



Fusion of U-Net and CNN model for segmentation and classification of skin lesion from dermoscopy images

Vatsala Anand^a, Sheifali Gupta^a, Deepika Koundal^{b,*}, Karamjeet Singh^c

^a Chitkara University Institute of Engineering and Technology, Chitkara University, Punjab, India

^b School of Computer Science, University of Petroleum & Energy Studies, Dehradun, Uttarakhand, India

^c Thapar University, Patiala, India



ARTICLE INFO

Keywords:

Skin diagnosis
Classification
Convolution Neural Network (CNN)
Segmentation
U-Net architecture
Optimizer
Fusion model
Skin cancer
Dermoscopy images

ABSTRACT

Skin is one of the most significant organs, which serves as a barrier to the outside surroundings of the human body. To improve mortality, skin disease detection at a prior stage is necessary else it may convert to skin cancer. But its diagnosis at the prior stage that may increase life expectancy is a great experiment since it has a similar look to skin diseases. To deal with biomedical images, a new innovative automated system is required that can quickly and precisely identify skin lesions. Deep Learning is attracting a lot of attention in the treatment of numerous disorders. A fusion model is proposed here with the integration of the U-Net and Convolution Neural Network model. For this, the U-Net model has been utilized to segment the diseases using skin images and the Convolution Neural Network model has been proposed for the multi-class classification of segmented images. The model is simulated and analyzed using the HAM10000 dataset having 10,015 dermoscopy images of seven different classifications of skin diseases. The proposed model has been analyzed using two optimizers i.e. Adam and Adadelta on 20 epochs and 32 batch sizes for the skin disease classification. The model has outperformed on Adadelta optimizer with an accuracy value of 97.96%.

1. Introduction

Humans' skin serves as a barrier between their bodies and the rest of their surroundings. It is the body's largest organ and performs main functions such as protection, regulation, and sensation (Deng and Xu, 2018). Regular skin check-up is necessary to detect skin disease at earlier stages. To avoid the growth of skin cancer, skin disease must be diagnosed as early as possible. Skin disease diagnosis can be performed with many procedures. The first procedure is dermatoscopy, which uses a dermatoscopy device to assess a mole's pigment and blood vessel layout without removing it. The most common skin tests include patch tests, skin biopsies, and cultures. The ABCDE rule is also used for the detection of skin disease proposed by Murphree (2017), Van Onselen (2011), White-Chu and Reddy (2011), Chang et al. (2018). ABCDE means Asymmetry, Border, Color, Diameter, and Evolution. It is difficult to detect skin disease in its earliest stages because of the resemblance between different skin diseases. When growths or lesions are segmented during dermoscopy, it is possible to classify the images more precisely. Due to uneven boundaries, inter and intra-class similarities or

dissimilarities, and several other reasons, nodule segmentation can be a challenging process. The images without segmentation include all the outside background, borders of a lesion, and skin texture. Due to this, unwanted features are extracted from background regions also. To deal with biomedical images, a new innovative automated system is required that can quickly and precisely identify skin lesions. Therefore, a fusion model has been presented with the integration of the U-Net model and CNN model. For this, the U-Net model has been proposed to segment the diseases from skin images, and the CNN model has been proposed for classifying the segmented images into seven different skin disease classes.

To avoid this problem, ROI is extracted from the original image using U-Net architecture so that only useful features can be obtained from the lesion part and classification accuracy can be enhanced.

2. Literature review

Researchers had worked using various Deep Learning (DL) techniques in the past to study skin diseases using dermoscopic images. Even

* Corresponding author.

E-mail addresses: vatsala.anand@chitkara.edu.in (V. Anand), sheifali.gupta@chitkara.edu.in (S. Gupta), dkoundal@ddn.upes.ac.in (D. Koundal), karamjeetsingh@thapar.edu.in (K. Singh).

though the dermoscopic imaging system has the potential to amplify lesions, due to the intricate structures of lesions, visual examination is highly challenging. This problem can be solved with the help of automatic skin disease segmentation and classification techniques. With the help of a literature review, it is concluded that skin disease detection can be performed using different segmentation and classification techniques sometimes with pre-processing or sometimes without pre-processing. With the help of dermatoscopy devices, skin lesions can be seen. Some authors are working only on a segmentation-based approach while other authors are working to classify skin lesions based on DL techniques. Very few authors are working on both segmentation and classification-based approaches. In this proposed work, a fusion model has been presented for the segmentation and classification of skin diseases using dermoscopy images.

[Al-Antari et al. \(2018\)](#) used DL models for skin disease classification. They used the classification challenge that includes 10,015 dermoscopy images. The evaluation results showed the multi-class classification accuracy values as 84.70 %, 84.60 %, 90.10 %, and 85.80 % with Inception-V3, DenseNet-201, ResNet-50, and InceptionResNet-V2 respectively. [Dutta et al. \(2021\)](#) proposed an automatic framework with the integration of augmentation of data, a convolutional network. Testing and data training is performed on dataset and has obtained values as 0.73, 0.76, and 0.74 for recall, precision, and F1 score. [Majtner et al. \(2016\)](#) presented a DL model with a combination of Local Binary Patterns and RSurf features. They obtained images from the ISIC dataset and the model is simulated on 1279 dermoscopic images of skin and obtained an accuracy of 82.6 %. [Hekler et al. \(2020\)](#) performed a binary classification test between melanoma and nevus, to observe the influence of label noise on the performance of CNNs. The achieved accuracy is 75.03 % for dermatological and 73.80 % for biopsy. [Huang et al. \(2021\)](#) performed binary classification and multi-class classification on Kaohsiung Chang Gung Memorial Hospital (KCGMH) and HAM10000 dataset. They used EfficientNet and DenseNet for the training and validation of the model. For binary classifications, the accuracy was 89.5 percent; for 7-class classification, it was 85.8 %, and for five-class classification in the KCGMH dataset, it was 72.1 percent.

[Khan et al. \(2020, a\)](#) proposed a model for skin cancer recognition with a combination of a deep learning model and a feature selection method. The contrast stretching approach is followed in the localization step. They used the DenseNet201 for the extraction of features. Training and testing were carried out on the ISIC2017 dataset and ISBI2016 dataset with an accuracy of 93.4 % and 94.5 % respectively. [Anand et al. \(2022a\)](#) performed classification using the VGG16 model. They had worked on the HAM10000 dataset. They obtained an accuracy of 89.09 % by changing the layers of the transfer learning model. Authors ([Anand et al. \(2022b\)](#), [Anand et al. \(2022c\)](#)) worked on transfer learning-based Xception and MobileNet models and obtained different values of accuracy using the HAM10000 dataset with seven different skin disease classes. [Khan et al. \(2021, b\)](#) reported a multi-class skin lesion segmentation and classification scheme consisting of the Deep Saliency Segmentation method, and a CNN consisting of ten layers. They utilized the ISBI 2016 dataset consisting of 1279 dermoscopic images, the ISBI 2017 dataset with 2750 images, the ISIC 2018 dataset with 3694 images, and PH2 datasets. For classification purposes, they used the HAM10000 dataset with 10,015 skin dermoscopy images. They obtained an accuracy of 95.38 %, 95.79 %, 92.69 %, and 98.70 %, on ISBI 2016, ISBI 2017, ISIC 2018, and PH2 datasets respectively. They obtained a classification accuracy of 90.67 %. [Al-Masni et al. \(2020\)](#) presented a combination of segmentation and classification. In the first step, segmentation is performed using deep learning Full Resolution Convolutional Network (FrCN). In the next step, classification is performed on segmented images using transfer learning models. The model was evaluated using ISIC 2016, 2017, and 2018. [Polat and Onur Koc \(2020\)](#) introduced two methods for the classification of skin disease. In the first method, a CNN is presented and in the second method, a convolutional method is combined with a one-versus-all scheme. The HAM10000

