Multimodal Convolutional Neural Networks for Detection of Covid-19 Using Chest X-Ray and CT Images

Abdelwhab Ouahab[a, b, *]

[a] Mathematics and Computer Science Department Faculty of Science and Technology University of Adrar, Adrar, 01000 Algeria
[b] Laboratoire Signaux, Systèmes et Données (LSSD), Université des Sciences et de la Technologie d'Oran, Bir El Djir, 31000 Algeria
*e-mail: ouahab.abdelwhab@univ-usto.dz

Abstract—The Covid-19 was first appeared in 2019 in Wuhan, China. It widely and rapidly expanded all over the world. Since then, it has had a strong effect on people's daily lives, the world economy and the public health. The fast prediction of Covid-19 can assist the medicine to choose the right treatment. In this paper, we propose a classification of Covid-19 using Models based on a Convolutional Neural Network (CNN). We propose two models to detect Covid-19. The first one uses CNN with CT or X-ray images separately. The second uses CNN with both CT and X-ray images at the same time. The used datasets contain X-ray and CT images divided into three classes which are Covid-19, Normal and Pneumonia. Each type image class has 1045 images for training and 300 for testing. All these data sets are available in Kaggle repository. In order to evaluate the proposed models, we calculate the confusion matrix, the accuracy, precision, recall and F1 score. The model that uses CNN with both X-ray and CT images of 0.99 achieves the best accuracy. We deduced that using CT images is more efficient than using X-ray images to predict Covid-19. The combination of the CT and X-ray images to detect Covid-19 is more efficient than using only CT or X-ray images. The proposed models could effectively assist the radiologists in predicting Covid-19.

1. INTRODUCTION

In 2019, a corona virus, named (COVID-19) appeared in Wuhan, China. This contagion virus spread rapidly in China and in the world, causing severe acute respiratory syndrome (SARS) in those who exposed to the virus [1]. Most people with the Covid-19 have symptoms like tiredness, fiver, loss of smell or taste and dry cough. When a person is infected by this virus, the symptoms appear after an average of 5 to 6 days [17].

Globally, as of 10:43 am CEST, 16 July 2021, there have been 188 332 972 confirmed cases of COVID-19, among of them 4 063 453 deaths, reported by the World Health Organization (WHO). As of 15 July 2021, a total of 3 402 275 866 vaccine doses have been administered.

Under these circumstances, it is very important to use a variety of strategies to enhance the accuracy in controlling the expansion of the pandemic and the level of information and communication between various health sectors to facilitate the healthcare procedures [2]. The earlier detection of the Covid-19 helps in choosing the appropriate treatment and patient care.

Because the Covid-19 infects principally the respiratory system, usage of chest radiography images in COVID-19 diagnosis is very important. Due to the huge increase in the number of injured, it is difficult for radiologists to process these images manually. Deep learning and other Artificial Intelligence (AI) methods have enormous potential for processing and extracting insights from large amounts of data [18−21].

In this context, a lot of systems have been developed to detect and estimate the infection by COVID-19, including Body temperature monitoring, Chest Computed Tomography (CT) scans and Chest X-rays (CXR) [3−5].

The CT scan is a procedure that uses X-rays and a computer to produce cross-sectional images of various structures inside of the human body. The CT scan can assist a doctor to determine the position, form and size of a tumor before making the radiotherapy [6]. CXR, in particular, is a widely used, relatively

inexpensive, quick, and accessible diagnostic modality that can be quickly carried to the patient's bedside, even in emergency rooms. As a result, CXR may be preferred over CT not only in resource-constrained settings, but also in situations when it is necessary to minimize treatment of more compromised [2].

The most of detection systems have used deep learning to classify CT and CXR images which are generally grouped into two or three classes. In [7], it uses active deep learning with CT images to identify COVID-19. In this method, a weakly-supervised deep active learning was performed using a 2D U-Net. The results obtained give an accuracy of 86.6%. In [8], it uses a set of deep learning models on CT images to predict covid-19. The principle contributions of this method are to choose optimal features which are extracted from two pre-trained models. The results obtained give an accuracy of 95.6%. The authors in [9], used SVM and multi-level thresholding to detect Covid-19 on X-ray images. The average accuracy using this method gives an accuracy of 97.48%. In [10], the authors proposed a deep learning model using AlexNet architecture which uses features selection. The experimental results give an accuracy of 99.18%. In [22], the authors proposed a modified structure of ResNet50 by adding three new layers at the end of its structure to extract the best features. This work achieved an accuracy of 97.7% on CT images and an accuracy of 97.1% on X-ray images.

