SKIN CANCER CLASSIFICATION USING THE DEEP LEARNING MODEL OPTIMIZED RESNET50

Sainni Sai Kiran, Department of Computer Science Engineering, Indian Institute of Industry Interaction Education and Research, Chennai, Tamil Nadu 600066
Parvatham Vijay Kumar, Department of Computer Science Engineering, Indian Institute of Industry Interaction Education and Research, Chennai, Tamil Nadu 600066
Yetcherla Sri Priya, Department of Computer Science Engineering, Indian Institute of Industry Interaction Education and Research, Chennai, Tamil Nadu 600066
Vasa Sambasiva Rao, Department of Computer Science Engineering, Indian Institute of Industry Interaction Education and Research, Chennai, Tamil Nadu 600066

ABSTRACT

The skin acts as a barrier against several external threats and is the body's first line of defense for essential organs. Even though the skin serves as a barrier, it can still become infected due to bacteria, viruses, or other objects like dust. If treatment is not received, even little skin lesions might develop into serious health problems. Making a timely recovery from skin illnesses is dependent on receiving an accurate diagnosis. Thus, the goal of this research is to create a system that uses Optimized Resnet50 (O-Resnet50) technology to detect skin diseases. Here, the SSA (Salp swarm algorithm) executes the best procedure. In this work, we introduce a novel method for digitally eliminating hair from medical photos by combining Gaussian filtering for de-blurring and denoising with morphological filtering techniques like the Black Hat transformation and inpainting algorithm. To precisely segment out the impacted lesions, we also use the automatic Grabcut segmentation method. Subsequently, the OCNN classifies the lesions. Using the ISIC-2019 dataset as a test dataset, experiments produced improved accuracy of 98.6% and precision of 99%.

Keywords:
Skin diseases, Optimized Resnet50, Salp swarm algorithm, Black Hat transformation, Grabcut segmentation

I. Introduction

One common type of cancer that starts in the skin's cells is skin cancer. UV radiation from the sun or artificial sources, such as tanning beds, is the main cause of it. There are several forms of skin cancer, and each has unique traits and behaviors [1]. The most common type, known as BCC (Basal Cell Carcinoma), typically manifests as a flesh-colored scar-like lesion or a waxy lump. If treatment is not received, it might lead to local tissue damage but seldom spreads. Firm, red nodules or scaly lesions are the hallmarks of SCC (Squamous Cell Carcinoma), which develops from the squamous cells. Compared to BCC, it has a higher chance of metastasis. Melanoma is a cancer that starts in the melanocytes and can spread to any part of the body. It usually manifests as black spots or moles with irregular shapes. If melanoma is not identified and treated promptly, it can spread quickly and become extremely aggressive. Merkel cell carcinoma, or MCC, is an uncommon but severe kind of cancer. It grows quickly and has a strong potential for metastasis. It first appears as hard, painless nodules [2]. Depending on the kind and stage of the cancer, treatment for skin cancer usually consists of surgically removing the tumor in addition to additional treatments like chemotherapy, radiation therapy, or immunotherapy. Skin lesions include a broad range of abnormalities, such as melanoma, benign moles, and many types of dermatological illnesses. Ensuring that these lesions are accurately and promptly classified is crucial for optimizing patient care and enhancing clinical results. Dermatologists' manual classification, however, can be laborious and subjective, which can vary in diagnosis [3]. In the realm of dermatology, DL (deep learning) approaches for skin lesion classification have shown promise. As deep learning models have been developed, advances in computer vision and machine learning have enabled them to diagnose a wide range of skin lesions with surprising accuracy. Compared to conventional methods, this technology offers faster and more accurate assessments, which holds tremendous potential for improving the diagnosis and treatment of skin diseases. A type of artificial intelligence called deep learning (DL) has demonstrated potential to help with skin cancer diagnosis and therapy. Through extensive training of deep learning algorithms on vast collections of skin imaging data, scientists have created systems that can reliably identify skin lesions as benign or malignant. Dermatologists can make more precise diagnosis and treatment decisions with the help of these tools. Furthermore, by using DL approaches, treatment outcomes can be predicted and customized treatment regimens based on the unique characteristics of each patient can be created. Deep learning models can find patterns and forecast outcomes by examining patient data and therapy responses; these findings could help guide therapeutic decision-making [4]. In conclusion, DL has enormous potential to improve skin cancer diagnosis and therapy by increasing diagnosis accuracy, forecasting treatment results, and facilitating individualized treatment plans. The following is the arrangement of the forthcoming sections: Section 2 lays out the relevant literature, Section 3 discusses potential strategies, Section 4 offers the findings and a discussion, and Section 5 concludes the framework proposal.

