The confusion matrix of the three-class labels is investigated in Table 7 which indicates the system reliability in terms of precision, recall, F1-score, and accuracy for the tested 150 CXR COVID-19 images that are crawled and classified using GSEN architecture. The comparison between the proposed GSEN and recent transfer learning approaches using the same crawled CXR COVID-19 images are shown in Fig. 16. The comparison study is performed using the same hyperparameter values in Table 3. Furthermore, we utilized the same number of CXR COVID-19 images which are 113 Normal, and COVID-19 cases as well as 150

(a)

(b)

Fig. 13. The visualization of the classified CXR COVID-19 images including normal, and COVID-19 cases, (a) Activation channel for 2 CXR COVID-19 images, (b) CXR COVID-19 images at different pyramid scales.
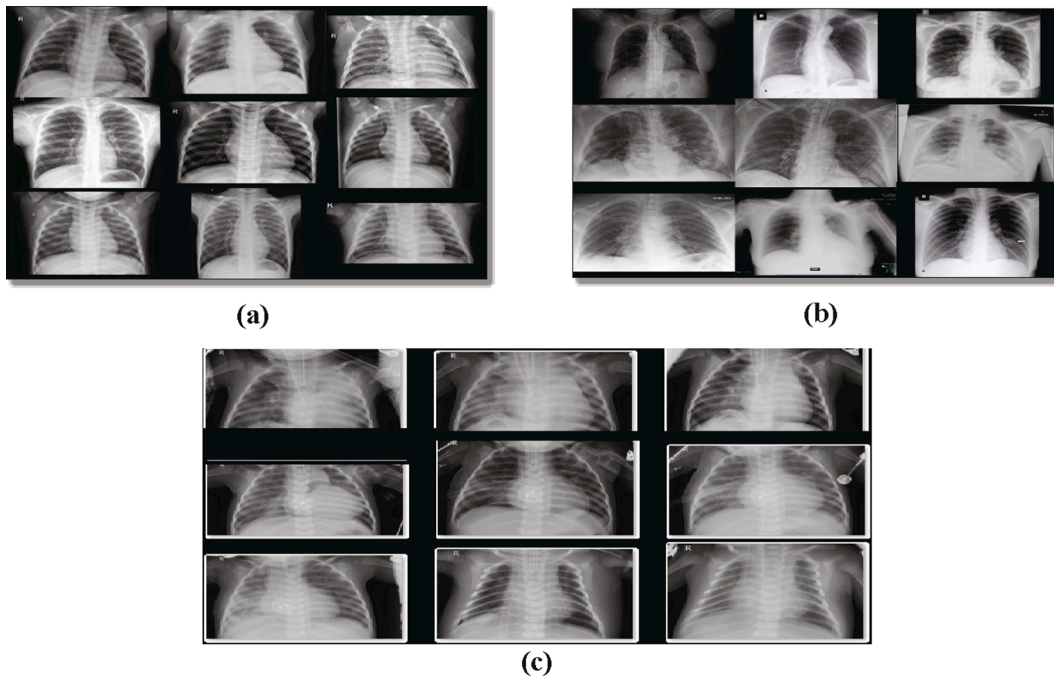
Fig. 14. (a) The samples of the crawled 603 CXR COVID-19 images including, (a) Normal Cases, (b) COVID-19 Cases, and (c) Pneumonia dataset found in Shams et al. [43].

Fig. 15. The confusion matrix of the 150 CXR COVID-19 images tested with the three-class labels (Pneumonia, Normal, and COVID-19).

Table 6
The validation results of the crawled 113 CXR using GSEN architecture.

Epochs	Iteration	Time Elapsed	Validation accuracy	Validation loss
60	600	91 min and 18 sec	92.67%	0.21543

Table 7
The precision, recall, F1-score, and accuracy of the tested 150 CXR COVID-19 images using GSEN for three class labels normal, pneumonia, and COVID-19.

