

Received February 2, 2020, accepted February 9, 2020, date of publication February 24, 2020, date of current version March 3, 2020.

Digital Object Identifier 10.1109/ACCESS.2020.2975198

# Discriminative Feature Learning for Skin Disease Classification Using Deep Convolutional Neural Network

**BELAL AHMAD<sup>1</sup>**, **MOHD USAMA<sup>1</sup>**, **CHUEN-MIN HUANG<sup>2</sup>**, **KAI HWANG<sup>3</sup>**, (Life Fellow, IEEE), **M. SHAMIM HOSSAIN<sup>4</sup>**, (Senior Member, IEEE), AND **GHULAM MUHAMMAD<sup>5</sup>**

<sup>1</sup>School of Computer Science and Technology, Huazhong University of Science and Technology, Wuhan 430074, China

<sup>2</sup>Department of Information Management, National Yunlin University of Science and Technology, Yunlin 64002, Taiwan, R.O.C.

<sup>3</sup>The Chinese University of Hong Kong, Shenzhen, Shenzhen 518172, China

<sup>4</sup>Department of Software Engineering, College of Computer and Information Sciences, King Saud University, Riyadh 11543, Saudi Arabia

<sup>5</sup>Department of Computer Engineering, College of Computer and Information Sciences, King Saud University, Riyadh 11543, Saudi Arabia

Corresponding author: Kai Hwang (hwangkai@cuhk.edu.cn)

This work was supported by a Research Grant from the Shenzhen Institute of Artificial Intelligence and Robotics for Society (AIRS), which was hosted at The Chinese University of Hong Kong, Shenzhen, China.

**ABSTRACT** Nowadays, skin disease among humans has been a common disease, especially in America millions of people are suffering from various kinds of skin disease. Usually, these diseases have hidden dangers which lead to human not only lack of self-confidence and psychological depression but also a risk of skin cancer. Diagnosis of these kinds of diseases usually required medical experts with high-level instruments due to a lack of visual resolution in skin disease images. Moreover, manual diagnosis of skin disease is often subjective, time-consuming, and required more human effort. Thus, there is a need to develop a computer-aided system that automatically diagnoses the skin disease problem. Moreover, most of the existing works in skin disease used convolutional neural networks (CNN) with classical loss functions, which limit the model to learn discriminative features from skin images. Thus to address the above mention problem we proposed a new framework by fine-tuning layers of ResNet152 and InceptionResNet-V2 models with a triplet loss function. In the proposed framework, first, we learning the embedding from input images into Euclidean space by using deep CNN ResNet152 and InceptionResNet-V2 model. Second, we compute L-2 distance among corresponding images from euclidean space to learn discriminative features of skin disease images by using triplet loss function. Finally, classify the input images using these L-2 distances. Human face skin disease images used in the proposed framework are acquired from the Hospital in Wuhan China. Experiment results and their analysis shows the effectiveness of the proposed framework which achieve better accuracy than many existing works in skin disease tasks.

**INDEX TERMS** Discriminative feature learning, convolutional neural networks, triplet loss function, skin disease.

## I. INTRODUCTION

At present, skin diseases are one of the most infectious diseases to see among people. Because of the physical structure affected by the direct exposure to ultraviolet radiation i.e. use of different types of high-frequency wireless equipment for a long time and it can develop skin cancer. In the United States, according to the statistical report of 2012, more than 3.3 million non-melanoma cases including basal cell carcinoma and squamous cell carcinoma are tested amongst 5.4 million individuals [1] whereas, the new

melanoma cases are diagnosed is estimated over 91,270 in 2018. The number of new occurrences of skin cancer such as breast, lung, prostate gland and colon increases every year [2]. Additionally, research analysis shows that 20% of Americans will produce skin cancer during their life [3]. The diagnosis of skin diseases is very challenging due to the symptom of skin disease is an extended and continuously transforming process with occurring in certain areas of the skin. In order to diagnose skin disease, a lot of clues such as specific lesion anatomy, physical assortment, scaling, shade, and arrangement can be used. The recognition process may be quite complicated by individually analyzing specific components [4], because of primarily four significant medical

The associate editor coordinating the review of this manuscript and approving it for publication was Marcin Woźniak<sup>1</sup>.

diagnosis methods such as ABCD rules, Menzies method, 7-Point Checklist and pattern analysis are used for correctly analyzed the melanoma skin cancer. All these methods achieve an effective result for the skin cancer diagnosis. To distinctness of skin lesions [5], the extreme level of skills is needed instead of the human specialist diagnoses, which relies on an individual opinion that could be barely reproducible [6], [7]. The computer-assisted diagnostic method is more effective and well-performing because of using the standard feature extraction method with well-known classifiers such as SVM and ANN. The diagnostic systems [8]–[10] achieve effective results on melanoma (specific skin cancer), whereas these are ineffective to execute on a wide range of skin disease categories.

Since hand-crafted features learning algorithms usually are committed to a single or a minimal number of subcategories and not effective in classification for a higher number of skin disease categories. On the other hand, a hand-crafted feature extraction algorithm impractical due to variation in the nature of skin diseases [11]. To address this issue, it requires to feature learning [12] rather than feature engineering to select essential features by machine. In past years, many feature learning-based classification approaches [13]–[17] are proposed. Whereas, they are limited to dermoscopy or even histopathology and especially emphasize the mitosis diagnosis that is a cancer indication [18], [19].

To deal with the problem, we proposed a discriminative feature learning approach based on transfer-learning for skin disease classification. The contributions of the paper are as follows:

- We proposed a new deep CNN based model for skin disease classification using triplet loss function.
- To learn discriminative features from skin disease images, we fine-tune CNN based model (i.e. ResNet152 and InceptionResNet-V2) with triplet loss function. To the best of our knowledge, no one used triplet loss function in the skin disease images.
- We perform layer-wise fine-tuning of pre-train deep CNN models, instead of block wise [20], to improve the performance of the end-to-end learning method.

## II. RELATED WORK

Recently, deep CNN gets incredibly preferred to feature learning and image classification. A network can train on a large-scale dataset using high-performance GPU to achieve effective results. Many types of research [21]–[23] on ImageNet [24] show that the traditional methods using deep CNN obtained better performance over humans in object classification.

Lately, Esteva *et al.* [25] designed a universal skin disease classification method by fine-tuning the VGG16 and VGG19 architectures to train the network. Their network obtained 60.0% and 80.3% Top-1 and Top-3 classification accuracy respectively, which is significantly surpassed the social skills in their experiments. They also encouraged to use a similar approach to obtain a better result.