In the aforementioned works, the authors have proposed models in which the features are extracted from one dataset (CT or X-ray images). This is explained by the lack of datasets that contain CT with X-ray images of the same patient and the CT or X-ray images were given in isolation.

In this paper, in order to enhance the accuracy of classification, we propose a multimodal deep learning approach for detection COVID-19 that explores both CT and X-ray images to extract the features. In order to highlight the importance of the proposed approach, we compare the results obtained by this model and those ones obtained by using only CT or X-ray images separately at same model.

2. PROPOSED FRAMEWORK

Two models of a Convolutional Neural Network (CNN) are proposed for the detection of COVID-19 images. The classification operation is used to differentiate between three classes which are normal, pneumonia and COVID-19. As shown in Table. 1, the first model uses CNN with only one datasets CT or X-ray images, whereas, as shown in Fig. 1, the second model uses CNN with both CT and X-ray images.

In the first model architecture, CNN consists of an input layer, convolutional layers, pooling layers, dropout layers and a classification layer. In the second model architecture, we applied the first model on CT image and X-ray image separately, but before the classification step, we concatenate the Flatten layers of the model when it is applied on CT and X-ray images. The input layer size is 128×128.

Convolutional layer. The convolutional layer is the important layer of CNN that uses convolution operations to extract the features from the input images [6]. In the proposed models, each convolutional layer has kernel of size 3×3.

Pooling layer. Pooling layer is applied to reduce the size of the feature maps. It summarizes the features extracted by the convolutional layer. Max pooling is the operation of selecting the maximum value in the feature maps where it is scanned by a filter. Max pooling is used in order to create features with less noise [11]. In the proposed models, each Max pooling layer has a pool size of 2×2.

Dropout layer. Dropout is a regularization developed by Geoff Hinton technique by randomly removing connections between selected nodes during the training phase. The role of this operation is to prevent overfitting [12]. In our case we have used a dropout of 0.2.

Flatten layer. Flatten layer is a layer where the entire input is converted into one dimension [11].

Classification layer. This layer converts the flatten layer into classes. The soft max activation function is applied to produce the classes. Adam optimizer is used to calculate the weights of the CNN according to learning data and sparse categorical cross entropy a cost function.

