Journal of Information & Optimization Sciences
ISSN 0252-2667 (Print), ISSN 2169-0103 (Online)
Vol. 46 (2025), No. 2, pp. 345–358
DOI : 10.47974/JIOS-1919

Deep models based skin disease classification for efficient cancer detection

Lakshmi Deepthi Nukala *
G. Madhukar Rao †
Department of Computer Science and Engineering
Koneru Lakshmaiah Education Foundation
Moinabad Road
Aziznagar 500075
Hyderabad
India

Abstract

A growing need for healthcare treatments for skin conditions, melanoma, etc, in the medical prognosis for skin illnesses. The timely identification or precise diagnosis of skin conditions or illnesses can avert subsequent health issues. For categorizing skin illnesses across high-quality photos, a variety of statical image processing techniques, histogram equalization (HE), Gabor filters, morphological operations, and Gray-level co-occurrence matrix (GLCM) are often utilized. Deep models effectively train or learn the skin images to reliably classify skin diseases for photos with moderate to standard quality. In terms of precision and accuracy performance metrics, deep learning-based skin healthcare solutions pay closer attention than classic image processing techniques. Convolutional neural networks (CNN) and their extended derivations are prompted for the substantial skin illness categorization findings in medical healthcare related to skin diagnostics. Current CNN models, such as Basic CNN, Visual Geometric Group Deep CNN (VGG), Dense CNN (DenseNet), and Residual Neural Network (ResNet), are experimented on benchmarked skin disease datasets to analyze classification performance in terms of accuracy and loss.

Subject Classification: *Primary 93A30, Secondary 49K15.*

Keywords: *CNN, Skin disease classification, Deep learning, Medical prognosis, Healthcare.*

* *E-mail:* **lakshmideepthi16@gmail.com (Corresponding Author)**
† *E-mail:* **madhusw511@klh.edu.in**

© Taru Publications

1. Introduction

The most important component of treating infections that millions of people worldwide suffer from is skin healthcare. An essential component of skin healthcare research is the diagnosis and classification of skin illnesses. A variety of image processing approaches are used in recent research to accurately extract features from images and apply deep learning models to the problem of skin disease categorization. In order to explain the main conclusions of skin healthcare, a preliminary investigation of skin healthcare technology is offered in this work. Skin image features that are helpful for disease analysis are effectively extracted by the statistical skin image processing algorithms SOM [1], GLCM [2], and HE [3]. To obtain image and numerical supported skin datasets, consult [4]; to extract skin image properties, apply the most interesting image processing techniques.

The main task in the category of skin diseases is skin lesion analysis [5]. This analysis is carried out over the vast collection of dermoscopic images using deep learning algorithms. Identifying cases of malignant melanoma for the patient at the time of diagnosis lowers the death rate and saves lives. It is quite challenging to find the dermoscopic images that are presented with the greatest intra-class similarity and the least inter-class similarity. According to recent advances, deep learning acts as a vital role in the diagnosis of skin conditions.

Classifying skin images is crucial for diagnosing early stages of skin cancer. Medical skin healthcare research has focused increasingly on making accurate predictions about skin diseases. The application of deep model and machine learning approaches is leading to significant and ongoing advancements in skin prognosis. This research examines several deep-learning models for practical applications in skin healthcare. Two of the most important problems in skin care are identifying features with low contrast in areas of skin images and addressing the notable discrepancy in skin image regions. In the majority of recent research, machine learning and computer vision algorithms have achieved remarkable success in such difficult problems [6].

Convolutional neural networks (CNNs) are used by researchers today [7] to accurately classify skin lesions. To classify skin lesions, four deep CNNs have been extensively researched: CNN, VGG 16 (which has 16 layers) [8], DenseNet [9], and ResNet [10]. These algorithms analyze the usefulness of skin image classifications for dermatologists to efficiently find the predictions of skin melanoma. First, the region of interest (ROI)

segmentation is obtained by pre-processing the dermoscopic color images. The deep models use the ROI segmentations as their input. The basic CNN is composed of five architectural building blocks: the convolutional layer, pooling layer, connected layer, loss layer, and nonlinear rectified linear units (ReLU). In the convolutional layer, the segmented ROI of the input image undergoes convolution operations; the outputs are then sent to the pooling layer. Applying the various pooling operation types—Average, Max, Sum, etc.—reduces the dimensionality of the picture features. The element-wise operations—that is, applying the ReLU operation to every pixel—are carried out by the ReLU layer.

