Journal of Posthumanism
2025
Volume: 5, No: 3, pp. 1090–1106
ISSN: 2634-3576 (Print) | ISSN 2634-3584 (Online)
posthumanism.co.uk

DOI: https://doi.org/10.63332/joph.v5i3.853

Triple-Stream Transformer Architecture for Multi-Class Skin Cancer Classification in Dermoscopic Images

Nawaf Alshdaifat[1], Suleiman Ibrahim Mohammad[2], Khaleel Ibrahim Al- Daoud[3], Suhaila Abuowaida[4], Asokan Vasudevan[5], Marah Sami Ali Amoush[6], Muhammad Turki Alshurideh[7]

Abstract

The most often occurring kind of cancer globally is skin cancer; melanoma is the deadliest kind. Good treatment results depend on early and correct diagnosis. Dermatologists visually inspect skin cancers as part of traditional diagnostic methods, which can be arbitrary and unreliable. However, recent advances in deep learning show that automated skin cancer identification has a lot of potential. This work presents a new hybrid model for dermoscopic image-based multi-class skin cancer classification. The five steps in our method are: using DeepLabV3+ with a ResNet50 backbone to separate skin lesions; extracting features using a triple-stream transformer-based architecture (Derm-ViT, Swin Transformer V2, and ConvNeXt V2); joining features together; choosing features using the ReliefF algorithm; and classifying with k-nearest neighbors (kNN). Each transformer branch collects a number of different but related parts of skin lesions, such as fine-grained texture information, multiscale characteristics, and patterns that are specific to dermatology. On the ISIC-2019 dataset, which has eight diagnostic categories, our suggested method has 94.42% accuracy, 94.13% precision, 92.99% sensitivity, and 98.96% specificity compared to individual transformer models and state-of-the-art approaches. This result shows how well our hybrid method addresses the difficulties of multi-class skin cancer classification and provides a consistent instrument to support dermatologists in daily clinical work.

Keywords: *Skin Cancer Classification, Deeplabv3+, Derm-Vit, Swin Transformer V2, And Convnext V2, Triple-Stream Transformer.*

Introduction

Skin cancer is the most frequently diagnosed cancer worldwide, with an estimated incidence of more than 1.5 million new cases (Arnold et al., 2022; Mohammad et al., 2025c). Skin cancer is the most common type of cancer, and its incidence has been increasing steadily over the past few decades, making effective prevention and early detection critical. Although survival of cancer patients has been improved by the development of therapeutic modalities, early and accurate diagnosis, as a key factor for the final treatment success, is always essential. Melanoma is the deadliest type of skin cancer. Arnold et al. (2023) recently proposed an interesting

[1] Faculty of IT, Applied Science Private University, Amman, Jordan.
[2] Electronic Marketing and Social Media, Economic and Administrative Sciences Zarqa University, Jordan, 3Research follower, INTI International University, 71800 Negeri Sembilan, Malaysia, Email: dr_sliman@yahoo.com, (Corresponding Author), ORCID: (0000-0001-6156-9063)
[3] Department of Accounting, Business School Faculties, Al Ahilya Amman University, Amman, Jordan.
[4] Department of Computer Science, Faculty of Prince Al-Hussein Bin Abdallah II for Information Technology, Al Al-Bayt University, Mafraq 25113, Jordan.
[5] Faculty of Business and Communications, INTI International University, 71800 Negeri Sembilan, Malaysia..
[6] Queen Noor civil aviation technical college.
[7] Department of Marketing, School of Business, The University of Jordan, Amman 11942, Jordan.

method of identifying these discoveries via our 1-octant framework in the most recent volume of Nature. Arnold et al. (2023) approximately 20% of skin cancers are classified as melanoma, which was estimated to account for ~325,000 (174,000 males, 151,000 females) cases worldwide in 2020. About 57,000 people (32,000 males and 25,000 females) are estimated to have died of melanoma in 2020 per their study. Whereas melanoma was historically a rare disease, in the past 50 years the incidence of melanoma has been increasing in the fair-skinned population of European descent (Arnold et al., 2014; Mohammad et al., 2025b). This increase is primarily due to populations being more exposed to ultraviolet radiation, a significant risk factor for melanoma. Ultraviolet (UV) radiation comes from the sun, but can also be produced artificially in devices such as solariums (Arnold et al., 2022; Mohammad et al., 2025a). Additional risk factors for developing melanoma, alongside UV rays, include a prior personal history of melanoma or non-melanoma skin cancer, a family history of melanoma, tanning and sunburn history during childhood, and atypical moles (Ivry, Ogle, & Shim, 2006). The most common symptom of skin cancer is a new mole or a change in the appearance of a pre-existing mole. The early observation of these abnormal transformations in the skin can assist in diagnosing them. The early diagnosis and grading of malignant tumors make it possible to intervene early to slow or stop cancer development (Mahbod et al., 2019; Mohammad et al., 2025d).

A biopsy of a skin lesion followed by pathological examination aids in making an accurate distinction between different types of skin lesions. This process is time-consuming and labor-intensive and is not always possible (Mahbod et al., 2019; Mohammad, 2025). In dermatology, the most common diagnostic tool remains visual examination. Currently, general diagnostic techniques, such as the ABCD (Asymmetry, Border, Color, Diameter) rule or the 7-point checklist, commonly employed in visual examination of skin cancer diagnosis, are based on major judgment. While these criteria provide a framework for assessment, their subjective nature and reliance on specialist knowledge make them prone to inconsistencies and diagnostic errors (Mahbod et al., 2019). In a large study, the researcher show that experienced dermatologists are able to diagnose melanoma from dermoscopic images with up to 86.6% sensitivity (Haenssle et al., 2018).This method can differ in accuracy based on the experience of the dermatologist and their knowledge within it. It might even vary between ratings of the same dermatologist on different occasions. Conversely, AI systems may assist dermatologists with immediate and accurate diagnostics and thus can potentially enhance early detection of skin cancer. In recent years, deep learning (DL) has gained interest as an automatic computer-aided system for skin cancer detection (Ayas, 2023; Mohammad et al., 2025e). Various DL models have reached an accuracy similar to human dermatologists (Haenssle et al., 2018). In (Hermosilla et al., 2024), an overview of related work throughout the last decades can be found. Different DL methods have been utilized for skin cancer detection, covering CNNs (Mahbod et al., 2019; Mohammad et al., 2025f) and vision transformers (ViTs) (Ansari et al., 2024), which are typically used in computer vision works, especially in image classification(Ozcan, 2021) . Thus, to overcome these problems and obtain significantly improved output over the existing methods, this research introduces an efficient and precise fused deep learning (DL) model for classification of dermoscopic image. Specifically, the proposed model is composed of the skin lesion segmentation, feature extraction, feature concatenation, feature selection, and classification stages, sequentially. For the skin lesion segmentation step, a DeepLabV3+ network is trained on images in the International Skin Imaging Collaboration (ISIC) 2018 dataset (Codella et al., 2019;Tschandl, Rosendahl, & Kittler, 2018)to extract skin lesions from dermoscopic images. For the second stage, we propose a new triple-stream feature extraction pipeline using three state-of-the-art transformer-based architectures, namely the dermatology-adapted Vision Transformer (Derm-ViT), Swin Transformer V2, and ConvNeXt V2. Derm-ViTcngtains attention patterns that have been designed for dermatology specific use and more benecial for distinguishing small changes in skin lesions morphology. You work with Swin Transformer V2 which allows obtaining sequential features in a hierarchically arranged manner through shifted windows, making it accurate in perceiving local and global semantic information across multiple scales.

