World Journal of Advanced Research and Reviews
eISSN: 2581-9615 CODEN (USA): WJARAI
Cross Ref DOI: 10.30574/wjarr
Journal homepage: https://wjarr.com/

(Research Article)

Check for updates

Skin Cancer Detection using VGG-16

Dilruba Shareen [1, 2, *] and Nazia Hossain [3]

[1] *Department of Computer Science and Engineering, Khulna University of Engineering and Technology (KUET), Khulna, Bangladesh.*
[2] *Department of Statistics and Data Science, Jahangirnagar University, Savar, Dhaka-1342, Bangladesh.*
[3] *Centre for Smart Analytics, Federation University Australia, Victoria, Australia.*

World Journal of Advanced Research and Reviews, 2024, 24(02), 2930-2937

Publication history: Received on 07 October 2024; revised on 23 November 2024; accepted on 28 November 2024

Article DOI: https://doi.org/10.30574/wjarr.2024.24.2.3520

Abstract

The recent increase in the prevalence of skin cancer, along with its significant impact on individuals' lives, has garnered the attention of many researchers in the field of deep learning models, especially following the promising results observed using these models in the medical field. This study aimed to develop a system that can accurately diagnose one of three types of skin cancer: basal cell carcinoma (BCC), melanoma (MEL), and nevi (NV). Additionally, it emphasizes the importance of image quality, as many studies focus on the quantity of images used in deep learning. In this study, transfer learning was employed using the pre-trained VGG-16 model alongside a dataset sourced from Kaggle.

Keywords: Skin cancer; Deep learning; Basal cell carcinoma (BCC); Melanoma (MEL); Nevi (NV), Dermoscopic images; VGG-16

1. Introduction

Cancer encompasses various types of malignant tumors, commonly referred to as neoplasms in medicine. Skin cancer starts in the cells of the skin [1].

Normally, skin cells grow and divide to make new cells. The old cells die, and new ones replace them. But sometimes, this process breaks. New cells might grow when they aren't needed, and old cells don't die when they should. This pile-up of extra cells forms a lump known as a tumor. This can be caused by many things, but a main cause is too much sun exposure, which can lead to a cancerous tumor. The tumor then harms the area around it and can even spread to other parts of the body [13].

There is a lot of new interest in finding and treating skin cancer. This is because it is very common and causes a lot of harm. Skin spots can be grouped in two ways: harmful (malignant) or harmless (benign) moles. Melanoma is one of the most deadly types of cancer. It causes about 70% of all skin cancer deaths. Skin cancer mainly shows up as serious damage to the outer layer of the skin.

Finding the disease early is very important to help patients recover. Because of this, a lot of work is being done to find good ways to diagnose the disease in its early stages. Older computer methods were used for this important task. But since people's lives are at risk, it is extremely important to be as accurate as possible. For this reason, deep learning (a type of artificial intelligence) is now being used to get the most accurate results.

* Corresponding author: Dilruba Shareen

In their research, Jayalakshmi et al achieved an accuracy of 89.3% by customizing and tuning the CNN model while using the PH2 dataset in a two-class classification scenario [14]. In general, the excellence of the convolutional neural network in image classification has been widely approved across various applications.

Hossain et al. [16] on ML-based Intrusion Detection Systems (IDS) can be adapted to the domain of IoMT-driven skin cancer detection. In this context, IDS techniques can be used to safeguard the data collected from IoMT devices like wearable sensors or imaging devices used for detecting skin cancer.

Brindha et al. unveiled the superiority of the CNN algorithm over the SVM algorithm in the classification of ISIC image dataset, resulting in a significant increase in accuracy from 61% to 83%. [4].

Pham and his colleagues achieved an accuracy of 79.5% and 87% in classifying the ISIC dataset by utilizing Transfer Learning methods, specially, Reznet50, and InceptionV3, respectively [19].