The three sets of convolution and max pooling layers, flatten, dense, and dropout, and final dense fully connected layers are defined successively in the CNN models' sequential layers approach. The final dense, fully connected layer is created at the CNN's end layer after neurons and their connections have been filtered. Another expanded CNN model with 16 layers intended for real-world data training is called VGG 16. The combination of completely connected layers known as DenseNet enhances the transmission of features. Hence, compared to CNN and VGG 16, it is more effective in training or learning the picture data. With the residual nature of the network, a more substantial number of layers are added to a current deep residual neural network (ResNet) model to increase classification accuracy. The ensuing subsections provide the specifics. This method was tested using the ISIC database [4], where about 7000 photos were utilized for testing and a total of 18000 distinct photographs of skin diseases were used to train eight different classifications. The first layer seeks to describe the height, width, and number of channels of the input skin image. The ISIC database's input photos are cropped to the appropriate dimensions.

Section 2 presents a thorough analysis of pertinent deep-learning methods for classifying skin diseases using picture data. The design elements of four deep models for the categorization were covered in Section 3. The performance of cutting-edge deep learning algorithms is reviewed in Section 4, and the study's findings and the direction of ongoing research are depicted in Section 5.

2. Literature Study of The Deep Models

With their dependable training, deep learning techniques enable the delivery of skin disease or cancer categorization. Deep learning has

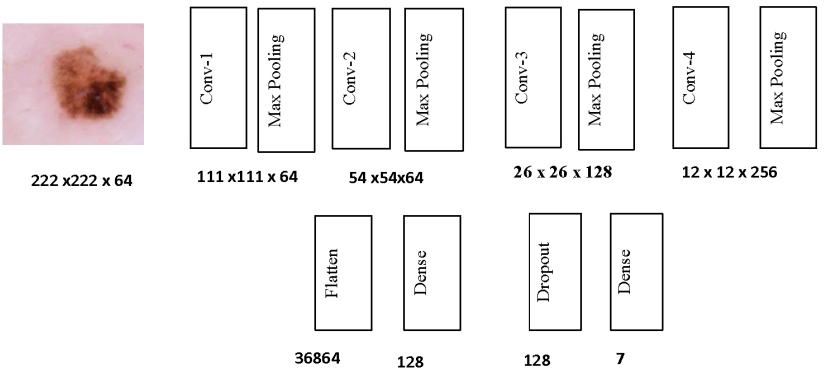


Figure 1

CNN with Stacked Layers

enabled careful, forward-thinking research to yield noteworthy results in skin healthcare in the last few years.

Finding signals or features for skin photos automatically without the need for feature extraction techniques is the alluring aspect of deep learning technology. Its design of the convolutional neural network (CNN) [11] with several stacked layers is a significant advancement. Fig. 1 displays the CNN stacked layers as well as its main building architectural diagram.

The skin images are taken in the first layer according to the specifications W x H x C. Pre-processing is then done to get all the test and trained images in the specified W, H, and C dimensions.

A convolutional neural network (CNN) normally consists of three layers: a convolutional layer, a fully connected layer, and a pooling layer. Convolution and pooling, the first two layers, are used to extract features. Convolutional neural networks (CNNs) have an architecture that includes a convolution layer, and a specific type of linear operation [12]. Convolution, pooling, and fully linked layers are some of the fundamental components of the CNN design [13]. This type of architecture typically consists of one or more fully connected layers, a pooling layer, and multiple convolution layers.

Convolution is a mathematical technique that involves applying an array of numbers, known as the "tensor," to a smaller array of numbers, known as the "kernel," with the goal of extracting patterns from the data [14]. The output value at each point of the output tensor, also referred to as a feature map, can be extracted by performing a feature-wise multiplication over each element of the kernel's value and the source tensor, then summing the results. The convolution operation is

demonstrated in this example [15]. The following methods can be used to compute the convolved features for a black-and-white image or the frames of a video clip:

1	1	1	0	0
0	1	1	1	0
0	0	1	1	1
0	0	1	1	0
0	1	1	0	0

Image or Frame (5 x5)

1	1	1
0	1	0
1	0	1

Kernel of Filter Matrix (3 x 3)

5	4	4
3	5	4
2	4	5

Convolved Features

The same process is applied to the R, G, and B channels if the image is color, resulting in three sums of convolved features for each channel. As demonstrated in the example below, it is possible to acquire the convolved features beneath the three channels.

