The Detection of COVID-19 in CT Medical Images: A Deep Learning Approach

Nour Eldeen M. Khalifa, Mohamed Hamed N. Taha, Aboul Ella Hassanien, and Sarah Hamed N. Taha

Abstract The COVID-19 coronavirus is one of the latest viruses that hit the earth in the new century. It was declared as a pandemic by the World Health Organization in 2020. In this chapter, a model for the detection of COVID-19 virus from CT chest medical images will be presented. The proposed model is based on Generative Adversarial Networks (GAN), and a fine-tuned deep transfer learning model. GAN is used to generate more images from the available dataset. While deep transfer models are used to classify the COVID-19 virus from the normal class. The original dataset consists of 746 images. The is divided into two parts; 90% for the training and validation phase, while 10% for the testing phase. The 90% then is divided into 80% percent for the training and 20% percent for the validation after using GAN as image augmenter. The proposed GAN architecture raises the number of images in the training and validation phase to be 10 times larger than the original dataset. The deep transfer models which are selected for experimental trials are Resnet50, Shufflenet, and Mobilenet. They were selected as they include a medium number of layers on their architectures if they are compared with large deep transfer models such as DenseNet, and Inception-ResNet. This will reflect on the performance of the proposed model in terms of reducing training time, memory and CPU usage.
experimental trials show that Shufflenet is selected to be the optimal deep transfer learning in the proposed model as it achieves the highest possible for testing accuracy and performance metrics. Shufflenet achieves an overall testing accuracy with 84.9, and 85.33% in all performance metrics which include recall, precision, and F1 score.

**Keywords** Coronavirus · COVID-19 CT images · Medical images · Generative adversarial networks · GAN · Deep transfer learning

1 Introduction

The coronavirus disease 2019 (COVID-19) is caused by severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2/2019-nCoV) [1, 2] and is considered one of the newest and most widespread viruses on the earth which was announced in December 2019. In March 2020 the World Health Organization has declared this new virus as a pandemic, meaning that the virus had spread globally and affected the entire world. COVID-19’s clinical symptoms range from asymptomatic to acute respiratory distress syndrome and multiorgan dysfunction. The common clinical features of this viral infection are similar to other respiratory infections which makes it indistinguishable. In several patients, the disease can progress pneumonia and respiratory failure [3]. The X-ray on the chest in cases of COVID 19 typically reveals bilateral infiltrations which may be normal in early illness. Computed tomography (CT) is considered a more sensitive and specific imaging tool that can show lung infiltration, ground-glass opacities and subsegmental consolidation in asymptomatic patients and patients with no clinical evidence of lower respiratory tract involvement.

In suspicious cases with negative molecular testing, CT scans were used to diagnose COVID-19; most of these patients had positive molecular results on repeated testing [3–5]. Here is the role of Artificial Intelligence and machine learning techniques would help doctors to detect and diagnose COVID 19 accurately and speedily. Over the last decade, various techniques in machine learning were developed and improved rapidly.

These techniques were used to improve the accuracy and time for diagnosis by CAD systems. Artificial intelligence (AI) has made tremendous progress in the area of analysis of medical images. Artificial intelligence including deep-learning systems for medical imaging has been developed especially in the extraction of image characteristics and including shape and spatial relation features.

Learning from a large number of training samples is the most important capability that AI models heavily rely on [6]. Deep learning algorithms achieved impressive and dependable results for computer vision tasks. To achieve high accuracy, huge datasets are needed by these algorithms for training. However, over-fitting is the main concern with deep learning algorithms when they are trained on small datasets, due to generalization lack. This issue is likely to occur with medical images. For supervised learning algorithms, the process of medical image annotation is a time-consuming issue since images would need to be annotated manually by experts [7].
Convolutional Neural Networks have achieved remarkable accuracy and performance in various applications through training on huge-scale annotated training data. Unfortunately, obtaining such huge annotated medical images is challenging. Classical Data Augmentation (DA) techniques such as rotation are geometric/intensity transformations of original images for accurate diagnosis [8]. Generative Adversarial Networks learn deep representations from data without the need for annotated training data. Learning is achieved through deriving backpropagation signals through a competitive process involving a pair of networks [9].