dataset with 10,015 total dermoscopy images is used for training and testing purposes. They attained 77 % accuracy in the classification of seven disease classes. Whereas, 92.9 % accuracy is obtained in the one-versus-all approach. [Salian et al. \(2020\)](#) classified melanoma from skin lesion images. They had worked on PH2 and HAM10000 datasets with and without augmentation. They had worked on VGG16 and MobileNet. The authors also designed a custom model for comparison of performance with pre-trained models. The authors obtained an accuracy of 80.07 % on the VGG16 model, 81.52 % on the MobileNet model, and 83.15 % on the custom model. [Yu et al. \(2016\)](#) gave a scheme for melanoma recognition using CNN and performed segmentation as well as classification. Firstly, the authors worked on residual learning and then they constructed a convolutional residual network for the formation of a two-stage framework. The model was evaluated on the ISBI 2016 dataset. The main drawback of segmentation algorithms is the lack of context. To segment an object out of the complex environment, it is important to understand the context in which the objects find themselves. In current deep-learning image segmentation methods, the feature map is converted into a vector and then the image is reconstructed from this vector. It may lose the context features during conversion from feature map to vector. Moreover, it is an immensely huge and tough task to convert a vector into an image and vice versa.

The proposed model has overcome the disadvantages of deep learning segmentation methods by using U-Net architecture for the segmentation of skin diseases. As U-Net also works on the same concept i.e. expansion and contraction of image feature map for which it uses an encoder and decoder. The advantage of U-Net is that it captures both the features of the context as well as the localization. U-Net uses the same feature maps that are used for contraction to expand a vector to a segmented image. This helps in preserving the structural integrity of the image which would reduce distortion enormously. The major offerings of the proposed work are as follows:

1. U-Net and CNN models have been integrated to propose a fusion model segmenting and classifying skin lesions.
2. In the first phase of the fusion model, the U-Net model is employed to extract the ROI from skin disease images. This architecture focuses and differentiates borders by categorizing every pixel; hence, the input and output layers of U-Net share the same size.
3. In the second phase, a CNN model has been proposed for the multi-class classification of segmented images obtained from the first phase of the fusion model. The CNN model includes three Convolution Blocks (CB) consisting of five convolution layers, three Max Pool (MP) layers, three dense layers, and one fully connected layer.
4. Analysis of the proposed model has been done with two Optimizers (OP) i.e. Adam and Adadelta on 20 Epochs (EP) and 32 Batch Size (BS) for the skin disease classification.
5. Various performance metrics, including precision, sensitivity, specificity, and accuracy have been employed to verify the proposed model.
6. Comparative analysis has been done with the state-of-art techniques which have shown that the proposed system has outperformed all other techniques.

The rest of the paper is designed as [Section 3](#) comprises the proposed fusion model, results and discussion is given in [Section 4](#), and the conclusion is given in [Section 5](#).

3. Proposed fusion model

[Fig. 1](#) illustrates the architecture of the fusion model consists of two phases. The first phase is the segmentation phase based on the U-Net model and the second phase is the multi-class classification phase based on the proposed CNN model. For this, the pre-processed images are provided as input to the U-Net model for mask generation. After that, the generated mask is multiplied by a pre-processed image for the

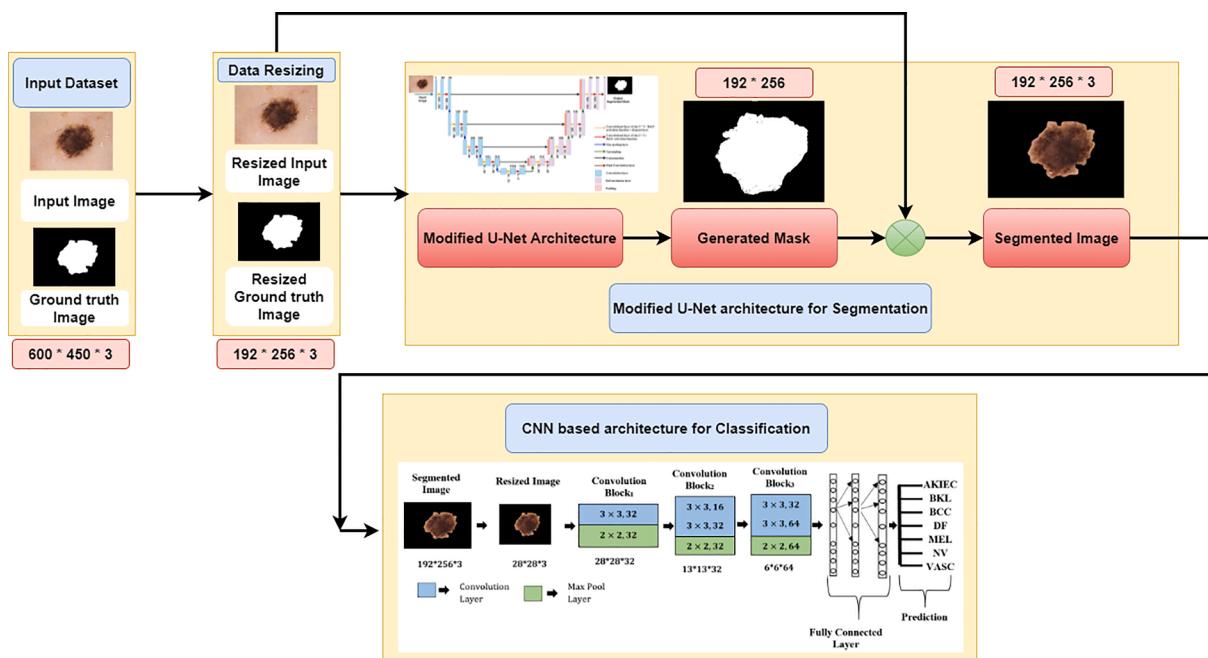


Fig. 1. Multi-class skin lesion classification on segmented images using U-net architecture.

generation of the segmented image. Segmented images are provided as input to the proposed CNN model for further classification. An analysis is done on the Google Colab platform using Python.

3.1. Input dataset and data resizing

Dermoscopy images of skin can be taken from different sources such as PH2, HAM10000, Interactive Atlas of Dermoscopy, Dermofit Image Library, and ISIC. In this paper, a HAM10000 dataset has been used that consists of a total of 10,015 dermoscopy images of skin Tschandl et al. (2018). This dataset includes seven skin disease classes named Melanoma (MEL), Vascular Lesions (VASC), Benign Keratosis - Lesions (BKL), Dermatofibroma (DF), Melanocytic Nevi (NV), Basal Cell Carcinoma (BCC), and Actinic Keratoses (AKIEC). The images are RGB and all the images are in jpg format. The complete dataset is split into training and testing. For the training purpose, 80 % of the data is used and for testing, 20 % of the data is used. All the Original (OG) and Ground Truth (GT) images have a size of 600 * 450 * 3 pixels. The size of both the GT and the OG images are resized to 192 * 256 for applying the segmentation technique using U-Net architecture for extraction of ROI. Fig. 2 shows the OG skin disease images whereas Fig. 3 illustrates the GT images/masks of respective OG images.