0	0	0	0	0	0	0
0	152	151	142	143	144	145
0	153	152	143	144	145	146
0	154	153	144	145	146	147
0	155	154	145	146	147	148
0	156	155	146	147	148	149
0	157	156	147	148	149	150

Frame Values - Red Channel

0	0	0	0	0	0	0
0	142	141	152	153	154	155
0	143	142	153	154	155	155
0	144	143	154	155	156	156
0	145	144	155	156	157	156
0	146	145	156	157	158	157
0	147	146	157	158	159	156

Frame Values - Green Channel

0	0	0	0	0	0	0
0	162	155	162	163	164	165
0	143	156	163	160	163	164
0	144	157	164	160	165	166
0	145	158	165	162	166	167
0	146	159	166	161	167	157
0	147	160	167	166	168	168

Frame Values - Blue Channel

-1	0	0
1	0	1
-1	0	-1

Kernel or Filter Matrix for Red Channel

1	0	1
0	0	0
0	0	0

Kernel or Filter Matrix for Green Channel

0	1	1
0	1	1
0	0	1

Kernel or Filter Matrix for Blue Channel

472	478	483	490	491
772	940	949	980	987
757	946	951	980	991
761	952	955	985	986
767	958	964	989	989

CNN Output for the Convolved RGB Inputs

CNN Output for the Convolved RGB Inputs

CNN frequently uses the two traditional pooling techniques, max pooling and average pooling [16]. The following displays these findings. produced a 5 x 5 convolution matrix, and then used the filter matrix 'F', which has a size of 2 x 2, and stride S=1, to apply either maximum pooling (max pooling) or average pooling (avg pooling). The size of the resulting features matrix is then calculated using the formula (I-F)/S + 1, where I is the number of rows or columns of the resulting convolution matrix. As a result, 4 × 4 is the maximum (or average) pooling. The following are the max pooling (or average pooling) matrices that were produced:

940	949	980	987
946	951	980	991
952	955	980	991
958	964	989	989

Output for the Convolved RGB Inputs (Method Applied: CNN Pooling Max)

665.5	712.5	726	737
853.75	946.5	965	984.5
854	951	967.75	985.4
859.5	957.25	973.25	987.25

Output for the Convolved RGB Inputs (Method Applied: CNN Pooling Average)

A neural network's output can vary from -1 to 1, or from 0 to 1, based on the activation function. Linear and nonlinear activation functions are the two varieties. Since its output is limited, the linear activation function has a linear form. Nonlinear activation functions come in three varieties: sigmoid, tanh, and ReLU (Rectified Linear Unit) [17]. The logistic or sigmoid activation function is the most widely used type of activation function. Its value is in the range of 0 and 1. In deep learning systems, feature maps from the last convolution or pooling layer are often flattened into a one-dimensional array of numbers (or vectors). This is followed by the connection of one or more dense layers, also known as fully linked layers [18], [19], [20], to this 1D array. Between each input and each output in these layers is a learnable weight. A subset of fully connected layers, for example, uses the down-sampled input from the pooling layers and the recovered data from the convolution layers to map the network's final

outputs, such as the probabilities for each class, in a classification task. Usually, the number of classes determines which output nodes are in the final completely linked layer. Some recent research [20], [21], and [22-23] has used VGG16, DenseNet, and ResNet as good deep learning models. These models have demonstrated strong classification results and are the most effective in evaluating skin lesions and early detection of skin cancers.

