PRAGUE UNIVERSITY OF ECONOMICS AND BUSINESS

Acta Informatica Pragensia
2023, Volume 12, Issue 1, pp. 141–159
Special Issue: Deep Learning Blockchain-enabled Technology for Improved Healthcare Industrial Systems
https://doi.org/10.18267/j.aip.211

Article Open Access

Multi-Class Skin Cancer Classification Using a Hybrid Dynamic Salp Swarm Algorithm and Weighted Extreme Learning Machines with Transfer Learning

Ramya Panneerselvam , Sathiyabhama Balasubramaniam

Department of Computer Science and Engineering, Sona College of Technology, Salem, India

Corresponding author: Ramya Panneerselvam (ramyapaneerselvam1@outlook.com)

Abstract

Skin cancer is a significant healthcare problem with a high mortality rate worldwide. Skin lesions occur due to the abnormal growth of skin cells in humans. Failure of early prediction and proper lesion diagnosis may lead to a malignant stage. In recent times, different skin lesion images have appeared with high similarity. Hence, classification is a more challenging task with imbalances in the dataset. The proposed work is implemented as a hybrid model with a dynamic salp swarm algorithm (DSSA) with a weighted extreme learning machine (DSSA-WELM) that addresses the imbalances in the dataset and performs the classification with higher accuracy. GoogleNet is a pre-trained network model used with the hybrid model, which helps converge faster with the optimization process. The extreme learning machine (ELM) is a multiclass classifier for accurate dermoscopic image classification. The DSSA, the best feature selection algorithm enhances the classification accuracy of the WELM. Image classification is accomplished with the International Skin Imaging Collaboration 2019 benchmark dataset. The proposed solution classifies images into eight classes: melanoma, melanocytic nevus, basal cell carcinoma, actinic keratosis, dermatofibroma, vascular lesion, squamous cell carcinoma and unknown lesion. The efficiency of the proposed solution is proved by comparing it with various state-of-the-art approaches such as support vector machine (SVM), ELM, and particle swarm optimization (PSO) methods. Results are evaluated using standard metrics of sensitivity, specificity and precision. The proposed solution outperforms all these older approaches.

Keywords

Skin cancer classification; Convolution neural network; Transfer learning; Extreme learning machine; Swarm algorithm; Metaheuristics.

Citation: Panneerselvam, R., & Balasubramaniam, S. (2023). Multi-Class Skin Cancer Classification Using a Hybrid Dynamic Salp Swarm Algorithm and Weighted Extreme Learning Machines with Transfer Learning. *Acta Informatica Pragensia*, 12(1), 141–159. https://doi.org/10.18267/j.aip.211

Special Issue Editors: Mazin Abed Mohammed, University of Anbar, Iraq
Seifedine Kadry, Noroff University College, Norway
Oana Geman, Ștefan cel Mare University of Suceava, Romania

Academic Editor: Zdenek Smutny, Prague University of Economics and Business, Czech Republic

1 Introduction

In recent years, skin cancer has outspreads broadly due to the uncontrollable and abnormal growth of cells in the epidermis. The outer layer of the skin triggers transmutation with deoxyribonucleic acid (DNA) damage that leads to skin cells multiplying rapidly and transforming into malignant tumours. The major reasons for the cause of skin cancer are harmful ultraviolet rays and ultraviolet (UV) tanning beds, and occupational hazards such as very long exposure to the sun. In general, skin cancer is broadly classified as basal cell carcinoma (BCC), melanoma and squamous cell carcinoma (SCC), which are the most common skin cancer types that put many people's lives at risk. Among various cancers that occurs in human beings, skin cancer is the fifth most common cancer in females and the sixth most common in males. Among these, melanoma is the deadliest cancer that leads to numerous deaths every year in the USA and India.

Various types of cancers affect individuals all over the world. One among those is skin cancer, a common human malignancy, which occurs due to overexposure to UV rays and sun. Skin cancers are seriously viewed in the USA and India; about 2 in 10 people there are suffering from this deadliest disease. Early prediction of skin cancer would certainly save people's lives by up to 50% (Hossin et al., 2020). However, the existing method of diagnosis involves detection by dermatologists or self-examination by individuals, or certain techniques such as biopsy, which is considered an invasive technique that only determines abnormal cells in a proliferated region.

This traditional approach to classifying the type of skin disease is a time-consuming process. Therefore, to overcome these clinical processes and to assist them, much research is done with machine learning techniques, which are used more effectively and intensively in the analysis of medical information to detect and predict disease automatically (Zhang et al., 2021). The results obtained prove to be an authentic and reliable tool that reduces human errors and certainly improves the quality of medical care, which provides more safety and comfort to patients at all, levels.

The BCC would pop up as a pinkish patch on the skin due to UV exposure, (Teng et al., 2021), which is mostly found on the face, head, neck, arms, legs, and abdomen. Moreover, it can spread over the body and grow in bones and nerves. The SCC usually occurs on the outer surface of the skin and appears as red firm bumps or scaly patches. It is due to over exposure to sun and is mostly found in people with pale light skin. It does not spread across the human body and is a slow-growing cancer cell, found in regions such as the ear, face, neck, and chest. The BCC and SCC carry UV rays and damage the DNA directly.

The death rate of the SCC is higher than that of the BCC but lesser than that of melanoma skin cancer (Hossin et al., 2020). As mentioned earlier, if these lesions are detected early, they can be treated e.g., by photodynamic therapy, chemotherapy, and electrodessication. The SCC is considered the second most common cancer disease in the world. Malignant melanoma is the most dangerous and deadliest skin cancer in the world. In 2020, the American Society of Clinical Oncology found that around 57,043 people worldwide died due to melanoma (ASCO, 2022). These cancers usually start or emerge from moles and are observed as dark spots on the outer layer of the skin. This fast-growing skin disease can spread easily over the body. Malignant melanoma is also due to UV rays but damages the DNA indirectly.

Dermatofibroma is a type of eczema skin cancer, which is found to be inflamed and it is one of the most common diseases. Actinic keratosis is a scaly bump, caused due to exposure to the sun. Actinic keratosis progresses to cancer. The SCC is a common form of skin cancer that begins as an ulcer, which is due to abnormal growth of cells. It usually develops in sun-exposed areas.

Cancer transforms the life of both the patients and their loved ones, bringing about physical, psychological, social, and spiritual anguish (De Tejada et al., 2016). The threat of receiving a cancer diagnosis to one's sense of security is compounded by the sentiments and emotions the disease causes, which disrupt daily life. Cancer sufferers must manage the stress brought on by the diagnosis as well as the physical side effects of the disease and its treatment as well as ongoing health issues, disabilities, weariness, and pain. Daily struggles brought on by emotional stress and mental health issues may include inability to work, financial challenges, and a lack of social support. This greatly influenced our decision to undertake our research and create a more accurate and economical model for disease prediction. Likewise, we intend to assist medical professionals in saving lives.

Early diagnosis of skin cancer is more significant in our day-to-day life with the expeditious growth of technologies that have involved more researchers in various machine learning and deep learning techniques to detect and classify cancer types at the inception stage. Many researchers have focused on image classification, an active problem of computer vision, and developed many computer aided design (CAD) techniques using machine learning and deep feature extraction techniques. In general, classification performance depends on exactness of the feature representation in an image. Dermatologists can detect and identify the type of skin lesions using these CAD techniques with higher accuracy and faster.

