I.J. Image, Graphics and Signal Processing, **2026, 3, 110-132**
Published Online on June 8, 2026 by MECS Press (http://www.mecs-press.org/)
DOI: 10.5815/ijigsp.2026.03.06

DS-MelNet: An Enhanced Dual Stream Semi-Supervised Mechanism for Melanoma Classification

Apurva S. Shinde*
Ramrao Adik Institute of Technology, D Y Patil Deemed to be University, Computer Engineering Department, Navi Mumbai, 400706, India
E-mail: apurva.shinde1812@gmail.com
ORCID iD: https://orcid.org/0000-0002-1479-9509

Sangita S. Chaudhari
Ramrao Adik Institute of Technology, D Y Patil Deemed to be University, Computer Science Engineering Department, Navi Mumbai, 400706, India
Email: sangita.chaudhari@rait.ac.in
ORCID iD: https://orcid.org/0000-0002-6174-7832
*Corresponding author

Received: July 8, 2025; Revised: September 20, 2025; Accepted: December 18, 2025; Published: June 8, 2026

Abstract: Melanoma skin disease is a major concern for skin cancer-related deaths worldwide. Early diagnosis and detection are crucial for improving patient outcomes. However, existing detection methods often result in false alarms, highlighting the need for more accurate and reliable approaches. This paper proposes a Dual-Stream Semi-Supervised Melanoma Network (DS-MelNet) for melanoma detection. The DS-MelNet utilizes a semi-supervised learning framework to incorporate both labeled and unlabeled data, enhancing detection accuracy. The model's performance is evaluated on the SIIM-ISIC Melanoma Classification Challenge dataset. The dataset undergoes hair detection and removal from skin lesion images using three algorithms proposed in literature viz. Modified Dull Razor, Modified E-shaver and Adaptive principle curvature with Modified dull razor fusion. Performance of the proposed models is assessed through commonly used metrics that include Accuracy, Recall, Precision, and F1-score. Comparative analysis of the DS-MelNet is performed against two benchmarks: Simple Convolutional Neural Network (SCNN) and a Fine-tuned VGG-16 model proposed in this paper. The results clearly indicate that the DS-MelNet demonstrates superior performance, achieving an accuracy of 86% and outperforming both SCNN (76%) and VGG-16 (82%) models. This exceptional performance underscores the potential of the DS-MelNet for effective melanoma classification. The study highlights the promise of semi-supervised learning frameworks and sophisticated neural networks in enhancing melanoma diagnostics. The ability of the proposed model to learn from a small set of labeled data makes it highly suitable for real-world applications where annotated datasets are limited.

Index Terms: Convolutional Neural Network, Dermatological Diagnostics, Melanoma Detection, Semi-Supervised Learning, VGG-16 Architecture

1. Introduction

Melanoma skin cancer is a growing concern worldwide, with increasing statistics that indicate a steady increase in incidence rates. In India, melanoma accounts for approximately 1–2% of the total incidence of skin cancer, with a higher prevalence among fair-skinned populations. Globally, the incidence of melanoma has increased by 53% over the past decade, with the highest rates reported in Australia, New Zealand, and North America. The above numbers highlights the need for early detection and effective treatment of melanoma skin cancer [1]. Traditionally, people have relied on visual examination and self-reporting to detect skin cancer. However, this approach has significant limitations, as melanoma can be asymptomatic and visually inconspicuous in its early stages. Furthermore, delayed detection can lead to poorer treatment outcomes [2] and reduced survival rates. In addition, visual examination can be subjective and clinicians may disagree on the diagnosis. Self-reporting is also unreliable, as patients may not notice changes in their skin or may not report them in the early stages. As a result, there is a need for more objective and reliable methods to

This work is open access and licensed under the Creative Commons CC BY 4.0 License.

detect skin cancer from melanoma. Researchers have made significant contributions for the detection of skin cancer, employing cutting-edge techniques such as dermoscopy [3], reflectance confocal microscopy, and optical coherence tomography. These methods enable clinicians to visualize the microscopic structure of the skin, facilitating the identification of suspicious lesions. Dermoscopy, in particular, has become a valuable tool for the diagnosis of melanoma, allowing clinicians to examine the surface and subsurface structures of the skin in detail. Reflectance confocal microscopy, on the other hand, provides high-resolution images of the skin's cellular structure, enabling clinicians to detect subtle changes that may indicate melanoma.The development of non-invasive diagnostic techniques, such as advanced image analysis and computational methods [4], presents promising opportunities for early detection of melanoma, enabling timely interventions and improved patient outcomes. The importance of non-invasive and early detection of skin cancer from melanoma cannot be overstated. Early intervention can significantly improve treatment outcomes, reduce morbidity, and improved patient survival rates. This paper aims to explore the potential of non-invasive and early detection methods for melanoma skin cancer [5], with a focus on developing innovative, image-based approaches for improved patient diagnosis.

Deep learning approaches [6] surpass the traditional machine learning methodology so researchers have focused on building Convolution Neural Network (CNN) [7] models that have higher diagnostic accuracy and perform well in all the scenarios. The incorporation of artificial intelligence and machine learning techniques [8, 9] has driven major advancements in the creation of computer-aid based diagnosis system, which can now analyze medical images with remarkable accuracy and detect melanoma with higher precision. Supervised CNN [8] have been explored and utilized for melanoma detection but Semi-supervised CNN has never been explored to a greater extent in literature. This paper presents Simple CNN architecture and Dual stream CNN model as contribution to the non-invasive methods for melanoma detection which are based on the semi-supervised learning mechanism. One of the early applications [10] presented a semi-supervised model applied on ISIC 2016 dataset. This is the first attempt best to our knowledge in terms of application of Semi-supervised learning for Melanoma skin cancer classification on the ISIC-2020 dataset which has more number of labeled images compared to the earlier attempt for semi-supervised learning.

Key contributions of the paper include:

1. Proposed Dual-Stream Semi-Supervised Melanoma Network (DS-MelNet) with parallel dropout layers for reducing and mitigating overfitting issue.
2. Proposed Simple Convolution Neural Network Architecture (SCNN) and Fine-tuned VGG-16 (FT-VGG) models for detection of melanoma using Sem-supervised approach.
3. Standard metrics are used to compare the proposed architectures to select the optimal model.

The following sections of paper describes related work done, proposed systems with their detailed explanation and result analysis in terms of the standard metrics.

2. Related Work

This section reviews previous research on melanoma skin cancer detection, focusing on non-invasive, image-based techniques developed using machine learning and deep learning approaches.

2.1. Machine Learning Based Methods

A novel computational approach for skin lesion analysis is proposed, leveraging anisotropic diffusion filtering, active contour modelling, and support vector machines to extract asymmetry, border, colour and texture features [11]. An automated method for detecting basal cell carcinoma tumours in skin histopathological images is proposed, utilizing Z-transform features. Experimental results show reasonable agreement with expert pathologists' diagnoses, demonstrating the method's potential for accurate tumour classification [12]. This study presents an automated approach for skin cancer classification, differentiating melanoma from non-melanoma by analysing colour and texture features. Lesion segmentation is performed using K-means clustering, and six different feature extraction techniques are utilized. The system's accuracy was evaluated using four classifiers, yielding promising results [13]. A CAD system has been proposed for classification of types of skin cancer by application of ML techniques from dermoscopic images. The system used active contour modelling to extract Region of Interest (ROI), combines texture and nonlinear features, and applies metaheuristic optimization algorithms for feature selection. The best results were achieved using a combination of selected features and a pattern recognition network classifier, yielding 92.24% accuracy and similar results for sensitivity and specificity [14].

2.2. Deep Learning based Methods

A novel framework for dermoscopic image recognition is proposed [15], combining deep learning and local descriptor encoding. A pre-trained residual neural network extracts features, which are then aggregate d using Fisher vector encoding and classification done with support vector machine. Experiments conducted on the ISBI 2016 dataset highlight the effectiveness and improved performance of the proposed method compared to existing state-of-the-art

techniques. Computer-aided diagnosis can provide objective scores, reducing dermatologist's workload. The authors introduced a system based on categorical generative adversarial networks (GANs) aimed at enabling unsupervised and semi-supervised feature learning for dermoscopic image analysis. The authors' novel approach attained an average precision score of 0.424, on a limited dataset of 140 labelled images from the ISIC 2016 dataset. Furthermore, the proposed method generated high-quality dermoscopic images [10]. Early detection is crucial in the diagnosis of melanoma, a potentially fatal skin cancer that can spread rapidly if left untreated.

A hybrid method for melanoma detection, combining CNN with two classical machine learning classifiers has been proposed. The CNN and classifiers are trained on features describing skin lesion borders, texture, and colour. Majority voting was used to combine the predictions, resulting in improved accuracy [16]. Agarwal, N. suggested feature learning technique works well in a number of tests. They proposed Deep Convolution GAN (DC GAN) for performing automated feature learning and then perform classification using standard algorithms. The proposed strategy bypassed the existing techniques in terms of the feature preparation time and execution with respect to some earlier work. Recall, precision, and accuracy scores were 0.7689, 0.7525, and 0.7384 with only 200 tagged images [17].

Gajera et al. [18] introduced an automated framework for melanoma classification, leveraging a pre-trained convolutional neural network (CNN) architecture. The authors approach combined multiple classifiers with eight prominent CNN models, including AlexNet, VGG-16, and DenseNet121, to analyse four datasets. The framework achieved remarkable results in terms of accuracy and F1-score which are 98.33% and 0.96 respectively, using DenseNet for the feature extraction and a Multilayer Perceptron for classification. Bassel et al. made a stacking classifier approach that achieved 90.9% precision to distinguish melanoma skin cancer from non-melanoma skin cancer [19]. A new methodology has been developed that integrates advanced analytical techniques to enhance the detection of skin cancer. It utilizes the Contourlet Transform and Local Binary Pattern Histogram to extract textural features, while deep learning employs advanced, state-of-the-art techniques. The model achieves 93% precision, with recall scores of 99.7% and 86% for benign and malignant cancers, respectively [20].

The workload of a dermatologist can be reduced by computer-aided diagnosis, which provides an objective score that is not related to clinical experience. Recent progress of deep learning algorithms in the space of general computer vision has motivated the successful utilization of supervised deep learning techniques in computer-assisted melanoma identification. Apart from supervised mechanism for detection semi-supervised methods can also contribute immensely in detection of unknown diseases once trained on the standard datasets, the next section presents the proposed semi-supervised approach for melanoma skin cancer detection.

3. Proposed Methodology

The paper explores the application of semi-supervised learning in melanoma classification, an approach that has not been widely investigated in this domain. Semi-supervised learning steps are shown in Fig. 1. The process mainly entails training the model using a small amount of labelled data and assessing its performance on a larger pool of unlabelled set. Model training is performed after completion of stage 1 and results obtained are evaluated against the results of the supervised learning process.

