Article
JOURNAL OF UNIVERSITY OF BABYLON
For Pure and Applied Sciences (JUBPAS)
Vol. 33 ; No.4 | 2025

A Hybrid Deep Learning Model for Skin Cancer Classification Using Elephant Herding Optimization

Suhad Hatim Jihad[1*], Wafaa Sallal abbood [2], Sumar Mohamed Khaleel [3]
Nisreen Saad Hadi[4]
1 Computer Center, University of Babylon, suhad.jihad@uobabylon.edu.iq, Hilla, Iraq
2 Computer Center, University of Babylon, wafaa.salal@uobabylon.edu.iq, Hilla, Iraq.
3 Computer Center , University of Babylon, sumar.moh@uobabylon.edu.iq, Hilla, Iraq.
4 Department of Administrative and Financial Affairs , University of Babylon , bsclec.nisreen.saad@uobabylon.edu.iq, Hilla, Iraq.

***Corresponding author email:** suhad.jihad@uobabylon.edu.iq; mobile: 07804066933

نموذج هجين للتعلم العميق لتصنيف سرطان الجلد باستخدام خوارزمية تحسين قطيع الفيلة

سهاد حاتم جهاد [1*]، وفاء صلال عبود [2]، سومار محمد خليل [3] ,نسرين سعد هادي[4]
مركز الحاسبة الالكترونية ، جامعه بابل ، suhad.jihad@uobabylon.edu.iq ، بابل ، العراق
مركز الحاسبة الالكترونية ، جامعه بابل ، wafaa.salal@uobabylon.edu.iq ، بابل ، العراق
مركز الحاسبة الالكترونية ، جامعه بابل ، sumar.moh@uobabylon.edu.iq ، بابل ، العراق
قسم الشؤون الادارية والمالية , جامعة بابل ,bsclec.nisreen.saad@uobabylon.edu.iq, Hilla, Iraq. , بابل, العراق

Accepted: 27/10/2025 **Published:** 31/12/2025

ABSTRACT

Background:

Skin cancer remains one of the most prevalent malignancies worldwide, and its prevention and early detection play a crucial role in reducing morbidity and mortality. Recent advances in deep learning have significantly improved the accuracy of automated skin lesion classification. However, several challenges remain in designing a computationally efficient hybrid model that integrates multiple lightweight architectures.

Materials and Methods:

This work proposes a hybrid deep learning model that combines the MobileNetV3+ and ShuffleNetV2 architectures. To enhance model performance, the Elephant Herd Optimization (EHO) algorithm was employed to optimize key hyperparameters, including learning rate, batch size, and the number of epochs. Skin image datasets were collected from SkinDataSe and the ISIC archive, encompassing both benign and malignant cases. Preprocessing involved contrast enhancement and various data augmentation techniques such as rotation, scaling, flipping, and translation to improve model generalization. All images were resized to 224×224×3 pixels and normalized to [0,1] range. The dataset was then divided into 70% for training, 15% for validation, and 15% for testing.

Results:

The hybrid model demonstrated exceptional performance, achieving a training accuracy of 99.72% and an F1-score of 0.93.

Conclusion:

The findings of this study confirm the potential of the proposed hybrid deep learning framework, optimized via EHO, in delivering accurate, swift, and early diagnosis of skin cancer.

Key words: skin cancer, Elephant Herding Optimization (EHO), transfer deep learning, Hybrid CNN, medical images.

ISSN: 2312-8135 | Print ISSN: 1992-0652
info@journalofbabylon.com | jub@itnet.uobabylon.edu.iq | www.journalofbabylon.com

INTRODUCTION

Skin cancer is a major public health, issue with disease incidence trends increasing gradually. It is reported as the 19th most common cancer worldwide with an unsettling increasing trend. The disease affects both genders, and it is most common among the whites in the United States. Over the past few decades, the number of diagnosed cases has increased drastically. Global incidence of skin cancer is on the rise, despite restrictions implemented on UV exposure and screening programs. The early detection of skin cancer is not only a mere figure. It is the very embodiment of all that relates to increased life expectancy and better outcomes in treatment. But the rapid and atypical proliferation of the distinct pathology of cancer makes the timely and accurate detection of skin cancer quite a challenge [1-6].

That is exacerbated by the normal modes of identifying it, such as dermoscopy. Even though dermoscopy substantially assists, its subjectivity, lengthy processes, and self-diagnosis have restrained its application and made individuals fabricate other treatments instead. One possible alternative is dual spectroscopic and imaging technologies like up-to-the-minute Raman spectroscopy when matched with accepted methods [7].

Most people cannot tell cancerous from non-cancerous tumors if they look alike-for example, melanomas and seborrheic keratoses[8]. An increase in the change of color, or outline within the same group of tumors makes diagnosis more difficult. The artificial intelligence model deep learning has redefined this area and gave a new impulse to one of the most long-standing diagnostic problems in medicine[9].

Machine learning incorporates image classification techniques with deep learning models. This has totally revolutionized skin cancer diagnosis and made it quick, inexpensive, and accurate [10].

It should be noted that the impetus for this study was the enormous impact that early detection can have on the patient. Indeed, the early diagnosis creates a critical point for the bettering of the prognosis, increasing the therapeutic effectiveness, and, in the end, saving human lives [11].

Recent developments in the pretrained deep learning models have provided a new scope for achieving accurate as well as efficient detection of skin cancer. Such models are highly computationally efficient and also possess strong feature extraction capabilities that suit practical application hence making them useful[8]. Therefore, they offer an avenue toward developing fast inexpensive yet dependable means of diagnosis thereby improving early detection together with clinical decision-making.

This paper presents a new hybrid convolutional neural network (Hybrid CNN) model of MobileNetV3+ and ShuffleNetV2 for skin lesion classification. The proposed model is optimally tuned in hyperparameters by the Elephant Herd Optimization (EHO) algorithm to deliver the best performance. Its effectiveness is evaluated on several benchmark datasets together with comparative analysis demonstrating improved accuracy and efficiency having potential real-time clinical as well as mobile-based diagnostic applications..

ISSN: 2312-8135 | Print ISSN: 1992-0652
info@journalofbabylon.com | jub@itnet.uobabylon.edu.iq | www.journalofbabylon.com

The rest of the manuscript is organized as follows. Relevant literature is discussed in Section 2, giving theoretical background in Section 3, and the suggested method for detecting skin cancer. Experiments are described and outcomes discussed in Section 5. Finally, Section 6 concludes the article.

MATERIALS AND METHODS

Several studies worked on diagnosing and analyzing brain tumor images via several intelligent artificial intelligence program designs and implementations.

Abdul Rahim Yilmaz et al. (2021) conducted research on brain tumor classification using multiple transfer learning deep learning models. The NasnetMobile model performed best among nine different models, with an accuracy of 82% and an F1 score of 0.8038. The results confirmed the significant role of transfer learning deep learning techniques in improving the accuracy of medical image classification[12].

On the ISIC2018 dataset [13] applied a deep learning algorithm using CNNs to classify benign and malignant skin tumors. The dataset consisted of 3,533 skin lesion images. The input image underwent image resizing, contrast enhancement, and normalization in the initial phase. Transfer learning from ResNet50, InceptionV3, and a combined version of both, Inception ResNet, were subsequently tested. This was further improved using ESRGAN to improve image quality. The aforementioned study yielded an accuracy ranging from approximately 83.2% to 85.8% depending on the model used, confirming the effectiveness of this approach[13].

The study presented a modified model of the data augmentation technique, which was subsequently used to improve melanoma skin cancer diagnosis. PH2 data was used to test the generation of synthetic images via oversampling in a low-dimensional nonlinear representation space. The SqueezeNet model was trained on the augmented images. However, this time the results showed improvements, with a classification accuracy of 92.18% (sensitivity 80.77%, specificity 95.1%, and F1 score 80.84%). The multi-class classification results were also better, even better than some state-of-the-art techniques [13]. In their review, Bill Cassidy et al. (2022) used images from the ISIC dataset from 2016 to 2020, noting the presence of a significant number of duplicate images, 14,310 of which were excluded in the first attempt. The results proved to be good, with the model's average area under the curve of 0.80 [14].

