Available online at www.ijournalse.org

Emerging Science Journal

(ISSN: 2610-9182)

Vol. 10, No. 3, June, 2026

Fractional White Smell Agent Optimization for CNN-Based Transfer Learning in Melanoma Classification

Vijaya P [1*], Basant Kumar [1], Joseph Mani [1], Satish Chander [2], Roshan Fernandes [3], Mohamed Sirajudeen Yoosuf [1]

[1] *Department of Mathematics and Computer Science, Modern College of Business and Science, Bowshar, Muscat, Oman.*

[2] *Department of Computer Science & Engineering, Birla Institute of Technology, Mesra, Ranchi, India.*

[3] *Department of Cybersecurity, NMAM Institute of Technology (NMAMIT), NITTE (Deemed to be University), Nitte 574 110, Karnataka, India.*

Abstract

Melanoma is the deadliest form of skin cancer, and early diagnosis and treatment can significantly reduce mortality rates. However, existing strategies for classifying melanoma from dermoscopic skin images still face significant challenges. Therefore, this study aims to develop an accurate method for melanoma classification using dermoscopic skin images. A novel melanoma classification framework, termed Fractional White Smell Agent Optimization-enabled Convolutional Neural Network-based Transfer Learning (FWSAO_CNN-based TL), is proposed. First, the input skin image is preprocessed using an Adaptive Kalman Filter. Subsequently, skin lesion segmentation is performed using LinkNet, where the network is trained using White Smell Agent Optimization (WSAO). Following segmentation, image augmentation is applied, and feature extraction is conducted. Finally, melanoma classification is performed using a CNN-based transfer learning model trained with the proposed Fractional White Smell Agent Optimization (FWSAO), which integrates the Fractional concept, Smell Agent Optimization (SAO), and White Shark Optimizer (WSO). The CNN utilizes hyperparameters derived from a pretrained GoogLeNet model. The performance of the proposed FWSAO_CNN-based TL framework was evaluated using accuracy, True Positive Rate (TPR), and True Negative Rate (TNR). The proposed method achieved values of 91.565%, 90.090%, and 91.269%, respectively. Furthermore, the proposed model demonstrated performance improvements of 18.4%, 8.1%, 17.5%, 12.55%, 8.2%, and 6.23% compared with conventional approaches.

Keywords:

Skin Cancer;
Melanoma Classification;
LinkNet;
Transfer Learning;
GoogleNet.

Article History:

Received:	22	November	2025
Revised:	26	April	2026
Accepted:	09	May	2026
Published:	01	June	2026

1- Introduction

Today's healthcare framework depends heavily on medical imaging to enable non-invasive diagnostic interventions. It consists of producing structured depictions of the human body's internal anatomy for diagnostic purposes. X-ray-based techniques, Computed Tomography (CT), and Magnetic Resonance Imaging (MRI) are the medical imaging approaches. In addition to these techniques, medical images are extensively employed for identifying more disorders, especially skin-based diseases [1]. The diagnostic images created by medical imaging techniques become a significant part in the diagnosis of disorders [2]. Skin cancer is the most prevalent form of cancer in humans. Furthermore, squamous cell carcinoma, basal cell carcinoma, and melanoma are the most common malignant skin lesions. Accurately differentiating among skin lesions and malignancies is difficult due to complexity. Skin cancer experts employ both eye inspection and hand-held dermoscopy to assess patient lesions. It can record microscopic, macroscopic, dermoscopic, or close-up digital pictures [3].

* **CONTACT**: vijaya.padmanabha@mcbs.edu.om

DOI: https://doi.org/10.28991/ESJ-2026-010-03-02

Dermoscopy utilizes polarization or immersion fluid for decreasing surface reflection when diagnosing the skin under intense light and magnification [3]. The first sign of melanoma is a skin lesion with an irregular shape. The mole is in an oval shape, which is minor in size with a constant shape. At the initial stage, irregular and enlarged moles should be assessed according to skin pigmentation to identify potential skin cancer. The melanoma entities identified by the medical experts and patients with irregular color diversification, border irregularity, and lesion development process [4]. The affected area or the entire skin surface is carefully examined to determine whether a lesion is benign or malignant. Oncologists can select to perform a biopsy procedure after the lesion identification in some cases [5]. The key purpose of melanoma identification is to facilitate automatic classification of pigmented skin lesions as benign or malignant. The process begins with the acquisition of pigmented skin lesion images, which are then labeled to train the classification approach. Proper training of classification techniques demands a substantial volume of labeled data, given the variability in lesion image distributions across multiple datasets. Nevertheless, data labeling is a high-cost process and needs high maintenance. Moreover, a learning model is learned on some pre-trained images to assist learning techniques for some other lesion images, which is needed to eliminate the purpose of melanoma detection [6].

Melanoma presently stands as the nineteenth most widespread disease in the world, with a high death toll. As a form of skin cancer, melanoma contributes to a high number of deaths annually. The survival rates of patients may be extensively increased even when the melanoma is predicted early and rapidly treated, since it increases slowly in the earlier phase [7, 8]. Skin lesions such as large brown moles with black flecks, which bleed irregularly or change color, may signal melanoma, a rapidly spreading and serious cancer. Melanoma has multiple subtypes, which include acral, nodular, superficial, and lentigo [9, 10]. Deep Learning (DL) is one of the ten ground-breaking techniques for melanoma classification. Convolutional Neural Networks (CNNs) can offer precise diagnostics for the relevant disease; thus, they improve patient care when utilized to diagnose tumors, especially melanoma skin cancer [11]. It has only a few available databases, like medical imaging applications for classification problems, and Transfer Learning (TL) is a fundamental module. Utilizing a pre-trained CNN module and merely fine-tuning its performance for faster operation is more computationally effective than training a CNN from the beginning, which requires a large amount of data and significant computational resources [12]. Several number of pre-trained CNNs, such as AlexNet, Inception, ResNet, as well as DenseNet, were trained on the ImageNet Large Scale Visual Recognition Challenge (ILSVRC) database. Some of the researchers employed several pre-trained CNNs and also described adequate skin cancer diagnosis performance [13].

1-1- Literature Survey

Loganathan et al. [5] designed Enhanced Deep CNN (E-DCNN) for melanoma classification. This module achieved faster classification with minimum computational time. Nevertheless, it failed to obtain maximum accuracy, and it was prone to overfitting, especially when the database was not well-balanced. Kaur et al. [9] introduced the Lesion Classification Network (LCNet) for classifying melanoma disease. This approach provided less complexity to automate melanoma diagnostic progression and an advanced detection method to save a life. However, it did not extend multi-class classification to identify other skin cancers. Ding et al. [7] designed a two-stage ensemble network model for melanoma classification. Even though the segmented images in this model increased the classification accuracy, it was unable to discover more classification methods by means of the features and diverse classes of dermoscopy images, particularly in preprocessing.

Salma & Eltrass [13] developed ResNet50 with a Support Vector Machine (SVM) for the classification of malignant melanoma. This approach had the ability to help medical experts and dermatologists to efficiently classify the skin lesions. Nonetheless, the performance of this technique was not precisely detected due to less volume of training and testing data. Aljohani & Turki [11] devised GoogleNet for automatic classification of melanoma skin cancer. It demonstrated the feasibility of determining whether the skin is affected by melanoma without the involvement of experts for feature engineering. Nevertheless, it neglected to include diverse augmentation methods to improve detection efficiency. Bhimavarapu & Battineni [14] presented Fuzzy GrabCut-stacked CNN (GC-SCNN) with SVM. Although this approach classified the lesions with maximum accuracy and minimum processing time, it eliminated the insecurities in boundary detection and reduced the processing time.

Qureshi et al. [15] established Transfer Learning based CNN (TL-based CNN). The developed approach had the capability to apply different databases based on better performance. But it failed to introduce a mobile application for detecting melanoma in the early stage. Shorfuzzaman [6] designed Explainable CNN-based stacked ensemble framework to classify melanoma. This model demonstrated enviable interpretability properties by effectively detecting several benign and malignant melanoma signs. Nonetheless, this approach was complex to maintain, and also it had issues in the training process. Garcia et al. [16] used a Vision Transformer ViT-L/16 model for the effective classification of melanoma. This method attained high accuracy in distinguishing malignant from benign lesions due to the global feature representation. However, this method is limited to binary classification and does not account for other types of skin lesions. Manikandan et al. [17] developed an efficient deep learning classification (EDLCS) for the classification of melanoma using dermascopic images. This method was robust and adaptable in handling different image representations. However, this method was evaluated on a particular dataset, and its diversity to other datasets was unknown.

1-2- Motivation

Melanoma classification is the process of categorizing skin lesions into melanoma and non-melanoma categories using visual or image-based analysis. Traditional melanoma classification methods faced some challenges, which are given below:

- Some existing methods have lower accuracy due to small and imbalanced datasets.
- The traditional methods suffered with computational complexity due to the deep and stacked architecture.
- Additionally, the existing methods were limited in handling complex and low-contrast lesion features due to the lack of advanced preprocessing and segmentation methods.

Thus, the FWSAO_CNN-based TL is designed to improve the melanoma classification by overcoming the limitations in previous studies. The proposed FWSAO_CNN-based TL employs FWSAO to optimally tune the hyperparameters for effective learning and uses data augmentation to handle small and imbalanced datasets, thereby improving the classification accuracy. Furthermore, FWSAO_CNN-based TL leverages transfer learning with pre-trained networks and FWSAO-based optimization, thereby reducing the computational complexity. Additionally, FWSAO_CNN-based TL employs WSAO_LinkNet for lesion segmentation and adaptive Kalman filtering for noise reduction, which improves the detection of low-contrast and complex lesion patterns.

1-3- Contribution

The principal contribution presented in this work is:

***Proposed FWSAO_CNN-based TL for melanoma classification*:** The novelty of the melanoma classification framework is to propose a module using CNN-based TL with a hybrid optimization algorithm. In this investigation, the melanoma classification is performed by CNN-based TL, where CNN-based TL is trained by utilizing FWSAO. The CNN network employs the hyperparameters from GoogleNet. The FWSAO is based on the combination of the fractional concept, SAO, and WSO.

2- Proposed Convolutional Neural Network-Based Transfer Learning-Enabled Fractional White Smell Agent Optimization for Melanoma Classification

From dermoscopic skin samples, the automatic detection of melanoma is a complex task. Nevertheless, the complexity of melanoma classification is reduced using the DL module. This research focuses on developing a melanoma classification approach leveraging FWSAO_CNN-based TL. Initially, input skin images acquired from the dataset [18] are allowed for image pre-processing that is carried out by an adaptive Kalman filter [19] for removing noises and artifacts from the input image. Followed by this process, skin lesion segmentation is done by LinkNet [20], which is trained by WSAO. This is formed by combining SAO [21] and WSO [22]. After the segmentation, image augmentation is performed where the augmentation techniques [23, 24] are shifting, shearing, rotation, random erasing, and cropping. Then, feature extraction like statistical features [25], Haralick texture features [26], CNN features [27], and Speeded-Up Robust Features (SURF) [28] will be extracted. Additionally, statistical features and Haralick texture descriptors are incorporated. Finally, Melanoma classification is accomplished using CNN-based TL [29], where CNN is used with the hyperparameters from the trained model GoogLeNet [30]. FWSAO is employed to optimize the performance of CNN-based TL. Figure 1 represents a schematic representation of Melanoma classification.

2-1- Image Acquisition

The skin lesion image database comprises 25,331 images categorized into nine classes. Assume the database as Υ, where the Υ is composed of Υ_{ν} number of images. The arithmetical computation of Υ is elucidated as,

$$\Upsilon = \{\Upsilon_1, \Upsilon_2, \ldots\ldots \Upsilon_{\mu}, \ldots\ldots \Upsilon_{\nu}\} \quad (1)$$

From this database, Υ_{μ} is taken as the input for the process below.

2-2- Image Pre-Processing by the Adaptive Kalman Filter

The image undergoes pre-processing to eliminate unwanted noise and artifacts. Moreover, it is deployed to increase image quality. This is accomplished using an Adaptive Kalman Filter with an input image Υ_{μ}.

2-2-1- Adaptive Kalman filter

An Adaptive Kalman Filter [19] is an enhanced type of Kalman Filter, which is established to handle uncertainties as well as variations in dynamic systems more efficiently. The measurement variance V_{ξ} becomes a significant factor for this filter. The value for V_{ξ} is improved to balance the errors. This approach seeks to represent random data points using

a framework that accounts for arbitrary noise by increasing the V_ξ such that the data points are not used for corrupting the measurement of parameters. The expression of V_ξ is calculated as,

$$V_\xi = \left(\frac{1}{\varphi}\right)\left(\sum_{\phi=1}^{\varphi} \psi_{\xi-\phi}.\psi_{\xi-\phi}^{T}\right) - \Im_\xi^t C_\xi \Im_\xi \quad (2)$$

Here, φ indicates overall points based on a pre-defined smoothing window,ϕ refers to index, C_ξ defines covariance matrix, ψ_ξ illustrates noise contribution to the measurement and $\Im_\xi^t$ implies row vector. Thus, the filtered image is signified as P_μ.

