

## Automated Deep Learning Based Melanoma Detection and Classification Using Biomedical Dermoscopic Images

Amani Abdulrahman Albraikan<sup>1</sup>, Nadhem NEMRI<sup>2</sup>, Mimouna Abdullah Alkhonaini<sup>3</sup>,  
Anwer Mustafa Hilal<sup>4,\*</sup>, Ishfaq Yaseen<sup>4</sup> and Abdelwahed Motwakel<sup>4</sup>

<sup>1</sup>Department of Computer Sciences, College of Computer and Information Sciences, Princess Nourah bint Abdulrahman University, P.O. Box 84428, Riyadh, 11671, Saudi Arabia

<sup>2</sup>Department of Information Systems, College of Science & Art at Mahayil, King Khalid University, Saudi Arabia

<sup>3</sup>Department of Computer Science, College of Computer and Information Sciences, Prince Sultan University, Saudi Arabia

<sup>4</sup>Department of Computer and Self Development, Preparatory Year Deanship, Prince Sattam bin Abdulaziz University, AlKharj, Saudi Arabia

\*Corresponding Author: Anwer Mustafa Hilal. Email: a.hilal@psau.edu.sa

Received: 23 December 2021; Accepted: 24 January 2022

**Abstract:** Melanoma remains a serious illness which is a common form of skin cancer. Since the earlier detection of melanoma reduces the mortality rate, it is essential to design reliable and automated disease diagnosis model using dermoscopic images. The recent advances in deep learning (DL) models find useful to examine the medical image and make proper decisions. In this study, an automated deep learning based melanoma detection and classification (ADL-MDC) model is presented. The goal of the ADL-MDC technique is to examine the dermoscopic images to determine the existence of melanoma. The ADL-MDC technique performs contrast enhancement and data augmentation at the initial stage. Besides, the k-means clustering technique is applied for the image segmentation process. In addition, Adagrad optimizer based Capsule Network (CapsNet) model is derived for effective feature extraction process. Lastly, crow search optimization (CSO) algorithm with sparse autoencoder (SAE) model is utilized for the melanoma classification process. The exploitation of the Adagrad and CSO algorithm helps to properly accomplish improved performance. A wide range of simulation analyses is carried out on benchmark datasets and the results are inspected under several aspects. The simulation results reported the enhanced performance of the ADL-MDC technique over the recent approaches.

**Keywords:** Biomedical images; dermoscopic images; deep learning; melanoma detection; machine learning

### 1 Introduction

Computation analysis for gathering the skin lesion information is the field of study is important due to its treatment planning and cancer diagnosis. Melanocytes are melanin-producing cells that



This work is licensed under a Creative Commons Attribution 4.0 International License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

primary objective is to offer color to eyes, skin, and hair [1]. Skin cancer is one of the common types of cancer [2], which is categorized into melanoma and non-melanoma. The considerable healthcare cost and higher morbidity related to the malignant types of lesion have inspired authors to propose more flexible and accurate models for earlier melanoma [3]. While abnormal melanocyte cells divide out of control and affect other parts of the body and it became high metastases and increases deaths rate. As per the information reported by American Cancer Society 2018, the rate of skin cancer has been increased over the past few years. In 2018, nearly 91,270 new melanoma cases would be detected and 9320 people are predicted to die of skin cancer in the US. But the skin cancer is extremely curable when it is detected at earlier stages.

Effective non-invasive diagnosis method assists dermatologists to gather information, providing insight into melanoma shape and structure. Macroscopic images, commonly called clinical images or non-dermoscopic since these images are taken by the traditional digital cameras that are non-invasive and inexpensive—utilized in computation analysis for patients of distinct ethnicities and ages [4,5]. Dermoscopic images are attained by dermoscopy, a special non-invasive imaging instrument, commonly have good invariance to contrast and illumination [6]. But, distinct image acquisition parameters and conditions like non-ideal illumination, image quality limits, variable distance, and poor resolutions like artefacts, noises, reflections, shadows, and hairs reduce the image quality which can compromise the strength of the melanoma analysis. Thus, a computer aided diagnosis (CAD) scheme is capable of analyzing dermoscopy-based images might be considered for assisting dermatologists in skin cancer diagnoses [7].

Recently, neural network has once again boomed in a wide-ranging of applications. By attaining reasonable outcomes, it becomes a promising tool for skin lesion classification. A convolutional neural network (CNN) is depending on the convolutional process of the image with various filters, offering varied and extensive data. Initially, CNN gets the edge and line features, whereas, in the deeper layer, they find more detailed information in the image. It is familiar that the convolutional process isn't invariant for geometric transformations such as scale, translation, and rotation [8]. For this reason, neural network requires a massive amount of training images for extracting each possible way that the image can have, so their higher computation cost. Additionally, images of skin lesions could have lower contrasts among the healthy skin and lesion that can interfere with the accurate segmentation of the lesion [9]. That might result in loss of data that is utilized for accurate classification of the injury.

This paper presents an automated deep learning based melanoma detection and classification (ADL-MDC) model. The ADL-MDC technique performs contrast enhancement and data augmentation at the initial stage. Followed by, the k-means clustering technique is applied for the image segmentation process. Next, Adagrad optimizer based Capsule Network (CapsNet) model is derived for effective feature extraction process. Finally, crow search optimization (CSO) algorithm with sparse autoencoder (SAE) model is utilized for the melanoma classification process. The exploitation of the Adagrad and CSO algorithm helps to properly accomplish improved performance. A wide range of simulation analyses is carried out on benchmark datasets and the results are inspected under several aspects.

## 2 Related Works

Daghrir et al. [10] established a hybrid technique for melanoma skin cancer recognition which is utilized for examining some suspicious lesions. The presented method depends on the forecast of 3 distinct approaches: A CNN and 2 typical machine learning (ML) techniques trained with group of features relating the border, texture, and color of skin lesion. These techniques are next related to

enhancing its efficiency utilizing popular voting. Kassani et al. [11] estimated the efficiency of many recent CNNs from dermoscopic images of skin lesions. The experiment was demonstrated on graphics processing unit (GPU) to speed up the trained and utilization procedure. For enhancing the quality of image, it can utilize distinct pre-processed stages. It also implements data augmentation techniques like horizontal as well as vertical flipping approaches for addressing the class skewness issue.

Thiyaneswaran et al. [12] purposed for detecting the symptom of melanoma primary with images of moles attained in image processing devices and classifying the varieties. The intensity of image was downsampled for decreasing the bit depth. When the illumination of downsampled image is not uniform, afterward gamma corrected was implemented for obtaining the uniform illumination. A textural feature was removed in the segmentation image. In ML approach was implemented for classifying the melanoma image. In [13], 2 approaches were planned for classifying melanoma cancer stage. A primary method classifies melanoma as stage 1 as well as stage 2. Secondary methods classify melanoma as stage 1, stage 2, or stage 3 melanoma. The presented technique utilizes CNN technique with Similarity Measure for Text Processing (SMTP) as loss functions.