Amal J. Diab et al. described a computerized clinical decision support system for skin cancer diagnosis that integrates deep learning models with non-invasive procedures. Using a CNN, the system performs feature extraction for the region of interest (ROI), which is performed by a support vector machine (SVM) in the classification phase. The models had been trained on datasets of CPTAC-CM and ISIC. Also, results for another version using GoogleNet, ResNet-50, AlexNet, and VGG19 are provided. The model gives an accuracy of 99.8% and 99.9% on both datasets; therefore, it is quite sensitive in the early diagnosis [15]. Emmanuel Karambinis et al. systemic oxidative stress markers in skin diseases like melanoma, basal cell carcinoma, squamous cell carcinoma, and benign tumors were measured. Abnormally low levels or high levels of antioxidant enzymes and oxidized proteins have been reported previously in patients depending upon their history as well as stage/treatment stage reflecting dermatological diagnosis involves biochemical analysis[16].

The researchers in[17]ran a comparison of four parametrically standardized YOLO models-YOLOv3, YOLOv4, YOLOv5, and YOLOv7-in the classification of skin lesions and registered best attainable results with the last one: average accuracy75.4%, F1 measure80%, inference time0.32 seconds per image with high precision in lesion localization(IoU=86.3%), thus making

ISSN: 2312-8135 | Print ISSN: 1992-0652
info@journalofbabylon.com | jub@itnet.uobabylon.edu.iq | www.journalofbabylon.com

it promising for early diagnosis.. In research, Muhammad Atiqur Rahman et al. (2024) described an improved DCNN model, based on the NASNet architecture, for classifying melanoma and non-melanoma skin cancers. The model included an additional base layer and better handling of partial and inconsistent data. This model was then evaluated using 2,637 skin images and achieved an accuracy of 85.62% using the Adam optimizer[18].

Nieuwoudt et al. conducted a study on the use of Raman spectroscopy in the differential diagnosis of benign, inflammatory, and malignant skin lesions. More than 867 spectra were obtained from 330 patients using a specially engineered fiber optic probe. Types were very accurately classified through partial discriminant analysis (PLS-DA), with an area under the curve (AUC) between 0.916 and 0.958[19]. A population-based retrospective study on the impact of a mobile health application for dermatology claims found one arm to have benign and pre-malignant tumor detection significantly better, wherein the cost for detecting another pre-malignant tumor was €2,567. This finding result is consistent with analysis evidence of artificial intelligence potential capability on skin cancer detection in digital health [20].

Literature clearly shows most previous works used single deep learning models individually without making use of model fusion or hybrid approaches to improve the results. Therefore, in this proposed system, the architecture of MobileNetV3+ and ShuffleNetV2 is integrated and further optimally tuned using Elephant Herd Optimization (EHO) algorithm. Table 1 shows the summary for related works.

Table 1: Summary for for related works

Study and Year	Model/Method	Dataset	Accuracy (%)	F1-score	Details
[12] 2021	Mobile Deep Learning (NasnetMobile)	Brain Tumors	82.0	0.8038	Focused on brain tumors; cited as an example of mobile-based deep learning
[13] 2022	CNN + Transfer Learning (ResNet50, InceptionV3, InceptionResNet)	ISIC2018 (3533 images)	83.2–85.8	–	Image quality enhanced with ESRGAN
[14] 2022	CNN	ISIC 2016–2020	–	AUC = 0.80	Excluded duplicate images; data quality review
[15] 2022	CNN + SVM Hybrid	CPTAC-CM & ISIC	99.8–99.9	–	Integrated deep learning with non-invasive procedures
[16]	Biochemical Analysis	–	–	–	Evaluated systemic oxidative stress indicators in various skin diseases

[17] 2024	YOLOv3/4/5/7	Skin Lesion Images	75.4	80	Best results with YOLOv7; high lesion localization accuracy
[18] 2024	Improved DCNN (NASNet)	2637 skin images	85.62	–	Better handling of partial and inconsistent data
[19] –	Raman Spectroscopy + PLS-DA	330 patients (867 spectra)	–	AUC 0.916–0.958	Differentiation of benign, inflammatory, and malignant lesions
[20] –	Mobile Health App + AI	–	–	–	Improved detection of benign and pre-malignant tumors; cost analysis included

MACHINE LEARNING

Unsupervised machine learning methods attempt to detect patterns, structures, or meaningful information within a set of data, without pre-defined outputs or outcomes—in simple terms, without a way to validate the outputs[21]. Unlike unsupervised techniques, which focus on modeling input attributes, supervised machine learning focuses on modeling the dependencies between input attributes and a target attribute. There are two main types of supervised learning: classification and regression. In classification, the target variable is within different classes, while in regression, it is continuous [22].

Classification is a popular technique in data mining and is used to define the class of each data object[23]. There are many different machine learning techniques, but classification is among the most popular and is of great importance in future applications and knowledge discovery. Therefore, the classification problem has received significant attention from data mining researchers and is one of the most researched problems [24].

Figure (2) shows a simple model of classification-based supervised learning methods. However, classification methods face the problem of missing data, which can degrade the training and classification process. Data may disappear due to incorrect and confusing data entry, some features are not considered at the time of entry because they seem less important, data is removed due to inconsistency with other data, or simply due to hardware failure [25]. To solve this problem, data mining experts can adopt different strategies, for example, not including records with missing values, replacing missing values with a global constant, taking the mean value of the attribute in the same class, manually investigating the affected samples, or assuming a possible or probable value to compensate for the missing value [26].

ISSN: 2312-8135 | Print ISSN: 1992-0652
info@journalofbabylon.com | jub@itnet.uobabylon.edu.iq | www.journalofbabylon.com

ISSN: 2312-8135 | Print ISSN: 1992-0652
info@journalofbabylon.com | jub@itnet.uobabylon.edu.iq | www.journalofbabylon.com

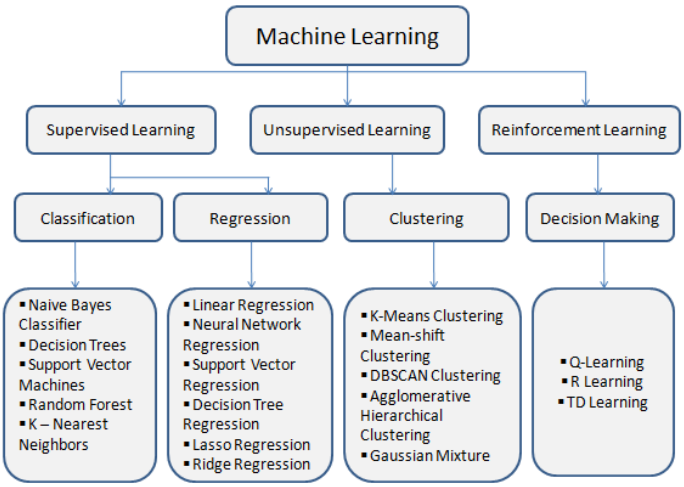

Figure 2: Machine learning techniques [23].

Recently, deep learning algorithms have become very popular and in fact, there has been great success achieved through the application of deep learning in many areas. One of the most popular methods of deep learning is convolutional neural networks (CNNs). CNNs are a type of neural network architecture with different layers performing specific functions such as convolution, loss calculation function pooling among others whose output acts as input to another layer[27].
CNNs had been widely used in computer vision but they really came to their full potential until the 2012 ImageNet competition. This transformation was enabled due to the fact that GPU s could be efficiently utilized along with effective data augmentation techniques[28].
Deep learning comprises several processing layers that build computer models able to detect complex patterns within large data sets by learning representations of the data at multiple levels of abstraction. The challenge lies in how the model adjusts the parameters required to generate each layer's representation based on the previous one, merging machine learning and artificial intelligence into a novel approach to machine learning .
Feature extraction, which is crucial in machine learning, faces limitations in processing and dealing with raw natural information. These limitations make it challenging to address problems in artificial intelligence. Deep learning was specifically designed to overcome these challenges [29].

TRANSFER LEARNING MODELS

Transfer learning, as shown in Figure 3, is a powerful paradigm in deep learning. A good model is first pre-trained on a large dataset on some task. The same model then adapts to various other related tasks with limited data availability. This method takes advantage of the learned weights and hierarchical features of deep neural networks, avoiding enormous hunger for computational power by training models de novo.