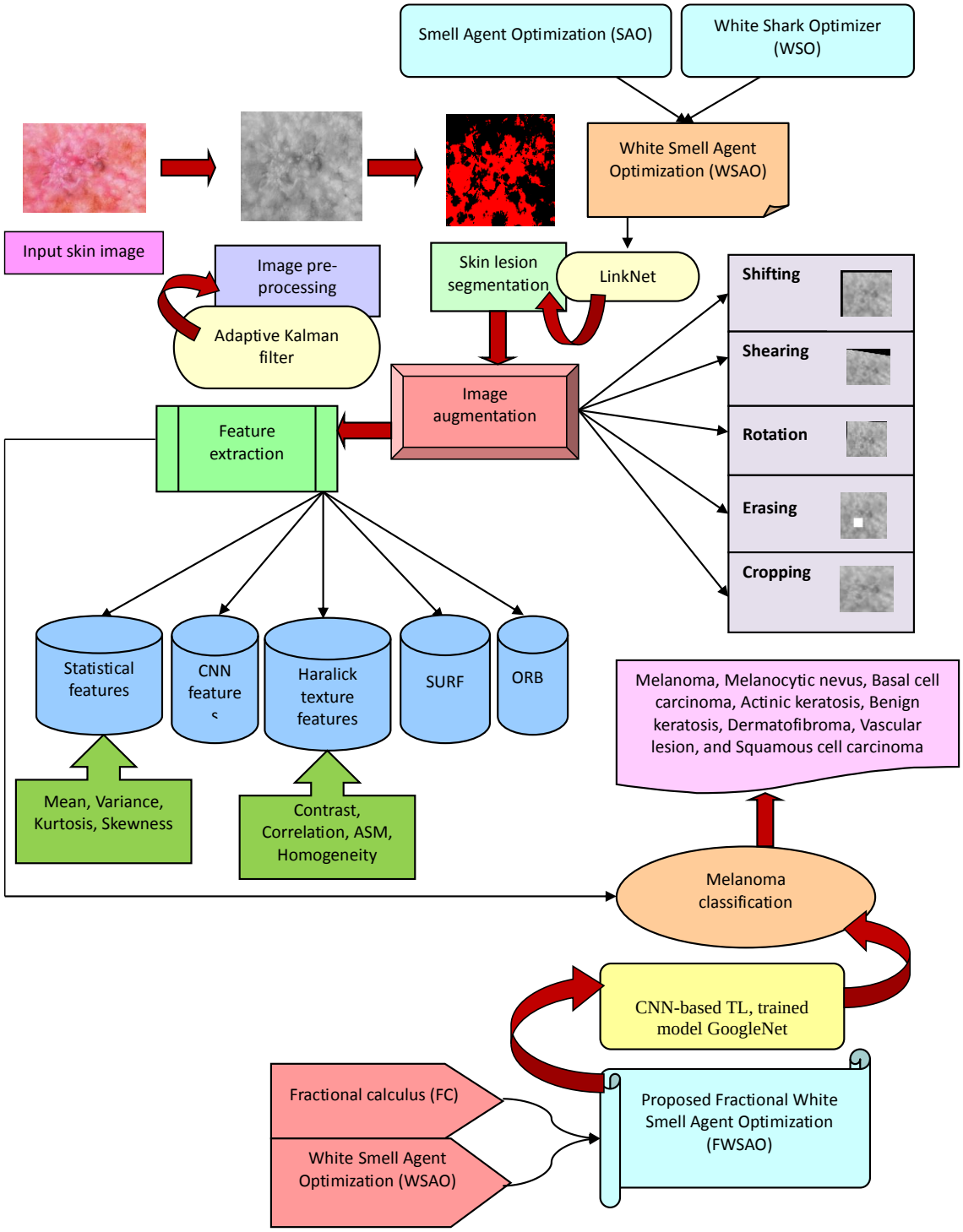


Figure 1. Schematic representation of Melanoma classification

2-3- Image Segmentation by LinkNet

This is the process of separating the image into non-overlapping regions. Moreover, it is a prime stage for image analysis and interpretation. In this technique, LinkNet is used to segment the filtered image P_μ into several regions. Here, LinkNet is trained using WSAO.

2-3-1- LinkNet Structure

Link-Net [20] comprises encoder and decoder blocks, and batch normalization (BN) is used within every convolution (conv) layer followed by Rectified Linear Unit (ReLU) non-linear function. The conv layer in the encoder is the first block with a kernel dimension of 7 × 7 with 2 strides. Moreover, the encoder also processes spatial max-pooling of 3 × 3 with 2 strides, which is followed by decoder blocks. Every input is given to the corresponding decoder output in the encoder block. Furthermore, LinkNet emphasizes the measurement of spatial information loss, while upsampling occurs in the decoder block. The decoder not only allocates the information obtained from the encoder at every layer but also integrates several factors. Hence, Link-Net output is expressed as O_μ. Figure 2 designs the LinkNet architecture.

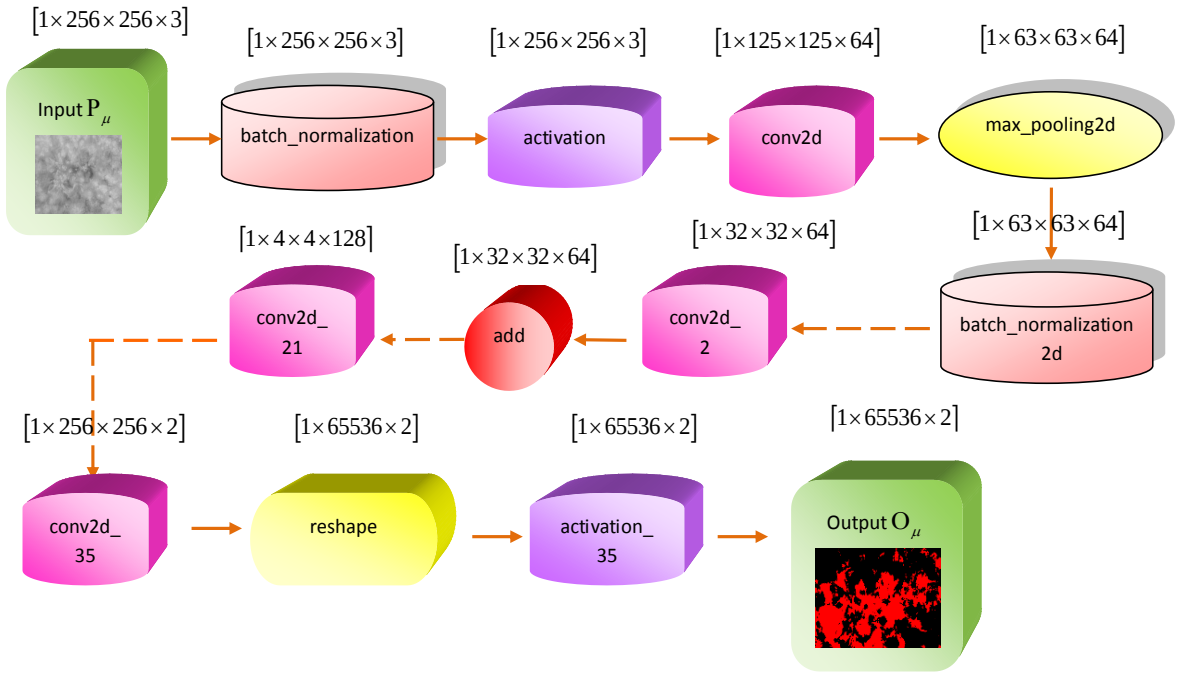


Figure 2. LinkNet structure

2-3-2- Training Algorithm of WSAO

WSAO is exploited for segmenting the skin lesions, where this new algorithm is designed by WSO and SAO. SAO [21] utilizes the sensation of smell and the spontaneous trailing phenomenon of an agent for finding the smell source. Sniffing and trailing modes form the core components of this algorithm. Trailing mode directs the agent's movement toward scent molecules, while sniffing mode simulates the release of molecules from the source. WSO [22] is based on the scholastic behavior of white sharks as they forage for food in their natural setting, demonstrating the importance of these habits for their survival in the deep ocean. It is designed to resolve unconstrained and constrained real-time optimization issues. Therefore, WSAO integrates SAO and WSO for efficient generalizability.

a) Solution Encoding

It is designed for allocating the suitable positions for every smell molecule in a search space (γ) that is originated by,

$$\gamma = [1 \times B] \tag{3}$$

Here, B implies the learning parameter of LinkNet.

b) Fitness Measure

It is determined for acquiring the utmost solution with favourable smell molecules, where its mathematical expression is articulated by,

$$Fit = \frac{1}{\nu}\sum_{\mu=1}^{\nu}\left[\Phi_\mu - O_\mu\right]^2 \tag{4}$$

Here, Fit refers to fitness value, O_μ implies LinkNet's output and Φ_μ defines target output.

c) Algorithm Steps

The steps for WSAO are described below:

Step 1: **Initialization**

SAO is initiated by an arbitrarily generated population of smell molecules, which is computed by,

$$b_{\omega}^{\tau} = \left\{b_{\aleph,1}^{\tau}, b_{\aleph,2}^{\tau}, \ldots\ldots, b_{\aleph,A}^{\tau}\right\} \quad (5)$$

Here, $\aleph$ elucidates smell molecules, A expresses evaporated smell molecules, and b_{ω}^{τ} refers to the population of smell molecules. Moreover, $\omega = \{1,2,\ldots\ldots,\aleph\}$ and τ enumerates smell molecules' position.

Step 2: Determine Fitness

The objective value of this population is estimated, and the most suitable fitness value is considered for further analysis using Equation 4.

Step 3: Sniffing Model

After the initialization of the population of smell molecules, the velocity for every molecule in SAO is determined by utilizing,

$$B_{(i,j,k)}^{\tau+1} = U_{(l,m,n)}^{\tau} + B_{(i,j,k)}^{\tau} \quad (6)$$

Here, $B_{(i,j,k)}^{\tau}$ and $U_{(l,m,n)}^{\tau}$ enumerates the position of every candidate solution and molecules velocity coordinates of smell molecules.

(i) Small Velocity and Position Update

Each small molecule has its own velocity for moving and updating its position. Due to the non-uniform motion of gas molecules, gas velocity is inferred from hydrostatic pressure, and it is expressed as,

$$\vartheta = \frac{1}{2}(acu^2) = \frac{3}{2}(\kappa aT) \quad (7)$$

Here, a elucidates the amount of substance in a gas and u is the velocity. Also, the velocity of non-uniform molecules moving through the search space is formulated is computed as,

$$u = \sqrt{\frac{3\kappa T}{c}} \quad (8)$$

The updated form of the smell molecules' velocity is enumerated as,

$$u_{\omega}^{(\tau+1)} = u_{\omega}^{\tau} + \upsilon_0 \times \sqrt{\frac{3\kappa T}{c}} \quad (9)$$

Here, $u_{\omega}^{(\tau+1)}$ specifies update velocity, u_{ω}^{τ} refers to present velocity, κ enumerates Boltzmann constant, T signifies temperature, c implies mass. The next phase computes the updated smell molecule solution based on the previous equation.

Step 4: Trailing Mode

When the search space is explored, the amount of smell molecules becomes greater than the present position of an agent. In this state, the agent navigates towards this target position. Then, the agent performs to track the position of all molecules with greater concentration till the molecule with the utmost objective measure is identified. In practical scenarios, the agent is expected to detect and track smell molecules, relying on intuition to follow the molecules in the hopes of identifying their source. Moreover, each agent has a unique capacity of olfaction based on the dimension of the olfactory lobes. SAO is generalized for overall agents; the accuracy might be influenced by the olfaction capacity of the agent. The control factor is referred to as olfaction capacity (OC).

The agent's olfactory capacity provides predictive understanding of the object emitting the smell, allowing it to track the molecule's trail. To efficiently follow the smell, the agent must recall its current location and the point with the least favourable position. By doing so, the agent can adapt its path to avoid unfavourable fitness and efficiently reach its target. Hence, the trailing behavior of the agent is enumerated by,

$$b_{\omega}^{\tau+1} = b_{\omega}^{\tau} + \upsilon_1 \times OC \times \left(b_{agent}^{\tau} - b_{\omega}^{\tau}\right) - \upsilon_2 \times OC \times (b_{worst}^{\tau} - b_{\omega}^{\tau}) \quad (10)$$

Here, υ_1 and υ_2 depicts random numbers with distinct intervals, ${b_\omega}^{r+1}$ is the updated position of the molecule and b_ω^r is the previous position of the molecule, b_{agent}^τ and b_{worst}^τ demonstrates the utmost and worst fitness, respectively.

$$b_\omega^{\tau+1} = b_\omega^\tau[1 - \upsilon_1 \times OC + \upsilon_2 \times OC] + \upsilon_1 \times OC \times b_{agent}^\tau - \upsilon_2 \times OC \times b_{worst}^\tau \quad (11)$$

The solution of WSO is considered to combine with SAO's solution for higher efficiency, thus the arithmetic expression is computed as,

$$\Psi_{\varsigma+1}^\zeta = \Psi_\varsigma^\zeta + \frac{\Re_\varsigma^\zeta}{\ell}; Rand \geq MF \quad (12)$$

Here, $\Psi_{\varsigma+1}^\zeta$ implies the new position vector of ζ^{th} white shark in $(\varsigma+1)^{th}$ iteration, $\Re_\varsigma^\zeta$ specifies the new velocity vector of ℓ ζ^{th} white shark in $(\varsigma)^{th}$ iteration indicates the frequency of the wavy motion of the white shark, $Rand$ enumerates an arbitrary value between $[0,1]$ and MF defines the movement force.

$$\Psi_{\varsigma+1}^\zeta = b_\omega^{\tau+1} \quad (13)$$

$$\Psi_\varsigma^\zeta = b_\omega^\tau \quad (14)$$

$$\Re_\varsigma^\zeta = \Re_\omega^\tau \quad (15)$$

By substituting Equation 13 to 15 in 12, then the expression becomes,

$$b_\omega^{\tau+1} = b_\omega^\tau + \frac{\Re_\omega^\tau}{\ell} \quad (16)$$

$$b_\omega^\tau = b_\omega^{\tau+1} - \frac{\Re_\omega^\tau}{\ell} \quad (17)$$

Apply Equation 17 in Equation 11, then the solution is written as,

$$b_\omega^{\tau+1} = \left(b_\omega^{\tau+1} - \frac{\Re_\omega^\tau}{\ell}\right)[1 - \upsilon_1 \times OC + \upsilon_2 \times OC] + \upsilon_1 \times OC \times b_{agent}^\tau - \upsilon_2 \times OC \times b_{worst}^\tau \quad (18)$$

$$b_\omega^{\tau+1} - b_\omega^{\tau+1}[1 - \upsilon_1 \times OC + \upsilon_2 \times OC] = \left[\upsilon_1 \times OC \times b_{agent}^\tau - \upsilon_2 \times OC \times b_{worst}^\tau\right] - \left(\frac{\Re_\omega^\tau}{\ell}\right)[\upsilon_1 \times OC - 1 - \upsilon_2 \times OC] \quad (19)$$

$$b_\omega^{\tau+1}(1 - 1 + \upsilon_1 \times OC - \upsilon_2 \times OC) = \frac{\ell[\upsilon_1 \times OC \times b_{agent}^\tau - \upsilon_2 \times OC \times b_{worst}^\tau]}{\frac{-\Re_\omega^\tau[\upsilon_1 \times OC - 1 - \upsilon_2 \times OC]}{\ell}} \quad (20)$$

$$b_\omega^{\tau+1} = \frac{\ell[\upsilon_1 \times OC \times b_{agent}^\tau - \upsilon_2 \times OC \times b_{worst}^\tau] - \Re_\omega^\tau[\upsilon_1 \times OC - 1 - \upsilon_2 \times OC]}{\ell \times (\upsilon_1 \times OC - \upsilon_2 \times OC)} \quad (21)$$

The frequency ℓ is expressed as,

$$\ell = \ell_{\min} + \frac{\ell_{\max} - \ell_{\min}}{\ell_{\max} + \ell_{\min}} \quad (22)$$

Here, $\ell_{\max}, \ell_{\min}$ expresses the maximal and minimal frequency.

