N. I. Md. Ashafuddula
Rafiqul Islam

MELANOMA SKIN CANCER AND NEVUS MOLE CLASSIFICATION USING INTENSITY VALUE ESTIMATION WITH CONVOLUTIONAL NEURAL NETWORK

Abstract

Melanoma skin cancer is one of the most dangerous and life-threatening cancers. Exposure to ultraviolet rays may damage the skin cell's DNA, which can cause melanoma skin cancer. However, detecting and classifying melanoma and nevus moles at their immature stages is difficult. In this work, an automatic deep-learning system has been developed based on intensity value estimation with a convolutional neural network model (CNN) for detecting and classifying melanoma and nevus moles more accurately. Since intensity levels are the most distinctive features for identifying objects or regions of interest, high-intensity pixel values have been selected from extracted lesion images. Incorporating these high-intensity features into CNN improves the overall performance of the proposed model than the state-of-the-art methods for detecting melanoma skin cancer. To evaluate the system, we used five-fold cross-validation. The experimental results showed that superior percentages of accuracy (92.58%), sensitivity (93.76%), specificity (91.56%), and precision (90.68%) were achieved.

Keywords   melanoma detection, medical imaging, image classification, convolutional neural network, intensity value estimation, canny edge detection

1. Introduction

Skin covers the entire outsides of our bodies and protects all of our internal body parts from environmental injuries. Nevertheless, it sometimes suffers from different diseases that could often be life-threatening dueto its location, and skin cancer is one of them [9]. There are three major types of skin cancer: basal cell carcinoma (BCC), squamous cell carcinoma (SCC), and melanoma and Merkel cell carcinoma (MCC) [5]. Among all skin cancers, melanoma is the most dangerous and life-threatening one [32]. It grows rapidly and can spread to other parts of the body. According to [5], about 99,780 new melanomas will be diagnosed (about 57,180 in men, and 42,600 in women), and 7,650 people are expected to die of melanoma skin cancer (about 5,080 men, and 2,570 women) in 2022. The World Health Organization (WHO) reports that one out of every three malignancies is affected by skin cancer [40]. From 2009 to 2019, there has been about a 55-percent increase in current skin cases examined annually [39]. The lifetime risk of being affected by melanoma is about 2.6% (1 in 38) for whites, 0.1% (1 in 1,000) for Blacks, and 0.6% (1 in 167) for Hispanics [5]. The early-stage detection of melanoma skin cancer can be beneficial in curing it [3, 11, 12, 37]. The early diagnosis survival rate of skin cancer is more than 90 percent [2].

Due to the complex nature of skin lesions, dermatologists relate many visual observations (such as the symmetry of a lesion area and its size, shape, color, and border) to diagnose malignant melanoma [1]. A sign of an ugly duckling lesion is another warning sign of melanoma [4]. There are also some popular scoring techniques to identify malignant melanoma, where a 7-point checklist [6], the Menzies method, a 3-point checklist, and ABCDE rules are the widely used methods [4]. Out of these, a common way for detecting melanoma skin cancer is the ABCDE rule: A for asymmetry (one half of the mole does not match the other), B for border irregularity, C for non-uniform color, D for a diameter that is greater than 6 mm (or $\frac{1}{4}$ inch – about the size of a pencil eraser), and E for evolving size, shape or color. To assess a mole's size, color, and texture, specialist doctors examine lesion areas in which the total dermatoscopic score (TDS) is estimated by using ABCDE and some weight factors. They conclude the diagnosis results from the TDS value; if the TDS is above a threshold value, they identify the lesion as a malignant one; otherwise, it is benign [29].

Since this detection procedure is time-consuming, it needs specialist doctors to ascertain the results [22]. There is a high possibility of being influenced by human subjectivity, which makes it inconsistent in certain conditions. A computerized artificial intelligence-based system can easily detect and differentiate between melanoma and normal skin to eliminate this weakness. Recently, deep neural networks (DNN) have been used in many different ways for medical imaging [17]. A convolutional neural network (CNN) is useful for image classification and recognition because of its capability to achieve high accuracy in a short amount of time [42, 43]. Compared to its predecessors, the main advantage of CNN is that it detects important features automatically without any human supervision [27]. In this work, an intelligent system is developed to predict melanoma and nevus moles at an early stage, considering the

complicated issue. Our innovative approach uses pre-processing steps, an intensity value estimation (IVE) model, and a CNN model.

To train and test the model, we used five-fold cross-validation. The proposed model is evaluated using five well-known quantitative metrics: sensitivity, specificity, positive predictive value (PPV), negative predictive value (NPV), and accuracy. Considering the issues, the major contributions through this work can be enlisted as follows:

- To develop an intensity value estimation (IVE) model.
- To present an efficient image-resizing technique to keep the lesion shape, minimizing possible data loss.
- To design the proposed methodology by combining the resizing technique, IVE model, and convolutional neural network (IVEwCNN) for the automatic classification and detection of melanoma skin cancer and nevus moles.
- A comparative histogram analysis is performed between our newly developed method and other state-of-the-art methods to evaluate the effectiveness of our newly developed model.

The rest of the paper is organized as follows: a related literature review is discussed in Section 2. The proposed methodology is categorized into subsections and briefly discussed in Section 3. Section 3.2 describes the high intense pixel value estimation and acquisition process. The experimental analysis is discussed in Section 5. Image acquisition and MED-NODE data set descriptions are stated in Section 4. In Section 5.2, the performance-evaluation metrics and experimental results are discussed concerning different methods and the MED-NODE data set. Finally, the conclusion is discussed in Section 7.

2. Background study

The authors in [14] worked with the MED-NODE data set. To segment regions of interest in healthy and lesioned areas, they used k-means clustering ($k = 2$). Before segmentation, a series of pre-processing steps were done to handle the noise and illumination effectively. They utilized Gaussian-smoothing ($\sigma = 5$) and Kuwahara-smoothing filters to remove noise and additional noisy features while preserving the edges. After removing the noise, they mapped each image into 50 sub-images of a pixel size of 15×15. A cluster-based adaptive metric classifier was developed to extract 675 features per image. The training feature vector size was 2250 (50×45), and the evaluated feature vector size was 6250 (50×125). The model was good, but the data set's size was one of its main limitations.