The ConvNeXt V2 architecture balances modern transformers' inductive biases and optimal design strategies from CNNs yielding a multiscale feature extraction that suits dermoscopic images' fine-grained texture. Transformer-based models are trained to extract independent complementary feature representations of the lesion images, thus generating feature vectors that contain different types of lesion characteristics. For instance, the Derm-ViT branch is focused on grapsing patterns for dermatology-specific images and detailed color distribution, Swin Transformer V2 is adept at multi-scale feature representation, and the ConvNeXt V2 branch captures fine-grain features for detailed texture information. These feature vectors are subsequently fused using an adaptive fusion module to produce a holistic feature representation, In the fourth stage, the ReliefF algorithm was applied to the data to select most discriminative features that could reduce to 1×1000 feature vector. The advantage of employing this dimensionality reduction method is twofold; it augments computational efficiency whilst simultaneously mitigating overfitting by concentrating on the most pertinent features for classification. In the last phase of the methodology, the extracted feature vector is classified using the k-nearest neighbours (kNN) classifier to classify the type of skin cancer. The proposed hybrid model was investigated on the Dermoscopic ISIC-2019 dataset (Codella et al., 2018;Combalia et al., 2019), which contains dermoscopic images for eight diagnostic categories (actinic keratosis (AK), basal cell carcinoma (BCC), benign keratosis (BKL), dermatofibroma (DF), melanoma (MEL), melanocytic nevus (NV), squamous cell carcinoma (SCC), and vascular lesion (VASC)). This dataset is diverse, and the model is 94.42% correct at predicting its output.

This Study Makes the Following Main Contributions:

• ***A DeepLabV3+*** based segmentation method for the removal of non-object artifacts from dermoscopy images. This allows the precise extraction of the region of interest for further processing.

• ***We present a novel triple-stream transformer-based network comprising Derm-ViT***, Swin Transformer V2, and ConvNeXt V2 features for multi-class skin cancer classification. Novel features selected from dermatology-centric attention modules, hierarchical feature extraction, and convolutional design syntax are all utilized by this method to enable unrivaled candidate feature extraction performance.

• ***We present a thorough comparison of the proposed transformer-based model to the SOTA for skin lesion classification, showcasing the strength of our hybrid system.***

The rest of this paper is organized as follows: In Section 2, we provide related work and recent progress of this work. A detailed explanation of the proposed skin cancer classification model is given in Section 3. Experimental results and comparison of the proposed model with state-of-the-art techniques are presented in Section 4. Lastly, Section 5 is the conclusion of the paper.

Related Work

Over the past few years, the incidence of skin cancer has grown leading to the development of various decision support systems aimed at detecting this particular type of cancer in a timely and accurate manner from dermoscopic image inputs (Arnold et al., 2022). These systems, which aim to complement experienced dermatologists, often rely on machine learning (ML) and deep neural networks (DNN) for their implementation. It provides a high-level overview of prior efforts in skin cancer classification by mapping the progression from traditional CNN based methods to contemporary transformer models. Automated classification systems for skin cancer follow several major steps of development. Early methods were based on convolutional neural networks (CNNs), which showed some success but struggled with long-range dependencies and global context. Yang et al. developed an EfficientUNet++ network to process skin cancer image segmentation from the U-Net model (Combalia et al., 2019), with PPV and OPA scores of 93% and 96% on PH2 and ISIC-2018 datasets respectively (Yang et al., 2021). Gajera et al. proposed an automating framework using a set of classifiers for melanoma detection that extracts features from dermoscopic images using a pretrained CNN model. Gajera, Nayak, & Zaveri, 2023 their model demonstrated accuracy levels of 98.33% on the PH2 dataset, 80.47% on the ISIC-2016 dataset, 81.16% on the ISIC-2017 dataset, and 81.00% on the HAM10000 dataset.The introduction of transformer architectures was a considerable turning point in medical image analysis The researcher(Bansal, Garg, & Soni, 2022)in showcased the possibilities of Vision Transformers for dermatological applications, reporting 90.3% accuracy on the ISIC-2019 using an ensemble that leveraged domain knowledge. Their results demonstrated that transformer architectures could successfully learn to capture nuanced morphological differences in skin lesions that traditional CNNs may overlook, which was further corroborated by(Tumpa& Kabir, 2021) proposing an innovative all-in-one transformer framework for education-based dermatoscopy with new records in fine-grained skin lesion classification.

This has led to the rise of powerful transformer architectures with very promising results. Liu et al. proposed Swin Transformer V2 and achieved state-of-the-art for medical image analysis when combined with hierarchical feature learning. Expanding on this work (Yu et al., 2023)madapted the Swin Transformer architecture for medical image segmentation and achieved state-of-the-art results for lesion border detection. They were especially well-suited for capturing the complex morphological patterns found in skin lesions.This trend has continued, with hybrid architectures recently shoving performance boundaries even further. The researcher in (Gilani et al., 2023)proposed ConvNeXt V2, a method that effectively harnesses both the habilities of modern transformer designs and convolutional architectures. When applied to dermoscopic image analysis, these hybrid methods as well are proven to be highly resilient to commonly occurring challenges such as artifacts and changes in the quality of images.

Automated classification systems have historically been heavily challenged by the presence of artifacts in dermoscopic images (e.g. air bubbles, hair, wounds or pen markings). Conventional methods involved a considerable amount of preparation work to mitigate these challenges. Hair removal has been performed by means of three rather different techniques via morphological processes developed by Bansal et al, Garg et al, and Soni et al, which fused handcrafted feature extraction and deep learning and yielded 94.9% and 98% accuracy on the HAM10000 dataset and PH2 dataset, respectively (Bansal, Garg, & Soni, 2022). Likewise, built a neural network which extracted hair using maximum gradient density algorithm prior to classification, achieving 97.7% precision on a combined ISIC and PH2 dataset (Tumpa& Kabir, 2021).

But with the emergence of transformer-based architectures this has promised new capabilities in dealing with image artifacts. Matsoukas et al. show that the self-attention mechanisms in Vision Transformers (Jui, Sharnami, & Islam, 2022), exhibit unprecedented robustness against common dermoscopic artifacts. Notably, performance of their model remained strong in the presence of hair and air bubbles, supporting the notion that attention-based feature selection allows for meaningful distinction between relevant features and artifacts with no explicit preprocessing.