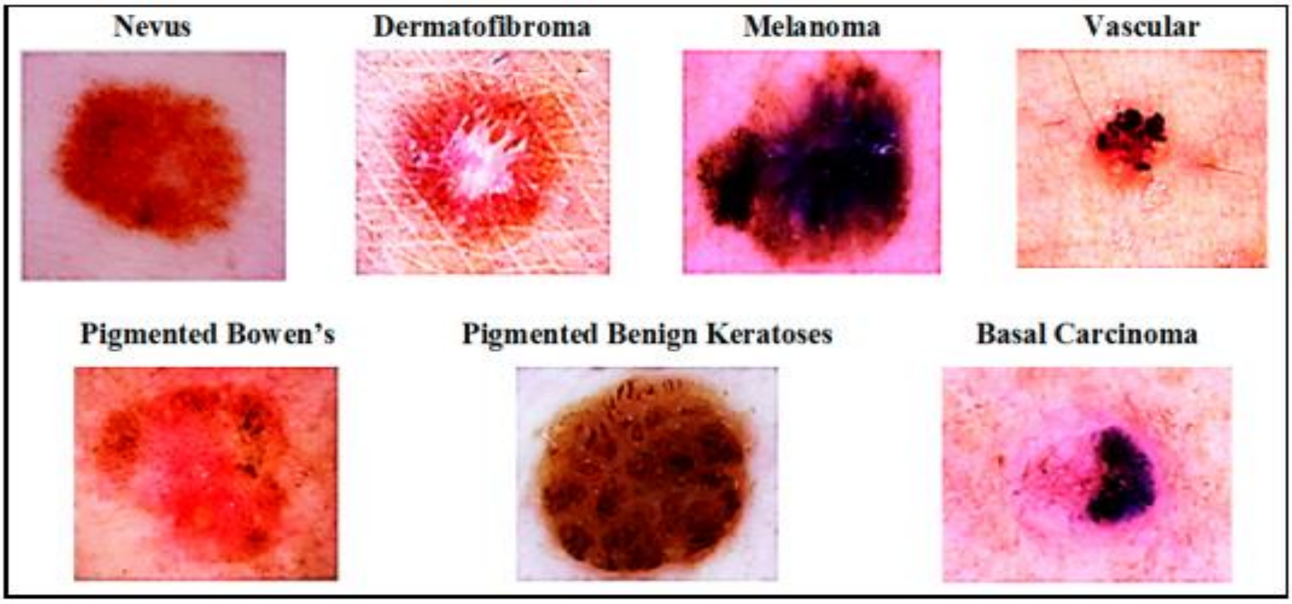

Figure 1. *ISIC 2019 skin lesion examples. Source: (ISIC, 2019).*

The hazard identification of skin lesions becomes easier for medical practitioners with these CAD models at an earlier stage without time consumption. This is the eventual aspect of proposing a model for early diagnosis of skin lesions and saving human lives and increasing the survival rate, see Figure 1.

The key steps involved in the CAD techniques are image acquisition from the dataset, feature selection, relevant features extraction, and classification. Deep feature extraction techniques have played a vital role in the last few decades compared to other traditional models. The convolutional neural network (CNN) is a deep learning technique that acts as the best feature extractor, done automatically without manual interference. It represents abstracted higher to lower features with its different layers (Hussain et al., 2018) without reducing the dimensionality of the image. After the work of (Krizhevsky et al., 2017), the CNN method has been very popular in the research community (Afza et al., 2022). Using the fully connected layer of the CNN, the most relevant and deep features are extracted that are used in the classification more effectively.

Compared to the traditional feature extraction models, the CNN model includes both the local and global information of the image using various layers such as the pooling layer, convolutional layer and fully connected layers that make the feature selection more effective and greatly enhance the accuracy of the image classification. In contrast, the traditional models include only the colour, texture, and shape and require manual intervention.

Although various techniques are used to make classification and feature extraction more effective, researchers still face problems related to redundant features that lead to image misclassification. Hence, it is essential to focus on the feature selection techniques that bring out more appropriate and relevant features for higher classification accuracy. Recently known popular feature selection techniques are ant colony optimization (ACO), firefly algorithm (FFA), butterfly optimization algorithm (BOA), and DSSA, which are also used for regression optimization and classification accuracy enhancement. When compared to binary classification, multiclass classification is more complex due to its intraclass similarity among lesions.

This paper proposes using a new framework gleaned from the hybridization of the dynamic salp swarm algorithm (DSSA) for feature selection and optimizing the weighted extreme learning machine WELM. The WELM is one of the variants of the ELM used to overcome the imbalances in dermoscopic images in the International Skin Imaging Collaboration (ISIC) 2019 dataset. The salp swarm algorithm (SSA) method implemented helps extract relevant features from the dermoscopic images and improves the accuracy of the classification of WELM using the optimization process, greatly reducing the computational complexity.

When the data available for training are large, the CNN would consume more time to learn. To reduce this time constraint, the CNN was used for feature extraction and the ELM was used as a multiclass classifier along with transfer learning pre-trained GoogleNet models for faster and effective classification of multiclass skin cancer images. The ELM results in simpler implementation, faster learning, and better generalization performance (Wang et al., 2022).

Our proposed work mainly focuses on higher accuracy in the classification of multiple types of skin lesions available in the dataset with a hybrid model using the DSSA-WELM multiclass classifier, which has the capability of fast learning, addresses imbalances in the dataset and elevates accuracy of classification ISIC 2019 dermoscopic images.

2 Related Work

Discrete research techniques are implemented to overcome the most challenging classification issues due to their similarity among various skin lesion types with many machine learning algorithms and deep learning methods for detecting and classifying skin lesions.

A model erythemato squamous disease (ESD) skin lesion was proposed by (Deepa et al., 2014) and analysed by the generation of action rules and fuzzy rules using-the action rule based diagnostic system (ARDS) and the adaptive neuro fuzzy inference system (ANFIS) classifiers based on severity levels. The knowledge management features enhance the correct disease identification. Furthermore, extensive research into automatic generation of accurate action rules with less human interference and computational intelligence techniques (Sriwong et al., 2019).

A model based on ELM to better identify the erythemato-squamous skin disease was proposed (Olatunji et al., 2014), which outperformed artificial neural networks (ANN). The ELM proves its unique ability to achieve excellent results in the field of biomedical diagnosis. Furthermore, research has expanded towards partitioning the data ratio for training and testing to increase the classification accuracy. since the performance increases as the percentage of the training set decreases.

A model that utilizes transfer learning and a pre-trained model with GoogleNet was proposed by (Kassem et al., 2020) for the accurate classification of skin lesions into eight different types and the accuracy was 94.92%. Another model ensemble with Bootstrap multiclass SVM was proposed by (Huang et al., 2006) that classifies even unknown images that are not found among the eight types of skin lesions.

The accuracy of this model is 81%. Although the model yielded promising results, it is time-consuming due to the repetitive fine-tuning of parameters.

A CNN algorithm proposed by (Sasikala et al., 2020) is implemented in four transfer learning techniques, namely AlexNet, ResNet 34, ResNet 50, and VGG 16, for image classification as malignant or benign. The CNN model was trained using 3700 clinical images and tested over 600 images. Among the four techniques, GoogleNet provides reliable results and a better approach to the early detection of skin cancer and treatment. In addition, the research can be extended with a multiclass classifier that classifies into more classes by increasing the convolutional neural network layers.

