Journal of Information Systems Engineering and Management
2025, 10(20s)
e-ISSN: 2468-4376
https://www.jisem-journal.com/

Research Article

Impact of Skin Lesion Descriptors and Deep Learning Architecture for the Effective Detection and Identification of Skin Diseases - A Systematic Review

Kavitha B[1], Dr. Kusuma Kumari B M[2]
[1]*Research Scholar, Department of Studies and Research in Computer Applications,Tumkur University, Tumakuru, Karnataka*
[2]*Assistant Professor, Department of Studies and Research in Computer Applications,Tumkur University, Tumakuru, Karnataka*
[1]*kavithab@tumkuruniversity.ac.in*, [2]*kusuma_bm@tumkuruniversity.ac.in*

ARTICLE INFO

Received: 21 Dec 2024

Revised: 27 Jan 2025

Accepted: 15 Feb 2025

ABSTRACT

In diagnosing skin diseases, segmentation of lesion region and classification of detected lesion type are the two major processes. This paper conducts a systematic review related to the lesion descriptors extracted from the skin region and the machine learning classifiers utilized in differentiating the descriptor types and also discusses the deep learning schemes that impact skin lesion diagnosis. More specifically, the paper focuses on lesion region detection-based deep learning schemes that are derived from the Mask Region-based CNN (RCNN), U-Net, and DeepLabV3+ architectures. In the case of skin lesion segmentation, the recent multi-attention scheme results in a recall, precision, and F1-score of 93.97%, 93.94%, and 93.73% respectively in the ISIC-2016 dataset which is higher than other lesion detection approaches. In the case of skin lesion classification, the Diverse CNN approach results in a maximum accuracy, recall, and precision of 96.12%, 93.11%, and 94.63% respectively when evaluated using the HAM-10000 dataset which is higher than other lesion classification approaches.

Keywords: Skin lesion detection, deep learning, lesion descriptors, Mask RCNN, DeepLab V3+.

1.Introduction

The skin lesion [1] is an abnormal appearance or growth of the skin region when compared to normal skin. The skin lesions can be classified as primary and secondary types [2]. The skin lesion that presents at birth, and can grow over time is termed to be primary type. The major reason for skin cancer death is due to delayed diagnosis, and treatment, A survey reports [3] that 2 to 3 million non-melanoma type skin lesion subjects are diagnosed across the globe, however only 130,000 melanoma-type subjects are diagnosed. Fig.1 shows the death rate [4] in the US for the year between 1992 and 2021. Early detection and treatment is an effective way of reducing skin cancer deaths. Due to the development of computer-aided medical diagnosis algorithms, early skin lesion detection is possible which increases the chance of early-stage treatment.

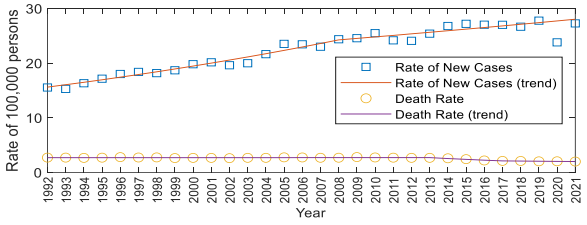
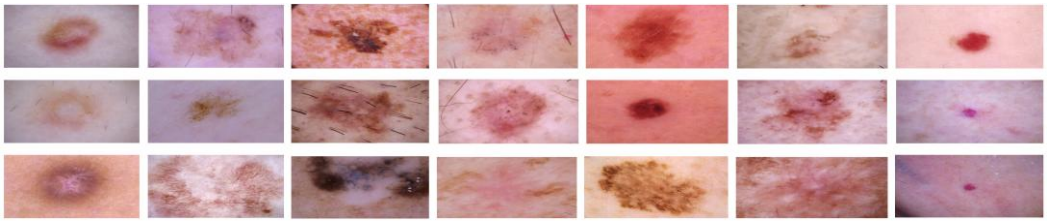


Fig.1: Death rate in the US for the years 1992 to 2021

Benign and Malignant type lesions [5] are another classification of lesion types. The benign skin lesions are non-cancerous, and their growth does not spread to other normal skin surfaces. The malignant type skin lesions are cancerous [6], and they can grow, invade nearby skin tissues, and spread to other body parts. The growth of such

lesions is abrupt and can change over time. The lesions region can be itchy and can bleed or ulcerated. Such lesions require treatments like targeted therapy (radiation), chemotherapy, or surgical removal.



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

Fig. 2: Dermatofibroma (b) Benign Keratosis (c) Melanoma (d) Basal cell carcinoma (e) Melanocytic Nevi (f) Actinic Keratoses (g) Vascular lesions

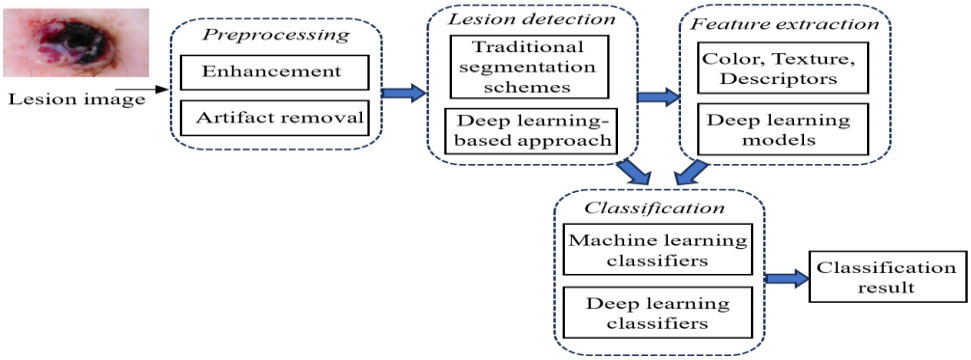


Fig.3: Representation of stages involved in detecting and identifying the lesion types

The dermoscopic images of different lesion types are illustrated in Fig. 2. approaches and classification of lesion types as illustrated in Fig. 3.

The contributions of the review paper are as follows,

(i) This study is on artifacts present in the skin lesion dermoscopic images and preprocessing approaches used for the removal of such artifacts.
(ii) The paper analyses different datasets to detect lesion and classify them.
(iii) The paper discusses about deep learning schemes to detect the tumor region.
(iv) Finally, the deep learning-based lesion classification approaches are discussed and analyzed with their performance measures.

2.Skin lesion artifact removal Enhancement Methods

This section discusses the different schemes that are used as preprocessing for the removal of different artifacts and to enhance the lesion images along with various artifacts.

2.1Artifacts in skin lesion images

The dermoscopic skin images contain different artifacts [7] that makes the detection of lesion regions or extraction of lesion features more challenging as discussed below in fig[4]

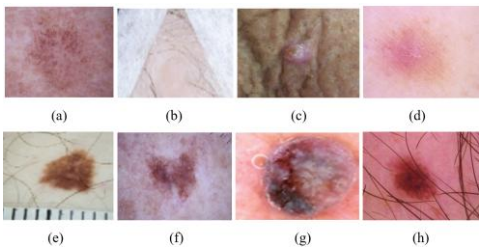


Fig. 4: Representation of different artifacts (a) Low contrast boundaries (b) Frame artifact (c) Wrinkle artifact (d) Irregular boundaries (e) Background artifact (f) Blood vessel artifact (g) Bubble (h) Hairs

2.2Approaches for the removal of lesion artifacts

Different artifact removal approaches and dermoscopic image enhancement approaches have been used in recent algorithms.

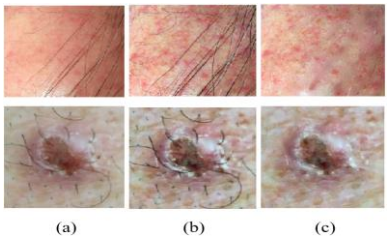


Fig. 5 Sample pre-processed results

Schemes such as DullRazor [8], Multi-scale decomposition [9], enhanced DullRazor [10], Frangi vesselness filter [11], fast line detector [12], Averaging filter, Threshold decomposition [13], and Morphological operation [14] are used in the removal of artifacts such as blood vessels, and hairs. Algorithms such as standard deviation normalization [15], Z-score transform [16], contrast stretching [17], log transform [18], Top bottom filtering [19], Contrast Limited Adaptive Histogram Equalization (CLAHE) [20], Gamma Correction [21], and Median filter are used for the enhancement of lesion region. Fig. 5 illustrates the combination of CLAHE lesion enhancement followed by the DullRazor approach for the removal of hair artifacts.

