Iraqi Journal for Computer Science and Mathematics

Volume 6 | Issue 1 Article 10

2025

# Double Dual Convolutional Neural Network (D2CNN): A Deep Learning Model Based on Feature Extraction for Skin Cancer Classification

Raya Sattar Shahadh
*Informatics Institute for Postgraduate Studies, University of Information Technology & Communications, Iraq*

Belal Al-Khateeb
*College of Computer Science and Information Technology, University of Anbar, Iraq*, belal-alkhateeb@uoanbar.edu.iq

Recommended Citation
Shahadh, Raya Sattar and Al-Khateeb, Belal (2025) "Double Dual Convolutional Neural Network (D2CNN): A Deep Learning Model Based on Feature Extraction for Skin Cancer Classification," *Iraqi Journal for Computer Science and Mathematics*: Vol. 6: Iss. 1, Article 10.
DOI: https://doi.org/10.52866/2788-7421.1240
Available at: https://ijcsm.researchcommons.org/ijcsm/vol6/iss1/10

IRAQI JOURNAL FOR COMPUTER SCIENCE AND MATHEMATICS 2025;6:139–148

ORIGINAL STUDY

# Double Dual Convolutional Neural Network (D2CNN): A Deep Learning Model Based on Feature Extraction for Skin Cancer Classification

**Raya Sattar Shahadh** [a], **Belal Al-Khateeb** [b,*]

[a] Informatics Institute for Postgraduate Studies, University of Information Technology & Communications, Iraq
[b] College of Computer Science and Information Technology, University of Anbar, Iraq

## ABSTRACT

Artificial intelligence, especially in the field of "deep learning", is still promising when it comes to skin cancer detection and diagnosis. Among deep learning algorithms, convolutional neural networks (CNNs) give a high level of accuracy in identifying and classifying different types of skin cancer. CNNs have a strong coordination due to understanding the important features from medical images that are extracted from convolutional layers. However, there is still a problem which is the high imbalance in the dataset with high noise in the images. This paper presents a new solution that combines different architectural structures of convolutional neural networks (CNNs) to extract features from skin cancer images with two forms of linear algebra methodologies, Principal Component Analysis (PCA) and Factor Analysis (FA) to eliminate the curse of combining the extracted features and removing the redundant features and passing them to a fully connected classifier to perform multiple classification, in this paper the HAM10000 dataset is used. The idea here is to retain the vital features of the image and extract the important features from it and eliminate the irrelevant features which leads to achieving higher classification accuracy for all classes. The proposed method was accurate and achieved a macro average test accuracy of 97.90% with a test loss of 0.14% and precision, recall and f1 score of 0.97%, 0.93% and 0.95% respectively and achieved a weighted average of 0.98% in precision, recall and f1 score in multi-class classification.

**Keywords:** Convolutional neural network (CNN), Feature extraction, Skin cancer, HAM 10000, Principal component analysis (PCA), Factor analysis (FA), Dual CNN, Fully connected classifier, Deep learning

## 1. Introduction

Skin cancer is a serious skin disease that affects millions of people every year. Early detection is expensive and difficult, but it can help save many lives. However, recent studies have shown that deep learning-based methods can help dermatologists effectively classify medical images [1]. Skin cancer has caused the death of millions of people, and over the past few decades, the rates of skin cancer diagnoses have risen dramatically. Skin cancer occurs when abnormal cells on the skin grow uncontrollably. This occurs when certain types of cells damage their DNA, causing mutations that cause skin cells to multiply rapidly and form a malignant tumor [2]. Dermatologists have extensive training in the different skin lesions that patients may experience and are therefore best positioned to correctly diagnose melanoma. However, melanoma diagnosis faces many problems because the boundaries between skin lesions and surrounding skin are not very clear, and because malignant and non-malignant skin lesions often look quite similar. Therefore, the creation of an accurate, automated skin tumor detection system that can analyze skin lesions and be an aid to dermatologists would be very useful [3].

Received 22 October 2024; revised 10 January 2025; accepted 10 January 2025.
Available online 8 March 2025

---

* Corresponding author.
E-mail address: belal-alkhateeb@uoanbar.edu.iq (B. Al-Khateeb).

https://doi.org/10.52866/2788-7421.1240

In the recent past, the development of advanced deep learning techniques has become useful in automating cancer classification processes. These algorithms involve the use of sophisticated techniques that are capable of learning important features from large datasets along with identifying complex patterns useful for the correct classification of cancers of different types. CNN techniques, among the deep learning algorithms, work effectively with large training datasets. However, the exact number of datasets that may be effective for a given problem varies somewhat depending on a number of factors, such as the nature of the problem or the structure of the network. Sometimes, a modest cache may be sufficient provided that the problem is not complex and the network in question maintains a relatively uncomplicated architecture. However, complex tasks require a significantly augmented dataset which will enable the development of a deep neural network. In the case of image classification primarily, if a larger dataset is available, it will significantly increase model accuracy [4]. The best example of this architecture is the convolutional neural network (CNN), which is an image recognition and classification tool that can learn directly from data. It has become one of the best machine learning algorithms for dealing with data structures such as networks or for dealing with images. CNN is used in various image processing and computer vision tasks, which include localization, segmentation, classification, and detection. CNNs, as a rule, include dozens or even hundreds of layers, each of which is trained to identify some specific aspect of an image. During training, convolution is applied to images at different scales, where the output from each layer is fed into the next layer. Initially, it first determines the baseline values of intensity and edge, but then continues to more precise definitions of the object feature present in the image. Between the input and output of a CNN, there are always several hidden layers, which can take action to learn features related to the data. Some of the layers are convolution, activation layer where ReLU is most common, and pooling. The convolution layer is the heart of CNN, it convolves the input images to run filters applied to activate aspects of the image. They help speed up training by setting all negative values to zero while keeping all positive values, thus only activated features are forwarded to the next layer. Pooling layers help reduce the number of parameters for the given problem in the situation and perform non-linear sampling of the output thus improving the performance of the network. These operations are essentially iterative across layers and with the experience gained in each layer, the ability to recognize newer features of the image improves. The exact classification result is determined by the classification layer at the top of the CNN design [5, 6].

