Distributed and Parallel Databases
https://doi.org/10.1007/s10619-021-07360-z

Deep learning-based computer aided diagnosis model for skin cancer detection and classification

Devakishan Adla[1,2] · G. Venkata Rami Reddy[3] · Padmalaya Nayak[4] · G. Karuna[5]

Received: 22 May 2021 / Accepted: 12 August 2021

Abstract

Skin cancer is a commonly occurring disease, which affects people of all age groups. Automated detection of skin cancer is needed to decrease the death rate by identifying the diseases at the initial stage. The visual inspection during the medical examination of skin lesions is a tedious process as the resemblance among the lesions exists. Recently, imaging-based Computer Aided Diagnosis (CAD) model is widely used to screen and detect the skin cancer. This paper is designed with automated Deep Learning with a class attention layer based CAD model for skin lesion detection and classification known as DLCAL-SLDC. The goal of the DLCAL-SLDC model is to detect and classify the different types of skin cancer using dermoscopic images. During image pre-processing, Dull razor approach-based hair removal and average median filtering-based noise removal processes take place. Tsallis entropy based segmentation technique is applied to detect the affected lesion areas in the dermoscopic images. Also, a DLCAL based feature extractor is used for extracting the features from the segmented lesions using Capsule Network (CapsNet) along with CAL and Adagrad optimizer. The CAL layer incorporated into the CapsNet is intended to capture the discriminative class-specific features to cover the class dependencies and effectively bridge the CapsNet for further process. Finally, the classification is carried out by the Swallow Swarm Optimization (SSO) algorithm based Convolutional Sparse Autoencoder (CSAE) known as SSO-CSAE model. The proposed DLCAL-SLDC technique is validated using a benchmark ISIC dataset. The proposed framework has accomplished promising results with 98.50% accuracy, 94.5% sensitivity and 99.1% specificity over the other methods interms of different measures.

Keywords Skin lesion · Dermoscopic images · Image classification · ISIC dataset · Deep learning · Image segmentation · Hair removal · Capsule Networks

✉ Devakishan Adla
it.devakishan@gmail.com

Extended author information available on the last page of the article

Published online: 23 August 2021

1 Introduction

Skin cancer is a lethal kind of cancer across the globe [1, 2]. Various kinds of skin cancer exist globally namely melanoma, basal cell carcinoma, squamous cell carcinoma, intraepithelial carcinoma and so on. The human skin comprises of three layers of tissues namely hypodermis, dermis and epidermis. The epidermis layer contains melanocytes that could generate high melanin at an abnormal rate in certain cases. For example, long-term disclosure of strong ultraviolet radiation from sunlight leads to melanin creation. The uncommon development of melanocytes results in melanoma that is a lethal kind of skin cancer [3]. Earlier diagnosis of melanoma is highly important for proper treatment. If the melanoma is identified at the earlier stage, the survival rate for 5 years is about 92%. The visual comparison between malign and benign skin lesions is the major problem in melanoma detection. Because of this, melanoma diagnostic process is complex even for the experienced professional. It is one of the difficult processes in determining what kind of lesion through naked eye. In recent times, various imaging approaches are utilized with Dermoscopy as part of them.

Dermoscopy is a noninvasive imaging method which involves picturing of skin surface using immersion fluid and light magnifying tool. It is most utilized imaging methods in dermatology that augmented the diagnosis efficiency of malignant cases by 50% [4–6]. Nevertheless, the utilization of human vision for detecting melanoma in skin lesion images may be subjective, imprecise and depends on the experienced dermatologists. To conquer these problems faced in the identification of melanoma, Computer Aided Diagnosis (CAD) system is required for assisting the professionals in diagnostic process. This segmentation phase is a troublesome procedure because of its huge variation in texture, size, color and position of skin lesion images. Moreover, the lower contrast of images prevents us from finding the difference of the nearby tissues. Additional features like ruler marks, air bubbles, color illumination, ebony frames, hair and blood vessels cause additional problems during the lesion segmentation process. Several techniques are employed for segmenting skin lesions using different imaging modalities. Recently, Convolutional Neural Networks (CNN) in Deep Learning (DL) approaches have attained an effective outcome in skin lesion detection [7]. The CNN accepts lower resolution images to reduce the number of estimations and variables in the network [8, 9]. This condition might result in loss of some significant features of the image.

1.1 Major contribution

- This research proposes a novel DL with class attention layer based CAD technique for skin lesion detection and classification known as DLCAL-SLDC.
- The DLCAL-SLDC model involves image pre-processing in two different stages namely, Dull razor approach based hair removal and average median filtering based noise removal.

- Tsallis entropy based segmentation technique is applied for detecting the segment of the infected lesion regions in the dermoscopic images.
- DLCAL model is used which extracts the features from the segmented lesions using Capsule Network (CapsNet) along with CAL and Adagrad optimizer.
- The Swallow Swarm Optimization (SSO) algorithm based Convolutional Sparse Autoencoder (CSAE) knowns as SSO-CSAE model is applied for classification.
- To investigate the effective skin lesion classification performance of the DLCAL-SLDC model, an extensive simulation analysis take place on the benchmark ISIC dataset.

2 Literature review

The recently developed automated DL based skin lesion detection and classification models are available in the literature. Hasan et al. [10] presented an automated and reliable architecture for skin lesion diagnosis called Dermoscopic Expert (DermoExpert). It comprises of pre-processing, Transfer Learning (TL) and hybrid CNN. The presented hybrid CNN classification comprises of three different feature extractors with similar input images that are combined to attain improved depth feature mapping of equivalent lesion. The fused and distinct feature mapping are categorized by distinct Fully Connected (FC) layers that are later ensembled for getting an ultimate predictive possibility. In pre-processing, they utilize class rebalancing, lesion segmentation and augmentation. To boost lesion detection, intensity-based augmentation, class rebalancing the geometric features are used to find the damage of major class and include additional images to the minor class. Acosta et al. [11] considered the early phase that automatically crops the affected area in a skin lesion image by utilizing Mask and Region based CNN method and the next phase depends on ResNet152 structure that categorizes the lesions as “benign” or “malignant”. Seeja and Suresh [12] focused on the design of DL based automated skin lesion segmentation techniques to achieve enhanced classification efficiency. Initially, a CNN based U-net method is utilized for processing segmentation. Later, the features like texture, shape and color gets extracted from the segmented image by Histogram of Oriented Gradients (HOG), Gabor Edge Histogram (GEH) and Local Binary Pattern (LBP) technique. At last, all the features are extracted by these techniques and are fed into KNN, RF, NB and SVM classification for diagnosing the skin image, i.e., melanoma or benign lesion. Li and Shen [13] presented two DL techniques for addressing three major processes carried out in the region of skin lesion image namely lesion dermoscopic feature extraction, lesion classification and lesion segmentation. A DL method comprising of 2 FCRN is projected for producing simultaneous segmentation and coarse classification results. Yap et al. [14] proposed a technique that integrates several imaging modalities along with person metadata to enhance the efficiency of automatic skin lesion diagnosis. It is employed on binary classification function for relating prior investigation and five class classification function depicting a real-time medical setting.