Mijwil exploited and compared three different architectures, namely, VGG19, ResNet, and Inception V3, to detect skin cancer using the ISIC2019 and ISIC2020 archives. The dataset consisted of a significant number of more than 24,000 images. They found an accuracy of 73.11%, and the best 86.9% for the mentioned architectures [17].

In their study, Nawaz et al. combined a region-based CNN technique with the support vector Machine (SVM) classifier and utilized the ISIC2016 dataset for melanoma classification. To increase the dataset size, they employed data augmentation techniques, resulting in more than 7,000 images. Their approach achieved an accuracy of 89.1% [18]. In their investigation, Alzubaidi and his colleagues achieved a classification accuracy of 97.5% for skin lesion images using a deep learning method. They employed a multi-phase training scenario and a multistage CNN model with the aim of surpassing the limitations posed by limited number of labeled data for medical applications [2]. In their paper, Ashraf et al. conducted an examination of skin lesion images with the help of deep learning method. They employed region of interest segmentation preprocessing and image augmentation. The initial result without region of interest segmentation and augmentation was approximately 81.3%. However, by implementing the segmentation and augmentation, they acquired an increase to 97.2% in the classification accuracy [3]. Rafi and coworkers achieved an accuracy of 98.7% by applying transfer learning architectures based on Efficient NET-B7. Their approach involved extensive image pre-processing, including resizing, conversion, augmentation, and in particular, a post scaling step [20]. Lafraxo and coworkers proposed a CNN architecture for recognizing malignancy in dermoscopic images. In their approach, they employed regularization, as well as geometric and color augmentations to enlarge the datasets. Specifically, they augmented the ISBI dataset to 18,000 images, the PH2 dataset to 2,880 images, and the MED-NODE dataset to 1,800 images. The achieved accuracies were 98.44%, 97.39%, and 87.77% respectively [15]. Rasel and his colleagues implemented a deep CNN model based on transfer learning, with the main ideas borrowed from LeNet. Their model consists of a total of 31 layers and utilized nonlinear variable Leaky ReLU activation function. The training was conducted over 250 epochs. They achieved accuracies of 75.50%, 97.50%, and 98.33% for PH2, augmented (rotated) PH2, and a smaller subset of images from ISIC archives, respectively [21]. Hassan et al. conducted a comprehensive literature survey to assess the performance of different optimization algorithms. Additionally, they demonstrated accuracies of 97.3% (92% up to 98%) and 99.07% for their deep learning model applied to the ISIC dataset (with 6000 iterations) and the COVIDx dataset (with 300 iterations), respectively. These impressive results were obtained by utilizing the Adam optimizer [10]. Furthermore, Hassan et al. achieved a superior accuracy of 97.47%, employing ResNet50 and Adam optimizer for the classification of retinal optical coherence tomography images with 84495 total number of images [9]. Alahmadi and coworkers presented a CNN/transformer coupled network, that incorporated both supervised and unsupervised training techniques. Their approach yielded accuracy rates of 95.51% and 97.11% for ISIC and PH2 datasets, respectively [1]. Wu et al. proposed and developed a novel two-stream network, that efficiently capture both local features and global long-range dependencies by combining a CNN with an additional transformer branch. They achieved accuracies of 95.78% (ISIC2018), 93.26% (ISIC2017), 96.04% (ISIC2016), and 97.03% (PH2) for the respective datasets. For a better model initialization, they used deit-tiny-distilled-patch16-224 and ResNet34. They also utilized dynamic polynomial learning rate decay [26]. In [27] Iqbal et al. presented a stacked CNN model for the classification of seven types of skin lesions along with preprocessing and data augmentation techniques. They examined the results with and without image preprocessing to assess the significance of this step in skin cancer diagnosis.

Rahman et al. proposed a weighted average ensemble model for the classification of seven types of skin lesions [28]. They considered five cutting-edge architectures for the model: ResNet, DenseNet, Xception, ResNeXt, and SeRes-NeXt. The problem of excessive data imbalance was solved by using cost-sensitive learning. In terms of performance, DenseNet came out on top, and the weighted average ensemble model outscored the individual models.

