INDERSCIENCE PUBLISHERS
Linking academia, business and industry through research

International Journal of Bioinformatics Research and Applications

ISSN online: 1744-5493 - ISSN print: 1744-5485
https://www.inderscience.com/ijbra

Skin cancer classification using ensemble classification model with improved deep joint segmentation

Jinu P. Sainudeen, S. Sathyalakshmi

DOI: 10.1504/IJBRA.2025.10065333

Article History:

Received:	20 December 2023
Last revised:	28 December 2023
Accepted:	11 March 2024
Published online:	21 January 2025

Skin cancer classification using ensemble classification model with improved deep joint segmentation

Jinu P. Sainudeen*

Department of Information Technology,
Hindustan Institute of Technology and Science,
Chennai, Rajiv Gandhi Salai (OMR),
Padur, Kelambakkam, Tamil Nadu 603103, India
Email: jinups99@gmail.com
*Corresponding author

S. Sathyalakshmi

Department of Computer Science and Engineering,
Hindustan Institute of Technology and Science,
Chennai, Rajiv Gandhi Salai (OMR), Padur, Kelambakkam,
Tamil Nadu 603103, India
Email: slakshmi@hindustanuniv.ac.in

Abstract: We present a six-phase skin cancer classification model based on improved deep joint segmentation (IDJS) in this work. The pre-processed image is segmented using IDJS in the second phase, after contrast enhancement with assistance from contrast limited adaptive histogram equalisation (CLAHE) in the first phase. The features of GLCM, CCF, LGIP, and median ternary pattern (MTP) are retrieved in the third phase. Data augmentation for the extracted features is carried out in the fourth phase. The fifth phase is ensemble classification using the deep maxout, LSTM, and CNN based on the enhanced data. To determine the final classified label, the enhanced score level fusion receives the output scores from these classifiers. While the RF is 0.9171, Deep Maxout is 0.9382, LSTM is 0.9362, Bi-GRU is 0.8150, RNN is 0.8687, CNN is 0.9382, TL-GOOGLENET is 0.9134, and KNN is 0.9328, respectively, the accuracy of the Ensemble approach is 0.9689.

Keywords: DL; skin cancer; segmentation; classification; recommendation.

Reference to this paper should be made as follows: Sainudeen, J.P. and Sathyalakshmi, S. (2025) 'Skin cancer classification using ensemble classification model with improved deep joint segmentation', *Int. J. Bioinformatics Research and Applications*, Vol. 21, No. 1, pp.72–101.

Biographical notes: Jinu P. Sainudeen received her Bachelor's degree in Information Technology from Cochin University of Science and Technology, Kerala in 2006, and MTech degree in Computer and Information Technology from M.S. University, Thirunelveli, in 2012. She is currently pursuing her PhD degree in Hindustan Institute of Technology and Science, Chennai. Her key research interests include medical, clinical, and biological-image data analysis, with the application of image processing, computer vision, machine learning and deep learning. Her profound interest in research has made her to publish

various papers and articles. She has also participated in various workshops and conferences. She has been with Mangalam College of Engineering since 2006. Currently, she is working as an Associate Professor in the Department of Computer Science and Technology. She is a dedicated and student-focused teaching professional and is committed in providing a well-balanced, supportive and learning environment to students.

S. Sathyalakshmi is working in Hindustan Institute of Technology and Science, Chennai as a Professor and has completed her BE in Computer Science and Engineering in 1989 and ME in Computer Science and Engineering in 2003. She received her PhD in 2017. Her research area is scheduling in grid computing. Her area of interests is machine learning, data analytics and theoretical computer science. She has published 68 papers in various national and international conferences and peer reviewed journals.

1 Introduction

The need for early diagnosis and the increased incidence of skin cancer (Adla et al., 2021; Mijwil, 2021) make it imperative to create a reliable system for automatically classifying skin cancer. The skin is more susceptible to illness since it is the biggest organ in the body and is responsible for shielding other bodily systems. The most frequent cancer in both women and men is melanoma, which had almost 300,000 new cases reported worldwide in that year. Basal cell carcinoma (BCC) (Kassem et al., 2019) and squamous cell carcinoma (SCC) (Balaha and Hassan, 2023), two more serious forms of skin cancer, have a reasonably high prevalence alongside melanoma, with 1 million occurrences each in 2018. Early skin cancer signs are not always obvious, and the accuracy of the diagnosis is frequently reliant (Filali et al., 2020; Thurnhofer-Hemsi and Domínguez, 2020) on the dermatologist's training. An automated diagnostic system is a crucial tool for more precise diagnoses for unskilled practitioners. In addition, making a diagnosis of skin cancer with the naked eye is extremely subjective and seldom generalisable. Thus, it is essential to create an autonomous categorisation system for skin cancer (Arora et al., 2020; Alphonse and Starvin, 2020) which is more precise, more affordable, and easier to detect. Additionally, using such automated diagnostic technologies can significantly reduce skin cancer mortality, which is advantageous for both patients and medical systems.

However, it is difficult to achieve automated categorisation of skin cancer (Kumar et al., 2020; Shorfuzzaman, 2021) due to the diversity and complexity of skin disease images. First of all, there are many similarities between various skin lesions, which might lead to a mistake. As an illustration, SCC and some other skin conditions can mirror BCC in histopathological scans (Qureshi and Roos, 2020). Because of this, it might be challenging for diagnosis methods to distinguish between skin cancers and their recognised mimics. The second difference is that different skin lesions under the same class vary with respect to colour, structure, feature, location and size. For instance, BCC and its subclasses virtually have distinct looks. As a result, it is challenging to identify many subcategories within a single category. The algorithms employed for categorisation are also quite sensitive to the kind of camera used to take the images. The testing images perform worse when they are from various domains.

For classifying skin cancer, machine learning (ML) techniques including support vector machine (SVM), K-nearest neighbour (KNN) (Hatem, 2022), neural network (NN), Naïve Bayes (NB) classifier, and DT have been applied. The demand for human-engineered characteristics in ML techniques is an issue. Due to their capacity for automated feature extraction, decision tree (DL) techniques like CNN gained popularity over the past ten years and have been widely applied in research. An ML technique known as the ensemble combines the decisions of multiple different individual learners to improve classification accuracy. It is anticipated that the ensemble technique would improve classification accuracy since it draws on a variety of individual models to make collective decisions. Numerous studies have attempted to classify skin cancer into binary classes but without greater success. In this paper, we propose an improved deep joint segmentation (IDJS)-based skin cancer detection model with major contributions given below:

1 For segmenting the contrast-enhanced image, we introduced the asymmetry, border, colour, diameter (ABCD)-based IDJS model which gives information about the segmented region contour accuracy and localisation.

2 ILGIP feature is extracted with the gradient magnitude computation process using Sobel Filter based on the segmented image.

3 Range-based score level fusion is done for DeepMaxout, long short term memory (LSTM), and convolutional neural network (CNN) scores fusion to attain the output class labels (0, 1, 2, 3, 4, 5, 6).

The structure of this research project is as follows: Section 1 embraces the foreword of our research. Section 2 embraces the existing work review. Section 3 embraces the presented research. Section 4 embraces an analysis of the presented work. Section 6 embraces the conclusion.

2 Literature review

In 2023, Adla et al. suggested a weight-optimised full-resolution convolutional network (FrCN) for the segmentation of dermoscopy images in this research. A dynamic graph cut algorithm (DGCA) optimised the weight of the network. Weight (hyperparameter) emphasised the appropriate balance amongst the exploration as well as exploitation by fusing the wolves' localised haunting methods with their broader global haunting efforts. The main goal of this project was to create a weight-optimised FrCN model that could accurately identify different forms of skin cancer by utilising dermoscopy images.

In 2023, Xin et al. developed a somewhat straightforward transformer model for classifying skin cancers and evaluating the framework's efficacy. The serialisation of images with multi-scale patch embedding in this study makes use of overlapping and multi-scale sliding windows. To ensure that the encoding outcomes of various datasets are as unlike as possible, this work uses contrast learning to create similar encoding of comparable data for skin cancer. The author would investigate the existing issue with the network's poor speed of processing for large images as well as the usage of transformers to create a combined multimodal classification of skin cancer in future research.

In 2023, Dandu et al. tackled the issue of the categorisation and segmentation of types of melanoma skin cancer. The sixth most prevalent type of skin cancer lesion was

melanoma. In recent years, bio-medical analysis and imaging have grown increasingly intriguing, helpful, and promising as a means of addressing potential difficulties with tissues of melanoma skin cancer that might form on skin surfaces. According to the research, attributes were chosen for classification using a colour layout filter technique.

In 2021, Abdar et al. employed three strategies for quantifying uncertainty to cope with it during the categorisation of skin cancer images. They were deep ensemble (DE), ensemble MC (EMC) dropouts, and Monte Carlo (MC) dropouts, in that order. They presented a unique hybrid dynamic BDL technique that accounted for uncertainty and was dependent on the three-way decision (TWD) theory to further resolve the ambiguity that remained after using the MC, EMC, and DE approaches. With the suggested dynamic model, they were able to use various UQ techniques and DNN in various classification stages. As a result, the components of each step could be changed depending on the dataset being considered.

In 2021, Ali et al. suggested a DL-based deep convolutional neural network (DCNN) model for accurately classifying malignant and benign skin lesions. The author first applied a kernel or filters to eliminate noise and then artifacts; then, normalised the given images and extracted features that aided in correct classification; lastly, data augmentation enhanced the total images, which enhanced the classification rate accuracy.

In 2021, Tumpa and Kabir presented a powerful artificial neural network (ANN) classifier image processing method based on hybrid texture characteristics. The classification findings were more accurate when the characteristics were combined. The hair pixels in the picture were located and then eliminated using a pre-processing method called gradient difference. Following segmentation using Otsu's global thresholding approach, numerous lesion characteristics were extracted using the grey-level co-occurrence matrix (GLCM), local binary pattern (LBP) and ABCD feature extraction methods. The NN was given target values for the input characteristics and used them to determine whether the lesion was benign or malignant.

In 2023, Yang et al. offered a new way to categorise skin cancer in medical skin images. To improve classification performance, class rebalancing was initially performed on the seven different forms of skin cancer images. In the second step of preprocessing, the given image is divided into similar-sized patches and flattened into a collection of tokens. The flattened patches were processed using a transformer encoder in the third step. A normalisation procedure was performed on the input of two different sub-layers, which comprised the fully connected (FC) feedback network unit and the multi-head self-attention unit, and a residual connection between the input as well as output was determined. The transformer encoder was followed by a classification block.