3.2. Skin disease segmentation using U-Net architecture

Fig. 4 displays an overview of the first phase in which U-Net architecture by Anand et al. (2022d) is used for the segmentation of skin disease. A new version of the U-Net architecture has been presented to

deal with biomedical images to identify the affected area. In the U-Net, the input image size is kept as 192 * 256 * 3. This U-Net architecture is our preliminary work that is already published, for details refer this Anand et al. (2022d).

The U-Net architecture is modified in terms of feature map size. Original U-Net architecture includes a feature map of square size starting from 572 * 572 in the first layer and 284 * 284 in the second layer. Again it is downsized to 140 * 140 in the third layer. Now, in the next step, it is downsized to 68 * 68 in the fourth layer and finally, it goes to 32 * 32 in the last layer. Now, the feature map size increases in the expansion path with 52 * 52 in the first layer from the bottom. It is upsized to 100 * 100 in the second layer and 196 * 196 in the third layer. Finally, the feature map size goes to 388 * 388 in the topmost layer.

Modified U-Net architecture includes feature maps rectangular in size starting from 192 * 256 in the first layer and 96 * 128 in the second layer. Again it is downsized to 48 * 64 in the third layer. Now, in the next step, it is downsized to 24 * 32 in the fourth layer and finally, it goes to 6 * 8 in the last layer. Now, the feature map size increases in the expansion path with 24 * 32 in the first layer from the bottom. It is upsized to 48 * 64 in the second layer and 96 * 128 in the third layer. Finally, the feature map size goes to 192 * 256 in the topmost layer.

Here, the resized images are provided as input to the U-Net architecture from which a mask of size 192 * 256 is generated. After this, the generated mask is multiplied by the resized image of size 192 * 256 * 3 to obtain the segmented image with size 192 * 256 * 3. For better classification accuracy, images with segmented ROI are used for classification.

Fig. 5 shows the U-Net architecture that consists of two pathways i.e.

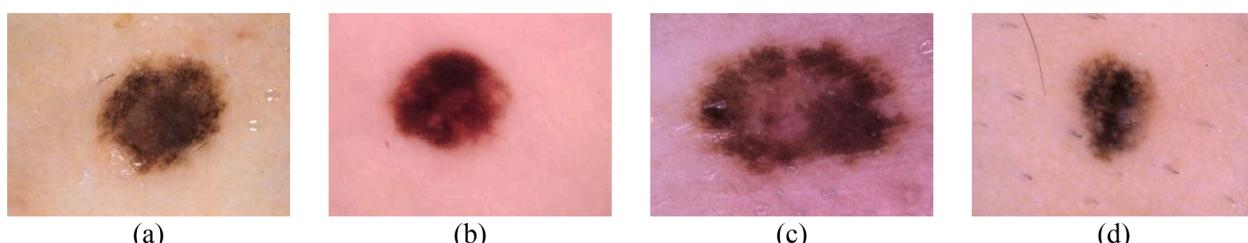


Fig. 2. Skin Disease OG Images: (a) Image 1; (b) Image 2; (c) Image 3; (d) Image 4.



Fig. 3. GT Masks for respective Skin Disease OG Images: (a) Image 1; (b) Image 2; (c) Image 3; (d) Image 4.

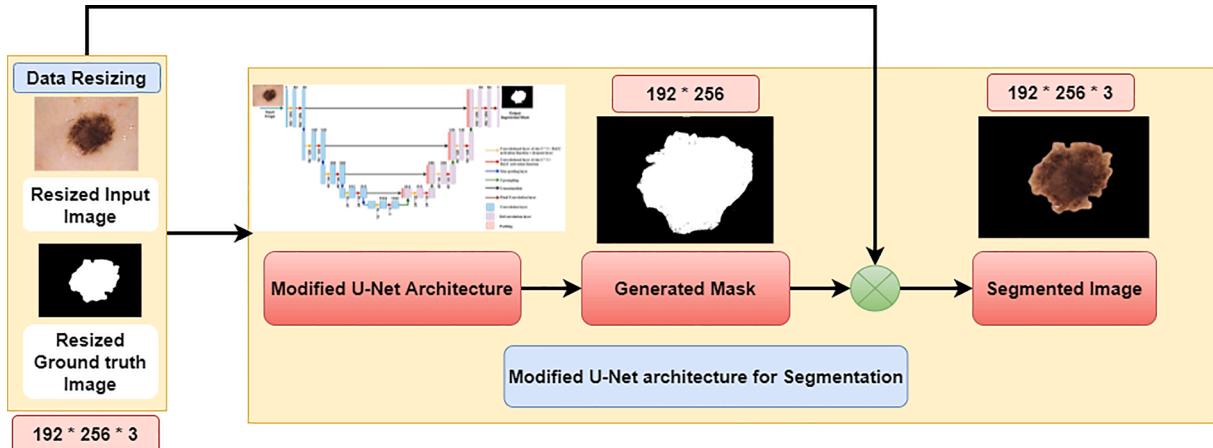


Fig. 4. Skin disease segmentation using U-Net architecture.

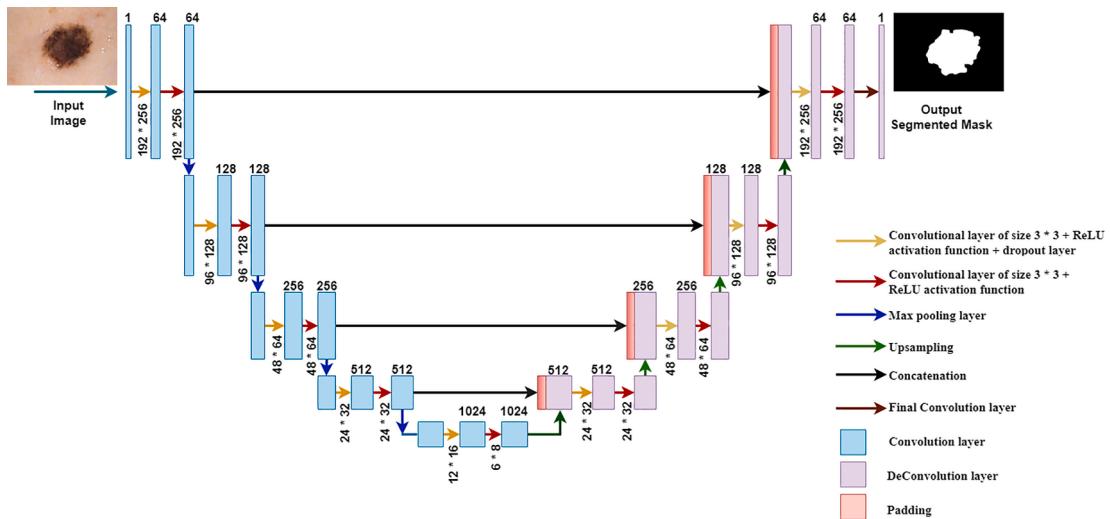


Fig. 5. U-Net Architecture Anand et al. (2022, d).

the contraction and expansion paths. An encoder is the contraction path, and the decoder is the symmetric expansion path. By classifying every pixel, the U-Net architecture localizes and identifies borders; as a result, the input and output sizes of both pathways are the same. Convolution and max pooling layer are used in the encoder for contraction whereas the transposed convolution layer and simple convolution layer are used in the decoder for expansion.