II. Correlated Publications

An enhanced version of Eff2Net, with the ECA (Efficient Channel Attention) block, was introduced by Karthik et al. [5]. This study attempts to replace the ECA block in the EfficientNetV2 architecture with the traditional SE (Squeeze and Excitation) block. Four instances, including melanoma, acne, psoriasis, and AK (actinic keratosis), were used in the experimentation. A system for digital hair removal was developed by Ahammed et al. [6] using morphological filtering methods like the inpainting algorithm and the Black Hat transformation. Additionally, they created an automatic Grabcut segmentation method using the Hue Saturation Value color space and k-means clustering to identify impacted lesions. Furthermore, the Gray Level Co-occurrence Matrix and statistical factors were utilized to extract essential features. Their study analyzed two commonly used datasets to choose the best model and assessed the performance of three machine learning algorithms for the classification of skin diseases. CNN models using the Imagenet dataset are refined to extract more pertinent characteristics, according to Sadik et al. [7]. The authors' goal was to improve the performance of our suggested method by using the lessons discovered from a large-scale dataset such as Imagenet. In order to evaluate the performance of our approach, we ran tests on a number of popular CNN models. These models are widely used in the research community and have shown high performance in a variety of computer vision applications. We sought to validate the significance of transfer learning and data augmentation and to set a benchmark by comparing the performance of our proposed technique with these well-known architectures. For the purpose of classifying skin lesions, Elashiri et al. [8] introduced HSBSO (Hybrid Squirrel Butterfly Search Optimization) with MLSTM (Modified Long Short Term Memory). At first, the histogram equalization procedure handled the contrast augmentation. Image segmentation is then carried out by applying the FCM (Fuzzy C Means) method. Three different approaches are used to extract deep features from the segmented images. These methods take features from the last levels and concatenate them.

III. Suggested Approach

Skin lesions include a broad spectrum of abnormalities, such as melanoma, benign moles, and many types of dermatological disorders. Prompt and precise identification of these lesions is essential for prompt intervention and efficacious therapy. Dermatologists' manual diagnosis, however, can be laborious and subjective, resulting in a range of diagnoses. Skin lesion categorization can be mechanized and objectively solved with DL using optimization-based techniques. Deep learning algorithms can identify intricate patterns and features suggestive of various skin disorders by utilizing extensive datasets of labeled photos of skin lesions. Afterwards, these models are able to categorize lesions with a high degree of efficiency and accuracy, which may help to lower diagnostic errors and enhance patient outcomes. The suggested classification of skin cancer is shown in Figure 1.

