Sensors and Materials, Vol. 36, No. 2 (2024) 683–699
MYU Tokyo

S & M 3555

Evaluating Feature Fusion Techniques with Deep Learning Models for Coronavirus Disease 2019 Chest X-ray Sensor Image Identification

Chih-Ta Yen,[1*] Jia-Xian Liao,[2] and Yi-Kai Huang[2]

[1]Department of Electrical Engineering, National Taiwan Ocean University, Keelung City 202301, Taiwan
[2]Department of Electrical Engineering, National Formosa University, Yunlin County 632, Taiwan

(Received July 2, 2023; accepted January 29, 2024)

Keywords: COVID-19, convolutional neural network, deep learning, chest X-ray (CXR), contrast-limited adaptive histogram equalization (CLAHE), feature fusion

Current diagnostic methods for coronavirus disease 2019 (COVID-19) mainly rely on reverse transcription polymerase chain reaction (RT-PCR). However, RT-PCR is costly and time-consuming. Therefore, an accurate, rapid, and inexpensive screening method must be developed for the diagnosis of COVID-19. In this study, we combined image processing technologies with deep learning algorithms to enhance the accuracy of COVID-19 identification from chest X-ray (CXR) sensor images. Contrast-limited adaptive histogram equalization (CLAHE) was used to improve the visibility level of unclear images. In addition, we examined whether our image fusion technique can effectively improve the performance of seven deep learning models (MobileNetV2, ResNet50, ResNet152V2, Inception-ResNet-v2, DenseNet121, DenseNet201, and Xception). The proposed feature fusion technique involves merging the features of an original image with those of an image subjected to CLAHE and then using the merged features to retrain, test, and validate deep learning models for identifying COVID-19 in CXR images. To avoid incidences of images not matching reality and to ensure high model stability, no data enhancement was conducted. The results of this study indicate that the proposed image fusion technique can improve the classification evaluation indicators, especially the sensitivity of deep learning models in two-class and three-class sortings. Sensitivity refers to a model's ability to detect an infection correctly. The highest accuracy in this study was achieved when combining Xception with the proposed feature fusion technique. In three-class sorting, the accuracy of this method was 99.74%, with the average accuracy of fivefold cross-validation being 99.19%. In two-class sorting, the accuracy of the aforementioned method was 99.74%, with the average accuracy of fivefold cross-validation being 99.50%. The results showed that the proposed image processing technologies with deep learning algorithms have exceptional generalization.

*Corresponding author: e-mail: chihtayen@gmail.com
https://doi.org/10.18494/SAM4685

ISSN 0914-4935 © MYU K.K.
https://myukk.org/

1. Introduction

Coronavirus disease 2019 (COVID-19) began spreading rapidly worldwide in December 2019.[1,2] As of February 1, 2021, over 100 million people have been infected by COVID-19, and numerous people have died because of it. Classic clinical symptoms of COVID-19 include fever, respiratory tract infections, pneumonia, multiple organ failure, fatigue, muscle soreness, and reductions in the numbers of white blood cells, platelets, and lymphocytes.[3] Currently, reverse transcription polymerase chain reaction (RT-PCR) is used to screen for COVID-19; however, this technique has some limitations. In certain situations, the ratio of false negatives detected with RT-PCR is high, which indicates that even if the test result is negative, a COVID-19 infection cannot be completely ruled out.[4] Moreover, RT-PCR is expensive, time-consuming, and complicated. Therefore, an accurate, rapid, and cheap screening method must be developed for diagnosing COVID-19.

Some countries are using medical imaging technologies, such as computed tomography (CT) and chest X-rays (CXRs), to screen for COVID-19.[5–8] CT scans are superior to CXRs in revealing early pathological changes in the lung. If diagnosis is performed by an experienced radiologist, CT-based diagnosis has a higher sensitivity than CXR-based diagnosis. However, CT scans have many cross sections; therefore, CT-based diagnosis is a lengthy and expensive process, and technicians must make numerous adjustments in this process. CT technicians might come into contact with the patient, and improper disinfection might lead to severe cross contaminations. Furthermore, in some areas, a lack of radiologists is a challenge. To overcome the aforementioned problems, in this study, we used CXRs for COVID-19 diagnosis because CXR imaging is faster and cheaper than RT-PCR.

In recent years, machine learning has been widely used in computer vision tasks,[9,10] which has driven the development of AI in medical diagnosis.[11–17] Deep learning algorithms have been widely examined because in contrast to conventional machine learning methods, which require manual feature extraction, deep learning algorithms can achieve automatic feature extraction. Deep learning techniques can identify fine image features that are not clearly visible in a raw image. Compared with conventional image diagnostic processes, which rely heavily on manual labor, AI-based image diagnostic processes have higher efficiency, safety, and accuracy; thus, AI-based image diagnostic processes enable rapid diagnoses and minimize the medical labor involved.

With the emergence of the deep Web and the broad application of transfer learning, deep learning methods that exhibit favorable results with a small quantity of training data have been developed. Apostolopoulos and Tzani[18] used transfer learning to divide data into two sets: one set contained 504 images of healthy lungs, 700 images of confirmed bacterial pneumonia, and 224 images of confirmed COVID-19, and the other set contained 504 images of healthy lungs, 714 images of confirmed bacterial and viral pneumonia, and 224 images of confirmed COVID-19. These images were scaled to 200 × 266 pixels, and the classification layers of the developed pretrained model were then fine-tuned. For the first data set, the highest two- and three-class accuracies of 98.75 and 93.48%, respectively, were achieved with VGG-19. For the second data set, the highest two- and three-class accuracies of 96.78 and 94.72%, respectively,