[31] proposed an automated system for early melanoma detection using novel deep transfer learning. The method used in [31] leveraged a pre-trained NASNet model, from which features were transferred to a new dataset for classification. They adapted the original network by incorporating global average pooling and customized classification layers. The system was trained and evaluated on skin images from the ISIC 2020 dataset [31]. [32] proposed a novel optimized NASNet architecture, in which the NASNet model was enhanced with additional data and an additional basic layer that was employed in CNNs. The strategy that was proposed enhanced the model's capacity to deal with incomplete and inconsistent data. A dataset of 2,637 skin images was used to demonstrate the benefits of the technique that was proposed. They analyzed the performance of the suggested method by examining its precision, sensitivity, specificity, F1-score, and area under the ROC curve. The accuracy of Optimized NASNet Mobile and NASNet Large provided accuracies of 85.62% and 83.98%, respectively, for the Adam optimizer.

2. Dataset

We chose datasets from Kaggle to train three models. For the first model, 14,454 images were used for training, with 1671 images used for validation. After the data collection stage, we moved on to the preprocessing stage where three main concepts were applied to obtain uniform data in terms of size, pixel values, and array dimensions. The following three methods were applied: converting to NumPy arrays, normalization, and resizing images [30].

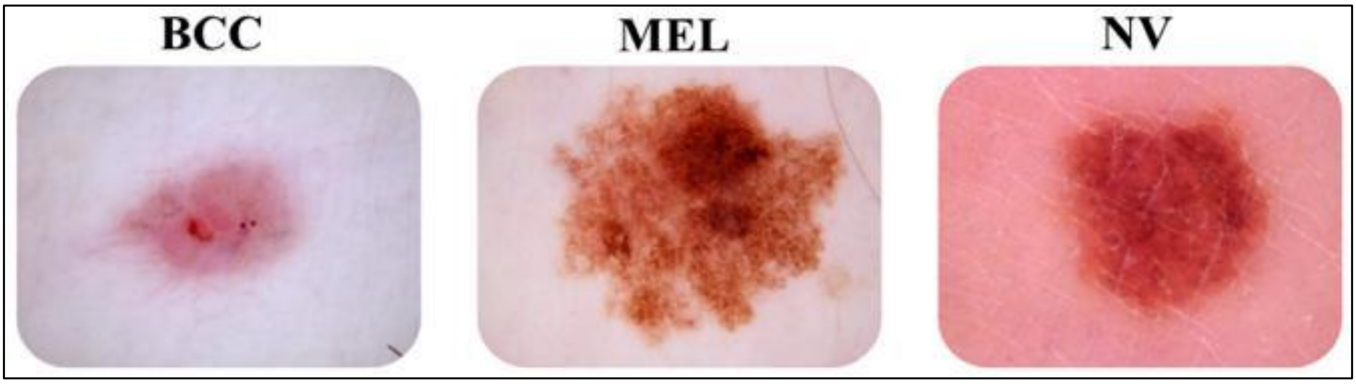

Figure 1 Images of "bcc" and "mel" skin cancer types and "nv" skin lesions

3. VGG 16

The pre-prepared VGG16 was prepared on a subset of the ImageNet dataset, and an assortment of more than 14 million pictures [34] having a place with 22,000 classifications was selected to create a custom skin anomaly diagnosis model for the following reasons: VGG has more convolution and pooling layers compared to previous architectures [34]. VGG16 achieved good results with only 7.0% of the top 5 test errors [34]. VGG 16 and VGG 19 were used in [34] for brain tumor classification. First, the images are pre-processed, and then the pre-processed images are classified using VGG-16 and VGG-19 in [34]. So, we used the same concept and framework here for skin cancer detection.

The main functions of each layer of VGG16 are as follows:

- Convolutional Layers

Convolutional layers extract features from the image, from basic edges to complex patterns [31], [32].

- Max-Pooling Layers

These layers reduce the spatial dimensions to control overfitting and reduce computational complexity [31], [32].

