check for updates

Enhancing skin lesion segmentation with U-Net++: Design, analysis, and performance evaluation

 Sheetal Nana Patil[1+]

 Hitendra D. Patil[2]

[1,2]Department of Computer Engineering, Shri Shivaji Vidya Prasarak Sanstha's, Bapusaheb Shivajirao Deore College of Engineering, Dhule, India.
[1]Email: Sheetalpatil2000@gmail.com
[2]Email: hitendradpatil@gmail.com

(+ Corresponding author)

ABSTRACT

The present research examines the enhancement of skin lesion segmentation with U-Net++. Achieving accurate classification of dermoscopy images is heavily contingent on the precise segmentation of skin lesions or nodules. However, this task is considerably challenging due to the elusive edges, irregular perimeters, and variations both within and across lesion classes. Despite numerous existing algorithms for segmenting skin lesions from dermoscopic images, they often fall short of industry benchmarks in terms of precision. To address this, our research introduces a novel U-Net++ architecture, implementing tailored adjustments to feature map dimensions. The aim is to significantly enhance automated segmentation precision for dermoscopic images. Our evaluation involved a comprehensive assessment of the model's performance, encompassing an exploration of various parameters such as epochs, batch size, and optimizer selections. Additionally, we conducted extensive testing using augmentation techniques to bolster the image volume within the HAM10000 dataset. A key innovation in our research is the integration of a hair removal process into the U-Net++ algorithm, significantly enhancing image quality and subsequently leading to improved segmentation accuracy. The results of our proposed method demonstrate substantial advancements, showcasing an impressive Mean Intersection over Union (IoU) of 84.1%, a Mean Dice Coefficient of 91.02%, and a Segmentation Test Accuracy of 95.10%. Our suggested U-Net++ algorithm does a better job of segmenting than U-Net, Modified U-Net, K-Nearest Neighbors (KNN), and Support Vector Machine (SVM). This shows that it could be used to improve dermoscopy image analysis. Our proposed algorithm shows remarkable improvement in both observational outcomes and classifier performance.

Contribution/Originality: The study presents a customized U-Net++ architecture for accurate segmentation of skin lesions in dermoscopy images. It emphasizes adjustments to feature map dimensions and integrates a hair removal step. Notably, contour prediction is highlighted for its crucial role in enhancing segmentation performance.

1. INTRODUCTION

Skin cancer, encompassing melanoma, stands as one of the prevailing forms of cancer on a global scale. Prompt and precise identification of melanoma holds immense importance, as it plays a critical role in ensuring successful treatment and better results for patients. Within the realm of machine learning, deep learning emerges as a subset with considerable promise, particularly in the realm of medical image interpretation. Melanoma, a skin condition that has seen a worldwide uptick, recorded nearly 3 million fresh instances in the year 2022 [1]. Tackling skin cancer has evolved into a significant issue within public health. While visually scrutinizing the skin's exterior might

appear uncomplicated for detection, medical professionals frequently miss the majority of melanoma cases. Approximately 60% of diagnoses stem from visual assessments conducted by proficient dermatologists.

Unfortunately, this also means that a significant number of potentially treatable melanomas remain unnoticed until they reach advanced stages. The accuracy and timelinessof segmenting skin lesions play a vital role in identifying and locating visual dermoscopic skin characteristics and categorizing skin-related ailments. Dermoscopy, an imaging method enhancing melanoma diagnosis and decreasing melanoma-related deaths, allows the visualization of deeper skin layers by minimizing surface reflections.
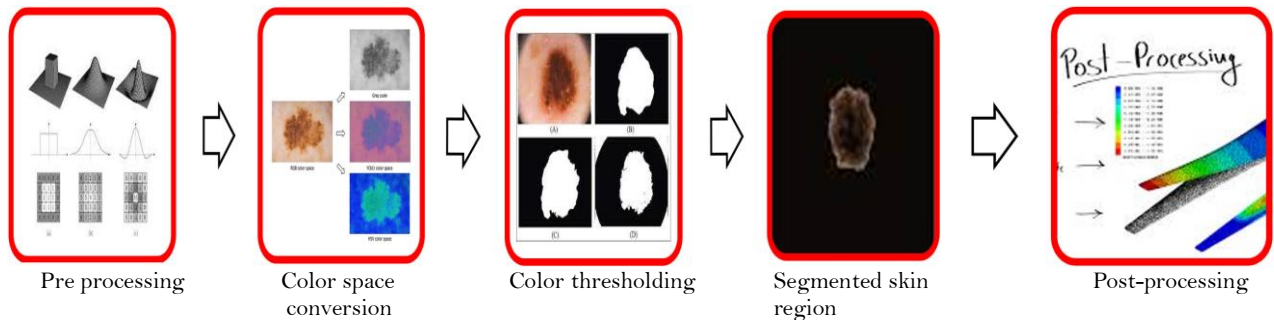

| Pre processing | Color space conversion | Color thresholding | Segmented skin region | Post-processing |

Figure 1. Generalized skin lesion segmentation block diagram.

Figure 1 shows a block diagram of skin lesions in dermoscopic images. Skin lesion segmentation becomes difficult due to variability in lesion types, complex and irregular boundaries, overlapping lesions, class imbalance, limited annotated data, varied imaging conditions, small lesions and fine details. An automated melanoma detection from the surrounding skin is a critical process in the computerized assessment of dermoscopic images. The presence of artefacts and inherent cutaneous elements such as hairs, borders, blood vessels, and air bubbles further compounds the complexity of automated segmentation. Diagnosing skin lesions becomes more intricate when comparing them based on color skin imaging, like distinguishing between nevi and melanoma. An imperative need exists for a robust automated technique to locate skin lesions, as early detection significantly shortens diagnostic timelines [2]. Prompt lesion identification remains crucial, particularly given the higher survival rate it offers when compared to individuals who have recently undergone treatment. Conventional visual assessment employed in dermoscopic image analysis proves inadequate for accurate lesion type assessment.

2. SKIN LESION SEGMENTATION USING DEEP LEARNING

Many scientists have expressed a keen interest in improving the diagnosis of severe skin conditions through various methods of segmenting skin diseases. These techniques include clustering, expanding regions, dividing, and utilizing both supervised and unsupervised learning methods. Skin lesion segmentation is of utmost importance, as it entails the process of separating a singular image into smaller sections [3]. Researchers have put forward a range of methods, which include both modern deep learning techniques and traditional segmentation approaches, to tackle this aspect of analyzing skin lesions.

Researchers have introduced a hybrid k-means segmentation [4], in which images are segmented using k-means clustering to locate the precise lesion region, and then the firefly technique is implemented to improve segmentation accuracy [5]. In histogram segmentation with the Genetic algorithm, initially computation of a set of image pixels using c-means clustering is employed, then graph cut methodology is employed to attain the segmentation of skin lesions [6, 7].

A different approach to segmentation involves using the region-growing technique, which incorporates fuzzy clustering segmentation [8]. This unsupervised clustering method is an extended form of k-means clustering. In this method, adjacent pixels are merged based on their spatial context to facilitate region growth and segmentation [9]. During the segmentation process, the image's similarity index is assessed based on important attributes such

as color and texture [10]. In situations involving K-means clustering and histogram computation, segmentation centers on the extraction of color features [11]. In the semi-supervised, shift mean segmentation technique, there is no need for an initial determination of the cluster count [12]. In threshold segmentation, the image is subdivided into various regions using a predefined threshold value as a reference. This strategy helps improve the distinctiveness of edges in areas affected by cancer, thereby supporting the segmentation process [13].

