

# Skin Lesions Classification and Segmentation: A Review

Marzuraikah Mohd Stofa, Mohd Asyraf Zulkifley, Muhammad Ammirul Atiqi Mohd Zainuri  
Department of Electrical, Electronic and Systems Engineering, Universiti Kebangsaan Malaysia, Bangi, Malaysia

**Abstract**—An automated intelligent system based on imaging input for unbiased diagnosis of skin-related diseases is an essential screening tool nowadays. This is because visual and manual analysis of skin lesion conditions based on images is a time-consuming process that puts a significant workload on health practitioners. Various machine learning and deep learning techniques have been researched to reduce and alleviate the workloads. In several early studies, the standard machine learning techniques are the more popular approach, which is in contrast to the recent studies that rely more on the deep learning approach. Although the recent deep learning approach, mainly based on convolutional neural networks has shown impressive results, some challenges remain open due to the complexity of the skin lesions. This paper presents a wide range of analyses that cover classification and segmentation phases of skin lesion detection using deep learning techniques. The review starts with the classification techniques used for skin lesion detection, followed by a concise review on lesions segmentation, also using the deep learning techniques. Finally, this paper examined and analyzed the performances of state-of-the-art methods that have been evaluated on various skin lesion datasets. This paper has utilized performance measures based on accuracy, mean specificity, mean sensitivity, and area under the curve of 12 different Convolutional Neural Network based classification models.

**Keywords**—*Lesion segmentation; lesion classification; machine learning; deep learning; skin lesions*

## I. INTRODUCTION

Skin cancer is one of the most dangerous types of cancer that infected humans regularly. In the field of dermatology, there are two types of skin cancers, which are melanocytic and non-melanocytic. For example, melanoma is a type of melanocytic cancer, which is found to be a riskier version of cancer compared to the non-melanocytic type. Therefore, diagnosis of the correct type of cancer at an early stage is important to reduce the mortality risk [1], [2]. Besides that, there are certain parts of the body that have a higher probability of infection such as the chest, back, and legs. Then, this paper observed that most research in recent years has focused on establishing an automated intelligent system for the unbiased diagnosis of pigmented skin lesions. The general framework of the system involves pre-processing, feature extraction, segmentation, and classification phases, which are necessary steps in obtaining accurate localization of the skin lesion map. Masood et al. [3] and Adeyinka et al. [4] is also found that diagnosis of skin cancer at an early stage using computer vision provides a significant improvement when machine learning techniques are implemented. First, the diagnosis process begins by removing unnecessary structures or artifacts

on the skin lesion image that might interfere during the segmentation process, such as air bubbles, hair, blood vessels, and oily surfaces. In general, skin lesions come in various colors, shapes, and sizes that limit the standard machine learning ability to obtain high levels of accuracy. This process involves complex annotations during manual screening even for dermatologists. Therefore, Al-Masni et al. [5] presented that an automated computerized diagnostic system is an important tool in skin lesion analysis that will be able to assist and support dermatologists in making timely decisions. Abdani et al. [6] show that deep learning has demonstrated its effectiveness in various applications, particularly in computer vision-related systems that use convolutional neural networks (CNN) as the base framework, even for a compact version. For example, a previous study in [7] has shown that the popular method in deep learning is through CNN utilization, which can process common and highly variable tasks in handling delicate objects. Krizhevsky et al. [8] and Lecun et al. [9] prove that this sophisticated and optimized model has better ability than handcrafted features in extracting outstanding features from the entire images of skin lesions.

The development of computer-aided algorithms is essential to address the increasing problem of global skin cancer cases where it is able to handle large amounts of data in real time and automatically. It is important to review the performance of deep learning algorithms in the classification and segmentation of skin lesions due to recent advances in deep learning paradigms, and particularly in medical imaging it shows excellent performance. So, in this study, an extensive investigation of the various approaches for analyzing skin lesions was conducted. In addition, the classification techniques are reviewed and compared in Section II, which is the process of categorizing the classes of skin lesions and other types of surfaces. A comparison between all the segmentation techniques is presented in the following Section III. In addition, a comparative analysis using deep learning methods for classification and segmentation of skin lesions was performed in Section IV to show the strengths and weaknesses of each method, and subsequently the conclusion section.

## II. CLASSIFICATION

The skin cancer detection system is made more accessible by categorizing images of lesions. This classification process can assist dermatologists in detecting the possibility of early skin cancer through visual-based sensing. According to the standard medical practice, skin lesions are often classified as benign or malignant cancer. Thence, each of the lesion types can be further classified into seborrheic keratosis, solar lentigo, squamous cell carcinoma, nevi, actinic keratosis, basal cell

carcinoma, melanoma, and others. In this paper, both the traditional and recent state-of-the-art methods were reviewed. Table I summarizes the differences between various general deep classification networks.

TABLE I. COMPARISON BETWEEN GENERAL CNN ARCHITECTURES FOR SKIN LESION CLASSIFICATION TASK

Techniques	Description	Advantages	Disadvantages
AlexNet [10]	Includes three fully connected layers and five convolutional layers. This model used various sizes of filter.	GPU are used as an accelerator to handle the complex architecture.	The probability of generating artifacts from feature maps is high because the filter's size is quite large.
VGGNet [11]	Uses only 3x3 of convolutional filters, placed on top of each other to increase the network depth.	Encouraging performance that uses up to 19 layers with significant improvement over the previous arrangements.	It is challenging to train the model, especially for the cases without transfer learning.
GoogleNet [12]	An architecture that has 22 layers of deep network.	There is no uncontrolled increment in computing complexity when more units are added at each level.	Difficulty in customizing the parameters due to the use of heterogeneous topology.
ResNet [13]	Applies feedforward neural network layers with skip connection by performing identity mapping and added them to the stacked-layer output.	Increase the network's depth and easier to optimize, while reducing zero diminishing gradient issues, which indirectly improves the accuracy.	Information on the features map is complicated and may be degraded throughout the feed-forward procedures.
Xception [14]	An "extreme" version of the Inception module that replaces the module with depth-wise separable convolution	Easy to define and modify with high accuracy performance.	High computational cost due to multiple layers of CNN with 728 filters.
DenseNet [15]	Have a complex connection to achieve maximum information flows between forward and backward layers.	Training is relatively easy due to the enhanced flows of gradient and information across the network.	The number of parameters increases a lot between shallow and deep configurations because of more feature maps in each layer.
EfficientNet [16]	The architecture model consists of eight configurations from B0 to B7, with each subsequent model refers to a variant with more parameters and higher accuracy.	Reduce computation cost and can produce faster classification inference.	In order to catch fine-grained patterns on huge images, the network requires additional layers that increase the receptive field size and uses more channels.