In Khan et al. [15], a fully automatic CAD method is presented depending upon DL architecture where the actual skin lesion images are primarily pre-processed by

the decorrelation formulation method which further permits the resulting image to the Mask-RCNN for segmenting lesions. During this phase, the Mask RCNN module is trained by the segmentation RGB image that is produced by the ground truth image of ISIC2017 and ISIC2016 datasets. Srinivasu et al. [16] presented a technical procedure for categorizing skin disease by DL based LSTM and MobileNet V2. The MobileNet V2 module is an effective method with improved accuracy that could process light weighted computation devices. The presented module is effective in preserving stateful data for accurate prediction. A GLCM is utilized to assess the development of the disease. Gazioğlu and Kamaşak [17] examined the impact of external objects (hairs and ruler) and image quality (contrast, noise, blur) on the melanoma classified images using CNN modules like AlexNet, VGG16, ResNet50 and DenseNet121. They employed data augmentation, trained 4 modules individually and investigated 6 datasets. In [26], a smart home system is presented via in-built sensors with AI techniques for diagnostic skin cancer of the inhabitants of the house. The projected model has been validated and elaborated owing to the significant utilization in real time. In [27], a novel segmentation-based classifier to detect skin cancer is derived by incorporating four main steps, namely pre-processing, grabcut based segmentation, Inception based feature extraction and ANFC based classification. The several models exist in the literature are tedious process because of its huge variances in texture, size, color and position of skin lesion images. Additional features like ruler marks, air bubbles, color illumination, ebony frames, hairs and blood vessels cause's additional problems to the lesion segmentation process. To resolve these issues, this paper presents a new DL model to accomplish enhanced diagnostic outcome.

3 The proposed CAD model

The proposed DLCAL-SLDC model aims to present a new CAD model to detect and classify different types of skin lesions with dermoscopic images. During image pre-processing, Dull razor approach-based hair removal and average median filtering-based noise removal processes take place. Besides, Tsallis entropy-based segmentation technique is employed to detect the affected lesion areas in the dermoscopic images. The block diagram of the DLCAL-SLDC model is shown in Fig. 1. The DLCAL-SLDC model involves distinct stages of operations namely image pre-processing, Tsallis entropy-based image segmentation, DLCAL based feature extraction and SSO-CSAE based classification. The detailed working of every module in the block diagram is discussed in the succeeding sections.

3.1 Image pre-processing

The pre-processing technique comprises of two phases. The first stage is the recognition of lesion position from the image and the next is the process of segmenting lesions. Each data in the trained dataset has been resized to 512×512 solution prior to CapsNet training process. Then, every image in the trained set is labeled based on

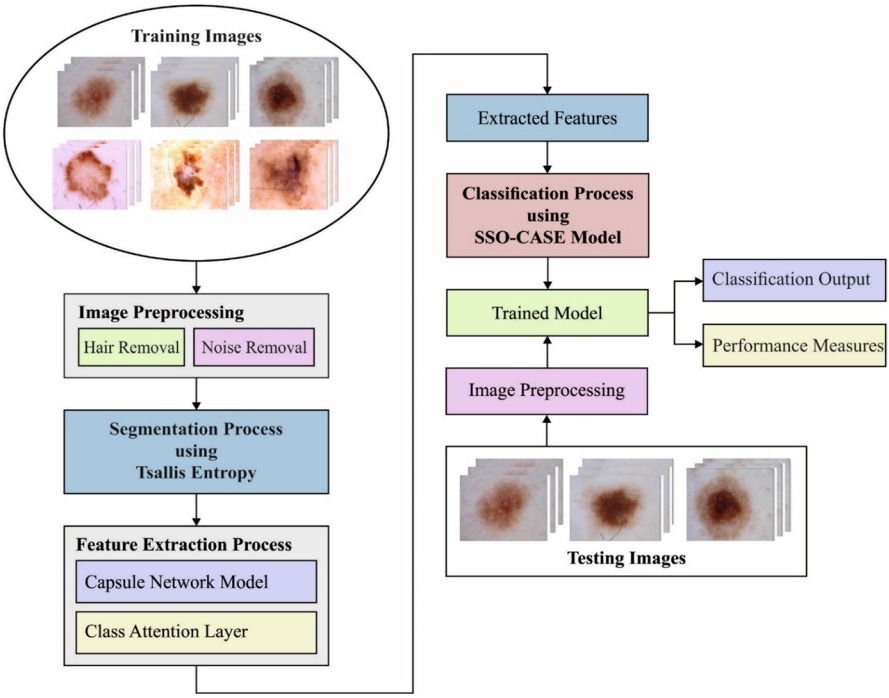

Fig. 1 Workflow of DLCAL-SLDC model

the training requirements of CapsNet and it requires few data regarding images and trained images. This data contains values of intermediate point coordinates (x, y), width (w) and height (h) of bounding box where class description of object should be identified. The upper left corner coordinate points ($\times 1$, y1) and bottom right corner coordinate points $(x2, y2)$ of bounding box are utilized for determining x and y coordinate points. Similarly, hairs have been removed from the dermoscopic images using the DullRazor [18] method for precise detection and segmentation. This technique removes the hair on lesion in three stages. At initial stage, it recognizes the hair positions by grayscale morphological closing function. In next stage, it ensures the hair positions by observing the length and thickness of the identified shapes and later, the ensured pixels are exchanged using bilinear interpolation technique. At last, it smoothens the exchanged pixel using an adaptive median filter.

3.2 Image segmentation

Once the input dermoscopic images are pre-processed, the next step is to segment the lesion areas using Tsallis entropy technique. The Entropic segmentation provides optimal outcome in several conditions that functions well while used for noisy images, where grayscale distribution usually consists of unimodal histogram [19]. The benefit of this technique is the usage of the global and objective function of the

histogram. Assume a grayscale image with L grayscale in the extent of $\{0, 1, 2, (L-1)\}$. Assume $h(i)$ denotes pixel count of the grayscale i for $0 \leq i < (L-1)$ and N denotes pixel count in the image. Thus,

$$N = \sum_{i=0}^{L-1} h(i) and P_i = \frac{h(i)}{N} \tag{1}$$

where P_i represents the likelihood of incidence of intensity level i in the image.

Then, the Tsallis multi-level thresholding is given as follows,

$$f(t) = \text{argmax}\,[S_q^A(t) + S_q^B(t) + \cdots + S_q^M(t) + (1-q) \cdot S_q^A(t) \cdot S_q^B(t) \cdot \cdots \cdot S_q^M(t)] \tag{2}$$

where $S_q^A(t) = \frac{1-\sum_{i=0}^{t_1-1}(\frac{P_i}{P^A})^q}{q-1}$; $P^A = \sum_{i=0}^{t_1-1} P_i$; $S_q^B(t) = \frac{1-\sum_{i=t_1}^{t_2-1}(\frac{P_i}{P^B})^q}{q-1}$; $P^B = \sum_{i=t_1}^{t_2-1} P_i$ and $S_q^M(t) = \frac{1-\sum_{i=t_M}^{L-1}(\frac{P_i}{P^M})^q}{q-1}$; $P^M = \sum_{i=t_M}^{L-1} P_i$

In above formula, $t_1, t_2\, t_M$ denotes threshold level where, $t_1 < t_2 < \cdots < t_M$.