The contributions of this paper can be summarized as follows:

- Dealing with unbalanced data by relying on a powerful structure to extract features from different CNN algorithm architectures by merging these features and using linear algebra techniques to extract the most important and relevant features for each class and removing duplicate features, thus eliminating overfitting instead of relying on synthetic images resulting from data augmentation
- Dealing with noise in images such as the presence of hair above the skin, which skin cancer images suffer from, through powerful and multiple CNN filters and getting rid of the curse of dimensions using linear algebra techniques with a powerful classifier instead of pre-processing operations for cleaning data or shaving hair or identifying edges.

The remaining sections of the paper are as follows: Previous work is mentioned in the Section 2. In Section 3, the dataset and proposed approach are explained in detail with the evaluation criteria. The obtained results are shown in Section 4, Comparative analysis is discussed in Section 5 and finally, the conclusion is presented in the Section 6.

## 2. Related works

In [1], the researchers presented a deep learning model for skin cancer classification using a five-layer CNN. The images are analyzed to estimate the tumor severity based on color prediction. The model achieved an accuracy of 86.68% on the HAM10000 dataset after data augmentation. However, the study lacks comprehensive reporting, as it does not provide other critical metrics such as accuracy and F1 score, which are essential for assessing the reliability of medical research.

The research in [2], developed a skin cancer classification system using CNN algorithm with Grad-CAM and Grad-CAM+ + and it is characterized by explaining the decisions taken by the model and the approach achieved a classification accuracy of 81.24% on the HAM10000 image set. This percentage is considered modest for medical applications, and it used accuracy only as a measure of the efficiency of the model while ignoring the rest of the measures since determining the number of false negatives and false positives is essential in medical classification. In [3], the combination of LSTM and CNN enhanced the extraction of complex features from medical images, helping to reduce errors and increase accuracy, especially when there are very similar classes (as in skin cancer). The researchers proposed a model that combines two deep learning algorithms CNN, LSTM with a machine learning classifier SVM on the HAM10000 dataset. The model achieved an accuracy of 88.24% on a reconnection of 88% and an F_SCORE of 88%, which are considered modest results in the medical field. A pre-trained models such as Densenet121, MobileNet, ResNet50, and VGG19 were used. In [4]. ImageNet weights were used by each of the above pre-trained networks. These models were then trained on the HAM10000 dataset. To determine which network is better, Densnet121 had the highest accuracy with an accuracy of 92.42. Despite achieving high results, the classification report was not mentioned to know if there is bias towards a particular class or if there is a low class criterion. In [5], the researchers proposed a DL model designed on VGG19 network with self-attention blocks integrated for skin cancer classification. The proposed model achieved an overall accuracy of 91.4% on HAM10000. The results showed that incorporating self-attention blocks into CNN improves the performance of skin cancer detection systems. When evaluated, its performance was lower for melanoma class. A combined system is described in [6], that incorporates DenseNet pre trained architecture with feature improvement by the aid of Convolutional Block Attention Module (CBAM) and HAM10000 dataset. The overall accuracy achieved with proposed approach is higher than dense net model averages of 93% for accuracy, 93% for precision and 93% for recall respectively. Using the LIME framework helps to go a step further in evaluating predictions and getting visual explanations. Although many pre-processing processes were used and use data augmentation, general evaluation criteria for the model are mentioned without specifying those criteria for each of the seven types of pests to know whether the criteria are biased towards a specific type or vice versa, i.e. there is one or more types of pests whose results are low. In [7] CNN was applied with data preprocessing like, down sampling, sharpen, and segment with autoencoder and decoder for dataset HAM 10000. Some pre-trained transfer learning models as DenseNet169 and Resnet 50 were used. under sampling technique used in DenseNet169 got the accuracy of 91%. On the second level, details referring to 2% of the vectors have a f1-measure of 91%. To be specific, the oversampling scheme used in Resnet50 enhanced accuracy of 7% and gave 83% accuracy in addition to an f1-measure of 84%, even though the accuracy of the model in general was good at Densenet169 (under sampling), this accuracy is biased as the classification report indicates bias in accuracy for certain types versus low accuracy and Precision, Recall for the rest of the types. The suggested method in [8] used the MobileNet architecture, small Depth Neural Network which is suitable to be implemented on computational constrained environments like handheld devices. The material under study comprises a set of skin diseases that contain seborrheic keratosis, benign nevi and melanoma. By using transfer learning, the model is trained on this dataset and achieves a high accuracy of over 95%, Precision 89.32%, Recall 85.21%, and F1-score 87.11%. In this research, the evaluation criteria such as accuracy and recall are not high, and the research did not include a classification report to know whether these results are biased towards one or more types without the rest of the types. In the work [9], the HAM 10000 dataset containing skin images was used. Models based on CNN, MobileNet V2, and Resnet50 were built and tested. A validation accuracy of 86% was achieved for CNN, 96% for MobileNet, and 89% for ResNet50. Despite the high accuracy of MobileNet, the rest of the evaluation criteria such as precision and recall are not high, which affects the accuracy of the medical results. In [10], the researchers presented an approach to classify melanoma images using the Xception model on the HAM10000 medical image set. This study achieved a classification accuracy of 90.24%. To demonstrate the classification model, the researchers generated heatmaps using Gradient-weighted Class Activation Mapping (Grad-CAM) and its upgrade, Grad-CAM+ +. Heatmaps allow the researchers to describe the role of each input region in the classification outcome. In this study, the classification is only binary, meaning it is either classified as melanoma or not.

