Segmentation and Classification of Skin Cancer in Dermoscopy Images Using SAM-Based Deep Belief Networks

Syed Ziaur Rahman[1]![ID], Tejesh Reddy Singasani[2]![ID], Khaja Shareef Shaik[3*]![ID]

[1] Faculty of Information Technology, Majan University College, 710 Muscat, Oman
[2] School of Computer and Information Sciences, University of Cumberlands, Louisville, 40769 KY, USA
[3] Department of Information Technology, MLR Institute of Technology, 500043 Hyderabad, India

* Correspondence: Khaja Shareef Shaik (shareef.sk@mlrinstitutions.ac.in)

Abstract: In the field of computer-aided diagnostics, the segmentation and classification of biomedical images play a pivotal role. This study introduces a novel approach employing a Self-Augmented Multistage Deep Learning Network (SAMNetwork) and Deep Belief Networks (DBNs) optimized by Coot Optimization Algorithms (COAs) for the analysis of dermoscopy images. The unique challenges posed by dermoscopy images, including complex detection backgrounds and lesion characteristics, necessitate advanced techniques for accurate lesion recognition. Traditional methods have predominantly focused on utilizing larger, more complex models to increase detection accuracy, yet have often neglected the significant intraclass variability and inter-class similarity of lesion traits. This oversight has led to challenges in algorithmic application to larger models. The current research addresses these limitations by leveraging SAM, which, although not yielding immediate high-quality segmentation for medical image data, provides valuable masks, features, and stability scores for developing and training enhanced medical images. Subsequently, DBNs, aided by COAs to fine-tune their hyper-parameters, perform the classification task. The effectiveness of this methodology was assessed through comprehensive experimental comparisons and feature visualization analyses. The results demonstrated the superiority of the proposed approach over the current state-of-the-art deep learning-based methods across three datasets: ISBI 2017, ISBI 2018, and the PH2 dataset. In the experimental evaluations, the Multi-class Dilated D-Net (MD2N) model achieved a Matthew's Correlation Coefficient (MCC) of 0.86201, the Deep convolutional neural networks (DCNN) model 0.84111, the standalone DBN 0.91157, the autoencoder (AE) model 0.88662, and the DBN-COA model 0.93291, respectively. These findings highlight the enhanced performance and potential of integrating SAM with optimized DBNs in the detection and classification of skin cancer in dermoscopy images, marking a significant advancement in the field of medical image analysis.

Keywords: Skin cancer classification; Dermoscopy image lesions; Deep belief network; Coot optimization algorithm; Segmentation network

1. Introduction

Skin cancer, particularly its melanoma and non-melanoma forms, stands as a prevalent health concern compared to other cancer types. Melanoma, characterized by the uncontrolled proliferation of pigmented cells (melanocytes) (Zhang et al., 2020), has been observed to increasingly contribute to mortality rates annually (Razmjooy et al., 2020). This upward trend in melanoma incidences poses a significant public health risk, underlining the necessity of early detection measures. It is well-established that early-stage melanoma can be effectively treated, thereby rendering early detection methodologies (Daghrir et al., 2020) increasingly vital. In the realm of melanoma skin cancer imaging, distinguishing lesioned from non-lesioned areas presents a formidable challenge (Jakkulla et al., 2023), often necessitating specialized expertise. Consequently, divergent opinions among dermatologists are not uncommon. In response to this, an automated analysis method has been proposed (Dildar et al., 2021), aiming to facilitate rapid and accurate diagnostic decisions by dermatologists. Central to this automated analysis is the

requirement for precise segmentation of the skin surrounding melanomas (Ali et al., 2021).

Recent technological advancements have heralded the widespread adoption of deep neural networks in computer vision tasks. Convolutional neural network (CNN) architectures, known for their adept feature extraction and classification capabilities (Macherla et al., 2023), have gained prominence in pattern recognition and segmentation tasks (Xu et al., 2020). In line with this, a number of studies have focused on the segmentation and classification of skin lesions, employing pre-trained CNN models augmented with transfer learning techniques (Kumar et al., 2020). Advanced deep learning methodologies, including UNet, Mask R-CNN, fully convolutional networks, feature pyramid networks, SegNet, and transformers, have been utilized to identify critical lesion locations at the pixel level (Ali et al., 2022; Pacheco & Krohling, 2020).

Building on the foundation of these prior studies (Toğaçar et al., 2021; Thomas et al., 2021), which predominantly revolve around lesion segmentation and classification in dermoscopy images (Wei et al., 2020), the current research introduces a multi-task learning network aimed at improving the segmentation and classification of skin lesions as melanoma. This network employs pre-processing techniques during segmentation, initially enhancing image clarity and removing hair details. Subsequently, SAMNet is utilized for lesion segmentation, followed by the application of a very deep super-resolution neural network to extract and enhance the quality of the segmented lesions. These high-resolution images then serve as inputs for the classifier model. The classifier model itself is a novel amalgamation of three robust, pre-trained deep models. Both the ISIC database and PH2 have been utilized to rigorously test each proposed method. Experimental results validate the superior performance of the developed classification system.

The advantages of the proposed deep network method are manifold:
- ❖ A very deep super-resolution neural network enhances the resolution of lesion images, refining them post-segmentation for input into the classifier model.
- ❖ A SAMNetwork approach effectively localizes lesions in a diverse range of dermoscopy images, as evidenced by numerical and visual results from experimental tests.
- ❖ The classifier model's design, based on an integration of diverse deep models, has demonstrated remarkable efficacy in melanoma classification, as verified through experimental studies.

The study is organized as follows: Section 2 presents the literature review, Section 3 describes the proposed model, Section 4 details the experimental analysis, and Section 5 offers the conclusion.

2. Related Works

In addressing the challenges of skin cancer detection and therapy planning, Subhashini & Chandrasekar (2023) proposed a distinctive deep learning-based framework. This approach incorporates preprocessing, segmenting the cancer regions, and Fast Adaptive Moth-Flame optimization with Cauchy mutation (FA-MFC) for data dimensionality reduction. The bidirectional Gated Recurrent Unit (BGRU)-quantum neural network (QNN) facilitates multi-class classification, while hidden features are extracted via the structural similarity with light U-structure (USSL). Tested on the Kaggle and ISIC-2019 datasets, this model demonstrated remarkable accuracy, achieving up to 96.458% on Kaggle and 94.238% on ISIC-2019. Such results indicate the potential of this hybrid deep learning methodology in enhancing skin cancer classification and, consequently, improving diagnostic, treatment, and survival outcomes.

Anand et al. (2023) explored the integration of Fully convolutional Neural Network as U-Net models for skin disease analysis. Their approach utilized U-Net for segmenting skin diseases in scans and introduced a convolutional model for multi-class classification of segmented images. The model was tested on the HAM10000 dataset, encompassing images across seven skin disease categories. Employing Adam and Adadelta optimizers for 20 epochs and a batch size of 32, the model demonstrated superior performance, especially with an Adadelta optimizer, achieving an accuracy of 97.96%.

Kumar et al. (2023) introduced the MD2N architecture for segmenting and classifying various types of skin cancer in screenings. The MD2N employs a downsampling ratio in its encoder phase to maintain feature information, crucial for differentiating small skin lesion patches. By utilizing dilated manifold parallel convolution, the network's receptive field is expanded, addressing the "grid issue" common in standard dilated convolutions. This approach enables the model to capture detailed feature information from regions with variably sized skin lesions. The system, powered by the proposed deep learning model, efficiently segments and identifies lesion locations from skin images provided by the International Skin Imaging Collaboration.

Singh et al. (2023) advocated for a modified deep learning model coupled with fuzzy logic-based image segmentation for skin cancer diagnosis. This model incorporates dermoscopy image enhancement techniques, including standard deviation methods, L-R fuzzy defuzzification, and mathematical logic infusion, aimed at augmenting lesion visibility by eliminating distractive artifacts such as hair follicles and dermoscopic scales. Following segmentation, the quality of the image is enhanced through histogram equalization before proceeding to detection. The modified model employs the You Look Only Once (YOLO) deep neural network approach for diagnosing melanoma lesions from digital and dermoscopic images. Enhancements to the YOLO model include

an additional convolutional layer and residual connections within its DCNN layer sequence, alongside feature concatenation across layers. Training on 2,000 and 8,695 dermoscopy images from the ISIC 2017 and ISIC 2018 datasets, with the PH2 dataset for testing, revealed that YOLO surpasses other classifiers in accuracy and speed.

