Intelligent Skin Cancer Detection from Dermoscopic Images with Machine and Deep Learning Approaches

Ali Abdul Karim Jawad and Asaad Abbas Hindi
Department of Geography, University of Basrah, 61004 Basrah, Iraq
pgs.ali.abdulkarim@uobasrah.edu.iq, asaad.hindy @uobasrah.edu.iq

Keywords: Skin Cancer Detection, Deep Learning, Dermoscopic Images, Machine Learning, Early Diagnosis.

Abstract: Melanoma, in particular, is one of the most common and dangerous cancers in the world, and early diagnosis is critical to improving survival rates. A traditional diagnostic method, such as visual examination or dermoscopy, often requires expert intervention, but it can be challenging to distinguish early-stage melanoma from benign lesions. Artificial intelligence (AI), particularly machine learning and deep learning techniques, are applied to dermoscopic images in this study to detect skin cancer more accurately. Results showed that deep learning models were more accurate, more recallable, and had higher F scores than traditional machine learning algorithms. We compare the performance of Logistic Regression, K-Nearest Neighbors, and advanced deep learning architectures such as Xception, VGG16, and ResNet50 on two public datasets containing dermatoscopic images, HAM10000 and PH2. As a result of the study, deep learning models, especially when fine-tuned, offer significant improvements in detecting skin lesions, including melanoma, allowing for early detection.

1 INTRODUCTION

The early detection of skin cancer is critical to improving the patient's outcome since it is one of the most common types of cancer worldwide. Dermatologists rely heavily on dermoscopic images to diagnose skin cancer, which is diagnosed through examining lesions and moles. In order to detect disease more efficiently, accurately, and scalable, new detection methods are needed to cope with the growing number of patients and the complexity of some cases. AI advances, specifically machine learning (ML) and deep learning (DL), have demonstrated promise in automating and improving the detection of skin cancer from dermoscopic images. The use of machine learning models can help us differentiate benign lesions from malignant lesions [1]. Using multiple layers of neural networks, deep learning is capable of extracting complex features from raw image data automatically.

An individual's skin covers their entire exterior, making it the major organ of the body. Among the human body's first lines of defence is its skin [2]. Protection against external environmental factors, regulation of body temperature, immunity against many diseases, and enhancement of beauty are some of its functions [3]. Sunlight exposes the skin to harmful ultraviolet rays, but it also produces vitamin D, which is imperative to human health. There are different skin colours (body pigmentation) and moisturizing levels (oily versus dry) depending on where you live in the world [4], [5]. In the presence of sunlight for a long time or when exposure to ultraviolet rays is prolonged, malignant skin diseases are more likely to develop and pigmentation declines. Detecting melanoma early is essential for its survival. As the cancer cells are similar to other skin cells, they are not detected when the disease is in its early stages. Malignant melanomas are caused by abnormal cancerous cells dividing rapidly under the surface of the skin and penetrating deep into the tissue [6].

Treatment of melanoma is challenging when the cancer spreads throughout the body. As a result, early diagnosis is crucial for saving lives. Dermatoscopic images reveal a wide variety of skin conditions. The two most common skin diseases are melanocytosis and nonmelanocytosis [7]. Skin colours vary from person to person, and air bubbles, hair, and artefacts all complicate matters. There are also many similarities between early-stage skin diseases, especially when they are in their early stages. Inadequate medical resources result in 420 million people suffering from skin diseases globally. A lack of early diagnosis also reduces the quality of life and social advancement in developing countries due to high treatment costs [8].

More than 70,000 cancer cases were reported in the United States in 2017, according to the American Cancer Society. A total of 100,350 cases of melanoma were diagnosed in that country in 2020, with 60,190 men and 40,160 women being diagnosed [9]. Early-stage melanoma and benign lesions are difficult to distinguish due to their similar appearance during the early stages. Computer-assisted diagnosis (AI) is, therefore, essential for early diagnosis. Using computer-aided diagnosis, physicians and experts can diagnose diseases earlier and more efficiently. In this study, we used artificial intelligence techniques to detect skin lesions at an early stage. Using artificial intelligence techniques, specific samples are trained to solve particular problems using many layers and complex neurons. In recent years, deep learning techniques have been used for diagnosing problems that cannot be solved using machine learning [10], [11].

