CNNs for Automatic Skin Cancer Classification

Ahmed J. Abougarair[1*], Hesham A. Enheba[1], Mosab J. Abugarir[2], Shada E. Elwefati[3]
[1]Electrical and Electronics Engineering, University of Tripoli, Tripoli, Libya.
[2]Higher Institute of Medical Sciences and Technologies, Algarabulli, Libya.
[3]Biomedical Engineering, University of Tripoli, Tripoli, Libya.

* Corresponding author. Tel: (+218)916094184; Email: a.abougarair@uot.edu.ly.
Manuscript submitted July 11, 2024; revised August 6, 2024; accepted August 22, 2024.
doi: 10.18178/JAAI.2024.2.2.218-234

Abstract: Skin cancer is a serious public health issue, and successful treatment depends on an early and precise diagnosis. The capacity of Convolutional Neural Networks (CNNs) to automatically learn and extract significant characteristics from skin lesion photos has made them an effective tool for classifying skin cancer.

This paper provides an abstract on the use of CNNs in skin cancer classification and discuss the importance of training CNN models on diverse and comprehensive datasets, the architecture of CNNs, and their capability to capture intricate patterns and features in skin lesions. Moreover, we highlight the potential of CNNs in aiding dermatologists in the early detection and diagnosis of skin cancer. Furthermore, we identify several future directions for research, including the expansion of datasets, integration of clinical information, enhancement of model interpretability, exploration of transfer learning and evaluation of robustness against adversarial attacks. Overall, CNNs have demonstrated considerable promise in advancing skin cancer classification, leading to improved diagnostic accuracy and patient care.

Keywords: Artificial intelligence, CNN, Skin cancer.

1. Introduction

The largest organ in the body, the skin shields all internal organs from the hostile environment outside. It guards against infections and aids in temperature regulation. When a clump of healthy skin cells begins to expand uncontrollably and takes on new characteristics, it is known as a tumor. This tumor may be benign, which implies that there is no risk to life and that it cannot grow or spread to other regions, or malignant, which indicates that it is interfering with normal processes and has the potential to grow and spread to other body parts [1]. Fig. 1 display Cancer Registration Statistics, England - Office for National Statistics. Skin conditions such as skin cancer are becoming more prevalent due to the air pollution, which is increasing at an exponential rate. The Ozone Layer serves as the sole barrier against the sun's dangerous UV radiation, which are the primary cause of skin cancer. This layer is shrinking geometrically as a result of rising air pollution levels, which is why there are more incidences of skin cancer. Melanoma (malignant) and non-melanoma (benign) are the two main types into which skin cancer can be generally divided. One of the worst types of cancer is melanoma. Nonetheless, increasing the likelihood of survival can be achieved by detecting this malignancy early [2]. Clinical evidence indicates that there may be racial differences in the outcomes of skin cancer cases: Although the risk of developing melanoma is roughly 20–30 times higher in those with darker skin tones than in those with lighter skin tones, it has been found that depending on their skin tone, those with darker skin tones can have a lower or higher mortality risk for particular kinds of melanoma. Accurately identifying a skin lesion is crucial to administering the right treatment. This technique increases

the survival rate by enabling the early diagnosis of melanoma in dermoscopy images and photographs [3].

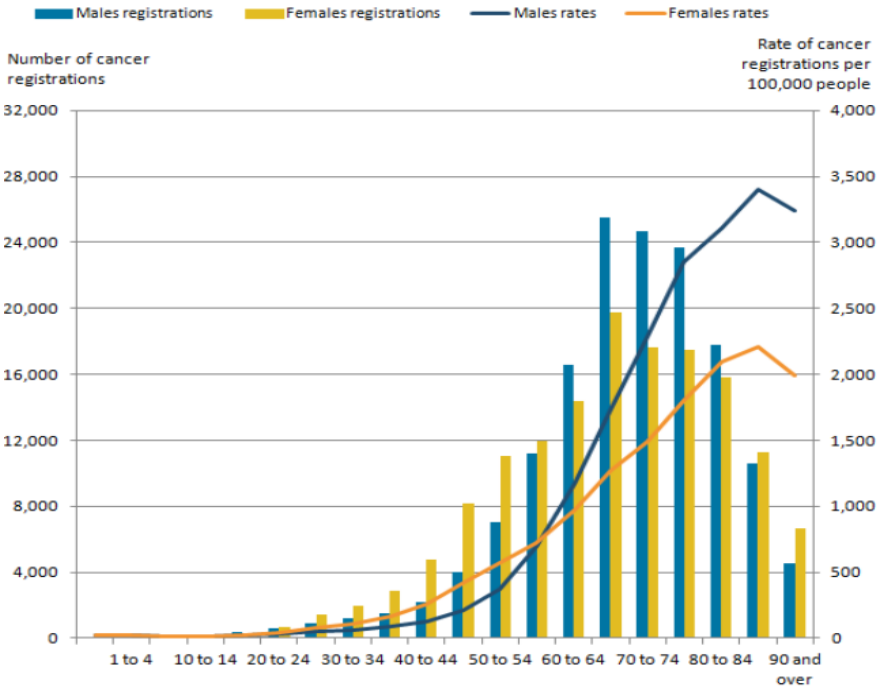

Fig. 1. Cancer Registration Statistics, England - Office for National Statistics.

Using deep learning to its full potential, it is possible to detect skin cancer through histopathology. Early detection and potential treatment of the disease can be aided by the classification of skin cancer lesions by specialized CNNs. Early symptoms of skin cancer include skin discoloration or inflammation, skin irritation and bleeding, moles, or red, waxy bumps on the skin caused by cancerous cells [4]. Uncontrolled cell division that has the potential to spread to other bodily areas is what defines cancer. It is divided into several kinds, such as liver, skin, breast, lung, and more. The most prevalent and often deadly type of cancer is skin cancer, which spreads quickly. The skin serves as the body's outermost layer of protection, thus any problems with it can have a substantial effect on the overall system. Skin cancer can arise from a variety of factors, including viruses, allergies, alcohol use, infections, physical activity, UV radiation exposure, changes in the environment, and odd growths. Skin cancer may be divided into two main categories: squamous cell carcinoma (SCC) and basal cell carcinoma (BCC) [5]. Sun exposure is the main source of these kinds, which frequently affect areas of the body such the hands, arms, and face that are exposed to the sun directly. While SCC can readily spread to adjacent organs and lymph nodes, BCC seldom affects these locations. Melanoma is another type of skin cancer that comes from melanocytes and is thought to be the deadliest kind. Early detection of melanoma improves treatment outcomes, while late detection might make survival more challenging. Though it may grow on any area of the body, melanoma is more frequent in white people. In men, it often appears on the trunk, and in women, it usually appears on the lower limbs [6]. Artificial intelligence (AI) holds promise in reducing the burden of skin diseases through early diagnosis. AI techniques like machine learning (ML) utilize various algorithms and models that can learn from training data to make predictions on new samples, a task that can be challenging for the human brain. Traditional ML models require the extraction of efficient features from skin images, and these models are often tailored for specific skin cancers rather than a generalized approach for different diseases. In contrast, DL models have demonstrated improved performance on large-scale datasets compared to ML models, without the need for extensive feature engineering. The key advantage of DL models is their ability to automatically learn relevant features from the data, eliminating the need for manual feature extraction, which can be a time-consuming and complex task. This makes DL models more

versatile and capable of handling a wider range of skin cancer types, rather than being restricted to specific cancer types like traditional ML models [7, 8].