Researchers introduced an auxiliary function-based technique [14], involving the creation of a smoothed auxiliary function using Bezier curves. This function is constructed using a local minimizer, effectively addressing global optimization concerns. The primary aim is focused on utilizing active contour fusion segmentation to segment pixel data with low contrast [15]. A novel method involving exponential neighborhood grey wolf optimization was employed to optimize a fully convolved network for the segmentation of dermoscopic data [16]. Improved CNNs such as Retina-Deep Lab [17], and recurrent CNN [18] are widely used by researchers for lesion segmentation on large dataset; alternatively combinations of ResNet and DenseNet inside a dense encoder-decoder-based system led to performance enhancement [19].

Furthermore, the integration of Atrous Spatial Pyramid Pooling (ASPP) involved in capturing contextual information with skip connections played a vital role in recovering image pixel-level information [20]. In Adaptive moment Estimation (ADAM) optimization [21], multi-scale fusion features and spatial image characteristics are considered to minimize segmentation loss. In improved fully convolutional networks (iFCN), lesion segmentation is performed without preprocessing. To address potential drawbacks of the segmentation process, the approach takes into account specific lesion location information along its perimeter [22]. In the current method, like the one involving a Gaussian radial basis kernel and SVM classifier [23, 24], the combination of texture and color features extracted from dermoscopic images faces challenges in generating segmented images of high quality and achieving superior classification results.

More recently, though, researchers have been looking more closely at different versions of Deep Learning (DL) and Convolutional Neural Network (CNN) architectures to try to solve the problems mentioned above in terms of segmentation [28, 29]. In a fully automated method for skin lesion segmentation [30], researchers employed a 19-layer deep CNN that was extensively trained without requiring any prior data-specific knowledge. Their approach yielded a Jacquard distance value of 0.963. Lately, the U-Net architecture was introduced for the segmentation of lesions and their characteristics, utilizing the framework of convolutional neural networks (CNN). Numerous researchers have engaged with the ISIC datasets from 2016, 2017, and 2018, encompassing a total of 2750 images. In a particular study by Yuan [30], the author focused on three distinct classes, namely melanocytic nevi (NV), melanoma (MEL), and intraepithelial carcinoma/Bowen disease (AKIEC), employing four pre-trained models: Inception-v3, ResNet-50, InceptionResNet-v2, and DenseNet20, for the purpose of segmentation across the 2750 images. The anticipated segmentation accuracies were 81.29%, 81.57%, 81.34%, and 73.44%, respectively.

U Net is an improved version of CNN that has flexible global aggregation blocks. These blocks are effective for preserving image pixel data during-up-and-down sampling. For the International Skin Imaging Collaboration (ISIC) 2018 dataset, a MultiRes U Net architecture [31] and cascaded fully convolutional neural networks [32] were developed, achieving dice score measurements of 87.9% and 92.3% respectively. The outcomes of the segmentation process show a noticeable improvement, and the achieved results are deemed satisfactory. Existing literature demonstrates that utilizing the segmented images for classification purposes leads to a notable enhancement in classification accuracy.

The significant outcomes from the study reveal and establish objectives to design novel algorithms to:

- Implement image enhancement techniques before segmentation for performance improvement.
- Data augmentation methods have been implemented to enhance the variability of images, thereby promoting improved stability in the overall process.
- The suggested model is subjected to validation using diverse optimizers, batch sizes, and epochs in order to

achieve enhanced accuracy.

- The proposed model has been analyzed with numerous performance parameters and observational results.

3. SEGMENTATION USING MODIFIED U-NET

3.1. Overview & Architecture

An improved iteration of the Convolutional Neural Network (CNN) architecture was created to handle biomedical images. A goal extends beyond categorizing the presence of an infection; it also involves objective measures.

It also involves identifying the precise area that the infection has affected. The Modified U-Net architecture comprises two fundamental paths: the initial path referred to as the contraction path, or encoder, and the expansion path or decoder. The encoder is dedicated to grasping the contextual information within the image, while the decoder employs transposed convolutions to facilitate accurate identification of localized regions [33].

Figure 2 shows a modified U-Net structure in which the input image undergoes a multi-tiered decomposition within the encoder pathway, and the feature maps are reduced using a max pooling layer.

Modified U Net architecture consists of 3X3 layer size Convolution [⟶] with ReLU (Rectified Linear Unit) activation function, Max pooling Layer [⟶], 2X2 size Upsampling [⟶] and concatenation [⟶] in the encoder-decoder section.

The details of Modified U-Net architecture are elaborated in Table 1.

Table 1. Layer wise modified U net model design summary.

| Layer/Operation	Input shape	Output shape	Parameters
Input_1	(None, 128, 128,3)	----	0
Conv2d	(None, 128, 128, 64)	(None, 128, 128, 64)	1792
Conv2d_1	(None, 128, 128, 64)	(None, 128, 128, 64)	36928
Max_pooling2d	(None, 128, 128, 64)	(None, 64, 64, 64)	0
Conv2d_2	(None, 64, 64, 64)	(None, 64, 64, 128)	73856
Conv2d_3	(None, 64, 64, 128)	(None, 64, 64, 128)	147584
Max_pooling2d_1	(None, 64, 64, 128)	(None, 32, 32, 128)	0
Conv2d_4	(None, 32, 32, 128)	(None, 32, 32, 256)	295168
Conv2d_5	(None, 32, 32, 256)	(None, 32, 32, 256)	590080
Max_pooling2d_2	(None, 32, 32, 256)	(None, 16, 16, 256)	0
Conv2d_6	(None, 16, 16, 256)	(None, 16, 16, 512)	1180160
Conv2d_7	(None, 16, 16, 512)	(None, 16, 16, 512)	2359808
Up_sampling2d	(None, 16, 16, 512)	(None, 32, 32, 512)	0
Concatenate	(None, 32, 32, 256)	(None, 32, 32, 768)	0
Conv2d_8	(None, 32, 32, 768)	(None, 32, 32, 256)	1769728
Conv2d_9	(None, 32, 32, 256)	(None, 32, 32, 256)	590080
Up_sampling2d_1	(None, 32, 32, 256)	(None, 64, 64, 256)	0
Concatenate_1	(None, 64, 64, 128)	(None, 64, 64, 384)	0
Conv2d_10	(None, 64, 64, 384)	(None, 64, 64, 128)	442496
Conv2d_11	(None, 64, 64, 128)	(None, 64, 64, 128)	147584
Up_sampling2d_2	(None, 64, 64, 128)	(None,128,128, 128)	0
Concatenate_2	(None, 128, 128, 128)	(None, 128, 128, 192)	0
Conv2d_12	(None, 128, 128, 192)	(None, 128, 128, 1)	193
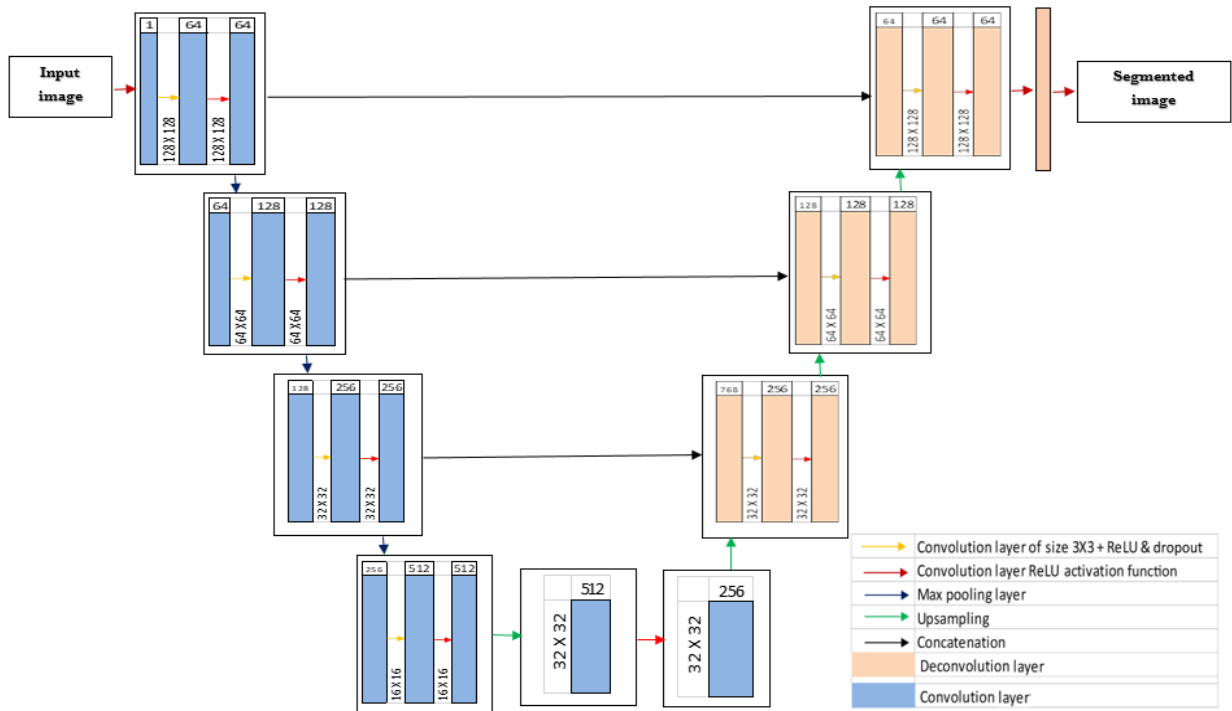
Figure 2. Modified U net architecture.