2.1. Dataset Descriptions

The training data set contains 1045 CT images and 1045 images X-ray images for each class (normal, pneumonia and COVID-19). The test data set contain 300 CT images and X-ray images for each class. The sizes of X-ray images and CT images are a 299×299 and 512×512 respectively. The X-ray and CT images have been collected by [14] and [15] respectively. All these data sets are available in Kaggle repository. Sample images of X-ray and CT images are shown in Fig. 2 and Fig. 3.
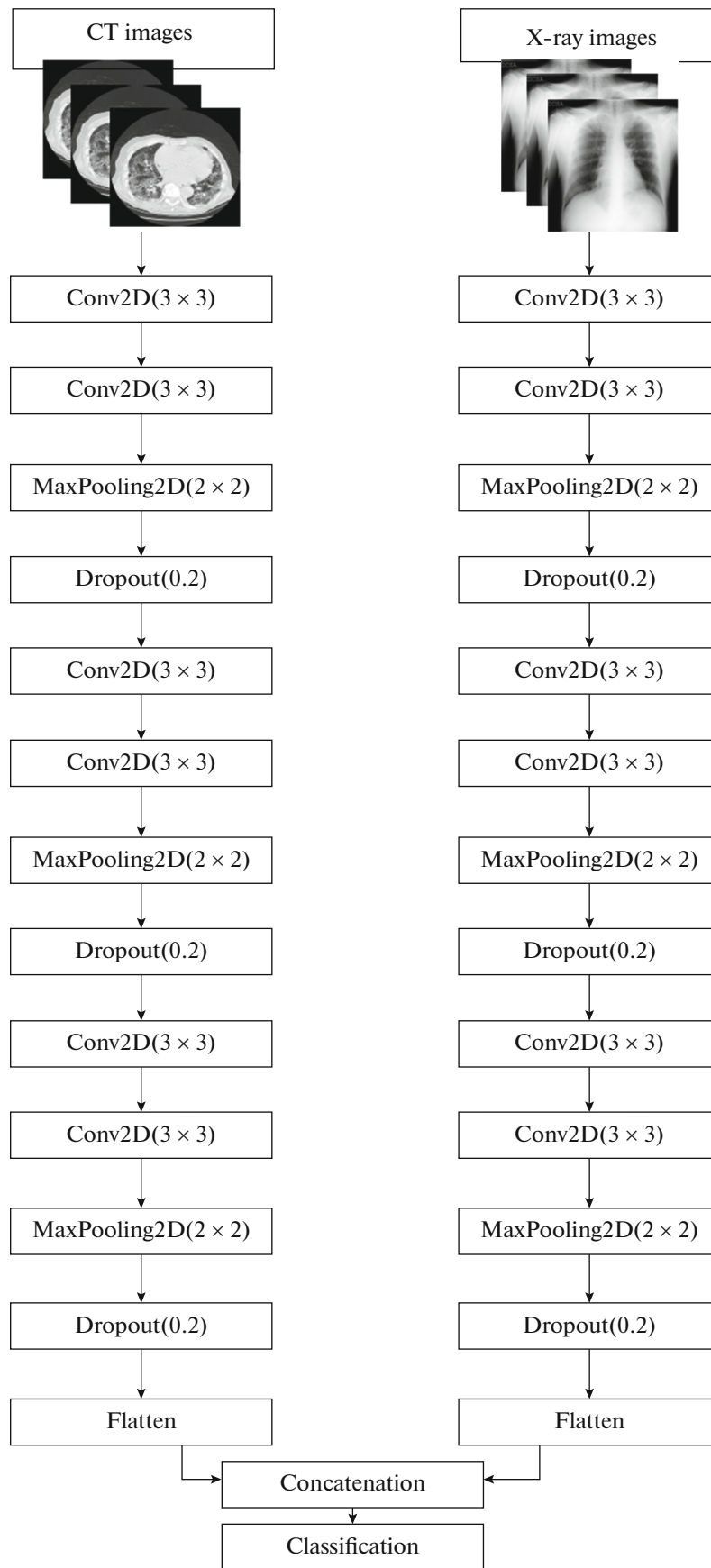
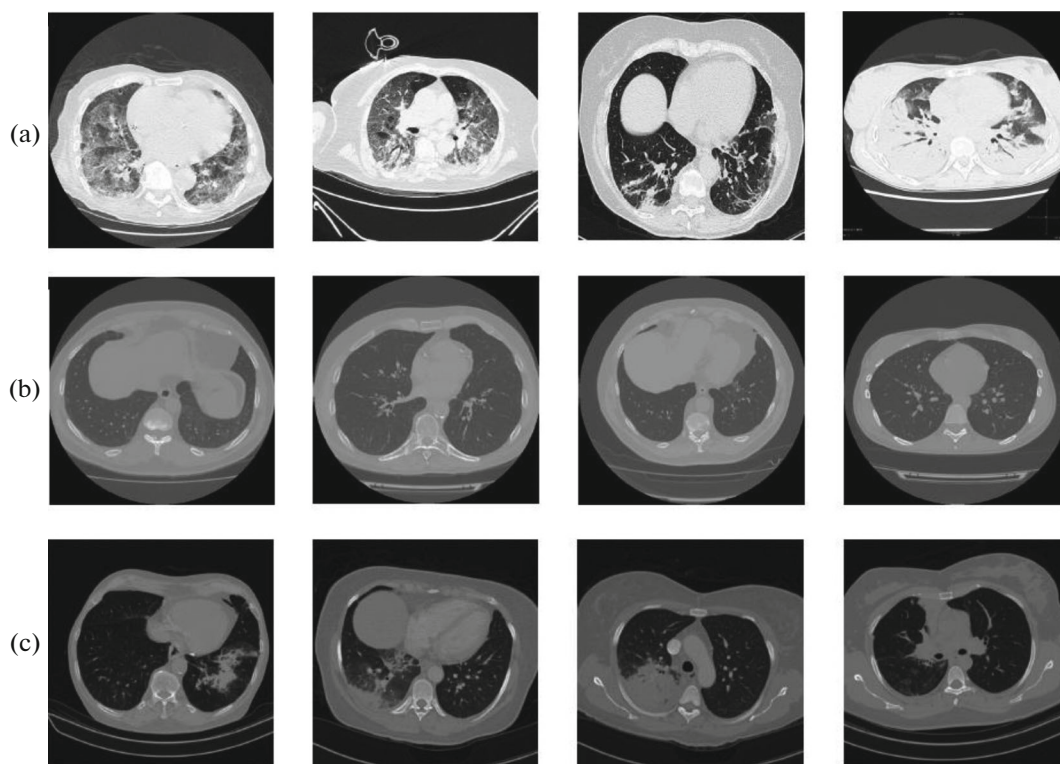**Fig. 1.** A Multimodal Convolutional Neural Network architecture with X-ray and CT images (Model 2).