- Fully Connected Layers

These layers combine and refine the highlights extricated by the convolutional layers into a last feature vector [33].

Between the layers that we used in our model, two activation functions were applied:

- ReLU: ReLU acquaints non-linearity with the organization, permitting it to learn complex examples in the information, also used in [31], [32].

- Softmax: This function converts the classification scores into probabilities, providing the final output for classification [31], [32].

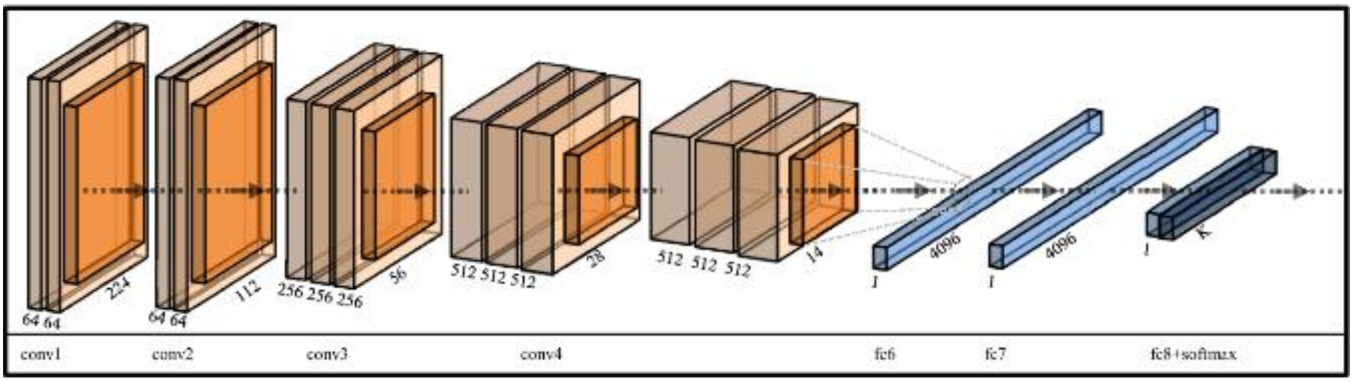

Figure 2 VGG16's pre-trained model architecture [29]

4. Training and Validation Model

We gave specific values to specific parameters needed for the training phase such as:

- Epochs = 40: The number of times the model will iterate over the entire training dataset.
- Batch_size = 32: The number of samples that will be propagated through the network at one time. After processing this batch, the model's weights will be updated.
- Callbacks: These are special functions that can be called during training at certain points.
- They are used for various purposes, such as saving the model after each epoch or stopping training early if the validation loss stops improving.
- Verbosity mode = 1: This controls the verbosity of the output, and "1" means progress updates will be shown during training.
- Optimizer = Adam: Adam is a popular optimization algorithm used for training machine learning models, especially deep learning models. "Adam" represents a versatile second assessment. It is an augmentation of the Stochastic Slope Plunge (SGD) calculation that determines versatile learning rates for every boundary [31], [32].

After defining all the needed parameters and their values, training and validation will begin, where the training phase is the period during which the model is exposed to the training dataset. During this phase, the model learns to map input features to output labels. [33] This is achieved by adjusting its weight based on the loss from predictions and true labels. The validation phase involves evaluating the model's execution on a different approval dataset that the model has not seen during preparation. This helps to monitor how well the model generalizes to new, unseen data [33].

Table 1 shows the CNN architecture, number of epochs, batch size, optimizer, dataset sizes, types of skin cancers/lesions, number of kernels, kernel size, padding, dropout rates, and activation functions.