3.2. Mathematical Modelling

In the encoder section of U Net X, the input image tensor is Conv2D represents the 2D convolutional operation, and MaxPooling2D represents the max-pooling operation. The encoder consists of 3 blocks; each block is represented by Equations 1 to 6.

$$F_1 = Conv2D(X) \qquad (1)$$

Pooled data by Ist block is
$$F_{1\,pooled} = Maxpooling\,2D\,(F_1) \qquad (2)$$

$$F_2 = Conv2D(F_{1\,pooled}) \qquad (3)$$

Pooled data by IInd block is
$$F_{2\,pooled} = Maxpooling\,2D\,(F_2) \qquad (4)$$

$$F_3 = Conv2D(F_{2\,pooled}) \qquad (5)$$

Pooled data by IIIrd block is
$$F_{3\,pooled} = Maxpooling\,2D\,(F_3) \qquad (6)$$

In the decoder section, upsampling & concatenation operations are performed by each block. The decoder consists of 3 blocks, in which F_i is the feature map from the ith encoder block [34].

Decoder-3 mathematically represented by Equation 7, 8 and 9

$$U_{3\,UP} = Upsampling\,2D(F_{3,pooled}) \qquad (7)$$

$$U_3 = Concatenate(U_{3,up}, F_3) \qquad (8)$$

$$D_{3,1} = Conv2D(U_3) \ \& \ D_{3,2} = Conv2D(D_{3,1}) \qquad (9)$$

Decoder-2 mathematically represented by Equation 10, 11 and 12

$$U_{2\,UP} = Upsampling\,2D(D_{3,2}) \qquad (10)$$

$$U_2 = Concatenate(U_{2,up}, F_2) \qquad (11)$$

$$D_{2,1} = Conv2D(U_2) \ \& \ D_{2,2} = Conv2D(D_{2,1}) \qquad (12)$$

Decoder-1 mathematically represented by Equation 13, 14 and 15

$$U_{1\,UP} = Upsampling\,2D(D_{2,2}) \qquad (13)$$

$$U_1 = Concatenate(U_{1,up}, F_1) \qquad (14)$$

$$D_{1,1} = Conv2D(U_1) \ \& \ D_{1,2} = Conv2D(D_{1,1}) \qquad (15)$$

Finally output of U Net is calculated by using Equation 16

$$Output = Conv2D(D_{1,2}) \qquad (16)$$

4. U NET++ WITH HAIR REMOVAL TECHNIQUES

4.1. Overview & Architecture

The core of our segmentation model is the U-Net++ architecture, an extension of the original U-Net. U-Net++ enhances the representation capability of the model by introducing multiple skip pathways at different depths in the network. This allows the model to capture multi-scale features and improve segmentation accuracy.

Our model consists of an encoder-decoder structure. The encoder extracts hierarchical features from the input image through a series of convolutional layers followed by max-pooling operations. This downsampling process reduces the spatial dimensions while increasing the depth of the features. The middle block further refines these features to capture more complex patterns [35]. In U-Net++, dense skip pathways are introduced by connecting each layer in the encoder with every layer in the corresponding decoder. This allows for dense and rich feature reuse, helping to capture both low-level and high-level features. Figure 3 depicts an enhanced U-Net++ encoder-decoder architecture, detailing the specifications for each layer in the structure.
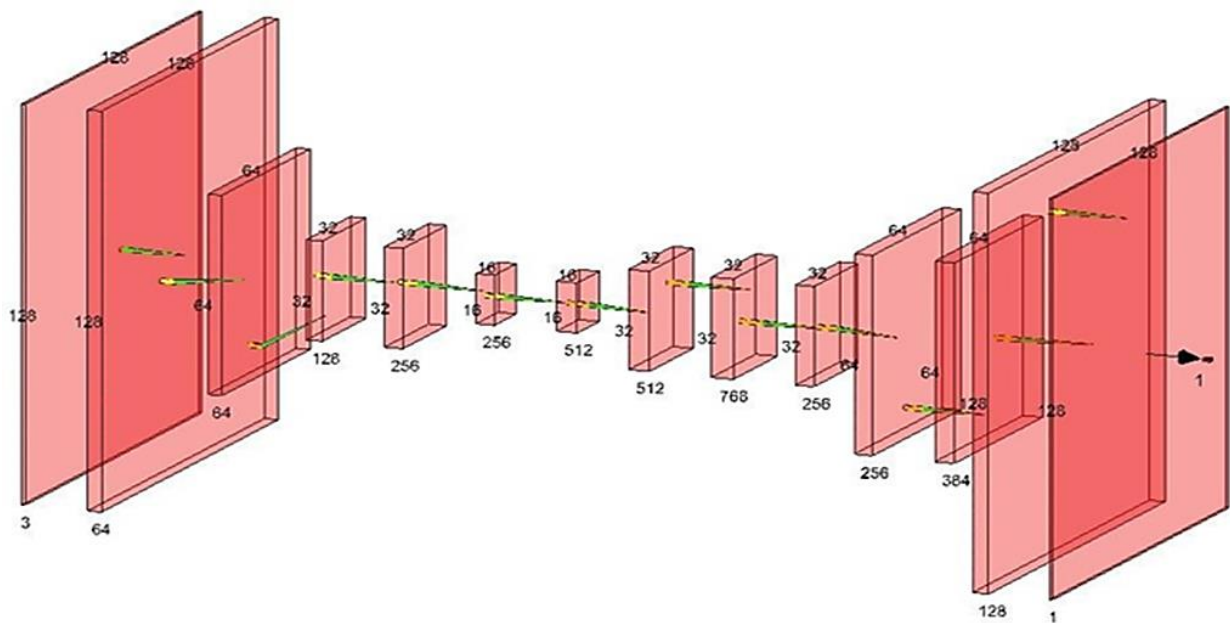

Figure 3. U Net++ Encoder-decoder architecture.

The decoder then reconstructs the segmented mask using a series of upsampling operations. Skip connections are established by concatenating the encoder's feature maps with the corresponding decoder's feature maps. This enables the model to leverage both low-level and high-level features for accurate segmentation. The details of the Modified U-Net architecture are elaborated in Table 2.