3.Datasets and evaluation measures

3.1Datasets: Various datasets are employed for the task of skin lesion detection and classification. ISIC-2016 has 899dermoscopicimages.

ISIC-2017 features 2000+ images distributed over three types of lesions.

ISIC-2018 or HAM 10000, contains 10,015 images spread over seven classes of lesion.

ISIC-2019expands further, with 25,331images and other squamous cell carcinoma types.

ISIC-2020 includes six classes of lesion and 27,124 images of unknown type. Other datasets include DermQuest (137 images), DermIS (69 images), PH2 (200 images with segmentation masks), and PAD-UFES-20 (2298 images across six lesion types). These represent rich benchmarks for deep learning-based skin lesion analysis. [27], [28],[29],[30],[31],[32],[33],[34]

Table I: Number of images in each dataset

Lesion type	**ISIC 2016**	**ISIC 2017**	**ISIC 2018**	**ISIC 2019**	**ISIC 2020**	**DERM QUEST**	**DERMIS**	**PH²**	**PAD-UFES-20**
Melanoma	173	374	1113	4522	584	76	43	40	52
Atypical Nevus	726	1372	6705	12875	46	61	26	80	244
Common nevus	-	-	-	-	5193	-	-	80	-
Dermatofibroma	-	-	115	239	-	-	-	-	-
Seborrheic Keratosis	-	254	1099	2624	135	-	-	-	235
Basal cell carcinoma	-	-	514	3323	7	-	-	-	845
Squamous cell carcinoma	-	-	-	628	-	-	-	-	192
Pyogenic Granuloma	-	-	-	-	-	-	-	-	-
Hemangioma	-	-	327	867	37	-	-	-	-
Actinic Keratosis	-	-	142	253	-	-	-	-	730

Unknown	-	-	-	-	27124	-	-	-	-
Total number of images	899	2000	10015	25331	33126	137	69	200	2298

Table I illustrates the number of images in each lesion class for different datasets that are commonly used for analysis. A few of the sample images from the datasets are illustrated in Fig. 6.

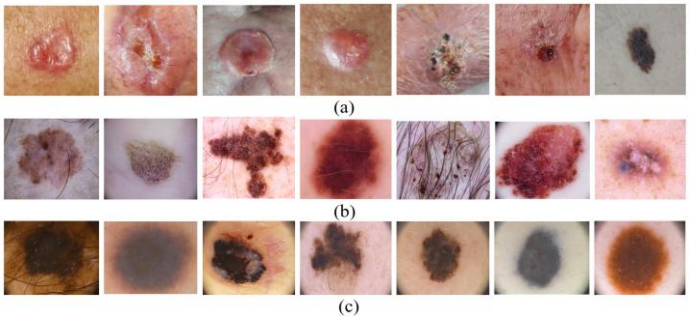


Fig.6: Sample images from the publicly available datasets (a) PAD-UFES-20 dataset (b) HAM-10000 dataset (c) PH^2 dataset

3.2 Evaluation measures

For the evaluation of the lesion detection and classification schemes, evaluation measures such as precision, recall, accuracy, specificity, and F1-score are commonly used as illustrated in Table II.

Table II: Evaluation metrics used for the analysis of lesion segmentation and classification

Process	*Precision*	*Recall*	*Accuracy*	*Specificity*	*F1 – Score*
Segmentation	$S_{pr} = \frac{S_{TP}}{S_{FP} + S_{TP}}$	$S_{re} = \frac{S_{TP}}{S_{FN} + S_{TP}}$	$S_{ac} = \frac{S_{TN} + S_{TP}}{S_{FP} + S_{TP} + S_{FN} + S_{TN}}$	$S_{sp} = \frac{S_{TN}}{S_{FN} + S_{FP}}$	$S_{F1} = \frac{S_{TP}}{\frac{1}{2}(S_{FN} + S_{FP}) + S_{TP}}$
Classification	$C_{pr} = \frac{C_{TP}}{C_{FP} + C_{TP}}$	$C_{re} = \frac{C_{TP}}{C_{FN} + C_{TP}}$	$C_{ac} = \frac{C_{TN} + C_{TP}}{C_{FP} + C_{TP} + C_{FN} + C_{TN}}$	$C_{sp} = \frac{C_{TN}}{C_{FN} + C_{FP}}$	$C_{F1} = \frac{C_{TP}}{\frac{1}{2}(C_{FN} + C_{FP}) + C_{TP}}$

In these equations S_{FP}, S_{TP}, S_{TN} and S_{FN} resemble the false positives, true positives, true negatives, and false negatives results obtained during segmenting. In classification, the terms C_{FP}, C_{TP}, C_{TN} and C_{FN} are estimated based on the classified result and the annotated lesion classes.

4.Detection of skin lesion region: Skin lesion detection schemes can broadly classified

4.1 Traditional Segmentation Approaches - The traditional approach uses schemes such as an edge detection-based approach [35], a region-based approach [36], and a threshold-based approach [37]. Clustering algorithms such as K-means [38], and fuzzy C-means [39] are also commonly used in differentiating the background skin region and foreground lesion region. The thresholding-based approach includes Adaptive thresholding [40], Otsu's thresholding [41], Renyi's entropy [42], etc. The region-based approach segments the group of related pixels to the lesion or skin region. The region-based approach includes region growing [43], the Mumford-Shah approach [44], Splitting, and merging approach [45]. Region growing approach includes the contour-based model such as active contour models. Different masks such as Canny, Laplacian, Robert, Sobel, and Prewitt operators [46] are utilized to detect the edge between the lesion and the skin surface.

4.2 Deep learning approach to detect lesion region

Few deep learning structures that provide better segmentation results are discussed below

(a) U-Net based approaches-The U-Net architecture [47] contains a bottleneck layer between the contraction path and the expansion path. The spatial dimension gets reduced while the feature channel increases in the contraction path. The attention U-Net [48] utilizes attention gates (AG) that give more weightage to the relevant area (lesion region), and less weightage to the irrelevant area (normal skin region) as illustrated in Fig. 7. The attention weights β_0 is estimated as, $\beta_0 = ReLU(c_g \times f_d + c_w \times f_e + c_b)$ (1)

Here, c_g, and c_w resembles the learnable weights The term f_d, and f_e resembles the features of encoder, and decoder, while c_b resembles the bias. Thus, the attention map is computed as,

$$\beta_1 = sigmoid(\beta_0) = \frac{1}{1+e^{-\beta_0}} \quad (2)$$

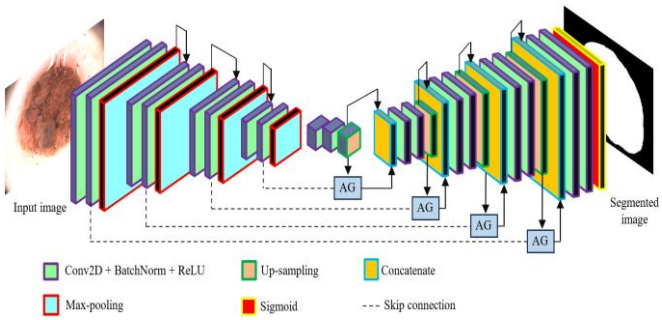


Fig.7: Attention U-Net structure in detecting lesion region

Xia et al. [49] combined U-Net with a dual discriminator to create a GAN using dilated convolution for fine-grain feature extraction. Multi-scale attention [50] in U-Net enhances feature aggregation while suppressing redundancy with bidirectional ConvLSTM. Variants like Psoriasis U-Net (29 layers), Anti-aliasing U-Net [52], and U-Net++ [54] improve spatial data collection, shift equivariance, and gradient disappearance issues. Combining CNN with U-Net achieves 97.96% accuracy on HAM-10000 for lesion classification and detection.

(b) Mask-RCNN based approach-The mask RCNN [55, 56] architecture can able to perform two simultaneous tasks such as skin lesion region detection as well as lesion type classification. The major components in mask RCNN include a backbone network for feature extraction, a region proposal network (RPN) to generate anchors. The mask RCNN also has mask prediction to construct the mask for the region that was detected as depicted in Fig. 8.