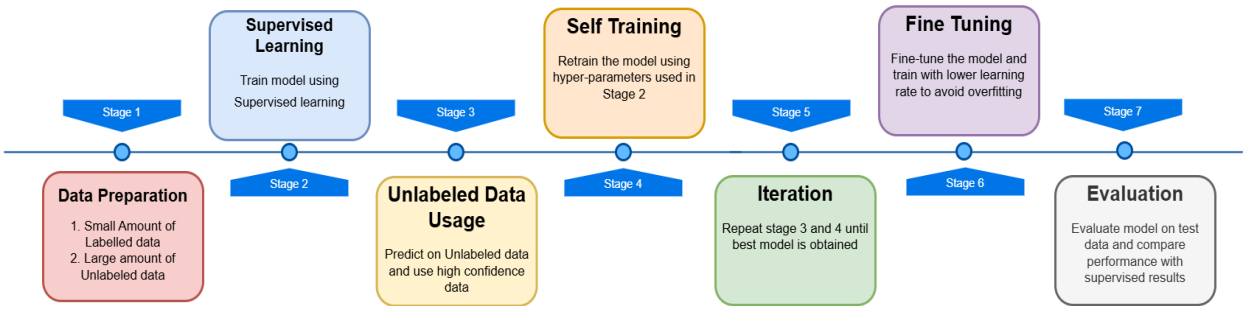


Fig. 1. Semi Supervised Steps.

3.1. Pre-processing Methodology

The ISIC 2020 dataset [21] is employed in this study, comprising 584 images diagnosed as melanoma & 32,542 as non-melanoma. Fig.2 presents few example images from the dataset. Since the dataset is imbalanced, data augmentation techniques are applied to achieve class balance. The augmentation methods are applied on the dataset using the ImageData Generator to increase the variations in the images to train the dataset on different variations and learn all possible patterns that can be obtained from the image. The augmentation performed includes rotation, width shift, horizontal and vertical flip and brightness variation, oversampling is applied to increase the count of melanoma class and balance the dataset. The class distribution of the dataset before balancing and after augmentation and balancing is shown in Fig. 3.

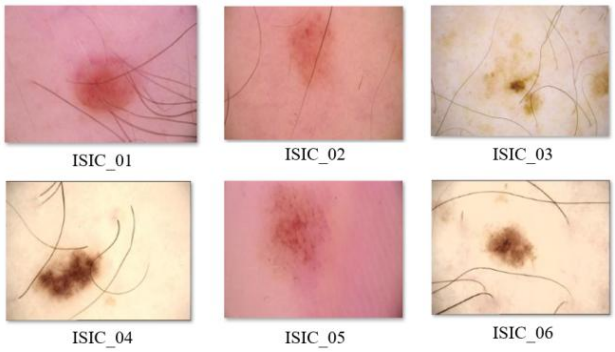


Fig. 2. Sample Dataset Images from ISIC-2020 Dataset.

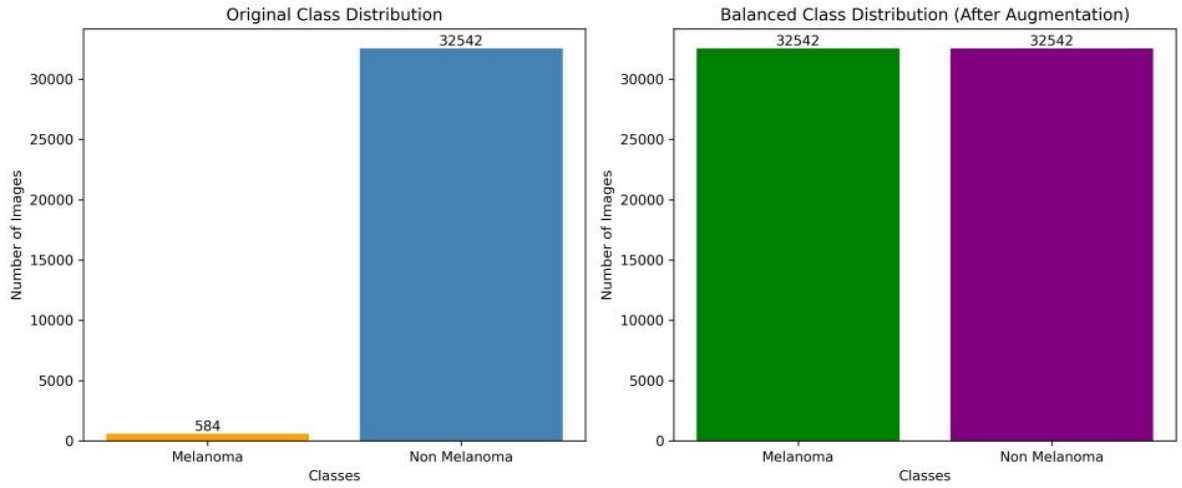


Fig. 3. ISIC-2020 class distribution before and after Augmentation.

Hair artefacts are identified and eliminated using different techniques to support better classification of skin lesions. Modified E-shaver, Modified Dull Razor with YCbCr plane (MDR-YCbCr) and Adaptive Principle Curvature with Modified dull razor with RGB plane (APC-MDR-RGB) these three pre-processing methods have been proposed in our previous work [22] which have been applied to the dataset images. These three methods are discussed with the output sample images for each of the algorithm.

1. Modified E-shaver algorithm [22]:

The existing E-shaver method is improved by introducing filtering and dilation operations that perform enhanced cleaning with detection and removal of the hair regions.

2. MDR-YCbCr algorithm [22]:

The existing dull razor operation is performed on the RGB image which is converted to grayscale with standard filtering and thresholding method, this is improved by converting colour space to YCbCr that detects hair regions using double thresholding.

3. APC-MDR-RGB algorithm [22]:

The Adaptive principle curvature focuses on hair removal on gray scale images by applying hessian matrix, the proposed APC-MDR-RGB performs hair detection for each channel of RGB colour space and then fuses it to perform cleaning of hairs from the image.

Sample output images after application of above methods for hair removal is shown in Fig. 4. All the three methods perform extremely well for hair removal task thus removing noise present in the skin lesion images.

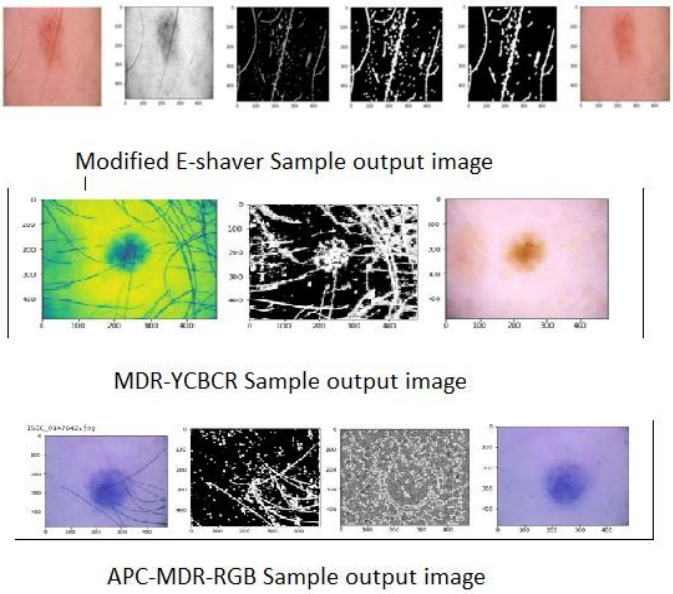


Fig. 4. Sample images after application of hair removal algorithms.

3.2. Dual Stream Semi-Supervised Melanoma Network (DS-MelNet)

Three models have been proposed and explained with its architectural diagram and working with their key features. The dataset was split as 10% of the data was used for training, 70% comprised unlabeled data for evaluation (pseudo label prediction), and 20% was reserved for validation. Although traditional semi-supervised learning leverages unlabeled data during training, this work focuses on evaluating model robustness in a low-label learning scenario, which we refer to as semi-supervised due to the limited availability of labeled samples. The unlabeled data is used for predictions and generating pseudo labels which are further added to the labeled data and used for re-training. The data partitioning strategy allowed the model to effectively learn from the limited labeled data while validating its performance on a larger unseen subset. The same dataset was applied for all the three models discussed in this work.

The proposed DS-MelNet architecture consists of convolution and max pooling layers attached in parallel to extract the low-level features mostly the edges, horizontal, vertical, diagonal lines, basic color information, simple textures, etc. The proposed architecture is a balanced model with similar block on both the streams, features extracted from both the streams are concatenated to improve the detection accuracy of the melanoma skin cancer. Semi-supervised learning (SSL) is particularly suited for melanoma detection due to two key factors: (i) high variability in lesion appearance, which requires large and diverse datasets for robust generalization, and (ii) the high cost and expertise required for annotation by dermatologists. SSL leverages abundant unlabeled data to improve performance while reducing annotation burden, making it clinically and economically advantageous

The architectural diagram of the DS-MelNet is presented in Fig. 5.

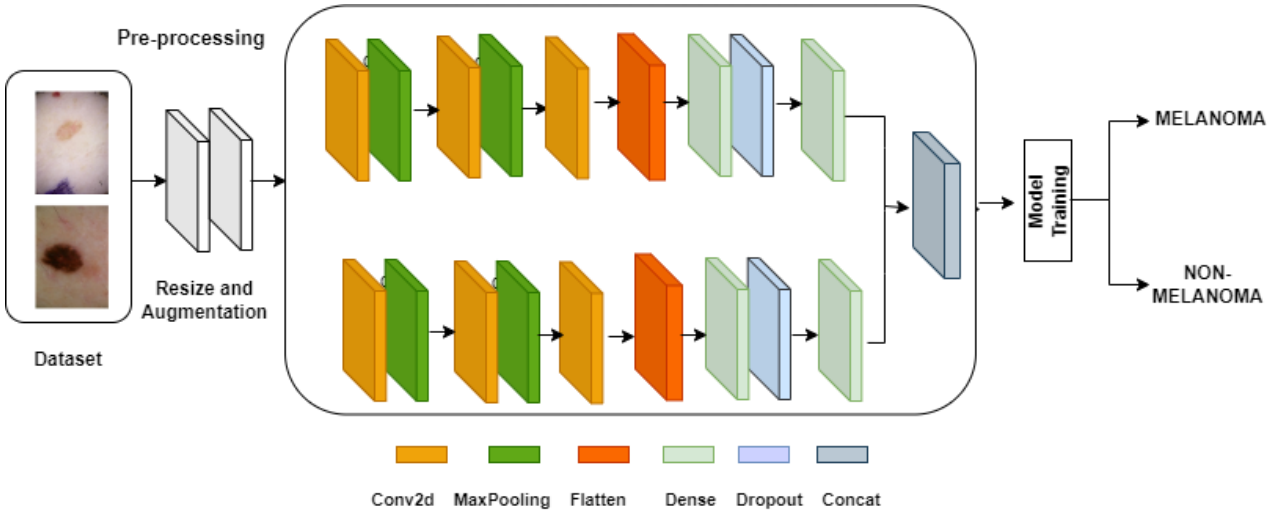


Fig. 5. Proposed DS-MelNet Architecture.