3-Class Labels	Precision	Recall	F1-score	Accuracy
Normal	93.10%	93.10%	93.10%	92.76%
Pneumonia	94.74%	92.31%	93.51%	
COVID-19	90.74%	92.45%	91.59%	

tested CXR COVID-19 images representing three class labels which are Normal, Pneumonia, and COVID-19 cases. The comparison study indicates that the proposed GSEN architecture is superior to other recent
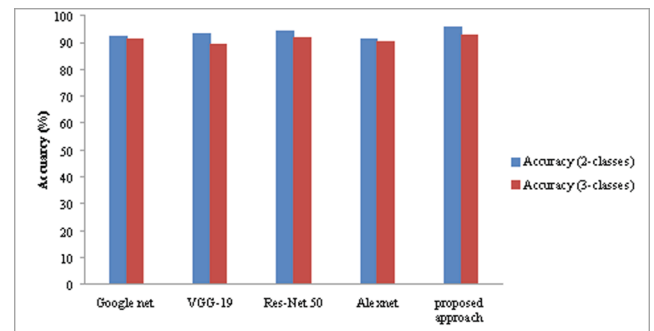

Fig. 16. The comparison study of the proposed GSEN and recent deep learning transfer learning approaches Googlenet, VGG-19, Res-Net 50, and Alexnet for both two and three class labels.

transfer learning approaches in both two and three class labels. For Fig. 16, using the same crawled dataset, we compared the proposed GSEN architecture with the most recent CNN models such as GoogleNet, VGG-19, Res-Net 50, and Alexnet for both two and three class labels. We utilized the same hyperparameter values in Table 3. The superiority of the proposed GSEN architecture results from a simple structure of the internal network. Furthermore, this architecture is specially designed to identify Gray X-ray images. The results obtained depend on the accuracy that is calculated from Equation 11. There is a slight enhancement of the proposed GSEN architecture compared with the GoogleNet and Res-Net 50. On the other hand, the loss error of the proposed GSEN were 0.5797, and 1.2264 for two and three classes respectively compared with the loss error of GoogleNet that reached the minimum values 0.2454 and 0.5873 for two and three class labels. The other compared VGG-19, Res-Net 50, and Alexnet achieved loss error 1.2545, 2.0054, 0.6987 for two class labels. While for three class labels the VGG-19, Res-Net 50, and Alexnet achieved loss error 2.0598, 3.9954, 1.2378 respectively.
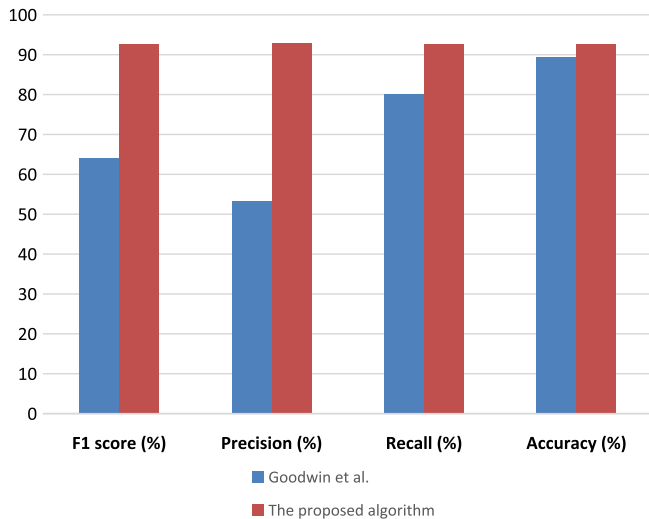
Fig. 17. CXR COVID-19 image diagnostic studies performance comparison for three-class labels.

5.8. Comparative analysis and discussion

As a result of collecting crawled CXR COVID-19 images from various dataset websites, the proposed GSEN architecture uses the highest number of CXR COVID-19 images and achieves better execution. The proposed GSEN architecture's accuracy of the crawler-based CXR COVID-19 images is 95.60 percent for two-class labels and 92.67 percent for three-class labels. The comparison between the proposed 3 classes COVID-19 CXR images based on crawling-GSEN and Goodwin et al. [44] based on 12 models ensembles shows that there are slightly enhanced results in terms of F1 score, precision, recall, accuracy as shown in Fig. 17.

