electronics

MDPI

Article

Noise Resilience in Dermoscopic Image Segmentation: Comparing Deep Learning Architectures for Enhanced Accuracy

Fatih Ergin [1], Ismail Burak Parlak [1,*], Mouloud Adel [1,2], Ömer Melih Gül [3,4] and Kostas Karpouzis [5,*]

[1] Department of Computer Engineering, Galatasaray University, NLPLAB, Ciragan Cad. No: 36, 34349 Istanbul, Turkey; erginfatih@yahoo.com (F.E.); mouloud.adel@univ-amu.fr (M.A.)
[2] Institut Fresnel, Aix Marseille University, CNRS, Centrale Marseille, 13013 Marseille, France
[3] Department of Computer Engineering, Bahçeşehir University, 34349 Istanbul, Turkey; omgul@itu.edu.tr
[4] Informatics Institute, Istanbul Technical University, 34485 Istanbul, Turkey
[5] Department of Communication, Media and Culture, Panteion University of Social and Political Science, 176 71 Athens, Greece
* Correspondence: bparlak@gsu.edu.tr (I.B.P.); kkarpou@panteion.gr (K.K.)

Abstract: Skin diseases and lesions can be ambiguous to recognize due to the similarity of lesions and enhanced imaging features. In this study, we compared three cutting-edge deep learning frameworks for dermoscopic segmentation: U-Net, SegAN, and MultiResUNet. We used a dermoscopic dataset including detailed lesion annotations with segmentation masks to help train and evaluate models on the precise localization of melanomas. SegAN is a special type of Generative Adversarial Network (GAN) that introduces a new architecture by adding generator and discriminator steps. U-Net has become a common strategy in segmentation to encode and decode image features for limited data. MultiResUNet is a U-Net-based architecture that overcomes the insufficient data problem in medical imaging by extracting contextual details. We trained the three frameworks on colored images after preprocessing. We added incremental Gaussian noise to measure the robustness of segmentation performance. We evaluated the frameworks using the following parameters: accuracy, sensitivity, specificity, Dice and Jaccard coefficients. Our accuracy results show that SegAN (92%) and MultiResUNet (92%) both outperform U-Net (86%), which is a well-known segmentation framework for skin lesion analysis. MultiResUNet sensitivity (96%) outperforms the methods in the challenge leaderboard. These results suggest that SegAN and MultiResUNet are more resistant techniques against noise in dermoscopic segmentation.

Keywords: artificial intelligence; image processing; machine learning; deep learning; neural networks

check for updates

Citation: Ergin, F.; Parlak, I.B.; Adel, M.; Gül, Ö.M.; Karpouzis, K. Noise Resilience in Dermoscopic Image Segmentation: Comparing Deep Learning Architectures for Enhanced Accuracy. *Electronics* **2024**, *13*, 3414. https://doi.org/10.3390/electronics13173414

Academic Editor: Luca Mesin

Received: 16 July 2024
Revised: 10 August 2024
Accepted: 15 August 2024
Published: 28 August 2024

Copyright: © 2024 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (https://creativecommons.org/licenses/by/4.0/).

1. Introduction

Over the past decade, computer-aided decision-making in medical deep learning has become a powerful tool that has outperformed traditional medical data processing tasks. New network architectures have been developed that allow machines to learn complex data structures and perform comprehensive data analysis. This has led to improved performance in several applications, such as enhanced medical image analysis, complex object characterization, and medical image segmentation on low-level morphological features. Deep learning networks (DNN) are improving the efficiency of lesion detection in medicine, from histological to radiological acquisitions.

Early cancer detection is considered one of the most complicated tasks of dermatology. Skin cancer is the most prevalent cancer globally. There are two primary types of skin cancer: melanoma and non-melanoma. While visual examination by a qualified dermatologist is the primary method for detection, many benign lesions can mimic these cancers. The repeated cases and follow-ups can also make it difficult to make accurate decisions. While the diagnosis is detailed by a specialist, DNNs would become a rapid tool for the patients at risk to reduce the time to diagnosis and the workload of physicians. Deep learning-based

medical segmentation is assessed using Ground Truth, a mask that outlines the entire area or volume in target images. The primary goal of deep learning methods is to iteratively learn computational model parameters using a training dataset to progressively enhance the model's ability to accomplish the desired goal. Once a computer is trained for a specific task, it can accurately perform the same task using numerous previously unseen data. Medical deep learning is distinguished from other medical computer-aided techniques due to its strong generalization ability [1–5].

Generative adversarial networks (GANs) point out a new paradigm in medical deep learning frameworks by addressing the synthesis of medical data, especially in medical imaging. GANs are composed of two principal sub-networks in competition; generator and discriminator. Medical data generation is attempted during the generator phase. The medical discriminator tries to understand real and fake medical data. Therefore, the competition leads to performing several tasks related to medical image understanding. GANs have also been integrated into cutting-edge applications of big data in computer vision. The synthesis of realistic facial images, the translation between images from one style to another, and the colorization of black-and-white images, which are considered hard tasks in computer vision, are accurately solved through GAN networks. Multilayered GAN hierarchies show promising results for skin lesion detection and evaluation. However, automatic lesion detection is complex due to factors such as inter-observer variability, inhomogeneity in image scale, and the challenge of obtaining annotated image corpora for deep learning. The International Skin Imaging Collaboration (ISIC) has shown promising results for clinical applications, but training datasets can cause variation in skin lesion detection. Therefore, GAN and transformer-based techniques are preferred due to their promising scores for automatic skin lesion analysis [6–8].

U-Net was introduced in 2015 and since then has become a benchmark in medical image segmentation. MultiResUNet has been developed as an enhancement of the U-Net architecture, which claims several key improvements that address some of the limitations of the original U-Net model. Our focus is comparing the enhanced U-Net model against the original one. Incorporating SegAN into our comparative study alongside U-Net and MultiResUNet offers a comprehensive analysis of medical image segmentation models on the ISIC 2017 dataset. SegAN, with its novel approach inspired by generative adversarial networks (GAN), brings a unique perspective to this comparison. SegAN's use of adversarial learning offers insights into how this technique performs in medical image segmentation, particularly in learning detailed and complex features. Including SegAN enables a direct comparison of adversarial learning with traditional and enhanced convolutional network approaches, offering a more comprehensive understanding of their respective strengths and weaknesses.

The imbalance problem is often more crucial in image classification than in image segmentation due to the differing nature of the tasks and the impact of class distribution on model performance. In image classification, where the goal is to categorize entire images into predefined classes, an imbalance in the dataset—where some classes are significantly underrepresented compared to others—can lead to biased models that favor the majority classes. This imbalance can cause the model to underperform on minority classes, resulting in poor generalization and reduced accuracy for less frequent categories. The challenge would be exaggerated in real-world scenarios; a recent ISIC classification challenge is where some classes, such as rare diseases or uncommon objects, may naturally occur infrequently. In contrast, image segmentation involves pixel-level classification, where each pixel is assigned a class label. While class imbalance can still impact segmentation performance, the problem is somewhat mitigated because segmentation models often operate within more localized regions of an image. Thus, while both tasks can suffer from class imbalance, the direct impact on overall classification accuracy and the risk of model bias make the imbalance problem particularly critical in image classification, demanding tailored strategies to ensure balanced and fair model performance across all classes. Moreover, in segmentation tasks, data augmentation techniques such as additive

noise, rotation, scaling, and cropping are commonly employed to enhance the diversity of training images and mitigate the risk of overfitting. These techniques not only increase the variability of the training data but also improve model generalization by exposing the model to various spatial transformations and augmentations. As a result, segmentation models benefit from these augmentations in terms of robustness and accuracy, making them less susceptible to the negative effects of class imbalance and data scarcity. Consequently, while these issues are still relevant, their impact is often less pronounced in segmentation compared to classification tasks in recent studies over the last decade.

We evaluated three DNN techniques under the same conditions in dermoscopic lesion segmentation. We remarked that lesion extraction is a relatively new challenge for computer-aided diagnosis in DNNs. The effective diagnosis is correlated with the quality of dermatologic images. The medical preprocessing routines grant complex tools such as enhancing filters to improve image quality. We measured the robustness of DNN techniques in a case of additive noise, which is frequently available in medical imaging, computer-aided diagnosis and telemedicine. Our study focused on the challenges posed by the characteristics of skin lesions, with the goal of developing promising results for diagnosing dermoscopic images. We measured the performance of cutting-edge medical DNN architectures based on hybrid features. We aimed to provide environmental Gaussian noise to simulate potential bottlenecks in skin lesion follow-ups. Even if lesion analysis focuses on image features, additional noise might distort initial ground truth during lesion follow-ups. We assessed the performance of DNN-based dermoscopic segmentation using ground truth, which is a binary image that provides the entire zone of the lesion in the images. Even though dermoscopic images might have intrinsic acquisition and quantization noise, they would be filtered in preprocessing steps. However, computer-aided diagnosis and telemedicine applications in hospital information systems might cause new additive noise (Gaussian) due to storage and indexing purposes. For this purpose, we pointed out the noise effect in skin lesion segmentation with DNN techniques. The rest of our study is given as follows. Section 2 reviews related works in dermatologic imaging techniques, image segmentation architectures and automatic diagnosis of skin lesions from dermoscopic images through DNN architectures. Section 3 details our proposed methodology given in Figure 1 and the ISIC 2017 image dataset on the basis of DNNs. Used neural network architectures, datasets, tools and deep learning frameworks are given with our corresponding representation through computational and statistical parameters. Section 4 shows our detection results with statistical evaluation based on Dice and Jaccard coefficients, sensitivity, specificity and accuracy. Section 5 highlights the scope of the study through the obtained results and presents a brief discussion. Finally, Section 6 concludes the assessment of examined neural networks in dermatologic lesion analysis using current cutting-edge and prospective trends.