To enhance the model's accuracy, we also implement a hair removal preprocessing step. By thresholding the input images in the HSV color space; we create a hair mask that highlights hair regions. Inverting this mask allows us to identify the background and hair regions. Applying this mask to the original image using bitwise operations removes the hair, which can otherwise interfere with accurate segmentation.

Table 2. Layer wise U net ++ Model design summary.

Layer/Operation	Input shape	Output shape	Parameters
Input	(None, 128, 128, 3)	0	
Conv2D_1	(None, 128, 128, 64)	1792	Input
Conv2D_2	(None, 128, 128, 64)	36928	Conv2D_1
MaxPooling2D_1	(None, 64, 64, 64)	0	Conv2D_2
Conv2D_3	(None, 64, 64, 128)	73856	MaxPooling2D_1
Conv2D_4	((None, 64, 64, 128)	147584	Conv2D_3
MaxPooling2D_2	(None, 32, 32, 128)	0	Conv2D_4
Conv2D_5	(None, 32, 32, 256)	295168	MaxPooling2D_2
Conv2D_6	(None, 32, 32, 256)	590080	Conv2D_5
MaxPooling2D_3	(None, 16, 16, 256)	0	Conv2D_6
Conv2D_7	(None, 16, 16, 512)	1180160	MaxPooling2D_3
Conv2D_8	(None, 16, 16, 512)	2359808	Conv2D_7
UpSampling2D_1	(None, 32, 32, 512)	0	Conv2D_8
Concatenate_1	(None, 32, 32, 768)	0	Conv2D_6, UpSampling2D_1
Conv2D_9	(None, 32, 32, 256)	1769728	Concatenate_1
Conv2D_10	(None, 32, 32, 256)	590080	Conv2D_9
UpSampling2D_2	(None, 64, 64, 256)	0	Conv2D_10
Concatenate_2	(None, 64, 64, 384)	0	Conv2D_4, UpSampling2D_2
Conv2D_11	(None, 64, 64, 128)	442496	Concatenate_2
Conv2D_12	(None, 64, 64, 128)	147584	Conv2D_11
UpSampling2D_3	(None, 128, 128, 128)	0	Conv2D_12
Concatenate_3	(None, 128, 128, 192)	0	Conv2D_2, UpSampling2D_3
Conv2D_13	(None, 128, 128, 64)	110656	Concatenate_3
Conv2D_14	(None, 128, 128, 64)	36928	Conv2D_13
Conv2D_15	(None, 128, 128, 1)	65	Conv2D_14

4.2. U Net++ Backgrounds with Mathematical Preliminaries

The U-Net++ architecture builds upon the U-Net architecture with additional skip connections and nested feature aggregation. The working of the encoder section in U-Net++ is similar to U-Net, only the decoder section was reformed. Decoder with Nested Skip Connections is as shown below:

Decoder-3 mathematically represented by Equation 17, 18 & 19.

$$U_{3\ UP} = Upsampling2D\ (F_{3,pooled}) \qquad (17)$$

$$U_3 = Concatenate\ (U_{3,up}, F_3) \qquad (18)$$

$$D_{3,1} = Conv2D\ (U_3)\ \&\ D_{3,2} = Conv2D\ (D_{3,1}) \qquad (19)$$

Decoder-2 mathematically represented by Equation 20, 21 & 22.

$$U_{2\ UP} = Upsampling2D\ (D_{3,2}) \qquad (20)$$

$$U_2 = Concatenate\ (U_{2,up}, F_2) \qquad (21)$$

$$D_{2,1} = Conv2D\ (U_2)\ \&\ D_{2,2} = Conv2D\ (D_{2,1}) \qquad (22)$$

Decoder-1 mathematically represented by Equation 23, 24 & 25

$$U_{1\ UP} = Upsampling2D\ (D_{2,2}) \qquad (23)$$

$$U_1 = Concatenate\ (U_{1,up}, F_1) \qquad (24)$$

$$D_{1,1} = Conv2D\ (U_1)\ \&\ D_{1,2} = Conv2D\ (D_{1,1}) \qquad (25)$$

The U Net++ architecture combines data from different levels of encoding through nested skip connections. This lets the model do more complete feature aggregation and better segmentation, especially when working with objects that are different sizes or appear at different scales in the image.

In U Net ++, up sampling & concatenate process is explained by Equation 26, 27 & 28.

$$U_{3,1} = Upsampling2D\ (D_{3,2}) \qquad (26)$$

$$U_{3,2} = Concatenate\ (U_{3,1}, F_2) \qquad (27)$$

$$U_{3,3} = Upsampling2D\ (D_{3,1}) \qquad (28)$$

While nesting process is expressed by using Equation 29, 30 & 31

$$U_{3,Nested} = Concatenate (U_{3,3}, U_{3,2}, F_1) \qquad (29)$$

$$D_{3,Nested,1} = Conv2D (U_{3,Nested}) \qquad (30)$$

$$D_{3,Nested,2} = Conv2D (D_{3,Nested,1}) \qquad (31)$$

Finally output of U Net++ is calculated by using Equation 32.

$$Output = Conv2 (D_{3,nested,2}) \qquad (32)$$

5. RESULTS & DISCUSSIONS

5.1. Materials and Methods

The suggested approach utilizes the Modified U-Net and U-Net++ architectures to perform segmentation of lesions in dermoscopy images of skin diseases. This novel model was tested on the Human against Machine (HAM 10000) dataset, which comprises 5000 dermoscopy images of skin diseases. The images were resized to a standardized resolution of 128 x 128 pixels. This step ensures uniformity in input dimensions and prepares the data formodel training. Training is conducted with a batch size of 16 images.

Using this smaller batch size ensures more frequent updates to the model's weights, facilitating quicker convergence. During each training iteration, 16 images along with their corresponding masks are processed simultaneously. In order to mitigate overfitting and improve the model's ability to generalize, dropout regularization is integrated. Dropout involves randomly disabling a portion of neurons during training, essentially generating an ensemble of neural networks with varying sets of active neurons in each iteration. A dropout rate of 0.2 is implemented, indicating that 20% of neurons are randomly deactivated during the training process. Both the U-Net and U-Net++ models undergo training for 10 epochs, where an epoch signifies a full iteration through the entire training dataset. Employing multiple epochs enables the models to progressively grasp intricate patterns and connections within the data. Observing the changes in training and validation losses across epochs aids in evaluating convergence and identifying the possibility of overfitting.
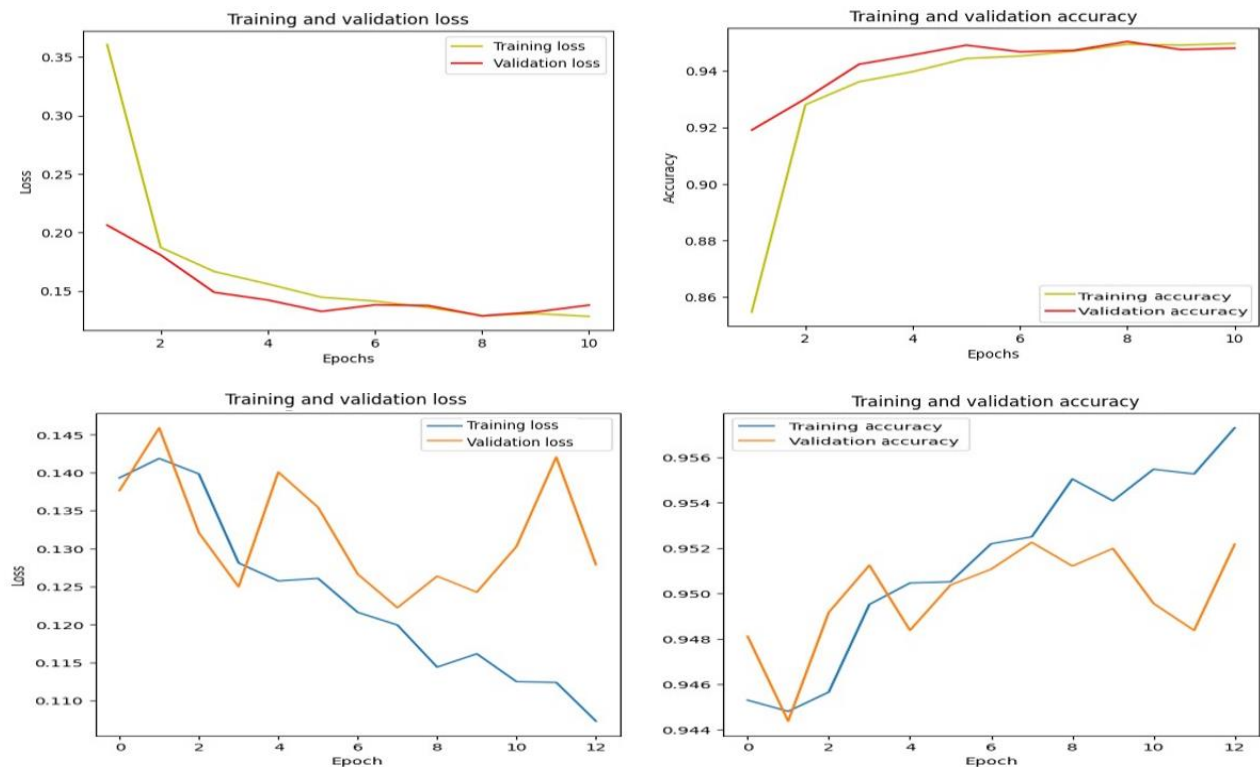