The architecture takes as input a skin lesion images that have been cleaned of hair artifacts using the techniques described in earlier section. The pre-processed images are passed to all the proposed architecture for further feature extraction and classification task and the results are evaluated using the standard performance metrics. The architecture in Fig. 4 is combination of Convolution [23] and Max pooling layers, which consists of 2 sets of Conv and Max Pooling and single standalone Conv layer before flattening. The output from convolution blocks is flattened and passed to the fully connected layer with Dropout connected to it. The features extracted from both the streams is concatenated in the concatenation block to form the feature vector to be passed to the classification block that used SoftMax classifier.

Stepwise mathematical representation of each layer is stated further. Equation (1), (3), (5) defines Convolutional Layer. Equation (2), (4), (6) refers the equation to evaluate Max Polling. Equation (7) defines Flatten Layer. Equation (8) & (10) defines Dense Layer used. Equation (9) defines dropout layer with 20%. Equation (11) defines concatenation layer and equation (12) defines SoftMax classifier.

Step 1: Convolutional Layer
Input: 32 × 32 × 1 (grayscale image)
Filters: 32, Dimension: 3 × 3
Output: 30 × 30 × 32

$$F_{\{a,b,c\}} = ReLU\left(\Sigma_{\{p=0\}}^{\{2\}} \Sigma_{\{q=0\}}^{\{2\}} \Sigma_{\{r=0\}}^{\{K-1\}} W_{\{p,q,r,c\}} \cdot X_{\{a+p,b+q,r\}} + b_c \right) \quad (1)$$

Step 2: Max Pooling Layer
Filter size: 2 × 2
Output: 15 × 15 × 32

$$P_{\{i,j,k\}} = max_{\{0 \le m,n < 2\}} F_{\{2i+m,2j+n,k\}} \quad (2)$$

Step 3: 2nd Convolutional Layer
Filters: 64, Dimension: 3 × 3
Output: 13 × 13 × 64

$$F_{\{a,b,c\}} = ReLU\left(\Sigma_{\{p=0\}}^{\{2\}} \Sigma_{\{q=0\}}^{\{2\}} \Sigma_{\{r=0\}}^{\{K-1\}} W_{\{p,q,r,c\}} \cdot X_{\{a+p,b+q,r\}} + b_c \right) \quad (3)$$

Step 4: Max Pooling Layer
Filter size: 2 × 2
Output: 6 × 6 × 64

$$P_{\{i,j,k\}} = max_{\{0 \le m,n < 2\}} F_{\{2i+m,2j+n,k\}} \quad (4)$$

Step 5: 3rd Convolutional Layer
Filters: 128, Dimension: 3 × 3
Output: 4 × 4 × 128

$$F_{\{a,b,c\}} = ReLU\left(\Sigma_{\{p=0\}}^{\{2\}} \Sigma_{\{q=0\}}^{\{2\}} \Sigma_{\{r=0\}}^{\{K-1\}} W_{\{p,q,r,c\}} \cdot X_{\{a+p,b+q,r\}} + b_c \right) \quad (5)$$

Step 6: Max Pooling Layer
Filter size: 2 × 2
Output: 2 × 2 × 128

$$P_{\{i,j,k\}} = max_{\{0 \le m,n < 2\}} F_{\{2i+m,2j+n,k\}} \quad (6)$$

Step 7: Flatten Layer
Input: 2 × 2 × 128
Output: 512-dimensional vector

$$v \in \mathbb{R}^{\{512\}} \quad (7)$$

Step 8: Dense Layer (1st)

$$h_1 = \sigma(W_1 \cdot v + b_1) \quad (8)$$

Step 9: Dropout Layer

Dropout rate: 20%

$$h1' = h1 \odot m, \qquad m_i \sim Bernoulli(0.8) \tag{9}$$

Step 10: Dense Layer (2nd)

$$h_2 = \sigma(W_2 \cdot h1' + b_2) \tag{10}$$

Step 11: Concatenation Layer

$$h_{concat} = concat\left(h_2^{(1)}, h_2^{(2)}\right) \tag{11}$$

Step 12: Final Classification Layer

$$\hat{y} = softmax(W_3 \cdot h_{concat} + b_3) \tag{12}$$

where SoftMax $(z_i) = e^{zi} / \sum_j e^{zi}$
where

- $F_{i,j,k}$ states Output Feature Map for application of convolution layer.
- $P_{i,j,k}$ represents pooling layer's output.
- h_i represents hidden layer's (dense layer) output.
- h_{concat} represents the feature vector obtained after concatenation of features from both streams.
- h_1' represents the output after application of dropout
- $\hat{y}$ represents the final output probability or predicted class.

The architecture consists of parallel dropout layers added to apply regularization and minimize the effect of over fitting. The dual stream model is an equally balanced architecture as it contains the same number of layers on both sides of the architecture with almost a similar set of parameters. Steps 1 to 10 of the above explanation will be the same for both streams and will generate a feature map as the output. The layers in the dual-stream architecture are chosen after evaluating various combinations of layers, each with different parameter settings, to determine the optimal configuration. The model underwent training for 30 iterations using a learning rate of 0.0001 and a binary cross-entropy loss function. During the training process, accuracy served as the primary evaluation metric. SoftMax classifier was used in the classification layer to perform the binary classification of images. Architecture is optimized using the Adam algorithm, which is the most used and efficient optimizer to converse the outputs faster and efficiently. The steps for the operation of the Adam optimizer are as follows.

1. Initialize the model parameters (weights and biases) at random.
2. Set the first-moment vector (m) and the second-moment vector (v) to 0 during initialization.
3. Initialize hyperparameters.

$\theta \leftarrow$ random initialization
$m \leftarrow 0$
$v \leftarrow 0$
$\alpha \leftarrow$ learning rate
$\beta^1 \leftarrow$ first moment decay rate
$\beta^2 \leftarrow$ second moment decay rate
$\varepsilon \leftarrow$ numerical stability term

4. Compute the gradients of the loss function with respect to the model parameters. The loss gradient can be calculated using equation 13.

$$g_t \leftarrow \nabla_\theta L(\theta) \tag{13}$$

5. Compute the first moment (m) at time step t using equation 14.

$$m_t \leftarrow \beta^1 m_t^{-1} + (1 - \beta^1) g_t \tag{14}$$

6. Compute the second order moment (v) at time step t using equation 15.

$$v_t \leftarrow \beta^2 v_t^{-1} + (1 - \beta^2) g_t^2 \tag{15}$$

7. Compute bias-corrected first order moment $\hat{m}_t$ and second moment $\hat{v}_t$ using equation 16 & 17 respectively

$$\widehat{m}_t \leftarrow \frac{m_t}{\left(1 - {\beta_1}^t\right)} \tag{16}$$

$$\hat{v}_t \leftarrow \frac{v_t}{(1 - \beta_2^t)} \tag{17}$$

8. Update the model parameters using the Adam update rule which is evaluated using equation 18.

$$\theta_t \leftarrow {\theta_t}^{-1} - \alpha \cdot \widehat{m}_t / (\sqrt{\hat{v}_t} + \varepsilon) \tag{18}$$

Stream 1 and Stream 2 comprises of multiple convolutional layers designed to capture low to mid-level features such as edges, textures, patterns, and shapes—attributes that are prominently visible in skin lesion images. The configuration of layers and the selection of inputs for each layer were finalized after extensive experimentation and multiple iterations. Based on this optimized setup, appropriate dropout and classification layers were subsequently integrated. The resulting architecture comprises approximately 9 million parameters, which is significantly lower compared to the parameter count of pre-trained and fine-tuned architectures such as VGG-16.

The key features of the proposed architecture that contribute to improved performance are as follows:

- Semantically aligned features due to similar set of layers on both the streams enabling better abstraction level.
- Avoids biasing of one stream over other and contributes to balanced feature concatenation.
- Consistent gradient flow and learning thus reducing the issue of vanishing gradient and overfitting.

The algorithm steps for semi-supervised implementation are provided in Algorithm 1.

During each training iteration, the model is optimized using both labeled and unlabeled samples. The labeled batch contributes to the supervised loss, while the unlabeled batch provides an auxiliary consistency-based loss using pseudo-labels. The total loss is computed as a weighted sum of supervised and unsupervised losses, allowing the model to leverage unlabeled data while maintaining stable supervised learning.

Algorithm 1. Semi-Supervised Training with Labelled and Unlabelled Batches

Input: Labelled dataset D_L , $Unlabeled\ dataset\ D_U$, $Model\ f_\theta$
Output: Trained model parameters θ
Step 1: Initialize model parameters (θ).
Step 2: Initialize the optimizer (Adam).
Step 3: For epoch = 1 to (E) **do**
  Step 4: For each iteration **do**
  Sample labelled batch (($x_{1,}$ y_1) ~ $D_{L;}$
  Sample unlabelled batch (Xu ~ D_U)
  Compute predictions: $\widehat{y_l} = f_\theta(x_l)$;
  Compute supervised loss: $L_{sup} = CE(\widehat{y_l}, y_l)$
  Generate pseudo-labels for unlabelled data: $\widehat{y_u} = f_\theta(x_u)$
  Compute unsupervised loss: $L_{unsup} = CE(\widehat{y_u}, stopgrad(\hat{y}_u))$
  Compute total loss: $L = L_{sup} + \lambda\, L_{unsup}$
  Update model parameters (θ) using backpropagation.
  End For

3.3. Simple Convolution Neural Network Architecture (SCNN)

This section introduces the SCNN model, which serves as a critical benchmark for evaluating the performance of CNN-based approaches in the detection of melanoma skin cancer. The model leverages the capabilities of Convolutional Neural Networks (CNNs) to effectively extract deep and meaningful features from the input image dataset. The overall SCNN architecture is depicted in Fig. 6.

The architecture comprises two sets of convolutional layers followed by max-pooling layers at the initial stages, designed to capture essential features such as edges, lines, and textures. While the convolutional layers are responsible for extracting both high-level and abstract features from the input images, the max-pooling layers perform down-sampling to reduce spatial dimensions, thereby retaining only the most significant information for further processing. Further the extracted features are flattened to 1D and passed to the Dense layer for further processing. These then go through dropout layer where few of the neurons are excluded and only the relevant and key features passed to the SoftMax for classification task.

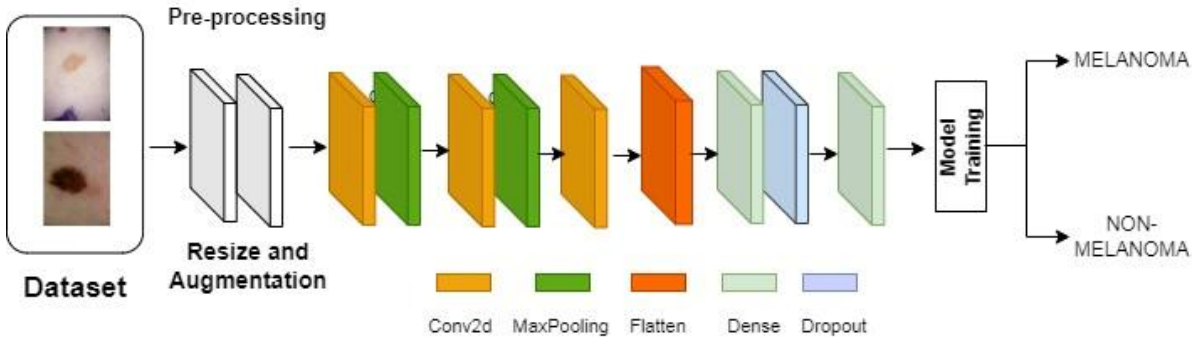


Fig. 6. Simple Convolution Neural Network Architecture (SCNN).