In [11], a pre-trained deep learning model was presented for skin lesion classification. Inception-Resnet-v2 which combines Inception and Resnet was chosen to classify skin cancer images. In this research, the main objective is to demonstrate the effect of increasing the number of images in the dataset by using data augmentation by applying affine transform technique to the skin cancer classification system. The highest accuracy reported in this study with 39787 imposed images in this study is 95.9% for image recognition and 83% for Inception-Resnet-v2 model for classification using the original dataset without data augmentation which contains 10015 images. In this study, data augmentation was used extensively in addition to using a hybrid model consisting of two pre-trained algorithms which increases the computational complexity. The research in [12] used a form of image preprocessing that comprised of hair removal, dataset enhancement as well as scaling to fit the various models. Transfer learning was applied with optimized CNNs, known as EfficientNets B0-B7 pre-trained by ImageNet weights in terms of performance in multi-class unbalanced classification on the HAM10000 dataset. As a result, Efficient Net B4 was the best-performing model with 87% accuracy achieved. Overall accuracy stands at 91%, and precision at 88%, the F1 score is 87%. Thus, though the results are quite reasonable, the accuracy itself is hardly satisfactory for practical use in clinics, as people can be classified both as true positives and negatives. In [13] study the authors used CNN with spatial attention as well as pre-trained DenseNet-201. This fusion of features obtained from both networks should enhance the prediction accuracy. The model achieved an overall accuracy of 82% out of the 1,000 samples used to test the model's performance. It achieved nearly 57% on the augmented HAM10000 dataset and a Cohen's kappa score of 0.654. Although the performance is quite reasonable, the study failed to provide other evaluation criteria such as precision and recall rate that affect the likelihood of medical outcomes. The study in [14] focused on comparing four deep learning techniques RNN (Recurrent Neural Networks), CNN, ResNet50 and Xception to conclude which algorithms perform better in predicting correctly and accurately on the dataset used for this research, which is HAM10000, which is available on Kaggle. The experimental results showed that the accuracy using CNN, RNN, ResNet50 and Xception was 72%, 69%, 79% and 93% respectively, with Xception performing the best. This paper didn't not include a report that explains the performance of the algorithm with each type of pest and whether the accuracy is high for all categories or there is a difference between categories. The work in [15] discussed the employment of deep learning for skin cancer detection. This type of learning was used to create five state-of-the-art convolutional neural networks including DenseNet201, GoogLeNetInception_ResNetV2, InceptionV3, and MobileNetV2. The classifiers include both regular and intermediate with two hierarchical levels of classification models Who can distinguish seven types of moles. The performance of the method in experiments was increased with the help of data augmentation on a large set of skin images (HAM10000 dataset). A brief analysis of the obtained results shows that the selected DenseNet201 network is effective for this task with a classification accuracy of 95.09% and F-measures with fewer false negatives. However, the classification report shows low scores for more than one type of pest, while high scores for other types of pests. The work in [16] distinguished the seven skin lesions found in HAM10000 by employing CNN. The proposed model was able to achieve the accuracy of 91.51%. The model was integrated with a web application and then assessed in a two-part process by seven expert dermatologists. In the first stage IT was concluded that the model is capable to diagnose the skin lesions with the accuracy of 90.28% accuracy in practice. In the second stage, the model corrected the experts' misdiagnoses with 11.14% accuracy. If considering the classification report on the individual lesion, it is observed that the accuracy is inflated in different lesion types (nv), and the other evaluation metrics on each lesion including Precision, Recall, F1-score are quite low. In [17], The transfer learning algorithm Inception_ResNet with boosting was used on the HAM10000 dataset to balance the high deviation among seven skin cancer classes where 450 samples were selected for each type and the accuracy obtained was 78.90%. The results showed that the accuracy is low, and the number of training samples is also low and also the accuracy was used only as a measure of the model efficiency. The research in [18] applied ensemble learning to a number of transfer learning algorithms including (DENSENET201_MOBILENETV2_XCEPTION_EFFIECIENTNETB the HAM10000 dataset and the performance comparison was made with each network individually. The results after ensemble learning were higher and led to significant improvement in accuracy as well as in other important medical diagnosis metrics such as precision, recall and f1 score, with the results being 96.4%, 90%, 92% and 89% respectively. In this work, the general evaluation criteria of the model were mentioned, and the work did not include a reference to the classification report or confusion matrix to know whether the criteria are biased towards a specific type or more of the seven skin lesions, or whether there is a decrease in the percentages of these criteria in some types. The fine-tuning was implemented in [19] on seven classes of HAM10000 and comparative experiment was conducted to compare five pre-trained CNNs and four ensemble models. The best accuracy one can ever obtain is of 93%. Among the models that were discussed ensemble method. Although the accuracy was high, the rest of the metrics such as reconnection and F1-score, Precision showed low performance.