Figure 4. (a) U net training, Validation loss & accuracy **(b)** Proposed U Net++ training, validation loss & accuracy.

5.2. Training, Validation Loss & Accuracy

For training, the binary cross-entropy loss function is utilized, which is well-suited for binary segmentation tasks. To adjust the model's weights during training, the Adam optimizer is employed. The role of the loss function is to steer the optimization procedure by minimizing the disparity between the predicted and actual masks. U-Net segmentation works better after the second epoch, where the model has stable loss and accuracy. Epoch count plays a vital role in U Net⁺⁺ segmentation. In Figure 4 (a), the training loss and accuracy variations for the existing U-Net segmentation model across epochs are depicted. The model attained a peak accuracy of 93.7% and a minimum loss of 15% after the second epoch. Figure 4 (b) shows how the training loss and accuracy changed over time for our proposed U-Net++ segmentation model that includes a hair removal component at different time points. Notably, this improved model achieved a maximum accuracy of 95.70% and a minimum loss of 10% after the tenth epoch.

5.3. Performance Parameters of Segmentation Model

Performance parameters for image segmentation tasks are crucial in evaluating the quality of the model's predictions. Our novel U-Net++ Segmentation approach, augmented with a hair removal algorithm, is evaluated by comparing its performance parameters with existing methodologies, including the modified U-Net (MU Net), conventional U-Net, the K-Nearest Neighbors Algorithm (KNN), and the Support Vector Machine (SVM).
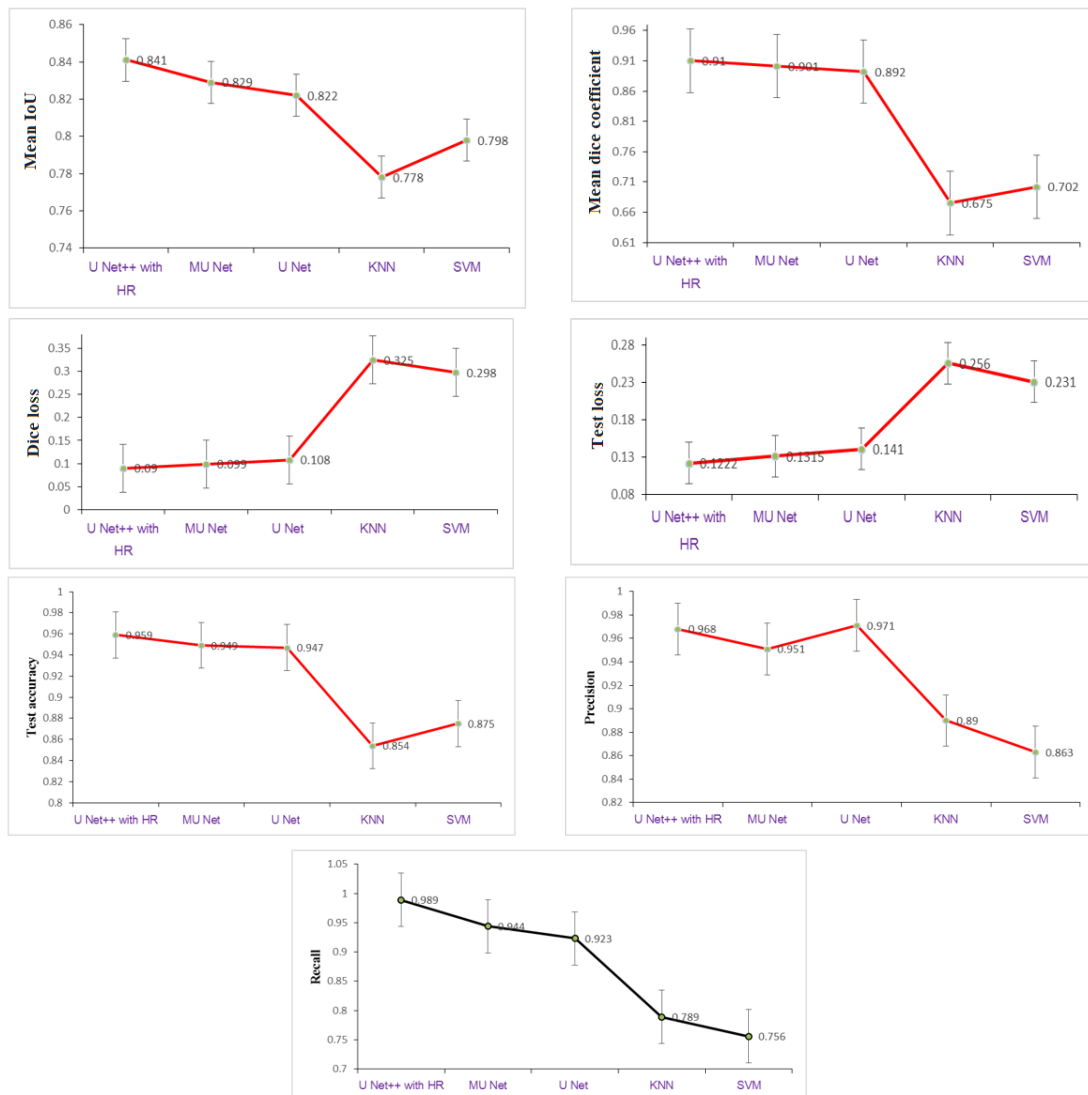

Figure 5. U net ⁺⁺, Modified U performance parameter comparison with U net [36], KNN [37] & SVM [24].

Above are some common performance metrics for segmentation, as illustrated in Figure 5. We can obtain insights into how effectively the model performs on the segmentation work by analyzing these performance characteristics and identifying areas for improvement, such as classes with lower IOU or Dice coefficient scores. These indicators are critical for evaluating the segmentation model's efficacy and directing the model's development and fine-tuning.