The complete layer-wise functioning of the SCNN model with the hyper parameters used in each of the layers is discussed as follows. Equation 19, 21, 23 defines Convolutional Layer. Equation 20, 22, 24 refers the equation to evaluate Max Polling. Equation 25 defines Flatten Layer. Equation 26 & 28 defines Dense Layer used. Equation 27 defines dropout layer with 20%. Equation 29 defines SoftMax classifier.

Step 1: Convolutional Layer
Input: 32 × 32 × 1 (grayscale image)
Filters: 32, Dimension: 3 × 3
Output: 30 × 30 × 32

$$F_{\{a,b,c\}} = ReLU\left(\Sigma_{\{p=0\}}^{\{2\}} \Sigma_{\{q=0\}}^{\{2\}} \Sigma_{\{r=0\}}^{\{K-1\}} W_{\{p,q,r,c\}} \cdot X_{\{a+p,b+q,r\}} + b_c \right) \quad (19)$$

Step 2: Max Pooling Layer
Filter size: 2 × 2
Output: 15 × 15 × 32

$$P_{\{i,j,k\}} = max_{\{0 \le m,n < 2\}} F_{\{2i+m,2j+n,k\}} \quad (20)$$

Step 3: 2nd Convolutional Layer
Filters: 64, Dimension: 3 × 3
Output: 13 × 13 × 64

$$F_{\{a,b,c\}} = ReLU\left(\Sigma_{\{p=0\}}^{\{2\}} \Sigma_{\{q=0\}}^{\{2\}} \Sigma_{\{r=0\}}^{\{K-1\}} W_{\{p,q,r,c\}} \cdot X_{\{a+p,b+q,r\}} + b_c \right) \quad (21)$$

Step 4: Max Pooling Layer
Filter size: 2 × 2
Output: 6 × 6 × 64

$$P_{\{i,j,k\}} = max_{\{0 \le m,n < 2\}} F_{\{2i+m,2j+n,k\}} \quad (22)$$

Step 5: 3rd Convolutional Layer
Filters: 128, Dimension: 3 × 3
Output: 4 × 4 × 128

$$F_{\{a,b,c\}} = ReLU\left(\Sigma_{\{p=0\}}^{\{2\}} \Sigma_{\{q=0\}}^{\{2\}} \Sigma_{\{r=0\}}^{\{K-1\}} W_{\{p,q,r,c\}} \cdot X_{\{a+p,b+q,r\}} + b_c \right) \quad (23)$$

Step 6: Max Pooling Layer
Filter size: 2 × 2
Output: 2 × 2 × 128

$$P_{\{i,j,k\}} = max_{\{0 \le m,n < 2\}} F_{\{2i+m,2j+n,k\}} \quad (24)$$

Step 7: Flatten Layer
Input: 2 × 2 × 128
Output: 512-dimensional vector

$$v \in \mathbb{R}^{\{512\}} \tag{25}$$

Step 8: Dense Layer (1st)

$$h_1 = \sigma(W_1 \cdot v + b_1) \tag{26}$$

Step 9: Dropout Layer
Dropout rate: 20%

$$h1' = h1 \odot m, \qquad m_i \sim Bernoulli(0.8) \tag{27}$$

Step 10: Dense Layer (2nd)

$$h_2 = \sigma(W_2 \cdot h1' + b_2) \tag{28}$$

Step 11: Concatenation Layer

$$\hat{y} = softmax(W_3 \cdot h_{concat} + b_3) \tag{29}$$

where $\text{SoftMax}(z_i) = e^{zi} / \sum_j e^{zi}$

The primary objective of the SCNN architecture is to deliver a lightweight deep convolutional neural network capable of effectively performing binary classification between melanoma and non-melanoma images at an early diagnostic stage. Overall, SCNN serves as a lightweight, simple, and effective baseline architecture upon which more complex CNN models can be developed for similar medical imaging applications. Layer-wise parameter comparison is done for the DS-MelNet and SCNN shown in the Table 1.

Table 1. Comparison of SCNN and DS-MelNet Architecture.

SCNN			**DS-MelNet**		
Layer	**Output Dimension**	**Parameters**	**Layer**	**Output Dimension**	**Parameters**
Input Layer	(32, 32, 3)	0	Input Layer	(None, 32, 32, 3)	0
Conv2D (32 filters, 3×3)	(30, 30, 32)	896	Conv2D (32 filters, 3×3)	(30, 30, 32)	896
MaxPooling2D (2×2)	(15, 15, 32)	0	Conv2D (32 filters, 3×3)	(30, 30, 32)	896
Conv2D (64 filters, 3×3)	(13, 13, 64)	18,496	MaxPooling2D (2×2)	(15, 15, 32)	0
MaxPooling2D (2×2)	(6, 6, 64)	0	MaxPooling2D (2×2)	(15, 15, 32)	0
Conv2D (128 filters, 3×3)	(4, 4, 128)	73,856	Flatten	(7200)	0
MaxPooling2D (2×2)	(2, 2, 128)	0	Flatten	(7200)	0
Flatten	(512)	0	Dense	(64)	460,864
Dense	(128)	65,664	Dense	(64)	460,864
Dense (SoftMax, 2 classes)	(2)	258	Dropout (20%)	(64)	0
			Dropout (20%)	(64)	0
			Dense	(2)	130
			Dense	(2)	130
			Concatenate	(4)	0
			Dense (SoftMax, 2 classes)	(2)	10
Total Parameters	**159,170 (621.76 KB)**		**Total Parameters**	**923,790 (3.52 MB)**	
Trainable Parameters	**159,170 (621.76 KB)**		**Trainable Parameters**	**923,790 (3.52 MB)**	
Non-trainable Parameters	**0 (0.00 B)**		**Non-trainable Parameters**	**0 (0.00 B)**	

3.4. Finetuned VGG-16 (FT-VGG) model

The FT-VGG architecture, illustrated in Fig. 7, has total 16 layers, with 13 convolutional and max pooling blocks and 3 Fully Connected layers.

This research explores the application of the VGG-16 convolutional neural network in the context of enhancing the melanoma detection accuracy via semi-supervised learning frameworks. The information gained from VGG-16 is applied to enhance the detection of melanoma cancer, using a semi-supervised learning method. The VGG-16 was originally trained during the ImageNet challenge; the dimension can be adjusted for specific problems as needed during implementation. The architecture begins with two convolutional layers, each with 64 channels, a filter size of 3x3, and padding value that preserve the original size of input image. This is followed with Max-Pooling Layer with 2x2 window is applied to down sample the spatial dimensions while preserving crucial information. Then, the model passes the information through two more convolutional layers. Feature extraction block utilizes convolutional layers with 256 and 512 channels, a channel size of 3x3, and the same padding. This block is followed by a max-pooling with pool size of 2x2 and stride set to 2 for spatial down sampling. This organized approach helps extraction of distinct features.

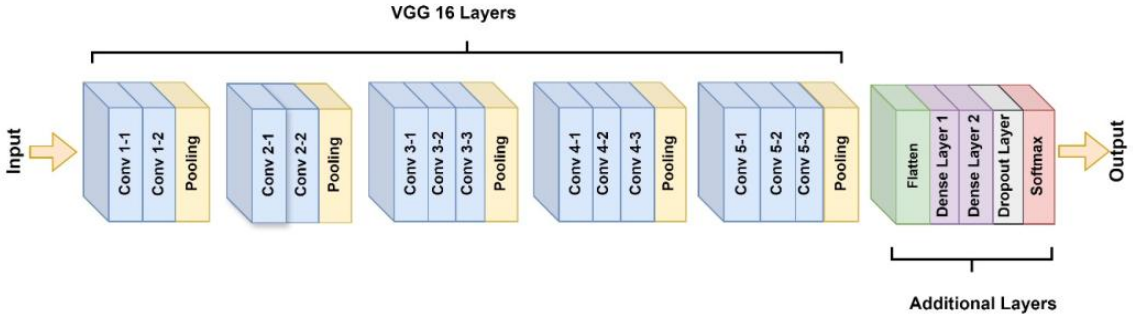


Fig. 7. Fine Tuned VGG-16 Architecture.

The mathematical representation for the fine-tuned VGG-16 architecture is discussed as follows.
Step 1: Input

Let the input image be denoted by equation 30:

$$X \in \mathbb{R}^{\{32\times32\times3\}} \tag{30}$$

Step 2: VGG-16 Feature Extractor

After down sampling, the final output feature map is shown in equation 31:

$$F = VGG16_{base}(X), \quad F \in \mathbb{R}^{\{2\times2\times512\}} \tag{31}$$

Step 3: Flatten Layer

Flatten Layer is evaluated using equation 32.

$$f = Flatten(F), \quad f \in \mathbb{R}^{\{2048\}} \tag{32}$$

Step 4: Dense Layer 1

Dense Layer 1&2 calculation are done using equation 33 & 34.

$$h1 = ReLU(W_1 \cdot f + b_1), \quad W_1 \in \mathbb{R}^{\{512\times2048\}}, \quad b_1 \in \mathbb{R}^{\{512\}} \tag{33}$$

Step 5: Dense Layer 2

$$h2 = ReLU(W_2 \cdot h_1 + b_2), \quad W_2 \in \mathbb{R}^{\{128\times512\}}, \quad b_2 \in \mathbb{R}^{\{128\}} \tag{34}$$

Step 6: Dropout Layer

Dropout layer is calculated using equation 35.

$$h_2^{drop} = Dropout(h_2, rate = 0.2) \tag{35}$$

Step 7: Output Layer (SoftMax)

Classification is done using SoftMax classifier which is obtained using equation 36 & 37.

$$\hat{y} = Softmax(Wo \cdot h_2^{drop} + bo),\ Wo \in \mathbb{R}^{\{2\times128\}},\ bo \in \mathbb{R}^{\{2\}} \tag{36}$$

$$\hat{y} = [P(Melanoma), P(Non-Melanoma)], \quad \Sigma\, y_i = 1 \tag{37}$$

- F represents the output after application of the pre-trained VGG-16.
- f represents the output from the Flatten layer.
- h_i represents the output from the dense layer.
- h_2^{drop} represents the output after application of the Dropout layer.
- $\hat{y}$ represents the final output probability or the predicted class.

The FT-VGG freezes the top 3 layers and extracts the unique features from the remaining layers. The fine-tuned models output layer is replaced by some additional layers as can be seen in the Fig. 7. The freezing of layers and updating the output layers is the key contribution to the fine-tuning of the pre-trained model.

The layer wise comparison for Standard VGG-16 and Finetuned VGG-16 is given in Table 2.