In 2023, Jaisakthi et al. aimed to promote digital skin lesion diagnostics development and explore the applicability of a novel CNN architecture called EfficientNet. They have suggested an automated technique for classifying skin lesions that makes use of dermoscopic images as well as information on patient meta-data. The dataset was very unbalanced and comprised images of varying resolutions, both of which might influence the conclusions. They have conducted a number of trials employing two distinct TL methodologies, such as fine-tuning and feature extractor, to overcome these concerns.

2.1 Review

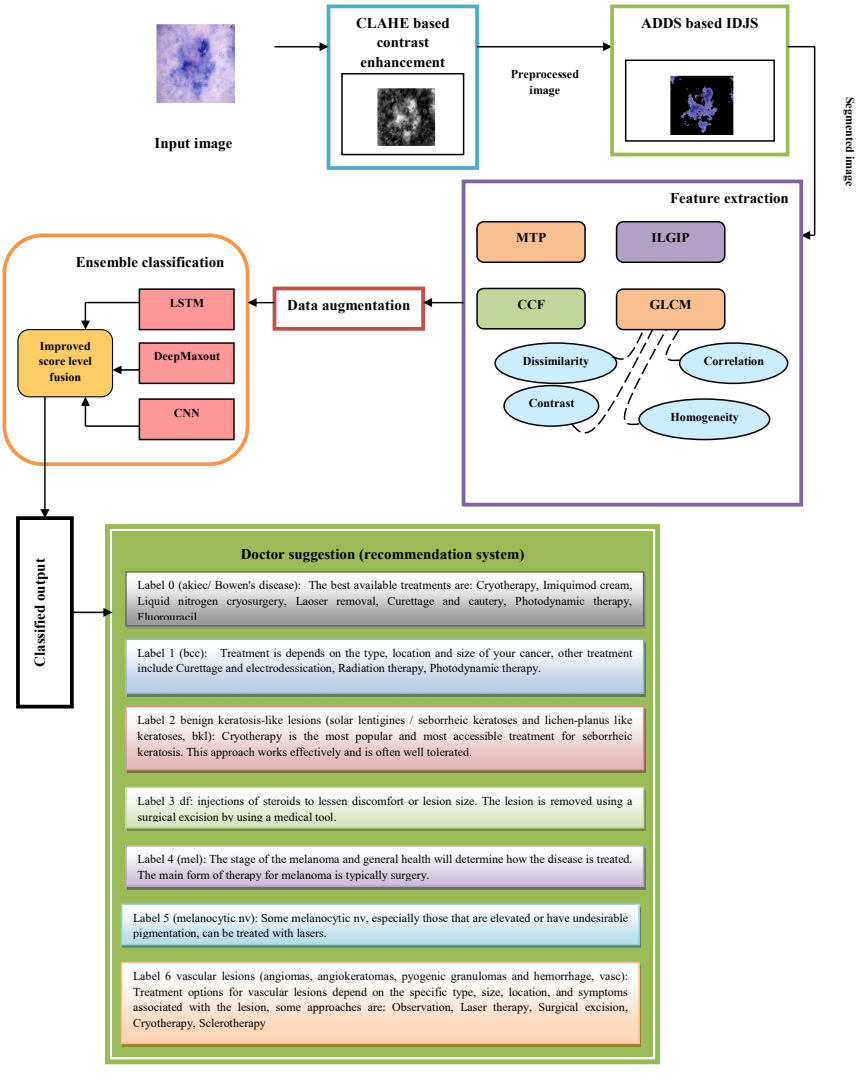
The bulk of skin disease images utilised for training are unbalanced and scarce, making it challenging to achieve automatic skin cancer classification. Additionally, the model

robustness and cross-domain adaptability were important obstacles. In this work, we analysed eight studies Table 1, and some of the difficulties we ran into include the following: limitations of FrCN-DGCA model (Adla et al., 2023) including generalisation and optimisation. In VIT network (Xin et al., 2023), it was challenging to get more data in a medical situation to verify the model performance. Negative transfer problem increased while using colour layout filter approach (Dandu et al., 2023). It was difficult to improve the various DL model knowledge of uncertainty while using TWDBDL model (Abdar et al., 2021). It was challenging to create a DNN (Ali et al., 2021) that could use CAD systems to find different kinds of skin lesions. ANN (Tumpa and Kabir, 2021) needed a lot of processing power. It was difficult to create additional models using pre-trained hyper parameters on a bigger dataset (Yang et al., 2023). Performance-wise, EfficientNets (Jaisakthi et al., 2023) were memory-bound since they were sparse.

Table 1 Traditional methods: features and challenges

Author [citation]	*Methods*	*Features*	*Challenges*
Adla et al. (2023)	Full-resolution convolutional networks-dynamic graph cut algorithm (FrCN-DGCA)	The use of augmented data in testing and training was beneficial for increasing efficiency	Limitations including generalisation and optimisation
Xin et al. (2023)	Vision transformer (VIT) network	Optimal outcomes in the vision field	It was challenging to get more data in a medical situation to verify the model performance
Dandu et al. (2023)	Colour layout filter approach	Efficient output with low accuracy, precision, recall, and ROC values	Negative transfer problem
Abdar et al. (2021)	Three-way decision Bayesian deep learning (TWDBDL)	Highest accuracy and F1 score	It was difficult to improve the various DL model knowledge of uncertainty
Ali et al. (2021)	Deep convolutional neural network (DCNN)	Maximum testing and training accuracy	It was challenging to create a DNN that could use CAD systems to find different kinds of skin lesions
Tumpa and Kabir (2021)	Artificial neural network (ANN)	Highest accuracy	Needed a lot of processing power
Yang et al. (2023)	ViT for skin cancer detection (ViTfSCD)	Provided soft attention on the similar testing as well as training dataset	It was difficult to create additional models using pre-trained hyper parameters on a bigger dataset
Jaisakthi et al. (2023)	EfficientNet	Highest F1 score	Performance-wise, EfficientNets were memory-bound since they were sparse

Figure 1 Broad view of the IDJS based skin cancer classification model (see online version for colours)



3 ASSD-based IDJS: ASSD-based improved deep joint segmentation based skin cancer classification

Both the ML and DL algorithms performance to conduct tasks like detection, segmentation and classification, as well as the medical pictures utilised to aid in diagnosis, have significantly improved recently.Accurate skin cancer diagnosis can be improved with image detection model and computer classification skills. The ASSD based IDJS-based skin cancer classification model is shown in broad view in Figure 1.We

describe the ASSD based IDJS-based skin cancer classification model in this work, which consists of six phases, including:

- preprocessing
- segmentation (improved)
- feature extraction
- data augmentation
- classification
- recommendation.

3.1 Preprocessing

The method of acquiring images must be non-uniform in a number of ways. As a result, the primary objective of the pre-processing stage is to improve the image's quality, by eliminating undesired elements from the foreground or background. In order to achieve its goals, contrast enhancement models have developed during the past few decades.In order to disperse the pixel intensity, our provided image I has been preprocessed using the contrast enhancement technique known as contrast limited adaptive histogram equalisation (CLAHE) (Nguyen et al., 2020). Instead of processing the full image, CLAHE works with discrete sections of it called tiles. Then, to eliminate the arbitrary borders, the adjacent tiles are blended employing the bilinear interpolation. The preprocessed image obtained by CLAHE is denoted as I^P.

3.2 Segmentation (ASSD based IDJS)

The process of dividing the input image ROI is known as segmentation. By assigning a comparable characteristic to each pixel in the input image, this separation may be achieved. The key benefit in this situation is that an input image that has been segmented into segments may be processed rather than the complete image. Figure 2 depicts ASSD based IDJS model. The preprocessed skin cancer image I^P is segmented using the IDJS model. Following are the IDJS steps:

3.2.1 Formation of the grids (Mahesh and Renjit, 2020)

Initially I^P is divided as numerous grids having the dimension of 2 × 2. The I^P divided grids are indicated as $g = (g_1, g_2,\ldots,g_n)$, in which n is the grid total. Once the grids have been generated, the intra grid points pixels are connected as $g_j = \frac{\sum_{i=1}^{N} p_i}{N} \pm t$ based on the mean and threshold computations. The average $g_j = \frac{\sum_{i=1}^{N} p_i}{N}$ pixel value in the grids is used to get the mean value. The next step is to identify which pixels should be connected

by applying a predetermined threshold value to the mean value. Here, the threshold value t in this case is set to 1.

3.2.2 Improved region fusion phase

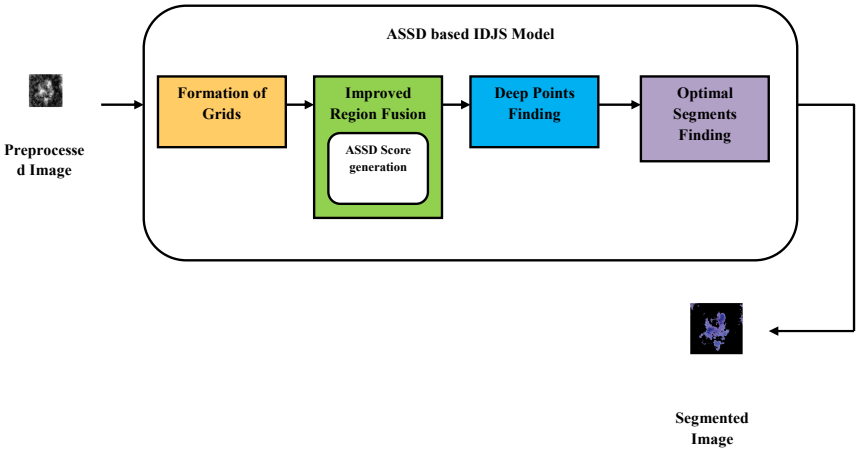
Depending on the pixel intensity along with the bi-constraint similarities, region fusion determines region similarity. If the mean $M_j < 3$ and just single grip point is selected for the each grid, we identify region similarity. The formula (1) is used to calculate region similarity, in which P_o^l is joined pixel having R count. Mapped points, which are the names given to the combined grids, are depicted as $H = (H_1, H_2, \ldots, H_m)$.

$$M_j = \frac{\sum_{o=l}^{W} P_o^Y}{R} \tag{1}$$

As per the improved logic, fuse the region having the least M_j for creating the mapped points. And also compute $ASSD_j$ value using formula (2) to obtain the mapped points. As a result, the contour accuracy as well as the localisation of segmented regions is given in more detail. The ASSD score identifies the separation between two regions borders.

$$ASSD = \frac{\sum d(x, \partial K) + \sum d(y, \partial G)}{|\partial K| + |\partial G|} \tag{2}$$

Figure 2 ASSD based IDJS model (see online version for colours)



3.2.3 Deep points finding

By taking into account both the mapped points and the lost pixel from the joining step, deep points are discovered in formula (3), in which Z is missed pixel total. As a result, formula (4) is used to calculate the deep points by combining the mapped and missing pixels.

$$E = \{P_H\}; 1 < H \le Z \tag{3}$$

$$Pts_{deep} = E + H_c \tag{4}$$

3.2.4 Optimal segments finding

First, a random selection is made among the segmented points *S*. From among the 20 deep points, three various segmented points are chosen at random. Following that, the deep, segmented points, and points with the smallest distance, as specified in formula (5), are picked as the current new segmented points. The procedure is repeated, and the best-segmented points are determined using the MSE score, which evaluates the two grids similarities. Finally, the segmented skin cancer disease image is denoted as I_S.

$$X_{Dis} = \sqrt{\sum_{c=1}^{20} \left(S_c - Pts_{deep}\right)^2} \tag{5}$$

3.3 Feature extraction

The feature extraction method is regarded to be the most crucial phase in the classification process. The process of extracting pertinent features from an input segmented dataset for use in subsequent calculations like classification and detection is referred to as feature extraction. In this phase, features such as median ternary pattern (MTP), ILGIP, GLCM, and CCF are extracted based on the I_S.