2 LITERATURE REVIEW

The deadly disease of melanomas is responsible for thousands of fatalities worldwide as a result of skin cancer. Throughout history, much has been done to detect and save the victims at the earliest stages of the disease. The first step is to introduce handcrafted features-based techniques for detecting it. In [12], edge-color histograms and local binary patterns (LBP) were used for melanoma lesion detection by using edge and color histograms alongside LBP.

Combining two different techniques to incorporate features of local and global images. The Author [13] proposed two different approaches for localizing melanoma skin cancer. Global image features such as colour, size, shape, texture, and shape of the melanoma lesion are determined using the Laplacian pyramid and a gradient histogram. On the basis of the data obtained, a binary classifier is trained. Bag-of-functions (BoF) classifiers are used to recognize local features in images [14]. Compared to texture features, colour features perform better in this study. Due to the variations in sizes, textures, shapes, and colours of melanoma moles, hand-coded feature-based approaches are often ineffective at detecting skin lesions. According to the Author [15], melanoma lesions can be classified according to the following methods. We calculated the scale-invariant feature transforms (SIFTs) and histograms of oriented gradients (HOGs) of representative features. Melanoma moles were classified using support vector machines (SVMs) and k-nearest neighbours (KNNs) [16].

Several image-processing techniques were used by the Author [17] to identify lesions on the skin. This approach involved obtaining a medical history before diagnosing. An expert examines the attributes and provides a set of them. Descriptors are derived automatically, and those observable are chosen as a final decision. Segmentation-based techniques are well suited to situations in which the chrominance distribution is uniform and the illumination of an image changes little. Still, in real-time applications, these lighting and illumination variations are usually inevitable. Using skin lesion recognition and segmentation, the Author [18] presented his approach. The input images were then preprocessed, and lesion segmentation was performed using seed region growing and graph-cut approaches. SURF and HOG descriptors were used to compute the features from the segmented lesions. In order to categorize lesions based on the extracted features, an SVM classifier was trained using the extracted features. Although the method is better at segmenting lesions and classifying them, it is computationally expensive.

Due to ML advancements, DL methods are becoming increasingly popular since they provide high accuracy. Radiologists and dermatologists are therefore exploring the use of these techniques for medical imaging as well, allowing them to detect fatal diseases sooner. As a result of deep learning approaches like CNN, it has proven possible to classify skin lesions with good accuracy. The Author [19] suggests that instead of training the CNN from scratch to detect melanoma lesions, pretrained ConvNets are used. The Author has developed a DL-based model for segmenting and categorizing melanomas. Feature calculation from the area of interest was performed using the area-average pooling (RAPooling) method in the first step. A classification framework was proposed using the lesion regions identified in the first step, as well as a segmentation process using RAPooling to guide the classification process. A RankOpt classifier was used for the final step of melanoma classification. Although it performs well for segmenting and categorizing skin moles, it is inefficient economically [20].

Key points from the computed key points were used to train SVMs, KNNs, and Random Forests (RF) to identify melanoma from benign moles. While the method is robust, it needs further improvement to improve its performance in detecting melanoma. In his paper [21], the Author introduced a method of localizing melanoma lesions automatically. For the first step, feature correlations were calculated using a linearly independent and linear prediction method (LP). The second step involved computing a representative set of keypoint vectors using RGB colour histograms and SIFT. The input samples were classified into melanoma and nonmelanoma categories based on the computed key points by SVM classifiers. While computationally efficient, this method may not be appropriate for samples whose chrominance changes are intense.