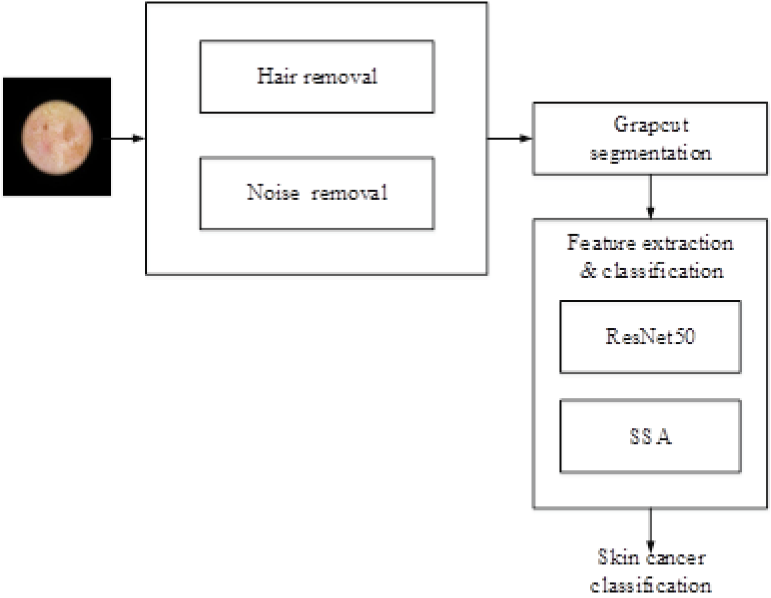

Figure 1: Proposed skin cancer classification

3.1 Pre-processing

To improve the quality and usefulness of skin images in a disease detection system, image preprocessing is crucial. Unwanted components in skin photos, such as hair, noise, or distortion, can have an impact on how well image processing algorithms work. Our goals in preprocessing the photos are to enhance image quality, decrease complexity, and boost accuracy in the processing stages that follow. Typical preprocessing actions consist of:

Image Resizing: To maintain consistency and ease further processing, the images are downsized to a standard size, like 512 × 512. This lowers processing times, boosts system speed, and increases the quality of features that are taken from the photos.

Hair Removal: In skin imaging, undesired hair might cause issues for algorithms used in illness identification. Hair can be successfully removed from photos using methods like morphological filtering or inpainting algorithms, which enhances the skin region's clarity.

Elimination of Noise: Noise in skin scans can mask significant information and reduce the precision of illness identification. Utilizing methods for reducing noise, including Gaussian blurring or median filtering, aids in image smoothing and increases signal-to-noise ratio.

3.2 Dividing

In many image processing applications, image segmentation—the division of a picture into discrete areas or groups—is essential. In order to distinguish items or elements of interest from the backdrop, it makes use of characteristics like texture, contrast, brightness, color, and gray level. Precise segmentation plays a critical role in medical imaging by enabling the differentiation of lesions from healthy tissue, which facilitates diagnosis and therapy planning. Nevertheless, because lesions vary widely in size, shape, and color and have minimal contrast with surrounding healthy tissue, segmenting microscopic pictures presents special difficulties. Different segmentation techniques, such as threshold-based, region-based, cluster-based, and edge-based methods, have been developed to tackle these difficulties. In this work, we utilize the HSV (Hue Saturation Value) color space in conjunction with k-means clustering to implement the Grabcut technique for picture segmentation. Grabcut is an automatic segmentation approach that estimates the color distributions of the background and foreground (target item) using the Gaussian Mixture Model (GMM). This allows the foreground to be extracted from the background with the least amount of user interaction. Grabcut simplifies and generates correct results from segmentation by automatically creating a bounding rectangle around the target object and iteratively fine-tuning the segmentation. By combining Grabcut with k-means clustering and the HSV color space, our method seeks to improve the precision and performance of image segmentation, especially when dealing with microscopic pictures that have complex lesion structures and low contrast. This methodological framework has the potential to make it easier to analyze and understand medical images, which will ultimately lead to better patient care and diagnostic results.

3.3 Classification of skin lesions

Classification, or putting a set of data into different groups or classes, is the last step in this investigation. The objective here is to use the attributes that were taken from the photos to forecast the kind of skin condition. Different approaches are used for categorization, depending on the application's requirements and the dataset's properties. These approaches comprise a variety of algorithms and strategies intended to efficiently identify patterns and connections in the data, hence enabling precise classification into pertinent classifications.