An ensemble hybrid CNN-ELM model for image classification was proposed by (Kannojia & Jaiswal, 2018). In this model, features are extracted automatically by the CNN more effectively and fed to the ELM for accurate classification. The CNN-ELM ensemble in parallel and the final output is computed on the majority-voting ensemble of the classifier's output. The model benchmarked on the Modified National Institute of Standards and Technology (MNIST) dataset with high accuracy and compared with core CNN, and core ELM. Hybrid CNN-ELM achieves higher accuracy, as proved by (Aditi et al., 2020). However, imbalances in the dataset make the proposed model less effective compared to other techniques.

Machine-learning techniques (Xu et al., 2020) are used to detect skin cancer at an early stage. The work proposed used an optimization method that was implemented using segmentation with CNN and satin bowerbird optimization (SBO). The features were extracted after the segmentation process and with the help of the SBO algorithm, the relevant features were extracted. Then the SVM is used for the two-fold classification of healthy and malignant skin lesions. Although the classification is done for unknown images, also the time complexity increases because of the usage of the CNN and SVM.

The ELM is best suited for multiclass classification and regression, when compared to traditional least square support vector machine (LS-SVM) and proximal support vector machine (PSVM), as proved by (Huang et al., 2011). It is widely used in binary classification applications but when the SVM is used for regression and multiclass classification, it would be a time-consuming process with higher computational complexity. Moreover, the SVM cannot be directly used as a multiclass classifier as in the case of the ELM.

The techniques stated above follow a standard procedure for skin lesion classification, namely pre-processing of dermoscopic images, segmenting of lesions, and extracting of relevant features for best classification using certain supervised learning algorithms. Note that the accuracy of features extracted using a pre-trained deep learning (DL) network is higher compared to feature extraction techniques of the traditional model. The crucial constraints of the methods used above are necessity of human intervention, redundant features that reduce the accuracy of the classifier, and a time-consuming process. Even though the ELM addresses the imbalance issues in the dataset, the imbalanced dataset remains a hurdle in the research field.

The proposed model overcomes the aforesaid limitations using ML-based algorithms with a meta-heuristic algorithm as a hybrid technique that automatically identifies skin lesions using dermoscopic images of either malignant or benign lesions. The ELM in general acts as the best multiclass classifier in image classification. There are several variants of the ELM, among which the WELM is considered for addressing the issue of imbalance in the dataset. Moreover, it applies a CNN for automatic feature extraction and uses transfer learning with GoogleNet pre-trained model to classify the ISIC 2019 dataset into eight types of skin lesions. The redundancy in the features is eliminated using the metaheuristic algorithm DSSA. Therefore, the ELM represents a simplified and unified approach, which has faster learning speed, can approximate target continuous function and classify any disjoint regions (Huang et al., 2006). The ELM is a thousand times faster than the SVM and LS-SVM.

The contribution of the proposed model is as follows:

- The model uses the GoogleNet pre-trained convolutional network model for feature extraction.
- The hybrid DSSA-WELM model optimizes the feature selection and retrieves only relevant features.
- The DSSA-WELM optimizes and enhances the accuracy of the classification of skin lesions effectively.
- The model is compared with other SOTA techniques, and it is proved that the proposed model outperforms all other models.

The paper is organized as follows: methods and materials are presented in Section 3, the proposed model in Section 4, results and discussions in Section 5, and conclusions in Section 6.

3 Methods and Materials

3.1 Convolutional neural network (CNN)

In the deep learning model, one of the most famous models is the convolutional neural network (CNN). The CNN aims to learn patterns and characteristics and generate features of images to recognize and classify. It can automatically learn a large number of filters in parallel to train datasets within specific predictive models termed image classification. The CNN involves three different layers in the network,

- **Convolutional layer:** the basic building block of the convolutional neural network, comprising many nodes used to extract important features from images. It applies filters to input and creates a feature map that summarizes the detected features of the input.
- **Pooling layer:** the one that appears after the convolutional layer. It focuses mainly on the spatial dimension reduction of the input, which in turn is sent to the next convolutional layer.
- **Fully connected layer:** occurs usually before the output layer.

Using the CNN architecture, the classification is done by creating a new deep network and training from scratch. Since the training is from scratch, it requires the size of the image and the exact number of layers used for the corresponding process. It requires millions of images for training, which is a time-consuming process. Thus, to overcome these issues, a better learning process was implemented, termed transfer learning, which requires pre-trained CNN models to classify a new set of images. Currently, there are numerous public pre-trained CNN models available for specific tasks. A few examples of these models are AlexNet, GoogleNet, VGGNet, and ResNet (Hosny et al., 2020).

These pre-trained CNN models require significantly less time than training from scratch and increase the accuracy of the model and size of the image in the classification process. Still, to speed up the process, feature extraction can be implemented which reduces large dimensions of the information to a smaller same as the original data.

Among various pre-trained CNN models, GoogleNet (Inception) is a deep convolutional neural network that approximates a sparse CNN by convolution-dense construction. A smaller number of neurons is sufficient to perform the task. GoogleNet uses various sizes of convolutions to extract data and features with different scales (5x5, 3x3, and 1x1). Moreover, this process helps greatly reduce the computation required. In GoogleNet, the fully connected layers are replaced with a simple pooling global average layer, which also helps reduce the total number of parameters significantly.

3.2 Transfer learning

In recent years, transfer learning (TL) has gained attention from researchers and academicians, who have successively implemented the TL process in various domains. Transfer learning builds accurate models in a time-saving way (Wang et al., 2022). The main process involved in transfer learning is that it deals with model pre-trained for a similar task, adds one or more layers at the end, and later re-trains the model with a large amount of labelled data more effectively with higher accuracy in a short time. In our proposed model, pre-trained GoogleNet was replaced with a few layers and combined with the ELM for higher classification accuracy.

Pre-trained model usage

The pre-trained models are used in most deep learning approaches. To implement the same, some research work is required to obtain the pre-trained model from the available research institutions that have released their training models. We must choose the pre-trained model based on a task similar to our predictive problem and make it easier and more effective using transfer learning. For example, Keras (2023) provides nine pre-trained models for transfer learning in the case of prediction, feature extraction and fine-tuning.

Feature extraction

Feature extraction is considered the best representation of a problem that is extracting the most important features of a problem, which simplifies the task and makes it easier. Also termed representation learning, it provides better performance when implemented. Many researchers and domain experts have already found that the features are extracted manually in machine learning but automatically with higher efficiency in deep learning.

It is more important to choose an algorithm effectively that discovers a good combination of features in a short period even for a complex task. Otherwise, it would require a lot of human effort, which would become a time-consuming process. Pre-trained models include Inception-v3, trained for the ImageNet "Large Visual Recognition Challenge". It uses a tensor flow framework to retrain the classifiers. Others are ResNet and Alex Net.

3.3 Extreme learning machine

The ELM is a popular machine-learning model that focuses on classification and prediction areas. The ELM is considered the best alternative for feedforward neural networks (FNN), which fix the issues in gradient-based machine learning algorithms. The ELM does not fine-tune all inner parameters as the in case of gradient-based algorithms (Kannojia & Jaiswal, 2018). The motivation to focus on the ELM is that it requires substantially lesser training time than the backpropagation (BP) algorithm and the (SVM) (Guo et al., 2015). Although the ELM possesses more hidden neurons, it is better than the BP algorithm and the SVM, as it is sufficient that it tunes only a few parameters. Since the feedforward network functions are slower than expected in various applications when implemented, the ELM is generally characterized as a single hidden layer feedforward neural network (SLFN). In the proposed model, the hidden nodes are selected at random, and their output weights are determined analytically. The ELM uses a wide range of feature mapping that includes kernels and hidden nodes. In the proposed model, the ELM, not only provides a unified solution but also has the capability of classifying disjoint regions. The output function of the ELM for the generalized SLFN is denoted as follows (Huang et al., 2006):

$$FL(X) = \beta i \; hi = h(x) \; \beta \tag{1}$$

Where $\beta = [\beta 1 \ldots\ldots\ldots\ldots \beta L]$ is the vector output weights between the hidden layer node and the output node, $h(x)$ indicates the output vector and maps the d-dimensional input space to L dimensional feature space; $h(x)$ indicates feature mapping. In general, the ELM not only reaches the lowest training error but also the smallest norm of output weights (Wang et al., 2022). According to Bartlett's theory, this is the condition possessed by the generally best-performing network.

