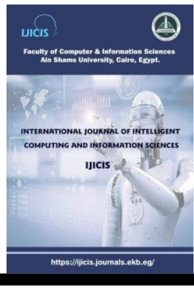
IJICIS, Vol.22, No.1, 88-97
DOI: 10.21608/ijicis.2022.104799.1136

International Journal of Intelligent Computing and Information Sciences

https://ijicis.journals.ekb.eg/



Experimental Comparative Study on Autoencoder Performance for Aided Melanoma Skin Disease Recognition

Zahraa E. Diame*
Department of Computer Science, Faculty of Computers and Information Sciences, Ain Shams University
Cairo, Egypt
zahraa2.edf@gmail.com

Maryam N. Al-Berry
Scientific Computing Department
Faculty of Computers and Information Sciences, Ain Shams University
Cairo, Egypt
maryam_nabil@cis.asu.edu.eg

Mohammed A.-M. Salem
Computer and Information Sciences, Ain Shams University,
Media Eng. and Technology, German University in Cairo
mohammed.salem@guc.edu.eg

Mohamed Roushdy
Faculty of Computers and Information Technology,
Future University in Egypt,
New Cairo, Cairo, Egypt.
Mohamed.Roushdy@fue.edu.eg

Received 2021- 11-07; Revised 2022-01-13; Accepted 2022-01-18

Abstract: *Melanoma is a dangerous and metastatic cancer that may be fatal and it has a high ability to invade other tissues and organs. Early diagnosis is an important reason to recover from melanoma and reduce mortality. So, automatic skin segmentation is considered an enthusiastic study at present. In this paper, we investigate the applicability of deep learning approaches to the segmentation of skin lesions by evaluating five architectures: Deeplabv3plus, Inception-ResNet-v2-unet, mobilenetv2_unet, Resnet50_unet, vgg19_unet by providing a comparative study of those methods. All methods were trained on the ISIC2017 dataset. The methods were trained on the original dataset, and then the dataset was pre-processed for use in training the five methods. We used quantitative evaluation metrics to evaluate the performance of the methods. The Deeplabv3+ architecture showed significant results compared to the rest of the architecture in F1 as high as 89%, Jaccard as high as 83% and Recall as high as 91%.*

Keywords: *Deep learning, Segmentation, Skin lesion, Melanoma detection, Lesion segmentation, Skin cancer, image processing.*

1. Introduction

The statistics of the American Cancer Society indicate that the incidence of skin cancer has increased significantly in the past few years. Melanoma is a dangerous and metastatic cancer that may be fatal if

* Corresponding author: Zahraa E. Diame
Faculty of Computers and Information Sciences, Ain Shams University, Cairo, Egypt
E-mail address: zahraa2.edf@gmail.com

not diagnosed early, and it has a high ability to invade other tissues and organs. Skin cancer is more than 20 times more common in whites than in African Americans. In general, the risk of developing skin cancer is about 2.6% (1 in 38) for whites, 0.1% (1 in 1,000) for blacks, and 0.6% (1 in 167) for Hispanics. Melanoma is more common in men overall, but before age 50 the rates are higher in women than in men. The risk of developing skin cancer increases with age. Sixty-five years is the average age at which people are diagnosed. But skin cancer is not common even amongst those under 30 years old. It is one of the several common types of cancer among young men (mainly young women) [1].

Initially, careful segmenting of the melanoma can help treat this life-threatening problem. But, manual segmenting needs good-trained experts [2]. The ability of automatic approaches to increase accuracy and performance of pathologists can be enhanced to segment skin lesions. The automatic segmentation of skin lesions faces significant challenges because of the low contrast and high degree of resemblance between normal skin areas and areas of skin lesions [3]. Per case has various skin issues of his own, which gives a distinctive appearance to the skin lesions such as the texture and color of the skin, in addition to the hair or veins [4].

So, automatic skin segmentation is considered an enthusiastic study at present. Knowing the boundaries of the lesion helps to accurately classify the lesion. In recent times, intelligence-based segmentation procedures like genetic algorithms, fuzzy logic, artificial neural networks (ANN), and deep learning methods have been used very recently for precise and reliable skin lesions segmentation Nasser et al. [5].

The paper is organized as follows: An overview of relevant previous work is given in Section 2. Section 3 presents the materials and methodologies we use which include dataset used in the research and pre-processing steps as well as evaluation metrics, deep network models, and architectures that were used for lesion segmentation. Section 4 presents our result and discussion, and finally, Section 5 provides a conclusion to this paper.