Tajbakhsh *et al.* [26] explained that a pre-trained network is better to use a deep CNN on scratch due to limited labeled input data; they solved the skin disease problem by a pre-trained network using images from other medical domain more effectively than train a deep CNN from the beginning. Giotis *et al.* [27] developed a decision support system using deep CNN, which utilized a different set of features such as color, visual and lesion texture. Haenssle [28] composed a method using deep CNN to classify the dermoscopy images with binary diagnostic categories. Dorj *et al.* [29] designed ECOC SVM based on deep CNN to organize skin cancer images of four diagnostic categories. Han *et al.* [30] proposed an image classification method using deep CNN, which classify the clinical images of 12 skin disease categories. Mohd *et al.* [31] proposed an image classification method to classify melanoma, the dataset of four types of skin diseases used in all experiments. Almansour and Jaffar [32] composed a technique using the k-means clustering and Support Vector Machine (SVM) to classify melanoma; they also demonstrate the comparative results. German *et al.* [33] explained the skin cancer diagnostic method by separately using AdaBoost MC. The dataset from a different category of skin lesions is used for cancer detection. Ioannis *et al.* [27] designed a support system using the image processing algorithm and the deep CNN for utilizing the different sets of features such as visual diagnostic attributes, color, lesion texture, affected area and degree of damage for melanoma.

## III. METHOD

To preserve the stability of the method, we used two well-trained CNNs, the implementation details are discussed in the next section. We fine-tune the networks to obtain  $128 - D$  embedding  $f(I)$  into the Euclidean space  $R^d$  then calculate L-2 distance between all input images using corresponding  $128 - D$  embedding, whereas L-2 distance is independent to each factor concerning the input images such as physical assortment, scaling, shade and arrangement. Although we have not directly compared the different class images, whereas computing the loss between these images then loss between same class images are projected at a single point into the Euclidean space. This triplet loss puts the margin between every image from the same class and differentiates them from other class images. The schematic view of the proposed method as shown in Figure 1.

## IV. TRIPLET LOSS

Currently, deep metric learning approaches using Triplet Loss get more consideration due to interacting capability of performance management with an extreme of labels. Prabhu *et al.* [34] presented an extreme multi-label classification. In general, conventional methods for multi-label classification linearly increase the number of parameters. Song *et al.* [35] and Weinberger *et al.* [36] learns an N-way softmax classifier with an extreme level of labels. Hence, the convolutional neural network using a triplet loss function learns a compact embedding to handle the classification

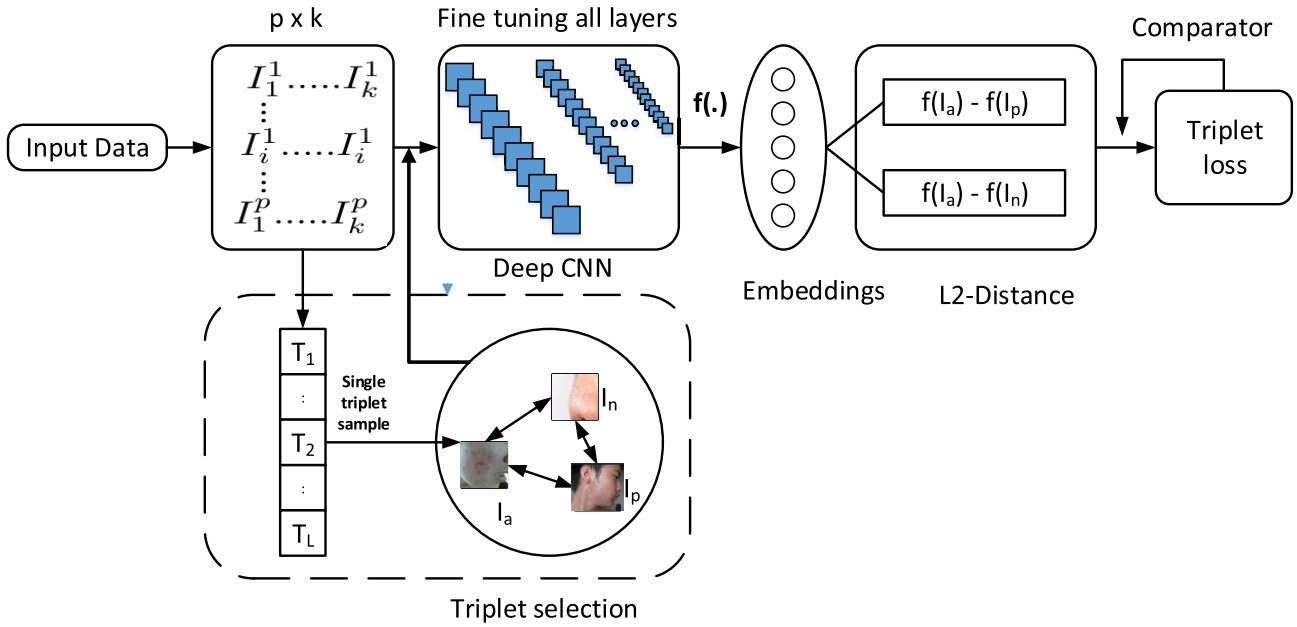


FIGURE 1. Proposed model architecture for skin disease classification.

problems effectively. Since triplet loss learns good embeddings  $f(I^i)$  of an image  $I^i$  into d-dimensional Euclidean space  $R^d$ , i.e.,  $f(I^i) \in R^d$ . Furthermore it ensure that a particular face image  $I_a$  (anchor) is closer to all images from same class  $I_p^i$  (positive) than images from the different class  $I_n^i$  (negative). Thus our objective is

$$\|f(I_a) - f(I_p^i)\|_2^2 + \alpha_t < \|f(I_a) - f(I_n^i)\|_2^2 \quad \forall f(I_a), f(I_p^i), f(I_n^i) \in R^d \quad (1)$$

where  $f(I_a)$ ,  $f(I_p^i)$  and  $f(I_n^i)$  are the embedding of a triplet  $(I_a, I_p^i, I_n^i)$  from a set of all possible triplets  $T$  in training set with major  $N$  and threshold  $\alpha_t$  is a predefined margin that imposes between images from different class. The Euclidean distance minimizes the triplet loss is defined as

$$L_t = \sum [ \|f(I_a) - f(I_p^i)\|_2^2 - \|f(I_a) - f(I_n^i)\|_2^2 + \alpha_t ] \quad (2)$$

These triplets would not be active during training in slower convergence. Hence it is crucial to select hard triplets, which are activated in training to improve the performance of the model. In the next section, we have discussed a triplet selection approach that we have used in this method.

### V. TRIPLET SELECTION

To assure fast convergence, triplet selection is essential that select the triplets to violet triplet constraint by generating bias in selection. Since triplet selection requires to adjust the trade-off between hard triplets to generating bias instead of select triplet properly that imbalance the trade-off, we directly minimize the bias to avoid this problem. In particular, make an equal contribution of all possible triplets in an unbiased situation. We compute argmax of metric values  $\|f(I_a) - f(I_p^i)\|_2^2$  and argmin of metric values  $\|f(I_a) - f(I_n^i)\|_2^2$  respectively,

where  $(I_p^i)$  (hard positive) and  $(I_n^i)$  (hard negative) for given image  $(I_a)$ . Since computing argmax and argmin values for the entire dataset is infeasible. Furthermore, the poor images would dominate the hard positives and negatives and mislabeled that can be caused by poor training. To overcome this issue, we online generate triplets and create mini-batch of few input samples to compute argmax and argmin within mini-batch [37].