A deep neural model in which the authors in [31] considered reducing the illumination and noise effect in the pre-processing step was proposed; then, they fed the enhanced images into a pre-trained CNN model. Hence, the data set had limited images, so they used cropping and rotation to expand the training data. They produced a segmentation mask by applying the k-means classifier ($k = 2$) to the

pre-processed image. The mask was enhanced by applying morphological operations. Based on the information of the segmentation mask, a Gaussian filter was used on the standard skin parts. In their CNN model, 20 feature maps were generated in the first convolution layer, and 50 were generated in the second convolution layer. After each convolution layer, the pooling layer and a two-layer fully connected stage were utilized. Finally, the diagnosis results were found from this two-layer network with a linear transfer function. The training data was fed to the network using a batch size of 64. They randomly split the data set into 20% testing and 80% training data with no overlapping. To train the model, they performed 20,000 iterations. They addressed the fact that the illumination correction increased their system's discrimination capability, which helped increase system accuracy.

The authors in [39] developed a combined model using multi-level segmentation, CNN, a support vector machine (SVM), and a back-propagation neural network that combined the Otsu, modified Otsu, and watershed segmentation methods for segmentation. Finally, CNN and SVM were used for the training and classification, respectively. Other authors [28] used MED-NODE and the international skin imaging collaboration (ISIC) data set together with 2,170 images for their classification model. To avoid noise and common lighting problems, they used a few pre-processing steps to improve the image texture and darken the perimeter of a lesion. They utilized contrast enhancement and an anisotropic diffusion filter to correct the image contrast, remove noise, and preserve the lesion edges. Then, the shape features method and principle component analysis (PCA) were used to extract features and reduce features dimension. A convolution deep neural network (CDNN) was used where the dropout was set to 0.5 to avoid overfitting, the learning rate tested 0.1 to 0.001, and the discovered minor variation worked well when tested on different batch sizes, but they got 32 and 64 (which work well in many situations). Their proposed system achieved 96.8% accuracy with a few noticeable epochs in 0.41 min.

Another skin cancer-detection technique was introduced in [13] that utilized the well-known HAM10000 data set. In the pre-processing steps, they removed noise, reduced image resolution, and applied image augmentation to avoid overfitting in order to increase the learnability of the system. The system produced several copies of existing images by applying translation, rotation, and zooming on images. For the classification task, they used CNN and the transfer-learning method along with other classification algorithms such as XGBoost, SVM, and random forest to classify and compare their analyses. Finally, they yielded an accuracy of 90.51% in the ResNet model's transfer-learning approach. The author's used the MED-NODE data set [30]. In the pre-processing step, Otsu's segmentation method was used on a grayscale image to segment the lesion part from the image. A total of 1900 features were extracted from each segmented lesion image. Twenty-five features were excluded from one thousand nine hundred features, as they were either too high, too low, or constant across the data set. With the different types of training and cost functions that were available in a multi-layer neural network (MLP), this feature set was tested by various

neurons. The relief method selected the best features rank-wise among all 1,875 features. Using the features that were found from a principal component analysis (PCA) of the extracted 1,875 features, neural network (MLP), linear SVM, medium KNN, and linear discriminant were the classifiers. Their work MLP with PCA featured only used 25 features showed an 87.18% accuracy. The authors in [44] addressed the effect of contrast-enhancing and image-texture analysis, and they considered pixel-intensity values in an image-classification model. They found that contrast-enhancing was designed to increase the discrimination between the intensity values of an image to help them be easily identifiable by human and computer vision. The primary pieces of information were stored in the intensity value of a pixel (this can be a single value for a gray-level image or three values for a color image). They also addressed the fact that image-texture analysis could be an essential factor for pattern recognition due to the power of its discrimination ability.

Noise was reduced from the skin lesions by using a Gaussian filter, then they applied an improved k-means clustering-based segmentation [7]. The authors in [7] extracted three different features by using a local binary pattern (LBP), a gray-level co-occurrence matrix (GLCM), and RGB color channel features from the ROI of skin lesions. A distinctive hybrid super feature vector was created by using the extracted textural and color features from a lesion. For classification, they used a support vector machine (SVM), k-nearest neighbor (KNN), naÛve Bayes (NB), and decision tree (DT). The DERMIS data set (146 melanoma and 251 nevus) was used in the experiment. The texture features, GLCM, and LBP features were merged with the color features to acquire a high classification accuracy.

3. Proposed methodology

The general process of detecting and diagnosing melanoma skin cancer has been summarized into key operating procedures such as image pre-processing, image segmentation, feature extraction and analysis, and the classification of lesion images [26, 38] as described in Figure 1.
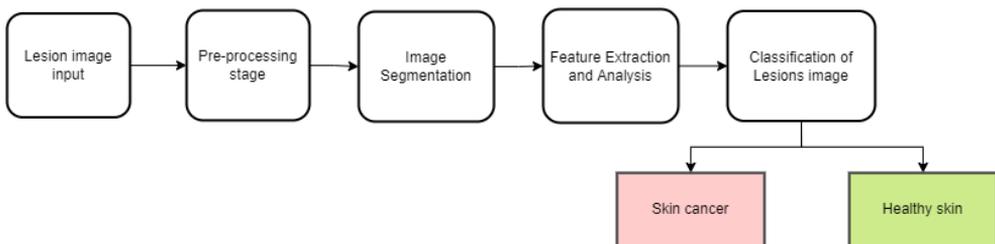
Figure 1. Process of diagnosing melanoma skin cancer

Our proposed methodology is divided into pre-processing, intensity value estimation, and convolutional neural network (CNN) steps. A few sub-parts in the

pre-processing step were removed to reduce any artifacts that could have misled the CNN model. The proposed methodology is described by the block diagram that is shown in Figure 2.
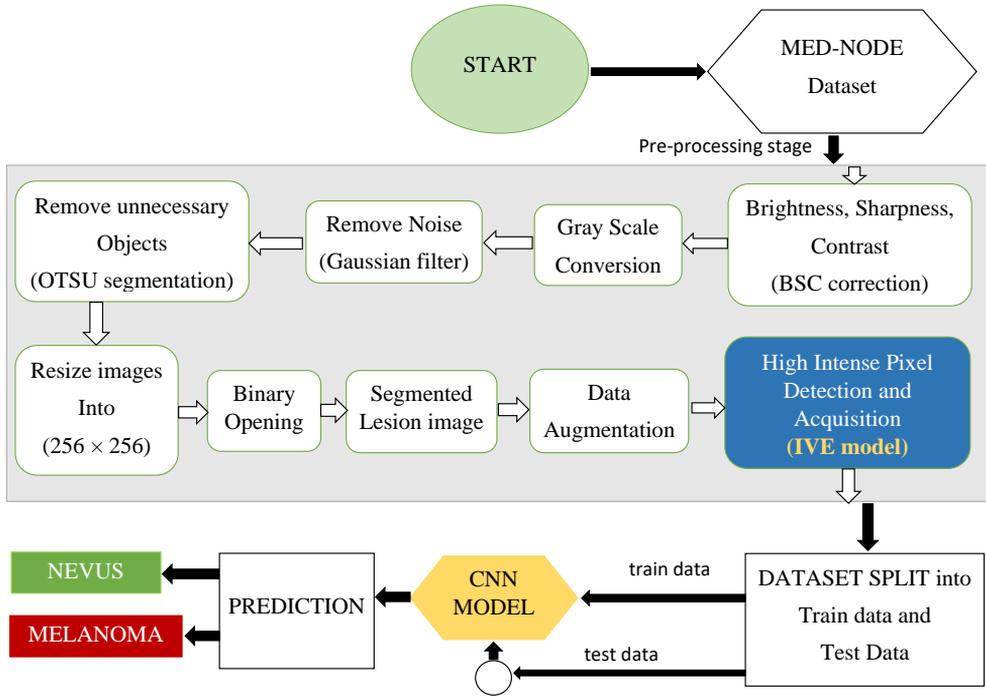

