CAE-COVIDX: automatic covid-19 disease detection based on x-ray images using enhanced deep convolutional and autoencoder

Hanafi [a,1,*], Andri Pranolo [b,c,2], Yingchi Mao [c,3]

[a] Universitas Amikom Yogyakarta, Jl. Ringroad Utara, Condongcatur, Depok, Sleman, Yogyakarta 55283, Indonesia
[b] Informatics Department, Universitas Ahmad Dahlan, Jl. Kapas No.9, Semaki, Umbulharjo, Yogyakarta 55166, Indonesia
[c] College of Computer and Information, Hohai University, 8 Fochengxi Road Jiangning District, Nanjing, Jiangsu Province, China
[1] hanafi@amikom.ac.id; [2] andri.pranolo@tif.uad.ac.id; [3] maoyingchi@gmail.com
* corresponding author

ARTICLE INFO

Article history
Received October 26, 2020
Revised January 1, 2021
Accepted January 18, 2021
Available online March 31, 2021

Keywords
Covid-19
CNN
Autoencoder
Deep learning
X-Ray image

ABSTRACT

Since the first case in 2019, Corona Virus has been spreading all over the world. World Health Organization (WHO) announced that COVID-19 had become an international pandemic. There is an essential section to handle the spreading of the virus by immediate virus detection for patients. Traditional medical detection requires a long time, a specific laboratory, and a high cost. A method for detecting Covid-19 faster compared to common approaches, such as RT-PCR detection, is needed. The method demonstrated that it could produce an X-ray image with higher accuracy and consumed less time. We propose a novel method to extract image features and to classify COVID-19 using deep CNN combined with Autoencoder (AE) dubbed CAE-COVIDX. We evaluated and compared it with the traditional CNN and existing framework VGG16 involving 400 normal images of non-COVID19 and 400 positive COVID-19 diseases. The performance evaluation was conducted using accuracy, confusion matrix, and loss evaluation. Based on experiment results, the CAE-COVIDX framework outperforms previous methods in several testing scenarios. This framework's ability to detect Covid-19 in various nonstandard image X-rays could effectively help medical employers diagnose Covid-19 patients. It is an important factor to decrease the spreading of Covid-19 massively.

This is an open access article under the CC–BY-SA license.

1. Introduction

The novel coronavirus has spread massively across countries all over the world. World Health Organization officially named the virus COVID-19. The widespread of the virus, which was first recognized in Wuhan China at the end of 2019, has changed billions of people's daily lives globally and has resulted in significant financial and non-financial losses. According to WHO reports (Fig. 1), COVID-19 has spread to most countries around the world. On October 5th, 2020, as shown by the official WHO website, 35,708,008 people had been infected by COVID-19, and 1,001,715 died [1]. Today, the world population is still combating the virus, and the spread of the virus throughout the world population cannot be reduced significantly. Confirmed Covid-19 cases are found in many parts of the world, which means the virus has spread in all countries.

An important approach to handle COVID-19 is the speed screening of infected people. For example, people infected should be able to receive treatment immediately, such as increasing their immunity through the provision of vitamins and nutrition. Patients infected by covid-19 should also be isolated to

prevent the spread of the virus. The common procedure to test the patients is using the Reverse Transcription Polymerase Chain Reaction (RT-PCR) test for COVID-19. This detection model requires a long time to obtain final results, needs several laborious tools, and involves complicated handcraft parts with limited stock [2]. Recent evidence reports that RT-PCR detection only had 30-60% accuracy. Moreover, the detection test is inconvenient, needs nasopharyngeal test swabs, and it is costly. On the other hand, the COVID-19 test using X-ray images is widely available, and the scanning process is inexpensive. In addition, this test produces the final results more quickly.

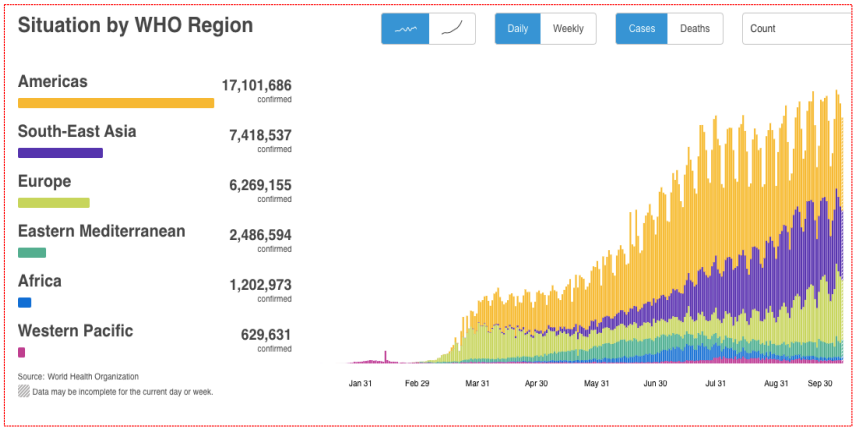

Fig. 1. Number of Covid-19 cases around the world

Patient infection indications include dyspnea, cough, respiratory symptoms, and fever. In most severe cases, the infection causes multi-organ failure, pneumonia, acute respiratory syndrome, septic shock, and death. Following the recent finding announced by Chinese government officials, the detection of COVID-19 can use blood sample testing or respiration as important identification key factors for RT-PCR. However, for common public health emergencies, RT-PCR's inaccuracy is what causes the late detection of most COVID-19 patients. COVID-19 is categorized as pneumonia infection, whereas pneumonia is a lung infection caused by bacteria, viruses, or fungi. The virus refers to an inflammation of the pulmonary airbags (alveoli) and fluid or pus. It can make it difficult for oxygen to enter the bloodstream. Fig. 2 is an example of a patient's lungs obtained by a photo scan.

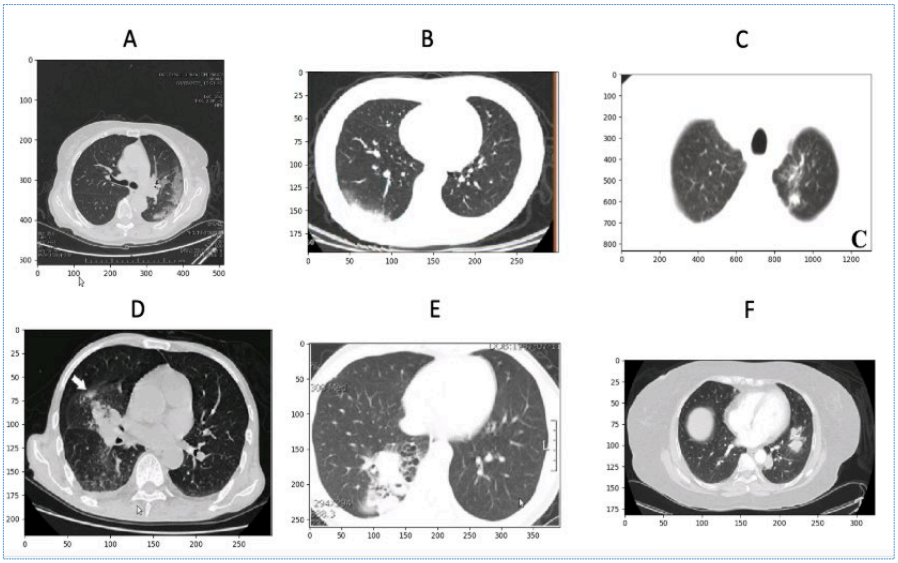

Fig. 2. Example of the CT scan image for COVID-19 infection (A, B, and C) and non-COVID-19 infection (D, E, and F). All of the images similar, making it very hard to generalize the classification algorithm.