Sethanan et al. (2023) developed the skin cancer classification system (SC-CS), a highly accurate method for skin cancer classification. This system targets specific skin anomalies such as melanoma, benign keratosis, and other carcinomas and skin moles, utilizing a dual artificial multiple intelligence system (AMIS) ensemble model. The SC-CS combines image segmentation and CNN algorithms, optimizing the weighting of findings for high-quality outcomes. Evaluated using accuracy, precision, Area Under Curve (AUC), and F1-score metrics, along with feedback from 31 specialists, including dermatologists and physicians experienced in skin cancer diagnosis, the SC-CS demonstrated remarkable accuracy on the HAM10000 and malignant vs. benign datasets. With an accuracy exceeding 99.4%, it outperformed state-of-the-art models by 2.1% for larger sizes and 15.7% for smaller sizes. The System Usability Scale (SUS) of 96.85% indicates high user satisfaction and likelihood of recommendation. Data security measures were stringently applied to ensure patient confidentiality, with processed data and images being promptly destroyed post-upload.

Adla et al. (2023) developed a hyper-parameter-optimized full resolution convolutional network-based model, demonstrating significant accuracy in diagnosing skin cancer types in dermoscopy images. This model presents a potential advancement in computer-aided diagnostics, potentially leading to faster and more precise diagnoses. The incorporation of a dynamic graph cut technique in this model addresses the prevalent issues of over-segmentation and under-segmentation commonly seen in graph cut methods. Moreover, the model adeptly handles the challenge of incorrectly segmented small areas in the grab cut technique. The importance of data augmentation in training and testing was underscored, with the model showing enhanced results compared to using new images alone. In various trials against other transfer models, this proposed model excelled, achieving an accuracy rate of 97.986% for skin lesion classification.

Hong et al. (2023) introduced a weakly supervised semantic segmentation approach (CNN-SRR) for dermoscopy images, an innovative adaptation of the unsupervised superpixel technique and deep learning-based classifiers. This method leverages the abundance of labeled data at the image level for fine-tuning, focusing on lesion areas. The process involves training a repurposed CNN classifier and back-propagating the top layer's peak values, followed by over-segmenting a test image into superpixels, which are then aggregated into region proposals. Lesion region responses are then identified via non-maximal suppression, leading to the selection of a segmented mask. Quantitative tests on the ISBI2017 and PH2 datasets demonstrated that the CNN-SRR method not only effectively discriminates lesion locations but also achieves competitive accuracy comparable to supervised segmentation algorithms. Specifically, on the ISBI2017 dataset, the CNN-SRR method outperformed the unsupervised superpixel segmentation algorithm by 12.4% on the Jaccard coefficient and 3.3% on segmentation accuracy.

Alenezi et al. (2023) reported on a multi-task learning approach for dermoscopic images. The approach begins with an efficient pre-processing strategy, employing max pooling, contrast, and shape filters to remove hair details and enhance images. These processed images are then segmented using a Fully Convolutional Network (FCN) layer architecture based on a Visual Geometry Group Network (VGGNet) model. Following segmentation, lesions are cropped and subjected to a very deep super-resolution neural network, aiming to adjust the cropped images to the classifier model's input size with minimal loss of detail. A deep learning network strategy, utilizing pre-trained convolutional neural networks, was then deployed for melanoma classification. Experiments conducted using the International available dermoscopy skin lesion dataset demonstrated impressive results. For lesion region segmentation, the method achieved 96.99% accuracy, 92.53% specificity, 97.65% precision, and 98.41% sensitivity. In classification tasks, it yielded 97.73% accuracy, 99.83% specificity, 99.83% precision, and 95.67% sensitivity.

3. Proposed System

3.1 Dataset Description

The proposed model's efficacy was evaluated using three distinct dermoscopy skin lesion datasets (Thapar et al., 2022). The ISIC-2017 dataset comprises 2,000 dermoscopy images for training, 150 for validation, and 600 for testing. Each image in this dataset is accompanied by separate ground truths for segmentation, distinguishing between foreground and background, and classification into categories such as melanoma, nevus, and seborrheic keratosis. Figure 1 exemplifies images from these datasets.

The ISBI additional dataset includes 1,320 annotated dermoscopy images, curated from the largest known repository of skin lesions in the ISIC archive. This dataset provides a comprehensive collection for training and validation purposes.

The PH2 dataset contains a total of 200 dermoscopy images, comprising 160 nevi and 40 melanomas, with annotations for both segmentation and classification. The dataset is categorized into two distinct types: melanoma

and nevus.

ISBI 2017: The International Symposium on Biomedical Imaging (ISBI) 2017 focuses on biomedical imaging, including medical image analysis and processing. This annual conference features challenges and competitions to foster the development of innovative algorithms and methodologies in biomedical imaging. The ISBI 2017 Challenge involved a specific biomedical imaging problem, encouraging researchers to develop and test algorithms using the provided dataset.

ISBI 2018: Similar to its predecessor, ISBI 2018 continued the tradition of hosting workshops, challenges, and sessions dedicated to advancing medical image analysis. The symposium serves as a platform for presenting the latest research and developments in the field.

PH2 dataset: The PH2 dataset, a cornerstone in dermatoscopic image analysis, is instrumental in skin cancer diagnosis research. Dermoscopy, the examination of the skin using a dermatoscope, is crucial for the early detection of skin cancer. The PH2 dataset includes high-quality dermoscopic images of various skin lesions, complete with clinical descriptions and ground truth annotations. This dataset is extensively used by researchers to develop and evaluate algorithms for image segmentation, feature extraction, and lesion classification.
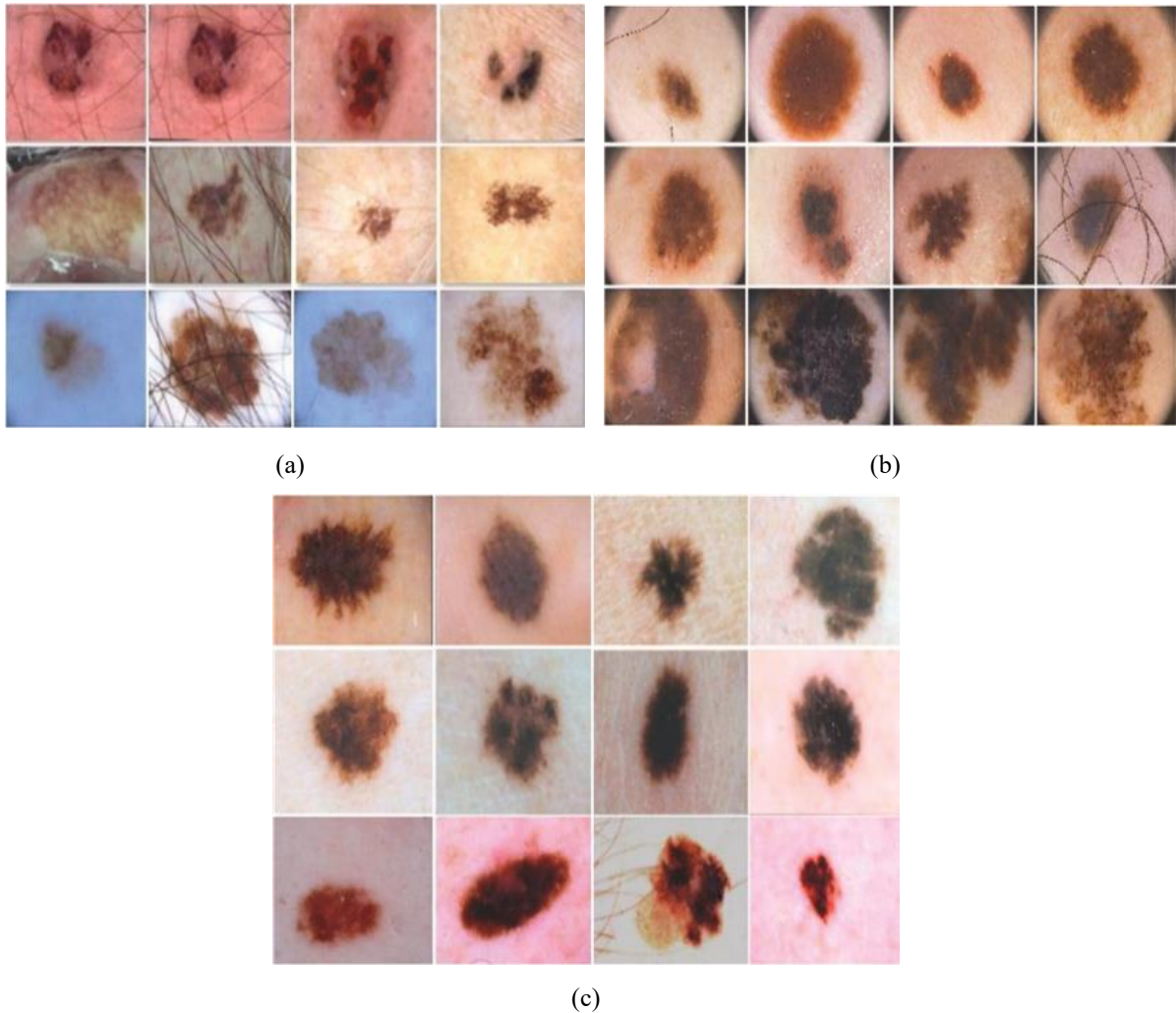

(a)   (b)

(c)

Figure 1. Sample dermoscopic images from (a) ISBI 2017, (b) PH2, and (c) ISBI 2018 datasets

3.2 Preprocessing

In the preprocessing stage, the primary objective was to enhance the quality of dermoscopic images and reduce noise in skin lesion images. This phase is critical for developing an effective skin lesion segmentation and classification model. The initial step involved applying a hair removal technique to the lesion area, followed by an intensity-based image quality enhancement process on the hairless dermoscopy images. This two-stage preparation was essential for accurately isolating the region of interest (ROI) in the images.