were achieved with MobileNetV2. Ferhat and Deniz[19] developed the COVIDiagnosis-Net model and applied it to 4209 images of pneumonia, 76 images of confirmed COVID-19, and 1583 images of healthy lungs. To solve the problem of insufficient data, they expanded the data set by increasing noise, rotating the viewing angle, increasing the brightness, and performing mirroring to enhance the original images. The aforementioned authors balanced the number of images of the aforementioned three conditions by deleting certain images. The accuracy rate of COVIDiagnosis-Net for the aforementioned data was 98.26%. Minaee *et al.*[20] used transfer learning to fine-tune the classification layers of four pretrained models—ResNet18, ResNet50, SqueezeNet, and DenseNet121—and to detect COVID-19 by using CXR images. Among the 520 images of confirmed COVID-19 and 5000 images not showing COVID-19 infection in the experiment, 336 images of COVID-19 from the training data were enhanced through flipping, rotation, and distortion. The results indicated that SqueezeNet had a two-class accuracy of 92.29% for the aforementioned data. Mesut *et al.*[21] used fuzzy color and stacking techniques to eliminate noise from 295 CXR images of COVID-19 conditions, 65 CXR images of healthy lungs, and 98 CXR images of pneumonia for enhancing the image quality. They used the MobileNetV2 and SqueezeNet deep learning models for feature extraction, then performed social mimic optimization, and finally extracted effective features from 1000 features with an accuracy of 99.27%. By referencing the Darknet-19 model,[22] Ozturk *et al.*[23] proposed the DarkCovidNet model for detecting COVID-19. The architecture of the DarkCovidNet model contains fewer network layers and convolution kernels than that of the Darknet-19 model. The aforementioned authors selected the leaky rectified linear unit as the activation function of the DarkCovidNet model to prevent neuron death. This model was trained using 500 images of healthy lungs, 500 images of pneumonia, and 125 images of confirmed COVID-19. The results indicated that the DarkCovidNet model had a three-class accuracy of 87.02% and a two-class accuracy of 98.08% for the aforementioned data. Chowdhury *et al.*[24] used a pretrained deep learning model based on transfer learning to detect COVID-19 from a set of CXR images, which comprised 423 images of COVID-19, 1579 images of healthy lungs, and 1485 images of viral pneumonia. They performed image rotation and translation for data enhancement. The results indicated that without data enhancement, the two- and three-class accuracies of the aforementioned model were 99.41 and 97.74%, respectively. Moreover, with data enhancement, the two- and three-class accuracies of the aforementioned model were 99.7 and 97.94%, respectively.[24]

Ohata *et al.*[25] used pretrained transfer learning models to extract features from CXR images. They then used the extracted features in a machine learning classifier. The performance of this classifier for two data sets was compared. Each data set comprised 194 CXR images of lungs with and without COVID-19, with the CXR images of healthy lungs in the two data sets being dissimilar. The results of the aforementioned study indicated that a classification accuracy of 98.62% was achieved when combining MobileNet with support vector machine (SVM) and that a classification accuracy of 95.64% was achieved when combining DenseNet201 with multilayer perceptron. After processing images through contrast-limited adaptive histogram equalization (CLAHE), Umri *et al.*[26] used VGG16 to classify 100 CXR images each of the lungs with and without COVID-19 and achieved a classification accuracy of 97.50%. Anunay *et*

al.[27] developed InstaCovNet-19, which is a stacked deep learning network that comprises pretrained models of ResNet101, Xception, InceptionV3, MobileNet, and NASNet. They used InstaCovNet-19 to apply transfer learning in the classification of 361 CXR images of COVID-19, 365 CXR images of healthy lungs, and 362 CXR images of pneumonia. These images were enhanced using blur color and stacking techniques. InstaCovNet-19 exhibited a three-class accuracy of 99.08% and a two-class accuracy of 99.53 for the aforementioned data. By adding a dropout layer and two fully connected layers to the Xception model pretrained through transfer learning, Khan *et al*.[28] developed the CoroNet model and classified 310 images of healthy lungs, 330 images of bacterial pneumonia, 327 images of viral pneumonia, and 284 images of COVID-19. The four-, three-, and two-class accuracies of the aforementioned model were 89.6, 95.02, and 99.00%, respectively. Wang and Wong[29] designed COVID-Net, which has a residual network architecture, to capture the features of CXR images and classify 5538 images of pneumonia, 358 images of COVID-19, and 8066 images of healthy lungs. Because the data volume was insufficiently large, the data were expanded through translation, rotation, horizontal flipping, scaling, and offsetting. The accuracy of COVID-Net for the aforementioned data was 93.3%. Bhadra and Kar[30] developed a multilayer convolutional neural network model for identifying CXR images and applied it to a combination of four databases containing 1130 CXR images of COVID-19, 1330 CXR images of pneumonia, and 1330 CXR images of healthy lungs. This model achieved an accuracy of 99.1% for the aforementioned data. Karim *et al*.[31] proposed an integrated method that involves combining VGG-19 and DenseNet161 to identify CXR images. The combination of VGG-19 and DenseNet161 outperformed combinations of other models and exhibited a prediction rate of 94.6% for a database comprising 8066 images of healthy lungs, 5538 images of pneumonia, and 358 images of COVID-19. Asmaa *et al*.[32] proposed the DeTraC deep convolution neural network. They enhanced the data of 80 CXR images of healthy lungs, 105 CXR images of COVID-19, and 11 CXR images of severe acute respiratory syndrome through flipping, translation, rotation, and modification. Moreover, the aforementioned authors used principal component analysis to project high-dimensional images into a low-dimensional space. This step was performed to reduce the dimensionality and thus increase the computational efficiency. They subsequently used transfer learning to adjust the weight training process of the model classification layer. The accuracy of the DeTraC model was 93.1% for the aforementioned database.

The studies discussed in the preceding section indicate that current network models for COVID-19 diagnosis mostly involve transfer learning. These networks have exhibited satisfactory results in the detection of COVID-19 in CXR images from small data sets. However, most of the aforementioned studies involved training with unbalanced data sets or data enhancement for increasing the number of images. Although data enhancement might yield superior experiment results, it can also lead to situations in which the enhanced images do not match reality. This may limit the generalizability of the model, and the model might need to be validated on a large data set. In this paper, we propose a feature integration technique for CXR images. The proposed technique allows deep learning models to obtain crucial features that increase the accuracy of deep learning networks. The combination of the proposed technique and a deep learning model can help medical staff by providing them a safe and highly effective

method for detecting COVID-19 during difficult times. Li *et al.* proposed MHA-CoroCapsule, a capsule network, and evaluated its performance in COVID-19 diagnosis.[33] In contrast to networks in other studies, the capsule network involved the use of a multihead attention-routing algorithm to quantify the relationship between two capsule layers. The traditional iterative dynamic routing process was replaced with a noniterative and parameterized multihead attention-routing algorithm, which enabled capsules in the model to jointly attend to information from different representational subspaces at different positions. More generalized features were extracted from CXRs, thus improving the detection of COVID-19.[33]

The remainder of this paper is structured as follows. In Sect. 2, we describe the data sets, preprocessing procedures, image processing techniques, feature integration process, and deep learning models used in this study. In Sect. 3, we present the experiment results and a comparison of these results with those in the literature. Finally, in Sect. 4, we present the conclusions of this study.