In this chapter, a model that is based on Generative Adversarial Networks and fine-tuned deep transfer learning model, is used for the detection of covid-19 from CT chest medical images. GAN is used to increase the number of images in the dataset by generating more images from the available dataset. Deep transfer models are used to differentiate between the COVID-19 virus and the normal class. The remaining of this chapter is organized as follows. In Sect. 2, related work and scope work will be explored. In Sect. 3, an overview of Generative Adversarial Networks and Deep Transfer Learning will be presented. The dataset used in the proposed model is discussed in Sect. 4. In Sect. 5, the proposed model’s architecture will be presented while Sect. 6 discusses our outcomes and discussion of the chapter. Finally, Sect. 7 provides conclusions and directions for further research.

2 Related Works

The coronavirus (Covid-19) draws other researchers’ attention to better explore the effects of this infectious disease [2]. One of those ways of investigation is the detection of pneumonia from X-ray chest images. There are numerous datasets for chest X-rays for pneumonia such as [10–12], but in this research, the dataset in [12] has been selected due to the availability of data and the dataset has been used in many research to compare our work with as it will be presented in the next paragraphs.

Saraiva et al. in [13] presented a classification of images of childhood pneumonia using convolutional neural networks. The authors proposed a deep learning model with 7 convolutional layers and 3 dense layers and achieved 95.30% testing accuracy. Liang and Zheng in [14] presented a transfer learning method with a deep residual network for pediatric pneumonia diagnosis. The authors a deep learning model with 49 convolutional layers and 2 dense layers and achieved 96.70% testing accuracy. Wu et al. in [15] presented a model to predict pneumonia with chest X-ray images based on convolutional deep neural learning networks and random forest, the authors achieved 97% testing accuracy.

Loey et al. in [16] proposed Generative Adversarial Networks with deep transfer learning for coronavirus detection in limited chest x-ray images. The lack of benchmark datasets for covid-19 especially in chest x-rays images was the main for authors. The main idea is to collect all the possible images for covid-19 and use the GAN network to generate more images to help in the detection of the virus from the available x-rays images. The authors claim that Googlenet is selected to be the main deep
transfer model as it achieves 100% in testing accuracy and 99.9% in the validation accuracy. Although all of the above works have achieved great accuracy in covid-19 classifications based on the patients’ x-rays, medical doctors have proven that covid-19 is probably not detected through x-ray images.

Huang et al. [17] used the convolutional neural network architecture of U-Net and trained it on an annotated dataset of COVID-19. A total of 842 patients (all confirmed to have COVID-19) were collected retrospectively for lung opacity segmentation training and testing, who underwent chest CT scans between 10 January 2020 and 25 January 2020 in Tongji Hospital, Wuhan, China.

Jin et al. [18] built an AI system that has can analyze CT images for the detection of COVID-19 pneumonia features. The system was trained on 1136 training cases (723 positives for COVID-19) from five hospitals. The system sensitivity of was 0.974 and the specificity was 0.922 on the test dataset.

Li et al. [19] developed a 3D deep learning framework COVID-19 detection using chest CT. It can extract both 2D local and 3D global representative features. The framework consists of a RestNet50 as the backbone. It takes a series of CT slices as input and generates features for the corresponding slices. The extracted features from all slices are then combined by a max-pooling operation. The final feature map is fed to a fully connected layer and the softmax activation function to generate a probability score for each type (COVID-19, CAP, and non-pneumonia). The framework archived 90% for sensitivity and 96% for Specificity. The collected dataset has about of 4356 chest CT exams for more than 3300 patients.

Sally et al. [20] has proposed Artificial Intelligence-inspired Model for COVID-19 Diagnosis and Prediction for Patient Response to Treatment (AIMDP). The Model has two function, the Diagnosis Module and Prediction Module. The Diagnosis Module is used for detecting the patients with COVID-19, while the Prediction Module is used for predicting the ability of the patient to respond to treatment based on different factors. The authors claim that the proposed model achieved accuracy 97.14%.

3 Generative Adversarial Networks and Deep Transfer Learning

GANs consist of two different types of networks. Those networks are trained simultaneously. The first network is trained on image generation while the other is used for discrimination. GANs are considered a special type of Deep Learning model. According to [21, 22], GANs have the ability and the power to generate reasonable new images from unlabelled or labelled original images with the capability to be used in various life applications.

Academic and industry fields have focused their attention on adversarial training scheme due to its efficiency and effectiveness in the process of new image generation. GANs have made significant improvements and marvelous performance in many
applications. These applications include image synthesis, semantic image editing, image super-resolution, style transfer and classification.