A. Recent Works on Conventional Classification Method

At the beginning of the study on skin lesion classification, the traditional machine learning approach was commonly used, whereby region-based or threshold-based approaches are utilized to extract the features. Some of the most popular conventional approaches nowadays are support vector machine (SVM), k-nearest neighbor, artificial neural network (ANN), and naive Bayesian algorithm [17], [18]. Then, the deep learning-based method was started to be developed to overcome the limitations of previously mentioned conventional approaches. Han et al. [19] presented the conventional methods require significant effort from humans to design the feature extractors, still they do not produce accurate multi-class skin lesions detection.

B. Recent Works on Deep Learning Classification Method

The deep learning method based on the CNN classifier has exceeded the general human capability in performing object classification tasks, whereby historically, it begins to gain popularity in 2012 [20]. Fig. 1 shows the general CNN architecture with standard major components such as CNN layers, activation function, and the trainable hyperparameters. Several previous studies have implemented CNN that was introduced in [7] to produce dermatologist-equivalent skin cancer classifiers [21], [22]. Compared to the traditional methods, the CNN-based methods proved to be more effective. Many CNN architectures are available for skin lesion classification such as AlexNet [10], GoogleNet/InceptionNet [12], VGG Net [11], ResNet [13], XceptionNet [14], DenseNet [15] or EfficientNet [16]. All these methods are discussed as follow:

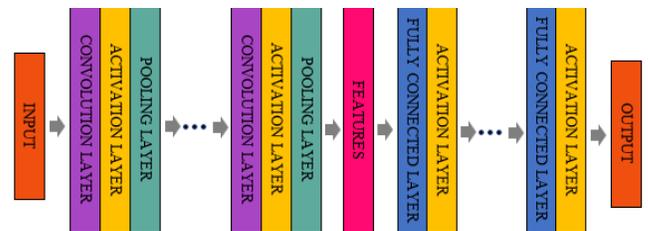


Fig. 1. General Architecture of CNN.

1) AlexNet: AlexNet is a system developed over 10 years ago [10]. It utilizes two operators, the convolutional network and the pooling layer, which will be the main building blocks of the network. The network starts with several layers of convolutional layers, followed by the fully connected layers, which are aligned through flatten operator. AlexNet has also been developed for implementing deep neural networks (DNNs) methodology in speech recognition and computer vision. Such as in 2019, the work in [23] has applied AlexNet to classify skin lesions using various configurations. The proposed method managed to overcome the overfitting problem by adjusting the weight values and enriching the data set with synthetic data generated from different rotation angles. The final classification layer is then replaced with the softmax layer to categorize more than two types of skin lesion categories. The experiment results have exceeded the initial performance expectations, whereby this model is still being

used as the benchmark in classifying the skin lesions. This general architecture of this network is illustrated in Fig. 2.

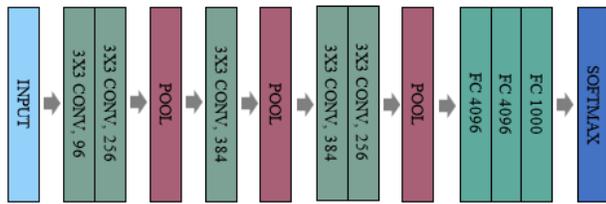


Fig. 2. Diagram of AlexNet Architecture.

2) *VGGNet*: In comparison to AlexNet, VGGNet is a deeper and more complicated network. This model has been further improved by lowering the number of parameters [11]. In fact, it has been used as the building block for many compact applications [24], [25]. This model has been tested in large deep CNN configurations that consist of many convolutional layers followed by pooling layers for huge image classification tasks. Besides that, the pre-trained VGG network is also commonly utilized in various transfer domain applications. However, this model uses a significant amount of processing resources, and hence makes the application of the VGG model a tiresome task. Sun et al. [26] have recommended the usage of VGGNet to diagnose 198 types of skin lesions which were trained until they reached an optimal set of hyperparameters. They have utilized the DermQuest data set, which included 6,584 clinical pictures, whereby they have managed to obtain 50.27% average accuracy. Fig. 3. shows the general VGG 16 architecture, which is one of the biggest VGG network variants.

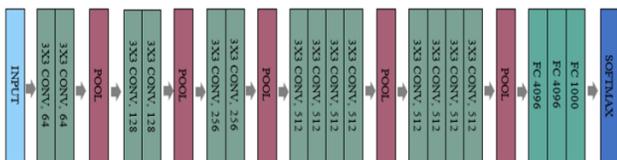


Fig. 3. Diagram of VGG 19 Architecture.

3) *GoogleNet*: GoogleNet is also popularly known as the InceptionNet [12]. It is composed of a 22-layer convolutional network structure. The primary goal of this design is to observe how the optimal local sparse structure can be handled and protected by the existing compact components. Most commonly, an Inception system is formed from modules that are stacked on top of one another. As Inception modules are stack on top of each other, their output correlation varies because deeper layers will capture better abstraction features, while the spatial concentration is expected to decrease accordingly. This reduction is done as such it will help the model to attain a faster training convergence. Thurnhofer-Hemsi et al. [27] have also used the CNN methodology to classify the skin lesion type based on the DermQuest database. The raw images were directly inserted into the CNN model to determine the presence of melanoma or not. They found out that GoogleNet and AlexNet produced the best results among the tested models. The authors have produced a highly accurate

system in terms of mean accuracy compared to the benchmarked models, even without utilizing any pre-processing step. The multiple inception modules are shown in Fig. 4.

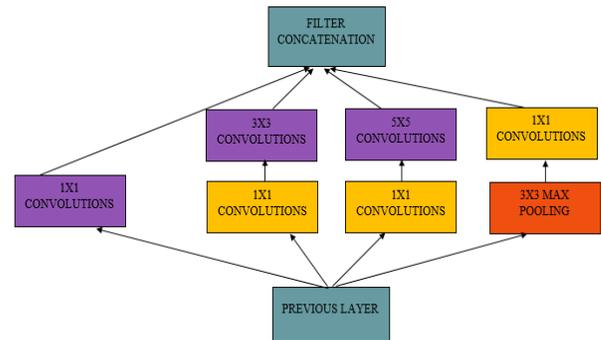


Fig. 4. Architecture of the GoogleNet Inception Module.

4) *ResNet*: As previously mentioned in [13], the residual modules in ResNet architecture can be used to train a very deep network effectively just by using the conventional stochastic gradient descent (SGD). He et al. [13] have shown that the residual networks with 152 layers can easily be trained and optimized to produce a model with good accuracy as its architecture becomes deeper. Moreover, they have also applied a feedforward neural network scheme with skip connection by performing identity mapping to combine the existing and skipped layers. This architecture is eight times deeper than VGGNet but it is still less complex and easier to train. Le et al. [28] have proposed a model that leverages the transfer learning method by using pre-trained models of ResNet 50, VGG 16, and MobileNet, coupled with weights and loss functions that focus on the classification process. Their results indicate that the ResNet 50 model produced the best performance with an average accuracy of 93% and total accuracy within the range [0.7, 0.94], which has surpassed the accuracy of dermatologists with an average accuracy of 84%.