Recent studies have compared the performance of ML-based systems to dermatologists for cancer identification, but further improvements are still required for effective healthcare applications. When designing DL models, considerations like dataset balancing and the use of high-resolution images are important, as they can increase training time and costs. Researchers have proposed various approaches combining computer vision and machine learning techniques for evaluating skin infections, achieving high accuracy. Additionally, studies have utilized advanced CNN architectures like GoogleNet V3 for classifying skin cancers using dermatoscopic and clinical skin images. Traditionally, early skin cancer detection has involved manually extracting various attributes from skin images, such as shape, texture, geometry, and other engineered features. These features are then used to train conventional machine learning models for the classification of skin cancer [9]. However, the manual feature engineering process can be labor-intensive and may not capture all the relevant information present in the skin images. To address this limitation, (CNNs) have emerged as a popular approach for identifying and analyzing medical images, including those related to skin cancer [10]. CNNs have demonstrated remarkable accuracy in the categorization of skin cancer types. The ability of CNNs to automatically learn relevant features directly from the input data, without the need for manual feature extraction, has made them a highly valuable tool in this field. The successful application of CNNs in skin cancer classification has highlighted their significant potential and value in improving early diagnosis and management of these conditions. The robust performance of CNN-based systems has underscored the advantages of this deep learning approach over traditional machine learning methods for medical image analysis tasks. The Objectives of the paper are:

1. Early detection: Creating a precise and effective categorization system that can identify skin cancer early on is one of the main goals. Skin cancer death rates can be decreased and treatment outcomes can be improved with early identification.

2. Accurate diagnosis: Accurate diagnosis of several forms of skin cancer, including melanoma, basal cell carcinoma, and squamous cell carcinoma, is the goal of CNNs. To obtain excellent precision and recall rates in skin cancer classification, the network is trained on large datasets of annotated skin lesion images.

3. Automated screening: By automating the screening of skin lesions, CNN-based categorization methods let medical personnel to evaluate a large number of cases with speed and accuracy. This goal is to shorten the time needed for diagnosis and increase dermatologists' efficiency.

4. Distinguishing benign lesions: CNNs can assist in distinguishing between benign and possibly cancerous skin lesions in addition to identifying malignant ones. By lowering pointless biopsies and treatments for benign cases, this goal will optimize healthcare resources.

5. Improving accessibility: CNN-based skin cancer classification can improve healthcare accessibility, especially in places where dermatologists and specialist medical facilities are few. People can receive initial examinations and assistance by utilizing these systems on mobile devices or remote diagnostic platforms, which enhances the potential for early intervention.

6. Personalized treatment recommendations: On the basis of the categorization outcomes, sophisticated CNN-based systems may be able to interact with patient information and medical records to offer individualized therapy suggestions. The purpose of this goal is to help medical practitioners choose the best course of action for each patient.

7. Continual learning and improvement: Accuracy and performance of CNN-based skin cancer classification systems can be continuously improved by allowing them to learn from new data. The goal is to create models with enhanced diagnostic capabilities that can adjust to emerging patterns, variances in skin lesions, and new information.

2. Related Work

By adding more convolutional layers to two pre-trained deep learning models, Zia *et al*. [11] improved the detection of skin cancer. The updated DenseNet201 model identified benign and malignant skin lesions with an astounding 95.50% accuracy rate. In the meantime, the enhanced MobileNetV2 model completed this assignment with an accuracy of 91.86%. Deep convolutional neural network (DCNN) models were used by Gouda *et al.* [12] to identify benign and malignant primary tumors. The accuracy percentage of the model they created was 83.2%. By way of contrast, the accuracy rates of other previous models were 84% for Inception ResNet, 85.8% for InceptionV3, and 83.7% for ResNet50. These stated performance metrics are the outcome of several experiment iterations. used a pre-trained AlexNet CNN to extract pertinent characteristics from the data in order to create a skin cancer prediction model [13]. Two methods for building a deep-learning architecture for hierarchical skin lesion identification were investigated by Barata *et al.* [14]. Where the accuracy and sensitivity were well. They trained, tested, and assessed their model using the ISIC 2017 dataset. Finding the three different kinds of skin lesions—nevus, keratosis, and melanoma—was the goal. Using the ResNet-18 deep learning algorithm and Explainable Artificial Intelligence (XAI) approaches, Natasha Nigar *et al.* [15] created a skin lesion categorization model. The International Skin Imaging Collaboration (ISIC) 2019 dataset was used to verify the model, and it performed remarkably well. The researchers employed the Local Interpretable Model-Agnostic Explanations (LIME) framework to produce visual explanations that complied with prior hypotheses and well recognized best practices for model interpretability in order to further evaluate the model's predictions. Using pictures of skin lesions, Ameri *et al.* [16] suggested a deep learning-based method for identifying skin cancer. A total of 3,400 photos from the HAM1000 dermoscopy image dataset were used in the development of their model. Cases of melanoma (860), dermatofibroma (115), benign keratoses (790), melanocytic nevi (795), actinic keratoses, and intraepithelial carcinoma (AKIEC - 327) are among those depicted in these photos. With the pre-trained AlexNet model, the researchers used transfer learning, and the outcome was an amazing Area Under the Receiver Operating Characteristic (ROC) curve of 0.91.

A CNN-based model for the identification and diagnosis of skin cancer was presented by Senthil Kumar *et al.* [17]. The ISIC dataset, which includes 2,637 photos of skin lesions, was used to train the algorithm. To prepare the data for the model, the researchers preprocessed the dataset using methods such picture resizing, reshaping, and transformation. The CNN model was then trained to identify and categorize the test pictures using the training dataset. The resultant model distinguished between malignant and normal skin cancer cells with an accuracy of 88%. A DNN-based classifier for the diagnosis of skin cancer from skin pictures was proposed by Stieler *et al.* [18]. Through their study, they showed how to better customize the AI system's explanations to the goal of diagnosing skin cancer, making them more domain-specific. To give interpretable explanations for the model's predictions, the researchers integrated the Local Interpretable Model-agnostic Explanations (LIME) framework with the well-established dermatological diagnosis approach known as the ABCD rule. The HAM10000 dataset, which includes microscopic pictures of typical pigmented skin lesions from many sources, was used to train the DNN-based classifier. Nevertheless, the accuracy that the suggested model achieves is not mentioned in the research. A fine-grained classification approach with feature discrimination served as the foundation for the lightweight skin cancer detection model created by Wei *et al.* [19]. The DenseNet77-based U-Net model is an enhanced deep learning-based method for skin lesion segmentation that was presented by Nawaz *et al.* [20]. In order to calculate a more representative collection of picture characteristics, the researchers used this method, introducing the DenseNet77 network as the encoder unit within the U-Net architecture. The U-Net model's decoder part then segments the important points that the encoder has retrieved. ISIC-2017 and ISIC-2018, two common datasets, were used to assess the suggested method, which produced segmentation accuracies of 99.21% and 99.51%, respectively. The