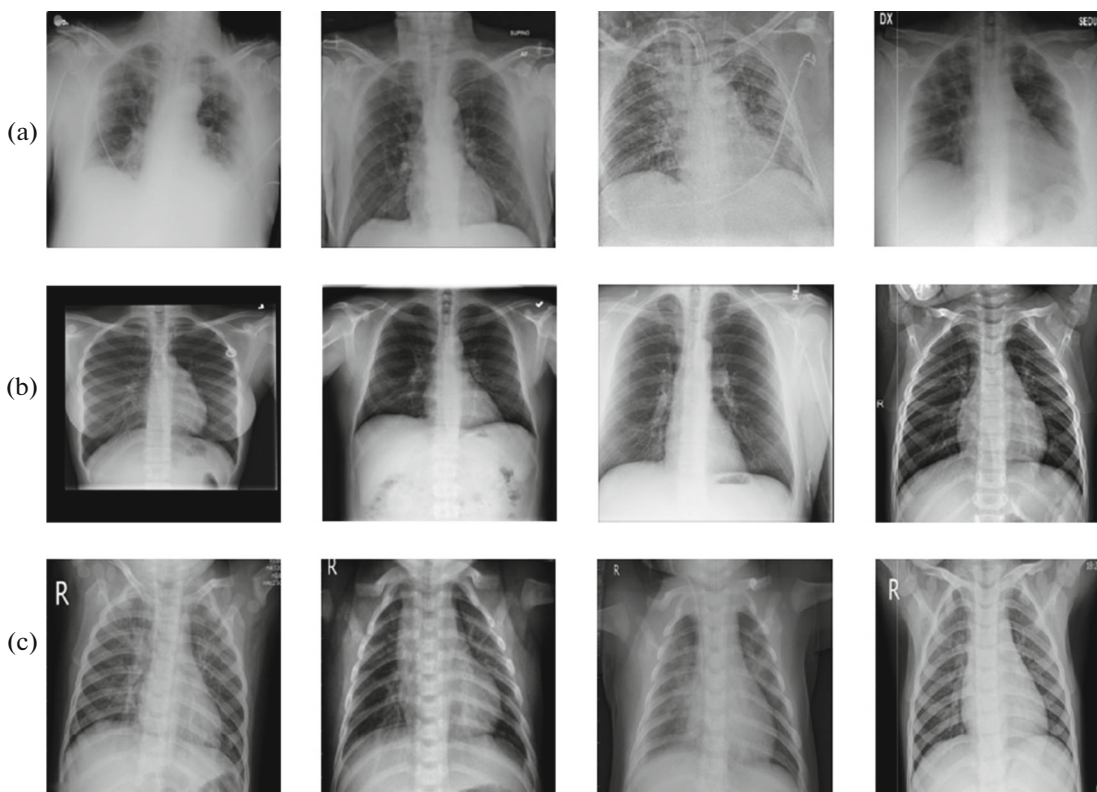**Fig. 2.** Samples of CT images: (a) Covid-19, (b) Normal, (c) Pneumonia.

Fig. 3. Samples of X-ray images: (a) Covid-19, (b) Normal, (c) Pneumonia.