VGG 16 Model

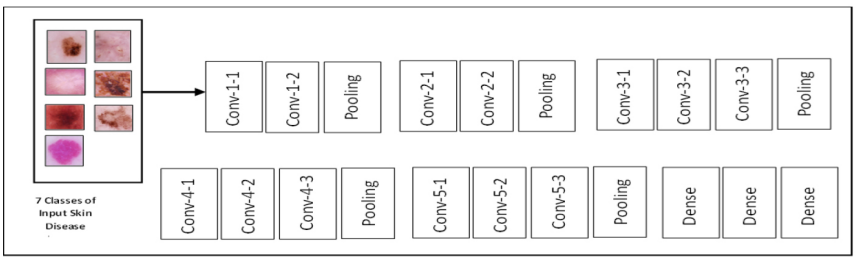


Figure 2

VGG 16 Deep Model for Skin Diseases Training with 16 Layers

One of the best methods for designing convolutional neural networks is the VGG 16 [22], which consists of sixteen layers total, of which three fully connected levels and thirteen convolution layers. Fig. 2 depicts the construction of the VGG 16. Seven skin disease photos plus sample images from seven classes make up the ISIC dataset. As seen in the VGG 16 architecture, these photos have undergone pre-processing to extract the region of interest. Every image should be captured at 224 x 224 x 3 in the VGG 16 model. The first two convolutions (1-1, 1-2) and one max pooling applied with 64 filters make it 112 x 112 x 64 in size; the subsequent sequence of two convolutions and one max pooling applied with 64 filters makes it 56 x 56 x 128 in size; the remaining layers and the skin image data's shape are finally described as 7 x 7 x 512 in its VGG Model

DenseNet Model

This work attempts to implement DenseNet-121, which includes all of the connections from layer to layer. The resulting concatenation of characteristics from the architecture's earlier levels becomes the input layer of DenseNet [21]. In comparison to CNN and VGG 16 models, it produced better accurate results for skin disease classification by

processing the features of the layers to any other layer in the DenseNet. The vanishing gradient problem arises because of the DenseNet's many layers. When the epoch numbers are increased utilizing the DenseNet deep model's more significant number of layers, accuracy declines. After multiple epochs, a small gradient—almost zero—has emerged. As a result, loss values rise and accuracy decreases. A residual neural network (ResNet) is made to solve the gradient vanishing problem in order to avoid the intricacy of network design [20].

ResNet Deep Learning Model

A residual deep learning model called ResNet [22] is essentially a pre-trained model with a larger number of layers and connections between them. Convolutional, ReLu, pooling, normalizing, addition, fully connected, and classifier layers make up its layers. Two convolution layers and one batch normalization layer make up ResNet's residual building piece. Even when applying skin disease categorization utilizing a higher number of layers, the residual property reduces the network's complexity.

3. Design Aspects of Image Processing Based Dcnn Models for The Classification Of Skin Diseases

For the DCNN to produce reliable skin disease diagnoses based on a large number of pertinent photos, certain design elements are necessary. Pre-processing transfer learning and dropout layer are essential components of any DCNN model's construction. The preceding sections contain architectural views of the CNN, VGG 16, DenseNet, and ResNet DCNN models. It was observed that effective DCNN model implementations, in conjunction with these important design elements, guarantee a higher classification rate for skin conditions or tumors.

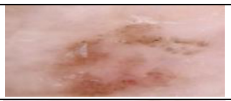
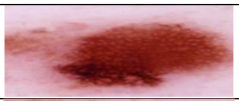
4. Performance of State-Of-The-Art of Deep Learning Techniques

The ISIC resource [21], [22], which contains seven classes of skin illness photos that can be used to train the four deep models covered in the previous sections, is where the skin image databases are gathered. These seven categories of photos are: vascular lesion, melanoma, melanocytic nevus, basal cell cancer, actinic keratosis, and benign keratosis. Sample images of these images are shown in Table 1.

There are three types of data in the ISIC dataset: nevus, seborrheic keratosis, and melanoma. For the experimental testing, there are a total of 326, 514, 1099, 115, 1113, 1521, and 432 photos in these seven classes. The performance of the most advanced deep learning method is examined to determine the loss and classification accuracy at each generation of the epoch. Figs. 3, 4, 5, and 6 illustrate the accuracy gains at each epoch and the loss reduction at each epoch using the CNN, VGG 16, DenseNet, and ResNet deep learning models, respectively.

Table 1

Skin Diseases Classes and its Sample Images

Name of the Skin Disease Class	Sample Image 1	Sample Image 2
Actinic keratosis		
Basal cell carcinoma		
Benign keratosis		
Dermatofibroma		
Melanoma		
Melanocytic Nevus		
Vascular Lesion		