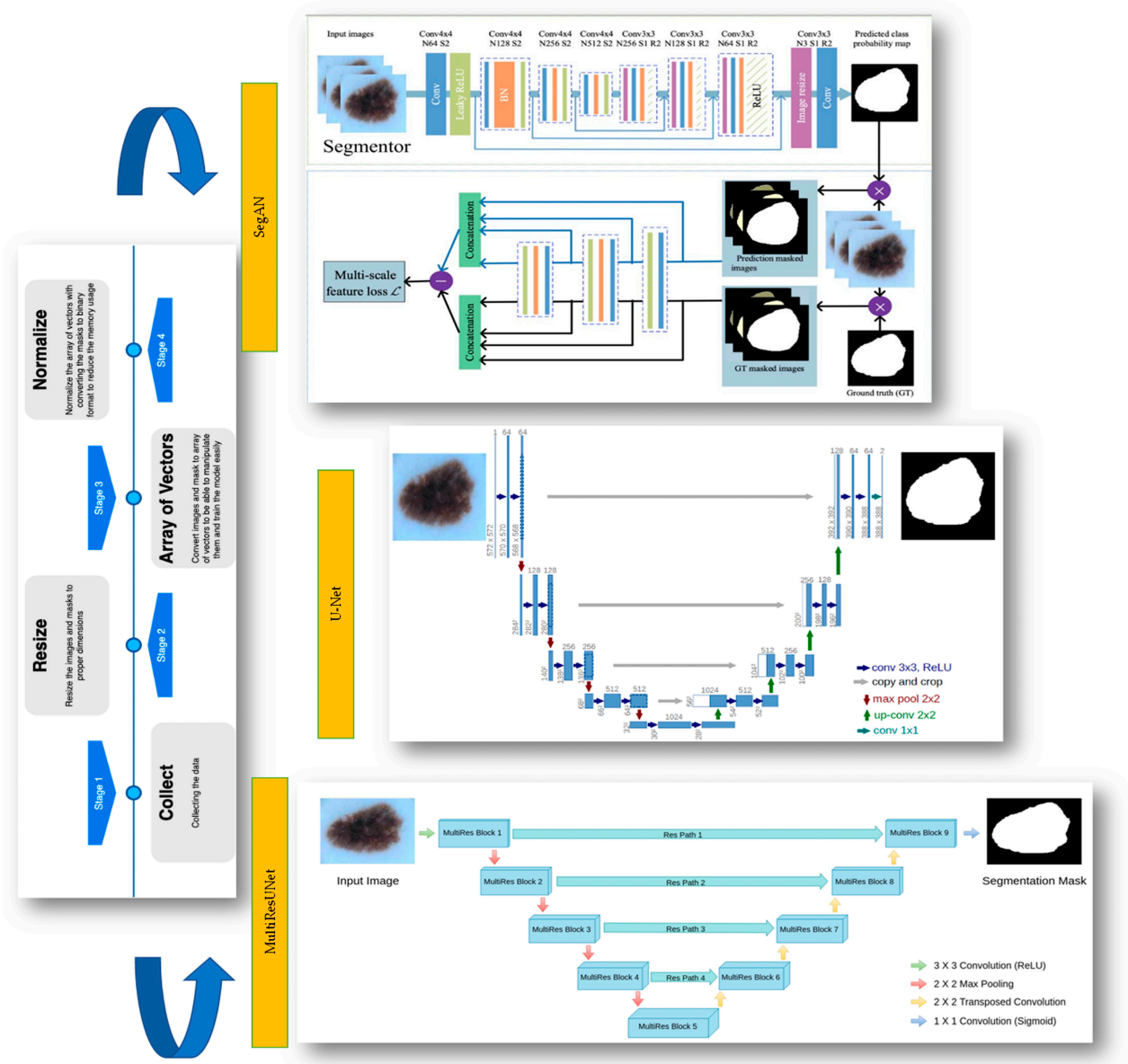


Figure 1. Flowchart of the proposed study.

2. Related Works

Medical image segmentation aims to determine the location and shape of the body part or structure within a 2D or 3D image automatically or semi-automatically [9]. The medical images are acquired using different modalities. A wide modality range and the high variability of human anatomy are the major differences in medical image segmentation. Medical images are divided into several interests related to the problem definition to detect or segment the tumor or mass. Irregularities, blurred vision borders, low contrast between lesions and skin, and air bubbles are some of the various artifacts that make segmentation medical imaging challenging [10].

Medical imaging techniques are concerned with creating medical images to be able to examine internal structures of the body without opening it up [11]. Traditional Photography (TP) is a well-known technique that makes visualizing and monitoring the top layer of the lesion possible [12]. The dermoscopy imaging technique is a real-time noninvasive diagnostic imaging technique that is more successful in distinguishing melanoma concentration than traditional photography [13]. Multispectral imaging provides information in both spectral and spatial domains. MI systems increase accuracy by calibrating image intensity

and controlling exposure time automatically with the help of a multispectral camera that includes different optical filters selected by the problem definition. MI is used in medical imaging to support detecting lesions about 2 mm in size [13]. Confocal Laser Scanning Microscopy (CLSM) is an imaging technique that provides real-time details of skin morphology and provides images with the same resolution as traditional microscopes [14]. CLSMs are very sensitive for clinical applications but they are relatively expensive to use there. Ultrasonography, which is also known as diagnostic sonography, is another imaging technique that is used to create medical imaging to create internal body parts using high-frequency broadband sound waves. Because different tissues behave differently under these sound waves, the images generated using the waves are reflected by tissue [15]. Calculating the depth of skin cancer is the focused usage of ultrasonography for this kind of project.

Fully convolutional networks (FCNs) indicate that convolutional neural networks are obtained by dismantling the fully connected layers from deep CNNs [16]. FCNs are built on traditional classification networks such as VGG [17], AlexNet [18], GoogLeNet [19], and ResNet [20]. Convolutional layers are used instead of fully connected layers to produce outputs with the same size inputs instead of classification scores, which are the outputs of CNNs. FCNs consist of two units: encoding and decoding. Convolution and subsampling operations are performed in the encoding unit to encode the lower dimensional latent space. Deconvolution and upsampling are performed in the decoding unit, which guarantees the same size output as the input. Since FCNs do not include fully connected layers, it is faster to obtain an image with respect to the classical CNNs.

The publication of AlexNet [18] in 2012 triggered a paradigm change in image segmentation, and since then, deep learning methods have provided prominent results and become the state-of-the-art in this area in recent years [21]. Long et al. [22] proposed an FCN from the CNNs known to be successful in semantic segmentation. They adapted well-known classification networks such as AlexNet, VGG, GoogleLeNet to fully convolutional networks. Then, to create a successful segmentation, they combined semantic details from a deep layer and the appearance details from a shallow layer to define a new skip architecture. The proposed architecture achieved remarkable results compared to state-of-the-art models on PASCAL VOC. Ronneberger et al. [23] built a new neural network aimed to be able to obtain accurate results with insufficient data by using them more effectively. U-Net, the proposed network, is based on classical FCNs and consists of two symmetric paths, namely contracting and expanding, which are responsible for capturing the context and enabling precise localization, respectively. The new neural network proved its success with very few images by winning the International Symposium on Biomedical Imaging (ISBI) 2015 Cell Tracking Challenge. In addition to being able to work with insufficient data, U-Net offers prominent results for training duration with images with relatively higher resolutions, such as 512 $\times$ 512. In the following years, new studies showed that the proposed U-shaped network was more successful than C-Means Clustering in the ISBI 2017 challenge dataset [24].

Yuan et al. [25] introduced an improved version of the FCN model using Jaccard distance as loss function. The aim of this network is to increase segmentation accuracy with solving common dermoscopic image problems such as imbalanced skin and lesion pixels, the existence of various artifacts, and irregular lesion borders. The proposed network achieved better results than the other state-of-the-art networks in the ISBI 2016 challenge and PH2 databases. Moreover, they presented a new skin lesion segmentation framework based on Fully Convolutional Deconvolutional Neural Networks (CDNN) [26]. Their main focus was to improve network architecture rather than additional pre- and post-processing steps. A Rectified Linear Unit (ReLU) was used for the activation of each layer in the network except the output layer. The internal covariate shift is reduced by adding batch normalization to the output of the CD layers. The proposed CDNN model won the ISBI 2017 challenge. They improved their other skin lesion segmentation architectures by using smaller kernels to optimize the discriminant capacity of their newly proposed

neural network. The improved version of the previous work is evaluated on the ISBI 2017 challenge dataset and placed among the top 21 in the ranking. Bi et al. [27] proposed a multistage FCN to increase the segmentation accuracy of classical FCNs. In this network, the first-stage FCN focused on learning localization information and coarse appearance, whereas the second-stage FCN focused on the subtle characteristics of the lesion boundaries. A parallel integration method is also introduced to combine the results of the first- and second-stage FCNs. Yu et al. [28] presented a novel deep neural network architecture consisting of two stages called segmentation and classification. The network combines a deep learning method with a local descriptor encoding strategy for dermoscopic image recognition. A pretrained large image dataset is used to extract deep representations of a rescaled image. After that, extracted descriptors are aggregated and encoded with a Fisher vector to obtain global features. In the end, the global features are used to classify images with the help of a support vector machine. The proposed network is a fully convolutional residual network (FCRN) and took second place in the segmentation category of the ISBI 2016 challenge. Al-Masni et al. [29] developed a framework for skin lesion segmentation via full-resolution convolutional networks (FrCN). This method eliminated subsampling layers and learned the full-resolution features directly. It was tested with ISBI 2017 challenge and PH2 datasets and has achieved better results against the well-known state-of-the-art segmentation networks, such as U-Net, SegNet and FCN.

Li et al. [30] introduced a new dense deconvolutional network (DDN) for skin lesion segmentation. The proposed network is based on residual learning. It consists of three main parts namely dense convolutional layer, hierarchical supervision (HS), and chained residual pooling (CRP). Dimensions of the input and output images remain unchanged in DDLs. CRP helps to capture contextual background features while HS is responsible for improving the prediction mask. They tested the network with the ISBI 2017 dataset, and it achieved 86.6% Dice coefficient indices. Xue et al. [31] proposed an Adversarial Neural Network (GAN), called SeGAN, which is a deep neural network aimed at increasing the accuracy of medical image segmentation. Classical GANs are not as good as expected in providing gradient feedback to the network, because their output is single, which may not represent pixel-level details of images. Segmentation label maps are created with the help of a newly created FCN-based segmentor network with a new activation function. Another significant improvement in the proposed network is the multi-scale L1 loss function aimed to extract both local and global features, which represent the relations between pixels. Peng et al. [32] introduced a new adversarial network-based segmentation architecture consisting of CNN-based discrimination and U-Net-based segmentation networks. This utilized generative adversarial network was evaluated on the ISBI 2016 challenge dataset and achieved a 97.0% accuracy rate. Tu et al. [33] proposed an adversarial network-based deep learning framework focused on solving the imbalanced lesion-background problem. The segmentation block of the proposed network is an encoder–decoder network with a dense-residual block. Deep supervision is utilized with a multi-scale loss function. The network was evaluated on the ISBI 2017 challenge dataset and gained better segmentation results than the other state-of-the-art methods participating in that challenge. Tschandl et al. [34] introduced a new FCN where pretrained ImageNet weights are being used to feed the network on ResNet34 layers, which are reused as encoding layers. The evaluation results showed that using pretrained weights improved the segmentation score on the ISBI 2017 challenge dataset.