3.3 Feature extraction

Next to image segmentation, feature extraction process is carried out to derive a collection of useful feature vectors from the segmented image. The CapsNet model is used as a feature extractor with the inclusion of CAL and Adagrad optimizer. The Fused Vector (FV) results may contain duplicate or unnecessary characteristics that are explicitly processed with an attribute or variable selection mechanism. Feature selection refers to the whole process of selecting a subset of the highly discriminant parameters. The idea of entropy is used in the proposed work to evaluate uncertain data and reveal the signal unpredictability by displaying the system disorder.

3.3.1 CapsNet architecture

CapsNet model is established for maintaining the object location and their features in the image and modeling the hierarchical relations. In CNN method, useful data in the information arrives in front of the pooling layer. As the information is transferred to the following pooling layer, it might not be probable for the network in learning small details. Figure 2 illustrates the structure of CapsNet Model. Additionally, the CNN method generates a scalar value in neural outcome. The CapsNets produce vector output in similar size however with distinct routing, capsules comprise several neurons. The vector routing denotes the variable of images [20]. The CNN utilizes scalar input activation functions like Tangent, ReLU and Sigmoid. Alternatively, the CapsNet utilizes vector activation function named squashing as defined in Eq. (3).

$$v_j = \frac{||S_j||^2}{1 + ||S_j||^2} \frac{S_j}{||S_j||} \tag{3}$$

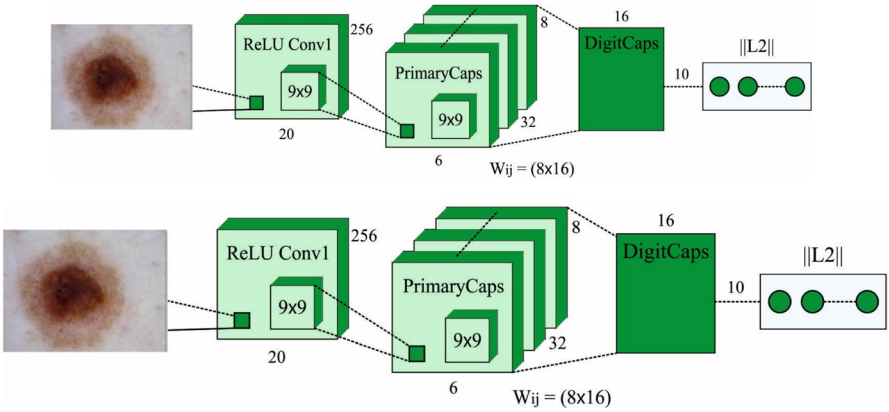

Fig. 2 Architecture of CapsNet model

where v_j denotes capsule output, j and s_j represents overall capsule input. v_j shrink long vector to one when there exists an object in the image and choke shorter vector to zero when there is no object in the image.

Apart from the initial layer of CapsNet, the overall input values of capsule s_j is estimated by the weighted number of the predictive vector ($U_{j|i}$) in the capsule as placed in lower layer. The predictive vector ($U_{j|i}$) is estimated by the product of a capsule from the lower layer with their output (O_i) and weight matrix (W_{ij}).

$$S_{j=}\sum_i b_{ij}u_{j|i} \tag{4}$$

$$u_{j|i=}W_{ij}O_i \tag{5}$$

where b_{ij} denotes coefficient defined by dynamic routing procedure and it is given by,

$$b_{ij} = \frac{\exp\,(a_{ij})}{\sum_k \exp\,(a_{ik})} \tag{6}$$

where a_{ij} indicates log-likelihood. The number of correlation coefficient between capsules i and capsules in top layer is one and log previous likelihood is defined by Softmax. In CapsNet, a margin loss is provided for determining the object of a certain class that exist and is estimated as follows,

$$L_k = T_k \max\,(O, m^+ - \|v_k\|)^2 + \lambda(1 - T_k \max\,(O, \|v_k\| - m^-)^2 \tag{7}$$

The value of T_k is one when class k is present. Besides, $m^+ = 0.9$ and $m^- = 0.1$ represents hyper variables and weight of the loss respectively. The vector length is estimated by the CapsNet that denotes the likelihood in the part of image, where vector direction comprises of the variable data like size, position, color, texture and so on.

3.3.2 Class attention layer mechanism

Though the extracted features from pretrained CapsNet are at higher level and straightaway fed into FC layer to generate multi-labelled prediction, it is impossible to learn the higher order probabilistic dependency by recurrent feeding with similar features. Hence, the extraction of discriminative class wise features acts as a major part in determining class dependency and effectively connecting CapsNet for multi-label classification processes. Therefore, CAL is employed for exploring features in every category. It comprises of two succeeding phases namely (1) creating class attention map with 1×1 convolution layer with stride 1 and (2) vectorising every class attention map to attain class specific features. Basically, for a given feature map X, in the feature extraction model, size is given by $W \times W \times K$ and $1v_l$ denotes the l-th convolutional filter in CAL layer. The attention mapping M_l for class l is obtained from the following equation:

$$M_l = X * w_l, \tag{8}$$

where l extents from one to several classes and $*$ denotes convolution function. Consider the size of convolution filter 1×1 and a class attention mapping M_l is basically a linear integration of whole channel in X. With this implementation, the presented CAL can learn distinct class attention mapping [21]. For instance, initially, a dermoscopic image is fed into the feature extraction model and the output of its convolutional block is assumed as feature mapping X in Eq. (8). It shows that class attention mapping emphasizes on distinct regions for various classification and demonstrates missing class. Consequently, class attention mapping M_l is converted to class wise feature vector v_l of W^2 dimensions by vectorization.

3.3.3 Adagrad optimizer

To optimally adjust the hyperparameters of the CapsNet model, Adam optimizer is applied on it. Adagrad optimizer is a gradient oriented optimization technique which functions well for sparse gradient [22]. The process of selecting a set of ideal hyperparameters for learning algorithm is known as hyperparameter optimization or tuning. A hyperparameter is the value that is used to influence the learning process. Other factor, such as node weight is learnt simultaneously. It would automatically adjust the learning rate depending upon the variables. The fundamental formula utilized for parameter updation is given in Eq. (9)

$$\theta_{t+1} = \theta_t - \frac{\alpha}{\sqrt{\varepsilon + \sum g_t^2}} \odot g_t \tag{9}$$

where θ_t represents the variable at time t, α indicates the learning rate, g_t denotes gradient estimation and $\odot$ represents component wise multiplication.

3.4 Image classification

At the final stage, the SSO-CSAE method is used for performing the classification process where the class labels are allotted in an effective way.