5) *XceptionNet*: As an extension of the Inception design, Xception uses a stack of depth-separable convolution schemes to replace the Inception modules. In the newer versions of the Inception, some of the modules have replaced the different spatial dimensions ( $1 \times 1$ ,  $3 \times 3$ , and  $5 \times 5$ ) with a single dimension ( $3 \times 3$ ), followed by a pointwise convolution ( $1 \times 1$  convolution) to manage the computational complexity [14]. The feature extraction layer in the Xception architecture has a total of 36 convolutional layers with a large filter utilization of 728. Chaturvedi et al. [29] have suggested an automatic multiclass skin cancer disease classification system by conducting training procedures to obtain the optimal hyperparameter for five CNN models including Xception, ResNext 101, NasNetLarge, Inception V3, InceptionResNet V2, and the ensemble model. The best accuracy for the individual model was obtained by ResNext 101 and the best accuracy for the ensemble model was obtained by the combined network of InceptionResNet V2 and ResNetXt 101. However, the individual Xception model and the ensemble

model that contain Xception architecture also obtained good accuracy performances.

6) *DenseNet*: DenseNet is quite similar to ResNet from the architecture perspective, but the integration format of the two incoming networks is different, which leads to different network behaviors. Huang et al. [15] have developed an architecture with a simple connection pattern to ensure the maximal information flows between forward and backward layers to resolve the vanishing gradient problem. DenseNet accommodates the additional input from all previous layers by using cross-layer connectivity through the concatenation operator. Then, it transmits its feature maps to all subsequent layers, again via cross-layer connectivity. For image recognition, down-sampling layers divided the whole architecture into several densely connected blocks. Transition layers are also inserted between the convolution and pooling layers of various blocks. Hassan et al. [30] have implemented DenseNet-121 architecture to classify seven different types of skin lesions based on the HAM10000 dataset. Their model was trained by using supplemented augmentation data that managed to reach 92% of categorical accuracy and 97% of top2 accuracy which is much better than other models. The illustrated DenseNet architecture is shown in Fig. 5.

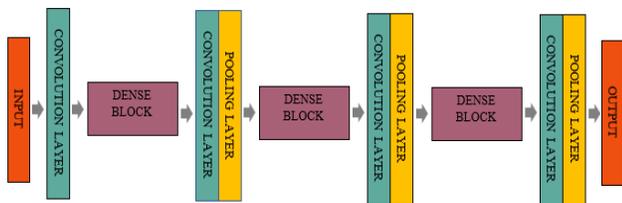


Fig. 5. Diagram of DenseNet Architecture.

7) *EfficientNet*: Recently, the EfficientNet model was introduced with a new up-and-down scaling strategy that scales uniformly the depth, width, and parameter resolution by using an effective compound coefficient [16]. One of the primary components of EfficientNet is MBConv or MobileNet, which follows an inverted bottleneck design. It is paired with an in-depth separable convolution by taking a shortcut between the bottlenecks, whereby it utilizes a considerably smaller number of channels. This model has achieved better classification accuracy compared to the existing models such as ResNet, DenseNet, Inception-V4, and NASNet when tested using a large ImageNet dataset. Gessert et al. [31] have implemented an ensemble of deep learning models to classify skin lesions using various EfficientNet architectures (B0-B6). They have boosted the training data by using meta information from ISIC 2019, whereby they have achieved the highest accuracy of 63.6% with AUC above 80% for the detection of skin lesions of eight different classes of skin.

### III. RECENT APPLICATION OF CNN MODELS FOR CLASSIFICATION TASKS

Many studies have indicated that CNN is a suitable method to be implemented for biomedical image applications, especially in the automated analysis of skin lesions. Esteva et

al. [7] have presented a method to classify malignant melanoma with significant accuracy by using a CNN architecture, which is through Google Inception V3 that has been pre-trained on 1.28 million images of general objects. Their automated system was then retrained with 129,450 clinical data of 2,032 different classes and managed to achieve 72.1% accuracy, whereas the two benchmark dermatologists only managed to obtain accuracy rates of 65.56% and 66%. Through the transfer learning scheme, the CNN classifier was able to achieve more or less a similar performance to those of 21 dermatologists in identifying malignant lesions, in which the CNN classifier produced an overall area under the curve (AUC) of >91%. Brinker et al. [32] then experimented on CNN deep learning architecture in categorizing skin lesions from 12,378 dermoscopic images, which were categorized into two classes of melanoma and atypical nevi. The findings were compared with the performances of dermatologists from various levels of competency and experience that also includes a few resident physicians from 12 German university hospitals. The CNN-based approach has managed to outperform the average accuracy of dermatologists.

Furthermore, Ratul et al. [33] have developed a computer-aided detection system for malignant skin lesions cases. Dilated convolution was used in four different architectures, namely InceptionV3, MobileNet, VGG16, and VGG19. The HAM10000 data set, which contains 10,015 dermoscopic images consisting of seven skin lesion classes were used to train, validate and test the algorithm with accuracy rates of 89.81%, 88.22%, 87.42%, and 85.02% for the previously mentioned models, respectively. Gessert et al. [34] researched further on the usage of patch-based techniques to extract fine-grain variations between different skin lesions using high image resolution input. Then, each image was divided into 5, 9, and 16 patches, which will be incorporated into a standard CNN architecture. Finally, three popular architectures were used to classify the skin lesions from high-resolution image patches, namely DenseNet, Inception V3, and SE-Resnext50.

Instead of using a fixed learning rate, Alqudah et al. [35] were integrated gradient descent with adaptive momentum learning rate and transfer learning approaches into two CNN architectures, which are AlexNet and GoogleNet for skin lesion classification tasks. Their method considered three types of skin lesions, which are benign, melanoma, and seborrheic keratosis, whereby the proposed classification approach was tested and evaluated using the International Skin Imaging Collaboration database. The system aimed to analyze input images to produce segmented and non-segmented skin lesions and reported accuracy rates of 92.2% and 89.8%, respectively. Contrary to the previous work, Akram et al. [36] developed a new framework for skin lesion classification that incorporates in-depth feature information to build the best discriminatory feature vectors while preserving the original feature space. To select discriminant features and reduce dimensionality, the authors have used the entropy-controlled neighborhood component analysis. The system employed several deep learning architectures, including Inception-V3, DenseNet 201, and Inception-ResNet-V2, as the classifiers. The proposed system was evaluated using different data sets, namely ISIC MSK, ISIC UDA, ISBI-2017, and PH2, with a common aim of

categorizing the skin lesions and obtained performance results of 98.8%, 99.2%, 97.1%, and 95.9%, respectively. The authors managed to cut off the features to less than 3% of the overall features, which resulted in improved classification accuracy by eliminating the redundancy and minimizing the computation time.