ResNet50: Also referred to as a residual network, it addresses the vanishing gradient issue by introducing the idea of skip connections, which helps to prevent distortion that could occur as the network gets deeper and more intricate. ResNet50, one of the ResNet model variants, is one of the models used. As seen in Figure 2, the ResNet-50 architecture is a network made up of 50 layers that were trained with ImageNet. Convolutional layers, four convolutional blocks, max pooling, and average pooling are some of its features that help prevent accuracy degradation and make it easier to create deeper CNNs without sacrificing accuracy. The ResNet-50 design made it possible to create CNNs with even greater depth without compromising accuracy. Interestingly, ResNet50 was one of the first CNNs to use batch normalizing features. An outline of the ResNet50 architecture is provided below:

Image input of a given size is accepted by the input layer.

Convolutional Layers: The rectified linear unit (ReLU) activation function, batch normalization layer, and convolutional layer are the order in which the network is constructed. To extract features from the input photos, this sequence is performed multiple times.

Residual Blocks: These are the fundamental components that make up ResNet designs. ResNet50 is made up of many residual blocks, each with several convolutional layers. Shortcut connections that bypass one or more layers are present within each residual block, which facilitates the easier gradient flow during training.

Pooling Layers: To minimize the spatial dimensions of the feature maps without sacrificing crucial information, periodic pooling layers—typically max pooling—are employed.

Fully Connected Layer: At the end of the network, one or more fully connected layers usually come after a global average pooling layer. By averaging all of the feature map values, the global average pooling layer lowers each feature map to a single value. Classification is carried out by the fully linked layers using the learned characteristics.

Output Layer: In multi-class classification tasks, the last layer of the network generates the output predictions, usually using a softmax activation function. Figure 2 shows the ResNet50 structure.

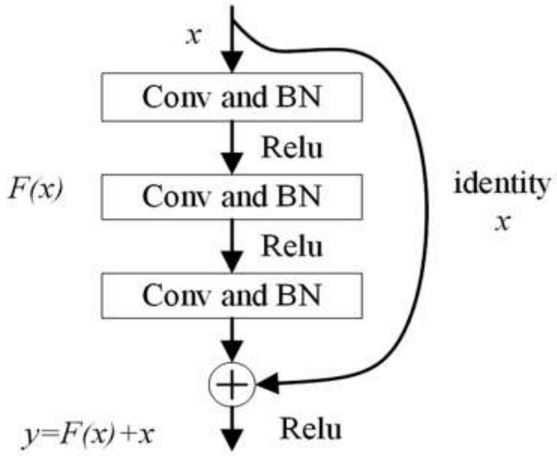

Figure 2: Structure of ResNet50

SSA is a metaheuristic algorithm that draws inspiration from nature and emulates the swarming behavior of salps, a kind of marine organism. To solve optimization issues optimally, a population of virtual salps collaborates to explore the solution space in SSA. The algorithm iteratively adjusts the placements of salps based on their interactions with surrounding salps and their own fitness. Each salp represents a possible solution. Salps optimize their placements by adhering to a set of guidelines derived from the behavior of salps seen in the wild. These guidelines include avoiding impediments (low-quality solutions), gravitating toward food sources (high-quality solutions), and moving in the direction of better options. The swarm of virtual salps converges towards promising areas of the solution space by repeatedly applying these principles, thus raising the caliber of solutions. In the SSA, the positions of salps are defined in an $n-$-dimensional search space, where n is the number of variables of the given problem. These positions are stored in a two-dimensional matrix denoted as y . Additionally, a target food source, referred to as Fs_k , is assumed to exist in the search space as the swarm's objective. To update the position of the leader salp y_k^1, the following equation is proposed:

$$y_k^1 = \begin{cases} Fs_k - c1\left(\left(ub_k - lb_k\right)c2 + lb_k\right) & c3 \geq 0 \\ Fs_k - c1\left(\left(ub_k - lb_k\right)c2 + lb_k\right) & c3 < 0 \end{cases} \quad (1)$$

where ub_k and lb_k are upper and lower bounds; $c1, c2, c3$ are random numbers. Equation (1) illustrates that the leader solely adjusts its position concerning the food source. The coefficient $c1$ holds paramount significance in SSA, as it governs the equilibrium between exploration and exploitation, delineated as follows:

$$c1 = 2\exp-\left(\frac{4t}{T}\right) \quad (2)$$

where t and T are present and maximum iterations.