A variety of factors have hindered previously conducted studies, including hair, air bubbles, artefacts, and light reflections. Furthermore, similar characteristics between diseases pose a major challenge when diagnosing and separating them. There have been no differences between this study and previous studies in terms of issues addressed. The images were cleaned with average and Laplacian filters in order to remove artefacts, air bubbles, skin lines, and reflections. A high degree of accuracy can also be achieved when removing hair from images using dull razor technology. Several diseases share similar features, so these features were extracted from each image and combined into one vector using three hybrid algorithms. Consequently, each disease has its characteristics. A deep and representative characteristic of each disease was extracted using ResNet-50 and AlexNet.

3 PROPOSED METHODOLOGY

3.1 Methods and Materials

In this study, not only melanoma and nevi are distinguished from pigmented lesions that commonly appear in clinics, but also no melanocytic pigments. With the help of two publicly available datasets, pre-trained CNNs are used to classify skin cancer images. First, we have the HAM1000 dataset [22]. This dataset is a benchmark for academic machine learning.

In addition, there is the PH2 dataset [23]. In this database, 200 dermatoscopic images of lesions with melanocytic cells are included. This database contains 80 samples of atypical nevi (Atyp NV) and 40 samples of melanoma (Mel), which we used to expand our diagnostic categories. A 768 x 560-pixel resolution RGB image in 8-bit colour. This paper contains 8 diagnostic classes in total.

Researchers have developed CNN in recent years to solve computer vision problems with greater accuracy. Our paper uses 4 deep CNN trained on ImageNet [24] using Tensor Flow [24], a framework developed by Google for deep learning [25].

3.2 Artificial Intelligence

In addition to planning, problem-solving, understanding natural language, and learning, machines can also perform some other tasks characteristic of human intelligence thanks to artificial intelligence. Our discussion of artificial intelligence will be divided into two main categories. There is also a discussion of pre-trained networks and transfer learning (TL).

3.2.1 Machine Learning

In machine learning, useful information is extracted from data by separate algorithms. After training, machine learning extracts salient features and natural patterns that are used to assign attributes to samples or cluster them based on cluster identifications. Models can be used to predict data that was previously unknown based on this information. We will outline the main ML models in the following sections:

- Decision Trees. DTs can be used for categorical as well as numerical variables in ML since they do not require assumptions regarding data distribution or classifier structure. Consequently, they are suitable for classification, regression, and multi-output analysis. Datasets of any size and complexity can be classified effectively and efficiently using them. Decision trees are assembled into random forests (RFs). When DTs are multiplied, the variance of each DT is reduced, resulting in RFs that are more robust and generalizable.
- Support Vector Machines. An SVM can be used to perform linear and nonlinear classifications (kernel methods) and regressions, as well as to detect outliers. In spite of this, it is important to keep in mind that SVMs perform best when the dataset is complex. As a result of using SVMs for classification, hyperplanes are constructed. As a result of every hyperplane, two distinct classes of features can be separated and differentiated from one another.
- K-Nearest Neighbors. Using KNN, new data is classified based on how similar it is to the nearest labelled data. As soon as a KNN classifier's parameters, namely the number of

nearest neighbours and the distance metric, are chosen (Euclidean and Manhattan distances are most relevant), a majority vote is used to assign a label to the new data set. Typically, K is set between 3 and 10 to prevent overfitting and underfitting.

- Artificial Neural Networks. ANNs were introduced as an alternative method of solving complex problems based on neuronal connections studies. Despite early attempts to mimic the human brain and synaptic connections between neurons, only a small amount of its mechanisms could be understood. The ANNs are implemented using simpler and more ordered architectures consisting of neurons (or nodes), which correspond to synaptic connections and layering. Using neural networks, data is analyzed based on the relevance of representations to the problem at hand. The process of learning occurs when neurons in a neural network change their connections between each other as they learn (e.g., as they gain experience). Methods for training neural networks are called learning paradigms.