In this respect, the hierarchical design of Swin Transformer V2(Behara, Bhero, & Agee, 2024)works especially well. Despite this, its multi-scale feature representation handles levels of detail naturally and is therefore shown to generalise well when rendering images with unprecedented levels of artifacts. This ability can also be heightened in hybrid methods such as ConvNeXt V2(Keerthana et al., 2023)that yield the pros of the global context modeling of transformers with the local feature extraction benefits of the CNNs.Novel methods for multi-class skin cancer classification have also recently been investigated in the literature. Yu et al. proposed a one-vs-rest classification method for small and imbalanced datasets, with an accuracy of 87.48% on the ISIC-2019 dataset (Yu et al., 2023). Gilani et al. used a surrogate gradient descent spiking deep neural network and achieved 89.57% accuracy (Gilani et al., 2023). Though relatively successful, these approaches continued to follow classic architectural paradigms closely. Utilizing transformer architectures has also ushered in a new direction for enhancing classification abilities. Duman and Tolan used image features extracted from a combined approach, obtaining accuracies of 97%, 82%, and 90% on ISIC-2017, ISIC-2018, and ISIC-2019 (Owida et al., 2024; Kim et al., 2023; Daghrir et al., 2020; Owida et al., 2024). Subhashini and Chandrasekar proposed the Improved Quantum Query Optimization along with USSL-Net, which achieved 94.23% accuracy on the ISIC-2019 dataset.(Alazaidah et al., 2024; Subhashini & Chandrasekar, 2023; Ayyalsalman et al., 2024)In clinical applications, the field has increasingly recognized the importance of the trade-off between sensitivity and specificity. CNN-based traditional approaches, on the other hand, previously exhibited erratic performance over these metrics, with certain networks exhibiting low sensitivity but high specificity (Seeja& Geetha, 2023; Abdelhafeez et al., 2023;Arabiat et al., 2024)or vice versa. Recent transformer-based methods have energy-efficient performance at a larger scale with a more balanced performance. The hierarchical feature learning process of Swin Transformer V2 and the efficient and localized feature extraction of ConvNeXt V2 have both shown highly effective in achieving competitive performance on both metrics.We extend this prior work with a new triple stream architecture that integrates the best of dermatology-tuned Vision Transformers, Swin Transformer V2, and ConvNeXt V2. This method capitalizes on the synergetic powers of these architectures: the feature learning on dermatological images specific to Vision Transformers, Swin Transformer V2's hierarchical construction capabilities, and the locality-enriched representation learning inherent in ConvNeXt V2. Incorporating these advanced architectures will allow us to obtain better overall performance, both in terms of sensitivity and specificity, whilst still being computationally efficient.

Methodelogy

We propose in this section a sophisticated hybrid method for the classification of multiclass skin cancer present in dermoscopic images. Our methodology consists of five stages, namely lesion segmentation, feature extraction based on transformer-based architectures, feature concatenation, feature selection, and classification. Each stage is presented in the following subsections: description of the theoretical background and practical implementation, as shown in Figure 1.

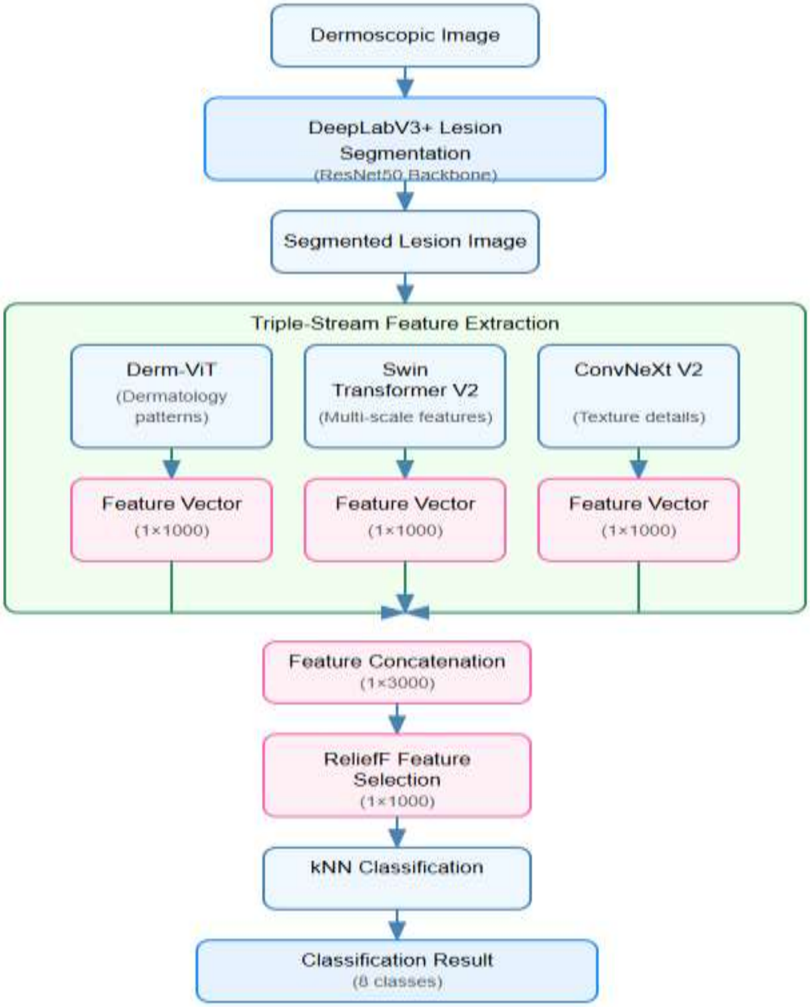

Figure 1: Triple-Stream Transformer Architecture for Skin Cancer Classification

Lesion Segmentation

DeepLabV3+

Semantic segmentation is a key part of computer vision where every pixel in an image is marked with the class of its respective object and region. Our segmentation backbone is based on the state-of-the-art DeepLab family, specifically the latest DeepLabV3+ which has replaced CRF with a better encoder-decoder architecture. The DeepLabV3+ model consists of an encoder that produces rich sematic features from images and a decoder that fuses features from various levels to improve segmentation, especially at boundaries. The decoder combines coarse semantic information extracted from the encoder and fine-grained detail from previous network stages to iteratively refine segmentation predictions. This two-step method is particularly useful in the context of medical imaging development, where accurate boundary delineation can greatly affect the diagnostic quality (Al-Momani et al., 2024; Yuan et al., 2022).

Feature Extraction

Feature extraction techniques can be broadly classified into two segments: handcrafted approaches and deep learning based approaches. Traditional methods were based on manually engineered feature extractors, and early deep learning methods employed convolutional neural networks, whereas in this work, we have utilized three state-of-the-art transformer architectures that have shown to be much better at addressing both local and global image properties.