Figure 2. Detailed architecture of proposed methodology

3.1. Pre-processing

In the pre-processing stage, images from the collected MED-NODE data set [14] were passed through a series of basic image-processing methods to reduce the effects of any misleading factors on the CNN model. As the images contained both healthy skin and lesion parts, we corrected the brightness (B), sharpness (S), and contrast (C) to enhance the images' lesion part to process further. This experiment used a Python "ImageEnhance" class from the "pillow" module to correct the BSC. The general intensity of the pixels refers to the brightness in an image, and the histogram gives a clearer indication of the brightness. The higher end in the histogram indicates a brighter image, and the image is darker when the histogram is confined to a small portion towards the lower end [19]. The number of details present in an image refers to its sharpness. Lossy compression, motion blur, de-noising, and out-of-focus filtering are some of the causes that affect perceived image sharpness [19]. In general, contrast refers to splitting up the dark and bright regions in an image [8]. The contrast-enhancement technique eliminates the anomaly that would otherwise occur between different regions in an image.

An RGB image contains three channels – red (R), green (G), and blue (B) – whereas the grayscale image contains only one channel. The three channels in an RGB image contain different intensity levels to represent an image in a color form. For various image-processing tasks such as morphological operation segmentation, using a grayscale image is more accessible than an RGB image because it has only one channel. So, we converted the images from RGB to grayscale. The grayscale pixel value is calculated as the weighted sum of the corresponding R (red), G (green), and B (blue) pixels as in Eq. (1) to convert an RGB channel image into a single grayscale channel image. Cathode-ray tube (CRT) phosphors use these weights to better represent the human perception of RGB than equal weights [33]. The texture in an image offers information about the spatial ordering of colors or intensities in an image or the selected region of an image [41]. Since the grayscale image is formed from the three-channel RGB image using Eq. (1), the image color and texture could be represented by the different levels of pixel intensity in the grayscale image [10, 23]:

$$G_{im} = 0.2125R + 0.7154G + 0.0721B \qquad (1)$$

Although professional high-resolution cameras take images, non-uniform light creates noise effects. We used the Gaussian filter to remove noise from the grayscale lesion images ($\sigma = 1.35$). The healthy skin part of the images was irrelevant to our model; so, we segmented the lesion part from the healthy skin part using Otsu segmentation before feeding the images into the CNN model.

Algorithm 1: Resize Image

Input: A 2D image $[oldIm_{r,c}]$, where $r \leq N_r$, $c \leq N_c$
Output: Resized 2D Image of size $(N_r \times N_c)$

1. Initialization
    (a) $reqSize = (N_r, N_c)$
    (b) A 2D image $[newIm_{r,c}]$, $r, c = reqSize$, where each element is initialized with integer 0.
    (c) $hr \leftarrow \lfloor (reqSize - len(oldIm_{rowsize}))/2 \rfloor$
    (d) $hc \leftarrow \lfloor (reqSize - len(oldIm_{colsize}))/2 \rfloor$
    (e) $hcUp \leftarrow hc$
2. **For** $r \leftarrow 0$ **to** $len(oldIm_{rowsize})$ **do**
    (a) **For** $c \leftarrow 0$ **to** $len(oldIm_{colsize})$ **do**
        i. $newIm[hr][hcUp] \leftarrow oldIm[r][c]$
        ii. $hcUp \leftarrow hcUp + 1$
    (b) **End For**
    (c) $hr \leftarrow hr + 1$
    (d) $hcUp \leftarrow hc$
3. **End For**
4. **Return** $newIm$

The use of a conventional resizing algorithm on a grayscale image would directly distort the skin session shape [35]. For convenience, we resized all of the segmented masks (SM) and Gaussian-filtered images (GFI) to a unique size ($N_r \times N_c$, where $N_r = 256$ and $N_c = 256$) by using Algorithm 1 in order to keep all of the lesion shapes the same as the original. We applied a binary opening on the resized mask images to remove hair, small objects, and unnecessary things to enhance the lesion area. The segmented mask is used to segment the lesion area (SLA) from the image that follows Eq. (2), where "r" and "c" represent the spatial (plane) coordinates of a 2D image, and the amplitude (r, c) is called the intensity or gray level at the point for that function:

$$SLA_{r,c} = GFI_{r,c} * SM_{r,c} \quad \text{where}, (r,c) = 1, 2, \cdots, (N_c, N_r) \tag{2}$$

Data augmentation is a strategy that significantly increases the diversity and amount of data from any available data set to train a model without collecting new data. Cropping, padding, flipping, zooming in and out, and varying an image size are the techniques that are used to augment the data. In this work, we used zoom in, zoom out, flipping, random contrast, random brightness, and rotation techniques to expand the data [24, 34]. Before feeding the images into the proposed CNN model, the proposed IVE model was applied to take high-intensity pixel values from segmented lesion images. The model was divided into two sub-steps: edge detection, and high-intensity pixel value estimation and acquisition. For edge detection, we used Canny edge detection; this eliminated the image noise by smoothing it and then found the image gradient to highlight those regions with high-spatial derivatives [20]. In the Canny edge detection, we multiplied the segmented lesion image by 255 in order to avoid losing precision while converting the image pixel values to unsigned 8-bit. After the data augmentation, the resized segmented lesion images were fed into the IVE model. The IVE model output a high intense pixel value that is stated in Figures 3g and 3h.