The Computed Tomography (CT) scan of the chest is a popular approach to detect pneumonia. Some proposed methods involve machine learning using CT photo examination equipment to conduct

virus detection, monitoring, and quantification of COVID-19. This tool plays an important role in detecting patients infected with COVID-19. X-ray machines are common tools to detect bone dislocations, tumors, cancer, lung diseases, and pneumonia. The CT scan (Fig. 3), particularly with an enhanced X-ray machine, identifies a very detailed human organ structure [3].

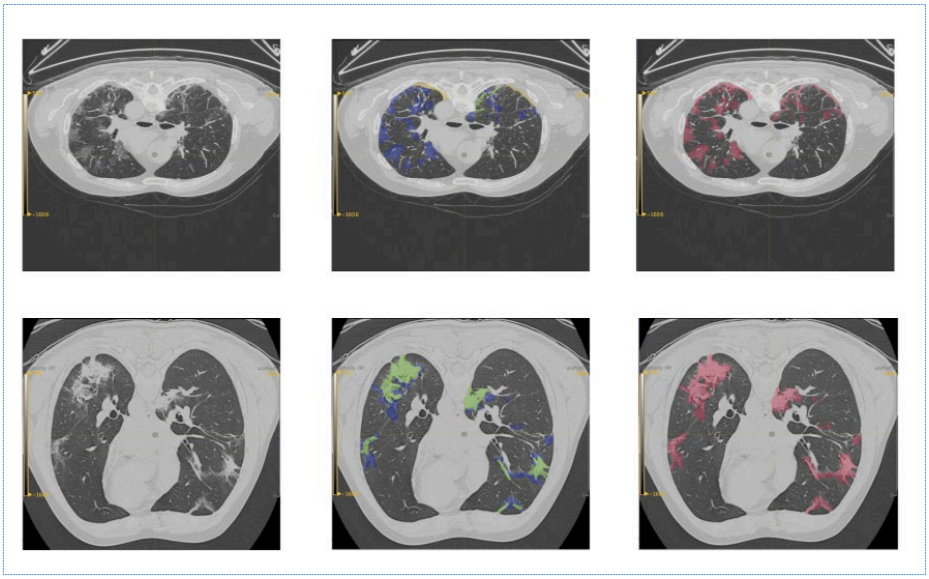

Fig. 3. Example of lung image obtained by a CT-Scan machine

According to the literature review from previous works, several experts have proposed an automatic COVID-19 detection model using the deep learning model. CNN is a sub-class of the deep learning approach with tremendous achievements in computer science application fields such as computer vision, image processing, voice recognition, text processing, and video processing [4][5]. Especially in COVID-19 detection, CNN has also become a popular method to develop the classification learning model to detect COVID-19 diseases based on X-ray chest images. One of the novels CNN models introduced is computer vision [6]. It becomes one of the modern deep learning methods based on CNN, widely used in several computer science applications with tremendous performance over existing traditional neural network models [7]. For automatic COVID-19 detection application specifically, we have collected data from a major computer science digital library that shows a detailed characteristic dataset in Table 1.

Some researchers have attempted to achieve better performance for the classification model following previous work using X-ray image classification. There is a slight difference in achievement between researchers. For example, [8] when they try to classify images to detect three categories of the condition, including bacteria pneumonia, virus pneumonia, and normal patients. The majority of researchers have also conducted research to achieve effectiveness in classification tasks. They consider adopting CNN as major machine learning to obtain COVID-19 detection. Unfortunately, the majority of CNN adoption only reaches low effectivity. There are several challenges open to improving the existing models.

Multi-layer CNN is also involved in this research category, such as [9], [10], [11] adopting Rest-net and VGG16 as a deep learning platform based on CNN. Their work achieves a better target over generic CNN. The ResNet adoption model reaches 86% accuracy, and ResNet achieves 90% accuracy in larger dataset adoption. DeTrac [12] involved CNN with transfer learning, decompose, and composing processes to classified chest X-ray images. A study adopted fuzzy to preprocessing dataset consist pneumonia, Covid-19 and normal. They implement deep learning platforms include MobileNetV2 and SqueezeNet [13]. A proposed model using deep hybrid learning consists of CNN, vanilla neural network, VGG, and capsule network [14]. A different model using multi-view fusion segmentation involves deep learning based on the Restnet50 framework [15]. A study to investigate several deep learning platforms considers ten frameworks: AlexNet, VGG-16, VGG-19, SqueezNet, GoogleNet, MobileNet-V2, ResNet-18 ResNet50, ResNet-101 success in classified COVID-19. According to their experiment

report where RestNet-101 success in achieved performance over another deep learning platform [16]. Several studies mentioned above, the using deep learning and preprocessing approach influences the performance of detection achievement. In some cases, the deep learning platform using ResNet has proven better performance than some deep learning platforms.

Table 1. The previous works using the ML approach

Ref.	Material	Datasets	Objective	Method	Result
[8]	X-ray image	45 COVID-19, 931 Bacteria Pneumonia, 660 Virus Pneumonia, 1203 Normal.	Category: Bacteria Pneumonia/Virus Pneumonia/Normal	CNN	83.5% Accuracy
[11]	CT-scan images	51 COVID-19, 55 Normal	Category: COVID-19/ Normal	UNET++	95.2% Accuracy
[17]	CT-scan images	496 COVID-19, 1385 Normal	Category: COVID-19/other	CNN	94.1% Sensitivity
[18]	CT-scan images	44 COVID-19, 55 Virus Pneumonia	Category: COVID-19/Virus Pneumonia.	CNN	82.9% Accuracy
[10]	CT-scan images	88 COVID-19, 86 Normal	Category: COVID-19/ Bacteria Pneumonia /Normal	ResNet-50	86% Accuracy
[19]	CT-scan images	219 COVID-19, 224 Influenza, 175 Normal	Category: COVID-19/ Influenza/Normal	CNN	86.7% Accuracy
[9]	CT-scan images	468 COVID-19, 1551 CAP, 1445 Non-pneumonia.	Category: COVID-19/CAP/ Non-pneumonia.	ResNet-50	90.0% (Sens.) 96.0% (Spec.)
[20]	CT-scan images	1658 COVID-19, 1027 CAP	Category: COVID-19/CAP	RF	87.9% Accuracy 90.7% Sensitivity 83.3 (Spec.)
[21]	CT-scan images	176 COVID-19 patients	Severity assessment	RF	87.5% Accuracy 93.3 (TPR) 74.5% (TNR)
[22]	X-ray images	70 COVID-19/normal	Category: COVID-19/normal	CNN	92.9% Accuracy

Automatic COVID-19 detection, firstly proposed using CNN, involves transfer learning [23]. The research involves specific datasets of 224 confirmed positive COVID-19 diseases, 700 confirmed common bacteria pneumonia, and 504 normal conditions. According to the experiment report, they achieved 96.46% accuracy in performance. Considering transfer learning with traditional CNN becomes the contribution to achieving better performance over the traditional CNN approach. Indeed, in the case of transfer learning adoption, it needs to collect larger datasets of the image of COVID-19 patients in which, until today, the availability of data is tough to collect.