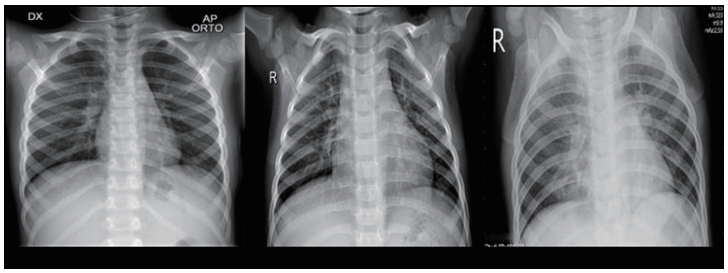
2. Image Fusion Techniques and Deep Learning Networks

2.1 CXR image data set

The main data set used in this study is publicly accessible and contains CXR images of healthy lungs, viral pneumonia, and COVID-19, as depicted in Fig. 1. Each image in this data set is in the PNG format. The COVID-19 images in the data set are scaled to 256 × 256 pixels, and the other images are scaled to 1024 × 1024 pixels. The adopted data set can be downloaded from the website of Kaggle. The present version of this data set contains 1200 images of COVID-19, 1341 images of healthy lungs, and 1345 images of viral pneumonia. The aforementioned data set is continually updated to assist scholars with COVID-19 research.[24]

2.2 Preprocessing of image data

To increase the rigorousness of the experiment and ensure that the image data contained no full duplicates, the obtained data were subjected to pixel-by-pixel image comparisons, and duplicate data were excluded. After exclusions, 3875 CXR images remained. These images



(a) (b) (c)

Fig. 1. CXR images in the adopted data set: (a) a chest with COVID-19, (b) a healthy chest, and (c) a chest with viral pneumonia.

comprised 1197 images of COVID-19, 1340 images of healthy lungs, and 1338 images of viral pneumonia. Deep learning models require a unified input format; therefore, after excluding the fully duplicate images, all the images were scaled to 224 × 224-pixel grayscale images. Subsequently, all the pixels were divided by 225 to yield an input value between 0 and 1 for model training and testing.

The adopted data set was divided into two data sets: data sets A and B. Data set A was used for three-class sorting and was formed by excluding duplicate images from the original data set. Data set B was used for two-class sorting. Each of these two data sets contained 1197 images of COVID-19 and 2678 images not showing COVID-19. The images not showing COVID-19 comprised images of healthy lungs and viral pneumonia. A total of 80, 10, and 10% of each data set were used as the training, testing, and validation sets, respectively. The numbers of images in data sets A and B are presented in Table 1.

2.3 Image processing techniques and feature fusion

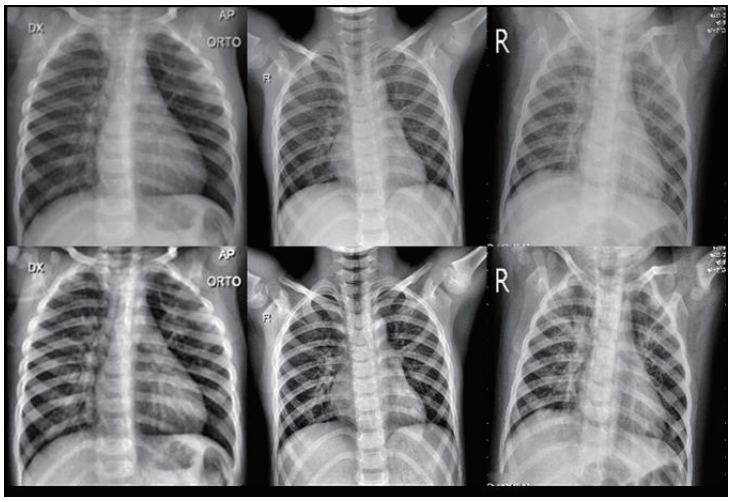
Although conventional histogram equalization (HE) can improve image contrast, this method involves processing the entire image. When the HE method is used on unevenly lit images, dark areas become darker and bright areas become brighter, which results in the severe loss of image data. This phenomenon is unacceptable in the processing of medical images. To retain image details while improving the image quality, CLAHE was performed in this study. CLAHE involves localized processing by dividing an image into many blocks and equalizing each block to even out the pixel distribution, enhance the image contrast, and control noise; this process results in an increase in the sharpness of details and contours.[34] In Fig. 2, the upper image is the original image, and the lower image is that obtained after conducting CLAHE. The processed image has sharper details and contours than the original image. The 3875 CXR images were subjected to CLAHE to obtain data that were more appropriate for training. The contrast threshold was set as 2, and the images were processed in 8 × 8-pixel blocks. After processing, the original images (224 × 224 × 1 pixels) were concatenated with the CLAHE-processed images (224 × 224 × 1 pixels) to enhance the image expressiveness. The image features of the data obtained after feature fusion (224 × 224 × 2 pixels) were increased to improve the model accuracy.

2.4 Deep learning models and parameters

Seven deep learning models were used in this study: MobileNetV2,[35] ResNet50,[36] ResNet152V2,[37] Inception-ResNet-v2,[38] DenseNet121,[39] DenseNet201,[38] and Xception.[40] These models were retrained, validated, and tested without using pretrained weights. The image input dimensions were 224 × 224 × 1 and 224 × 224 × 2 pixels. The size of both the original and CLAHE-processed images was 224 × 224 × 1 pixels, and the size of the feature fusion images was 224 × 224 × 2 pixels. Through the use of Adam as the optimizer of the deep learning model and the adoption of learning rate decay methods, the learning rate was decayed from 0.1 to

Table 1
Images in data sets A and B.

Data set	A			B	
Type	COVID-19	Normal conditions	Viral pneumonia	COVID-19	Without COVID-19
Quantity	1197	1340	1338	1197	Normal conditions: 1340 Viral pneumonia: 1338
Total	3875			3875	



(a) (b) (c)

Fig. 2. (Top) Original and (bottom) CLAHE-processed images: (a) a chest with COVID-19, (b) a healthy chest, and (c) a chest with viral pneumonia.

0.0000001. The rectified linear unit activation function was used to strengthen the nonlinear relationship between the layers of the neural network. The strengthening of this relationship resulted in originally negative values being output as 0, which prevented situations of overfitting to a certain extent and reduced occurrences of gradient disappearance. The number of epochs was set as 100 as an early stopping technique. When the model loss function and accuracy did not change significantly, the training was stopped and the training model was stored. The batch size was set as 16. The loss function for three-class sorting was sparse categorical cross entropy. For three-class sorting, softmax was the output function, whose value was set as 3. The probability of an image belonging to each image category was calculated, and the highest probability was determined as the output. Binary cross entropy was used as the loss function for two-class sorting. For two-class sorting, the loss function output was the sigmoid function, which was set as 1. The output function for two-class sorting was between 0 and 1. The two classes (presence or absence of COVID-19) were distinguished according to whether the output was lower or higher than 0.5. The closer the predicted and actual values were to each other, the smaller was the loss function.