The parameters $c2, c3$ are randomly generated numbers distributed uniformly within the interval [0, 1]. These parameters play a crucial role in determining whether the next position in the k^{th} dimension should move towards positive infinity or negative infinity, as well as determining the step size. To update the positions of the followers y_k^1, the following equations are employed, akin to Newton's law of motion:

$$y_k^m = \frac{1}{2}gl^2 + u_0 l \quad (1)$$

where g and u_0 are the acceleration and initial velocity.

In optimization, where time is represented by iterations l and the discrepancy between iterations is fixed at 1, and with the u_0 set to 0, this equation can be represented as follows:

$$y_k^m = \frac{1}{2}\left(y_k^m + y_k^{m-1}\right) \quad (1)$$

Algorithm 1: Pseudocode of the SSA

```
Initialize salp population randomly within search space
Evaluate fitness of each salp in the population
Set leader salp as the one with the best fitness
Repeat until termination criterion is met:
for each salp in the population:
Update position of salp based on leader and food source
      Evaluate fitness of updated salp
      Update leader salp if necessary
   End for
End loop
Return best solution found
```

The approach mentioned above initializes salps at random inside the search space.

- The objective function that needs to be optimized is the basis for evaluating each salp's fitness.
- The salp who is the most fit is chosen as the leader.
- Until a termination requirement (such as the maximum number of iterations) is satisfied, the main loop iterates.
- Every time the loop is repeated, each salp modifies its location in relation to the leader's and the food source's positions. This stage, which directs salps toward better solutions, is usually based on a mathematical equation.
- Each salp's fitness is reassessed when positions are updated, and if needed, the leader salp may also be altered.
- Lastly, the algorithm output returns the optimal solution that was discovered throughout the optimization process.

IV. Analysis of the results

4.1 Details of the dataset

In the fields of dermatology and medical image analysis, the ISIC 2019 dataset—also referred to as the International Skin Imaging Collaboration (ISIC) 2019 challenge dataset—is a commonly used benchmark dataset. It includes ground truth labels, accompanying information, and high-resolution dermoscopic images of skin lesions. The dataset was made available as a component of the ISIC 2019 challenge, which sought to further the study of automated machine learning and computer vision methods for the identification and categorization of skin lesions. A wide range of skin lesion types are represented in the dataset's photos, such as actinickeratosis (AK), melanoma (MEL), benign keratosis (BKL), melanocytic nevus (NV), vascular lesion (VASC), SCC, and BCC (DF). Figure 3 displays the ISIC 2019 dataset's input and scaled photos.

(a)

(b)

Figure 3: Input and resized images of ISIC 2019 dataset

4.2 Performance Analysis

You would normally take the following actions to undertake accuracy, precision, recall, and specificity analysis for skin cancer classification:

Accuracy: The percentage of cases among all instances that are accurately classified.

Precision is defined as the percentage of correctly classified positive instances among all positive instances.

Recall: The percentage of real positive instances among all correctly identified positive instances.

Specificity: The percentage of real negative instances among all correctly categorized negative instances

4.3 Analysis in comparison

The comparative study of several DL models, including CNN, SVM (support vector machine), NB (naïve bayes), and the suggested O-ResNet50 as indicated in Table 1, is presented in the part that follows.