ISSN: 2312-8135 | Print ISSN: 1992-0652
info@journalofbabylon.com | jub@itnet.uobabylon.edu.iq | www.journalofbabylon.com

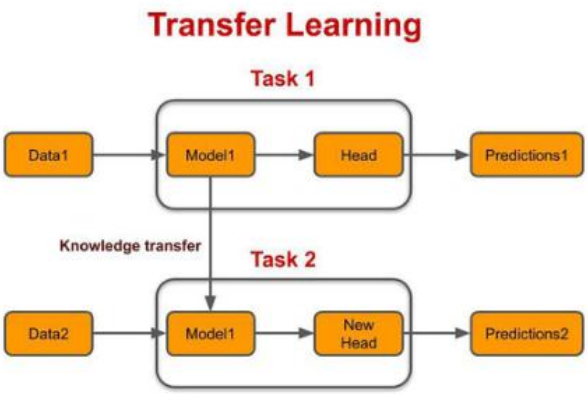

Figure 3: Transfer machine learning [30].

The key concept is to retrain part of the network and keep constant the part that has already been trained, especially when the new dataset has something in common with the original dataset but is much smaller. This way previously gained knowledge is applied again leading to great advances in training efficiency, convergence speed, and overall accuracy when resources or data are labeled sparingly [30].Some major pre-trained models have been taken greatly across different fields, including image sorting out, normal language processing, and temporal trend recognition since they have strong architectures and perform well in generalizing. One example of these models that is very good is DenseNet (or Densely Connected Convolutional Networks) [31–32] which deals with the disappearing gradient issue by setting up direct connections among all layers to ensure maximum flow of information and reuse of features. Unlike the conventional deep architectures that mostly depend on increasing depth or width, efficiency by connectivity is the main focus in DenseNet with versions such as DenseNet-121 and DenseNet-201 showing how deep the network is.Another major architecture, as an evolution of the MobileNet series by Google, is MobileNetV3, specifically optimized for real-time mobile and embedded settings. It unites lightweight design with high accuracy and low latency. Building on MobileNetV2, MobileNetV3 further introduces Inverted Residuals, Squeeze-and-Excitation (SE) modules, and the brand-new Hard-Swish activation function. It is discovered via Neural Architecture Search under AutoML two versions targeted for low-resource environments and high-performance requirements: MobileNetV3-Small and MobileNetV3-Large [33]. These structures, when used in the method of transfer learning, give a workable and expandable fix for medical picture problems like skin cancer finding, where fast and sure diagnosis is key.

MOBILENETV3 ALGORITHM

The Google team created MobileNetV3, the most recent iteration of the MobileNet series, to offer a deep model with excellent accuracy and execution speed while consuming minimal resources, which makes it perfect for embedded systems and mobile devices. In addition to incorporating powerful components like MobileNetV2 inverted residuals, SE (Squeeze-and-Excitation), and a new activation called Hard-Swish, a more effective alternative to Swish, MobileNetV3 combines a number of architectural improvements, building on the outcomes of automatic search using AutoML techniques (specifically, NAS - Neural Architecture Search). There are two iterations of the model: MobileNetV3-Small for devices with modest capabilities and MobileNetV3-Large for excellent performance [33].

ISSN: 2312-8135 | Print ISSN: 1992-0652
info@journalofbabylon.com | jub@itnet.uobabylon.edu.iq | www.journalofbabylon.com

ShuffleNet-V2

ShuffleNet v2 was offered as a direct improvement over ShuffleNet v1 after researchers realized that various theoretical notions employed in lightweight network design do not necessarily convert to actual performance on real-world devices such as smartphones. The paper outlines four guidelines for designing more efficient lightweight networks: minimizing the number of pointwise convolutions due to their high resource consumption; minimizing memory access overhead (MAC) due to its greater impact on model speed than the number of FLOPs; avoiding excessive channel grouping, which has a negative impact on channel communication; and minimizing non-parallel operations, such as summation in short paths, due to their slowdown[34].

Evaluation Metrices

The evaluation of how well a machine learning model works is a basic component in proving its efficiency most especially classification models. There are some standard metrics that are commonly used to determine the ability of a model to predict. Among these include Accuracy which is simply the proportion of correctly classified samples to the total number of predictions as expressed in equation(1).

$$\boldsymbol{accuracy} = \frac{\boldsymbol{number\ of\ correct\ prediction}}{\boldsymbol{total\ number\ op\ prediction}} \qquad (1)$$

But accuracy can never be enough, especially in the case of unbalanced datasets. To add more perspective, a confusion matrix is mostly used which gives a detailed account of the results of predictions in terms of true positives, false positives, true negatives and false negatives as shown in Figure 4.

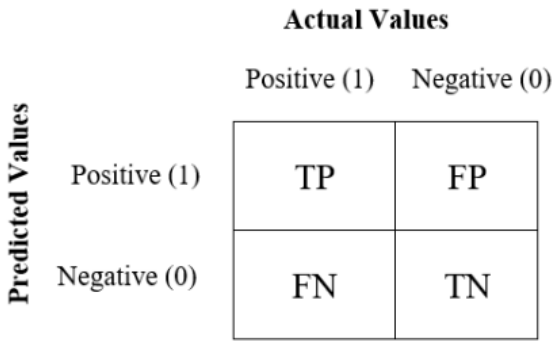

Figure 4: The confusion matrix.

The accuracy in this context can also be computed using Equation (2).

$$\text{acuraccy} = \frac{\boldsymbol{TP+PN}}{\boldsymbol{TP+PN+FN+FP}} \qquad (2)$$

Recall is given as the ratio between True Positive and False Negative added to True Positive

$$\textbf{Recall} = \frac{\text{TP}}{\text{TP+FN}} \qquad (3)$$

Precision (Equation 4) is the ratio of true positive predictions to all predicted positives. Thus, it highlights the reliability of positive classifications. In order to balance both recall and precision as harmonic means, a single metric that contains both completeness and exactness of the model's performance is taken into consideration by the F1-score (Equation 5). These metrics make for a very strong bi-dimensional evaluation of classification systems [35].

$$\textbf{precision} = \frac{TP}{TP+FP} \qquad (4)$$

$$\textbf{F1score} = \frac{2*Precision*Recall}{Precision+Recall} \qquad (5)$$

Performance and efficiency have to be checked. There are several measures commonly around the assessment of the model to make sure that performance is properly checked when analyzing machine learning models. In the analysis of classification tasks, such evaluation measures as Accuracy, Confusion Matrix, Recall, Precision, and F1 Score will be described in the following sub-sections[36].

PROPOSED MODEL METHODOLOGY

The section contains basic steps of the suggested model for skin cancer detection under binary (benign/malignant) and multi-class classification modes. The full workflow of the proposed system is illustrated in Figure 5.

It starts at the stage of data collection where images are collected from the ISIC dataset that contains dermoscopic images with cases, both benign and malignant, of skin cancer.

Next, the data passes a preprocessing step of several enhancement operations including augmenting the data by rotating and mirroring images, besides optimizing illumination to get better quality images. Good quality images are always suitable for generalization capability in models and reduce overfitting. The dataset is split into 70% training, 15% validation, and 15% testing. In the next step, a hybrid deep learning model is developed by integrating the architectures of MobileNetV3 and ShuffleNetV2. Therefore, an appropriate balance between classification accuracy and computational efficiency can be attained. The Elephant Herding Optimization(EHO) algorithm is used in the optimization phase to make this model more effective. EHO dynamically updates major hyperparameters like learning rate, batch size, and several epochs to reach the best accuracy of the model. After getting the best parameters, the model is again trained to fine-tune its performance. During the testing phase, the classification report and confusion matrix distinctly mark benign and malignant skin lesions. Visual performance charts and CSV files give a summary of metrics and results of the proposed hybrid skin cancer detection system in test valuation phase..

