COVID-19 Detection Using Integration of Deep Learning Classifiers and Contrast-Enhanced Canny Edge Detected X-Ray Images

Stefanus Tao Hwa Kieu, Abdullah Bade, Mohd Hanafi Ahmad Hijazi, and Hoshang Kolivand

***Abstract*—COVID-19 is a deadly disease, and should be efficiently detected. COVID-19 shares similar symptoms with pneumonia, another type of lung disease, which remains a cause of morbidity and mortality. This study aims to demonstrate an ensemble deep learning approach that can differentiate COVID-19 and pneumonia based on chest x-ray images. The original x-ray images were processed to produce two sets of images with different features. The first set was images enhanced with contrast limited adaptive histogram equalization. The second set was edge images produced by Contrast-Enhanced Canny Edge Detection. Convolutional neural networks were used to extract features from the images and train classifiers which were able to classify COVID-19, pneumonia and healthy lungs cases. Results show that the classifiers were able to differentiate x-rays of different classes, where the best performing ensemble achieved an overall accuracy of 97.90%, with a sensitivity of 99.47% and specificity of 98.94% for COVID-19 detection.**

***Index Terms*— Canny edge detector, COVID-19 detection, deep learning, medical image analysis, pneumonia detection**

I. Introduction

CORONAVIRUS Disease (COVID-19) is an infectious disease caused by a recently discovered coronavirus [1]. Senior citizens and those with medical problems like chronic respiratory disease, cardiovascular disease, and diabetes are more susceptible to develop severe sickness [2].

Pneumonia is a lung infection that causes inflammation in the air sacs within the lungs. Symptoms include severe shortness of breath, cough, fever, or fatigue. Community-acquired pneumonia is still a recurrent cause of morbidity and mortality [3].

Machine learning has long been used in image classification [4]. To teach an algorithm to recognize images, a Convolutional Neural Network (CNN) is used. CNN is especially useful for image classification and recognition as it automatically learn features without needing manual feature extraction [5].

Medical image analysis is an active field of research for machine learning. Machine learning has been used in many disease detection, such as age-related macular degeneration using retinal images [6], tuberculosis using x-ray images [5] and cervical cancer detection on pap smear images [7].

Several works that used CNN for pneumonia detection in x-rays can be found in [8], [9]. Overall, CNN has achieved remarkable accuracy when detecting pneumonia in chest x-ray images.

Using image classification techniques on CT scan image to screen COVID-19 patients have been clinically proven to be feasible [10], [11]. This could expedite the screening process that could take days to be completed using the existing Reverse Transcription Polymerase Chain Reaction (RT-PCR) screening method. However, CT scan devices are expensive. Hence, more work has focused on using chest x-ray images to detect COVID-19

Submitted for review on 18th June.

The work presented is funded by Universiti Malaysia Sabah (UMS) through grant SDK0191-2020.

Stefanus Tao Hwa Kieu is with the Faculty of Science and Natural Resources, UMS, Malaysia (e-mail: stefanuskieu@gmail.com).

Abdullah Bade is with the Faculty of Science and Natural Resources, UMS, Malaysia (e-mail: abb@ums.edu.my).

Mohd Hanafi Ahmad Hijazi is with the Faculty of Computing and Informatics, UMS, Malaysia (e-mail: hanafi@ums.edu.my).

[12][13]. Deep learning was shown to be able to perform multiclass classification between healthy lungs, pneumonia infected lungs and COVID-19 infected lungs [14]–[16]. The symptoms of pneumonia are similar to those of COVID-19; therefore, it is important to quickly and accurately screen and differentiate patients with COVID-19 from pneumonia.

This paper presents an integration approach to classify x-ray images of COVID-19 and pneumonia patients using ensemble deep learning and image processing. Since deep leaning performed well in classifying images, two types of images were investigated. One was contrast-enhanced images; another was edge images. Edge images were considered because edge detected images provide essential features for image analysis [17]. Edges represents shape feature of an image, and shape is an important and powerful feature used for image classification [18]. Also, ensemble was performed to investigate if predictions made by models trained on different images are more accurate than the prediction made when the models were trained on just one type of image. With deep learning, classifiers can be trained to differentiate between the three classes based on the images it is presented with. The contributions of this paper are thus:

1. Using deep learning on medical images to detect COVID-19 and pneumonia.
2. Generate ensemble CNN classifiers for COVID-19 and pneumonia detection using two sets of images, namely the contrast-enhanced images and edge images of chest x-rays.