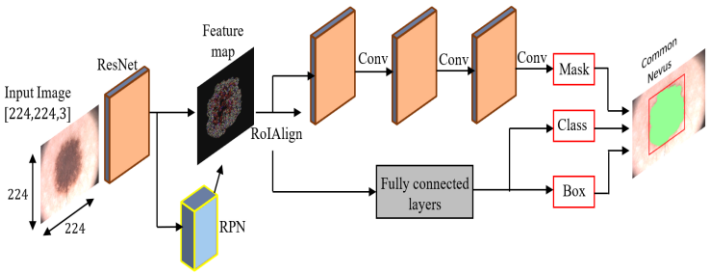


Fig.8: Mask R-CNN structure in detecting lesion region and identification of lesion type

The author Bagheri et al. combined Mask R-CNN structure [57] with DeepLab to obtain two different segmented lesion regions. Transfer learning architecture was deployed in mask R-CNN [58], in which the mask R-CNN segmented lesion is fed to DenseNet architecture to collect the descriptors. This approach results in 92.7%, and 93.6% accuracy when analyzed using ISIC-2017, and ISIC-2016 datasets.

(c) DeepLabV3+ based approaches -The DeepLabV3+ [59] has two major blocks namely encoder, and decoder. The descriptors are dilated at different levels to obtain multi-scale descriptors. The decoder refines the descriptors by performing up-sampling that improves the accuracy of the segmented lesion region as depicted in Fig. 9.

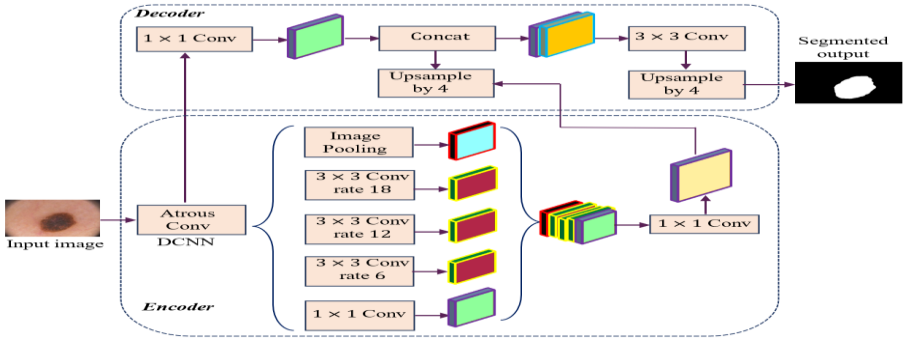


Fig.9: DeepLabV3+ structure in the detection of lesion region

A three-step process was proposed by Bagheri et al. [60] for segmenting the lesion region. The DeepLabV3+ architecture uses the MobileNetV2 model [61] to detect the lesion. This approach can simultaneously classify, and detect the lesions. Saeed et al. used DeepLabV3+ along with VGG-19, and VGG-16 architectures to develop generative AI-based architecture [62]. This architecture when evaluated using the ISIC-2018 dataset results in 97% segmentation accuracy. The actual classification was done by a vision transformer [63] that has 8 blocks, where each block has 8 multi-head attention. This hybrid architecture results in an accuracy of 97.73%, and 96.97% when evaluated using the ISIC-20, and ISIC-19 dataset.

4.3 Performance analysis

The performance of the deep learning-based lesion detection algorithms was evaluated using measures such as recall, precision, and F1-score using the datasets namely PH^2, ISIC-2016, and ISIC-2017 datasets.

Table III: Comparison of recall, precision, and F1-score in segmenting the lesion region

Method	PH^2			ISIC-2016			ISIC-2017		
	Recall	Precision	F1	Recall	Precision	F1	Recall	Precision	F1
U-Net [64]	88.51	84.43	86.43	91.17	90.48	88.66	70.18	89.66	78.73
Attention U-Net++ [65]	93.05	91.66	92.35	90.68	92.94	88.34	74.34	92.54	80.66
U-Net++ [66]	93.74	93.40	93.57	87.32	93.77	90.20	70.70	88.15	80.16
DCSAU-Net [67]	93.23	95.00	94.11	91.42	91.32	92.72	88.01	83.93	85.93
CASF-Net [68]	93.78	95.44	94.60	92.26	92.12	91.46	84.51	85.14	84.20
MS-Net [69]	94.18	94.93	94.55	91.54	91.36	89.40	76.79	91.07	83.31
Multi-attention [70]	94.11	95.97	95.03	93.97	93.94	93.73	91.22	93.92	87.15

Table III illustrates the performance comparison of different deep learning architectures in detecting the lesion region. The comparison was made between the architectures U-Net [64], attention U-Net++ [65], U-Net++ [66], DCSAU-Net [67], CASF-Net [68], MS-Net [69] and multi-attention [70]. The recall, precision, and F1-score are higher for the multi-attention approach than other deep learning architectures. The recall, precision, and F1-score were estimated as 94.11%, 95.97%, 95.03% for the multi approach when evaluated using the PH^2 dataset. The precision and recall estimated using Mult-attention architecture is 0.53% and 0.43% higher than the CASF-Net [68] approach. But the recall of multi-attention approach is 0.07% less than the MS-Net [69] approach. In case of ISIC-2016 dataset the recall, precision and F1-score was estimated as 93.97% 93.94% and 93.73% respectively using the Multi-attention approach. The same approach results in recall, precision and F1-score of 91.22%, 93.92% and 87.15% when evaluated using the ISIC-2017 dataset. Fig.10 illustrates the graphical comparison chart for the performance between the deep learning-based lesion detection approaches. In case of ISIC-2016 dataset the recall

and precision estimated using the multi-attention is 1.71% and 0.17% greater than the CASF-Net [80] and U-Net++ [66] approach respectively, while the F1-score is 1.01% higher than the DCSAU-Net architecture [67]. In case of ISIC-2017 dataset, the recall and F1-score estimated using the multi-attention is 3.21% and 1.22% greater than the DCSAU-Net [67] approach, while the precision is 1.38% higher than the attention-U-Net++ architecture [66].

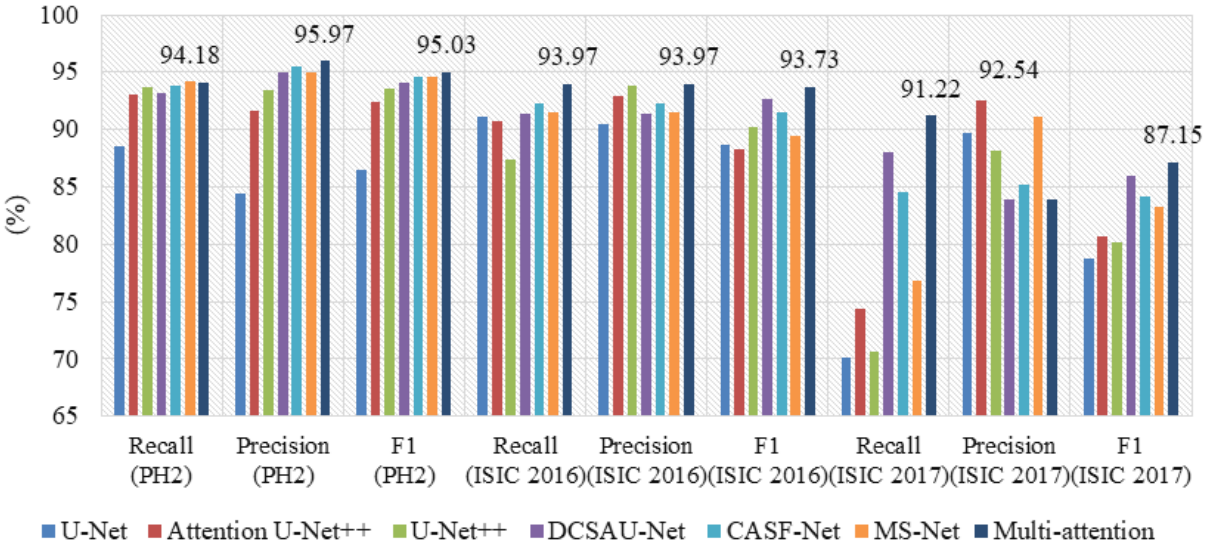


Fig.10: Comparison of performance by different architectures in segmenting the lesion region

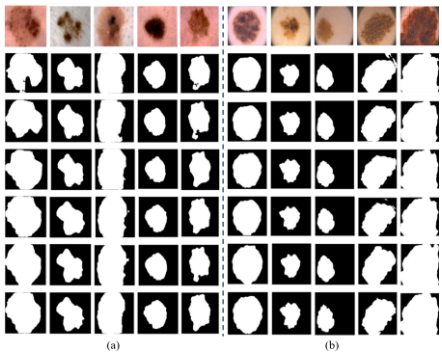


Fig.11: Segmentation result obtained by different deep learning architectures on datasets

5.Lesion descriptor and machine learning approaches: Few of the handcrafted and deep features are provided in Table IV.