In [14], a comparative analysis to classify the set of melanoma utilizing the supervised ML technique. The group of melanoma in dermoscopic data was presented for helping the medical consumption of dermatoscopy imaging techniques skin sores classification. The melanoma has been segmentation in the background utilizing adaptive k-means clustering technique with 2 clusters then feature extraction techniques were dependent upon intensity and texture feature in the segmentation data that is succeeded by trained of classifier and eventually testing on unknown dermoscopic data. Ali et al. [15] examined an automated technique for skin lesion border irregularity recognition. This technique contains removing the skin lesion in the image, identifying the skin lesion border, calculating the border irregularity, trained a CNN and Gaussian naive Bayes (GNB) ensemble, to automatic recognition of border irregularity that outcomes in an objective decision if the skin lesion border was regarded as regular/irregular. ALEnezi [16] presented an image processing based technique for identifying skin disease. This technique gets the digital image of disease affects skin area, afterward utilize image analysis for identifying the kind of diseases. This technique work on input of color image. Afterward, resize of image for extracting features utilizing pre-trained CNN. Then the classified feature utilizes Multiclass support vector machine (SVM).

### **3 The Proposed Model**

In this study, a novel ADL-MDC technique has been derived for melanoma detection and classification using dermoscopic images. The ADL-MDC technique encompasses several stages including pre-processing, K-means segmentation, CapsNet based feature extraction, Adagrad optimizer based hyperparameter tuning, SAE based classification, and CSO based parameter optimization. The exploitation of the Adagrad and CSO algorithm helps to properly accomplish improved performance.

#### **3.1 Image Pre-processing**

At the initial stage, the pre-processing of the dermoscopic images takes place using two ways namely contrast enhancement and data augmentation. In this study, Contrast Limited Adaptive Histogram Equalization (CLAHE) is used for image equalization [17], which processes on smaller areas in the image instead of the whole image. Besides, the adjacent tiles are integrated by the use of bilinear interpolation approach for the removal of artificial boundaries. It is mainly applied to boost the contrast level of the image. Next, data augmentation takes place in two ways as given below.

- Flipping: flipping the image vertically or horizontally
- Rotation: rotates the image by a specified degree.

### 3.2 Image Segmentation

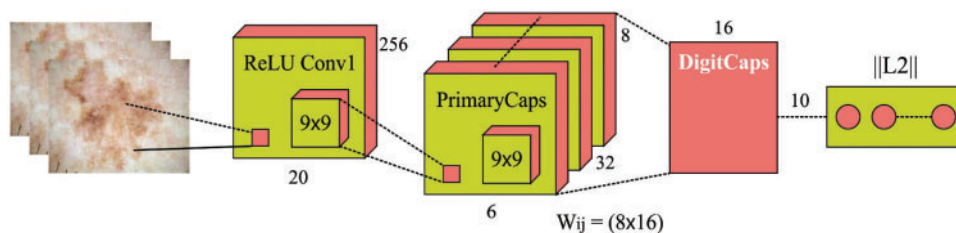
Next, the K-means algorithm is used to segment the dermoscopic images and thereby determine the infected regions. Flow level clustering needs effectively subdividing gathered flows for all applications into groups according to the transferred NetFlow attribute. The most leading unsupervised clustering methods are the  $k$ -means clustering method chosen on another approach like hierarchical clustering, because of its improved computation efficacy [18].  $k$ -means minimize a provided amount of vectors through selecting  $k$  random vector as initial cluster center and assign all vectors to clusters as defined as a distance metrics compared to the cluster centers (a squared error function) as shown in Eq. (1). Then, Cluster center is recalculated as the average (or mean) of the cluster member. This process repeats continuously, terminating whether the cluster converges or a certain amount of processes can be defined as:

$$J = \sum_{j=1}^k \sum_{i=1}^n \|x_i^j - c_j\|^2 \quad (1)$$

where  $c_j$  represent cluster centers,  $n$  equal the sizes of sample space (gathered flow), as well as  $k$  represents the selected values for amount of single cluster (flow class). Therefore, with  $k$ -means,  $n$  flow is separated as  $k$  class.  $k$  value is very important since it influences directly the amount of flow class that affects overfitting.

### 3.3 Feature Extraction

During feature extraction process, the CapsNet model is derived for the useful extraction of the images. The CNN is widely employed technique for 2D object classifications. But data for example pose and position in the object is removed by the CNN because of its data routing process. In order to compensate for the limitations of CNN, a network structure named the capsule network has been introduced [19]. The activation neuron represents the characteristics of components in the object. All the capsules are accountable for describing an individual component in the object, and each capsule collectively describes the general architecture of the object. Unlike some DL models, this architecture preserves objects component and spatial data. Like CNN, a CapsNet is made up of a multilayer network. Fig. 1 demonstrates the architecture of a capsule and how data is routed from the low-level capsule to the high-level capsule.



**Figure 1:** Structure of CapsNet model

The inputs and outputs of the capsule are vectors. The length of the output  $u_j$  signifies the likelihood of presence of its respective components, and the direction of the vector  $u_i$  encodes several features (for example., position, size) of its respective components. The prediction vector  $\hat{u}$  denotes

the belief, that encodes the relationships amongst the  $i$ -th capsules in the low-level capsule and  $j$ -th capsules in the high-level capsule with a linear transformation matrix  $W_{ij}$ :

$$\hat{u}_{j|i} = W_{ij} \cdot u_i \quad (2)$$

Specifically, the pose data and detected presence of components are utilized for predicting the pose data and whole existence. In the training model, the network learns to alter the conversion matrix of the capsule pair by the respective relationships among the whole in the object and components [20]. In high level capsule,  $s_j$  and  $v_j$  indicates the input and output of capsule  $j$ , correspondingly  $s_j$  indicates the sum values of the predictive vectors  $\hat{u}_{j|i}$  with respective weight  $c_{ij}$  in lower-level capsules  $i$ . In Eq. (3),  $c_{ij}$  shows the coupling coefficient and is described by an iterative dynamic routing model, whereas  $\sum_j c_{ij} = 1$  and  $c_{ij} \geq 0$ . If  $c_{ij} = 0$ , there is no data transmission among capsule  $i$  and  $j$ , but if  $c_{ij} = 1$ , each data of capsule  $i$  is transferred to the higher-level capsules  $j$ . The squash function can be expressed in Eq. (4):

$$s_j = \sum_i c_{ij} \cdot \hat{u}_{j|i} \quad (3)$$

$$v_j = \frac{\|s_j^2\|}{1 + \|s_j^2\|} \frac{s_j}{\|s_j\|} \quad (4)$$

$$c_{ij} = \frac{\exp(b_{ij})}{\sum_k \exp(b_{ik})} b_{ij} \leftarrow b_{ij} + \hat{u}_{j|i} \cdot v_j \quad (5)$$

If the low- and high-level capsules are reliable with the prediction, the values of  $c_{ij}$  become large by Eq. (5), and get small if they are unreliable. By altering the routing coefficients, the dynamic routing method guarantees that the low-level capsule transmits the predictive vector to the high-level capsule which is reliable with the prediction, thus the output of the sub-capsule is transmitted to the accurate parent capsule.