The image context is captured by the encoder, and exact localization is achieved by the decoder using transposed convolutions. Input image with image size $192 \times 256 \times 3$ is provided to U-Net architecture and mask is generated with size 192×256 . After that, the generated mask is multiplied by the resized image taken from the pre-processed dataset. In the last step, a segmented image is generated. Now, this segmented

image is taken as input in the proposed CNN model.

Fig. 6(a) displays the GT mask, and Fig. 6(b) displays the resized image that is given as input to the U-Net architecture. From this U-Net, a predicted mask is generated which is shown in Fig. 6(c). Now this predicted mask is multiplied with the resized image to obtain the segmented image that is shown in Fig. 6(d). Fig. 6(e) shows the cropped image of the size 28×28 .

3.3. Proposed CNN model for classification of segmented images

In the second phase of the fusion model, a CNN-based architecture has been proposed to classify the multi-class skin diseases using segmented images obtained from the first phase. The segmented image

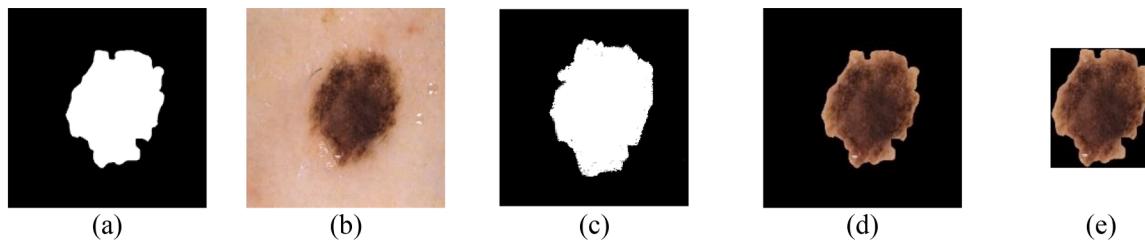


Fig. 6. Segmented image on Adam OP using BS of 8 and 75 EP (a) GT mask of the OG image, (b) Resized Image, (c) Predicted mask of the segmented image, (d) Segmented Image, (e) Cropped Image.

from the last step is provided as input to the proposed CNN model for better classification accuracy. The problem with images without segmentation is that the image includes all the outside backgrounds, borders of the lesion, and skin texture. Due to this, unwanted features are extracted from background regions also. As a solution, the Region of Interest (ROI) can be segmented from the original image so that only important features from the lesion section can be recovered.

Fig. 7 displays the organization of the proposed CNN model having different CB consisting of convolution layers and max pooling layers. In this work, no padding is used and the default stride value is used. The default stride is (1,1), meaning that the filter will move one pixel right for each horizontal movement of the filter and one pixel down for each vertical movement of the filter. The output image obtained from the first phase is of size $192 * 256 * 3$ which is resized to $28 * 28 * 3$. The resized image is provided as input to the first CB of the proposed CNN model. The first CB includes one convolution layer and one MP layer. The first convolution layer includes 32 filters of size $(3 * 3)$ and the MP layer of size $(2 * 2)$. In the second CB, there are two convolution layers and one MP layer. The filters of both the convolution layers are of size $(3 * 3)$ with 16 and 32 filters respectively and the MP layer with size $(2 * 2)$. The third CB again consists of two convolution layers and one MP layer. The filters of the third convolution layers are of $(3 * 3)$ size with 16 and 32 filters respectively and the MP layer with size $(2 * 2)$. After the third CB, the flattening layer flattens all of the features by converting feature space into a single feature vector. In the last step, three dense layers are applied to classify feature vectors into seven skin disease classes.

Table 1 lists the parameters of the fusion model, consisting of various sizes of input, output, filter, and parameters of all the layers.

4. Results and discussion

The proposed model is simulated using Adam Kingma and Ba (2014)

and Adadelta Zeiler (2012) optimizer with batch size values of 32 and 20 epochs. The author has calculated results firstly using batch sizes 8, 16, and 32. The results are best on batch size 32. Therefore, the results are shown using batch size 32. In this paper, epoch value 20 is considered for the simulation of results. The number of epochs can be increased. If a lot of data is present in the dataset, then only the increase of epochs will help in increasing accuracy. The model will reach a threshold point after which the value of accuracy cannot be improved. The value of epochs and batch size is kept fixed and taken as a reference from the author Anand et al. (2022e) because Anand et al. (2022e) used an almost similar CNN network for the classification of original skin disease images. Although, in this proposed work, the CNN model has been implemented on the segmented images. Table 2 shows the filtered images of individual blocks after every convolution layer and MP layer used in the CNN model. As all the blocks contain a different number of filtered images but only a single filtered image has been listed in this table.

4.1. Model accuracy (MA) and model loss (ML) analysis

Fig. 8 illustrates the MA and loss curves on Adam and Adadelta optimizers. From this figure, it has been analyzed that with an increase in the number of epochs, the value of accuracy increases although the value of ML decreases. Fig. 8 (a) and (b) display the MA and ML respectively for the Adam optimizer. It is analyzed from the curve that the initial value of the loss is 1.2 which is decreasing with the rise in the EP. At the 20th epoch, the value of training loss is reduced to 0.1 whereas the value of validation loss is reduced to 0.2. Fig. 8 (c) and (d) shows the model accuracy and loss respectively for the Adadelta optimizer. It has been noticed from the results that the value of training accuracy is above 96 % and training loss is reduced to 0.01 on the 20th epoch value.

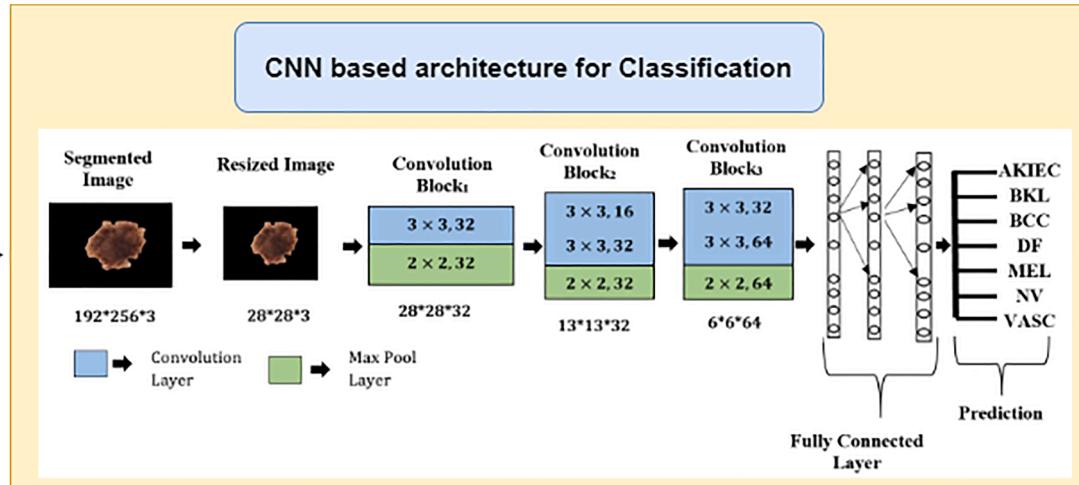


Fig. 7. Structure of proposed CNN model.