Step 5: Re-Evaluate the Fitness

It is recomputed for obtaining an optimal solution with less error using Equation 4.

Step 6: Termination

The above steps will be performed till the solution obtains the finest population of smell molecules. Table 1 expresses the pseudocode of WSAO.

Table 1. Pseudocode of WSAO

SL. No.	Pseudocode of WSAO
1	**Input: Number of iterations (Itr), population ($\aleph$)**
2	Output: $b_{\omega}^{\tau+1}$
3	Begin
4	Initialize the parameters
5	Determine fitness using Equation 4
6	Generate the initial position of smell molecules in the search space and allocate the velocity using Equation 6.
7	Determine velocity by Equation 8 and conduct sniffing mode by Equation 9.
8	Evaluate the molecule's position with the worst sniffing fitness and conduct the trailing mode using Equation 10
9	Compute the update solution by utilizing Equation 21
10	Re-evaluate the fitness measure
11	End
12	Return
13	Terminate

2-4- Image Augmentation based on the Segmented Image

This method is used to generate new training samples from prior ones by applying transformations to the actual images. Here, the augmentation process takes place with the segmented image O_{μ}, where it is carried out by using the following techniques.

2-4-1- Shifting

Shifting the image or particular object by a fixed number of pixels in any direction, like up, down, left, right, or diagonally [24]. Hence, shifting is specified as N_1.

2-4-2- Shearing

It is the alteration of the actual image along $\hbar$ and λ direction [23]. The shearing in $\hbar$ - direction is determined as,

$$\begin{bmatrix} \partial_{\hbar} \\ \partial_{\lambda} \end{bmatrix} = \begin{pmatrix} 1 & shR \\ 0 & 1 \end{pmatrix} \cdot \begin{bmatrix} \hbar \\ \lambda \end{bmatrix} \tag{23}$$

Moreover, the shearing in λ - direction is determined as,

$$\begin{bmatrix} \partial_{\hbar} \\ \partial_{\lambda} \end{bmatrix} = \begin{pmatrix} 1 & 0 \\ shS & 1 \end{pmatrix} \cdot \begin{bmatrix} \hbar \\ \lambda \end{bmatrix} \tag{24}$$

Here, $\partial_{\hbar}$ and ∂_{λ} enumerates the new position of every pixel after shearing $\hbar$ and λ image coordinates. Thus, shearing is indicated as N_2.

2-4-3- Random Erasing

This technique employs to random erase of one square in a square region of an image, which is denoted as N_3.

2-4-4- Cropping

It is defined as the process of expanding the original image, and it is also known as zooming or scaling [23], which is calculated using,

$$\begin{bmatrix} \partial_{\hbar} \\ \partial_{\lambda} \end{bmatrix} = \begin{pmatrix} R_{crop} & 0 \\ 0 & S_{crop} \end{pmatrix} \cdot \begin{bmatrix} \hbar \\ \lambda \end{bmatrix} \tag{25}$$

2-4-5- Rotation

It is performed to rotate the image around an axis in either a left or right direction [23] at an angles $[1,359]$. It is computed as,

$$\begin{bmatrix} \partial_{\hbar} \\ \partial_{\lambda} \end{bmatrix} = \begin{pmatrix} cos\,\phi & -sin\,\phi \\ sin\,\phi & cos\,\phi \end{pmatrix} \cdot \begin{bmatrix} \hbar \\ \lambda \end{bmatrix} \tag{26}$$

Here, ϕ indicates rotation angle and rotation is referred as N_5.

Hence, the augmented image N_μ is modelled as,

$$N_\mu = \{N_1, N_2, N_3, N_4, N_5\} \quad (27)$$

2-5- Feature Extraction based on an Augmented Image

It generates numerical representations of the data and applies dimensionality reduction. The augmented image N_μ is taken to extract suitable features.

2-5-1- SURF

SURF [28] eliminates the local robust feature through the hessian detector as well as the distribution descriptor, where the Hessian detector automatically detects the interest points. Thus, it is articulated by,

$$K(p,\varepsilon) = \begin{bmatrix} E_{pp}(p,\varepsilon)E_{pq}(p,\varepsilon) \\ E_{pa}(p,\varepsilon)E_{aa}(p,\varepsilon) \end{bmatrix} \quad (28)$$

Moreover, the distribution-based descriptor computes the interest region. It is separated into 4×4 sub-regions that are determined by wavelet response, and the expression of the vector I for every sub-region is calculated as,

$$I = \{\sum dp, \sum|dp|, \sum dq, \sum|dq|\} \quad (29)$$

Here, dp, dq wavelet response in p, q direction, $E_{pp}(p,\varepsilon)$ elucidates conv of the image with the second derivative of the Gaussian. Here, T_μ specifies the textural feature obtained from the SURF feature.

Moreover, textural feature T_μ is applied over statistical features and also Haralick texture features to acquire the resultant feature vector.

2-5-2- Statistical features

The statistical features are expressed below.

a) Mean

It describes the distribution's central value, corresponding to the average pixel intensity. Thus, mean [25] is originated as,

$$X_1 = \sum_{x=0}^{\mathbb{S}-1} x \times \mathbb{Z}(N_\mu) \quad (30)$$

Here, X_1 depicts the mean, $\mathbb{S}$ enumerates the overall gray levels, $\mathbb{Z}(N_\mu)$ implies the probability of the input image, and x elucidates each gray level of an image.

b) Variance

It computes the spread or dispersion of grey level values in an image based on the mean grey level [25], and it is determined as,

$$X_2 = \sum_{x=0}^{\mathbb{S}-1}(x-o)^2 * \mathbb{Z}(N_\mu) \quad (31)$$

Here, X_2 defines variance, o indicates the mean.

c) Skewness

It explains and measures the degree of asymmetry of the distribution of a particular feature around the mean [25], and it is enumerated as,

$$X_3 = o^{-3}\left[\sum_{x=0}^{\mathbb{S}-1}(x-o)^3 * \mathbb{Z}(N_\mu)\right] \quad (32)$$

d) Kurtosis

It evaluates the levelling of a distribution based on the normal distribution, and it is known as the fourth normalized moment [25]. Kurtosis is calculated as,

$$X_4 = o^{-4}\left[\sum_{x=0}^{\mathbb{S}-1}(x-o)^4 * \mathbb{Z}(N_\mu)\right] \quad (33)$$

Here, X_4 explicates kurtosis.

The features extracted from this phase X_μ are depicted as,

$$X_\mu = \{X_1, X_2, X_3, X_4\} \tag{34}$$

2-5-3- Haralick Features

The Haralick features considered for this research are,

a) Contrast

It evaluates local intensity variations, yielding smaller values when neighboring pixels share similar gray levels. Contrast [26] is modeled as,

$$Z_1 = \sum_x^D \sum_y^E (x-y)^2 W(N_\mu) \tag{35}$$

Here, Z_1 implies contrast, and D, E represents the dimension of an image $W(N_\mu)$ for $x = 1,2,\ldots\ldots,D; y = 1,2,\ldots\ldots,E$.

b) Correlation

It represents the level of correlation between the two pixels in each pixel pair. When the gray levels of the pixel pairs are largely interrelated, the Correlation [26] is predicted to be large. Thus, it is enumerated by,

$$Z_2 = \sum_x^D \sum_y^E \frac{(x-o)(y-o)W(N_\mu)}{\wp^2} \tag{36}$$

Here, Z_2 enumerates correlation, o and $\wp$ indicates the mean and standard deviation.

c) ASM

It determines the count of repeated pairs. When the presence of repeated pixel pairs is high, the ASM [26] is expected to be large. ASM is modelled as,

$$Z_3 = \sum_x^D \sum_y^E W^2(N_\mu) \tag{37}$$

Here, Z_3 refers to ASM.

d) Homogeneity

It evaluates the degree of smoothness within the pixel pair. When the gray levels of each pixel pair are the same, it is predictable to be maximum [26] and it is given by,

$$Z_4 = \sum_x^D \sum_y^E \frac{W(N_\mu)}{1+|x-y|} \tag{38}$$

Here, Z_4 depicts homogeneity.

The Haralick features Z_μ attained from this phase are enumerated as,

$$Z_\mu = \{Z_1, Z_2, Z_3, Z_4\} \tag{39}$$

Therefore, the feature vector $M_{\mu 1}$ is modelled as,

$$M_{\mu 1} = \{X_\mu, Z_\mu\} \tag{40}$$

Here, X_μ denotes statistical features and Z_μ indicates Haralick features.

2-5-4- CNN Feature

CNN [27] comprises interconnected layers, wherein every layer has a set of neurons. Here, N_μ is given as the input to CNN structure. The activations of neurons in a layer arise from applying a function to the outputs from the preceding layer, which is determined as $\lambda = \mathbb{k}(N_\mu)$. The layers of CNN are Fully connected (FC), conv, activation and pooling, where the conv layer is extracted from the CNN network. A conv layer uses parameter sharing, where identical weights and a bias are applied across neurons, making up the filter. Assume filter dimension as $N_n \times N_n$, then every neuron is linked to $N_n \times N_n$region of neurons in the prior layer. Moreover, the resultant for $(J_j, K_k)^{th}$ neuron is determined as,

$$M_{\mu 2} = \sum_{L_l=0}^{N_n-1} \sum_{M_m=0}^{N_n-1} W w_{L_l,M_m} N_{\mu(J_j+L_l,K_k+M_m)} + B_b \quad (41)$$

Here, $M_{\mu 2}$ indicates CNN feature, $W w_{L_l,M_m}$ elucidates the weight connection between ${L_l}^{th}$and ${M_m}^{th}$ neuron, and B_b enumerates bias. The configuration of the CNN features is presented in Figure 3.

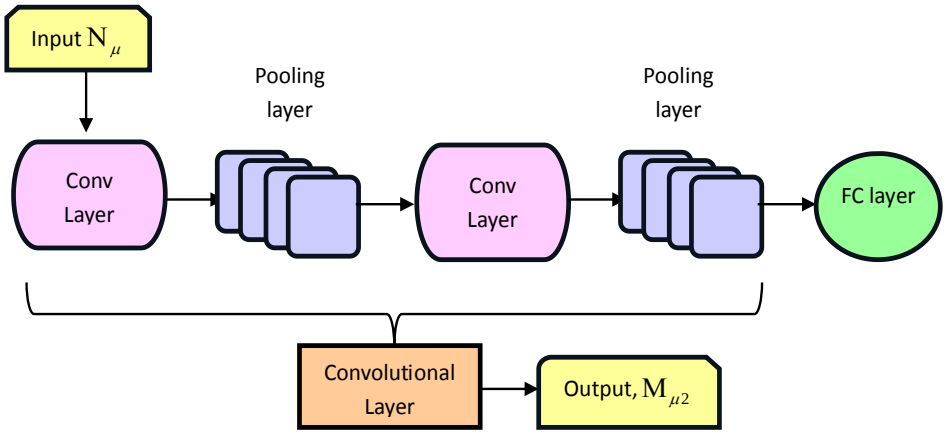


Figure 3. Outlook of CNN feature

Finally, the resultant feature vector M_μ is depicted as,

$$M_\mu = \{M_{\mu 1}, M_{\mu 2}\} \quad (42)$$

2-6- Melanoma Classification Using FWSAO_CNN-Based TL

Melanoma is considered a highly deadly type of skin cancer. However, distinguishing it from a nevus is challenging because both share similar features and symptoms. This malignancy exhibits a mortality rate that outstrips other dermatological cancers. When the occurrence of melanoma is increasing amongst children, early diagnosis can enhance the survival rate. However, the cost and time to evaluate all patients for melanoma are considerable. Therefore, the melanoma classification is accomplished by FWSAO_CNN-based TL. The structure of this newly developed module is explained below.

2-6-1- GoogleNet with CNN-Based TL

Figure 4 enumerates GoogleNet with CNN-based TL representation. Here, the feature vector M_μ is given to train the GoogleNet module. After performing the training process, the hyperparameters are fetched from the trained GoogleNet. Here, the fetched hyperparameters are signified as $\mathbb{F}_\mu$. Thereafter, the feature vector M_μ and fetched hyperparameters are subjected to the CNN module for the training process. Finally, the classified outcome is attained.