3.2.2 Deep Learning

Due to the fact that it mimics the human brain in its architecture, deep learning has proven to be the most successful solution in DL. Computer vision algorithms based on CNNs gained popularity in the early 2000s. Developed on the basis of neural networks in the visual cortex, they have excellent results. Each CNN layer has a distinct function, and there are several types. Some of them can indeed be trained, while others only need to implement established functions. CNN architectures commonly use the following layers:

- Convolutional layers. In addition to learning local patterns, convolutional layers can learn spatial hierarchies of patterns as well as translation-invariant patterns. With the increasing depth of the network, CNN is able to learn increasingly complex visual concepts efficiently. In addition to being translation-invariant, convolutional layers can learn spatial hierarchies of patterns, and these two properties allow them to learn local patterns efficiently. Layers that perform convolution over input images produce feature maps to be sent to the following layers as a result of each layer's convolution operation.
- Normalization layers. There is no trainable parameter in the normalization function, and forward propagation is the only available option. In recent years, these layers have become less popular.
- Regularization layer. During training, these layers ignore a portion of neurons randomly in order to reduce overfitting. Dropout is the most common regularization technique.
- Pooling layers. By pooling layers, feature maps can be subsampled while maintaining the main information presented therein, thus reducing model parameters and computation costs. Feature maps are convolutioned using average pooling and max pooling, like convolutional layers, but no trainable parameters are included.
- Fully connected layers. There is a connection between a neuron and an activation function in another layer. Using a CNN classifier, the classification results are determined after a fully connected layer (FC).

3.2.3 Transfer Learning and Pre-Trained Modelsgoogle's Inception V3

Based on Google's Inception v3 architecture, we were able to categorize 8 diagnostic categories by connecting two fully connected layers, one average pooling layer, and a softmax layer [26]. To make this model compatible with input images, they were all resized to (299, 299). Based on a decay and momentum of 0.9 and a learning rate of 0.0007, the stochastic gradient descent (SGD) algorithm was used to optimize:

- InceptionResNet v2. Using our dataset, the Inception ResNet v2 architecture [27] was retrained by replacing the top layer with a global-average pooling layer, followed by a fully connected layer, followed by a softmax layer, allowing the classification of eight diagnostic categories.
- ResNet 152. Using our dataset, ResNet 152 architecture [28] was retrained by replacing top layers with one average-pooling, one fully connected, and finally, one softmax layer, which allowed the organization of 8 diagnostic categories.
- DenseNet 201. An eight-category classification system was built using DenseNet 201 architecture [29] and a global average pooling layer plus an SL.

3.3 Model Performance Evaluation

The effectiveness of the proposed model was assessed through a comprehensive performance evaluation using a confusion matrix framework. Prior to model development, the dataset was partitioned into distinct training and testing sets to ensure unbiased evaluation. After model training, performance assessment was conducted exclusively on the test set to measure generalization capability.

Following the methodology presented in [30], the skin cancer detection effectiveness of the proposed model was evaluated using four key performance metrics derived from the confusion matrix: accuracy, precision, recall, and F1-score. These metrics collectively provide a comprehensive assessment of the model's classification performance across all test instances.

4 RESULT ANALYSIS AND DISCUSSION

The five machine learning algorithms presented in Figure 1 – LR, LDA, KNN, Classification, and RT, and GNB – are compared based on four metrics: accuracy, precision, recall, and F-score. True positives are more likely to be detected using KNN and CART, particularly in terms of recall and F-score. There is, however, a lower recall for LR and GNB, which indicates that they miss more positive results. Each algorithm for detecting skin cancer is ranked according to its strengths and weaknesses in the chart below.