3.2. Intensity value estimation

The intensity value estimation (IVE) model is one of the significant contributions of the work, as intensity levels act as one of the most distinctive features for identifying objects or regions of interest [25]. In the IVE model, a multiplication between every pixel of a segmented lesion image and a constant value is performed using Eq. (3). Here, f is a 2D light intensity function, and (x, y) denotes the spatial coordinates of the image. Then, the normalization on the calculated image is used to distribute the intensity of the lesion region pixels using Eq. (4). The normalization process converts the gray-level intensity of the image by the range of 0 to 255 (Figs. [3e] and [3f]). The histogram shows the pixel value between 0 and 255, whereas the edge detection of the segmented lesion image histogram shows the pixel intensity to be within 0 or 255 (Figs. [3c] and [3d]). So, there are no intensity values between 0 and 255 in the edge-detection image, which cuts off a few features from the image. Here, our proposed IVE model helps retain the lesion image's different intensity levels. These varieties

of the intensity levels of lesion images contain more information (color and texture) in the intensity level form than can be found in normal edge-detection techniques.

a)

b)

c)

d)

e)

f)

g)

h)
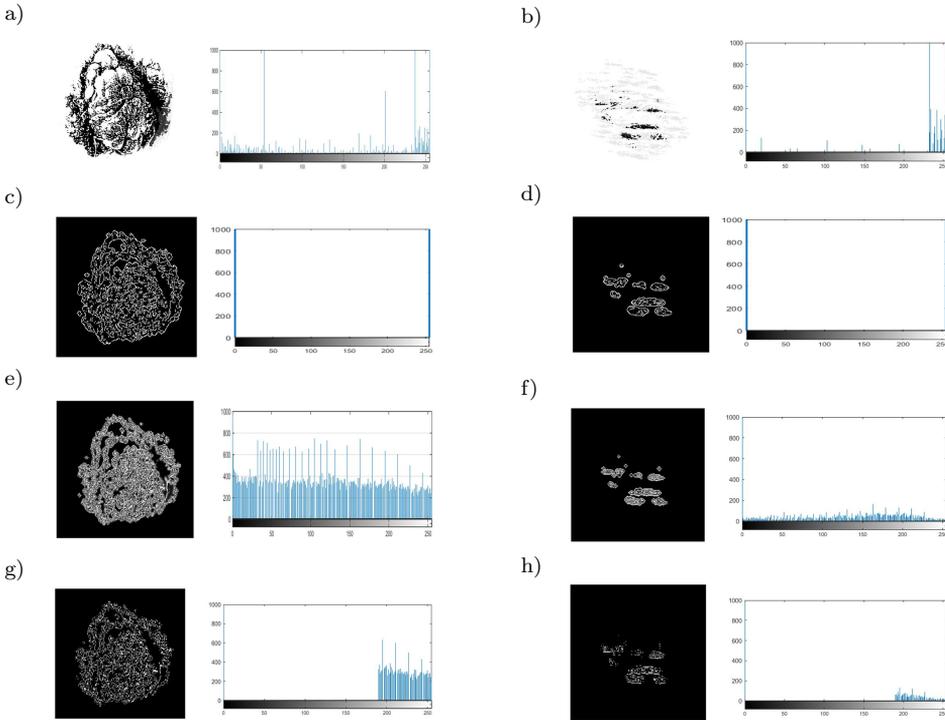
Figure 3. Melanoma and nevus lesion histogram comparison: a) gray-level image (melanoma); b) gray-level image (nevus); c) canny edge detection (melanoma); d) canny edge detection (nevus); e) unsigned intensity (melanoma); f) unsigned intensity (nevus); g) high intense pixel collection (melanoma); h) high intense pixel collection (nevus)

Table 2 shows the experimental results of our proposed model using Canny edge detection. Since the improvement of the effect was not significant, we resolved the model to distribute the gray-level intensity of the image using Eq. (4). It can be observed that this modification could select the pixel values in the segmented lesion images that were greater or equal to a threshold value Eq. (5):

$$f(x, y) = f(x, y) \times \text{Constant}, \; for \; every \; x, y; \tag{3}$$

$$f(x, y) = normalize(f(x, y)); \tag{4}$$

$$f(x, y) = \begin{cases} f(x, y), & \text{if } f(x, y) > Threshold \\ 0, & \text{otherwise.} \end{cases} \tag{5}$$

Using the proposed high-intensity value-estimation model produced more features in different intensity levels than the only the edge detection of the segmented lesion

(Figs. [3g] and [3h]). The model extracted the color and texture features in the forms of the different intensity levels of the pigmentary lesion. In this aspect, choosing intense high pixels helped our model to differentiate melanoma cells from nevus cells. This decision experiment is depicted in Figure 3.

3.3. Convolutional neural network

It has been proven that a simple artificial neural network (ANN) fails at a certain point – especially, an over-fitting may arise due to the size of the image [15]; so, CNN has tremendous advantages over ANN in image-classification problems. CNN is a class of deep neural networks that has made a massive breakthrough in image classification, recognition, object detection, face recognition, and many other areas [18]. CNN helps to auto-detect important features and extract features from images, which may be helpful in the image-classification task [34]. It extracts features from images using filters and reduces the number of learnable parameters by using the pooling technique. The CNN model is comprised of five stages of neural layers in its structure: input, convolutional (Convo + ReLu), pooling, fully connected, and output. The proposed CNN model workflow is depicted in Figure 4.
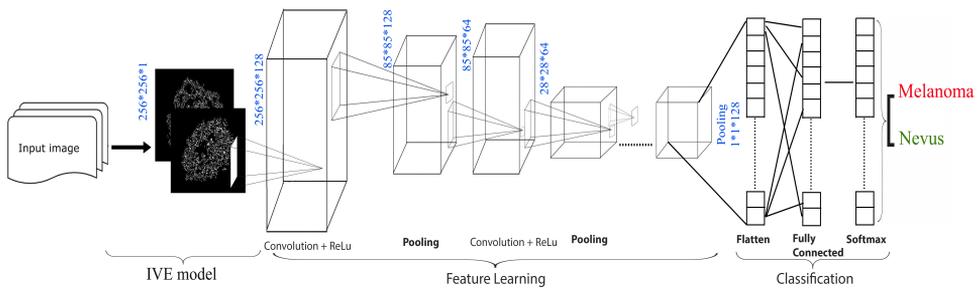
Figure 4. Proposed IVEwCNN model workflow