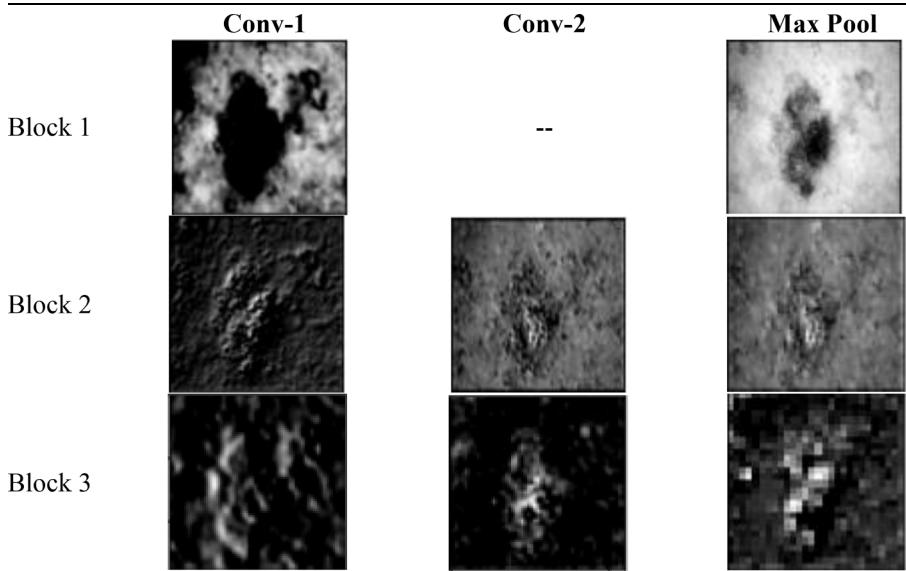
Table 1

Parameters of the Fusion Model.

S. No.	Layers	Size of Input Image	Size of Filter	No. of Filter	Activation function	Output	Parameters
1	Input Image	28*28*3	—	—	—	—	—
2	Convolutional	28*28*3	3*3	32	ReLU	28*28*32	896
3	MP	28*28*32	Poolszie (2*2)	—	—	14*14*32	0
4	Convolutional	14*14*32	3*3	16	ReLU	14*14*16	4624
5	Convolutional	14*14*16	3*3	32	ReLU	12*12*32	4640
6	MP	12*12*32	Poolszie (2*2)	—	—	6*6*32	0
7	Convolutional	6*6*32	3*3	32	ReLU	6*6*32	9248
8	Convolutional	6*6*32	3*3	64	ReLU	4*4*64	18,496
9	MP	4*4*64	Poolszie (2*2)	—	—	2*2*64	0
10	Flatten	2*2*64	—	—	—	2*2*64	0
11	Dense	2304	64	—	ReLU	64	16,448
12	Dense	16,448	32	—	ReLU	32	2080
13	Dense	2080	7	—	Softmax	7	231

Table 2

Filtered Images after convolution and MP layer.



It can be analyzed from the results that the proposed CNN model is better on the Adadelta optimizer as compared to the Adam optimizer in terms of MA and ML. Training accuracy is higher and the value of the loss is less in the case of the Adadelta optimizer. The same applies to validation accuracy and loss. Moreover, the next part assesses the Adam and Adadelta optimizer's performance by examining the parameters of the confusion matrix.

4.2. Confusion matrix (CM) parameters analysis

Fig. 9 displays the CM of the proposed model on the Adam and Adadelta optimizer. The matrix offers a clear vision of true and predicted labels. The class name is used to label each column and its matching row. Akiec, bcc, bkl, df, vasc, and mel are represented in this study by classes 0, 1, 2, 3, 4, 5, and 6. In a matrix, the number of images classed by a given model can be determined by the diagonal values of the matrices. Fig. 9 (a) and Fig. 9 (b) display the confusion matrix of the proposed model on the Adam optimizer and Adadelta optimizer, respectively.

Fig. 10 illustrates the results of the proposed CNN model on the Adam and Adadelta optimizer. The results are analyzed for all seven skin disease classes.

Fig. 10 (a) and (b) show the results of precision, sensitivity, and specificity with the Adam and Adadelta optimizer. In the case of precision, the results on AKIEC, BKL, DF, NV, and VASC disease classes with values as 81 %, 79 %, 88.88, 97.8 %, and 91.66 % respectively with

Adadelta optimizer whereas, in the case of Adam optimizer the values are 85 %, 83 %, 80 % and 80 % on MEL, BCC, VASC, and NV disease class respectively. Fig. 10 (c) shows the average values of confusion matrix parameters for the Adam and Adadelta optimizer. The value of accuracy in the case of the Adam optimizer is 96 % and 97.96 % in the case of the Adadelta optimizer.

Similarly, other confusion matrix parameters are also compared from which, it can be concluded that the Adadelta optimizer is outperforming better in terms of all CM parameters. Hence, the Adadelta optimizer is preferred over the Adam optimizer for the proposed model in terms of MA, ML, and all CM parameters.

4.3. Visual analysis of classification results using Adadelta optimizer

From the confusion matrix parameter analysis, the results are obtained best on the Adadelta optimizer. Therefore, Fig. 11 and Fig. 12 show classification and misclassification results on the Adadelta optimizer. Fig. 11 shows the true and predicted class of some skin disease images which are correctly classified.

Fig. 12 shows how the proposed approach misclassifies some of the images. This happens as a result of seven distinct skin diseases sharing the same characteristics.

Classification using Proposed CNN with Adam optimizer

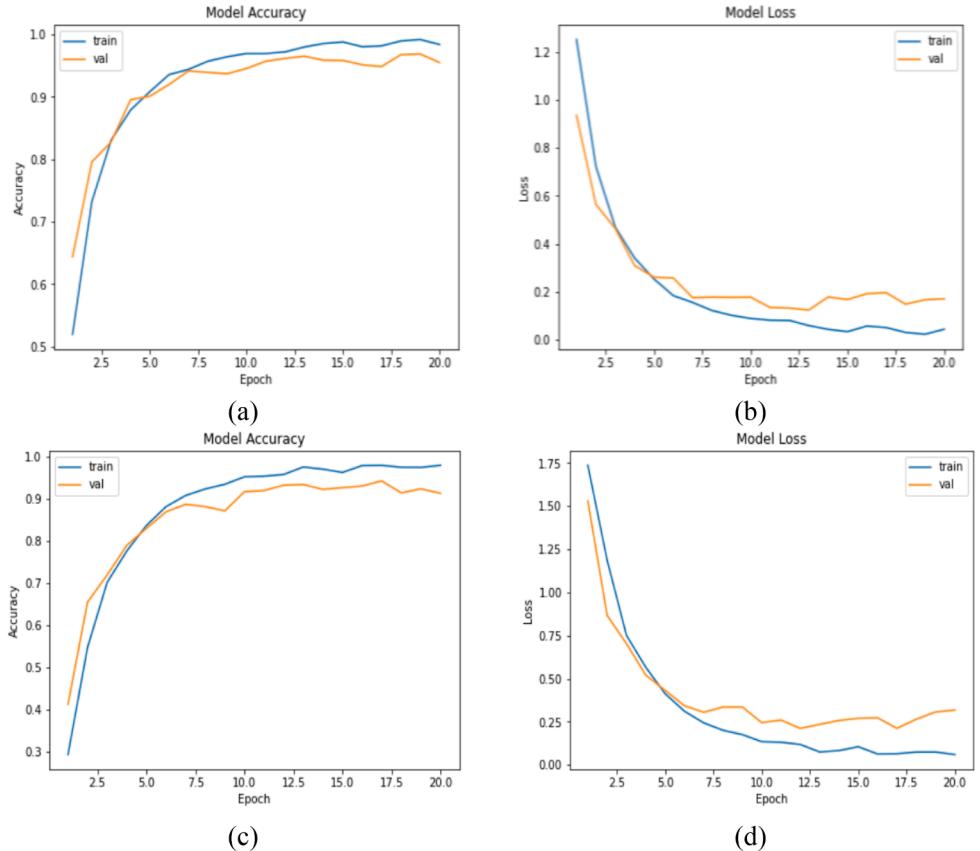


Fig. 8. MA and ML curves of Proposed Model: (a) (b) Classification using Proposed CNN with Adam optimizer and (c) (d) Classification using Proposed CNN with Adadelta optimizer.