Table 1. Convolutional Neural Network architecture with X-ray or CT images (Model 1)

	Layer	Kernel size	Activation
Input (CT or X-ray) images	28 × 28 × 3		
Conv2D	16	3 × 3	Relu
Conv2D	16	3 × 3	Relu
MaxPooling2D		2 × 2	
Dropout (0.2)			
Conv2D	32	3 × 3	Relu
Conv2D	32	3 × 3	Relu
MaxPooling2D		2 × 2	
Dropout(0.2)			
Conv2D	64	3 × 3	Relu
Conv2D	64	3 × 3	Relu
MaxPooling2D		2 × 2	
Dropout (0.2)			
Flatten()			

3. RESULTS

In this section, we discuss and explain the results obtained by the implemented models. We have implemented the proposed models for COVID-19 detection using Python programming language with Tensorflow framework, processor of Intel Core i5 and RAM of 8 GB running on Windows 7.

The aforementioned data sets are used with two experiments.

In the first experiment the data set is decomposed into three classes. In the second experiment the data set is decomposed into two classes. Each experiment has two scenarios which are:

Scenario 1: Model 1 with CT images (Table 1).

Scenario 2: Model 1 with X-ray images (Table 1).

Scenario 3: Model 2 with CT and X-ray images (Fig. 1).

The required preprocessing is applied to the dataset before it is used in the training and testing steps. Since the images in the dataset have different dimensions, we cannot process them. Hence, all images are scaled to a size of 128 × 128.

3.1. Evaluation Metrics

The results are reported according to confusion matrix, precision, recall F1 score and accuracy.

3.1.1. Accuracy. Accuracy is the fraction of samples that got correct to the total number of predictions. It has the following definition [16]:

$$\text{Accuracy} = \frac{(TP + TN)}{TP + TN + FP + FN}. \tag{1}$$

3.1.2. Precision. Precision is the ratio of positive samples that predicted corrects to the total number of the samples predicted positives.

It is represented as:

$$\text{Precision} = \frac{(TP)}{TP + FP}. \tag{2}$$

3.1.3. Recall (or Sensitivity). Recall is the ratio of positive samples that predicted corrects to the total number of positive samples. It is defined as:

$$\text{Recall} = \frac{(TP)}{TP + FN}. \tag{3}$$

3.1.4. F1 score. F1 score is the weighted mean of Recall and Precision. It takes into consideration both false negatives and false positives. It is defined as:

Table 2. The confusion matrixes of the used models with three classes

Classes	Model-1 with X-ray images (scenario 1)			Model 1 with CT images (scenario 2)			Model-2 with CT and X-ray images(scenario 3)		
Covid-19	260	30	10	300	0	0	300	0	0
	87%	10%	3%	100%	0%	0%	100%	0%	0%
Normal	33	265	2	0	300	0	0	300	0
	11%	88.33%	0.67%	0%	100%	0%	0%	100%	0%
Pneumonia	3	4	288	22	0	278	1	0	299
	1%	1.33%	96%	7.33%	0%	92.67	0.33%	0%	99.67%

Table 3. The accuracy value of each model with three classes

	Model-1 with X-ray images	Model 1 with CT images	Model-2 with CT and X-ray images
Accuracy	90%	98.37%	99%
Precision	93.89%	97.66%	99%
Recall	91%	97.66%	99%
F1 score	91.67%	97.33%	99%

$$\text{F1 score} = \frac{(2TP)}{2TP + FP + FN}, \tag{4}$$

where if the tested class is Covid-19 than:

TP (true positive): Person with covid-19 detected positive.

TN (true negative): Person without covid-19 detected negative.

FP (False positive): Person with covid-19 detected negative.

FN (false negative): Person with covid-19 detected negative.

3.1.5. The experiment 1. In this experiment, the used datasets is decomposed into three classes (Covid-19, normal and Pneumonia). The proposed models are trained for 100 epochs.

Table 3 shows:

The average accuracy, precision, recall and F1 score for scenario 1 were 90, 93.89, 91 and 91.67% respectively.

The average accuracy, precision, recall and F1 score for scenario 2 were 98.37, 97.66, 97.66 and 97.33 respectively.

The average accuracy, precision, recall and F1 score for scenario 3 where 99%.

The best results are achieved when we have used the model-2 with CT and X-ray images. This is explained by the benefit of multi-model classification.

From Table 2, it can be seen that the model 1 with CT image gives a few confusion between the Covid-19 and pneumonia classes. This confusion is highly reduced when using the model 2 with CT and X-ray images. The model 1 with X-ray images gives the lowest values of the accuracy. This is explained by the fact that the use of X-ray images to detect Covid-19 is less efficient than those of CT images.

The combination of the CT and X-ray images to detect Covid-19 is more efficient than using only CT or X-ray images.