2. Related Works

In the last years, simultaneously with improvements in the computational power of GPUs, methodologies based on deep learning are mainly characterized as one of the most powerful approaches in image processing, such as [6-10]. Image segmentation has been a hot topic since its emergence and since then many researchers have worked on the detection and segmentation of melanoma. Researchers have already used a variety of techniques for the image processing task to help classify skin lesions. For example, methods such as edge detection and border have been utilized to bifurcate images of skin lesions [11]. In [12] they used DeepLab 3, the atrous convolution method for image segmentation. The results from this run are not perfect, with an average Jaccard index of 0.498. In [13] they used the Deeplabv3+ method for segmentation of skin lesions and obtained a score of 0.951 in the Jaccard index and accuracy. In [14] they introduced two new architectures for deep learning networks, Inception-Resnet and W-net, to resolve the segmentation and classification problems respectively. W-net consists of a ResNet decoder, an ConvNet decoder and a feature hierarchy network. The results show an accuracy rate of 98.1% in the Inception-Resnet method. In [15] they presented the Inception-ResNet-v2 method, which showed good results with an Accuracy rate of 0.9623. In [16], they performed an automatic segmentation of skin lesions utilizing a lightweight decoder, MobileNetV3-UNet, which can attain great accuracy with low resources and this was evident in the results that showed superiority in the Jaccard index rate of 0.9108. In [17] they implemented an experiment that implied a DeepLab-based method with MobileNetV2 being a feature extracted with a display multiplier of 0.35 and optimized with soft Dice loss. Obtained a Jaccard coefficient threshold of 72.97% and 78.51%, respectively. In [18] they used an automated lesion boundary segmentation technique that integrates two methods, ResNet and U-Net, together named ResUnet. The method got 0.772 in the Jaccard index. In [19] they improved the implementation of U-Net in different segmentation functions, a new structure named DoubleU-Net, which consists of two U-Net structures. The primary U-Net utilized the pre-trained VGG-19 being an encoder. The lesion segmentation shows that DoubleU-Net outperforms U-Net. In [20] they used CNN network VGG19, Inception V3 and VGG16 which provides the best execution. These models provide the better execution of test data with an accuracy of Inception V3 74%, VGG16 77%, and VGG19 76%.

3. Material and Methodology

In this part, we describe the data sets used, the preprocessing step, the quantitative evaluation metrics, the models of the architectures used, and the implementation steps.

3.1 Data set

The International Skin Imaging Collaboration (ISIC) The Skin Cancer Project is an academic and industry partnership that aims to support the application of digital skin imaging to help reduce skin cancer deaths. In addition, ISIC has developed and expanded an open source archive for public access to skin images to test and verify proposed standards. This archive serves as a public image resource for teaching, developing, and testing automated diagnostic systems. ISIC 2017 [21] was utilized. It contains of 2000 lesion images in JPEG format with segmentation masks. The dataset images have been divided into 1200 for training, 400 for validation and 400 for testing. Shown in Fig. 1 are some samples for the ISIC2017 dataset.

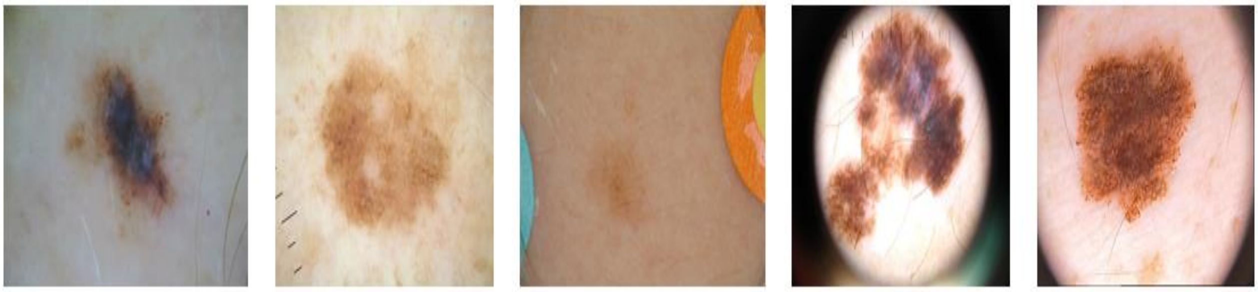


Figure.1 samples for the ISIC2017 dataset.