revised U-Net technique is resilient and can properly segment skin lesions, including moles of different hues and sizes, as confirmed by the quantitative and qualitative results. The DenseNet77-based U-Net model is an enhanced deep learning-based method for skin lesion segmentation that was presented by Nawaz *et al.* [20]. In order to calculate a more representative collection of picture characteristics, the researchers used this method, introducing the DenseNet77 network as the encoder unit within the U-Net architecture. The U-Net model's decoder part then segments the important points that the encoder has retrieved. ISIC-2017 and ISIC-2018, two common datasets, were used to assess the suggested method, which produced segmentation accuracies of 99.21% and 99.51%, respectively. Maqsood [21] study proposes a novel deep learning-based framework for multiclass skin lesion localization and classification. The framework combines feature extraction from multiple pre-trained deep learning models, followed by a novel feature fusion and selection approach to identify the most discriminative features for skin lesion analysis. The revised U-Net technique is resilient and can properly segment skin lesions, including moles of different hues and sizes, as confirmed by the quantitative and qualitative results. In order to overcome the issues of data scarcity and class imbalance in skin cancer diagnosis, Abayomi *et al.* [22] suggested a unique data augmentation approach based on the covariant Synthetic Minority Oversampling approach (SMOTE), the findings demonstrated a noteworthy enhancement in melanoma detection.

3. Skin Cancer

In general, aberrant states or illnesses that interfere with the body's regular functioning and cause physical or mental impairments are referred to as diseases. Numerous variables can contribute to its cause, such as immune system abnormalities, genetic mutations, environmental factors, lifestyle decisions, and infections like bacteria, viruses, fungus, and parasites [4]. Since there are billions of cells in the human body, cancer may start almost anywhere. Normally, when the body needs new cells, human cells divide to produce them. We refer to this process as growth and proliferation. New cells replace old ones when they age or get damaged [3]. Sometimes this well-ordered process goes awry, allowing damaged or abnormal cells to multiply and expand when they shouldn't. Benign tumors do not invade or spread to neighboring tissues. Cancerous tumors can regrow after excision, although benign ones typically don't. However, benign tumors can occasionally develop to be rather large. Fig. 2 present the different between normal cells and cancer cells. Malignant cells separate from the original tumor and move to different areas of the body via a complicated process known as metastasis, or the spread of cancer. Via the lymphatic system, circulation, or direct invasion of surrounding tissues, this can occur.

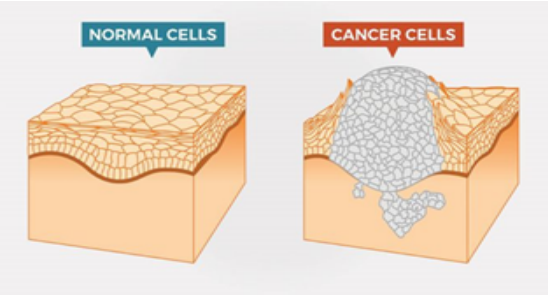

Fig. 2. Different between normal cells and cancer cells.

There are several main types of skin cancer as shown if Fig. 3, which are classified based on the specific type of skin cell where the cancer originates. The are most common types of skin cancer are [4]:

1. Basal Cell Carcinoma (BCC): BCC is the most often occurring kind of skin cancer. Although it can occur anywhere on the body, the face, scalp, neck, and arms—areas that are exposed to the sun—are where it typically manifests. The cause of BCC is the basal cells, which are found in the deepest layer of the skin's

epidermis, which is its outermost layer. Prolonged exposure to ultraviolet (UV) light from the sun or tanning beds is the main cause of BCC. UV rays cause DNA damage to skin cells, which promotes unchecked proliferation and the development of malignant lesions.

2. Melanoma (Mel): Melanocytes, the cells that make the pigment melanin, are the source of melanoma, a kind of skin cancer. While it is less prevalent than basal cell carcinoma and squamous cell carcinoma, it is more aggressive and has the potential to spread to other body areas if left untreated. Though it can sometimes form in other regions, such the eyes and mucous membranes, melanoma often develops on skin that has been exposed to the sun.

3. Melanomic Neves (NV): Atypical moles, often called melanoma nevi (NV), are moles that resemble melanoma, the most dangerous form of skin disease. They are thought to be melanoma precursors, which means that if treatment is not received, they may progress to melanoma.

4. Benign Keratosis (BKL): Benign keratosis (BKL) is a skin growth that is not malignant and can develop on different body areas. The cells that comprise the skin's outermost layer, known as keratinocytes, are the cause of it because of their unchecked proliferation. BKLs are extremely prevalent, particularly in the elderly. They typically don't cause any symptoms and are not communicable.

5. Actinic Keratosis (AKIEC): Actinic keratosis (AKIEC), sometimes referred to as sun keratosis, is a precancerous skin growth that, in the absence of treatment, can progress to squamous cell carcinoma (SCC). Prolonged exposure to ultraviolet (UV) light from tanning beds or the sun is the cause. People with fair skin, those with a history of sunburns, and those who reside in sunny climates are more likely to develop AKIECs.

6. Dermatofibroma (DF): Dermatofibroma (DF) is a slow-growing, benign skin tumor that usually affects the legs but can affect other body areas as well. The dermis, or main layer of the skin, has an excessive amount of fibrous tissue, which is the source of it. DFs typically show no symptoms and are not malignant.

7. Vascular Lesion (VASC): A localized region of aberrant blood vessels in the skin is called a vascular lesion (VASC). Numerous causes can contribute to these lesions, such as trauma, sun exposure, and heredity. Though they can occur anywhere on the body, VASCs are most frequently seen on the face, neck, and chest.

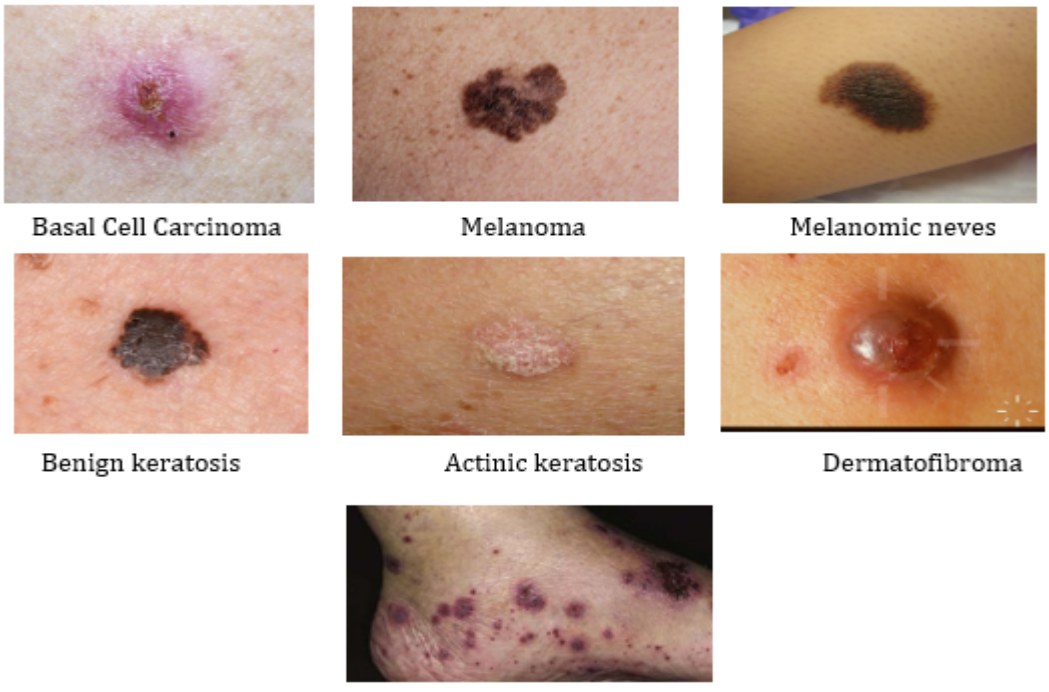

Fig. 3. Types of skin cancer.