The hair removal process, executed using morphological operations, targeted the precise identification of the

ROI. Post hair removal, an intensity-based picture quality enhancement was performed, aimed at improving the pixel quality of the now hairless dermoscopic images. The resultant images, devoid of hair and enhanced in quality, proved instrumental in the accurate segmentation of the ROI.

The entropy of the processed images serves as a key metric for validation. Entropy, in this context, quantifies the degree of randomness or disorder within the image data. Mathematically, entropy is defined as:

$$S = kb \ln W \tag{1}$$

where, kb represents the Boltzmann constant, measured at 1.38064852 1023 m^2 kgs^2 K^1, and W denotes the microscopic distribution. The closer the entropy of the final processed image is to that of the original, the more effective the preprocessing is considered. The mask's preprocessing threshold is adjusted based on each entropy calculation, and it is maintained constant once an increase in entropy difference is observed.

Figure 2 displays examples of pre-processed images, illustrating the outcomes of the HR-IQE method utilized in the preprocessing stage.
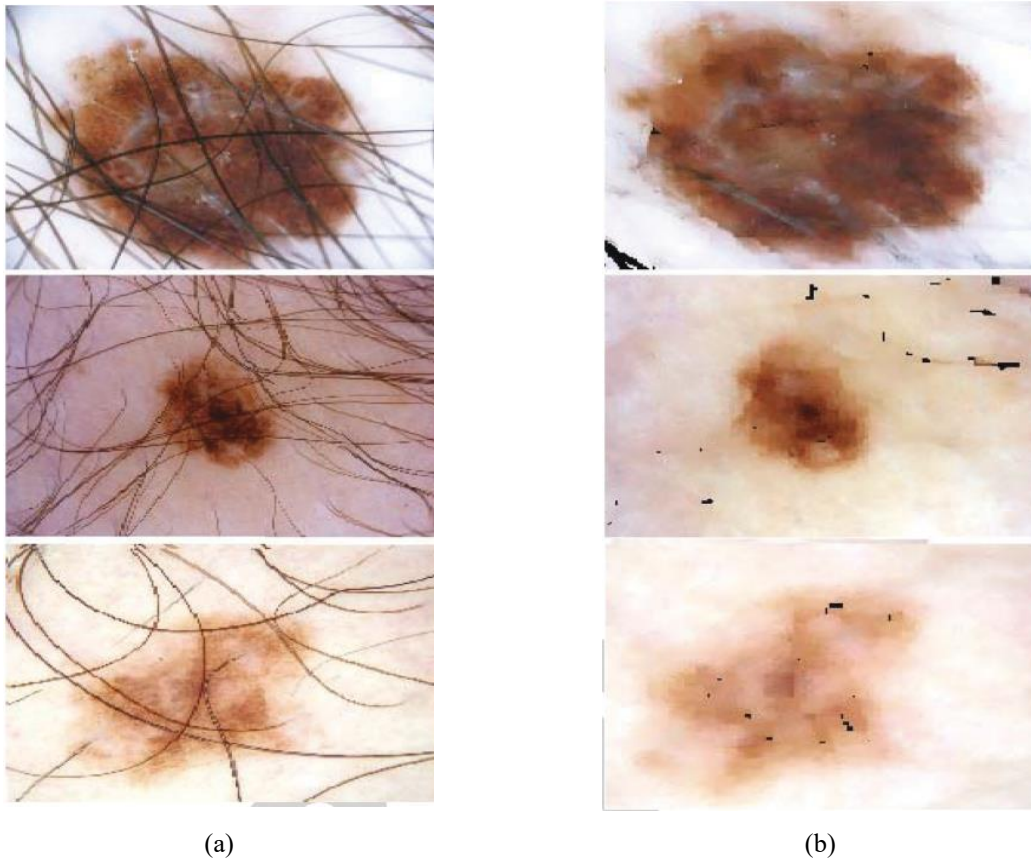

(a)   (b)

Figure 2. Preprocessing outcomes: (a) original image and (b) preprocessed image

3.3 Segmentation Methodology

The SAM mechanism, operated in grid prompt mode, was utilized to generate segmentation masks for an input image. In this mode, segmentation masks were created across all logical sections of the image, with each mask subsequently stored in a database. A segmentation prior map was then produced for each mask, using the stability score provided by SAM to determine the drawing level. In addition to the segmentation prior map, a boundary prior map was also generated based on the external borders of each mask from the list. Thus, for a given image x, two prior maps, namely prior seg and prior boundary, were created.

3.3.1 Augmenting input images

Post-creation of the prior maps, these were employed to augment the input image x. The augmentation method chosen for this study involved adding the prior maps directly to the raw image. Medical image segmentation tasks often entail a three-class segmentation job: the background, ROIs, and boundaries between ROIs and the

19

background. In the augmented image, the segmentation prior map was superimposed on the second channel, while the boundary prior map was placed on the third. For grayscale images, a three-channel image was constructed with the raw image, segmentation prior map, and boundary prior map as the first, second, and third channels, respectively. This process resulted in an augmented version $x_{aug} = Aug(prior_{seg}, prior_{boundary}, x)$ of each image x in the training set.

3.3.2 Model training with SAM-augmented images

The self-attention mechanism is an integral component of deep neural networks, notably in Transformer models, which have extended their applications beyond natural language processing to include computer vision. SAM is pivotal in capturing long-range dependencies and contextual information within images, especially for image segmentation tasks. It serves as an alternative to convolutional layers, effectively capturing spatial relationships between pixels or regions. The central element of SAM is the computation of attention weights, signifying the significance of various regions or pixels relative to a query. These weights are calculated using a similarity function, typically the dot product, between the query and key for all query-key pairs, followed by the application of a softmax function to yield the attention weights.

Once attention weights are determined, they are utilized to aggregate information from the values, proportionally to the relevance of each region to the query.

With the augmented training set $\{(x_1{}^{aug}, y_1), (x_2{}^{aug}, y_2),..., (x_n{}^{aug}, y_n)\}$, where $x_i{}^{aug} \in R^{w \times h \times C}$, $y_i \in \{0,1\}^{w \times h \times C}$ represents the annotation set, the learning objective was formulated considering the limits of M:

$$\sum_{i=1}^{n} \text{loss}\left(M\left(x_i^{aug}\right), y_i\right)$$

(2)

This objective focused solely on training with augmented images. During model testing, the trained model received only SAM-augmented images. In scenarios where SAM did not provide plausible prior maps, consideration was given to training a segmentation model using both raw and SAM-augmented images. The new objective function for the model M was defined as:

$$\sum_{i=1}^{n} \beta \, \text{loss}\left(M\left(x_i\right), y_i\right) + \lambda \, \text{loss}\left(M\left(x_i^{aug}\right), y_i\right)$$

(3)

where, β and λ regulate the importance of training loss for samples with raw and augmented images, respectively. By default, both β and λ were set to 1. A Stochastic Gradient Descent (SGD) -optimization function was employed.