3.3.1 MTP feature (Khan et al., 2013)

When there is noise, MF still works. Therefore, a strategy that takes use of the median is more resistant to noise and so obtains reliable data. Due to noise or changes in light, a pixel's intensity might change. As a result, a window is specified to create a 3-valued code known as a ternary pattern rather than a 2-valued code. In almost uniform locations, noise is defeated by it. Based on the median value of pixel intensities in a ternary pattern, MTP is a texture descriptor that describes local texture patterns. It is helpful for recording texture information in pictures, especially when patterns show variations in strength. MTP is appropriate for a variety of image analysis applications, including texture classification and segmentation, since it can offer robustness against noise and variations in light. The MTP in formula (6) combines the benefits of median as well as intensity values quantisation, in which M_c refers to local median, v refers to grey level of neighbour, and t refers to threshold.

$$MTP(v) = \begin{cases} 1; v > M_c + t \\ 0; M_c - t \le v \le M_c + t \\ -1; v < M_c - t \end{cases} \tag{6}$$

3.3.2 ILGIP feature

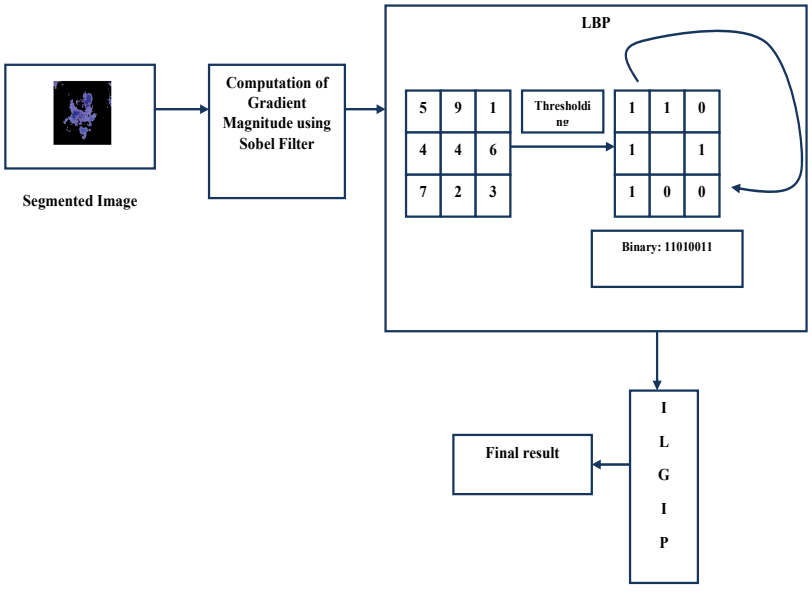
The goal of the feature extraction technique known as LGIP is to locate areas of an image that have an increasing gradient. This capability is especially useful for jobs like object detection and recognition when edge or gradient information is crucial. Localised gradient intensity and direction fluctuations are captured by LGIP, which facilitates

object or region differentiation within an image. ILGIP encode the pixel intensity growing nature at each pixel in eight different directions by employing the eight different binary bits and assign the decimal code for describing the total increasing growth. Figure 3 represents the ILGIP feature descriptor. The following are the ILGIP steps:

- *Step 1 (computation of gradient magnitude using sobel filter):* Apply the gradient filtering (Holder and Tapamo, 2017) for the segmented image I_S using *Sobel*() function given in formula (7) from the open CV library, in which $\frac{\partial f}{\partial x}, \frac{\partial f}{\partial y}$ refers to the derivative in terms of *x*, *y* (Gradient in direction of *x*, *y*). At each pixel, compute the gradient vector via convolving the segmented image with the vertical as well as horizontal derivative filters. i.e., gradient of *x* and gradient of *y* by the *cv*2 · *Sobel*() function

$$\nabla f = \begin{bmatrix} g_x \\ g_y \end{bmatrix} = \begin{bmatrix} \frac{\partial f}{\partial x} \\ \frac{\partial f}{\partial y} \end{bmatrix}; \frac{\partial f}{\partial x} = S_x \otimes f, \frac{\partial f}{\partial y} = S_y \otimes f \tag{7}$$

Figure 3 ILGIP feature descriptor (see online version for colours)



- *Step 2:* At each pixel, calculation of gradient magnitude using formula (8).

$$MOG = \sqrt{Gx^2 + Gy^2} \tag{8}$$

- *Step 3:* Assign the nonlinear enhancement operation.
- *Step 4:* Normalisation of the enhanced magnitude within the range of [0 to 255].

- *Step 5:* The LBP operator receives the result of the computation of gradient magnitude when it has been completed. LBP feature given in formula (9) apply each pixel having the eight different binary bits via the comparison of the pixel intensity having its eight different neighbours and finally converting it to the decimal vale, in which i_p, i_c refers to neighbour of p^{th} and centre pixel intensity.

$$LBP(x_c, y_c) = \sum_{p=0}^{7} s(i_p - i_c) 2^p; s(x) = \begin{cases} 1, x \geq 0 \\ 0, x < 0 \end{cases} \tag{9}$$

- *Step 6:* The output of LBP (Zhou and Wang, 2012) is then given to the ILGIP as an input for further processing. ILGIP performs better than LBP and LDP because to its numerous advantageous characteristics. ILGIP, which is simple, quick, and robust to monotonic intensity variations initially, inherits the advantages of traditional LBP. Second, LGIP uses uniform and gradient patterns to its advantage. One thing to note is that a bigger LGIP code suggests a more skewed gradient in the OTV direction for the local intensity. Additionally, the table's centre does not contain any significant LGIP patterns. They often form part of regular patterns and depict the most useful microstructures, such as corners and edges. Third, LGIP maintains a fair level of stability even in the face of non-monotonic intensity and white noise variations, such as LDP. ILGIP's dimensionality is smaller than that of LBP and LDP. The computation of image-to-image distance can be sped faster.

Table 2 GLCM feature (*GLCM* = [*ASM Ct C D H*])

GLCM feature	*Description*	*Formula*
ASM (Mohanaiah et al., 2013)	It is the squared sum of all GLCM entries. ASM assesses the homogeneity of the given image. When the homogeneity of the image is excellent or the similarity of the pixels, ASM is high	$ASM = \sum_{i=0}^{N_g-1} \sum_{j=0}^{N_g-1} p_{ij}^2$
Contrast (Zubair and Alo, 2019)	It measures the difference among the lower and higher values in the adjacent group of pixels	$Ct = \sum_{i=0}^{N_g-1} \sum_{j=0}^{N_g-1} (i-j)^2 p(i,j)$
Correlation (Mohanaiah et al., 2013)	The linear relationship between the grey levels of nearby pixels is measured through correlation	$C = \frac{\sum_{i=0}^{N_g-1} \sum_{j=0}^{N_g-1} (i,j) p(i,j) - \mu_x \mu_y}{\sigma_x \sigma_y}$
Dissimilarity (Zubair and Alo, 2019)	Dissimilarity is a metric for the distance of pixel pairs inside the ROI	$D = \sum_{i=0}^{N_g-1} \sum_{j=0}^{N_g-1} \lvert i-j \rvert p(i,j)$
Homogeneity (Zubair and Alo, 2019)	Homogeneity scales the local variations in texture of image and represents the of image textures homogeneity	$H = \sum_{i=0}^{N_g-1} \sum_{j=0}^{N_g-1} \frac{p(i,j)}{1+(i-j)^2}$

3.3.3 GLCM feature

A popular method for texture analysis and feature extraction is GLCM. Based on the grey-level values of the pixels in a pair, it describes the spatial relationships between them in an image. Texture pattern characteristics including homogeneity, contrast, and

entropy are captured by GLCM features and are crucial for a variety of image processing applications like segmentation, texture classification, and object recognition. GLCM feature which are extracted in this paper are represented in Table 2.

3.3.4 CCF (Said and Khurshid, 2017)

Specific colour information may be derived by estimating the likelihood that a certain pixel will appear. The colour co-occurrence matrix is another way to express the spatial data in an input image. It is determined from the input image, and then CCF is determined in formula (10) using the matrix of colour co-occurrences. CCF uses a variety of colour spaces, such as RGB, grey, LAB, etc. The CCF feature is computed as L-CCF mean, a-CCF mean, b-CCF mean, L-CCF SD, a-CCF SD, b-CCF SD. By taking into account the co-occurrence of colour values within predetermined distances and angles, CCF expands the idea of GLCM to colour images. By capturing the spatial correlations between colour value pairs, it adds to our understanding of the colour texture and structure of an image. Colour image analysis activities including material detection, content-based image retrieval, and image retrieval can benefit from the usage of CCF characteristics.

$$CCF\left(p_1, p_2\right) = p_r \left\{ \begin{array}{l} \tilde{j}_{\min}(a,b) = p_1, \tilde{j}_{\max}(a,b) = p_2 \middle| a = 1,2,...,\dfrac{M^i}{m}; \\ b = 1,2,...,\dfrac{N}{n} \end{array} \right\} \tag{10}$$

Finally, the extracted features are denoted as:

$$F\left[MTP \quad ILGIP \quad GLCM \quad CCF\right]$$

3.4 Data augmentation

Based on the extracted features *F*, image data augmentation is carried out, which is the process of creating new, altered versions of the various images from the provided image dataset to boost its variety. We use the 'randomised subset' to create the random sample. The array's random subset of elements is chosen, and they are modified within predetermined constraints. The augmented data is denoted as *D*.