## 3. Methodology

The details of the proposed model are shown in this section, this including the dataset description, proposed approach. and the evaluation criteria.

### 3.1. Dataset

The HAM1000 dataset ("Human vs. Machine with 10,000 training images"), a publicly available image collection of skin lesions, is a comprehensive resource

Fig. 1. Distribution of the seven diagnostic categories.

obtained from two main sites: the Queensland Skin Cancer Clinic in Australia and the department of Dermatology at the Medical University of Vienna, Austria. This dataset includes skin images classified into seven distinct categories: malignant: melanoma (MEL), basal cell carcinoma (BCC), actinic keratosis and intraepithelial carcinoma (AKIEC) and benign: pigmented nevi (NV), benign keratinizing lesions (BKL), vascular lesion (VASC) and dermatofibroma (DF). Its primary use is to determine whether a patient has skin cancer and, if so, what type it is. There are a total of 10015 color images in the dataset which have been used in several studies to characterize and evaluate different diagnostic tools and model performance [20] Fig. 1 represents samples from the HAM10000 data set.

### 3.2. Proposed approach

The proposed method includes a dual convolutional CNN architecture, which uses different structural designs to capture a wide range of features. Each design includes four convolutional layers, where features are extracted through these layers, followed by an activation function with each of the four layers. After each convolutional layer, there is a max pooling layer and after extracting the features from all layers, these features pass through a flattening layer. In each model, PCA is applied to the features extracted from the first dual CNN to reduce the dimensionality and retain the important features and FA is applied to the features extracted from the second dual CNN to reduce the dimensionality and retain the important features. The idea of using two types of linear algebra to reduce the dimensionality is to obtain as many important and diverse features as possible that help in identifying the features of the pests. After that, the features are merged so that the FA features are part of the PCA columns. After this merger, the correlation matrix is used to remove the duplicate features double dual CNN. The features are then passed to a fully connected classifier to classify these merged features. The final classification step uses a three-layer fully connected network, including two dropout layers, followed by a SoftMax layer, to classify the features into seven categories of skin lesions, as shown in Fig. 2. Table 1 shows the architecture of the proposed approach.

#### Preprocessing

A. Resizing images: In deep learning, images are typically represented as arrays of pixel values. Larger images require more memory to store and process, which can become a bottleneck when working with limited computational resources. Resizing an image to a smaller size reduces the memory footprint, which makes it easier to work with, so the images were resized to (224, 224).
B. Feature extraction: The process of extracting features from medical images using CNNs aims to identify relevant patterns or features in medical images such as color, edges, shapes, and texture, which can be later used in classification or analysis. Here, features are extracted using four different CNN architectures, which helps to obtain powerful and diverse features.
C. Reduce dimensions: Both PCA and FA are data analysis algorithms that aim to reduce dimensionality or extract hidden patterns in data. PCA relies on the variance in data to convert it into principal components [21], while FA looks for common factors that influence features and is used to understand the underlying structure of data [22]. This helps in reducing the dimensionality and hence reducing the computational complexity while improving the performance of the model.

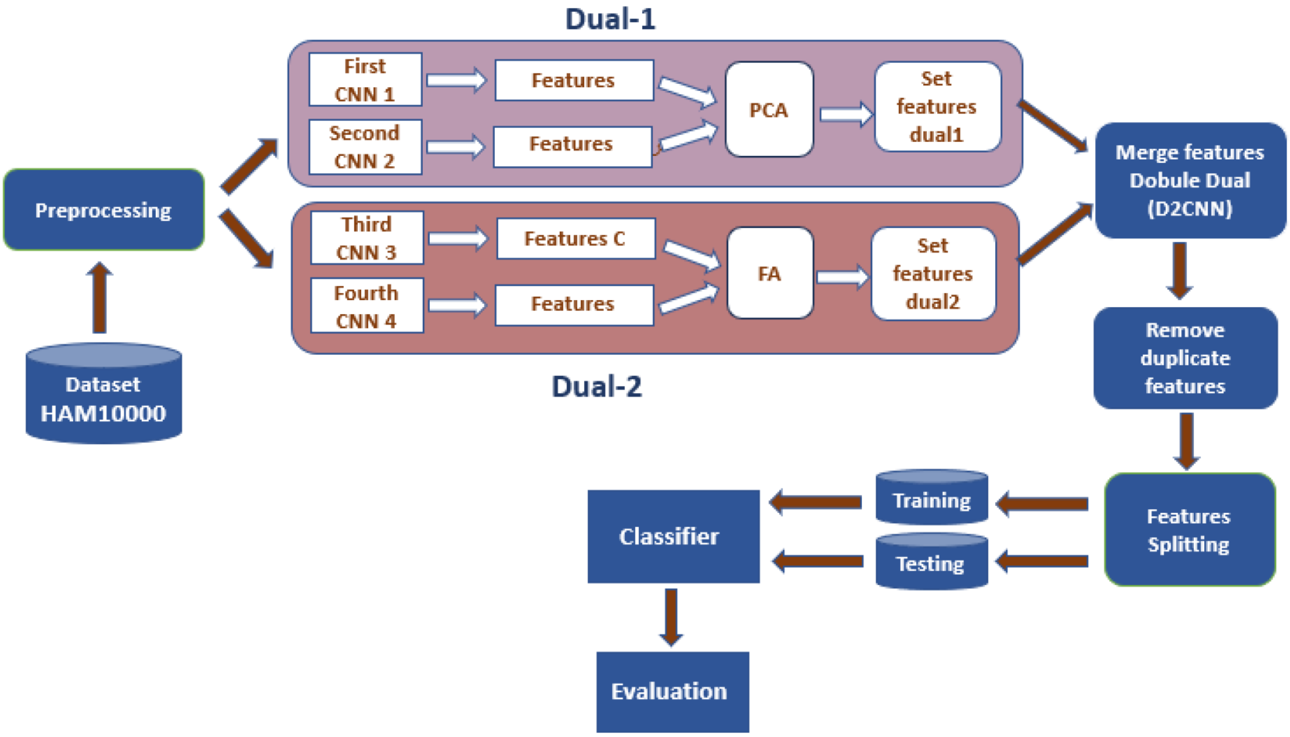

Fig. 2. Block diagram of the proposed approach (double dual CNN).