To keep a notable presentation for positive L-2 distance, we need to create a mini-batch of a minimum number of input images from individual class. In our case, we choose 200 images from each class including a few random negative images to make a mini-batch. In case of an obtained indecisive result due to generate the triplets, for small mini-batch, choosing the hardest negative leads to bad local minima at the initial stage of training, especially for a poor network (i.e.,  $f(I) = 0$ ). Hence, we reduce this problem by selecting  $I_n^i$  such that

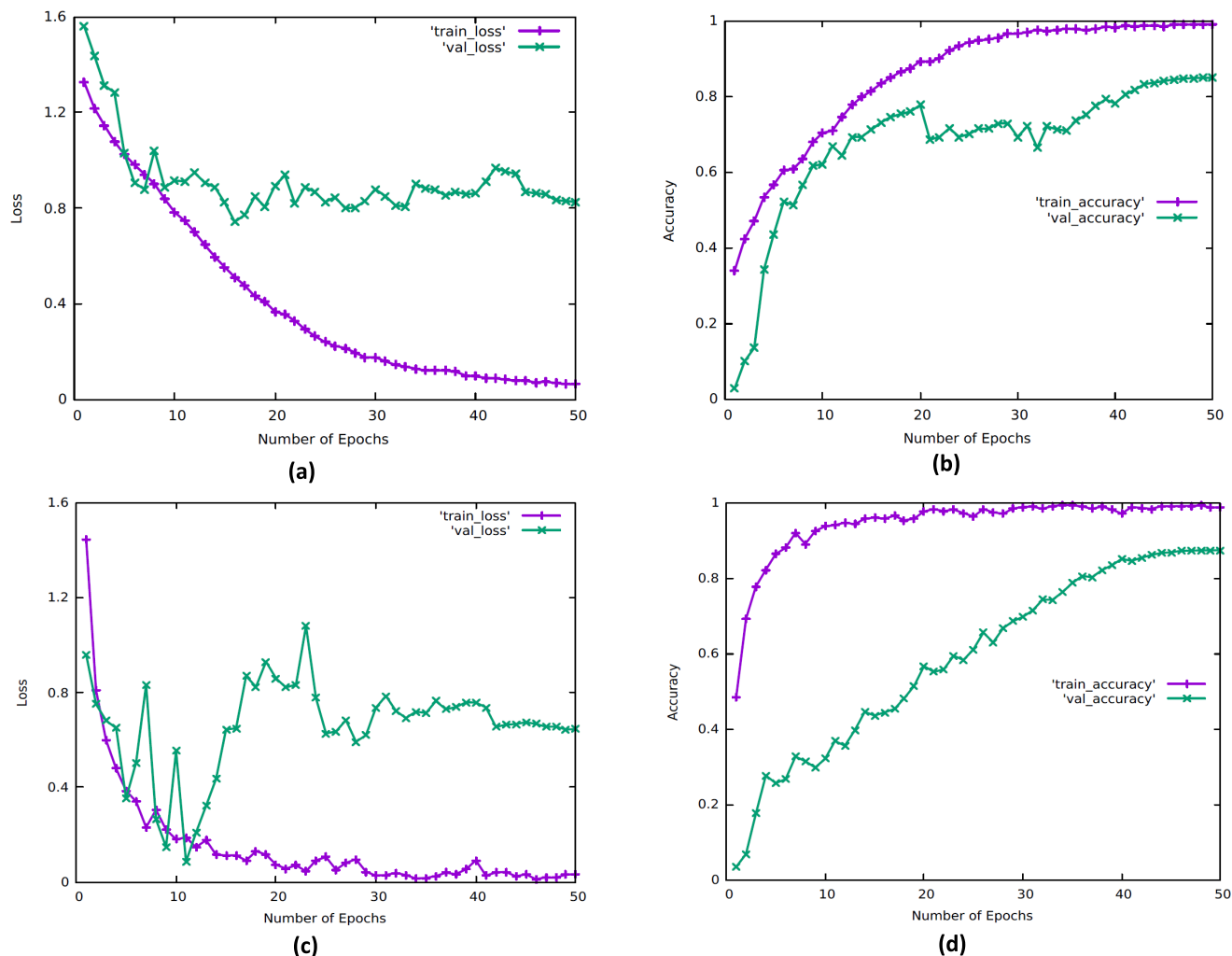
$$\|f(I_a) - f(I_p^i)\|_2^2 - \|f(I_a) - f(I_n^i)\|_2^2 \quad (3)$$

The distance of anchor from a negative image is greater than the positive image, whereas L-2 distance is closer to anchor positive distance. Thus we called semi-hard to these negative images and the range of these images is under the value of pre-defined margin  $\alpha_t$ .

In order to improve the convergence rate, we need to take a small mini-batch. In our case, we choose a mini-batch size of 800 samples to reduce the obstacle of appropriate hard triplet selection from each mini-batch.

### VI. DEEP CONVOLUTIONAL NEURAL NETWORK

Deep CNN can be effectively trained in many cases of transfer learning [38], [39]. Researchers choose a pre-trained model



**FIGURE 2.** Trend of loss and accuracy of proposed method. Figures (a) & (b) shows the loss and accuracy curve against number of epochs for networks1 (i.e., fine tune ResNet152 with triplet loss) and Figures (c) & (d) shows the loss and accuracy curve against number of epochs for networks2 (i.e., fine tune InceptionResNet-V2 with triplet loss).

for fine-tuning the weights by backpropagation, instead of training the network using randomly initialized parameters. Since previous layers of the pre-trained network have some generic features, which can be quite useful in decision function. Especially, those features are directly used, when network train on the new dataset.

In our method, we changed and fine-tuned two well-trained CNNs architectures: ResNet152 [40] and InceptionResNet V2 [41] for feature extraction. We customized both networks by removing the last fully connected layer. First, flatten the output from the feature map then take one FC layer with 512 neurons afterward use dropout layer with rate of 0.3 and finally the 12 regularization layer with 128 –  $D$  features to learn embedding vectors of dimension 128 as an identity descriptor, which avoids the connections to preserve the power of error signals for improve the network performance. Moreover, it helps to reduce the vanishing gradient problem.