3.1 GAN Architecture

The important aspect of GAN is the min-max two-player zero-sum game. In this game, one player takes advantage of the equivalent loss of the other player. Here, the players correspond to different networks of GAN called discriminator and generator. The discriminator denoted as D, the main objective is to determine whether a sample belongs to a real distribution or fake distribution [23]. On the other hand, the generator denoted as G, generates a fake sample of images to deceive the discriminator.

Discriminator generates the probability of a given sample to be a real sample one. A real sample is likely to have a higher probability value. Fake samples are indicated by the near-zero probability value. The generator may have an optimal solution when the discriminator loses its ability to differentiate real and fake samples when the probability value is near to 0.5 [23]. The general architecture of GAN is shown in Fig. 1. A multi-dimensional random sample z is given as an input to the Generator to generate samples [23].

The Generator: The Generator is a neural network that uses random noise Z for image generation. The images generated from the Generator using noise are recorded as G(z). Gaussian noise is considered the input, which is a random point in latent space. During the training process, the parameter of both neural networks G and D are updated iteratively.

The Discriminator: The Discriminator is a neural network that works on determining whether a given image belongs to a real distribution or not. It receives image X as input and produces the output D(x) [24]. The objective function of a two-person minimax game is illustrated in Eq. 1

\[
\min_{G} \max_{D} V(D, G) = \mathbb{E}_{x \sim p_{data}(x)} \left[ \log (d(x)) \right] + \mathbb{E}_{z \sim p_{z}} \left[ \log (1 - D(G(z))) \right]
\]  

(1)

![Fig. 1 Graphical representation of the generative adversarial network](image-url)
The Analysis for time complexity (time performance), any explicit measurements of classification or training time could be a misleading. As the hardware capabilities (CPU, GPU, RAM), in addition to the software libraries (Matlab, Tensorflow) and the size of used dataset may reflect a lot about secondary elements, not on the GAN itself [25]. So the time complexity for GANs could be deceptive due to the large number of parameters that may affect the discrimination and training process.

3.2 Deep Transfer Learning Networks

Deep Learning is considered a branch machine learning that depends on algorithms for data processing and thinking process simulation, or for developing abstractions [26–28]. Deep Learning maps inputs to outputs by using layers of methods to process and analyze hidden patterns in data and visually objects detection [29–31]. Data is passed through each layer of a deep network, with the output of the previous layer providing input for the next layer. The first layer is the input layer in the deep neural network, while the output layer is the final layer in the deep network. All the hidden layers are located between input layers and output layers [26, 32].

Years after, various advances in deep convolutional neural networks further reduced the error rate on the image classification competition tasks. CNN models demonstrated significant improvements in succeeding in the ImageNet Large Scale Visual Recognition Competition (ILSVRC) annual challenges. The Visual Geometry Group at Oxford (VGG) developed the VGG-16 and VGG-19 model for the ILSVRC-2014 competition with a 7.3% Top-5 error rate [33]. The winner of the ILSVRC 2014 competition was GoogleNet with a 6.7% Top-5 error rate [34]. In 2015, Residual Neural Network (ResNet) is the winner ILSVRC 2015 competition with a 3.6% Top-5 error rate [35].

4 Datasets Characteristics

The COVID-19 CT scan images dataset [36] used in this research was created and published by Zhao et al. (https://github.com/UCSD-AI4H/COVID-CT). The dataset contains subfolders for each image category (COVID/NonCOVID). There are 742 CT images and 2 categories (COVID/NonCOVID). Figure 2 presents samples of the used data set in this research.
5 Proposed Model Architecture

The proposed model consists of three main blocks, the first block is responsible for generating a new image from the original dataset. The second block is the training and the validation process based on different deep transfer models. The final block is the testing process which calculates the overall testing accuracy with the performance metrics of the proposed model. Figure 3 presents the graphical structure for the proposed model.

In the first block, GAN is used to produce new images that will be used in the second block for the training and validation. The structure of the proposed GAN architecture is presented in Fig. 4. The proposed GAN architecture consists of three main phases: The first is the generator, the augmentation strategy and the third is the discriminator. The generator network consists of 5 transposed convolutional layers, 4 ReLU layers, 4 batch normalization layers, and Tanh Layer at the end of the model, while the discriminator network consists of 5 convolutional layers, 4 leaky ReLU, and 3 batch normalization layers. All the convolutional and transposed convolutional layers used the same window size of 4*4* pixel. Samples of generated images are presented in Fig. 5.

