The American Journal of Medical Sciences and Pharmaceutical Research
ISSN 2689-1026 Volume 08 - 2026

A Robust Skin Cancer Detection and Classification Model Based on Improved Convolutional Neural Networks

Gona Mohammed Hasan
High Institute Health in Kirkuk, Iraq

Received: 26 Feb 2026 | Received Revised Version: 12 Mar 2026 | Accepted: 18 Apr 2026 | Published: 23 May 2026

Volume 08 Issue 05 2026 | DOI: 10.37547/tajmspr/Volume08Issue05-17

Abstract

Skin cancer (SC) is a major health problem that demands early detection and classification for the best chances of survival. Existing computer vision techniques lack the capacity to address the fine-grained variability exhibited by skin lesion features across surfaces. It is important to identify the SC signs early as the incidence and death rates of this disease are rising rapidly, along with the cost of care. Techniques for diagnosing SC often rely on visual analysis or techniques that do not have accurate results; this can be dangerous for the patient. Hence, deep learning (DL) methods have proved helpful for researchers to construct different methods for the early detection of SC. To distinguish SC from melanoma, these techniques used the lesion's color, size, form, symmetry, and other features. Using the HAM10000 dataset, a sizable and varied dataset, this work suggests a novel DL-based SC detection and diagnosis scheme (DL-SCDCM) that employs DL techniques, particularly CNN, to guarantee an accurate but efficient diagnosis. The proposed CNN model was trained before testing, and it demonstrated impressive performance, correctly diagnosing seven distinct skin lesion types with an accuracy of 96.9 %. Furthermore, the results were compared to those of other research that proposed a slightly different approach, and the proposed model outperformed the others in these comparisons.

Keywords: Skin cancer Classification, Medical Image Modification, Deep learning, CNN, Normalization, and Augmentation

Cite This Article: Hasan, G. M. (2026). A Robust Skin Cancer Detection and Classification Model Based on Improved Convolutional Neural Networks. The American Journal of Medical Sciences and Pharmaceutical Research, 8(05), 79–93. https://doi.org/10.37547/tajmspr/Volume08Issue05-17

1. Introduction

Among the cancer types that are on the rise worldwide, skin cancer (SC) comes in at number six [1]. Tissue-forming cells make up the majority of the skin; cancer results when cells or nearby tissues proliferate abnormally or uncontrollably. Some of the variables that lead to the development of cancer include exposure to UV radiation, a family history of cancer, and a weakened immune system [2]. There are two types of abnormal cellular proliferation: benign and malignant; malignant SC is the most deadly kind, even though it is rare [3][4]; it spreads fast all across the body. Neural crest neoplasia gives rise to malignant melanocytes [5]. SC causes about 55,500 deaths annually, or 0.7 % of all cancer-related deaths [6][7]. The incidence of melanoma differs per nation due to racial and cultural diversity; living in low-latitude areas, consuming large amounts of alcohol, eating fatty foods, and having melanocytic or dysplastic naevi, exposure to UV radiation [8], history of melanoma, and fair skin, hair, and eyes, are risk factors. The incidence of melanoma is correlated with the human development index [9]. By facilitating early identification, treatment, and access to healthcare, a rise

in the HDI index reduces mortality rates. Patients with early-stage melanoma have a 5-year survival rate of about 98 % [10]. However, five years following diagnosis, only 20–50 % of patients with advanced-stage melanoma survive [11].

Furthermore, a benign tumor can grow but does not spread to other body parts; a knowledgeable clinician managing a benign skin lesion should be aware of the typical signs and symptoms of a potentially cancerous lesion. If any growth on the skin occurs suspiciously, it is also important [12] to check up. Skin growths that are not cancerous include seborrhoeic keratoses, pyrogenic granulomas, etc. When cutaneous lesions [13] are diagnosed early, therapy can start early, and complete healing will occur. The first step in diagnosis is the dermoscopic examination, which is a non-invasive way of inspecting a skin lesion's colour and shape at roughly 10–20 times magnification. Automatic SC detection [14] can speed up diagnosing and therapy, giving time for treatment to start quickly; in addition, it helps to determine if the skin lesion is cancerous or not and if it should be removed. The extraction method is less intrusive, although the procedure is likely still uncomfortable, if not dangerous, in some respects.

In the past couple of decades, there has been much interest in AI, along with active research in its medical applications, a branch of computer-assisted technology. SC remains the only form of cancer whose incidence is increasing globally. It is also the only form of cancer that has a consecutively rising death rate, even though SC is also the only form of cancer whose treatment is the least expensive. It is not only an easily treatable form of cancer; its traditional form of diagnosis, which is merely looking at the skin, has been shown to have zero accuracy. So, the traditional methods of diagnosis have proven to be dangerously ineffective, which is why there is a phenomenon of a large gap between diagnosis and treatment. The research gap has been easily filled with Deep Learning, which specialises in Artificial Intelligence and is the most promising in detecting early SC found at its lesions. This study aims to bridge the gap in SC diagnosis and detection by proposing a new Deep Learning model called DL-SCDCM that focuses on SC detection and classification. It is a model based on the HAM10000 and uses CNN; its target is to achieve a reliable, efficient, and accurate detection of melanoma and other skin lesions.

1.1 Research Contributions

This study has several contributions to accurate SC detection and classification; these include:

i. Investigation of the relevant state-of-the-art publications to define the research gap.

ii. Analyzing the HAM10000 dataset and preparing the image before the model-building stage, utilizing several stages like data normalization, data augmentation, and image resizing.

iii. Training the proposed model utilizing an improved CNN to accurately classify SC.

iv. Benchmak the achieved results with relavant satae-of-the-arts.

2. Literature Review

The recent studies in DL-based SC detection are reviewed in this section of the paper. SCaLiNG was presented by Attallah et al., [15] for the extraction of features for the categorization of SC. SCaLiNG achieved a more detailed feature representation by merging features from several CNNs and processing images through GW sub-bands. The model's performance is further optimized through a feature selection stage. The efficacy of SCaLiNG is demonstrated by its accuracy of 0.9170, which surpasses that of conventional single-CNN models. The combination of DL feature fusion with ELM has been presented by Afza et al., [15] as a novel skin lesions classification method. The approach outperformed current methods in terms of performance and computing efficiency, with superior accuracy of 93.40% on the HAM10000 dataset and 94.36% on the ISIC2018 dataset. Several DL models were merged with an "Information-Theoretic Fusion strategy" by Akram et al., [16] to improve and simplify feature selection; noise removal was achieved with an entropy-driven binary Bat selection algorithm. The feature set was optimized by fusion and refinement. The method demonstrated effectiveness when evaluated on different datasets, proving its ability to generate unique and informative feature representations.