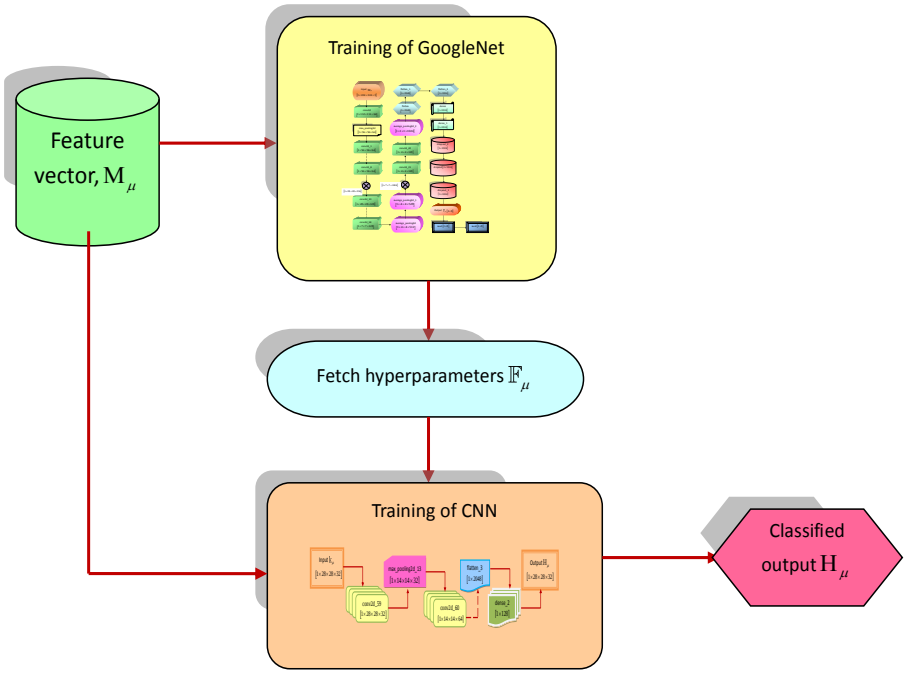


Figure 4. Block diagram of GoogleNet with CNN-based TL

a) GoogleNet

GoogLeNet [30] is the advanced version of the Inception architecture. It employs a deeper and wider Inception network with utmost quality, yet it adds to the ensemble-based technique for slightly improving the outcomes. This architecture is trained with diverse image patch sampling techniques.

The entire conv layer, along with the Inception model, utilize the ReLU function. The dimension of the receptive field in this structure is 224×224 in the RGB color space with zero mean. $[3 \times 3]$ eradicate and $[5 \times 5]$ reduce enumerates number of 1×1 filters in the reduction layer employed before $[3 \times 3]$ as well as $[5 \times 5]$ conv functions. GoogleNet has 22 deep layers with parameters, or 27 layers with parameters and pooling layers. The layers employed for constructing the network are 100. The linear layer allows for convenient adjustment of the networks to different labels, but it is primarily utilized for ease of use and does not anticipate it to significantly impact performance. Moreover, it is identified that a transmission from FC layers to average pooling enhanced the accuracy, yet the usage of dropout remained even when FC layers are eliminated.

Owing to network depth, the capability of gradients back through all layers raises some issues. The superior performance in shallower networks suggests that the features generated by intermediate layers are quite discriminative. For improving discrimination in a small-scale classifier, auxiliary classifiers are designed and connected to the intermediate layers. It intends to conquer the vanishing gradient issue by acting as a form of regularization. These auxiliary classifiers have small conv networks positioned on top of the results of Inception models. During training, its loss information is included in the overall network loss with a discount weight. Nevertheless, these auxiliary networks are removed at inference time. This analysis shows that the influence of these auxiliary networks is relatively minimal, which is enough to attain better outcomes.

Therefore, the classified output is depicted as H_{μ}. Figure 5 shows the GoogleNet structure.

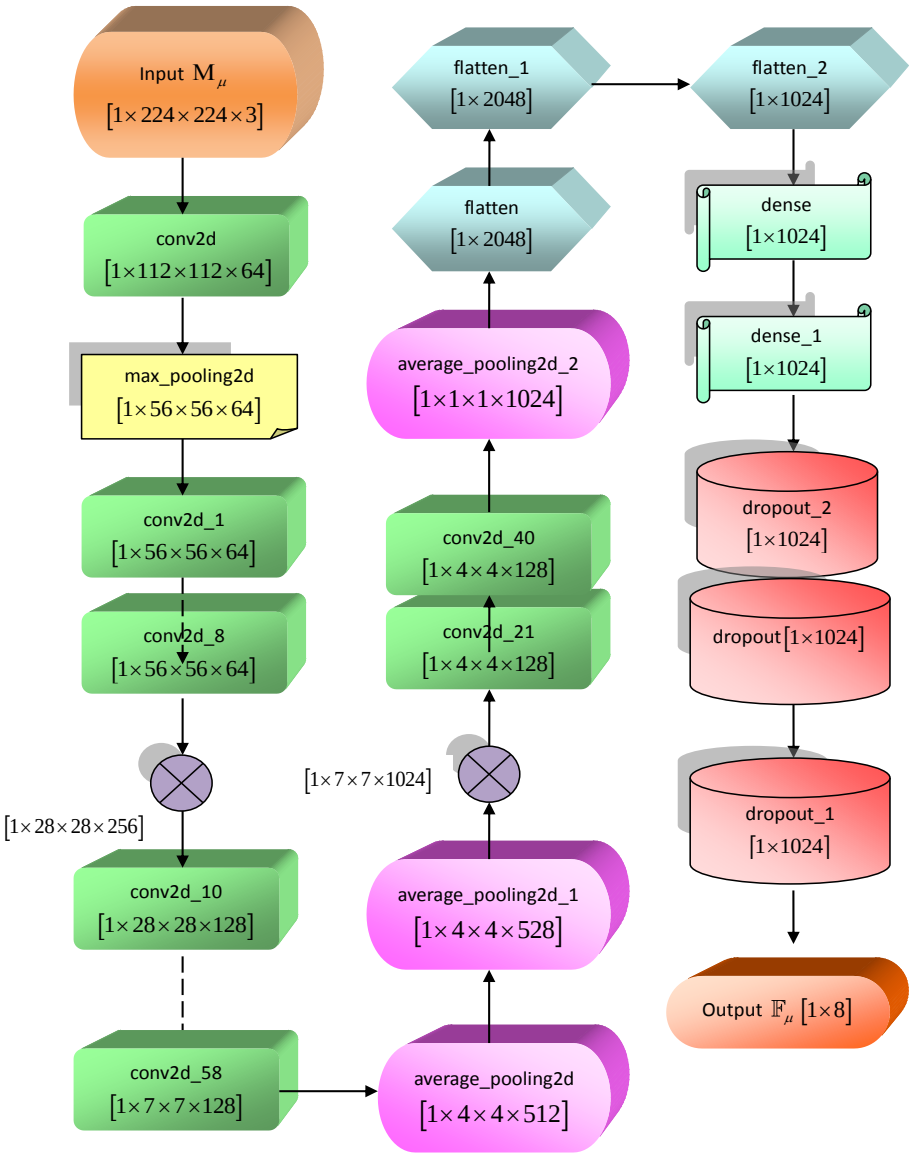


Figure 5. Structure of GoogleNet

b) CNN module

CNN [29] has conv layers, maxpooling layers, activation functions like Rectified Linear Unit (ReLU) or sigmoid and a softmax layer that creates a well-determined probability distribution for classification tasks. Here, the M_μ and $\mathbb{F}_\mu$ are given as the input for the CNN structure, which is commonly implied as $\mathbb{K}_\mu$.

Generally, the initial conv layer extracts subsequent layers, that serves as higher-level feature extractors. The pooling layer is deployed for determining the mean of every patch of the feature map from the preceding layer. ReLU is employed for addressing the saturation issue. The arithmetic form of CNN is specified as A that is the configuration of a sequence of functions, and it is determined as,

$$G = g_I g_{I-1} g_{I-2} g \cdots g_2 g_1 \tag{43}$$

Every function A exploits a layer that considers the result of the former layer. Once the training process is performed, CNN is used as a classifier based on TL, which is deployed to enhance the performance. The classified output is specified as H_μ. Figure 6 displays the CNN structure.

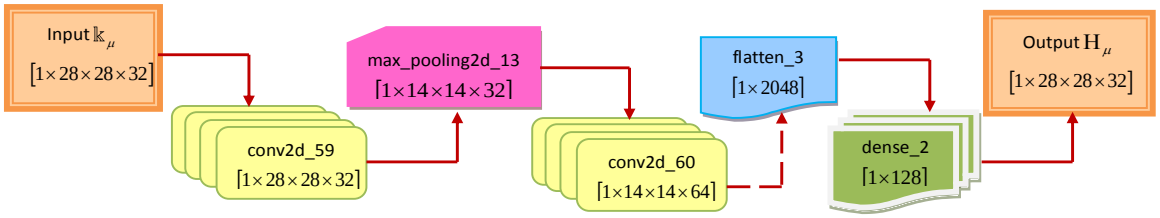


Figure 6. CNN structure

2-6-2- Training Algorithm of GoogleNet with CNN-Based TL

FWSAO is designed for training GoogleNet with CNN-based TL, and it is designed based on the fractional calculus (FC) and WSAO. The algorithm of WSAO is already elaborated in section 2.3.2. In addition to that, the updated solution of WSAO is considered for training GoogleNet with CNN-based TL by applying FC.

a) Fitness value

It is computed to achieve the utmost outcomes with the appropriate candidate solution, where its arithmetic expression is determined in the form of,

$$Fit = \frac{1}{v}\sum_{\mu=1}^{v}\left[Z_\mu - H_\mu\right]^2 \tag{44}$$

b) Algorithm steps

The steps followed for FWSAO are described in section 2.3.2 c). The updated solution of WSAO is modelled as,

$$b_\omega^{\tau+1} = \frac{\ell\left[\upsilon_1 \times OC \times b_{agent}^{\tau} - \upsilon_2 \times OC \times b_{worst}^{\tau}\right] - \Re_\omega^{\tau}\left[\upsilon_1 \times OC - 1 - \upsilon_2 \times OC\right]}{\ell \times (\upsilon_1 \times OC - \upsilon_2 \times OC)} \tag{45}$$

To apply FC, subtract b_ω^{τ} on both sides; therefore, the above expression is rewritten as,

$$b_\omega^{\tau+1} - b_\omega^{\tau} = \frac{\ell\left[\upsilon_1 \times OC \times b_{agent}^{\tau} - \upsilon_2 \times OC \times b_{worst}^{\tau}\right] - \Re_\omega^{\tau}\left[\upsilon_1 \times OC - 1 - \upsilon_2 \times OC\right]}{\ell \times (\upsilon_1 \times OC - \upsilon_2 \times OC)} - b_\omega^{\tau} \tag{46}$$

Apply the FC to the above expression for better stability, then the updated solution is determined as,

$$b_\omega^{\tau+1} - \delta b_\omega^{\tau} - \frac{1}{2} b_\omega^{\tau-1} - \frac{1}{6} b_\omega^{\tau-2} - \frac{1}{24}\delta(1-\delta)(2-\delta) b_\omega^{\tau-3} = \frac{\ell\left[\upsilon_1 \times OC \times b_{agent}^{\tau} - \upsilon_2 \times OC \times b_{worst}^{\tau}\right]}{\frac{-\Re_\omega^{\tau}\left[\upsilon_1 \times OC - 1 - \upsilon_2 \times OC\right]}{\ell \times (\upsilon_1 \times OC - \upsilon_2 \times OC)}} - b_\omega^{\tau} \tag{47}$$

$$b_\omega^{\tau+1} = (\delta - 1) b_\omega^{\tau} + \frac{1}{2} b_\omega^{\tau-1} + \frac{1}{6} b_\omega^{\tau-2} + \frac{1}{24}\delta(1-\delta)(2-\delta) b_\omega^{\tau-3} + \frac{\ell\left[\upsilon_1 \times OC \times b_{agent}^{\tau} - \upsilon_2 \times OC \times b_{worst}^{\tau}\right]}{\frac{-\Re_\omega^{\tau}\left[\upsilon_1 \times OC - 1 - \upsilon_2 \times OC\right]}{\ell \times (\upsilon_1 \times OC - \upsilon_2 \times OC)}} \tag{48}$$

Here, δ is defined as constant.

3- Results and Discussion

The results of FWSAO_CNN-based TL are evaluated in corresponding sub-sections based on traditional algorithms.

3-1- Experimental Setup

The optimized FWSAO_CNN-based TL framework is efficiently implemented in Python.

3-2- Dataset Description

The proposed FWSAO_CNN-based TL is evaluated using two datasets which are described in the below section.

3-2-1- Skin Lesion Images for Melanoma Classification

Skin Lesion Images for Melanoma Classification database (dataset 1) [18] is composed of the training data for the ISIC 2019 challenge. Moreover, the database for ISIC 2019 has 25,331 images which belong to nine classes. It is considerably employed for the process of dermoscopic image classification. The size of this database is 1.29 MB.

3-2-2- Melanoma Skin Cancer Dataset of 10000 Images

The Melanoma Skin Cancer Dataset of 10000 images (dataset 2) [31] contains 10000 images, of which 9600 images are used for training and 1000 images are used for testing. The training set consists of two directories of where one directory consists of 5000 benign files and 4605 malignant files. The test set has 2 directories, one of which has 500 benign files and another has 500 malignant files.

3-3- Experimental Outcomes

Figure 7 demonstrates the Image results of melanoma classification. Figure 7 a) input image-1, Figure 7 b) filter image-1, Figure 7 c) segmented image-1, Figure 7 d) segmented color image-1, Figure 7 e) sheared image-1, Figure 7 f) shifted image-1, Figure 7 g) rotated image-1, Figure 7 h) erased image-1, Figure 7 i) cropped image-1, Figure 7 j) classified image-1, Figure 7 k) input image-2, Figure 7 l) filter image-2, Figure 7 m) segmented image-2, Figure 7 n) segmented color image-2, Figure 7 o) sheared image-2, Figure 7 p) shifted image-2, Figure 7 q) rotated image-2, Figure 7 r) erased image-2, Figure 7 s) cropped image-2, and Figure 7 t) classified image-2.