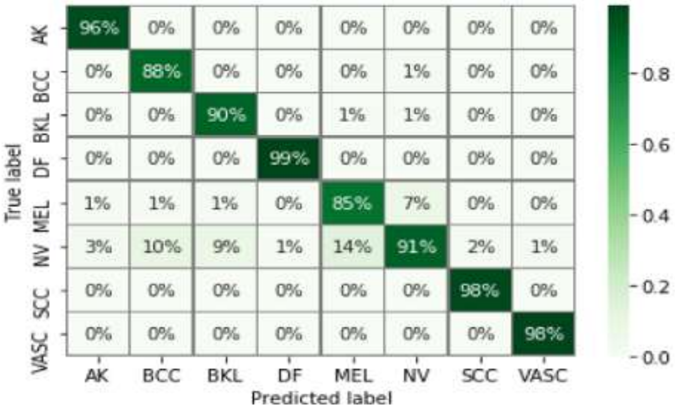

Figure 4: Confusion matrix of the proposed O-ResNet50

The suggested O-ResNet50's confusion matrix is shown in Figure 4. 95 percent of the samples are classed as VASC, 88 percent are classified as BCC, 90 percent are classified as BKL, 99% are classified as DF, 85 percent are classified as MEL, 91% are classified as NV, 88 percent are classified as AK, and 98 percent are classified as SCC.

Table 1: comparative analysis of different DL models

Methods	Accuracy (%)	Precision (%)	Recall (%)	Specificity (%)
CNN	80.2	97.9	89.4	93.1
SVM	78.4	97.5	90.2	94.5
NB	79.1	97.2	91.3	95.2
proposed O-ResNet50	98.6	99	97.2	96.9

The accuracy-loss curves for the suggested O-ResNet50 model are shown in Figure 4. This image probably illustrates how, as training goes on, the accuracy performance of the model changes in proportion to the loss function. Typically, accuracy-loss curves show how a model's accuracy and loss function change over the course of training iterations. The loss function often drops as the model is trained, showing that the model's predictions are getting closer to the actual values. The accuracy usually increases at the same time, indicating the model's growing capacity to accurately categorize or forecast data items. The accuracy-loss curves offer important insights into the model's performance and training dynamics.

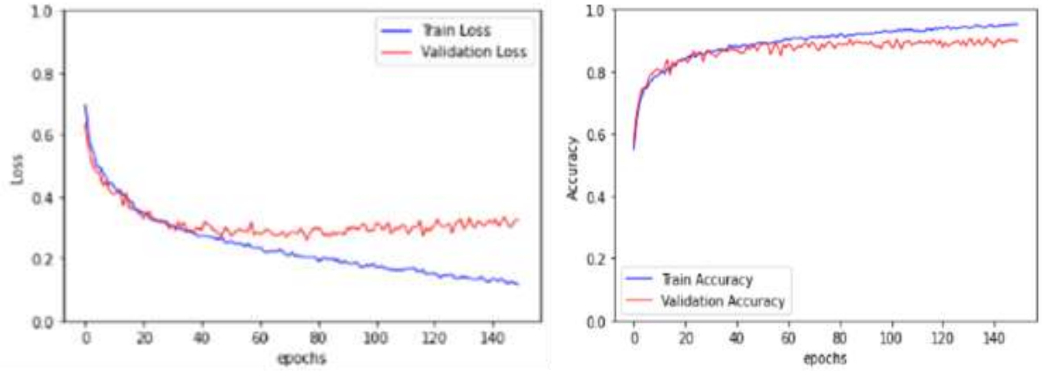

Figure 4: Accuracy-loss curves of the proposed O-ResNet50

The ROC-AUC curve sheds important light on the proposed model's capacity to distinguish between positive and negative samples. Better performance is indicated by a larger AUC value; a value of 1 denotes perfect classification, while a value of 0.5 denotes random guessing. As seen in Figure 5, analyzing the ROC-AUC curve of the suggested model aids in determining how well it can distinguish between various classes and make accurate predictions.