### 5.4. Observational Results on Training Dataset

In Figure 6 (a), the simulation results of the proposed segmentation method during the model training are presented. The comparison involves the input image being compared with both the ground truth image and the predicted mask for segmentation.
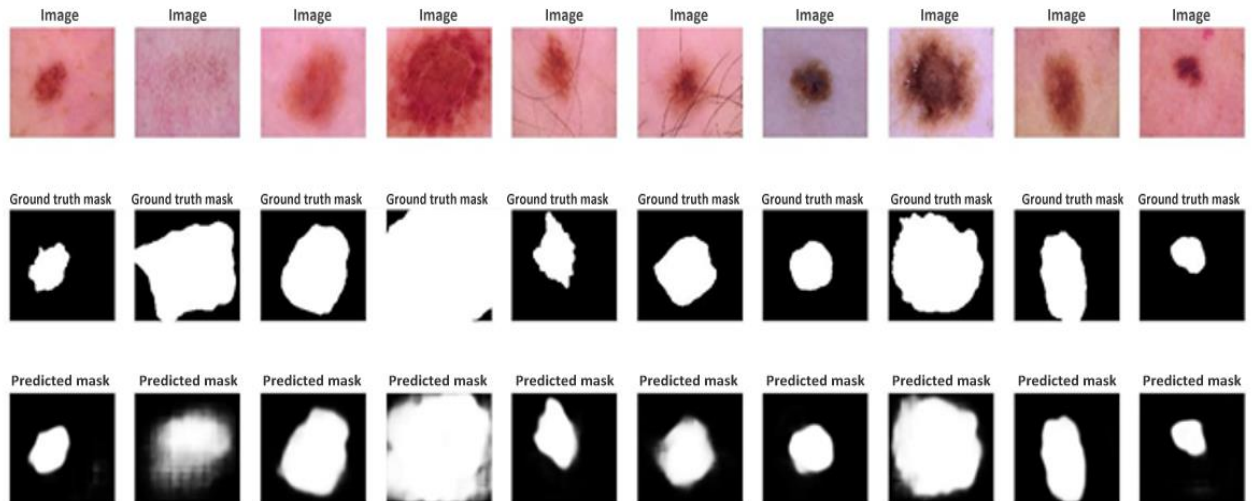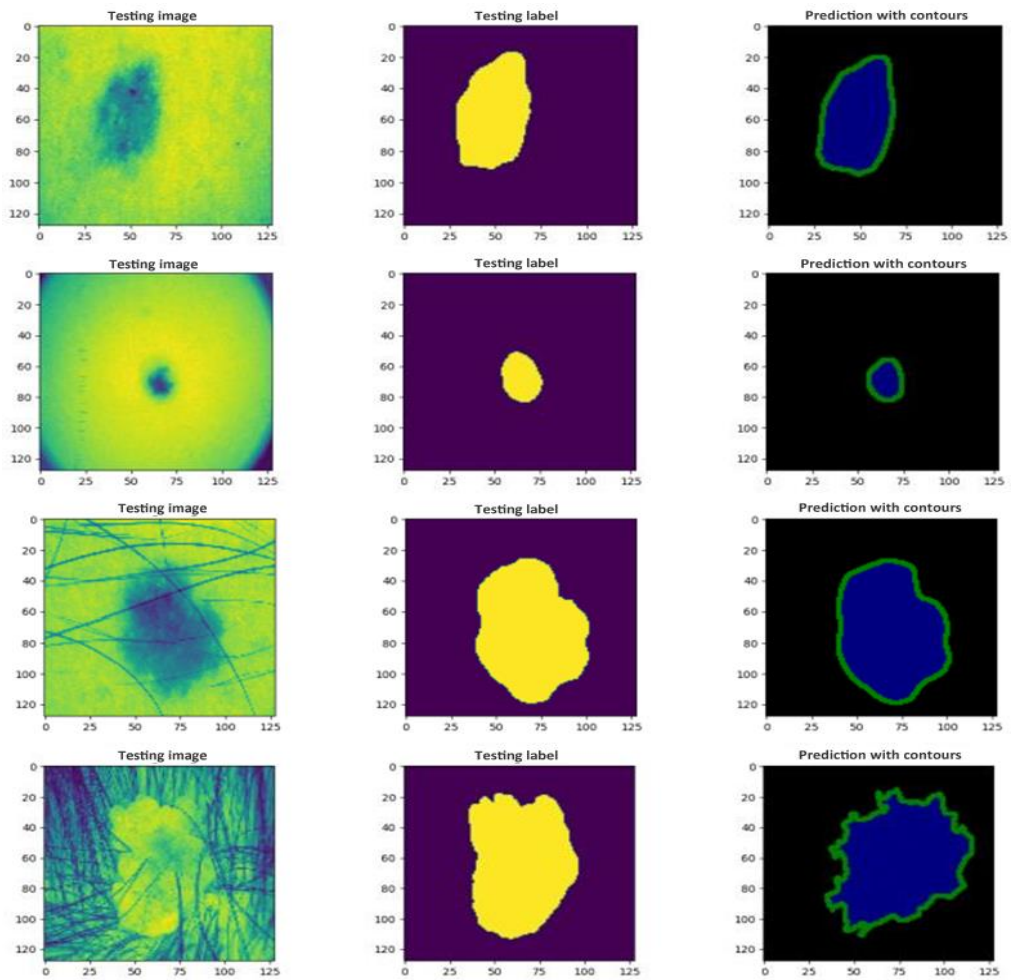

(a): U Net ++ Training images with their ground truth masks and predicted masks.
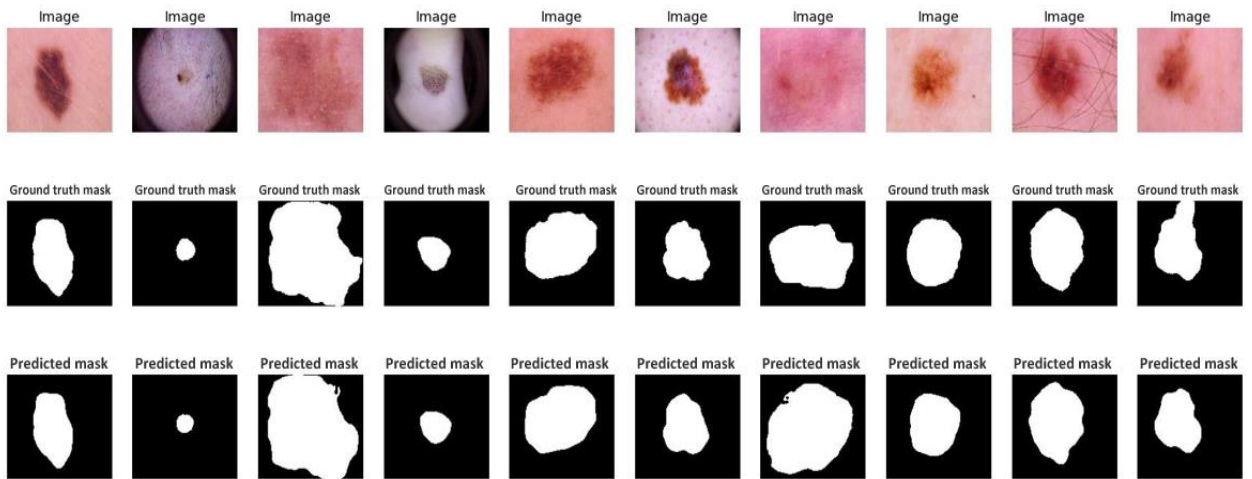
(b): U Net ++ Testing images with their ground truth masks and predicted masks.
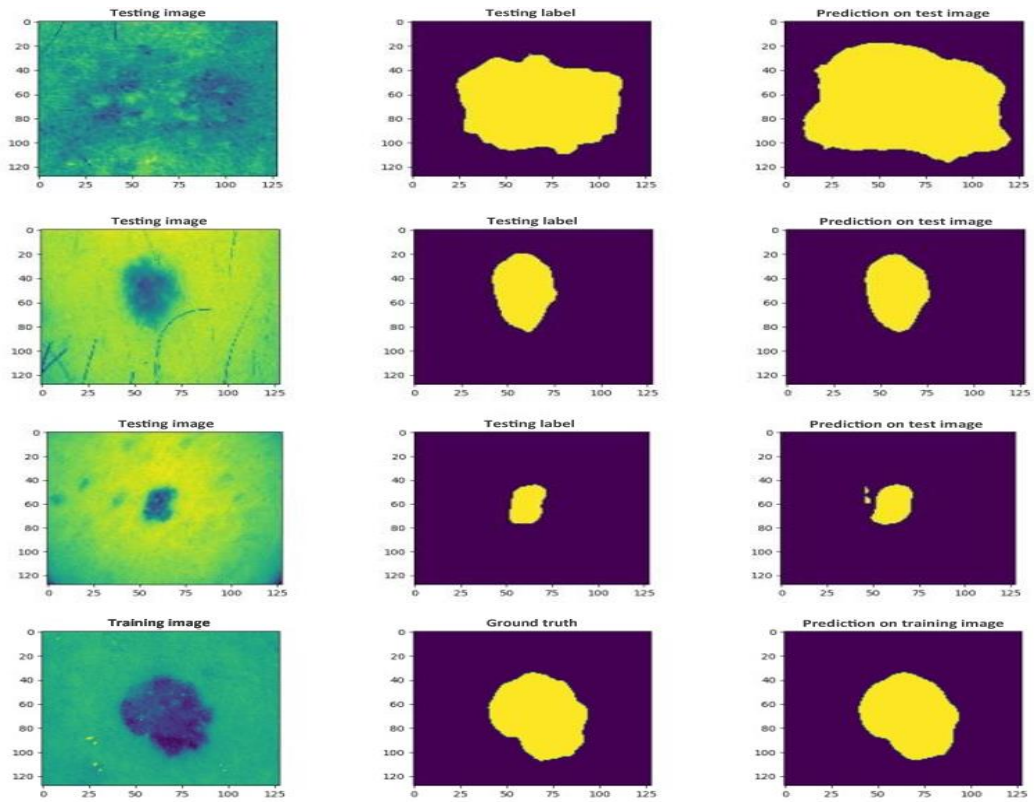
(c): Prediction on testing images without contours.