3.3.3 Model deployment with SAM-augmented images

In scenarios where the model is exclusively trained on SAM-augmented images, the deployment (testing) phase mandates the use of SAM-augmented images as input. The deployment of the model can be mathematically represented as:

$$\hat{y} = \tau\left(M\left(x^{aug}\right)\right)$$

(4)

In this equation, τ denotes an output activation function, such as a sigmoid or softmax function, and x^{aug} is a SAM-augmented image. However, when the segmentation model M is trained on both raw and SAM-augmented images, additional opportunities emerge at inference time to harness the full potential of the trained model. A practical approach involves dual inference for each test sample: initially using the raw image x as input, followed by the SAM-augmented image. The final segmentation output is derived as an average ensemble of these two outputs, formally expressed as:

$$\hat{y} = \tau\left(M(x) + M\left(x^{aug}\right)\right)$$

(5)

Another strategy for leveraging the two output candidates, $M(x)$ and $M(x^{aug})$, is to select a plausible segmentation output from these candidates:

$$\hat{y} = \tau\left(M\left(x^*\right)\right)$$

(6)

where, x^* is chosen by solving the optimization problem:

$$x^* = \operatorname{argmin}_{x' \in \{x, x^{avg}\}} \text{Entropy}\left(\tau\left(M\left(x'\right)\right)\right) \qquad (7)$$

This selection is based on the entropy, or prediction certainty, of the segmentation output. An output with lower entropy indicates a higher certainty in the model's prediction, which often correlates positively with segmentation accuracy (Wang et al., 2020). Figure 3 displays sample images for segmentation, showcasing the input and segmented images.
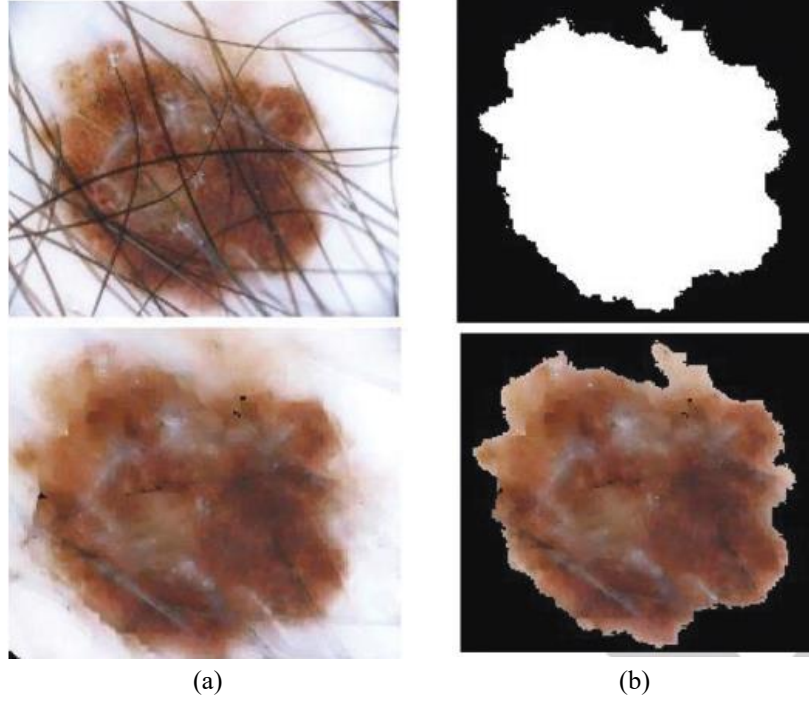

(a)   (b)

Figure 3. (a) Sample input image and (b) segmented image

3.4 Proposed DBN for Classification

The implementation of DBNs in the proposed skin cancer detection system leverages recent advances in deep learning. DBNs, an energy-based model for generating probabilities, are constructed by stacking multiple Restricted Boltzmann Machines (RBMs). The DBN model operates in two phases: training and testing. In the training phase, the DBN layers are trained utilizing a series of RBM layers. These layers comprise neurons located in the hidden layer, situated above the input layer. Neurons between different layers communicate freely, whereas intra-layer communication is restricted. The architecture of the surrounding visible layer is represented as follows:

$$vis = \left\{ vis_1, vis_2, \ldots, vis_j, \ldots, vis_m \right\}; \left(vis_j \in \{0,1\} \right) \qquad (8)$$

Similarly, the expression for any one of the hidden layers in the DBN is denoted as:

$$hid = \left\{ hid_1, hid_2, \ldots, hid_j, \ldots, hid_m \right\}; \left(hid_j \in \{0,1\} \right) \qquad (9)$$

In this model, visible layers are denoted by 'vis', and hidden layers are indicated by 'hid'. The classification network (CN) technique, embedded within the DBN, determines the number of layers based on the pre-processed input result $N(y,z)$. Adjusted characteristics are transmitted to the visible layer, which in turn communicates them to the hidden layer via a weight relationship, rather than direct connections. The primary parameters of this architecture are defined as $= W,b,c$, where W represents the weight matrix, B the bias in the hidden layer, and C the bias in the visible layer. vis_j indicates the visible unit in the j-th layer and hid_j denotes the hidden unit in the j-th layer. The RBM architecture, consisting of n hidden neurons and m input neurons, is mathematically structured as:

$$W = \left\{ w_{j,k} \in S^{m \times n} \right\} \tag{10}$$

where, $w_{j,k}$ denotes the weight value between neurons in the j-th and k-th layers. The bias function for the hidden layer B is articulated as:

$$B = \left\{ b_j \in S^n \right\} \tag{11}$$

The term b_j in the equation represents the bias threshold for the hidden neuron at the j-th position. A similar methodology is employed for determining the bias function of the visible layer:

$$C = \left\{ c_k \in S^m \right\} \tag{12}$$

In this context, c_k signifies the bias function's threshold for the k-th observable neuron. The RBM model's energy function between the hidden and visible layers is employed to learn probability distributions, as calculated by:

$$E(vis, hid| \ \theta) = -\sum_{j=1}^{m} b_j vis_j - \sum_{k=1}^{n} c_k hid_k - \sum_{j=1}^{m} \sum_{k=1}^{n} vis_j W_{jk} hid_k \tag{13}$$

where, $\theta = \left\{ W_{jk}, b_j, c_k \right\}$ consists of the RBM model's parameters, and E, which defines the energy function between hidden and visible nodes. In the DBN architecture under consideration, the total count of neurons within the visible layer is denoted as m, while the aggregate of neurons across the hidden layers is represented as n. This configuration underpins the structural framework of the network.

Upon incorporating the exponential and regularization of the energy function, the probability function is defined, establishing the overlap between the probabilities of the RBM model's visible and hidden layers:

$$P(vis, hid| \ \theta) = \frac{1}{Z(\theta)} e^{-E(\text{ vis, hid} \ | \ \theta)} \tag{14}$$

This equation is grounded in the Gibbs distribution function of the RBM model. From this, the partition function is derived:

$$Z(\theta) = \sum_{vis, hid} e^{-E(\text{ vis,hid} \ | \ \theta)} \tag{15}$$

where, the function $Z(\theta)$ represents an energy state that is either uniformly distributed or normalized. This function encapsulates the total energy of all layers in the DBN, both visible and hidden. It is defined as the sum of energy evaluations conducted on these layers. The probability function in the DBN model is utilized to regulate the parameters' standards. The notation $P(vis,hid/)$ represents the joint distribution of the visible and hidden layers, while $P(vis/)$ symbolizes the marginal distribution function of the visible layer:

$$P(vis| \ \theta) = \frac{1}{Z(\theta)} \sum_{hid} e^{-E(\text{ vis,hid} \ | \ \theta)} \tag{16}$$

The marginal layer is evaluated by summarizing the conditions of the entire network. For computing the marginal distribution function of the hidden layer, the following equation is used:

$$P(hid| \ \theta) = \frac{1}{Z(\theta)} \sum_{vis} e^{-E(vis,hid| \ \theta)} \tag{17}$$

Given the binary nature of its components, a sigmoid activation function is deemed suitable for the proposed RBM architecture. The conditional probability values within the RBM are calculated as follows, given the independence of its visible and hidden layers:

$$P(vis| \ hid) = \prod_i P(vis_j| \ hid) \tag{18}$$

$$P(hid \mid vis) = \prod_k P(hid_k \mid vis) \tag{19}$$

The probabilities for the RBM structure's visible and hidden layers are then derived using this activation function:

$$sigmoid(y) = \frac{1}{\left(1 + e^{-y}\right)} \tag{20}$$

With this activation function as a basis, the probabilities for the RBM structure's visible and hidden layers are as follows:

$$P\left(vis_j = 1 \mid hid\right) = \frac{1}{1 + \exp\left(-b_j - \sum_{k=1}^{n} w_{jk} hid_k\right)} \tag{21}$$

$$P\left(hid = 1 \mid vis\right) = \frac{1}{1 + \exp\left(-c_k - \sum_{j=1}^{m} w_{jk} vis_j\right)} \tag{22}$$

The next phase involves updating the rules for relevant parameters $=W,b,c$. The Gibbs distribution function, foundational to the RBM model, necessitates an efficient learning methodology due to its complexity. Thus, contrast divergence is employed as a rapid learning technique to minimize time consumption. The updated parameter values, based on this learning method, are quantitatively presented in the following equations:

$$W^{time+1} = W^{time} + \varepsilon\left(P\left(hid \mid vis^{(0)}\right)\left[vis^{(0)}\right]^{time} - P\left(hid \mid vis^{(1)}\right)\left[vis^{(1)}\right]^{time}\right) \tag{23}$$

$$b^{time+1} = b^{time} + \varepsilon\left(vis^{(0)} - vis^{(1)}\right) \tag{24}$$

$$c^{time+1} = c^{time} + \varepsilon\left(P\left(hid \mid vis^{(0)}\right) - P\left(vis^{(1)}\right)\right) \tag{25}$$

In these equations, 'time' denotes the magnitude of iterative learning steps. This process continues until the parameter values have been adjusted to yield a feature representation that is more abstract and representable than that produced by the lower layer. The DBN model utilized these equations for efficient feature extraction. The training method of the coot optimization process allows the proposed detection model to bypass specialized capabilities, producing a more sophisticated set of features overall, thereby enhancing the effectiveness of the detection system.