To properly adjust the hyperparameters of the CapsNet model, the Adagrad optimizer is applied. During the Adagrad optimizing process, the gradient and accumulated squared gradient to all the variables are calculated at the round [20]:

$$G_t = \sum_{\tau=1}^t g_\tau \odot g_\tau, \quad (6)$$

where  $\odot$  implies the elementwise multiplication and  $g_\tau \in \mathbb{R}^{|\theta|}$  stands for the gradient of existing variables at  $\tau$  round. The variable from the Adagrad is upgraded in Eq. (7):

$$\Delta\theta_t = -\frac{\alpha}{\sqrt{G_t + \varepsilon}} \odot g_\tau \quad (7)$$

where  $\alpha$  defines the rate of learning and  $\varepsilon$  represents the smoothing modules that remove the division by 0. As the rate of learning was set earlier to trained method, Eq. (8) is represented as:

$$\Delta\theta_t = -\alpha \left( \frac{1}{\sqrt{G_t + \varepsilon}} \odot g_\tau \right) \quad (8)$$

It can be recognized that  $G_t$  refers the calculation of previous gradient and  $g'_t$  is signified as:

$$g'_t = \frac{1}{\sqrt{G_t + \varepsilon}} \odot g_\tau \quad (9)$$

So, the Adagrad is upgraded in Eq. (10):

$$\Delta\theta_i = -\alpha g'_i, \quad (10)$$

It can be same as the upgrading method of the typical gradient descent (GD). Thus, Adagrad optimizing is regarded as hyperparameter tuning utilizing the gradient.

### 3.4 Image Classification

At the final stage, the SAE model is applied to classify the dermoscopic images into various class labels. AE is a feed forward neural network (NN), comprises input, hidden, and output layers. It receives the input  $x$  and performs encoding process to a latent representation  $h$  by the use of encoding unit. The final layer includes the identical node count as the input layer and the input can be retrieved at the final layer. The issue of trivial identity mapping can be resolved by the use of sparsity constraints. The AEs including the sparsity regularization constraints to the hidden units are termed as SAE [21]. The neuron can be treated as firing when the outcome is closer to 2 and it is considered inactive when the outcome lies in the range of [42, 43]. Another limitation is included and therefore, the outcome can come nearer to 0 in several cases. Consider the average activation of hidden layer  $j$  be

$$\hat{\rho}_j = \frac{1}{k} \sum_{i=1}^k h_j \quad (11)$$

The approximation takes place  $\hat{\rho}_j = \rho$  where  $\rho$  is the sparsity proportion variable indicating a small integer value, roughly equals to 0. For imposing the constraints, a penalty term is included using the Kullback-Leibler (KL) divergence that can be defined by  $\sum_{j=1}^k KL(\rho||\hat{\rho}_j)$  where

$$KL(\rho||\hat{\rho}_j) = \rho \log \frac{\rho}{\hat{\rho}_j} + (1 - \rho) \log \frac{1 - \rho}{1 - \hat{\rho}_j} \quad (12)$$

Through the application of KL divergence term, the cost function can be derived as follows.

$$J_{SAE} = \frac{1}{2} \sum_{i=1}^N \|x_j - \hat{z}_i\|^2 + \frac{\lambda}{2} (\|W\|^2 + \|W'\|^2) + \beta \sum_{j=1}^N KLN(\rho||\hat{\rho}_j) \quad (13)$$

where  $\beta$  denotes the weight of the sparsity regularization component. The cost function minimization can result in the optimization of variables such as  $W$ ,  $W'$ ,  $b$  and  $b'$ .

For determining the parameters involved in the SAE model, the CSO algorithm is applied. CSO algorithm is a novel search approach stimulated by the behaviour of crows stealing and hiding food [22]. It is non-greedy and could increase the generated solution. The principle of CSO algorithm is given in the following: (1) with some possibility, the crows keep their food from being stolen by randomly flying and aware of being tracked, (2) crows life in the form of flock, and (3) all the crows could remember the location of their hiding place and steal food from another crow. The location of all the crows represent a solution  $x'_i$ . In the iteration, crow  $i$  tracks a random crow  $j$ . When crow  $j$  isn't aware of being traced (that is.,  $r_j \geq AP$ ), crow  $i$  approaches the hiding place  $pb'_j$  of crow  $j$ ; if the  $j$  crows know that  $i$  crow is tracing it, crow  $j$  fools  $i$  crow by going to random place in the searching space to keep its hiding places from being detected [23].

$$\begin{cases} \{x_i^{t+1} = x_i^t + r_3 \cdot fl.(pb'_j - x_i^t), & r_j \geq AP \\ a \text{ random position}, & \text{else} \end{cases} \quad (14)$$

Whereas  $r_j$  and  $r_3$  represent arbitrary values with uniform distribution in  $[0, 1]$ ,  $fl$  indicates the length of a crow flight,  $AP$  indicates the perceptual possibility of crow  $j$ , and  $pb_j^t$  denotes the position in which the existing crow  $j$  stores food, i.e., equal to the past optimum solution. CSO algorithm is a population-based optimization method i.e., quite simple with two adaptable variables (perceived probability  $AP$  and flight length  $fl$ ), that sequentially make it attractive for engineering applications. In CSO algorithm, the parameter of perceived probability is used directly for controlling the diversity of the method. In comparison to other metaheuristics, CSO algorithm is easy to execute and has less variables to alter. Furthermore, the individuals in the CSO algorithm have the probability of reaching an overall random location, and therefore, it has strong capability of exploring unknown regions. But the CSO algorithm has insufficient conditions to choose the terminus and the selection is arbitrarily made among crows; as well, the flight length is a constant value. This characteristic of CSO algorithm results in a weak capacity of exploitation of existing data in comparison to PSO leads to premature convergence, lower searching accuracy, and higher chance of getting into the local optimal, particularly for multi-dimension optimization problems.

#### 4 Experimental Validation

This section investigates the experimental validation of the ADL-MDC technique using ISIC dataset. The dataset comprises various images under distinct classes [24,25]. Fig. 2 shows the sample test images. The dataset contains different classes of skin lesion (SL) such as Actinic keratosis (AK-SL), Basal Cell Carcinoma (BCC-SL), Seborrheic Keratosis (SCC-SL), Melanoma (MEL-SL), Nevus (NEV-SL), BKL-SL, and Vascular lesion (VASC-SL).