ISSN: 2312-8135 | Print ISSN: 1992-0652
info@journalofbabylon.com | jub@itnet.uobabylon.edu.iq | www.journalofbabylon.com

ISSN: 2312-8135 | Print ISSN: 1992-0652
info@journalofbabylon.com | jub@itnet.uobabylon.edu.iq | www.journalofbabylon.com

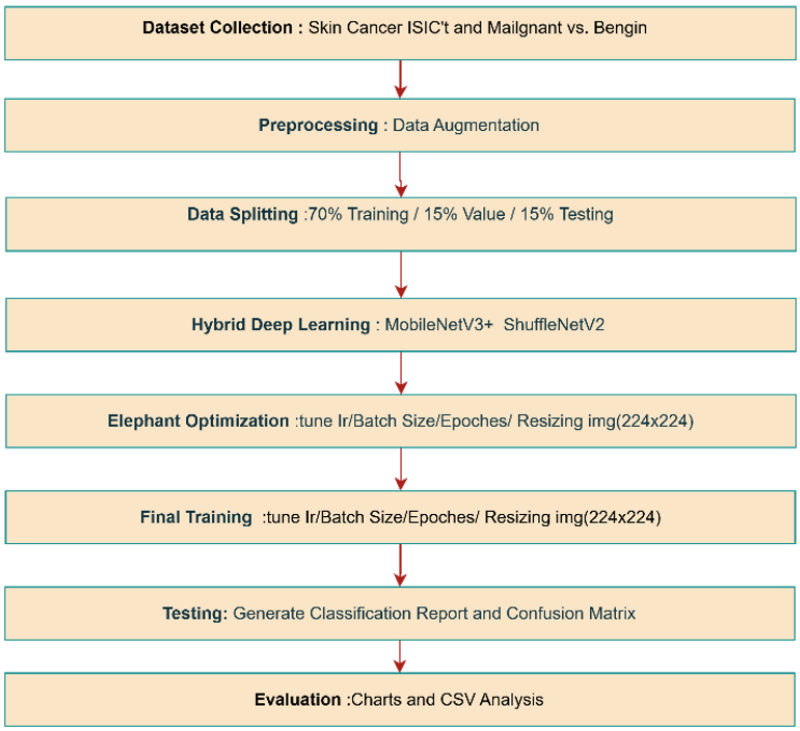

Figure 5: Proposed Model Architecture

DATASET

Data is what powers AI applications, as all experiments and analyses depend on its availability and quality. In this study, two widely known and reliable skin cancer detection datasets, available on the Kaggle Research platform, were used. The first dataset, *Skin Cancer: Malignant vs. Benign, was used for binary classification tasks and contains 1,800 skin images classified as malignant and benign. This dataset is: https://www.kaggle.com/datasets/fanconic/skin-cancer-malignant-vs-benign[37].

The next dataset is the *2020 ISIC Skin Cancer Classification*, a multi-class dataset compiled by the ISIC International Consortium for Dermatological Imaging. It contains 2,357 medical images classified into nine skin disease groups, both cancerous and non-cancerous, including: actinic keratosis, basal cell carcinoma, dermatofibroma, melanoma, nevus, benign pigmented keratosis, seborrheic keratosis, squamous cell carcinoma, and vascular lesions. The images were distributed according to the ISIC classification, with a slight increase in number for some groups, such as melanoma and nevi. This dataset can be accessed via this link: https://www.kaggle.com/datasets/nodoubttome/skin-cancer9-classesisic](https://www.kaggle.com/datasets/nodoubttome/skin-cancer9-classesisic) [38-40].

ISSN: 2312-8135 | Print ISSN: 1992-0652
info@journalofbabylon.com | jub@itnet.uobabylon.edu.iq | www.journalofbabylon.com

Figure 6: dataset selection .

A methodology was applied to combine the two datasets and preprocess the new dataset, as shown in Figure 8, using the Elephant Grazing Optimization (EHO) algorithm. The preprocessing process involved three main steps:

1. Resizing: All images were resized to 224×224×3 pixels, ensuring that this matched the inputs of the deep model.
2. Data Augmentation: Using ImageDataGenerator, we created four copies of each image, applying spatial and color transformations. This will help increase the data diversity and, consequently, improve the model's learning.
3. Data Normalization: Pixel values were converted to a scale of 0 to 1 to accelerate the model's learning rate and make it more stable.

The data was divided into two parts, with the majority allocated to the training set to improve the model's learning ability, and the remainder allocated to the test set to verify the model's accuracy. Table (2) illustrates the direct impact of the initialization process on the data size of each category. For example, the number of images for the melanoma category in the ISIC 2020 dataset increased to 1,718 images, and for the basal cell carcinoma category to 1,656 images, as shown in Table (2). In the binary classification group, the number of images for malignant tumors increased to 4,647 images, and for benign tumors to 4,950 images.

ISSN: 2312-8135 | Print ISSN: 1992-0652
info@journalofbabylon.com | jub@itnet.uobabylon.edu.iq | www.journalofbabylon.com

Table 2: Pre-processing Summary for proposed Skin Cancer Dataset.

Details	Pre-processing (Before)	Pre-processing (After)
Actinic keratosis	129	650
Basal cell carcinoma	392	1656
Dermatofibroma	111	555
Melanoma	454	1718
Nevus	363	1637
Pigmented benign keratosis	478	1742
Seborrheic keratosis	83	400
Squamous cell carcinoma	197	985
Vascular lesion	145	710
Malignant (binary dataset)	1800	4647
Benign (binary dataset)	1800	4950

After the pre-processing as shown in the figure 7 phase, the dataset was randomly shuffled and then partitioned into training and testing sets using the train_test_split function. Specifically, 70% of the data was allocated for training and 15% for testing. To ensure reproducibility of the data split, a fixed random seed (random_state=101) was applied.

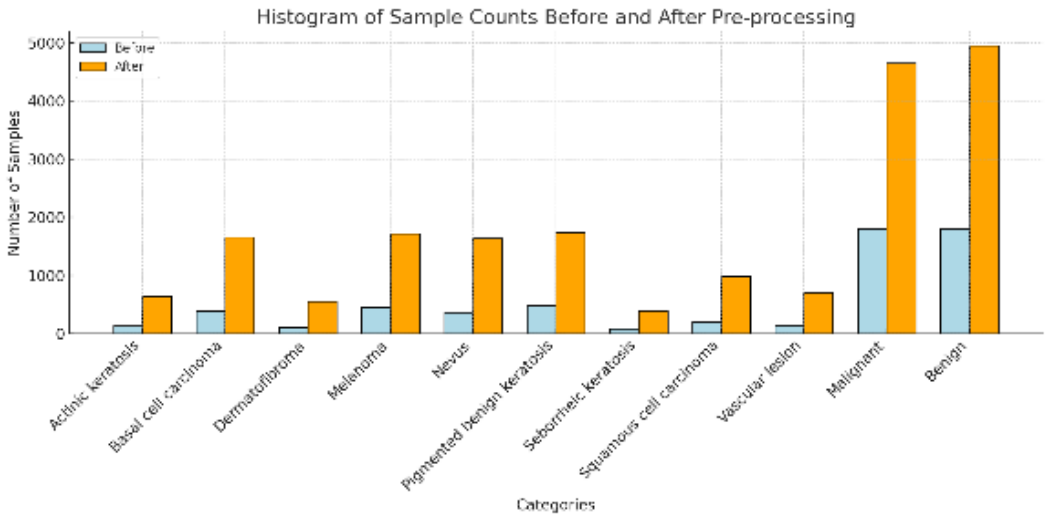

Figure 7: dataset selection

ISSN: 2312-8135 | Print ISSN: 1992-0652
info@journalofbabylon.com | jub@itnet.uobabylon.edu.iq | www.journalofbabylon.com

RESULTS AND DISCUSSION

The proposed hybrid model performs two main tasks in diagnosing skin lesions and tumors. The training dataset applied contains ten different classes, out of which one class is for benign normal skin conditions and the other nine classes represent different types of skin diseases.To reduce the high complexity of massive hyperparameterization in deep learning architectures, efficient hyperparameter tuning was performed using an evolutionary optimization algorithm. The algorithm is inspired by the social behavior of elephant herds that allow effective exploration of a search space to find optimal parameter settings. Elephant Herding Optimization (EHO) delivers a good solution long before consuming all the allocated function evaluations.EHO was used to optimize key training parameters associated with learning rate scheduling and validation accuracy monitoring. After 20 epochs, the hybrid model records an accuracy of 99.5% on the training set and 95.77% on the validation set—proof enough about how effective EHO-based hyperparameter tuning is in enhancing not just model performance but also its generalization capability.