2.5 Research process

After the CXR images were preprocessed, the original and CLAHE-processed images were subjected to feature fusion. Training, validation, and comparisons were performed for the deep learning models after the data required for deep learning had been obtained. Figure 3 shows the process of creating the three-class data set A. Unlike in the creation process for data set A, the non-COVID-19 images were not divided further when creating data set B.

3. Experimental Results and Discussion

The results obtained with the adopted seven models for data sets A and B were compared. Subsequently, a fivefold cross-validation was conducted for the model with the best results. Each time, 80 and 20% of all the data were used for training and testing, respectively. The training data were verified to not contain any testing data. The adopted image processing techniques and proposed feature fusion method were implemented on the CXR images, which were input into

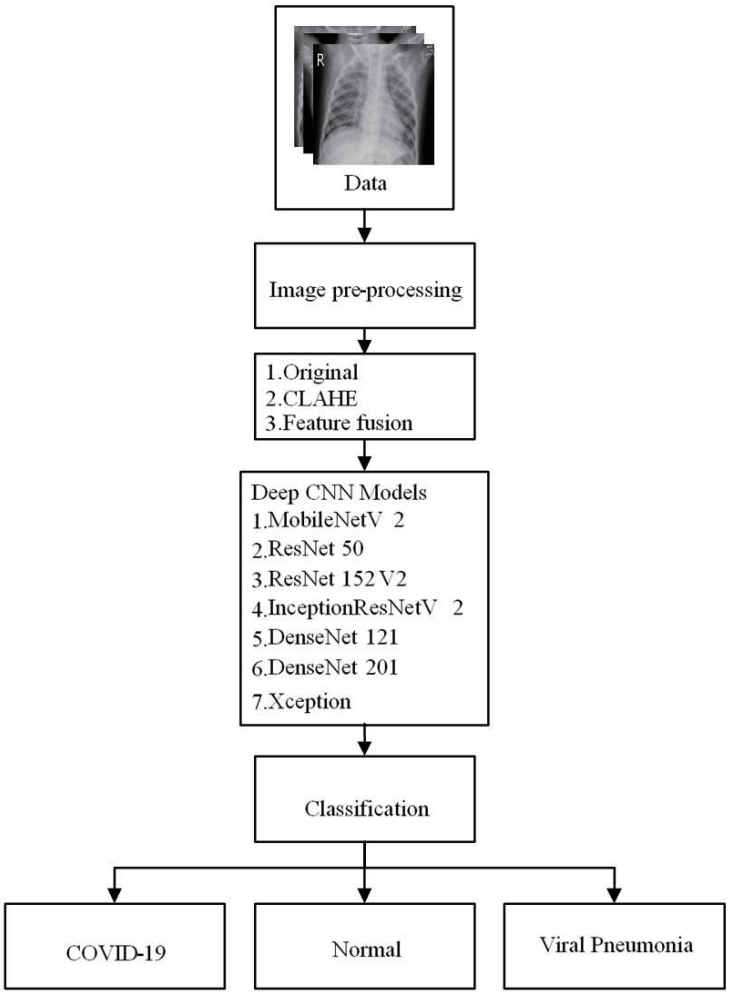


Fig. 3. Research process.

the deep-learning-based COVID-19 identification models. The COVID-19 identification performance of the models was evaluated using different classification indicators. The three-class sorting performance was evaluated using the following indicators: accuracy, macro average precision, macro average sensitivity, macro average $F1$ score, and macro average specificity. These indicators are defined as

$$Accuracy\ (\%) = \frac{TP + TN}{TP + FP + FN + TN} \times 100\%, \tag{1}$$

$$Precision_{macro}\ (\%) = \frac{1}{n} \sum_{i=1}^{n} \frac{TP_i}{TP_i + FP_i} \times 100\%, \tag{2}$$

$$Sensitivity_{macro}\ (\%) = \frac{1}{n} \sum_{i=1}^{n} \frac{TP_i}{TP_i + FN_i} \times 100\%, \tag{3}$$

$$F1 - Score_{macro}\ (\%) = 2\frac{Precision_{macro} \times Sensitivity_{macro}}{Precision_{macro} + Sensitivity_{macro}}, \tag{4}$$

$$Specificity_{macro}\ v\ (\%) = \frac{1}{n} \sum_{i=1}^{n} \frac{TN_i}{TN_i + FP_i} \times 100\%, \tag{5}$$

where i is the CXR category and has a range of 1–3.

The two-class sorting performance was examined using the following indicators: accuracy [Eq. (1)], precision, sensitivity, $F1$ score, and specificity. Precision, sensitivity, $F1$ score, and specificity are defined as

$$Precision\ (\%) = \frac{TP}{TP + FP} \times 100\%, \tag{6}$$

$$Sensitivity\ (\%) = \frac{TP}{TP + FN} \times 100\%, \tag{7}$$

$$F1 - Score\ (\%) = 2\frac{Precision \times Sensitivity}{Precision + Sensitivity}, \tag{8}$$

$$Specificity\ (\%) = \frac{TN}{TN + FP} \times 100\%, \tag{9}$$

where TP, TN, FP, and FN represent the numbers of true positives, true negatives, false positives, and false negatives, respectively.

3.1 Results for data set A

Tables 2–4 present the accuracy of each model. When the proposed image fusion technique was employed, the accuracy of MobileNetV2 increased from 92.50 and 92.76 to 98.19%; the accuracy of ResNet50 increased from 96.64 to 98.96%; the accuracy of ResNet152V2 increased from 97.41 and 96.38 to 98.44%; the accuracy of Inception-ResNet-v2 increased from 97.15 and 95.60 to 98.70%; the accuracy of DenseNet121 increased from 97.15 and 97.67 to 98.96%; the accuracy of DenseNet201 increased from 97.41 and 96.89 to 98.70%; and the accuracy of Xception increased from 97.41 to 99.74%. The aforementioned high accuracies indicate that the adopted models have low classification error rates for data set A. The three-class sorting results obtained for the images subjected to feature fusion were more favorable than those obtained for the original images or the images subjected to CLAHE processing only. The aforementioned result was obtained because the use of concatenated images resulted in the identification of a high number of image features and high image expressiveness, which allowed the extraction of sharp details and contour features by the adopted models. All the models exhibited an upward

Table 2
Model indicators of three-class deep learning for the original images.