In Fig. 17, the proposed GSEN architecture is compared with the Goodwin et al. [44] based on three-class labels COVID-19, Pneumonia, and Normal cases. They utilized the twelve ensembles model and the average accuracy of the Goodwin combined architecture reached 89.40% while the proposed architecture accuracy achieved 92.76%. On the other hand, in terms of the loss function, the proposed GSEN architectures achieved loss error 1.2264, while the Goodwin architecture achieved an error loss less than 0.4.

Nevertheless, the proposed approach employs a higher number of

CXR COVID-19 images. In general, the proposed approach outperformed well than other techniques. Table 8 contains information on the studies that used deep learning to diagnose CXR COVID-19 images. The comparative study [21,23,27,32,45–57] discussed the confusion matrices of the state-of-the-art approaches with different methods are presented in Table 8. We noticed that the proposed crawler GSEN approach is not the top result but on the other hand, the crawling of CXR images from different sites is still the major contribution of this work. This is because it provides updated versions of images simultaneously as well as an enhancement of the classification accuracies will be validated in using large-scale datasets.

6. Conclusion and future work

A web crawler is considered one of the main sources of dynamic and updated information. Focused web crawling is a more specialized crawler that focuses on a specific domain or topic of information, so this manuscript is designed based on topical web crawling to collect and update the data set from a specific source for a specific content that makes the proposed system is continuously up to date automatically. The variability and instability of the enrolled datasets uploaded to the host websites such as Kaggle, GitHub, Google Drive, Mendeley, and the other host websites are more confused for the researchers who are interested in diagnosis and prognosis CXR images. Currently, machine learning approaches, particularly regression and classification issues, are considered the most significant tools to fight COVID-19 spread. In this work, we presented an architecture based on crawling the updated CXR COVID-19 images from different websites using a cloud computing environment. Furthermore, the updated images are simultaneously adapted to be enrolled in the feature extraction and classification stages. The feature extraction and classification stages are based on GSEN to classify the input CXR images. The extracted feature can be handled based on a convoluted layer and the results indicate the robustness and superiority of the proposed system compared with the state-of-the-art. We proposed two experiments that demonstrate and prove the ability of the proposed GSEN using large-scale standard datasets. We used an elastic web crawler over cloud computing to collect different sources of datasets and the experimental results investigated that the classification accuracies were 95.60%, and 92.76% for 113 and 150 CXR COVID-19 images such that 113 CXR COVID-19 images are two class labels; COVID-19 and Normal cases, and 150 are three class labels Normal, Pneumonia, and COVID-19 cases, respectively. As presented in the experimental results, although the utilization of the limited number of

Table 8
CXR COVID-19 image diagnostic studies performance comparison for two-class labels.

Studies	Methods	Accuracy (%)	Sensitivity (%)	Specificity (%)	Recall (%)	Precision (%)	F1 score (%)
Narin et al. [27]	CNN– ensemble	94.00	–	–	–	–	–
Abbas et al. [45]	DCNN	–	97.91	91.87	–	93.36	–
Afshar et al. [21]	COVID-CAPS	95.70	90.00	95.80	–	–	–
Hemdan et al. [46]	COVIDX-Net	90.00	–	–	–	–	91.00
Khan et al. [23]	CoroNet	89.60	–	–	98.20	93.00	–
Maghdid et al. [47]	CNN-AlexNet	94.00	–	–	–	–	–
Xu et al. [48]	CNN ResNet18	86.70	–	–	–	–	–
Song et al. [49]	CNN-ResNet50	93.00	–	–	–	–	–
Zheng et al. [50]	CNN	–	90.70	91.10	–	–	–
Ghoshal et al. [51]	CNN	92.90	–	–	–	–	–
Wang et al. [52]	CNN	93.30	–	–	–	–	–
Apostolopoulos, et al. [53]	VGG-19	93.48	–	–	–	–	–
Ozturk et al. [32]	DarkCovidNet	87.02	–	–	–	–	–
Murugan et al. [54]	E-DiCoNet	94.07	98.15	91.48	85.21	98.15	91.22
Shi et al. [55]	iSARF (Random Forest)	87.90	90.70	83.30	–	–	–
Hall et al. [56]	3 Models Ensembled	91.24	78.79	93.12	–	–	–
He et al. [57]	DenseNet169 (Self-supervised Transfer Learning)	86.00	–	–	–	–	85.00
The proposed algorithm	Crawling + GSEN	95.60	93.40	98.10	–	98.30	95.50