II. Methodology

The proposed approach has four stages: (i) data acquisition, (ii) image processing, (iii) classifiers generation, and (iv) ensemble classification. Fig. 1 shows the process of the proposed approach in this study. Each of the phases is described in details in subsections A to D.

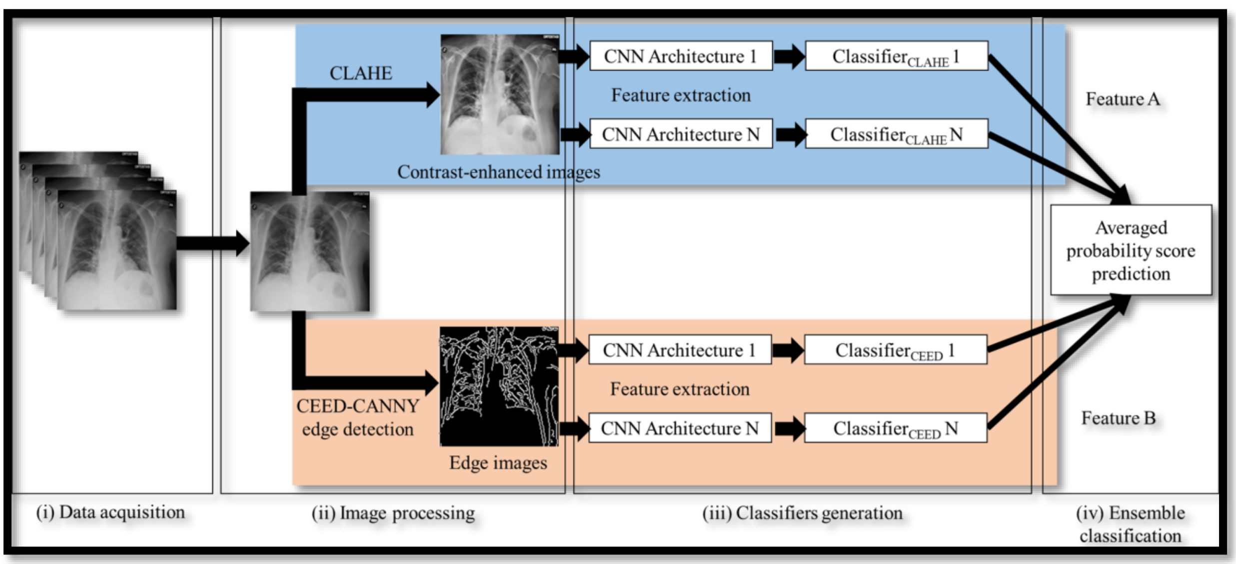


Fig. 1. Process flow chart of current proposed approach

A. *Data Acquisition*

A total of three datasets were used in this paper. Two of them were used to obtain COVID -19 positive chest x-rays, while the third was used to obtain pneumonia positive chest x-rays and normal (healthy) chest x-rays. The first dataset was acquired from Joseph Cohen's GitHub repository [19]. This repository contains CT scans and chest x-rays of patients with various diseases, including COVID-19, acute respiratory distress syndrome and severe acute respiratory syndrome. The second dataset consists of CT scans and chest x-rays extracted from articles publicly available in RSNA, Radiopedia and SIRM (https://www.kaggle.com

/andrewmvd/convid19-X-rays). The third dataset contains x-rays images of healthy patients (normal lungs) and patients infected with pneumonia [9]. To avoid class imbalance, the number of images for COVID19, pneumonia and normal lungs were the same, at 196 images. Hence, 588 images were used in this paper.

B. Image Processing

For image processing, three processes were conducted. The first is image resizing, followed by contrast limited adaptive histogram equalization (CLAHE), and Contrast-Enhanced Canny (CEED-Canny) edge detection. First, all x-rays images were resized to 150 x 150 pixels. Afterwards, CLAHE and CEED-Canny edge detection were performed to produce two very distinct types of images; the contrast-enhanced x-rays and the edge images of the x-rays.