Another researcher proposed a novel model with 17 layer CNN called Darknet [24]. The model achieves 88.02% accuracy for multi-class cases. COVID-NET, a novel model of deep learning with an autoencoder (AE) model mechanism, has been involved in detecting this virus disease [25]. The experiment was conducted on datasets of 1044 patients with 449 COVID-19 patients confirmed, 397 different types of pathology, 98 with lung cancer, and 100 normal conditions. Their model achieves higher than 93% under the ROC curve. Unlike the sub-class of CNN, AE relies on feature extraction to obtain the classification of COVID-19 disease.

Combining deep learning class using CNN and transfer learning based on Alex Net was proposed [26]. The model is very simple where it just adopts 1 Convolutional layer and then incorporates it with the preprocessing image from the Alex Net model. Their model claimed to reach 94% accuracy [27]. On the other hand, the Alex Net framework's involvement reaches better enrichment over traditional CNN that most researchers in this field apply.

A new model of COVID-19 detection based on the image that obtained CT-Scan using pre-train using CNN [28], Another study also using optimization deep learning machine based on CNN incorporated with Bayesian, They called Bayesian Convolutional Neural Network (BCNN) [22]. A study involves anomaly detection based on deep learning [29]. This model adopts 18-layer CNN, ImageNet pre-trained image datasets. According to the evaluation report, they reach a sensitivity of 96.00%. [30] A study applied DenseNet deep learning platform aims to classify COVID-19 (+) or COVID (-). ImageNet considers adopting to utilized image feature extraction. According to the experiment report, this model superior to the competitor in previous work.

According to the reference mentioned above, two important factors will become challenges in the future: the sensitivity of the amount of data and the feature learning platform in calculating image layers. It is used to produce classification tasks that would achieve better accuracy performance to detect COVID-19 disease using lung X-ray images. Therefore, We proposed a novel model to generate automatic COVID-19 detection by applying the hybrid dual deep learning approach based on the convolutional neural network (CNN) and autoencoder (AE) model, which is called CAE-COVIDX. Our design considers the hierarchical model of CNN and AE to detect cases more accurately. CNN leads to dimensional reduction, which has shortcomings in data feature representation. Meanwhile, AE has specific characteristic advantages in feature extraction. On the other hand, our proposed model combines both models to achieve better accuracy when detecting people infected by COVID-19. Our research contributions propose a novel method to extract image features based on X-ray chest images using lung image feature characteristics and a novel method to classify COVID-19 diseases using hybrid CNN and AE to determine disease infection.

2. Method

Our proposed model is called CAE-COVIDX, which uses different machine learning models which involve dimensional reduction features by CNN and feature extraction representation by AE platform. We called our framework CAE (Convolutional Autoencoder) as the core feature selection to learning classification tasks to distinguish COVID-19 patients from normal patients. This model really inspired the enhancement of the CNN platform to classify datasets in the case of MNIST image characteristics [31]. Our complete name, CAE-COVIDX, consists of 4 essential steps to obtain final results based on machine learning which are as follows: 1) image lung COVID-19 data acquisition, 2) data image preprocessing, 3) proposed hybridization model of CNN and AE, and 4) training process and evaluation.

2.1. Convolutional neural network

CNN categories of neural network sub-class achieved success in the computer science field research and application [32]. CNNs are hierarchical models in which convolutional layers alternate with subsampling layers, evocative of the primary visual cortex's basic and complex cells (Fig. 4). Three fundamental building blocks comprise the network architecture, which must be stacked and compiled as needed. A convolutional layer, a max-pooling layer, and a classification layer are all included. CNNs are one of the most widely used models for supervised image classification, and they set the bar high for a variety of benchmarks. Additionally, convolutional neural networks employ multi-stage, multi-stage trainable architectures. Function maps are collections of array values that serve as the input and output of each point factor.

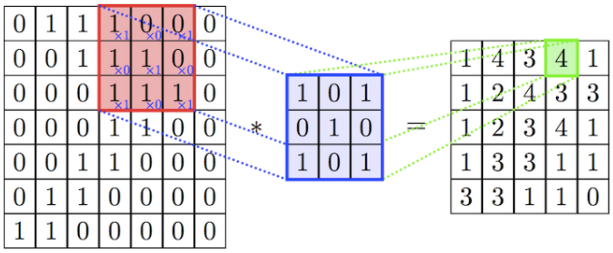

Fig. 4. The basic concept of the convolutional neural network

2.1.1. Autoencoder

We note the underlying concepts of auto-encoder models where the basic autoencoder (Fig. 5) begins with input $x \in \mathcal{R}^d$ and the second essential aspect associated with the first map of latent representation $h \in \mathcal{R}^{d'}$ applying deterministic approach given the formula $h = f\theta = \sigma(W\bar{x}a^2 + b^2 = c^2 + b)$ using parameter scenario $\theta = \{W, b\}$. This mechanism adopted to extract the input uses convert mapping with $f: y = f\theta'(h) = \sigma(W'h + b')$ where $\theta' = \{W'h + b'\}$. These two formulations must be constrained with $W' = W^T$, by applying a similar value for encoding as an input layer, and extraction becomes the latent representation of the decoding process.

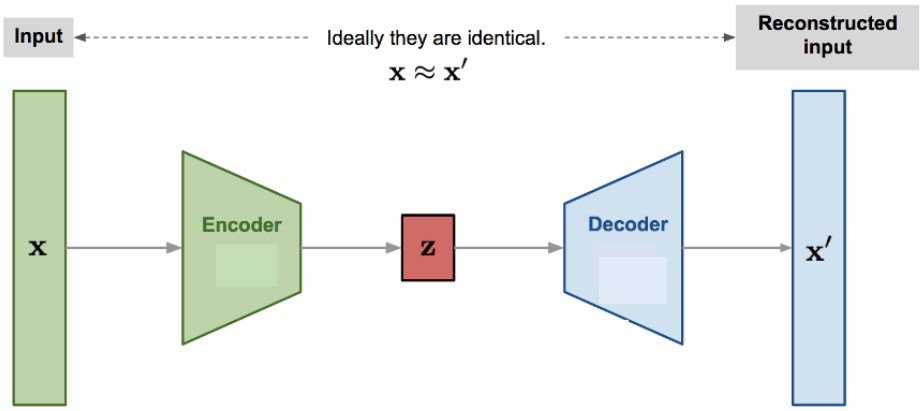

Fig. 5. Basic concept autoencoder

2.1.2. A hybrid of convolutional and autoencoder (CAE)

Both AE and DAE ignore the structure of the two-dimensional image. This is not only considered when working with practical inputs, but it also ensures parameter consistency, resulting in each attribute being global. However, vision and object recognition are toward discovering localized features that reinforce one another in the input. CAEs are distinguished from conventional AEs because their weights are exchanged between all input locations, preserving spatial localization. Restoration is then attributed to a linear combination of simple image patches dependent on the latent code. CAE architecture is conceptually identical to segment architecture, except for shared weights. For a single-channel input x, the k-*th* function map's latent representation is given (1).