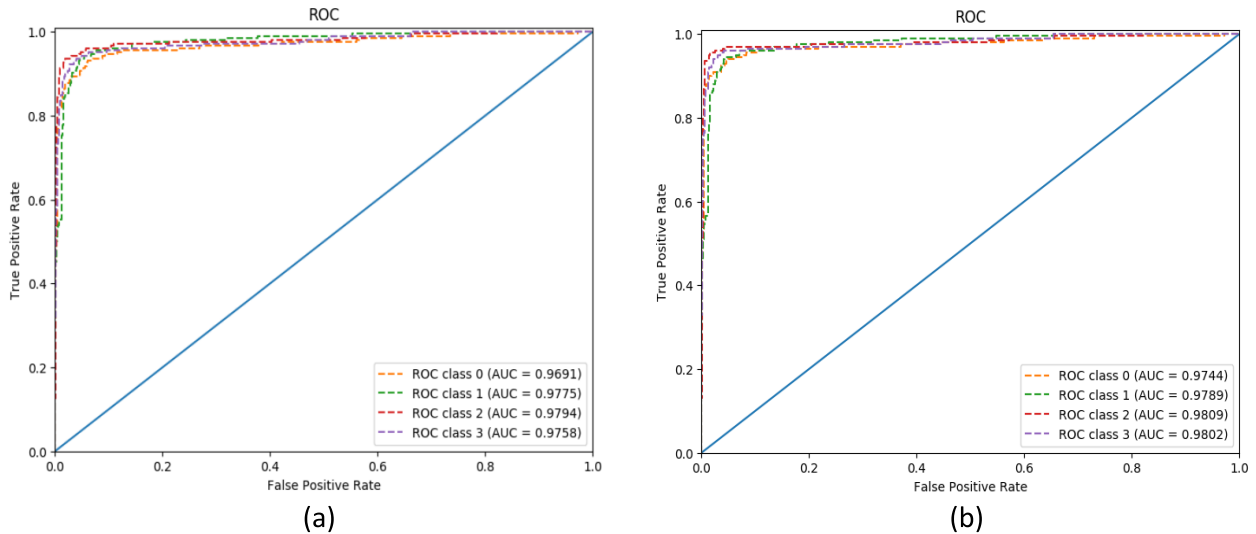
In this method, we fine-tune two well-trained networks with a skin disease dataset by re-initializing the data layers. In this process, we use the NVIDIA GTX 980 Ti

GPU to accelerate the computation for full implementations and Tensorflow library to improve the network definition (train-validation protocol) of the pre-trained models. In the data layer, the input data specified by a text file (source-parameter), each line of text file indicates the individual file name with labels. At the training time, we set the new height, new width with the number of input images and batch size. As mentioned in the next section, we augment and resized the input images of the same value (224, 224) as the pre-trained networks to be compatible with this model. Thus, our networks have the same scale as the pre-trained networks, which are used in our implementation. To optimize the networks, we use stochastic gradient descent (SGD) [42] with a learning rate of 0.0001 and a momentum of 0.8.

## VII. DATASET AND EVALUATION METHODS

### A. DATASET

The dataset for the human face skin disease classification is built [43], the images are acquired from web pages by keyword searches. A total of 6144 images are obtained with



**FIGURE 3.** Area Under Curve (AUC) values for all category of skin disease for proposed methods (a)-fine tune model ResNet152 with triplet loss and (b)-fine tune InceptionResNet-V2 with triplet loss respectively.

labels of 14 classes, then dermatologists from Wuhan Union Hospital classify all skin images and unified these images of 14 classes into five categories: acne, spots, blackheads, dark circles and clean face, which is categorized based on characteristics of diseases. Clean faces are considered as the negative samples and each type of skin disease such as acne, spots, blackheads and dark circles are selected as positive samples. Afterward, make the ratio of positive and negative samples is 1:1 to avoid the sample imbalance problem. Finally, we preprocessed these images under few arguments such as rescale, zoom with range 0.3, rotation with range 0.3, horizontal flip, and resized the input with the size of width and height is 224 then augmented the dataset to 14000 samples to improve the results.

**B. EVALUATION METHODS**

In this experiment, deep CNNs using a triplet loss function classify extracted features into four categories, i.e., acne, blackheads, dark circles and spots.

The accuracy, sensitivity (recall) and specificity are computed in this analysis by using the following methods.

Accuracy (The probability of this method):

$$Accuracy = \frac{TP + TN}{TP + FP + TN + FN} \tag{4}$$

Sensitivity or Recall shows the true prediction value with respect to wrongly predicted as negative ailments (true positive rate):

$$Sensitivity(Recall) = \frac{TP}{TP + FN} \tag{5}$$

Specificity shows the true negative rate:

$$Specificity = \frac{TN}{TN + FP} \tag{6}$$

where TP is true positive or correctly predicted as positive ailments (i.e., same diseases), TN is true negative or correctly predicted as negative ailments (i.e., different diseases),

FP is false positive or wrongly predicted as positive ailments (i.e., same diseases), FN is false negative or wrongly predicted as negative ailments (i.e., different diseases).

We also used AOC (area under curve) for skin disease classification. The AOC curve graphically illustrates the classification ability of our method. The curve is plotted TPR (true positive rate) against FPR (false positive rate). The TPR and FPR are computed using following methods:

$$TPR = \frac{TP}{(TP + FN)} \tag{7}$$

$$FPR = \frac{FP}{(FP + TN)} \tag{8}$$

The greater value of AOC indicates that the model is better (can see in Figure 3).

**VIII. EXPERIMENTS**

In all our experiments, we use a total of 12000 skin disease images of four categories to extract the 128-D features. 10% of training data are used for validation by splitting the dataset during training. In the next section, we describe the feature interpretation process in detail.

**A. DEEP LEARNING FEATURE INTERPRETATION**

In order to classify the skin diseases, we used the deep CNNs to extract the embeddings  $f(I)$  from the input image  $I$  into an Euclidean space  $R^d$ . Here, we explore the different dimensions of embeddings and finally take 128 for all our experiments; it is possible to achieve better results by taking a more significant size after prolonged training.

It has been noted that training with a 128 –  $D$  floating vector that can be quantized to 128-bytes without accuracy loss, which is better for large-scale image recognition and clustering. Whereas smaller embeddings size can get minor loss of accuracy, which can be better to implement for mobile devices.