In this research, CLAHE was used to enrich features by improving the visibility of abnormalities. CEED-Canny method combines local morphological contrast enhancement and the Canny edge detection technique. It was first applied in [20], and the algorithm is shown as follows:

i) Get original pixel value, local maximum and local minimum

ii) Perform morphological contrast enhancement.

iii) Perform noise reduction using Gaussian smoothing.

iv) Find intensity gradient of the image.

v) Perform non-maximum suppression.

vi) Perform hysteresis thresholding.

Edges consist of meaningful, significant information and features. By applying edge detector to an image, the amount of data needed to be processed can be reduced, and the information regarded as less relevant can be filtered [17]. In this research, we conjectured that lungs with abnormalities have more uncommon edges compared to healthy lungs. These features can help to differentiate lung conditions.

C. Classifiers Generation

There are many CNN architectures available. In this paper, two CNN architectures were employed, namely VGG16 and InceptionV3. The pre-trained CNNs were used as fixed features extractor. The extracted features became the input in for a new model. VGG16 and InceptionV3 were only used as feature extractors, therefore, no training or freezing was done on their layers. Dropout layers were added to prevent overfitting [16]. Classifiers were trained for CLAHE and CEED-Canny images. Using these two CNN architectures, four classifiers were generated.

D. Ensemble Classification

In machine learning, the ensemble is the application of more than one model to make prediction [5]. The ensemble usually reduces the variance of predictions, and this can produce predictions that are more accurate than any single model. In this paper, the ensemble was performed by averaging the probability score of the individual models. Let P be the probability score, let C be a classifier and let n be the number of classifier, then the final label is calculated as shown in (1). When predicting the label of a new image, a model produces a probability score. By averaging the probability scores of multiple models, the final label of the image is determined.

$$final\ label = \frac{\sum_1^n P(C_1) + P(C_2) + \cdots + P(C_n)}{n} \quad (1)$$

E. Advantages of Using Ensemble from Various Features

In this study, edge detection was used to provide a new dataset represented by a new feature. These two datasets, namely the enhanced x-ray dataset and the edge x-ray dataset, were used, because these two datasets represents two distinct features. The enhanced x-ray dataset represent image feature, while the edge dataset represents shape feature [18]. Edge x-ray images are considerably different from normal x-ray images, because after edge detection was applied on an x-ray image, information that may be regarded as less relevant were filtered out and the structural properties of the image is preserved [17]. If two classifiers can classify x-rays based on different features, they will complement each other. Consider two classifiers, one was trained on enhanced x-ray images, the other was trained on edge x-ray images. The first classifier can differentiate lung conditions based on the enhanced x-ray images, while the second classifier can differentiate lung conditions based on the edges of the x-ray images. When classifying a new image, the first classifier will look at the enhanced x-ray to give a prediction. Likewise, the second classifier will look at the edges of the x-ray to give a prediction. This is how different classifiers can classify an x-ray image by studying it from different features.

The fusion of so many traditional techniques, which is the usage of edge detection and CNN using ensemble, has various purposes. During ensemble, using classifiers trained from significantly different features serves to reduce correlation of error between base classifiers. This in turns increases generalizability. This also helps reduce overfitting and improve accuracy.

III. Results And Discussion

In this section, detailed explanations of the followings are presented: (i) the metrics used to evaluate the performance of the proposed approach, (ii) the performance of individual models, (iii) the performance of ensembles, and (iv) comparison with other works.

A. Metrics

The metrics used to evaluate the proposed method were accuracy, sensitivity and specificity. Sensitivity is used to measure the ability of the model to identify positive cases, while specificity measures how well the model to identify negative cases. The overall performance of the model is indicated by the accuracy.

In this paper, for the case of COVID-19, positive cases are images with symptoms of COVID-19 present in the lungs, while negative cases are images without symptoms of COVID-19. The same goes for pneumonia and normal cases.

B. Performance of Individual Models

The CNN architectures, InceptionV3 and VGG16, were pre-trained using the ImageNet dataset. The batch size was set to 50, the learning rate was set to 0.0001, and the number of the epoch was set to 100. These settings were selected because they produced optimal results after multiple initial experiments, similar to the work presented in [20]. The train-test split was in a ratio of 9:1. Tenfold cross validation was also performed to prevent overfitting, similar to the work in [6] and [16] . Table I shows the performance of the individual models.