In the work of Bibi et al., [17] to detect multiclass SCs, a deep learning system has been developed comprising image pre-processing, feature extraction, selection, and classification. The quality of the image is enhanced using a new luminance-based image contrast enhancement, and features are extracted using modified versions of DarkNet-53 and DenseNet-201, which have been genetically tuned. A serial harmonic mean method is

applied to feature fusion, and irrelevant features are eliminated using Rényi Entropy and marine predator optimisation. Ozdemir and Pacal proposed a hybrid model, which is intended for mobile deployment to tackle data imbalance and model complexity. In the initial phases, the model captures critical features; in the latter phases, it improves sensitivity by concentrating on crucial regions for diagnosis [18]. Dillshad et al.,[19] have further advanced the DL method by multiclass SC classification. The methodology includes innovative data augmentation, suppressing any other data augmentation technique, as well as hybrid contrast enhancement for pre-processing. Transfer learning for feature extraction is performed from MobileNetV2 and NasNet Mobile, and the features are then fused using a dual-threshold serial technique. The variance-controlled Marine Predator algorithm is used to optimise feature selection, and the classification is performed using machine learning, providing an accurate and effective solution.

To categorize eight forms of SC, Chopade et al.,[20] presented SNC_Net that merged handmade characteristics and DL; CNN was utilized for training and validation on the ISIC 2019 dataset. SNC_Net demonstrated its efficacy in the detection of SC by outperforming recent classifiers and a number of baseline models, such as EfficientNetB0 and ResNet-101. DVFNet, a deep learning method for detecting SC from dermoscopy images, was presented by UdriȘtoiu.,[21]. The technique combines VGG19 and Histogram of Oriented Gradients (HOG) to extract information and improve image quality through anisotropic diffusion. Class imbalance was resolved in the ISIC 2019 dataset using SMOTE Tomek. Damaged skin patches are highlighted by segmentation, and a CNN uses a feature vector map created from HOG and VGG19 features to carry out multiclass classification.

Chanda et al.,[22] presented DCENSnet, an ensemble of three DCNNs with specially designed dropout layers to improve feature learning and get the best possible bias–variance trade-off. The model performed quite well on the HAM10000 dataset, exhibiting great precision, recall, F1 score, and AUC for every class, as well as a mean accuracy of 99.53 %. In an effort to improve diagnostic accuracy, this approach performed better than the recent networks and exhibits great promise for trustworthy detection and classification of skin lesions.

Brancaccio et al., [23] stress that working with human specialists yields the greatest results for AI-based diagnostic tools. Although the majority of research evaluates AI's performance in controlled settings, its practical clinical efficacy and implementation issues are yet unknown. The paper examines the possible advantages and drawbacks of AI for dermatologists, general practitioners, and consumers. Pacal et al.used HSW-MSA to improve the Swin Transformer architecture's ability to manage issues of overlapping in SC regions and handle long-range dependencies. To attain better accuracy, quicker training, and better parameter optimization, they also replaced the traditional MLP with a SwiGLU-based MLP, which is a refined form of the Gated Linear Unit (GLU).

3. The proposed Methodology

Deep learning (DL) methods can learn intricate features and patterns and establish connections between data; hence, they are selected in this work for the task of detecting and diagnosing SC in this study; this allows the knowledge to be applied to previously unseen data to produce precise results. The selected model (CNN) was trained and tested on the HAM10000 dataset. Figure 1 shows the specifics of the procedures involved in creating and implementing the suggested DL-SCDCM model.