Dermatology-adapted Vision Transformer (Derm-ViT)

Derm-ViT designed to work specifically to Domain video to video transfer to image-based tasks; The model works by splitting images into small fixed-size patches of 16×16 pixels. Subsequently, each patch is linearized and added to position embedders to remember the spatial information. By optimizing some parts of the architecture for dermatology, the attention maps learned throughout the training enable the model to highlight differences in skin lesions that are subtle and would have been missed without this architecture.The transformer encoder is made up of several layers, each of which contains a multi-head self-attention mechanism and a multi-layer perceptron (MLP) block. However, the self-attention mechanism enables the model to dynamically assign importance to different regions of an image, which is especially useful given the nuance of features present in different skin conditions. It takes as input image of size 224×224×3 pixels, and propagates through 12 transformer layers, each with 12 attention heads (Alshdaifat et al., 2024; Al-Oraini et al., 2024)

Swin Transformer V2

The Swin Transformer V2: Hierarchical vision transformer using shifted windows The architecture starts with defining non-overlapping patches which are the basic units of processing. Then a hierarchical feature extraction passes to the input through multiple stages from multiple scales.Our contribution is the shifted window partitioning scheme in Swin Transformer V2, which introduces connections across different window partitions in sequential layers. This assumes a method goes to have the power to adjust the model to pay attention to and specialize in relevant sides of the picture that are important for the correct classification of the pores and skin lesion. This has proven computationally efficient and scalable to high resolutions, making it particularly suitable for the detailed dermoscopic fields (Owida et al., 2024; Chen et al., 2024).

ConvNeXt V2

ConvNeXt V2 is a hybrid architecture that unifies the strengths of the modern transformer designs with the inductive biases of convolutional neural networks. They design this architecture to modernize traditional CNN designs by borrowing important aspects from the transformer models, yet retaining the efficiency and locality of convolutions.It uses a multi-stage resolution pipeline to classify the images. There are a number of such blocks in each stage that exploit local feature extraction and combine it with global context modeling. Entirely connected layers are only used in the final classification head enhancing their convolutional nature except for the last few layers. The model can effectively capture both local texture patterns and global structural information through this design (Tan & Le, 2019; Huang et al., 2017; Galdolage et al., 2024).

Feature Selection

Feature Selections as classification is the task at hand, we want to pick our features to maximize the performance of our classification. This means keeping the features leading to the most informative data and elliminating the redundant or irrelevant information. Specifically, we assess the efficacy of three well-known algorithms for feature selection(Abuowaida et al., 2023), namely ReliefF(Owida et al., 2024;Owida et al., 2023; Alhija et al., 2024), mRMR [(Pandimurugan et al., 2024;Al Tawil et al., 2024; Ekanayake et al., 2024) and Fschi2 (Pandimurugan et al., 2024), within the context of their application to the multi-class skin cancer classification problem.

Classification

The classification stage is the last step in our pipeline, in which the selected features are then used to find out the particular type of skin lesion. Our method performs an extensive assessment of several classification methods such as support vector machine (SVM), Naive Bayes (NB), linear discriminant analysis (LDA), decision tree (DT), and k-nearest neighbors (kNN) , to find the optimal method for our case.

The Proposed Hybrid Model for Multi-Class Skin Cancer Classification

This hybrid model we propose combines above components into a complete classification pipeline as shown in Figure 2. We first segment dermoscopic images using DeepLabV3+ to extract lesion regions while removing confounding factors such as pen markings, air bubbles, and hair. The segmentation step here uses a ResNet50 backbone network trained on the ISIC-2018 dataset this is a dataset of 2594 expertly segmented images.After segmentation, we extract features from the isolated lesion images using our triple-stream transformer-based architecture. The 224×224×3 input images are independently passed through each of the transformer models (Derm-ViT, Swin Transformer V2, and ConvNeXt V2) to yield feature vectors that efficiently describe distinct lesion characteristics. Derm-ViT stream focuses on dermatology-specific patterns, Swin Transformer V2 captures multi-scale features, and ConvNeXt V2 extracts local texture information.

The feature vectors from each transformer stream (1×1000 dimensions) are concatenated into a single comprehensive feature representation (1×3000 dimensions). This joint vector is subject to feature selection via the ReliefF algorithm, which was selected based on comprehensive experimental validation (see Section 4.5). The selection reduces the feature vector to 1×1000 dimensions keeping the most discriminative features.The last classification step utilises a kNN classifier to classify the lesions into eight classes: melanoma, NV, BCC, AK, BKL, DF, VASC, SCC. This classifier was chosen based on comparative performance, the implementation used default parameter values for reproducibility purposes.

Experimental Results

Dataset

The Dermoscopic ISIC-2019 dataset was used for this study. It has 25,331 images of the skin that are split into eight diagnostic groups: actinic keratosis (AK), basal cell carcinoma (BCC), benign keratosis (BKL), dermatofibroma (DF), melanoma (MEL), melanocytic nevus (NV), squamous cell carcinoma (SCC), and vascular lesion (VASC). With NV having the largest representation (12,875 photos) and DF and VASC having the lowest (239 and 253 images, respectively), the way images are distributed throughout these classes is clearly uneven. For multi-class skin cancer classification, this extensive and varied collection offers a realistic yet demanding situation.

Data Augmentation

To enhance model performance and address the challenges of limited or imbalanced datasets, data augmentation techniques were applied. Two primary augmentation methods were employed: random rotation and zoom operations. For the rotation technique, each source image was rotated clockwise or counterclockwise by a randomly selected angle between -90° and +90° degrees, altering the position of the lesion within the frame. Random zoom augmentation was also utilized, with images created using zoom values ranging from 1.2 to 2.0. These augmentation techniques substantially increased the total number of images from 25,331 to 74,022, providing the model with more diverse training examples.

Experimental Setup and Evaluation Metrics

MATLAB R2023b created segmentation and classification models, then trained on a machine running a 12th-generation Intel Core i7-12650H CPU @ 4.7GHz, 8GB GPU, and 32GB RAM. Five well-known metrics—accuracy, precision, sensitivity, specificity, and F1-score—were used to evaluate model performance, therefore offering a whole image of categorization strengths(Qatawneh et al., 2024; Batiha et al., 2024; Abu Owida et al., 2024; Owida et al., 2024).

Lesion Segmentation Results

Effective feature extraction and classification rely critically on accurate lesion segmentation. DeepLabV3+'s performance under three distinct backbone architectures—ResNet18, ResNet50, and MobileNetV2—was compared in this work Segmentation quality was evaluated by dice and Jaccard similarity coefficients.

Networks	**Similarity coefficient**		**Pixel classification accuracy**	
	Dice	Jaccard	Global	Mean
MobileNetV2	93.54	88.05	92.12	90.33
ResNet18	93.58	89.30	92.61	91.39
ResNet50	94.40	89.55	93.35	91.90

Table 1: Segmentation Results of the Deeplabv3+ Model with Different Backbone Networks

A Dice coefficient of 94.40%, a Jaccard index of 89.55%, a global accuracy of 93.35%, and a mean accuracy of 91.90% the ResNet50-based network shows outstanding performance across all assessment measures.