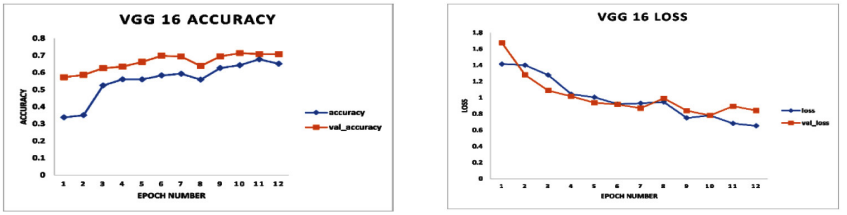


Figure 3

Accuracy Improvement and Loss Reduction at Each Epoch in CNN Model

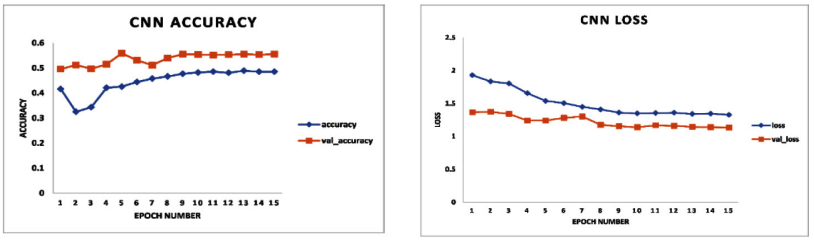


Figure 4

Accuracy Improvement and Loss Reduction at Each Epoch in VGG 16 Model

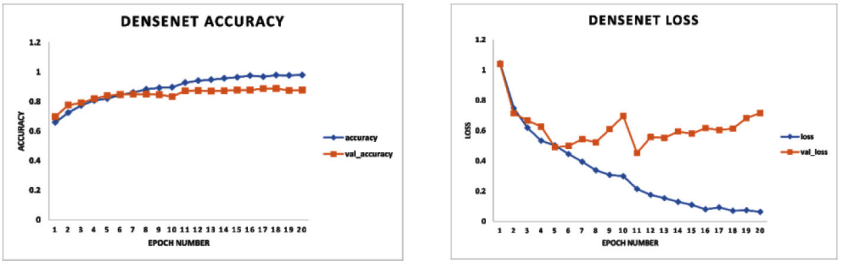


Figure 5

Accuracy Improvement and Loss Reduction at Each Epoch in DenseNet Model

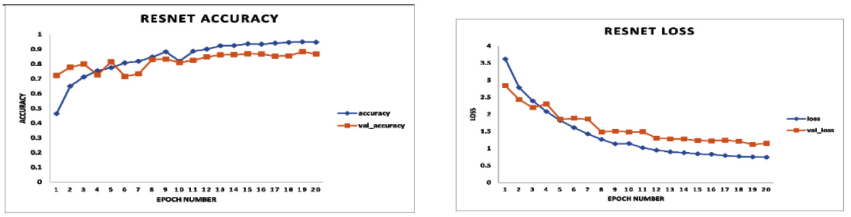


Figure 6

Accuracy Improvement and Loss Reduction at Each Epoch in ResNet Model

It is seen that the accuracy levels are greatly increased for both the provided skin image data and test instances of image data after the CNN model is trained and identifies the skin illness picture throughout the course of 15 epochs. Considering that test case image data accuracy is defined as val_accuracy in the line graphs (defined with orange color), and skin image data accuracy is read as accurate in the graph (blue color). Comparably, the loss function's two parameters, loss and val_loss, are defined by the line graph. Loss and val_loss are progressively decreased for each epoch when CNN is used to learn the data for the classification of skin diseases. The VGG 16 uses 16 layers, which comprise 13 convolutional layers and three fully connected layers. Using the 16 layers, it has successfully taught the skin illness images.

In comparison to CNN and VGG 16, the DenseNet and ResNet models' extensive topologies and efficient learning led to their remarkable success in categorizing photos of skin diseases. The effectiveness of deep learning is increased with several layers, resulting in decreased loss values and excellent accuracy. According to the accuracy comparison in Fig. 7, both the DenseNet and ResNet models performed better than the others for the classification of skin disease images in medical settings.