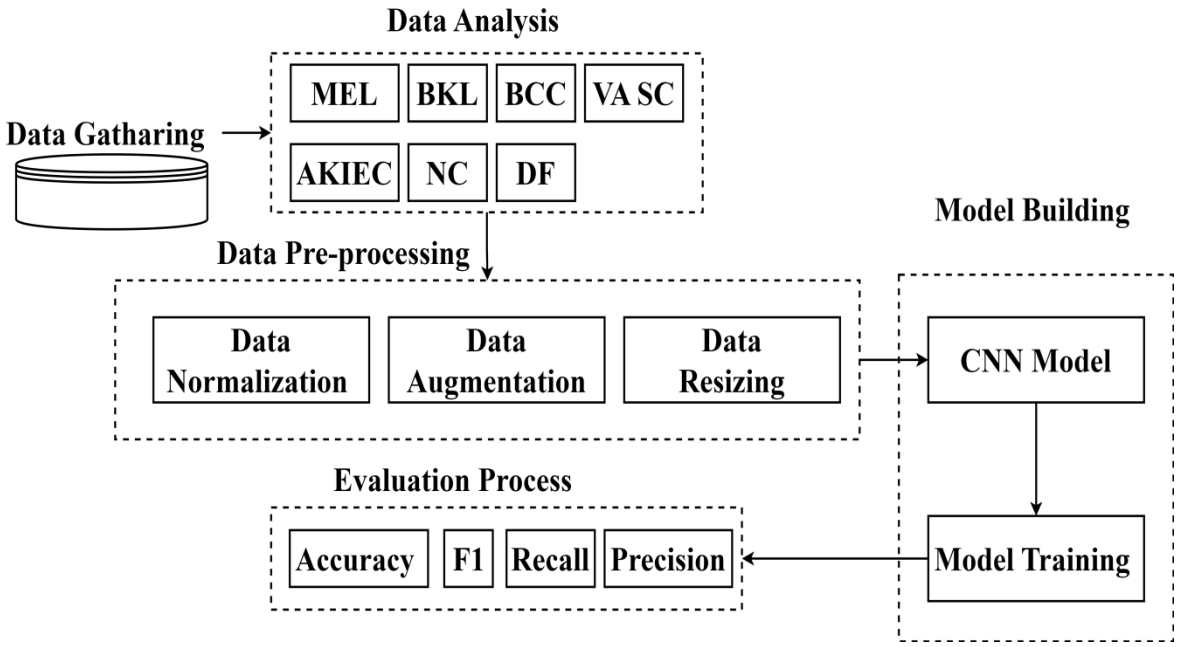

Figure 1. The proposed DL-SCDCM flowchart

The HAM1000 dataset, which includes seven distinct classifications of SC (MEL, BKL, BCC, VASC, AKIEC, NV, and DF), is first obtained; data normalization was the first of several steps in the data pre-processing process that was carried out before feeding the collected data into the suggested DL-SCDCM [12] (See equation no:1); it involves standardization and scaling of values in the dataset into a uniform scale that ranges between 0 and 1; it ensures that data is not biased and suitable for use in the scheme. The second process is data augmentation [14]; data augmentation involves horizontal flipping (Equations 2 and 3), vertical flipping (Equations 4 and 5), and rotation (Equation 6) of the images in the dataset to create multiple variations of the SC lesions and to increase the image quality. Data resizing is the last process [14] (See equation no. 7); it involves changing the size and dimensions of images in the dataset to improve compatibility with the proposed scheme requirements. Following the data pre-processing step, the processed dataset is fed into the suggested DL-SCDCM, which is then used to train and learn features of different types of SC. This will allow the model to identify these features during the testing phase. The suggested DL-SCDCM model's classification performance is evaluated during the testing phase by presenting the trained DL-SCDCM with new, unidentified datasets for classification. The assessment considers a number of variables, including F1 score, accuracy, recall, and precision.

$$\text{Normalized\_pixel\_value} = \frac{\text{original\_pixel\_value}}{255} \quad (1)$$

$$\begin{bmatrix} p' \\ q' \\ 1 \end{bmatrix} = \begin{bmatrix} -1 & 0 & 0 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{bmatrix} \times \begin{bmatrix} p \\ q \\ 1 \end{bmatrix} \quad (2)$$

$$p' = -p, q' = q \quad (3)$$

$$\begin{bmatrix} p' \\ q' \\ 1 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & -1 & 1 \\ 0 & 0 & 1 \end{bmatrix} \times \begin{bmatrix} p \\ q \\ 1 \end{bmatrix} \quad (4)$$

$$p' = p, q' = -q \quad (5)$$

$$R = \begin{bmatrix} \cos(q) & \sin(q) & 0 \\ -\sin(q) & \cos(q) & 0 \\ 0 & 0 & 1 \end{bmatrix} \quad (6)$$

$$\text{image resized} = \text{resize}(\text{image}, (224, 224)) \quad (7)$$

3.1. Dataset

To identify the particular features of each SC category, the classifier must first be trained on a suitable dataset. The dataset HAM10000, which started out with 10,000 images, has been chosen for this analysis [12]. The HAM10000 dataset contains 10,000 high-resolution SC images, the preparation of which is credited to the Department of Dermatology, Medical University of Vienna. The dataset now includes a total of 10,015 images of pigmented SC tissues, 15 of which were added to the original collection. Seven distinct forms of pigmented SC are covered by the images in the dataset, which were gathered using several dermoscopes. Each image in the dataset additionally displays metadata, such as the patient's age, gender, and annotation pertaining to SC type. SC can be categorized into seven classes: "vascular lesions (VASC), benign keratosis-like lesions (BKL); actinic keratoses (AKIEC); melanoma (MEL), melanocytic nevi (NV), dermatofibroma (DF), and basal cell carcinoma (BCC). In Table 1, these seven forms of SC are described.

Table 1. The skin lesion types with possible risk factors and treatment priority.

Type of Skin Lesion	Description	Risk Factors	Common Locations	Treatment Priority	References
MEL (Melanoma)	-Has potential to metastasize, which makes it dangerous. -Manifests on the face or trunk of males.	Excessive exposure to the sun, existing moles	Any body part	High	[3,4]
NV (Melanocytic Nevus)	0-Is benign and not dangerous.	Genetics, Excessive exposure to the sun	Any body part	Low	[8]
BCC (Basal Cell Carcinoma)	Very common. -Manifests as a flesh-colored growth, pearl-shaped bump, or pinkish skin patch.	Having fair colored skin, Excessive exposure to the sun, and indoor tanning	Any body part, including the head, neck, and arms	High	[3,4]
AKIEC (Actinic Keratosis / Bowen's Disease)	-It is linked to sun-damaged skin.	Excessive exposure to the sun	Exposed skin, head, neck, hands, forearms	Moderate	[11,4]

BKL (Benign Keratosis)	Includes seborrheic keratoses, lichen planus-like keratoses, and solar lentigines.	Excessive exposure to the sun	Any body part	Low	[20]
DF (Dermatofibroma)	-Benign and not dangerous -It is characterized by firm nodules.	Unknown	Legs, arms	Low	[21]
VASC (Vascular Lesion)	-It is related to a blood vessel abnormality. -Includes angiomas and angiokeratomas.	Unknown	Any body part	Low	[18]

Figure 2 presents 16 randomly selected images from the HAM10000 dataset.

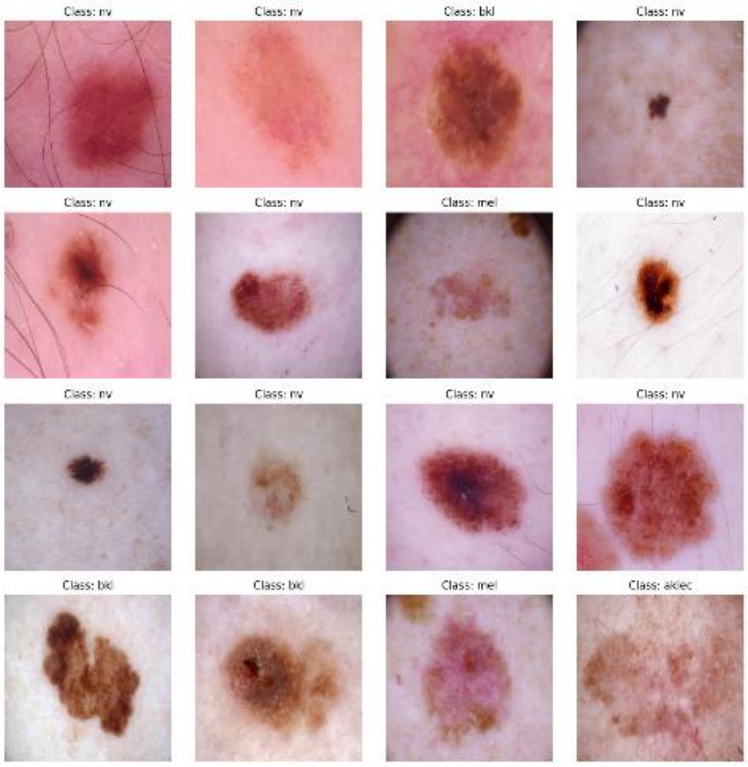

Figure 2. Randomly selected skin cancer images from the HAM10000 dataset [12]