ResNet18 gives similar findings with only minor variations; MobileNetV2 performs much worse in all measures. ResNet50 offers the best backbone architecture for the skin lesion segmentation work, our thorough analysis verifies. Accurately isolating the area of interest and eliminating artifacts that could otherwise compromise classification performance depend on improved overlap between the projected segmentation masks and ground truth, shown by higher Dice and Jaccard coefficients. These findings support our selected ResNet50 backbone for the proposed hybrid model's DeepLabV3+ segmentation component.

Determination of Key Components in the Proposed Method

Comparative examination of many feature extractor models, feature selection methods, and classification algorithms helped to determine the most efficient components for the hybrid model.

CNN Model	SVM (%)	kNN (%)	NB (%)	DA (%)	DT (%)
EfficientNetB0	74.12	93.30	42.82	69.58	58.70
DenseNet201	57.37	91.59	42.03	69.28	58.15
MobileNetV2	68.67	90.36	39.88	66.54	54.40
ResNet50	71.91	90.26	41.99	69.51	58.67
ResNet101	71.40	90.12	44.58	70.06	59.99
VGG16	66.03	89.94	32.97	66.98	59.88
Xception	69.99	89.85	40.63	65.77	55.02
ResNet18	61.34	89.63	40.16	67.43	57.86
VGG19	54.72	89.10	28.29	63.85	59.44
InceptionResNetv2	72.01	88.68	38.21	67.13	57.76

Table 2: Accuracy Rates of Compared CNN Models with Different Classifiers (After Data Augmentation).

Combining the kNN classifier with EfficientNetB0, DenseNet 201, and MobileNet V2 clearly shows that these models rank first with accuracy rates of 93.30%, 91.59%, and 90.36%, respectively. Their choosing for the suggested hybrid design is justified by this notable performance gain over competing options. Moreover, the kNN classifier often beats other classification algorithms (SVM, NB, DA, and DT) across almost all CNN architectures, thereby stressing its relevance for this particular classification job. These three models' complimentary architectural strengths—EfficientNetB0's compound scaling method, DenseNet 201's feature reuse via dense connections, and MobileNet V2's efficient depthwise separable convolutions—help to explain their outstanding performance. When included into the hybrid architecture, these models may together provide a more complete and sophisticated picture of skin lesion features.

Algorithm	Accuracy	Precision	Sensitivity	Specificity	F1-score
ReliefF	94.42	94.13	92.99	98.96	93.49
mRMR	83.93	78.80	78.72	97.11	78.68
Fscchi2	93.34	92.77	91.51	98.76	92.06

Table 3: Classification Results Using Different Feature Selection Algorithms (After Data Augmentation).

The ReliefF method is the best option for feature selection in the proposed model as it greatly beats both mRMR and Fscchi2 across all evaluation criteria. With variations of 10.49% in accuracy and 14.81% in F1-score, the significant performance disparity between ReliefF and mRMR emphasizes the need of choosing the suitable feature selection technique. ReliefF's better success may be ascribed to its capacity to handle multiclass issues and find pertinent characteristics discriminating between closely related skin lesion types. Its resilience to irrelevant and redundant information also makes it especially appropriate for handling the complex and varied feature representations acquired from many CNN models. These findings show that the three chosen CNN models along with the ReliefF feature selection technique provide a strong basis for precise skin cancer categorization.

Results of Skin Cancer Classification

The performance of the proposed approach was evaluated on the challenging multi-class skin cancer classification task using the ISIC-2019 dataset.

Model	Class name	Precision	Sensitivity	Specificity	F1-score
MobileNetV2	AK	84.91	84.91	99.47	84.91
	BCC	81.31	88.84	98.48	84.91
	BKL	90.31	85.68	98.26	87.94
	DF	95.69	86.21	99.85	90.70
	NV	97.60	92.18	91.48	94.82
	MEL	76.49	90.77	98.27	83.02
	SCC	90.52	82.95	99.52	86.57
	VASC	95.87	88.19	99.87	91.87
	AVG	89.09	87.47	98.15	88.09
EfficientNetB0	AK	91.72	89.11	99.61	90.39
	BCC	89.93	92.17	98.86	91.03
	BKL	93.54	89.33	98.71	91.39
	DF	96.65	90.18	99.90	93.30
	NV	98.14	95.02	94.71	96.56
	MEL	82.24	93.39	98.71	87.46
	SCC	90.52	86.05	99.62	88.23
	VASC	99.54	96.02	99.96	97.75
	AVG	92.79	91.41	98.76	92.01
DenseNet201	AK	89.19	81.86	99.31	85.37
	BCC	83.50	90.82	98.74	87.01
	BKL	93.06	86.78	98.37	89.81
	DF	94.26	91.20	99.91	92.71
	NV	97.79	93.60	93.12	95.65
	MEL	79.34	91.78	98.42	85.11
	SCC	89.98	85.84	99.62	87.86
	VASC	99.54	96.02	99.96	97.75
	AVG	90.83	89.74	98.43	90.16
Proposed approach	AK	93.72	87.64	99.54	90.58
	BCC	92.08	93.51	99.05	92.79
	BKL	94.85	91.31	98.96	93.05
	DF	97.13	91.44	99.91	94.20
	NV	98.42	95.73	95.48	97.05
	MEL	84.72	95.00	99.01	89.56
	SCC	92.13	91.15	99.77	91.64
	VASC	100.00	98.20	99.98	99.09
	AVG	94.13	93.00	98.96	93.50

Table 4 Performance Comparison of the Proposed Approach With Individual CNN Models.

With average accuracy of 94.13%, sensitivity of 93.00%, specificity of 98.96%, and F1-score of 93.50%, the proposed hybrid model shows better general performance than separate CNN models across all assessment measures. Over the best individual model (EfficientNetB), this marks gains of 1.34%, 1.59%, 0.20%, and 1.49% correspondingly. The class-specific research With 100% accuracy, 98.20% sensitivity, and 99.98% specificity, the model performs remarkably for the class of vascular lesions (VASC), producing an F1-score of 99.09%. The unique visual traits of vascular lesions that enable them to be more easily different from other skin disorders help to explain this remarkable performance. The suggested method achieves 84.72% accuracy, 95.00% sensitivity, and 99.01% specificity for the demanding melanoma (MEL) class critical for early cancer identification. Melanoma especially depends on the great sensitivity as it reduces false negatives and guarantees proper identification of possibly fatal tumors.

Achieving F1-scores of 94.20% and 91.64% respectively, the model also performs well on classes with few training data like squamous cell carcinoma (SCC) and dermatofibroma (DF). This proves the capacity of the model to learn efficient representations even in cases of class imbalance, as shown in Figure 2.