4. Convolutional Neural Network (CNN)

One kind of artificial neural network that excels at interpreting visual data is the convolutional neural network (CNN). It draws inspiration from the structure and operation of the animal visual cortex, particularly from the neurons' receptive fields. CNNs are frequently utilized for computer vision tasks such as object detection, picture and video recognition, and others [23]. Convolutional, pooling, and fully linked layers are among the layers that make up the network. Convolutional layers process input data by applying filters, sometimes referred to as kernels, in order to extract local features and produce feature maps. The feature maps are downsampled by pooling layers, which lowers their dimensionality [24]. High-level predictions are made using learnt features by fully linked layers, which link all neurons from the previous layer to the subsequent layer. The main benefit of CNNs is that they can automatically learn hierarchical representations from unprocessed pixel data, doing away with the requirement for human feature engineering. CNNs are particularly good at image classification tasks because they can recognize local patterns and spatial correlations in an image by utilizing convolutional and pooling procedures [25, 26]. The procedure is shown in Fig. 4, which covers every stage from gathering the picture database to the last classification stage. Every step of the process is intended to solve certain issues related to skin lesion photos and enhance the CNN model's functionality.

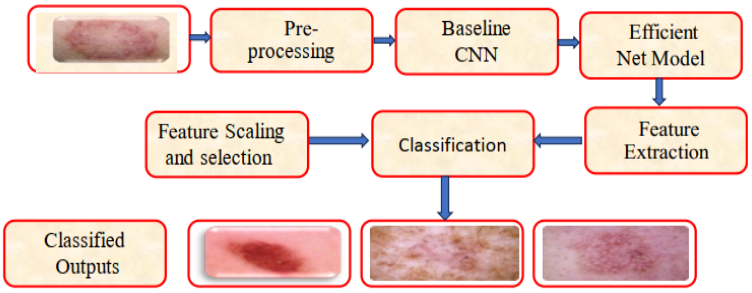

Fig. 4. Architecture Diagram of Methodology.

4.1. Architecture of Convolutional Neural Network

A specific kind of neural network called a CNN is made to analyze data having a grid-like layout, such pictures [27, 28]. Because they can automatically learn to extract pertinent features from the input data without the need for manual feature engineering, they are very useful in computer vision applications. As seen in Fig. 5 the standard architecture of a CNN comprises the subsequent essential elements.

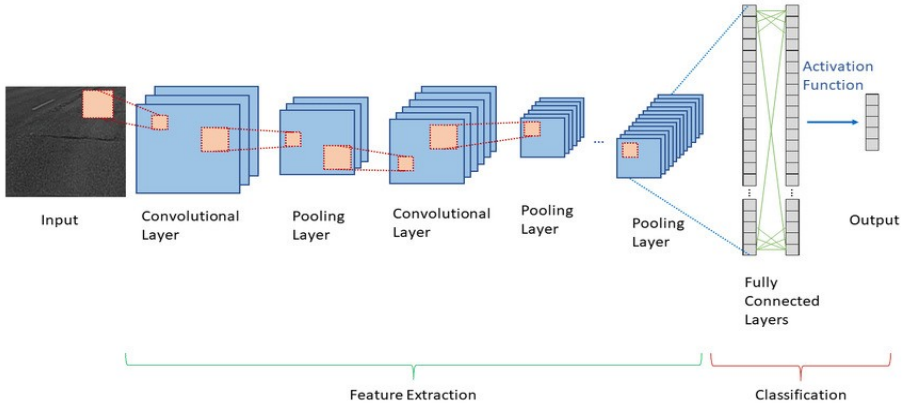

Fig. 5. Architecture of convolutional neural network.

Input Layer: The input data, which is usually an image, is received by this layer. The input data is represented as a three-dimensional tensor, with the first dimension denoting the image's height, the second

the image's width, and the third the number of channels (such as red, green, and blue) [24].

Convolutional Layers: The convolution layer is the core building block of the CNN. It carries the main portion of the network's computational load. This layer does a dot product between two matrices: the confined area of the receptive field is one matrix, and the other matrix is the set of learnable parameters, also referred to as a kernel. Compared to a picture, the kernel is more in-depth but less in space. This indicates that the depth fills all three channels if the image consists of three (RGB) channels, but the kernel height and breadth will be spatially little.

The kernel builds the image representation of the receptive area by sliding over the height and width of the image during the forward pass. This results in the creation of an activation map, a two-dimensional representation of the image that shows the kernel's response at each spatial location in the image. A stride is the name given to the kernel's sliding size. The output volume may be computed using the following formula given an input of dimensions $W \times W \times D$, a Dout number of kernels with a spatial size of F, stride S, and padding amount P:

$$W_{out} = \frac{W-F+2P}{S} + 1 \quad (1)$$

Pooling Layers: The pooling layer replaces the network's output at some points by computing a summary statistic from the outputs nearby. This helps reduce the spatial size of the representation, which in turn reduces the amount of computation and weights required. For the purpose of the pooling procedure, each slice of the representation is handled separately. Some examples of pooling functions include the L2 norm of the rectangle neighborhood, the average of the rectangular neighborhood, and a weighted average based on the distance from the center pixel. However, max pooling—which yields the neighborhood's maximum output—is the most often employed technique. With an activation map of size $W \times W \times D$, a pooling kernel of spatial size F, and stride S, we can use the following formula [28] to calculate the output volume size:

$$W_{out} = \frac{W-F}{S} + 1 \quad (2)$$

Fully Connected Layers: Every neuron in this layer is fully linked to every other neuron in the layer above and below, just like in a traditional FCNN. This means that the normal approach of matrix multiplication and bias effect may be used to compute it.

Output Layer: This layer generates the CNN's final output. A single neuron for binary classification, several neurons for multi-class classification, or a collection of bounding boxes for object identification can all be found in the output layer.

Activation Functions: To provide non-linearity to the feature maps, an activation function is used following the convolution procedure. CNNs frequently employ the sigmoid, tanh, and ReLU (Rectified Linear Unit) activation functions [29–33].