Ninh et al. [35] proposed a SegNet architecture-based FCN framework, which aimed to decrease the number of upsampling and downsampling layers of classical SegNet architecture to reduce the learned parameters. The proposed network was evaluated on the ISBI 2017 challenge dataset and gained sufficient results in terms of the Jaccard Index and Dice coefficient. Mirikharaji et al. [36] proposed a deep CNN framework that focused on segmenting skin lesions. The main focus of the proposed network was the use of two different annotation sets consisting of reliable and unreliable annotations. The reliable annotations were marked by experts and showed reliable segmentation results.

This reweighting was performed by a newly deployed meta-learning approach. The proposed network shows that using different levels of annotation noise on weighting affects the segmentation results and model robustness positively. Sarker et al. [37] proposed a lightweight GAN framework, called MobileGAN, aiming to reduce the number of training parameters while keeping the segmentation accuracy high. They combined the channel attention module with the 1D non-bottleneck factorization networks for the generator part of the GAN. MobileGAN was trained with the ISIC 2018 training dataset and was evaluated with the ISBI 2017 challenge dataset. Compared to state-of-the-art models such as FCN, U-Net, or SegNet, the results showed that the proposed network had fewer parameters, about 2.3 million, and achieved considerable scores. Lei et al. [38] proposed a GAN framework aiming to increase skin lesion segmentation accuracy and won the first part of the ISBI 2017 challenge. The segmentation part of the proposed GAN was constructed with a skip connection and dense convolution U-Net, while the discrimination part consisted of a dual discriminator module. One of the discriminators was responsible for increasing the detection of boundaries, while the other one was responsible for learning the contextual information. Zafar et al. [39] proposed an automated neural network architecture aimed at segmenting skin lesions accurately. Res-Unet, the proposed network, is a combination of two well-known neural networks in image segmentation, namely U-Net and ResNet. The other major improvement in this network is using image inpainting for hair removal. It was evaluated on the ISBI 2017 challenge and PH2 datasets and obtained Jaccard Indices of 77.2% and 85.4%, respectively. Xie et al. [40] introduced a CNN variant called MB-DCNN, which consisted of three sub-CNNs, namely a coarse segmentation network, a mask-guided segmentation network, and an enhanced segmentation network, respectively. The first network was responsible for creating coarse masks, which had been used on the next network to classify the lesions. The third network was a segmentation network fed from the second classification network. There were learning transfers between networks to increase the segmentation accuracy. MB-DCNN was tested with the ISBI 2017 challenge and PH2 datasets, and it achieved Jaccard indices of 80.4% and 89.4%.

In recent years, several advanced methods have emerged for dermoscopic image segmentation that build on or extend the foundational techniques provided by models like U-Net, SegAN, and MultiResUNet. These newer methods often incorporate innovations in deep learning architectures, attention mechanisms, and transfer learning to improve segmentation performance. DeepLabV3+ uses dilated convolutions to capture multi-scale contextual information and a depthwise separable convolution for efficient feature extraction. It provides detailed and accurate segmentation by integrating context from different scales to deal with handling variations in lesion sizes and shapes [41]. Attention U-Net enhances the standard U-Net by incorporating attention gates that help the model focus on the relevant parts of the image. Attention gates selectively emphasize the features that are important for the segmentation task, improving the model's ability to differentiate between lesions and the background [42]. UNet++ introduces nested skip pathways and deep supervision to improve model performance by refining feature extraction and enhancing the learning of multi-scale features [43]. Attention-based Residual U-Net combines U-Net with attention mechanisms and residual connections. It uses attention blocks to focus on relevant features and residual connections to improve training stability and performance [43]. Attention Residual U-Net integrates residual learning with U-Net architecture [44]. ResUNet uses residual blocks within the U-Net framework to address vanishing gradient issues and improve model training. It enhances feature extraction and improves segmentation of complex lesions [45]. V-Net applies volumetric (3D) convolutions to handle data with three-dimensional context to capture spatial context in three dimensions [46]. DenseNet-UNet combines DenseNet with U-Net architecture to uses dense connections to improve feature reuse and gradient flow, enhancing segmentation accuracy [47]. SWIN-UNet uses transformer blocks integrated with U-Net for segmentation to handles complex patterns and fine details effectively [48]. TransUNet combines Transformer-based architecture with

U-Net and integrates the Transformer's self-attention mechanism with the U-Net structure to capture both local and global features [49].

Shehzad et al. [50] presented an innovative method for diagnosing skin cancer by leveraging deep ensemble learning techniques to enhance diagnostic accuracy. Their ensemble method integrates various convolutional neural networks (CNNs) to improve the robustness and generalization of skin cancer detection systems. By aggregating predictions from different models, the approach aims to reduce individual model biases and errors, ultimately leading to more reliable and precise skin cancer diagnoses. Almuayqil et al. [51] explored a hybrid deep learning approach to enhance early diagnosis of skin diseases. The study introduces a method that fuses multiple types of features, including both visual and clinical data, to improve diagnostic accuracy. By combining various feature extraction techniques with deep learning models, the authors aim to capture a comprehensive set of characteristics from skin images, leading to more precise identification of early signs of skin conditions. Gouabou et al. [52] introduced a novel deep learning technique designed to tackle the challenges of classifying skin lesions in imbalanced datasets. The proposed method, called End-to-End Decoupled Training (EEDT), addresses the problem of long-tailed distributions, where certain classes are underrepresented compared to others. By decoupling the training process into separate stages for feature extraction and classification, EEDT improves the model's ability to learn from minority classes without being overwhelmed by the majority classes. The approach enhances classification performance and robustness in dermoscopic image analysis, demonstrating significant improvements over conventional methods in handling class imbalance. Ibraheem et al. [53] explored a method for staging melanocytic skin neoplasms by leveraging high-level pixel-based features extracted from dermoscopic images. The authors propose a novel approach that utilizes advanced image processing techniques to identify and analyze detailed pixel-level characteristics, which are then used to determine the stage of skin neoplasms. By focusing on these high-level features, the method aims to enhance the accuracy and precision of skin cancer staging, providing more detailed and informative assessments compared to traditional approaches.

3. Materials and Methods

The International Skin Imaging Collaboration (ISIC) is a global initiative aimed at improving the diagnosis and treatment of skin diseases, particularly skin cancer, through the use of advanced imaging technologies and machine learning. ISIC provides various datasets for research and development purposes. These datasets differ in several key aspects. ISIC Challenge Datasets are specific to annual challenges organized by ISIC. Each challenge may focus on different aspects of skin imaging, such as automated diagnosis or segmentation of lesions. The datasets for these challenges are tailored to the specific goals and evaluation metrics of each year's competition. Image types are dermatoscopic images for examining skin lesions, clinical images; standard photographs of skin lesions taken under regular lighting conditions and histopathological images; microscopic images of skin tissue samples, usually obtained from biopsies. The datasets contain classification labels, segmentation masks and metadata, including additional information about the images, such as patient demographics (age, gender), lesion location, and clinical history. ISIC has organized annual challenges from 2016 to 2020, each focusing on different aspects of skin lesion analysis. ISIC 2016 was centered on skin lesion classification classified into various categories such as melanoma, basal cell carcinoma, squamous cell carcinoma, and benign conditions. ISIC 2017 (also known as ISBI 2017) focused on skin lesion analysis towards melanoma detection and included detailed lesion annotations with segmentation masks to help train and evaluate models on the precise localization of melanoma. ISIC 2018 was on skin lesion segmentation and classification tasks with segmentation masks offering a more comprehensive set of annotations to support both tasks. ISIC 2019 emphasized skin lesion segmentation and classification with a focus on melanoma. Finally, ISIC 2020 was on the semantic segmentation of skin lesions. In a nutshell, each year's challenge built on the

previous ones, gradually incorporating more complex tasks and annotations to advance the field of dermatological image analysis. In segmentation studies, both ISIC 2017 and 2019 were used for different purposes. ISIC 2019 generally provides more advanced and detailed segmentation masks, specifically benefiting from higher resolution and greater precision in annotations. ISIC 2017 includes melanoma data but also covers a broader range of lesions, which might dilute the focus on melanoma-specific segmentation. In skin lesion segmentation, there are no studies with additive noise to ensure the quality of segmentation for melanomas using these datasets.

We preferred ISBI Challenge 2017 [54]—Skin Lesion Analysis Towards Melanoma Detection: Lesion Segmentation dataset in this study as it covers a broader range of lesions, which might dilute the focus on melanoma-specific segmentation. This dataset has separate training, validation and test data. The training dataset consists of 2000 dermoscopic JPEG images and related masks in PNG format. The dataset includes various types of lesions such as malignant melanoma, nevus and seborrhoeic keratosis. Sample images are given with corresponding masks where the first row represents the original images, and the second row shows the ground truth aka the corresponding masks. The masks were generated by a medical expert. This expert employed a combination of manual techniques and semi-automated methods for accuracy. Each mask is presented in a grayscale format, where the pixel values are designated as black to represent the background and white to indicate the lesion areas. Figure 1 illustrates the general mechanism. There are also validation and test datasets, which contain 150 and 600 images, respectively. The results are based on several common image similarity metrics, which are given in a related section. The images are of various dimensions and the neural network model cannot handle relatively big images because of the inner constraints in the architecture and memory. Therefore, all images have been resized into the same dimension to reduce memory consumption and to increase the accuracy as a preprocessing stage. Arrays of mask files have been converted to uint8 to reduce the size of the masks.

Irregularity and images in different scales are common conditions in medical imaging samples. Neural networks aiming to obtain accurate results in medical imaging should be able to overcome these kinds of problems. Dealing with images of different scales is an ongoing situation for medical imaging even if there are some studies about it, and because of that, it is not possible to say that this issue has been definitively resolved. Szegedy et al. [19] proposed Inception architecture built on convolutional layers with various kernel sizes to minimize the difference in the scales between images. MultiResUNet has an improvement similar to Inception architecture. In addition to the 3×3 convolution layer in the classic U-Net, MultiResUnet has convolution layers in different kernels such as 5×5 and 7×7. Figure 1 shows the evolution of the MultiRes blocks with different attempts, resulting from the different uses of these kernels. These MultiRes blocks have replaced the sequences of two convolutional layers in the vanilla U-Net.