TABLE I
Performance of Individual Models (A: Model trained on VGG16 extracted features from CLAHE Images, B: Model trained on InceptionV3 extracted features from CLAHE Images, C: Model trained on VGG16 extracted features from ceed-canny images, D: Model trained on InceptionV3 extracted features from ceed-canny images)

Model	Accuracy	COVID-19		Normal		Pneumonia	
		Sensitivity	Specificity	Sensitivity	Specificity	Sensitivity	Specificity
A	0.9772	0.9947	0.9947	0.9842	0.9815	0.9527	0.9895
B	0.9386	0.9737	0.9751	0.9421	0.9672	0.9000	0.9631
C	0.9369	0.9684	0.9746	0.9474	0.9625	0.8948	0.9657
D	0.9035	0.9526	0.9579	0.9158	0.9379	0.8421	0.9543

All of the models managed to achieve accuracy higher than 90%. There was little difference between the training loss and testing loss, suggesting that overfitting did not occur [14]. The models were able to differentiate the x-rays of different lung conditions. This shows that the classifiers were able to differentiate lung conditions based on lung images alone. Results also show that the CNNs performed better when trained on CLAHE images than CEED-Canny images.

The results in Table I also shows that model A achieved a high sensitivity of 0.9947 for COVID-19 cases. This is especially important because we want to avoid false negatives. A false negative for COVID-19 will mean that a patient infected with COVID-19 is classified as not infected and would go on possibly infecting others.

C. Performance of Ensembles

An ensemble must consists of more than one model. From four initial models, eleven combinations of ensembles were formed. Table II shows the performance of the ensembles.

From Table II, most of the ensemble combination achieved better accuracy than the individual models. This showed that the combination of classifiers could increase the detection performance. Results also suggest that in some cases, the ensembles do improve performance, although not much. For example, ensemble ABC has the highest accuracy among all the ensemble, at 0.9790, but, model A has an accuracy of 0.9772. The ensemble only increased the accuracy by 0.0018. Nevertheless, these results supported our hypothesis that using classifiers trained from significantly different features during ensemble is able to improve accuracy. Model ABC also managed to attain the highest sensitivity for COVID-19 and pneumonia and also the highest specificity for normal and pneumonia which further supports our hypothesis.

TABLE II
PERFORMANCE OF ENSEMBLES (A: MODEL TRAINED ON VGG16 EXTRACTED FEATURES FROM CLAHE IMAGES, B: MODEL TRAINED ON INCEPTIONV3 EXTRACTED FEATURES FROM CLAHE IMAGES, C: MODEL TRAINED ON VGG16 EXTRACTED FEATURES FROM CEED-CANNY IMAGES, D: MODEL TRAINED ON INCEPTIONV3 EXTRACTED FEATURES FROM CEED-CANNY IMAGES)

Model	Accuracy	COVID-19		Normal		Pneumonia	
		Sensitivity	Specificity	Sensitivity	Specificity	Sensitivity	Specificity
AB	0.9579	0.9842	0.9863	0.9737	0.9678	0.9158	0.9816
AC	0.9772	0.9895	**0.9921**	**0.9895**	0.9815	0.9527	0.9921
AD	0.9474	0.9895	0.9723	0.9632	0.9594	0.8895	0.9868
BC	0.9527	0.9842	0.9784	0.9579	0.9733	0.9158	0.9763
BD	0.9316	0.9790	0.9352	0.9474	0.9773	0.8684	0.9813
CD	0.9281	0.9737	0.9668	0.9368	0.9494	0.8737	0.9733
ABC	**0.9790**	**0.9947**	0.9894	0.9842	**0.9842**	**0.9579**	**0.9947**
ABD	0.9702	0.9947	0.9839	0.9842	0.9787	0.9316	0.9921
ACD	0.9632	0.9895	0.9838	0.9790	0.9683	0.9211	0.9921
BCD	0.9562	0.9895	0.9730	0.9790	0.9707	0.9000	0.9895
ABCD	0.9737	0.9947	0.9866	0.9790	0.9815	0.9474	0.9920

D. *Comparison with Other Works*

To the best of our knowledge, no direct comparison can be made between the work presented in this paper and other work because different datasets were used in these studies. The presented results in this section aims to show the current state, with respect to the detection performance, of deploying ensemble deep learning in detecting COVID-19 on x-ray images. The presented comparison also considers three-class classification only.