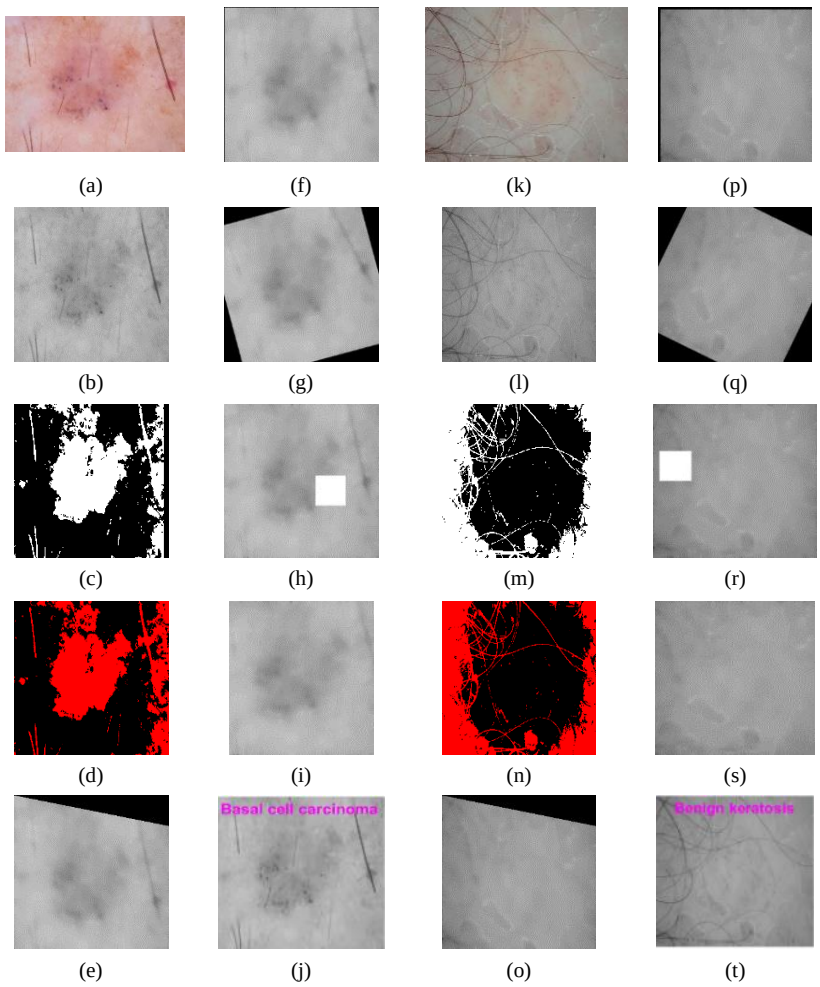


Figure 7. Image results of melanoma classification, a) input image-1, b) filter image-1, c) segmented image-1, d) segmented color image-1, e) sheared image-1, f) shifted image-1, g) rotated image-1, h) erased image-1, i) cropped image-1, j) classified image-1, k) input image-2, l) filter image-2, m) segmented image-2, n) segmented color image-2, o) sheared image-2, p) shifted image-2, q) rotated image-2, r) erased image-2, s) cropped image-2, t) classified image-2.

3-4- Evaluation Metrics

The metrics analyzed for FWSAO_CNN-based TL are expressed below.

3-4-1- Accuracy

It computes the accurate classification of test cases to the overall test cases [32], which is enumerated as,

$$M_1 = \frac{J_p+J_n}{J_p+J_n+L_p+L_n} \tag{49}$$

3-4-2- TPR

The percentage of accurate classification from the overall number of positive cases is known as TPR [32], which is determined as,

$$M_2 = \frac{J_p}{J_p+L_n} \tag{50}$$

3-4-3- TNR

The percentage of accurate negative classification from the overall number of negative cases is known as TNR [32], which is computed as,

$$M_3 = \frac{J_n}{J_n+L_p} \tag{51}$$

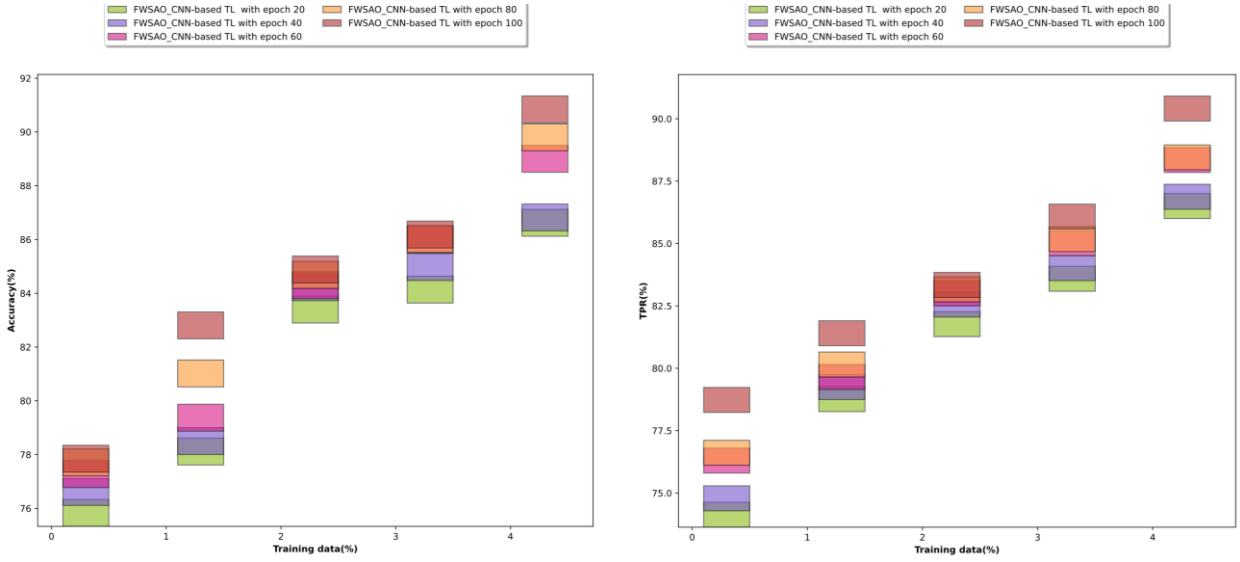
Here, L_p, L_n, specifies false positive and negative as well as J_p, J_n, implies true positive and negative.

3-5- Performance Analysis

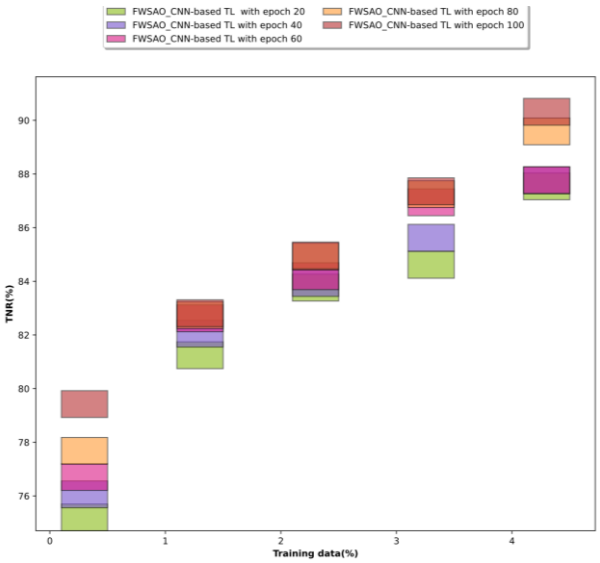
The FWSAO_CNN-based TL evaluation regarding epoch values is carried out with training data and the K-value.

3-5-1- Analysis of FWSAO_CNN-based TL with Training Data

The FWSAO_CNN-based TL analysis with training data is demonstrated in Figure 8. Here, the epoch values are from 20 to 100 and also the training data is taken as 90%. Figure 8 a) depicts the FWSAO_CNN-based TL assessment regarding accuracy. The performance values of FWSAO_CNN-based TL with accuracy gained 77.847%, 82.804%, 84.887%, 86.186% and 90.840% for epochs 20 to 100. In Figure 8 b), the FWSAO_CNN-based TL evaluation relating to TPR is represented. In this assessment, TPR achieved for the FWSAO_CNN-based TL model is 78.730%, 81.403%, 83.341%, 86.079% and 90.409% for epochs from 20 to 100. Figure 8 c) elucidates FWSAO_CNN-based TL valuation based on TNR. Here, the FWSAO_CNN-based TL with respect to TNR gained 79.423%, 82.811%, 84.964% and 87.356% and 90.318% for epochs 20 to 100. The evaluation results demonstrate that the performance of the proposed FWSAO_CNN-based TL model improves consistently with increasing epochs. This indicates that the model benefits from extended training, thereby proving its reliability and effectiveness in melanoma classification.



(a) (b)

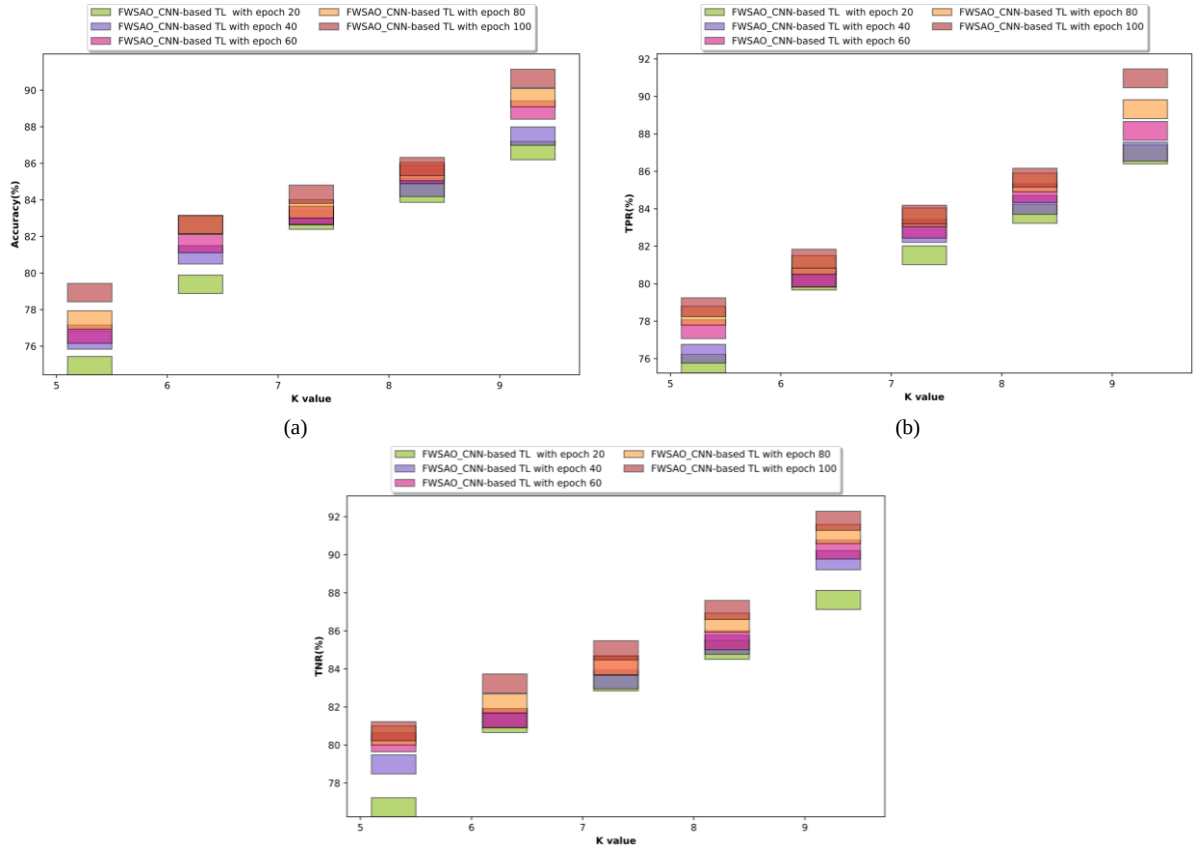


(c)

Figure 8. Analysis of FWSAO_CNN-based TL with training data, a) Accuracy, b) TPR, c) TNR

3-5-2- Analysis of FWSAO_CNN-based TL with K-value

The FWSAO_CNN-based TL analysis is illustrated in Figure 9. The K-value is considered as 9 across different epochs. In Figure 9 a), the results of FWSAO_CNN-based TL assessment of accuracy are presented, showing the performance values of 78.929%, 82.649%, 84.308%, 85.822% and 90.644% for epochs 20 to 100, respectively. Figure 9 b) provides an evaluation of the FWSAO_CNN-based TL model regarding TPR. The FWSAO_CNN-based TL based on TPR obtained 78.744%, 81.338%, 83.688%, 85.660% and 90.964% across the epochs from 20 to 100. Figure 9 c) represents FWSAO_CNN-based TL performance in terms of TNR, which gained 80.726%, 83.233%, 84.978%, 87.097%, and 91.785% for epochs 20 to 100. The assessment results indicate that the proposed FWSAO_CNN-based TL model exhibits consistent improvement across accuracy, TPR, and TNR as the number of training epochs increases from 20 to 100. This highlights the effectiveness and robustness of the proposed model in melanoma classification



(a) (b)

(c)

Figure 9. Analysis of FWSAO_CNN-based TL with K-value, a) Accuracy, b) TPR, c) TNR

3-6- Comparative Methods

The FWSAO_CNN-based TL is evaluated with the traditional methods such as LCNet [9], GoogleNet [11], GC-SCNN+SVM [14], Explainable CNN-based stacked ensemble [6], ViT-L/16 [16] and EDLCS [17].