Table IV: Accuracy comparison of machine learning classifiers utilizing different descriptors

Type	**Descriptor**	**# Descriptor**	**ISIC-2017**				**ISIC-2016**			
			SVM	**MLP**	**KNN**	**RF**	**SVM**	**MLP**	**KNN**	**RF**
Texture	LBP	256	25.96	19.47	43.67	46.95	44.41	38.47	44.73	48.25
Texture	Moments	4	24.62	22.15	28.84	33.64	24.73	21.28	34.92	41.68
Color	LCH	264	41.38	28.63	48.07	67.04	46.84	38.17	63.66	69.60
Color	CEDD	144	49.98	44.74	64.38	64.63	50.38	45.11	75.09	75.62
Color	GCH	66	40.53	29.32	76.95	84.13	46.31	37.25	78.52	87.44
Color	Haralickcolor	15	34.62	31.51	45.62	69.00	24.40	20.32	40.89	75.53

Texture, color	FCTH	192	45.99	41.42	64.53	67.34	41.46	37.84	71.21	75.09
Texture, color	JCD	168	50.66	44.26	69.56	71.34	50.84	44.72	78.8	79.59
Generic	Xception	2048	61.81	48.77	62.75	55.02	70.90	59.35	76.81	64.93
Generic	InceptionV3	2048	59.95	48.20	63.59	54.04	70.08	58.51	79.83	64.78
Generic	DenseNet121	1024	67.10	56.87	75.85	66.72	72.82	61.72	85.65	75.77
Generic	VGG19	512	62.51	52.23	67.85	59.72	71.33	61.44	84.02	71.46
Generic	VGG16	512	62.98	52.77	69.28	60.80	71.50	60.79	85.34	72.16
Generic	MobileNet	1024	64.67	53.44	73.15	58.43	74.78	63.22	89.03	71.68
Generic	ResNet50	2048	67.06	53.16	74.25	64.68	77.51	64.72	88.80	76.40

Table IV illustrates the accuracy comparison of different machine learning classifiers namely SVM [71], MLP [72], KNN [73], and RF. Texture features such as local binary pattern (LBP) [74] and moments and Color descriptors such as local color histogram (LCH) [75], Color and edge directivity descriptors (CEDD) [76], Global color histogram (GCH) [77], and Haralickcolor features [78] are used for analysis. Also, the hybrid color and texture descriptors such as fuzzy color and texture histogram (FCTH) [79] and joint composite descriptors are used in the analysis. Deep learning-based feature-extracting models [80] are also considered for analysis.

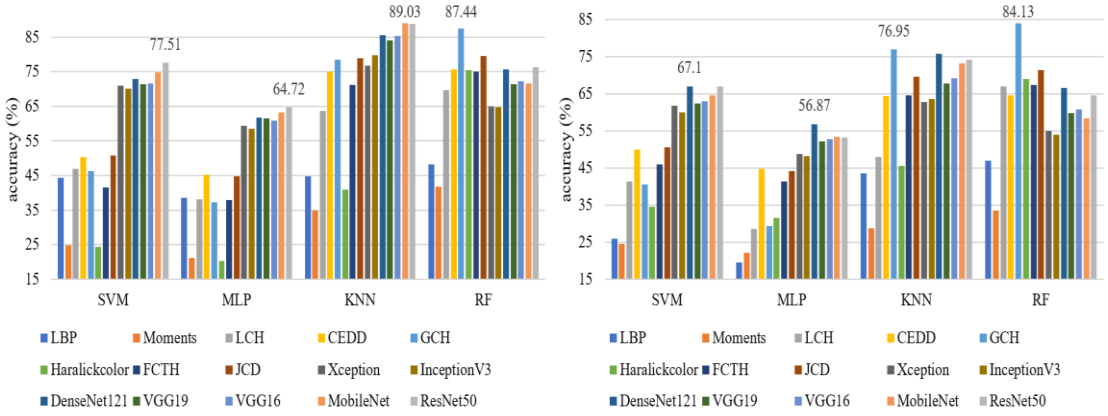


Fig.12: Comparison of classification accuracy for the different machine learning classifiers

Fig.12 shows the graphical illustration of classification accuracy and Among the four classifiers, the SVM classifier provides moderate results when utilizing Generic features extracted by deep learning architectures.

6.Deep learning-based lesion classification approaches: Deep learning approaches that are commonly used in classification of lesion types.

6.1 ResNet based approaches

The ResNet-based architecture [81] minimizes problems like vanishing gradients. The ResNet utilizes the residual blocks that can bypass layers with the use of skip connections as depicted in Fig. 13.

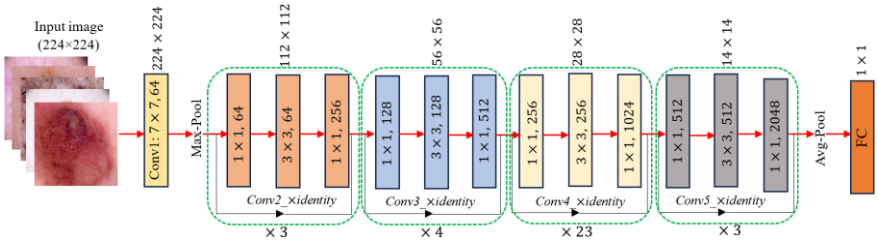


Fig.13: ResNet50 Model in classifying skin lesions

The ResNet architecture was modified to derive multi-channel ResNet [82] that attains 91.7% accuracy in classifying the lesion types such as ResNet18 [83], VGG-16, and AlexNet are used to extract the deep features which are classified using SVM to differentiate the lesion type Seborrheic keratosis, and melanoma. Support vector machine [84] is utilized to train the multiple descriptors collected by inter, and intra architecture. Pre-trained models are used to classify the features from the segmented lesion region which was detected by full-resolution CNN [85]. When comparing EfficientNet-B0, and ResNet-50, models the EfficientNet-B0 [86] shows better performance in classifying the skin lesion. To minimize a residual learning CNN was proposed by Zhang et al. [87].

6.2 EfficientNet based approaches

The EfficientNet [88] architecture provides a higher classification performance. The EfficientNet performs uniform width, depth, and resolution represented as,

$$(s_d, s_w, s_r) = \left(\alpha_1^{\tau} \cdot s_d^{(0)}, \alpha_2^{\tau} \cdot s_w^{(0)}, \alpha_3^{\tau} \cdot s_r^{(0)}\right) \quad (3)$$

Here, α_1, α_2, and α_3 resembles the constant to control the scaling. τ resembles the parameter used to control the network size. $s_d^{(0)}$, $s_w^{(0)}$, and $s_r^{(0)}$ resembles the initial, depth, width, and resolution. These MBConv blocks are represented as the block in EfficientNet architecture provided in Fig. 14. The architecture uses switch activation that scales the input v using the transform to generate the output as.

$$S(v) = sigmoid(v) \times v = \frac{v}{1+e^{-v}} \quad (4)$$

The feature vector is passed through a dense network to estimate the final score.

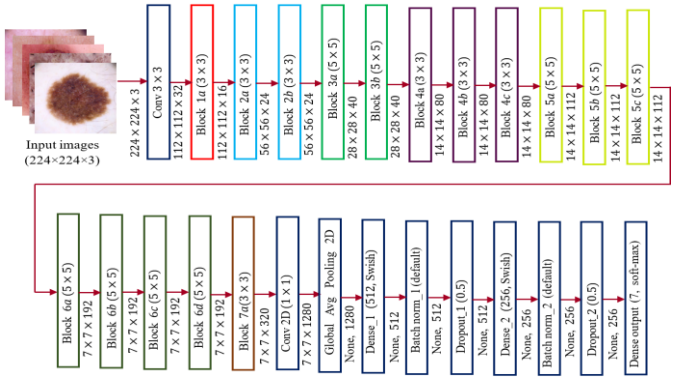


Fig.14: Structure of EfficientNet in the classification of skin lesions

Multi-resolution EfficientNet [89] is derived from the EfficientNet architecture by Alexandar et al. Different EfficientNet models namely EfficientNet B0-B7 [90] are used to fine tune the weights using HAM10000 datasets. The EfficienNetV2-M model is proposed by Stephen et al. [91]. A wide variety of skin lesions are classified by the EfficientNet model [92] by Abdual et al. The architecture EfficientNet results in higher accuracy than ResNet, and VGG networks. Vipin et al. [93] used the regression layer in the EfficientNet instead of using the classification layer. EfficientNet architecture is derived [94] that attains an AUC (area under the ROC curve) of 98.65%

6.3 Vision transformer-based approach

The vision transformer structure [95] initially subdivides an image into non-overlapping patches. The positional encodings along with patches embedding are applied to a transformer encoder that has multi-head attention, feed-forward network, and residual connections as illustrated in Fig. 15. To minimize the class imbalance problem GAN-based vision transformers [96] consist of four major steps Synthetic image generation, augmentation, pattern identification, followed by classification. A deep bottleneck transformer was proposed by Han et al. [97] that extracts deep features and uses these features for global correlation. The dual encoder was further modified to a bidirectional encoder [98] that minimizes the intra-class variation problems caused by inconsistent lesion shapes.