3.5 Ensemble classification

In ensemble modeling, many distinct base models are employed to forecast a result. To lower the forecast generalisation error, ensemble models are used. When using the ensemble technique, the prediction error lowers as long as the basis models are varied and independent. In this paper, we use the ensemble classifiers including DeepMaxout, LSTM, and CNN based on the augmented data *D*. The output scores of these classifiers is subjected to the improved score level fusion process to achieve the final classified output.

3.5.1 DeepMaxout (Peta and Koppu, 2022)

The network classifier uses the input value to categorise the skin cancer condition using clinical information. It makes efficient use of a dropout and a maxout layer to improve classification performance. The suggested model's resilience is increased by the classifier's activation function. Input, and dropout, embedding layers as well as max-pooling, convolution, maxout functions and activation, dense layers are among the layers that make up the network design. Formula (11) illustrates how this network's maxout unit works, in which η is the feature map, λ is weight, and γ is bias.

$$Q(D) = \underset{z\in[1,\eta]}{Max}\ R_{yz};\ R_{yz} = D\cdot\lambda_{yz} + \gamma_{yz} \tag{11}$$

The embedding layer computes the output after receiving the input D from the input layer and processing it. The convolution layer, third convolution layer output, and dropout layer are all placed after each other. The result is produced by a max-pooling layer that comes after the convolution layer. The dropout layer follows the dense layer in order. The preceding dropout layer's input is used by the maxout unit to output computation, which is then contributed to the dense layer to produce the final product. The activation function then calculates the classification result DM^{out} using the dense layer's output.

Advantages of deep maxout network

- The requirement for human feature engineering is diminished by Deep Maxout networks' ability to automatically extract meaningful representations or features from raw data.
- Natural language processing, image identification, and audio recognition are just a few of the fields in which DeepMaxout networks have been effectively used.

3.5.2 LSTM (Zhou et al., 2019)

In tasks requiring sequence prediction, LSTM, a subclass of RNNs, does exceptionally well at learning the long-term dependencies. Because it enables the learning of several parameters, the LSTM cell increases long-term memory in a way that is more performing. The input gate, forget gate, output gate and candidate cell value are the four crucial components that make up an LSTM unit. Formula (12) may be used to compute memory cell along with the output based on these components, in which σ states logistic sigmoid, b states bias, (^{o}) states dot product, h states hidden layer, C states cell, and W is weight. The output score obtained from the LSTM classifier is denoted as L^{out}.

$$\left.\begin{aligned}
f_t &= \sigma\left(W_{hf}\cdot h_{t-1} + W_{Df}\cdot D_t + b_f\right)\\
i_t &= \sigma\left(W_{hi}\cdot h_{t-1} + W_{Di}\cdot D_t + b_i\right)\\
\tilde{C}_t &= \tanh\left(W_{hC}\cdot h_{t-1} + W_{DC}\cdot D_t + b_C\right)\\
C_t &= f_t^{\circ} C_{t-1} + i_t^{\circ}\tilde{C}_t\\
O_t &= \sigma\left(W_{hO}\cdot h_{t-1} + W_{DO}\cdot D_t + b_O\right)\\
h_t &= O_t^{\circ}\tanh\left(C_t\right)
\end{aligned}\right\} \tag{12}$$

Advantages of LSTM

- LSTMs are excellent for a variety of applications, including sentiment analysis, machine translation, speech recognition, and more.
- They can handle input sequences of different lengths.
- Natural language processing applications like text creation, machine translation, sentiment analysis, and audio recognition frequently make use of LSTMs.
- They are also used in time-series prediction jobs such as weather and stock market forecasting.

3.5.3 CNN (Fu'adah et al., 2020)

The MLP has been developed into CNN, which is intended to handle two-dimensional input. Due to its extensive use with input data and high network depth, CNN is considered a form of DNN. CNN's design is similar to that of NN in general; its neurons contain functions for activation, bias, and weight. The pooling layer serves as the layer responsible for feature extraction in CNN design, while the convolution layer having the ReLU activation serves as the classification layer.

- *Convolution layer:* Convolution is the primary procedure that supports CNN in the Convolution layer. The input augmented data will be processed as the first source system model by the convolution layer. A feature map, also known as an extractor filter, will be applied to the input for extracting the input features.
- *Activation:* ReLU is an activation layer used in CNN to speed up the training phase of the NN, which has the benefit of reducing errors. When ReLU activation occurs if the image pixel has a value less than zero, all of the pixels become zero.
- *Pooling layer:* In the CNN approach, pooling layers are often placed after a number of convolutional layers. The pooling layer has a number of benefits and may gradually lower the output volume size on the feature map to control over-fitting. Utilising maximum or mean pooling, the pooling layer is utilised to minimise data.While the primary pooling determines the average value, the max-pooling will choose the highest value.
- *FC layer:* The MLP uses a layer at the very end of its design called FC. This layer will link every neuron from the activation layer that came before it. At this point, all input layer neurons must undergo a flattening procedure, or conversion to 1D data. After that, a variant of the logistic regression approach called Softmax activation may be used to categorise more than the two different labels. The output score obtained from the CNN classifier is denoted as C^{out}.

Advantages of CNN

- CNNs are widely used in computer vision tasks such as object identification, facial recognition, picture classification, and medical image analysis.
- They have also found utility in natural language processing applications such as text classification and sentiment analysis.

3.5.4 Range based score level fusion

In order to identify the class label, improved score level fusion may be able to successfully fuse the match scores produced by the LSTM, DeepMaxout, and CNN. Conventionally, Min-Max normalisation is employed for fusing the scores as defined in formula (13). Improved score level fusion is defined as in equation (14), in which $New_{\min}$, $New_{\max}$ refers to the maximum and minimum desired value of normalised range. We get the output score (accuracy) of each classifier. From the output accuracy score, the highest output is selected. i.e.,

$$Q(x) = \arg\max_{k=1}^{K} y_i(x),$$

in which k is classifier count, y_i is i^{th} classifier output score, $Q(x)$ is maximum score index. Assign the weight function

$$W = \left\lfloor v_1 = 1/3, v_2 = 1/3, v_3 = 1/3 \right\rfloor$$

for three classifiers initially. Based on $Q(x)$, the weights of index are set as $[v_1 + v_2]$. Remaining index weights are $v_{1/2}$ and $v_{1/2}$. For example, second index has maximum score means weight may be changed to

$$W = \left[v_1 = \frac{1/3}{2}, v_2 = 1/3 + 1/3, v_3 = \frac{1/3}{2} \right].$$

i.e.,

$$\sum_{i=1}^{k} W = 1$$

$$SC_i = \frac{SC_i - Min_{SC_i}}{Max_{SC_i} - Min_{SC_i}}$$

$$Fusion\_Score = \sum_{i=1}^{M} \left(SC_{DM^{out}} + SC_{L^{out}} + SC_{C^{out}} \right) \tag{13}$$

$$SC_i = \frac{SC_i - Min_{SC_i}}{Max_{SC_i} - Min_{SC_i}} * \left(New_{\max} - New_{\min} \right) + New_{\min}$$

$$Fusion\_Score = \sum_{i=1}^{M} \left(W_1 * SC_{DM^{out}} + W_2 * SC_{L^{out}} + W_3 * SC_{C^{out}} \right) \tag{14}$$

3.6 Recommendation

Doctors offer suggestions (recommendations) to patients depending on the input dataset based on the categorised output.

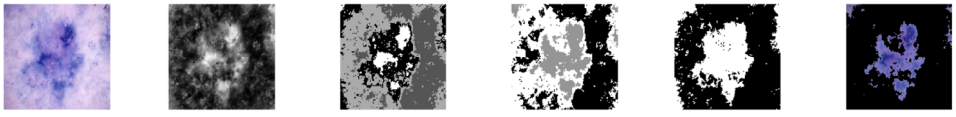
- *Label 0 means (akiec/ Bowen's disease):* The best available treatments are: cryotherapy, imiquimod cream, liquid nitrogen cryosurgery, laoser removal, curettage and cautery, photodynamic therapy, fluorouracil.
- *Label 1 means (bcc):* Treatment is depending on the type, location and size of your cancer, other treatment includes curettage and electrodessication, radiation therapy, photodynamic therapy.
- *Label 2 means benign keratosis-like lesions (solar lentigines/seborrheic keratoses and lichen-planus like keratoses, bkl):* Cryotherapy is the most popular and most accessible treatment for seborrheic keratosis. This approach works effectively and is often well tolerated.
- *Label 3 means df:* injections of steroids to lessen discomfort or lesion size. The lesion is removed using a surgical excision by using a medical tool.
- *Label 4 means (mel):* The stage of the melanoma and general health will determine how the disease is treated. The main form of therapy for melanoma is typically surgery.
- *Label 5 means (melanocytic nv):* Some melanocytic nv, especially those that are elevated or have undesirable pigmentation, can be treated with lasers.
- *Label 6 means vascular lesions (angiomas, angiokeratomas, pyogenic granulomas and hemorrhage, vasc):* Treatment options for vascular lesions depend on the specific type, size, location, and symptoms associated with the lesion, some approaches are: observation, laser therapy, surgical excision, cryotherapy, sclerotherapy.

4 Result analyses on ensemble model for skin cancer classification

4.1 Simulation procedure

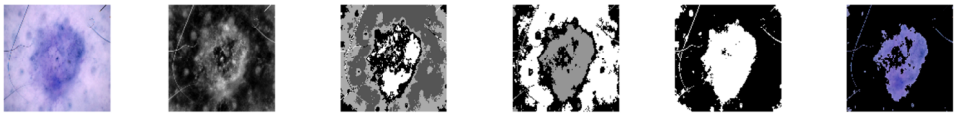
The proposed skin cancer classification approach was simulated in PYTHON. Moreover, the python version is 'PYTHON 3.7' and processor used is 'AMD Ryzen 5 3450U with Radeon Vega Mobile Gfx 2.10GHz and the installed Ram size is '16.0 GB (13.9 GB usable)'. Further, the Skin Cancer MNIST: HAM10000 dataset was accumulated from https://www.kaggle.com. In order to classify the skin cancer accurately, the Ensemble model is contrasted over the RF, Deep Maxout, LSTM, Bi-GRU, RNN, CNN, TL-GOOGLENET (Kassem et al., 2019) and KNN (Hatem, 2022). Additionally, it was evaluated with respect FNR, F-measure, accuracy and other metrics. Further, the skin cancer classification images for seven varied types of classes are shown from Figure 4 to Figure 10.