3-7- Comparative Analysis

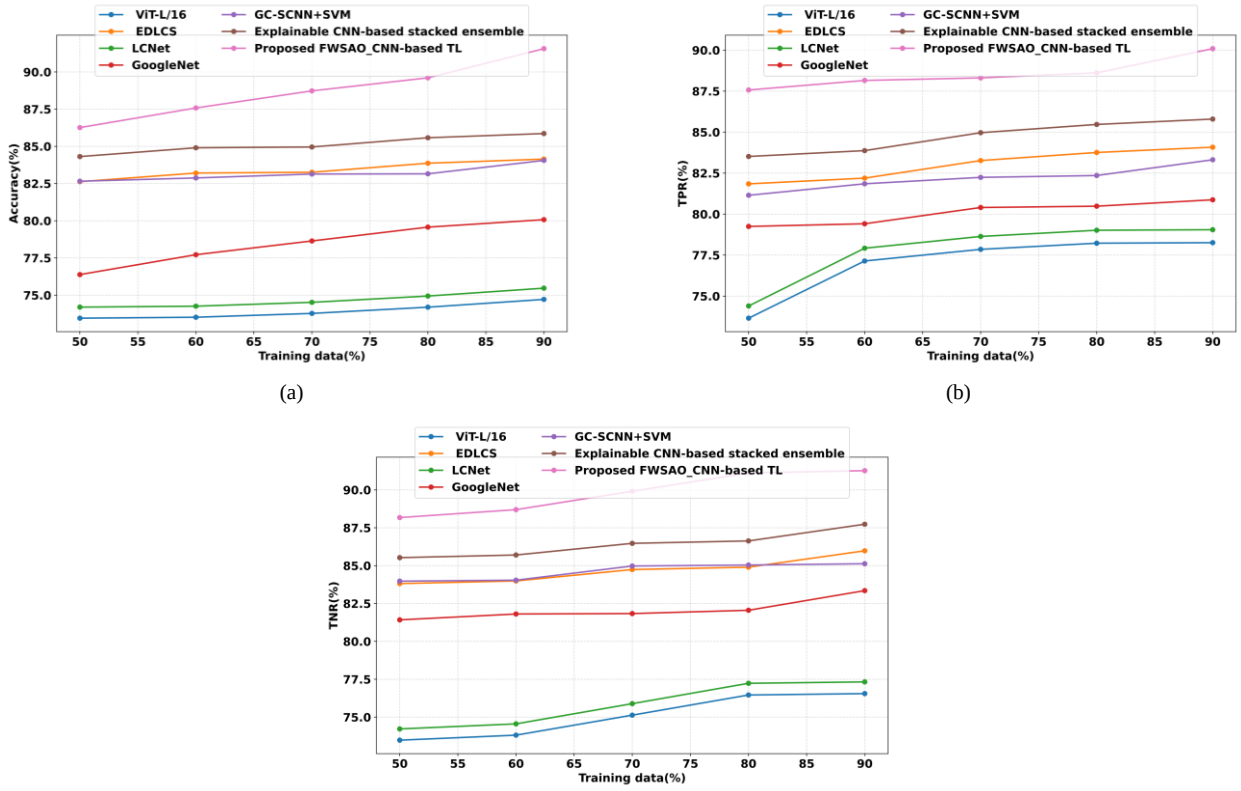
The FWSAO_CNN-based TL regarding training data and K-value is carried out in the following sections.

3-7-1- Evaluation of FWSAO_CNN-based TL Employing Dataset 1

The proposed FWSAO_CNN-based TL is assessed using training data and K-value, which is given in the section below.

a) Analysis of FWSAO_CNN-based TL with Training Data

Figure 10 designates FWSAO_CNN-based TL valuation regarding training data, where its value is assumed as 90%. In Figure 10 a), FWSAO_CNN-based TL assessment is carried out, based on accuracy is illustrated. The FWSAO_CNN-based TL is evaluated based on the accuracy obtained 91.565% where the conventional techniques, like ViT-L/16, EDLCS, LCNet, GoogleNet, GC-SCNN+SVM, and Explainable CNN-based stacked ensemble, achieved accuracy of 74.715%, 84.144%, 75.469%, 80.074%, 84.055% and 85.861%. Figure 10 b) enumerates the FWSAO_CNN-based TL assessment regarding TPR. Here, the TPR obtained for the FWSAO_CNN-based TL module is 90.090%, while the prior methods like ViT-L/16, EDLCS, LCNet, GoogleNet, GC-SCNN+SVM, and Explainable CNN-based stacked ensemble acquired TPRs of 78.258%, 84.081%, 79.048%, 80.873%, 83.315%, and 85.797%. In Figure 10 c), the FWSAO_CNN-based TL analysis concerning TNR is designated. This shows the better value of FWSAO_CNN-based TL in terms of TNR of 91.269%, where the traditional methodologies, namely, ViT-L/16, EDLCS, LCNet, GoogleNet, GC-SCNN+SVM, and Explainable CNN-based stacked ensemble, accomplished TNRs of 77.325%, 83.350%, 77.325%, 83.35%, 85.125% and 87.734%. The analysis results demonstrate that the proposed FWSAO_CNN-based TL consistently outperforms conventional methods across all the considered metrics, thereby indicating its superior effectiveness and reliability in melanoma classification.



(a) (b)

(c)

Figure 10. Valuation of FWSAO_CNN-based TL with training data, a) Accuracy, b) TPR, c) TNR

b) Analysis of FWSAO_CNN-based TL with K-value

Figure 11 displays the analysis of FWSAO_CNN-based TL utilizing K-values of 5-9. In Figure 11 a), the accuracy of the FWSAO_CNN-based TL method is explicated. The evaluation of FWSAO_CNN-based TL with accuracy gained

91.358%, while previous approaches such as ViT-L/16, EDLCS, LCNet, GoogleNet, GC-SCNN+SVM, and Explainable CNN-based stacked ensemble achieved 78.542%, 85.984%, 79.335%, 82.189%, 84.594% and 87.738%. Figure 11 b) shows the assessment of FWSAO_CNN-based in accordance with TPR. In this analysis, the TPR attained by the FWSAO_CNN-based TL module is 89.084%, where existing techniques, namely, ViT-L/16, EDLCS, LCNet, GoogleNet, GC-SCNN+SVM, and Explainable CNN-based stacked ensemble, acquired TPR of 79.407%, 85.984%, 80.209%, 81.467%, 83.920% and 86.700%. In Figure 11 c), the FWSAO_CNN-based TL based on TNR is demonstrated. The TNR obtained by the FWSAO_CNN-based TL framework is 89.530%, where the TNRs attained by ViT-L/16, EDLCS, LCNet, GoogleNet, GC-SCNN+SVM, and Explainable CNN-based stacked ensemble are 77.075%, 83.623%, 7.854%, 81.31%, 82.761% and 85.329%. The assessment results confirm that the proposed FWSAO_CNN-based TL method achieves superior performance in melanoma classification compared to existing methods. This demonstrates the robustness and effectiveness of the proposed approach in effective melanoma classification.

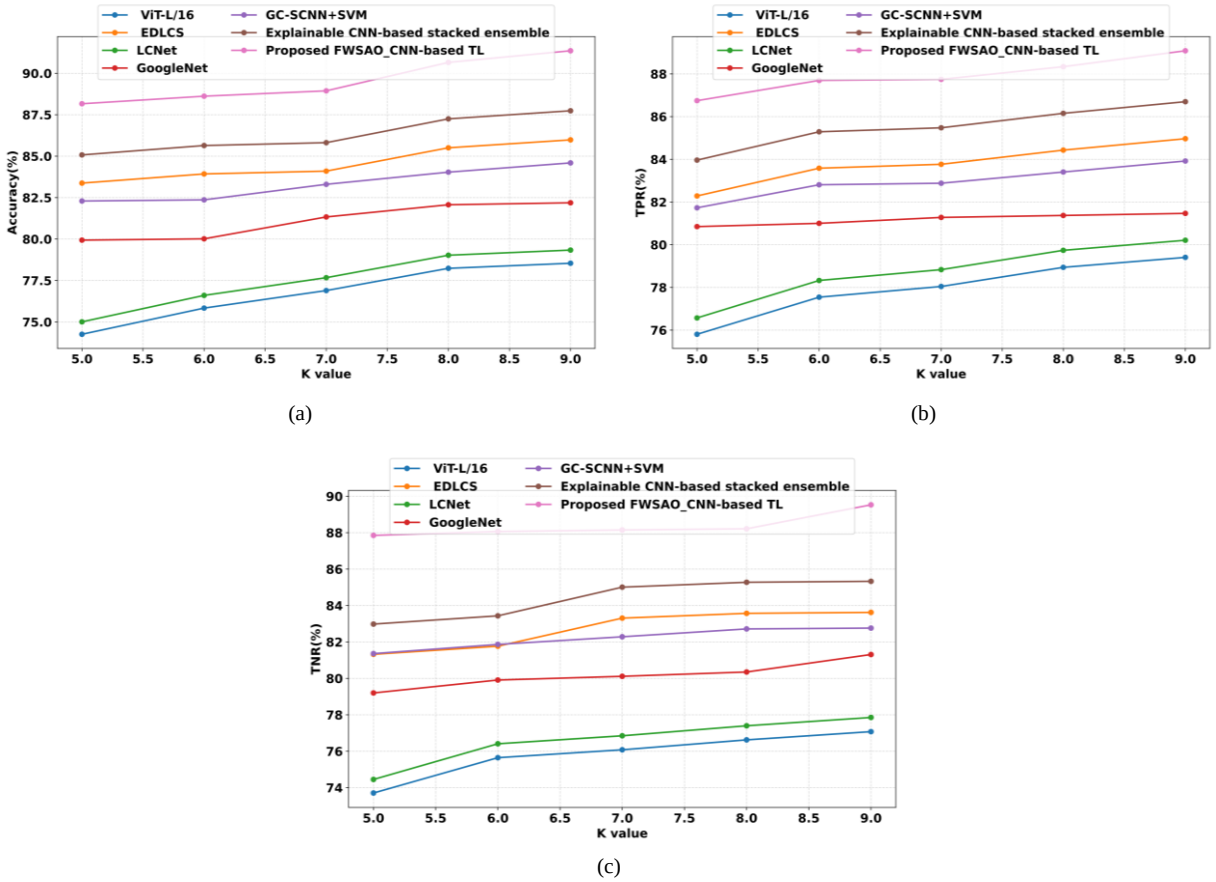


Figure 11. Analysis of FWSAO_CNN-based TL with K-value, a) Accuracy, b) TPR, c) TNR

3-7-2- Evaluation of FWSAO_CNN-based TL Employing Dataset 2

Figure 12 displays the evaluation of FWSAO_CNN-based TL valuation regarding training data 50-90% and employing dataset 2. In Figure 12 a), the FWSAO_CNN-based TL assessment based on accuracy is illustrated. The FWSAO_CNN-based TL obtained an accuracy of 90.191% where the conventional techniques, like ViT-L/16, EDLCS, LCNet, GoogleNet, GC-SCNN+SVM, and Explainable CNN-based stacked ensemble, achieved an accuracy of 72.107%, 73.594%, 75.337%, 78.872%, 82.794% and 84.573% for the training data 90%. Figure 12 b) enumerates FWSAO_CNN-based TL assessment regarding TPR. Here, the TPR obtained for the FWSAO_CNN-based TL module is 88.738%, while the prior methods like ViT-L/16, EDLCS, LCNet, GoogleNet, GC-SCNN+SVM, and Explainable CNN-based stacked ensemble acquired TPRs of 75.526%, 77.084%, 77.862%, 79.66%, 82.065%, 84.51%, when the training data is 90%. In Figure 12 c), the FWSAO_CNN-based TL analysis concerning TNR is designated. For the training data 90%, the FWSAO_CNN-based TL attained a TNR of 89.9%, where the traditional methodologies, namely, ViT-L/16, EDLCS, LCNet, GoogleNet, GC-SCNN+SVM, and Explainable CNN-based stacked ensemble, accomplished TNRs of 73.88%, 75.404%, 76.165%, 82.1%, 83.848% and 86.418%. The assessment results demonstrate that the FWSAO_CNN-based TL model delivers consistently high performance across different training data proportions on dataset 2. This proves that FWSAO_CNN-based TL model is capable of effectively classifying melanoma.

