Ref	Techniques	Modalities	Task	Accuracy
[19]	CNN (VGG16)	1248 chest X-ray	3-classes classification: (Healthy/COVID-19pneumonia/non-COVID-19 pneumonia)	83.6% accuracy, 90.9% sensitivity
[20]	SqueezeNet	5184 Chest X-rays	Binary classification of images as COVID-19 or not	98% sensitivity, 92.9 specificity
[21]	CNN model	13,824 X-ray images	Binary classification of images as COVID-19 or normal	96.71% accuracy, 97% F1-score
[22]	CNN called COVID-Ne	13,975X-ray	Prediction of cases as (normal/pneumonia/COVID-19)	93.3% accuracy
[23]	CNN called nCOVnet	284 X-ray images	(Normal/COVID-19)	88.1% accuracy
[24]	EfficientNet	16,634 X-ray images	Classify cases as: (Normal/COVID-19/other)	93.48% accuracy
[25]	VGG-16	700 X-ray images	binary classification (Normal/COVID-19)	94.3% accuracy, 94.3% F1-score
Proposed model	VGG19	1600 X-ray images	Binary classification (Normal/COVID-19)	95% accuracy, 0.9505% F1-score, 0.96% sensitivity, 0.94% specificity

Table 6. Comparison of proposed method with the relevant researches.

Dataset	Model	F1 score	Test acc
Histeq	EfficientNetB0	0.93495	0.935
CLAHE	VGG19	0.9339	0.93375
Complement	VGG19	0.9337865	0.93375

Table 7. The results of best achieved models for 4-classes classification.

datasets used different number of samples. Some research performed binary class classification, where others performed multi-class classification. Thus, the proposed work has been compared with others that used the same modality which is X-ray with mentioning the number of used samples and the task.

By comparing the results of the proposed framework with recent literature, it was found that the proposed framework outperforms most of the state-of-the-art works. However,[21] slightly outperforms the proposed framework. Where the accuracy and F1-score of[21] are 96.71% and 97% respectively, the corresponding values of the proposed framework are 95% and 0.9505 respectively. Taking in consideration that in[21] un-balanced data has been used; the number of COVID images used is 3626 where the number of normal images is 10,198 as mentioned in the manuscript.

For more validation, different classes of the datasets have been used for training the CNN models. To check the capability of the proposed framework for 4-classes classification. The dataset versions that achieved the highest binary classification have been utilized for multi-classes classification. Since the performance of all models that trained using enhanced full image versions is close to each other, therefore, these versions have been utilized for 4-classes classification. A set of 3200 X-ray images (800 of each class) have been used to train CNN models. The newly added classes are Viral Pneumonia and Lung Opacity in addition to COVID-19 and normal classes.

It was found that the best-achieved accuracy of 4-classes classification using the full image versions reached 0.935 for histeq version by EfficientNetB0. While it reached 0.93375 for both CLAHE, and complement versions using VGG19. Table 7 shows the results of the best-achieved models for 4-classes classification.

Conclusion and future work

In this research, a framework has been developed for automatically classifying chest X-ray images as COVID-19 positive cases or normal cases. Different techniques such as histeq, CLAHE, and complement have been applied to enhance the original X-ray images and therefore, both the original and enhanced versions have been introduced to the selected CNN pre-trained models. Two pre-trained CNN models which are VGG19 and EfficientNetB0 have been used to train different versions with the last dense layer set to (2/4) according to the number of classification classes.

Two approaches have been utilized to train pre-trained CNN models which are using whole chest X-ray images and using lung segmented images with their enhanced versions. The best binary classification accuracy reached 95% for the model trained using CLAHE full images version utilizing VGG19. The best achieved accuracy for a model trained using a segmented dataset is 91% for the model trained using Histeq version utilizing VGG19. By testing the framework for 4-classes classification, it achieved promising results which reached 0.935 accuracy.

It is obvious from the results that, the proposed framework can be employed in the future to support physicians and decrease the effect of doctors' shortages in the struggle against the disease. However, extra validations are required before applying any system, as more accuracy and more careful experiments are needed when things are related to human life. In the future, the authors are willing to try the proposed model on local data.

Data availability

The used data has been obtained from an available online database and it has been referenced in the manuscript. The link to the database used in the study: https://www.kaggle.com/tawsifurrahman/covid19-radiography-database