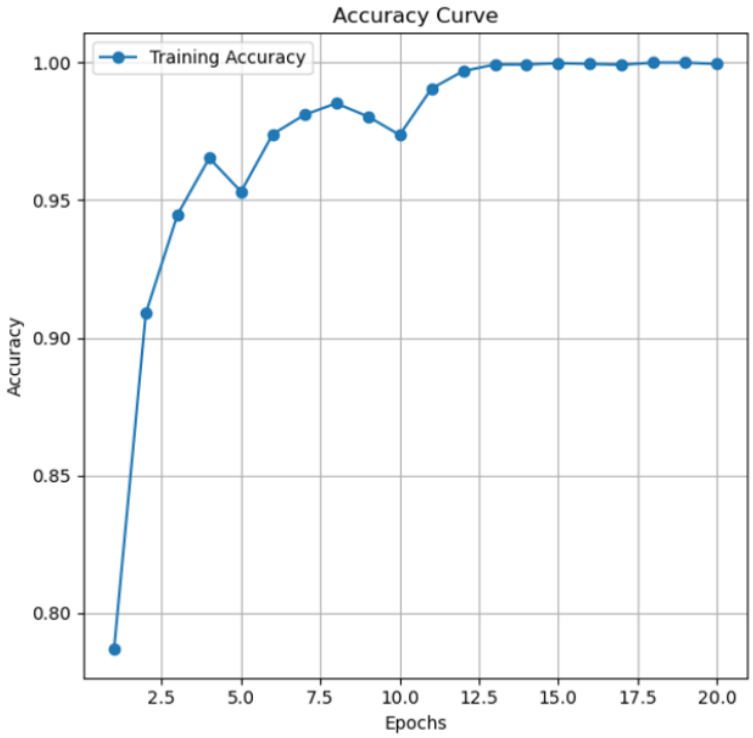

Figure 8: Training accuracy curve of the hybrid model

ISSN: 2312-8135 | Print ISSN: 1992-0652
info@journalofbabylon.com | jub@itnet.uobabylon.edu.iq | www.journalofbabylon.com

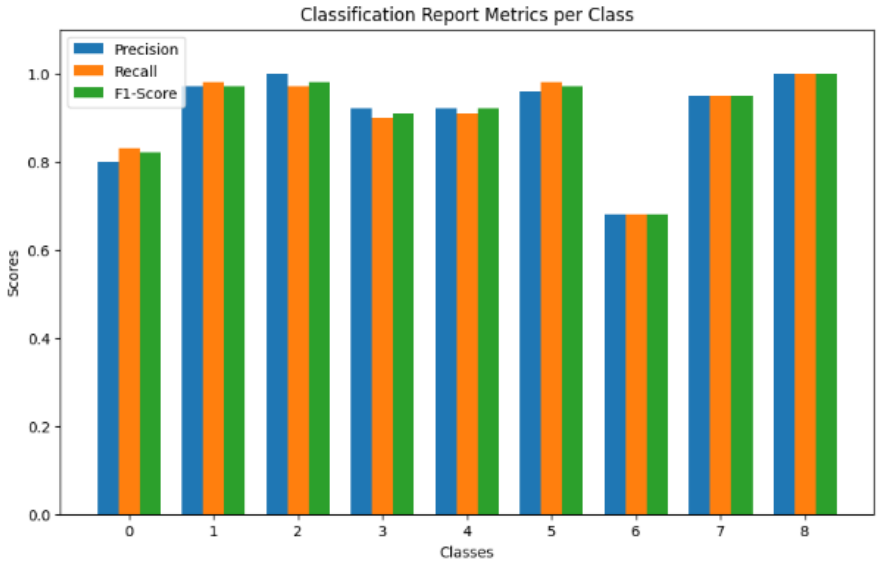

Figure 9: Evaluation of Classification Performance for the hybrid model

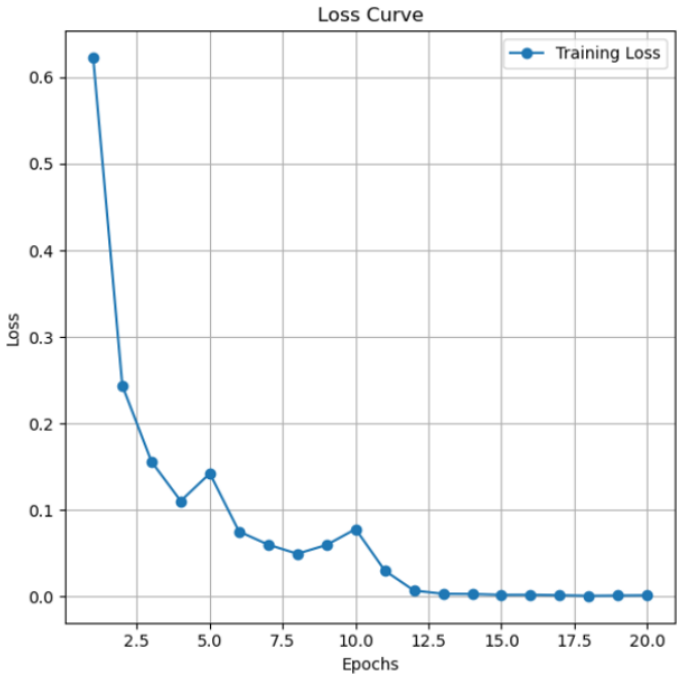

Figure 10: Loss curve of the hybrid model

ISSN: 2312-8135 | Print ISSN: 1992-0652
info@journalofbabylon.com | jub@itnet.uobabylon.edu.iq | www.journalofbabylon.com

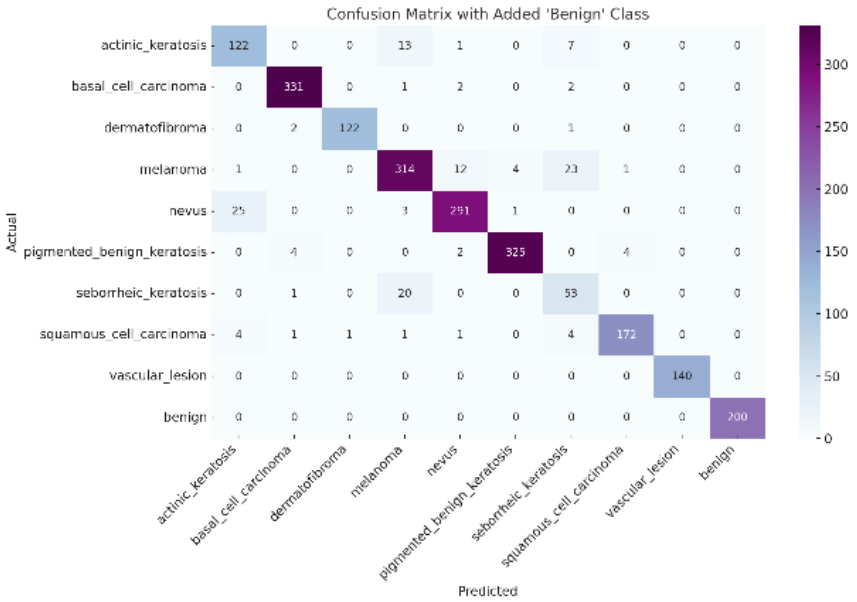

Figure 11: confusion matrix for the hybrid model.

The confusion matrix presented in Figure 11 illustrates the model's classification performance across all ten classes, highlighting its ability to accurately distinguish between different types of skin conditions. The variation in classification accuracy across classes underscores the importance of maintaining a balanced number of samples in the dataset for each class.