The second block of the proposed model is the training and the validation process using deep transfer models. Three deep transfer models have been selected for investigation in this research. The three models are Resnet50, ShuffleNet, and Mobilenet. Those deep transfer models were selected as the contains a medium number of layers.
| Generator Network                      | Discriminator Network            |
|---------------------------------------|---------------------------------|
| • Input                               | • Input                          |
| • Transposed Convolution 1            | • Convolution 1                  |
| • Batch Normalization 1               | • Leaky ReLU 1                   |
| • ReLU 1                              | • Convolution 2                  |
| • Transposed Convolution 2            | • Batch Normalization 2          |
| • Batch Normalization 3               | • Leaky ReLU 2                   |
| • ReLU 2                              | • Convolution 3                  |
| • Transposed Convolution 3            | • Batch Normalization 3          |
| • Batch Normalization 3               | • Leaky ReLU 3                   |
| • ReLU 3                              | • Convolution 4                  |
| • Transposed Convolution 4            | • Batch Normalization 4          |
| • Batch Normalization 4               | • Leaky ReLU 4                   |
| • ReLU 4                              | • Convolution 5                  |
| • Transposed Convolution 5            |                                 |
| • Tanh                                |                                 |

Fig. 4 GAN proposed architecture

Fig. 5 Samples of GAN generated images for COVID-19 class
if it is compared with large deep transfer model such as DenseNet [37], and InceptionResNet [38], which consist of 201 and 164 layers. While the selected models only contain 50 layers for Resnet50 [38], 50 layers for Shufflenet [39], and 53 layers for Mobilenet [40] as illustrated in Fig. 6.

The authors of research tried first to build their deep neural networks based on the works presented in [41–43] but the testing accuracy wasn’t acceptable. So, the proposed alternative way is to use deep transfer learning models to transfer the learning weights to reduce the training time, mathematical calculations and the consumption of the available hardware resources. This alternative method was adapted in similar research in [44, 45]. The selection of 80% for the training and 20% for validation proved it is efficient in many types of research such as [44, 45].

The selected deep transfer models have been fine-tuned in the last fully connected layer to adapt the number of classes in the data set. The number of classes is two (COVID-19, Normal) class.

6 Experimental Results

The proposed model was implemented using a software package (MATLAB). The development was GPU specific. All experiment trials were conducted on a workstation with Intel Core i9 (2 GHz) as a processor and 32 GB of RAM with Titan X GPU. All experiment trials were carried out by dividing the original dataset into two parts. The first part for the training which contains 90% percent of the original dataset while the second part for testing and contains 10% of the original dataset. The 90% percent represents 673 images while 10% represents 73 images from the original dataset.

The 90% percent of the original data was augmented using GAN to be 10 times larger than the original training dataset. The number of images has been raised to 6730 images. The 6730 images were divided into two parts, the first part contains 80% percent of the data which contains 5384 for the training phase. The second part contains 20% percent which contains 1346 for the validation phase. The hyperparameters for the proposed model were set to as following (i) learning rate = 0.001, (ii) mini-batch size: 64, (iii) number of iterations = 30, (iv) early-stopping = 3 epochs, and (v) optimizer: AdaBoost [46].
6.1 Deep Transfer Model’s Accuracy and Performance Metrics Without GAN

To measure the effectiveness of using GAN as image augmenters. Experimental trials were conducted directly on the original dataset. Table 1 presents the testing accuracy and performance metrics such as recall, precision, and F1 score [47]. The deep transfer models which are investigated in this research are Resnet50, Shufflenet, and Mobilenet.

The testing accuracy and performance metrics which include recall, precision, and F1 score were calculated according to Eqs. (2–5):

\[
\text{Testing Accuracy} = \frac{(TN + TP)}{(TN + TP + FN + FP)} \tag{2}
\]

\[
\text{Precision} = \frac{TP}{TP + FP} \tag{3}
\]

\[
\text{Recall} = \frac{TP}{TP + FN} \tag{4}
\]

\[
\text{F1 Score} = 2 \times \frac{\text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}} \tag{5}
\]

where TP is the count of True Positive samples, TN is the count of True Negative samples, FP is the count of False Positive samples, and FN is the count of False Negative samples from a confusion matrix.