$$h^k = \sigma(X * X^k + b^k) \tag{1}$$

Where the bias is uniformly distributed across the map, it is a combination of an activation function (in all of our experiments, we used the scaled hyperbolic tangent) and a two-dimensional convolution. We use a single latent map distraction because any filter to specialize in the entire input set of features (one distraction per pixel is too free) by acquiring reconstruction (2).

$$y = \sigma \sum_{k \in H} h^k * \widehat{W}^k + c \tag{2}$$

Where c represents bias per input channels, H is a group of latent feature maps, $\widehat{W}^k$ represents an operation of both dimensions of weight. The backpropagation algorithm is used to measure the error function gradient to the parameters for regular networks. Convolution operations can conveniently do this by using (3).

$$\frac{\vartheta E(\theta)}{\vartheta W^k} = x * \delta h^k + \tilde{h}^k * \delta y \tag{3}$$

Where The symbol of δh and δy represent the delta of the hidden layer and reconstruction layer, the mechanism updates of shared weights adopt the stochastic gradient descent.

A max-pooling layer is commonly used to accomplish translation invariants throughout hierarchical networks in general and CNNs specifically. Max extracts samples of latent representations via a steady component, generally take the highest value across non-supervalent subregions. This contributes to the filter's selectivity by determining the "fit" between the function and the input field across the region of interest in order to activate each neuron in a latent representation. Initially, max-pooling was reserved for architectures with fully supervised feed forwarding.

A max-pooling sheet was added to spares the representation secret of non-overlapping sub-regions, removing all non-maximum values. It forces the detectors to implement more extensively, ignoring trivial solutions like just one "on" weight (identity function). Reducing the average number of filters leading to each pixel's decoding during the reconstruction period requires filters to be more general. Also, It does not need *L1* and/or *L2* over hidden units regularization and/or weights with a max-pooling layer.

Many AEs can be designed as a stack to form a profound hierarchy; for example, from the layer latent representation, and each layer receives its input. Deep belief networks adequate to adopt in veracious, then weights can be modified with backpropagation, or the SVMs could use the highest stage activations as attribute vectors or other classification instruments. In another case, a CNN with the same topology can be initialized by the CAE model. The detailed work mechanism of the convolutional layer process and autoencoder process are shown in Fig. 6.

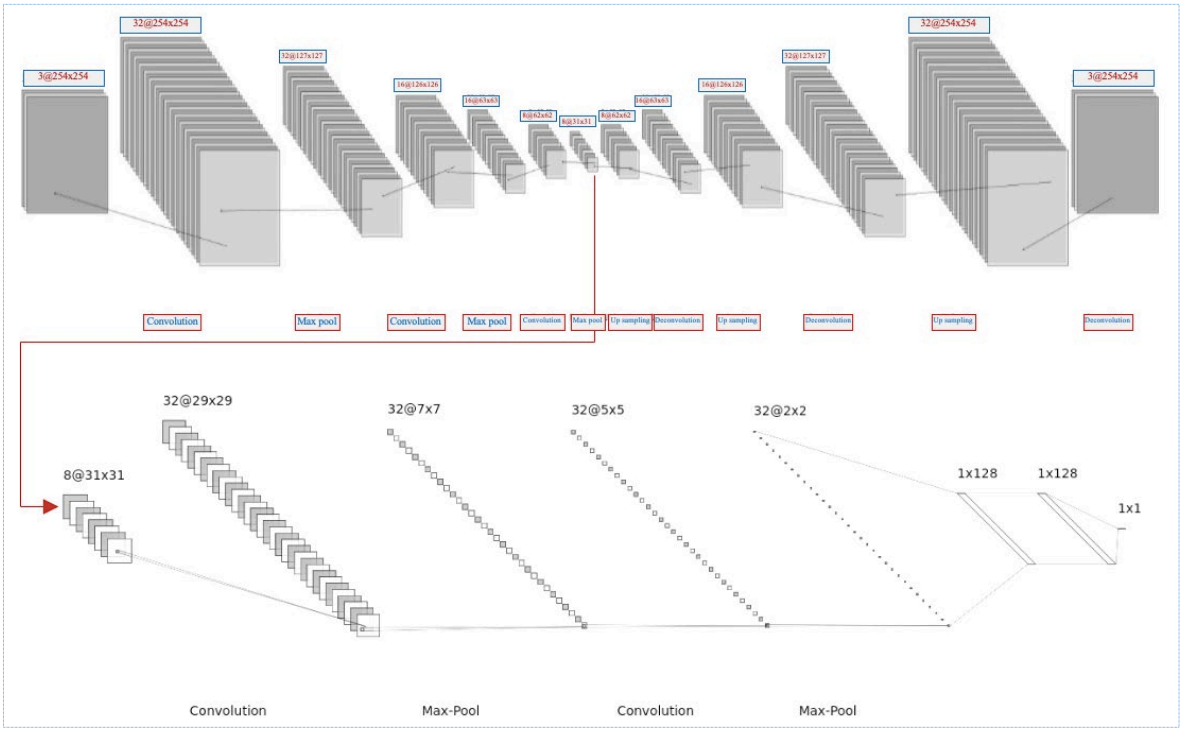

Fig. 6. The hybridization of CNN and AE model (CAE)

The illustration in Fig. 6 explains the initial stage of taking an X-ray image collected from GitHub datasets. Then the image dimensions are converted into dimensions 254x254. The convolution process transforms the images into 32 with image dimensions of 254x254, and the dimensional reduction stage transforms the image into 127x127. That means the filtering process uses a 2x2 mechanism. Then the sequential process transforms the image into 126x126, 63x63, 62x62, 31x31. The last convolutional process taken is 31x31, and this dimension is the smallest image convolutional result. The next stage is transforming the image dimension into larger dimension sizes with the deconvolutional process. The deconvolutional process is similar to a convolutional process, but the main difference is that this aims to enlarge the dimensions of the image, in contrast to the convolutional process aiming to reduce image dimensions.

The second essential task is putting the last convolutional process results in terms of 8@31x31 image dimensions and the transforming process into larger slicing and reducing the dimensional image into 29x29 dimensions. The further process is to reduce the image dimension into a smaller one to generate only one possibility, including 32@7x7, 32@5x5, 32@2x2, and then the flattening process to convert image dimensions into single vectors as a result of classification tasks.