The CNN model accepts the input as a three-dimensional matrix ($width \times height \times dimension$) to the input layer. Dimension defines the number of channels that are contained in an image; for a grayscale image, this is 1, and for RGB, it is 3. The first layer of the CNN model is the input layer, where an input image with a size of $256 \times 256 \times 1$ is given. Then, the input image is passed to the convolutional layer, where filters are applied to the original images. The filters are slid over the receptive fields of the same input image by a stride and continue through the whole image. The convolutional layer uses the ReLu activation function to zero all negative values. A pooling layer is used after each convolutional layer to reduce the spatial volume of the input image. The fully connected layer takes the output from the previous layer and flattens the inputs to convert them into a single vector. This layer involves weights, biases, neurons, and activation functions, and it is responsible

for updating the weights in the training session. Here, we used the ReLu activation function to produce outputs from each layer. The output layer is the final layer that uses the Softmax activation function to estimate melanoma skin cancer and nevus mole probability.
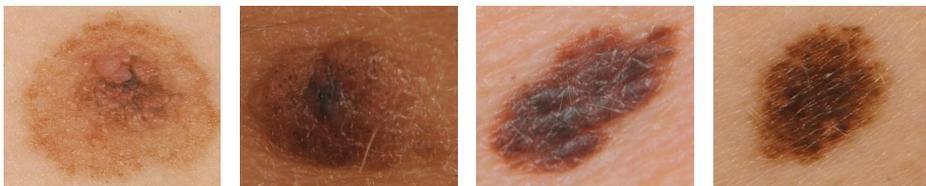
4. Data set

The proposed system was evaluated on the publicly available standard MED-NODE melanoma data set that contains high-resolution skin lesion images [14]. The MED-NODE data set is a subset of the digital image archives of 50,000 images that were collected by the Department of Dermatology at the University Medical Center Groningen (UMCG). A dermatologist examined and assessed all of these pictures as being of the highest caliber [14]. These photographs were captured in the JPEG format with a Nikkor lens on Nikon D3 and D1x cameras from a distance of around 33 cm from the skin lesion areas. Figures 5 and 6 display some images from the MED-NODE data set.

The soundness of the data set was ensured by considering the following criteria [14]:

- MED-NODE data set (170 images) is created by randomly chosen images (relevant patient instances cannot possibly be identified);
- superficial spreading melanoma and nevi are included in data set;
- each picture originated from different patient (aside from image, which shows how disease significantly differs in various bodily areas);
- each picture is sharp and properly exposed to annotate correctly;
- each picture has represented group to which it belongs.
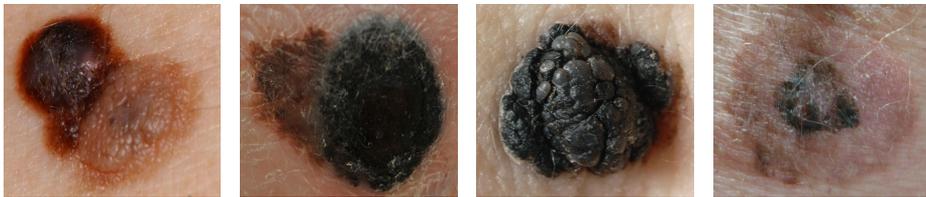
Figure 5. Melanoma skin-cancer images

Figure 6. Nevus mole images

5. Experimental analysis

Some of the data set image resolutions were above $(N_r \times N_c)$ (where $N_r = 256$ and $N_c = 256$), which needs a high cost of computing power. Since rescaled lesion images would be better for deep neural networks [35], we used pre-processing steps to resize the images to $(N_r \times N_c)$ and removed any unwanted artifacts. As the low amount of data could mislead and overfit the deep neural network model, the augmentation technique was applied to increase the amount of data and its diversity [36] to reduce the misleading and overfitting of the data. Then, the IVE model was applied to acquire high intense pixel values from each segmented lesion image, which allowed us to keep important information about the melanoma and nevus mole features. Feeding the significant feature value to CNN achieved a better result.

The authors in [16] discovered that, for a training/validation set, the number of available adjustable parameters for the data set should be inversely proportional to the square of a significant portion of the patterns. As a result, the 80/20 split was chosen to avoid overtraining the deep neural network. Most of the state-of-the-art methods [14, 21, 29, 31] that use the same-sized (or smaller) data sets used 80% of the data for training and 20% for testing purposes. In the experiment, the data-augmentation technique was used to enhance the data set. After the data augmentation, 80% of the images were employed to train the algorithm and the remaining 20% for validation and testing purposes. To avoid biased performance results, we chose average results from the five-fold cross-validation on our proposed model.

5.1. Training and testing

At the beginning of our system, it starts to train our model with 80% training images, and the remaining 20% is used for testing purposes. Our system split the data set such that no overlapping occurred in the training and test data sets. We set 40 as the number of epochs for training our model. For optimization purposes, we used Adam optimization and sparse categorical cross entropy loss to calculate the loss of the CNN model. We implemented our model in the Google cloud platform using a Colab notebook, which provided a 12 GB NVIDIA Tesla K80 GPU that could be used continuously for up to 12 hours and was highly integrated with Google Drive. It also offered TPU recently for free. The full CNN model is depicted in Table 1.

Table 1
Architecture of CNN Layers

Layers	Output Size	Kernel Size	Activation Function
Input Layer	$256 \times 256 \times 1$	–	–
Conv2D	$256 \times 256 \times 128$	7×7	ReLu Activation
MaxPooling2D	$85 \times 85 \times 128$	3×3 (strides = 3)	–
Conv2D	$85 \times 85 \times 64$	4×4	ReLu Activation
MaxPooling2D	$28 \times 28 \times 64$	3×3 (strides = 3)	–

Table 1 cont.

Conv2D	$28 \times 28 \times 32$	5×5	ReLu Activation
MaxPooling2D	$14 \times 14 \times 32$	2×2 (strides = 2)	–
Conv2D	$14 \times 14 \times 128$	6×6	ReLu Activation
MaxPooling2D	$4 \times 4 \times 128$	3×3 (strides = 3)	–
Conv2D	$4 \times 4 \times 32$	5×5	ReLu Activation
MaxPooling2D	$2 \times 2 \times 32$	2×2 (strides = 2)	–
Conv2D	$2 \times 2 \times 128$	6×6	ReLu Activation
MaxPooling2D	$1 \times 1 \times 128$	2×2 (strides = 2)	–
Flatten	128	–	–
Dense182	512	–	ReLu Activation
Dense183	128	–	ReLu Activation
Dense184	64	–	ReLu Activation
Dense185	512	–	ReLu Activation
Dense186	512	–	ReLu Activation
Dense187	64	–	ReLu Activation
Dense188	2	–	Softmax Activation

The proposed CNN architecture is comparable with the other well-known DNN models that are implemented for skin-cancer detection. The DNN model must have the minimum number of these proposed layers with the IVE model in the processed image in order to achieve the best experimental results.

5.2. Performance evaluation

The proposed IVE model was used to estimate the high-intensity pixel values from the training data set, which were incorporated into our developed CNN model to detect and classify melanoma skin cancer and nevus moles. Table 2 shows the five-fold cross-validation experimental results of our proposed model using Canny edge detection, and Table 3 shows the five-fold cross-validation experimental results of our proposed model using intensity value estimation with CNN (IVEwCNN).