(d): Prediction on Testing images with contours.

Figure 6. U Net ++ Training, testing & prediction simulation results.

In Figure 6 (b), the simulation results of the proposed segmentation method during model testing are showcased. The comparison entails presenting the input image alongside both the ground truth image and the predicted mask for segmentation.

Figure 6 (c), illustrates the proposed model segmentation results without contour prediction. It may result in imprecise delineation of melanoma boundaries. This lack of precise boundary information can affect the accuracy of lesion size, shape, and margin assessment, which is crucial for diagnosis and treatment planning. Also, it is hard to accurately capture Melanoma lesions irregular and complex shapes. Also, pixel-level features for segmentation may not capture all the discriminative characteristics of melanoma.

Figure 6 (d) shows the proposed model segmentation of testing images with contours. Predicting contours makes it possible to keep track of how the lesion's form and size vary over time, which is crucial for monitoring the melanoma's growth or retreatment, which helps medical professionals decide on the best course of action. The boundaries of melanoma lesions can be properly detected and defined using contour prediction, which also it helps to differentiate between the melanoma lesion and surrounding healthy skin tissue.

6. CONCLUSION & FUTURE WORK

To conclude, this study introduces a highly resilient deep learning network tailored for segmentation purposes. This network holds the promise of delivering heightened accuracy compared to alternative models, particularly when confronted with the intricate task of delineating skin lesion regions, even in scenarios where hair is present in the input images. Mean Intersection Over Union (IOU) evaluates overlapping between the predicted and ground truth masks.

By using U Net $^{++}$ we get a maximum IoU of 0.841 (+0.012 mod U Net, +0.019 U Net, 0.063 KNN, +0.043 SVM), which is better than all existing and modified methods. The mean Dice coefficient computes the Dice coefficient that is averaged over all samples in the dataset. It returns a single result that represents the model's overall segmentation performance. Maximum Dice coefficient 0.910 is obtained from the U-net [proposed] method. Figure 4 shows that all performance parameter behavior is efficient for U Net, Modified U-Net, and U Net++ with hair removal methods.

The hair removal preprocessing step eliminates the interference that hair regions, which are frequently present in medical images, cause. By explicitly identifying and removing these regions, we reduce noise and increase the accuracy of the segmentation.

To avoid overfitting and select the best-performing model, we implement early stopping and model checkpointing. The hierarchical feature extraction improves the model's ability to segment complex structures accurately. Nested skip connections are a key architectural feature of the U Net++ model that enhances its ability to capture hierarchical and multi-scale features in an image. These connections allow the model to fuse information from different levels of the encoding and decoding paths, leading to improved feature representation and segmentation accuracy.

The U Net++ architecture enables the model to perform more comprehensive feature aggregation and achieve improved segmentation performance, especially when dealing with objects that vary in size or are present at different scales within the image.

REFERENCES

[1]   World Cancer Reasearch Fund International, "Skin cancer statistics," Retrieved: https://www.wcrf.org/cancer-trends/skin-cancer-statistics/. 2021.

[2]   Z. A. Nazi and A. A. Tasnim, "Automatic skin lesion segmentation and melanoma detection: Transfer learning approach with u-net and dcnn-svm," in *Proceedings of International Joint Conference on Computational Intelligence: IJCCI 2018, Springer Singapore*, 2020.

[3]   S. M. Thomas, J. G. Lefevre, G. Baxter, and N. A. Hamilton, "Interpretable deep learning systems for multi-class segmentation and classification of non-melanoma skin cancer," *Medical Image Analysis*, vol. 68, p. 101915, 2021. https://doi.org/10.1016/j.media.2020.101915

[4]   M. Nawaz *et al.*, "Skin cancer detection from dermoscopic images using deep learning and fuzzy k-means clustering," *Microscopy Research and Technique*, vol. 85, no. 1, pp. 339-351, 2022. https://doi.org/10.1002/jemt.23908

[5]   M. Tehseen *et al.*, "The role of machine learning and deep learning approaches for the detection of skin cancer," *Healthcare, MDPI*, vol. 11, no. 3, p. 415, 2023. https://doi.org/10.3390/healthcare11030415

[6]   S. S. Devi, N. H. Singh, and R. H. Laskar, "Fuzzy C-means clustering with histogram based cluster selection for skin lesion segmentation using non-dermoscopic images," *International Journal of Interactive Multimedia & Artificial Intelligence*, vol. 6, no. 1, pp. 26-31, 2020. https://doi.org/10.9781/ijimai.2020.01.001

[7]   J. Ramya, H. Vijaylakshmi, and H. M. Saifuddin, "Segmentation of skin lesion images using discrete wavelet transform," *Biomedical Signal Processing and Control*, vol. 69, p. 102839, 2021. https://doi.org/10.1016/j.bspc.2021.102839

[8]   G. B. Loganathan *et al.*, "Melanoma classification using enhanced fuzzy clustering and DCNN on dermoscopy images," *Neuro Quantology*, vol. 12, pp. 196-213, 2022. https://doi.org/10.3390/info8030089

[9]   N. Nida, A. Irtaza, A. Javed, M. H. Yousaf, and M. T. Mahmood, "Melanoma lesion detection and segmentation using deep region based convolutional neural network and fuzzy C-means clustering," *International Journal of Medical Informatics*, vol. 124, pp. 37-48, 2019. https://doi.org/10.1016/j.ijmedinf.2019.01.005

[10]   A. Chakkaravarthy, Prabhu, and A. Chandrasekar, "Automatic detection and segmentation of melanoma using fuzzy c-means," presented at the 2019 Fifth International Conference on Science Technology Engineering and Mathematics (ICONSTEM), IEEE, 2019.