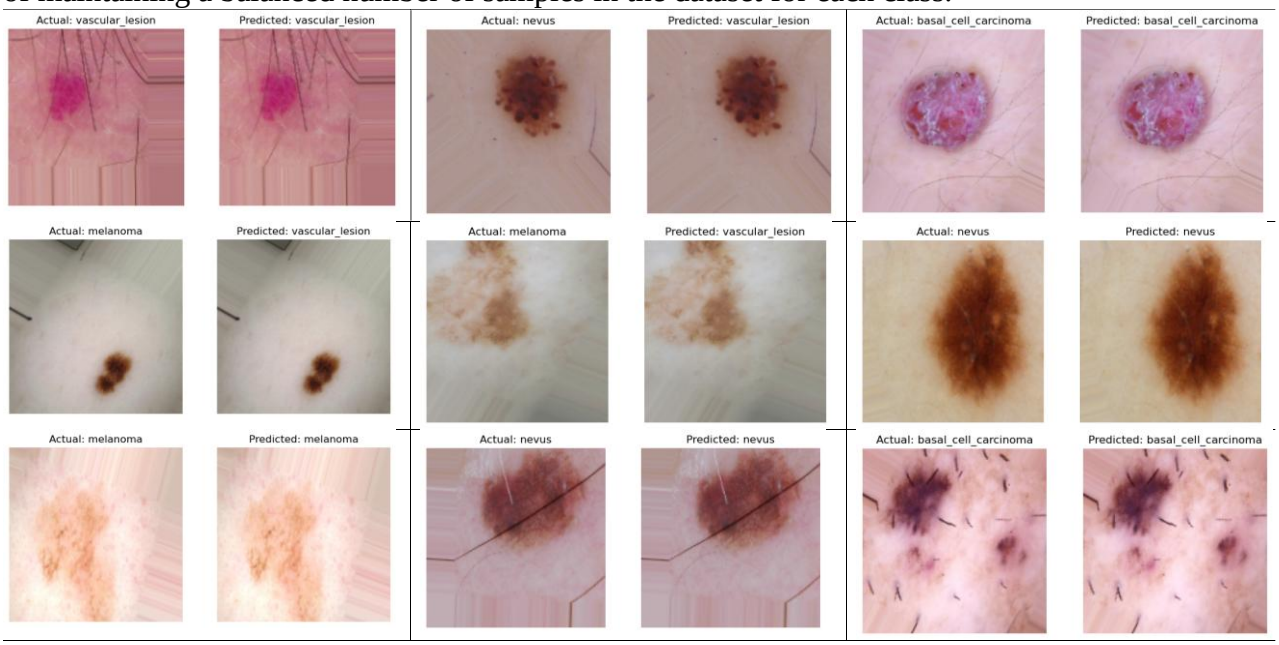

Figure 12 Comparison of actual versus predicted samples for the Hybrid model.

To assess the model's performance in real-world scenarios, it was evaluated on a test set comprising 15% of the dataset, which was not used during training. The model demonstrated high detection accuracy on a set of randomly selected test images (Figure 22). A detailed comparison

between the model's predictions and the actual classes confirmed its generalization capability and suitability for real-world deployment.

The differences in accuracy evaluation between the nine types as shown in Figure 20 prove the varied numbers of images from one category to another. This, therefore, confirms the need for providing a sufficient number of images in the datasets.

to improve the model's ability to fit real-world data. The next section presents the three tests for evaluating performance using test data.

The test data, which represents 15% of the dataset, is not displayed on the proposed model. Therefore, the high detection accuracy of the 15 randomly-generated images is considered important to estimate the suitability of this model for real-world data.

The proposed model was subjected to comparison with three images taken randomly from the test data, which were predicted and compared with the actual class, as shown in Figure 12. Table 3 summarizes the performance comparison of the proposed hybrid model with previous studies.

Table 2: Pre-processing Summary for proposed Skin Cancer Dataset.

Study [Ref]	Year	Model/Method	Dataset	Accuracy (%)	F1-score	Notes
Proposed Model	2025	MobileNetV3+ + ShuffleNetV2 Hybrid (optimized via EHO)	Multiple Skin Datasets	**99.5 (Training), 95.77 (Validation)**	–	Superior accuracy, lightweight, suitable for real-time and mobile deployment
[12]	2021	Mobile Deep Learning (NasnetMobile)	Brain Tumors	82.0	0.8038	Brain tumor classification; cited as general DL example
[13]	2022	CNN + Transfer Learning (ResNet50, InceptionV3)	ISIC2018	83.2–85.8	–	Skin lesion classification; image quality enhanced with ESRGAN
[14]	2021	SqueezeNet + Modified Data Augmentation	PH2 Dataset	92.18	0.8084	Improved melanoma classification using synthetic images
[18]	2024	Improved DCNN (NASNet)	2637 skin images	85.62	–	Better handling of partial/inconsistent data
[17]	2024	YOLOv7	Skin Lesion Images	75.4	80	Lesion detection; high localization accuracy
[15]	2022	CNN + SVM Hybrid	CPTAC-CM & ISIC	99.8	–	Integrated DL and non-invasive clinical support

Results in the table 3 clearly show how differently the models perform on medical image classification tasks depending on network architecture, nature of the dataset, and type of preprocessing applied. Traditional models yield moderate accuracies and are quite effective in

ISSN: 2312-8135 | Print ISSN: 1992-0652
info@journalofbabylon.com | jub@itnet.uobabylon.edu.iq | www.journalofbabylon.com

ISSN: 2312-8135 | Print ISSN: 1992-0652
jub@itnet.uobabylon.edu.iq | www.journalofbabylon.com
info@journalofbabylon.com

locating lesions or regions of interest within images. However, such YOLO or DCNN-based frameworks most often operate under a tradeoff condition between accuracy and speed of computation.Meanwhile, the training accuracy achieved by our proposed hybrid model-MobileNetV3+ and ShuffleNetV2 architectures integrated with Elephant Herd Optimization (EHO)- was 99.5% while its validation accuracy stood at 95.77%. This result proves that an intelligent combination of lightweight architectures and evolutionary optimization methods can deliver a working compromise on computational efficiency and classification exactness in real-time clinical or mobile diagnostic applications.The results in general point out that hybrid and algorithmically optimized models are a promising direction toward making diagnostic systems intelligent, efficient, and suitable for practical deployment.

CONCLUISION:

This paper proposed a deep learning hybrid based automatically detecting and distinguishing benign and malignant skin lesions, which comprises MobileNetV3 and ShuffleNetV2 architectures optimized in terms of classification accuracy as well as computational efficiency by the Elephant Herd Optimization (EHO) algorithm. Experimental results revealed that this framework outperformed several contemporarily state-of-the-art frameworks with high training and validation accuracies within a lightweight structure suitable for real implementations. The binary and multi-class datasets help improve generalization and stability because of less class imbalance. Practically, several recommendations can be drawn from this study.

First, the model can be deployed as a CDSS or teledermatology application in helping dermatologists identify and prioritize suspected high-risk lesions. Second, because of its lightweight architecture, HerLeS can serve well from mobile-based applications running at the edge to remote resource-constrained settings where quick accessible screening is required. Third, future works should consider hybrid GANs for synthetic data generation to allow further fine-tuning of the network parameters while making it more robust to imbalance datasets; not ruling out real-time optimization which could provide instant diagnostic feedback during dermoscopic imaging sessions on cross-dataset validation accompanied with clinical metadata such as age , lesion history , biochemical indicators among others that would increase reliability hence better clinical applicability.

ISSN: 2312-8135 | Print ISSN: 1992-0652
info@journalofbabylon.com | jub@itnet.uobabylon.edu.iq | www.journalofbabylon.com

مجلة جامعة بابل للعلوم الصرفة والتطبيقية مجلة جامعة بابل للعلوم الصرفة والتطبيقية

Conflict of interests.

There are non-conflicts of interest.

References

[1] R. S. Eapen, et al., "PSMA PET applications in the prostate cancer journey: from diagnosis to theranostics," World Journal of Urology, vol. 37, pp. 1255–1261, 2019. doi:10.1007/s00345-018-2524-z.

[2] Z. Dong, K. Xue, A. Verma, J. Shi, Z. Wei, X. Xia, et al., "Photothermal therapy: a novel potential treatment for prostate cancer," Biomaterials Science, vol. 12, no. 10, pp. 2480-2503, 2024, doi: 10.1039/D4BM00057A.