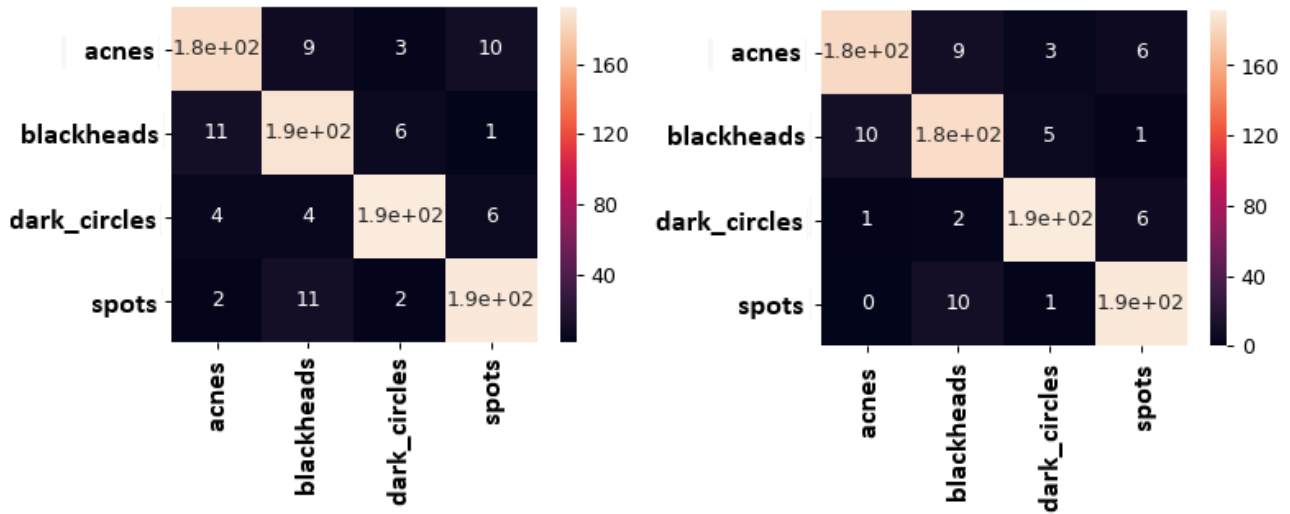


FIGURE 4. Confusion matrices of proposed method shows the prediction values for all categories of skin disease.

TABLE 1. Classification performance of proposed method. Total of 200 skin disease images from each class are used for validation.

Categories	ResNet152 + Triplet				Inception ResNet-V2 + Triplet			
	Accuracy (%)	Sensitivity (recall) (%)	Specificity (%)	no. of samples	Accuracy (%)	Sensitivity (recall) (%)	Specificity (%)	no. of samples
acnes	-	91.54	94.30	200	-	96.29	96.90	200
blackheads	-	88.67	89.75	200	-	96.92	95.57	200
dark-circles	-	95.52	95.50	200	-	97.57	98.40	200
spots	-	91.82	93.56	200	-	97.40	98.06	200
Average	84.91	91.88	93.28	200	87.42	97.04	96.48	200

**B. CLASSIFICATION**

Here, we classify the feature vector of four classes of input images. After interpreting the features, we use triplet loss function with deep CNN to learn good embeddings and computes L-2 distance between all input images into d-dimensional Euclidean space ( $R^d$ ), which directly corresponds to image similarity. The L-2 distance between the same class images is small, whereas large for different class images are shown in Equation. The distance between images is independent of different imaging factors.

We used two pre-trained networks; ResNet152 and InceptionResNet-V2 to classify the preprocessed images. Figure 3 shows the area under the curve (AUC) for each skin disease category of proposed networks i.e. fine-tune the ResNet152 and fine-tune the InceptionResNet-V2. The loss function optimizes the training of networks. We used 70% of images as a training set, 10% for validation to obtain the quick stopping epoch and 20% of data as the testing set is used to evaluate the method. We augmented each class’s images due to the visual similarity between different types of skin diseases with some arguments that are mentioned before. We summarized the proposed network’s accuracy with their classification performance with a number of

samples using for prediction mentioned in Table 1. The next section described the results of the proposed method.

**IX. RESULTS AND DISCUSSION**

In this process, we use the NVIDIA GTX 980 Ti GPU to accelerate the computation for full implementations and Tensorflow library to improve the network definition (train-validation protocol) of the pre-trained models. To record the running time of fine-tuned ResNet152 with triplet loss and fine-tuned with InceptionResNet-V2 architectures, we selected 188 iterations and extracted the features of dimension 128. The running time of fine-tuned ResNet152 with triplet loss on the system with specification of 4 GB CUDA cores GPU is 34142 second while the fine-tuned with InceptionResNet-V2 architectures takes 43714 seconds for training. We recorded the performance of the proposed method based on several aspects such as different parameter value effects, training error, and accuracy (true accept rate).

**A. TRAINING LOSS AND ACCURACY**

We have plotted the loss and accuracy curve against the number of epochs, which is shown in Figure 2.

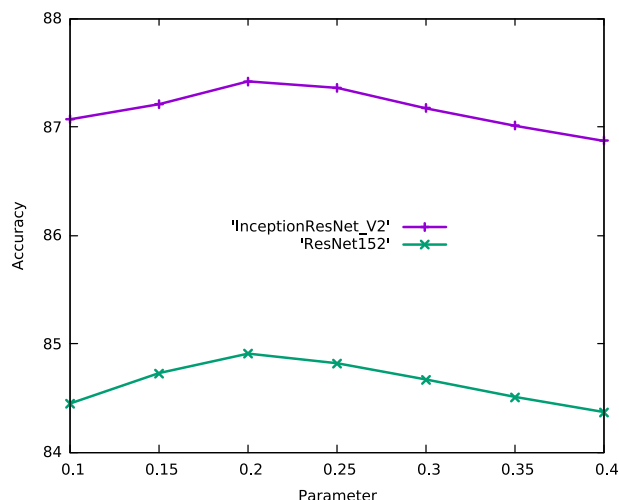


FIGURE 5. Performance of proposed model on different parameter ( $\alpha_t$ ) values. Models achieved best result at 0.2.

TABLE 2. Comparison of proposed method with some traditional methods. The value in bold indicates the best result.

Method	Accuracy (%)	Sensitivity (%)	Specificity (%)
Esteva et al. [25]	72.1	96	-
Anas et al. [31]	83.33	-	-
German et al. [33]	-	90	77
Ioannis et al. [27]	81	-	-
Daniel et al. [44]	-	78.43	<b>97.87</b>
Qaisar et al. [45]	-	89.28	93.75
Blum et al. [46]	87.2	92.6	87.2
Modified GoogLeNet [47]	81.03	-	-
Fine-tuned AlexNet [48]	77.55	-	-
Amelard et al. [49]	81.17	96.64	65.06
Proposed method	<b>87.42</b>	<b>97.04</b>	97.23

In addition, the confusion matrices of proposed networks shown in Figure 4. The row values of confusion metrics denote the corresponding true label and column indices denote the corresponding predicted labels. The value that appears in each cell shows the prediction labels. Here, we can see that the diagonal cell of confusion metrics acquiring a high level of prediction, which indicates the low error rate (high probability of accurate prediction) for each category of skin diseases. To test the efficiency of the proposed method, we computed some evaluation factors, namely, sensitivity (recall), specificity and accuracy using confusion matrices, which are demonstrated in Table 1. Additionally, we computed AUC (true accept rate against false accept rate) values for all categories of skin diseases.

**B. EFFECT OF PARAMETERS**

To balance the training and validation loss measure the impact on classification results with the convergence rate, we tested the network performance on different  $\alpha_t$  values, which shown in Figure 5, the classification accuracy improved with the threshold, and started to drop when the threshold was taken

about 0.20. Therefore, we could assume that we achieved the best result on a threshold value of 0.20.