Table 1 illustrated that the testing accuracy with other related performance measurements is quite low. The highest accuracy was achieved by Resnet50 and Shufflenet with 80.82%. The performance metrics strengths that Shufflenet achieved better performance measurements than Resnet50. Nevertheless, all the achieved accuracies are quite low. Figure 7 presents the validation accuracy for shufflenet during the training process. The figure reflects that the data is limited to be trained on, as

| Metric/Model   | Resnet50 | Shufflenet | Mobilenet |
|----------------|----------|------------|-----------|
| Testing Accuracy | 80.82%   | 80.82%     | 74.08%    |
| Precision      | 80.11%   | 80.78%     | 74.05%    |
| Recall         | 80.02%   | 80.92%     | 74.92%    |
| F1 Score       | 80.06%   | 80.85%     | 74.48%    |
the validation accuracy is away below the training accuracy. The need for GAN is mandatory in this case to generate more images.

6.2 Confusion Matrices for Deep Transfer Models with GAN

The confusion matrix illustrates the testing accuracy of every class and the overall testing accuracy for any proposed models. It also helps in deciding the most appropriate deep transfer model that fits the nature of the original dataset. Figures 8, 9, and

![Fig. 8](image_url) Confusion matrix for Resnet50

(a) validation, and (b) testing
Figures 8, 9, and 10 illustrated that the deep transfer model achieved in validation accuracy 97.8, 97.3, and 96.9% for Resnet50, Shufflenet, and Mobilenet accordingly. The highest validation accuracy achieved by Resnet50 with 97.8%. Moreover, the figures also illustrated that deep transfer models achieved testing accuracy of 84.9, 84.9, and 76.7% for Resnet50, Shufflenet, and Mobilenet accordingly. The highest validation accuracy achieved by Resnet50 and Shufflenet with 84.9%.

Another measurement to test the proposed model accuracy is the measurement of testing accuracy for every class. Table 2 presents the validation accuracy for every class for the different deep transfer models while Table 3 illustrates the testing accuracy.

Table 2 illustrates that on the validation accuracy for every class Resnet50 achieved the highest accuracy possible for Covid-19, and Normal class. While Table 3 illustrates that on the testing accuracy for every class Resnet50 achieved the highest accuracy possible for Covid-19 class with 81.1%. Shufflenet deep transfer model
Table 2  Validation accuracy for every class using different deep transfer models

| Accuracy/Model | Resnet50 | Shufflenet | Mobilenet |
|----------------|----------|------------|-----------|
| Covid-19       | 97.0%    | 96.0%      | 95.4%     |
| Normal         | 98.5%    | 98.2%      | 98.3%     |
| Total Accuracy | 97.8%    | 97.3%      | 96.9%     |

Table 3  Testing accuracy for every class using different deep transfer models

| Accuracy/Model | Resnet50 | Shufflenet | Mobilenet |
|----------------|----------|------------|-----------|
| Covid-19       | 81.1%    | 79.5%      | 69.3%     |
| Normal         | 88.9%    | 91.2%      | 68.7%     |
| Total Accuracy | 84.9%    | 84.9%      | 76.7%     |

achieved the highest accuracy possible for Normal class with 91.2%. As the main target of this research is to detect the Covid-19, so Resent50 will be the optimal model until now. The final decision will be introduced after the investigation of performance metrics in the following subsection. Figure 11 illustrates the progress

Fig. 11  Validation accuracy progress for the training phase for Shufflenet with GAN
of the validation accuracy for Resnet50. Notably, the validation accuracy has been improved when the training was conducted without GAN as presented in Fig. 7.

6.3 Performance Metrics for Deep Transfer Models with GAN

The performance metrics which are investigated in this research are the recall, precision, and F1 score. Table 4 presents the recall, precision, and F1 score metrics for the different deep transfer model using GAN in the validation phase.

Table 4 illustrated that all the deep transfer model achieved competitive results ranging from 96.84 to 98.41% for all the performance metrics in the validation phase. Resnet50 achieved the highest percentage in the performance metrics with 98.30% in precision, 98.31% in the recall metric, and 98.30% in the F1 score metric in the validation phase.

Table 5 shows the recall, precision, and F1 score metrics for the different deep transfer model using GAN in the testing phase.