Classification using Proposed CNN with Adadelta optimizer

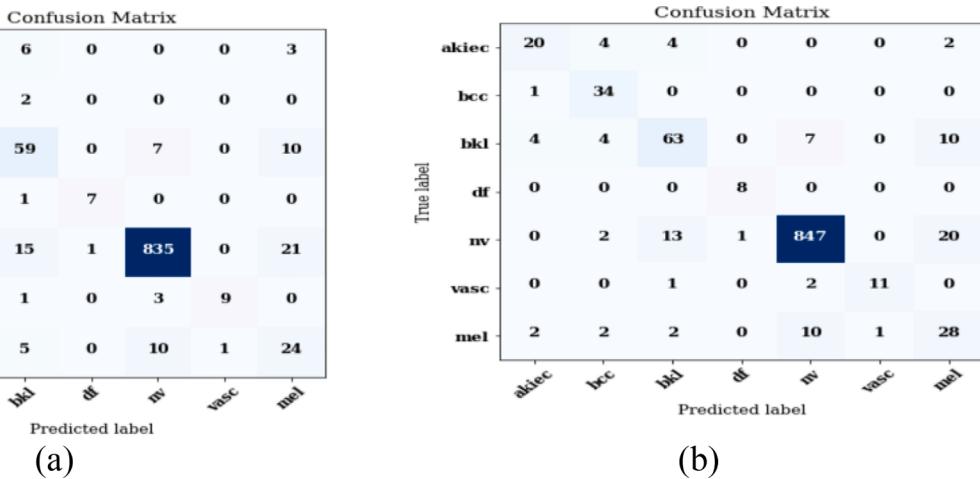


Fig. 9. Confusion Matrix (a) Classification using Proposed CNN with Adam optimizer, (b) Classification using Proposed CNN with Adadelta optimizer.

4.4. Comparative analysis with the state-of-art techniques

Table 3 illustrates that the suggested framework outperforms the current state-of-art techniques in terms of overall accuracy. Different datasets (ISIC-2016, 2017, ISBI-2017, and HAM10000) have been used. Some authors are working only on a segmentation-based approach while other authors are working on the classification of skin lesions based on deep-learning techniques. Very few authors are functioning on both segmentation and classification-based approaches. In this work, a fusion model has been presented that is based on the segmentation and classification of skin disease using dermoscopy images.

5. Conclusion and future scope

Nowadays the most frequently diagnosed disease in human beings is a skin disease. Although, the disease can be treated well if it is diagnosed at an early stage. In this work, a fusion model is presented with the combination of the U-Net and CNN model. For this, the U-Net model has been employed to segment the original images, and the CNN model has been presented to classify the segmented images into seven different skin disease classes. The models are simulated and analyzed using the HAM10000 dataset having 10,015 dermoscopy images. Simulation of the model is carried out by taking into account several key parameters,

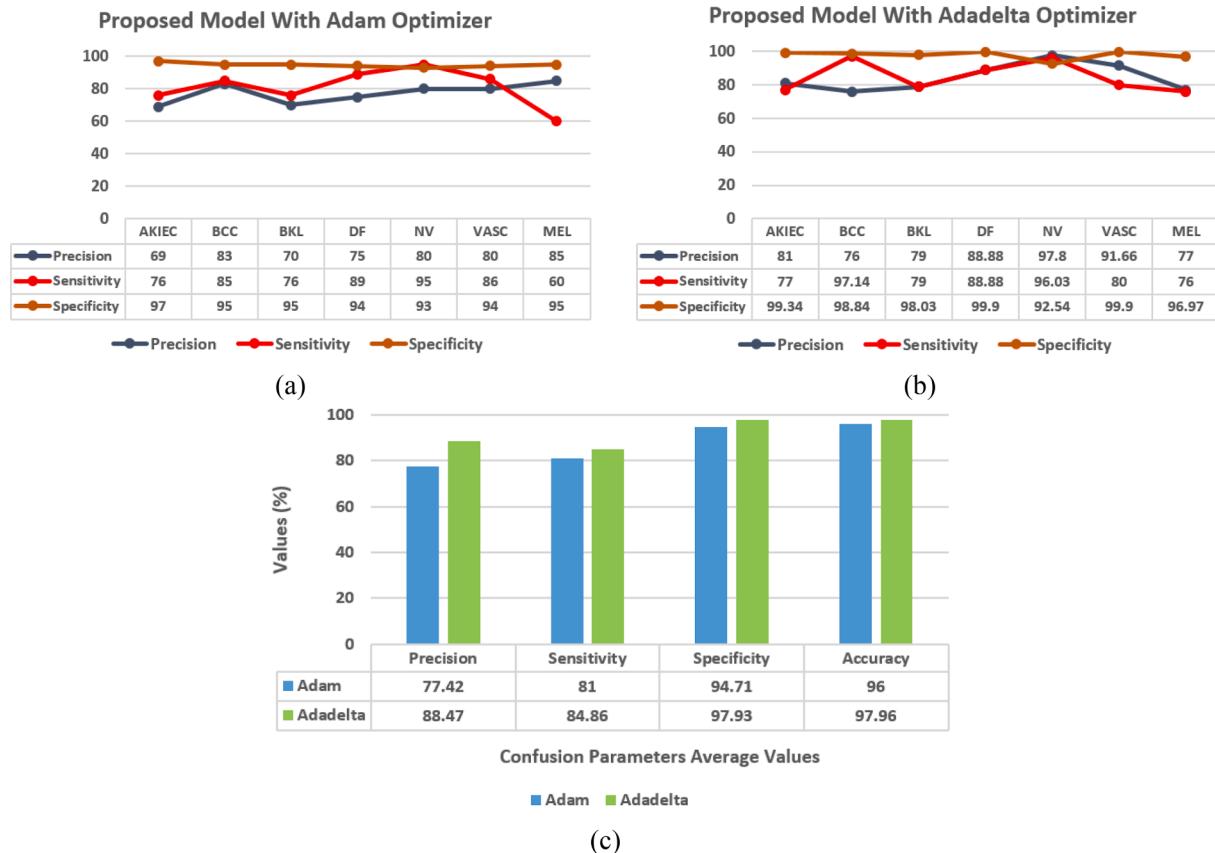


Fig. 10. CM Parameters Analysis (a) Proposed Model Results using Adam Optimizer, (b) Proposed Model Results using Adadelta Optimizer, (c) Confusion Matrix Parameters Average Values.