Modes	Accuracy (%)	Precision (%)	Sensitivity (%)	*F*1 score (%)	Specificity (%)
MobileNetV2	92.50	92.62	92.66	92.57	96.23
ResNet50	96.64	96.69	96.70	96.68	98.30
ResNet152V2	97.41	97.53	97.44	97.48	98.68
InceptionResNetV2	97.15	97.24	97.20	97.21	98.55
DenseNet121	97.15	97.25	97.20	97.21	98.55
DenseNet201	97.41	97.56	97.44	97.46	98.68
Xception	97.41	97.57	97.48	97.50	98.68

Table 3
Model indicators of three-class deep learning for the CLAHE-processed images.

Modes	Accuracy (%)	Precision (%)	Sensitivity (%)	*F*1 score (%)	Specificity (%)
MobileNetV2	92.76	92.84	92.81	92.82	96.35
ResNert50	96.64	96.63	96.73	96.66	98.32
ResNet152V2	96.39	96.46	96.48	96.46	98.17
InceptionResNetV2	95.60	95.63	95.73	95.68	97.78
DenseNet121	97.67	97.63	97.72	97.71	98.82
Denset201	96.89	96.95	96.95	98.43	98.68
Xception	97.41	97.49	97.41	97.45	98.68

Table 4
Model indicators of three-class deep learning for the images subjected to feature fusion.

Modes	Accuracy (%)	Precision (%)	Sensitivity (%)	*F*1 score (%)	Specificity (%)
MobileNetV2	98.19	98.20	98.25	98.20	99.09
ResNert50	98.96	98.94	99.00	98.97	99.48
ResNet152V2	98.44	98.48	98.41	98.41	99.21
InceptionResNetV2	98.70	98.76	98.72	98.74	99.34
DenseNet121	98.96	99.09	99.00	99.00	99.47
Denset201	98.70	98.72	98.75	98.73	99.34
Xception	99.74	99.75	99.75	99.75	99.86

trend in the evaluation indicators following the application of the feature fusion method. The Xception model exhibited the best evaluation indicators. Its macro average precision increased by approximately 2% after feature fusion, which indicates that this model has a low likelihood of erroneously classifying images of healthy lungs or viral pneumonia as COVID-19. The Xception model's macro average sensitivity increased by 2% after feature fusion, which indicates that it can accurately identify COVID-19 images. The macro average F1 score of the Xception model increased by 2% after feature fusion, which indicates that this model has high precision and sensitivity in COVID-19 detection. Moreover, the macro average specificity of the Xception model increased by 1% after feature fusion, which indicates that this model can accurately classify images that neither exhibit COVID-19 nor viral pneumonia.

Figures 4 and 5 indicate that after feature fusion data and 30 epochs of training were employed, the accuracies of the validation and training sets were ≥97% for the Xception model. The loss functions of the validation and training sets decreased significantly with increasing number of epochs; thus, the Xception model approached convergence. The Xception model misclassified one CXR image of viral pneumonia as an image of healthy lungs (Fig. 6). However, it did not misclassify any images of viral pneumonia or COVID-19 as COVID-19 or healthy lungs, respectively. Thus, feature fusion data can improve the multiclass accuracy of models.

The fivefold cross-validation of the model with the highest accuracy for data set A indicated

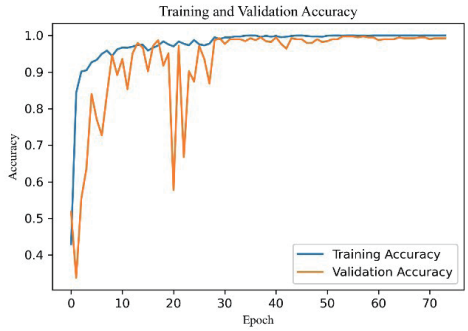


Fig. 4. (Color online) Three-class accuracy of the Xception model for the feature fusion data.

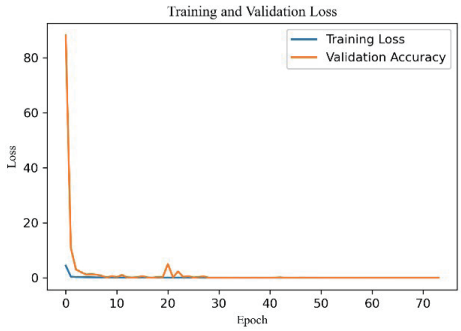


Fig. 5. (Color online) Three-class loss function of the Xception model for the feature fusion data.

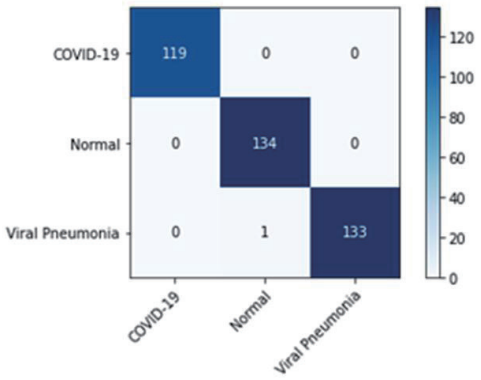


Fig. 6. (Color online) Three-class confusion matrix of the Xception model for the feature fusion data.

that after the data were divided into five folds for cyclic evaluations, the accuracy of each evaluation was higher than 98% (Table 5). This result verifies that when using feature fusion data, the Xception model can accurately detect the conditions displayed in CXR images.

3.2 Results obtained for data set B

As presented in Tables 6–8, high accuracy and sensitivity were achieved in two-class sorting when using the preprocessed original images as the input. When the feature fusion data were used, the sensitivity of most models reached 100%. A higher sensitivity indicates a lower

Table 5
Results of fivefold cross-validation for the three-class accuracy of the Xception model when using feature fusion data.

Indicator	Fold					
	First	Second	Third	Fourth	Fifth	Average
Accuracy (%)	98.83	99.35	99.35	99.48	98.96	99.1

Table 6
Model indicators of two-class deep learning for the original images.

Modes	Accuracy (%)	Precision (%)	Sensitivity (%)	*F*1 Score (%)	Specificity (%)
MobileNetV2	98.70	97.50	98.31	97.90	98.88
ResNert50	98.19	96.66	97.47	97.07	98.50
ResNet152V2	98.19	95.90	98.31	97.09	98.13
InceptionResNetV2	99.48	98.34	100	99.16	99.25
DenseNet121	99.48	99.15	99.15	99.15	99.62
Denset201	98.71	97.50	98.32	97.90	98.88
Xception	98.70	98.30	97.47	97.89	99.25

Table 7
Model indicators of two-class deep learning for the CLAHE-processed images.