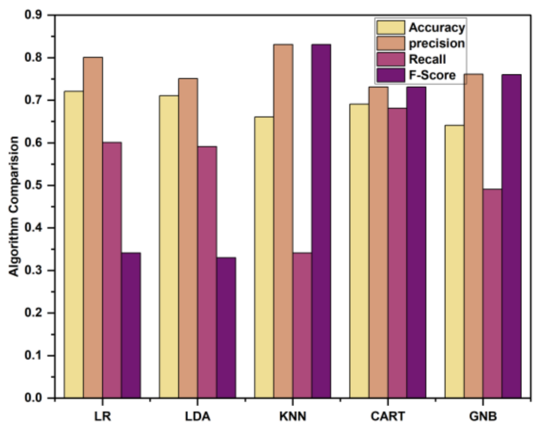

Figure 1: Kaggle-based machine learning models for skin cancer prediction.

In Figure 2, three digital learning models are compared with the frozen base approach to detect skin cancers (Xception, VGG16, and ResNet50). Xception, ResNet50, and VGG16 achieve the best accuracy and F-score. There is good true positive detection in every model, but Xception performs the best overall in terms of recall, indicating it is the most reliable.

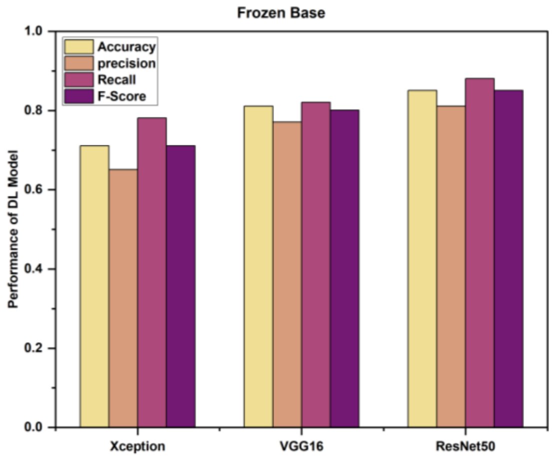

Figure 2: Performance of DL Model versus Frozen Base.

Figure 3 compares the performance of three deep learning models, VGG16 and ResNet50-on skin cancer detection, using a "Fine-Tuning" approach, in which the pre-trained models are more extensively modified to improve results. The Xception model performs better in this scenario compared to the "Frozen Base" scenario, followed closely by ResNet50 and VGG16. There is strong precision and recall across all three models, indicating their ability to detect true positives effectively. As a result of this fine-tuning configuration, Xception remains the most effective model, demonstrating increased performance in all metrics.

In Figure 4, accuracy values for a deep learning model in both frozen and tuned configurations are shown over 25 epochs. It is evident that Tuned Accuracy and Tuned Validation Accuracy are improving steadily, especially after the 10th epoch, surpassing the Frozen Accuracy and Frozen Validation Accuracy, which remain constant. Comparing fine-tuning to using frozen layers highlights the benefits of fine-tuning.

For a frozen deep learning model and a tuned deep learning model, Figure 5 shows loss values over 25 epochs. Losses for Tuned and Tuned Validation are steadily decreasing, reaching lower values than for Frozen and Frozen Validation Losses, which fluctuate more. Using fine-tuned layers is more efficient than using frozen layers for convergence and performance.

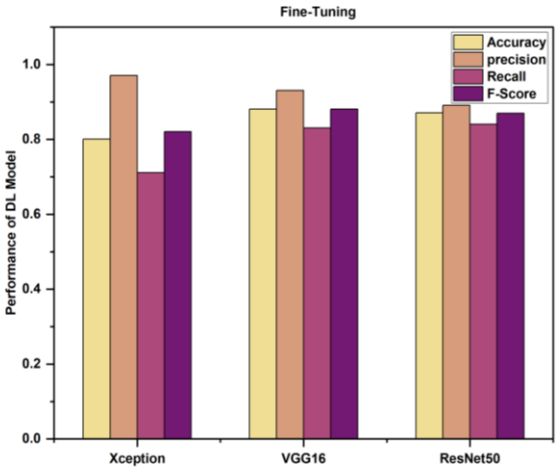

Figure 3: Performance of DL Model versus Fine-Tuning.