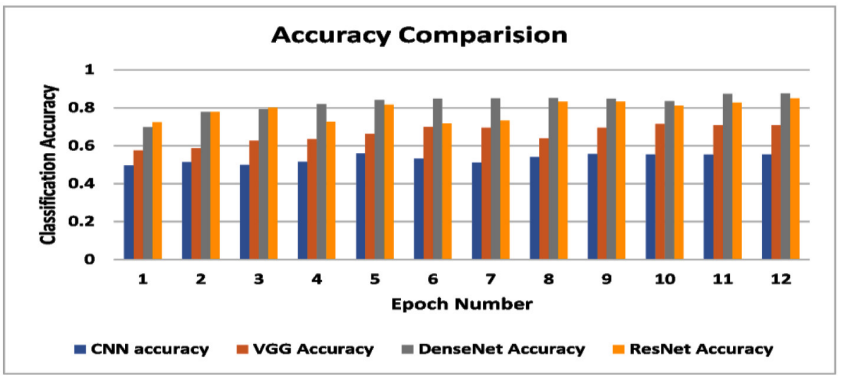


Figure 7

Accuracy Comparative Analysis for CNN, VGG 16, DenseNet, and ResNet Models

5. Conclusion and scope of the future research work

In the field of medicine, deep learning methods are being used to diagnose skin conditions. For the initial study of skin disease image categorization, ROI pre-processing of images is required. Moreover, ROI

skin disease photos undergo deep learning. Massive photos of ISIC skin illnesses are used for research and experimentation with the CNN, VGG 16, DenseNet, and ResNet models as well as their architectural design processes. To determine the accuracy and loss levels, each deep learning model runs approximately 15 to 20 epochs. In contrast to the CNN and VGG 16 models, the experimental study found that the DenseNet and ResNet models increased the accuracy rate from 8 to 16. For practical training in seven types of skin illnesses, DenseNet and ResNet both use architectures with a higher number of layers. Therefore, in comparison to previous deep-learning models, they significantly lower the loss values and increase the skin disease image categorization rate. Single-view modelling is used for the deep learning process. Future research could focus on creating multi-view deep modelling methods for the most precise classification of skin diseases in skin care applications.

References

[1] Z. Abbas, M. U. Rehman, S. Najam, and S. M. Danish Rizvi, "An Efficient Gray-Level Co-Occurrence Matrix (GLCM) based Approach Towards Classification of Skin Lesion," 2019 Amity International Conference on Artificial Intelligence (AICAI), Dubai, United Arab Emirates, pp. 317-320 (2019), doi: 10.1109/AICAI.2019.8701374.

[2] T. Goswami, V. K. Dabhi, and H. B. Prajapati, "Skin Disease Classification from Image - A Survey," 2020 6th International Conference on Advanced Computing and Communication Systems (ICACCS), Coimbatore, India, pp. 599-605 (2020), doi: 10.1109/ICACCS48705.2020.9074232.

[3] M. Rana and M. Bhushan, "Machine learning and deep learning approach for medical image analysis: diagnosis to detection," Multimedia Tools and Applications, vol. 82, pp. 26731–26769 (2023). doi: 10.1007/s11042-022-14305-w.

[4] A, Joy Christy, A, Umamakeswari, V, Ramya Madhuri and A, Srilakshmi, "A novel CNN framework for skin disease classification using adaptive percentage filter for image binarization and fast-marching inpainting method," *Multimedia Tools and Applications* (2024). doi: 10.1007/s11042-023-17967-2.

[5] Y. S. Alsahafi, M. A. Kassem, and K. M. Hosny, "Skin-Net: a novel deep residual network for skin lesions classification using multilevel feature extraction and cross-channel correlation with detection of

outlier," *Journal of Big Data,* vol. 10, p. 105 (2023). doi: 10.1186/s40537-023-00769-6.

[6] N. Girdhar, A. Sinha, and S. Gupta, "DenseNet-II: an improved deep convolutional neural network for melanoma cancer detection," *Soft Computing,* vol. 27, pp. 13285–13304 (2023). doi: 10.1007/s00500-022-07406-z.

[7] B. Zhang, X. Zhou, Y. Luo, L. Li, and Z. Wang, "Opportunities and Challenges: Classification of Skin Disease Based on Deep Learning," *Chinese Journal of Mechanical Engineering*, vol. 34, p. 112 (2021). doi: 10.1186/s10033-021-00629-5.

[8] Z. Wu, Y. Zhang, L. Liu, and S. Zhao, "Studies on Different CNN Algorithms for Face Skin Disease Classification Based on Clinical Images," *IEEE Access,* vol. 7, pp. 66505-66511 (2019). doi: 10.1109/ACCESS.2019.2918221.

[9] S. Rahman, M. R. Hossain, and A. F. Alam, "Attention-Based CNN Model for Burn Severity Assessment," 2023 IEEE EMBS International Conference on Biomedical and Health Informatics (BHI), Pittsburgh, PA, USA, pp. 1-4 (2023). doi: 10.1109/BHI58575.2023.10313435.