Table 1 Models during the training and validation steps using some factors

Comparison Factor	**Model 1**
CNN model	VGG16
Number of epochs	40
Batch size	32
Optimizer	Adam
Number of images in training dataset	14,454

Number of images in validation dataset	1671
Types of skin cancers/lesions	bcc/mel/NV
Number of kernels	100
Kernel size	3 × 3
Padding	valid
Dropout	0.75
Activation function	Softmax

Figure 3 shows the validation accuracy curve. The validation accuracy is 84.72% Figure 4 shows the confusion matrix of VGG-16. The accuracy of Optimized NASNet Mobile and NASNet Large provided an accuracy of 85.62% and 83.98%, respectively, for Adam optimizer in [32] for malignant and benign skin lesions classification. We have improved this using VGG for basal cell carcinoma (bcc), melanoma (mel), and nevus detection. [31] used a NASNet-based deep transfer learning approach for an automated melanoma detection system. By integrating global average pooling and customized classification layers, the model was able to extract and classify dermoscopic image features. But it was binary classification, which also improves in this work because the number of images in the training and validation datasets is 14,454 and 1671, respectively, which is large.

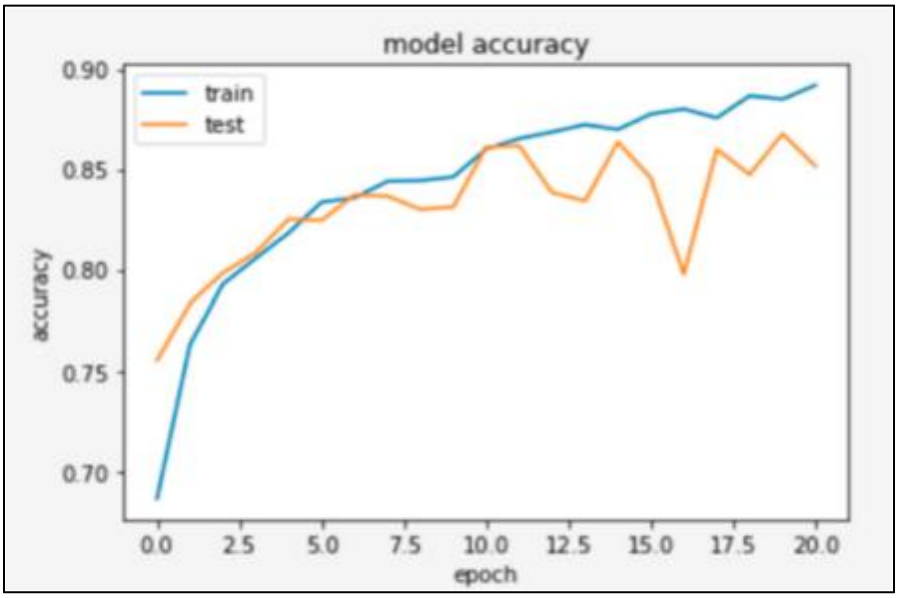

Figure 3 The validation accuracy curve

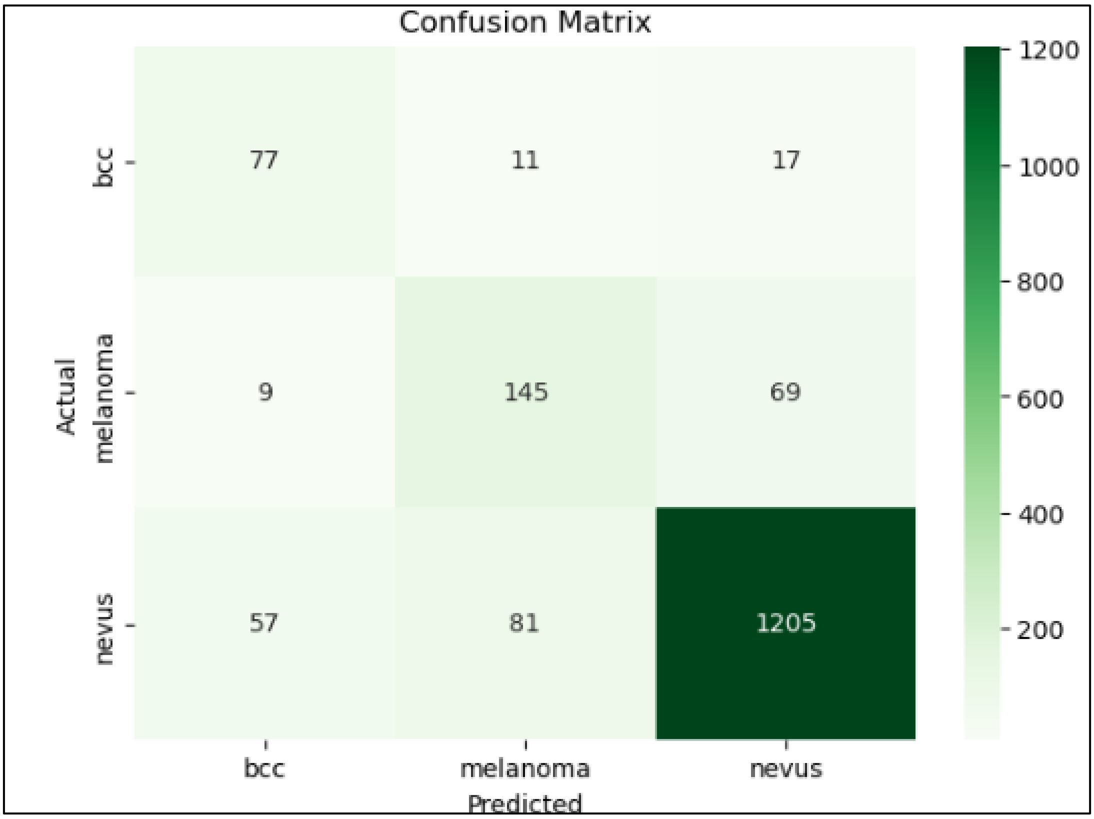

Figure 4 Confusion matrix of VGG-16

5. Conclusion

This research shows the effectiveness of the VGG-16 model in skin cancer diagnosis, and several avenues for future research could enhance its performance and applicability. Exploring alternative architectures, such as ResNet and VGG-19, may improve accuracy and robustness due to their advanced feature extraction capabilities. Incorporating ensemble methods that combine predictions from multiple models could further enhance diagnostic accuracy. However, it is crucial to note that these models must be trained on pre-processed data to ensure optimal performance.