The ELM is ultimately concerned with solving a linear system in the case of classification and regression. Moreover, it mainly focuses on achieving a global optimum.

The ELM was introduced by (Huang et al., 2006). In general, the ELM is used to overcome the over-fitting problem, training time and a dataset imbalance. The ELM is also called a neural network (NN) assigned with random weights (RW). The ELM as NNRW is, preferable to traditional neural networks such as the ANN. The ANN mechanism, based mainly on the backpropagation method, results in slow convergence, time-consuming process, and local minima problems. In contrast, the ELM addresses all these drawbacks of the ANN.

The ELM implements a training method such that the weights are randomly initialized between the input and hidden layers and biases. This makes the ELM faster than all other traditional mechanisms. Many researchers had focused on the random parameter initialization that is relevant to the optimal performance of the ELM classifier. Paper (Cao et al., 2019) specified the relationship between the random parameters and the best performance of the ELM.

3.3.1 Weighted ELM (WELM)

The WELM is a variant of the ELM, that mainly focuses on the imbalance in data for binary/multiclass classification (Guo et al., 2015). The WELM is capable of balanced data generalization (Zong et al., 2013). The uneven distribution of data between the classes is balanced by supplementing additional weights to the matrix W and downgrading the impact of the majority class and upgrading the impact of the minority class.

The ELM is rewritten as:

Minimize:

$$L_{ELM} = \frac{1}{2} \| \beta \| 2 + \frac{1}{2} CW \sum_{i=1}^{n} \xi i^2 \tag{2}$$

Subject to:

$$H(Xi)\, \beta = tiT - \xi iT \text{ , } i= 1\ldots.N \tag{3}$$

Where $h(xi)$ denotes the feature mapping vector for the hidden layer with respect to xi and β, which indicates the output weight vector connecting the hidden layer and the output layer. One of the important advantages of the ELM is that as a single classifier, it can be used for multiclass classification tasks (Zong et al., 2013). One interesting point to be noted here is that the popular SVM machine learning algorithm is basically a binary classifier, and it fails to act as a multiclass classifier directly without any modification. Usually a multiclass problem is broken into binary subproblems and implemented with multiple SVM's, which makes the process highly computationally complex.

The general multiclass optimization problem for a weighted ELM is given as:

Minimize:

$$L_{WELM} = \frac{1}{2} \|\beta\| 2 + CW \frac{1}{2} \sum_{i=0}^{N} \xi_i^2 - \sum_{i=0}^{N} \alpha_i \left(h(x_i)\beta - t_i + \xi_i\right), \tag{4}$$

However, the learning method of many machine learning algorithms such as neural networks, or the SVM is comparatively easier than the learning method of the ELM. Thus, to enhance the efficiency and improve the learning rate of the ELM, we use metaheuristic techniques such as the butterfly optimization algorithm, firefly optimization algorithm, mayfly algorithm, SSA and DSSA. Among these, the DSSA is used to optimize feature selection and optimize parameters that enhance the accuracy of the WELM multiclass classifier. The proposed DSSA-WELM hybrid model outperforms the others and provides determinately generalized performance with a simplified structure.

3.4 Dynamic salp swarm optimization algorithm

Classification is a key role in many real-time applications based on machine learning classifiers. There are situations where these classifiers are influenced by the features of the images that possess irrelevant or noisy image segments. Therefore, there is a need for feature selection (FS), which mainly focuses on retrieving informative features, eliminating irrelevant or noisy features, reducing the learning time of the classifier, and increasing the classification accuracy (Huang et al., 2006). Specifically, two methods are involved in the FS, the filter method, and the wrapper method. In the case of the filter method, the feature selection is independent of the classifier whereas in the wrapper method, an optimization algorithm is used to select the features and it is dependent on the classifier. Hence, the wrapper method and these optimization algorithms are used in our proposed work as well as in many real-time applications to solve biomedical and engineering problems.

There are several, meta-heuristic optimization algorithms, among which the SSA was introduced by (Mirjalili et al., 2017). It is used because of its simplicity, efficiency, and flexibility. The SSA is easy to implement and consists of only a few parameters compared to other algorithms. Although it is very simple and easy to implement, there are a few limitations such as population diversity and local optima stagnation (Zong et al., 2013). Thus, an improved version of the algorithm called the DSSA was introduced by (Tubishat et al., 2021), which addresses the above issues of the SSA. The hybridization of the DSSA with the WELM is the best model which is used for FS and optimizes the accuracy of the weighted ELM multiclass classifier.

4 Proposed Algorithm

The proposed work comprises five different processes. Initially, the ISIC 2019 dataset is retrieved and pre-processed to remove irrelevant images and produce appropriate data as input for the CNN network. The CNN now extracts the features from the dermoscopic input images as the next step. The extracted features are then fed to the FS stage using the DSSA to retrieve the most relevant features for classification. In the final stage, these retrieved relevant features are sent to the WELM multiclass classifier, where the classification is performed.

4.1 Pre-processing

The dataset must be pre-processed before being fed into the CNN pre-trained models. As the initial requirement, the dataset is resized to an acceptable size by the pre-trained model, i.e., 224*224 pixels, and then later transformed to RGB format with a 24-bit depth.

4.2 Transfer learning in CNN feature extraction

ResNet 18, ResNet 20, and GoogleNet are used for feature extraction in the proposed model. The fully connected layer is completely responsible for extracting the features with 1000 features of the pre-trained models. The total number of features extracted from the pre-trained models is n*1000, where n denotes the total number of dermoscopic images.

The CNN has the ability to automatically acquire knowledge even with a huge amount of filters in parallel based on the trained dataset under the constraint of the specific modelling problem. In general, the images are static in nature, which indicates that the features learned from one image can be applied to any other part of an image. Moreover, the CNN is widely used in large applications nowadays due to its computational efficiency.

Simply taking a model that has been trained on a large dataset and applying it to a smaller dataset is the fundamental tenet of transfer learning. In order to use a CNN for object identification, we freeze the network's early convolutional layers and only train the most predictive layers. The concept is that the convolutional layers extract common, low-level properties that apply to all images such edges, patterns, and gradients, while the later layers detect particular aspects inside an image, such as eyes or wheels. So the use of transfer learning along with the CNN enhances our model and efficiency.

4.3 Dynamic salp swarm optimization algorithm

As discussed above, multiclass classification is more complex than binary classification, which deals only with values 0 and 1. Thus to optimize a multiclass classifier the optimization technique chosen is also multi-objective. The significant goal to be achieved by this optimization technique is that it must provide maximum recognition accuracy with a few most relevant features. In our proposed model (see Figure 2), with a limited number of relevant features, the DSSA optimization technique has proven to be the best for feature selection. Moreover, the computational complexity is reduced, and the skin lesions are identified much faster with this hybrid model (DSSA-WELM).