Researchers have also developed an integrated approach for skin lesion segmentation and multi-class lesion classification [37]. In this work, full-resolution convolutional network models have been applied for segmenting the lesion regions using popular CNN backbones of Inception, Densenet201, and ResNet-Inception to classify the segmented skin lesions. The Inception-ResNet model provided the most remarkable results out of the tested techniques. This model performed the best if it is trained with balanced data rather than imbalanced data. Using a similar approach, Purnama et al. [38] have tested two pre-trained CNN models, Inception V3 and MobileNet V1 for skin lesion classification. Then, they introduced an innovation through a web classifier. Their proposed method used a benchmark dataset of MNIST HAM 1000, where the results showed that Inception V3 had 72% accuracy, whereas MobileNet V1 only had 58% accuracy.

A unique multiple CNN models approach was proposed in [39] for solving challenging classification tasks due to the presence of artifacts, low-contrast images, and high intraclass differences in dermoscopic images. Multiple pre-trained CNN architectures were explored that include AlexNet, ResNet, GoogleNet, and VGG16 to speed up the training process using the dataset of ISIC 2016. This approach achieved an accuracy of 97.78% with an AUC of 0.98 for the training dataset and 85.22% with an AUC of 0.81 for the testing dataset. Instead of four models, Miglani and Bhatia [40] compared only two deep CNN models, ResNet-50 and EfficientNet-B0 for skin lesion classification purposes. The models were tested using the HAM10000 dataset, which resulted in the EfficientNet-B0 outperforming the ResNet-50 by achieving mean macro and micro AUC of 0.93 and 0.97, respectively. Their test has concluded that the recent CNN model is better in extracting richer, more complex, and fine-grain features of dermoscopic skin lesion images. Table II shows a comparison of the recent methods for skin lesion classification using the deep CNN methods.

TABLE II. COMPARISON OF RECENT CNN METHODS USED FOR SKIN LESION DETECTION

References	Datasets	Skin lesion classes	CNN architectures	Performance measures
[7]	ASIC, Edinburgh Dermofit Library, Stanford Hospital [7]	Benign and Malignant	Google Inception V3	Accuracy: 72.1%
[32]	HAM 10000 [41]	Melanoma and Nevi	ResNet50	Mean Specificity: 64.4% Mean Sensitivity: 89.4% ROC: 0.769
[33]	HAM 10000 [41]	Melanocytotic nevus, basal cell carcinoma, vascular lesions, dermatofibroma, benign keratosis, melanoma, and actinic keratosis	VGG16 VGG19 MobileNet InceptionV3	Accuracy; VGG16: 90.10% VGG19: 86.39% MobileNet: 89.48% InceptionV3: 90.95%
[34]	HAM 10000 [41]	Melanocytotic nevus, benign keratosis, vascular lesions, dermatofibroma, basal cell carcinoma, melanoma, and actinic keratosis	Inception V3 DenseNet121 SE-Resnext50	MC-Sensitivity; Inception V3: 64.0% DenseNet121: 67.8% SE-Resnext50: 66.9%
[35]	ISIC 2017 [42]	benign, melanoma, and seborrheic keratosis	AlexNet GoogleNet	Non-segmented accuracy; AlexNet: 92.2% GoogleNet: 92.2% Segmented accuracy AlexNet: 89.8% GoogleNet: 86.0%
[36]	PH2 [43] ISIC 2017 [42] ISIC-UDA, ISIC-MSK [44]	PH2: benign and melanoma ISBI 2017: melanoma, keratosis and benign ISIC-UDA, ISIC-MSK: benign and melanoma	Inception-V3 Inception-ResNet-V2 DenseNet-201	Accuracy; PH2: 98.80% ISBI-2017: 95.90% ISIC-UDA: 97.10% ISIC-MSK: 99.20%
[37]	ISIC 2016 [44] ISIC 2017 [42] HAM 10000 [41]	ISIC 2016: benign and melanoma ISIC 2017: benign, seborrheic keratosis, and melanoma HAM 10000: Melanocytotic nevus, basal cell carcinoma, vascular lesions, dermatofibroma, benign keratosis, melanoma, and actinic keratosis	DenseNet-201 ResNet-50 Inception-v3 Inception-ResNet-v2	ResNet-50 accuracy; ISIC 2016: 81.79% ISIC 2017: 81.57% ISIC 2017: 89.28%
[38]	MNIST HAM 10000 [41]	Melanocytotic nevus, basal cell carcinoma, vascular lesions, dermatofibroma, benign keratosis, melanoma, and actinic keratosis	MobileNet v1 Inception V3	Accuracy; Inception V3: 72% MobileNet v1: 58%
[39]	HAM 10000 [41]	Melanocytotic nevus, basal cell carcinoma, vascular lesions, dermatofibroma, benign keratosis, melanoma, and actinic keratosis	AlexNet VGG16 GoogleNet ResNet	AUC; Training: 0.99 Validation: 0.72
[40]	HAM 10000 [41]	Melanocytotic nevus, basal cell carcinoma, vascular lesions, dermatofibroma, benign keratosis, melanoma, and actinic keratosis	EfficientNet	Averaged AUC; macro: 0.93 micro: 0.97

#### IV. SEGMENTATION

Image segmentation is needed in a large-scale approach to diagnosing skin lesions automatically. It is an important step where the images will undergo pattern recognition or the utilization of a rule-based method to segment the region of interest (ROI). Al-Masni et al. [5] have defined ROI as the lesion areas that are separated from the non-lesion region. Generally, identifying the ROI requires a module to detect gaps in the images, before applying the similarity criteria to segment the lesions together [45]. The conventional approach involves handcrafted feature-based methods such as the edge [46], region [47], threshold [48], and intelligence-based methods [5]. The machine learning methods include both deep learning and conventional techniques, which will be discussed in the following section to examine and compare the segmentation performance. Table III shows the comparison between all segmentation techniques.

##### A. Recent Works on Skin Lesion Segmentation using Conventional Intelligence-Based Method

Conventional artificial intelligence allows rapid implementation of the skin lesion segmentation without much training requirement with a much lesser dataset compared to the deep learning approach. It can be easily implemented as such it allows a wider spread of the application to help with skin-related disease diagnosis. Researchers have analyzed the ability of artificial intelligence-based approaches by performing image analyses based on perception, reasoning, and learning using the existing medium-sized image datasets. The most recent popular artificial intelligence-based segmentation methods are ANN models [49], Fuzzy C-Means [52], and genetic algorithms [50], [51]. Artificial intelligence usually utilizing an analytical planning to make machine learn without program it especially using existing dataset while deep learning further with neural networks that imitate the neurons in human brain and enclose with multiple architecture layers. However, starting from 2015, the deep learning approach starts to be implemented due to the introduction U-Net, which changes the research direction in many bio-medical applications.