3.4.1 Network structure of CSAE model

The Auto-Encoder (AE) unit includes an encoder as well as decoder. The encoding unit maps the input to hidden representation using deterministic non-linear function and the decoding unit reconstructs the input where the hidden representation is again mapped into the actual input space. For a provided dataset $\mathrm{x}^{(i)} \in \mathcal{R}^d$, $i = 1 \cdots N$, the encoding unit will encode $\mathrm{x}^{(i)}$ by $\mathrm{z}^{(i)} = f\left(\mathrm{W}_e \mathrm{x}^{(i)} + \mathrm{b}\right)$. In this case, $f(\bullet)$ is a component wise non-linear activation function. $\mathrm{W}_e \in \mathcal{R}^{s \times d}$ and $\mathrm{b} \in \mathcal{R}^s$ are the encoded weight matrices and biases, respectively. The code $\mathrm{z}^{(i)}$ is later decoded using $\tilde{\mathrm{x}} = f\left(\mathrm{W}_d \mathrm{z}^{(i)} + \mathrm{c}\right)$, where $\mathrm{W}_d \in \mathcal{R}^{d \times s}$ denotes the weight matrix and $\mathrm{c} \in \mathcal{R}^d$ indicates the bias of the decoding unit. Figure 3 demonstrates the architecture of CSAE.

The AE boosts $\tilde{\mathrm{x}}$ to be closer to $\mathrm{x}^{(i)}$ with a provided distance measure. Every individual row of the weight matrices W_e links to a feature. For preventing the AE from learning disintegrated features, the associated weight matrix $\mathrm{W}_e = \mathrm{W}_d^T$ is promoted. Besides, a regularization term, $(\bullet)$ is executed on the activation of hidden units for achieving features like sparse response [23]. Further, the regularized AE gets trained using the minimization of the loss function, as given in Eq. (10):

$$\begin{aligned} \mathcal{J} &= \min_{W,b,c} \frac{1}{N} \sum_{i=1}^{N} \left\| \tilde{x}^{(i)} - x^{(i)} \right\|_2^2 + \phi\left(\mathrm{z}^{(i)}\right) \\ &= \min \frac{1}{N} \sum_{i=1}^{N} \left\| f\left(\mathrm{W}^T f\left(\mathrm{W}\mathrm{x}^{(i)} + \mathrm{b}\right) + \mathrm{c}\right) - \mathrm{x}^{(i)} \right\|_2^2 + \varphi\left(\mathrm{z}^{(i)}\right). \end{aligned} \tag{10}$$

The classical AE can be extended to design CSAE for handling large sized images. Particularly, for the input image $X \in \mathcal{R}^{m \times n}$, the CSAE encodes it by using the convolution of the bank of filters $\mathrm{W}_e \in \mathcal{R}^{k \times k \times d}$, resultant to d response map.

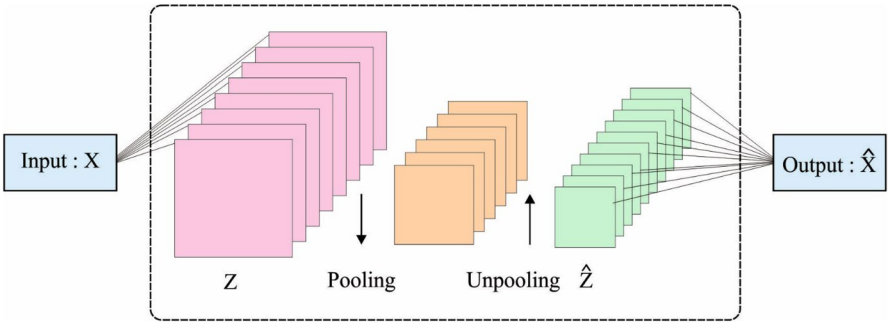

Fig. 3 Structure of CSAE

They are again passed via a component wise nonlinear transformation for generating the respective feature map $Z_i \in \mathcal{R}^{(m-k+1)\times(n-k+1)}$, where $i = 1 \cdots d$. For obtaining sparse feature map, the CSAE sparsify the feature map by retaining the highest value and corresponding location in every local sub-region and next is the unpooling of map with the saved locations.

At last, the decoding unit outputs the reconstruction $\hat{X}$ by convolution of the unpooled sparse feature map with filters bank $W_d \in \mathcal{R}^{k\times k\times d}$. To achieve learning competence, the tied filters are employed in AEs, representing $W_i = W_{ei} = \mathrm{rot}(W_{di}, 180)$ in CSAE. The CSAE training process can be defined as follows.

$$\mathcal{J}(W,\ b,\ c) = \min_{W,b,c} \frac{1}{2} \sum\nolimits_{l=1}^{L} \|\hat{X}^{(l)} - X^{(l)}\|_F^2 \tag{11}$$

where,

$$\hat{X}^{(l)} = \sum\nolimits_{i=1}^{d} \mathrm{rot}(W_i, 180) * \hat{Z}_i^{(l)} + c_i \tag{12}$$

$$\hat{Z}_i^{(l)} = \mathcal{G}_{p,s}\left(Z_i^{(l)}\right) = \mathcal{G}_{p,s}\left(f\left(W_i * X^{(l)} + b_i\right)\right) \tag{13}$$

where $\hat{Z}_i^{(l)}$ denotes ith sparsified feature map, f is the component wise non-linear function, b and c are the encoding and decoding biases respectively. $\mathcal{G}_{p,s}(X)$ implies the sparsifying operator, it initially max-pools X with pooling-size p and stride s and archives the position of the maximum values in every sub-region of X and then it unpools the map with values kept in the stored locations.

3.4.2 Parameter tuning using SSO algorithm

For optimal adjustment of the hyperparameters exist in the CSAE model, the SSO algorithm is employed to increase the classification accuracy. SSO is a metaheuristic approach, mainly depends upon the swarming nature of swallows. In total, eight distinct breeds of swallows are available. They encompass social life, migrates together and determine proper locations to rest, breed and feed. The swarming nature makes them immune to other birds' attacks. A pair of swallows searches a broad region for finding food, whereas in collective living when one swallow identifies the food, soon the other swarms fly towards that region. They are highly intellectual birds. They fly rapidly and there exists robust communication among the members of the group [24]. They use distinct sounds in diverse scenarios (cautioning, seeking help, inviting for food, ready to breed). Initially, in every round, the population is arranged according to objective function values. Later, the succeeding parts are allocated:

1. Head leader (HL) is a particle with optimum value of objective function.
2. The local leader (LL) represents l particles which follows the HL based on the value of objective function
3. Drifting particle is denoted by k with reduced value of objective function

4. Explorer is the other particles.

In present round, HL does not move, performing as beacons for exploring particle, to the search space between adjacent head and LLs. The explorer particle can modify their locations by the following equation:

$$\theta_e(t+1) = \theta_e(t) + V(t+1)$$

$$V(t+1) = VHL(t+1) + VLL(t+1)$$

$$\mathrm{VHL}(t+1) = VHL(t) + rand(0,1)\left(\theta_e^{besi}(t) - \theta_e(t)\right) + rand(0,1)\left(\theta_{HL}(t) - \theta_e(t)\right)$$

$$VLL(t+1) = VLL(t) + rand(0,1)(\theta_e^{best}(t) - \theta_e(t)) + rand(0,1)(\theta_{LL}(t) - \theta_e(t))$$

where θ_e represents explorer location, θ_{HL} indicates HL location, θ_{LL} represents location of LL adjacent to the explorer, θ_e^{besi} indicates optimum location, V, VHL, and VLL indicates velocity vector of particle, velocity vector of particle moves to HL and velocity vector of particle moves to the adjacent LL respectively.