Additionally, a visual depiction of the SC class distribution within the dataset may be obtained, which aids in determining if the dataset is balanced or unbalanced and whether the class exhibits dominance over other classes in terms of image allocation. Additionally, it can assist in identifying biases in the dataset that may impact the fairness of the classification model's performance. Determining the data augmentation approach can also be aided by visual representations of class distribution. The distribution of the seven groups of cutaneous legion types in the dataset is depicted in Figure 3.

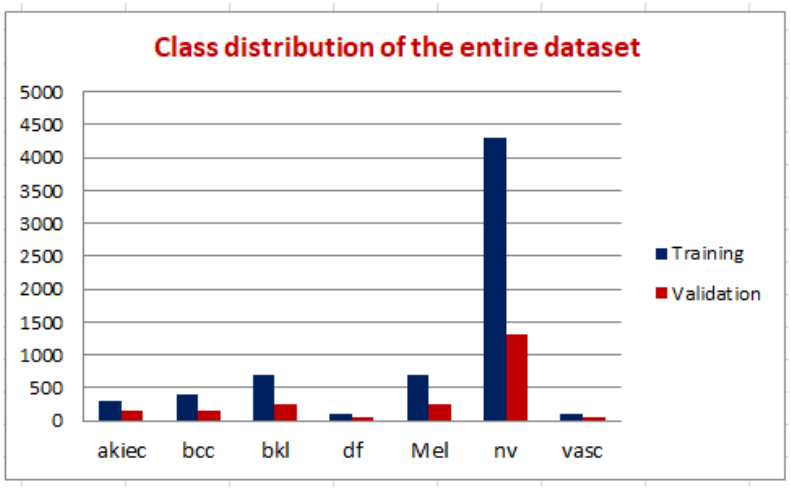

Figure 3. Class distribution of skin cancer types in the HAM10000 dataset

A bar chart illustrating each class and the number of allocated images in the training and testing datasets is produced in Figure 3. As can be seen from Figure 3, there are initially significantly more images in the training dataset than in the testing dataset across all seven classes. However, with almost 4000 images in the training dataset, it is evident that the NV class has the most images, much outnumbering all other classes. The VASC skin lesion class is the least common, followed by the MEL and BKL classes, each of which has about 1000 training images.

2.2 Data Pre-Processing

Image pre-processing is a crucial step in data preparation for modelling; it makes the images better by removing distortions and enhancing their features, which makes it simpler for the model to recognize and classify the features. In addition to guaranteeing that the images in the dataset are of excellent quality, data pre-processing aids in the identification of significant features that will be discovered through analysis and the development of correlations between the features. As was previously noted, data pre-processing includes data resizing, data augmentation, and data standardization.

Data augmentation on the HAM10000 dataset was achieved via random transformations of the training dataset's images, such as rotation_range, width_shift, and height_shift, which represent random rotations and translations, respectively, and shear_range and zoom_range, which represent random shearing transformations and random zooming [2]. The next process is data normalization, which involves data conversion into values between 0 and 1; this process helps prevent one feature from controlling the learning process, thereby preventing feature dominance. Furthermore, data standardization improves the model's performance and encourages convergence. Next comes the data scaling stage, which allocates the resolution of the enhanced dataset as 28 x 28 pixels.

In CNN models, data of a fixed size is required so that they can be compatible and can use batch processing effectively. Data resizing ensures uniformity in the size of all images fed to the CNN model. In the end, 80 percent of the images from the pre-processed dataset were used for the training of the CNN model, while 20 percent were used for testing of the model. This is to ensure that there is enough data to train the model and maintain enough data to evaluate its performance.

2.3 Proposed DL-SCDCM Model

CNNs [13] were selected in this work for skin lesion diagnosis; CNNs are renowned for their ability to recognize and classify images. The two main layers in the CNN are the convolutional and the pooling layers; however, their architectures may differ depending on the task at hand. Each CNN layer performs an important role in extracting hierarchical input image features. The CNN employed in this study had two Conv2D layers, which aid in the extraction of spatial attributes through convolutional operations. The convolutional layers of the model contribute non-linearity by using 3 x 3 kernels and the ReLU activation function. The 'padding' parameter was set to 'same' to ensure that the feature maps had consistent spatial dimensions. Next are the MaxPooling2D layers, which aid in preserving the important features while reducing the feature maps' spatial dimensions. By downsampling the feature maps, the pooling procedure reduces computing cost and

attempts to achieve translation invariance. To achieve a hierarchical structure that enables the suggested model to capture complex patterns at various sizes and promotes efficient feature learning, the Conv2D-MaxPooling2D layers sequence is repeatedly performed to increase the number of filters.

To prepare the data for the fully linked layers, the Flatten layer in the suggested model helps convert the 3D spatial information into a 1D vector. These layers, represented by Dense, introduce global dependencies and classify data using the learnt features. Two dense layers, one with 32 neurons and the other with 512 neurons, are part of the proposed paradigm. Utilizing the ReLU activation function, both layers are coded. The Dense layer is the last and contains 7 neurons, encoded with the Softmax activation function that enables multi-class classification by generating probability distributions for each of the 7 classes.

CNNs are proficient since they learn their own features from raw pixel values; hence, no manual feature engineering is necessary. CNNs are very useful for imaging tasks, like the classification of SC. The complexity and sophistication of the proposed model, which contains a large number of convolutional layers and pooling layers, along with appropriate classifiers, will capture the detailed structural or texture information regarding melanoma dermatoscopic images for accurate and robust classification of SC. The layers of the proposed CNN model architecture are specifically as follows, as shown in Figure 4: Features are extracted using convolutional and pooling layers, classification is done using fully connected layers, and non-linearity is added to the model using ReLU activation functions. The softmax activator in the output layer produces class probabilities for multiple classifications.