3.4.1 Training DBN model with COA

The optimization of the DBN model is achieved through the application of the COA, an innovative optimization technique derived from the behavioral patterns of coots, as detailed in (Houssein et al., 2022). The COA emulates the synchronized activities of a coot flock on water, encompassing behaviors such as aimless roving, chain walking, position adjustments relative to leaders, and guiding the flock to optimal areas. These behaviors are mathematically modeled to enable their implementation in the optimization process.

Initially, a population of coots is generated randomly, given a problem of D-dimensions. Utilizing Eq. (26), a population (N) of coots is created within the constraints of upper boundaries (UB):

$$PosCoot(i) = random(1, D) \times (UB - LB) + LB, i = 1, 2, \ldots, N \tag{26}$$

This equation ensures that the initial positions of the coots in the multidimensional space are randomly determined, respecting the predefined upper boundaries. The fitness of this initial population is assessed as per Eq. (27):

$$F(i) = \text{Fitness}\left(PosCoot(i)\right), i = 1, 2, \ldots, N \tag{27}$$

Subsequently, random locations for the coots are generated using Eq. (28). This is followed by recalculating these locations as per Eq. (29):

$$R = \text{random}(1, D) \times (UB - LB) + LB \tag{28}$$

$$\text{PosCoot}(i) = \text{PosCoot}(i) + A \times RN2 \times (R - \text{PosCoot}(i)) \tag{29}$$

In Eq. (29), RN2 denotes a random number between 0 and 1.

$$A = 1 - \left(T(i) \times \frac{1}{IterMax} \right), B = 2 - \left(T(i) \times \frac{1}{IterMax} \right) i = 1, 2, \ldots, IterMax \tag{30}$$

During the current iteration $T(i)$, the maximum number of iterations *IterMax* is used in Eq. (30). Chain movement, a critical behavior of coots, is simulated as one coot moving towards another, as described in Eq. (31):

$$\text{PosCoot}(i) = 0.5 \times (\text{PosCoot}(i-1) + \text{PosCoot}(i)) \tag{31}$$

Leader selection among the coots is governed by Eq. (32):

$$L_{\text{ind}} = 1 + \left(i MOD \, N_L \right) \tag{32}$$

In Eq. (32), N_L represents the parameterized total number of leaders, while Lind signifies the index of a leader. The probability p is also defined in this context. The ranking of leaders is then conducted as per the criteria in Eq. (33):

$$\text{LeaderPos}(i) = \begin{cases} B \times R3 \times \cos(2R\pi) \times (gBest - LeaderPos(i) + gBest \cdots R4 < P) \\ B \times R3 \times \cos(2R\pi) \times (gBest - LeaderPos(i) + gBest \cdots R4 \geq P) \end{cases} \tag{33}$$

Eq. (33) incorporates random integers *R3* and *R4* within the range [0, 1], the current global best *gBest*, and a constant value of 3.14. The pseudocode for the COA is detailed in Algorithm 1.

Procedure 1. Pseudocode of the COA
1. *Initialize the first population of coots randomly by Eq.* (26)
2. *Initialize the termination criteria, probability p, number of leaders and number of coots*
3. *Ncoot = Number of coots − Number of leaders*
4. *Random selection of leaders from the coots*
5. *Calculate the fitness of coots and leaders*
6. *Find the best coot or leader the global optimum while the end criterion is not satisfied*
7. *Calculate A, B parameters by Eq.* (30)
8. *If rand < P*
9. *R, R1, and R3 are random vectors along the dimensions of the problem*
10. *Else*
11. *R, R1, and R3 are random number*
12. *End*
13. *For i = 1 to the number of the coot*
14. *Calculate the parameter of K by Eq.* (32)
15. *If rand > 0.5*
16. *Update the position of the coot by Eq.* (33)
17. *Else*
18. *If rand < 0.5 i~ = 1*
19. *Update the position of the coot by Eq.* (33)

```
20.  Else
21.  Update the position of the coot by Eq. (31)
22.  End
23.  End
24.  Calculate the fitness of coot
25.  If the fitness of coot
26.  If the fitness of coot < the fitness of leader (k)
27.  Temp = Leader(k); leader(k) = coot; coot = Temp;
28.  End
29.   For number of Leaders
30.  Update the position of the leader using the rules given in Eq. (33)
31.  If the fitness of leader < gBest
32.  Temp = gBest; gBest = leader; leader = Temp; (update global optimum)
33.  End
34.  End
35.  Iter = iter + 1;
36.  End
37.  Postprocessor results
```

4. Results and Discussion

The experimental studies were conducted on a system equipped with an Intel(R) Core (TM) i5-7200u processor, ranging from 2.50 to 2.7 GHz, complemented by 16 GB RAM and an 8 GB graphics card, utilizing MATLAB 2020b.

4.1 Segmentation Analysis

In evaluating the segmentation performance, metrics such as accuracy (ACC), specificity (SPE), and sensitivity (SEN), alongside the Jaccard and Dice similarity coefficients (JSC and DSC, respectively), were employed. These metrics are pivotal in assessing the effectiveness of the segmentation method. The true positive (TP) rate was determined by identifying the overlapping area between the segmented region and the ground truth, denoted as y and x, respectively. This intersection is represented mathematically as $TP = yx$. Conversely, the false positive (FP) rate was defined as $FP = yx$, which represents the segmented region in x not corresponding to y. The false negative (FN) rate, indicating the proportion of incorrect predictions, was calculated as $FN = yx$. The true negative (TN) set, representing the complement of the segmented and ground truth regions, is described as $TN = \bar{y} \cap \bar{x}$. The following equations precisely articulate the calculations of the aforementioned metrics:

These metrics were integral in quantitatively assessing the segmentation's accuracy and reliability, thereby providing a comprehensive understanding of the model's performance in skin lesion segmentation.

$$ACC = \frac{TP+TN}{TP+TN+FP+FN} \tag{34}$$

$$DSC = \frac{2.TP}{2.TP+FP+FN} \tag{35}$$

$$JSC = \frac{TP}{TP+FP+FN} \tag{36}$$

$$SEN = \frac{TP}{TP+FN} \tag{37}$$

$$SPE = \frac{TN}{TP + FP} \tag{38}$$

4.2 Performance Analysis of Segmentation Models on Three Datasets

The segmentation performance of the proposed model was evaluated across three different datasets, with the findings detailed in Table 1, Table 2, and Table 3.

Table 1 shows the validation investigation of the proposed model on Dataset 1. In the analysis of Dataset 1, the performance of the proposed model was evaluated using various statistical measures. For Image 1, it was observed that the dice coefficient reached 84.90, the Jaccard index was calculated at 76.5, sensitivity was found to be 82.5, specificity stood at 97.5, and accuracy was measured at 93.40. The analysis of subsequent images in the dataset revealed variations in these metrics, indicating the model's differentiated performance across various images. Specifically, Image 2 demonstrated a dice coefficient of 76.27, Jaccard index of 61.64, sensitivity of 67.15, specificity of 97.24, and accuracy of 90.14. In contrast, Image 3 showed improved outcomes with a dice coefficient of 87.08, Jaccard index of 77.11, sensitivity of 85.40, specificity of 96.69, and accuracy of 94.03. Continuing this trend, Image 4 exhibited a dice coefficient of 87.80, Jaccard index of 78.20, sensitivity of 81.60, specificity of 98.30, and accuracy of 93.60. Image 5 further reflected the model's robustness with a dice coefficient of 88.06, Jaccard index of 82.03, sensitivity of 89.80, specificity of 96.44, and accuracy of 94.70.