Modes	Accuracy (%)	Precision (%)	Sensitivity (%)	*F*1 Score (%)	Specificity (%)
MobileNetV2	97.41	92.91	99.15	95.93	96.64
ResNert50	98.19	94.44	100	97.14	97.38
ResNet152V2	98.44	97.47	97.47	97.47	98.80
InceptionResNetV2	98.19	94.44	100	97.14	97.38
DenseNet121	98.96	97.52	99.16	98.33	98.80
Denset201	99.22	99.15	98.31	98.73	99.62
Xception	99.22	98.33	99.15	98.74	99.25

Table 8
Model indicators of two-class deep learning for the feature fusion data.

Modes	Accuracy (%)	Precision (%)	Sensitivity (%)	*F*1 Score (%)	Specificity (%)
MobileNetV2	98.44	97.47	97.47	97.47	98.88
ResNert50	99.48	98.34	100	99.16	99.25
ResNet152V2	99.22	97.54	100	98.75	98.88
InceptionResNetV2	99.22	98.33	99.15	98.74	99.25
DenseNet121	99.48	98.34	100	99.16	99.25
Denset201	99.74	99.16	100	99.58	99.62
Xception	99.74	99.16	100	99.58	99.62

possibility of a model misclassifying a COVID-19 image as a non-COVID-19 image; thus, a model with a higher sensitivity has a lower likelihood of misidentifying patients with COVID-19 as having healthy lungs or viral pneumonia. The aforementioned finding indicates that the proposed image feature fusion technique can effectively increase the COVID-19 identification accuracy of a deep learning model.

As depicted in Figs. 7 and 8, after 25 epochs of training with feature fusion data, the Xception model had validation and training accuracies higher than 98%. Furthermore, the validation and training losses of this model approached convergence.

As displayed in Fig. 9, the Xception model misclassified one non-COVID-19 CXR image as a COVID-19 image. However, this model did not misclassify any COVID-19 image as a non-COVID-19 image. These results indicate that the proposed image feature fusion technique can effectively increase the COVID-19 identification accuracy of a deep learning model.

The fivefold cross-validation of the model with the highest accuracy for data set B indicated that after the data were divided into five folds for cyclic evaluations, the accuracy of each evaluation was higher than 98% (Table 9). This result verifies that when using feature fusion data, the Xception model can accurately detect the conditions displayed in CXR images.

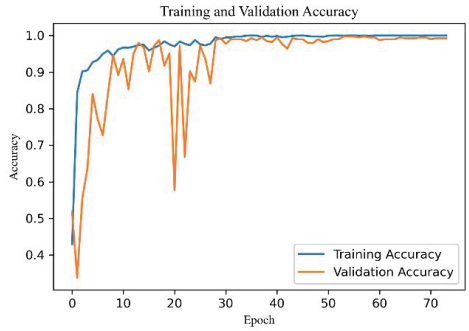


Fig. 7. (Color online) Two-class accuracy of the Xception model for the feature fusion data.

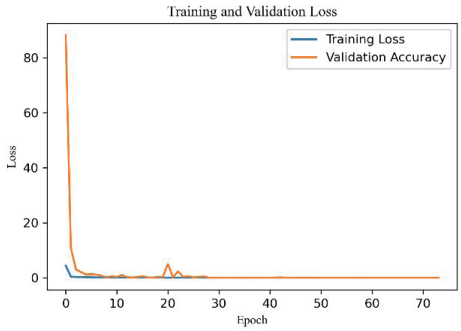


Fig. 8. (Color online) Two-class loss of the Xception model for the feature fusion data.

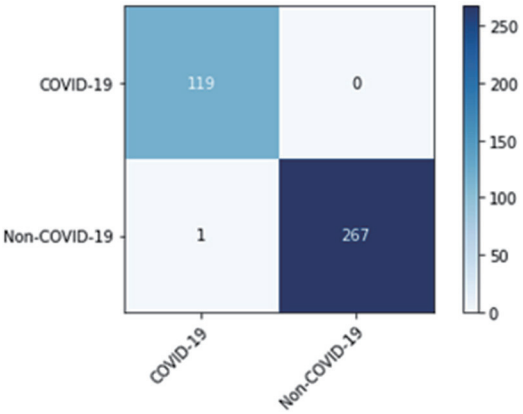


Fig. 9. (Color online) Two-class confusion matrix of the Xception model for the feature fusion data.

Table 9
Results of fivefold cross-validation for the two-class accuracy of the Xception model for the feature fusion data.

Indicator	Fold					
	First	Second	Third	Fourth	Fifth	Average
Accuracy (%)	99.22	99.61	99.87	99.48	99.35	99.50

Table 10
Three-class accuracies obtained for CXR image classification in this study and other studies.

Study	Method	Accuracy (%)
Ferhat *et al.*	COVIDiagnosis-Net[19]	98.26
Mesut *et al.*	MobileNetV2+SqueezeNet+SVM[21]	99.27
Ozturk *et al.*	DarkCovidNet[23]	87.02
Chowdhury *et al.*	DenseNet201[24]	97.94
Anunay *et al.*	InstaCovNet-19[27]	99.08
Apostolopoulos and Tzani	MobileNetV2[18]	97.42
Khan *et al.*	CoroNet[28]	95.02
Wang and Wong	COVID-Net[29]	93.33
Bhadra and Kar	Multi-layered CNN[30]	99.10
Ours	Xception+Feature Fusion	99.74

Table 11
Two-class accuracies obtained for CXR image classification in this study and other studies.

Study	Method	Accuracy (%)
Minaee *et al.*	SqueezeNet[20]	92.29
Ozturk *et al.*	DarkCovidNet[23]	98.08
Chowdhury *et al.*	DenseNet201[24]	99.70
Ohata *et al.*	MobileNet+SVM[25]	98.62
Apostolopoulos and Tzani	MobileNetV2[18]	96.78
Umri *et al.*	VGG16[26]	97.50
Anunay *et al.*	InstaCovNet-19[27]	99.53
Khan *et al.*	CoroNet[28]	99.00
Ours	Xception+Feature Fusion	99.74

3.3 Comparison of the results of this study with those of other studies

Tables 10 and 11 show the COVID-19 identification accuracies obtained in the current study and other studies. Among the compared methods, the proposed method exhibits the highest two- and three-class accuracies. To avoid incidences of images not matching reality and ensure high model stability, data enhancement was not conducted in this study.