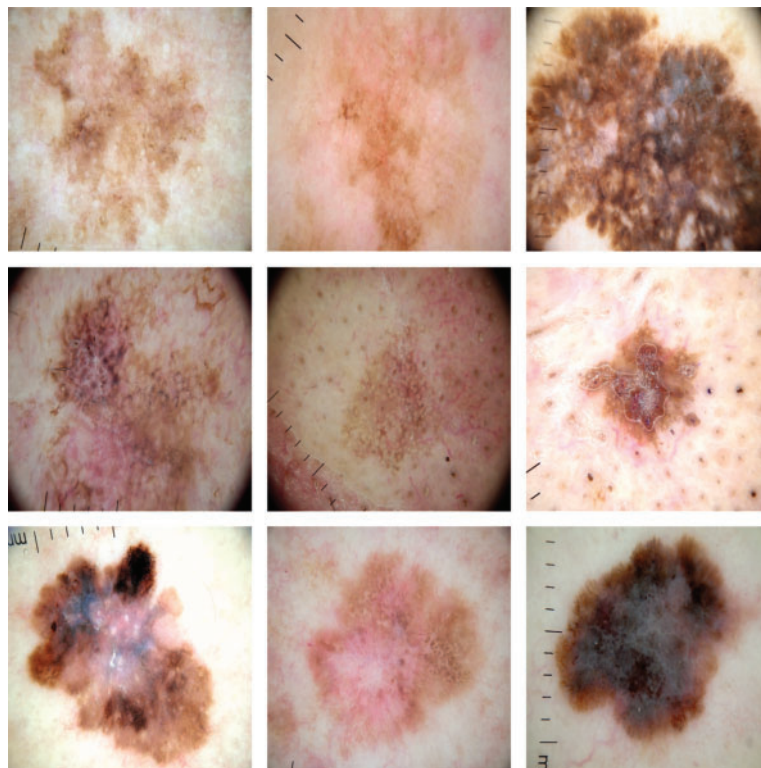
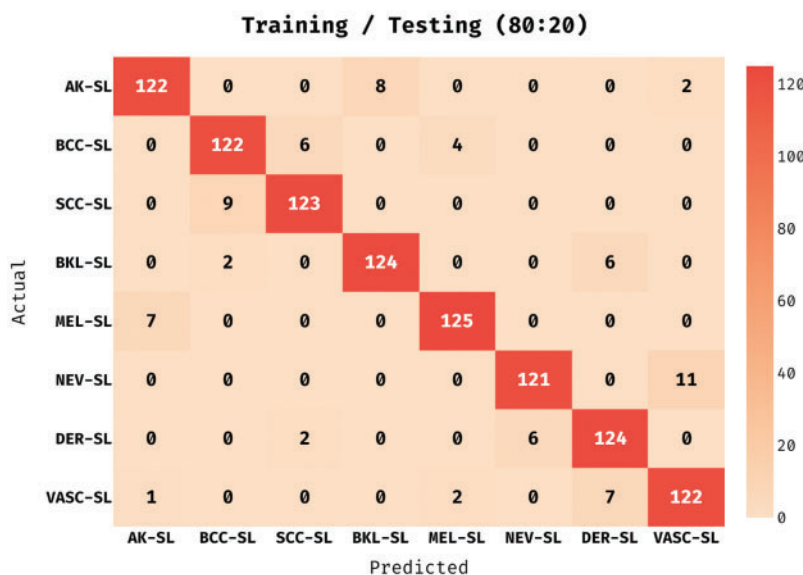


Figure 2: Samples images

The confusion matrix obtained by the ADL-MDC technique on melanoma classification with the training/testing data of 80:20 in Fig. 3. The figure reported that the ADL-MDC technique has classified 122 images under AK-SL class, 122 images under BCC-SL class, 123 images under SCC-SL class, 124 images under BKL-SL class, 125 images under MEL-SL class, 121 images under NEV-SL class, and 122 images under VASC-SL class.



**Figure 3:** Confusion matrix of ADL-MDC technique (Training/Testing of 80:20)

A brief melanoma classification result analysis of the ADL-MDC technique with the training/testing set of 80:20 is shown in Tab. 1 and Fig. 4. The results show that the ADL-MDC technique has classified all the images effectively. For instance, with AK-SL class, the ADL-MDC technique has obtained  $PRE_N$ ,  $REC_L$ ,  $SPEC_Y$ ,  $ACC_Y$ , and  $F_{SCORE}$  of 93.85%, 92.42%, 99.13%, 98.30%, and 93.13% respectively. In line with BCC-SL class, the ADL-MDC technique has attained  $PRE_N$ ,  $REC_L$ ,  $SPEC_Y$ ,  $ACC_Y$ , and  $F_{SCORE}$  of 91.73%, 92.42%, 98.81%, 98.01%, and 92.08% respectively. Along with that, with SCC-SL class, the ADL-MDC technique has offered  $PRE_N$ ,  $REC_L$ ,  $SPEC_Y$ ,  $ACC_Y$ , and  $F_{SCORE}$  of 93.89%, 93.18%, 99.13%, 98.39%, and 93.54% respectively.

**Table 1:** Classification results of ADL-MDC technique (Training/Testing of 80:20)

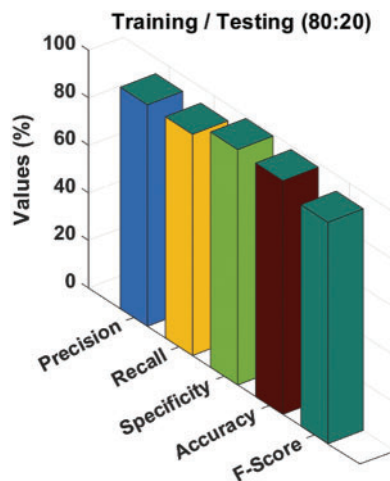
Training/Testing (80:20)					
Classes	Precision	Recall	Specificity	Accuracy	F-Score
AK-SL	93.85	92.42	99.13	98.30	93.13
BCC-SL	91.73	92.42	98.81	98.01	92.08
SCC-SL	93.89	93.18	99.13	98.39	93.54
BKL-SL	93.94	93.94	99.13	98.48	93.94
MEL-SL	95.42	94.70	99.35	98.77	95.06
NEV-SL	95.28	91.67	99.35	98.39	93.44

(Continued)



**Table 1:** Continued

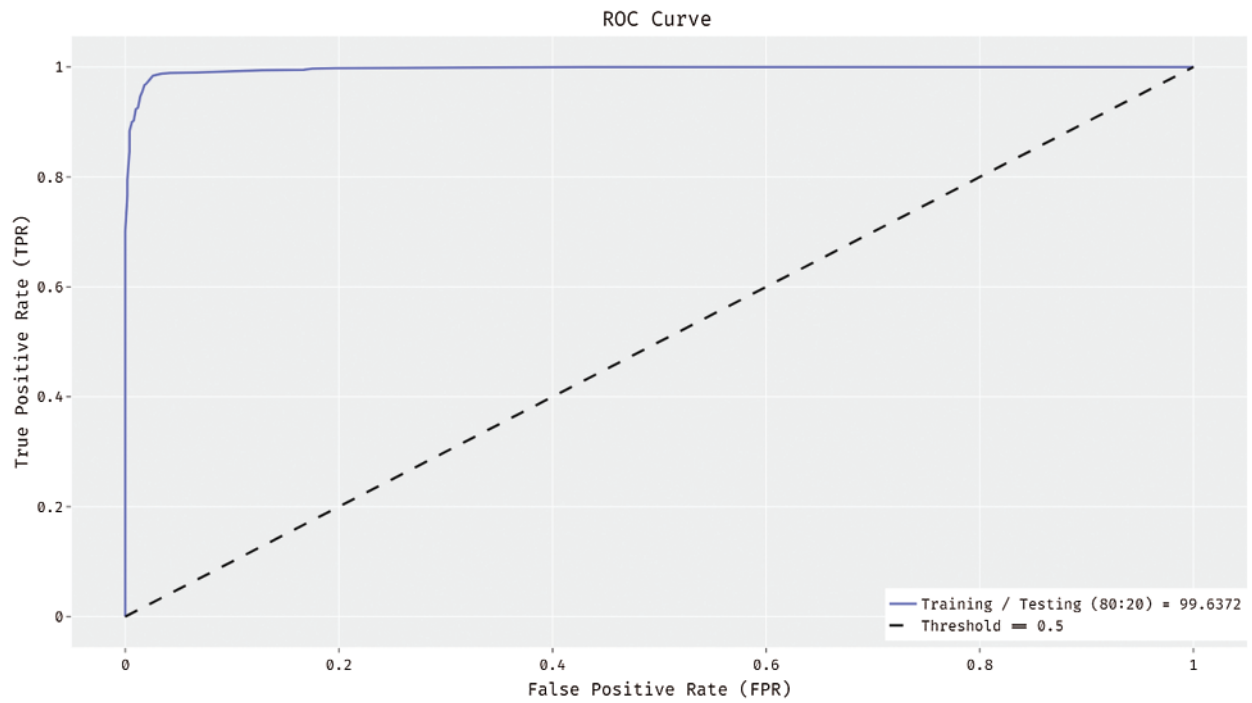
Training/Testing (80:20)					
Classes	Precision	Recall	Specificity	Accuracy	F-Score
DER-SL	90.51	93.94	98.59	98.01	92.19
VASC-SL	90.37	92.42	98.59	97.82	91.39
<b>Average</b>	<b>93.12</b>	<b>93.09</b>	<b>99.01</b>	<b>98.27</b>	<b>93.10</b>