Table 2
Five-fold cross-validation experimental results using Canny edge detection with CNN

Metrics	Fold-1	Fold-2	Fold-3	Fold-4	Fold-5	Average
Sensitivity	0.500	0.750	0.375	0.438	0.688	0.550
Specificity	0.789	0.684	0.842	0.737	0.842	0.779
PPV	0.667	0.667	0.667	0.583	0.786	0.674
NPV	0.652	0.765	0.615	0.609	0.762	0.680
Accuracy	0.657	0.714	0.629	0.600	0.771	0.674

Table 3

Five-fold cross-validation experimental results using intensity value estimation
with CNN (IVEwCNN)

Metrics	Fold-1	Fold-2	Fold-3	Fold-4	Fold-5	Average
Sensitivity	1.000	1.000	1.000	0.875	0.813	0.9376
Specificity	0.894	1.000	1.000	0.789	0.895	0.9156
PPV	0.889	1.000	1.000	0.778	0.867	0.9068
NPV	1.000	1.000	1.000	0.894	0.850	0.9488
Accuracy	0.943	1.000	1.000	0.829	0.857	0.9258

It has been demonstrated that pixel intensity levels serve as one of the distinguishing features for identifying an object or region of interest – particularly when considering skin cancer. Textural features are regularly used in image classification because they enhance the classification of nevus and melanoma by computing the irregularities of their structures [25]. Since choosing high-intensity pixels gives more information, the model can easily differentiate melanoma cells from nevus cells more preciously; therefore, the proposed model shows better results when compared to others.

From Tables 2 and 3, we can see that our proposed (IVEwCNN) model performed significantly better in terms of each described performance evaluation metric than the proposed model with Canny edge detection. We used the same threshold value for all of the performance evaluation metrics in this experiment. For an experimental evaluation of the algorithm, the proposed model was compared with some existing works that have been described in [14, 21, 29–31], as they all worked with the same-sized (or smaller) data sets and evaluated their system performance. For performance measurement, the proposed model was evaluated on five commonly employed metrics (sensitivity, specificity, PPV, NPV, and accuracy) that are widely used in classification problems. The performance evaluation metrics that were considered in this work can be defined as follows:

$$Sensitivity = \frac{true\ detected\ melanoma\ cases}{all\ melanoma\ cases} \tag{6}$$

$$Specificity = \frac{true\ detected\ non\ melanoma\ cases}{all\ non\ melanoma\ cases} \tag{7}$$

$$PPV = \frac{true\ detected\ melanoma\ cases}{detected\ melanoma\ cases} \tag{8}$$

$$NPV = \frac{true\ detected\ non\ melanoma\ cases}{detected\ non\ melanoma\ cases} \tag{9}$$

$$Accuracy = \frac{true\ detected\ cases}{all\ cases} \qquad (10)$$

The proposed IVE model performance and the Canny edge-detection technique performances are compared with other state-of-the-art methods that were evaluated on the same-sized (or smaller) data sets that are described in Table 4. The values in Table 4 were updated to two decimal points to compare with other state-of-the-art methods. In Table 4, the results in bold show the experimental results; as can be seen, our proposed methodology (IVEwCNN) shows better performance evaluation metrics. We can conclude that the effective use of the CNN model with a well-processed image generates a superior result when using high-intensity pixel values from segmented lesion skin images.

Table 4

Experimental result evaluation of proposed methodology with the state-of-the-art methods

| Methods	Metrics				
	Sensitivity (Recall)	Specificity	PPV (Precision)	NPV	Accuracy
Texture descriptor [14]	0.62	0.85	0.74	0.77	0.76
Color descriptor [14]	0.74	0.72	0.64	0.81	0.73
Illumination correction [31]	0.81	0.80	0.75	0.86	0.81
Optimized NN using PSO [29]	0.86	0.86	–	–	0.86
S. R. S. Jianu et al. [21]	0.72	0.89	0.87	0.76	0.81
S. Mukherjee et al. (raw feature only) [30]	0.87	0.73	–	–	0.83
S. Mukherjee et al. (PCA feature only) [30]	0.87	0.87	–	–	0.87
Proposed Methodology (Canny Edge detection)	**0.65**	**0.78**	**0.67**	**0.68**	**0.67**
Proposed Methodology (IVEwCNN)	**0.94**	**0.92**	**0.91**	**0.95**	**0.93**

6. Discussion

Instead of conventional visual observations, an efficient expert system was developed to assist expert physicians in the early-stage detection and classification of melanoma skin cancer. The proposed IVEwCNN model adopts the new intensity value estimation (IVE) technique in which high-intensity pixel values are calculated from each segmented lesion image after rescaling the image. Most of the work is done by conducting lesion segmentation and detecting the edges of the segmented lesion, which is then fed into a deep neural network. This contrast enhancement and texture analysis increase the discrimination between the intensity of the values of an image, which improves the overall performance [44]. Our proposed method performs this more precisely and accurately using three different stages. First, a new image is generated in

the pre-processing step that consists of a shape of $(N_r \times N_c)$, where the lower dimension of the segmented image pixels are mapped to retain the original lesion shape. Keeping the lesion shape the same helps retain information about the lesion area, size, and border.

a)   b)   c)   d)
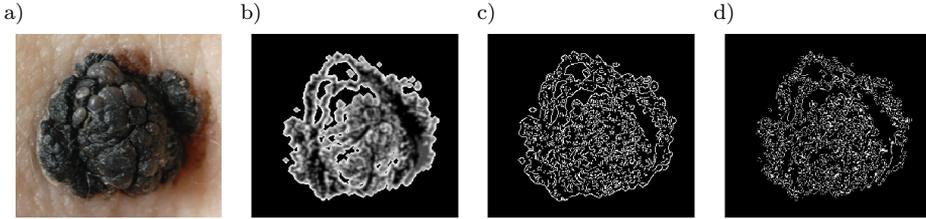
Figure 7. Melanoma skin-cancer images: a) melanoma skin lesion;
b) resized segmented lesion; c) canny edge detection; d) IVE model output

a)   b)   c)   d)
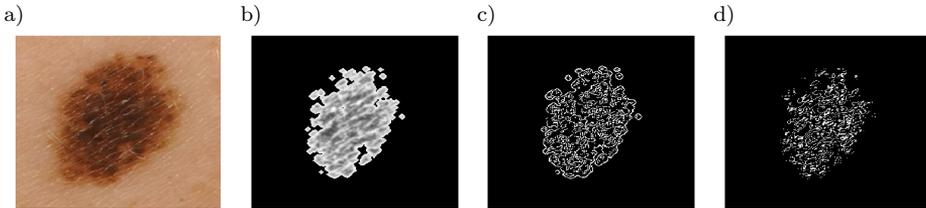
Figure 8. Nevus mole images: a) nevus skin lesion; b) resized segmented lesion;
c) canny edge detection; d) IVE model output