Table 1. Validation analysis of the proposed model on Dataset 1

Image	Dice	Jaccard	Sensitivity	Specificity	Accuracy
Image 1	84.90	76.5	82.5	97.5	93.40
Image 2	76.27	61.64	67.15	97.24	90.14
Image 3	87.08	77.11	85.40	96.69	94.03
Image 4	87.80	78.20	81.60	98.30	93.60
Image 5	88.06	82.03	89.80	96.44	94.70

These results demonstrate the model's varying degrees of accuracy in segmenting and identifying lesions in dermoscopic images across different cases within the dataset. The variation in metrics such as dice, Jaccard, sensitivity, specificity, and accuracy across different images is indicative of the model's adaptability and effectiveness in handling diverse image characteristics.

Table 2. Experimental analysis on Dataset 2

Image	Dice	Jaccard	Sensitivity	Accuracy	Specificity
Image 1	84.90	76.50	82.50	93.40	97.50
Image 2	84.70	76.20	82.00	93.20	97.80
Image 3	87.08	77.11	85.40	94.03	96.69
Image 4	88.13	79.54	83.63	92.99	94.02
Image 5	84.26	74.81	90.82	93.39	92.68

Table 2 characterises the experimental analysis on Dataset 2. In the experimental analysis conducted on Dataset 2, the performance of the proposed model was systematically evaluated across multiple images. The analysis revealed that for Image 1, the dice score was recorded at 84.90, the Jaccard index was 76.50, sensitivity was 82.50, accuracy stood at 93.40, and specificity was 97.50. These metrics illustrate the model's effectiveness in accurately segmenting and classifying skin lesions. Subsequent images within the dataset displayed similar levels of performance, albeit with some variations. For instance, Image 2 demonstrated a dice score of 84.70 and a Jaccard index of 76.20, while Image 5 showed a dice score of 84.26 and a Jaccard index of 74.81. Notably, sensitivity, accuracy, and specificity also varied across the images, indicating the model's capacity to handle diverse image characteristics. In particular, Image 3 exhibited a dice score of 87.08 and a Jaccard index of 77.11, combined with a sensitivity of 85.40 and an accuracy of 94.03. Image 4, on the other hand, achieved a dice score of 88.13 and a Jaccard index of 79.54, highlighting the model's consistent performance in terms of segmentation accuracy. Overall, the results from Dataset 2 underscore the proposed model's capability in effectively segmenting and classifying dermoscopic images. The consistency observed in key metrics such as dice, Jaccard, sensitivity, specificity, and accuracy across different images further reinforces the model's applicability in clinical diagnostic settings.

Table 3 presents a comprehensive validation analysis of the proposed model on Dataset 3. In this analysis, the performance metrics for each image were carefully evaluated. For Image 1, it was observed that the dice coefficient achieved a value of 94.21, and the Jaccard index was 87.63. sensitivity was recorded at 90.17, with a corresponding specificity of 87.63. These metrics collectively indicate the model's high accuracy in lesion detection and

segmentation. Subsequent images within this dataset exhibited similar levels of performance, albeit with variations indicative of the model's adaptability to different image characteristics. For example, Image 2 displayed a dice score of 92.68 and a Jaccard index of 93.91, along with sensitivity and specificity scores of 92.56 and 91.26, respectively. This trend continued with Image 3, where the dice score rose to 94.83, and the Jaccard index reached 94.57, further demonstrating the model's effectiveness. Image 4 and Image 5 maintained this pattern of high performance, with dice scores of 91.34 and 92.12, and Jaccard indices of 92.04 and 92.44, respectively. Sensitivity and specificity scores for these images were also within a high range, underscoring the robustness of the proposed model in accurately classifying and segmenting skin lesions in dermoscopic images. Overall, the results from Dataset 3 reinforce the proposed model's proficiency in handling diverse dermoscopic images. The consistency in high scores across key metrics such as accuracy, dice, Jaccard, sensitivity, and specificity across different images further cements the model's potential for clinical application in the accurate diagnosis of skin lesions.

Table 3. Validation analysis of the proposed model on Dataset 3

Image	Accuracy	Dice	Jaccard	Sensitivity	Specificity
Image 1	94.21	87.63	90.17	83.34	87.63
Image 2	92.68	93.91	92.56	88.72	91.26
Image 3	94.83	94.57	94.83	89.33	92.61
Image 4	91.34	92.04	91.29	85.27	88.18
Image 5	92.12	92.44	92.36	86.67	89.24

4.3 Experiment on Classification Results

To assess the performance of the proposed model in classification, various metrics were evaluated. These included accuracy, sensitivity, specificity, precision, false positive rate (FPR), false negative rate (FNR), negative predictive value (NPV), false discovery rate (FDR), F1-score, and MCC.

a. Accuracy: Defined as the ratio of correctly predicted observations to the total number of observations (Eq. (39)). This metric is crucial in determining the overall effectiveness of the model.

$$T^{accuracy} = \frac{\left(Tr^p + Tr^n\right)}{\left(Tr^p + Tr^n + Fa^p + Fa^n\right)} \tag{39}$$

b. Sensitivity: Measured as the number of true positives accurately identified (Eq. (40)). Sensitivity is indicative of the model's ability to correctly detect the presence of a condition.

$$Se = \frac{Tr^p}{Tr^p + Fa^n} \tag{40}$$

c. Specificity: Calculated as the number of true negatives that are precisely determined (Eq. (41)). This metric assesses the model's accuracy in identifying the absence of a condition.

$$sp = \frac{Tr^n}{Fa^n} \tag{41}$$

d. Precision: Represents the ratio of accurately predicted positive observations to the total number of positive predictions made (Eq. (42)). Precision is vital for understanding the model's exactness.

$$Pr = \frac{Tr^p}{Tr^p + Tr^p} \tag{42}$$

e. FPR: The ratio of false positive predictions to the total count of negative predictions (Eq. (43)). FPR provides insights into the instances where the model incorrectly predicts a positive outcome.

$$FPR = \frac{Fa^P}{Fa^P + Tr^n} \tag{43}$$

f. FNR: The proportion of positive cases that result in negative test outcomes (Eq. (44)). FNR is crucial for assessing the model's ability to avoid missed diagnoses.

$$FNR = \frac{Fa^n}{Tr^n + Tr^p} \tag{44}$$

g. NPV: The likelihood that subjects with a negative test result genuinely do not have the condition (Eq. (45)). NPV is essential in evaluating the model's reliability in negative predictions.

$$NPV = \frac{Fa^n}{Fa^n + Tr^n} \tag{45}$$

h. FDR: Calculated as the number of false positives in all rejected hypotheses (Eq. (46)). FDR is critical for understanding the rate of error in the model's predictions.

$$FDR = \frac{Fa^p}{Fa^p + Tr^p} \tag{46}$$

i. F1-score: Defined as the harmonic mean between precision and recall (Eq. (47)). The F1-score is a crucial statistical measure for evaluating the model's performance balance.

$$F1score = \frac{Se.Pr}{Pr + Se} \tag{47}$$

j. MCC: A correlation coefficient computed using four values (Eq. (48)). MCC is a comprehensive measure that considers all four quadrants of the confusion matrix, providing a holistic view of the model's performance.

$$MCC = \frac{Tr^p \times Tr^n - Fa^p \times Fa^n}{\sqrt{\left(Tr^p + Fa^p\right)\left(Tr^p + Fa^n\right)\left(Tr^n + Fa^p\right)\left(Tr^p + Fa^n\right)}} \tag{48}$$

Each of these metrics plays an integral role in providing a detailed and nuanced understanding of the model's capabilities, strengths, and areas for improvement in the context of medical image analysis.

The performance of existing models such as MD2N (Kumar et al., 2023) and DCNN (Singh et al., 2023; Sethanan et al., 2023), alongside the proposed model, was scrutinized. The results are presented in Table 4, Table 5 and Table 6.