4.2. Mechanism of Action

Fig. 6 displays an activity diagram that illustrates the steps in the methodology's sequential order. The visual augmentation, image processing pipeline flow, evaluation of the EfficientNet model, and ultimate classification are all depicted in the figure. The gathering of the picture database is the first step. The intricate method employed in this investigation is briefly summarized in this illustration.

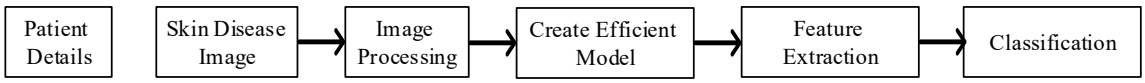

Fig. 6. Activity diagram.

1. **Feature Extraction**: CNNs are built with the ability to automatically identify and extract pertinent information from input photos. The network applies a series of learnable filters (kernels) to the input image

through convolutional layers, capturing local patterns and characteristics at various scales. These filters go across the picture, applying convolution operations to create feature maps that emphasize significant aspects of the image.

2. **Hierarchical Representation**: CNNs are trained to create feature representations that are hierarchical. Low-level elements such as edges, textures, and color gradients are captured by the network's lower layers. The network gains the ability to discriminate between various kinds of skin lesions and other more intricate and abstract qualities that are unique to the task at hand as the data moves through deeper levels. As the network develops, this hierarchical representation enables it to capture features that are increasingly discriminative.

3. **Non-linear Activation**: At each convolutional layer's output, activation functions like the Rectified Linear Unit (ReLU) are applied. The network can model complicated relationships and identify nonlinear patterns in the data thanks to these non-linearities, which also bring non-linear changes to the feature maps [34].

4. **Pooling and Downsampling**: CNNs often incorporate pooling layers, such as max pooling or average pooling, to downsample the feature maps while retaining the most salient information. Pooling helps reduce the spatial dimensions of the feature maps, making the network more computationally efficient and providing some degree of translation invariance.

5. **Classification**: In CNN design, fully connected layers are typically employed toward the end. These layers create a one-dimensional feature vector from the flattened and pooled feature maps. They allow the network to understand the relationship between the extracted features and the appropriate class labels by connecting each neuron from the previous layer to the next. The output layer delivers the final classification probabilities for each class, usually using a softmax activation function.

6. **Training and Optimization**: CNNs are taught via a technique known as backpropagation on labeled data. The network discovers the ideal combination of weights and biases during training in order to minimize a predetermined loss function. The parameters are iteratively updated using optimization algorithms, like stochastic gradient descent (SGD), to enhance the network's effectiveness in classifying skin cancer.

5. Dataset and Preprocessing

Data Source

1. Obtain a well-labeled dataset of skin cancer images, such as the ISIC 2019 Challenge dataset or the HAM10000 dataset [35].
2. To expand the dataset size and enhance model generalization, preprocess the photos by shrinking them to a consistent size and using data augmentation techniques (e.g., random rotation, flipping, scaling).
3. Divide the dataset, making sure there is no overlap between the training, validation, and test sets.

Model Architecture

1. Use a pre-trained (CNN) as the backbone, such as ResNet, VGG, or Inception, which have been successful in image classification tasks.
2. Fine-tune the pre-trained model on the skin cancer dataset by replacing the final classification layer with a new layer that matches the number of skin cancer classes.
3. Optimize the model's hyperparameters, such as learning rate, batch size, and number of epochs, using the validation set.

Training and Evaluation

1. Train the model using the training set, monitoring the performance on the validation set to prevent overfitting.

2. Evaluate the model's performance on the test set using appropriate metrics, such as accuracy, precision, recall, and F1-score, for each skin cancer class.
3. Analyze the model's performance and identify any class-specific errors or biases.

Interpretability and Visualization

1. Use techniques like Grad-CAM or Guided Backpropagation to visualize the regions of the input image that the model is focusing on to make its predictions.
2. Analyze the visual patterns and features that the model has learned to distinguish different skin cancer types.

Deployment and Clinical Integration

1. Develop a web or mobile application that allows users to upload skin images and receive automated skin cancer classification predictions.
2. Integrate the trained model into a clinical workflow, where it can be used as a decision support tool for dermatologists and physicians.
3. Monitor the model's performance in the real-world setting and continuously update the model with new data to maintain high accuracy.

6. Results

When a CNN is used to classify skin cancer, the predicted class labels for the input photos are usually displayed. Each class is given a probability or confidence score by CNN, which indicates the possibility that the image belongs to that particular class. The manner in which the outcome is displayed can vary based on the demands and particular use case. It can contain further information like the probability scores for each class, or it can be as straightforward as a binary classification result (benign or malignant). These probability scores can reveal information about the classification's degree of confidence. It is noteworthy that the precision and dependability of the findings obtained from the skin cancer classification are contingent upon the caliber of the training data, the CNN model's performance, and the variety of skin lesions included in the dataset. It is advised to assess the model's performance using suitable assessment metrics, such as recall, accuracy, precision, and F1-score, and to confirm the findings with dermatologists or other medical specialists .

The model's performance can be assessed by computing the following performance measures:

1. Accuracy: This is the overall proportion of accurate forecasts made by the model. Accuracy gives a broad idea of the model's ability to classify bacterial samples correctly.

2. Precision: Precision is defined as the proportion of actual positive predictions among all positive predictions made by the model. Precision is especially important when false positive forecasts are costly or undesirable.

3. Recall (Sensitivity): Recall quantifies the proportion of true positive events that are accurate positive forecasts out of all actual positive events.

4. F1-Score: This balanced statistic is computed as the harmonic mean of recall and accuracy, accounting for both metrics. The F1-score is useful when you want a single metric that accounts for the model's accuracy and recall.

5. Confusion Matrix: By displaying the quantity of true positives (TP), true negatives (TN), false positives (FP), and false negatives (FN), a confusion matrix gives a thorough assessment of the model's performance. This can assist in pinpointing the precise regions in which the model is having difficulty or excelling.

6. The receiver operating characteristic's area under the curve (ROC-AUC): The ROC-AUC score measures the model's ability to distinguish between positive and negative classifications. This statistic is useful when

dealing with imbalanced datasets or when adjusting the decision threshold. The five performance indicators listed below were used for the suggested classification model, as shown in Table 1 and Table 2. Table 3 typically display the performance classification for the skin lesion classification task .