##### B. Recent Works on Deep Learning Approach to Skin Lesion Segmentation

The deep learning approach has been proven to be state-of-the-art in supervised image segmentation applications. Despite the heavy complexity of deep learning models, more information from the raw images can be learned optimally rather than being designed by a human designer. Researchers have utilized various deep learning models to segment skin lesions, including U-Net [64], fully CNN (FCNN) [65], deep fully convolutional residual network (FCRN), and SegNet [66].

1) *FCNN architecture*: FCNN is a segmentation module with deeper encoder parts compared to the decoder parts. This model has been used in [53] to segment the skin lesions automatically. Besides that, researchers have also developed multi-stage FCNN for skin lesion segmentation by using a parallel integration method. An evaluation of the suggested technique has been tested using the ISBI 2016 dataset, which has revealed a high segmentation performance with a dice coefficient score of 91.18% and an accuracy of 95.51%. Yuan

et al. [54] then presented a modified deep FCNN-based method for skin lesion segmentation, which was evaluated using two different databases; one is from the ISBI 2016 database and the other one is from the PH2 database. The modified method was found to outperform the previously mentioned techniques. Jafari et al. [55] then used a pre-processing approach to start the image analysis as such the pixels are smoothed so that the extracted edges will be larger with reduced noise artifacts such as hair. Then, each pixel of the pre-processed image is fed into the FCNN to obtain 98.5% accuracy and 95% sensitivity performance.

TABLE III. COMPARISON BETWEEN CONVENTIONAL METHODS AND INTELLIGENCE-BASED METHODS FOR SKIN LESION SEGMENTATION

Techniques	Description	Advantages	Disadvantages
Edge and region-based method [46], [47]	This method performs edge detection first, followed by detecting and linking the edge pixels to form connected contour. Then, region identification is done by comparing surrounding pixels that have comparable grey levels.	This method performs very well, whereby the algorithm rapidly converges when a massive number of dermoscopy images is used for training.	Requires extensive computational resources since this technique utilizes pixel level information with the presence of noises.
Threshold-based method [48]	The lesion will be removed from the background skin in the image using the thresholding approach, followed by analysis on blue channel image.	Easy to implement and extremely fast.	This method does not perform well since it is sensitive to noises.
Intelligence-based method: - ANN models [49] - Genetic Algorithms [50], [51] - Fuzzy C-Means (FCM) [52]	The system is based on artificial intelligence approach, which is the most popular approach in automated dermatology field. This approach can better inform the patients on the state of skin lesion by increasing the sensitivity and accuracy of skin lesion examination.	Less costly and fast computational time during inference stage.	Difficult to obtain prior information on the number of clusters due to skin lesions complexity.
Deep Learning: - Fully Convolutional Neural Network (FCN) [53]–[55] - U-Net [56]–[58] - Deep Residual Network [59]–[61] SegNet [62], [63]	Deep learning can learn optimal hierarchical features from the raw images directly rather than hand-crafted features by the network designer.	Perform end-to-end learning where raw data will be processed into network, whereby the network learns for optimal automation of the task.	Requires large memory storage and more expensive computation compared to other methods.

2) *U-Net architecture*: U-Net was inspired by the FCNN, which consisted of equal distribution of encoder and decoder paths, coupled with few feedforward layers. In addition, pooling layers are used to down-scale the encoder feature maps, while standard interpolation is used to up-scale the decoder parts. Moreover, in between the encoder and decoder paths, there are shortcut skip connections. U-Net has become a well-known architecture after it has achieved outstanding segmentation performances in various medical applications with limited training datasets. Many recent studies have also been conducted based on this architecture for skin lesion segmentation purposes. Skin lesion segmentation performance was improved by adding dilated convolution and batch normalization layers to the U-Net architecture as proposed in [56]. Moreover, Iranpoor et al. [57] have proven that the modified U-Net has significantly improved the architecture efficiency by utilizing a pre-trained ResNet model in the encoder path. According to [58], the SkinNet system is also based on a modified U-Net architecture but uses dilated convolution to improve the encoder branch. Fig. 6 shows the modified U-Net architecture for improved segmentation performance.

3) *Deep FCRN*: The deep residual network is a unique network invention that uses skip connections to jump over some convolutional layers to build a pyramid-like structure. Generally, it consists of multiple feedforward convolutional

layers. A fully convolutional residual network was developed in [59] to segment skin lesions in dermoscopic images. This method proposed a deep CNN model with an effective training process that can be used to evaluate complex medical images. Li and Shen [60] have used the fully convolutional residual network to develop the lesion index calculation unit in 2018 for skin lesions segmentation. Besides that, Nathan and Kansal [61] have suggested a base of U-Net architecture with deep residual units as the backbone of encoders and decoders. Each downsampling block consists of one convolutional layer and two deep residual units to improve skin lesion segmentation performance.

4) *SegNet architecture*: This architecture is based on a deep neural network with a straight flow of an encoder network followed by a corresponding decoder network, whereby the final layer is formed for pixel-wise classification tasks. The feature maps are produced by implementing convolution with a filter bank in each encoder network. Additionally, a recent study in [62] has proposed a modified SegNet for skin lesion segmentation. The authors have reduced the total learned parameter of the architecture by lowering the downsampling and upsampling layers of the original SegNet. Similarly, the work in [63] has utilized the SegNet architecture in skin lesion segmentation application and has been found to be accurate based on PH2 dataset testing.