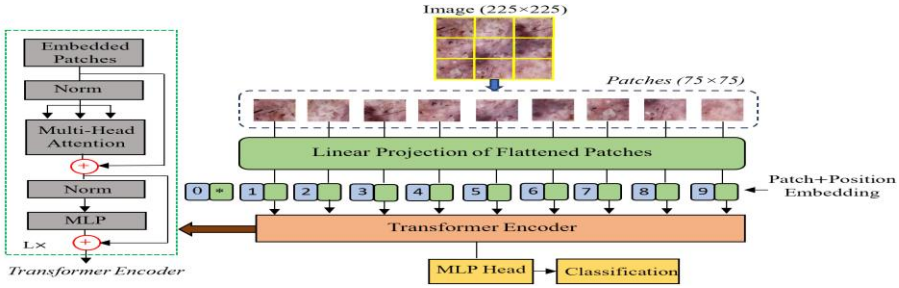


Fig.15: Structure of Vision transformer approach in the classification of skin lesions

The vision transformer in combination with attention-based U-Net termed hybrid TransUNet results in a 92.11% accuracy which is higher than other transformer models. A full transformer network [99] was utilized to extract the contextual features in classifying the skin lesion. An improved transformer was proposed [100] that uses a three-phase process. The traditional loss function in the CNN model is replaced by the focal loss function and the resultant architecture is combined with a vision transformer [101] to extract high-level descriptors.

6.4 Performance analysis

This section shows the performance comparison between different schemes that show better results that are used in the classification of skin lesion types. The evaluation measures such as F1-score, precision, recall, and accuracy are used for analysis. The schemes such as Ensemble DNN [102], multi-crop [103], Ensemble CNN [104], GAN [105], Synthetic GAN [106], MobileNetV2 with LSTM [107], Soft attention [108], EfficientNet [88], Lightweight network [109], Modified LSTM [110], Multi-scale attention [111], CNN with DenseNet [112], Transformer model [97], CNN with machine learning [113], U-Net with CNN [53], DenseNet with ConvNeXt [114], Diverse CNN [115], and TResNet [116] are used for analysis. When comparing these schemes, the diverse CNN results in an F1-score, precision, recall, and accuracy of 93.86%, 94.63%, 93.11%, and 96.12% respectively as illustrated in Table V. This evaluation was done using the HAM-10000 dataset.

Table V: Comparison of performance for different skin lesion classification schemes

Method	F1-Score	Precision	Recall	Accuracy
Ensemble DNN (2018) [102]	83.0	86.1	80.4	96.8
Multi-crop (2018) [103]	84.1	88.8	80.9	97.2
Ensemble CNN (2018) [104]	82.3	82.6	83.3	95.8
GAN (2019) [105]	83.44	84.43	82.86	-
Synthetic GAN (2020) [106]	73.2	76.9	74.3	95.2
MobileNetV2+LSTM (2021) [107]	-	-	92.24	90.21
Soft attention (2021) [108]	-	93.7	-	93.4
EfficientNet (2022) [90]	87.9	88	88	87.9
Lightweight network (2022) [109]	72.61	75.15	70.71	95.66
Modified LSTM (2022) [110]	90.48	90.38	90.58	91.03
Multi-scale Attention (2022) [111]	-	-	73.5	91.6
CNN+DenseNet (2022) [112]	93.27	-	-	92.25
Transformer model (2022) [97]	-	96.1	-	95.84
CNN+ML (2022) [113]	86	88	85	95.19
U-Net+CNN (2023) [53]	-	88.47	84.86	97.96
DenseNet+ConvNeXt (2023) [114]	83.45	83.75	83.81	-
TResNet (2024) [116]	89.48	90.23	88.85	98.23
Diverse CNN (2024) [115]	93.86	94.63	93.11	96.12

The Diverse CNN [115] results in an increase in F1-score and recall of 0.59% and 0.87% than the CNN with DenseNet architecture [112] and MobileNetV2 with LSTM architecture [107] respectively. The transformer model [97] results in a maximum precision of 96.1% and the TResNet [116] architecture results in a maximum accuracy of 98.23%.

7.Discussion and future perspectives

Skin lesion classification achieves high performance with Diverse CNN (F1: 93.86%, Acc: 96.12%), TResNet (Acc: 98.23%), and transformers (Precision: 96.1%, Acc: 95.84%). For segmentation, attention-based models like multi-attention (F1: 93.73%) outperform others. CNN-derived architectures with attention mechanisms, Vision Transformers, and EfficientNet excel, especially on large datasets, while data augmentation aids small datasets.Fine-tuning pretrained models enhances generalization across different demographics. Hyperparameter tuning via Bayesian optimization, grid search, and random search improves performance. U-Net-based models enhance lesion detection accuracy, crucial for effective treatment. Hardware deployment using ASIC, FPGA, or DSP enables real-time diagnosis with dermoscopic images.

8.Conclusion

The paper provides a systematic review related to the detection and classification of skin lesions. In the case of machine learning classifier, the use of an SVM classifier with generic deep features that are extracted by ResNet or DenseNet architecture provides a moderate result in the classification of lesions. In detecting the lesion region, the attention-based deep learning models and U-Net++ models provide better performance. Recently the approach multi-attention [70] resulted in recall, precision, and F1-score of 93.97%, 93.94%, and 93.72% respectively when evaluated utilizing the ISIC-2016 dataset. In the classification of lesion types, the usage of hybrid architectures and transformer models provides a better performance. Recently the Diverse CNN [115] results in an F1-score, precision, recall, and accuracy of 93.86%, 94.63%, 93.11%, and 96.12% respectively.

References

[1]. Gloster Jr, H. M., & Neal, K. (2006). Skin cancer in skin of color. *Journal of the American Academy of Dermatology*, *55*(5), 741-760.

[2]. Madan, V., Lear, J. T., & Szeimies, R. M. (2010). Non-melanoma skin cancer. *The lancet*, *375*(9715), 673-685.

[3]. Leiter, U., Keim, U., & Garbe, C. (2020). Epidemiology of skin cancer: update 2019. *Sunlight, Vitamin D and Skin Cancer*, 123-139.

[4]. Mangione, C. M., Barry, M. J., Nicholson, W. K., Chelmow, D., Coker, T. R., Davis, E. M., ... & US Preventive Services Task Force. (2023). Screening for skin cancer: US Preventive Services Task Force recommendation statement. *JAMA*, *329*(15), 1290-1295.

[5]. Sharma, S., Guleria, K., Kumar, S., & Tiwari, S. (2023, January). Benign and malignant skin lesion detection from Melanoma skin cancer images. In *2023 International Conference for Advancement in Technology (ICONAT)* (pp. 1-6). IEEE.

[6]. Mortaja, M., & Demehri, S. (2023). Skin cancer prevention–Recent advances and unmet challenges. *Cancer Letters*, 216406.

[7]. Hosny, K. M., Elshoura, D., Mohamed, E. R., Vrochidou, E., & Papakostas, G. A. (2023). Deep learning and optimization-based methods for skin lesions segmentation: a review. *IEEE Access*, *11*, 85467-85488.

[8]. Shinde, A., & Chaudhari, S. (2022, November). Statistical Analysis of Hair Detection and Removal Techniques Using Dermoscopic Images. In *International Conference on Computer Vision and Image Processing* (pp. 402-414). Cham: Springer Nature Switzerland.

[9]. Tadmor, E., Nezzar, S., & Vese, L. (2008). Multiscale hierarchical decomposition of images with applications to deblurring, denoising, and segmentation.

[10]. Sa'idah, S., Suparta, I. P. Y. N., & Suhartono, E. (2022). Modification of convolutional neural network GoogLeNet architecture with dull razor filtering for classifying skin cancer. *Jurnal Nasional Teknik Elektro dan TeknologiInformasi*, *11*(2).