Table 1. Performance Measurements

Measurement	Formula/ Description
Accuracy	$\frac{TN+TP}{TN+FP+TP+FN}$
Precision	$\frac{TP}{FP+TP}$
Recall	$\frac{TP}{FN+TP}$
F1 Score	$2 \times \frac{Precision * Recall}{Precision + Recall}$
AUC_ROC	Area under the receiver operating characteristic curve

Table 2. Model Training

Epoch	Accuracy	Val_accuracy	Loss	Val_loss
1	0.3727	0.3653	1.7076	1.5897
10	0.6117	0.7190	1.0220	0.8008
20	0.6894	0.7863	0.8263	0.5708
30	0.7452	0.8413	0.6820	0.4406
40	0.7850	0.8819	0.5784	0.3476
50	0.8079	0.8999	0.5162	0.3066
60	0.8305	0.9055	0.4687	0.2756
70	0.8467	0.9182	0.4173	0.2535
80	0.8618	0.9254	0.3802	0.2246
90	0.8736	0.9258	0.3509	0.2271
100	0.8826	0.9354	0.3244	0.2114

Table 3. Performance Classification

Class label	Support	Precision	Recall	F1-score
0	1359	0.95	1.00	0.97
1	1318	0.92	0.99	0.95
2	1262	0.89	0.91	0.90
3	1351	0.98	0.98	0.98
4	1374	0.94	0.67	0.78
5	1358	0.97	1.00	0.99
6	1365	0.86	0.96	0.91

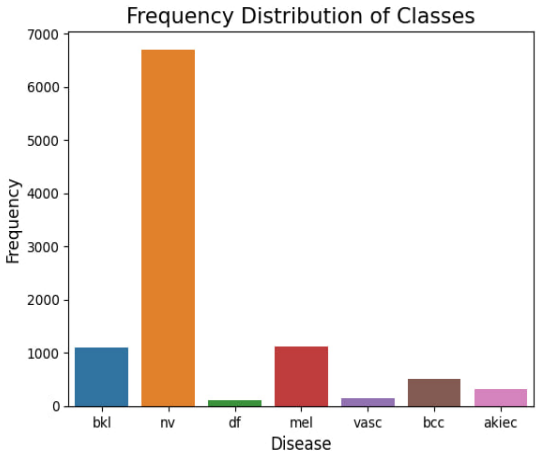

Fig. 7. Distribution of classes.

After taking the dataset and training the neural network, the following results were obtained. The Fig. 7 shows the frequency distribution of classes in a dataset. The x-axis shows the class labels, and the y-axis shows the frequency of each class. The most frequent class is nv, followed by bkl, mel, bcc, akiec, vasc, and df.

The Fig. 8 shows the distribution of age of patients. The x-axis is age and the y-axis is the number of patients. The histogram is centered around 50 years old, with a few outliers on either side.

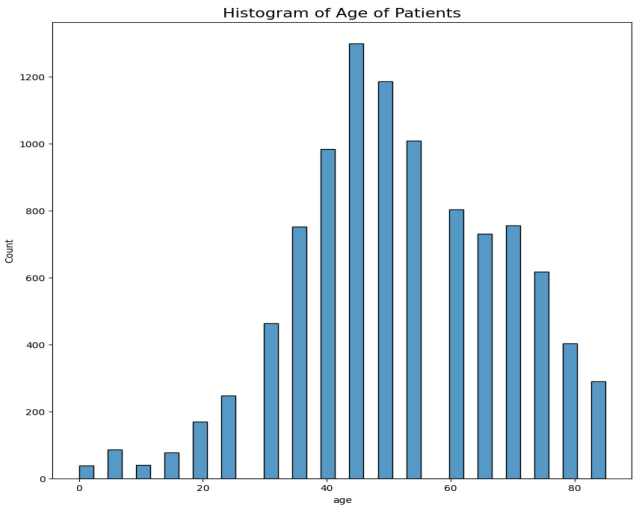

Fig. 8. Age of patient.

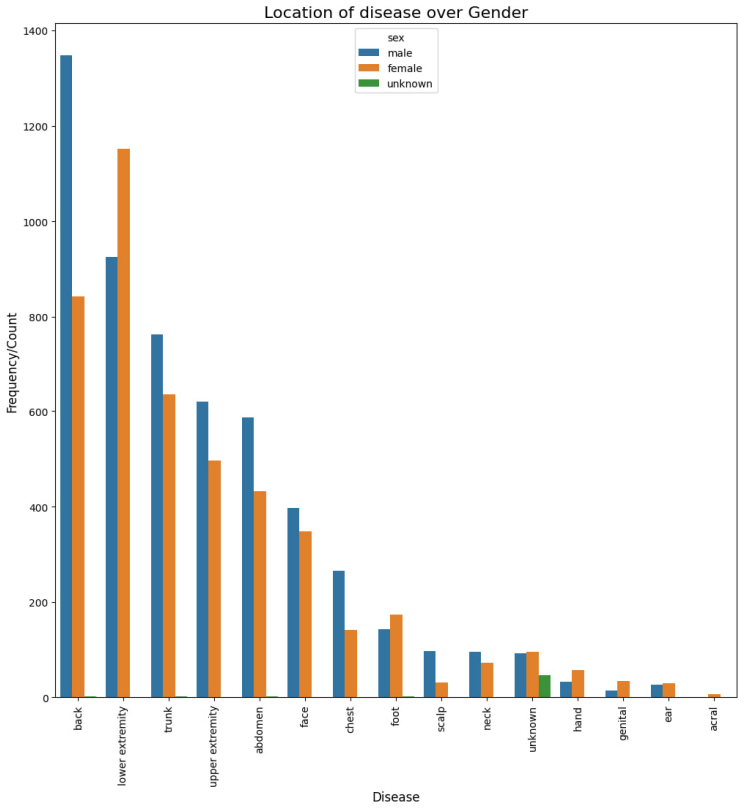

Fig. 9. Disease over gender.

The Fig. 9 showing the frequency of diseases by location on the body, and broken down by gender and whether the gender is known. We can see that the most common location for diseases is the back, followed by the lower extremity and the trunk. We can also see that males are more likely to have diseases on the back,

while females are more likely to have diseases on the lower extremity. Fig. 10 showing the gender of patients. 54% of patients are male, 45.5% are female, and 0.5% are unknown. The accuracy of a model across 100 epochs is displayed in Fig. 11. The orange line represents the validation accuracy, while the blue line represents the training accuracy. The model's accuracy rises with time, but after 40 epochs, the validation accuracy begins to fall. This indicates that the training data is being overfitted by the model. A neural network model's training and validation losses are depicted in Fig. 12. The model's loss on training data is referred to as the training loss, while the model's loss on validation data is referred to as the validation loss. There are 100 training epochs for the model. The model is learning as evidenced by the decreasing training loss over time. Additionally, as time goes on, the validation loss drops, indicating that the model is successfully generalizing to fresh data. The matrix of confusion seen in Fig. 13.

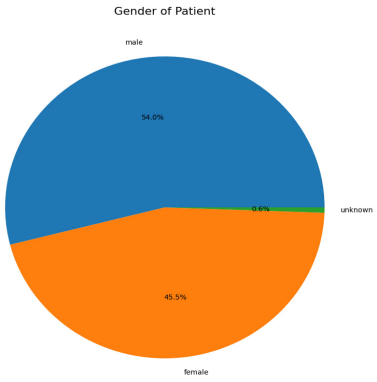

Fig. 10. Gender of patient.

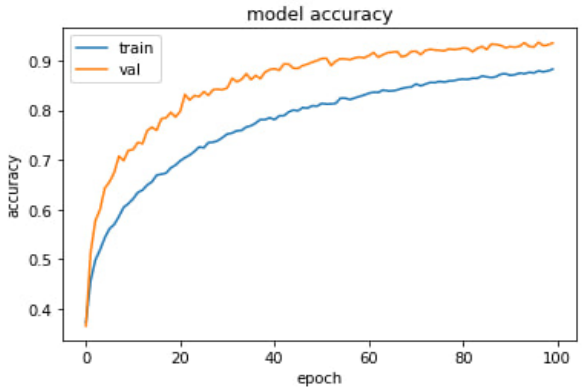

Fig. 11. Model accuracy.