Table 2. Comparison of Standard VGG-16 and Fine-tuned VGG-16 Architectures (Top 3 Layers Frozen).

Layer	Standard VGG-16 Output	Standard Params	Fine-tuned VGG-16 Output	Fine-tuned Params
Input Layer	(32, 32, 3)	0	(32, 32, 3)	0
Conv2D (64, 3×3)	(32, 32, 64)	1,792	(32, 32, 64)	1,792 (Frozen)
Conv2D (64, 3×3)	(32, 32, 64)	36,928	(32, 32, 64)	36,928 (Frozen)
MaxPooling2D (2×2)	(16, 16, 64)	0	(16, 16, 64)	0 (Frozen)
Conv2D (128, 3×3)	(16, 16, 128)	73,856	(16, 16, 128)	73,856
Conv2D (128, 3×3)	(16, 16, 128)	147,584	(16, 16, 128)	147,584
MaxPooling2D (2×2)	(8, 8, 128)	0	(8, 8, 128)	0
Conv2D (256, 3×3)	(8, 8, 256)	295,168	(8, 8, 256)	295,168
Conv2D (256, 3×3)	(8, 8, 256)	590,080	(8, 8, 256)	590,080
Conv2D (256, 3×3)	(8, 8, 256)	590,080	(8, 8, 256)	590,080
MaxPooling2D (2×2)	(4, 4, 256)	0	(4, 4, 256)	0
Conv2D (512, 3×3)	(4, 4, 512)	1,180,160	(4, 4, 512)	1,180,160
Conv2D (512, 3×3)	(4, 4, 512)	2,359,808	(4, 4, 512)	2,359,808
Conv2D (512, 3×3)	(4, 4, 512)	2,359,808	(4, 4, 512)	2,359,808
MaxPooling2D (2×2)	(2, 2, 512)	0	(2, 2, 512)	0
Flatten	(2048)	0	(2048)	0
Dense	(4096)	8,396,800	(1072)	2,198,016
Dropout (20%)	–	0	–	0
Dense	(4096)	16,781,312	(4096)	4,398,592
Dense (SoftMax, 2 classes)	(2)	8,194	(2)	8,194
Total Parameters	—	**29,418,370**	—	**14,898,370**
Trainable Parameters	—	**29,418,370**	—	**14,598,370**
Non-trainable Parameters	—	**0**	—	**300,000 (Frozen)**

Table 3. Comparison of Proposed Architecture.

Architecture	Depth	Size	Optimizer	No of. Parameters
Proposed DS-MelNet	9	3.52 MB	Adam	923,790
Proposed SCNN	9	621.76 KB	Adam	159,170
Finetuned VGG-16	19	56.14 MB	Adam	14,898,370

Comparison of the architectures discussed in this paper is shown in Table 3. The proposed dual-stream model has fewer layers than FT-VGG but operates in a parallel structure. The proposed model is small in size with fewer parameters that generalize well and have higher results compared to the VGG-16 and SCNN model.

4. Experimental Results

The implementation was performed using Kaggle platform which provided 2 GPUs suitable for implementation of the proposed models. The DS-MelNet and SCNN architecture took less time for the training process compared to fine-tuned VGG-16 due to lesser number of parameters and simple architecture and flow.

Results are shown in terms of different performance parameters and obtained for images after applying three different pre-processing methods mentioned in the previous section. The Modified E-shaver, Modified Dull Razor and Adaptive Principal Curvature and Modified Dull Razor fusion methods are applied on the ISIC-2020[21] dataset for hair removal. These pre-processed images as passed through the proposed architectures for further evaluation.

4.1. Cross Validation Results

Five-fold cross validation was employed to evaluate model's performance and ensure robustness of the proposed model. The Modified Dull Razor-YCbCr pre-processed output is considered for cross-validation results. The Accuracy results are obtained for the DS-MelNet and SCNN Baseline model and comparison is provided in the Table 4. The results clearly indicate DS-MelNet outperforming the baseline in all the folds with performance improvement ranging from 12 to 14.7 pp. The mean result difference is also quite high with a value of 13.10 pp for the baseline model.

Table 4. Comparison of semi-supervised vs baseline accuracy per fold, with mean, standard deviation, 95% confidence intervals, and paired t-test result. Δ denotes percentage-point (pp) difference.

Fold	DS-MelNet Accuracy (%)	SCNN-Baseline Accuracy (%)	Δ (pp)
1	87.2	75.2	+12.0
2	85.8	73.5	+12.3
3	88.1	74.3	+14.2
4	84.2	72.5	+12.3
5	87.5	72.8	+14.7
Mean	**86.56**	**73.67**	**+13.10**
Standard Deviation (pp)	**1.57**	**1.02**	**1.14**
95% Confidence Interval	**[84.61, 88.51]**	**[72.40, 74.94]**	**[11.68, 14.52]**

Paired t-test	**t(4) = 25.7, p < 0.001 (significant)**		

These results demonstrate that the semi-supervised approach consistently improves classification accuracy over the baseline, with relatively low variance, suggesting robustness across different data splits.

Statistical Analysis:

Across-fold 95% confidence intervals (t-based, n=5, df=4, t^*=2.776) were computed as

$$CI_{95} = \bar{x} \pm t \cdot (s / \sqrt{n})^* \tag{38}$$

Semi-supervised accuracy: [84.61%, 88.51%]; Baseline accuracy: [72.40%, 74.94%];

Improvement Δ: [11.68, 14.52] percentage points. A paired t-test on fold-wise differences yielded t(4)=25.7, p<0.001, indicating a highly significant improvement of the semi-supervised model over the baseline.

These intervals reflect variability across folds; folds are not fully independent, so CIs should be interpreted as cross-validation variability rather than population-level uncertainty.

4.2. Performance Metrics

The proposed system is analysed using the standard performance metrics discussed below, which are obtained from the parameters of the confusion matrix.

1. Accuracy [23]:

The accuracy of processed images is obtained as per the equation 39.

$$Accuracy = C_T / (C_T + I_T) \tag{39}$$

where
C_T = Correct prediction (TP + TN)
I_T = Incorrect prediction (FP + FN)

2. Recall (Re) [23]:

Recall calculates the ratio of correctly recognized actual positive cases. Recall is obtained as per equation 40.

$$Re = C_P / (C_P + M_N) \tag{40}$$

where
C_P = Correctly predicted positive cases (True Positives)
M_N = Missed positive cases (False Negatives)

3. Precision (Pre) [23]:

Precision denotes the ratio of correctly recognized positive cases among total predicted positive cases. The precision is obtained as per equation 41.

$$Pre = C_P / (C_P + M_P) \tag{41}$$

where
CP = Correctly predicted positive cases (True Positives)
MP = Misclassified positive cases (False Positives)

4. F1-score [24]:

F1-score represents the harmonic mean between recall and precision it provides a unique metric that has equal balance between the two measures of model's performance. F1-score is calculated using the equation 42.

$$F1 = \frac{(2 \times Pre \times Re)}{(Pre + Re)} \tag{42}$$

Pseudo label accuracy with confidence interval of 95% is shown in Table 5, the accuracy increases with increase in epochs and presents upward graph in terms of model performance.

The consistency loss steadily decreases as training progresses, indicating improved agreement between model predictions on perturbed unlabelled samples. This trend confirms that the model learns stable and invariant representations, which is essential for effective semi-supervised learning. The same is shown in the Fig. 8.

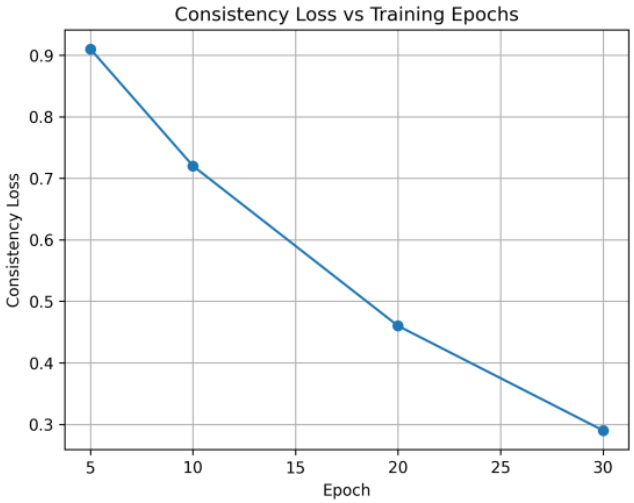


Fig. 8. Consistency Loss Vs Epochs for Pseudo labels.

Table 5. Pseudo-label accuracy across training epochs with 95% confidence interval (CI).

Epoch	Pseudo-Label Accuracy (%) pm 95% CI
5	68.4 ±1.9
10	74.9 ±1.6
20	82.6 ±1.3
30	88.1 ±1.1

To ensure reproducibility and align with standard semi-supervised learning practices, FixMatch [25] was implemented, which combines pseudo-labelling and consistency regularization. Pseudo-labels were generated from weakly augmented images when the model's confidence exceeded 0.95, and consistency loss was computed as cross-entropy between pseudo-labels and predictions on strongly augmented views. This approach follows established SSL benchmarks and provides a reproducible framework for evaluating pseudo-label accuracy and consistency loss trends. The detailed graph for consistency loss and pseudo label accuracy table is presented above.

Ablation Study:

The ablation study performed for the selection of the proposed DS-MelNet architecture and its selection of parameters is discussed in the Table 6.

Although the dual-stream architecture incurs a moderate increase in computational complexity, the proposed DS-MelNet remains significantly more lightweight than conventional deep models such as VGG-16. Specifically, DS-MelNet achieves comparable accuracy with approximately 16× fewer parameters, 13× lower FLOPs, and 4× faster inference time, making it suitable for real-time and resource-constrained applications.

Table 6. Ablation study evaluating single-stream and dual-stream architectures in terms of performance and computational efficiency.

Model Variant	Single	Dual	Depth	Params	FLOPs (G)	Time (ms)	Accuracy (%)
SCNN (Single Stream Baseline)	Yes	No	9	122,050	0.42	3.1	76.23
DS-MelNet (Dual Stream, No Fusion Ablation)	No	Yes	9	857,338	1.11	7.1	81.79
DS-MelNet (Dual Stream + Feature Fusion)	No	Yes	9	923,790	1.12	7.3	**86.57**
Fine-Tuned VGG-16	Yes	No	19	14,715,714	15.3	28.9	82.78

Ablation Study for Supervised 10% data:

To evaluate the contribution of unlabelled data, the supervised baseline was trained on 10% labelled samples with the semi-supervised approach using the same labelled subset plus unlabelled data.

As shown in Fig. 9, the supervised baseline achieved 74.89% accuracy, whereas the semi-supervised model reached 86.57%, yielding an improvement of +11.68 percentage points.