One of the significant improvements in U-Net is using the skip connections between the encoder and decoder. Thus, features that are lost during pooling are recovered and transferred from an encoder block to a decoder block. It is expected that the features sent by the encoder to the decoder are low level while the features in the decoder are expected to be high level. They thought that this might cause a semantic gap between the encoder and decoder and proposed another improvement called Res path, which can be seen in Figure 1. The proposed Res path consists of convolutional layers connected by residual connections to make learning easier [55]. The features being sent from an encoder to decoder are transmitted over the Res paths instead of classical skip connections of U-Net. The proposed MultiResUNet framework is shown in Figure 1 with all improvements. MultiResUNet has been tested and evaluated through several datasets including Murphy lab, ISBI 2012, ISIC 2018, CVC-ClinicDB, and BraTS17 with different modalities such as fluorescence microscopy, electron microscopy, dermoscopy, endoscopy, and MRI, respectively. Their results show that the MultiResUNet offers more accurate results than the classical U-Net for all 5 different datasets especially in dermoscopy and endoscopy images.

SegAN consists of two networks, segmentor and critic, which can be seen in Figure 1, similar to the generator and discriminator networks of conventional GANs. It looks like a game in GAN, where the segmentor tries to fool the critic with the samples it creates. The main difference arises with the multi-scale loss function. While two separate loss functions are defined for the generator and discriminator in GAN, segmentor and critic use a common multi-scale loss function to force both networks of SegAN to learn local and global features, which acquire relations between pixels. SegAN is trained with the BRATS 2015 dataset and achieved remarkable results compared to other state-of-the-art models, including U-Net, in the field of semantic segmentation.

Our study is composed of three steps: preprocessing, implementations of networks, and evaluation, as given in Figure 1. During the preprocessing, image normalization procedures have been applied to data, including image resizing and the conversion of file formats. Moreover, data size has been augmented by creating additional image files in different noise levels. Because our main focus is comparing the proposed network under the same conditions, the preprocessing stages were kept the same for a fair comparison of U-Net, MultiResUNet and SegAN. U-Net and MultiResUnet have been trained for 200 epochs with a batch size of 8 and binary cross entropy loss function. As the performance did not improve, the epoch size has been kept as 200. An Adam optimizer has been used as an optimizer with the default parameters stated in the original paper. Furthermore, SegAN has been trained for 200 epochs with a batch size of 200 and an adaptive learning rate for the Adam optimizer, which started from 2.0×10^{-4} and multiplied by a decay rate of 0.5 every 25 epochs. Several learning and decay rates have been tried but the given parameters were found optimal like the original article. Early stopping has been used for all networks. If the performances of models stopped improving after a certain number of epochs, 30 was set to stop the training.

Additive noise is frequently applied in image classification studies rather than image segmentation studies due to its impact on model training and performance. In classification tasks, the primary goal is to recognize and categorize entire images based on their overall content, and additive noise can effectively simulate a range of real-world distortions and variations that might affect image quality. However, in image segmentation, where the objective is to precisely delineate and classify pixel-level details and boundaries within an image, the introduction of additive noise can disrupt the fine-grained spatial information crucial for accurate segmentation. The noise can distort boundaries and small features, making it challenging for segmentation algorithms to maintain precision. While additive noise is valuable for training robust classification models, its application in segmentation studies requires careful consideration due to the potential degradation of crucial spatial information needed for accurate pixel-wise analysis. Thus, its application and comparison in deep learning architectures are less studied in dermoscopic studies.

Our experiments have been designed by providing additional Gaussian noise into dermoscopic data. Five different noise experiments have been designed on DNNs using Gaussian noise distribution given as in the initial image I_i;

$$I_f = I_i + I_n \tag{1}$$

I_f and I_n denote the final image and noise level, respectively. The Gaussian noise is generated as follows

$$I_n(z) \quad \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(z-\mu)^2}{2\sigma^2}} \tag{2}$$

$I_n(z)$ represents the noise level in a single-channel image. Our dermoscopic data have been represented as color images. Therefore, the additive expression noise has been used for all RGB channels.

The Dice coefficient measures the overlap between two samples. For image segmentation, it compares the pixels in the ground truth mask (actual segmentation) and the predicted mask (segmentation predicted by the model). The Dice coefficient ranges from 0

to 1, where 1 indicates perfect overlap. The Jaccard coefficient, also known as Intersection over Union, compares the similarity and diversity of sample sets. For segmentation tasks, it measures the overlap between the predicted mask and the ground truth. The Jaccard coefficient also ranges from 0 to 1, with 1 representing a perfect match. They are particularly effective for evaluating how well the model segments an image, considering both the true positives and the size of both the predicted and actual segments. This is especially important in medical image analysis, where the precise delineation of an area similar to a tumor is critical.

4. Results

The experiments show that both SegAN and MultiResUNet achieved almost the same Dice coefficient result for the noise-free images, but vanilla U-Net did not achieve similar results in terms of the same similarity metrics. It is not as successful as the others. MultiResUNet is slightly more successful than SegAN if they are compared using the Jaccard coefficient. The detailed results are given in Tables 1–3 through statistical metrics. Although the Dice results of SegAN (Figures 2 and 3) and MultiResUNet (Figures 4 and 5) are very close for the noiseless datasets, the Dice results differ for all models as the noise level increases. Both MultiResUnet and SegAN achieved their best results around the epoch size of 50. The results were found similar after this point. Figures 6–8 point out the Dice results of models at different levels of epoch size. As the number of epochs increases, we note that the increase in the Dice score looks similar in U-Net and SegAN. However, MultiResUNet differs with its ability through epoch size. Tables 1–3 show the evaluation results of dermoscopic images from the evaluation dataset with different statistical rates. For all DNNs, the scores of the Dice and Jaccard indices decrease, and the noise level increases. However, statistical parameters are less affected by noise due the nature of the melanoma properties in the ISIC 2017 dataset. Figure 9 points out how Dice coefficient rates drop with Gaussian noise levels. Furthermore, Figure 10 shows the results of detection variations with different models. We remark that SegAN is more robust than vanilla U-Net and MultiResUNet at increased levels of Gaussian noises. When the noise level is 50%, the Dice results of MultiResUNet decreased up to 28%, U-Net's decreased up to 23%, while SegAN's decreased up to 53%. SegAN introduces fake skin lesions during the generator level, and the discriminator makes a decision after the training as to whether the test image is a lesion. The noise added during the training phase of SegAN makes the model more successful against noisy data. Figure 10 shows the change in segmentation accuracy through three DNNs where additive noise levels increase. We note that SegAN is more robust than vanilla U-Net and MultiResUNet at increased levels of Gaussian noises. Although it is not possible to create a model that fits all dataset, the main objective is to present a model that best generalizes them. Figures 9 and 10 are the outputs obtained by evaluating two pictures with two models with different levels of noise. While SegAN gives more successful results for the image in Figure 9, MultiResUNet is more successful with the image in Figure 10. As can be seen from that comparison, there is no precise superiority of the models to each other for certain data.

Table 1. Evaluation of U-Net segmentation at different Gaussian noise levels.

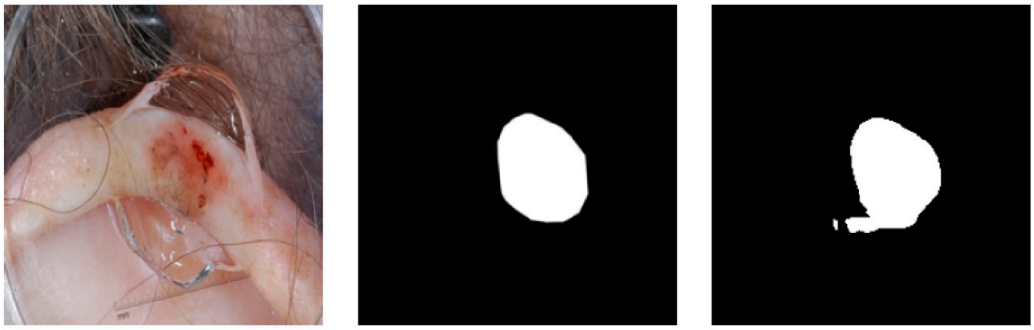
Gaussian Noise	Accuracy	Dice	Jaccard	Specificity	Sensitivity
Challenge Winner	0.934	0.849	0.765	0.975	0.825
Li and Chen [56]	0.950	0.839	0.753	0.974	0.855
0%	0.861	0.643	0.534	0.882	0.735
10%	0.855	0.615	0.508	0.871	0.756
20%	0.842	0.558	0.448	0.850	0.768
30%	0.828	0.474	0.370	0.831	0.781
40%	0.785	0.256	0.173	0.794	0.623
50%	0.795	0.234	0.163	0.786	0.757

Table 2. Evaluation of SegAN segmentation at different Gaussian noise levels.

Gaussian Noise	Accuracy	Dice	Jaccard	Specificity	Sensitivity
Challenge Winner	0.934	0.849	0.765	0.975	0.825
Li and Chen [56]	0.950	0.839	0.753	0.974	0.855
0%	0.923	0.811	0.696	0.924	0.899
10%	0.812	0.557	0.400	0.844	0.623
20%	0.813	0.551	0.393	0.841	0.632
30%	0.813	0.545	0.387	0.839	0.633
40%	0.809	0.537	0.379	0.844	0.608
50%	0.811	0.536	0.378	0.835	0.636

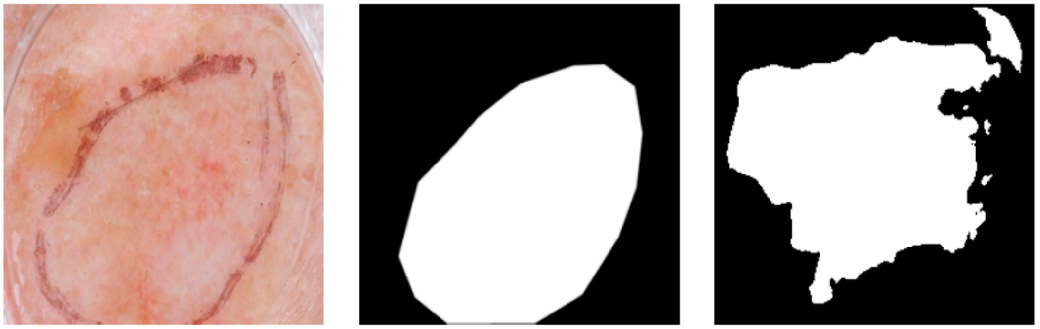
Table 3. Evaluation of MultiResUNet segmentation at different Gaussian noise levels.

Gaussian Noise	Accuracy	Dice	Jaccard	Specificity	Sensitivity
Challenge Winner	0.934	0.849	0.765	0.975	0.825
Li and Chen [56]	0.950	0.839	0.753	0.974	0.855
0%	0.922	0.816	0.722	0.948	0.964
10%	0.905	0.770	0.674	0.902	0.892
20%	0.882	0.739	0.624	0.870	0.878
30%	0.855	0.605	0.478	0.794	0.840
40%	0.806	0.434	0.306	0.810	0.703
50%	0.787	0.285	0.196	0.784	0.772



(a) Original image (b) Ground truth (c) Result(dice=0.82)

Figure 2. SegAN high Dice score; 0.82 at 0% of Gaussian noise level.