They decide not to choose the variables α_{HL}, β_{HL}, α_{LL} and β_{LL} that are utilized for computing the velocity vectors regarding local and HLs. This study demonstrates that the equivalent process might increase the execution time of this method without substantial enhancement of the outcome. In this process, all variables are fixed to one.

The equation for modifying the aimless particle location is changed due to its actual equation that could affect particles at the boundary of the search space that exceeds. This equation decreases the likelihood of its performance and enables the explorer particle to cause the performance of aimless particles. For changing the aimless particle's location, the following equations are utilized.

$$\theta_O(t+1) = rand(0.5,2).VSS,$$

$$VSS = \frac{\sum_{j=1}^{N-k} \theta_e^j(t)}{N-k},$$

where θ_O represents aimless particle location, θ^j denotes location of jth particle, N indicates overall number of particles from population and k represents aimless particles count. After the end criteria is encountered, this method provides the location of the HL as novel solution.

3.5 Statistical testing using ANOVA & Z test

The Central Limit Theorem claims that, as the number of observations increases, any distribution with finite variance will approach normality (that is, the mean will approach the midpoint of the distribution). In our study and estimation, the samples are assumed to be distributed naturally. Figuring out how people perceive their opinions

and contrasting the binary value distribution to the distribution of positive, negative, and neutral views (only positive & negative). The null hypothesis in this study means the positive and negative numbers are identical where Value = 1 or 'PASS'. Alternative hypothesis H1 describes that it is impossible to compare a positive number and a negative number. If value = 0 or 'FAIL' was implied. The value of P is set to 0.5, indicating the value = 0 or 'FAIL'. These data are used in the Chi-Squared test. According to the hypothesis H0, opinions are equally true. Observed and expected are same where opinions have no value (biased). Expected and seen are different. Expected values are randomly produced simulated data, whereas observed values are the real data. Of the 100 samples in Whole Data, 21 do not pass the test, 21 out of 100 people in Bluebottoms failed the test. The overall sample was found to have a majority likelihood of passing the Chi-squared test. Therefore, we agree to accept the null hypothesis defining zero variation. An analysis of variance (ANOVA) was performed to see if there is a significant difference on the data. The theory goes like this: The null hypothesis, which asserts that opinions are essentially identical, is stated as H0: Opinions are indistinguishable. Hypothesis H1: It is argued that people have diverse beliefs. All the samples were tested. The conclusion: Because F cal = 52.213 > F tab (0.05) (23.142) = 1.357, which is greater and rejects the null hypothesis and accepts the alternative hypothesis.

4 Performance validation

The skin lesion classification efficiency of the presented technique is validated against the benchmark skin lesion ISIC 2017 dataset [25]. The dataset holds images with dimensions of 640*480 pixels. The dataset comprises of images with distinct class labels such as Angioma, Nevus, Lentigo NOS, Solar Lentigo, Melanoma, Seborrheic Keratosis and Basal Cell Carcinoma (BCC) as shown in Fig. 4. The parameter setting of the presented technique is defined here: batch size: 148, number of epochs: 300, learning rate: 0.001, filter size: 32, kernel size: 3, number of hidden feature layer: 128 and L2 normalization term: 0.001. A set of measures used to examine the performance are sensitivity, specificity, accuracy, precision and G-Measure. They are defined as follows.

$$\text{Sensitivity} = \frac{\text{TP}}{\text{TP} + \text{FN}} \tag{14}$$

$$\text{Specificity} = \frac{\text{TN}}{\text{TN} + \text{FP}} \tag{15}$$

$$\text{Accuracy} = \frac{\text{TP} + \text{TN}}{\text{TP} + \text{TN} + \text{FP} + \text{FN}} \tag{16}$$

$$\text{Precision} = \frac{\text{TP}}{\text{TP} + \text{FP}} \tag{17}$$

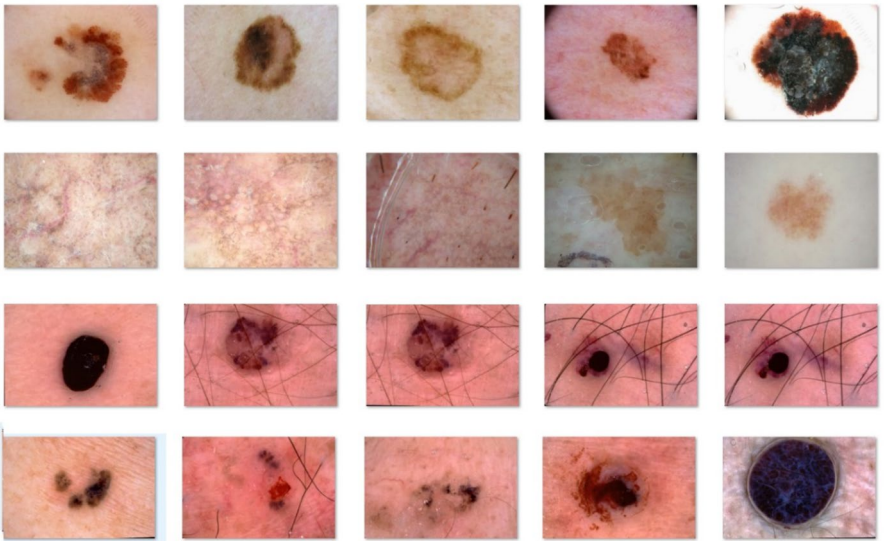

Fig. 4 Sample test images

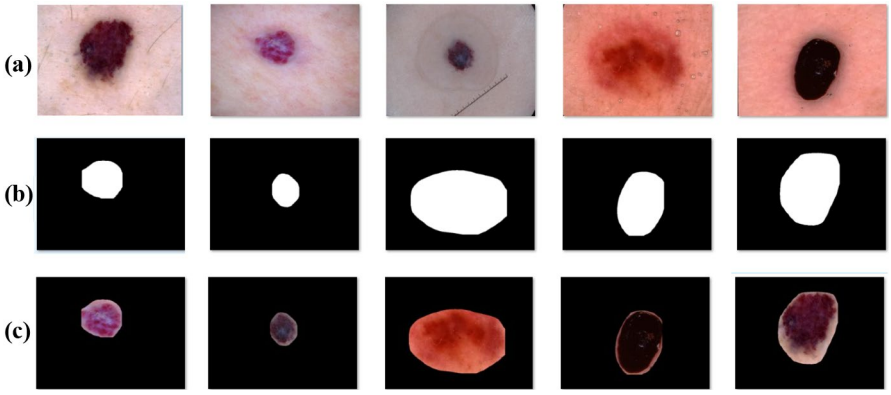

Fig. 5 **a** Original images, **b** masked images, **c** segmented portions

$$\text{G - measure} = \sqrt{\text{Sensitivity} \times \text{Precision}} \tag{18}$$

where, TP, TN, FP and FN denote true positive, true negative, false positive and false negative respectively.

An illustration of the sample results attained by the DLCAL-SLDC model on the test images is presented in Fig. 5. From the figure, it is evident that the DLCAL-SLDC model proficiently masked and segmented the lesion regions from the test dermoscopic images.