[3] C. L. Kok, C. K. Ho, R. Vicknesh, C. Lee and Y. Y. Koh, "Optimizing Deep Learning on Sustainable Embedded Systems: A Study of Handwritten Digit Recognition with CNN and OpenCV," TENCON 2024 - 2024 IEEE Region 10 Conference (TENCON), Singapore, Singapore, 2024, pp. 1865-1868, doi: 10.1109/TENCON61640.2024.10903079.

[4] F. G. Abdiwi, "Hybrid Machine Learning and Blockchain Technology for Early Detection of Cyberattacks in Healthcare Systems," International Journal of Safety & Security Engineering, vol. 14, no. 6, 2024, doi: 10.18280/ijsse.140622.

[5] World Health Organization, Report of the fifth WHO stakeholders meeting on gambiense and rhodesiense human African trypanosomiasis elimination, Geneva, Switzerland, 7-9 June 2023. Geneva: World Health Organization, 2024.

[6] R. Mao, L. Xie, X. Lu, J. Pei, X. Xu and S. Chang, "Harnessing Multiple Level Features to Improve Segmentation Performance of Deep Neural Network: A Case Study in Magnetic Resonance Imaging of Nasopharyngeal Cancer," in IEEE Access, vol. 12, pp. 82469-82481, 2024, doi: 10.1109/ACCESS.2024.3411099.

[7] S. Jégou, M. Drozdzal, D. Vazquez, A. Romero and Y. Bengio, "The One Hundred Layers Tiramisu: Fully Convolutional DenseNets for Semantic Segmentation," 2017 IEEE Conference on Computer Vision and Pattern Recognition Workshops (CVPRW), Honolulu, HI, USA, 2017, pp. 1175-1183, doi: 10.1109/CVPRW.2017.156.

[8] A. B. Tufail, I. Ullah, W. U. Khan, M. Asif, I. Ahmad, Y.-K. Ma, R. Khan, Kalimullah, and M. S. Ali, "Diagnosis of diabetic retinopathy through retinal fundus images and 3D convolutional neural networks with limited number of samples," Wireless Commun. Mobile Comput., vol. 2021, Art. no. 6013448, Nov. 2021, doi: 10.1155/2021/6013448.

[9] L. N. Fuzzell, R. B. Perkins, S. M. Christy, P. W. Lake, and S. T. Vadaparampil, "Cervical cancer screening in the United States: Challenges and potential solutions for underscreened groups," Preventive Medicine, vol. 144, Mar. 2021, Art. no. 106400. [Online]. Available: https://doi.org/10.1016/j.ypmed.2020.106400.

[10] R. Mehrotra, M. A. Ansari, R. Agrawal, and R. S. Anand, "A Transfer Learning approach for AI-based classification of brain tumors," Machine Learning with Applications, vol. 2, Art. no. 100003, Dec. 2020. [Online]. Available: https://doi.org/10.1016/j.mlwa.2020.100003.

[11] A. Yilmaz, M. Kalebasi, Y. Samoylenko, M. E. Guvenilir, and H. Uvet, "Benchmarking of Lightweight Deep Learning Architectures for Skin Cancer Classification using ISIC 2017 Dataset," arXiv preprint arXiv:2110.12270, Oct. 2021. [Online]. Available: https://doi.org/10.48550/arXiv.2110.12270.

[12] W. Gouda, N. U. Sama, G. Al-Waakid, M. Humayun, and N. Z. Jhanjhi, "Detection of Skin Cancer Based on Skin Lesion Images Using Deep Learning," Healthcare, vol. 10, no. 7, p. 1183, 2022. [Online]. Available: https://doi.org/10.3390/healthcare10071183.

[13] O. O. Abayomi-Alli, R. Damasevicius, S. Misra, R. Maskeliunas, and A. Abayomi-Alli, "Malignant skin melanoma detection using image augmentation by oversampling in nonlinear lower-dimensional embedding manifold," Turkish Journal of Electrical Engineering and Computer Sciences, vol. 29, no. 8, pp. 2600–2614, 2021. [Online]. Available: https://doi.org/10.3906/elk-2101-133.

[14] B. Cassidy, C. Kendrick, A. Brodzicki, J. Jaworek-Korjakowska, and M. H. Yap, "Analysis of the ISIC image datasets: Usage, benchmarks and recommendations," Medical Image Analysis, vol. 75, 102305, Jan. 2022. [Online]. Available: https://doi.org/10.1016/j.media.2021.102305

[15] A. G. Diab, N. Fayez, and M. M. El-Seddek, "Accurate skin cancer diagnosis based on convolutional neural networks," Indonesian Journal of Electrical Engineering and Computer Science, vol. 25, no. 3, pp. 1429–1441, Mar. 2022. [Online]. Available: https://doi.org/10.11591/ijeecs.v25.i3.pp1429-1441.

[16] E. Karampinis, P.-M. Nechalioti, K. E. Georgopoulou, G. Goniotakis, A. V. Roussaki Schulze, E. Zafiriou, and D. Kouretas, "Systemic oxidative stress parameters in skin cancer patients and patients with benign lesions," Stresses, vol. 3, no. 4, pp. 785–812, 2023. [Online]. Available: https://doi.org/10.3390/stresses3040054.

[17] N. A. AlSadhan, S. A. Alamri, M. M. Ben Ismail, and O. Bchir, "Skin cancer recognition using unified deep convolutional neural networks," *Cancers*, vol. 16, no. 7, p. 1246, 2024. [Online]. Available: https://doi.org/10.3390/cancers16071246.

[18] M. A. Rahman, E. Bazgir, S. M. S. Hossain, and M. Maniruzzaman, "Skin cancer classification using NASNet," International Journal of Science and Research Archive, vol. 11, no. 1, pp. 775–785, 2024. [Online]. Available: https://doi.org/10.30574/ijsra.2024.11.1.0106.

[19] M. Nieuwoudt, P. Jarrett, H. Matthews, M. Locke, M. Bonesi, B. Burnett, H. Holtkamp, C. Aguergaray, I. Mautner, T. Minnee, and M. C. Simpson, "Portable system for in-clinic differentiation of skin cancers from benign skin lesions and inflammatory dermatoses," JID Innovations, vol. 4, no. 1, p. 100238, Jan. 2024. [Online]. Available: https://doi.org/10.1016/j.xjidi.2023.100238.

[20] A. M. Smak Gregoor, T. E. Sangers, L. J. Bakker, et al., "An artificial intelligence based app for skin cancer detection evaluated in a population based setting," npj Digital Medicine, vol. 6, p. 90, 2023. [Online]. Available: https://doi.org/10.1038/s41746-023-00831-w.

[21] P. Varshney, C. Gupta, P. Girdhar, A. Mohan, P. Agrawal, and V. Madaan, "Data analysis using machine learning: An experimental study on UFC," in Machine Learning and Data Science: Fundamentals and Applications, P. Agrawal, C. Gupta, A. Sharma, V. Madaan, and N. Joshi, Eds. Wiley, 2022, ch. 2. [Online]. Available: https://doi.org/10.1002/9781119776499.

[22] R. W. Pettit, R. Fullem, C. Cheng, and C. I. Amos, "Artificial intelligence, machine learning, and deep learning for clinical outcome prediction," *Emerging Topics in Life Sciences*, vol. 5, no. 6, pp. 729–745, Dec. 2021. [Online]. Available: https://doi.org/10.1042/ETLS20210246.

[23] Q. An, S. Rahman, J. Zhou, and J. J. Kang, "A comprehensive review on machine learning in healthcare industry: Classification, restrictions, opportunities and challenges," *Sensors*, vol. 23, no. 9, p. 4178, 2023. [Online]. Available: https://doi.org/10.3390/s23094178

[24] S. Naeem, A. Ali, S. Anam, and M. M. Ahmed, "An unsupervised machine learning algorithms: Comprehensive review," International Journal of Computing and Digital Systems, vol. 13, no. 1, Apr. 2023. [Online]. Available: http://dx.doi.org/10.12785/ijcds/130172.

[25] A. Khan, M. Qureshi, M. Daniyal, and K. Tawiah, "A novel study on machine learning algorithm-based cardiovascular disease prediction," Health & Social Care in the Community, Feb. 2023. [Online]. Available: https://doi.org/10.1155/2023/1406060.