Table 5 illustrates that Resnet50 and Shufflenet achieved a close result in all performance metrics in the testing phase. Shufflenet achieved the highest percentage in the performance metrics with 85.34% on all performance metrics in the testing phase. According to the achieved results in the testing phase for overall testing

| Table 4 | The recall, precision, and F1 score metrics for different deep transfer models using GAN for the validation phase |
|---------|--------------------------------------------------|
| Metric/Model | Resnet50 | Shufflenet | Mobilenet |
| Precision | 98.30% | 97.25% | 96.97% |
| Recall | 98.31% | 97.34% | 96.84% |
| F1 Score | 98.30% | 97.30% | 96.90% |

| Table 5 | The recall, precision, and F1 score metrics for different deep transfer models using GAN for the testing phase |
|---------|--------------------------------------------------|
| Metric/Model | Resnet50 | Shufflenet | Mobilenet |
| Precision | 84.98% | 85.33% | 78.22% |
| Recall | 85.14% | 85.33% | 77.45% |
| F1 Score | 85.06% | 85.33% | 77.83% |
Shufflenet is selected to be the optimal deep transfer learning model as it achieved the highest possible testing accuracy and performance metrics. Figure 12 presents samples of testing images with the achieved testing accuracy.

### 7 Conclusions and Future Works

The COVID-19 coronavirus is one of the latest viruses in the new century. It was declared as a pandemic by the World Health Organization in 2020. In this chapter, a model for the detection of the COVID-19 virus from CT chest medical images is presented. The proposed model was based on Generative Adversarial Networks (GAN), and a fine-tuned deep transfer learning model. GAN has been used to generate more images from the available dataset. While deep transfer models have been used to classify the COVID-19 virus from the normal class. The original data sets consisted of 746 images. The data set has been divided into two parts, 90% for training and
validation while 10% for testing. The 90% was divided into 80% percent for training and 20% percent for validation after using GAN. Proposed GAN architecture raised the number of images in the training and validation process to be 10 times larger than the original dataset. The deep transfer models selected for experimental trials are Resnet50, Shufflenet, and Mobilenet. They were selected as they included a medium number of layers if it is compared with a large deep transfer model such as DenseNet, and InceptionResNet. That will reflect on the performance of the proposed model in terms of reducing training time, memory and CPU usage. The experimental trials showed that Shufflenet was selected to be the optimal deep transfer learning in the proposed model as it achieved the highest possible testing accuracy and performance metrics. Shufflenet achieved an overall testing accuracy with 84.9, and 85.33% in all performance metrics which included recall, precision, and F1 score. One of the potential future works is to select deeper deep transfer learning such as DenseNet, and InceptionResNet to improve the testing accuracy of the proposed model.

Acknowledgements We gratefully acknowledge the support of NVIDIA Corporation, which donated the Titan X GPU used in this research.

References

1. Liu, J., et al.: Hydroxychloroquine, a less toxic derivative of chloroquine, is effective in inhibiting SARS-CoV-2 infection in vitro. Cell Discov. 6(1), 6–9 (2020). https://doi.org/10.1038/s41421-020-0156-0
2. Hageman, J.R.: The coronavirus disease 2019 (COVID-19). Pediatr. Ann. (2020). https://doi.org/10.3928/19382359-20200219-01
3. Singhal, T.: A review of coronavirus disease-2019 (COVID-19). Indian J. Pediatr. 87(4), 281–286 (2020). https://doi.org/10.1007/s12098-020-03263-6
4. Huang, P., et al.: Use of chest CT in combination with negative RT-PCR assay for the 2019 novel coronavirus but high clinical suspicion. Radiology 295(1), 22–23 (2020). https://doi.org/10.1148/radiol.2020200330
5. Cinkooglu, A., Bayraktaroglu, S., Savas, R.: Lung changes on chest CT during 2019 novel coronavirus (COVID-19) pneumonia. Eur. J. Breast Heal. 16(2), 89–90 (2020). https://doi.org/10.5152/ejbh.2020.010420
6. Hosny, A., Parmar, C., Quackenbush, J., Schwartz, L.H., Aerts, H.J.W.L.: Artificial intelligence in radiology. Nat. Rev. Cancer (2018). https://doi.org/10.1038/s41568-018-0016-5
7. Iqbal, T., Ali, H.: Generative adversarial network for medical images (MI-GAN). J. Med. Syst. (2018). https://doi.org/10.1007/s10916-018-1072-9
8. Shorten, C., Khoshgoftaar, T.M.: A survey on image data augmentation for deep learning. J. Big Data (2019). https://doi.org/10.1186/s40537-019-0197-0
9. Yi, X., Walia, E., Babyn, P.: Generative adversarial network in medical imaging: a review. Med. Image Anal. (2019). https://doi.org/10.1016/j.media.2019.101552
10. Oakden-Rayner, L.: CheXNet: an in-depth review. Luke Oakden-Rayner Blog (2019)
11. Irvin, J., et al.: CheXpert: a large chest radiograph dataset with uncertainty labels and expert comparison (2019)
12. Kermany, D., Zhang, K., Goldbaum, M.: Labeled optical coherence tomography (OCT) and chest X-ray images for classification (2018)
13. Saraiva A.A., et al.: Classification of images of childhood pneumonia using convolutional neural networks. In: BIOIMAGING 2019—6th International Conference on Bioimaging, Proceedings; Part 12th International Joint Conference on Biomedical Engineering Systems and Technologies BOSTEC 2019, pp. 112–119 (2019). \url{https://doi.org/10.5220/0007404301120119}