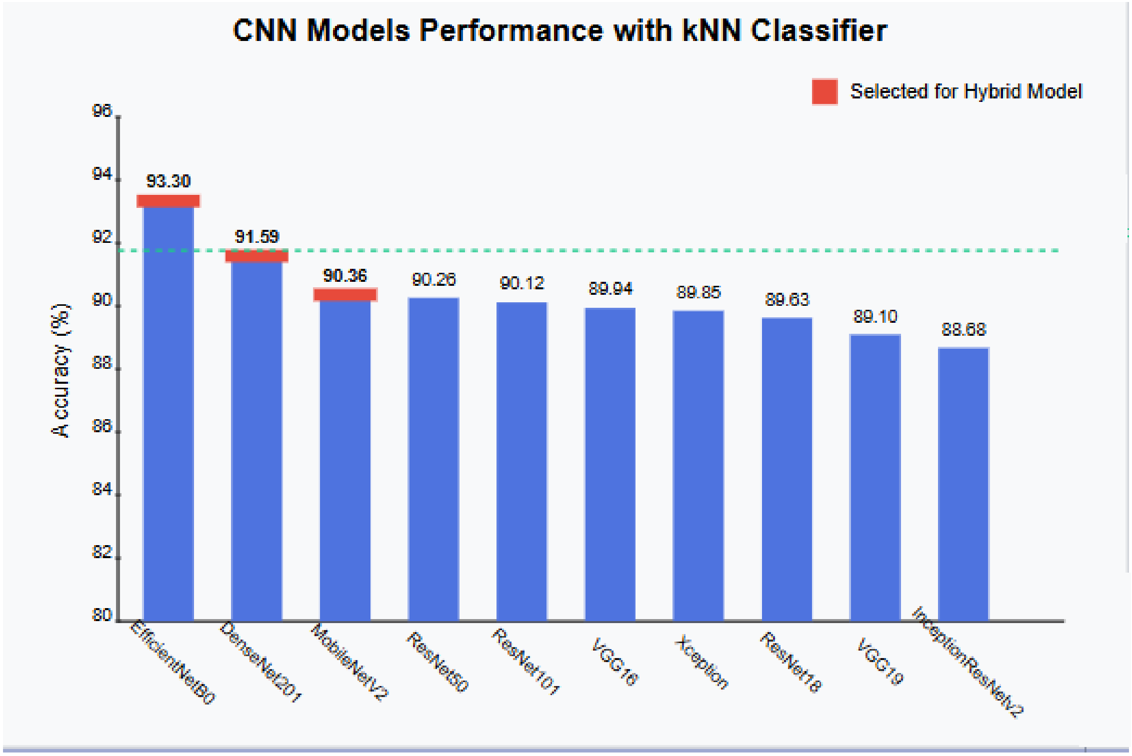

Figure 2: CNN Models Performance with kNN Classifier

Conclusion

This work developed a new hybrid method using dermoscopic pictures for multi-class skin cancer categorization. The suggested model combines MobileNetV2, EfficientNetB0, and DenseNet 201 with DeepLabV3+ for lesion segmentation in a framework for hybrid feature extraction. Feature selection made use of the ReliefF method; final classification made use of the kNN algorithm.

The proposed hybrid model clearly beats individual CNN models and current state-of-the-art techniques, as shown by its 94.42% accuracy on the difficult ISIC-2019 dataset. The study showed how important lesion segmentation is for improving classification performance and how useful it is to combine complementary CNN architectures to get a lot of different traits. For dermatologists, the suggested method presents a useful tool that will help them classify skin lesions from dermoscopic pictures more precisely and consistently. The development of ensemble learning models capable of enhancing classification performance and enabling more general clinical use would be the main emphasis of the next studies.

Acknowledgment

This work was supported by Zarqa University.

References

Abdelhafeez, A., Mohamed, H. K., Maher, A., & Khalil, N. A. (2023). A novel approach toward skin cancer classification through fused deep features and neutrosophic environment. Frontiers in Public Health, 11, 1123581.

Abu Al-Haija, Q., & Al-Fayoumi, M. (2023). An intelligent identification and classification system for malicious uniform resource locators (URLs). Neural Computing and Applications, 35(23), 16995–17011.

Abu Owida, H., Turab, N., Al-Nabulsi, J. I., & Al-Ayyad, M. (2024). Progress in self-powered medical devices for breathing recording. Bulletin of Electrical Engineering and Informatics, 13(5), 3590-3600. https://doi.org/10.11591/eei.v13i5.5253

Al Tawil, A., Al-Shboul, L., Almazaydeh, L., &Alshinwan, M. (2024). Fortifying network security: Machine learning-powered intrusion detection systems and classifier performance analysis. International Journal of Electrical and Computer Engineering (IJECE), 14(5), 5894-5905.

Alazaidah, R., Owida, H. A., Alshdaifat, N., Issa, A., Abuowaida, S., & Yousef, N. (2024). A comprehensive analysis of eye diseases and medical data classification. TELKOMNIKA (Telecommunication Computing Electronics and Control), 22(6), 1422-1430.

Alhija, M. A., Al-Baik, O., Hussein, A., & Abdeljaber, H. (2024). Optimizing blockchain for healthcare IoT: A practical guide to navigating scalability, privacy, and efficiency trade-offs. Indonesian Journal of Electrical Engineering and Computer Science, 35(3), 1773-1785.

Alhusenat, A. Y., Owida, H. A., Rababah, H. A., Al-Nabulsi, J. I., &Abuowaida, S. (2023). A secured multi-stages authentication protocol for IoT devices. Mathematical Modelling of Engineering Problems, 10(4), 1-10.

Alkhdour, T., Almaiah, M. A., Alahmed, M. A., Al-Shareeda, M. A., Lutfi, A., &Alrawad, M. (2024). Cybersecurity risk management in IoT systems: A systematic review. Journal of Theoretical and Applied Information Technology, 102(13), 1-15.

Al-Momani, A., Al-Refai, M. N., Abuowaida, S., Arabiat, M., Alshdaifat, N., & Rahman, M. N. A. (2024). The effect of technological context on smart home adoption in Jordan. Indonesian Journal of Electrical Engineering and Computer Science, 33(2), 1186-1195.

Alomoush, W., Houssein, E. H., Alrosan, A., Abd-Alrazaq, A., Alweshah, M., &Alshinwan, M. (2024). Joint opposite selection enhanced Mountain Gazelle Optimizer for brain stroke classification. Evolutionary Intelligence, 17(4), 2865-2883.

Al-Oraini, B., Khanfar, I. A., Al-Daoud, K., Mohammad, S. I., Vasudevan, A., Fei, Z., & Al-Azzam, M. K.