2.1.3. CAE-COVIDX framework

The explanation above is the essence of the CAE engine's basic mechanism work, which is a combination of CNN and AE. CAE machines can normally work in order to detect a patient is infected with COVID-19 or not. Indeed, it requires several stages and some elements to complete the tasks. Following Fig. 7, several steps include collecting datasets. This experiment uses datasets of patient lung images from GitHub, involving 400 normal patients and 400 images of patients infected with COVID-19. Pre-processing is needed so that the image can be calculated by the feature learning that the CAE model will carry out.

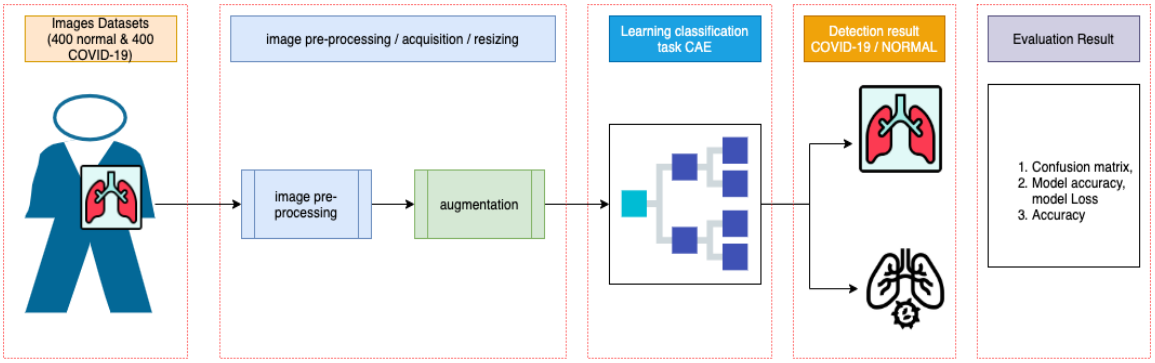

Fig. 7. CAE-COVIDX architecture

The final process is to evaluate the detection results of CAE-COVIDX by adopting three famous evaluation methods that are often used by researchers in classification cases, such as confusion matrix, accuracy. The second section of the evaluation step compares the performance of the CAE-COVIDX model with several existing state-of-the-art methods.

2.2. Evaluation Metrics

Specific metrics were recorded for the CNNs classification task: (a) positive diseases correctly classified (True Positives, TP), (b) negative diseases incorrectly classified (False Negatives, FN), (c) negative diseases correctly classified (True Negatives, TN), and (d), positive diseases incorrectly classified (False Negatives, FP). According to the procedure, TP denotes cases of COVID-19 that are appropriately expected, FP denotes normal or pneumonia cases classified as COVID-19 cases by CNN, and TN denotes normal or pneumonia cases categorized as non-COVID-19 cases. On the other hand, FN refers to COVID-19 cases that are classified as normal or common pneumonia. Due to the study's primary objective of detecting COVID-19, two distinct accuracies are measured. The first accuracy is referred to as the accuracy of three levels and refers to the model's total precision when the three divisions (normal pneumonia-COV) are distinguished. The second precision refers exclusively to COVID-19 accuracy. In other words, if a case is considered typical and pneumonia is classified as such by CNN, then the COVID-19 existence is deemed appropriate. As mentioned previously, precision is referred to as 2-class precision.

We developed two task models in practice that distinguish COVID-19-infected patients from bacteria-infected patients and healthy controls, respectively. For each mission, the patient-level was divided randomly into 60% to train the models, 10% validation sets to fine-tune the hyperparameter for optimal performance, and 30% of the test sets to independently tested the final optimal models. Furthermore, the evaluations were measured based on the Precision, Accuracy, Recall, and confusion matrix. Equations (4), (5), and (6) represent the evaluation model scenario's symbols.

$$\text{Accuracy} = \frac{TP+TN}{TP+FP+TN+FN} \quad (4)$$

$$Precision = \frac{TP}{TP+FP} \quad (5)$$

$$Recall = \frac{TP}{TP+FN} \quad (6)$$

TP represents the true positives, TN for the true negatives, FP as the false positives, and FN for the false negatives. We have also measured computation level efficiency to achieve the accuracy using the loss model and accuracy, including the number of literacies to reach convergence.

2.3. Datasets of Chest X-ray

Several X-ray references have been accessed for the experiments. Next, the GitHub repository for the associated datasets was examined. A set of Cohens X-ray images [33] has been chosen. Both photos and data are currently published in https://github.com/ieee8023/covid-chestxray-dataset. According to the explanation above, the collected images are ready for public consumption. The data collected comprises 400 images of the COVID-19 patients, 400 images of common pneumonia confirmed patients, and 504 normal patients.

2.4. Pre-processing

The X-ray images have been resized to 254x254 pixels. To obtain flawless rescaling to 200x266, a black background ratio of 1:1.5 was applied to the images of various pixel rations to prevent distortions. The reader should note that the CNNs can ignore a small positional variance, which means they look for patterns in particular image positions and moving patterns.

The basic X-ray necessary resized and rescaled before processing to the input layer of the network. There is different characteristic of every deep learning platform, for instance, SqueezeNet needs to be resized into 227x227 pixels. Other platforms, such as VGG16, VGG19, ResNet, and DenseNet require 224 x 224 pixels, while Inceptionv3 requires to be rescaled into 299 x 299 pixels size. Consider COVID-19 positive chest X-ray includes 400, also 400 set number of normal X-ray images in this experiment.

2.5. Tools and Library

Image processing is categorized as high computation research. In recent years, deep learning application leads to the use of GPU to calculate classification tasks. An NVIDIA I7-8700k central processing unit (CPU) is used as a trial server, but GPU GeForce GTX 1080ti has also been included. The time to stop relied primarily on the number of image layers in one set of CTs. On average, for a CT set with 70 layers, it took less than 30 seconds to go from preprocessing data to report output. Additional deep-learning modules including Keras, TensorFlow, and Python 3 are essential tools to achieve this research successfully.

3. Results and Discussion

As a critical tool for diagnosis, deep learning techniques for pneumonia identification and classification can accurately distinguish between non-COVID-19 pneumonia and infectious pneumonia COVID-19; designing a suitable approach for reducing time and unnecessary resources is interesting. In this research, We developed a pneumonia classification model based on 800 datasets, which are divided into 400 COVID-19 infected and 400 COVID-19 non-infected samples and 800 X-ray images obtained from GitHub. To demonstrate the utility of our proposed CAE-COVIDX architecture, we compared it to traditional CNN and VGG16.

As seen in Table 2, our CAE-COVIDX outperforms traditional CNN and popular deep learning frameworks based on CNN (VGG16). According to the experiment report, CAE-COVIDX also successfully increases accuracy by around 2-4 percent. To ensure performance stability, we conducted the training experiment 5 times, including CNN, VGG16, and CAE-COVIDX. Table 2 shows that CAE-COVIDX is superior to previous popular work in 5 training scenarios where our method produces fairly stable accuracy results at 0.98.