[26] Z. Wu, H. Jiang, K. Zhao, and X. Li, "An adaptive deep transfer learning method for bearing fault diagnosis," Measurement, vol. 151, p. 107227, Feb. 2020. [Online]. Available:

ISSN: 2312-8135 | Print ISSN: 1992-0652
info@journalofbabylon.com | jub@itnet.uobabylon.edu.iq | www.journalofbabylon.com

https://doi.org/10.1016/j.measurement.2019.107227

[27] C. Yu, R. Han, M. Song, C. Liu, and C. I. Chang, "A simplified 2D-3D CNN architecture for hyperspectral image classification based on spatial-spectral fusion," IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens., vol. 13, pp. 2485–2501, 2020, doi: 10.1109/JSTARS.2020.2983224.

[28] M. H. Hesamian, W. Jia, X. He, and P. Kennedy, "Deep learning techniques for medical image segmentation: Achievements and challenges," *J. Digit. Imaging*, vol. 32, pp. 582–596, 2019, doi: 10.1007/s10278-019-00227-x.

[29] B. Yang, C.-G. Lee, Y. Lei, N. Li, and N. Lu, "Deep partial transfer learning network: A method to selectively transfer diagnostic knowledge across related machines," *Mechanical Systems and Signal Processing*, vol. 156, p. 107618, Jul. 2021. https://doi.org/10.1016/j.ymssp.2021.107618.

[30] B. Yang, Y. Lei, F. Jia, and S. Xing, "An intelligent fault diagnosis approach based on transfer learning from laboratory bearings to locomotive bearings," *Mechanical Systems and Signal Processing*, vol. 122, pp. 692–706, May 2019. [Online]. Available: https://doi.org/10.1016/j.ymssp.2018.12.051

[31] J. Yin, M. Xu, and H. Zheng, "Fault diagnosis of bearing based on symbolic aggregate approximation and Lempel-Ziv," *Measurement*, vol. 138, pp. 206–216, May 2019. [Online]. Available: https://doi.org/10.1016/j.measurement.2019.02.011.

[32] M. G. Tektonidou, "Cardiovascular disease risk in antiphospholipid syndrome: Thrombo-inflammation and atherothrombosis," *Journal of Autoimmunity*, vol. 128, p. 102813, Apr. 2022. [Online]. Available: https://doi.org/10.1016/j.jaut.2022.102813

[33] S. A. Albelwi, "Deep architecture based on DenseNet-121 model for weather image recognition," *International Journal of Advanced Computer Science and Applications (IJACSA)*, vol. 13, no. 10, 2022.

[34] G. Alwakid, W. Gouda, M. Humayun, and N. Z. Jhanjhi, "Diagnosing melanomas in dermoscopy images using deep learning," *Diagnostics*, vol. 13, no. 10, p. 1815, 2023. [Online]. Available: https://doi.org/10.3390/diagnostics13101815.

[35] J. N. Saeed and S. R. M. Zeebaree, "Skin lesion classification based on deep convolutional neural networks architectures," Journal of Applied Science and Technology Trends (JASTT), vol. 2, no. 1, pp. 41–51, Mar. 2021. [Online]. Available: https://doi.org/10.38094/jastt20189.

[36] W. Gouda, N. U. Sama, G. Al-Waakid, M. Humayun, and N. Z. Jhanjhi, "Detection of skin cancer based on skin lesion images using deep learning," *Healthcare*, vol. 10, no. 7, p. 1183, 2022. [Online]. Available: https://doi.org/10.3390/healthcare10071183

[37] T. Mazhar, I. Haq, A. Ditta, S. A. H. Mohsan, F. Rehman, I. Zafar, J. A. Gansau, and L. P. W. Goh, "The role of machine learning and deep learning approaches for the detection of skin cancer," *Healthcare*, vol. 11, no. 3, p. 415, 2023. [Online]. Available: https://doi.org/10.3390/healthcare11030415.

[38] R. A. Naqvi, D. Hussain, and W.-K. Loh, "Artificial intelligence-based semantic segmentation of ocular regions for biometrics and healthcare applications," Computers, Materials & Continua, vol. 66, no. 1, pp. 715–732, 2021. [Online]. Available: https://doi.org/10.32604/cmc.2020.013249.

[39] S. S. U. Hassan, S. Q. Abbas, F. Ali, M. Ishaq, I. Bano, M. Hassan, H.-Z. Jin, and S. G. Bungau, "A comprehensive in silico exploration of pharmacological properties, bioactivities, molecular docking, and anticancer potential of vieloplain F from Xylopia vielana targeting B-Raf kinase," Molecules, vol. 27, no. 3, p. 917, 2022. [Online]. Available: https://doi.org/10.3390/molecules27030917.

[40] G. Alwakid, W. Gouda, M. Humayun, and N. U. Sama, "Melanoma detection using deep learning-based classifications," Healthcare, vol. 10, no. 12, p. 2481, 2022. [Online]. Available: https://doi.org/10.3390/healthcare10122481.

[41]

ISSN: 2312-8135 | Print ISSN: 1992-0652

info@journalofbabylon.com | jub@itnet.uobabylon.edu.iq | www.journalofbabylon.com

مجلة جامعة بابل للعلوم الصرفة والتطبيقية مجلة جامعة بابل للعلوم الصرفة والتطبيقية

ISSN: 2312-8135 | Print ISSN: 1992-0652
info@journalofbabylon.com | jub@itnet.uobabylon.edu.iq | www.journalofbabylon.com

الخلاصة

المقدمة:

لا يزال سرطان الجلد أحد أكثر الأورام الخبيثة انتشارًا في جميع أنحاء العالم، وتلعب الوقاية منه والكشف المبكر عنه دورًا حاسمًا في تقليل معدلات الإصابة بالأمراض والوفيات. أدت التطورات الحديثة في التعلم العميق إلى تعزيز دقة التصنيف الآلي لآفات الجلد بشكل كبير. ومع ذلك، لا تزال هناك العديد من التحديات المتمثلة في تصميم نموذج هجين ذي كفاءة حسابية عالية، يدمج بنىً متعددة خفيفة الوزن.

المواد وطرق العمل :

يقترح هذا العمل نموذج تعلم عميق هجين يجمع بين معماريتي **MobileNetV3+** و **ShuffleNetV2.** ولتعزيز أداء النموذج، تم استخدام خوارزمية **تحسين قطيع الفيلة (EHO)** لضبط فرط المعلمات الأساسية مثل معدل التعلم، حجم الدفعة، وعدد العصور التدريبية. تم جمع بيانات صور الجلد من **SkinDataSe** وأرشيف **ISIC**، وشملت حالات حميدة وخبيثة. تضمنت المعالجة المسبقة تحسين التباين وتقنيات متنوعة لزيادة البيانات مثل التدوير، التحجيم، الانعكاس، والإزاحة بهدف تحسين التعميم. كما تم تغيير حجم جميع الصور إلى **3×224×224بكسل** وتطبيعها إلى المجال [0,1]. بعد ذلك، قُسمت البيانات إلى 70% للتدريب، 15% للتحقق، و15% للاختبار.

النتائج:

أظهر النموذج الهجين أداءً متميزًا، حيث حقق دقة تدريب بلغت **99.72%** ودرجة **F1** بمقدار **0.93**، مما يشير إلى انخفاض كبير في الإيجابيات الكاذبة، ويؤكد فعالية النموذج في التمييز بين الآفات الجلدية الحميدة والخبيثة.

الإستنتاجات:

تؤكد نتائج هذه الدراسة إمكانات إطار التعلم العميق الهجين المقترح، والذي تم تحسينه عبر **EHO**، في تقديم تشخيص دقيق وسريع ومبكر لسرطان الجلد. يمكن لهذا النظام أن يعزز بشكل كبير عمليات اتخاذ القرار السريري ويسهم في تحسين نتائج المرضى من خلال التدخل في الوقت المناسب.

الكلمات المفتاحية: سرطان الجلد، تحسين قطيع الفيلة(EHO) ، التعلم العميق بالنقل(Transfer Deep Learning) ، الشبكات العصبية التلافيفية الهجينة(Hybrid CNN) ، الصور الطبية.