Table 4. Classifier analysis on Dataset 1

Measures	MD2N	DCNN	DBN	AE	DBN-COA
Accuracy	0.92969	0.91927	0.95573	0.94271	0.965
Sensitivity	0.89063	0.88021	0.94792	0.91667	0.942
Specificity	0.96875	0.95833	0.96354	0.96875	0.988
Precision	0.9661	0.9548	0.96296	0.96703	0.988
FPR	0.03125	0.041667	0.036458	0.03125	0.0155
FNR	0.10938	0.11979	0.052083	0.083333	0.0523
NPV	0.96875	0.95833	0.96354	0.96875	0.988
FDR	0.033898	0.045198	0.037037	0.032967	0.0116
F1-score	0.92683	0.91599	0.95538	0.94118	0.9652
MCC	0.86201	0.84111	0.91157	0.88662	0.9321

Table 4 in the study provides an insightful classifier analysis for Dataset 1, evaluating various models including MD2N, DCNN, DBN, AE, and DBN-COA. The evaluation encompassed several key metrics. In the analysis of Accuracy, it was observed that the MD2N model achieved a score of 0.92969, while the DCNN model recorded 0.91927, and the DBN model reached 0.95573. The AE model demonstrated an accuracy rate of 0.94271, with the DBN-COA model surpassing all with a score of 0.96615. When assessing sensitivity, the MD2N model was found to have a rate of 0.89063, and the DCNN model's rate stood at 0.88021. The DBN model achieved 0.94792, paralleled by the AE model at 0.91667, and the DBN-COA model also reached 0.94792. Specificity calculations revealed that the MD2N model achieved 0.96875, the DCNN model attained 0.95833, and the DBN model scored 0.96354. The AE model's specificity was noted as 0.96875, with the DBN-COA model achieving the highest at 0.98438. In precision, the MD2N model scored 0.9661, and the DCNN model attained 0.9548. The DBN model reached 0.96296, closely followed by the AE model at 0.96703, and the DBN-COA model achieved 0.98378. The

FPR analysis showed the MD2N model with 0.03125, the DCNN model with 0.041667, and the DBN model with 0.036458. The AE model recorded an FPR of 0.03125, while the DBN-COA model had a lower rate of 0.015625. In the FNR metric, the MD2N model recorded 0.10938, the DCNN model 0.11979, and the DBN model 0.052083. The AE model scored 0.083333, closely followed by the DBN-COA model at 0.052083. For the NPV, the MD2N model attained 0.96875, the DCNN model reached 0.95833, and the DBN model scored 0.96354. The AE model had an NPV of 0.96875, with the DBN-COA model achieving 0.98438. In the FDR, the MD2N model scored 0.033898, the DCNN model 0.045198, and the DBN model 0.037037. The AE model recorded an FDR of 0.032967, and the DBN-COA model had a rate of 0.016216. The F1-score calculations showed the MD2N model with 0.92683, the DCNN model with 0.91599, and the DBN model with 0.95538. The AE model achieved 0.94118, and the DBN-COA model excelled with 0.96552. Finally, in the MCC, the MD2N model achieved 0.86201, the DCNN model 0.84111, and the DBN model 0.91157. The AE model recorded 0.88662, with the DBN-COA model reaching the highest at 0.93291. These results from Dataset 1 demonstrate the varying degrees of efficacy of the different models, with the DBN-COA model generally showing superior performance across most metrics.

Table 5. Analysis of different models on Dataset 2

Measures	MD2N	DCNN	DBN	AE	DBN-COA
Accuracy	0.95517	0.94828	0.95172	0.95862	0.9786
Sensitivity	0.94483	0.92414	0.92414	0.93103	0.9741
Specificity	0.96552	0.97241	0.97931	0.98621	0.9731
Precision	0.96479	0.97101	0.9781	0.9854	0.97917
FPR	0.034483	0.027586	0.02069	0.013793	0.0269
FNR	0.055172	0.075862	0.075862	0.068966	0.0286
NPV	0.96552	0.97241	0.97931	0.98621	0.9731
FDR	0.035211	0.028986	0.021898	0.014599	0.0233
F1-score	0.9547	0.947	0.95035	0.95745	0.9778
MCC	0.91054	0.8976	0.90483	0.91864	0.9575

Table 5 in the study delineates the analysis of various models' performance on Dataset 2, employing a range of metrics to evaluate their effectiveness. The accuracy metric revealed that the MD2N model achieved a score of 0.95517, while the DCNN model recorded 0.94828, and the DBN model reached 0.95172. The AE model demonstrated an accuracy rate of 0.95862, with the DBN-COA model surpassing all with a score of 0.97586. In terms of sensitivity, the MD2N model was found to have a rate of 0.94483, and the DCNN model's rate stood at 0.92414. The DBN model achieved 0.92414, paralleled by the AE model at 0.93103, and the DBN-COA model also reached 0.97241. Specificity calculations revealed that the MD2N model achieved 0.96552, the DCNN model attained 0.97241, and the DBN model scored 0.97931. The AE model's specificity was noted as 0.98621, with the DBN-COA model achieving the highest at 0.97931. In precision, the MD2N model scored 0.96479, and the DCNN model attained 0.97101. The DBN model reached 0.9781, closely followed by the AE model at 0.9854, and the DBN-COA model achieved 0.97917. The FPR analysis showed the MD2N model with 0.034483, the DCNN model with 0.027586, and the DBN model with 0.02069. The AE model recorded an FPR of 0.013793, while the DBN-COA model had a lower rate of 0.02069. In the FNR metric, the MD2N model recorded 0.055172, the DCNN model 0.075862, and the DBN model 0.075862. The AE model scored 0.068966, closely followed by the DBN-COA model at 0.027586. For the NPV, the MD2N model attained 0.96552, the DCNN model reached 0.97241, and the DBN model scored 0.97931. The AE model had an NPV of 0.98621, with the DBN-COA model achieving 0.97931. In the FDR, the MD2N model scored 0.035211, the DCNN model 0.028986, and the DBN model 0.021898. The AE model recorded an FDR of 0.014599, and the DBN-COA model had a rate of 0.020833. The F1-score calculations showed the MD2N model with 0.9547, the DCNN model with 0.947, and the DBN model with 0.95035. The AE model achieved 0.95745, and the DBN-COA model excelled with 0.97578. Finally, in the MCC, the MD2N model achieved 0.91054, the DCNN model 0.8976, and the DBN model 0.90483. The AE model recorded 0.91864, with the DBN-COA model reaching the highest at 0.95175. These results from Dataset 2 demonstrate the varying degrees of efficacy of the different models, with the DBN-COA model generally showing superior performance across most metrics.

Table 6 presents a comparative analysis of various models on Dataset 3, employing a multitude of metrics to ascertain their performance efficacy. The accuracy metric revealed that the MD2N model achieved a score of 0.91467, the DCNN model recorded 0.89069, and the DBN model reached 0.91733. The AE model demonstrated an accuracy rate of 0.91467, with the DBN-COA model surpassing all with a score of 0.9466. In terms of FDR, the MD2N model recorded a rate of 0.13559, the DCNN model 0.1791, and the DBN model 0.14583. The AE model had an FDR of 0.15625, while the DBN-COA model achieved a lower rate of 0.0769. The F1-score calculations showed the MD2N model with 0.76119, the DCNN model with 0.77465, and the DBN model with 0.66667. The AE model achieved 0.77698, and the DBN-COA model excelled with 0.85714. Sensitivity assessments indicated that the MD2N model achieved 0.68, the DCNN model 0.73333, and the AE model 0.54667.

The DBN model scored 0.72, and the DBN-COA model attained the highest at 0.8. Specificity analysis showed the MD2N model with 0.97333, the DCNN model with 0.96, and the AE model with 0.97678. The DBN model reached 0.966, and the DBN-COA model achieved 0.983. In precision, the MD2N model scored 0.86441, and the DCNN model attained 0.8209. The DBN model reached 0.85425, closely followed by the AE model at 0.843, and the DBN-COA model achieved 0.923. The FPR for the MD2N model was 0.026667, the DCNN model 0.04, and the AE model 0.023333. The DBN model recorded an FPR of 0.0333, while the DBN-COA model had a lower rate of 0.0166. In the FNR metric, the MD2N model recorded 0.32, the DCNN model 0.26667, and the AE model 0.45333. The DBN model scored 0.28, closely followed by the DBN-COA model at 0.2. For the NPV, the MD2N model attained 0.97333, the DCNN model reached 0.96, and the AE model scored 0.97667. The DBN model had an NPV of 0.96667, with the DBN-COA model achieving 0.983. Finally, in the MCC, the MD2N model achieved 0.71772, the DCNN model 0.72397, and the AE model 0.62658. The DBN model recorded 0.73007, with the DBN-COA model reaching the highest at 0.82775. These results from Dataset 3 demonstrate the varying degrees of efficacy of the different models, with the DBN-COA model generally showing superior performance across most metrics.