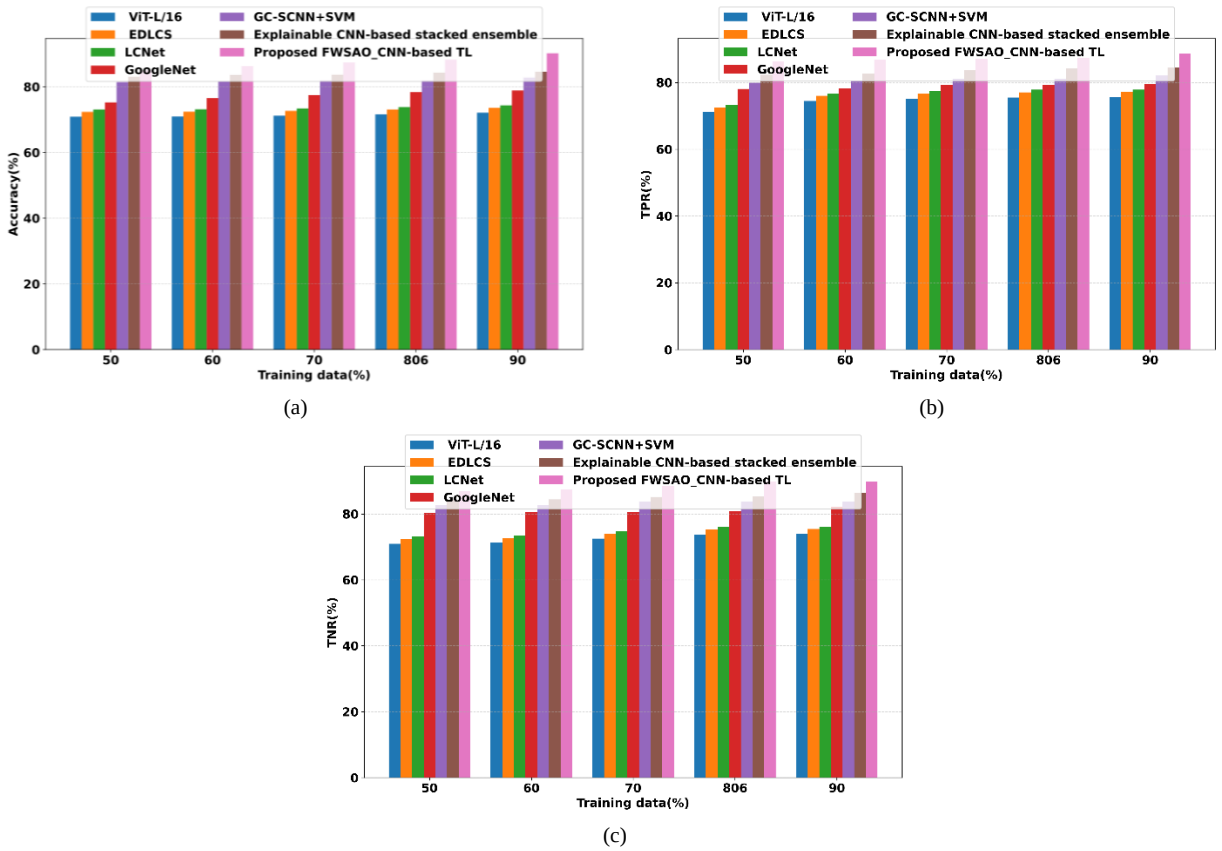


Figure 12. Assessment of FWSAO_CNN-based TL with training data employing dataset 2, a) Accuracy, b) TPR, c) TNR

3-8- Segmentation Accuracy

The WSAO_LinkNet is evaluated for segmentation accuracy with the conventional segmentation techniques such as U-Net [33], K-Net [34], Deepjoint segmentation [35] and Deep Fuzzy Clustering (DFC) [36]. Figure 13 enumerates the segmentation accuracy of WSAO_LinkNet based on training data employing dataset 1. When training data is 90%, the WSAO_LinkNet obtained segmentation accuracy of 90.899%. Meanwhile, the traditional approaches, namely, U-Net, K-Net, Deepjoint segmentation and DFC, achieved 79.668%, 81.398%, 84.524% and 87.659%. While comparing the WSAO_LinkNet with the classical segmentation models like U-Net, K-Net, Deepjoint segmentation and DFC, the performance enhancement of WSAO_LinkNet obtained 12.355%, 10.452%, 7.013% and 3.564%. The evaluation results demonstrate that the proposed WSAO_LinkNet achieves superior segmentation performance compared to conventional methods. This indicates that WSAO_LinkNet provides more precise and reliable lesion segmentation, which is crucial for accurate melanoma classification.

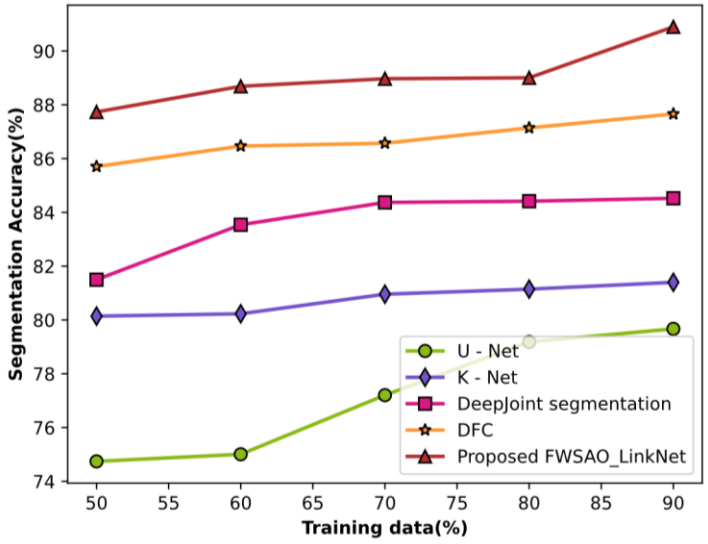
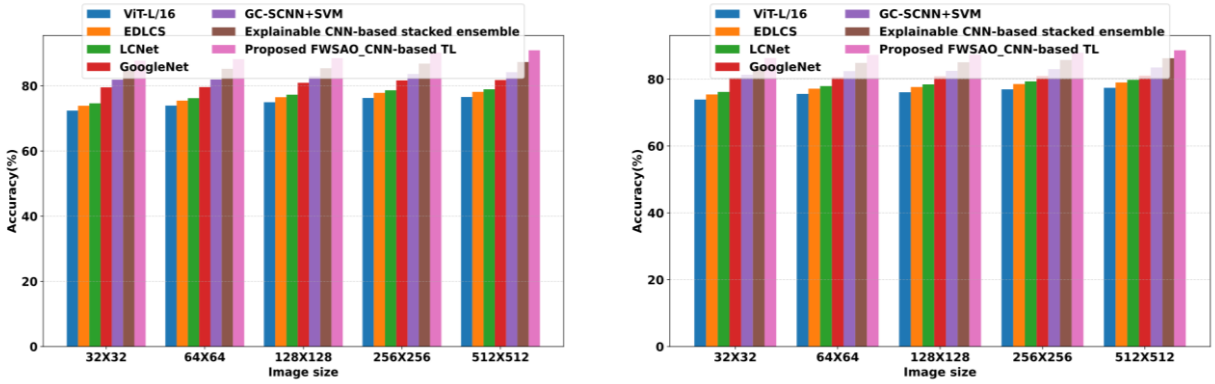


Figure 13. Segmentation Accuracy of training data employing dataset 1

3-9- Scalability Analysis

Figure 14 illustrates the scalability assessment of the FWSAO_CNN-based TL by varying the image size and using accuracy metrics, employing the two datasets. Image sizes used for this evaluation are 32x32, 64x64, 128x128, 256x256 and 512x512. Figure 14 a) depicts the scalability analysis of the FWSAO_CNN-based TL and the prior methods utilizing dataset 1. When the image size is 256x256, the accuracy attained by ViT-L/16 is 76.27%, EDLCS is 77.84%, LCNet is 78.63%, GoogleNet is 81.66%, GC-SCNN+SVM is 83.62%, and Explainable CNN-based stacked ensemble is 86.81%, while the proposed FWSAO_CNN-based TL attained an accuracy of 90.2%. The scalability analysis regarding dataset 2 is displayed in Figure 14 b). The accuracy attained by ViT-L/16, EDLCS, LCNet, GoogleNet, GC-SCNN+SVM, Explainable CNN-based stacked ensemble and FWSAO_CNN-based TL are 77.41%, 79.01%, 79.81%, 81.06%, 83.5%, 86.27% and 88.64%, when the image size is 512x512. The scalability assessment demonstrates that the proposed FWSAO_CNN-based TL model maintains high accuracy across varying image sizes and datasets. This indicates that the model is robust and scalable in effectively handling different image resolutions without significant performance degradation in melanoma classification tasks.

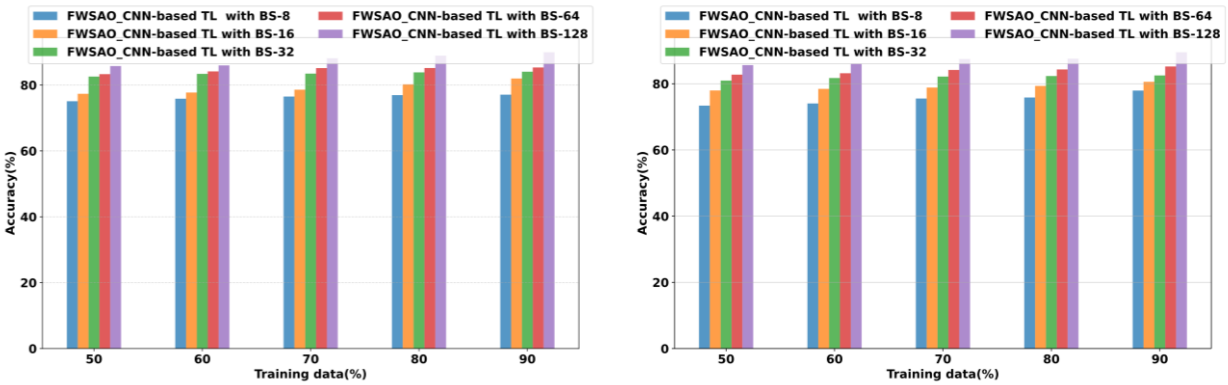


(a) (b)

Figure 14. Scalability analysis utilizing a) dataset 1 and b) dataset 2

3-10- Sensitivity Analysis

The sensitivity analysis of the FWSAO_CNN-based TL with accuracy by varying the training data from 50-90% and batch size from 8-128 employing the two datasets is displayed in Figure 15. The sensitivity assessment utilizing dataset 1 is illustrated in Figure 15 a). Here, the accuracy attained by FWSAO_CNN-based TL with batch size 8 is 77.07%, batch size 16 is 81.98%, batch size 32 is 84.04%, batch size 64 is 85.32% and batch size 128 is 89.93%. Figure 15 b) displays the sensitivity analysis employing dataset 2. The accuracy attained by FWSAO_CNN-based TL with batch size 8 is 77.94%, FWSAO_CNN-based TL with batch size 16 is 80.59%, FWSAO_CNN-based TL with batch size 32 is 82.51%, FWSAO_CNN-based TL with batch size 64 is 85.22% and FWSAO_CNN-based TL with batch size 128 is 89.5%. The sensitivity analysis indicates that the performance of the FWSAO_CNN-based TL model improves consistently with increasing batch size across both datasets. This highlights the robust behavior of FWSAO_CNN-based TL under varying hyperparameter conditions, for efficient melanoma classification.



(a) (b)

Figure 15. Sensitivity analysis employing a) dataset 1 and b) dataset 2

3-11- Analysis based on Computational Time and Memory

Table 2 displays the computational time and memory analysis of the proposed FWSAO_CNN-based TL employing the two datasets. In this assessment, the proposed FWSAO_CNN-based TL requires the processing time of 3.0 s and memory of 20.0 MB with dataset 1 and processing time of 5.55 s and memory of 22.5 MB of memory with dataset 2, which are the lowest among all evaluated models, such as ViT-L/16, EDLCS, LCNet, GoogleNet, GC-SCNN+SVM, and Explainable CNN-based stacked ensemble. The computational analysis demonstrates that the proposed FWSAO_CNN-based TL model is highly efficient compared to existing methods in melanoma classification.

Table 2. Computational time and memory assessment

Metrics/ Methods	ViT-L/16	EDLCS	LCNet	GoogleNet	GC-SCNN+SVM	Explainable CNN-based stacked ensemble	Proposed FWSAO_CNN-based TL
Dataset 1							
Computational time (s)	10.0	8.80	7.70	6.50	5.30	4.20	3.00
Computational Memory (MB)	45.0	40.8	36.7	32.5	28.3	24.2	20.0
Dataset 2							
Computational time (s)	12.5	11.6	10.3	9.54	8.60	7.56	5.55
Computational Memory (MB)	51.6	50.6	49.3	47.5	36.5	29.6	22.5

3-12- Analysis based on Confidence Interval

The confidence interval assessment of FWSAO_CNN-based TL regarding the accuracy metrics and utilizing the two datasets is illustrated in Table 3. In this analysis, the proposed FWSAO_CNN-based TL achieves a confidence interval of 86.1–92.1 with dataset 1 and 83.3-90.9 with dataset 2, which are highest range among all compared methods, such as ViT-L/16, EDLCS, LCNet, GoogleNet, GC-SCNN+SVM, and Explainable CNN-based stacked ensemble. This analysis highlights that the proposed FWSAO_CNN-based TL model is robust, stable, and trustworthy for melanoma classification tasks.

Table 3. Confidence Interval analysis

Methods	Confidence Intreval	
	Dataset 1	Dataset 2
ViT-L/16	[72.4-76.1]	[69.5-74.3]
EDLCS	[81.2-85.4]	[71.2-75.1]
LCNet	[74.1-76.7]	[72.8-75.9]
GoogleNet	[75.2-81.6]	[73.4-79.4]
GC-SCNN+SVM	[81.9-85.3]	[80.2-83.5]
Explainable CNN-based stacked ensemble	[83.2-86.7]	[81.9-86.2]
Proposed FWSAO_CNN-based TL	[86.1-92.1]	[83.3-90.9]

3-13- Comparative Discussions

Table 4 articulates comparative values of FWSAO_CNN-based TL. Here, the novel framework is compared with the existing models like LCNet, GoogleNet, GC-SCNN+SVM, and Explainable CNN-based stacked ensemble. The accuracy of the proposed approach obtained 91.565%, while the former methods observed 74.715%, 84.144%, 75.469%, 80.074%, 84.055% and 85.861%. This shows that the high accuracy of FWSAO_CNN-based TL shows better representation of the model's ability to differentiate between the two classes due to the balanced dataset. The FWSAO_CNN-based TL based on TNR gained 90.090%, where the prior methods acquired 78.258%, 84.081%, 79.048%, 80.873%, 83.315% and 85.797%. This expresses that the proposed technique is efficient in detecting melanoma cases, which is important for preventing undiagnosed cases of cancer. Moreover, the TNR of FWSAO_CNN-based TL obtained 91.269%, where the traditional approaches accomplished 77.325%, 83.35%, 77.325%, 83.350%, 85.125% and 87.734%. This demonstrates that the designed framework reduces the risk of psychological distress and unnecessary medical procedures. Therefore, this analysis based on the evaluation measures demonstrates that the proposed approach is easy to understand and interpret, avoids irrelevant treatments, and provides superior generalizability.