Table 2. Accuracy performance results

Section	CNN	VGG16	CAE-COVIDX
1	0.96	0.95	**0.98**
2	0.94	0.95	**0.98**
3	0.96	0.96	**0.98**
4	0.94	0.96	**0.98**
5	0.96	0.96	**0.98**

As demonstrated in Table 3, CAE-COVIDX achieved better performance and is stable enough compared to CNN and VGG16 in several schematic training model scenarios due to efficient information extraction. These deep learning models' high output also revealed well-distinguished CT images of patients with pneumonia and healthy people. The generic CNN model and VGG16 achieved unstable performance with lower effectivity compared to the CAE-COVID19 model. It can be inferred that two training scenarios inform different results where CNN and VGG16 reach low performances compared to the proposed model. Comparing CNN and VGG16, both of them have similar achievements in this training result. In all sections, CAE-COVIDX outperform both models. CNN and VGG16 share performance results where they almost produce the same performance.We conducted three training evaluation scenarios based on the confusion matrix to evaluate random effect performance in the image selection process. The experiment reported in Fig. 8 shows that our proposed model achieves significant performance over previous work based on confusion matrix evaluation.

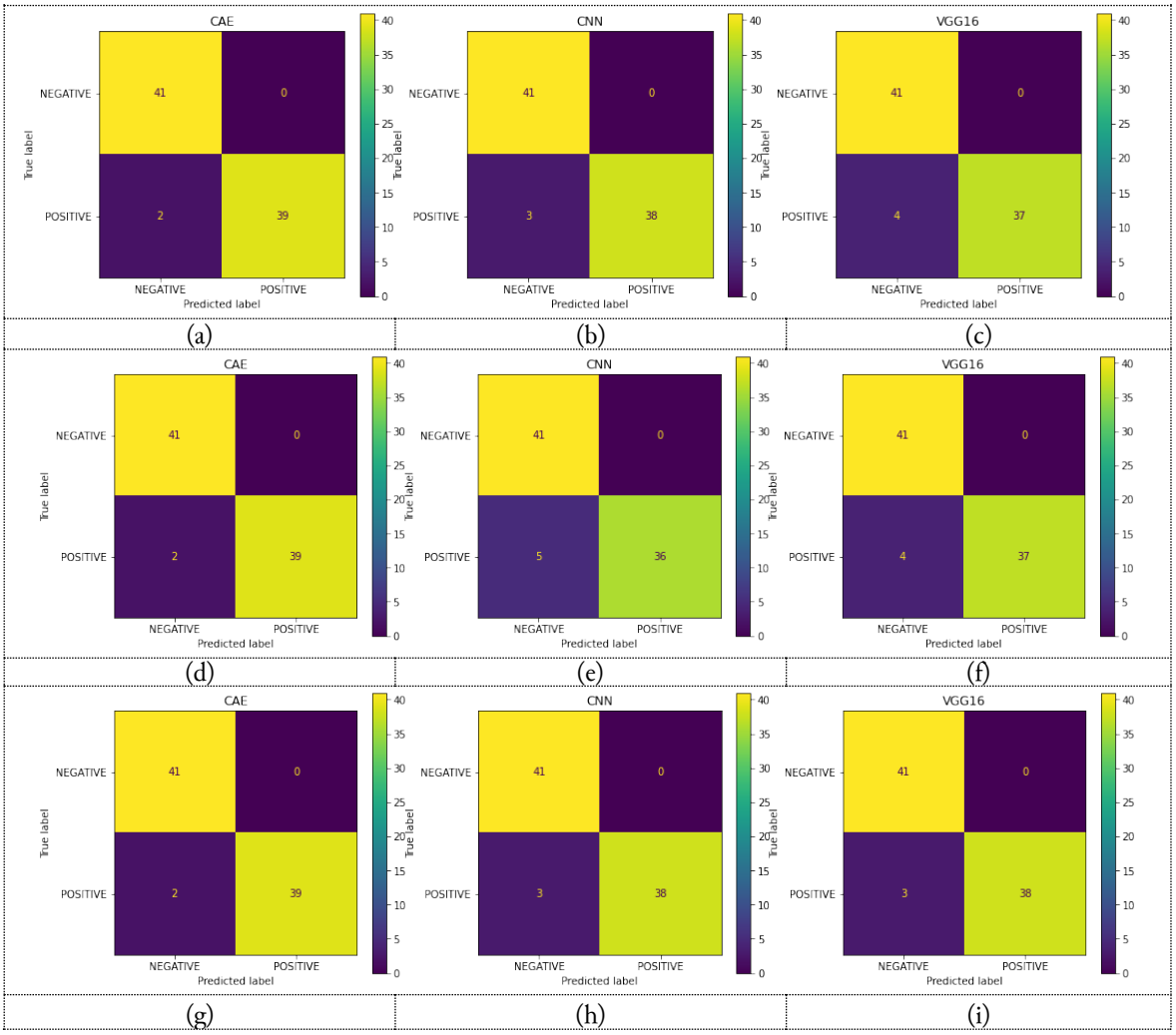

Fig. 8. Evaluation of confusion matrix training result. (a), (b), and (C) are first evaluation; (d), (e), and (f) are the second evaluation; and (g), (h), and (i) are the third evaluation.

Following the confusion matrix test section, consider repeating until three times testing with random image selection. Our model achieves good performance, and this is demonstrated with the level of error result are very low. The minimum error matrix of this model shows in Fig. 8 (g). Compared to CNN (Fig. 8 (h)) and VGG16 (Fig. 8 (i)), CAE-COVIDX stays outperform with significant results in three testing sections. We believe that the impact of minimum error result due to feature extraction (AE) mixed with dimensional reduction (CNN). The lost information obtained by dimensional reduction (CNN platform) applied in common deep learning models can be eliminated by using the AE feature.

In the third evaluation model, we considered applying model accuracy and model loss to compare the proposed algorithm with previous works. The details of the results are shown in Fig. 9 and Fig. 10. According to this performance evaluation, CAE-COVIDX achieved the best performance compared to CNN and VGG16. However, our model requires more iteration to reach convergence. The number of iterations between VGG16 and CAE-COVIDX is almost similar. On the other hand, both models are categorized as high computation even though they achieved different performances where CAE-COVIDX performed better in 2-time training scenarios.