The confusion matrix generated by the DLCAL-SLDC model on the classification of skin lesions is given in Fig. 6. The figure portrayed that the DLCAL-SLDC

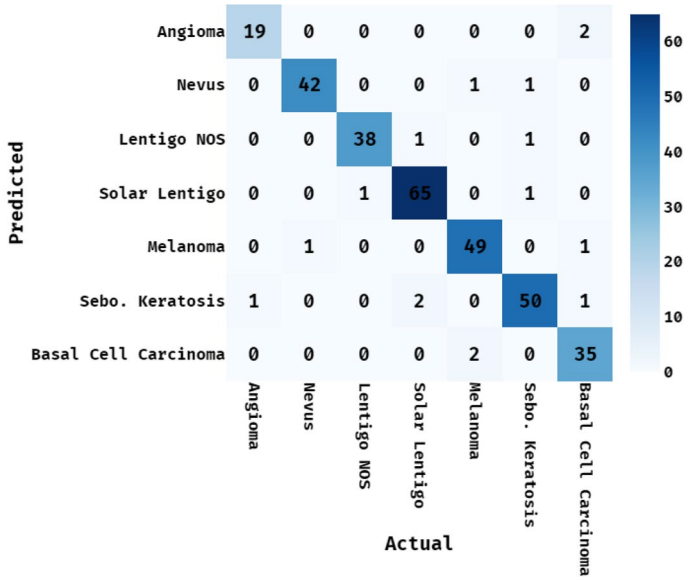

Fig. 6 Confusion matrix for proposed DLCAL-SLDC method

model has effectually categorized a collection of 19 images to Angioma, 42 images to Nevus, 38 images to Lentigo NOS, 65 images to Solar Lentigo, 49 images to Melanoma, 50 images to Seborrheic Keratosis and 35 images to BCC class respectively. These values tells that the DLCAL-SLDC technique has effectually classified the images into distinct class labels.

A detailed skin lesion classification results of the DLCAL-SLDC model is provided in Table 1 and Figs. 7, 8. The table values demonstrates that the DLCAL-SLDC model has obtained improved performance on the applied skin lesion images. From the obtained results, it ensures that the DLCAL-SLDC model has resulted in an effective classification of images under distinct class labels. For instance, the DLCAL-SLDC model has proficiently identified the Angioma class

Table 1 Classification results of DLCAL-SLDC Model under different classes

Different classes	Sensitivity	Specificity	Accuracy	Precision	G-measure
Angioma	0.905	0.997	0.990	0.950	0.927
Nevus	0.955	0.996	0.990	0.977	0.966
Lentigo NOS	0.950	0.996	0.990	0.974	0.962
Solar Lentigo	0.970	0.988	0.984	0.956	0.963
Melanoma	0.961	0.989	0.984	0.942	0.952
Seborrheic Keratosis	0.926	0.989	0.978	0.943	0.935
BCC	0.946	0.986	0.981	0.897	0.921
Average	0.945	0.991	0.985	0.949	0.946

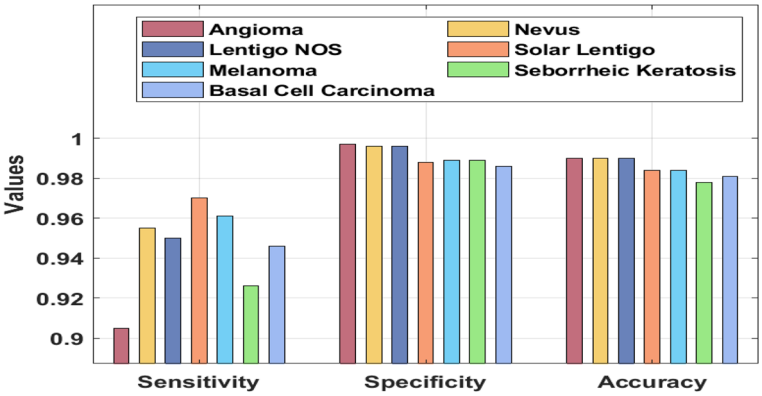

Fig. 7 Result analysis of DLCAL-SLDC model

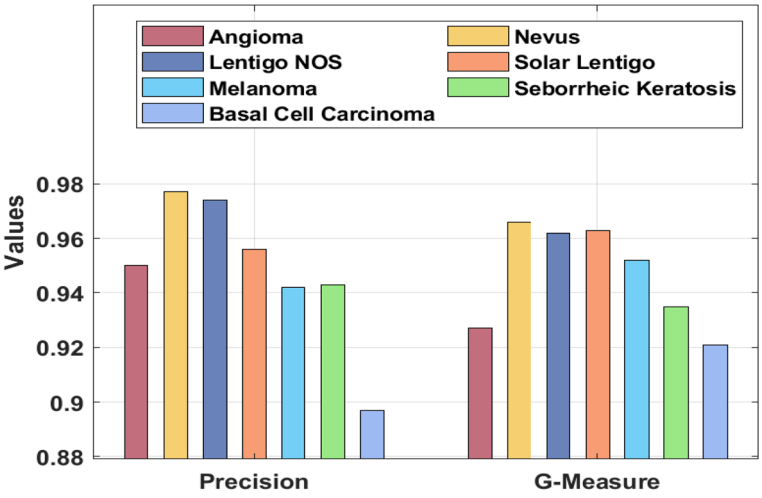

Fig. 8 Precision and G-measure analysis of DLCAL-SLDC model

with 90.5% of sensitivity, 99.7% of specificity, 99.0% of accuracy, 95.0% of precision and 92.7% of G-measure. Followed by, the DLCAL-SLDC model has proficiently identified the Nevus class with 95.5% of sensitivity, 99.6% of specificity, 99.0% of accuracy, 97.7% of precision and 96.6% of G-measure. Eventually, the DLCAL-SLDC model has proficiently identified the Lentigo NOS class with 95.0% of sensitivity, 99.6% of specificity, 99.0% of accuracy, 97.4% of precision and 96.2% of G-measure. Meanwhile, the DLCAL-SLDC model has proficiently identified the Solar Lentigo class with 97.0% of sensitivity, 98.8% of specificity, 98.4% of accuracy, 95.6% of precision and 96.3% of G-measure. Simultaneously, the DLCAL-SLDC model has proficiently identified the Melanoma class with 96.1% of sensitivity, 98.9% of specificity, 98.4% of accuracy, 94.2% of precision

Table 2 Performance of DLCAL-SLDC with state of art models

Methods	Sensitivity	Specificity	Accuracy
Proposed DLCAL-SLDC	94.50	99.10	98.50
DL-ANFC	93.40	98.70	97.90
Conv. NN	81.70	82.90	82.40
SVM	73.20	75.40	74.30
High level features	83.50	81.30	81.10
CDNN	82.50	97.50	93.40
DLN	82.00	97.80	93.20
ResNets	80.20	98.50	93.40
DCCN-GC	90.80	92.70	93.40