Figure 4 Skin cancer classification images for Actinic keratoses and intraepithelial carcinoma, (a) input image (b) preprocessed image (c) FCM image (d) K-means image (e) conventional DJS image (f) IDJS image (see online version for colours)



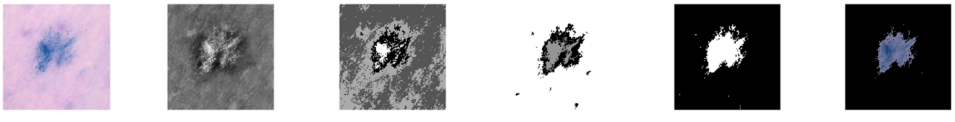
(a) (b) (c) (d) (e) (f)

Figure 5 Skin cancer classification images for BCC, (a) input image (b) preprocessed image (c) FCM image (d) K-means image (e) conventional DJS image (f) IDJS image (see online version for colours)



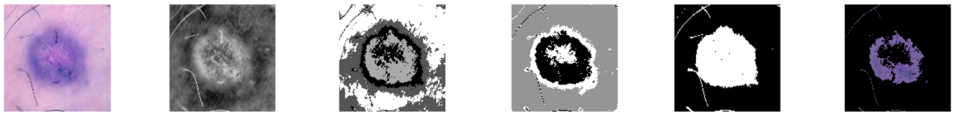
(a) (b) (c) (d) (e) (f)

Figure 6 Skin cancer classification images for benign keratosis-like lesions, (a) input image (b) preprocessed image (c) FCM image (d) K-means image (e) conventional DJS image (f) IDJS image (see online version for colours)



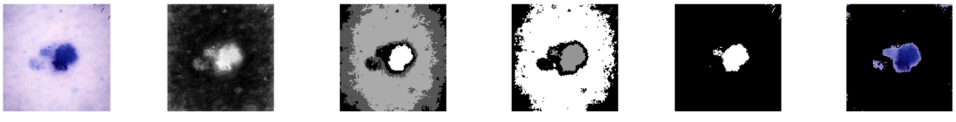
(a) (b) (c) (d) (e) (f)

Figure 7 Skin cancer classification images for dermatofibroma, (a) input image (b) preprocessed image (c) FCM image (d) K-means image (e) conventional DJS image (f) IDJS image (see online version for colours)



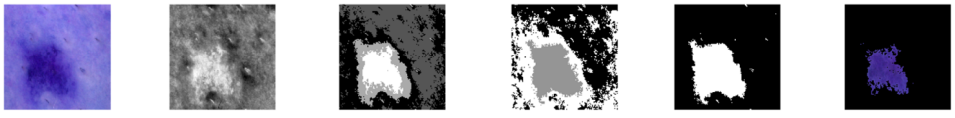
(a) (b) (c) (d) (e) (f)

Figure 8 Skin cancer classification images for melanoma, (a) input image (b) preprocessed image (c) FCM image (d) K-means image (e) conventional DJS image (f) IDJS image (see online version for colours)



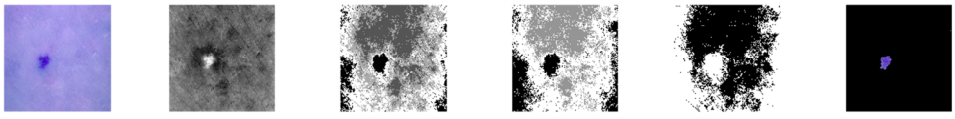
(a) (b) (c) (d) (e) (f)

Figure 9 Skin cancer classification images for melanocytic nevi, (a) input image (b) preprocessed image (c) FCM image (d) K-means image (e) conventional DJS image (f) IDJS image (see online version for colours)



(a) (b) (c) (d) (e) (f)

Figure 10 Skin cancer classification images for vascular lesions, (a) input image (b) preprocessed image (c) FCM image (d) K-means image (e) conventional DJS image (f) IDJS image (see online version for colours)



(a) (b) (c) (d) (e) (f)

4.2 Skin Cancer MNIST: HAM10000 dataset description

Skin Cancer MNIST: HAM10000 dataset consists of 10,015 images. Here, the total number of images was split in to two types a) Part1 and b) Part2. Further, the part1 has 5000 images and the part2 has 5015 images. Moreover, 70% of images used for testing and 30% of images used for training. This dataset includes 7 different classes, they are described below

- Actinic keratoses and intraepithelial carcinoma / Bowen's disease (akiec)
- BCC
- benign keratosis-like lesions (solar lentigines / seborrheic keratoses and lichen-planus like keratoses, bkl)
- dermatofibroma (df)
- melanoma (mel)
- melanocytic nevi (nv)
- vascular lesions (angiomas, angiokeratomas, pyogenic granulomas, hemorrhage, vasc).

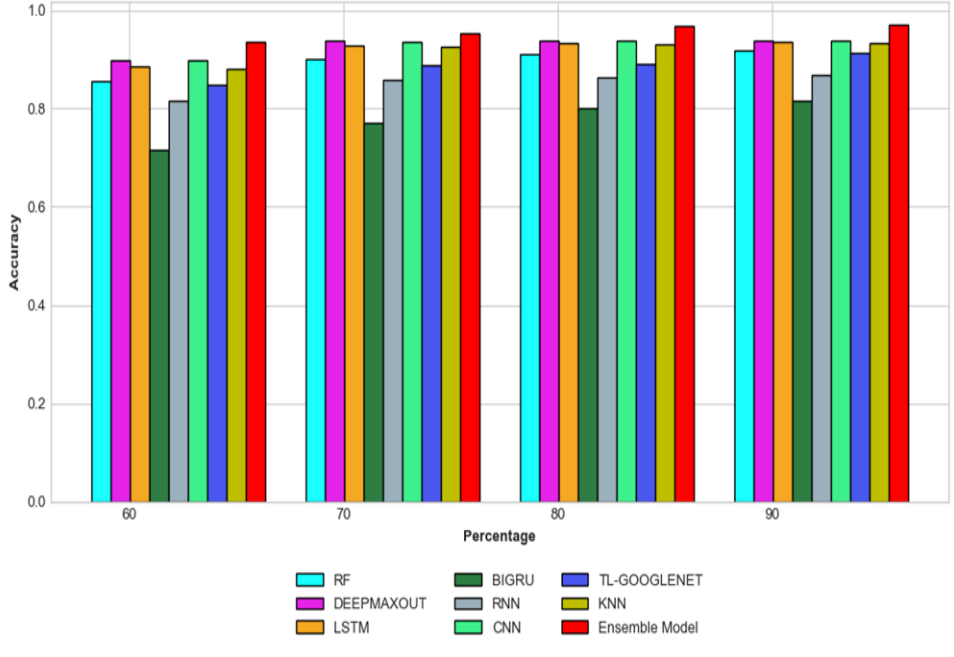
4.3 Examination on Ensemble model and conventional schemes regarding positive metric for skin cancer classification

The evaluation on positive metric of the Ensemble model over the RF, deep maxout, LSTM, Bi-GRU, RNN, CNN, TL-GOOGLENET and KNN for the classification of skin cancer is exposed in Figure 11. Further, it is assessed while modifying the training rate from 60–90. Here, the model should attain greater positive metric ratings for the exact classification of skin cancer. Similarly, the Ensemble model yielded better results over

the conventional strategies. In particular, the Ensemble model accomplished obtained the maximal accuracy rate of 96.8863 at the training rate 90%, whilst the traditional schemes scored the least accuracy value, notably, RF = 91.7141, deep maxout = 93.8157, LSTM = 93.6246, Bi-GRU = 81.4996, RNN = 86.8703, CNN=93.8238, TL-GOOGLENET = 91.3383 and KNN = 93.2773, correspondingly. In addition, at the training rate 70%, the Ensemble model generated the highest precision (95.4393) than the RF, Deep Maxout, LSTM, Bi-GRU, RNN, CNN, TL-GOOGLENET and KNN, respectively.

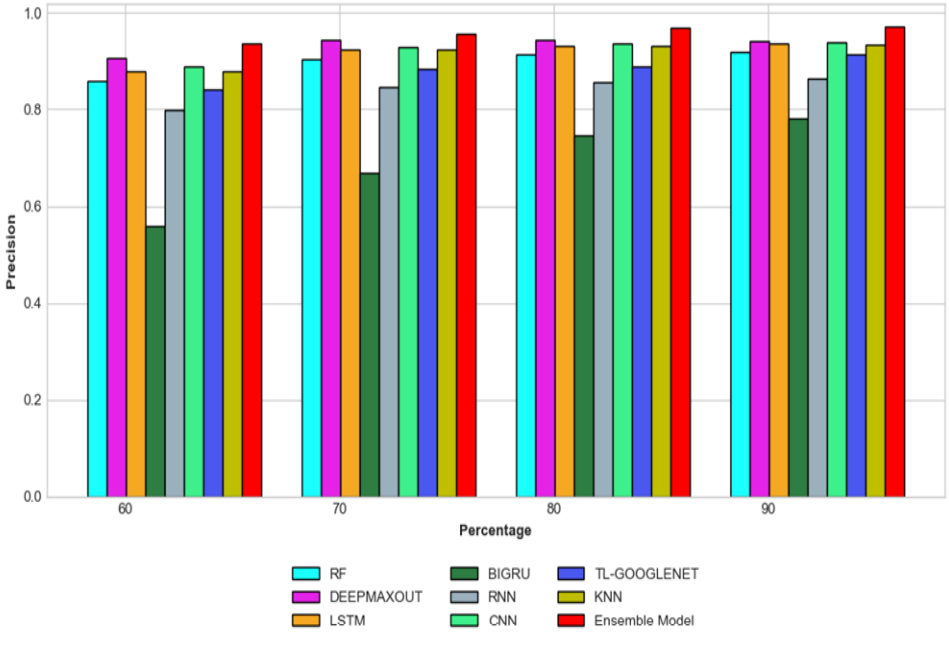
Furthermore, the Ensemble model and the traditional strategies are examined with respect to sensitivity and specificity is depicted in Figure 11(a) and Figure 11(b). For instance, the Ensemble model attained the highest sensitivity is attained by the RF, deep maxout, LSTM, CNN and KNN (learning rate =80%), i.e. greater than 90%, however, the Ensemble model recorded the maximised sensitivity of 96.7486. Finally, for the 90% training percentage, the specificity of the Ensemble model is 96.8427, mean while the RF (91.6762), deep maxout (93.9166), LSTM (93.3579), Bi-GRU (83.9288), RNN (87.2514), CNN (93.5474), TL-GOOGLENET (91.0632) and KNN (93.1127) attained minimised specificity value. The introduction of improved segmentation process paves way to the better feature extraction approach, which has produced greater outcomes. Moreover, the ensemble model (deep maxout, LSTM and CNN) also aided in improved classification of skin cancer.