3.1.6. The experiment 2. In this experiment the used data set is decomposed into two classes (Covid-19 and non-Covid).We combined the classes Normal and Pneumonia into one class.

Table 5 shows:

The average accuracy, precision, recall and F1 score for scenario 1 were 88.33, 88.8, 88.5 and 88% respectively.

The average accuracy, precision, recall and F1 score for scenario 2 were 88.3, 89, 91 and 88.5% respectively.

The average accuracy, precision, recall and F1 score for scenario 3 were 93.83, 94, 94.5 and 93.5% respectively.

Table 4. The confusion matrixes of the used models with two classes

	Model-1 with X-ray images		Model-1 with CT images		Model-2 with CT and X-ray images	
Covid-19	262 87.33%	38 12.66%	300 100%	0 0%	300 100%	0 0%
Non-Covid	34 11.33%	266 88.66%	67 22.33	233 77.66	37 12.33%	263 87.66%

Table 5. The accuracy value of each model with two classes

	Model-1 with X-ray images	Model 1 with CT images	Model-2 with CT and X-ray images
Accuracy	88.83%	88.3%	93.83%
Precision	88.83%	89%	94%
Recall	88.5%	91%	94.5%
F1 score	88	88.5%	93.5%

The best results are achieved when we have used the model-2 with CT and X-ray images. This is explained by the benefit of combining CT and X-ray images for classification.

Table 4 shows the percentage of Covid-19 samples that were predicted correctly is 100% for scenario 2 and scenario 3. This value is more interesting, because in fact, detecting Covid-19 cases is more important than detecting Non-Covid cases.

4. CONCLUSIONS

In this work, we proposed two models to detect Covid-19. The first one uses CNN with CT or X-ray images separately. The second uses CNN with both CT and X-ray images at the same time. To evaluate the performance of each model, we calculated the confusion matrix, accuracy, precision, recall and F1 score. The use of the multi-model CNN achieves an accuracy of 99 and 93.83% for three classes and two classes respectively. These accuracies are better than other models. The proposed models could effectively assist the radiologic in predicting Covid-19 specially the second model.

REFERENCES

1. Chaddad, A., Hassan, L., and Desrosiers, C., Deep CNN models for predicting COVID-19 in CT and x-ray images, *J. Med. Imaging*, 2021, vol. 8, no. S1, pp. 014502.
   https://doi.org/10.1117/1.JMI.8.S1.014502
2. Signoroni, A., Savardi, M., Benini, S., Adami, N., Leonardi, R., Gibellini, P., Vaccher, F., Ravanelli, M., Borghesi, A., Maroldi, R., and Farina, D., BS-Net: Learning COVID-19 pneumonia severity on a large chest X-ray dataset, *Med. Image Anal.*, 2021, vol. 71, pp. 102046.
   https://doi.org/10.1016/j.media.2021.102046
3. Brunese, L., Martinelli, F., Mercaldo, F., and Santone, A., Machine learning for coronavirus covid-19 detection from chest x-rays, *Procedia Comput. Sci.*, 2020, vol. 176, pp. 2212−2221.
   https://doi.org/10.1016/j.procs.2020.09.258
4. Ingacheva, A.S. and Buzmakov, A.B. Methods of preprocessing tomographic images taking into account the thermal instability of the X-ray tube, *Optoelectron. Instrum. Proc.,* 2019, vol. 55, pp. 138−147.
   https://doi.org/10.3103/S8756699019020055
5. Zolotov, D.A., Asadchikov, V.E., Buzmakov, A.V., et al., X-ray diffraction tomography using laboratory sources for studying single dislocations in a low absorbing silicon single crystal, *Optoelectron. Instrum. Proc.*, 2019, vol. 55, pp. 126−132.
6. El-Shafai, W., El-Hag, N., El-Banby, G., and Soliman, A., An efficient cnn-based automated diagnosis framework from covid-19 ct images, *Comput., Mater. Continua*, 2021, vol. 69, no. 1, pp. 1323−1341.
7. Wu, X., Chen, C., Zhong, M., Wang, J., and Shi, J., COVID-AL: The diagnosis of COVID-19 with deep active learning, *Med. Image Anal.*, 2021, vol. 68, 101913.
   https://doi.org/10.1016/j.media.2020.101913
8. Majid, A., Khan, M., Nam, Y., Tariq, U., and Roy, S., Covid19 classification using ct images via ensembles of deep learning models, *Comput., Mater. Continua*, vol. 69, no. 1, pp. 319−337.