**Figure 4:** Melanoma classification results of ADL-MDC technique (Training/Testing of 80:20)

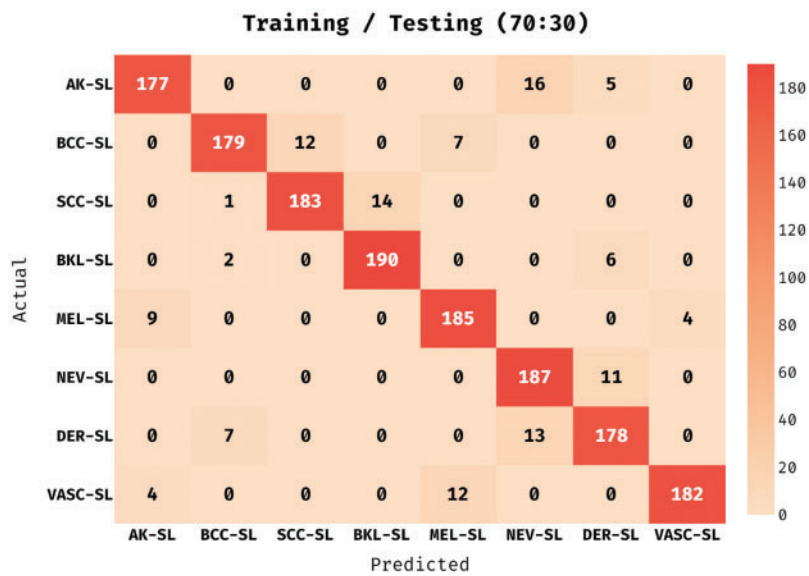
In Fig. 5, the ROC analysis of the ADL-MDC technique takes place on the training/testing data of 80:20. The figure exhibited that the ADL-MDC technique has the ability to attain improved melanoma classification performance with the maximum ROC of 99.6372.

The confusion matrix achieved by the ADL-MDC technique on melanoma classification with the training/testing data of 70:30 in Fig. 6. The figure portrayed the ADL-MDC technique has categorized 177 images into AK-SL class, 179 images into BCC-SL class, 183 images into SCC-SL class, 190 images into BKL-SL class, 185 images into MEL-SL class, 187 images into NEV-SL class, 178 images into DER-SL class, and 122 images into VASC-SL class.

Tab. 2 and Fig. 7 give a comprehensive melanoma classification result of the ADL-MDC technique on the training/testing set of 70:30 is shown in. The results exposed that the ADL-MDC technique has categorized all the images efficiently. For instance, with AK-SL class, the ADL-MDC technique has resulted to  $PRE_N$ ,  $REC_L$ ,  $SPEC_Y$ ,  $ACC_Y$ , and  $F_{SCORE}$  of 93.16%, 89.39%, 99.06%, 97.85%, and 91.24% respectively. Meanwhile, under BCC-SL class, the ADL-MDC technique has gained  $PRE_N$ ,  $REC_L$ ,  $SPEC_Y$ ,  $ACC_Y$ , and  $F_{SCORE}$  of 94.71%, 90.40%, 99.28%, 98.71%, and 92.51% respectively. Eventually, with SCC-SL class, the ADL-MDC technique has accomplished  $PRE_N$ ,  $REC_L$ ,  $SPEC_Y$ ,  $ACC_Y$ , and  $F_{SCORE}$  of 93.85%, 92.42%, 99.13%, 98.30%, and 93.13% respectively.



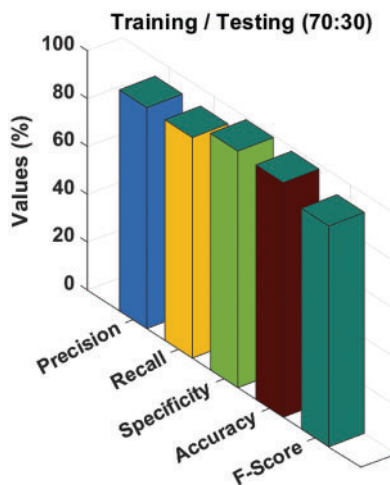
**Figure 5:** ROC curve of ADL-MDC technique (Training/Testing of 80:20)



**Figure 6:** Confusion matrix of ADL-MDC technique (Training/Testing of 70:30)

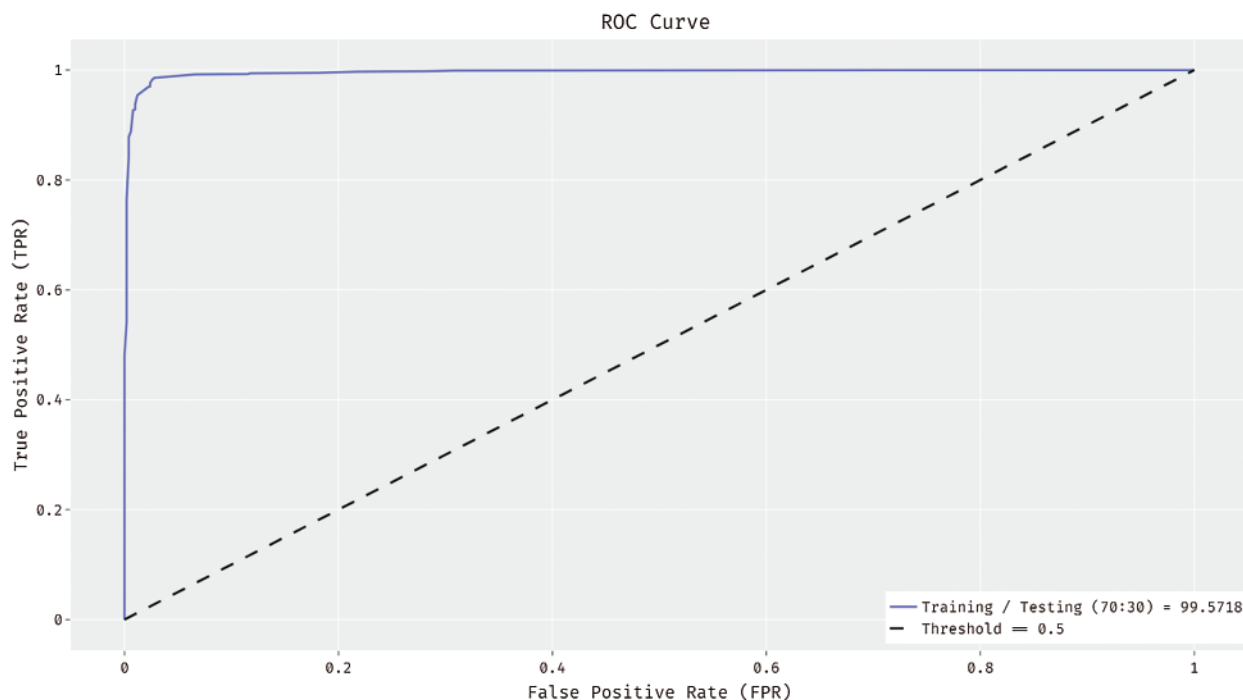
**Table 2:** Classification results of ADL-MDC technique (Training/Testing of 70:30)