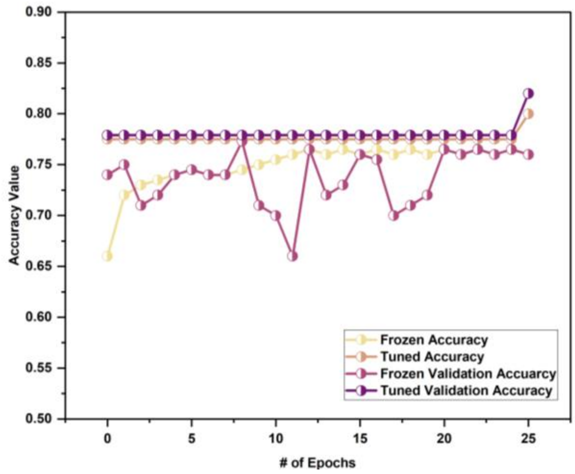

Figure 4: Model accuracy comparison across epochs for frozen and tuned models.

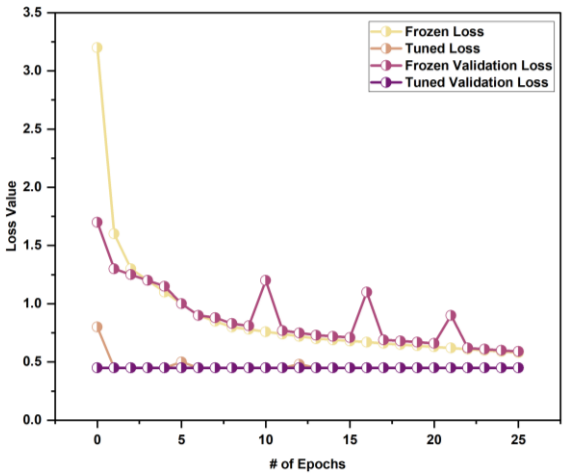

Figure 5: Epoch-by-epoch comparison of frozen and tuned models.

5 CONCLUSIONS

Using machine learning and deep learning, this study demonstrated that skin cancer can be effectively detected in the early stages from dermoscopic images. The comparative analysis clearly showed that pre-trained deep learning models significantly outperformed traditional machine learning approaches, particularly when enhanced through advanced fine-tuning techniques. Among the evaluated models, Xception consistently exhibited superior performance over ResNet50 and VGG16, achieving higher accuracy, precision, and sensitivity across all tested metrics. This highlights that fine-tuning deep learning models results in more precise and dependable skin cancer detection, positioning these tools as highly valuable resources for dermatologists, clinicians, and patients, enabling faster and more accurate diagnoses. Future research could further explore how variations in skin types, pigmentation, demographic factors, and environmental influences impact the reliability and efficacy of these AI models when integrated into real-time clinical diagnostic systems.

6 REFERENCES

[1] P. Rani, S. Verma, S. P. Yadav, B. K. Rai, M. S. Naruka, and D. Kumar, "Simulation of the lightweight blockchain technique based on privacy and security for healthcare data for the cloud system," Int. J. E-Health Med. Commun. (IJEHMC), vol. 13, no. 4, pp. 1–15, 2022.

[2] S. Rosenbaum, "Data Governance and Stewardship: Designing Data Stewardship Entities and Advancing Data Access," Health Serv. Res., vol. 45, no. 5p2, pp. 1442–1455, Oct. 2010, doi: 10.1111/j.1475-6773.2010.01140.x.

[3] E. Proksch, J. M. Brandner, and J. Jensen, "The skin: an indispensable barrier," Exp. Dermatol., vol. 17, no. 12, pp. 1063–1072, Dec. 2008, doi: 10.1111/j.1600-0625.2008.00786.x.

[4] T. H. Sakuma and H. I. Maibach, "Oily Skin: An Overview," Skin Pharmacol. Physiol., vol. 25, no. 5, pp. 227–235, 2012, doi: 10.1159/000338978.