4. Conclusions

A feature fusion technique is proposed in this paper. In this study, this technique was applied in the identification of COVID-19 in CXR images. The proposed technique was used in multiple deep learning models, which were then trained, validated, and compared. The proposed method involves image preprocessing, image processing, and feature fusion. The collected data were

divided into three- and two-class sorting data for research. No additional data enhancement techniques were used; therefore, the processed images matched reality. In three-class sorting, the Xception model achieved the following results when using feature fusion data: accuracy = 99.74%, precision = 99.75%, sensitivity = 99.75%, *F*1 score = 99.58%, and specificity = 99.62%. The average fivefold cross-validation accuracy of this model was 99.50%. The aforementioned results verify that the proposed feature fusion technique can improve the performance of deep learning models in identifying COVID-19 in CXR images.

Acknowledgments

We thank the National Science and Technology Council NSTC (grant nos. MOST 111-2221-E-019-074, NSTC 111-2634-F-019-001-, and NSTC 112-2221-E-019-023) and the National Taiwan Ocean University for supporting this research. We also thank the editor for the kind coordination. Moreover, we are grateful to the reviewers for their constructive suggestions.

References

1 J. F.-W. Chan, S. Yuan, K.-H. Kok, K. K.-W. To, H. Chu, J. Yang, F. Xing, J. Liu, C. C.-Y. Yip, R. W.-S. Poon, H.-W. Tsoi, S. K.-F. Lo, K.-H. Chan, V. K.-M. Poon, W.-M. Chan, J. D. Ip, J.-P. Cai, V. C.-C. Cheng, H. Chen, C. K.-M. Hui, and K.-Y. Yuen: Lancet **395** (2020) 514. https://doi.org/10.1016/S0140-6736(20)30154-9

2 N. Zhu, D. Zhang, W. Wang, X. Li, B. Yang, J. Song, X. Zhao, B. Huang, W. Shi, R. Lu, P. Niu, F. Zhan, X. Ma, D. Wang, W. Xu, G. Wu, G. F. Gao, D. Phil., and W. Tan: N. Engl. J. Med. **382** (2020) 723. https://doi.org/10.1056/NEJMoa2001017

3 W. Guan, Z. Ni, Yu Hu, W. Liang, C. Ou, J. He, L. Liu, H. Shan, C. Lei, D.S.C. Hui, B. Du, L. Li, G. Zeng, K.-Y. Yuen, R. Chen, C. Tang, T. Wang, P. Chen, J. Xiang, S. Li, Jin-lin Wang, Z. Liang, Y. Peng, L. Wei, Y. Liu, Ya-hua Hu, P. Peng, Jian-ming Wang, J. Liu, Z. Chen, G. Li, Z. Zheng, S. Qiu, J. Luo, C. Ye, S. Zhu, and N. Zhong: J. Emerg. Med. **58** (2020) 711. https://doi.org/10.1016/j.jemermed.2020.04.004

4 T. Ai, Z. Yang, H. Hou, C. Zhan, C. Chen, W. Lv, Q. Tao, Z. Sun, and L. Xia: Radiology **296** (2020) E32. https://doi.org/10.1148/radiol.2020200642

5 G. D. Rubin, C. J. Ryerson, L. B. Haramati, N. Sverzellati, J. P. Kanne, S. Raoof, N. W. Schluger, A. Volpi, J.-J. Yim, I. B. K. Martin, D. J. Anderson, C. Kong, T. Altes, A. Bush, S. R. Desai, J. Goldin, J. M. Goo, M. Humbert, Y. Inoue, H.-U. Kauczor, F. Luo, P. J. Mazzone, M. Prokop, M. R.-Jardin, L. Richeldi, C. M. S.-Prokop, N. Tomiyama, A. U. Wells, and A. N. Leung,: Chest **158** (2020) 106. https://doi.org/10.1016/j.chest.2020.04.003

6 H. Y. F. Wong, H. Y. S. Lam, A. H.-T. Fong, S. T. Leung, T. W.-Y. Chin, C. S. Y. Lo, M. M.-S. Lui, J. C. Y. Lee, K. W.-H. Chiu, T. W.-H. Chung, E. Y. P. Lee, E. Y. F. Wan, I. F. N. Hung, T. P. W. Lam, M. D. Kuo, and M.-Y. Ng: Radiology **296** (2020) E72. https://doi.org/10.1148/radiol.2020201160

7 A. Bernheim, X. Mei, M. Huang, Y. Yang, Z. A. Fayad, N. Zhang, K. Diao, B. Lin, X. Zhu, K. Li, S. Li, H. Shan, A. Jacobi, and M. Chung: Radiology **295** (2020) 685. https://doi.org/10.1148/radiol.2020200463

8 X. Xie, Z. Zhong, W. Zhao, C. Zheng, F. Wang, and J. Liu: Radiology **296** (2020) E41. https://doi.org/10.1148/radiol.2020200343

9 A. Krizhevsky, S. Ilya, and E. H. Geoffrey: Commun. ACM **60** (2017) 84. https://doi.org/10.1145/3065386

10 S. Ren, K. He, R. Girshick, and J. Sun: arXiv (2016) arXiv:1506.01497. https://doi.org/10.48550/arXiv.1506.01497

11 H. Almutairi, G. M. Hassan, and A. Datta: Proc. 2020 28th European Signal Processing Conf. (2020 EUSIPCO) 1382–1386. https://doi.org/10.23919/Eusipco47968.2020.9287360

12 M. Abdelazez, S. Rajan and A. D. C. Chan: IEEE Trans. Instrumentation and Measurement **70** (2021) 1. https://doi.org/10.1109/TIM.2020.3027930

13 E. Essa, and X. Xie: Proc. 2020 28th European Signal Processing Conf. (2020 EUSIPCO) 1085. https://doi.org/10.23919/Eusipco47968.2020.9287520.