Training/Testing (70:30)					
Classes	Precision	Recall	Specificity	Accuracy	F-Score
AK-SL	93.16	89.39	99.06	97.85	91.24
BCC-SL	94.71	90.40	99.28	98.17	92.51
SCC-SL	93.85	92.42	99.13	98.30	93.13
BKL-SL	93.14	95.96	98.99	98.61	94.53
MEL-SL	90.69	93.43	98.63	97.98	92.04
NEV-SL	86.57	94.44	97.91	97.47	90.34
DER-SL	89.00	89.90	98.41	97.35	89.45
VASC-SL	97.85	91.92	99.71	98.74	94.79
Average	92.37	92.23	98.89	98.06	92.25

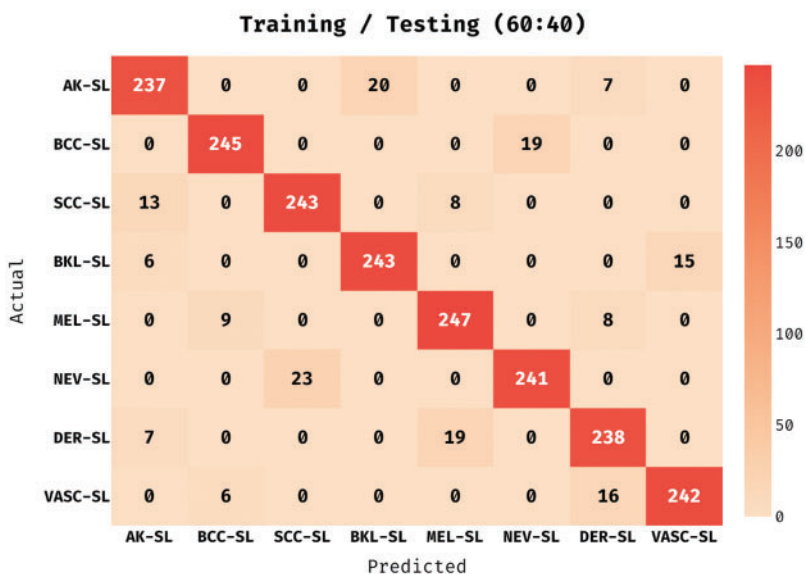
**Figure 7:** Melanoma classification results of ADL-MDC technique (Training/Testing of 70:30)

In Fig. 8, the ROC analysis of the ADL-MDC technique is carried out on the training/testing data of 70:30. The figure displayed that the ADL-MDC technique has the capability of accomplishing enhanced melanoma classification performance with the higher ROC of 99.5718.

The confusion matrix derived by the ADL-MDC technique on melanoma classification with the training/testing data of 60:40 in Fig. 9. The figure stated that the ADL-MDC technique has categorized 237 images under AK-SL class, 245 images under BCC-SL class, 243 images under SCC-SL class, 243 images under BKL-SL class, 247 images under MEL-SL class, 241 images under NEV-SL class, and 242 images under VASC-SL class.



**Figure 8:** ROC curve of ADL-MDC technique (Training/Testing of 70:30)



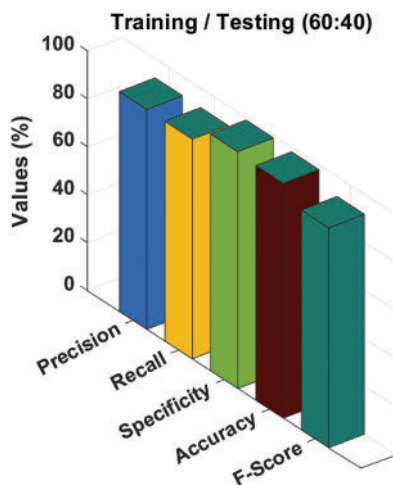
**Figure 9:** Confusion matrix of ADL-MDC technique (Training/Testing of 60:40)

A detailed melanoma classification result analysis of the ADL-MDC technique with the training/testing set of 60:40 takes place in [Tab. 3](#) and [Fig. 10](#). The experimental values denoted that the ADL-MDC technique has categorized the skin lesion images effectually. For instance, with AK-SL class, the ADL-MDC technique has reached to  $PRE_N$ ,  $REC_L$ ,  $SPEC_Y$ ,  $ACC_Y$ , and  $F_{SCORE}$  of

90.11%, 89.77%, 98.59%, 97.49%, and 89.94% respectively. In the meantime, under BCC-SL class, the ADL-MDC technique has gained  $PRE_N$ ,  $REC_L$ ,  $SPEC_Y$ ,  $ACC_Y$ , and  $F_{SCORE}$  of 94.23%, 92.80%, 99.19%, 98.39%, and 93.51% respectively. At last, with SCC-SL class, the ADL-MDC technique has accomplished  $PRE_N$ ,  $REC_L$ ,  $SPEC_Y$ ,  $ACC_Y$ , and  $F_{SCORE}$  of 91.35%, 92.05%, 98.76%, 97.92%, and 91.70% respectively.

**Table 3:** Classification results of ADL-MDC technique (Training/Testing of 60:40)

Training/Testing (60:40)					
Classes	Precision	Recall	Specificity	Accuracy	F-Score
AK-SL	90.11	89.77	98.59	97.49	89.94
BCC-SL	94.23	92.80	99.19	98.39	93.51
SCC-SL	91.35	92.05	98.76	97.92	91.70
BKL-SL	92.40	92.05	98.92	98.06	92.22
MEL-SL	90.15	93.56	98.54	97.92	91.82
NEV-SL	92.69	91.29	98.97	98.01	91.98
DER-SL	88.48	90.15	98.32	97.30	89.31
VASC-SL	94.16	91.67	99.19	98.25	92.90
Average	91.70	91.67	98.81	97.92	91.67