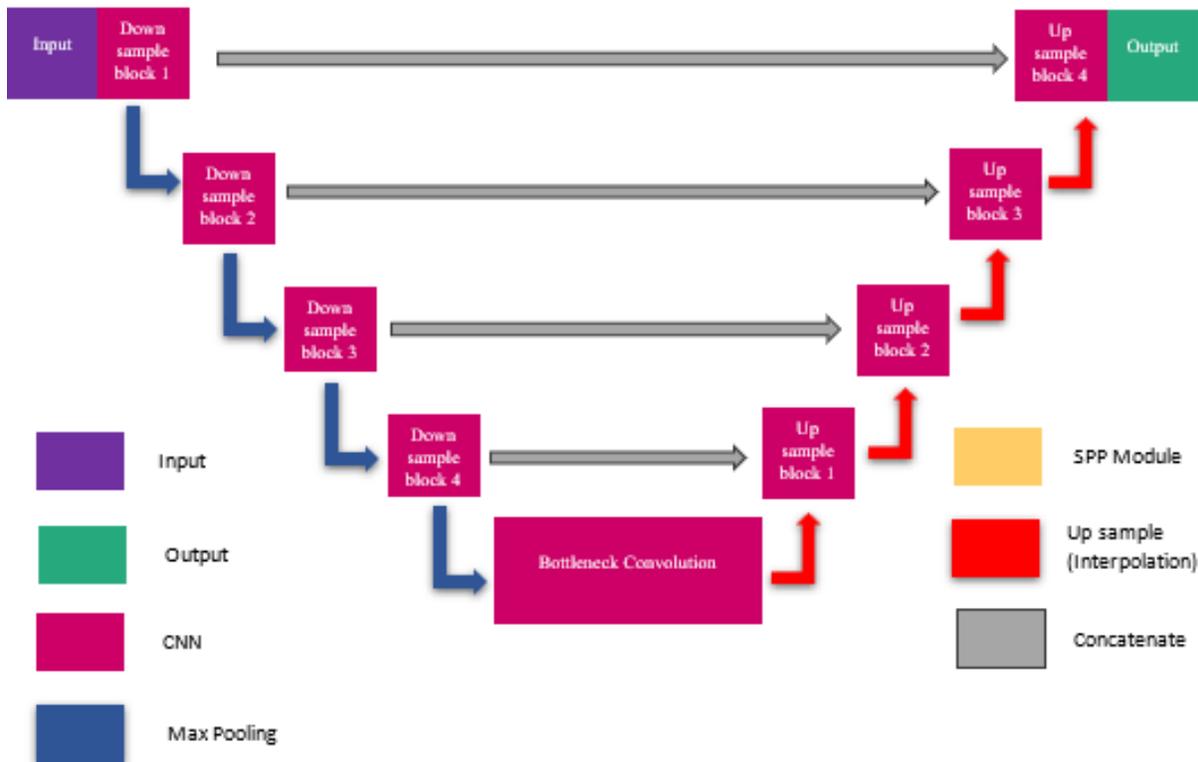


Fig. 6. Workflow Diagram of the U-Net Architecture.

## V. CONCLUSION

This paper presents a comparative study of state-of-the-art techniques, models, and methodologies for analyzing skin lesion images. This study has described the analysis process of skin lesions from segmentation to classification. Recently, researchers have paid more attention and effort into improving the accuracy of the diagnosis of skin lesions. However, some challenges remain difficult, especially in the interpretation of dermoscopic skin lesion images that contain noises such as bubbles, blood vessels, and hair. On the other hand, owing to the newest advancements in deep learning and its exceptional achievement in medical imaging, performance assessment of the deep learning approach for skin lesion segmentation and classification is worth to be reviewed. In this regard, state-of-the-art CNN models such as ResNet, Inception, Xception, DenseNet, and EfficientNet have generally shown excellent performances. However, these models require extensive computational resources with a long time to reach an optimal convergence state. This paper summarises the most important developments in this field and provides a complete discussion of the current approaches. Deep learning framework capabilities combined with pre and post-processing approaches are expected to improve future results and open the path for trustworthy screening and diagnostic systems.

Further research works should be tested using multiple open datasets to allow for better comparison. Besides that, more variations on the skin tone should be validated as most of the existing datasets are focusing on individuals with fair skin tones. Therefore, skin lesions with dark skin tone datasets also should be developed to produce a more robust testing platform that consists of all skin color types. A comprehensive analysis of various segmentation algorithms must be performed on the same dataset to achieve better accuracy, so that reliable results can be obtained. After that, a performance comparison of classification and segmentation models should also be tested on the same data set to produce a fair baseline model comparison.

## ACKNOWLEDGMENT

The authors would like to acknowledge funding from Universiti Kebangsaan Malaysia (Geran Universiti Penyelidikan: GUP-2019-008) and Ministry of Higher Education Malaysia (Fundamental Research Grant Scheme: FRGS/1/2019/ICT02/UKM/02/1).

## REFERENCES

- [1] S. R. Abdani, M. A. Zulkifley, and N. Hani Zulkifley, "A Lightweight Deep Learning Model for COVID-19 Detection," in 2020 IEEE Symposium on Industrial Electronics and Applications, ISIEA 2020, 2020, pp. 1–5.
- [2] N. A. Mohamed, M. A. Zulkifley, and A. Hussain, "On analyzing various density functions of local binary patterns for optic disc segmentation," in ISCAIE 2015 - 2015 IEEE Symposium on Computer Applications and Industrial Electronics, 2015, pp. 37–41.
- [3] A. Masood and A. A. Al-Jumaily, "Computer aided diagnostic support system for skin cancer: A review of techniques and algorithms," International Journal of Biomedical Imaging. 2013.
- [4] Adeyinka A.A., Viriri S. "Skin Lesion Images Segmentation: A Survey of the State-of-the-Art," In: Groza A., Prasath R. (eds) Mining Intelligence and Knowledge Exploration. MIKE 2018. Lecture Notes in Computer Science, vol 11308. Springer, Cham.
- [5] M. A. Al-masni, M. A. Al-antari, M. T. Choi, S. M. Han, and T. S. Kim, "Skin lesion segmentation in dermoscopy images via deep full resolution convolutional networks," vol. 162, pp. 221–231, 2018.
- [6] S. R. Abdani, M. A. Zulkifley, and A. Hussain, "Compact Convolutional Neural Networks for Pterygium Classification using Transfer Learning," in Proceedings of the 2019 IEEE International Conference on Signal and Image Processing Applications, ICSIPA 2019, 2019, pp. 140–143.
- [7] Esteva A, Kuprel B, Novoa RA, Ko J, Swetter SM, Blau HM, Thrun S, "Dermatologist-level classification of skin cancer with deep neural networks," Nature 542, pp. 115–118, 2017.
- [8] A. Krizhevsky, I. Sutskever, and G. E. Hinton, "ImageNet classification with deep convolutional neural networks," Commun. ACM, 2017.
- [9] Y. Lecun, Y. Bengio, and G. Hinton, "Deep learning," Nature. 2015.
- [10] A. Krizhevsky, I. Sutskever, and G. E. Hinton, "ImageNet classification with deep convolutional neural networks," 2012.
- [11] K. Simonyan and A. Zisserman, "Very deep convolutional networks for large-scale image recognition," 2015.
- [12] Szegedy, C., Liu, W., Jia, Y., Sermanet, P., Reed, S.E., Anguelov, D., Erhan, D., Vanhoucke, V., & Rabinovich, A., "Going deeper with convolutions," in Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition, 2015, pp. 1–9.
- [13] K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," in Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition, 2016, pp. 770–778.
- [14] F. Chollet, "Xception: Deep learning with depthwise separable convolutions," in Proceedings - 30th IEEE Conference on Computer Vision and Pattern Recognition, CVPR 2017, 2017, pp. 1800–1807.
- [15] G. Huang, Z. Liu, L. Van Der Maaten, and K. Q. Weinberger, "Densely connected convolutional networks," in Proceedings - 30th IEEE Conference on Computer Vision and Pattern Recognition, CVPR 2017, 2017, pp. 2261–2269.
- [16] M. Tan and Q. V. Le, "EfficientNet: Rethinking model scaling for convolutional neural networks," 2019.
- [17] U. Jamil and S. Khalid, "Comparative study of classification techniques used in skin lesion detection systems," in 17th IEEE International Multi Topic Conference: Collaborative and Sustainable Development of Technologies, IEEE INMIC 2014 - Proceedings, 2014, pp. 266–271.
- [18] M. Q. Khan, A. Hussain, S. U. Rehman, U. Khan, M. Maqsood, K. Mehmood, and M. A. Khan, "Classification of Melanoma and Nevus in Digital Images for Diagnosis of Skin Cancer," IEEE Access, vol. 7, pp. 90132–90144, 2019.
- [19] S. S. Han, W. Lim, M. S. Kim, I. Park, G. H. Park, and S. E. Chang, "Interpretation of the Outputs of a Deep Learning Model Trained with a Skin Cancer Dataset," The Journal of investigative dermatology, 138(10), pp. 2275–2277, 2018.
- [20] Russakovsky, O., Deng, J., Su, H., Krause, J., Satheesh, S., Ma, S., Huang, Z., Karpathy, A., Khosla, A., Bernstein, M., Berg, A. C., & Fei-Fei, L., "ImageNet Large Scale Visual Recognition Challenge," International Journal of Computer Vision, 2015, 115(3), pp. 211–252.
- [21] Brinker, T. J., Hekler, A., Enk, A. H., Klode, J., Hauschild, A., Berking, C., Schilling, B., Haferkamp, S., Schadendorf, D., Fröhling, S., Utikal, J. S., & von Kalle, C., "A convolutional neural network trained with dermoscopic images performed on par with 145 dermatologists in a clinical melanoma image classification task. European," Journal of Cancer (Oxford, England 1990), 111, pp. 148–154, 2019.
- [22] S. S. Han, M. S. Kim, W. Lim, G. H. Park, I. Park, and S. E. Chang, "Classification of the Clinical Images for Benign and Malignant Cutaneous Tumors Using a Deep Learning Algorithm," The Journal of investigative dermatology, 138(7), pp. 1529–1538, 2018.
- [23] K. M. Hosny, M. A. Kassem, and M. M. Foad, "Classification of skin lesions using transfer learning and augmentation with Alex-net," PLoS One, 2019.
- [24] M. A. Zulkifley and N. Trigoni, "Multiple-Model Fully Convolutional Neural Networks for Single Object Tracking on Thermal Infrared Video," IEEE Access, vol. 6, pp. 42790–42799, 2018.
- [25] M. A. Zulkifley, "Two streams multiple-model object tracker for thermal infrared video," IEEE Access, vol. 7, pp. 32383–32392, 2019.