Future research should focus on expanding the dataset to include a wider variety of skin types, lesions, and imaging conditions to improve generalizability. Implementing techniques such as data augmentation and synthetic image generation could simulate diverse scenarios and strengthen model adaptability. Additionally, integrating explainable AI methods would provide insights into the model's decision-making processes, increasing trust among healthcare professionals. Lastly, conducting longitudinal studies and real-world clinical trials to evaluate the model's performance over time and in diverse healthcare settings would be invaluable for assessing its practical utility and efficacy in improving patient outcomes.

In the future, it may be possible to achieve better results using other deep learning networks like ResNET, VGG-19, and RCNN. Our work was conducted using the VGG-16 model, which is regarded as an effective architecture. Using a faster, more robust model is not enough to diagnose skin cancer reliably from complex images. However, this does not preclude the possibility that other networks may yield superior results. The model with the highest validation score was selected for further testing using a separate test dataset that it had not previously encountered to accurately evaluate the model's performance. In context related to training medical models, we hope there will be wider interest and a greater focus on the importance of data quality, whether in the preprocessing phases or the acquisition phase of images, i.e., the acquisition method and the tools used.

Our research aims to advance general deep learning models, specifically for medical applications. Additionally, we want to improve the model's ability to identify more abnormalities and malignancies and determine whether they are connected to other illnesses or the skin.

References

[1] Alahmadi MD, Alghamdi W (2022) Semi-supervised skin lesion segmentation with coupling CNN and transformer features. IEEE Access 10:122560–122569. https://doi.org/10.1109/ACCESS.2022.3224005

[2] Alzubaidi L, Al-Amidie M, Al-Asadi A et al (2021) Novel transfer learning approach for medical imaging with limited labeled data. Cancers 13(7):1590. https://doi.org/10.3390/cancers13071590

[3] Ashraf R, Afzal S, Rehman AU et al (2020) Region-of-interest based transfer learning assisted framework for skin cancer detection. IEEE Access 8:147858–147871.