[11]. Jerman, T., Pernuš, F., Likar, B., & Špiclin, Ž. (2015, March). Beyond Frangi: an improved multiscale vesselness filter. In *Medical imaging 2015: Image processing* (Vol. 9413, pp. 623-633). SPIE.

[12]. Abbas, Q., Garcia, I. F., Emre Celebi, M., & Ahmad, W. (2013). A feature-preserving hair removal algorithm for dermoscopy images. *Skin Research and Technology*, *19*(1), e27-e36.

[13]. Chen, Y., & Leedham, G. (2005). Decompose algorithm for thresholding degraded historical document images. *IEE Proceedings-Vision, Image and Signal Processing*, *152*(6), 702-714.

[14]. Louverdis, G., Vardavoulia, M. I., Andreadis, I., & Tsalides, P. (2002). A new approach to morphological color image processing. *Pattern recognition*, *35*(8), 1733-1741.

[15]. Sepasian, M., Balachandran, W., & Mares, C. (2008, October). Image enhancement for fingerprint minutiae-based algorithms using CLAHE, standard deviation analysis and sliding neighborhood. In *Proceedings of the World congress on Engineering and Computer Science* (pp. 22-24).

[16]. Sharma, T., & Verma, N. K. (2021). Single Image Dehazing and Non-uniform Illumination Enhancement: AZ-Score Approach. *SN Computer Science*, *2*(6), 488.

[17]. Negi, S. S., & Bhandari, Y. S. (2014, May). A hybrid approach to image enhancement using contrast stretching on image sharpening and the analysis of various cases arising using histogram. In *International conference on recent advances and innovations in engineering (ICRAIE-2014)* (pp. 1-6). IEEE.

[18]. Hossain, F., & Alsharif, M. R. (2007, November). Image enhancement based on logarithmic transform coefficient and adaptive histogram equalization. In *2007 International conference on convergence information technology (ICCIT 2007)* (pp. 1439-1444). IEEE.

[19]. Kim, H., Choi, S. M., Kim, C. S., & Koh, Y. J. (2021). Representative color transform for image enhancement. In *Proceedings of the IEEE/CVF international conference on computer vision* (pp. 4459-4468).

[20]. Reza, A. M. (2004). Realization of the contrast limited adaptive histogram equalization (CLAHE) for real-time image enhancement. *Journal of VLSI signal processing systems for signal, image and video technology*, *38*, 35-44.

[21]. Guan, X., Jian, S., Hongda, P., Zhiguo, Z., & Haibin, G. (2009, December). An image enhancement method based on gamma correction. In *2009 Second international symposium on computational intelligence and design* (Vol. 1, pp. 60-63). IEEE.

[27]. Gutman, D., Codella, N. C., Celebi, E., Helba, B., Marchetti, M., Mishra, N., & Halpern, A. (2016). Skin lesion analysis toward melanoma detection: A challenge at the international symposium on biomedical imaging (ISBI) 2016, hosted by the international skin imaging collaboration (ISIC). *arXiv preprint arXiv:1605.01397*.

[28]. Codella, N. C., Gutman, D., Celebi, M. E., Helba, B., Marchetti, M. A., Dusza, S. W., ... & Halpern, A. (2018, April). Skin lesion analysis toward melanoma detection: A challenge at the 2017 international symposium on biomedical imaging (isbi), hosted by the international skin imaging collaboration (isic). In *2018 IEEE 15th international symposium on biomedical imaging (ISBI 2018)* (pp. 168-172). IEEE.

[29]. Tschandl, P., Rosendahl, C., & Kittler, H. (2018). The HAM10000 dataset, a large collection of multi-source dermatoscopic images of common pigmented skin lesions. *Scientific data*, *5*(1), 1-9.

[30]. Combalia, M., Codella, N. C., Rotemberg, V., Helba, B., Vilaplana, V., Reiter, O., ... & Malvehy, J. (2019). Bcn20000: Dermoscopic lesions in the wild. *arXiv preprint arXiv:1908.02288*.

[31]. Rotemberg, V., Kurtansky, N., Betz-Stablein, B., Caffery, L., Chousakos, E., Codella, N., ... & Soyer, H. P. (2021). A patient-centric dataset of images and metadata for identifying melanomas using clinical context. *Scientific data*, *8*(1), 34.

[32]. Hosny, K. M., Kassem, M. A., & Foaud, M. M. (2019). Classification of skin lesions using transfer learning and augmentation with Alex-net. *PloS one*, *14*(5), e0217293.

[33]. DermIS. Dermatology Information System. Accessed: Apr. 7, 2024. [Online]. Available: https://www.dermis.net/dermisroot/en/home/ index.htm

[34]. Mendonça, T., Ferreira, P. M., Marques, J. S., Marcal, A. R., & Rozeira, J. (2013, July). PH 2-A dermoscopic image database for research and benchmarking. In *2013 35th annual international conference of the IEEE engineering in medicine and biology society (EMBC)* (pp. 5437-5440). IEEE.

[35]. Gupta, D., & Anand, R. S. (2017). A hybrid edge-based segmentation approach for ultrasound medical images. *Biomedical Signal Processing and Control*, *31*, 116-126.

[36]. Gould, S., Gao, T., & Koller, D. (2009). Region-based segmentation and object detection. *Advances in neural information processing systems*, *22*.

[37]. Mustaqeem, A., Javed, A., & Fatima, T. (2012). An efficient brain tumor detection algorithm using watershed & thresholding based segmentation. *International Journal of Image, Graphics and Signal Processing*, *4*(10), 34.

[38]. Dhanachandra, N., Manglem, K., & Chanu, Y. J. (2015). Image segmentation using K-means clustering algorithm and subtractive clustering algorithm. *Procedia Computer Science*, *54*, 764-771.

[39]. Yang, J., Ke, Y. S., & Wang, M. Z. (2017). An adaptive clustering segmentation algorithm based on FCM. *Turkish Journal of Electrical Engineering and Computer Sciences*, *25*(6), 4533-4544.

[40]. Bhattacharyya, S., Maulik, U., & Dutta, P. (2011). Multilevel image segmentation with adaptive image context based thresholding. *Applied soft computing*, *11*(1), 946-962.

[41]. Goh, T. Y., Basah, S. N., Yazid, H., Safar, M. J. A., & Saad, F. S. A. (2018). Performance analysis of image thresholding: Otsu technique. *Measurement*, *114*, 298-307.

[42]. Khehra, B. S., Singh, A., Pharwaha, A. P. S., & Kaur, P. (2016). Image segmentation using two-dimensional Renyi entropy. In *Proceedings of the International Congress on Information and Communication Technology: ICICT 2015, Volume 1* (pp. 521-530). Springer Singapore.

[43]. Abbas, Q., Fondón, I., Sarmiento, A., & Emre Celebi, M. (2014). An improved segmentation method for non-melanoma skin lesions using active contour model. In *Image Analysis and Recognition: 11th International Conference, ICIAR 2014, Vilamoura, Portugal, October 22-24, 2014, Proceedings, Part II 11* (pp. 193-200). Springer International Publishing.

[44]. Kim, B., & Ye, J. C. (2019). Mumford–Shah loss functional for image segmentation with deep learning. *IEEE Transactions on Image Processing*, *29*, 1856-1866.

[45]. Bansal, S., & Maini, R. (2013). Performance analysis of color based region split and merge and otsu's thresholding techniques for brain tumor extraction. *International Journal of Engineering Research and Applications*, *3*(4), 1640-1643.

[46]. Ahmed, A. S. (2018). Comparative study among Sobel, Prewitt and Canny edge detection operators used in image processing. *J. Theor. Appl. Inf. Technol*, *96*(19), 6517-6525.

[47]. Iranpoor, R., Mahboob, A. S., Shahbandegan, S., & Baniasadi, N. (2020, December). Skin lesion segmentation using convolutional neural networks with improved U-Net architecture. In *2020 6th Iranian Conference on Signal Processing and Intelligent Systems (ICSPIS)* (pp. 1-5). IEEE.

[48]. Wu, F., Liu, S., Li, B., & Tang, J. (2022, December). Increase Channel Attention Based on Unet++ Architecture for Medical Images. In *International Conference on Machine Learning for Cyber Security* (pp. 516-520). Cham: Springer Nature Switzerland.

[49]. Lei, B., Xia, Z., Jiang, F., Jiang, X., Ge, Z., Xu, Y., ... & Wang, S. (2020). Skin lesion segmentation via generative adversarial networks with dual discriminators. *Medical Image Analysis*, *64*, 101716.