PCA is applied to reduce the dimensions extracted from the first dual CNN, and FA is applied to the second dual CNN. The purpose of using more than one dimensionality reduction technique is to obtain the greatest diversity of features.

D. Combined features and remove duplicate features: The reduced features are combined after applying PCA and FA (Where FA features are added as additional columns to PCA features) to create a unified and efficient representation that can be used for image classification. During fusion, duplication of features may occur. A correlation matrix is used to remove this duplication where the correlation between all features is calculated [21] and the highly correlated features are identified and removed. The resulting features are then saved for use in the classification process.

E. Classifier: The fully connected classifier is used with ReLU and SoftMax activation function for multi-classification.

Table 1 shows the proposed CNN architectures for feature extraction and classification.

Hyperparameters are chosen according to testing many options. Adam optimizer is used, Batch Size is 128, Number of Epochs is 80 and Loss Function is the Categorical Cross Entropy, the extracted features were divided into 70% training sets, 10% validation sets, and 20% testing sets.

### 3.3. Evaluation criteria

The Classification performance was evaluated using several criteria, including accuracy, precision, recall, validation loss, and F1 score [21]. Accuracy was calculated by dividing the number of predictions that were correct by the total number of samples as in Eq. (1).

$$Accurcy = \frac{TP + TN}{TP + TN + FP + FN} \tag{1}$$

Precision is determined by dividing the number of correctly predicted positive cases by the sum of all correctly and incorrectly predicted positive cases. This measure is particularly useful when reducing false positives is more important than reducing false negatives, as shown in Eq. (2)

$$Precision = \frac{TP}{TP + FP} \tag{2}$$

Recall, also known as sensitivity, is calculated by dividing the number of correctly predicted positive cases by the total number of actual positive cases. This measure is especially important when false negatives

**Table 1.** Proposed CNN architectures for feature extraction.

| CNN architecture | Layer | Kernal | Filter | activation |
|---|---|---|---|---|
| First CNN | Conv 2D | (3, 3) | 32 | Relu |
| | Max Pooling 2D | (2, 2) | | |
| | Conv 2D | (3, 3) | 64 | Relu |
| | MaxPooling2D | (2, 2) | | |
| | Conv 2D | (3, 3) | 128 | Relu |
| | Max Pooling 2D | (2, 2) | | |
| | Conv 2D | (3, 3) | 256 | Relu |
| | Max Pooling 2D | (2, 2) | | |
| | Flatten | | | |
| | | PCA | | |
| Second CNN | Conv 2D | (7, 7) | 256 | Relu |
| | AveragePooling2D | (2, 2) | | |
| | Conv 2D | (5, 5) | 128 | Relu |
| | AveragePooling2D | (2, 2) | | |
| | Conv 2D | (3, 3) | 96 | Relu |
| | AveragePooling2D | (2, 2) | | |
| | Conv 2D | (3, 3) | 96 | Relu |
| | AveragePooling2D | (2, 2) | | |
| | Flatten | | | |
| | | PCA | | |
| Third CNN | Conv 2D | (3, 3) | 32 | Relu |
| | MaxPooling2D | (2, 2) | | |
| | Conv 2D | (5, 5) | 96 | Relu |
| | MaxPooling2D | (2, 2) | | |
| | Conv 2D | (5, 5) | 128 | Relu |
| | MaxPooling2D | (2, 2) | | |
| | Conv 2D | (7, 7) | 256 | Relu |
| | MaxPooling2D | (2, 2) | | |
| | Flatten | | | |
| | | FA | | |
| Fourth CNN | Conv 2D | (3, 3) | 32 | Relu |
| | AveragePooling2D | (3, 3) | | |
| | Conv 2D | (3, 3) | 32 | Relu |
| | AveragePooling2D | (3, 3) | | |
| | Conv 2D | (5, 5) | 64 | Relu |
| | AveragePooling2D | (3, 3) | | |
| | Conv 2D | (5, 5) | 128 | Relu |
| | AveragePooling2D | (3, 3) | | |
| | Flatten | | | |
| | | FA | | |
| Fully Connected Classifier | Dense | 256 | Relu | |
| | Dropout0.2 | | | |
| | Dense | 128 | Relu | |
| | Dropout0.2 | | | |
| | Dense | 64 | Relu | |
| | Dropout0,2 | | | |
| | Dense | 7 | SoftMax | |

**Table 2.** Performance scores (double dual CNN).

| Lesion | Precision | Recall | F1-score |
|---|---|---|---|
| Akiec | 0.98 | 0.91 | 0.96 |
| Bcc | 0.95 | 0.95 | 0.95 |
| Bkl | 0.97 | 0.99 | 0.98 |
| Df | 1.00 | 0.88 | 0.92 |
| Mel | 0.99 | 0.97 | 0.99 |
| Nv | 1.00 | 1.00 | 1.00 |
| Vasc | 0.94 | 0.86 | 0.90 |
| macro avg | 0.97 | 0.93 | 0.95 |
| Weighted avg | 0.98 | 0.98 | 0.98 |