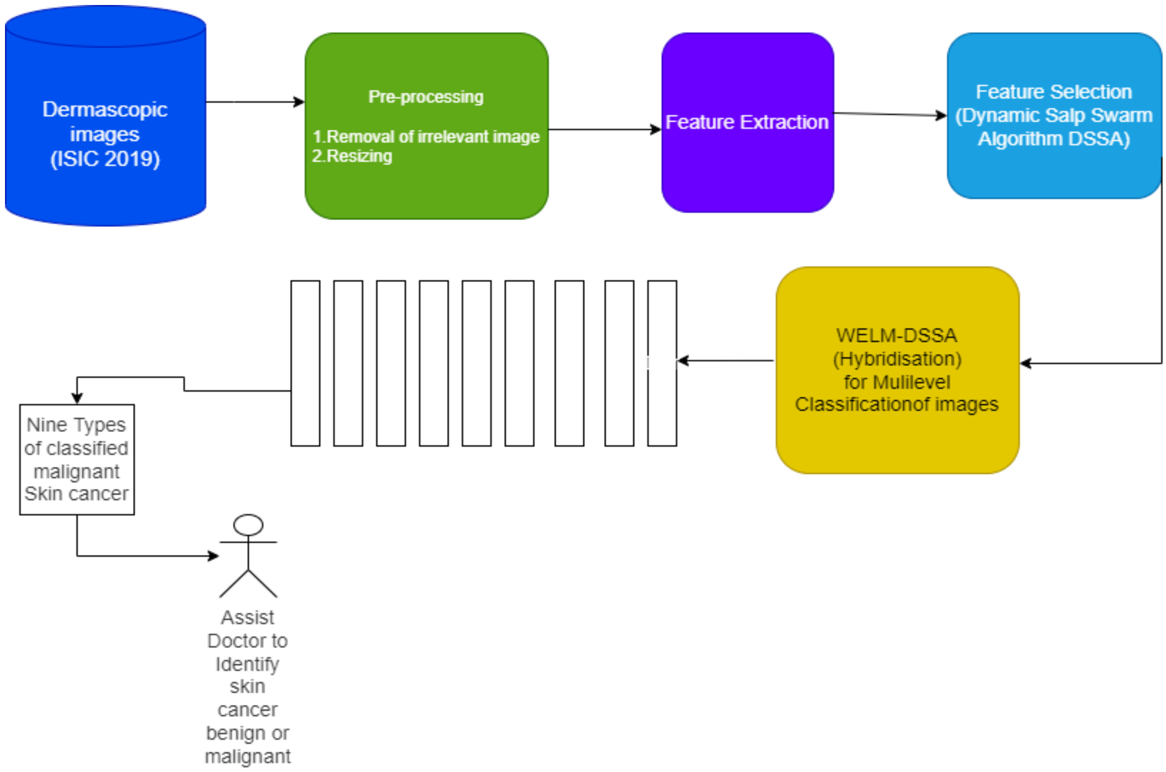

Figure 2. *Proposed DSSA-WELM model architecture.*

The DSSA uses the local search algorithm (LSA) to find the best solution. The LSA fixes the value as 2 or 5 features from the best solution randomly. The main purpose of using this LSA algorithm in the DSSA is to comprise the dataset with different dimensionality from low to high and make the classification accuracy of the DSSA very high (Tubishat et al., 2021).

4.4 Chaotic map

To generate an unbiased random number or random weights many researchers to date have implemented chaotic maps (Tschandl et al., 2018). The performance of metaheuristic algorithms is improved with different mathematical approaches. One of the approaches is the chaos theory, a nonlinear system with dynamic behaviour. Chaotic maps can improve the diversity of search algorithms due to their ergodic property (Lu et al., 2014). They are bounded by a nonlinear system that has random properties.

Chaos in general is combined with many optimization algorithms to improve their classification accuracy. In practice, there are many types of chaotic maps, among which the DSSA uses the singer chaotic map, which is used to update the leader position.

The equation is given as:

$$X_{k+1} = \mu\,(7.86x_k - 23.31x^2_k + 28.75x^3_k - 13.302875x^4_k) \tag{5}$$

Where μ is the control parameter with the value between [0.9, 1.08] and $x0 \in (0,1)$ and $xk \in (0,1)$.

Using the SSA, the following equation is used in the DSSA to update the leader position of the salp.

$$X^1{}_j = \begin{cases} Fj + c1\big((ubj - lbj)c2 + lbj\big)c3 \geq 0.5 \\ Fj - c1\big((ubj - lbj)c2 + lbj\big)c3 < 0.5 \end{cases} \tag{6}$$

In order to update the leader position, Equation (7) is used

$$X^i_j = \frac{1}{3}\,(x1 + x2 + x*) \tag{7}$$

The solution diversity is enhanced with these improvements to establish a balance between exploitation and exploration. Moreover, the LSA is used to determine the best solution, where 2 or 5 features are selected randomly to cover different dataset dimensionality. The DSSA initiates by computing a random number of the population (salp). The solutions of each salp are represented using the binary values either 0 or 1. The DSSA would calculate the food fitness of every solution based on the fitness function. Accordingly, the solution with minimum fitness is set as the best solution (food source).

Currently, the WELM classifier along with the DSSA acts as a fitness objective function to determine the best solution (food source/weights). After repeated iterations, the DSSA would update the salp position using Equations (6) and (7). Thus, the time complexity of the DSSA is generally given as *O (t(d * n + cof * n)* (Tschandl et al., 2018), where *t* represents the number of iterations, *d* is dimensions or a number of variables, and *n* is the population size.

Pseudocode 1

Local search algorithm (LSA)

```
t1 = f
S = 1
While (S< max_no_of_iterations)
      relevant_features_chosen = rand of 3 or 6
      LSA choose randomly 3 or 6 features from t1 based on relevant_features_chosen value
      for each feature chosen in t1
            if f_v = = 1 (1 denotes chosen feature and 0 not chosen feature)
                  f_v = 0
            else
                  f_v =1
            end if
      compute the fitness function of t1
if s (t1 ) < s (f)
       f= t1
end if
S=S+1
End while
Return f
```

Pseudocode 2

DSSA

```
Initiate the initial Salps Si (i=1,2…n) assuming the upper bound Ub and Lower bound Lb
Compute each fitness solution
Fa   best fitness value assumed on salp position
While (t lower than maximum iterations)
       Update C
For every salp Si and if i==1 update leader salp position
Else
SC_value < 0.5 update Follower salp Position
Else
Update Leader salp position
       End if
End For
Rearrange the salps that go beyond the search limit
Determine every salp fitness value in the chain
Identify the value of the IC variable
If IC >=2
       Call lsa of F to upgrade the value
End if
T=t+1
End while
Return best WELM network with minimum Mean square error (MSEtd)
Return F
```

Pseudocode 3

Hybrid optimization algorithm using dynamic salp swarm algorithm (DSSA) with weighted extreme learning machine (WELM)

```
Input:       Input Features of imbalance Training dataset and Target Class
             Total Number of Hidden weights Wh
             Activation function-Sigmoid
             Sa Search Agents
             Ti Total number of Iterations
             SC_value taken as Singer_Chaotic_map
             IC improvement counter variable
             F best fitness value
Output: Fb-Best-hidden bias Bi and weights Wi
Fitness: MSEtd as the fitness function for training data
Initialization:
       Step 1: Diagonal matrix W is computed, and weights are generated for each instance of a
       class in weighted ELM (WELM). Every matrix element is considered the diagonal vector of
       a particular instance xi.
       Step 2: The activation function opted for WELM is the sigmoid function.
       Step 3: Randomly compute the input weights w and bias b of the hidden neuron.
```

Step 4: up to the maximum number of iterations reached repeat steps 5 and 6.