[5] P. Rani, S. P. Yadav, P. N. Singh, and M. Almusawi, "Real-World Case Studies: Transforming Mental Healthcare With Natural Language Processing," in Demystifying the Role of Natural Language Processing (NLP) in Mental Health, A. Mishra, S. P. Yadav, M. Kumar, S. M. Biju, and G. C. Deka, Eds., IGI Global, 2025, pp. 303–324, doi: 10.4018/979-8-3693-4203-9.ch016.

[6] M. Nasir et al., "Melanoma Detection and Classification using Computerized Analysis of Dermoscopic Systems: A Review," Curr. Med. Imaging Former. Curr. Med. Imaging Rev., vol. 16,

no. 7, pp. 794–822, Sep. 2020, doi: 10.2174/1573405615666191223122401.

[7] P. Rani, K. Ur Rehman, S. P. Yadav, and L. Hussein, "Deep Learning and AI in Behavioral Analysis for Revolutionizing Mental Healthcare," in Demystifying the Role of Natural Language Processing (NLP) in Mental Health, A. Mishra, S. P. Yadav, M. Kumar, S. M. Biju, and G. C. Deka, Eds., IGI Global, 2025, pp. 263–282, doi: 10.4018/979-8-3693-4203-9.ch014.

[8] B. Zhang et al., "Opportunities and Challenges: Classification of Skin Disease Based on Deep Learning," Chin. J. Mech. Eng., vol. 34, no. 1, p. 112, Dec. 2021, doi: 10.1186/s10033-021-00629-5.

[9] C. Herman, "Emerging technologies for the detection of melanoma: achieving better outcomes," Clin. Cosmet. Investig. Dermatol., p. 195, Nov. 2012, doi: 10.2147/CCID.S27902.

[10] Y. LeCun, Y. Bengio, and G. Hinton, "Deep learning," Nature, vol. 521, no. 7553, pp. 436–444, May 2015, doi: 10.1038/nature14539.

[11] G. Ansari, P. Rani, and V. Kumar, "A novel technique of mixed gas identification based on the group method of data handling (GMDH) on time-dependent MOX gas sensor data," in Proc. Int. Conf. Recent Trends Comput. (ICRTC 2022), Springer, 2023, pp. 641–654.

[12] N. C. F. Codella et al., "Deep learning ensembles for melanoma recognition in dermoscopy images," IBM J. Res. Dev., vol. 61, no. 4/5, pp. 5:1–5:15, Jul. 2017, doi: 10.1147/JRD.2017.2708299.

[13] C. Barata, M. Ruela, M. Francisco, T. Mendonca, and J. S. Marques, "Two Systems for the Detection of Melanomas in Dermoscopy Images Using Texture and Color Features," IEEE Syst. J., vol. 8, no. 3, pp. 965–979, Sep. 2014, doi: 10.1109/JSYST.2013.2271540.

[14] P. Rani, D. S. Mohan, S. P. Yadav, G. K. Rajput, and M. A. Farouni, "Sentiment Analysis and Emotional Recognition: Enhancing Therapeutic Interventions," in Demystifying the Role of Natural Language Processing (NLP) in Mental Health, A. Mishra, S. P. Yadav, M. Kumar, S. M. Biju, and G. C. Deka, Eds., IGI Global, 2025, pp. 283–302, doi: 10.4018/979-8-3693-4203-9.ch015.

[15] J. Daghrir, L. Tlig, M. Bouchouicha, and M. Sayadi, "Melanoma skin cancer detection using deep learning and classical machine learning techniques: A hybrid approach," in Proc. 5th Int. Conf. Adv. Technol. Signal Image Process. (ATSIP), Sousse, Tunisia: IEEE, Sep. 2020, pp. 1–5, doi: 10.1109/ATSIP49331.2020.9231544.

[16] P. Rani and R. Sharma, "Intelligent transportation system for internet of vehicles based vehicular networks for smart cities," Comput. Electr. Eng., vol. 105, p. 108543, 2023.