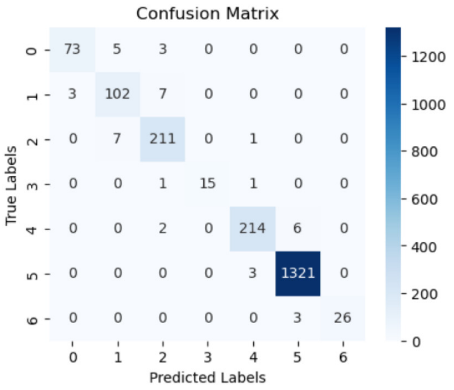

**Fig. 3.** Confusion matrix for dual CNN.

are more problematic than false positives as shown in Eq. (3).

$$Recall = \frac{TP}{TP + FN} \tag{3}$$

The F1 score (Eq. (4)) provides a balanced assessment of precision and recall, which often have an inverse relationship—improving one can negatively impact the other. It is the harmonic mean of precision and recall and provides an overall measure of both. The F1 score ranges from 0 to 1, with 1 indicating the best possible performance and 0 indicating the worst possible performance.

$$F1_{score} = 2\frac{Precision.Recall}{Precision + Recall} \tag{4}$$

## 4. Results

The classification report for the dual CNN, shown in Table 2, highlights the method's effectiveness in multi-class skin cancer lesion classification. The report indicates that the highest accuracy was obtained for melanocytic nevi (nv), with a precision, recall, and F1 score are 1.00. In contrast, the lowest performance was recorded for actinic keratoses (vasc), where the precision was 0.94, the recall was 0.86, and the F1 score was 0.90.

We also evaluated the model's performance by calculating the accuracy, validation accuracy, accuracy loss, and validation loss. The results, as illustrated in the Figs. 3 to 5, indicate strong model performance, with the validation accuracy reaching 98% and the validation loss recorded at 0.09. Fig. 3 displays the confusion matrix for the dual CNN, Fig. 4 presents

**Table 3.** Comparing the proposed methodology with previous studies that used HAM 10000 dataset.

| Year | Ref | Techniques | Accuracy | Precision | Recall | F1-score |
|---|---|---|---|---|---|---|
| 2023 | [1] | CNN | 86.68% | NA | NA | NA |
| 2023 | [2] | CNN +Grad-CAM++. | 82% | NA | NA | NA |
| 2023 | [3] | CNN&LSTM&SVM | 88.24% | 88% | 88% | 88% |
| 2023 | [4] | DenseNet121 | 92.42% | NA | NA | NA |
| | | MobileNet | 90.29% | NA | NA | NA |
| | | Resnet50 | 89.93% | NA | NA | NA |
| | | VGG19 | 84.74% | NA | NA | NA |
| 2023 | [5] | VGG16 | 79.1 | 72% | 67% | 69% |
| | | ResNet50 | 87.9% | 83% | 82% | 82% |
| | | AlexNet | 86.2% | 81% | 76% | 78% |
| | | Dense Net | 86.9% | 80% | 81% | 81% |
| | | Mobile Net | 84.3% | 80% | 83% | 81% |
| | | InceptionV3 | 86.2% | 81% | 82% | 82% |
| | | VGG19 with Attention (Proposed) | 91.4% | 90% | 89% | 89% |
| 2023 | [6] | Dense Net +CBAM | 93% | 93% | 93% | NA |
| 2023 | [7] | Densenet169 (under sampling) | 91.20% | NA | NA | 91.70% |
| | | Resnet50 (over sampling) | 83% | | | 84% |
| 2023 | [8] | Mobile Net | 95% | 89.32% | 85.21% | 87.11% |
| 2023 | [9] | CNN | 86% | 66% | 48% | 55% |
| | | Mobile Net V2 | 96% | 63% | 47% | 53% |
| | | Resnet50 | 89% | 62% | 31% | 39% |
| 2023 | [10] | Xception, &Grad CAM++, & Grad CAM (XAI) | 90.24% | 91.83 | 89.18 | NA |
| 2023 | [11] | Inception&Resnet&v2 (with augmentation) | 95.09% | 95.49% | 95.16% | 95.27% |
| | | Inception&Resnet&v2 (without augmentation) | 83.59% | 81.17% | 67.14% | 72.31% |
| 2022 | [12] | EfficientNet | 87.9% | 88% | 88% | 87% |
| 2022 | [13] | CNN+DenseNet-201 | 82.58% | NA | NA | NA |
| 2022 | [14] | CNN | 0.72 % | 0.87% | 0.72% | NA |
| | | RNN | 0.69% | 0.72 % | 0.88% | NA |
| | | ResNet50 | 0.93% | 0.89% | 0.93% | NA |
| | | Xception | 0.79% | 0.89% | 0.89% | NA |
| 2021 | [15] | DenseNet201 | 95.09% | 93.46% | 91.57% | 92.51% |
| | | GoogleNet | 89.80% | 84.13% | 85.2% | 84.66% |
| | | Inception_ ResNetV2 | 89.02% | 78.90% | 91.15% | 84.58% |
| | | InceptionV3 | 93.02% | 93.88% | 84.38% | 88.88% |
| | | MobileNetV2 | 80.88% | 92.41% | 91.48% | 86.26% |
| 2021 | [16] | CNN | 91.28 | NA | NA | NA |
| 2020 | [17] | Transfer Learning Inception_ ResNet | 78.90% | NA | NA | NA |
| 2020 | [18] | ensemble learning (DenseNet201_ MobileNetV2_Xception_ EffiecientNetB0) | 96.4% | 90% | 92% | 89.1% |
| | | DenseNet 201 | 88.5% | 86.9% | 87.1 % | 84.7% |
| | | Xception | 91.3% | 89.8% | 88.3% | 89.5% |
| | | EffiecientNetB0 | 89.6% | 88.3% | 90.1% | 86.8% |
| | | MobileNetV2 | 85.9% | 81.4% | 86.2% | 86.3% |
| 2020 | [19] | Ensemble learning InceptionV3 + Xception | 91.56% | 82% | 84% | 83% |
| | | InceptionResNetV2 + Xception | 88.66% | 80% | 82% | 81% |
| | | InceptionResNetV2 + ResNeXt101 | 92.83% | 83% | 84% | 84% |
| | | InceptionResNetV2 + ResNeXt101 + Xception | 98.66% | 83% | 85% | 84% |
| 2024 | | Proposed approach (double dual CNN) | 98% | 97% | 93% | 95% |
| | | macro avg Weighted avg | 97.90% | 98% | 98% | 98% |