When comparing our work with [14]–[16], we found that all works recorded accuracies of greater than 85%. Our method produced the highest sensitivity for COVID-19 detection. The result demonstrates that using ensemble classifiers, trained on different image features, coupled with transfer learning could produce better detection result. Variation of image features could enrich the features learned by the CNN, and subsequently, generate better classifiers.

There are some limitations to the presented work. First, the number of images used should be significantly increased to improve classification accuracy through better model generated from more images. Second, the image resolution could be increased because essential details might be lost in smaller images. Our future directions also include the employment of feature fusion and segmentation, as latest advancement shows the potential of using these methods in improving the identification of significant features on the images. Concerning the feature fusion, we are looking at how the features learned from the images can be fused with features extracted from non-image data.

IV. CONCLUSION

The work presented in this paper shows that it is highly possible to use our integrated or ensemble deep learning on medical images to detect COVID-19, in particular, and other lung-related disease like pneumonia. Ensemble CNN classifiers were generated for COVID-19 and pneumonia detection using two sets of images, namely the contrast-enhanced images and edge images of chest x-rays. The results demonstrate that features extracted from different types of images can be used by CNN to identify a disease. We do not claim the diagnostic performance of our model without a clinical study, but merely demonstrated the potential of using ensemble deep learning to differentiate COVID-19 and pneumonia based on x-ray images.

References

[1] C. Huang *et al.*, "Clinical features of patients infected with 2019 novel coronavirus in Wuhan, China," *Lancet*, vol. 395, no. 10223, pp. 497–506, 2020.

[2] T. P. Velavan and C. G. Meyer, "The COVID-19 epidemic," *Trop. Med. Int. Heal.*, vol. 25, no. 3, pp. 278–280, 2020.

[3] R. G. Wunderink and G. Waterer, "Advances in the causes and management of community acquired pneumonia in adults," *BMJ*, pp. 1–13, 2017.

[4] C. K. On, T. K. Yao, R. Alfred, A. A. A. Ibrahim, W. Cheng, and T. T. Guan, "A Comparison of BPNN, RBF, and ENN in Number Plate Recognition," 2016, pp. 37–47.

[5] S. T. H. Kieu, M. H. A. Hijazi, A. Bade, R. Yaakob, and S. Jeffree, "Ensemble deep learning for tuberculosis detection using chest X-Ray and canny edge detected images," *IAES Int. J. Artif. Intell.*, vol. 8, no. 4, pp. 429–435, 2019.

[6] M. H. A. Hijazi, F. Coenen, and Y. Zheng, "Data mining for AMD screening: A classification based approach," *Int. J. Simul. Syst. Sci. Technol.*, vol. 15, no. 2, 2014.

[7] K. Bora, M. Chowdhury, L. Mahanta, K. K. Malay, and K. D. Anup, "Automated classification of pap smear image to detect cervical dysplasia," *Comput. Methods Programs Biomed.*, vol. 138, pp. 31–47, 2017.

[8] E. Ayan and H. M. Ünver, "Diagnosis of Pneumonia from Chest X-Ray Images using Deep Learning," in *2019 Scientific Meeting on Electrical-Electronics & Biomedical Engineering and Computer Science*, 2019, pp. 1–5.

[9] D. S. Kermany, M. Goldbaum, W. Cai, and M. A. Lewis, "Identifying Medical Diagnoses and Treatable Diseases by Image-Based Deep Learning," *Cell*, vol. 172, no. 5, pp. 1122-1131.e9, 2018.

[10] T. Ai *et al.*, "Correlation of Chest CT and RT-PCR Testing in Coronavirus Disease 2019 (COVID-19) in China: A Report of 1014 Cases," *Radiology*, p. 200642, Feb. 2020.

[11] O. Gozes, M. Frid, H. Greenspan, and D. Patrick, "Rapid AI Development Cycle for the Coronavirus (COVID-19) Pandemic: Initial Results for Automated Detection & Patient Monitoring using Deep Learning CT Image Analysis Article," *arXiv*, 2020.