14. Liang G., Zheng, L.: A transfer learning method with deep residual network for pediatric pneumonia diagnosis. Comput. Methods Programs Biomed. (2020), \url{https://doi.org/10.1016/j.cmpb.2019.06.023}

15. Wu, H., Xie, P., Zhang, H., Li, D., Cheng, M.: Predict pneumonia with chest X-ray images based on convolutional deep neural learning networks. J. Intell. Fuzzy Syst. (2020). \url{https://doi.org/10.3233/jifs-191438}

16. Loey, M., Smarandache, F., Khalifa, N.E.M.: Within the lack of COVID-19 benchmark dataset: a novel GAN with deep transfer learning for corona-virus detection in chest X-ray images (2020)

17. Huang, L., et al.: Serial quantitative chest CT assessment of COVID-19: deep-learning approach. Radiol. Cardiothorac. Imaging 2(2), e200075 (2020). \url{https://doi.org/10.1148/ryct.2020200075}

18. Jin S., et al.: AI-assisted CT imaging analysis for COVID-19 screening: building and deploying a medical AI system in four weeks. medRxiv p. 2020.03.19.20039354 (2020). \url{https://doi.org/10.1101/2020.03.19.20039354}

19. Li L., et al.: Artificial intelligence distinguishes COVID-19 from community acquired pneumonia on chest CT. Radiology, p. 200905 (2020). \url{https://doi.org/10.1148/radiol.2020200905}

20. ELGhamrawy S.M., Hassanien, A.E.: Diagnosis and prediction model for COVID19 patients response to treatment based on convolutional neural networks and whale optimization algorithm using CT images. medRxiv p. 2020.04.16.20063990 (2020). \url{https://doi.org/10.1101/2020.04.16.20063990}

21. Sandfort, V., Yan, K., Pickhardt, P.J., Summers, R.M.: Data augmentation using generative adversarial networks (CycleGAN) to improve generalizability in CT segmentation tasks. Sci. Rep. 9(1), 16884 (2019). \url{https://doi.org/10.1038/s41598-019-52737-x}

22. Antoniou, A., Storkey, A., Edwards, H.: Data augmentation generative adversarial networks (2017) [Online]. Available: \url{http://arxiv.org/abs/1711.04340}

23. Alqahtani, H., Kavakli-Thorne, M., Kumar, G.: Applications of generative adversarial networks (GANs): an updated review. Arch. Comput. Methods Eng. (2019). \url{https://doi.org/10.1007/s11831-019-09388-y}

24. Goodfellow, I.J., et al.: Generative adversarial nets. Adv. Neural Inf. Process. Syst. (2014). \url{https://doi.org/10.3156/jsof.29.5_177_2}

25. Sokolova, M., Lapalme, G.: A systematic analysis of performance measures for classification tasks. Inf. Process. Manag. 45(4), 427–437 (2009). \url{https://doi.org/10.1016/j.ipm.2009.03.002}

26. LeCun, Y., Bengio, Y., Hinton, G.: Deep learning. Nature 521(7553), 436–444 (2015). \url{https://doi.org/10.1038/521.436}

27. Eraslan, G., Avsec, Ž., Gagneur, J., Theis, F.J.: Deep learning: new computational modelling techniques for genomics. Nat. Rev. Genet. 20(7), 389–403 (2019). \url{https://doi.org/10.1038/s41576-019-0122-6}

28. Voulodimos, A., Doulamis, N., Doulamis, A., Protopapadakis, E.: Deep learning for computer vision: a brief review. Comput. Intell. Neurosci. 2018, 7068349 (2018). \url{https://doi.org/10.1155/2018/7068349}