- [26] X. Sun, J. Yang, M. Sun, and K. Wang, "A Benchmark for Automatic Visual Classification of Clinical Skin Disease Images," *Proc. ECCV*, pp. 206-222, 2016.
- [27] K. Thurnhofer-Hemsi and E. Domínguez, "Analyzing Digital Image by Deep Learning for Melanoma Diagnosis," In: *Proceedings of the 15th international work-conference on artificial neural networks (IWANN)*, pp. 270-279, 2019.
- [28] D. N. T. Le, H. X. Le, L. T. Ngo, and H. T. Ngo, "Transfer learning with class-weighted and focal loss function for automatic skin cancer classification," Sep. 2020.
- [29] S. S. Chaturvedi, J. v. Tembhumbe, and T. Diwan, "A multi-class skin Cancer classification using deep convolutional neural networks," *Multimedia Tools and Applications*, vol. 79, no. 39-40, Oct. 2020.
- [30] S. R. Hassan, S. Afroge, and M. Binte Mizan, "Skin Lesion Classification Using Densely Connected Convolutional Network," *IEEE Region 10 Symposium (TENSYP)*, pp. 750-753, 2020.
- [31] N. Gessert, M. Nielsen, M. Shaikh, R. Werner, and A. Schlaefler, "Skin lesion classification using ensembles of multi-resolution EfficientNets with meta data," *MethodsX*, vol. 7, 2020.
- [32] Brinker, T. J., Hekler, A., Enk, A. H., Klode, J., Hauschild, A., Berking, C., Schilling, B., Haferkamp, S., Schadendorf, D., Holland-Letz, T., Utikal, J. S., von Kalle, C., & Collaborators, "Deep learning outperformed 136 of 157 dermatologists in a head-to-head dermoscopic melanoma image classification task," *European journal of cancer (Oxford, England : 1990)*, 113, pp. 47-54, 2019.
- [33] M. Ratul, M. H. Mozaffari, W.-S. Lee, and E. Parimbelli, "Skin Lesions Classification Using Deep Learning Based on Dilated Convolution," *BioRxiv*, Jan. 2019, Art. no. 860700.
- [34] Gessert, N., Sentker, T., Madesta, F., Schmitz, R., Knip, H.C., Baltruschat, I.M., Werner, R., & Schlaefler, A., "Skin Lesion Classification Using CNNs With Patch-Based Attention and Diagnosis-Guided Loss Weighting," *IEEE Transactions on Biomedical Engineering*, 67, pp. 495-503, 2020.
- [35] A. M. Alqudah, H. Alquran, and I. A. Qasmieh, "Segmented and non-segmented skin lesions classification using transfer learning and adaptive moment learning rate technique using pretrained convolutional neural network," In *Journal of Biomimetics, Biomaterials and Biomedical Engineering 2019*; 42, pp. 67-78.
- [36] Akram, T., Lodhi, H.M.J., Naqvi, S.R. et al., "A multilevel features selection framework for skin lesion classification," *Human-centric Comput. Inf. Sci.*, vol. 10, pp. 1-26, 2020.
- [37] M. A. Al-masni, D. H. Kim, and T. S. Kim, "Multiple skin lesions diagnostics via integrated deep convolutional networks for segmentation and classification," *Comput. Methods Programs Biomed.*, vol. 190, 2020.
- [38] I. K. E. Purnama et al., "Disease Classification based on Dermoscopic Skin Images Using Convolutional Neural Network in Tele dermatology System," *International Conference on Computer Engineering, Network, and Intelligent Multimedia (CENIM)*, 2019, pp. 1-5.
- [39] Y. Guo, A. S. Ashour, L. Si, and D. P. Mandalaywala, "Multiple Convolutional Neural Network for Skin Dermoscopic Image Classification," in *2018 IEEE International Symposium on Signal Processing and Information Technology, ISSPIT 2018*, 2019, pp. 365-369.
- [40] V. Miglani and M. Bhatia, "Skin Lesion Classification: A Transfer Learning Approach Using EfficientNets," in *Advances in Intelligent Systems and Computing*, 2021, pp. 315-324.
- [41] N. Codella et al., "Skin lesion analysis toward melanoma detection 2018: A challenge hosted by the international skin imaging collaboration (ISIC)," *arXiv*. 2019.
- [42] N. C. F. Codella et al., "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 *Proceedings - International Symposium on Biomedical Imaging*, 2018, pp. 168-172.
- [43] T. Mendonca, P. M. Ferreira, J. S. Marques, A. R. S. Marcal, and J. Rozeira, "PH2 - A dermoscopic image database for research and benchmarking," in *Proceedings of the Annual International Conference of the IEEE Engineering in Medicine and Biology Society, EMBS*, 2013, pp. 5437-5440.
- [44] Gutman, David; Codella, Noel C. F.; Celebi, Emre; Helba, Brian; Marchetti, Michael; Mishra, Nabin; Halpern, Allan., "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)," pp. 3-7, 2016.
- [45] M. Filho, Z. Ma, and J. M. R. S. Tavares, "A Review of the Quantification and Classification of Pigmented Skin Lesions: From Dedicated to Hand-Held Devices," *Journal of Medical Systems*. 2015.