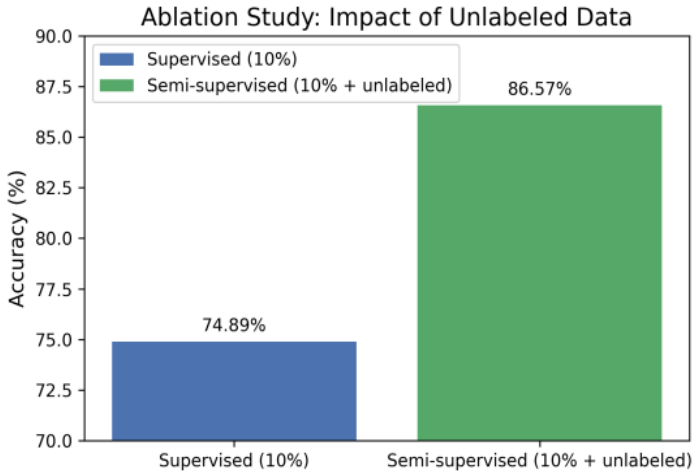


Fig. 9. Comparison of Supervised and Semi-Supervised on 10% labeled data.

4.3. Proposed DS-MelNet Results

The results for proposed DS-MelNet are shown in Fig. 10. The metrics wise comparison of the three methods is shown clearly to indicate the dominance of the best method of pre-processing and final classification. In terms of the accuracy metrics, the DS-MelNet achieved the highest accuracy of 86.57% on the Modified Dull Razor processed images. For the other two methods the DS-MelNet achieved quite good accuracy of 84% and 83% respectively. Similarly, for Precision, Recall and F1-score the DS-MelNet achieved highest results for the Modified Dull Razor YCbCr method, it achieved precision score of 92.23%, Recall of 84.22% and F1-score of 88.45%. The average accuracy achieved by the DS-MelNet model on the all the three methods is approximately 85% which is quite good for the proposed model with fewer parameters and lesser computational cost. Similarly, it achieved an average Precision value of 88%, Recall of 82% and F1-score of 86%.

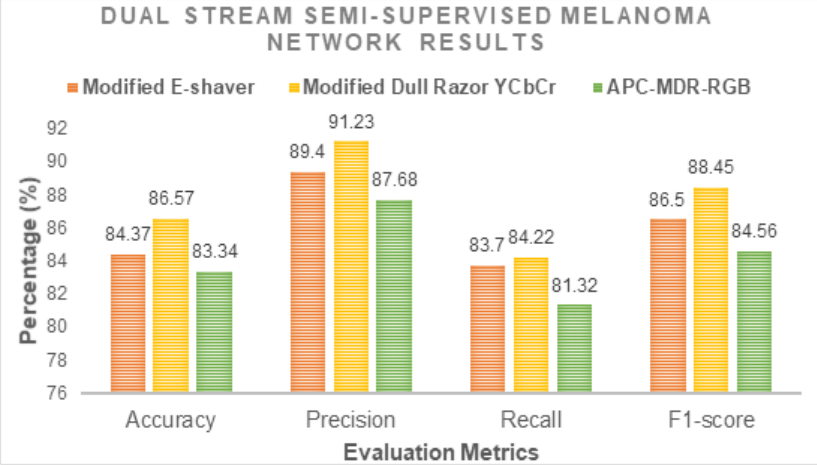


Fig. 10. DS-MelNet results for evaluation metrics on ISIC-2020 dataset

PH2 Dataset [26] Results for DS-MelNet:

The dataset consists of dermoscopic images which are obtained at the Dermatology Service of Hospital Pedro Hispano [26] under standard condition's using a Mole Analyzer with an efficient 20x magnification. The images are 8-bit RGB images having a standard resolution of 768x560 pixels. It consists of 200 images including melanoma and non-melanoma, the dataset is imbalanced with only 40 melanoma images and remaining are non-melanoma.

The rotation range was set to 30, width and height shift range value was set to 0.1 and similarly for shear and zoom the value was set to 0.2. The TensorFlow utility ImageDataGenerator was used to perform the augmentation, to balance the dataset 200 images per class were generated along with the original images. The sample images from PH2 dataset are shown in Fig. 11.

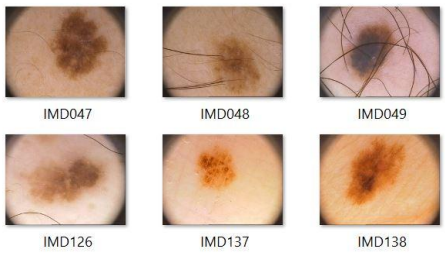


Fig. 11. Sample Images from PH2 Dataset.

The PH2 dataset is split in similar proportion (10:70:20) as ISIC-2020 so that it can be used for semi-supervised learning. The results for performance metrics obtained on the PH2 dataset are shown in the Fig. 12. The proposed model achieves satisfactory results on the PH2 dataset with an accuracy of 81.27% on the Modified Dull Razor images for YCbCr plane.

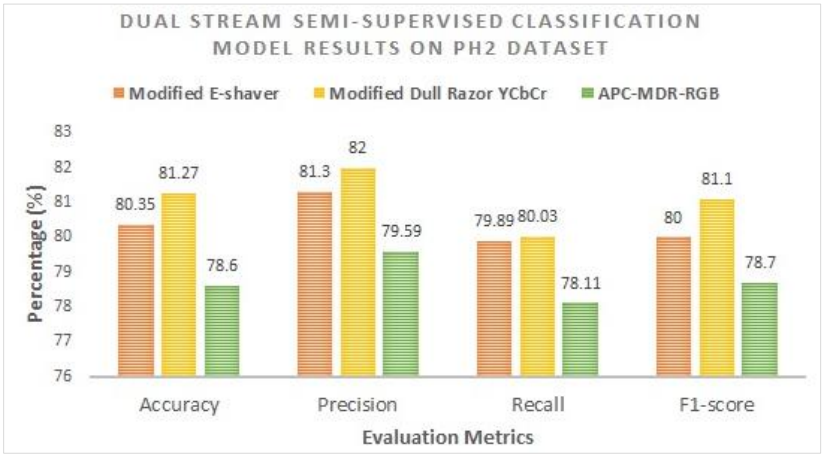


Fig. 12. DS-MelNet Results on the PH2 Dataset

DS-MelNet Visual Explanation using Grad-CAM

DS-MelNet being the major contribution of the paper here the Grad-CAM analysis is presented for the architecture. Grad-CAM visualization highlights the regions contributing most to the model's decision. As shown in Fig. 13, the model focuses on the lesion area, indicating that predictions are based on clinically relevant features rather than background noise. This enhances interpretability and trust in the model. The last two samples where hair artefacts and background noise are present have issues with classification and do not present relevant features.

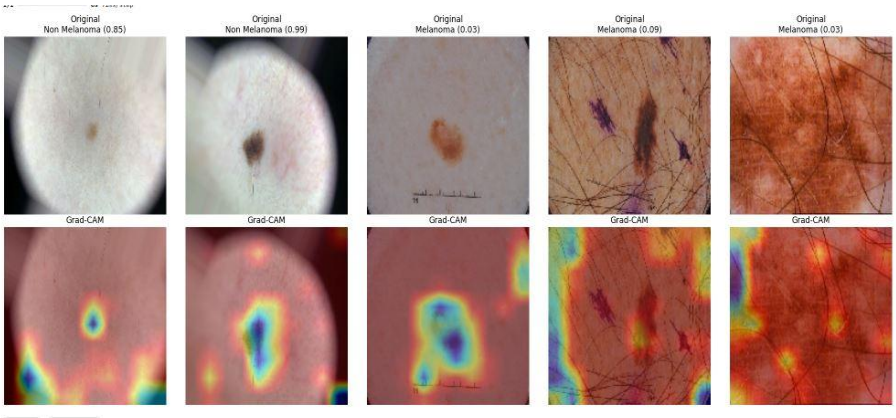


Fig. 13. Grad-CAM output on sample images by application of DS-MelNet architecture.

Although the proposed DS-MelNet model has not been implemented on edge devices, an approximate analysis of memory usage and inference time to assess feasibility for mobile deployment is provided. These estimates are based on the model's size (3.52 MB) and parameter count (923,790) and are summarized in Table 7.

Table 7. Approximate Edge Deployment Metrics for DS-MelNet

Metric	Value
Model Size	3.52 MB
Parameters	923,790
Memory Usage (Inference)	~80 MB
Inference Time (Snapdragon)	~120ms
Optimizer	Adam

4.4. Proposed SCNN Results

The results for proposed SCNN model are shown in Fig. 14. In terms of accuracy metrics, the SCNN achieved the highest accuracy of 76.23% on the APC-MDR-RGB pre-processed images. For the other two methods the DS-MelNet achieved quite good accuracy of 75% and 73% respectively. Similarly, for Precision, Recall and F1-score the SCNN achieved highest results for the same method, it achieved precision score of 81.78%, Recall of 71% and F1-score of 75.67%. The average accuracy achieved by the SCNN model on all the three methods is approximately 75% which is quite good for the proposed model with fewer parameters and lesser computational cost. Similarly, it achieved an average Precision value of 79%, Recall of 70% and F1-score of 74%.

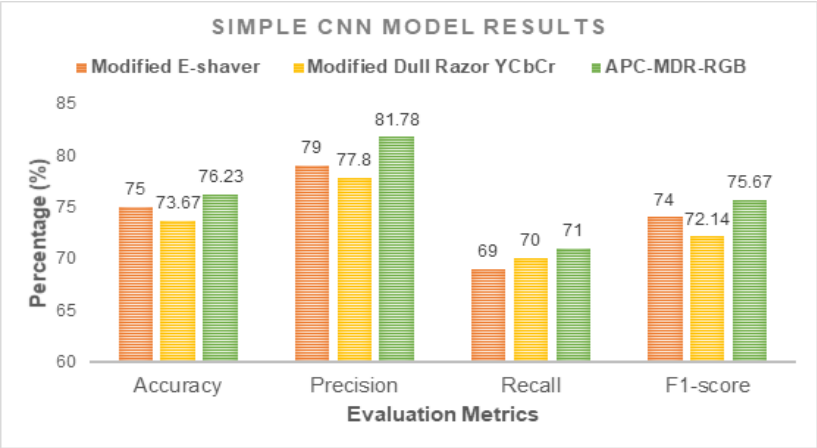


Fig. 14. Simple CNN model results for evaluation metrics on ISIC-2020 dataset.