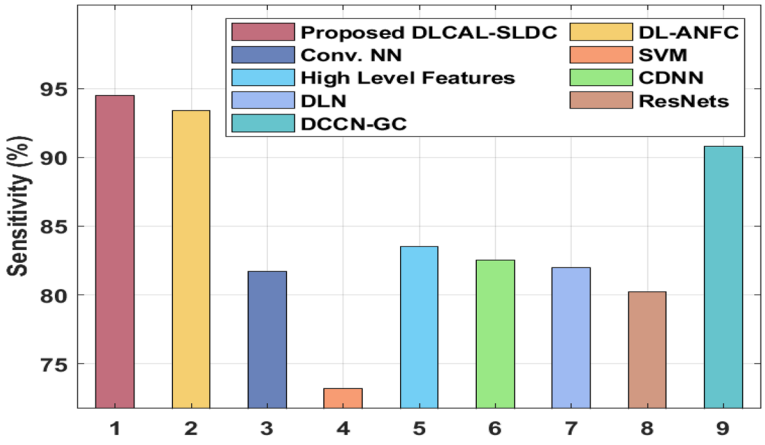

Fig. 9 Comparative analysis of DLCAL-SLDC model interms of sensitivity

and G-measure of 95.2%. Concurrently, the DLCAL-SLDC model has proficiently identified the Seborrheic Keratosis class with the sensitivity of 92.6%, specificity of 98.9%, accuracy of 97.8%, precision of 94.3% and G-measure of 93.5%. At last, the DLCAL-SLDC model has proficiently identified the MCC class with the sensitivity of 94.6%, specificity of 98.6%, accuracy of 98.1%, precision of 89.7% and G-measure of 92.1%.

To guarantee the superior classification performance of the DLCAL-SLDC model, a brief comparative result analysis takes place in Table 2 and Figs. 9, 10, 11 [26, 27]. The results depict that the SVM technique has resulted in ineffective outcomes with a sensitivity of 73.2%, specificity of 75.4% and accuracy of 74.3%.

In Fig. 9, the high-level features model has attained slightly increased outcome with a sensitivity of 83.5%, specificity of 81.3% and accuracy of 81.1%. Followed by the Conv. NN model has demonstrated with improved result such as sensitivity of 81.7%, specificity of 82.9% and accuracy of 82.4%. Besides, the DLN model has accomplished moderate efficiency with a sensitivity of 82%, specificity

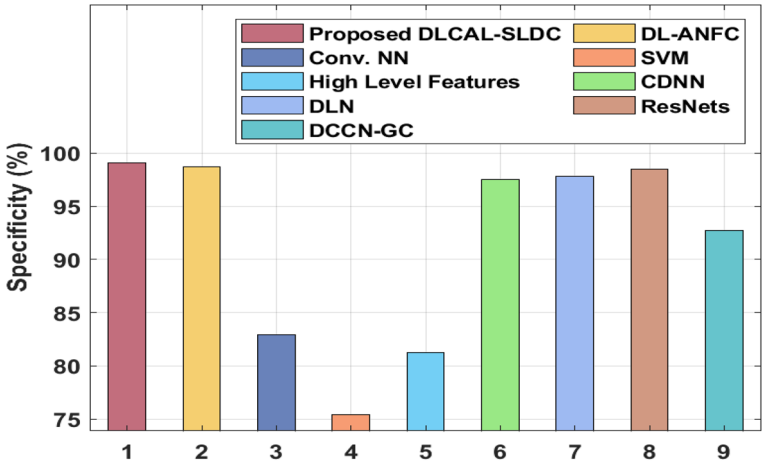

Fig. 10 Comparative analysis of DLCAL-SLDC model interms of specificity

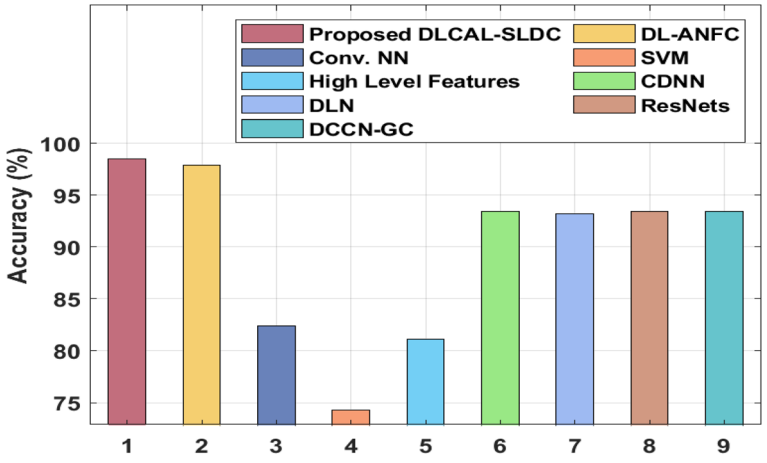

Fig. 11 Comparative accuracy analysis of DLCAL-SLDC model

of 97.8% and accuracy of 93.2%. Finally the proposed DLCAL-SDLC model attained highest sensitivity of 94.50% than existing models.

Figures 10 and 11 shows the CDNN model has attained enhanced outcome with a sensitivity of 82.5%, specificity of 97.5% and accuracy of 93.4%. Furthermore, the ResNets model has tried to attain reasonable outcome with a sensitivity of 80.2%, specificity of 98.5% and accuracy of 93.4%. Next, the DCCN-GC model has demonstrated nearly acceptable outcome with a sensitivity of 90.8%, specificity of 92.7% and accuracy of 93.4%. Eventually, the DL-ANFC model has accomplished an optimal performance with sensitivity of 93.4%, specificity of 98.7% and accuracy of 97.9%. However, the presented DLCAL-SLDC technique has surpassed the

compared techniques by accomplishing a maximum sensitivity of 94.5%, specificity of 99.1% and accuracy of 98.5%.

By observing the aforementioned tables and figures, it is obviously clear that the DLCAL-SLDC technique has exhibited superior outcome over the other models, in recent times. The improved performance of the DLCAL-SLDC model is due to the inclusion of CAL, Adagrad optimizer and parameter tuning. Therefore, it is used as an appropriate tool for detecting and classifying skin lesions using dermoscopic images.

5 Conclusion

A novel DLCAL-SLDC model to detect and categorize the existence of skin lesions from dermoscopic images is developed in this paper. The DLCAL-SLDC model involves encompasses image preprocessing, Tsallis entropy based image segmentation, DLCAL based feature extraction and SSO-CSAE based classification. When the input dermoscopic images are pre-processed, the next step is to segment the lesion areas using Tsallis entropy technique. Then, the CapsNet model is applied as a feature extractor with inclusion of CAL and Adagrad optimizer. Though the extracted features from pre-trained CapsNet are at higher level and straightaway fed into the FC layer to generate multi-labelled prediction, it is impossible to learn the higher order probabilistic dependency by recurrent feeding with similar features. Hence, the CAL is utilized in CapsNet model for exploring features regarding every category. Finally, the SSO-CSAE model is used for performing the classification process where the class labels are allotted in an effectual way. The presented model attains maximum outcomes due to the inclusion of CAL, Adam optimizer and SSO based parameter optimization of the CSAE model. The proposed framework has accomplished promising results with 98.50% of accuracy, 94.5% of sensitivity and 99.1% of specificity over existing models. In future, Skin cancer at early-stage detection techniques can be implemented using machine learning algorithm with optimization model and the performance of the proposed model can also be tested on a large scale dataset and IoT enabled environment.