29. Riordon, J., Sovijl, D., Sanner, S., Sinton, D., Young, E.W.K.: Deep learning with microfluidics for biotechnology. Trends Biotechnol. 37(3), 310–324 (2019). \url{https://doi.org/10.1016/j.tibtech.2018.08.005}

30. You, J., McLeod, R.D., Hu, P.: Predicting drug-target interaction network using deep learning model. Comput. Biol. Chem. 80, 90–101 (2019). \url{https://doi.org/10.1016/j.compbiolchem.2019.03.016}
31. Jaganathan, K., et al.: Predicting splicing from primary sequence with deep learning. Cell 176(3), 535–548.e24 (2019). https://doi.org/10.1016/j.cell.2018.12.015
32. Cao, C., et al.: Deep learning and its applications in biomedicine. Genomics, Proteomics Bioinf. 16(1), 17–32 (2018). https://doi.org/10.1016/j.gpb.2017.07.003
33. Liu, S., Deng, W.: Very deep convolutional neural network based image classification using small training sample size. In: 2015 3rd IAPR Asian Conference on Pattern Recognition (ACPR), pp. 730–734 (2015) https://doi.org/10.1109/acpr.2015.7486599
34. Szegedy, C., et al.: Going deeper with convolutions. In: Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition, 2015, pp. 1–9. https://doi.org/10.1109/cvpr.2015.7298594
35. He, K., Zhang, X., Ren, S., Sun, J.: Deep residual learning for image recognition. In: 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2016, pp. 770–778. https://doi.org/10.1109/CVPR.2016.90
36. Zhao, J., Zhang, Y., He, X., Xie, P.: COVID-CT-dataset: a CT scan dataset about COVID-19. pp. 1–5 (2020)
37. Huang, G., Liu, Z., van der Maaten, L., Weinberger, K.Q.: Densely connected convolutional networks. In: 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2017, pp. 2261–2269. https://doi.org/10.1109/cvpr.2017.243
38. Szegedy, C., Ioffe, S., Vanhoucke, V., Alemi, A.A.: Inception-v4, inception-ResNet and the impact of residual connections on learning In: 31st AAAI Conference on Artificial Intelligence, AAAI 2017 (2017)
39. Zhang, X., Zhou, X., Lin, M., Sun, J.: ShuffleNet: an extremely efficient convolutional neural network for mobile devices. In: Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition, 2018. https://doi.org/10.1109/cvpr.2018.00716
40. Qin, Z., Zhang, Z., Chen, X., Wang, C., Peng, Y.: Fd-Mobilenet: improved Mobilenet with a fast downsampling strategy. In: Proceedings—International Conference on Image Processing, ICIP, 2018, https://doi.org/10.1109/icip.2018.8451355
41. Khalifa, N.E.M., Taha, M.H.N., Hassanien, A.E., Hemedan, A.A.: Deep bacteria: robust deep learning data augmentation design for limited bacterial colony dataset. Int. J. Reason. Intell. Syst. 11(3), 256 (2019). https://doi.org/10.1504/IJRIS.2019.102610
42. Khalifa, N.E.M., Taha, M.H.N., Ali, D.E., Slowik, A., Hassanien, A.E.: Artificial intelligence technique for gene expression by tumor RNA-Seq data: a novel optimized deep learning approach. IEEE Access (2020). https://doi.org/10.1109/access.2020.2970210
43. Khalifa, N.E.M., Taha, M.H.N., Hassanien, A.E., Selim, I.M.: Deep Galaxy: classification of Galaxies based on deep convolutional neural networks (2017). arXiv:1709.02245
44. Khalifa, N., Loey, M., Taha, M., Mohamed, H.: Deep transfer learning models for medical diabetic retinopathy detection. Acta Inform. Medica 27(5), 327 (2019). https://doi.org/10.5455/aim.2019.27.327-332
45. Khalifa, N.E.M., Loey, M., Taha, M.H.N.: Insect pests recognition based on deep transfer learning models. J. Theor. Appl. Inf. Technol. 98(1), 60–68 (2020)
46. Žižka, J., Dařená, F., Svoboda, A., Žižka, J., Dařená, F., Svoboda, A.: AdaBoost. In: Text Mining with Machine Learning (2019)
47. Goutte C., Gaussier, E.: A probabilistic interpretation of precision, recall and F-score, with implication for evaluation. In: Losada D.E., Fernández-Luna J.M. (eds.) Advances in Information Retrieval, pp. 345–359. Berlin, Heidelberg: Springer Berlin Heidelberg (2005)