Step 5: Determine the population fitness value and enumerate the salp population.

Step 6: For every salp Si if i==1, then update the leader salp position else SC_value<0.5 then update follow position as shown in Figure 3.

Step 7: Identify the incremental variable Ic and if Ic> = 2 calls upon the LSA algorithm to determine the global best solution. Otherwise, go to step 5.

Step 8: Based on the global best solution fitness value, the relevant salp positions such as the input weights and bias of the training data are obtained.

Step 9: Set the optimum weights and bias to the WELM for high classification accuracy.

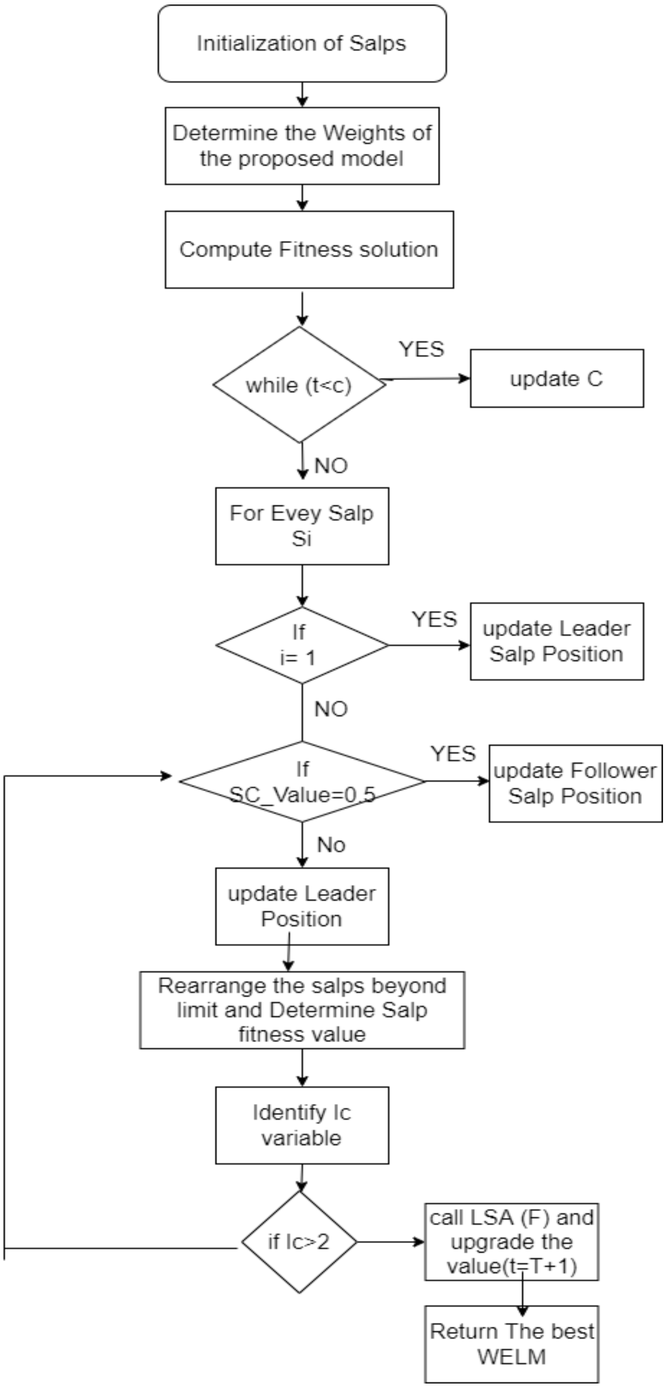

Figure 3. *Proposed model workflow.*

5 Results and Discussion

In this proposed ensemble model, the DSSA-WELM is implemented using Python with the Keras deep learning framework with a Theano backend (O'Reilly, 2023), and the ELM is also implemented with Python. The experiment was performed on a cluster with 16 CPU cores and 64 GB RAM. For assembling purposes, the CNN was initially trained for feature extraction that contains a 3x3 convolutional layer and three max-pooling layers with a pooling size (2, 2) that extract low-level, mid-level and high-level features accurately.

5.1 Dataset

The well-known dataset from (ISIC, 2019) was used to test and evaluate the model. This dataset contains images of Human Against Machine HAM10000 and BCN_20000. Here, the GoogleNet model is used for more effective automatic skin lesion detection with transfer learning. The dataset contains the training data for the ISIC 2019 challenge, note that it includes data from previous years. The dataset for ISIC 2019 contains 25,331 images available for classification of dermoscopic images among 8 different diagnostic categories. After training and validating the pre-trained GoogleNet model by applying transfer learning. 80% of the dataset consisting of 20,265 images and 10% of 2,531 images respectively were used for training and the remaining 10% of the data (2,531) images were used for testing.

5.2 WELM-DSSA classification performed

The dataset is taken as 80% for training, 10% for validation and 10% for testing. The ISIC 2019 dataset consists of 8 types of skin lesions namely melanoma (MEL), melanocytic nevus (NV), basal cell carcinoma (BCC), actinic keratosis (AKIEC), benign keratosis (BKL), dermatofibroma (DF), vascular lesion (VASC) and squamous cell carcinoma (SCC). The dataset comprises 25,331 images, consisting of 4,522 MEL images, 12,875 NV images, 3,323 BCC images, 867 AKIEC images, 2,624 BKL images, 239 DF images, 253 VASC images, and 628 SCC images.

Since the images of different skin lesions in the ISIC 2019 dataset are imbalanced classes, the classification is performed more accurately with the hybridization process. Initially, the images are pre-processed to remove unwanted images and remove hair or other irrelevant scars in the images, and later on, using the DSSA as the feature selection algorithm, the relevant features are extracted. The ELM is an ensemble with the pre-trained CNN model. Using transfer learning, the features extraction from the images is performed more effectively, and later image classification of each type of skin lesion is done using the hybridization technique DSSA with the WELM for higher accuracy. The WELM acts as a multiclass classifier that resolves imbalances in the dataset with a replacement of the last layers in the pre-trained model. The ELM classifies even unknown images in the group, which are considered normal without any lesions.

Table 1. *Parameter setting of the proposed model.*

Algorithm	Parameter	Value
DSSA-WELM algorithm *(proposed algorithm)*	Search agent number	20
	Total number of iterations	100
	Input node	20
	Hidden node	20
	Activation function	Sigmoid
	Output node	1
	Learning rate	0.01

5.3 Confusion matrix

The confusion matrix implemented in the proposed model to calculate the performance measure is given in Table 2 below, where the true values are known for a set of test data. For example, in our proposed model, suppose it has to be correctly predicted as vascular lesion and the actual image is also a vascular lesion image. In that case, the possibilities are computed. The confusion matrix helps us identify the performance of the classifier, with a higher ratio of correct classification of images indicating the accuracy of the classifier.

True positive

The true positive value indicates that the actual value and the predicted value are the same.

False positive

False positive indicates that the actual value is positive but the predicted value from the model is negative, i.e., wrong prediction.

False negative

False negative means that the actual value is false, i.e., for example in our proposed model it is actinic keratosis (wrong or negative actual value) but the predicted value is a vascular lesion, which is the right prediction.

True negative

True negative means that the predicted and the actual value are the same, i.e., actinic keratosis as the actual and predicted value.