9. Hassanien, A., Mahdy, L., Ezzat, K., Elmousalami, H., and Ella, H., Automatic x-ray covid-19 lungimage classification system based on multi-level thresholding and support vector machine, *medRxiv*, 2020.
https://doi.org/10.1101/2020.03.30.20047787

10. Turkoglu, M., COVIDetectioNet: COVID-19 diagnosis system based on X-ray images using features selected from pre-learned deep features ensemble, *Appl. Intell.*, 2021, vol. 51, no. 3, pp. 1213−1226.

11. Alla, S. and Adari, S., Beginning anomaly detection using python-based deep learning: With Keras and Py-Torch, *APRESS*.
https://doi.org/10.1007/978-1-4842-5177-5

12. Chollet, F., Deep Learning with Python, ISBN 9781617294433.

13. Chowdhury, M., Rahman, T., Khandakar, A., Mazhar, R., Kadir, M., Mahbub, Z., Islam, K., Khan, M., Iqbal, A., Al-Emadi, N., Reaz, M., and Islam, M., Can AI help in screening Viral and COVID-19 pneumonia, *IEEE Access.*, 2020, vol. 8, pp. 132665−132676.

14. Rahman, T., Khandakar, A., Qiblawey, Y., Tahir, A., Kiranyaz, S., Kashem, S.B.A., Islam, M.T., Maadeed, S.A., Zughaier, S.M., Khan, M.S., and Chowdhury, M.E., Exploring the effect of image enhancement techniques on COVID-19 detection using chest X-ray images, *Comput. Biol. Med.*, 2020, vol. 132, pp. 104319.
https://doi.org/10.1016/j.compbiomed.2021.104319

15. Maftouni, M., Law, A.C, Shen, B., Zhou, Y., Yazdi, N., and Kong, Z.J., A robust ensemble-deep learning model for COVID-19 diagnosis based on an integrated CT scan images database, *Proceedings of the 2021 Industrial and Systems Engineering Conf., Virtual Conf.*, May 22−25, 2021.

16. Naveen, A., Manoj, B., Akhila, G., Nakarani, M., Sreekar, J., Beriwal, P., Gupta, N., and Narayanan, S., Deep learning techniques for detection of COVID-19 using chest X-rays, *Adv. Syst. Sci. Appl.*, 2021, vol. 21, no. 2, pp. 42−57.
https://doi.org/10.25728/assa.2021.21.2.1009

17. Singh, V.K. and Kolekar, M.H., Deep learning empowered COVID-19 diagnosis using chest CT scan images for collaborative edge-cloud computing platform, *Multimed Tools Appl.*, 2021.
https://doi.org/10.1007/s11042-021-11158-7

18. Alsharif, M.H., Alsharif, Y.H., Yahya, K., Alomari, O.A., Albreem, M.A., and Jahid A., Deep learning applications to combat the dissemination of COVID-19 disease, *Eur. Rev. Med. Pharmacol. Sci.*, 2020, vol. 24, no. 21, pp. 11455−11460.
https://doi.org/10.26355/eurrev_202011_23640

19. Georgakopoulos, S.V., Kottari, K., and Delibasis, K., Improving the performance of convolutional neural network for skin image classification using the response of image analysis filters, *Neural Comput. Appl.*, 2019, vol. 31, pp. 1805−1822.
https://doi.org/10.1007/s00521-018-3711-y

20. Nan, S., Feature extraction and segmentation processing of images based on convolutional neural networks, *Opt. Mem. Neural Networks*, 2021, vol. 30, pp. 67−73.
https://doi.org/10.3103/S1060992X21010069

21. Changgeng, Yu., Liu, K., and Zou, W., A method of small object detection based on improved deep learning, *Opt. Mem. Neural Networks*, 2020, vol. 29, pp. 69−76.
https://doi.org/10.3103/S1060992X2002006X

22. Elpeltagy, M. and Sallam, H., Automatic prediction of COVID-19 from chest images using modified ResNet50, *Multimed Tools Appl.*, 2021, vol. 80, pp. 26451−26463.
https://doi.org/10.1007/s11042-021-10783-6