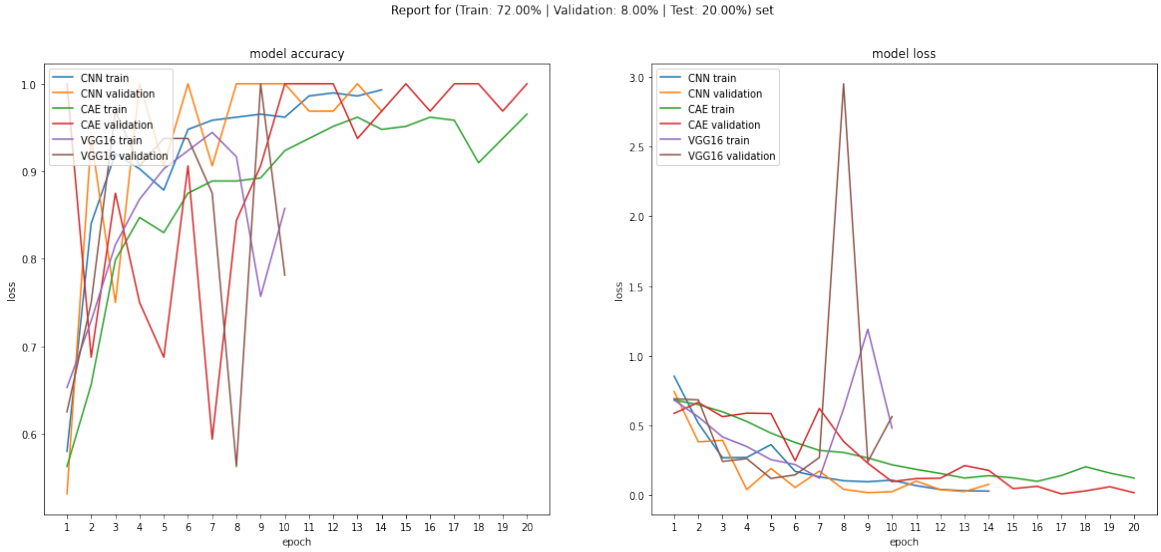

Fig. 9. First evaluation of accuracy and loss model

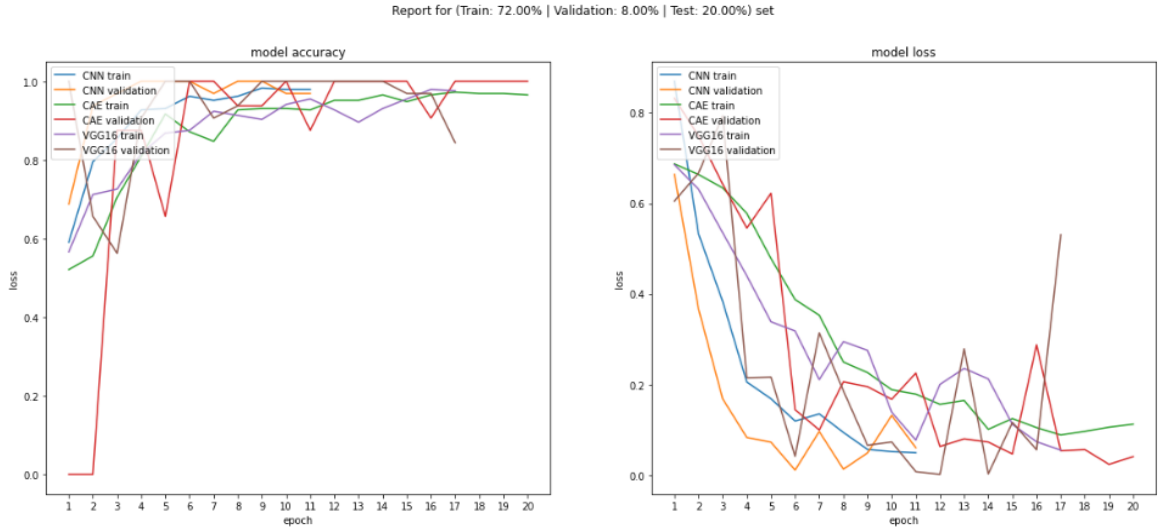

Fig. 10. The Second evaluation of accuracy and loss model

The second scenario (Fig. 10) shows that they reached almost similar results with previous evaluation work, including performance achievement and the number of iterations needed for all models. It can be

concluded that even though the image in training is randomly chosen, the proposed model performance achieves stability and also the better performance with the famous algorithm.

4. Conclusion

This research shows the deep learning approach viability to help medical employees, and doctors diagnose COVID-19 patients and automatically recognize lesions from X-ray and CT-scan images. Additional feature extraction features by AE into CNN have proven in accuracy performance. Moreover, another advantage of CAE-COVIDX is the result of the classification tasks more consistent even the input of the X-ray images is not standard. The proposed system achieved better performance to detect and classify both images of common pneumonia and Covid-19 images. Our analysis has some drawbacks due to the relatively large number of variable objects, especially those outside the lungs that are irrelevant for pneumonia diagnosis; it caused the classification of CT images to be complicated. In our research, only one radiologist worked out of the ROI region. The training data collection is comparatively limited. This system's efficiency is expected to improve as training intensity increases. It should also be remembered that, at later stages of disease progression, we examined the characteristics of the CT images from severe lung lesions patients. Research to correlate this with improvement and all COVID-19 pathological phases is required to refine the diagnostic method. It is challenging to conduct future research in hierarchical links between CT image features and future considerations such as epidemiological and clinical multi-omics, genomics, and multi-modeling for improved diagnostic accuracy.

Acknowledgments

Thanks to Universitas Amikom Yogyakarta to fully support this research, Chinese Government Scholarship (CGS) and Hohai University, and all institutions have helped us complete this research. We would also thank Cohen, who has collected the valuable datasets of medical images that share in the public online repository through GitHub.

Declarations

Author contribution. All authors contributed equally to the main contributor to this paper. All authors read and approved the final paper.
Funding statement. None of the authors have received any funding or grants from any institution or funding body for the research.
Conflict of interest. The authors declare no conflict of interest.
Additional information. No additional information is available for this paper.

References

[1] WHO, "WHO Coronavirus Disease (COVID-19) Dashboard," 2020. Available at: Google Scholar

[2] M. E. H. Chowdhury *et al.*, "Can AI Help in Screening Viral and COVID-19 Pneumonia?," *IEEE Access*, vol. 8, pp. 132665–132676, 2020, doi: 10.1109/ACCESS.2020.3010287.

[3] J. L. Strunk, H. Temesgen, H. Andersen, and P. Packalen, “In Pr es s In es,” vol. 80, no. 2, pp. 1–8, 2020. Available at: Google Scholar

[4] Hanafi, N. Suryana, and A. S. B. H. Basari, "An understanding and approach solution for cold start problem associated with recommender system: A literature review," *J. Theor. Appl. Inf. Technol.*, vol. 96, no. 9, 2018. Available at: Google Scholar

[5] I. Goodfellow, Y. Bengio, and A. Courville, "Deep Learning," *ICML2013 Tutor.*, pp. 1–800, 2011, doi: 10.1038/nmeth.3707.

[6] Y. LeCun, K. Kavukcuoglu, and C. Farabet, "Convolutional Networks and Applications in Vision," *Environ. Sci. Technol.*, vol. 10, no. 3, pp. 256–262, 1997, doi: 10.1021/es60114a003.