CXR images to be classified using GSEN, it is expected that the accuracy will be increased for large-scale crawled datasets. For future direction, we plan to use CT images and study different updated cases of the COVID-19 x-ray image. In the future, the proposed approach would be used as transfer learning.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

References

[1] S.-Q. Deng, H.-J. Peng, Characteristics of and public health responses to the coronavirus disease 2019 outbreak in China, J. Clin. Med. 9 (2020) 575.

[2] T. Ai, Z. Yang, H. Hou, C. Zhan, C. Chen, W. Lv, Q. Tao, Z. Sun, L. Xia, Correlation of chest CT and RT-PCR testing for coronavirus disease 2019 (COVID-19) in China: a report of 1014 cases, Radiology 296 (2020) (2019) E32–E40.

[3] M.Y. Shams, O.M. Elzeki, M. Abd Elfattah, T. Medhat, A.E. Hassanien, Why Are Generative Adversarial Networks Vital for Deep Neural Networks? A Case Study on COVID-19 Chest X-Ray Images, in: Big Data Anal. Artif. Intell. COVID-19 Innov. Vis. Approach, Springer, 2020: pp. 147–162.

[4] Y.H. Bae, K. Park, Advanced drug delivery 2020 and beyond: perspectives on the future, Adv. Drug Deliv. Rev. (2020).

[5] P. Hassanzadeh, Nanotheranostics against COVID-19: From multivalent to immune-targeted materials, J. Controlled Release. 328 (2020) 112–126.

[6] N.A. Awad, A. Mahmoud, Improving reconstructed image quality via hybrid compression techniques, CMC-Comput. Mater. Contin. 66 (2021) 3151–3160.

[7] A. Arasu, J. Cho, H. Garcia-Molina, A. Paepcke, S. Raghavan, Searching the web, ACM Trans. Internet Technol. TOIT. 1 (1) (2001) 2–43.

[8] M.E. ElAraby, M.Y. Shams, Face retrieval system based on elastic web crawler over cloud computing, Multimed. Tools Appl. 80 (8) (2021) 11723–11738.

[9] F. Homayounieh, O. Holmberg, R.A. Umairi, S. Aly, A. Basevičius, P.R. Costa, A. Darweesh, V. Gershan, P. Ilves, D. Kostova-Lefterova, Variations in CT utilization, protocols, and radiation doses in COVID-19 Pneumonia: results from 28 countries in the IAEA study, Radiology. 298 (2021) E141–E151.

[10] Y.-L. Wan, U.J. Schoepf, C.C. Wu, D.P. Giovagnoli, M.-T. Wu, H.-H. Hsu, Y.-C. Chang, C.-T. Yang, W.-J. Cherng, Preparedness and best practice in Radiology Department for COVID-19 and other future pandemics of severe acute respiratory infection, J. Thorac. Imaging 35 (2020) 239–245.

[11] C.M. Hamper, F.N. Fleckenstein, L. Büttner, B. Hamm, N. Thieme, H.-M. Thiess, O. Scholz, F. Döllinger, G. Böning, Submillisievert chest CT in patients with COVID-19-experiences of a German Level-I center, Eur. J. Radiol. Open 7 (2020), 100283.