[50]. Alahmadi, M. D. (2022). Multiscale attention U-Net for skin lesion segmentation. *IEEE Access*, *10*, 59145-59154.

[51]. Dash, M., Londhe, N. D., Ghosh, S., Semwal, A., & Sonawane, R. S. (2019). PsLSNet: Automated psoriasis skin lesion segmentation using modified U-Net-based fully convolutional network. *Biomedical Signal Processing and Control*, *52*, 226-237.

[52]. Le, P. T., Pham, B. T., Chang, C. C., Hsu, Y. C., Tai, T. C., Li, Y. H., & Wang, J. C. (2023). Anti-aliasing attention U-net model for skin lesion segmentation. *Diagnostics*, *13*(8), 1460.

[53]. Anand, V., Gupta, S., Koundal, D., & Singh, K. (2023). Fusion of U-Net and CNN model for segmentation and classification of skin lesion from dermoscopy images. *Expert Systems with Applications*, *213*, 119230.

[54]. Zhao, C., Shuai, R., Ma, L., Liu, W., & Wu, M. (2022). Segmentation of skin lesions image based on U-Net++. *Multimedia Tools and Applications*, *81*(6), 8691-8717.

[55]. Bharati, P., & Pramanik, A. (2020). Deep learning techniques—R-CNN to mask R-CNN: a survey. *Computational Intelligence in Pattern Recognition: Proceedings of CIPR 2019*, 657-668.

[56]. Suchetha, M., Ganesh, N. S., Raman, R., & Dhas, D. E. (2021). Region of interest-based predictive algorithm for subretinal hemorrhage detection using faster R-CNN. *Soft Computing*, *25*(24), 15255-15268.

[57]. Bagheri, F., Tarokh, M. J., & Ziaratban, M. (2021). Skin lesion segmentation from dermoscopic images by using Mask R-CNN, Retina-Deeplab, and graph-based methods. *Biomedical Signal Processing and Control*, *67*, 102533.

[58]. Khan, M. A., Akram, T., Zhang, Y. D., & Sharif, M. (2021). Attributes based skin lesion detection and recognition: A mask RCNN and transfer learning-based deep learning framework. *Pattern Recognition Letters*, *143*, 58-66.

[59]. Peng, H., Xue, C., Shao, Y., Chen, K., Xiong, J., Xie, Z., & Zhang, L. (2020). Semantic segmentation of litchi branches using DeepLabV3+ model. *Ieee Access*, *8*, 164546-164555.

[60]. Bagheri, F., Tarokh, M. J., & Ziaratban, M. (2022). Skin lesion segmentation by using object detection networks, DeepLab3+, and active contours. *Turkish Journal of Electrical Engineering and Computer Sciences*, *30*(7), 2489-2507.

[61]. Zafar, M., Amin, J., Sharif, M., Anjum, M. A., Mallah, G. A., & Kadry, S. (2023). DeepLabv3+-based segmentation and best features selection using slime mould algorithm for multi-class skin lesion classification. *Mathematics*, *11*(2), 364.

[62]. Masood, H., Naseer, A., & Saeed, M. (2024). Optimized Skin Lesion Segmentation: Analysing DeepLabV3+ and ASSP Against Generative AI-Based Deep Learning Approach. *Foundations of Science*, 1-25.

[63]. Masood, H., Naseer, A., & Saeed, M. (2024). Optimized Skin Lesion Segmentation: Analysing DeepLabV3+ and ASSP Against Generative AI-Based Deep Learning Approach. *Foundations of Science*, 1-25.

[64]. Ronneberger, O., Fischer, P., & Brox, T. (2015). U-net: Convolutional networks for biomedical image segmentation. In *Medical image computing and computer-assisted intervention–MICCAI 2015: 18th international conference, Munich, Germany, October 5-9, 2015, proceedings, part III 18* (pp. 234-241). Springer International Publishing.

[65]. Li, W., Qin, S., Li, F., & Wang, L. (2021). MAD-UNet: A deep U-shaped network combined with an attention mechanism for pancreas segmentation in CT images. *Medical Physics*, *48*(1), 329-341.

[66]. Zhou, Z., Rahman Siddiquee, M. M., Tajbakhsh, N., & Liang, J. (2018). Unet++: A nested u-net architecture for medical image segmentation. In *Deep Learning in Medical Image Analysis and Multimodal Learning for Clinical Decision Support: 4th International Workshop, DLMIA 2018, and 8th International Workshop, ML-CDS 2018, Held in Conjunction with MICCAI 2018, Granada, Spain, September 20, 2018, Proceedings 4* (pp. 3-11). Springer International Publishing.

[67]. Xu, Q., Ma, Z., Na, H. E., & Duan, W. (2023). DCSAU-Net: A deeper and more compact split-attention U-Net for medical image segmentation. *Computers in Biology and Medicine*, *154*, 106626.

[68]. Zheng, J., Liu, H., Feng, Y., Xu, J., & Zhao, L. (2023). CASF-Net: Cross-attention and cross-scale fusion network for medical image segmentation. *Computer Methods and Programs in Biomedicine*, *229*, 107307.

[69]. Zhao, X., Zhang, L., & Lu, H. (2021). Automatic polyp segmentation via multi-scale subtraction network. In *Medical Image Computing and Computer Assisted Intervention–MICCAI 2021: 24th International Conference, Strasbourg, France, September 27–October 1, 2021, Proceedings, Part I 24* (pp. 120-130). Springer International Publishing.

[70]. Dong, Z., Li, J., & Hua, Z. (2024). Transformer-based multi-attention hybrid networks for skin lesion segmentation. *Expert Systems with Applications*, *244*, 123016.

[71]. Melbin, K., & Raj, Y. J. V. (2021). Integration of modified ABCD features and support vector machine for skin lesion types classification. *Multimedia Tools and Applications*, *80*(6), 8909-8929.

[72]. Ali, A. R., Li, J., Kanwal, S., Yang, G., Hussain, A., & Jane O'Shea, S. (2020). A novel fuzzy multilayer perceptron (F-MLP) for the detection of irregularity in skin lesion border using dermoscopic images. *Frontiers in medicine*, *7*, 297.

[73]. Fisher, R. B., Rees, J., & Bertrand, A. (2020). Classification of ten skin lesion classes: Hierarchical knn versus deep net. In *Medical Image Understanding and Analysis: 23rd Conference, MIUA 2019, Liverpool, UK, July 24–26, 2019, Proceedings 23* (pp. 86-98). Springer International Publishing.

[74]. Banerjee, A., Sarkar, S., Nasipuri, M., & Das, N. (2023). Skin Diseases Detection Using LBP and WLD: An Ensembling Approach. *SN Computer Science*, *5*(1), 72.

[75]. Aydin, Y. (2023). A Comparative Analysis of Skin Cancer Detection Applications Using Histogram-Based Local Descriptors. *Diagnostics*, *13*(19), 3142.

[76]. George, Y., Aldeen, M., & Garnavi, R. (2019). Automatic scale severity assessment method in psoriasis skin images using local descriptors. *IEEE Journal of Biomedical and Health Informatics*, *24*(2), 577-585.

[77]. Malinga, B., Raicu, D., & Furst, J. (2006). Local vs. global histogram-based color image clustering. *University of Depaul, Technical Reports: TR06-010*.

[78]. Mukadam, S. B., & Patil, H. Y. (2024). Machine learning and computer vision based methods for cancer classification: A systematic review. *Archives of Computational Methods in Engineering*, 1-36.

[79]. Garcia-Arroyo, J. L., & Garcia-Zapirain, B. (2019). Segmentation of skin lesions in dermoscopy images using fuzzy classification of pixels and histogram thresholding. *Computer methods and programs in biomedicine*, *168*, 11-19.

[80]. Han, X., Zhang, Z., Ding, N., Gu, Y., Liu, X., Huo, Y., ... & Zhu, J. (2021). Pre-trained models: Past, present and future. *AI Open*, *2*, 225-250.

[81]. Sharma, M., Jain, B., Kargeti, C., Gupta, V., & Gupta, D. (2021). Detection and diagnosis of skin diseases using residual neural networks (RESNET). *International Journal of Image and Graphics*, *21*(05), 2140002.

[82]. Guo, S., & Yang, Z. (2018). Multi-Channel-ResNet: An integration framework towards skin lesion analysis. *Informatics in Medicine Unlocked*, *12*, 67-74. Skin Lesion Classification Using Hybrid Deep Neural Networks.