Table 2. *Confusion matrix prediction.*

Actual		Predicted values
True positive (TP)	False positive (FP)	
False negative (FN)	True negative (TN)	

The performance measures for the model are evaluated using five quantitative measures: accuracy, sensitivity, specificity, precision and F1 score. The computation is done as follows:

$$\text{Accuracy} = \frac{TP + TN}{TP + FP + FN + TN}$$

$$\text{Sensitivity (TPR)} = \frac{TP}{TP + FN}$$

$$\text{Specificity (TNR)} = \frac{TN}{FP + TN}$$

$$\text{Precision} = \frac{TP}{TP + FP}$$

$$F1\ score = 2 * \frac{Precision * Sensitivity}{Precision + Sensitivity} * 100$$

Where TN, TP, FN and FP are true negative, true positive, false negative, and false positive, respectively. TPR, TNR and PPV refer to true positive rate, true negative rate and positive predictive value. Here, the true negative should be high whereas the false positive should be small.

Table 3. *9x9 matrix of classification process.*

	AKIEC	DF	MEL	NV	BCC	SCC	BKL	VASC	UK
AKIEC	305	04	73	17	45	16	18	40	1
DF	90	68	102	20	52	23	09	65	0
MEL	136	186	486	12	46	55	13	42	0
NV	41	52	86	45	68	14	08	23	0
BCC	47	56	502	13	71	305	56	42	4
SCC	23	68	271	56	42	108	62	98	0
BKL	46	246	142	48	53	175	271	46	0
VASC	474	312	33	16	94	56	35	118	1
UK	5	6	0	1	1	4	0	1	327

5.4 Comparison of proposed model with other SOTA

The pre-trained model after applying transfer learning classifies the image into eight types and an unknown image lesion that is normal is also obtained from the DSSA-WELM classification accurately. The WELM multiclass classifier hybridized with DSSA metaheuristic optimization technique has an accuracy of 97.25%, which is slightly better compared to the classification of the bootstrap multiclass SVM classifier and also higher than the other compared models. Thus, our proposed model works effectively with better accuracy and scalability, low computational complexity, and less time consumption, which are the important factors considered during the classification process.

Table 4. *Comparison of proposed model with other SOTA techniques.*

Model	Accuracy	Sensitivity	Specificity	Precision
Bootstrap SVM-GoogleNet	81%	74%	84%	77%
Deep CNN	95.98%	91.97%	97.33%	91.96%
AlextNet-ELM (FORFO)	97.14%	96.38%	96.64%	87.51%
GoogleNet-WELM (DSSA) *(proposed model)*	97.25%	96.40%	96.68%	87.49%

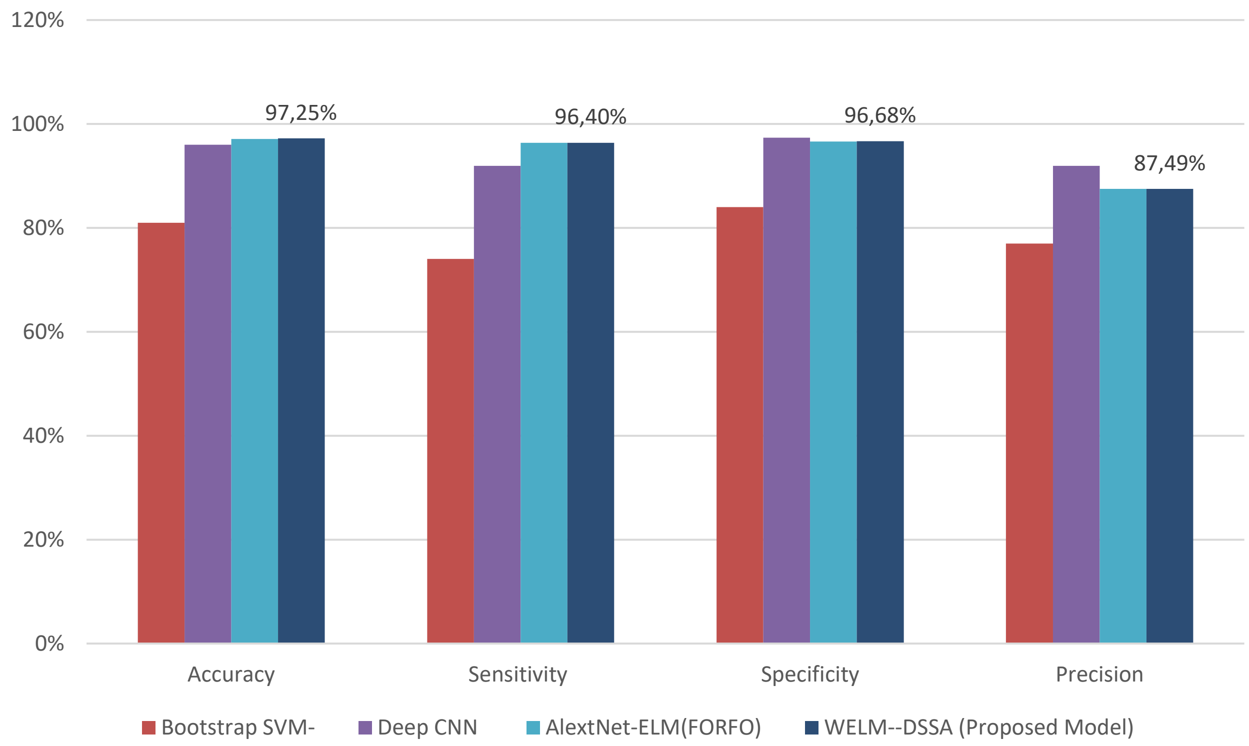

Figure 4. *Graphic representation of performance metrics of proposed model.*

6 Conclusion

As image classification in skin cancer has become a challenging task with imbalanced datasets and high similarities found in various types of skin lesions, we developed a model with the pre-trained CNN model GoogleNet using transfer learning for classification of images done with the most challenging ISIC 2019 dataset that comprises 8 types of skin lesions. Our proposed model classified the types of skin lesions with not only better accuracy but less computational complexity by hybridizing the model with the DSSA-WELM. The model classified unknown images that do not fall among the types, meaning that the patient is recognized as normal more effectively. Compared to the bootstrap multiclass SVM classifier and other models, our proposed model shows better accuracy and scalability. These techniques and models developed by researchers assist greatly in the biomedical field. Early disease diagnoses were done more effectively with our proposed model in a simpler way and saved human lives. In the future, we would like to implement early skin cancer detection with deep learning techniques such as the generative adversarial network (GAN).

Additional Information and Declarations

Conflict of Interests: The authors declare no conflict of interest.

Author Contributions: R.P.: Conceptualization, Methodology, Software, Supervision, Data curation, Writing – Original draft preparation. S.B.: Visualization, Investigation, Software, Validation, Writing – Reviewing and Editing.

Institutional Review Board Statement: Ethical review and approval were waived for this study due to the use of the external and publicly available dataset.

Data Availability: The data supporting this study's findings are openly available from (ISIC, 2019).