[12] H.G. Atakla, K. Condé, M.M.U.D. Noudohounsi, M.S.S. Dongmo, A.H. Garba, D. S. Houinato, F.A. Cissé, Interest of the thoracic scanner in the diagnosis of COVID-19: study of 35 cases in the Republic of Guinea, Pan Afr. Med. J. 35 (2020).

[13] D. Sha, Y. Liu, Q. Liu, Y. Li, Y. Tian, F. Beaini, C. Zhong, T. Hu, Z. Wang, H. Lan, Y. Zhou, Z. Zhang, C. Yang, A spatiotemporal data collection of viral cases for COVID-19 rapid response, Big Earth Data 5 (1) (2021) 90–111.

[14] D. Sha, A.S. Malarvizhi, Q. Liu, Y. Tian, Y. Zhou, S. Ruan, R. Dong, K. Carte, H. Lan, Z. Wang, A state-level socioeconomic data collection of the United States for COVID-19 research, Data 5 (2020) 118.

[15] A.S.R. Srinivasa Rao, J.A. Vazquez, Identification of COVID-19 can be quicker through artificial intelligence framework using a mobile phone–based survey when cities and towns are under quarantine, Infect. Control Hosp. Epidemiol. 41 (7) (2020) 826–830.

[16] A.E. Hassanien, L.N. Mahdy, K.A. Ezzat, H.H. Elmousalami, H.A. Ella, Automatic x-ray covid-19 lung image classification system based on multi-level thresholding and support vector machine, MedRxiv (2020).

[17] G.R. Shinde, A.B. Kalamkar, P.N. Mahalle, N. Dey, J. Chaki, A.E. Hassanien, Forecasting models for coronavirus disease (COVID-19): a survey of the state-of-the-art, SN Comput. Sci. 1 (2020) 1–15.

[18] M. Loey, G. Manogaran, N.E.M. Khalifa, A deep transfer learning model with classical data augmentation and cgan to detect covid-19 from chest ct radiography digital images, Neural Comput. Appl. 1–13 (2020).

[19] N.N. Das, N. Kumar, M. Kaur, V. Kumar, D. Singh, Automated deep transfer learning-based approach for detection of COVID-19 infection in chest X-rays, Irbm (2020).

[20] D. Singh, V. Kumar, V. Yadav, M. Kaur, Deep neural network-based screening model for COVID-19-infected patients using chest X-ray images, Int. J. Pattern Recognit. Artif. Intell. 35 (2021) 2151004.

[21] P. Afshar, S. Heidarian, F. Naderkhani, A. Oikonomou, K.N. Plataniotis, A. Mohammadi, Covid-caps: a capsule network-based framework for identification of covid-19 cases from x-ray images, Pattern Recognit. Lett. 138 (2020) 638–643.

[22] F. Ucar, D. Korkmaz, COVIDiagnosis-Net: deep bayes-SqueezeNet based diagnosis of the coronavirus disease 2019 (COVID-19) from X-ray images, Med. Hypotheses 140 (2020), 109761.

[23] A.I. Khan, J.L. Shah, M.M. Bhat, CoroNet: a deep neural network for detection and diagnosis of COVID-19 from chest x-ray images, Comput. Methods Programs Biomed. 196 (2020), 105581.

[24] O.M. Elzeki, M. Shams, S. Sarhan, M. Abd Elfattah, A.E. Hassanien, COVID-19: a new deep learning computer-aided model for classification, PeerJ Comput. Sci. 7 (2021) e358.

[25] J. Zhang, Y. Xie, G. Pang, Z. Liao, J. Verjans, W. Li, Z. Sun, J. He, Y. Li, C. Shen, Viral Pneumonia Screening on Chest X-rays Using Confidence-Aware Anomaly Detection, IEEE Trans. Med. Imaging. (2020).