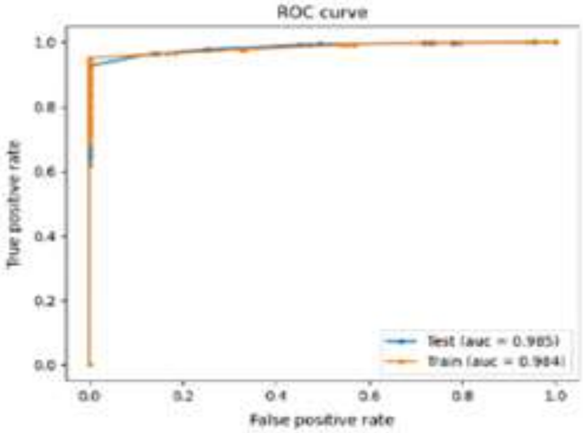

Figure 5: ROC-AUC curve of the proposed O-ResNet50

V. Conclusion

Deep learning has demonstrated encouraging results in the identification of skin cancer and has great potential to increase diagnostic speed and accuracy. Deep learning models can be trained to recognize complex patterns and features that correspond to many kinds of skin lesions by analyzing dermatoscopic pictures. This allows for automated categorization and detection of skin lesions. According to this work, O-ResNet50 can perform on par with dermatologists in some situations, if not better. These models are useful for correctly categorizing skin lesions into benign and malignant groups, which helps in skin cancer early identification and treatment. These models are also accessible and scalable, which makes them useful resources for medical professionals in a variety of contexts, including impoverished places with limited access to dermatologists with advanced training. Nonetheless, there is still a need for study and development to address issues such the lack of data, the interpretability of the model, and the generalization to different populations. To guarantee the dependability and efficiency of deep learning systems in clinical practice, these issues must be resolved. The use of O-ResNet50 in skin cancer diagnostics opens up interesting new possibilities for enhancing patient outcomes in the future. To fully utilize deep learning in the fight against skin cancer, further research and cooperation between dermatologists, doctors, and computer scientists are required.

References

[1]Singh, J., Sandhu, J. K., & Kumar, Y. An analysis of detection and diagnosis of different classes of skin diseases using artificial intelligence-based learning approaches with hyper parameters. Archives of Computational Methods in Engineering, 2024, 31(2), 1051-1078.

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

[3]Haggenmüller, S., Maron, R. C., Hekler, A., Utikal, J. S., Barata, C., Barnhill, R. L et.al. Skin cancer classification via convolutional neural networks: systematic review of studies involving human experts. European Journal of Cancer, 2021, 156, 202-216.

[4]Dorj, U. O., Lee, K. K., Choi, J. Y., & Lee, M. The skin cancer classification using deep convolutional neural network. Multimedia Tools and Applications, 2018, 77, 9909-9924.

[5]Karthik, R., Vaichole, T. S., Kulkarni, S. K., Yadav, O., & Khan, F. Eff2Net: An efficient channel attention-based convolutional neural network for skin disease classification. Biomedical Signal Processing and Control, 2022, 73, 103406.

[6]Ahammed, M., Al Mamun, M., & Uddin, M. S. A machine learning approach for skin disease detection and classification using image segmentation. Healthcare Analytics, 2022, 2, 100122.

[7]Sadik, R., Majumder, A., Biswas, A. A., Ahammad, B., & Rahman, M. M. An in-depth analysis of Convolutional Neural Network architectures with transfer learning for skin disease diagnosis. Healthcare Analytics, 2023, 3, 100143.

[8]Elashiri, M. A., Rajesh, A., Pandey, S. N., Shukla, S. K., & Urooj, S. Ensemble of weighted deep concatenated features for the skin disease classification model using modified long short term memory. Biomedical Signal Processing and Control, 2022, 76, 103729.

[9]https://www.kaggle.com/datasets/andrewmvd/isic-2019

[10] Mirjalili, S., Gandomi, A. H., Mirjalili, S. Z., Saremi, S., Faris, H., & Mirjalili, S. M. Salp Swarm Algorithm: A bio-inspired optimizer for engineering design problems. Advances in engineering software, 2017, 114, 163-191.