**Fig. 4.** Train and test accuracy.

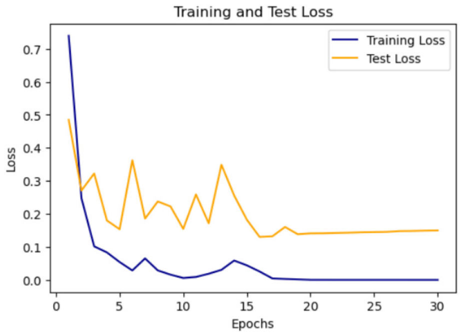

**Fig. 5.** Train and test loss.

the train and test accuracy, and Fig. 5 illustrates the train and test loss.

## 5. Comparative analysis

Compared to previous studies using the HAM10000 dataset, there are several limitations, as only accuracy was used as a measure of model efficiency. Other limitations include weak evaluation criteria and the lack of a classification report or confusion matrix to determine the percentage of false positives and false negatives. Also, some results were biased towards a certain class with classes having low classification scores. Our method showed superior classification results in all the used evaluation metrics, namely accuracy, reconnection, and F1_score. The classification report and confusion matrix also show the high performance of the proposed model with all the variables without bias. This improvement is due to the effectiveness of feature extraction from different CNN architectures that produced diverse features, in addition to the efficiency of linear algebra techniques that were used to eliminate the curse of dimensionality, which confirms the success of our approach despite dealing with data suffering from high imbalance. Table 3 compares our methodology with previous studies using the same dataset.

## 6. Conclusion and future work

In this study, a method for skin cancer classification based on features extracted from CNN algorithms with dimensionality reduction using linear algebra techniques with a fully connected classifier is presented. In the results of this study, significant progress has been made over previous related works using other techniques, thus providing a very good case for the proposed method. It has also been shown that the proposed method is capable of solving the problems inherent in big data, namely the most sensitive drawbacks, without using data augmentation methods that have been used in previous works to generate synthetic images. Furthermore, the challenges specific to skin cancer images, such as the presence of hair covering the skin, are effectively addressed using CNN filters and layers, eliminating the need for preprocessing steps such as hair removal, noise reduction, data cleaning, and other similar operations. The main limitation of this work is the time as there were four CNNs. Further research will attempt to integrate this feature extraction and classification technique into a comprehensive automated diagnostic system. Furthermore, the impact of this approach will be examined when compared to other critical performance evaluation settings in order to further validate the results and evaluate its applicability in actual clinical practice.

## Funding statement

This research received no external funding.

## Ethical approval

Not applicable.

## Conflict of interest

The author declares no conflict of interest.

## Data availability

The datasets generated during and/or analyzed during the current study are available in the [Kaggle] repository, https://www.kaggle.com/datasets/kmader/skin-cancer-mnist-ham10000.

## Author contribution

Raya Sattar Shahadh: Implementation of codes, analysis of results and comparison with previous works Belal Al-Khateeb: Design of the method and review of the paper.