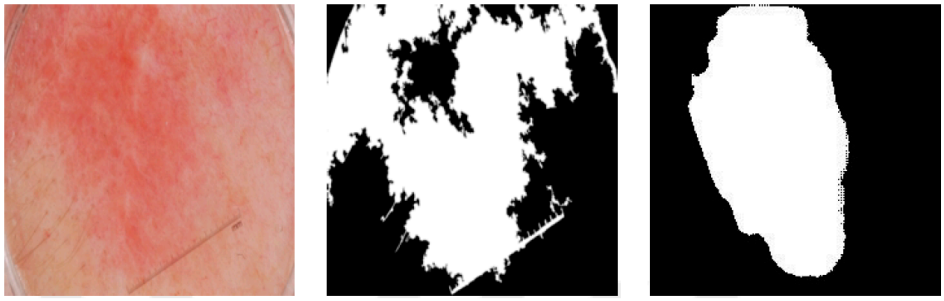
(a) Original image (b) Ground truth (c) Result(dice=0.50)

Figure 3. SegAN low Dice score; 0.5 at 0% of Gaussian noise level.



(a) Original image (b) Ground truth (c) Result(dice=0.90)

Figure 4. MultiResUNet high Dice score; 0.90 at 0% of Gaussian noise level.



(a) Original image (b) Ground truth (c) Result(dice=0.54)

Figure 5. MultiResUNet low Dice score; 0.54 at 0% of Gaussian noise level.

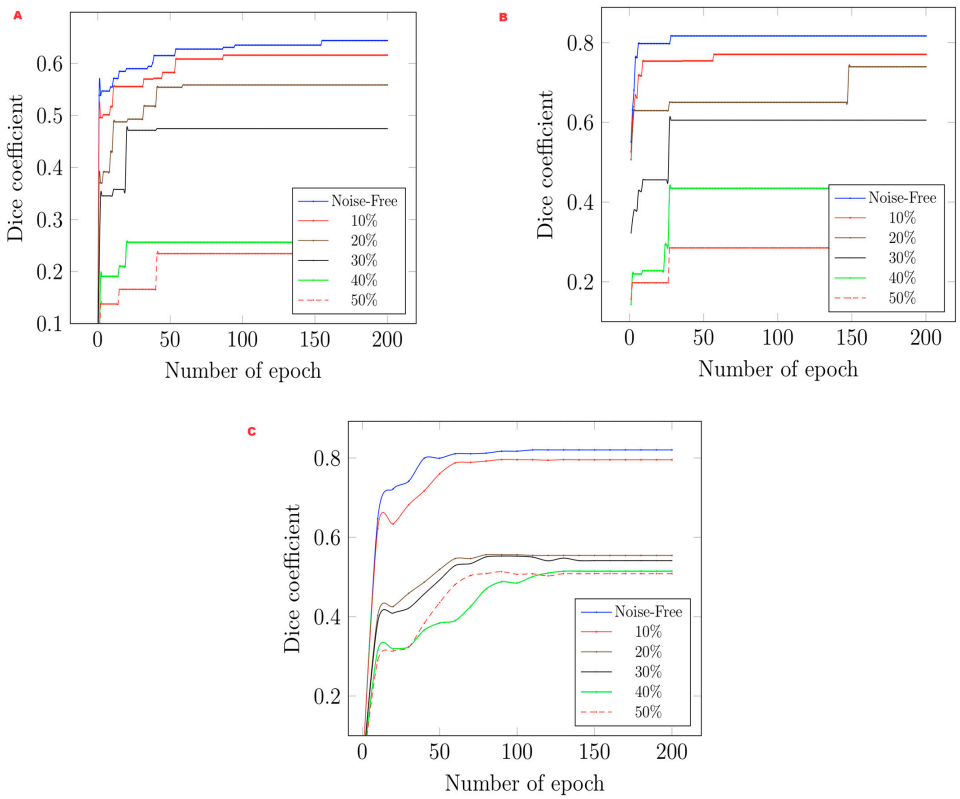


Figure 6. Dice results at different Gaussian noise levels through the number of epochs for (**A**) U-Net, (**B**) SegAN, (**C**) MultiResUNet DNN techniques.

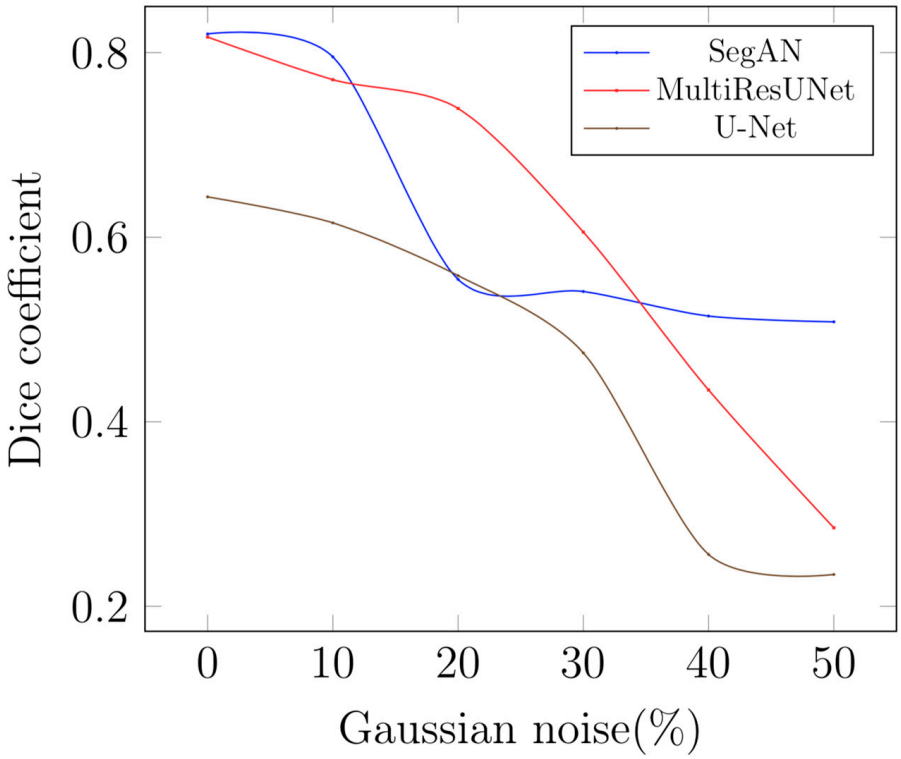


Figure 7. Comparison of DNN models at different noise levels.

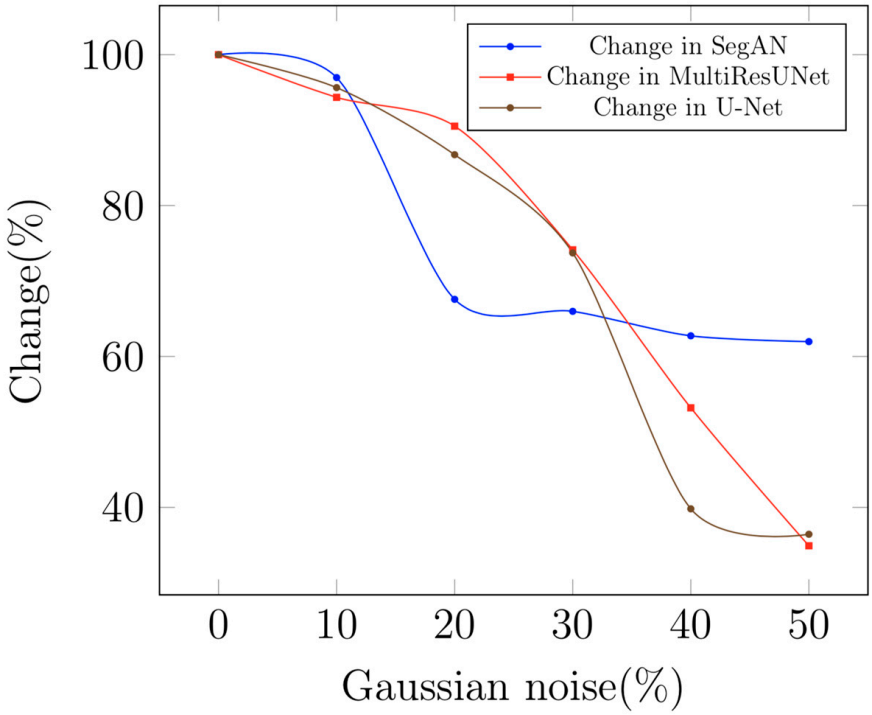


Figure 8. Variation in the Dice coefficient through additive Gaussian noise.