Table 4. Comparative Discussions

Setups	Metrics/ Methods	ViT-L/16	EDLCS	LCNet	GoogleNet	GC-SCNN+SVM	Explainable CNN-based stacked ensemble	Proposed FWSAO_CNN-based TL
Dataset 1								
Training data=90%	Accuracy (%)	74.715	84.144	75.469	80.074	84.055	85.861	91.565
	TPR (%)	78.258	84.081	79.048	80.873	83.315	85.797	90.090
	TNR (%)	77.325	83.350	77.325	83.350	85.125	87.734	91.269
K-value=9	Accuracy (%)	78.542	85.984	79.335	82.189	84.594	87.738	91.358
	TPR (%)	79.407	84.966	80.209	81.467	83.920	86.700	89.084
	TNR (%)	77.075	83.623	77.854	81.310	82.761	85.329	89.530
Dataset 2								
Training data=90%	Accuracy (%)	72.107	73.594	74.337	78.872	82.794	84.573	90.191
	TPR (%)	75.526	77.084	77.862	79.660	82.065	84.510	88.738
	TNR (%)	73.880	75.404	76.165	82.100	83.848	86.418	89.900

4- Conclusion

Skin cancer is the most serious health threat experienced by individuals, with melanoma being a certain type that develops in the melanocytes, where the cells of melanocytes are responsible for skin pigmentation. Early detection of melanoma is essential for lowering mortality rates associated with this disease. In this research, a novel FWSAO_CNN-based TL is established for melanoma classification. From the given dataset, a particular skin image is taken to classify melanoma disease. After that, the input image is initially pre-processed by the adaptive Kalman filter, and LinkNet_WSAO is considered for segmenting skin lesions from the pre-processed image. Thereafter, the segmented image is then given to the augmentation process, which is done by employing augmentation techniques. Followed by this process, feature extraction is conducted to obtain the utmost feature vectors. Finally, the melanoma classification is done by employing FWSAO_CNN-based TL. The evaluation metrics of FWSAO_CNN-based TL, like accuracy, TPR, and TNR, achieved 91.565%, 90.090%, and 91.269%. The proposed FWSAO_CNN-based TL model aids in early melanoma detection, supports clinicians with accurate diagnosis, enables remote screening via teledermatology, and assists in medical training with annotated images. However, FWSAO_CNN-based TL is trained for dermoscopic skin images and not assessed on other medical imaging modalities such as histopathology slides, CT scans, and MRI images. Future work involves evaluating the FWSAO_CNN-based TL model on different imaging modalities to further assess the generalizability of the model.

5- Declarations

5-1- Author Contributions

Conceptualization, V.P. and B.K.; methodology, V.P.; software, V.P., and M.S.Y.; validation, V.P., J.M., and S.C.; formal analysis, V.P. and J.M.; investigation, V.P.; resources, V.P.; data curation, V.P.; writing—original draft preparation, V.P. and S.C.; writing—review and editing, V.P.; visualization, V.P., and R.F.; project administration, V.P. All authors have read and agreed to the published version of the manuscript.

5-2- Data Availability Statement

Publicly available datasets were used in this study. These datasets are available in References [18] and [31].

5-3- Funding

This research was funded by The Research Council (TRC), Ministry of Education, Oman, under a Research Grant (RG) project, with support from Modern College of Business and Science, Bowshar, Muscat, Oman.

5-4- Acknowledgments

The authors would like to acknowledge that this work is the outcome of a Research Grant (RG) project funded by The Research Council (TRC Oman), Ministry of Education, Oman, with support from Modern College of Business and Science, Bowshar, Muscat, Oman.

5-5- Institutional Review Board Statement

Not applicable.

5-6- Informed Consent Statement

Not applicable.

5-7- Conflicts of Interest

The authors declare that there is no conflict of interest regarding the publication of this manuscript. In addition, the ethical issues, including plagiarism, informed consent, misconduct, data fabrication and/or falsification, double publication and/or submission, and redundancies have been completely observed by the authors.

6- References

[1] Haque, Intisar Rizwan I., and Jeremiah Neubert. "Deep learning approaches to biomedical image segmentation." Informatics in Medicine Unlocked 18 (2020): 100297. doi:10.1016/j.imu.2020.100297.

[2] Vijaya, P. (2023). Adaptive jellyfish search optimization trained deep learning for breast cancer classification using histopathological images. Mapana Journal of Sciences, 22(3), 59. doi:10.12723/mjs.66.3.

[3] Goyal, M., Oakley, A., Bansal, P., Dancey, D., & Yap, M. H. (2020). Skin Lesion Segmentation in Dermoscopic Images with Ensemble Deep Learning Methods. IEEE Access, 8, 4171–4181. doi:10.1109/ACCESS.2019.2960504.

[4] Breeze, S., Peterson, C., Garioch, J., Nobes, J., & Moncrieff, M. (2025). A Simplified Classification System for In-Transit Melanoma Metastases. Annals of Surgical Oncology, 33(3), 2579–2590. doi:10.1245/s10434-025-18542-9.

[5] Loganathan, G. B., Hamadamen, N. I., Yasin, E. T., Yasin, A. T., Mohammad, A. A., Adil, I. N., ... & Hamadameen, S. F. (2022). Melanoma classification using enhanced fuzzy clustering and DCNN on dermoscopy images. NeuroQuantology, 12, 196-213.

[6] Shorfuzzaman, M. (2022). An explainable stacked ensemble of deep learning models for improved melanoma skin cancer detection. Multimedia Systems, 28(4), 1309–1323. doi:10.1007/s00530-021-00787-5.

[7] Ding, J., Song, J., Li, J., Tang, J., & Guo, F. (2022). Two-Stage Deep Neural Network via Ensemble Learning for Melanoma Classification. Frontiers in Bioengineering and Biotechnology, 9, 1355. doi:10.3389/fbioe.2021.758495.

[8] Aksoy, S. (2025). Multi-Input Melanoma Classification Using Mobilenet-V3-Large Architecture. Journal of Automation, Mobile Robotics and Intelligent Systems, 19(1), 73–84. doi:10.14313/jamris-2025-008.

[9] Kaur, R., Gholamhosseini, H., Sinha, R., & Lindén, M. (2022). Melanoma Classification Using a Novel Deep Convolutional Neural Network with Dermoscopic Images. Sensors, 22(3), 1134. doi:10.3390/s22031134.

[10] Ghosh, S., Singh, A., & Kumar, S. (2025). Multiplier leadership optimization algorithm (MLOA): unconstrained global optimization approach for melanoma classification. Discover Internet of Things, 5(1). doi:10.1007/s43926-025-00168-8.

[11] Aljohani, K., & Turki, T. (2022). Automatic Classification of Melanoma Skin Cancer with Deep Convolutional Neural Networks. AI (Switzerland), 3(2), 512–525. doi:10.3390/ai3020029.

[12] Broekaert, S. M. C., Roy, R., Okamoto, I., Van Den Oord, J., Bauer, J., Garbe, C., Barnhill, R. L., Busam, K. J., Cochran, A. J., Cook, M. G., Elder, D. E., McCarthy, S. W., Mihm, M. C., Schadendorf, D., Scolyer, R. A., Spatz, A., & Bastian, B. C. (2010). Genetic and morphologic features for melanoma classification. Pigment Cell & Melanoma Research, 23(6), 763–770. doi:10.1111/j.1755-148X.2010.00778.x.

[13] Salma, W., & Eltrass, A. S. (2022). Automated deep learning approach for classification of malignant melanoma and benign skin lesions. Multimedia Tools and Applications, 81(22), 32643–32660. doi:10.1007/s11042-022-13081-x.

[14] Bhimavarapu, U., & Battineni, G. (2022). Skin Lesion Analysis for Melanoma Detection Using the Novel Deep Learning Model Fuzzy GC-SCNN. Healthcare (Switzerland), 10(5), 962. doi:10.3390/healthcare10050962.

[15] Qureshi, M. N., Umar, M. S., & Shahab, S. (2022). A Transfer-Learning-Based Novel Convolution Neural Network for Melanoma Classification. Computers, 11(5), 64. doi:10.3390/computers11050064.

[16] Garcia, A., Zhou, J., Pinero-Crespo, G., Beachkofsky, T., & Huang, X. (2025). Clinical Application of Vision Transformers for Melanoma Classification: A Multi-Dataset Evaluation Study. Cancers, 17(21), 3447. doi:10.3390/cancers17213447.

[17] Manikandan, S. P., Narani, S. R., Karthikeyan, S., & Mohankumar, N. (2025). Deep learning for skin melanoma classification using dermoscopic images in different color spaces. International Journal of Electrical and Computer Engineering, 15(1), 319. doi:10.11591/ijece.v15i1.pp319-327.

[18] Kaggle. (2019). Skin Lesion Images for Melanoma Classification dataset. Kaggle, San Francisco, United States. Available online: https://www.kaggle.com/datasets/andrewmvd/isic-2019 (accessed on May 2026).

[19] Rutan, S. C. (1991). Adaptive Kalman Filtering. Analytical Chemistry, 63(22), 1103A-1109A. doi:10.1021/ac00022a002.

[20] Chaurasia, A., & Culurciello, E. (2017). LinkNet: Exploiting encoder representations for efficient semantic segmentation. 2017 IEEE Visual Communications and Image Processing (VCIP), 1–4. doi:10.1109/VCIP.2017.8305148.

[21] Salawudeen, A. T., Mu'azu, M. B., Sha'aban, Y. A., & Adedokun, E. A. (2019). on the Development of a Novel Smell Agent Optimization (Sao) for Optimization Problems. I-Manager's Journal on Pattern Recognition, 5(4), 13. doi:10.26634/jpr.5.4.15677.

[22] Braik, M., Hammouri, A., Atwan, J., Al-Betar, M. A., & Awadallah, M. A. (2022). White Shark Optimizer: A novel bio-inspired meta-heuristic algorithm for global optimization problems. Knowledge-Based Systems, 243, 108457. doi:10.1016/j.knosys.2022.108457.

[23] Khalifa, N. E., Loey, M., & Mirjalili, S. (2022). A comprehensive survey of recent trends in deep learning for digital images augmentation. Artificial Intelligence Review, 55(3), 2351–2377. doi:10.1007/s10462-021-10066-4.

[24] Shorten, C., & Khoshgoftaar, T. M. (2019). A survey on Image Data Augmentation for Deep Learning. Journal of Big Data, 6(1). doi:10.1186/s40537-019-0197-0.

[25] Lessa, V., & Marengoni, M. (2016). Applying Artificial Neural Network for the Classification of Breast Cancer Using Infrared Thermographic Images. Computer Vision and Graphics, ICCVG 2016, Lecture Notes in Computer Science, vol 9972, Springer, Cham, Switzerland. doi:10.1007/978-3-319-46418-3_38.

[26] Mahmood, F. H., & Abbas, W. A. (2016). Texture features analysis using gray level co-occurrence matrix for abnormality detection in chest CT images. Iraqi Journal of Science, 57(1A), 279-288.

[27] Huang, K., Liu, X., Fu, S., Guo, D., & Xu, M. (2021). A Lightweight Privacy-Preserving CNN Feature Extraction Framework for Mobile Sensing. IEEE Transactions on Dependable and Secure Computing, 18(3), 1441–1455. doi:10.1109/TDSC.2019.2913362.

[28] Hany, U., & Akter, L. (2015). Speeded-Up Robust Feature extraction and matching for fingerprint recognition. 2015 International Conference on Electrical Engineering and Information Communication Technology (ICEEICT), 1–7. doi:10.1109/ICEEICT.2015.7307439.

[29] Shi, Z., Hao, H., Zhao, M., Feng, Y., He, L., Wang, Y., & Suzuki, K. (2019). A deep CNN based transfer learning method for false positive reduction. Multimedia Tools and Applications, 78(1), 1017–1033. doi:10.1007/s11042-018-6082-6.

[30] Szegedy, C., Liu, W., Jia, Y., Sermanet, P., Reed, S., Anguelov, D., Erhan, D., Vanhoucke, V., & Rabinovich, A. (2015). Going deeper with convolutions. Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition, 07-12-June-2015, 1–9. doi:10.1109/CVPR.2015.7298594.

[31] Kaggle. (2026). The Melanoma Skin Cancer Dataset of 10000 images. Kaggle, San Francisco, United States. Available online: https://www.kaggle.com/datasets/hasnainjaved/melanoma-skin-cancer-dataset-of-10000-images (accessed on May 2026).

[32] Sharma, M., Monika, Kumar, N., & Kumar, P. (2021). Badminton match outcome prediction model using Naïve Bayes and Feature Weighting technique. Journal of Ambient Intelligence and Humanized Computing, 12(8), 8441–8455. doi:10.1007/s12652-020-02578-8.

[33] Ronneberger, O., Fischer, P., & Brox, T. (2015). U-Net: Convolutional Networks for Biomedical Image Segmentation. Medical Image Computing and Computer-Assisted Intervention – MICCAI 2015. MICCAI 2015, Lecture Notes in Computer Science, vol 9351, Springer, Cham, Switzerland. doi:10.1007/978-3-319-24574-4_28.

[34] Zhang, W., Pang, J., Chen, K., & Loy, C. C. (2021). K-Net: Towards Unified Image Segmentation. Advances in Neural Information Processing Systems, 13, 10326–10338.

[35] Michael Mahesh, K., & Arokia Renjit, J. (2020). DeepJoint segmentation for the classification of severity-levels of glioma tumour using multimodal MRI images. IET Image Processing, 14(11), 2541–2552. doi:10.1049/iet-ipr.2018.6682.

[36] Feng, Q., Chen, L., Philip Chen, C. L., & Guo, L. (2020). Deep Fuzzy Clustering-A Representation Learning Approach. IEEE Transactions on Fuzzy Systems, 28(7), 1420–1433. doi:10.1109/TFUZZ.2020.2966173.