## References

1. A. Mankawade, A. Bodhankar, A. Mahajan, D. Prasad, S. Mahajan, and R. Dhakalkar, "Skin cancer detection and intensity calculator using deep learning," *2023 Int. Conf. Adv. Technol. ICONAT*, pp. 1–4, 2023. doi: 10.1109/ICONAT57137.2023.10080001.
2. K. Mridha, M. M. Uddin, J. Shin, S. Khadka, and M. F. Mridha, "An interpretable skin cancer classification using optimized convolutional neural network for a smart healthcare system," *IEEE Access*, vol. 11, pp. 41003–41018. (April) 2023. doi: 10.1109/ACCESS.2023.3269694.
3. I. N. Nawrin, T. T. Trina, R. Rahman, and A. A. Rasel, "A comparative analysis of deep learning and hybrid models to diagnose multi-class skin cancer," *Int. Conf. Next-Generation Comput. IoT Mach. Learn. NCIM*, June 2023, pp. 1–6. doi: 10.1109/NCIM59001.2023.10212670.
4. B. Gunasundari, R. Thiagarajan, D. Bhavitha, R. Dharshini, and T. R. Madhunisha, "Deep convolutional neural networks and transfer learning for classifying malignant and benign skin cancer," *Proc. 2nd IEEE Int. Conf. Adv. Comput. Commun. Appl. Informatics, ACCAI.* pp. 1–8, 2023. doi: 10.1109/ACCAI58221.2023.10199438.
5. H. Singh, Suganya Devi K, S. S. Gaur, and R. Bhattacharjee, "Automated skin cancer detection using deep learning with self-attention mechanism," *Proc. Int. Conf. Comput. Intell. Sustain. Eng. Solut. CISES.*, pp. 464–469, 2023. doi: 10.1109/CISES58720.2023.10183586.
6. A. Jha and V. S. Ananthanarayana, "*An Interpretable Deep Learning Model for Skin Lesion Classification*," Springer Nature Switzerland. vol. 1848, CCIS 2023. doi: 10.1007/978-3-031-37940-6_44.
7. H. L. Gururaj, N. Manju, A. Nagarjun, Manjunath Aradhya VN, and F. Flammini, "DeepSkin: A deep learning approach for skin cancer classification," *IEEE Access*, vol. 11, no. May, pp. 50205–50214, 2023. doi: 10.1109/ACCESS.2023.3274848.
8. P. Shourie, V. Anand, and S. Gupta, "Seven class skin disease classification using pre-trained mobilenet network," *Proc. - 2023 Int. Conf. Innov. Eng. Technol. ICIET 2023*, pp. 1–5. doi: 10.1109/ICIET57285.2023.10220683.
9. M. N. Sandhua, A. Hussain, Al-Jumeily D, B. M. Mahmmod, and S. H. Abdulhussain, "Deep learning-based skin cancer identification," *Proc. - Int. Conf. Dev. eSystems Eng. DeSE*, vol. 2023-Janua, pp. 404–410. doi: 10.1109/DeSE58274.2023.10100194.
10. L. Gamage, U. Isuranga, De Silva S, and D. Meedeniya, "Melanoma skin cancer classification with explainability," *ICARC 2023-3rd Int. Conf. Adv. Res. Comput. Digit. Transform. Sustain. Dev.*, pp. 30–35, 2023. doi: 10.1109/ICARC57651.2023.10145622.
11. F. Bozkurt, "Skin lesion classification on dermatoscopic images using effective data augmentation and pre-trained deep learning approach," *Multimed. Tools Appl.*, vol. 82, no. 12, pp. 18985–19003, 2023. doi: 10.1007/s11042-022-14095-1.
12. K. Ali, Z. A. Shaikh, A. A. Khan, and A. A. Laghari, "Multiclass skin cancer classification using EfficientNets–a first step towards preventing skin cancer," *Neurosci. Informatics*, vol. 2, no. 4, p. 100034, 2022. doi: 10.1016/j.neuri.2021.100034.
13. M. R. Uddin and T. I. Mahmud, "Dense-par-AttNet: An attention based deep learning model for skin lesion classification by transfer learning approach," *4th IEEE Int. Conf. Artif. Intell. Eng. Technol. IICAIET*, pp. 1–5, 2022. doi: 10.1109/IICAIET55139.2022.9936758.
14. P. Singh, M. Kumar and A. Bhatia, "A comparative analysis of deep learning algorithms for skin cancer detection," *Proc. - 2022 6th Int. Conf. Intell. Comput. Control Syst. ICICCS*, no. Iciccs, pp. 1160–1166, 2022. doi: 10.1109/ICICCS53718.2022.9788197.
15. K. Thurnhofer-Hemsi and E. Domínguez, "A convolutional neural network framework for accurate skin cancer detection," *Neural Process. Lett.* vol. 53, no. 5, pp. 3073–3093, 2021. doi: 10.1007/s11063-020-10364-y.
16. O. Sevli, "A deep convolutional neural network-based pigmented skin lesion classification application and experts evaluation", *Neural Comput. Appl.*, vol. 33, no. 18, pp. 12039–12050, 2021. doi: 10.1007/s00521-021-05929-4.
17. P. BB, D. PB, and M. SB, "Inception network and transfer learning," vol. 1, pp. 536–545. doi: 10.1007/978-3-030-58799-4.
18. M. Kumar, S. Batra, G. S. Chakraborty, P. Kaur, and M. Mahajan, "An efficient deep learning based ensemble approach for multiclass skin cancer classification using medical imaging," *Proc. 2023 2nd Int. Conf. Augment. Intell. Sustain. Syst. ICAISS. 2023*, no. Icaiss, pp. 417–423. doi: 10.1109/ICAISS58487.2023.10250499.
19. S. S. Chaturvedi, J. V. Tembhurne, and T. Diwan, "A multiclass skin Cancer classification using deep convolutional neural networks," *Multimed. Tools Appl.*, vol. 79, no. 39–40, pp. 28477–28498, 2020. doi: 10.1007/s11042-020-09388-2.
20. A. A. Mukhlif, Al-Khateeb B, and M. A. Mohammed, "An extensive review of state-of-the-art transfer learning techniques used in medical imaging: Open issues and challenges," *J. Intell. Syst.*, vol. 31, no. 1, pp. 1085–1111, 2022. doi: 10.1515/jisys-2022-0198.
21. P-N Tan, M. Steinbach, and V. Kumar, "*Introduction to Data Mining (New International Editon)*," no. September. 2014.
22. M. Tavakol and A. Wetzel, "Factor analysis: A means for theory and instrument development in support of construct validity," *Int. J. Med. Educ.*, vol. 11, pp. 245–247, 2020. doi: 10.5116/ijme.5f96.0f4a.