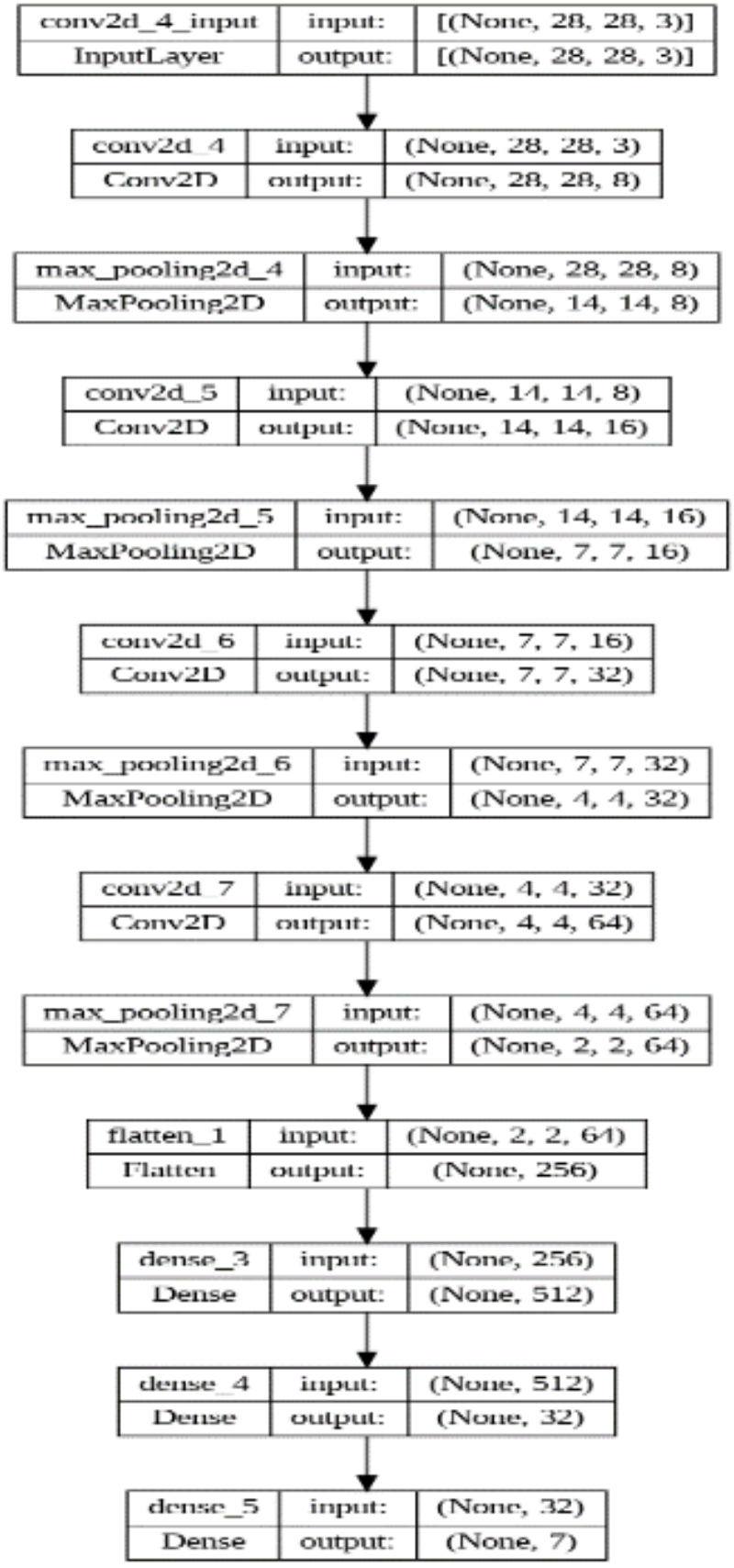

Figure 4. CNN Architecture Visual Representation

4. Results and Discussion

4.1. Training environment

A Tesla T4 GPU with 15 GB of VRAM was employed as the training environment on Google Colab. The system's 15 GB RAM aided the training process. The Keras API inside the TensorFlow library was used for the Dl model development and assessment.

4.2. Testing Evaluation Metrics

Following CNN architecture construction and training, 20 % of the dataset is used to test the suggested model's performance. The evaluation process allows for assessment of the tested model's reliability in SC classification. Numerous parameters are used for model performance evaluation; for instance, the model was assessed during a classification task using accuracy and loss measures. The number of occurrences the model accomplishes in relation to the total number of categorization incidents is known as the accuracy. This implies that the model's ability to classify the various classes is directly determined by the accuracy metric. On the other hand, the model's inconsistencies are represented numerically by the loss metric. It evaluates the model's fit to the data as well. The accuracy and loss of the CNN model are shown in Figure 5; the graph indicates that the training and validation accuracy gradually increase in tandem with the ephocs number. The model reached peak training and validation accuracy (almost 100% and 97.9% respectively) at 15 epochs. However, an increasing number of epochs drastically reduced the training and validation loss, peaking at 15 epochs. Because there was little more progress, the training was terminated early in period 26.

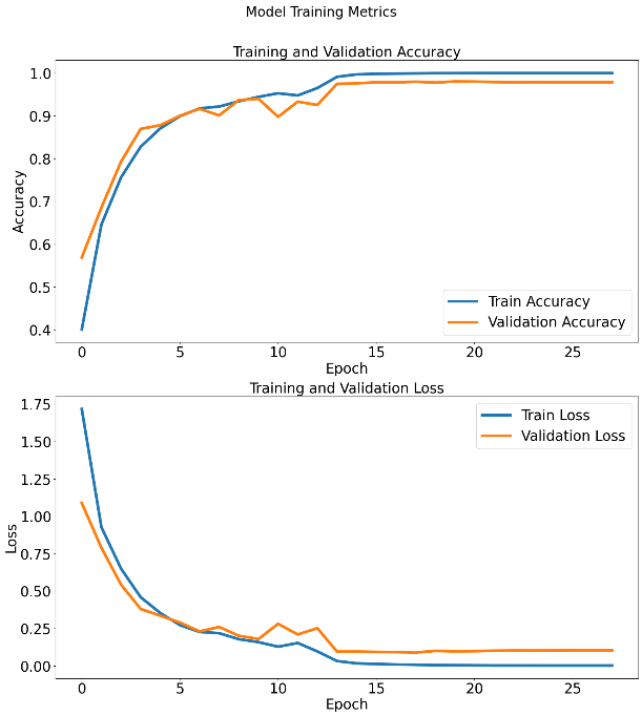

Figure 5. Training and validation accuracy and loss metrics

Model performance is evaluated using metrics like recall, accuracy, f1-score, and precision [24] [25]:

- Accuracy: A commonly used metric for assessing models is accuracy, which expresses the proportion of samples that the model correctly identified. It is calculated by dividing the total number of predictions the model made by the number of accurate predictions. The calculation for accuracy is as follows:

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

- Precision: Precision is a metric that evaluates how well the model predicts positive outcomes and makes sure that negative samples are not mistakenly labeled as positive. It can be computed in this way:

:

$$\text{Precision} = \frac{\text{TP}}{\text{TP} + \text{FP}}$$

- Recall: Recall is a frequently employed evaluation value that computes the ratio of correctly identified actual positive samples by the model. It can be computed as follows:

$$\text{Recall} = \frac{\text{TP}}{\text{TP} + \text{FN}}$$

- F1 score provides a balanced measure between recall and precision, and is computed as follows:

$$\text{f1} - \text{score} = \frac{2 \times (\text{Precision} \times \text{Recall})}{\text{Precision} + \text{Recall}}$$

The recall, accuracy, f1-score, and precision scores are presented in Figure 6; the proposed models' accuracy was 94.95%, which indicates that 94.9 percent of the sampled cases were properly identified by the CNN classifier. The precision and recall values were 1.00 and 0.96, respectively. A recall value of 0.96 indicates the classifiers' ability to correctly identify 96 % of the true positive cases in the dataset; however, the projected positive examples did not contain any false positives. A good degree of performance in the classification of different skin lesions is shown by the acquired F1 score of 0.98. An F1 score of 0.98 suggests a high accuracy of the model in identifying the positive cases correctly (precision) and gathering a sizable portion of the true positive samples (recall); it demonstrates the ability of the classifier to achieve low rates of false positives and false negatives reports.

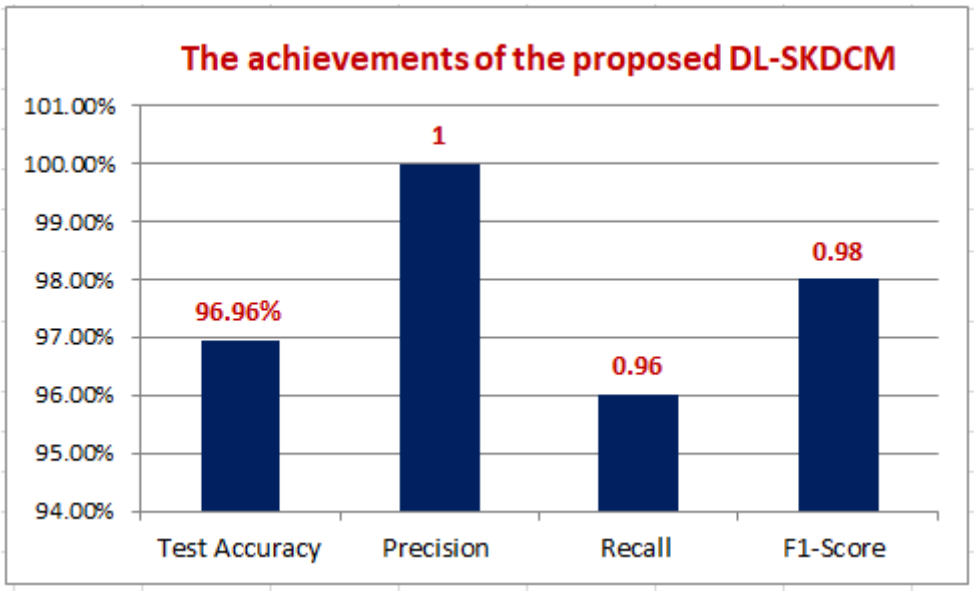

Figure 6. Performance of the proposed model in terms of accuracy, F1, precision, and recall.

A confusion matrix, which provides a visual representation of the performance of the model per class, was also produced, displaying the false positives, true positives, false negatives, and true negatives in this matrix.

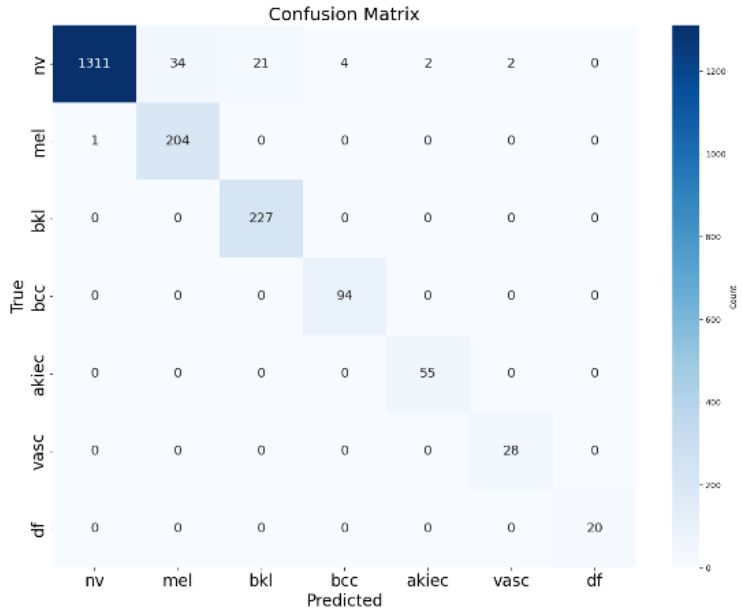

Figure 7. Confusion matrix of the proposed CNN models' performance.