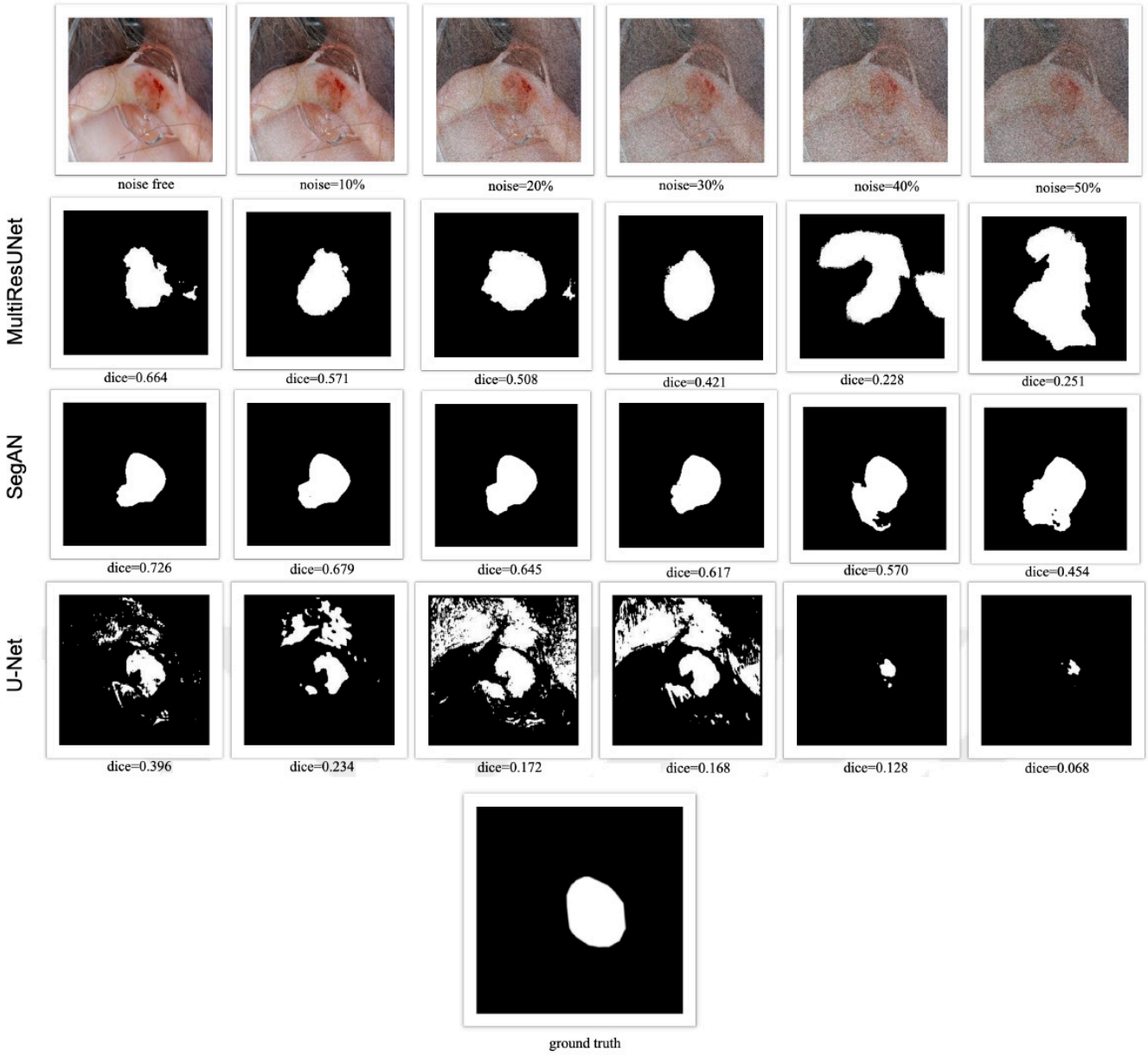


Figure 9. Dice results of the same image for all networks at different Gaussian noise levels. The images in a column from top to bottom show the input and segmentation results for MultiResUnet, SegAN, and U-Net, respectively. SegAN shows more accurate outputs.

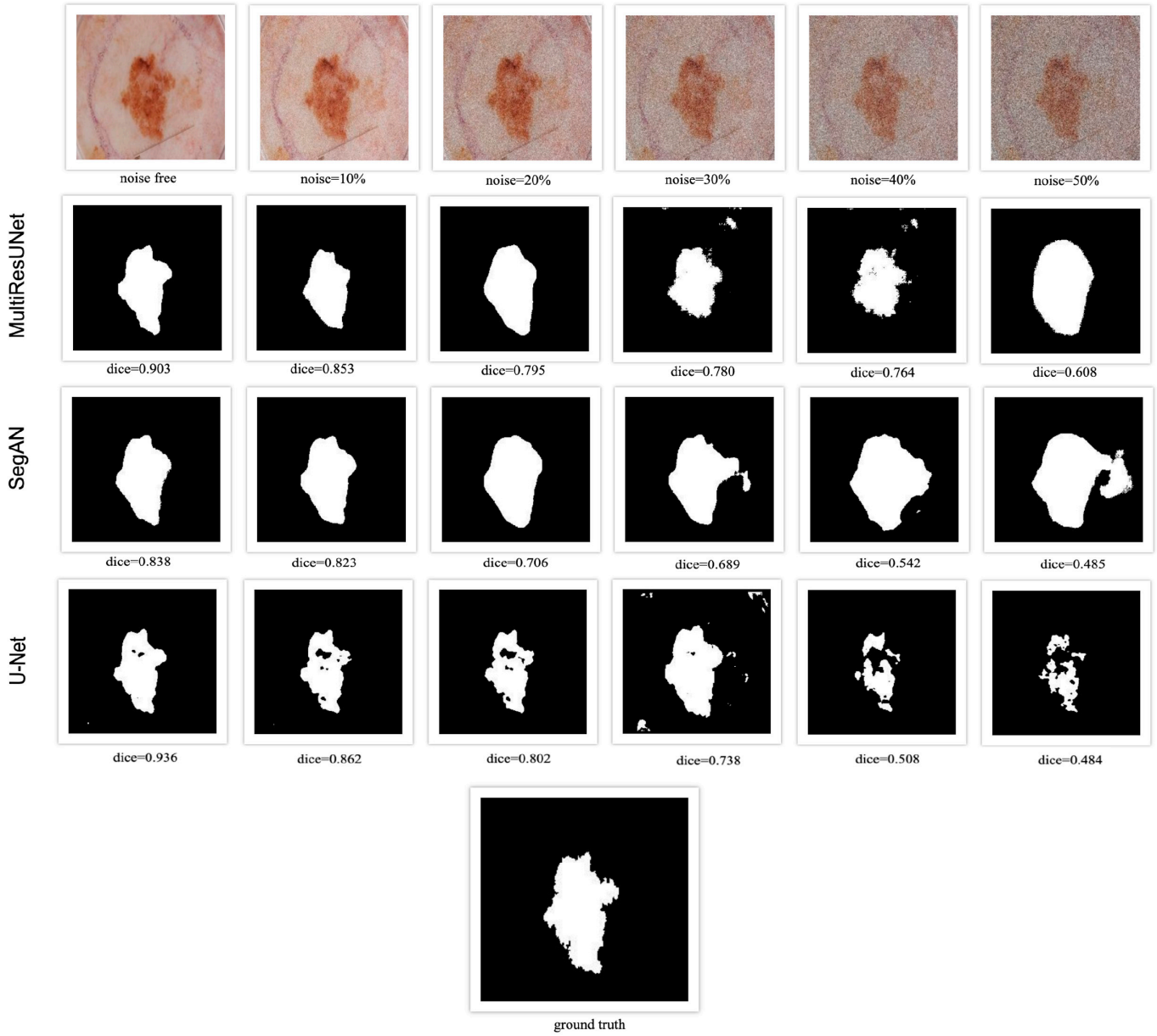


Figure 10. Dice results of the same image for all networks at different Gaussian noise levels. The images in each column from top to bottom show the input and segmentation results for MultiResUnet, SegAN, and U-Net, respectively. U-Net shows more accurate outputs.

5. Discussion

Skin lesions or tumors may have harmful impacts on human health. The early analysis of potential moles can increase the survival rate by using appropriate detection paradigms. Advanced technologies such as deep learning are actually used in several fields of medicine to increase the diagnosis of illnesses in the early stages. Image-based analysis can help oncologists or surgeons when detecting skin tumors. Since the dermatologist makes the final medical decision regarding skin lesions, DNNs would serve to speed up the diagnosis of at-risk patients to reduce the time to diagnosis and the workload of physicians [1–5].

Using grayscale images reduces the input dimensionality, leading to faster training and testing of our models. This efficiency simplifies the computational workload and speeds up both the learning and application phases of the models. Grayscale images remove potentially redundant color information, facilitating the development of more general models. By focusing on structural and textural features rather than color, our

models become more adept at identifying essential patterns relevant to the segmentation task. This results in models that are less likely to overfit to color-specific features and more capable of generalizing across various medical imaging scenarios where color may not be a distinguishing factor.

U-Net was introduced in 2015, and since then has become a benchmark in medical image segmentation. MultiResUNet has been developed as an enhancement of the U-Net architecture, which claims several key improvements that address some of the limitations of the original U-Net model. Our focus is comparing the enhanced U-Net model against the original one. Incorporating SegAN into our comparative study alongside U-Net and MultiResUNet offers a comprehensive analysis of medical image segmentation models on the ISBI 2017 dataset. Including SegAN enables a direct comparison of adversarial learning with traditional and enhanced convolutional network approaches, offering a more comprehensive understanding of their respective strengths and weaknesses.

The advantage of MultiResUNet is its architecture, and it outperforms U-NET. MultiResUNet addresses discrepancies between the encoder and decoder features in U-Net by introducing Res paths for more homogeneous feature maps. Additionally, the MultiRes block in MultiResUNet better captures multiscale features, which is crucial for varied medical images. These are the key enhancements in the MultiResUNet architecture that allow for better accuracy against U-Net in certain medical imaging tasks. MultiResUNet and SegAN achieve similar accuracy despite having completely different architectures for noise-free images. MultiResUNet achieves slightly better results in terms of the Dice and Jaccard metrics, which are common metrics in image segmentation tasks, as shown in Tables 2 and 3. MultiResUNet is more resistant to noisy conditions and more robust in medical digital imaging, computer-aided diagnosis and telemedicine. Even if MultiResUNet performance drops down gradually, we note that additive noise has a drastic effect on SegAN.

We addressed the segmentation of skin lesions, especially melanoma by providing a unified hierarchy to compare several deep-learning methods. Medical image segmentation has been performed using U-Net, SegAN and MultiResUNet. The dataset was created along ISIC for the ISBI 2017 Challenge and has been enriched by adding Gaussian noises at different levels. In image segmentation tasks, the issues of imbalance and data augmentation often have less critical impact compared to other areas like image classification due to the nature of segmentation challenges and techniques. Unlike the classification problem in recent dermoscopic challenges, which relies on categorizing entire images into one of several classes, segmentation involves labeling each pixel in an image, which can inherently distribute the class labels more evenly across different regions. This pixel-wise labeling reduces the likelihood of severe class imbalance affecting model performance. Moreover, in segmentation tasks, data augmentation techniques such as additive noise, rotation, scaling, and cropping are commonly employed to enhance the diversity of training images and mitigate the risk of overfitting. These techniques not only increase the variability of the training data but also improve model generalization by exposing the model to various spatial transformations and augmentations. As a result, segmentation models benefit from these augmentations in terms of robustness and accuracy, making them less susceptible to the negative effects of class imbalance and data sparsity. Consequently, while these issues are still relevant, their impact is often less pronounced in segmentation compared to classification tasks in recent studies over the last decade.

In image analysis, Gaussian noise is often preferred over other noise types due to its statistical properties and its impact on model robustness. Gaussian noise, characterized by its bell-shaped probability distribution, is mathematically well-defined and closely resembles the types of noise typically encountered in real-world imaging scenarios. Unlike salt-and-pepper noise or speckle noise, which introduce synthetic or irregular artifacts, Gaussian noise affects pixel values with a smooth, continuous variation. This makes it particularly useful for simulating realistic variations and perturbations in dermoscopic images, which helps in training more robust models and allows generalizing the vulnerability of

models based on light sensors. Moreover, Gaussian noise's predictable distribution allows for easier modeling and incorporation into data augmentation strategies, enabling effective and controlled noise injection that can enhance a model's ability to generalize across different conditions. Its widespread use in standard image processing algorithms and statistical models further facilitates compatibility and integration, making Gaussian noise a preferred choice for improving the performance and reliability of image analysis systems.