References

Afza, F., Sharif, M., Khan, M. S., Tariq, U., Yong, H., & Cha, J. (2022). Multiclass Skin Lesion Classification Using Hybrid Deep Features Selection and Extreme Learning Machine. *Sensors*, 22(3), 799. https://doi.org/10.3390/s22030799

ASCO. (2022). Melanoma: Statistics. American Society of Clinical Oncology. https://cancer.net/cancer-types/melanoma/statistics

Cao, W., Patwary, M. J. A., Yang, P., Wang, X., & Ming, Z. (2019). An Initial Study on the Relationship Between Meta Features of Dataset and the Initialization of NNRW. In *International Joint Conference on Neural Network*. IEEE. https://doi.org/10.1109/ijcnn.2019.8852219

Deepa, T., Sathiyabhama, B., Akilandeswari, J., & Gopalan, N. (2014). Action fuzzy rule based classifier for analysis of dermatology databases. *International Journal of Biomedical Engineering and Technology*, 15(4), 360–379. https://doi.org/10.1504/ijbet.2014.064828

De Tejada, M. G., Bilbao, A., Baré, M., Briones, E., Sarasqueta, C., Quintana, J. M., & Escobar, A. (2016). Association of social support, functional status, and psychological variables with changes in health-related quality of life outcomes in patients with colorectal cancer. *Psycho-Oncology*, 25(8), 891–897. https://doi.org/10.1002/pon.4022

Guo, L., & Ding, S. (2015). A hybrid deep learning CNN-ELM model and its application in handwritten numeral recognition. *Journal of Computational Information Systems*, 11(7), 2673–2680. https://doi.org/10.12733/jcis13987

Hosny, K. M., Kassem, M. a. E., & Fouad, M. M. M. (2020). Classification of Skin Lesions into Seven Classes Using Transfer Learning with AlexNet. *Journal of Digital Imaging*, 33(5), 1325–1334. https://doi.org/10.1007/s10278-020-00371-9

Hossin, M. A., Rupom, F. F., Mahi, H. R., Sarker, A., Ahsan, F., & Warech, S. (2020). Melanoma Skin Cancer Detection Using Deep Learning and Advanced Regularizer. In *International Conference on Advanced Computer Science and Information Systems*. IEEE. https://doi.org/10.1109/icacsis51025.2020.9263118

Huang, G., Zhu, Q., & Siew, C. K. (2006). Extreme learning machine: Theory and applications. *Neurocomputing*, 70(1–3), 489–501. https://doi.org/10.1016/j.neucom.2005.12.126

Huang, G., Zhou, H., Ding, X., & Zhang, R. (2011). Extreme Learning Machine for Regression and Multiclass Classification. *IEEE Transactions on Systems, Man, and Cybernetics*, 42(2), 513–529. https://doi.org/10.1109/tsmcb.2011.2168604

Hussain, M., Bird, J. T., & Faria, D. R. (2018). A Study on CNN Transfer Learning for Image Classification. In *Advances in Intelligent Systems and Computing*, (pp. 191–202). Springer. https://doi.org/10.1007/978-3-319-97982-3_16

ISIC. (2019). Skin Lesion Images for Melanoma Classification. https://www.kaggle.com/datasets/andrewmvd/isic-2019

Kannojia, S. P., & Jaiswal, G. (2018). Ensemble of Hybrid CNN-ELM Model for Image Classification. In *International Conference on Signal Processing* (pp. 538–541). IEEE. https://doi.org/10.1109/spin.2018.8474196

Kassem, M. a. E., Hosny, K. M., & Fouad, M. M. M. (2020). Skin Lesions Classification Into Eight Classes for ISIC 2019 Using Deep Convolutional Neural Network and Transfer Learning. *IEEE Access*, 8, 114822–114832. https://doi.org/10.1109/access.2020.3003890

Keras. (2023). Keras Applications. http://keras.io/api/applications

Krizhevsky, A., Sutskever, I., & Hinton, G. E. (2017). ImageNet classification with deep convolutional neural networks. *Communications of the ACM*, 60(6), 84–90. https://doi.org/10.1145/3065386

Lu, H., Wang, X., Fei, Z., & Qiu, M. (2014). The Effects of Using Chaotic Map on Improving the Performance of Multiobjective Evolutionary Algorithms. *Mathematical Problems in Engineering*, 2014, Article ID 924652. https://doi.org/10.1155/2014/924652

Mirjalili, S., Gandomi, A. H., Mirjalili, S., Saremi, S., Faris, H., & Mirjalili, S. (2017). Salp Swarm Algorithm: A bio-inspired optimizer for engineering design problems. *Advances in Engineering Software*, 114, 163–191. https://doi.org/10.1016/j.advengsoft.2017.07.002

Olatunji, S. O., & Arif, H. (2014). Identification of Erythemato-Squamous Skin Diseases Using Support Vector Machines and Extreme Learning Machines: A Comparative Study towards Effective Diagnosis. *Transactions on Machine Learning and Artificial Intelligence,* 2(6), Article no. 124. https://doi.org/10.14738/tmlai.26.812

O'Reilly. (2023). Using the Theano backend with Keras. https://www.oreilly.com/library/view/keras-deep-learning/9781788621755/c89ce815-17a2-4fde-b8b6-97c677c6494f.xhtml

Sasikala, S., Kumar, S. A., Shivappriya, S. N., & Priyadharshini, T. (2020). Towards Improving Skin Cancer Detection Using Transfer Learning. *Bioscience Biotechnology Research Communications*, 13(11), 55–60. https://doi.org/10.21786/bbrc/13.11/13

Sriwong, K., Bunrit, S., Kerdprasop, K., & Kerdprasop, N. (2019). Dermatological Classification Using Deep Learning of Skin Image and Patient Background Knowledge. *International Journal of Machine Learning and Computing*, 9(6), 862–867. https://doi.org/10.18178/ijmlc.2019.9.6.884

Teng, Y., Yu, Y., Li, S., Huang, Y., Xu, D., Tao, X., & Fan, Y. (2021). Ultraviolet Radiation and Basal Cell Carcinoma: An Environmental Perspective. *Frontiers in Public Health*, 9. https://doi.org/10.3389/fpubh.2021.666528

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). https://doi.org/10.1038/sdata.2018.161

Tubishat, M., Jaafar, S., Alswaitti, M., Mirjalili, S., Idris, N., Ismail, M. A., & Omar, M. S. (2021). Dynamic Salp swarm algorithm for feature selection. *Expert Systems With Applications*, 164, 113873. https://doi.org/10.1016/j.eswa.2020.113873

Wang, J., Lu, S., Wang, S., & Zhang, Y. (2022). A review on extreme learning machine. *Multimedia Tools and Applications*, 81(29), 41611–41660. https://doi.org/10.1007/s11042-021-11007-7

Xu, Z., Sheykhahmad, F. R., Ghadimi, N., & Razmjooy, N. (2020). Computer-aided diagnosis of skin cancer based on soft computing techniques. *Open Medicine*, 15(1), 860–871. https://doi.org/10.1515/med-2020-0131

Zhang, B., Zhou, X. S., Yichen, L., Zhang, H., Yang, H., Ma, J., & Ma, L. (2021). Opportunities and Challenges: Classification of Skin Disease Based on Deep Learning. *Chinese Journal of Mechanical Engineering*, 34(1). https://doi.org/10.1186/s10033-021-00629-5

Zong, W., Huang, G., & Chen, Y. (2013). Weighted extreme learning machine for imbalance learning. *Neurocomputing*, 101, 229–242. https://doi.org/10.1016/j.neucom.2012.08.010

Editorial record: The article has been peer-reviewed. First submission received on 29 November 2022. Revisions received on 5 January 2023 and 10 February 2023. Accepted for publication on 9 March 2023. The editors coordinating the peer-review of this manuscript were Mazin Abed Mohammed, Seifedine Kadry, and Oana Geman. The editor in charge of approving this manuscript for publication was Zdenek Smutny.

Special Issue: Deep Learning Blockchain-enabled Technology for Improved Healthcare Industrial Systems.

Acta Informatica Pragensia is published by Prague University of Economics and Business, Czech Republic.

ISSN: 1805-4951