3.2 Pre-processing

Data augmentation techniques are the use of some resampling schemes to increase the number of training samples. Usually, the skin lesion is located in the center of the dermatoscopy images. But due to the difference in the condition of the photo and the measurement, the skin lesion in the photos can be located in different positions. We used different types of original image transformations, such as center crop, random rotation 90, grid distortion, horizontal flip and vertical flip which increased the number of images corresponding to the training iteration shown in Fig.2.

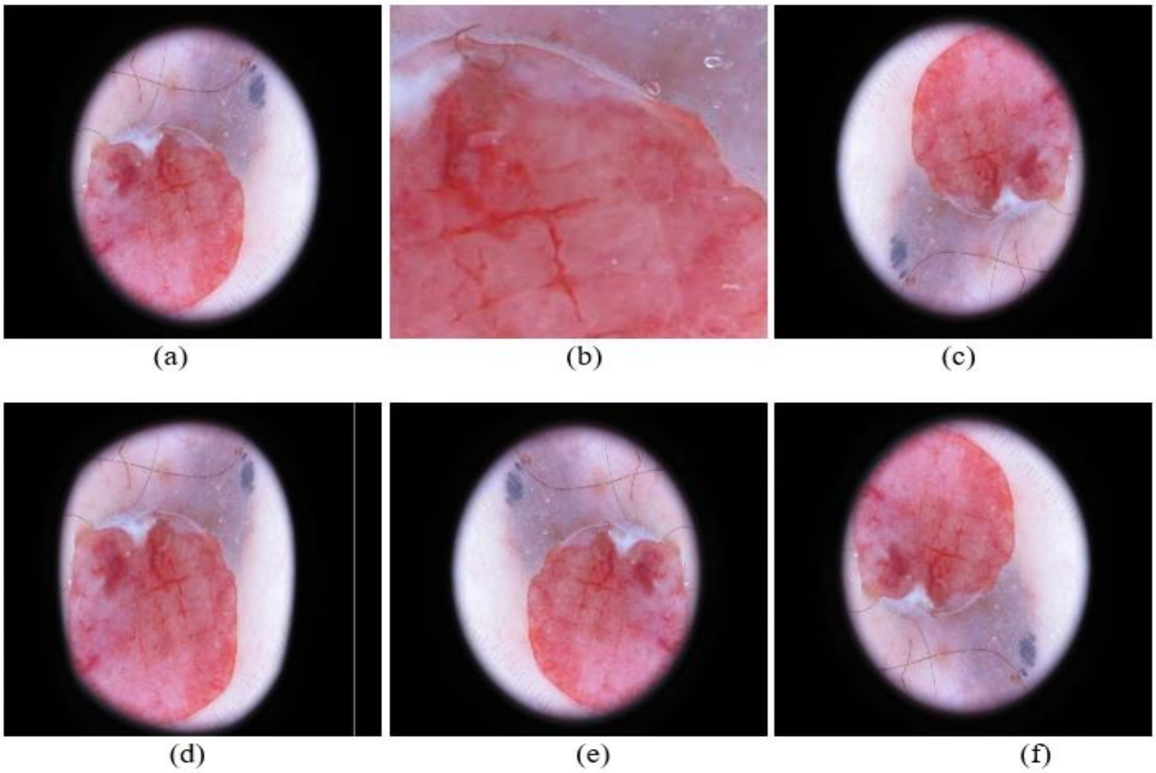


Figure.2 Data augmentation (a) original image (b) center crop (c) random rotation 90 (d) grid distortion (e) horizontal flip (f) vertical flip.

3.3 Model architecture

3.3.1 Deeplabv3+

The DeepLabV3+ [22] architecture was presented by Google. It is an improvement over the existing DeepLabV3 architecture. It is an encoder-decoder architecture with Atrous Spatial Pyramid Pooling (ASPP) and bilinear up-sampling. The network begins with a pre-trained ResNet50 [23] as the encoder, which is followed by the ASPP. The ASPP consists of dilated convolution which helps to encode multi-scale contextual information. Next, it is followed by a bilinear up-sampling by a factor of 4 and then concatenated with the low-level information from the encoder. After that, a few 3×3 convolutions are applied and again it is followed by a bilinear up-sampling by a factor of 4. Finally, at last we get the output mask. The results show that the DeepLabv3+ model identifies new cutting-edge performance on different datasets.

3.3.2 Inception-ResNet-v2-unet

Inception-ResNet-v2-unet is a combination of Inception-Resnet-v2 [24] and U-Net [25] architectures. In Inception-ResNet-v2-unet architecture was used the Inception ResNetV2 as the pre-trained encoder for UNET. Inception-Resnet-v2 is developed based on a fusion of the Inception structure and the Residual connection. In the Inception-Resnet block, multiple sized convolutions are merged by using residual connections. With the residual connections, the degradation problem was avoided, and the training time was reduced.