Our results showed that MultiResUNet and SegAN give more accurate results compared to vanilla U-Net through Gaussian noise, and MultiResUNet and SegAN provide high scores for all of the datasets. Nawaz et al. proposed a multi-stage approach using deep learning and fuzzy k-means clustering on noise-free ISBI 2017 dataset [57]. Their segmentation accuracy and specificity were found to be 95% and 98%. The winner of the ISBI 2017 challenge, Yuan et al. [25,26], measured a Jaccard index of 78.4% in the segmentation results. Li and Shen measured 75% accuracy in ISBI 2017 segmentation and they provided a summary table for challenge leaderboards. We note that the sensitivity score of 96.4% outperforms all approaches in the challenge. The Jaccard Index and Dice coefficients together bring complementary aspects to image segmentation. Even if our MultiResUNet Jaccard score is better than SegAN, we note that they have almost equal Dice scores in validation.

6. Conclusions

Recent studies on image segmentation and classification used different datasets for deep learning architectures. ISIC datasets have become the gold standard for ensuring better benchmark analysis through statistical scores. In conclusion, ISIC 2017 and 2019 datasets were used in the image segmentation analysis. We focused on noise resilience in melanoma segmentation. While there may not be direct references asserting ISIC 2017 as superior to ISIC 2019, ISIC 2019 offers more advanced features in image resolution. On the other hand, ISIC 2017 emphasized melanoma detection, providing valuable data for developing algorithms specifically aimed at identifying melanoma. Moreover, ISIC 2017 featured a wide range of skin lesions, which can be beneficial for developing models that generalize across different types of skin conditions. ISIC 2019 reduced emphasis on a broader range of lesions. Deep neural networks applied to dermoscopic images are promising for the segmentation of skin diseases and the follow-ups for post-operative treatments. The classification results were trained using a non-invasive dermoscopic imaging modality, widely available in clinics. Color images improve the segmentation accuracy and render lesion boundaries according to low-level image features.

MultiResUNet and SegAN generally offer superior performance over U-Net, especially due to their advanced features and enhancements. U-Net remains a strong baseline, but SegAN and MultiResUNet can provide additional improvements in segmentation accuracy, particularly in complex scenarios like melanoma detection. While U-Net, SegAN, and MultiResUNet are not the newest architectures in the field of image segmentation, they remain influential and widely used due to their foundational contributions and effectiveness. Newer DNN models build upon and extend the concepts introduced by these methods, aiming to improve performance and handle more complex segmentation tasks in instance or semantic medical image segmentation. However, SegAN, and MultiResUNet are considered robust techniques in melanoma segmentation.

Additive noise is a common technique used in data augmentation and robustness testing in machine learning, including in medical image segmentation tasks. It is worth noting that specific studies applying additive noise directly to ISIC segmentation datasets are less common. Researchers often apply noise in a general context of data augmentation or robustness testing and may not always detail specific datasets in the context of noise. Data augmentation is preferred in image classification problems due to imbalance challenges. Current limitations are the training steps for high-resolution images, environmental noise during dermoscopic acquisition, limited data availability for large-scale dermatologic lesion analysis and inconsistent acquisitions for follow-ups. Even though dermoscopic images

might have intrinsic acquisition and quantization noise, they would be filtered in the preprocessing steps. However, computer-aided Diagnosis and telemedicine applications in hospital information systems might cause new additive noise (Gaussian) due to storage and indexing purposes. For this purpose, we pointed out the noise effect in skin lesion segmentation with DNN techniques. In dermoscopic images, which are used for skin lesion analysis and melanoma detection, various types of noise can affect image quality and the performance of segmentation and classification algorithms. DNN techniques such as adding synthetic noise during training can help models become more robust to real-world noise. In future steps, we will address the problem by creating a database where we will locate the melanoma features through different color features such as texture and contour with a follow-up paradigm. We will compare the segmentation performance using ISIC 2019 and ISIC 2020 by generalizing segmentation aspects of melanoma and their features. Therefore, melanoma prediction would serve to explore the spatial characteristics of skin lesions.

Author Contributions: Conceptualization, F.E., I.B.P., M.A. and Ö.M.G.; methodology, F.E., I.B.P. and M.A.; validation, F.E., I.B.P., Ö.M.G. and K.K.; investigation, F.E., I.B.P., M.A., Ö.M.G. and K.K.; resources, F.E. and I.B.P.; data curation, F.E., I.B.P., M.A. and Ö.M.G.; writing—original draft preparation, F.E., I.B.P., M.A., Ö.M.G. and K.K.; writing—review and editing, I.B.P., M.A., Ö.M.G. and K.K.; supervision, I.B.P., M.A., Ö.M.G. and K.K. project administration, F.E., I.B.P., M.A., Ö.M.G. and K.K. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: No new data were created as a consequence of this investigation. Data can be downloaded from: https://biomedicalimaging.org/2017/challenges/ (accessed on 1 June 2024).

Conflicts of Interest: The authors declare no conflicts of interest.

References

1. Cirrincione, G.; Cannata, S.; Cicceri, G.; Prinzi, F.; Currieri, T.; Lovino, M.; Militello, C.; Pasero, E.; Vitabile, S. Transformer-Based Approach to Melanoma Detection. *Sensors* **2023**, *23*, 5677. [CrossRef] [PubMed]
2. Mukhlif, A.A.; Al-Khateeb, B.; Mohammed, M.A. Incorporating a Novel Dual Transfer Learning Approach for Medical Images. *Sensors* **2023**, *23*, 570. [CrossRef]
3. Bandy, A.D.; Spyridis, Y.; Villarini, B.; Argyriou, V. Intraclass Clustering-Based CNN Approach for Detection of Malignant Melanoma. *Sensors* **2023**, *23*, 926. [CrossRef] [PubMed]
4. Singh, S.K.; Abolghasemi, V.; Anisi, M.H. Skin Cancer Diagnosis Based on Neutrosophic Features with a Deep Neural Network. *Sensors* **2022**, *22*, 6261. [CrossRef] [PubMed]
5. Yang, S.; Wang, L. HMT-Net: Transformer and MLP Hybrid Encoder for Skin Disease Segmentation. *Sensors* **2023**, *23*, 3067. [CrossRef]
6. Giacopelli, G.; Migliore, M.; Tegolo, D. NeuronAlg: An Innovative Neuronal Computational Model for Immunofluorescence Image Segmentation. *Sensors* **2023**, *23*, 4598. [CrossRef]
7. Li, Y.; Xu, C.; Han, J.; An, Z.; Wang, D.; Ma, H.; Liu, C. MHAU-Net: Skin Lesion Segmentation Based on Multi-Scale Hybrid Residual Attention Network. *Sensors* **2022**, *22*, 8701. [CrossRef]
8. Dong, Y.; Wang, L.; Cheng, S.; Li, Y. FAC-Net: Feedback Attention Network Based on Context Encoder Network for Skin Lesion Segmentation. *Sensors* **2021**, *21*, 5172. [CrossRef]
9. Merjulah, R.; Chandra, J. Classification of myocardial ischemia in delayed contrast enhancement using machine learning. In *Intelligent Data Analysis for Biomedical Applications*; Elsevier: Amsterdam, The Netherlands, 2019; pp. 209–235.
10. Guo, Y.; Ashour, A.S. Neutrosophic sets in dermoscopic medical image segmentation. In *Neutrosophic Set in Medical Image Analysis*; Elsevier: Amsterdam, The Netherlands, 2019; pp. 229–243.
11. Kasban, H.; El-Bendary, M.; Salama, D. A comparative study of medical imaging techniques. *Int. J. Inf. Sci. Intell. Syst.* **2015**, *4*, 37–58.
12. Feit, N.E.; Dusza, S.W.; Marghoob, A.A. Melanomas detected with the aid of total cutaneous photography. *Br. J. Dermatol.* **2004**, *150*, 706–714. [CrossRef]