Declarations

Conflict of interest The authors declared that they have no conflict of interest.

References

1. Ünver, H.M., Ayan, E.: Skin lesion segmentation in dermoscopic images with combination of YOLO and grabcut algorithm. Diagnostics **9**(3), 72 (2019)
2. Karimkhani, C., Green, A., Nijsten, T., Weinstock, M., Dellavalle, R., Naghavi, M., Fitzmaurice, C.: The global burden of melanoma: results from the Global Burden of Disease Study 2015. Br. J. Dermatol. **177**, 134–140 (2017)

3. Jemal, A., Siegel, R., Ward, E., Hao, Y., Xu, J., Thun, M.J.: Cancer statistics, 2019. CA Cancer J. Clin. **69**, 7–34 (2019)
4. Neelakandan, S., Paulraj, D.: A gradient boosted decision tree-based sentiment classification of twitter data. Int. J. Wavelets Multiresolut. Inf. Process. (2020). https://doi.org/10.1142/S0219691320500277
5. Paulraj, D.: An automated exploring and learning model for data prediction using balanced CA-Svm. J. Ambient Intell. Humaniz. Comput. (2020). https://doi.org/10.1007/s12652-020-01937-9
6. Kamalraj, R., Neelakandan, S., Ranjith Kumar, M., Chandra Shekhar Rao, V., Anand, R., Singh, H.: Interpretable filter based convolutional neural network (IF-CNN) for glucose prediction and classification using PD-SS algorithm. Measurement (2021). https://doi.org/10.1016/j.measurement.2021.109804
7. Okur, E., Turkan, M.: A survey on automated melanoma detection. Eng. Appl. Artif. Intell. **73**, 50–67 (2018)
8. Satpathy, S., Das, S., Debbarma, S.: A new healthcare diagnosis system using an IoT-based fuzzy classifier with FPGA. J. Supercomput. **76**(8), 5849–5861 (2020). https://doi.org/10.1007/s11227-019-03013-2
9. Berlin, M.A., Tripathi, S., et al.: IoT-based traffic prediction and traffic signal control system for smart city. Soft. Comput. (2021). https://doi.org/10.1007/s00500-021-05896-x
10. Hasan, M.K., Elahi, M.T.E., Alam, M.A., Jawad, M.T.: DermoExpert: skin lesion classification using a hybrid convolutional neural network through segmentation, transfer learning and augmentation. medRxiv (2021). https://doi.org/10.1101/2021.02.02.21251038
11. Acosta, M.F.J., Tovar, L.Y.C., Garcia-Zapirain, M.B., Percybrooks, W.S.: Melanoma diagnosis using deep learning techniques on dermatoscopic images. BMC Med. Imaging **21**(1), 1–11 (2021)
12. Seeja, R.D., Suresh, A.: Deep learning based skin lesion segmentation and classification of melanoma using support vector machine (SVM). Asian Pac. J. Cancer Prev.: APJCP **20**(5), 1555 (2019)
13. Li, Y., Shen, L.: Skin lesion analysis towards melanoma detection using deep learning network. Sensors **18**(2), 556 (2018)
14. Yap, J., Yolland, W., Tschandl, P.: Multimodal skin lesion classification using deep learning. Exp. Dermatol. **27**(11), 1261–1267 (2018)
15. Khan, M.A., Akram, T., Zhang, Y.D., Sharif, M.: Attributes based skin lesion detection and recognition: a mask RCNN and transfer learning-based deep learning framework. Pattern Recogn. Lett. **143**, 58–66 (2021)
16. Srinivasu, P.N., SivaSai, J.G., Ijaz, M.F., Bhoi, A.K., Kim, W., Kang, J.J.: Classification of skin disease using deep learning neural networks with MobileNet V2 and LSTM. Sensors **21**(8), 2852 (2021)
17. Gazioğlu, B.S.A., Kamaşak, M.E.: Effects of objects and image quality on melanoma classification using deep neural networks. Biomed. Signal Process. Control **67**, 102530 (2021)
18. Lee, T., Ng, V., Gallagher, R., Coldman, A., McLean, D.: Dullrazor®: a software approach to hair removal from images. Comput. Boil. Med. **27**, 533–543 (1997)
19. Manikantan, K., Arun, B.V., Yaradoni, D.K.S.: Optimal multilevel thresholds based on Tsallis entropy method using golden ratio particle swarm optimization for improved image segmentation. Procedia Eng. **30**, 364–371 (2012)
20. Toraman, S., Alakus, T.B., Turkoglu, I.: Convolutional CapsNet: a novel artificial neural network approach to detect COVID-19 disease from X-ray images using capsule networks. Chaos Solitons Fract. **140**, 110122 (2020)
21. Hua, Y., Mou, L., Zhu, X.X.: Recurrently exploring class-wise attention in a hybrid convolutional and bidirectional LSTM network for multi-label aerial image classification. ISPRS J. Photogramm. Remote. Sens. **149**, 188–199 (2019)
22. Saravanan, S., Hailu, M., Gouse, G.M., Lavanya, M., Vijaysai, R.: Optimized secure scan flip flop to thwart side channel attack in crypto-chip. Adv. Sci. Technol. (2019). https://doi.org/10.1007/978-3-030-15357-1_34
23. Luo, W., Li, J., Yang, J., Xu, W., Zhang, J.: Convolutional sparse autoencoders for image classification. IEEE Trans. Neural Netw. Learn. Syst. **29**(7), 3289–3294 (2017)
24. Hodashinsky, I., Sarin, K., Shelupanov, A., Slezkin, A.: Feature selection based on swallow swarm optimization for fuzzy classification. Symmetry **11**(11), 1423 (2019)
25. Subbulakshmi, P.: Mitigating eavesdropping by using fuzzy based MDPOP-Q learning approach and multilevel Stackelberg game theoretic approach in wireless CRN. Cogn. Syst. Res. **52**, 853–861 (2018). https://doi.org/10.1016/j.cogsys.2018.09.021

26. Połap, D., Winnicka, A., Serwata, K., Kęsik, K., Woźniak, M.: An intelligent system for monitoring skin diseases. Sensors **18**(8), 2552 (2018)
27. Sikkandar, M.Y., Alrasheadi, B.A., Prakash, N.B., Hemalakshmi, G.R., Mohanarathinam, A., Shankar, K.: Deep learning based an automated skin lesion segmentation and intelligent classification model. J. Ambient Intell. Humaniz. Comput. (2020). https://doi.org/10.1007/s12652-020-02537-3

Authors and Affiliations

Devakishan Adla[1,2] · G. Venkata Rami Reddy[3] · Padmalaya Nayak[4] · G. Karuna[5]

1 JNTUH, Hyderabad, India

2 Department of IT, VJIT, Hyderabad, India

3 School of Information Technology, JNTUH, Hyderabad, India

4 Department of CSE, GRIET, Hyderabad, India

5 Department of CSE, GRIET, JNTUH, Hyderabad, India