[12] A. Narin, C. Kaya, and Z. Pamuk, "Automatic Detection of Coronavirus Disease (COVID-19) Using X-ray Images and Deep Convolutional Neural Networks," *arXiv e-print*, 2020.

[13] F. M. Salman, S. S. Abu-naser, E. Alajrami, B. S. Abu-nasser, and B. A. M. Ashqar, "COVID-19 Detection using Artificial Intelligence," *Int. J. Acad. Eng. Res.*, vol. 4, no. 3, pp. 18–25, 2020.

[14] X. Xu *et al.*, "A Deep Learning System to Screen Novel Coronavirus Disease 2019 Pneumonia," *Engineering*, pp. 0–7, 2020.

[15] L. Wang, Z. Q. Lin, and A. Wong, "COVID-Net: A Tailored Deep Convolutional Neural Network Design for Detection of COVID-19 Cases from Chest X-Ray Images," *arXiv*, 2020.

[16] 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.*, pp. 1–6, 2020.

[17] I. Kumar, J. Rawat, and H. S. Bhadauria, "A Conventional Study of Edge Detection Technique in Digital Image Processing," *Int. J. Comput. Sci. Mob. Comput.*, vol. 3, no. 4, pp. 328–334, 2014.

[18] B. Ramamurthy and K. R. Chandran, "Content based Image Retrieval for Medical Images using Canny Edge Detection Algorithm," *Int. J. Comput. Appl.*, vol. 17, no. 6, pp. 32–37, 2011.

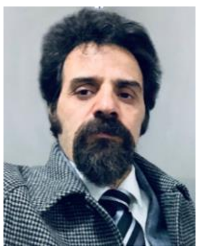
[19] J. P. Cohen, P. Morrison, and L. Dao, "COVID-19 Image Data Collection," *arXiv*, 2020.

[20] S. T. H. Kieu, M. H. A. Hijazi, A. Bade, and M. Saffree Jeffree, "Tuberculosis detection using deep learning and contrast-enhanced canny edge detected x-ray images," *IAES Int. J. Artif. Intell.*, vol. 9, no. 4, 2020.

Stefanus Tao Hwa Kieu received the B. S. degree in mathematics and computer graphics from the Faculty of Science and Natural Resources, Universiti Malaysia Sabah (UMS), Malaysia, in 2018. He is currently pursuing the M. S. degree in mathematics and computer graphics in UMS. His research interests include image processing and machine learning.

Abdullah Bade was born in Sabah, Malaysia in 1977. He received the BSc. in Computer Science in 2000, MSc. in Computer Science (Computer Graphics) in 2003 both from Universiti Teknologi Malaysia, and PhD in Industrial Computing from National University of Malaysia in 2008. From 2000 to 2010, he was a Senior Lecturer at the Faculty of Computing, Universiti Teknologi Malaysia. Currently, he is an Associate Prof. in Computer Graphics and Image Processing at the Faculty of Science and Natural Resources, UMS. Dr. Abdullah is a professional member of ACM SIGGRAPH, Malaysian Society of Mathematics and Science (PERSAMA), IEEE and serve to the government as Malaysia Qualification Agency panel of assessor in the field of Computing and Multimedia since 2005. His research interest includes Computer Graphics and Image Processing algorithms and all aspects of image generation. He had published numerous articles in refereed journals, conference proceedings, books and technical papers.

Mohd Hanafi Ahmad Hijazi (M'14) received the B.Sc. and M.Sc. degrees in computer science at the Universiti Teknologi Malaysia, in 2001 and 2005 and the Ph.D. degree in computer science from the University of Liverpool, United Kingdom, in 2012.
From 2012 to 2018 he was a Senior Lecturer with the Faculty of Computing and Informatics, UMS. Since 2018, he has been an Associate Professor at the same Faculty. His research interest includes data mining and artificial intelligence with the application on healthcare and biometrics.



Hoshang Kolivand is Senior Lecturer in Liverpool John Moores University, UK. He received his MS degree in applied mathematics and computer from Amirkabir University of Technology, Iran, in 1999, and his PhD from Universiti Teknologi Malaysia (UTM) in 2013. He has published numerous articles in international journals, conference proceedings and technical papers, including chapters in books.