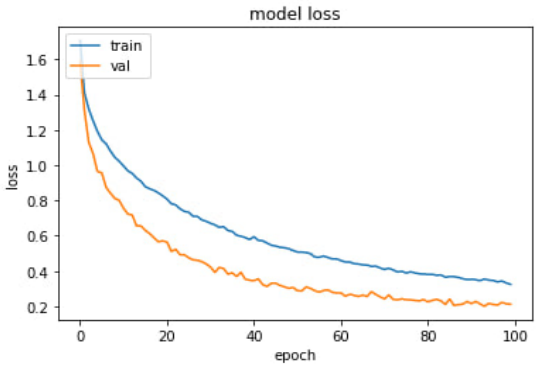

Fig. 12. Model loss.

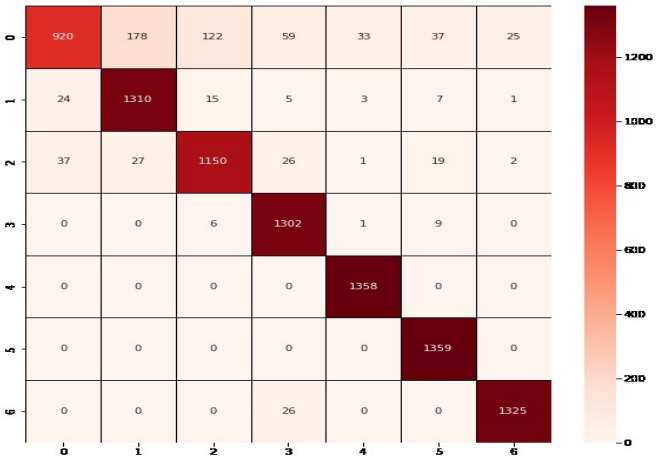

Fig. 13. Confusion matrix.

In the field of medical image analysis, CNNs have shown promise in the categorization of skin cancer. CNNs are very useful for image analysis because of their capacity to automatically discover and extract relevant characteristics from unprocessed pixel data. CNN models are trained on enormous datasets of photos of skin lesions that have been classified as benign or malignant in order to classify skin cancer. The CNN picks up on patterns and characteristics that set it apart from the other classes during training. These might include asymmetrical forms, varying hues, dissimilar textures, and other visual traits suggestive of skin cancer. CNNs' capacity to record hierarchical representations of pictures is one of their main features. The network is composed of several layers: fully connected layers carry out the final classification, pooling layers downsample the feature maps, and convolutional layers use filters to identify local patterns. Because of its hierarchical structure, the CNN is able to acquire ever more intricate representations of the incoming data, which helps it provide precise predictions. A big and varied dataset is needed to train a CNN for the classification of skin cancer. This dataset needs to comprise a variety of skin lesion kinds, representing a range of states and looks. The amount and diversity of the dataset can be increased by using data augmentation techniques like rotation, scaling, and flipping, which will enhance the network's capacity to generalize to new data.

The CNN can be trained to identify fresh photos of skin lesions. After the input image is fed into the network, the probability distribution over the various classifications (benign or malignant) is provided by the output layer. For that particular image, the class with the highest probability is regarded as the anticipated class. Analyzing the CNN's performance is essential to determining how reliable it is. Evaluation criteria including accuracy, precision, recall, and F1-score are frequently used. While CNNs have demonstrated encouraging results in the classification of skin cancer, it is crucial to remember that they should not take the position of medical professionals. The goal of using CNNs is to support dermatologists in the early identification and detection of skin cancer.

7. Conclusion

In summary, the application of CNNs for skin cancer classification has demonstrated significant potential in precisely differentiating between benign and malignant skin lesions. CNNs are ideal for this task because of their capacity to automatically learn and extract pertinent features from unprocessed image data. CNN models may be trained on vast and varied datasets of photos of skin lesions to enable them to identify patterns and traits that may be suggestive of skin cancer. Because of their hierarchical structure, CNNs may accurately classify data by capturing progressively sophisticated representations of the input. It's crucial to remember, though, that CNNs shouldn't take the role of medical professionals' knowledge. Rather, they can be an important aid to dermatologists in the early identification and detection of skin cancer. CNNs can be used to classify skin cancer by analyzing vast amounts of photos, which could increase efficiency and help identify cases that need more investigation or treatment. CNNs will be much more effective at classifying skin cancer if developments in model interpretability, dataset quality, and CNN designs continue. The goal of this field's ongoing research and development is to improve and optimize these models in order to improve patient outcomes, initiate therapies earlier, and manage skin cancer more effectively overall.

8. Future Work

Develop more sophisticated explainable AI techniques to provide insights into how CNNs make decisions and identify the specific features they use for classification.

A robotic arm or platform could be used to capture high-quality, standardized images of skin lesions. This automated imaging system could help reduce variability and improve the quality of the input data for the deep learning-based classification model [36–38].

Explore the use of attention mechanisms to highlight the most relevant regions of the image for the model's prediction.

Develop CNN models that can continuously learn and adapt to new data and changing environments, improving their performance over time.

Explore the use of active learning techniques to select the most informative data points for further training and improve the efficiency of the learning process.

Encourage collaboration between researchers, clinicians, and industry experts to accelerate progress in skin cancer classification using CNNs.

The control system could be integrated with a deep learning-based classification model to create a closed-loop control system [39–43].

References

[1] AlSalman, S., Alkaff, T., Alzaid, T. & Binamer, Y. (2018). Nonmelanoma skin cancer in Saudi Arabia: Single center experience, *Ann. Saudi Med.*, *38(1)*, 42–45.

[2] Nehal, Y. & Bichakjian, C. (2018). Update on keratinocyte carcinomas, *New England J. Med., 379(4)*, 363–374.

[3] Abougarair, A., Oun, A., Sawan, S., Abougard, T. & Maghfiroh, H. (2024), Cancer treatment precision strategies through optimal control theory, *Journal of Robotics and Control, 5(5)*,1261-1290.

[4] *Key Statistics for Melanoma Skin Cancer*, American Cancer Society, Atlanta, GA, USA, 2022.

[5] Albahar, M. (2019). Skin lesion classification using convolutional neural network with novel regularizer, *IEEE Access, 7,* 38306–38313.

[6] Asan, M. et al. (2021). Comparative analysis of skin cancer (benign vs. malignant) detection using convolutional neural networks, *J. Health Eng*.

[7] Abougarair, A. & Elwefati, S. (2023). Identification and control of epidemic disease based neural networks and optimization technique, *International Journal of Robotics and Control Systems, 3 (4)*, 780-803.

[8] Abougarair, A., Oun, A., Sawan, S., & Ma'arif, A. (2024). *Deep learning-based automated approach for classifying bacterial images, international journal of robotics and control systems, 4(2)*, 849-876.

[9] Barata, C., Marques, J., & Celebi, M. (2019). Deep attention model for the hierarchical diagnosis of skin lesions, *Proceedings of IEEE/CVF Conf. Comput. Vis. Pattern Recognit. Workshops (CVPRW)), (pp. 1–9)*.