[26] O.M. Elzeki, M. Abd Elfattah, H. Salem, A.E. Hassanien, M. Shams, A novel perceptual two layer image fusion using deep learning for imbalanced COVID-19 dataset, PeerJ Comput. Sci. 7 (2021).

[27] A. Narin, C. Kaya, Z. Pamuk, Automatic detection of coronavirus disease (covid-19) using x-ray images and deep convolutional neural networks, Pattern Anal. Appl. 24 (3) (2021) 1207–1220.

[28] Z. Xuefang, H. Zhanxiao, Design and implementation of a web crawler for images [J], J. Nanjing Norm. Univ. Eng. Technol. Ed. 4 (2008).

[29] M.E. ElAraby, H.M. Moftah, S.M. Abuelenin, M.Z. Rashad, Elastic web crawler service-oriented architecture over cloud computing, Arab. J. Sci. Eng. 43 (12) (2018) 8111–8126.

[30] M.E. ElAraby, S.M. Abuelenin, H.M. Moftah, M.Z. Rashad, A new architecture for improving focused crawling using deep neural network, J. Intell. Fuzzy Syst. 37 (1) (2019) 1233–1245.

[31] S. Sarhan, A.A. Nasr, M.Y. Shams, Multipose face recognition-based combined adaptive deep learning vector quantization, Comput. Intell. Neurosci. 2020 (2020) 1–11.

[32] T. Ozturk, M. Talo, E.A. Yildirim, U.B. Baloglu, O. Yildirim, U.R. Acharya, Automated detection of COVID-19 cases using deep neural networks with X-ray images, Comput. Biol. Med. 121 (2020), 103792.

[33] J.M. Stokes, K. Yang, K. Swanson, W. Jin, A. Cubillos-Ruiz, N.M. Donghia, C. R. MacNair, S. French, L.A. Carfrae, Z. Bloom-Ackermann, V.M. Tran, A. Chiappino-Pepe, A.H. Badran, I.W. Andrews, E.J. Chory, G.M. Church, E. D. Brown, T.S. Jaakkola, R. Barzilay, J.J. Collins, A deep learning approach to antibiotic discovery, Cell 180 (4) (2020) 688–702.e13.

[34] Y. Kiarashinejad, S. Abdollahramezani, A. Adibi, Deep learning approach based on dimensionality reduction for designing electromagnetic nanostructures, NPJ Comput. Mater. 6 (2020) 1–12.

[35] Q. Ni, Z.Y. Sun, L. Qi, W. Chen, Y. Yang, L. Wang, X. Zhang, L. Yang, Y. Fang, Z. Xing, Z. Zhou, Y. Yu, G.M. Lu, L.J. Zhang, coronavirus disease (COVID-19) pneumonia in chest CT images, Eur. Radiol. 30 (12) (2020) 6517–6527.

[36] O. Shamir, T. Zhang, Stochastic gradient descent for non-smooth optimization: Convergence results and optimal averaging schemes, in: Int. Conf. Mach. Learn., PMLR, 2013, 71–79.

[37] X. Cui, W. Zhang, Z. Tüske, M. Picheny, Evolutionary stochastic gradient descent for optimization of deep neural networks, ArXiv Prepr. ArXiv181006773. (2018).

[38] Y. Liu, Y. Gao, W. Yin, An improved analysis of stochastic gradient descent with momentum, ArXiv Prepr. ArXiv200707989. (2020).

[39] Y. Lei, Y. Ying, Fine-grained analysis of stability and generalization for stochastic gradient descent, in: Int. Conf. Mach. Learn., PMLR, 2020: pp. 5809–5819.

[40] R. Bassily, V. Feldman, C. Guzmán, K. Talwar, Stability of stochastic gradient descent on nonsmooth convex losses, ArXiv Prepr. ArXiv200606914. (2020).

[41] Y. Drori, O. Shamir, The complexity of finding stationary points with stochastic gradient descent, in: Int. Conf. Mach. Learn., PMLR, 2020: pp. 2658–2667.