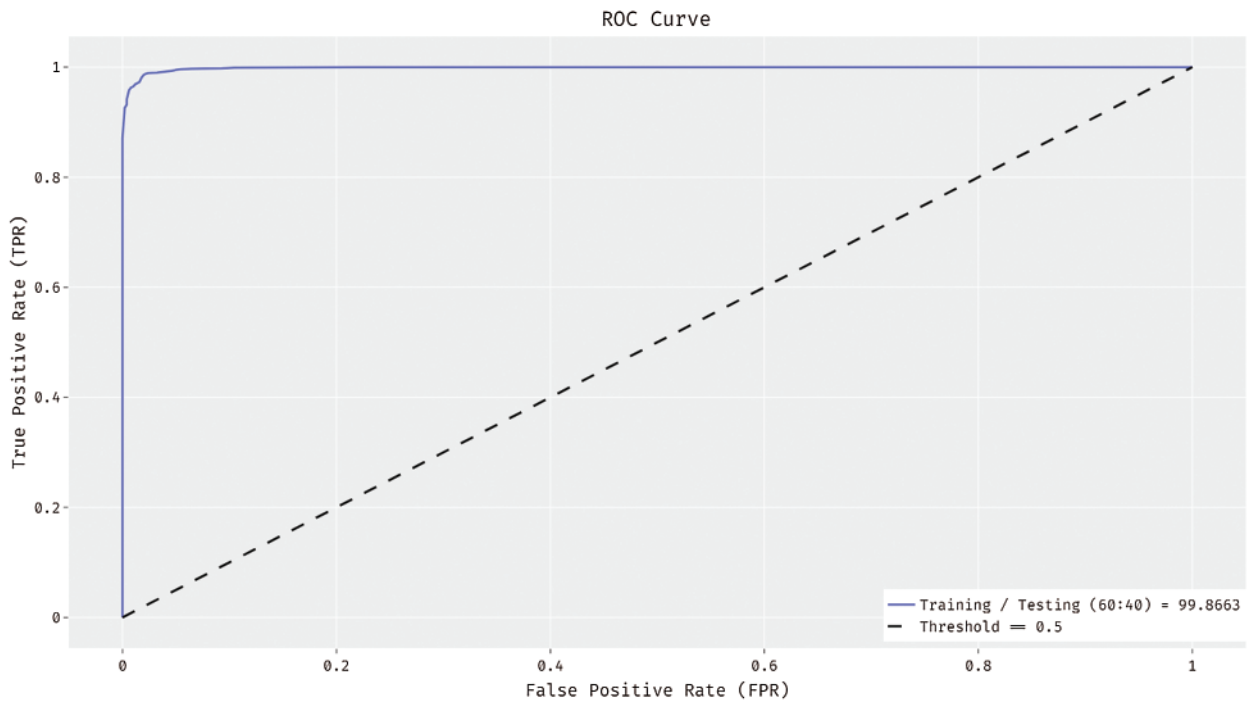


**Figure 10:** Melanoma classification results of ADL-MDC technique (Training/Testing of 60:40)

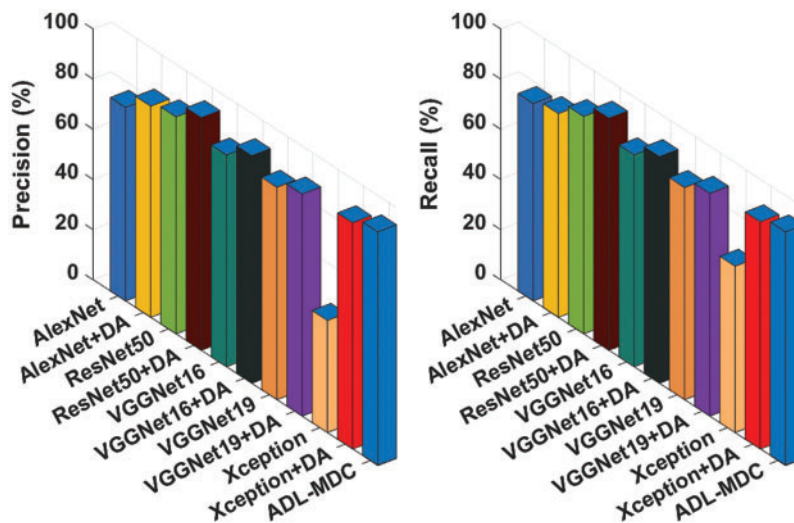
Fig. 11 exhibits the ROC analysis of the ADL-MDC technique on the training/testing data of 60:40. The results showcased that the ADL-MDC technique has the capacity of reaching better melanoma classification performance with a maximum ROC of 99.8663.

Fig. 12 illustrates the  $PRE_N$  and  $REC_L$  analysis of the ADL-MDC technique with recent methods [12]. The figure exhibited that the AlexNet and Xception models have obtained lower values of  $PRE_N$  and  $REC_L$ . Followed by, the AlexNet + DA, VGGNet16, VGGNet19, ResNet50, and VGGNet19 + DA models have obtained moderately closer values of  $PRE_N$  and  $REC_L$ . In line with, the

Xception + DA, VGGNet16 + DA and ResNet50 + DA techniques have resulted in reasonable values of  $PRE_N$  and  $REC_L$ . However, the ADL-MDC technique has outperformed the other techniques with the maximum  $PRE_N$  and  $REC_L$  of 93.12% and 93.09% respectively.



**Figure 11:** ROC curve of ADL-MDC technique (Training/Testing of 60:40)



**Figure 12:** Comparative  $PRE_N$  and  $REC_L$  analysis of ADL-MDC technique

Fig. 13 demonstrates the  $ACC_Y$ , and  $F_{SCORE}$  analysis of the ADL-MDC technique with existing techniques. The results reported that the AlexNet and Xception models have accomplished minimal

values of  $ACC_Y$ , and  $F_{SCORE}$ . Moreover, the AlexNet + DA, VGGNet16, VGGNet19, ResNet50, VGGNet16 + DA, and VGGNet19 + DA models have gained reasonably increased values of  $ACC_Y$ , and  $F_{SCORE}$ . Furthermore, the Xception + DA and ResNet50 + DA techniques have resulted in near optimal values of  $ACC_Y$ , and  $F_{SCORE}$ . However, the ADL-MDC technique has showcased improved performance with the higher  $ACC_Y$ , and  $F_{SCORE}$  of 98.27% and 93.09% respectively. From the above mentioned tables and figures, it is evident that the ADL-MDC technique has reached maximum melanoma classification performance compared to other techniques.

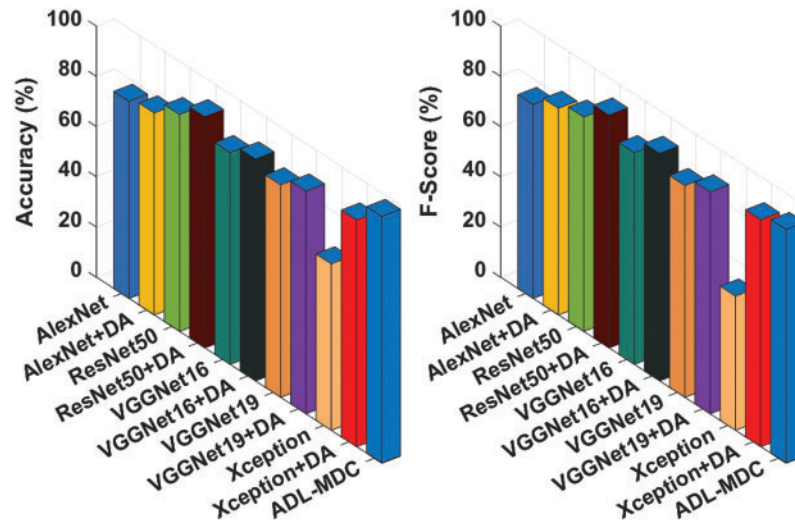


Figure 13: Comparative  $ACC_Y$  and  $F_{Score}$  analysis of ADL-MDC technique

## 5 Conclusion

In this study, a novel ADL-MDC technique has been derived for melanoma detection and classification using dermoscopic images. The ADL-MDC technique encompasses several stages including pre-processing, K-means segmentation, CapsNet based feature extraction, Adagrad optimizer based hyperparameter tuning, SAE based classification, and CSO based parameter optimization. The exploitation of the Adagrad and CSO algorithm helps to properly accomplish improved performance. A wide range of simulation analyses is carried out on benchmark datasets and the results are inspected under several aspects. The simulation results reported the enhanced performance of the ADL-MDC technique over the recent approaches. As a part of future extension, the classification performance of the ADL-MDC technique can be improved by the design of DL based image segmentation techniques.