Figure 11 Assessment on Ensemble model and the traditional strategies with respect to positive metric for skin cancer classification (see online version for colours)

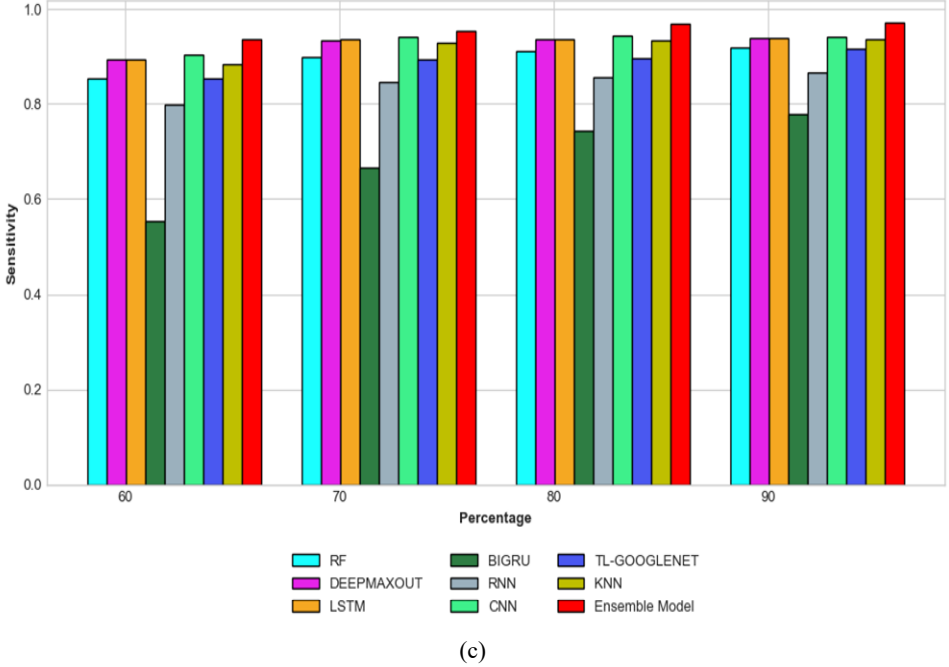


(a)

Figure 11 Assessment on Ensemble model and the traditional strategies with respect to positive metric for skin cancer classification (continued) (see online version for colours)

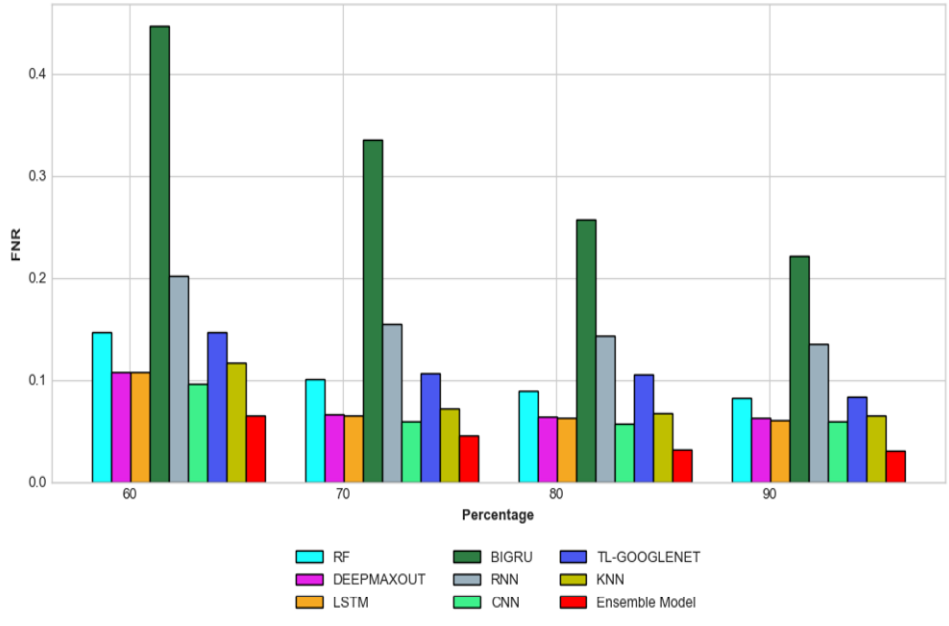


(b)

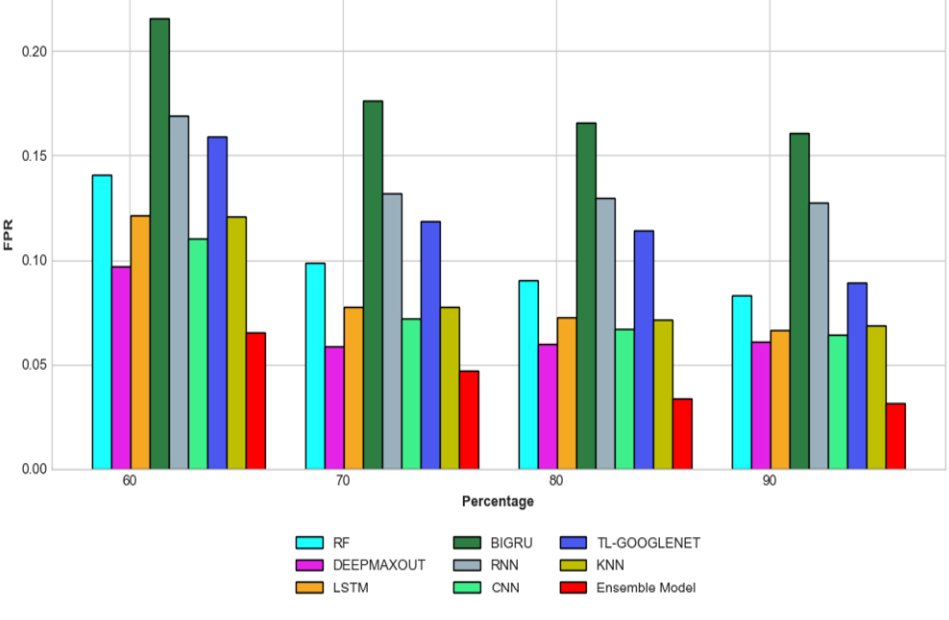


(c)

Figure 12 Assessment on Ensemble model and the traditional strategies with respect to negative metric for skin cancer classification

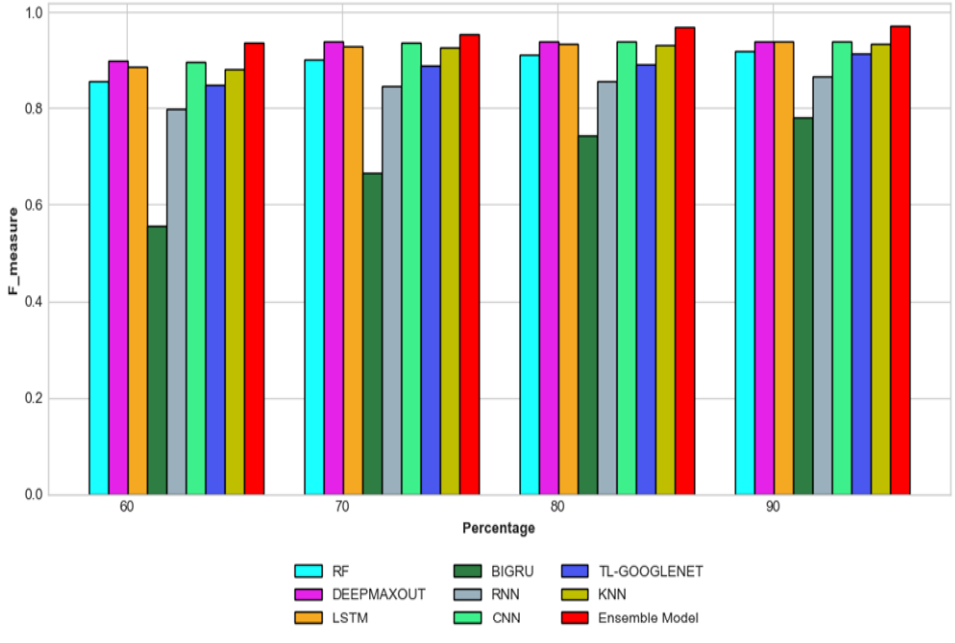


(a)

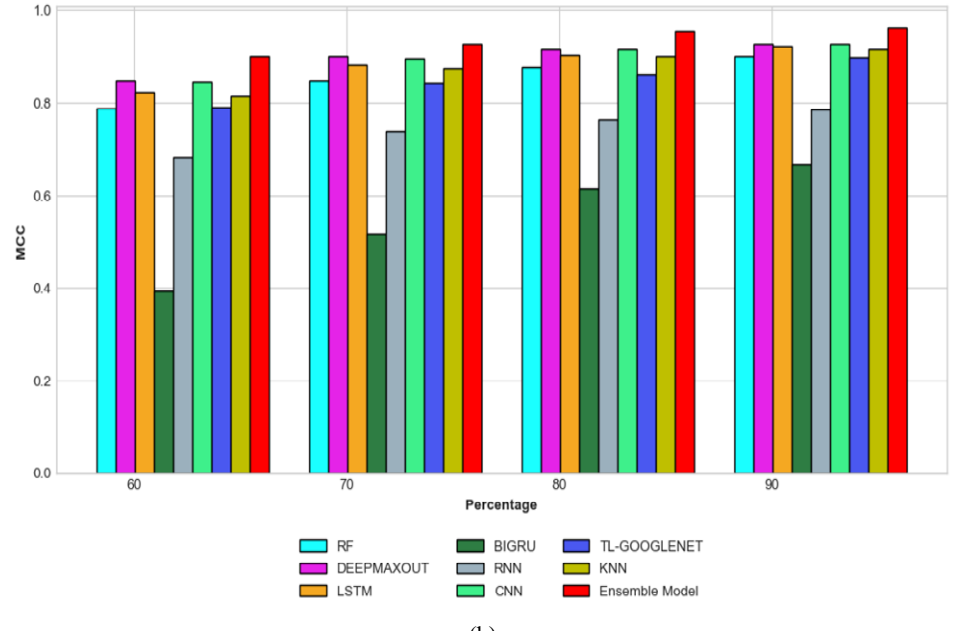


(b)