[10] Soyer, H., Argenziano, G., Ruocco, V., & Chimenti, S. (2001). Dermoscopy of pigmented skin lesions (Part II), *Eur. J. Dermatol., 11(5)*, 483–498.

[11] Rehman, M. *et al.* (2022). Classification of skin cancer lesions using explainable deep learning, *Sensors, 22(18.*

[12] Gouda, W. *et al.* (2022). Detection of skin cancer based on skin lesion images using deep learning, *Healthcare, 10(7)*.

[13] Dorj, U. *et al.* (2018). The skin cancer classification using deep convolutional neural network, *Multimedia Tools Appl., 77(8)*, 9909–9924.

[14] Barata, C. & Marques, J. (2019). Deep learning for skin cancer diagnosis with hierarchical architectures, *Proceedings of IEEE 16th Int. Symp. Biomed. Image (ISBI)) (pp. 841–845).*

[15] Nigar, N. *et al.* (2022). A deep learning approach based on explainable artificial intelligence for skin lesion classification, *IEEE Access, 10*, 113715–113725.

[16] Ameri, A. (2020). A deep learning approach to skin cancer detection in der-moscopy images, *J. Biomed. Phys. Eng., 10(6)*, 801–806.

[17] Kumar, R. *et al.* (2022). Skin cancer detection using deep learning, *Proceedings of Int. Conf. Electron.*

Renew. Syst. (ICEARS)), (pp. 1724–1730).

[18] Stieler, F., Rabe, F., & Bauer, B. (2021). Towards domain-specific explain- able AI: Model interpretation of a skin image classifier using a human approach, *Proceedings of IEEE/CVF Conf. Comput. Vis. Pattern Recognit. Work- shops (CVPRW)), (pp. 1802–1809).*

[19] Wei, L., Ding, K., & Hu, K. (2020). Automatic skin cancer detection in der-moscopy images based on ensemble lightweight deep learning network, *IEEE Access, (8*), 99633–99647.

[20] Nawaz, M. *et al.* (2022). A framework of improved DenseNet77 and UNET convolutional neural network, *Int. J. Imag. Syst. Technol., 32(6)*, 2137–2153

[21] Maqsood, S., & Damaševičius, R. (2023). Multiclass skin lesion localization and classification using deep learning-based features fusion and selection framework for smart healthcare, *Neural Netw., 160*, 238–258.

[22] Abayomi, S. *et al.* (2021). Malignant skin melanoma detection using image augmentation by oversampling in nonlinear lower-dimensional embedding manifold, *TURKISH J. Electr. Eng. Comput. Sci.*, *29(SI-1)*, 2600–2614.

[23] Elwefati, S. *et al.* (2021). Control of epidemic disease based optimization technique, *Proceedings of 2021 IEEE 1st International Maghreb Meeting of the Conference on Sciences and Techniques of Automatic Control and Computer*, MISTA 2021.

[24] M. Alshaibi, M. *et al.* (2024). Blood cells cancer detection based on deep learning, *Journal of Advances in Artificial Intelligence (JAAI), 2(1)*.

[25] Edardar, M. *et al.* (2021). Adaptive neural networks based robust output feedback control for linearized system, *WSEAS Transactions on Computer Research, 9*, 125-136.

[26] Ajagbe, S. *et al.* (2021). Multi-classification of Alzheimer disease on magnetic resonance images (MRI) using deep convolutional neural network (DCNN) approaches, I*nt. J. Adv. Comput. Res., 11(53)*, 51–60.

[27] Ganan, H. *et al.* (2021). Implementation of a Brain-Computer Interface for Robotic Arm Control, 2021 IEEE 1st International Maghreb Meeting of the Conference on Sciences and Techniques of Automatic Control and Computer Engineering (MI-STA2021).

[28] Can, O. *et al.* (2024). Blood cells cancer detection based on deep learning, *Journal of Advances in Artificial Intelligence, 2*(*1*).

[29] Tilamon, H. *et al.* (2023). Artificial pancreas control using optimized fuzzy logic based genetic algorithm, *International Robotics & Automation Journal, 9(2).*

[30] Aburakhis, M. *et al.* (2022). Design and implementation of smart voice assistant and recognizing academic words, *International Robotics & Automation Journal, 8(1).*

[31] Abougarair, A. (2020). Neural networks identification and control of mobile robot using adaptive neuro fuzzy inference system, *Proceedings of the 6th International Conference on Engineering & MIS.*

[32] Arebi, W. *et al.* (2022). Smart glove for sign language translation. *International Robotics & Automation Journal, 8(3)*.

[33] Abougarair, A. (2022). Optimal control synthesis of epidemic model, *IJEIT International Journal on Engineering and Information Technology, 8(2)*.

[34] Abougarair, A. (2023). adaptive neural networks based optimal control for stabilizing nonlinear system, *Proceedings of 2023 IEEE 3rd International Maghreb Meeting of the Conference on Sciences and Techniques of Automatic Control and Computer Engineering (MI-STA2023)*.

[35] Tschandl, P., Rosendahl, C., & Kittler, H. (2018). The HAM10000 dataset, a large collection of multi-source dermatoscopic images of common pigmented skin lesions, *Scientific Data*, *5*.

[36] Oun, A. *et al.* (2024). Comparative evaluation of pid controller tuning through conventional and genetic algorithm, *Proceedings of 2024 IEEE 4th International Maghreb Meeting of the Conference on Sciences and*

Techniques of Automatic Control and Computer Engineering (MI-STA).

[37] I. Attawil *et al.* (2024). Lateral control of autonomous vehicles through adaptive model predictive control, model predictive control for optimizes battery charging process, *Proceedings of 2024 IEEE 4th International Maghreb Meeting of the Conference on Sciences and Techniques of Automatic Control and Computer Engineering (MI-STA)*.

[38] Abougarair, A. *et al.* (2023). Analysis of mobile accelerometer sensor movement using machine learning algorithms, *Proceedings of 2023 IEEE 3rd International Maghreb Meeting of the Conference on Sciences and Techniques of Automatic Control and Computer Engineering*.

[39] Arif, A. *et al.* (2022). Sliding mode control design for magnetic levitation system, *Journal of Robotics and Control (JRC), 3(6)*.

[40] Emhemmed A. *et al.* (2022). Simulation analysis and control of wireless power transfer for implantable medical devices, *Proceedings of IEEE 21st International Conference on Sciences and Techniques of Automatic Control and Computer Engineering*.

[41] Abougarair, A. (2022). Real time classification for robotic arm control based electromyographic signal, *Proceedings of 2022 IEEE 2st International Maghreb Meeting of the Conference on Sciences and Techniques of Automatic Control and Computer Engineering*.

[42] Aburakhis, M. *et al.* (2022). Adaptive neural networks based robust output feedback controllers for nonlinear systems, *International Journal of Robotics and Control Systems, 2(1)*, 37-56.

[43] Elmolihi, A. (2020). robust control and optimized parallel control double loop design for mobile robot, *IAES International Journal of Robotics and Automation (IJRA), 9(3)*.