[7] Hanafi, N. Suryana, and A. S. B. H. Basari, "Deep learning for recommender system based on application domain classification perspective: A review," *J. Theor. Appl. Inf. Technol.*, vol. 96, no. 14, 2018. Available at:

[8] L. Wang and A. Wong, "COVID-Net: A Tailored Deep Convolutional Neural Network Design for Detection of COVID-19 Cases from Chest X-Ray Images," pp. 1–12, 2020, doi: 10.1038/s41598-020-76550-z

[9] H. Hou *et al.*, "Pr es s In Pr," *Appl. Intell.*, vol. 2019, pp. 1–5, 2020, doi: 10.1007/s10489-020-01714-3.

[10] Y. Song *et al.*, "Deep learning Enables Accurate Diagnosis of Novel Coronavirus (COVID-19) with CT images," *medRxiv*, p. 2020.02.23.20026930, 2020, doi: 10.1101/2020.02.23.20026930.

[11] J. Chen *et al.*, "Deep learning-based model for detecting 2019 novel coronavirus pneumonia on high-resolution computed tomography: a prospective study," 2020, doi: 10.1101/2020.02.25.20021568.

[12] A. Abbas, M. M. Abdelsamea, and M. M. Gaber, "Classification of COVID-19 in chest X-ray images using DeTraC deep convolutional neural network," pp. 1–9, 2020, doi: 10.1101/2020.03.30.20047456

[13] M. Toğaçar, B. Ergen, and Z. Cömert, "COVID-19 detection using deep learning models to exploit Social Mimic Optimization and structured chest X-ray images using fuzzy color and stacking approaches," *Comput. Biol. Med.*, vol. 121, no. May, 2020, doi: 10.1016/j.compbiomed.2020.103805.

[14] S. Bharati, P. Podder, and M. R. H. Mondal, "Hybrid deep learning for detecting lung diseases from X-ray images," *Informatics Med. Unlocked*, vol. 20, p. 100391, 2020, doi: 10.1016/j.imu.2020.100391.

[15] X. Wu *et al.*, "Deep learning-based multi-view fusion model for screening 2019 novel coronavirus pneumonia: A multicentre study," *Eur. J. Radiol.*, vol. 128, no. April, pp. 1–9, 2020, doi: 10.1016/j.ejrad.2020.109041.

[16] A. A. Ardakani, A. R. Kanafi, U. R. Acharya, N. Khadem, and A. Mohammadi, "Application of deep learning technique to manage COVID-19 in routine clinical practice using CT images: Results of 10 convolutional neural networks," *Comput. Biol. Med.*, vol. 121, no. April, p. 103795, 2020, doi: 10.1016/j.compbiomed.2020.103795.

[17] C. Jin *et al.*, "Development and Evaluation of an AI System for COVID-19 Diagnosis," *medRxiv*, p. 2020.03.20.20039834, 2020, doi: 10.1101/2020.03.20.20039834.

[18] S. Wang *et al.*, "A deep learning algorithm using CT images to screen for corona virus disease (COVID-19)," *medRxiv*, p. 2020.02.14.20023028, 2020, doi: 10.1101/2020.02.14.20023028.

[19] X. Xu *et al.*, "A Deep Learning System to Screen Novel Coronavirus Disease 2019 Pneumonia," *Engineering*, no. xxxx, pp. 1–8, 2020, doi: 10.1016/j.eng.2020.04.010.

[20] F. Shi *et al.*, "Large-Scale Screening of COVID-19 from Community Acquired Pneumonia using Infection Size-Aware Classification," 2020, doi: 10.1088/1361-6560/abe838

[21] Z. Tang *et al.*, "Severity Assessment of Coronavirus Disease 2019 (COVID-19) Using Quantitative Features from Chest CT Images," vol. 2019, pp. 1–18, 2020. Available at: Google Scholar

[22] B. Ghoshal and A. Tucker, "Estimating Uncertainty and Interpretability in Deep Learning for Coronavirus (COVID-19) Detection," pp. 1–14, 2020. Available at: Google Scholar

[23] I. D. Apostolopoulos and T. A. Mpesiana, "Covid-19: automatic detection from X-ray images utilizing transfer learning with convolutional neural networks," *Phys. Eng. Sci. Med.*, vol. 43, no. 2, pp. 635–640, 2020, doi: 10.1007/s13246-020-00865-4.

[24] T. Ozturk, M. Talo, E. A. Yildirim, U. B. Baloglu, O. Yildirim, and U. Rajendra Acharya, "Automated detection of COVID-19 cases using deep neural networks with X-ray images," *Comput. Biol. Med.*, vol. 121, no. April, p. 103792, 2020, doi: 10.1016/j.compbiomed.2020.103792.

[25] S. R. A. Amyar , R. Modzelewski, "Multi-Task Deep Learning Based CT Imaging Analysis for Covid-19 : Classification and Segmentation General Electric Healthcare , Buc , France LITIS - EA4108 - University of Rouen , Rouen , France Nuclear Medicine Department , Henri Becquerel Center , Rouen," 2020, doi: 10.1101/2020.04.16.20064709

[26] H. S. Maghdid, A. T. Asaad, K. Z. Ghafoor, A. S. Sadiq, and M. K. Khan, "Diagnosing COVID-19 Pneumonia from X-Ray and CT Images using Deep Learning and Transfer Learning Algorithms," pp. 1–8, 2020. Available at: Google Scholar

[27] O. Gozes, M. Frid-Adar, N. Sagie, H. Zhang, W. Ji, and H. Greenspan, "Coronavirus Detection and Analysis

on Chest CT with Deep Learning," pp. 1–8, 2020. Available at: Google Scholar

[28] R. M. Pereira, D. Bertolini, L. O. Teixeira, C. N. Silla, and Y. M. G. Costa, "COVID-19 identification in chest X-ray images on flat and hierarchical classification scenarios," *Comput. Methods Programs Biomed.*, p. 105532, 2020, doi: 10.1016/j.cmpb.2020.105532.

[29] J. Zhang, Y. Xie, Y. Li, C. Shen, and Y. Xia, "COVID-19 Screening on Chest X-ray Images Using Deep Learning based Anomaly Detection," 2020. Available at: Google Scholar

[30] A. Jaiswal, N. Gianchandani, D. Singh, V. Kumar, and M. Kaur, "Classification of the COVID-19 infected patients using DenseNet201 based deep transfer learning," *J. Biomol. Struct. Dyn.*, vol. 0, no. 0, pp. 1–8, 2020, doi: 10.1080/07391102.2020.1788642.

[31] J. Masci, U. Meier, D. Cireşan, and J. Schmidhuber, "Stacked convolutional auto-encoders for hierarchical feature extraction," *Lect. Notes Comput. Sci. (including Subser. Lect. Notes Artif. Intell. Lect. Notes Bioinformatics)*, vol. 6791 LNCS, no. PART 1, pp. 52–59, 2011, doi: 10.1007/978-3-642-21735-7_7.

[32] Hanafi, N. Suryana, and A. Samad, "Dynamic convolutional neural network for eliminating item sparse data on recommender system," *IJAIN*, vol. 4, no. 3, pp. 226–237, 2018, doi: 10.26555/ijain.v4i3.291

[33] J. P. Cohen, P. Morrison, and L. Dao, "COVID-19 Image Data Collection," 2020. Available at: arxiv.org