Figure 13 Assessment on ensemble model and the traditional strategies with respect to other metric for skin cancer classification (see online version for colours)

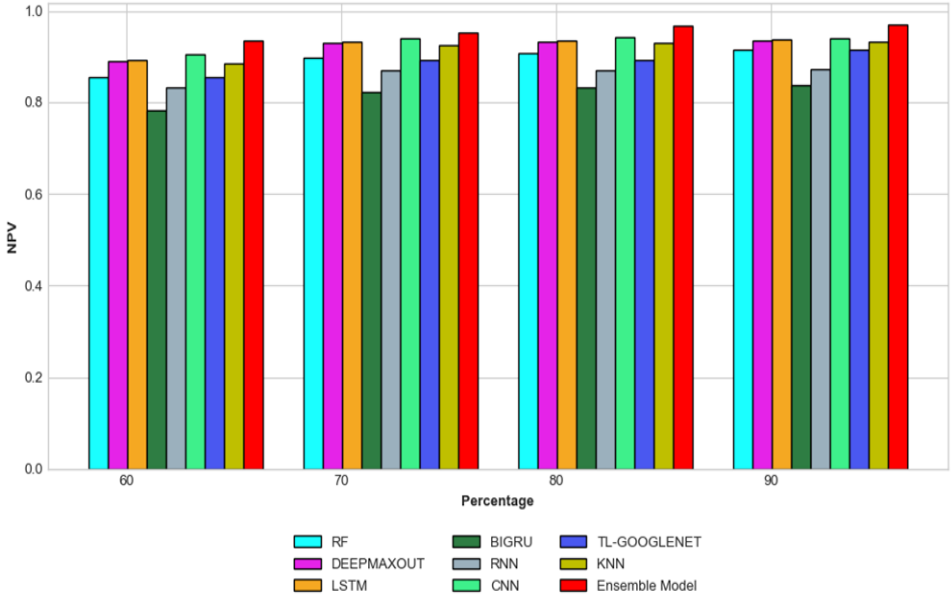


(a)



(b)

Figure 13 Assessment on ensemble model and the traditional strategies with respect to other metric for skin cancer classification (see online version for colours)



(c)

4.4 *Examination on ensemble model and conventional schemes regarding negative metric for skin cancer classification*

Figure 12 explains the negative metric assessment on Ensemble model is compared with the RF, Deep Maxout, LSTM, Bi-GRU, RNN, CNN, TL-GOOGLENET and KNN for skin cancer classification. For the efficacious classification of skin cancer, the model should attain lower negative metric ratings. Moreover, the FNR of the Ensemble model at the training percentage 90 is 3.0749, though the RF, Deep Maxout, LSTM, Bi-GRU, RNN, CNN, TL-GOOGLENET and KNN yielded the greatest FNR. In addition, learning percentage = 60, the highest FPR is acquired by the Bi-GRU algorithm (21.5403), accompanied by RNN (16.8969) and TL-GOOGLENET (15.8951), even though the Ensemble model scored the lowest FPR of 6.5242. Simultaneously, evaluating the training rate 90%, the Ensemble model generated the minimised FPR of 3.1572, while the RF, Deep Maxout, LSTM, Bi-GRU, RNN, CNN, TL-GOOGLENET and KNN maintained the highest FPR error rate. Thus, Figure 12 indicates that the Ensemble model is more effectual for skin cancer classification with diminished error rate.

4.5 *Examination on ensemble model and conventional schemes regarding other metric for skin cancer classification*

The other metric evaluation on Ensemble approach is contradicted with the RF, deep maxout, LSTM, Bi-GRU, RNN, CNN, TL-GOOGLENET and KNN for skin cancer classification is displayed in Figure 13. The other metric needed to be maximised for the exact classification of skin cancer. The F-measure of the Ensemble model for the training

percentage 80 is 96.7615, whilst the extant methodologies gained the minimal F-measure ratings, notably, RF = 91.0983, deep maxout = 93.8923, LSTM = 93.2954, Bi-GRU = 74.4107, RNN = 85.6006, CNN = 93.8439, TL-GOOGLENET = 89.1448 and KNN = 93.1671, correspondingly. At the 90th training rate, the Ensemble model yielded the highest MCC (96.0537) than the 60th training rate. Finally, evaluating the NPV metric, the Ensemble model accomplished maximised rate of 96.1572, this is superior to RF, deep maxout, LSTM, Bi-GRU, RNN, CNN, TL-GOOGLENET and KNN, correspondingly. Thus, the ensemble method's successful training allowed it to maximise other metrics through an improved LGIP feature extraction method, which is much successful in classifying the skin cancer.

4.6 Impact on ensemble model, model with conventional deep joint segmentation, model with conventional LGIP and model with conventional score level fusion for skin cancer classification

Table 3 illustrates the impact on Ensemble model, model with conventional deep joint segmentation, model with conventional LGIP and model with conventional score level fusion for skin cancer classification. Here, the Ensemble scheme with all the enhancements can offer superior results with accurate classification of skin cancer. Further, the FNR of the Ensemble model is 0.0461, model with conventional deep joint segmentation is 0.0833, model with conventional LGIP is 0.1020 and model with conventional score level fusion is 0.0694, correspondingly. Additionally, the accuracy of the Ensemble model is 0.9535, model with conventional deep joint segmentation = 0.9162, model with conventional LGIP = 0.8975 and model with conventional score level fusion = 0.9302, correspondingly. Therefore, the Ensemble model exhibited its superiority by accomplishing highest positive and lowest negative measure ratings. This has facilitated the ability to accurately classify the skin cancer.

Table 3 Impact on ensemble model, model with conventional deep joint segmentation, model with conventional LGIP and model with conventional score level fusion for skin cancer classification

Metrics	*Model with conventional deep joint segmentation*	*Model with conventional LGIP*	*Model with conventional score level fusion*	*Ensemble model*
FNR	0.0833	0.1020	0.0694	0.0461
Accuracy	0.9162	0.8975	0.9302	0.9535
NPV	0.9153	0.8967	0.9292	0.9526
Sensitivity	0.9167	0.8980	0.9306	0.9539
FPR	0.0844	0.1030	0.0702	0.0469
Specificity	0.9156	0.8970	0.9298	0.9531
F-measure	0.9168	0.8982	0.9309	0.9541
Precision	0.9170	0.8984	0.9312	0.9544
MCC	0.8998	0.8813	0.9014	0.9280

4.7 Statistical evaluation on ensemble and conventional schemes with regard to accuracy for skin cancer classification

The statistical study of the ensemble model is contradicted with the RF, deep maxout, LSTM, Bi-GRU, RNN, CNN, TL-GOOGLENET and KNN for skin cancer classification is summarised in Table 4. Moreover, it is assessed in terms of accuracy under distinctive type of statistical metrics. Concerning the maximum statistical metric, the accuracy of the Ensemble technique is 0.9689, whilst the RF is 0.9171, Deep Maxout is 0.9382, LSTM is 0.9362, Bi-GRU is 0.8150, RNN is 0.8687, CNN is 0.9382, TL-GOOGLENET is 0.9134 and KNN is 0.9328, correspondingly. In addition, the greatest accuracy of the Ensemble algorithm is 0.9560 under the mean statistical metric, though the RF, deep maxout, LSTM, Bi-GRU, RNN, CNN, TL-GOOGLENET and KNN generated minimised accuracy value. It is obvious from the statistical study that the ensemble model is superior to the established schemes in improving the classification accuracy.

Table 4 Statistical assessment on ensemble and the traditional schemes with respect to accuracy for skin cancer classification

Methods	*Median*	*Maximum*	*Standard deviation*	*Minimum*	*Mean*
RF	0.9050	0.9171	0.0273	0.8563	0.8959
Deep Maxout	0.9376	0.9382	0.0202	0.8974	0.9277
LSTM	0.9302	0.9362	0.0237	0.8853	0.9205
Bi-GRU	0.7856	0.8150	0.0436	0.7162	0.7756
RNN	0.8608	0.8687	0.0239	0.8165	0.8517
CNN	0.9360	0.9382	0.0201	0.8967	0.9267
TL-GOOGLENET	0.8888	0.9134	0.0275	0.8472	0.8845
KNN	0.9277	0.9328	0.0244	0.8810	0.9173
Ensemble model	0.9603	0.9689	0.0158	0.9345	0.9560

4.8 Comparison on ensemble model and conventional strategies for skin cancer classification

Table 5 describes the comparative study on ensemble model and AOHBA (Sainudeen, 2023) for the classification of skin cancer. Moreover, it is evaluated for varied types of performance metrics and the ensemble scheme obtained outstanding performance than the AOHBA. More particularly, the ensemble model acquired the NPV = 0.9526, FNR = 0.0461, accuracy = 0.9535 and FPR = 0.0469, whereas the AOHBA obtained the NPV = 0.8879, FNR = 0.0594, accuracy = 0.9179 and FPR = 0.1243, correspondingly. Altogether, the improved LGIP based feature extraction with ensemble classifiers allows the proposed approach to detect the skin cancer more appropriately.

Table 5 Comparative study on ensemble model and AOHBA for skin cancer classification

Metrics	*AOHBA*	*Ensemble model*
FNR	0.0594	0.0461
Accuracy	0.9179	0.9535
NPV	0.8879	0.9526
Sensitivity	0.9406	0.9539
F-measure	0.9372	0.9541
Specificity	0.8757	0.9531
FPR	0.1243	0.0469
Precision	0.9337	0.9544
MCC	0.8190	0.9280

4.9 Segmentation performance assessment on IDJS and traditional segmentation methods

The segmentation performance evaluation on IDJS is examined over the FCM, K-means and conventional DJS regarding Dice, Jaccard and Segmentation accuracy is summarised in Table 6. While assessing Table 6, the IDJS generated the maximised ratings in all the metrics. In a similar manner, the IDJS acquired the dice value of 0.8660, although the FCM is 0.7611, K-means is 0.6625 and conventional DJS is 0.8326. Additionally, the segmentation accuracy and Jaccard rate attained using the IDJS is 0.8315 and 0.9123, whereas the FCM, K-means and conventional DJS recorded minimal values. Hence, Table 6 highlights the excellent ability of the IDJS for the skin cancer classification.