[42] G.B. Orr, K.-R. Müller, Neural Networks: Tricks of the Trade, Springer, 2003.

[43] M. Shams, O. Elzeki, M. Abd Elfattah, A. ella Hassanien, Chest X-ray images with three classes: COVID-19, Normal, and Pneumonia, 3 (2020). https://doi.org/10.17632/fvk7h5dg2p.3.

[44] B.D. Goodwin, C. Jaskolski, C. Zhong, H. Asmani, Intra-model variability in covid-19 classification using chest x-ray images, ArXiv Prepr. ArXiv200502167. (2020).

[45] A. Abbas, M.M. Abdelsamea, M.M. Gaber, Classification of COVID-19 in chest X-ray images using DeTraC deep convolutional neural network, Appl. Intell. 51 (2) (2021) 854–864.

[46] E.E.-D. Hemdan, M.A. Shouman, M.E. Karar, Covidx-net: A framework of deep learning classifiers to diagnose covid-19 in x-ray images, ArXiv Prepr. ArXiv200311055. (2020).

[47] H.S. Maghdid, A.T. Asaad, K.Z. Ghafoor, A.S. Sadiq, S. Mirjalili, M.K. Khan, Diagnosing COVID-19 pneumonia from X-ray and CT images using deep learning and transfer learning algorithms, in: Multimodal Image Exploit. Learn. 2021, International Society for Optics and Photonics, 2021: p. 117340E.

[48] X. Xu, X. Jiang, C. Ma, P. Du, X. Li, S. Lv, L. Yu, Q. Ni, Y. Chen, J. Su, G. Lang, Y. Li, H. Zhao, J. Liu, K. Xu, L. Ruan, J. Sheng, Y. Qiu, W. Wu, T. Liang, L. Li, A deep learning system to screen novel coronavirus disease 2019 pneumonia, Engineering 6 (10) (2020) 1122–1129.

[49] Y. Song, S. Zheng, L. Li, X. Zhang, X. Zhang, Z. Huang, J. Chen, R. Wang, H. Zhao, Y. Zha, Deep learning enables accurate diagnosis of novel coronavirus (COVID-19) with CT images, IEEE/ACM Trans. Comput. Biol. Bioinform. (2021).

[50] C. Zheng, X. Deng, Q. Fu, Q. Zhou, J. Feng, H. Ma, W. Liu, X. Wang, Deep learning-based detection for COVID-19 from chest CT using weak label, MedRxiv (2020).

[51] B. Ghoshal, A. Tucker, Estimating uncertainty and interpretability in deep learning for coronavirus (COVID-19) detection, ArXiv Prepr. ArXiv200310769. (2020).

[52] 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, Sci. Rep. 10 (2020) 1–12.

[53] I.D. Apostolopoulos, S.I. Aznaouridis, M.A. Tzani, Extracting possibly representative COVID-19 biomarkers from X-ray images with deep learning approach and image data related to pulmonary diseases, J. Med. Biol. Eng. 40 (3) (2020) 462–469.

[54] R. Murugan, T. Goel, E-DiCoNet, Extreme learning machine based classifier for diagnosis of COVID-19 using deep convolutional network, J. Ambient Intell. Humaniz. Comput. (2021) 1–12.

[55] F. Shi, L. Xia, F. Shan, B. Song, D. Wu, Y. Wei, H. Yuan, H. Jiang, Y. He, Y. Gao, Large-scale screening to distinguish between COVID-19 and community-acquired pneumonia using infection size-aware classification, Phys. Med. Biol. 66 (2021), 065031.

[56] L.O. Hall, R. Paul, D.B. Goldgof, G.M. Goldgof, Finding covid-19 from chest x-rays using deep learning on a small dataset, ArXiv Prepr. ArXiv200402060. (2020).

[57] X. He, X. Yang, S. Zhang, J. Zhao, Y. Zhang, E. Xing, P. Xie, Sample-efficient deep learning for covid-19 diagnosis based on ct scans, MedRxiv (2020).