[4] Brindha PG, Rajalaxmi R, Kabhilan S et al (2020) Comparative study of svm and cnn in identifying the types of skin cancer. J Crit Rev 7(11):640–643. https://doi.org/10.31838/jcr.07.11.117

[5] Coskun M, YILDIRIM Ö, Ay¸segül U et al (2017) An overview of popular deep learning methods. Eur J Techn (EJT) 7(2):165–176

[6] Fahad NM, Sakib S, Khan Raiaan MA et al (2023) Skinnet-8: an efficient CNN architecture for classifying skin cancer on an imbalanced dataset. In: 2023 International conference on electrical, computer and communication engineering (ECCE), pp 1–6. https://doi.org/10.1109/ECCE57851.2023.10101527

[7] Gao Y, Mosalam KM (2018) Deep transfer learning for image-based structural damage recognition. Comput Aided Civ Infrastruct Eng 33(9):748–768. https://doi.org/10.1111/mice.12363

[8] Giotis I, Molders N, Land S et al (2015) MED-NODE: a computer-assisted melanoma diagnosis system using non-dermoscopic images. Expert Syst Appl 42:6578–6585. https://doi.org/10.1016/j.eswa.2015. 04.034

[9] Hassan E, Elmougy S, Ibraheem MR et al (2023) Enhanced deep learning model for classification of retinal optical coherence tomography images. Sensors 23(12):5393. https://doi.org/10.3390/s23125393

[10] Hassan E, Shams MY, Hikal NA et al (2023) The effect of choosing optimizer algorithms to improve computer vision tasks: a comparative study. Multimedia Tools Appl 82(11):16591–16633. https://doi.org/ 10.1007/s11042-022-13820-0

[11] Hinton GE, Srivastava N, Krizhevsky A et al (2012) Improving neural networks by preventing coadaptation of feature detectors. arXiv:1207.0580

[12] Hoefler T, Alistarh D, Ben-Nun T et al (2021) Sparsity in deep learning: pruning and growth for efficient inference and training in neural networks. J Mach Learn Res 22(241):1–124. https://doi.org/10.1145/ 3578356.3592583

[13] Jain S, Pise N et al (2015) Computer aided melanoma skin cancer detection using image processing. Procedia Comput Sci 48:735–740. https://doi.org/10.1016/j.procs.2015.04.209

[14] Jayalakshmi G, Kumar VS (2019) Performance analysis of convolutional neural network (CNN) based cancerous skin lesion detection system. In: 2019 International conference on computational intelligence in data science (ICCIDS), IEEE, pp 1–6. https://doi.org/10.1109/ICCIDS.2019.8862143

[15] Lafraxo S, Ansari ME, Charfi S (2022) MelaNet: an effective deep learning framework for melanoma detection using dermoscopic images. Multimedia Tools Appl 81(11):16021–16045. https://doi.org/10. 1007/s11042-022-12521-y

[16] Hossain, M. B., & Hoque, K. (2022). Machine learning approaches in Intrusion Detection Systems. International Journal of Science and Research Archive, 7(2), 706-715. https://doi.org/10.30574/ijsra.2022.7.2.0303

[17] Mijwil MM (2021) Skin cancer disease images classification using deep learning solutions. Multimedia Tools Appl 80(17):26255–26271. https://doi.org/10.1007/s11042-021-10952-7

[18] Nawaz M, Masood M, Javed A et al (2021) Melanoma localization and classification through faster region-based convolutional neural network and SVM. Multimedia Tools Appl 80(19):28953–28974. https://doi. org/10.1007/s11042-021-11120-7

[19] Pham TC, Tran CT, Luu MSK et al (2020) Improving binary skin cancer classification based on best model selection method combined with optimizing full connected layers of deep CNN. In: 2020 International conference on multimedia analysis and pattern recognition (MAPR), IEEE, pp 1–6.