[10] Y. Weng, T. Zhou, L. Liu, and C. Xia, "Automatic Convolutional Neural Architecture Search for Image Classification Under Different Scenes," *IEEE Access*, vol. 7, pp. 38495-38506 (2019). doi: 10.1109/ACCESS.2019.2906369.

[11] L. Yu, J. Guo, and C. Yu, "Application and Improvement of Convolution Operation in Image Restoration," 2022 IEEE 6th Advanced Information Technology, Electronic and Automation Control Conference (IAEAC), Beijing, China, pp. 1008-1012 (2022). doi: 10.1109/IAEAC54830.2022.9929846.

[12] P. Dileep, D. Das, and P. K. Bora, "Dense Layer Dropout Based CNN Architecture for Automatic Modulation Classification," 2020 National Conference on Communications (NCC), Kharagpur, India, pp. 1-5 (2020), doi: 10.1109/NCC48643.2020.9055989.

[13] H. Zhang, Y. Zhao, Z. Liu, and L. Wu, "Research on the Classification of Benign and Malignant Parotid Tumors Based on Transfer Learning and a Convolutional Neural Network," *IEEE Access,* vol. 9, pp. 40360-40371 (2021). doi: 10.1109/ACCESS.2021.3064752.

[14] V. Bhatnagar, S. D. Sharma, and R. Kumar, "Descriptive analysis of COVID-19 patients in the context of India," *Journal of Interdisciplinary Mathematics*, vol. 24, no. 3, pp. 489-504 (2021).

[15] C. Szegedy, V. Vanhoucke, S. Ioffe, J. Shlens, and Z. Wojna, "Rethinking the inception architecture for computer vision," Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, pp. 2818–2826 (2016), doi: 10.1109/CVPR.2016.308.

[16] E. T. Hastuti, A. Bustamam, P. Anki, R. Amalia, and A. Salma, "Performance of True Transfer Learning using CNN DenseNet121 for COVID-19 Detection from Chest X-Ray Images," 2021 IEEE International Conference on Health, Instrumentation & Measurement, and Natural Sciences (InHeNce), Medan, Indonesia, pp. 1-5 (2021). doi: 10.1109/InHeNce52833.2021.9537261.

[17] L. Alzubaidi, J. Zhang, A. J. Humaidi, Y. L. Zhang, X. Z. Fu, and M. X. S. S. T. Zhang, "Review of deep learning: concepts, CNN architectures, challenges, applications, future directions," *Journal of Big Data*, vol. 8, p. 53 (2021). doi: 10.1186/s40537-021-00444-8.

[18] N. Gouda and J. Amudha, "Skin Cancer Classification using ResNet," 2020 IEEE 5th International Conference on Computing Communication and Automation (ICCCA), Greater Noida, India, pp. 536-541 (2020). doi: 10.1109/ICCCA49541.2020.9250855.

[19] M. Suleman Basha, S. K. Mouleeswaran, and K. Rajendra Prasad, "Hybrid visual computing models to discover the clusters assessment of high dimensional big data," *Soft Computing*, vol. 27, pp. 4249–4262 (2023). doi: 10.1007/s00500-022-07092-x.

[20] K. Rajendra Prasad, G. R. Kamatam, M. B. Myneni, and N. R. Reddy, "A novel data visualization method for the effective assessment of cluster tendency through the dark blocks image pattern analysis," *Microprocessors and Microsystems*, vol. 93, p. 104625 (2022). doi: 10.1016/j.micpro.2022.104625.

[21] G. Saini, N. B. Kumar, and R. D. Singh, "Structural equation based model to investigate the moderating effect of fear of COVID using partial least square method," *Journal of Interdisciplinary Mathematics*, vol. 25, no. 3, pp. 703-720 (2022).

[22] M. A. Kassem, K. M. Hosny, and M. M. Fouad, "Skin Lesions Classification into Eight Classes for ISIC 2019 Using Deep Convolutional Neural Network and Transfer Learning," *IEEE Access*, vol. 8, pp. 114822-114832 (2020). doi: 10.1109/ACCESS.2020.3003890.

[23] G. M. Rao, D. Ramesh, P. Gantela, and K. Srinivas, "A hybrid deep learning strategy for image based automated prognosis of skin disease," *Soft Computing* (2023). doi: 10.1007/s00500-023-08007-0.

Received November, 2024