13. Aljanabi, M.; Jumaa, F.; Aftan, A.; Salah, M.; Alkafaji, S.; Alanı, N.; Al-Tameemi, Z.; Al-mamoori, D. Various types of skin tumors lesion medical imaging (stlmi) of healthy and unhealthy moles a review and computational of: Segmentation, classification, methods and algorithms various types of skin tumors lesion medical imaging (stlmi) of healthy and unhealthy moles a review and computational of: Segmentation, classification, methods and algorithms. In *IOP Conference Series Materials Science and Engineering*; IOP Publishing: Bristol, UK, 2019.
14. Gerger, A.; Koller, S.; Kern, T.; Massone, C.; Steiger, K.; Richtig, E.; Kerl, H.; Smolle, J. Diagnostic applicability of in vivo confocal laser scanning microscopy in melanocytic skin tumors. *J. Investig. Dermatol.* **2005**, *124*, 493–498. [CrossRef]
15. Sahuquillo, P.; Tembl, J.I.; Parkhutik, V.; Vázquez, J.F.; Sastre, I.; Lago, A. The study of deep brain structures by transcranial duplex sonography and imaging resonance correlation. *Ultrasound Med. Biol.* **2013**, *39*, 226–232. [CrossRef]
16. Ulku, I.; Akagunduz, E. A survey on deep learning-based architectures for semantic segmentation on 2d images. *arXiv* **2019**, arXiv:1912.10230. [CrossRef]
17. Simonyan, K.; Zisserman, A. Very deep convolutional networks for largescale image recognition. *arXiv* **2014**, arXiv:1409.1556.
18. Krizhevsky, A.; Sutskever, I.; Hinton, G.E. Imagenet classification with deep convolutional neural networks. In Proceedings of the Advances in Neural Information Processing Systems 25 (NIPS 2012), Lake Tahoe, NV, USA, 3–6 December 2012; pp. 1097–1105.
19. Szegedy, C.; Liu, W.; Jia, Y.; Sermanet, P.; Reed, S.; Anguelov, D.; Erhan, D.; Vanhoucke, V.; Rabinovich, A. Going deeper with convolutions. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Boston, MA, USA, 7–12 June 2015; pp. 1–9.
20. He, K.; Zhang, X.; Ren, S.; Sun, J. Deep residual learning for image recognition. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Las Vegas, NV, USA, 27–30 June 2016; pp. 770–778.
21. Quang, N.H.; Thao, L.T. Automatic skin lesion analysis towards melanoma detection. In Proceedings of the 2017 21st Asia Pacific Symposium on Intelligent and Evolutionary Systems (IES), Hanoi, Vietnam, 15–17 November 2017; IEEE: Piscataway, NJ, USA, 2017; pp. 106–111.
22. Long, J.; Shelhamer, E.; Darrell, T. Fully convolutional networks for semantic segmentation. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Boston, MA, USA, 7–12 June 2015; pp. 3431–3440.
23. Ronneberger, O.; Fischer, P.; Brox, T. U-net: Convolutional networks for biomedical image segmentation. In *International Conference on Medical Image Computing and Computer-Assisted Intervention*; Springer: Berlin/Heidelberg, Germany, 2015; pp. 234–241.
24. Lin, B.S.; Michael, K.; Kalra, S.; Tizhoosh, H.R. Skin lesion segmentation: U-nets versus clustering. In Proceedings of the 2017 IEEE Symposium Series on Computational Intelligence (SSCI), Honolulu, HI, USA, 27 November–1 December 2017; IEEE: Piscataway, NJ, USA, 2017; pp. 1–7.
25. Yuan, Y.; Chao, M.; Lo, Y.-C. Automatic skin lesion segmentation using deep fully convolutional networks with jaccard distance. *IEEE Trans. Med. Imaging* **2017**, *36*, 1876–1886. [CrossRef]
26. Yuan, Y.; Lo, Y.-C. Improving dermoscopic image segmentation with enhanced convolutional-deconvolutional networks. *IEEE J. Biomed. Health Inform.* **2017**, *23*, 519–526. [CrossRef] [PubMed]
27. Bi, L.; Kim, J.; Ahn, E.; Kumar, A.; Fulham, M.; Feng, D. Dermoscopic image segmentation via multistage fully convolutional networks. *IEEE Trans. Biomed. Eng.* **2017**, *64*, 2065–2074. [CrossRef]
28. Yu, Z.; Jiang, X.; Zhou, F.; Qin, J.; Ni, D.; Chen, S.; Lei, B.; Wang, T. Melanoma recognition in dermoscopy images via aggregated deep convolutional features. *IEEE Trans. Biomed. Eng.* **2018**, *66*, 1006–1016. [CrossRef] [PubMed]
29. Al-Masni, M.A.; Al-antari, M.A.; Choi, M.-T.; Han, S.-M.; Kim, T.-S. Skin lesion segmentation in dermoscopy images via deep full resolution convolutional networks. *Comput. Methods Programs Biomed.* **2018**, *162*, 221–231. [CrossRef] [PubMed]
30. Li, H.; He, X.; Zhou, F.; Yu, Z.; Ni, D.; Chen, S.; Wang, T.; Lei, B. Dense deconvolutional network for skin lesion segmentation. *IEEE J. Biomed. Health Inform.* **2018**, *23*, 527–537. [CrossRef]
31. Xue, Y.; Xu, T.; Huang, X. Adversarial learning with multi-scale loss for skin lesion segmentation. In Proceedings of the 2018 IEEE 15th International Symposium on Biomedical Imaging (ISBI 2018), Washington, DC, USA, 4–7 April 2018; IEEE: Piscataway, NJ, USA, 2018; pp. 859–863.
32. Peng, Y.; Wang, N.; Wang, Y.; Wang, M. Segmentation of dermoscopy image using adversarial networks. *Multimed. Tools Appl.* **2019**, *78*, 10965–10981. [CrossRef]
33. Tu, W.; Liu, X.; Hu, W.; Pan, Z.; Xu, X.; Li, B. Segmentation of lesion in dermoscopy images using dense-residual network with adversarial learning. In Proceedings of the 2019 IEEE International Conference on Image Processing (ICIP), Taipei, Taiwan, 22–25 September 2019; IEEE: Piscataway, NJ, USA, 2019; pp. 1430–1434.
34. Tschandl, P.; Sinz, C.; Kittler, H. Domain-specific classification-pretrained fully convolutional network encoders for skin lesion segmentation. *Comput. Biol. Med.* **2019**, *104*, 111–116. [CrossRef]
35. Ninh, Q.C.; Tran, T.-T.; Tran, T.T.; Tran, T.A.X.; Pham, V.-T. Skin lesion segmentation based on modification of segnet neural networks. In Proceedings of the 2019 6th NAFOSTED Conference on Information and Computer Science (NICS), Hanoi, Vietnam, 12–13 December 2019; IEEE: Piscataway, NJ, USA, 2019; pp. 575–578.
36. Mirikharaji, Z.; Yan, Y.; Hamarneh, G. Learning to segment skin lesions from noisy annotations. In *Domain Adaptation and Representation Transfer and Medical Image Learning with Less Labels and Imperfect Data*; Springer: Berlin/Heidelberg, Germany, 2019; pp. 207–215.

37. Sarker, M.; Kamal, M.; Rashwan, H.A.; Abdel-Nasser, M.; Singh, V.K.; Banu, S.F.; Akram, F.; Chowdhury, F.U.; Choudhury, K.A.; Chambon, S.; et al. Mobilegan: Skin lesion segmentation using a lightweight generative adversarial network. *arXiv* **2019**, arXiv:1907.00856.
38. Lei, B.; Xia, Z.; Jiang, F.; Jiang, X.; Ge, Z.; Xu, Y.; Qin, J.; Chen, S.; Wang, T.; Wang, S. Skin lesion segmentation via generative adversarial networks with dual discriminators. *Med. Image Anal.* **2020**, *64*, 101716. [CrossRef]
39. Zafar, K.; Gilani, S.O.; Waris, A.; Ahmed, A.; Jamil, M.; Khan, M.N.; Sohail Kashif, A. Skin lesion segmentation from dermoscopic images using convolutional neural network. *Sensors* **2020**, *20*, 1601. [CrossRef] [PubMed]
40. Xie, Y.; Zhang, J.; Xia, Y.; Shen, C. A mutual bootstrapping model for automated skin lesion segmentation and classification. *IEEE Trans. Med. Imaging.* **2020**, *39*, 2482–2493. [CrossRef]
41. Chen, L.C.; Zhu, Y.; Papandreou, G.; Schroff, F.; Adam, H. DeepLabV3+. 2018. Available online: https://arxiv.org/abs/1802.02611 (accessed on 14 August 2024).
42. Oktay, O.; Schlemper, J.; Folgoc, L.L.; Lee, M.; Heinrich, M.; Misawa, K.; Mori, K.; McDonagh, S.; Hammerla, N.Y.; Kainz, B.; et al. Attention U-Net. 2018. Available online: https://arxiv.org/abs/1804.03999 (accessed on 14 August 2024).
43. Zhou, Z.; Rahman Siddiquee, M.M.; Tajbakhsh, N.; Liang, J. UNet++. 2018. Available online: https://arxiv.org/abs/1807.10165 (accessed on 14 August 2024).
44. Li, R.; Zheng, S.; Duan, C.; Su, J.; Zhang, C. Multistage Attention ResU-Net for Semantic Segmentation of Fine-Resolution Remote Sensing Images. *IEEE Geosci. Remote Sens.* **2021**, 1–5. [CrossRef]
45. Diakogiannis, F.I.; Waldner, F.; Caccetta, P.; Wu, C. Resunet-a: A deep learning framework for semantic segmentation of remotely sensed data. *ISPRS J. Photogramm. Remote Sens.* **2020**, *162*, 94–114. [CrossRef]
46. Milletari, F.; Navab, N.; Ahmadi, S.A. V-Net. 2016. Available online: https://arxiv.org/abs/1606.04797 (accessed on 14 August 2024).
47. Hasan, M.J.; Ahmad, W.S.H.M.W.; Fauzi, M.F.A.; Abas, F.S. Hybrid Deep Learning Architectures for Histological Image Segmentation. In Proceedings of the 2024 International Conference on Artificial Intelligence in Information and Communication (ICAIIC), Osaka, Japan, 19–22 February 2024; pp. 075–080.
48. Cao, H.; Wang, Y.; Chen, J.; Jiang, D.; Zhang, X.; Tian, Q.; Wang, M. Swin-UNet. 2021. Available online: https://arxiv.org/abs/2105.05537 (accessed on 14 August 2024).
49. Chen, J.; Lu, Y.; Yu, Q.; Luo, X.; Adeli, E.; Wang, Y.; Lu, L.; Yuille, A.L.; Zhou, Y. TransUNet. 2021. Available online: https://arxiv.org/abs/2102.04306 (accessed on 14 August 2024).
50. Shehzad, K.; Zhenhua, T.; Shoukat, S.; Saeed, A.; Ahmad, I.; Sarwar Bhatti, S.; Chelloug, S.A. A Deep-Ensemble-Learning-Based Approach for Skin Cancer Diagnosis. *Electronics* **2023**, *12*, 1342. [CrossRef]
51. Almuayqil, S.N.; Abd El-Ghany, S.; Elmogy, M. Computer-Aided Diagnosis for Early Signs of Skin Diseases Using Multi Types Feature Fusion Based on a Hybrid Deep Learning Model. *Electronics* **2022**, *11*, 4009. [CrossRef]
52. Foahom Gouabou, A.C.; Iguernaissi, R.; Damoiseaux, J.-L.; Moudafi, A.; Merad, D. End-to-End Decoupled Training: A Robust Deep Learning Method for Long-Tailed Classification of Dermoscopic Images for Skin Lesion Classification. *Electronics* **2022**, *11*, 3275. [CrossRef]
53. Ibraheem, M.R.; El-Sappagh, S.; Abuhmed, T.; Elmogy, M. Staging Melanocytic Skin Neoplasms Using High-Level Pixel-Based Features. *Electronics* **2020**, *9*, 1443. [CrossRef]
54. Codella, N.C.; Gutman, D.; Celebi, M.E.; Helba, B.; Marchetti, M.A.; Dusza, S.W.; Kalloo, A.; Liopyris, K.; Mishra, N.; Kittler, H.; 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 of the 2018 IEEE 15th International Symposium on Biomedical Imaging (ISBI 2018), Washington, DC, USA, 4–7 April 2018; IEEE: Piscataway, NJ, USA, 2018; pp. 168–172.
55. Drozdzal, M.; Vorontsov, E.; Chartrand, G.; Kadoury, S.; Pal, C. The importance of skip connections in biomedical image segmentation. In *Deep Learning and Data Labeling for Medical Applications*; Springer: Berlin/Heidelberg, Germany, 2016; pp. 179–187.
56. Li, Y.; Shen, L. Skin Lesion Analysis towards Melanoma Detection Using Deep Learning Network. *Sensors* **2018**, *18*, 556. [CrossRef] [PubMed]
57. Nawaz, M.; Mehmood, Z.; Nazir, T.; Naqvi, R.A.; Rehman, A.; Iqbal, M.; Saba, T. Skin cancer detection from dermoscopic images using deep learning and fuzzy k-means clustering. *Microsc. Res. Tech.* **2022**, *85*, 339–351. [CrossRef]

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.