Table 6 Segmentation performance examination on IDJS and traditional segmentation methodologies

Metrics	*FCM*	*K-means*	*Conventional DJS*	*IDJS*
Dice	0.7611	0.6625	0.8326	0.8660
Segmentation accuracy	0.7747	0.7421	0.8879	0.9123
Jaccard	0.7601	0.6371	0.7924	0.8315

5 Conclusions

We presented an IDJS based skin cancer classification model with six major stages in this research. In initial stage, CLAHE assisted contrast enhancement was performed and in the second phase, IDJS based segmentation was done for the preprocessed image. In third phase, MTP, LGIP, GLCM (ASM, Contrast, Correlation, Dissimilarity, Homogeneity), and CCF feature were extracted. In the fourth phase, data augmentation was performed depending on extracted features. Based on the augmented data, ensemble classification combining the DeepMaxout, LSTM, and CNN was performed in the fifth phase. The output scores of these classifiers were fed to the improved score level fusion process for attaining final classified label. Based on these classified labels, doctor suggests treatments (recommendation) to the patients in the final stage. Finally, evaluating the NPV metric, the Ensemble model accomplished maximised rate of 96.1572, this is

superior to RF, Deep Maxout, LSTM, Bi-GRU, RNN, CNN, TL-GOOGLENET and KNN, correspondingly.

References

[online] https://www.kaggle.com/datasets/kmader/skin-cancer-mnist-ham10000.

Abdar, M., Samami, M., Mahmoodabad, S.D., Doan, T., Mazoure, B., Hashemifesharaki, R., Liu, L., Khosravi, A., Acharya, U.R., Makarenkov, V. and Nahavandi, S. (2021) 'Uncertainty quantification in skin cancer classification using three-way decision-based Bayesian deep learning', *Computers in Biology and Medicine*, August, Vol. 135, pp.1–17.

Adla, D., Reddy, G.V.R., Nayak, P. and Karuna, G. (2021) 'Deep learning based computer aided diagnosis model for skin cancer detection and classification', *Distributed and Parallel Databases*, Vol. 40, No. 4, pp.717–736.

Adla, D., Reddy, G.V.R., Nayak, P. and Karuna, G. (2023) 'A full-resolution convolutional network with a dynamic graph cut algorithm for skin cancer classification and detection', *Healthcare Analytics*, Vol. 3, No. 11, pp.1–12.

Ali, M.S., Miah, M.S., Haque, J., Rahman, M.M. and Islam, M.K. (2021) 'An enhanced technique of skin cancer classification using deep convolutional neural network with transfer learning models', *Machine Learning with Applications*, Vol. 5, No. 5, pp.1–9.

Alphonse, A.S. and Starvin, M.S. (2020) 'A novel and efcient approach for the classifcation of skin melanoma', *Journal of Ambient Intelligence and Humanized Computing*, Vol. 11.

Arora, G., Dubey, A.K., Jaffery, Z.A. and Rocha, A. (2020) 'Bag of feature and support vector machine based early diagnosis of skin cancer', *Neural Computing and Applications*, Vol. 34, pp.8385–8392.

Balaha, H.M. and Hassan, A.E-S. (2023) 'Skin cancer diagnosis based on deep transfer learning and sparrow search algorithm', *Neural Computing and Applications*, Vol. 35, pp.815–853.

Dandu, R., Murthy, M.V. and Ravi Kumar, Y.B. (2023) 'Transfer learning for segmentation with hybrid classification to detect melanoma skin cancer', *Heliyon*, Vol. 9.

Filali, Y., Khoukhi, H.E., Sabri, M.A. and Aarab, A. (2020) 'Efficient fusion of handcrafted and pre-trained CNNs features to classify melanoma skin cancer', *Multimedia Tools and Applications*, Vol. 79, No. 4, pp.1–20.

Fu'adah, Y.N., Pratiwi, N.K.C., Pramudito, M.A. and Ibrahim, N. (2020) 'Convolutional neural network (CNN) for automatic skin cancer classification system', *IOP Conf. Series: Materials Science and Engineering*, Vol. 982, p.012005, DOI: 10.1088/1757-899X/982/1/012005.

Hatem, M.Q. (2022) 'Skin lesion classification system using a K-nearest neighbor algorithm', *Hatem Visual Computing for Industry, Biomedicine, and Art*, Vol. 5, p.7.

Holder, R.P. and Tapamo, J.R. (2017) 'Improved gradient local ternary patterns for facial expression recognition', *Holder and Tapamo EURASIP Journal on Image and Video Processing*, Vol. 42, DOI: 10.1186/s13640-017-0190-5.

Jaisakthi, S., Mirunalini, P., Aravindan, C. and Appavu, R. (2023) 'Classification of skin cancer from dermoscopic images using deep neural network architectures', *Multimedia Tools and Applications*, Vol. 82, pp.15763–15778.

Kassem, M.A., Hosny, K.M. and Fouad, M.M. (2019) 'Skin lesions classification into eight classes for ISIC 2019 using deep convolutional neural network and transfer learning', *Digital Object Identifier*, Vol. 8, pp.114822–114832.

Khan, A., Bashar, F., Ahmed, F. and Kabir, M.H. (2013) 'Median ternary pattern (MTP) for face recognition', *2013 International Conference on Informatics, Electronics and Vision (ICIEV)*.

Kumar, M., Alshehri, M., AlGhamdi, R., Sharma, P. and Deep, V. (2020) 'A DE-ANN inspired skin cancer detection approach using fuzzy C-means clustering', *Mobile Networks and Applications*, Vol. 25, pp.1319–1329.

Mahesh, M. and Renjit, A. (2020) 'DeepJoint segmentation for the classification of severity-levels of glioma tumour using multimodal MRI images', *IET Image Processing*, Vol. 14, No. 11, pp.2541–2552.

Mijwil, M.M. (2021) 'Skin cancer disease images classification using deep learning solutions', *Multimedia Tools and Applications*, Vol. 80, pp.26255–26271.

Mohanaiah, P., Sathyanarayana, P. and Kumar, L.G. (2013) 'Image texture feature extraction using GLCM approach', *International Journal of Scientific and Research Publications*, May, Vol. 3, No. 5, pp.1–5.

Nguyen, T.P.H., Cai, Z., Nguyen, K., Keth, S., Shen, N. and Park, M. (2020) 'Pre-Processing Images Using Brightening, Clahe and Retinex'.

Peta, J. and Koppu, S. (2022) 'An IoT-based framework and ensemble optimized deep maxout network model for breast cancer classification', *Electronics*, Vol. 11, p.4137, https://doi.org /10.3390/electronics11244137.

Qureshi, A.S. and Roos, T. (2020) 'Transfer learning with ensembles of deep neural networks for skin cancer detection in imbalanced data sets', *Neural Processing Letters*, Vol. 55, pp.4461–4479.

Said, L. and Khurshid, K. (2017) 'Using co-occurrence and granulometry features for content based image retrieval', Conference Paper, October [online] https://www.researchgate.net /publication/320034654.

Sainudeen, J. (2023) 'Skin cancer detection: improved deep belief network with optimal feature selection', in *Multiagents and Grid Systems*, Vol. 19, No. 2, pp.187–210.

Shorfuzzaman, M. (2021) 'An explainable stacked ensemble of deep learning models for improved melanoma skin cancer detection', *Multimedia Systems*, Vol. 28, pp.1309–1323.

Thurnhofer-Hemsi, K. and Domínguez, E. (2020) 'A convolutional neural network framework for accurate skin cancer detection', *Neural Processing Letters*, Vol. 53, No. 5, pp.3073–93.

Tumpa, P.P. and Kabir, A. (2021) 'An artificial neural network based detection and classification of melanoma skin cancer using hybrid texture features', *Sensors International*, Vol. 2, pp.1–8.

Xin, C., Liu, Z., Zhao, K., Miao, L., Ma, Y., Zhu, X., Zhou, Q., Wang, S., Li, L., Yang, F., Xu, S. and Chen, H. (2023) 'An improved transformer network for skin cancer classification', *Computers in Biology and Medicine*, Vol. 149, pp.1–10.

Yang, G., Luo, S. and Greer, P. (2023) 'A novel vision transformer model for skin cancer classification', *Neural Processing Letters*, Vol. 55, No. 7, pp.9335–9351.

Zhou, F., Hang, R., Liu, Q. and Yuan, X. (2019) 'Hyperspectral image classification using spectral-spatial LSTMs', *Neurocomputing*, Vol. 328, pp.39–47.

Zhou, L. and Wang, H. (2012) 'Local gradient increasing pattern for facial expression recognition', *2012 19th IEEE International Conference on Image Processing*, DOI: 10.1109/ICIP .2012.6467431.

Zubair, A.R.A. and Alo, O.A. (2019) 'Grey level co-occurrence matrix (GLCM) based second order statistics for image texture analysis', *International Journal of Science and Engineering Investigations*, October, Vol. 8, No. 93, pp.64–73.

Nomenclature

Abbreviation	*Description*
FrCN-DGCA	Full-resolution convolutional networks-dynamic graph cut algorithm
VIT	Vision transformer
MC	Monte Carlo
EMC	Ensemble MC
TWD	Three-way decision
DE	Deep ensemble
UQ	Uncertainty quantification
BDL	Bayesian deep learning
DCNN	Deep convolutional neural network
ABCD	Asymmetry, border, colour, diameter
GLCM	Grey-level co-occurrence matrix
LBP	Local binary pattern
ViTfSCD	ViT for skin cancer detection
CNN	Convolutional neural network
NN	Neural network
SVM	Support vector machine
KNN	K-nearest neighbour
NB	Naïve Bayes
ML	Machine learning
DT	Decision tree
DL	Deep learning
SCC	Squamous cell carcinoma
BCC	Basal cell carcinoma
DNN	Deep neural network
FC	Fully connected
TL	Transfer learning
CLAHE	Contrast limited adaptive histogram equalisation
ASSD	Average symmetric surface distance
MTP	Median ternary pattern
MF	Median filter
ILGIP	Improved local gradient increasing pattern
GLCM	Grey-level co-occurrence matrix
CCF	Co-occurrence colour feature
LDP	Local directional pattern
ASM	Angular second moment
ROI	Region of interest
SD	Standard deviation
RGB	Red green blue

Nomenclature (continued)

Abbreviation	*Description*
MLP	Multilayer perceptron
ReLU	Rectified linear unit
LSTM	Long short term memory
RNN	Recurrent neural network
df	dermatofibroma
akiec	Actinic keratoses and intraepithelial carcinoma
bcc	basal cell carcinoma
mel	melanoma
nv	nevi
RF	Random forest
Bi-GRU	Bidirectional gated recurrent unit