14 M. Hammad, A. M. Iliyasu, A. Subasi, E. S. L. Ho, and A. A. A. El-Latif: IEEE Trans. Instrument. Meas. **70** (2020) 1. https://doi.org/10.1109/TIM.2020.3033072

15 W. Liu, X. Liu, H. Li, M. Li, X. Zhao, and Z. Zhu: IEEE J. Biomed. Health Inf. **25** (2021) 3073. https://doi.org/10.1109/JBHI.2021.3053023
16 H. Hu, Q. Li, Y. Zhao, and Y. Zhang: IEEE Trans. Ind. Inf. **17** (2021) 4. https://doi.org/10.1109/TII.2020.3022912.
17 T.-H. Lin, J.-Y. Jhang, C.-R. Huang, Y.-C. Tsai, H.-C. Cheng, and B.-S. Sheu: IEEE J. Biomed. Health Inf. **25** (2020) 1. https://doi.org/10.1109/JBHI.2020.2999731
18 I. D. Apostolopoulos, and A. M. Tzani: Phys. Eng. Sci. Med. **43** (2020) 2. https://doi.org/10.1007/s13246-020-00865-4
19 U. Ferhat, and K. Deniz: Med. Hypotheses **140** (2020) 109761. https://doi.org/10.1016/j.mehy.2020.109761
20 S. Minaee, R. Kafieh, M. Sonka, S. Yazdani, and G. J. Soufi: Med. Image Anal. **65** (2020) 101794. https://doi.org/10.1016/j.media.2020.101794
21 T. Mesut, B. Ergen, and Z. Cömert: Comput. Biol. Med. **121** (2020) 103805. https://doi.org/10.1016/j.compbiomed.2020.103805
22 J. Redmon and A. Farhadi: Proc. 2017 IEEE Conf. Computer Vision and Pattern Recognition (2017 CVPR) 6517–6525. https://doi.org/10.1109/CVPR.2017.690.
23 T. Ozturk, M. Talo, E. A.. Yildirim, U. B. Baloglu, O. Yildirim, and U. R. Acharya: Comput. Biol. Med. **121** (2020) 103792. https://doi.org/10.1016/j.compbiomed.2020.103792
24 M. E. H. Chowdhury, T. Rahman, A. Khandakar, R. Mazhar, M. A. Kadir, Z. B. Mahbub, K. R. Islam, M. S. Khan, A. Iqbal, N. A. Emadi, M. B. I. Reaz, and M. T. Islam: IEEE Access **8** (2020) 132665. https://doi.org/10.1109/ACCESS.2020.3010287
25 E. F. Ohata, G. M. Bezerra, J. V. S. d. Chagas, A. V. Lira Neto, A. B. Albuquerque, V. H. C. d. Albuquerque, and P. P. Reboucas Filho: IEEE/CAA J. Automatica Sinica **8** (2020) 239. https://doi.org/10.1109/JAS.2020.1003393
26 B. K. Umri, M. Wafa Akhyari, and K. Kusrini: Proc. 2020 2nd Int. Conf. Cybernetics and Intelligent System (2020 ICORIS) 1–5. https://doi.org/10.1109/ICORIS50180.2020.9320806.
27 G. Anunay, S. Gupta, and R. Katarya: Appl. Soft Comput. **99** (2020) 106859. https://doi.org/10.1016/j.asoc.2020.106859
28 A. I. Khan, J. L. Shah, and M. M. Bhat: Comput. Methods Programs Biomed. **196** (2020) 105581. https://doi.org/10.1016/j.cmpb.2020.105581
29 Z. Q. L. Wang and A. Wong: Sci. Rep. **10** (2020) 1. https://doi.org/10.1038/s41598-020-76550-z
30 R. Bhadra and S. Kar: Proc. 2020 IEEE Bombay Section Signature Conf. (2020) 214–218. https://doi.org/10.1109/IBSSC51096.2020.9332210
31 M. R. Karim, T. Döhmen, M. Cochez, O. Beyan, D. Rebholz-Schuhmann, and S. Decker: Proc. 2020 IEEE Int. Conf. Bioinformatics and Biomedicine (2020 BIBM) 1034–1037. https://doi.org/10.1109/BIBM49941.2020.9313304.
32 A. Asmaa, M. M. Abdelsamea, and M. M. Gaber: Appl. Intell. **51** (2021) 2. https://doi.org/10.1007/s10489-020-01829-7
33 F. Li, X. Lu, and J. Yuan: IEEE Trans. Med. Imaging **41** (2022) 1208. https://doi.org/10.1109/TMI.2021.3134270
34 K. Zuiderveld: Academic Press Professional, Inc., USA (1994). https://www.tamps.cinvestav.mx/~wgomez/material/AID/CLAHE.pdf
35 M. Sandler, A. Howard, M. Zhu, A. Zhmoginov, and L.-C. Chen: Proc. 2018 IEEE/CVF Conf. Computer Vision and Pattern Recognit. (IEEE 2018). 4510–4520. https://doi.org/10.1109/CVPR.2018.00474.
36 K. He, X. Zhang, and S. Ren, J. Sun: Proc. 2016 IEEE Conf. Computer Vision and Pattern Recognition (2016 CVPR) 770–778. https://doi.org/10.1109/CVPR.2016.90.
37 K. He, X. Zhang, S. Ren, and J. Sun: Computer Vision-ECCV 2016, B. Leibe, J. Matas, N. Sebe, and M. Welling Eds. (2016) 630. https://doi.org/10.1007/978-3-319-46493-0_38
38 C. Szegedy, S. Ioffe, V. Vanhoucke, and A. A. Alemi: arXiv:1602.07261 (2016). https://doi.org/10.48550/arXiv.1602.07261
39 G. Huang, Z. Liu, L. Van Der Maaten, and K. Q. Weinberger: Proc. 2017 IEEE Conf. Computer Vision and Pattern Recognition (2017 CVPR) 2261–2269. https://doi.org/10.1109/CVPR.2017.243.
40 F. Chollet: Proc. 2017 IEEE Conf. Computer Vision and Pattern Recognition (2017 CVPR) 1800–1807. https://doi.org/10.1109/CVPR.2017.195.

About the Authors

Chih-Ta Yen received his Ph.D. degree from the Department of Electrical Engineering, National Cheng Kung University, Taiwan, in 2008. He is currently a professor in the area of artificial intelligence applications in engineering at the Department of Electrical Engineering, National Taiwan Ocean University, Keelung, Taiwan. His major interests are in the areas of multiuser optical communications, wireless communication systems, machine learning, deep learning, and optical design. (chihtayen@gmail.com)

Jia-Xian Liao was born in Taichung, Taiwan, in September 1996. He has already graduated with a master's degree at National Formosa University, Taiwan, majoring in electrical engineering. His major interests are in the areas of deep learning, machine learning, and embedded systems. (10865120@gm.nfu.edu.tw)

Yi-Kai Huang was born in Yunlin, Taiwan, in October 1996. He has already graduated with a master's degree at the Department of Electrical Engineering, National Formosa University. His major interests are in the area of deep learning. (10865122@gm.nfu.edu.tw)