**C. COMPARISON**

Here, we discussed the comparative study with some other traditional methods to examine the concernment of the proposed method. While four types of skin diseases i.e., acne blackheads, dark circles, and spots are classified in our method, skin diseases are classified in many traditional methods. Hence, we compared the results of the proposed method (accuracy, specificity, and sensitivity) with some previous works are shown in Table 2. The accuracy and specificity of our method is better than [25], [27], [31], [33], [44]–[49]. Even though the specificity of [44] is more than the proposed method but their sensitivity is comparatively very low.

**X. CONCLUSION AND FUTURE WORKS**

In this paper, we proposed a model using deep CNN with triplet loss function to improve the skin disease classification. We fine-tune all layers of ResNet152 and InceptionResNet-V2 to address the problem of facial skin disease images. First, we extract the 128-D features (embeddings) from training samples into Euclidean space then computes L-2 distances between corresponding images using learned embeddings. After that, we perform skin disease classification task considering L2 distance among images. The dataset used in the experiment consists of four types of skin diseases, i.e., acne, blackheads, dark circles, and spots. A total of 12000 input images were used for training, 2000 for testing, and 10% of training data for the validation set to evaluate the method. Our method outperforms then state-of-art works in skin disease classification. The proposed method can also be used for other disease classification tasks. Since the biological taxonomy organizes the dermnet images, thus the performance of the proposed method can be improved by designing a dataset with the help of a dermatologist to visually organized taxonomy.

**ACKNOWLEDGMENT**

The authors extend their appreciation to the Researchers Supporting Project number (RSP-2019/34), King Saud University, Riyadh, Saudi Arabia for funding this work.

**REFERENCES**

- [1] H. W. Rogers, M. A. Weinstock, S. R. Feldman, and B. M. Coldiron, "Incidence estimate of nonmelanoma skin cancer (Keratinocyte Carcinomas) in the US population, 2012," *JAMA Dermatol.*, vol. 151, no. 10, p. 1081, Oct. 2015.
- [2] R. L. Siegel, K. D. Miller, and A. Jemal, "Cancer statistics, 2015," *CA, A Cancer J. Clinicians*, vol. 65, no. 1, pp. 5–29, Jan. 2015.
- [3] J. K. Robinson, "Sun exposure, sun protection, and vitamin D," *J. Amer. Med. Assoc.*, vol. 294, no. 12, pp. 1541–1543, Sep. 2005.
- [4] N. Cox and I. Coulson, *Diagnosis of Skin Disease. Rook's Textbook of Dermatology*, vol. 5, 7th ed. Oxford, U.K.: Blackwell Science, 2004.
- [5] J. D. Whited, "Does this patient have a mole or a melanoma?" *J. Amer. Med. Assoc.*, vol. 279, no. 9, pp. 696–701, Mar. 1998.
- [6] A. Masood and A. Ali Al-Jumaily, "Computer aided diagnostic support system for skin cancer: A review of techniques and algorithms," *Int. J. Biomed. Imag.*, vol. 2013, pp. 1–22, 2013.