**Funding Statement:** The authors extend their appreciation to the Deanship of Scientific Research at King Khalid University for funding this work under Grant Number (RGP 1/80/43). Princess Nourah bint Abdulrahman University Researchers Supporting Project Number (PNURSP2022R191), Princess Nourah bint Abdulrahman University, Riyadh, Saudi Arabia.

**Conflicts of Interest:** The authors declare that they have no conflicts of interest to report regarding the present study.

## References

- [1] E. O. M. Molina, S. S. Calderón and J. Á. Borrego, “Classification of dermoscopy skin lesion color-images using fractal-deep learning features,” *Applied Sciences*, vol. 10, no. 17, pp. 5954, 2020.
- [2] T. J. Brinker, L. Kiehl, M. Schmitt, T. B. Jutzi, E. I. K. Henning *et al.*, “Deep learning approach to predict sentinel lymph node status directly from routine histology of primary melanoma tumours,” *European Journal of Cancer*, vol. 154, pp. 227–234, 2021.
- [3] M. A. Al-masni, M. A. Al-antari, M. T. Choi, S. M. Han and T. S. Kim, “Skin lesion segmentation in dermoscopy images via deep full resolution convolutional networks,” *Computer Methods and Programs in Biomedicine*, vol. 162, pp. 221–231, 2018.
- [4] M. Q. Khan, A. Hussain, S. U. Rehman, U. Khan, M. Maqsood *et al.*, “Classification of melanoma and nevus in digital images for diagnosis of skin cancer,” *IEEE Access*, vol. 7, pp. 90132–90144, 2019.
- [5] A. Marka, J. B. Carter, E. Toto and S. Hassanpour, “Automated detection of nonmelanoma skin cancer using digital images: A systematic review,” *BMC Medical Imaging*, vol. 19, no. 1, pp. 21, 2019.
- [6] M. A. Wahba, A. S. Ashour, Y. Guo, S. A. Napoleon and M. M. A. Elnaby, “A novel cumulative level difference mean based GLDM and modified ABCD features ranked using eigenvector centrality approach for four skin lesion types classification,” *Computer Methods and Programs in Biomedicine*, vol. 165, pp. 163–174, 2018.
- [7] N. Amoroso, D. Diacono, A. Fanizzi, M. L. Rocca, A. Monaco *et al.*, “Deep learning reveals Alzheimer’s disease onset in MCI subjects: Results from an international challenge,” *Journal of Neuroscience Methods*, vol. 302, pp. 3–9, 2018.
- [8] C. Barata, J. S. Marques and M. E. Celebi, “Improving dermoscopy image analysis using color constancy,” in *2014 IEEE Int. Conf. on Image Processing (ICIP)*, Paris, France, pp. 3527–3531, 2014.
- [9] N. Gessert, M. Nielsen, M. Shaikh, R. Werner and A. Schlaefer, “Skin lesion classification using ensembles of multi-resolution EfficientNets with meta data,” *MethodsX*, vol. 7, pp. 100864, 2020.
- [10] J. Daghrir, L. Tlig, M. Bouchouicha and M. Sayadi, “Melanoma skin cancer detection using deep learning and classical machine learning techniques: A hybrid approach,” in *2020 5th Int. Conf. on Advanced Technologies for Signal and Image Processing (ATSIP)*, Sousse, Tunisia, pp. 1–5, 2020.
- [11] S. H. Kassani and P. H. Kassani, “A comparative study of deep learning architectures on melanoma detection,” *Tissue and Cell*, vol. 58, pp. 76–83, 2019.
- [12] B. Thiyaneswaran, K. Anguraj, S. Kumarganesh and K. Thangaraj, “Early detection of melanoma images using gray level co-occurrence matrix features and machine learning techniques for effective clinical diagnosis,” *International Journal of Imaging Systems and Technology*, vol. 31, no. 2, pp. 682–694, 2021.
- [13] R. Patil and S. Bellary, “Machine learning approach in melanoma cancer stage detection,” *Journal of King Saud University-Computer and Information Sciences*, 2020. <http://doi:10.1016/j.jksuci.2020.09.002>.
- [14] J. B. Janney and S. E. Roslin, “Classification of melanoma from Dermoscopic data using machine learning techniques,” *Multimedia Tools and Applications*, vol. 79, no. 5–6, pp. 3713–3728, 2020.
- [15] A. R. Ali, J. Li, S. Kanwal, G. Yang, A. Hussain *et al.*, “A novel fuzzy multilayer perceptron (f-mlp) for the detection of irregularity in skin lesion border using dermoscopic images,” *Frontiers in Medicine*, vol. 7, pp. 297, 2020.
- [16] N. S. A. ALenezi, “A method of skin disease detection using image processing and machine learning,” *Procedia Computer Science*, vol. 163, pp. 85–92, 2019.
- [17] A. M. Reza, “Realization of the contrast limited adaptive histogram equalization (clahe) for real-time image enhancement,” *The Journal of VLSI Signal Processing-Systems for Signal, Image, and Video Technology*, vol. 38, no. 1, pp. 35–44, 2004.
- [18] N. Dhanachandra, K. Mangleam and Y. J. Chanu, “Image segmentation using k-means clustering algorithm and subtractive clustering algorithm,” *Procedia Computer Science*, vol. 54, pp. 764–771, 2015.
- [19] H. Chao, L. Dong, Y. Liu and B. Lu, “Emotion recognition from multiband EEG signals using CapsNet,” *Sensors*, vol. 19, no. 9, pp. 2212, 2019.



- [20] C. Zhang, M. Yao, W. Chen, S. Zhang, D. Chen *et al.*, “Gradient descent optimization in deep learning model training based on multistage and method combination strategy,” *Security and Communication Networks*, vol. 2021, pp. 1–15, 2021s.
- [21] B. Abraham and M. S. Nair, “Computer-aided classification of prostate cancer grade groups from MRI images using texture features and stacked sparse autoencoder,” *Computerized Medical Imaging and Graphics*, vol. 69, pp. 60–68, 2018s.
- [22] A. Askarzadeh, “A novel metaheuristic method for solving constrained engineering optimization problems: Crow search algorithm,” *Computers & Structures*, vol. 169, pp. 1–12, 2016.
- [23] Y. H. Jia, J. Qiu, Z. Z. Ma and F. F. Li, “A novel crow swarm optimization algorithm (CSO) coupling particle swarm optimization (PSO) and crow search algorithm (CSA),” *Computational Intelligence and Neuroscience*, vol. 2021, pp. 1–14, 2021.
- [24] P. Tschandl, C. Rosendahl and H. Kittler, “The HAM10000 dataset, a large collection of multi-source dermatoscopic images of common pigmented skin lesions,” *Scientific Data*, vol. 5, no. 1, pp. 180161, 2018.
- [25] N. C. F. Codella, D. Gutman, M. E. Celebi, B. Helba, M. A. Marchetti *et al.*, “Skin lesion analysis toward melanoma detection: A challenge,” in *2018 IEEE 15th Int. Symp. on Biomedical Imaging (ISBI 2018)*, Washington, DC, pp. 168–172, 2018.