[11]   T. Saba, "Computer vision for microscopic skin cancer diagnosis using handcrafted and non-handcrafted features," *Microscopy Research and Technique*, vol. 84, no. 6, pp. 1272-1283, 2021. https://doi.org/10.1002/jemt.23686

[12]   A. Murugan, S. A. H. Nair, A. A. P. Preethi, and K. S. Kumar, "Diagnosis of skin cancer using machine learning techniques," *Microprocessors and Microsystems*, vol. 81, p. 103727, 2021. https://doi.org/10.1016/j.micpro.2020.103727

[13]   B. Sreedhar, S. B. Manjunath, and K. M. Sunil, "A comparative study of melanoma skin cancer detection in traditional and current image processing techniques," presented at the 2020 Fourth International Conference on I-SMAC (IoT in Social, Mobile, Analytics and Cloud)(I-SMAC), IEEE, 2020.

[14]   L. Liu, Y. Y. Tsui, and M. Mandal, "Skin lesion segmentation using deep learning with auxiliary task," *Journal of Imaging*, vol. 7, no. 4, p. 67, 2021. https://doi.org/10.3390/jimaging7040067

[15]   Q. Ding and N. Razmjooy, "An optimal diagnosis system for melanoma dermoscopy images based on enhanced design of horse herd optimizer," *International Journal of Imaging Systems and Technology*, vol. 33, no. 3, pp. 1092-1107, 2023. https://doi.org/10.1002/ima.22852

[16]   A. Daniel, S. Venkatraman, S. Doss, B. C. Balusa, A. Maseleno, and K. Shankar, "IoT-based automated skin lesion detection and classification using gray Wolf optimization with deep Neural Network," *Artificial Intelligence Techniques in IoT Sensor Networks*, pp. 197-212, 2020.

[17]   F. Bagheri, M. J. Tarokh, and M. Ziaratban, "Skin lesion segmentation from dermoscopic images by using Mask R-CNN, Retina-Deeplab, and graph-based methods," *Biomedical Signal Processing and Control*, vol. 67, p. 102533, 2021. https://doi.org/10.1016/j.bspc.2021.102533

[18] M. A. Khan, Y.-D. Zhang, M. Sharif, and T. Akram, "Pixels to classes: Intelligent learning framework for multiclass skin lesion localization and classification," *Computers & Electrical Engineering*, vol. 90, p. 106956, 2021. https://doi.org/10.1016/j.compeleceng.2020.106956

[19] S. Qamar, P. Ahmad, and L. Shen, "Dense encoder-decoder–based architecture for skin lesion segmentation," *Cognitive Computation*, vol. 13, no. 2, pp. 583-594, 2021. https://doi.org/10.1007/s12559-020-09805-6

[20] M. M. Stofa, M. A. Zulkifley, M. A. A. M. Zainuri, and A. A. Ibrahim, "U-net with atrous spatial pyramid pooling for skin lesion segmentation," in *Proceedings of the 6th International Conference on Electrical, Control and Computer Engineering: InECCE2021, Kuantan, Pahang, Malaysia, 23rd August. Singapore: Springer Singapore*, 2022.

[21] H. Wu, J. Pan, Z. Li, Z. Wen, and J. Qin, "Automated skin lesion segmentation via an adaptive dual attention module," *IEEE Transactions on Medical Imaging*, vol. 40, no. 1, pp. 357-370, 2020. https://doi.org/10.1109/tmi.2020.3027341

[22] H. Wu, S. Chen, G. Chen, W. Wang, B. Lei, and Z. Wen, "FAT-Net: Feature adaptive transformers for automated skin lesion segmentation," *Medical Image Analysis*, vol. 76, p. 102327, 2022. https://doi.org/10.1016/j.media.2021.102327

[23] P. Thapar, M. Rakhra, G. Cazzato, and M. Hossain, "A novel hybrid deep learning approach for skin lesion segmentation and classification," *Journal of Healthcare Engineering*, vol. 2022, p. 1709842, 2022.

[24] R. Seeja and A. Suresh, "Deep learning based skin lesion segmentation and classification of melanoma using support vector machine (SVM)," *Asian Pacific Journal of Cancer Prevention*, vol. 20, no. 5, p. 1555, 2019.

[25] L. Liu, L. Mou, X. X. Zhu, and M. Mandal, "Automatic skin lesion classification based on mid-level feature learning," *Computerized Medical Imaging and Graphics*, vol. 84, p. 101765, 2020. https://doi.org/10.1016/j.compmedimag.2020.101765

[26] Y. Li and L. Shen, "Skin lesion analysis towards melanoma detection using deep learning network," *Sensors*, vol. 18, no. 2, p. 556, 2018. https://doi.org/10.3390/s18020556

[27] V. K. Singh *et al.*, "FCA-Net: Adversarial learning for skin lesion segmentation based on multi-scale features and factorized channel attention," *IEEE Access*, vol. 7, pp. 130552-130565, 2019. https://doi.org/10.1109/access.2019.2940418

[28] X. Yang, Z. Zeng, S. Y. Yeo, C. Tan, H. L. Tey, and Y. Su, "A novel multi-task deep learning model for skin lesion segmentation and classification," *arXiv preprint arXiv:1703.01025*, 2017.

[29] Y. Xie, J. Zhang, Y. Xia, and C. Shen, "A mutual bootstrapping model for automated skin lesion segmentation and classification," *IEEE Transactions on Medical Imaging*, vol. 39, no. 7, pp. 2482-2493, 2020. https://doi.org/10.1109/tmi.2020.2972964

[30] Y. Yuan, "Automatic skin lesion segmentation with fully convolutional-deconvolutional networks," *arXiv preprint arXiv:1703.05165*, 2017.

[31] N. Ibtehaz and M. S. Rahman, "MultiResUNet: Rethinking the U-Net architecture for multimodal biomedical image segmentation," *Neural networks*, vol. 121, pp. 74-87, 2020. https://doi.org/10.1016/j.neunet.2019.08.025

[32] P. F. Christ *et al.*, "Automatic liver and lesion segmentation in CT using cascaded fully convolutional neural networks and 3D conditional random fields," in *In Proceedings of the International Conference on Medical Image Computing and Computer-Assisted Intervention, Athens, Greece, 17–21 October 2016; Springer: Cham, Switzerland*, 2016, pp. 415–423.

[33] M. Dash, N. D. Londhe, S. Ghosh, A. Semwal, and R. S. Sonawane, "PsLSNet: Automated psoriasis skin lesion segmentation using modified U-Net-based fully convolutional network," *Biomedical Signal Processing and Control*, vol. 52, pp. 226-237, 2019. https://doi.org/10.1016/j.bspc.2019.04.002

[34] M. Nawaz *et al.*, "Melanoma segmentation: A framework of improved DenseNet77 and UNET convolutional neural network," *International Journal of Imaging Systems and Technology*, vol. 32, no. 6, pp. 2137-2153, 2022. https://doi.org/10.1002/ima.22750

[35] X. Wu, D. Hong, and J. Chanussot, "UIU-Net: U-Net in U-Net for infrared small object detection," *IEEE Transactions on Image Processing*, vol. 32, pp. 364-376, 2022. https://doi.org/10.1109/tip.2022.3228497

[36]   R. L. Araújo, F. H. D. Araújo, and R. R. E. Silva, "Automatic segmentation of melanoma skin cancer using transfer learning and fine-tuning," *Multimedia Systems,* vol. 28, no. 4, pp. 1239-1250, 2022. https://doi.org/10.1007/s00530-021-00840-3

[37]   A. Murugan, S. Anu, H. Nair, and K. Sanal Kumar, "Detection of skin cancer using SVM, random forest and kNN classifiers," *Journal of Medical Systems,* vol. 43, pp. 1-9, 2019.