- [46] J. H. Jaseema Yasmin, M. Mohamed Sathik, and S. Zulaiyka Beevi, "Robust segmentation algorithm using LOG edge detector for effective border detection of noisy skin lesions," in *2011 International Conference on Computer, Communication and Electrical Technology, ICCCT 2011*, 2011, pp. 60-65.
- [47] H. Zhou, G. Schaefer, M. E. Celebi, F. Lin, and T. Liu, "Gradient vector flow with mean shift for skin lesion segmentation," *Comput. Med. Imaging Graph.*, vol. 35, no. 2, pp. 121-127, 2011.
- [48] M. Emre Celebi, Q. Wen, S. Hwang, H. Iyatomi, and G. Schaefer, "Lesion Border Detection in Dermoscopy Images Using Ensembles of Thresholding Methods," *Ski. Res. Technol.*, vol. 19, 2013.
- [49] Hogarty, D. T., Su, J. C., Phan, K., Attia, M., Hossny, M., Nahavandi, S., Lenane, P., Moloney, F. J., & Yazdabadi, A., "Artificial Intelligence in Dermatology—Where We Are and the Way to the Future: A Review," *American journal of clinical dermatology*, 21(1), pp. 41-47, 2020.
- [50] R. B. Aswin, J. A. Jaleel, and S. Salim, "Hybrid genetic algorithm - Artificial neural network classifier for skin cancer detection," in *2014 International Conference on Control, Instrumentation, Communication and Computational Technologies, ICCICT 2014*, 2014, pp. 1304-1309.
- [51] F. Xie and A. C. Bovik, "Automatic segmentation of dermoscopy images using self-generating neural networks seeded by genetic algorithm," *Pattern Recognit.*, vol. 46, no. 3, pp. 1012-1019.
- [52] A. A. I. Mohamed, M. M. Ali, K. Nusrat, J. Rahebi, A. Sayiner, and F. Kandemirli, "Melanoma Skin Cancer Segmentation with Image Region Growing Based on Fuzzy Clustering Mean," *Int. J. Eng. Innov. Res.*, vol. 6, no. 2, pp. 91-95, 2017.
- [53] L. Bi, J. Kim, E. Ahn, A. Kumar, M. Fulham, and D. Feng, "Dermoscopic Image Segmentation via Multistage Fully Convolutional Networks," *IEEE Trans. Biomed. Eng.*, vol. 64, no. 9, pp. 2065-2074, 2017.
- [54] Y. Yuan, M. Chao, and Y. C. Lo, "Automatic Skin Lesion Segmentation Using Deep Fully Convolutional Networks with Jaccard Distance," in *IEEE Transactions on Medical Imaging*, vol. 36, no. 9, pp. 1876-1886, Sept. 2017.
- [55] Jafari, M., Karimi, N., Nasr-Esfahani, E., Samavi, S., Sorousmehr, S.M., Ward, K., & Najarian, K., "Skin lesion segmentation in clinical images using deep learning," *23rd International Conference on Pattern Recognition (ICPR)*, pp. 337-342, 2016.
- [56] L. Liu, L. Mou, X. X. Zhu, and M. Mandal, "Skin Lesion Segmentation Based on Improved U-net," in *2019 IEEE Canadian Conference of Electrical and Computer Engineering, CCECE 2019*, 2019, pp. 1-4.
- [57] R. Iranpoor, A. S. Mahboob, S. Shahbandegan, and N. Baniasadi, "Skin lesion segmentation using convolutional neural networks with improved U-Net architecture," in *6th Iranian Conference on Signal Processing and Intelligent Systems, ICSPIS 2020*, 2020, pp. 1-5.
- [58] S. Vesal, N. Ravikumar, and A. Maier, "SkinNet: A deep learning framework for skin lesion segmentation," *arXiv*. 2018.
- [59] L. Yu, H. Chen, Q. Dou, J. Qin, and P. A. Heng, "Automated Melanoma Recognition in Dermoscopy Images via Very Deep Residual Networks," in *IEEE Transactions on Medical Imaging*, vol. 36, no. 4, pp. 994-1004, April 2017.
- [60] Y. Li and L. Shen, "Skin lesion analysis towards melanoma detection using deep learning network," *Sensors (Switzerland)*, vol. 18, 2018.
- [61] S. Nathan and P. Kansal, "Lesion Net-Skin Lesion Segmentation Using Coordinate Convolution and Deep Residual Units," pp. 1-8, 2012.
- [62] Q. C. Ninh, T. T. Tran, T. T. Tran, T. Anh Xuan Tran, and V. T. Pham, "Skin lesion segmentation based on modification of SegNet neural

- networks,” in Proceedings - 2019 6th NAFOSTED Conference on Information and Computer Science, NICS 2019, 2019, pp. 575–578.
- [63] P. Brahmhatt and S. N. Rajan, “Skin Lesion Segmentation using SegNet with Binary Cross-Entropy | Papers with Code,” International Conference on Artificial Intelligence and Speech Technology, 2019.
- [64] O. Ronneberger, P. Fischer, and T. Brox, “U-net: Convolutional networks for biomedical image segmentation,” In: Navab N., Hornegger J., Wells W., Frangi A. (eds) Medical Image Computing and Computer-Assisted Intervention – MICCAI 2015. MICCAI 2015. Lecture Notes in Computer Science, vol 9351. Springer, Cham.
- [65] J. Long, E. Shelhamer and T. Darrell, "Fully convolutional networks for semantic segmentation," in 2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Boston, MA, USA, 2015 pp. 3431-3440.
- [66] V. Badrinarayanan, A. Kendall, and R. Cipolla, “SegNet: A Deep Convolutional Encoder-Decoder Architecture for Image Segmentation,” in IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 39, no. 12, pp. 2481-2495, 1 Dec. 2017.