Figure 7 displays the actual class labels in rows and the predicted labels in columns. The numbers inside each cell represent the number of incidents. Greater numbers are shown by lighter hues in the matrix, which is color-coded as a heatmap. A confusion matrix functions as a precise study of each class's accuracy and probable faults in this particular setting. More precisely, the classes that the model was able to accurately predict are DF, VASC, AKIEC, BCC, and BKL, as shown in the confusion matrix in Figure 7. On one occasion, though, the MEL class was mistakenly categorized as NV. Out of 1311 cases that were classified correctly, 34 were not correctly classified as MEL, 21 as BKL, 4 as BCC, 2 as AKIEC, and 2 as VASC. Out of these, NV had the highest rate of misclassifications. These numbers are highlighted, and the rest are clearly marked on the accompanying heatmap, which gives a good sense of the models' capabilities. Such graphics are essential in estimating the overall performance of the classification and determining the shortcomings in the model's ability to predict the results.

4.3 Discussion

A classification accuracy of 96.9 % indicates that the proposed CNN model performed commendably in classifying different types of SC. The models' effectiveness is further emphasized through high performance scores during testing. The model actually achieves perfect scores in a variety of class classifications, such as the BCC, VASC, and DF classes. The CNN model was able to detect all seven distinct classes of skin lesions well, as evidenced by achieving 1.00 precision and 0.96 recall of difficult classes, for instance, the NV class. The proposed CNN models' performance was benchmarked against other models for a better insight into their performance; for instance, the examination done by [20] used an incremental CNN model to classify skin lesions. To guarantee that the model is properly trained, the procedure consists of four parts. Similar to the HAM10000 dataset, the ISIC 2018 challenge dataset, which also contains the seven skin lesion classes, was used to train their model in order to complete the classification job. Their incremental CNN model technique produced good results that outperformed the normal CNNs' performance on the same dataset. Their accuracy was 90 % as opposed to 64 % for the standard CNN model. The task of classifying skin lesions using a CNN architecture and the HAM10000 dataset was also covered in another paper by [21]. Classifying all seven forms of skin lesions was another goal of the study. Four layers (4-CNN) comprising two max-pooling layers (Pool1, Pool2), six batch normalization levels, four dropout layers, and four fully linked layers comprised their developed CNN model. Overall, their model produced good results with an accuracy of 93 %. The suggested CNN model

outperforms the results of these two investigations in terms of recall, accuracy, precision, and F1 score, as seen in Figure 8.

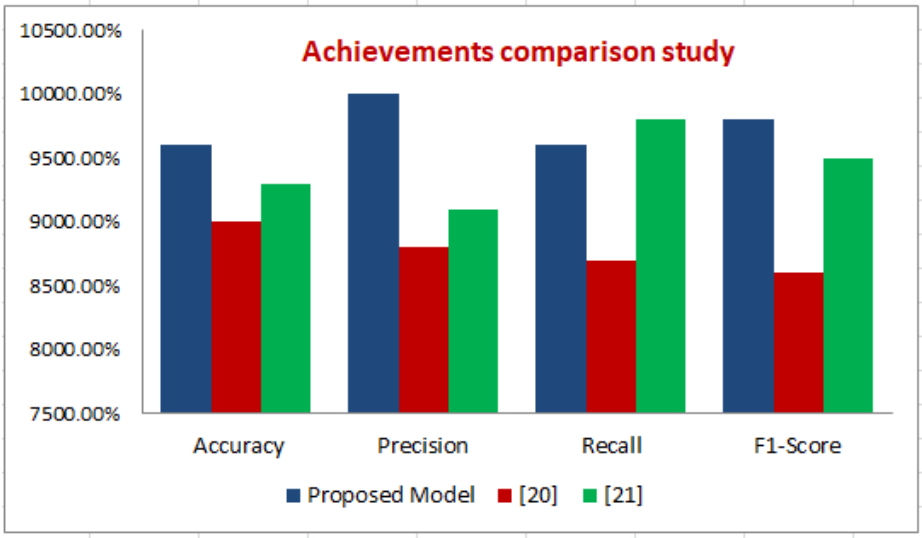

Figure 8. Bar graph Comparing the Accuracy, Precision, Recall, and F1-score for the Proposed Model and Other Models

The suggested CNN model performed better than the benchmark models in terms of accuracy, recall, precision, and F1 score, as shown in Figure 8 [20], [21]. The developed model had the highest accuracy of 96.9 %, followed by the Ankir model at 90 % and the UdriOtoiu model at 93.6 %. In addition, the precision score of the suggested model was greater (1) than that of Ankir (0.87) and UdriOtoiu (0.91). The proposed model had a slightly lower memory value (0.96) than the UdriOtoiu model, which had a higher score (0.98). The Ankir model had the lowest recall value (0.87). Lastly, the suggested model had the greatest F1-score (0.98) compared to the benchmark models. Collectively, the proposed model achieved better performance than most of the benchmark models.

1. Conclusion

Early identification and diagnosis of SCs allow faster and better treatment, thus proper management of SCs. CNNs have been recognized as be state-of-the-art method for the diagnosis of skin lesions because of their ability to learn complex data and feature patterns within skin images, and are trained mostly on large datasets for accurate performance. At the same time, this training will aid in the efficient and automated screening of lesions while minimizing human error. The suggested model showed notable performance, achieving 96.95 % test accuracy on the HAM10000 dataset. In this work, a DL model, more precisely, the CNN model, was created for the classification of SC using dermatoscopic images. The test and validation performance of the proposed model in terms of accuracy, precision, and recall scores was better relative to the benchmark models against various skin lesion classes. The model has a good F1 score, meaning that it could be used accurately to diagnose SC.

The future works of this study include:

i. Expanding the dataset with multi-source and multi-ethnic dermatoscopic images to enhance model generalization.

ii. Applying transfer learning and domain adaptation techniques to improve robustness across various imaging conditions.

iii. Developing hybrid architectures that combine CNNs with transformer or attention-based models for better feature extraction.

iv. Integrating explainable AI methods such as Grad-CAM to improve interpretability for clinical decision-making.

v. Designing real-time diagnostic systems deployable on mobile or web platforms.

vi. Incorporating patient metadata like age, lesion history, and location for multi-modal learning.

vii. Conducting extensive clinical validation in collaboration with dermatologists to assess the real-world applicability and reliability of the proposed model.