Then, the IVE model is applied (as shown in Figs. 7 and 8) to compute the high-intensity pixels that store the descriptive features of object or region-of-interest identification [25]. The proposed model (IVEwCNN) showed a significant difference between the outcome of Canny edge detection and existing works. Table 2 depicts that the best outcome of Canny edge detection with CNN was 0.75 for sensitivity at Fold-2, 0.842 for specificity at Fold-3 and Fold-5, 0.786 for PPV at Fold-5, 0.765 for NPV at Fold-2, and 0.771 for accuracy at Fold-5. The results of the other existing works are shown in Table 4, where the best results that were achieved were 0.8744 for sensitivity [30], 0.89 for specificity [21], 0.8674 for PPV [21], 0.86 for NPV [31], and 0.8718 for accuracy [30]. Our proposed model exhibited average IVEwCNN outcomes as follows: 0.936 for sensitivity, 0.9156 for specificity, 0.9068 for PPV, 0.9488 for NPV, and 0.9258 for accuracy (as shown in Table 3). Thus, it outperformed the conventional edge detection-based techniques.

7. Conclusion

This paper presents intensity value estimation with a convolutional neural network (IVEwCNN) algorithm for detection and classification. The method led the system to achieve high accuracy, sensitivity, specificity, precision, and NPV when detecting

and classifying melanoma skin cancer and nevus moles. In addition, pre-processed images increase the learnability of any system. Here, we chose to take pixels with higher intensities than a threshold value from a segmented lesion image instead of edge detection. The technique preserved more features than only edge detection; so, this helped our system increase learnability and predict melanoma, skin cancer, and nevus moles more accurately. Our proposed system takes 39 seconds on average to detect and predict melanoma skin cancer and nevus moles. Finally, we considered the five-most-popular performance-evaluation metrics to evaluate the proposed system performance and compared them with some notable existing works on the same-sized (or smaller) data sets. Due to machine limitations, large data sets were not considered in this experiment when evaluating the proposed IVE model; hence, the proposed model's performance was only evaluated by comparing it to the state-of-the-art models that used the same-sized (or smaller) data sets for their models. An experimental comparison shows that the proposed algorithm had better results than all of the state-of-the-art models. The proposed automatic deep-learning system can be implemented in dermatological diagnosis to aid doctors in early melanoma detection. Before going for a clinical application, our proposed system needs to train the model with an improved large data set. Future work could include feature selection, feature dimension reduction, and optimizing the CNN model to increase efficiency and decrease model-training time.

Acknowledgements

References

[1] Adegun A., Viriri S.: An Enhanced Deep Learning Framework for Skin Lesions Segmentation. In: *International Conference on Computational Collective Intelligence*, pp. 414–425, Springer, 2019.

[2] Agilandeeswari L., Sagar M.T., Keerthana N.: Skin Lesion Detection using Texture Based Segmentation and Classification by Convolutional Neural Networks, *International Journal of Innovative Technology and Exploring Engineering*, vol. 9(2), pp. 2117–2121, 2019. doi: 10.35940/ijitee.B7085.129219.

[3] Al-Masni M.A., Al-Antari M.A., Choi M.T., Han S.M., Kim T.S.: 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] American Cancer Society: Melanoma Warning Signs, https://www.skincancer.org/skin-cancer-information/melanoma/melanoma-warning-signs-and-images/, 2021.

[5] American Cancer Society: Key Statistics for Melanoma Skin Cancer, https://www.cancer.org/cancer/melanoma-skin-cancer/about/key-statistics/, 2022.

[6] Argenziano G., Fabbrocini G., Carli P., De Giorgi V., Sammarco E., Delfino M.:
    Epiluminescence microscopy for the diagnosis of doubtful melanocytic skin
    lesions: comparison of the ABCD rule of dermatoscopy and a new 7-point
    checklist based on pattern analysis, *Archives of Dermatology*, vol. 134(12),
    pp. 1563–1570, 1998.

[7] Baldi S., Michailidis I., Ntampasi V., Kosmatopoulos E.B., Papamichail I., Papa-
    georgiou M.: Simulation-based synthesis for approximately optimal urban traffic
    light management. In: *2015 American Control Conference (ACC)*, pp. 868–873,
    IEEE, 2015.

[8] Bora D.J.: Importance of Image Enhancement Techniques in Color Image Seg-
    mentation: A Comprehensive and Comparative Study, *arXiv*, 2017. https:
    //arxiv.org/abs/1708.05081.

[9] Byrd A.L., Belkaid Y., Segre J.A.: The human skin microbiome, *Nature Reviews
    Microbiology*, vol. 16(3), pp. 143–155, 2018.

[10] Ĉadík M.: Perceptual evaluation of color-to-grayscale image conversions. In:
    Computer Graphics Forum, vol. 27, pp. 1745–1754, Wiley Online Library, 2008.

[11] Dong H., Farid F.: A Deep learning based patient care application for skin cancer
    detection, Research Square, 2022. doi: 10.21203/rs.3.rs-1582255/v1.

[12] Eltayef K., Li Y., Liu X.: Detection of melanoma skin cancer in dermoscopy
    images. In: *Journal of Physics: Conference Series*, vol. 787, p. 012034, IOP
    Publishing, 2017.

[13] Garg R., Maheshwari S., Shukla A.: Decision Support System for Detection and
    Classification of Skin Cancer using CNN, *arXiv preprint arXiv:191203798*, 2019.

[14] Giotis I., Molders N., Land S., Biehl M., Jonkman M.F., Petkov N.: MED-NODE:
    a computer-assisted melanoma diagnosis system using non-dermoscopic images,
    Expert Systems with Applications, vol. 42(19), pp. 6578–6585, 2015.

[15] Gogul I., Kumar V.S.: Flower species recognition system using convolution neural
    networks and transfer learning. In: *2017 Fourth International Conference on
    Signal Processing, Communication and Networking (ICSCN)*, pp. 1–6, 2017.

[16] Guyon I.: A scaling law for the validation-set training-set size ratio, *AT&T Bell
    Laboratories*, vol. 1(11), 1997.

[17] Haenssle H.A., Fink C., Schneiderbauer R., Toberer F., Buhl T., Blum A.,
    Kalloo A., *et al.*: Man against machine: diagnostic performance of a deep learning
    convolutional neural network for dermoscopic melanoma recognition in compari-
    son to 58 dermatologists, *Annals of Oncology*, vol. 29(8), pp. 1836–1842, 2018.