4.5. Proposed Fine-tuned VGG-16 Results

The results for proposed fine-tuned model are shown in Fig. 15. The Fine-tuned VGG-16 performed well compared to the SCNN architecture and satisfactory compared to the DS-MelNet model. It utilized the saved weights for feature extraction with newly defined output layer or top layer that helped it achieved these results. In terms of the accuracy metrics it achieved the highest accuracy of 82.78% on the Modified Dull Razor YCbCr pre-processed images. For the other two methods the fine-tuned model achieved quite good accuracy of 80.25% and 78.6% respectively. Similarly, for Precision, Recall and F1-score the highest results were achieved for the same method, it achieved precision score of 88.34%, Recall of 82.54% and F1-score of 83.97%. The average accuracy achieved by the model on

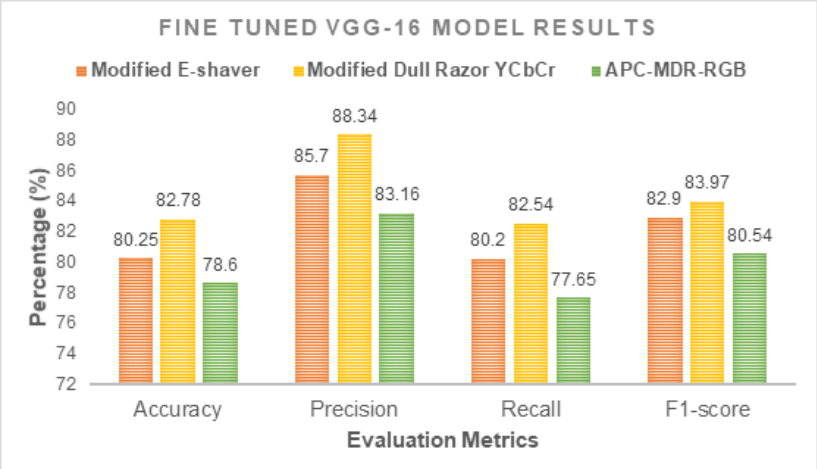


Fig. 15. Fine-tuned VGG-16 results for evaluation metrics on ISIC-2020 dataset.

the all the three methods is approximately 80% which is quite good but utilized large number of parameters and adds to the computational cost. Similarly, it achieved an average Precision value of 85%, Recall of 80% and F1-score of 82%.

Fig. 10, 14, 15 have discussed the results in terms of performance metrics for all the three proposed architectures showing the best pre-processing method that yields best results for the each of the proposed architectures. Metrics wise result comparison for the proposed architectures is discussed in the following section.

4.6. Metrics wise comparison of the results for the proposed systems

1. Accuracy:

In terms of Accuracy the DS-MelNet model achieved the best result of 86.57% for the pre-processed images by applying the Modified Dull Razor YCbCr method compared to the other two models proposed in the paper as shown in Fig. 16. The Fine-tuned architecture utilized the pre-trained weights and performed satisfactorily achieving accuracy of 82% which is closer to the results of DS-MelNet for the same pre-processing method. Similarly, the DS-MelNet architecture showed best results for other two pre-processing methods too compared with the SCNN and Fine-tuned models. Thus, the proposed DS-MelNet is quite effective and efficient in skin cancer classification with smaller model size and fewer parameters which makes it reliable for implementation in small devices too. The SCNN model performed best with highest accuracy of 76% on the APC-MDR-RGB images which is quite different compared to the other two model results.

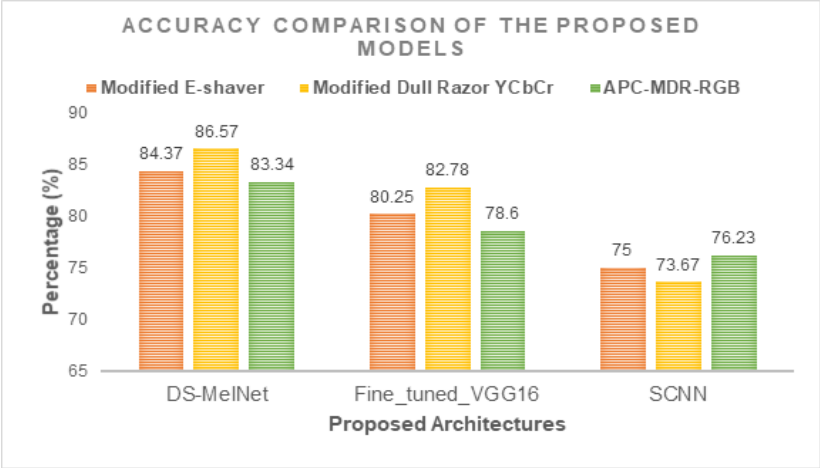


Fig. 16. Accuracy comparison of proposed system.

2. Precision and Recall:

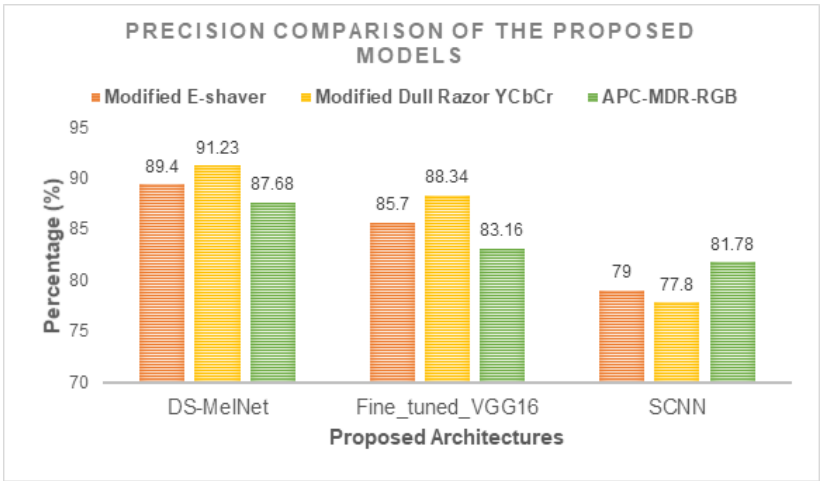


Fig. 17. Precision comparison of proposed system.

Precision and recall are another important metrics that represent the overall performance of the proposed methodology considering the positive and negative classes. The DS-MelNet achieved an excellent Precision of 91.23% and Recall of 84.22% for the Modified Dull Razor YCbCr pre-processed skin lesion images, whereas it achieved satisfactory results of 89% and 87% for precision for the other two methods. Recall achieved was quite good around 81% for APC-MDR-RGB and 83% for Modified E-shaver for the DS-MelNet, the other proposed models also showed

satisfactory results for precision of 88% and 81% respectively. The Recall values for other proposed methods also are satisfactory ranging between 69% to 83%. Precision comparison is shown in Fig. 17 and recall comparison shown in the Fig. 18.

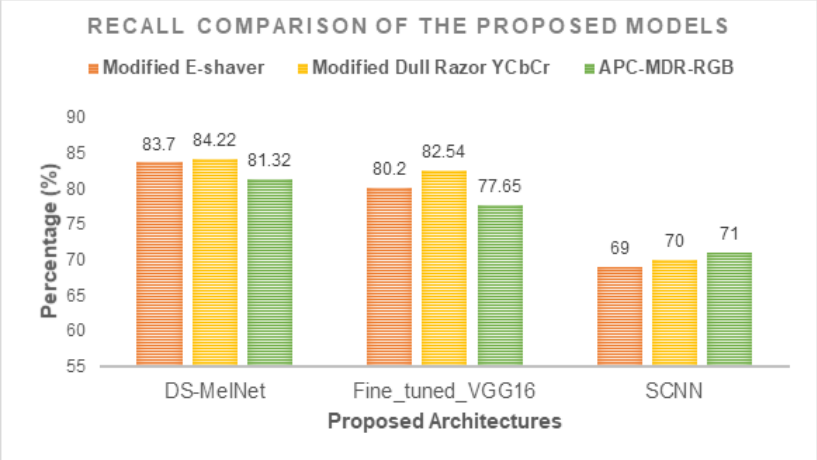


Fig. 18. Recall comparison of proposed system.

3. F1-score:

F1-score is an important metrics as it considers all the four parameters of the confusion matrix and gives result by balanced mechanism reducing the bias of any one class during the classification task. The F1-score graph shown in Fig. 19 indicates the superiority of the proposed DS-MelNet model on the other two architectures indicating its generalization in detection of skin cancer images. The DS-MelNet achieved F1-score of 88.45% for the Modified Dull Razor YCbCr method whereas the fine-tuned architecture achieved best value of 83.97% for the same pre-processing method. The SCNN presented best results on the APC-MDR-RGB images with F1-score of 75.67%.

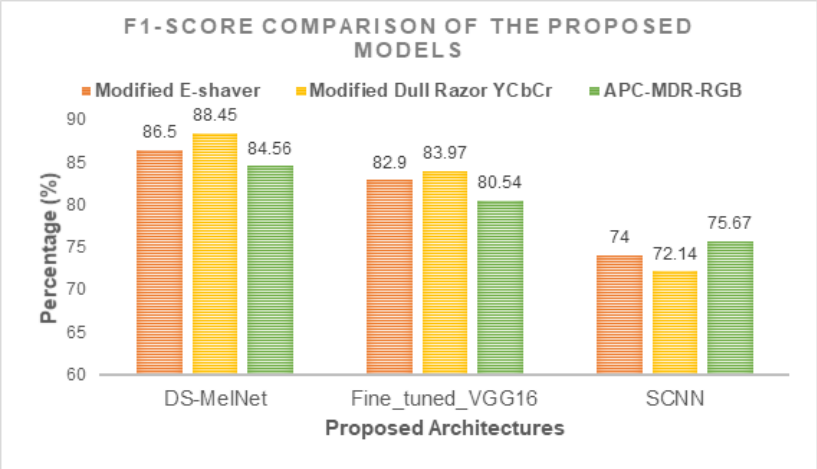


Fig. 19. F1-Score comparison of proposed system.

Recall graphs presented above give the sensitivity representation of the performance of the DS-MelNet model. The results for False Negative Rate (FNR) and Specificity are shown in Table 8. The model achieves a favourable balance between low false negatives for melanoma detection and high specificity for benign lesion identification.

Table 8. False Negative Rate (FNR) and Specificity for Melanoma Lesion Classification.

FNR (%)	Specificity (%)
16.30	85.04
15.78	88.92
18.68	85.36

The architectural design evolved through multiple experimental iterations. Starting from a single-stream baseline (SCNN), a second parallel stream was introduced to enhance representational diversity. The final configuration was chosen on the basis of consistent improvements in performance across multiple runs while maintaining computational efficiency. The various ablation study results and the cross validation presented in this section confirms and confronts the novelty of the proposed DS-MelNet.

4.7. Comparison of proposed architecture results with Existing systems results

Comparison with existing systems that have used ISIC-2020 [21] dataset is shown in Table 9. The proposed architecture DS-MelNet and fine-tuned VGG perform superior than the existing architectures in literature. The architectures available in literature have higher number of parameters compared to the DS-MelNet architecture, LCNet uses 3.3M parameters, DSCC_Net uses 1.1M parameters and InSiNet uses 2M parameters for classification, the DS-MelNet model just utilizes 0.92M parameters.

Table 9. Comparison of Proposed System and Existing Systems.

Methodology	Accuracy	Precision	Recall
LCNet [27]	83.67%	85.92%	82.67%
DSCC_Net [28]	82.29%	87.23%	81.89%
InSiNet [29]	80.23%	82.48%	79.50%
Fine-tuned VGG-16	82.78%	88.34%	82.54%
Proposed DS-MelNet	**86.57%**	**91.23%**	**84.22%**

The Table 10 shows the results obtained for Fix Match and Fully supervised model and details in terms of labels utilized, accuracy, model size and data utilization. The results clearly show that SSL is the need and presents a promising method for melanoma classification while handling negative cases.

Table 10. Comparison of Learning Paradigms for Melanoma Classification.