- [7] A. Steiner, H. Pehamberger, and K. Wolff, "Improvement of the diagnostic accuracy in pigmented skin lesions by epiluminescent light microscopy," *Anticancer Res.*, vol. 7, pp. 433–434, 1987.
- [8] C. Barata, J. Marques, and T. Mendonça, "Bag-of-features classification model for the diagnose of melanoma in dermoscopy images using color and texture descriptors," in *Proc. Int. Conf. Image Anal. Recognit.* Lecture Notes in Computer Science, vol. 7950, M. Kamel and A. Campilho, Eds. Berlin, Germany: Springer, 2013, pp. 547–555.
- [9] G. Fabbrocini, V. De Vita, S. Cacciapuoti, G. Di Leo, C. Liguori, A. Paolillo, A. Pietrosanto, and P. Sommella, "Automatic diagnosis of melanoma based on the 7-point checklist," in *Computer Vision Techniques for the Diagnosis of Skin Cancer* (Series in BioEngineering), J. Scharcanski and M. E. Celebi, Eds. Berlin, Germany: Springer, 2014, pp. 71–107.
- [10] M. Chen, Y. Hao, K. Hwang, L. Wang, and L. Wang, "Disease prediction by machine learning over big healthcare data," *IEEE Access*, vol. 5, pp. 8869–8879, 2017.
- [11] A. Sáez, B. Acha, and C. Serrano, "Pattern analysis in dermoscopic images," in *Computer Vision Techniques for the Diagnosis of Skin Cancer* (Series in BioEngineering), J. Scharcanski and M. E. Celebi, Eds. Berlin, Germany: Springer, 2014, pp. 23–48.
- [12] Y. Bengio, A. Courville, and P. Vincent, "Representation learning: A review and new perspectives," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 35, no. 8, pp. 1798–1828, Aug. 2013.
- [13] H. Chang, Y. Zhou, A. Borowsky, K. Barner, P. Spellman, and B. Parvin, "Stacked predictive sparse decomposition for classification of histology sections," *Int. J. Comput. Vis.*, vol. 113, no. 1, pp. 3–18, Dec. 2014.
- [14] M. Chen, X. Shi, Y. Zhang, D. Wu, and M. Guizani, "Deep features learning for medical image analysis with convolutional autoencoder neural network," *IEEE Trans. Big Data*, to be published, doi: 10.1109/TBDDATA.2017.2717439.
- [15] A. Cruz Roa, J. E. A. Ovalle, A. Madabhushi, and F. A. G. Osorio, "A deep learning architecture for image representation, visual interpretability and automated basal-cell carcinoma cancer detection," in *Medical Image Computing and Computer-Assisted Intervention-MICCAI 2013*. Springer, 2013, pp. 403–410.
- [16] A. Cruz-Roa, A. Basavanahally, F. González, H. Gilmore, M. Feldman, S. Ganesan, N. Shih, J. Tomaszewski, and A. Madabhushi, "Automatic detection of invasive ductal carcinoma in whole slide images with convolutional neural networks," *Proc. SPIE*, vol. 9041, Mar. 2014, Art. no. 904103.
- [17] J. Arevalo, A. Cruz-Roa, V. Arias, E. Romero, and F. A. González, "An unsupervised feature learning framework for basal cell carcinoma image analysis," *Artif. Intell. Med.*, vol. 64, no. 2, pp. 131–145, Jun. 2015.
- [18] H. Wang, A. Cruz-Roa, A. Basavanahally, H. Gilmore, N. Shih, M. Feldman, J. Tomaszewski, F. Gonzalez, and A. Madabhushi, "Cascaded ensemble of convolutional neural networks and handcrafted features for mitosis detection," *Proc. SPIE*, vol. 9041, Mar. 2014, Art. no. 90410.
- [19] M. Chen, Y. Hao, H. Gharavi, and V. C. M. Leung, "Cognitive information measurements: A new perspective," *Inf. Sci.*, vol. 505, pp. 487–497, Dec. 2019.
- [20] T. Schlegl, J. Ofner, and G. Langs, "Unsupervised pre-training across image domains improves lung tissue classification," in *Medical Computer Vision: Algorithms for Big Data*. Springer, 2014, pp. 82–93.
- [21] J. Deng, W. Dong, R. Socher, L.-J. Li, K. Li, and L. Fei-Fei, "ImageNet: A large-scale hierarchical image database," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, Jun. 2009, pp. 248–255.
- [22] S. Ioffe and C. Szegedy, "Batch normalization: Accelerating deep network training by reducing internal covariate shift," 2015, *arXiv:1502.03167*. [Online]. Available: <https://arxiv.org/abs/1502.03167>
- [23] K. Simonyan and A. Zisserman, "Very deep convolutional networks for large-scale image recognition," 2014, *arXiv:1409.1556*. [Online]. Available: <https://arxiv.org/abs/1409.1556>
- [24] O. Russakovsky, J. Deng, H. Su, J. Krause, S. Satheesh, S. Ma, Z. Huang, A. Karpathy, A. Khosla, M. Bernstein, A. C. Berg, and L. Fei-Fei, "ImageNet large scale visual recognition challenge," *Int. J. Comput. Vis.*, vol. 115, no. 3, pp. 211–252, Apr. 2015.
- [25] A. Esteva, B. Kuprel, R. A. Novoa, J. Ko, S. M. Swetter, H. M. Blau, and S. Thrun, "Dermatologist-level classification of skin cancer with deep neural networks," *Nature*, vol. 542, no. 7639, pp. 115–118, Jan. 2017.
- [26] N. Tajbakhsh, J. Y. Shin, S. R. Gurudu, R. T. Hurst, C. B. Kendall, M. B. Gotway, and J. Liang, "Convolutional neural networks for medical image analysis: Full training or fine tuning?" *IEEE Trans. Med. Imag.*, vol. 35, no. 5, pp. 1299–1312, May 2016.
- [27] I. Giotis, N. Molders, S. Land, M. Biehl, M. F. Jonkman, and N. Petkov, "MED-NODE: A computer-assisted melanoma diagnosis system using non-dermoscopic images," *Expert Syst. Appl.*, vol. 42, no. 19, pp. 6578–6585, Nov. 2015.
- [28] L. Oakden-Rayner, "Reply to 'Man against machine: Diagnostic performance of a deep learning convolutional neural network for dermoscopic melanoma recognition in comparison to 58 dermatologists' by haenssle et al.," *Ann. Oncol.*, vol. 30, no. 5, pp. 1836–1842, May 2019.
- [29] U.-O. Dorj, K.-K. Lee, J.-Y. Choi, and M. Lee, "The skin cancer classification using deep convolutional neural network," *Multimedia Tools Appl.*, vol. 77, pp. 9909–9924, Feb. 2018.
- [30] S. S. Han, M. S. Kim, W. Lim, G. H. Park, I. Park, and S. E. Chang, "Classification of the clinical images for benign and malignant cutaneous tumors using a deep learning algorithm," *J. Investigative Dermatol.*, vol. 138, no. 7, pp. 1529–1538, Jul. 2018.
- [31] M. Anas, K. Gupta, and S. Ahmad, "Skin cancer classification using K-means clustering," *Int. J. Tech. Res. Appl.*, vol. 5, no. 1, pp. 62–65, 2017.
- [32] E. Almansour and M. A. Jaffar, "Classification of Dermoscopic skin cancer images using color and hybrid texture features," *Int. J. Comput. Sci. Netw. Secur.*, vol. 16, no. 4, pp. 135–139, 2016.
- [33] G. Capdehourat, A. Corez, A. Bazzano, R. Alonso, and P. Musé, "Toward a combined tool to assist dermatologists in melanoma detection from dermoscopic images of pigmented skin lesions," *Pattern Recognit. Lett.*, vol. 32, no. 16, pp. 2187–2196, Dec. 2011.
- [34] Y. Prabhu and M. Varma, "FastXML: A fast, accurate and stable tree-classifier for extreme multi-label learning," in *Proc. SIGKDD*, 2014, pp. 263–272.
- [35] H. O. Song, Y. Xiang, S. Jegelka, and S. Savarese, "Deep metric learning via lifted structured feature embedding," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jun. 2016, pp. 4004–4012.
- [36] K. Q. Weinberger, J. Blitzer, and L. K. Saul, "Distance metric learning for large margin nearest neighbor classification," in *Proc. NIPS*. Cambridge, MA, USA: MIT Press, 2006, pp. 1473–1480.
- [37] B. Yu, T. Liu, M. Gong, C. Ding, and D. Tao, "Correcting the triplet selection bias for triplet loss," in *Proc. ECCV*, 2018, pp. 71–87.
- [38] A. S. Razavian, H. Azizpour, J. Sullivan, and S. Carlsson, "CNN features off-the-shelf: An astounding baseline for recognition," 2014, *arXiv:1403.6382*. [Online]. Available: <https://arxiv.org/abs/1403.6382>
- [39] J. Yosinski, J. Clune, Y. Bengio, and H. Lipson, "How transferable are features in deep neural networks?" 2014, *arXiv:1411.1792*. [Online]. Available: <https://arxiv.org/abs/1411.1792>
- [40] K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jun. 2016, pp. 770–778.
- [41] C. Szegedy, S. Ioffe, V. Vanhoucke, and A. A. Alemi, "Inception-v, inception-resnet and the impact of residual connections on learning," in *Proc. AAAI*, vol. 4, 2017, p. 12.
- [42] D. R. Wilson and T. R. Martinez, "The general inefficiency of batch training for gradient descent learning," *Neural Netw.*, vol. 16, no. 10, pp. 1429–1451, Dec. 2003.
- [43] M. Chen, P. Zhou, D. Wu, L. Hu, M. M. Hassan, and A. Alamri, "AI-skin: Skin disease recognition based on self-learning and wide data collection through a closed-loop framework," *Inf. Fusion*, vol. 54, pp. 1–9, Feb. 2020.
- [44] D. Ruiz, V. Berenguer, A. Soriano, and B. Sánchez, "A decision support system for the diagnosis of melanoma: A comparative approach," *Expert Syst. Appl.*, vol. 38, no. 12, pp. 15217–15223, Nov. 2011.
- [45] Q. Abbas, M. E. Celebi, C. Serrano, I. Fondón García, and G. Ma, "Pattern classification of dermoscopy images: A perceptually uniform model," *Pattern Recognit.*, vol. 46, no. 1, pp. 86–97, Jan. 2013.
- [46] A. Blum, H. Luedtke, U. Ellwanger, R. Schwabe, G. Rassner, and C. Garbe, "Digital image analysis for diagnosis of cutaneous melanoma. Development of a highly effective computer algorithm based on analysis of 837 melanocytic lesions," *Brit. J. Dermatol.*, vol. 151, no. 5, pp. 1029–1038, Nov. 2004.
- [47] S. Koitka and C. M. Friedrich, "Traditional feature engineering and deep learning approaches at medical classification task of imageCLEF 2016," in *Proc. Work. Notes, CEUR Workshop (CLEF)*, vol. 1609, 2016, pp. 304–317.
- [48] A. Kumar, J. Kim, D. Lyndon, and D. Feng, "Subfigure and multilabel classification using a fine-tuned convolutional neural network," in *Proc. Work. Notes, CEUR Workshop (CLEF)*, vol. 1609, 2016, pp. 318–321.
- [49] R. Amelard, J. Glaister, A. Wong, and D. A. Clausi, "High-level intuitive features (HLIFs) for intuitive skin lesion description," *IEEE Trans. Biomed. Eng.*, vol. 62, no. 3, pp. 820–831, Mar. 2015.





and healthcare data, and image processing.