[18] Islam R., Imran S., Ashikuzzaman M., Khan M.M.A.: Detection and
    Classification of Brain Tumor Based on Multilevel Segmentation with Convolu-
    tional Neural Network, *Journal of Biomedical Science and Engineering*, vol. 13(4),
    pp. 45–53, 2020.

[19] Jaya V., Gopikakumari R.: IEM: a new image enhancement metric for contrast
    and sharpness measurements, *International Journal of Computer Applications*,
    vol. 79(9), 2013.

[20] Jemal A., Siegel R., Xu J., Ward E.: Cancer Statistics, 2010, *CA: A Cancer Journal for Clinicians*, vol. 60(5), pp. 277–300, 2010. doi: 10.3322/caac.20073.

[21] Jianu S.R.S., Ichim L., Popescu D.: Automatic diagnosis of skin cancer using neural networks. In: *2019 11th International Symposium on Advanced Topics in Electrical Engineering (ATEE)*, pp. 1–4, IEEE, 2019.

[22] Kadampur M.A., Al Riyaee S.: Skin cancer detection: Applying a deep learning based model driven architecture in the cloud for classifying dermal cell images, *Informatics in Medicine Unlocked*, vol. 18, 100282, 2020.

[23] Kanan C., Cottrell G.W.: Color-to-grayscale: does the method matter in image recognition?, *PloS one*, vol. 7(1), e29740, 2012.

[24] Kassani S.H., Kassani P.H.: A comparative study of deep learning architectures on melanoma detection, *Tissue and Cell*, vol. 58, pp. 76–83, 2019.

[25] Khan M.Q., Hussain A., Rehman S.U., Khan U., Maqsood M., Mehmood K., Khan M.A.: Classification of melanoma and nevus in digital images for diagnosis of skin cancer, *IEEE Access*, vol. 7, pp. 90132–90144, 2019.

[26] Koundal D., Sharma B.: Advanced neutrosophic set-based ultrasound image analysis. In: *Neutrosophic Set in Medical Image Analysis*, pp. 51–73, Elsevier, 2019.

[27] Lakshminarayanan A.R., Bhuvaneshwari R., Bhuvaneshwari S., Parthasarathy S., Jeganathan S., Sagayam K.M.: Skin Cancer Prediction Using Machine Learning Algorithms. In: *Artificial Intelligence and Technologies,* Springer, pp. 303–310, 2022.

[28] Maiti A., Chatterjee B.: Improving detection of Melanoma and Naevus with deep neural networks, *Multimedia Tools and Applications*, pp. 1–20, 2019.

[29] Mukherjee S., Adhikari A., Roy M.: Malignant Melanoma Detection Using Multi Layer Perceptron with Optimized Network Parameter Selection by PSO. In: *Contemporary Advances in Innovative and Applicable Information Technology*, pp. 101–109, Springer, 2019.

[30] Mukherjee S., Adhikari A., Roy M.: Malignant melanoma detection using multi layer preceptron with visually imperceptible features and PCA components from MED-NODE dataset, *International Journal of Medical Engineering and Informatics*, vol. 12(2), pp. 151–168, 2020.

[31] Nasr-Esfahani E., Samavi S., Karimi N., Soroushmehr S.M.R., Jafari M.H., Ward K., Najarian K.: Melanoma detection by analysis of clinical images using convolutional neural network. In: *2016 38th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*, pp. 1373–1376, IEEE, 2016.

[32] Peram M.R., Jalalpure S., Kumbar V., Patil S., Joshi S., Bhat K., Diwan P.: Factorial design based curcumin ethosomal nanocarriers for the skin cancer delivery: *in vitro* evaluation, *Journal of Liposome Research*, vol. 29(3), pp. 291–311, 2019.

[33] Poynton C.: Frequently asked questions about color, *Retrieved June*, vol. 19(449), 2004, 1997.

[34] Refianti R., Mutiara A.B., Priyandini R.P.: Classification of Melanoma Skin Cancer using Convolutional Neural Network, *International Journal of Advanced Computer Science and Applications*, vol. 10(3), pp. 409–417, 2019. doi: 10.14569/ IJACSA.2019.0100353.

[35] Salih O., Viriri S.: Skin lesion segmentation using local binary convolution-deconvolution architecture, *Image Analysis & Stereology*, vol. 39(3), pp. 169–185, 2020.

[36] Shorten C., Khoshgoftaar T.M.: A survey on image data augmentation for deep learning, *Journal of Big Data*, vol. 6(1), pp. 1–48, 2019.

[37] Silpa S.R., Chidvila V.: A review on skin cancer, *International Research Journal of Pharmacy*, vol. 4(8), pp. 83–88, 2013.

[38] Ünver H.M., Ayan E.: Skin lesion segmentation in dermoscopic images with combination of YOLO and GrabCut algorithm, *Diagnostics*, vol. 9(3), 72, 2019.

[39] Vijayalakshmi M.: Melanoma Skin Cancer Detection using Image Processing and Machine Learning, *International Journal of Trend in Scientific Research and Development (IJTSRD)*, vol. 3(4), pp. 780–784, 2019.

[40] World Health Organization (WHO): Radiation: Ultraviolet (UV) radiation and skin cancer, https://www.who.int/uv/faq/skincancer/en/index1.html, 2022.

[41] Yadav A.K., Roy R., Kumar V., Kumar A.P.: Survey on content-based image retrieval and texture analysis with applications, *International Journal of Signal Processing, Image Processing and Pattern Recognition*, vol. 7(6), pp. 41–50, 2014.

[42] Zaman K., Bangash J.I., Maghdid S.S., Hassan S., Afridi H., Zohaib M.: Analysis and Classification of Skin Cancer Images using Convolutional Neural Network Approach. In: *2020 4th International Symposium on Multidisciplinary Studies and Innovative Technologies (ISMSIT)*, pp. 1–8, IEEE, 2020.

[43] Zaman K., Maghdid S.S.: Medical Images Classification for Skin Cancer Using Convolutional Neural Network Algorithms, *Advances in Mechanics*, vol. 9(3), pp. 526–541, 2021.

[44] Zheng C., Sun D.W.: Image Segmentation Techniques. In: *Computer Vision Technology for Food Quality Evaluation*, pp. 37–56, Elsevier, 2008. doi: 10.1016/ b978-012373642-0.50005-3.

Affiliations

N. I. Md. Ashafuddula
Department of Computer Science and Engineering, Dhaka University of Engineering & Technology, Gazipur-1707, Bangladesh, md.ashafuddula@gmail.com

Rafiqul Islam
Department of Computer Science and Engineering, Dhaka University of Engineering & Technology, Gazipur-1707, Bangladesh, rafiqul.islam@duet.ac.bd