[83]. Mahbod, A., Schaefer, G., Wang, C., Ecker, R., & Ellinge, I. (2019, May). Skin lesion classification using hybrid deep neural networks. In *ICASSP 2019-2019 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)* (pp. 1229-1233). IEEE.

[84]. Gouda, N., & Amudha, J. (2020, October). Skin cancer classification using ResNet. In *2020 IEEE 5th International Conference on Computing Communication and Automation (ICCCA)* (pp. 536-541). IEEE.

[85]. Al-Masni, M. A., Kim, D. H., & Kim, T. S. (2020). Multiple skin lesions diagnostics via integrated deep convolutional networks for segmentation and classification. *Computer methods and programs in biomedicine*, *190*, 105351.

[86]. Miglani, V., & Bhatia, M. P. S. (2020, February). Skin lesion classification: A transfer learning approach using efficientnets. In *International Conference on Advanced Machine Learning Technologies and Applications* (pp. 315-324). Singapore: Springer Singapore.

[87]. Zhang, J., Xie, Y., Xia, Y., & Shen, C. (2019). Attention residual learning for skin lesion classification. *IEEE transactions on medical imaging*, *38*(9), 2092-2103.

[88]. Hridoy, R. H., Akter, F., & Rakshit, A. (2021, July). Computer vision based skin disorder recognition using EfficientNet: A transfer learning approach. In *2021 International conference on information technology (ICIT)* (pp. 482-487). IEEE.

[89]. Gessert, N., Nielsen, M., Shaikh, M., Werner, R., & Schlaefer, A. (2020). Skin lesion classification using ensembles of multi-resolution EfficientNets with meta data. *MethodsX*, *7*, 100864.

[90]. Ali, K., Shaikh, Z. A., Khan, A. A., & Laghari, A. A. (2022). Multiclass skin cancer classification using EfficientNets–a first step towards preventing skin cancer. *Neuroscience Informatics*, *2*(4), 100034.

[91]. Venugopal, V., Raj, N. I., Nath, M. K., & Stephen, N. (2023). A deep neural network using modified EfficientNet for skin cancer detection in dermoscopic images. *Decision Analytics Journal*, *8*, 100278.

[92]. Rafay, A., & Hussain, W. (2023). EfficientSkinDis: An EfficientNet-based classification model for a large manually curated dataset of 31 skin diseases. *Biomedical Signal Processing and Control*, *85*, 104869.

[93]. Venugopal, V., Joseph, J., Das, M. V., & Nath, M. K. (2022). An EfficientNet-based modified sigmoid transform for enhancing dermatological macro-images of melanoma and nevi skin lesions. *Computer Methods and Programs in Biomedicine*, *222*, 106935.

[94]. Hosny, K. M., Kassem, M. A., & Foaud, M. M. (2018, December). Skin cancer classification using deep learning and transfer learning. In *2018 9th Cairo international biomedical engineering conference (CIBEC)* (pp. 90-93). IEEE.

[95]. Flosdorf, C., Engelker, J., Keller, I., & Mohr, N. (2024). Skin Cancer Detection utilizing Deep Learning: Classification of Skin Lesion Images using a Vision Transformer. *arXiv preprint arXiv:2407.18554*.

[96]. Krishna, G. S., Supriya, K., & Sorgile, M. (2023). LesionAid: vision transformers-based skin lesion generation and classification. *arXiv preprint arXiv:2302.01104*.

[97]. Nakai, K., Chen, Y. W., & Han, X. H. (2022). Enhanced deep bottleneck transformer model for skin lesion classification. *Biomedical Signal Processing and Control*, *78*, 103997.

[98]. Sarker, M. M. K., Moreno-García, C. F., Ren, J., & Elyan, E. (2022, July). TransSLC: Skin lesion classification in dermatoscopic images using transformers. In *Annual Conference on Medical Image Understanding and Analysis* (pp. 651-660). Cham: Springer International Publishing.

[99]. He, X., Tan, E. L., Bi, H., Zhang, X., Zhao, S., & Lei, B. (2022). Fully transformer network for skin lesion analysis. *Medical Image Analysis*, *77*, 102357.

[100]. Xin, C., Liu, Z., Zhao, K., Miao, L., Ma, Y., Zhu, X., ... & Chen, H. (2022). An improved transformer network for skin cancer classification. *Computers in Biology and Medicine*, *149*, 105939.

[101]. Nie, Y., Sommella, P., Carratù, M., O'Nils, M., & Lundgren, J. (2022). A deep CNN transformer hybrid model for skin lesion classification of dermoscopic images using focal loss. *Diagnostics*, *13*(1), 72.

[102]. Li, Y., & Shen, L. (2018). Skin lesion analysis towards melanoma detection using deep learning network. *Sensors*, *18*(2), 556.

[103]. Gessert, N., Sentker, T., Madesta, F., Schmitz, R., Kniep, H., Baltruschat, I., ... & Schlaefer, A. (2018). Skin lesion diagnosis using ensembles, unscaled multi-crop evaluation and loss weighting. *arXiv preprint arXiv:1808.01694*.

[104]. Nozdryn-Plotnicki, A., Yap, J., & Yolland, W. (2018). Ensembling convolutional neural networks for skin cancer classification. *International Skin Imaging Collaboration (ISIC) Challenge on Skin Image Analysis for Melanoma Detection. MICCAI*.

[105]. Rashid, H., Tanveer, M. A., & Khan, H. A. (2019, July). Skin lesion classification using GAN based data augmentation. In *2019 41St annual international conference of the IEEE engineering in medicine and biology society (EMBC)* (pp. 916-919). IEEE.

[106]. Qin, Z., Liu, Z., Zhu, P., & Xue, Y. (2020). A GAN-based image synthesis method for skin lesion classification. *Computer Methods and Programs in Biomedicine*, *195*, 105568.

[107]. Srinivasu, P. N., SivaSai, J. G., Ijaz, M. F., Bhoi, A. K., Kim, W., & Kang, J. J. (2021). Classification of skin disease using deep learning neural networks with MobileNet V2 and LSTM. *Sensors*, *21*(8), 2852.

[108]. Datta, S. K., Shaikh, M. A., Srihari, S. N., & Gao, M. (2021). Soft attention improves skin cancer classification performance. In *Interpretability of Machine Intelligence in Medical Image Computing, and Topological Data Analysis and Its Applications for Medical Data: 4th International Workshop, iMIMIC 2021, and 1st International Workshop, TDA4MedicalData 2021, Held in Conjunction with MICCAI 2021, Strasbourg, France, September 27, 2021, Proceedings 4* (pp. 13-23). Springer International Publishing.

[109]. Hoang, L., Lee, S. H., Lee, E. J., & Kwon, K. R. (2022). Multiclass skin lesion classification using a novel lightweight deep learning framework for smart healthcare. *Applied Sciences*, *12*(5), 2677.

[110]. Elashiri, M. A., Rajesh, A., Pandey, S. N., Shukla, S. K., & Urooj, S. (2022). Ensemble of weighted deep concatenated features for the skin disease classification model using modified long short term memory. *Biomedical Signal Processing and Control*, *76*, 103729.

[111]. Qian, S., Ren, K., Zhang, W., & Ning, H. (2022). Skin lesion classification using CNNs with grouping of multi-scale attention and class-specific loss weighting. *Computer Methods and Programs in Biomedicine*, *226*, 107166.

[112]. Kousis, I., Perikos, I., Hatzilygeroudis, I., & Virvou, M. (2022). Deep learning methods for accurate skin cancer recognition and mobile application. *Electronics*, *11*(9), 1294.

[113]. Shetty, B., Fernandes, R., Rodrigues, A. P., Chengoden, R., Bhattacharya, S., & Lakshmanna, K. (2022). Skin lesion classification of dermoscopic images using machine learning and convolutional neural network. *Scientific Reports*, *12*(1), 18134.

[114]. Wei, M., Wu, Q., Ji, H., Wang, J., Lyu, T., Liu, J., & Zhao, L. (2023). A skin disease classification model based on densenet and convnext fusion. *Electronics*, *12*(2), 438.

[115]. Tabibi, S. T., Nikravanshalmani, A., & Saboohi, H. (2024). An Ensemble Classifier Based on Diverse Convolutional Neural Networks for Skin Lesions Classification. *IEEE Access*.

[116]. Su, Q., Hamed, H. N. A., Isa, M. A., Hao, X., & Dai, X. (2024). A GAN-based data augmentation method for imbalanced multi-class skin lesion classification. *IEEE Access*.