**BELAL AHMAD** received the B.Sc. degree (Hons.) from the Department of Statistics and Operation Research, Aligarh Muslim University (AMU), Aligarh, India, in 2009, and the MCA degree from the Department of Computer Science, AMU, in 2013. He is currently pursuing the Ph.D. degree with the School of Computer Science and Technology, Huazhong University of Science and Technology (HUST), Wuhan, China. His research interests include deep learning, machine learning



and healthcare. His research interests include deep learning, machine learning and sentiment analysis, and healthcare.

**MOHD USAMA** received the B.Sc. degree (Hons.) in statistics and the master's degree in computer science from Aligarh Muslim University (AMU), Aligarh, India. He is currently pursuing the Ph.D. degree with the School of Computer Science and Technology, Huazhong University of Science and Technology (HUST), Wuhan, China. His research interests include deep learning, machine learning and sentiment analysis, and healthcare.



**CHUEN-MIN HUANG** received the M.D. and B.A. degrees from National Taiwan University, Taiwan, and the Ph.D. degree in library and information science from the University of Pittsburgh, Pittsburgh, PA, USA. She was appointed as the Library Director from 1991 to 1993 and from 2002 to 2006, respectively. She is currently a Professor with the Department of Information Management, National Yunlin University of Science and Technology (NYUST), Taiwan. Her research interests focus on machine learning, image annotation, natural language processing, software engineering, and design pattern, and has published dozens of articles. She currently teaches courses, including system analysis, object-oriented software engineering, and seminar on deep learning. She has obtained four invention patents, including one US patent and three Taiwan patents. Her invention patent won two international patent competition gold medals in 2016 and 2017, respectively. In addition to research, she devotes herself to teaching and received two times' Excellent Teacher Award recognition, in 2007 and 2012, respectively. She serves many prominent journals' reviewer, including the *Information Processing and Management*, the *Transactions on Asian Language Information Processing*, the *Information Sciences*, the *Journal of Internet Technology*, and the *Journal of Software Engineering Studies*.



**KAI HWANG** (Life Fellow, IEEE) received the Ph.D. degree from the University of California, Berkeley, CA, USA, in 1972. In 1986, he has taught at Purdue University for 11 years. He is currently a Professor of electrical engineering and computer science with the University of Southern California (USC). He has served as the founding Editor-in-Chief of the *Journal of Parallel and Distributed Computing* from 1983 to 2011. He has published eight books and 250 scientific articles. According to Google Scholars, his work was cited over 15 000 times with an H-index of 52. His most cited book on Computer Architecture and Parallel Processing was cited more than 2300 times and his PowerTrust (IEEE-T-PDS, April 2007) article was cited over 540 times. He received Lifetime Achievement Award from the IEEE Cloudcom—2012 for his pioneering contributions in the field of computer architecture, parallel, distributed and cloud computing, and cyber security.

**M. SHAMIM HOSSAIN** (Senior Member, IEEE) received the Ph.D. degree in electrical and computer engineering from the University of Ottawa, Canada. He is currently a Professor with the Department of Software Engineering, College of Computer and Information Sciences, King Saud University, Riyadh, Saudi Arabia. He is also an Adjunct Professor with the School of Electrical Engineering and Computer Science, University of Ottawa. His research interests include cloud networking, smart environment (smart city, smart health), social media, the IoT, edge computing and multimedia for health care, deep learning approach to multimedia processing, and multimedia big data. He has authored and coauthored approximately 200 publications, including refereed journals, conference papers, books, and book chapters. He is a Senior Member of the ACM. He has served as a member of the organizing and technical committees of several international conferences and workshops. He has served as the co-chair, general chair, workshop chair, publication chair, and TPC for over 12 IEEE and ACM conferences and workshops. He is currently the Co-Chair of the 2nd IEEE ICME Workshop on Multimedia Services and Tools for Smart-Health (MUST-SH 2019). He is a recipient of a number of awards, including the Best Conference Paper Award and the 2016 ACM Transactions on Multimedia Computing, Communications and Applications (TOMM) Nicolas D. Georganas Best Paper Award and the Research in Excellence Award from the College of Computer and Information Sciences (CCIS), King Saud University (three times in a row). He is on the editorial board of the IEEE TRANSACTIONS ON MULTIMEDIA, IEEE NETWORK, the IEEE MULTIMEDIA, the IEEE WIRELESS COMMUNICATIONS, IEEE ACCESS, the *Journal of Network and Computer Applications* (Elsevier), and the *International Journal of Multimedia Tools and Applications* (Springer). He also presently serves as the Lead Guest Editor of IEEE NETWORK and IEEE ACCESS. Previously, he served as a Guest Editor of the *IEEE Communications Magazine*, the IEEE TRANSACTIONS ON INFORMATION TECHNOLOGY IN BIOMEDICINE (currently JBHI), the IEEE TRANSACTIONS ON CLOUD COMPUTING, the *International Journal of Multimedia Tools and Applications* (Springer), *Cluster Computing* (Springer), the *Future Generation Computer Systems* (Elsevier), the *Computers and Electrical Engineering* (Elsevier), *Sensors* (MDPI), and the *International Journal of Distributed Sensor Networks*.



**GHULAM MUHAMMAD** received the Ph.D. degree in electrical and computer engineering from the Toyohashi University and Technology, Japan, in 2006. He is currently a Professor with the Department of Computer Engineering, College of Computer and Information Sciences, King Saud University (KSU), Riyadh, Saudi Arabia. His research interests include image and speech processing, cloud and multimedia for healthcare, biometrics, and security. He has authored and coauthored more than 200 publications, including IEEE/ACM/Springer/Elsevier journals, and flagship conference papers. He holds a U.S. patent on audio processing. He received the Best Faculty Award of the Computer Engineering Department, KSU, from 2014 to 2015. He has supervised more than ten Ph.D. and Master Theses. He is involved in many research projects as a principal investigator and a co-principal investigator. He was a recipient of the Japan Society for Promotion and Science (JSPS) Fellowship from the Ministry of Education, Culture, Sports, Science and Technology, Japan.

...