A. (2024). Determinants of Customer Intention to Adopt Mobile Wallet Technology. Appl. Math, 18(6), 1331-1344.
Alshdaifat, N., Owida, H. A., Mustafa, Z., Aburomman, A., Abuowaida, S., Ibrahim, A., &Alsharafat, W. (2024). Automated blood cancer detection models based on EfficientNet-B3 architecture and transfer learning. Indonesian Journal of Electrical Engineering and Computer Science, 36(3), 1731-1738.
Alshinwan, M., Khashan, O. A., Khader, M., Tarawneh, O., Shdefat, A., Mostafa, N., &AbdElminaam, D. S. (2024). Enhanced Prairie Dog Optimization with Differential Evolution for solving engineering design problems and network intrusion detection system. Heliyon, 10(17), e12345.
Ansari, S. A., Agrawal, A. P., Wajid, M. A., Wajid, M. S., & Zafar, A. (2024). MetaV: A pioneer in feature augmented meta-learning based vision transformer for medical image classification. Interdisciplinary Sciences: Computational Life Sciences, 16, 469-488.
Arabiat, M., Abuowaida, S., Al-Momani, A., Alshdaifat, N., & Chan, H. Y. (2024). Depth estimation method based on residual networks and SE-Net model. Journal of Theoretical and Applied Information Technology, 102(3).
Arnold, M., Holterhues, C., Hollestein, L., et al. (2014). Trends in incidence and predictions of cutaneous melanoma across Europe up to 2015. Journal of the European Academy of Dermatology and Venereology, 28(9), 1170-1178.
Arnold, M., Singh, D., Laversanne, M., et al. (2022). Global burden of cutaneous melanoma in 2020 and projections to 2040. JAMA Dermatology, 158(5), 495-503.
Atoum, I., &Otoom, A. A. (2024). Enhancing software effort estimation with pre-trained word embeddings: A small-dataset solution for accurate story point prediction. Electronics, 13(23), 4843.
Ayas, S. (2023). Multiclass skin lesion classification in dermoscopic images using Swin Transformer model. Neural Computing and Applications, 35(9), 6713-6722.
Ayyalsalman, K. M., Alolayyan, M. N., Alshurideh, M. T., Al-Daoud, K., & Al-Hawary, S. I. S. (2024). Mathematical Model to Estimate The Effect of Authentic Leadership Components on Hospital Performance. Appl. Math, 18(4), 701-708.
Bansal, P., Garg, R., & Soni, P. (2022). Detection of melanoma in dermoscopic images by integrating features extracted using handcrafted and deep learning models. Computers and Industrial Engineering, 168, 108060.
Batiha, I. M., Jebril, I. H., Anakira, N., Al-Nana, A. A., Batyha, R., & Momani, S. (2024). Two-dimensional fractional wave equation via a new numerical approach. International Journal of Innovative Computing, Information & Control, 20(4), 1045-1059.
Behara, K., Bhero, E., & Agee, J. T. (2024). An improved skin lesion classification using a hybrid approach with active contour snake model and lightweight attention-guided capsule networks. Diagnostics, 14(6), 636.
Chen, W., Vasudevan, A., Al-Daoud, K. I., Mohammad, S. I. S., Arumugam, V., Manoharan, T., & Foong, W. S. (2024). Integrating cultures, enhancing outcomes: Perceived organizational support and its impact on Chinese expatriates' performance in Dubai. Herança, 7(3), 25-39.
Codella, N. C., Gutman, D., Celebi, M. E., et al. (2018). Skin lesion analysis toward melanoma detection: A challenge at the 2017 International Symposium on Biomedical Imaging (ISBI), hosted by the International Skin Imaging Collaboration (ISIC). In 2018 IEEE 15th International Symposium on Biomedical Imaging (ISBI) (pp. 168-172). IEEE.
Codella, N., Rotemberg, V., Tschandl, P., et al. (2019). Skin lesion analysis toward melanoma detection 2018: A challenge hosted by the International Skin Imaging Collaboration (ISIC). arXiv preprint arXiv:1902.03368.
Combalia, M., Codella, N. C., Rotemberg, V., et al. (2019). BCN20000: Dermoscopic lesions in the wild.

arXiv preprint arXiv:1908.02288.
Daghrir, J., Tlig, L., Bouchouicha, M., & Sayadi, M. (2020). Melanoma skin cancer detection using deep learning and classical machine learning techniques: A hybrid approach. In 2020 5th International Conference on Advanced Technologies for Signal and Image Processing (ATSIP) (pp. 1-5). IEEE.
Ekanayake, E. A., Al-Daoud, K. I., Vasudevan, A., Wenchang, C., Hunitie, M. F. A., & Mohammad, S. I. S. (2024). Leveraging Aquaculture and Mariculture for Sustainable Economic Growth in Sri Lanka: Challenges and Opportunities. Journal of Ecohumanism, 3(6), 1229-1247.
Gajera, H. K., Nayak, D. R., & Zaveri, M. A. (2023). A comprehensive analysis of dermoscopy images for melanoma detection via deep CNN features. Biomedical Signal Processing and Control, 79, 104186.
Galdolage, B. S., Ekanayake, E. A., Al-Daoud, K. I., Vasudevan, A., Wenchang, C., Hunitie, M. F. A., & Mohammad, S. I. S. (2024). Sustainable Marine and Coastal Tourism: A Catalyst for Blue Economic Expansion in Sri Lanka. Journal of Ecohumanism, 3(6), 1214-1228.
Gilani, S. Q., Syed, T., Umair, M., & Marques, O. (2023). Skin cancer classification using deep spiking neural network. Journal of Digital Imaging, 36(3), 1137-1147.
Haenssle, H., Fink, C., Schneiderbauer, R., et al. (2018). Man against machine: Diagnostic performance of a deep learning convolutional neural network for dermoscopic melanoma recognition in comparison to 58 dermatologists. Annals of Oncology, 29(8), 1836-1842.
Hermosilla, P., Soto, R., Vega, E., Suazo, C., & Ponce, J. (2024). Skin cancer detection and classification using neural network algorithms: A systematic review. Diagnostics, 14(4), 454.
Huang, G., Liu, Z., Van Der Maaten, L., & Weinberger, K. Q. (2017). Densely connected convolutional networks. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (pp. 4700-4708). IEEE.
Ivry, G. B., Ogle, C. A., & Shim, E. K. (2006). Role of sun exposure in melanoma. Dermatologic Surgery, 32(4), 481-492.
Jui, A. H., Sharnami, S., & Islam, A. (2022). A CNN based approach to classify skin cancers using transfer learning. In 2022 25th International Conference on Computer and Information Technology (ICCIT) (pp. 1063-1068). IEEE.
Keerthana, D., Venugopal, V., Nath, M. K., & Mishra, M. (2023). Hybrid convolutional neural networks with SVM classifier for classification of skin cancer. Biomedical Engineering Advances, 5, 100069.
Kim, C., Jang, M., Han, Y., Hong, Y., & Lee, W. (2023). Skin lesion classification using hybrid convolutional neural network with edge, color, and texture information. Applied Sciences, 13(9), 5497.
Mahbod, A., Schaefer, G., Ellinger, I., Ecker, R., Pitiot, A., & Wang, C. (2019). Fusing fine-tuned deep features for skin lesion classification. Computerized Medical Imaging and Graphics, 71, 19-29.
Mohammad, A. A. S. (2025). The impact of COVID-19 on digital marketing and marketing philosophy: evidence from Jordan. International Journal of Business Information Systems, 48(2), 267-281.
Mohammad, A. A. S., Al-Daoud, K. I., Rusho, M. A., Alkhayyat, A., Doshi, H., Dey, P., ... & Kiani, M. (2025b). Modeling polyethylene glycol density using robust soft computing methods. Microchemical Journal, 210, 112815.
Mohammad, A. A. S., Mohammad, S. I. S., Al Oraini, B., Vasudevan, A., &Alshurideh, M. T. (2025c). Data security in digital accounting: A logistic regression analysis of risk factors. International Journal of Innovative Research and Scientific Studies, 8(1), 2699-2709.
Mohammad, A. A. S., Mohammad, S. I. S., Al-Daoud, K. I., Al Oraini, B., Vasudevan, A., & Feng, Z. (2025a). Optimizing the Value Chain for Perishable Agricultural Commodities: A Strategic Approach for Jordan. Research on World Agricultural Economy, 6(1), 465-478.
Mohammad, A. A., Shelash, S. I., Saber, T. I., Vasudevan, A., Darwazeh, N. R., &Almajali, R. (2025e). Internal audit governance factors and their effect on the risk-based auditing adoption of commercial