Approach	Label Utilized	Accuracy (%)	Model Size	Data Utilization
SSL (FixMatch)	10% labelled + unlabelled	86.57	3.52 MB	High (leverages unlabelled data)
Fully Supervised	10% labelled	74.89	3 MB	High

5. Conclusion & Future Scope

The detection of melanoma in the early stages is crucial to reduce future complexity and problems associated with skin cancer. The paper presents two architectures: SCNN and DS-MelNet based on the basic concept of the convolution neural network and one fine-tuned VGG-16 model. The proposed models have very fewer parameters and are computationally effective compared to the pre-trained and fine-tuned model discussed in literature. The dual stream model achieves an accuracy of 86%, which is based on the semi-supervised learning framework, whereas the SCNN model achieves an accuracy of 76% which is satisfactory compared to the DS-MelNet. The FT-VGG model presents an excellent test accuracy of 82%, indicating its improved ability to accurately classify data and effectively reduce errors. These results highlight the potential of pre-trained models for dermatological diagnosis, but they involve a more complex architecture and require more processing time compared to the CNN model. The semi-supervised learning-based architectures proposed in this study achieve high accuracy and robust performance on the ISIC-2020 skin lesion dataset, thus reducing the false alarms in skin cancer detection.

This paper utilized the VGG-16 architecture and fine-tuned it for improved results which have 14M parameters, the researchers can take this work and perform comparison of various existing pre-trained models in terms of their semi-supervised learning capability and analyse their performance. Various other datasets can also be explored for a similar framework for their generalization in the early detection of melanoma skin cancer.

All the Declarations and Statements

Author Contributions Statement

Apurva Shinde – Conceptualization, Methodology, Implementation , Data Curation , Original Draft Writing and Update, Model training and Validation.

Sangita Chaudhari – Conceptualization, Methodology, Review and Supervision: Proposed research ideas, Constructed the overall framework, Reviewed and edited the manuscript and supervised research execution.

All authors have read and agreed to the published version of the manuscript.

Conflict of Interest Statement
The authors declare no conflicts of interest.

Funding Declaration
Not applicable as this research was not supported by any grant.

Data Availability Statement
This study analyzed publicly available datasets. The datasets can be found in reference: 21 and 26.

Ethical Declarations
Not applicable as this study did not involve human subjects and/or animals.

Acknowledgments
Authors would like to thank Dr. D.Y.Patil University's Ramrao Adik Institute of Technology for their support and providing the research environment.

Declaration of Generative AI in Scholarly Writing
AI-assisted technologies are used at very few places to strengthen the language and readability.

Abbreviations
The following abbreviations are used in this manuscript:

APC-MDR-RGB - Adaptive Principal Curvature with Modified dull razor with RGB plane
CNN - Convolutional Neural Network
DS-MelNet - Dual-Stream Semi-Supervised Melanoma Network
DL - Deep Learning
FNR - False Negative Rate
MDR-YCbCr - Modified Dull Razor with YCbCr plane
SCNN - Simple Convolutional Neural Network
FT-VGG - Fine-tuned VGG16 Network
SSL - Semi-Supervised Learning

Appendix
Not applicable.

References

[1] Labani, S., Asthana, S., & Jain, M., "Incidence of melanoma and other skin cancers in India," *Asian Pacific Journal of Cancer Prevention*, vol. 21, no. 3, pp. 707–712, 2020.
[2] Halladja, M., et al., "Impact of delayed diagnosis on melanoma outcomes," *European Journal of Dermatology*, vol. 35, no. 1, pp. 45–53, 2025.
[3] Yang, J., Sun, J., & Li, Y., "Sampling-based dermoscopic image analysis for melanoma detection," *Computerized Medical Imaging and Graphics*, vol. 75, pp. 23–32, 2019.
[4] Huang, C., et al., "Skin lesion analysis using deep learning techniques," *IEEE Reviews in Biomedical Engineering*, vol. 12, pp. 123–137, 2019.
[5] Wang, Z., et al., "Non-invasive image-based techniques for melanoma detection," *Biomedical Signal Processing and Control*, vol. 71, 103162, 2022.
[6] Hussein, S., et al., "Binary classification of melanoma using deep learning," *Artificial Intelligence in Medicine*, vol. 148, 102745, 2024.
[7] Dakhli, A., et al., "Convolutional neural networks for skin cancer classification," *Expert Systems with Applications*, vol. 212, 118657, 2023.
[8] Baykal, B., et al., "CNN-based melanoma detection using dermoscopic images," *Diagnostics*, vol. 14, no. 2, p. 180, 2024.
[9] Ngo, T., et al., "Machine learning approaches for melanoma diagnosis," *Computers in Biology and Medicine*, vol. 142, 105220, 2022.
[10] Yi, X., et al., "Semi-supervised learning for skin lesion classification," *IEEE Journal of Biomedical and Health Informatics*, vol. 22, no. 2, pp. 434–445, 2018.
[11] Oliveira, R. B., Papa, J. P., Pereira, A. S., & Tavares, J. M. R. S., "Computational methods for pigmented skin lesion classification in images: Review and future trends," Neural Computing and Applications, vol. 27, no. 8, pp. 2101–2117, 2016.

[12] Noroozi, N., et al., “Computer-aided detection of basal cell carcinoma in histopathological images using Z-transform features,” Journal of Medical Imaging and Health Informatics, vol. 6, no. 4, pp. 1060–1066, 2016.
[13] Anas, A., et al., “Skin cancer classification using color and texture features,” International Journal of Computer Applications, vol. 170, no. 3, pp. 1–6, 2017.
[14] Ghahfarrokhi, B. S., et al., “Malignant skin cancer detection using machine learning from dermoscopic images,” Biomedical Signal Processing and Control, vol. 79, 104108, 2023.
[15] Yu, Z., Jiang, X., Zhou, F., Qin, J., Ni, D., Chen, S., Lei, B., & Wang, T., “Melanoma recognition in dermoscopy images via aggregated deep convolutional features,” IEEE Transactions on Biomedical Engineering, vol. 66, no. 4, pp. 1006–1016, 2018.
[16] Daghrir, J., Tlig, L., Bouchouicha, M., & Sayadi, M., “Melanoma skin cancer detection using deep learning and classical machine learning techniques: A hybrid approach,” in Proceedings of the 5th International Conference on Advanced Technologies for Signal and Image Processing (ATSIP), IEEE, pp. 1–5, 2020.
[17] Agarwal, N., Singh, V., & Singh, P., “Semi-supervised learning with GANs for melanoma detection,” in Proceedings of the 6th International Conference on Intelligent Computing and Control Systems (ICICCS), IEEE, pp. 141–147, 2022.
[18] Gajera, H. K., Nayak, D. R., & Zaveri, M. A., “A comprehensive analysis of dermoscopy images for melanoma detection via deep CNN features,” Biomedical Signal Processing and Control, vol. 79, p. 104186, 2023.
[19] Bassel, A., Abdulkareem, A. B., Alyasseri, Z. A. A., Sani, N. S., & Mohammed, H. J., “Automatic malignant and benign skin cancer classification using a hybrid deep learning approach,” Diagnostics, vol. 12, no. 10, p. 2472, 2022.
[20] Tembhurne, J. V., Hebbar, N., Patil, H. Y., & Diwan, T., “Skin cancer detection using ensemble of machine learning and deep learning techniques,” Multimedia Tools and Applications, vol. 82, no. 18, pp. 27501–27524, 2023.
[21] Rotemberg, V., Kurtansky, N., Betz-Stablein, B., Caffery, L., Chousakos, E., Codella, N., Combalia, M., Dusza, S., Guitera, P., Gutman, D., et al., “A patient-centric dataset of images and metadata for identifying melanomas using clinical context,” Scientific Data, vol. 8, no. 1, p. 34, 2021.
[22] Apurva Shinde and Sangita Chaudhari. Statistical analysis of hair detection and removal techniques using dermoscopic images. In International Conference on Computer Vision and Image Processing, pages 402–414. Springer,2022.
[23] Namrata Verma and Pankaj Kumar Mishra. Design of an efficient unet-based transfer learning model for enhancing skin cancer segmentation and classification performance. International Journal of Image, Graphics and Signal Processing, 2025.
[24] Momina Shaheen, Usman Saif, Shahid M Awan, Faizan Ahmad, and Aimen Anum. Classification of images of skin lesion using deep learning. International Journal of Intelligent Systems and Applications, 15(2):23, 2023.
[25] Kihyuk Sohn, David Berthelot, Chun-Liang Li, Zizhao Zhang, Nicholas Carlini, Ekin D. Cubuk, Alex Kurakin, Han Zhang, and Colin Raffel. Fixmatch: Simplifying semi-supervised learning with consistency and confidence, 2020.
[26] Teresa Mendonc ̧a, M Celebi, T Mendonca, and J Marques. Ph2: A public database for the analysis of dermoscopic images. Dermoscopy image analysis, 2, 2015.
[27] Rajdeep Kaur and Sukhjeet Kaur Ranade. Du-net+: a fully convolutional neural network architecture for semantic segmentation of skin lesions. Signal, Image and Video Processing, 19(1):152, 2025.
[28] Maryam Tahir, Ahmad Naeem, Hassaan Malik, Jawad Tanveer, Rizwan Ali Naqvi, and Seung-Won Lee. Dscc net: multi-classification deep learning models for diagnosing of skin cancer using dermoscopic images. Cancers, 15(7):2179, 2023.
[29] Hatice Catal Reis, Veysel Turk, Kourosh Khoshelham, and Serhat Kaya. Insinet: a deep convolutional approach to skin cancer detection and segmentation. Medical & Biological Engineering & Computing, pages 1–20, 2022.

Authors’ Profiles



Apurva Shinde is working as an Assistant Professor in the Department of Computer Engineering at Ramrao Adik Institute of Technology, D. Y. Patil Deemed to be University, Nerul, Navi Mumbai. She completed her B.E. and M.E. in Computer Engineering in 2013 and 2016, respectively. She is currently pursuing her Ph.D. from Ramrao Adik Institute of Technology in the domain of Medical Image Processing. She has over 12 years of teaching and research experience and has published several research papers in reputed journals and conferences. Her research interests include Image Processing, Machine Learning, and Data Science.



Sangita Chaudhari is working as Professor and Head in Computer Science and Engineering Department at Ramrao Adik Institute of Technology, D. Y Patil Deemed to be University, Nerul, Navi Mumbai. She has completed her UG and PG in Computer Engineering in 1999 and 2008 respectively. She has earned PhD from IIT Bombay in Geo spatial Data Security in 2016. She has 26 years of teaching and research experience. She has published 100+ research papers in reputed Journals and Conferences. Her research interests Satellite Image Processing, Pattern Recognition, Information Security, Advanced Databases, Geographic Information System and Data Analytics.

How to cite this paper:

Apurva S. Shinde, Sangita S. Chaudhari, "DS-MelNet: An Enhanced Dual Stream Semi-Supervised Mechanism for Melanoma Classification", International Journal of Image, Graphics and Signal Processing(IJIGSP), Vol.18, No.3, pp. 110-132, 2026. DOI:10.5815/ijigsp.2026.03.06