[17] J. Fernandez Alcon et al., "Automatic Imaging System With Decision Support for Inspection of Pigmented Skin Lesions and Melanoma Diagnosis," IEEE J. Sel. Top. Signal Process., vol. 3, no. 1, pp. 14–25, Feb. 2009, doi: 10.1109/JSTSP.2008.2011156.

[18] A. Rehman, M. A. Khan, Z. Mehmood, T. Saba, M. Sardaraz, and M. Rashid, "Microscopic melanoma detection and classification: A framework of pixel-based fusion and multilevel features reduction," Microsc. Res. Tech., vol. 83, no. 4, pp. 410–423, Apr. 2020, doi: 10.1002/jemt.23429.

[19] J. Kawahara, A. BenTaieb, and G. Hamarneh, "Deep features to classify skin lesions," in Proc. IEEE 13th Int. Symp. Biomed. Imaging (ISBI), Prague, Czech Republic: IEEE, Apr. 2016, pp. 1397–1400, doi: 10.1109/ISBI.2016.7493528.

[20] S. Verma et al., "An automated face mask detection system using transfer learning based neural network to preventing viral infection," Expert Syst., p. e13507, 2024.

[21] K. Hu et al., "Classification of melanoma based on feature similarity measurement for codebook learning in the bag-of-features model," Biomed. Signal Process. Control, vol. 51, pp. 200–209, May 2019, doi: 10.1016/j.bspc.2019.02.018.

[22] P. Tschandl, C. Rosendahl, and H. Kittler, "The HAM10000 dataset, a large collection of multi-source dermatoscopic images of common pigmented skin lesions," Sci. Data, vol. 5, no. 1, pp. 1–9, 2018.

[23] T. Mendonça, P. M. Ferreira, J. S. Marques, A. R. Marcal, and J. Rozeira, "PH² - A dermoscopic image database for research and benchmarking," in Proc. 35th Annu. Int. Conf. IEEE Eng. Med. Biol. Soc. (EMBC), 2013, pp. 5437–5440. [Online]. Available: https://ieeexplore.ieee.org/document/6610779/.

[24] O. Russakovsky et al., "ImageNet Large Scale Visual Recognition Challenge," Int. J. Comput. Vis., vol. 115, no. 3, pp. 211–252, Dec. 2015, doi: 10.1007/s11263-015-0816-y.

[25] M. Abadi et al., "TensorFlow: A system for large-scale machine learning," in Proc. 12th USENIX Symp. Operating Systems Design and Implementation (OSDI 16), 2016, pp. 265–283. [Online]. Available: https://www.usenix.org/conference/osdi16/technical-sessions/presentation/abadi.

[26] C. Szegedy, V. Vanhoucke, S. Ioffe, J. Shlens, and Z. Wojna, "Rethinking the inception architecture for computer vision," in Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR), 2016, pp. 2818–2826. [Online]. Available: https://www.cv-foundation.org/openaccess/content_cvpr_2016/html/Szegedy_Rethinking_the_Inception_CVPR_2016_paper.html.

[27] C. Szegedy, S. Ioffe, V. Vanhoucke, and A. Alemi, "Inception-v4, inception-resnet and the impact of residual connections on learning," in Proc. AAAI Conf. Artif. Intell., 2017. [Online]. Available: https://ojs.aaai.org/index.php/aaai/article/view/11231.

[28] K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," in Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR), 2016, pp. 770–778. [Online]. Available: http://openaccess.thecvf.com/content_cvpr_2016/html/He_Deep_Residual_Learning_CVPR_2016_paper.html.

[29] G. Huang, Z. Liu, L. Van Der Maaten, and K. Q. Weinberger, "Densely Connected Convolutional Networks," in Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR), Honolulu, HI, USA, 2017, pp. 4700–4708, doi: 10.1109/CVPR.2017.243.

[30] D. M. W. Powers, "Evaluation: From precision, recall and F-measure to ROC, informedness, markedness and correlation," *J. Mach. Learn. Technol.*, vol. 2, pp. 37–63, 2011, doi: 10.9735/2229-3981.