banks in Jordan. Data and Metadata, 4, 464.

Mohammad, A.A.S., Al-Hawary, S.I.S., Hindieh, A., Vasudevan, A., Al-Shorman, M. H., Al-Adwan, A.S., Turki Alshurideh, M., & Ali, I. (2025d). Intelligent Data-Driven Task Offloading Framework for Internet of Vehicles Using Edge Computing and Reinforcement Learning. Data and Metadata, 4, 521.

Mohammad, S. I. S., Al-Daoud, K. I., Al Oraini, B. S., Alqahtani, M. M., Vasudevan, A., & Ali, I. (2025f). Impact of Crude Oil Price Volatility on Procurement and Inventory Strategies in the Middle East. International Journal of Energy Economics and Policy, 15(2), 715-727.

Owida, H. A., AlMahadin, G., Al-Nabulsi, J. I., Turab, N., Abuowaida, S., &Alshdaifat, N. (2024). Automated classification of brain tumor-based magnetic resonance imaging using deep learning approach. International Journal of Electrical & Computer Engineering, 14(3).

Owida, H. A., Alnaimat, F., Al-Nabulsi, J. I., Al-Ayyad, M., & Turab, N. M. (2024). Application of smart hydrogels scaffolds for bone tissue engineering. Bulletin of Electrical Engineering and Informatics, 13(6), 4388-4393. https://doi.org/10.11591/eei.v13i6.7608

Owida, H. A., Alshdaifat, N., Almaghthawi, A., Abuowaida, S., Aburomman, A., Al-Momani, A., ... & Chan, H. Y. (2024). Improved deep learning architecture for skin cancer classification. Indonesian Journal of Electrical Engineering and Computer Science, 36(1), 501-508.

Owida, H. A., Hassan, M. R., Ali, A. M., Alnaimat, F., Al Sharah, S., Abuowaida, S., &Alshdaifat, N. (2024). The performance of artificial intelligence in prostate magnetic resonance imaging screening. International Journal of Electrical and Computer Engineering, 14(2), 2234-2241.

Owida, H. A., Moh'd, B. A.-h., Turab, N., Al-Nabulsi, J., &Abuowaida, S. (2023). The evolution and reliability of machine learning techniques for oncology. International Journal of Online & Biomedical Engineering, 19(8).

Ozcan, T. (2021). A new composite approach for COVID-19 detection in X-ray images using deep features. Applied Soft Computing, 111, 107669.

Pandimurugan, V., Ahmad, S., Prabu, A. V., Rahmani, M. K. I., Abdeljaber, H. A., Eswaran, M., & Nazeer, J. (2024). CNN-based deep learning model for early identification and categorization of melanoma skin cancer using medical imaging. SN Computer Science, 5(7), 911.

Qatawneh, A. M., Lutfi, A., & Al Barrak, T. (2024). Effect of artificial intelligence (AI) on financial decision-making: Mediating role of financial technologies (Fin-Tech). HighTech and Innovation Journal, 5(3), 759-773.

Sagi, O., & Rokach, L. (2018). Ensemble learning: A survey. Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery, 8(4), e1249.

Sánchez-Paniagua, M., Fernández, E. F., Alegre, E., Al-Nabki, W., & Gonzalez-Castro, V. (2022). Phishing URL detection: A real-case scenario through login URLs. IEEE Access, 10, 42949–42960.

Sawaneh, I. A. (2020). Cybercrimes: Threats, challenges, awareness, and solutions in Sierra Leone. Asian Journal of Interdisciplinary Research, 3(2), 185–195.

Selvaganapathy, S., Nivaashini, M., & Natarajan, H. (2018). Deep belief network based detection and categorization of malicious URLs. Information Security Journal: A Global Perspective, 27(3), 145–161.

Singh, P., & Ranga, V. (2021). Attack and intrusion detection in cloud computing using an ensemble learning approach. International Journal of Information Technology, 13(2), 565–571.

Singh, R., & Patel, M. (2023). Adaptive ensemble methods in cybersecurity applications. Journal of Network Security, 15(3), 234-249.

Siyal, R., Long, J., Asim, M., Ahmad, N., Fathi, H., &Alshinwan, M. (2024). Blockchain-enabled secure data sharing with honey encryption and DSNN-based key generation. Mathematics, 12(13), 1956.

Subhashini, G., & Chandrasekar, A. (2023). Hybrid deep learning technique for optimal segmentation and classification of multi-class skin cancer. Imaging Science Journal (Early View), 1-22. https://doi.org/10.1080/13682199.2023.2241794

Tabassum, T., Alam, M. M., Ejaz, M. S., & Hasan, M. K. (2023). A review on malicious URLs detection using machine learning methods. Journal of Engineering Research and Reports, 25(12), 76–88.

Tan, M., & Le, Q. (2019). EfficientNet: Rethinking model scaling for convolutional neural networks. In International Conference on Machine Learning (pp. 6105-6114). PMLR.

Taylor, D., & Brown, N. (2023). Random forest applications in cybersecurity. Machine Learning, 112(3), 234-249.

Tian, Z., Luo, C., Qiu, J., Du, X., & Guizani, M. (2019). A distributed deep learning system for web attack detection on edge devices. IEEE Transactions on Industrial Informatics, 16(3), 1963–1971.

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(1), 1-9.

Tumpa, P. P., & Kabir, M. A. (2021). An artificial neural network based detection and classification of melanoma skin cancer using hybrid texture features. Sensors International, 2, 100128.

Yu, L., Wang, Y., Zhou, L., Wu, J., & Wang, Z. (2023). Residual neural network-assisted one-class classification algorithm for melanoma recognition with imbalanced data. Computational Intelligence, 39(6), 1004-1021.

Yuan, H., Zhu, J., Wang, Q., Cheng, M., & Cai, Z. (2022). An improved DeepLab v3+ deep learning network applied to the segmentation of grape leaf black rot spots. Frontiers in Plant Science, 13, 795410.