[20] Rafi TH, Shubair RM (2021) A scaled-2D CNN for skin cancer diagnosis. In: 2021 IEEE conference on computational intelligence in bioinformatics and computational biology (CIBCB), IEEE, pp 1–6. https:// doi.org/10.1109/CIBCB49929.2021.9562888

[21] Rasel M, Obaidellah UH, Kareem SA (2022) convolutional neural network-based skin lesion classification with variable nonlinear activation functions. IEEE Access 10:83398–83414. https://doi.org/10.1109/ACCESS.2022.3196911

[22] Rotemberg V, Kurtansky N, Betz-Stablein B et al (2021) A patient-centric dataset of images and metadata for identifying melanomas using clinical context. Sci Data 8(34):1–8. https://doi.org/10.34970/2020- ds01

[23] Shahidi Zandi M, Rajabi R (2022) Deep learning based framework for Iranian license plate detection and recognition. Multimedia Tools Appl 81(11):15841–15858. https://doi.org/10.1007/s11042-022-12023-x

[24] Sonsare PM, Gunavathi C (2021) Cascading 1D-convnet bidirectional long short term memory network with modified COCOB optimizer: a novel approach for protein secondary structure prediction. Chaos Solitons Fractals 153:111446. https://doi.org/10.1016/j.chaos.2021.111446

[25] Wu H, Gu X (2015) Towards dropout training for convolutional neural networks. Neural Netw 71:1–10.

[26] Wu H, Chen S, Chen G et al (2022) FAT-Net: feature adaptive transformers for automated skin lesion segmentation. Med Image Anal 76:102327.

[27] I. Iqbal, M. Younus, K. Walayat, M. U. Kakar, and J. Ma, "Automated multi-class classification of skin lesions through deep convolutional neural network with Dermoscopic Images," Computerized Medical Imaging and Graphics, vol. 88, p. 101843, 2021.

[28] Z. Rahman, M. Hossain, M. Islam, M. Hasan and R. Hridhee, "An approach for multiclass skin lesion classification based on ensemble learning", Informatics in Medicine Unlocked, vol. 25, p. 100659, 2021. Available: 10.1016/j.imu.2021.100659.Gong, M., Feng, J., and Xie, Y. (2020). Privacy-enhanced multi-party deep learning. Neural Networks, 121, 484–496.

[29] Tabrizchi, H., Parvizpour, S. & Razmara, J. An Improved VGG Model for Skin Cancer Detection. Neural Process Lett 55, 3715–3732 (2023). https://doi.org/10.1007/s11063-022-10927-1World Health Organization . Global Health Observatory. World Health Organization; Geneva, Switzerland: 2022.

[30] https://www.kaggle.com/datasets/kmader/skin-cancer-mnist-ham10000

[31] Md Masum Billah, Amit Deb Nath, Denesh Das, Tanvir Mahmud and Rashedur Rahman, "Skin Cancer Classification using NasNet", World Journal of Advanced Research and Reviews, 2023, 19(01), 1652-1658.

[32] Rahman, M. A., Bazgir, E., Hossain, S. S., & Maniruzzaman, M. (2024). Skin cancer classification using NASNet. International Journal of Science and Research Archive, 11(1), 775-785.

[33] Rahman, M.S., Hoque, K., Hossain, M.B., Das, D., Wu, T. (2025). Detection of Coal Miner with a Comprehensive Dataset Using Transfer Learning Techniques. In: Kumar, A., Swaroop, A., Shukla, P. (eds) Proceedings of Fourth International Conference on Computing and Communication Networks. ICCCN 2024. Lecture Notes in Networks and Systems, vol 1292. Springer, Singapore. https://doi.org/10.1007/978-981-96-3250-3_41

[34] Podder, P., Bharati, S., Rahman, M. A., & Kose, U. (2021). Transfer learning for classification of brain tumor. In Deep learning for biomedical applications (pp. 315-328). CRC Press.