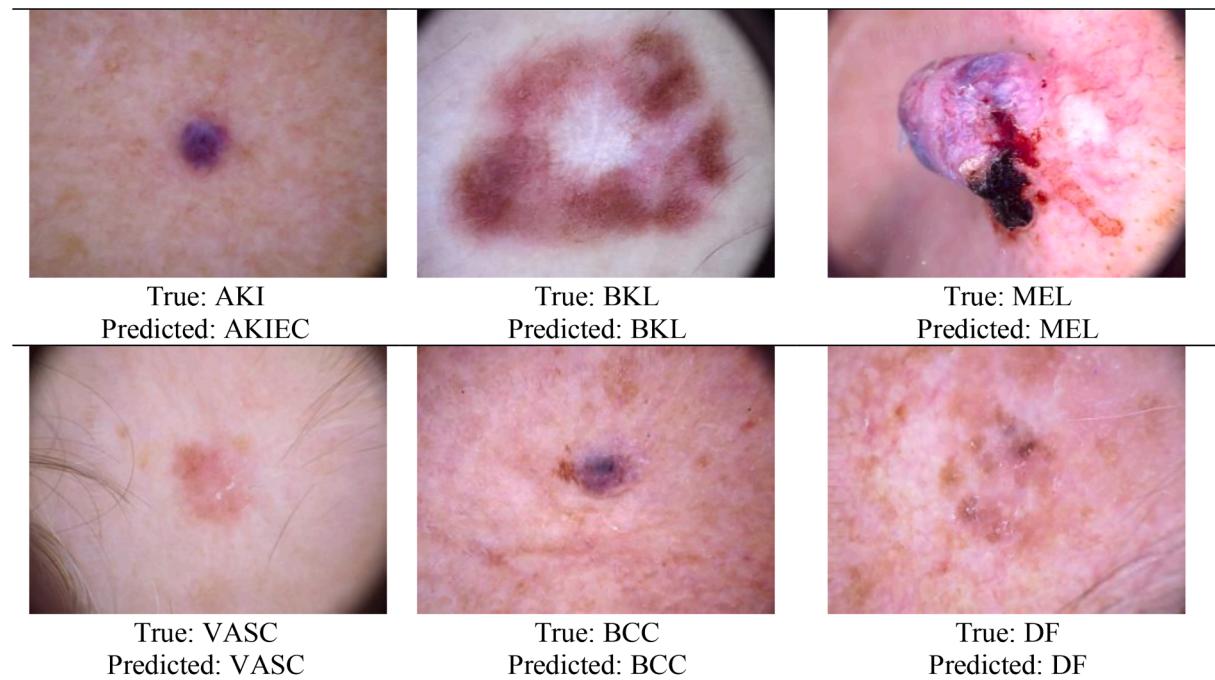


Fig. 11. Classification Results of Proposed Fusion Model.

including EP, BS, and the various types of OP. The results are best on the Adadelta optimizer as compared to the Adam optimizer. The proposed fusion model performs better as compared to state-of-art approaches

because classification is performed on segmented images. The images without segmentation include all the outside background, borders of a lesion, and skin texture. Due to this, unwanted features are extracted

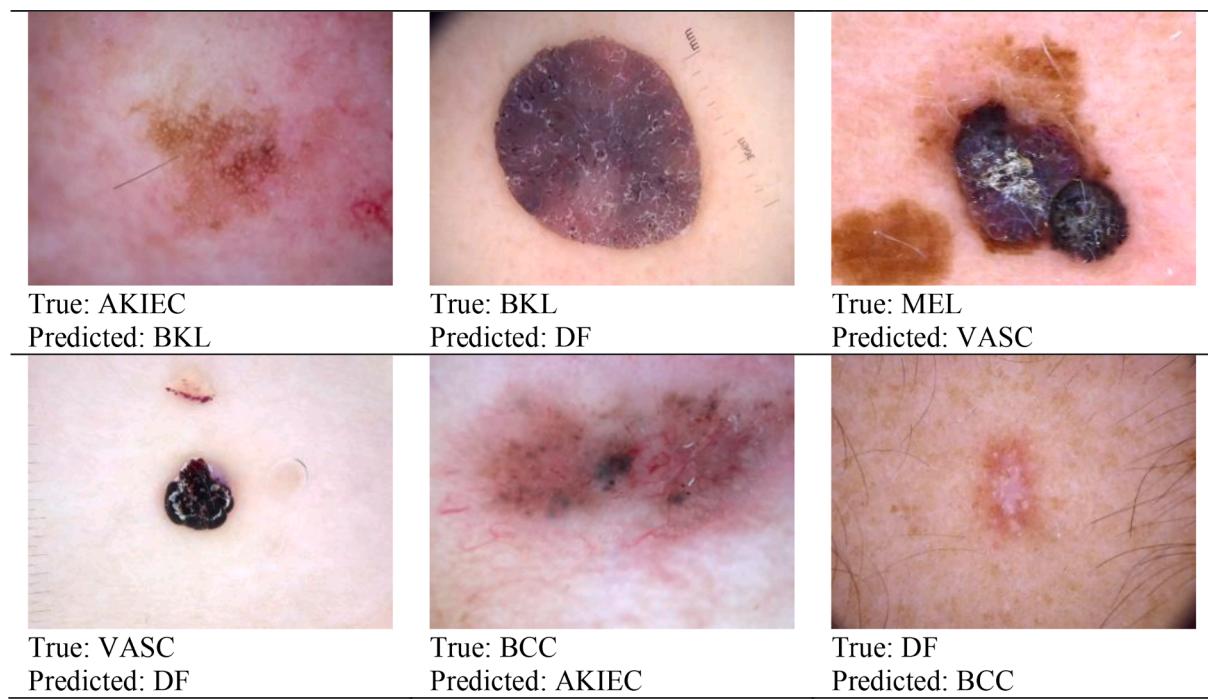


Fig. 12. Misclassification results using the Proposed Fusion Model.

Table 3
Comparative analysis of the fusion Model with State-of-Art Techniques.

Ref / Year	Technique Used	Dataset Used	Segmentation	Classification	Accuracy (%)
Majtner et al. (2016)	CNN with RSurf features and Local Binary Patterns	ISIC-2016 / 900	✗	✓	82.6
Zhang et al. (2019)	(ARL-CNN) model	ISIC-2017 / -	✗	✓	87.4
Antari et al. (2011)	DenseNet-201	HAM10000	✗	✓	84.60
	Inception-V3				84.70
	ResNet-V2				85.80
Masni et al. (2020)	Inception-ResNet-v2	HAM10000	✗	✓	87.74
Huang et al. (2020)	EfficientNet-B4	HAM10000	✗	✓	85.8
Polat et al. (2020)	CNN (Conv layers and max pool layers)	HAM10000	✗	✓	77.00
Hekler et al. (2020)	ResNet50	HAM10000	✗	✓	75.03
Salian et al. (2020)	VGG16	HAM10000	✗	✓	80.07
	MobileNet				81.52
	Custom-Model				83.15
Yu et al. (2016)	Fully Convolutional Residual Network (FCRN)-38, 50,101 layers	ISIC-2016	✓	✓	94.9
Khan et al. (2020)	Newton-Raphson (IcNR) based feature selection method	ISIC-2016 + ISBI-2017 / 4029	✓	✓	93.4
Khan et al. (2021)	Classification	HAM10000	✓	✓	90.7
Proposed Fusion Model	Classification	HAM10000	✓	✓	97.96

from background regions also. To avoid this problem, ROI is extracted from the original image using U-Net architecture so that only useful features can be obtained from the lesion part and classification accuracy can be enhanced. The proposed fusion model can be assisted by dermatologists as a second opinion tool for the diagnosis and treatment of various skin lesions. The proposed approach is validated for the skin. This approach can be generalized by using different biomedical images. Also, by using a combination of different epoch values, various optimizers, and different batch size values, the efficiency of the proposed model can be improved for different applications.

For further study in the area of skin disease detection, images can be enhanced in the pre-processing stage to improve the results. Furthermore, the ensemble approach can also be used for skin disease classification.