Table 6. Comparative analysis of various models on Dataset 3

Measures	MD2N	DCNN	DBN	AE	DBN-COA
Accuracy	0.91467	0.91467	0.89069	0.91733	0.9466
FDR	0.13559	0.1791	0.14583	0.15625	0.0769
F1-score	0.76119	0.77465	0.66667	0.77698	0.85714
Sensitivity	0.68	0.73333	0.54667	0.72	0.8
Specificity	0.97333	0.96	0.97678	0.966	0.983
Precision	0.86441	0.8209	0.85425	0.843	0.923
FPR	0.026667	0.04	0.023333	0.0333	0.0166
FNR	0.32	0.26667	0.45333	0.28	0.2
NPV	0.97333	0.96	0.97667	0.96667	0.983
MCC	0.71772	0.72397	0.62658	0.73007	0.82775

5. Conclusion

This research has successfully developed a lesion detection model based on dermoscopy images, integrating a pre-trained lightweight network for feature extraction. A novel augmentation approach, SAMAug, was introduced, enhancing medical image segmentation by leveraging the SAM to augment the inputs for established medical image segmentation models. The efficacy of this innovative strategy was empirically validated across three distinct segmentation tasks. In the classification phase, DBNs were employed, with the COA optimizing the selection of DBN weights. The proposed model's performance was rigorously tested and validated using three datasets, evaluated against a comprehensive set of metrics.

Looking ahead, the exploration of a more advanced and sophisticated augmentation function warrants consideration. Future research will focus on identifying and analyzing the variables and factors influencing the performance of the proposed method. Additionally, the transformer architecture within this context will be scrutinized, with necessary modifications to the proposed method being contemplated. Another critical avenue for future work lies in the collection and curation of larger and more diverse datasets of dermoscopy images. The development of a substantial dataset is imperative for enhancing the robustness and generalizability of the model. Collaborative efforts in creating annotated datasets will prove immensely beneficial for the research community, facilitating advancements in the field of medical image analysis.

Data Availability

The data used to support the research findings are available from the corresponding author upon request.

Conflicts of Interest

The authors declare no conflict of interest.

References

Adla, D., Reddy, G. V. R., Nayak, P., & Karuna, G. (2023). A full-resolution convolutional network with a dynamic graph cut algorithm for skin cancer classification and detection. *Healthc. Anal., 3*, 100154. https://doi.org/10.1016/j.health.2023.100154.

Alenezi, F., Armghan, A., & Polat, K. (2023). A novel multi-task learning network based on melanoma

segmentation and classification with skin lesion images. *Diagnostics, 13*(2), 262. https://doi.org/10.3390/diagnostics13020262.

Ali, K., Shaikh, Z. A., Khan, A. A., & Laghari, A. A. (2022). Multiclass skin cancer classification using EfficientNets-a first step towards preventing skin cancer. *Neurosci. Inf., 2*(4), 100034. https://doi.org/10.1016/j.neuri.2021.100034.

Ali, M. S., Miah, M. S., Haque, J., Rahman, M. M., & Islam, M. K. (2021). An enhanced technique of skin cancer classification using deep convolutional neural network with transfer learning models. *Mach. Learn. Appl., 5*, 100036. https://doi.org/10.1016/j.mlwa.2021.100036.

Anand, V., Gupta, S., Koundal, D., & Singh, K. (2023). Fusion of U-Net and CNN model for segmentation and classification of skin lesion from dermoscopy images. *Expert Syst. Appl., 213*, 119230. https://doi.org/10.1016/j.eswa.2022.119230.

Daghrir, J., Tlig, L., Bouchouicha, M., & Sayadi, M. (2020). Melanoma skin cancer detection using deep learning and classical machine learning techniques: A hybrid approach. In 2020 5th international conference on advanced technologies for signal and image processing (ATSIP), Sousse, Tunisia, pp. 1-5. https://doi.org/10.1109/ATSIP49331.2020.9231544.

Dildar, M., Akram, S., Irfan, M., Khan, H. U., Ramzan, M., Mahmood, A. R., Alsaiari, S. A., Saeed, A. H. M., Alraddadi, M. O., & Mahnashi, M. H. (2021). Skin cancer detection: a review using deep learning techniques. *Int. J. Environ. Res. Publ. Health, 18*(10), 5479. https://doi.org/10.3390/ijerph18105479.

Hong, Y., Zhang, G., Wei, B., Cong, J., Xu, Y., & Zhang, K. (2023). Weakly supervised semantic segmentation for skin cancer via CNN superpixel region response. *Multimedia Tool Appl., 82*(5), 6829-6847. https://doi.org/10.1007/s11042-022-13606-4.

Houssein, E. H., Hashim, F. A., Ferahtia, S. & Rezk, H. (2022). Battery parameter identification strategy based on modified coot optimization algorithm. *J. Energy Storage, 46*, 103848. https://doi.org/10.1016/j.est.2021.103848.

Jakkulla, P. K., Ganesh, K. M., Jayapal, P. K., Malla, S. J., Chandanapalli, S. B., & Sandhya, E. (2023). Selection of features using adaptive tunicate swarm algorithm with optimized deep learning model for thyroid disease classification. *Ing. Syst. Inf., 28*(2), 299-308. https://doi.org/10.18280/isi.280205.

Kumar, M. D., Sivanarayana, G. V., Indira, D. N. V. S. L. S., & Raj, M. P. (2023). Skin cancer segmentation with the aid of multi-class dilated D-net (MD2N) framework. *Multimed Tools Appl., 82*, 35995–36018. https://doi.org/10.1007/s11042-023-14605-9.

Kumar, M., Alshehri, M., AlGhamdi, R., Sharma, P., & Deep, V. (2020). A DE-ANN inspired skin cancer detection approach using fuzzy c-means clustering. *Mobile Netw Appl., 25*, 1319–1329. https://doi.org/10.1007/s11036-020-01550-2.

Macherla, H., Kotapati, G., Sunitha, M.T., Chittipireddy, K.R., Attuluri, B., & Vatambeti, R. (2023). Deep learning framework-based chaotic hunger games search optimization algorithm for prediction of air quality index. *Ing. Syst. Inf., 28*(2), 433-441. https://doi.org/10.18280/isi.280219.

Pacheco, A. G. & Krohling, R. A. (2020). The impact of patient clinical information on automated skin cancer detection. *Comput. Biol. Med., 116*, 103545. https://doi.org/10.1016/j.compbiomed.2019.103545.

Razmjooy, N., Ashourian, M., Karimifard, M., Estrela, V. V., Loschi, H. J., Do Nascimento, D., França, R. P., & Vishnevski, M. (2020). Computer-aided diagnosis of skin cancer: A review. *Curr. Med. Imaging, 16*(7), 781-793. https://doi.org/10.2174/1573405616666200129095242.

Sethanan, K., Pitakaso, R., Srichok, T., Khonjun, S., Thannipat, P., Wanram, S., Boonmee, C., Gonwirat, S., Enkvetchakul, P., Kaewta, C., & Nanthasamroeng, N. (2023). Double AMIS-ensemble deep learning for skin cancer classification. *Expert Syst. Appl., 234*, 121047. https://doi.org/10.1016/j.eswa.2023.121047.

Singh, S. K., Abolghasemi, V., & Anisi, M. H. (2023). Fuzzy logic with deep learning for detection of skin cancer. *Appl. Sci., 13*(15), 8927. https://doi.org/10.3390/app13158927.

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

Thapar, P., Rakhra, M., Cazzato, G., & Hossain, M. S. (2022). A novel hybrid deep learning approach for skin lesion segmentation and classification. *J. Healthcare Eng., 2022*, 1709842. https://doi.org/10.1155/2022/1709842.

Thomas, S. M., Lefevre, J. G., Baxter, G., & Hamilton, N. A. (2021). Interprefigure deep learning systems for multi-class segmentation and classification of non-melanoma skin cancer. *Med. Image Anal., 68*, 101915. https://doi.org/10.1016/j.media.2020.101915.

Toğaçar, M., Cömert, Z., & Ergen, B. (2021). Intelligent skin cancer detection applying autoencoder, MobileNetV2 and spiking neural networks. *Chaos, Solitons Fractals, 144*, 110714. https://doi.org/10.1016/j.chaos.2021.110714.

Wang, D., Shelhamer, E., Liu, S., Olshausen, B., & Darrell, T. (2020). Tent: Fully test-time adaptation by entropy minimization. *arXiv preprint,* arXiv:2006.10726. https://doi.org/10.48550/arXiv.2006.10726.

Wei, L., Ding, K., & Hu, H. (2020). Automatic skin cancer detection in dermoscopy images based on ensemble lightweight deep learning network. *IEEE Access, 8*, 99633-99647. https://doi.org/10.1109/ACCESS.2020.2997710.

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

Zhang, N., Cai, Y. X., Wang, Y. Y., Tian, Y. T., Wang, X. L., & Badami, B. (2020). Skin cancer diagnosis based on optimized convolutional neural network. *Artif. Intell. Med., 102*, 101756. https://doi.org/10.1016/j.artmed.2019.101756.