References

1. Swapno, S. M. R., Nobel, S. N., Meena, P. K., Meena, V. P., Bahadur, J., & Appaji, A. (2025). Accelerated and precise skin cancer detection through an enhanced machine learning pipeline for improved diagnostic accuracy. *Results in Engineering*, *25*, 104168.
2. Bappi, M. B. R., Swapno, S. M. R., & Rabbi, M. F. (2024, May). Skin cancer disease detection using MCD-GRU: A deep learning approach. In *2024 6th International Conference on Electrical Engineering and Information & Communication Technology (ICEEICT)* (pp. 445-450). IEEE.
3. Jojoa Acosta, M. F., Caballero Tovar, L. Y., Garcia-Zapirain, M. B., & Percybrooks, W. S. (2021). Melanoma diagnosis using deep learning techniques on dermatoscopic images. *BMC Medical Imaging*, *21*(1), 6.
4. Khodadadi, H., Sedigh, A. K., Ataei, M., Motlagh, M. R. J., & Hekmatnia, A. (2017). Nonlinear analysis of the contour boundary irregularity of skin lesion using Lyapunov exponent and KS entropy. *Journal of Medical and Biological Engineering*, *37*(3), 409-419.
5. Khazaei, Z., Ghorat, F., Jarrahi, A. M., Adineh, H. A., Sohrabivafa, M., & Goodarzi, E. J. W. C. R. J. (2019). Global incidence and mortality of skin cancer by histological subtype and its relationship with the human development index (HDI); an ecology study in 2018. *World Cancer Res J*, *6*(2), e13.
6. Ibrahim, O., Gastman, B., & Zhang, A. (2014). Advances in diagnosis and treatment of nonmelanoma skin cancer. *Annals of plastic surgery*, *73*(5), 615-619.
7. Melarkode, N., Srinivasan, K., Qaisar, S. M., & Plawiak, P. (2023). AI-powered diagnosis of skin cancer: a contemporary review, open challenges and future research directions. *Cancers*, *15*(4), 1183.
8. Chaurasia, B. K., Raj, H., Rathour, S. S., & Singh, P. B. (2023). Transfer learning–driven ensemble model for detection of diabetic retinopathy disease. *Medical & Biological Engineering & Computing*, *61*(8), 2033-2049.
9. Tripathi, A., Misra, A., Kumar, K., & Chaurasia, B. K. (2023). Optimized machine learning for classifying colorectal tissues. *SN Computer Science*, *4*(5), 461.
10. Singh, P. B., Singh, P., Dev, H., & Chaurasia, B. K. (2023). Glaucoma classification using enhanced deep transfer learning models with hybrid ROI cropped optic disc technique. *SN Computer Science*, *4*(6), 722.
11. Alam, T. M., Shaukat, K., Khelifi, A., Aljuaid, H., Shafqat, M., Ahmed, U., ... & Luo, S. (2023). A fuzzy inference-based decision support system for disease diagnosis. *The Computer Journal*, *66*(9), 2169-2180.
12. Hasan, M. R., Fatemi, M. I., Monirujjaman Khan, M., Kaur, M., & Zaguia, A. (2021). Comparative analysis of skin cancer (benign vs. malignant) detection using convolutional neural networks. *Journal of Healthcare Engineering*, *2021*(1), 5895156.
13. Kumar, A., & Chaurasia, B. K. (2024). Detection of SARS-CoV-2 virus using lightweight convolutional neural networks. *Wireless Personal Communications*, *135*(2), 941-965.
14. asan, M., Barman, S. D., Islam, S., & Reza, A. W. (2019, April). Skin cancer detection using convolutional neural network. In *Proceedings of the 2019 5th international conference on computing and artificial intelligence* (pp. 254-258).
15. Afza, F., Sharif, M., Khan, M. A., Tariq, U., Yong, H. S., & Cha, J. (2022). Multiclass skin lesion classification using hybrid deep features selection and extreme learning machine. *Sensors*, *22*(3), 799.
16. Akram, T., Alsuhaibani, A., Khan, M. A., Khan, S. U., Naqvi, S. R., & Bilal, M. (2024). Dermo-Optimizer: Skin Lesion Classification Using Information-Theoretic Deep Feature Fusion and Entropy-Controlled Binary Bat Optimization. *International Journal of Imaging Systems and Technology*, *34*(5), e23172.
17. Bibi, S., Khan, M. A., Shah, J. H., Damaševičius, R., Alasiry, A., Marzougui, M., ... & Masood, A. (2023). MSRNet: multiclass skin lesion recognition using additional residual block based fine-tuned deep model's information fusion and best feature selection. *Diagnostics*, *13*(19), 3063.
18. Ozdemir, B., & Pacal, I. (2025). An innovative deep learning framework for skin cancer detection employing ConvNeXtV2 and focal self-attention mechanisms. *Results in Engineering*, *25*, 103692.

19. Dillshad, V., Khan, M. A., Nazir, M., Saidani, O., Alturki, N., & Kadry, S. (2025). D2LFS2Net: Multi-class skin lesion diagnosis using deep learning and variance-controlled Marine Predator optimisation: An application for precision medicine. *CAAI Transactions on Intelligence Technology*, *10*(1), 207-222.

20. Chopade, A. (2019, April). Seven class classification of skin lesions by using incremental convolutional neural network in Python. In *Proceedings of the International Conference on Advances in Electronics, Electrical & Computational Intelligence (ICAEEC)*.

21. UdriȘtoiu, A. L., Stanca, A. E., Ghenea, A. E., Vasile, C. M., Popescu, M., UdriȘtoiu, Ș. C., ... & Gruionu, G. (2020). Skin diseases classification using deep leaning methods. *Current Health Sciences Journal*, *46*(2), 136.

22. Chanda, D. et al. A new deep convolutional ensemble network for skin cancer classification. Biomed. Signal. Process. Control. 89, 105757 (2024).

23. Brancaccio, G., Balato, A., Malvehy, J., Puig, S., Argenziano, G., & Kittler, H. (2024). Artificial intelligence in skin cancer diagnosis: a reality check. *Journal of Investigative Dermatology*, *144*(3), 492-499.

24. Obaid, A. L., Haddad, N. M., & Taha, M. S. (2024, November). DL-SCDDS: Accurate Skin Cancer Detection and Diagnosis Scheme Based on an Improved Convolutional Neural Networks Model. In *International Human-Centered Technology Conference* (pp. 201-214). Cham: Springer Nature Switzerland.

25. abbar, N. K., Naderan, M., & Taha, M. S. (2025). HybridIoMT: A Dual-Phase Machine Learning Framework for Robust Cybersecurity in Internet of Medical Things. *International Journal of Intelligent Engineering & Systems*, *18*(4).