CRediT authorship contribution statement

Vatsala Anand: Conceptualization, Methodology, Validation, Formal analysis, Data curation. **Sheifali Gupta:** Conceptualization,

Validation, Writing – review & editing, Supervision. **Deepika Koundal:** Writing – review & editing, Supervision. **Karamjeet Singh:** Formal analysis, Data curation, Writing – review & editing, Supervision.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

Data will be made available on request.

References

- Al-Antari, M. A., Rivera, P., Al-Masni, M. A., Valarezo, E., Gi, G., Kim, T.-Y., et al. (2018). An Automatic Recognition of Multi-class Skin Lesions via Deep Learning Convolutional Neural Networks. In *In Proceedings of the ISIC 2018 skin image analysis workshop and challenge* (pp. 1–4).

- Anand, V., Gupta, S., Altameem, A., Nayak, S. R., Poonia, R. C., & Saudagar, A. K. J. (2022a). a). An Enhanced Transfer Learning Based Classification for Diagnosis of Skin Cancer. *Diagnostics*, 12(7), 1628. <https://doi.org/10.3390/diagnostics12071628>
- Anand, V., Gupta, S., Koundal, D., Nayak, S. R., Nayak, J., & Vimal, S. (2022b). b). Multi-class Skin Disease Classification Using Transfer Learning Model. *International Journal on Artificial Intelligence Tools*, 31(02), 2250029. <https://doi.org/10.1142/S0218213022500294>
- Anand, V., Gupta, S., & Koundal, D. (2022c). c). Detection and Classification of Skin Disease Using Modified Mobilenet Architecture. *ECS Transactions*, 107(1), 5059. <https://doi.org/10.1149/10701.5059ecst>
- Anand, V., Gupta, S., Koundal, D., Nayak, S. R., Barsocchi, P., & Bhoi, A. K. (2022d). d). Modified U-NET Architecture for Segmentation of Skin Lesion. *Sensors*, 22(3), 867. <https://doi.org/10.3390/s22030867>
- Anand, V., Gupta, S., Koundal, D., Mahajan, S., Pandit, A.K. and Zagaria, A. (2022,e). Deep learning based automated diagnosis of skin diseases using dermoscopy. *CMC-Computer Materials & Continua*, 71(2), 3145-3160. 10.32604/cmc.2022.022788.
- Al-Masni, M. A., Kim, D. H., & Kim, T. S. (2020). Multiple skin lesions diagnostics via integrated deep convolutional networks for segmentation and classification. *Computer methods and programs in biomedicine*, 190, Article 105351. <https://doi.org/10.1016/j.cmpb.2020.105351>
- Chang, A. L. S., Chen, S. C., et al. (2018). A daily skincare regimen with a unique ceramide and filaggrin formulation rapidly improves chronic xerosis, pruritus, and quality of life in older adults. *Geriatric Nursing*, 24–28. <https://doi.org/10.1016/j.gerinurse.2017.05.002>
- Dutta, A., Kamrul Hasan, M., & Ahmad, M. (2021). Skin Lesion Classification Using Convolutional Neural Network for Melanoma Recognition. In *Proceedings of International Joint Conference on Advances in Computational Intelligence*, 55–66. https://doi.org/10.1007/978-981-16-0586-4_5
- Deng, L., & Xu, S. (2018). Adaptation of human skin color in various populations. *Hereditas*, 155(1), 1–12. <https://doi.org/10.1186/s41065-017-0036-2>
- D.P. Kingma, J. Ba, 2014. Adam: A method for stochastic optimization, *arXiv preprint arXiv:1412.6980*. <https://doi.org/10.48550/arXiv.1412.6980>.
- Hekler, A., Kather, J. N., Krieghoff-Henning, E., Utikal, J. S., Meier, et al. (2020). Effects of Label Noise on Deep Learning-Based Skin Cancer Classification. *Frontiers in Medicine*, 7. <https://doi.org/10.3389/fmed.2020.00177>
- Huang, H. W., Hsu, B. W. Y., Lee, C. H., & Tseng, V. S. (2021). Development of a light-weight deep learning model for cloud applications and remote diagnosis of skin cancers. *Journal of Dermatology*, 48(3), 310–316. <https://doi.org/10.1111/1346-8138.15683>
- Khan, M. A., Sharif, M., Akram, T., Bukhari, S. A. C., & Nayak, R. S. (2020). a). Developed Newton-Raphson based deep features selection framework for skin lesion recognition. *Pattern Recognition Letters*, 129, 293–303. <https://doi.org/10.1016/j.patrec.2019.11.034>
- Khan, M. A., Sharif, M., Akram, T., Damaševičius, R., & Maskeliūnas, R. (2021). b). Skin lesion segmentation and multiclass classification using deep learning features and improved moth flame optimization. *Diagnostics*, 11(5), 811. <https://doi.org/10.3390/diagnostics11050811>
- Murphree, R. W. (2017). Impairments in skin integrity. *Nursing Clinics*, 52(3), 405–417. <https://doi.org/10.1016/j.cnur.2017.04.008>
- Majtner, T., Yıldırım-Yayılgan, S. and Hardeberg, J.Y. (2016, December). Combining deep learning and hand-crafted features for skin lesion classification. In 2016 Sixth International Conference on Image Processing Theory, Tools and Applications (IPTA), 1–6. IEEE. 10.1109/IPTA.2016.7821017.
- Polat, K., & Onur Koc, K. (2020). Detection of Skin Diseases from Dermoscopy Image Using the combination of Convolutional Neural Network and One-versus-All. *Journal of Artificial Intelligence and Systems*, 2(1), 80–97. <https://doi.org/10.33969/AIS.2020.021006>
- Salian, A.C., Vaze, S., Singh, P., Shaikh, G.N., Chapaneri, S. and Jayaswal, D. (2020, April). Skin lesion classification using deep learning architectures. In 2020 3rd International conference on communication system, computing and IT applications (CSCITA), 168–173. IEEE. 10.1109/CSCITA47329.2020.9137810.
- Tschandl, P., Rosendahl, C., & Kittler, H. (2018). The HAM10000 dataset, a large collection of multi-source dermatoscopic images of common pigmented skin lesions. *Scientific Data*, 14(5), 1–9. <https://doi.org/10.1038/sdata.2018.161>
- Van Onselen, J. (2011). Skin care in the older person: Identifying and managing eczema. *British Journal of Community Nursing*, 16(12), 578–580. <https://doi.org/10.7748/ns1999.08.13.48.47.c2664>
- White-Chu, E. F., & Reddy, M. (2011). Dry skin in the elderly: Complexities of a common problem. *Clinical Dermatology*, 29(1), 37–42. <https://doi.org/10.1016/j.clindermatol.2010.07.005>
- Yu, L., Chen, H., Dou, Q., Qin, J., & Heng, P. A. (2016). Automated melanoma recognition in dermoscopy images via very deep residual networks. *IEEE Transactions on Medical Imaging*, 36(4), 994–1004. <https://doi.org/10.1109/TMI.2016.2642839>
- Zeiler, M. D. (2012). Adadelta: An adaptive learning rate method, *arXiv preprint arXiv:1212.5701*. <https://doi.org/10.48550/arXiv.1212.5701>.
- Zhang, J., Xie, Y., Xia, Y., & Shen, C. (2019). Attention residual learning for skin lesion classification. *IEEE Transactions on Medical Imaging*, 38(9). <https://doi.org/10.1109/TMI.2019.2893944>