3.3.3 Mobilenetv2_unet

Mobilenetv2_unet is a combination of MobileNetV2 [26] and U-Net [25] architectures. In Mobilenetv2_unet architecture was used pre-trained MobileNetV2 as the encoder for the U-Net architecture. U-Net is a fully convolutional neural network. MobileNetV2 is an architecture that is optimized for mobile devices. It develops the modern performance of mobile models in multiple functions and criteria as well as through a range of various form sizes. The MobileNetV2 is used for the encoder/downsampling path of the U-Net (the left half of the U). The advantage of using MobileNetV2 as an Encoder is MobileNetV2 has fewer parameters, due to which it is easy to train. Using a pre-trained encoder helps the model to converge much faster in comparison to the non-pretrained model. A pre-trained encoder helps the model to achieve high performance as compared to a non pre-trained model.

3.3.4 Resnet50_unet

Resnet50_unet is a combination of Resnet50 [23] and U-Net [25] architectures. In Resnet50_unet chitecture was used pre-trained Resnet50 as the encoder for the U-Net architecture. U-Net is a fully convolutional neural network. Resnet50 is a variant of ResNet model which has 48 Convolution layers along with 1 MaxPool and 1 Average Pool layer. It has 3.8 x 10^9 Floating points operations. The Resnet50 is used for the encoder/downsampling path of the U-Net (the left half of the U). Resnet50_unet helps in reducing the training time as the encoder weights are not initialized from scratch.

3.3.5 Vgg19_unet

VGG19-UNet, is a deep fully convolutional network known as VGG19 [27] is embedded at the encoder part of UNET [25] architecture. In VGG19-UNet was used pre-trained VGG19 as the encoder for the U-Net architecture. VGG19 is a variant of VGG model which in short consists of 19 layers (16 convolution layers, 3 Fully connected layer, 5 MaxPool layers and 1 SoftMax layer). The VGG19 used as a good classification architecture for many other datasets.

3.4 Evaluation Metrics

To evaluate segmentation performance, we used the Jaccard index, Precision, Accuracy, F-measure, and Recall where TP is true positives, TN is true negatives, FP is false positives, and FN is false negatives. The scales are calculated utilizing equations (1) - (5):

Jaccard index, is essentially a method to quantify the percent overlap between the ground truth mask with the mask we created. Quite simply, the Jaccard index measures the number of pixels common between the target and prediction masks divided by the total number of pixels present across both masks.

$$\boldsymbol{Jaccard\ index = \frac{TP}{TP + FP + FN}} \qquad (\mathbf{1})$$

Accuracy an alternative metric to evaluate a semantic segmentation is to simply report the percent of pixels in the image which were correctly classified. The pixel accuracy is commonly reported for each class separately as well as globally across all classes. When considering the per-class pixel accuracy we're essentially evaluating a binary mask; a true positive represents a pixel that is correctly predicted to belong to the given class (according to the target mask) whereas a true negative represents a pixel that is correctly identified as not belonging to the given class.

$$\boldsymbol{accuracy = \frac{TP + TN}{TP + TN + FP + FN}} \qquad (\mathbf{2})$$

Precision effectively describes the purity of our positive detections relative to the ground truth.

$$\boldsymbol{Precision = \frac{TP}{TP + FP}} \qquad (\mathbf{3})$$

Recall effectively describes the completeness of our positive predictions relative to the ground truth.

$$\boldsymbol{Recall = \frac{TP}{TP + FN}} \qquad (\mathbf{4})$$

The F-score is a way of combining the precision and recall of the model, and it is defined as the harmonic mean of the model's precision and recall.

$$\boldsymbol{F1 = \frac{2TP}{2TP + FN + FP}} \qquad (\mathbf{5})$$

3.5 Implementation

Network architectures implemented in Python using Keras [28] package with Tensorflow backend on Dell Latitude 5480 with Intel(R) Core(TM) i7-7820HQ CPU 2.90 GHz NVIDIA GeForce 930MX GPU with Intel(R) HD Graphics 630. The training images in the data set are often of different sizes and pixels, but the deep neural network model requires input images of a fixed size. The form of input images for common deep learning models are often square matrices. We used Keras Image Data

Generator to wrap lesion images to a fixed size. We use fixed-size (256 x 256) training images to train the models. We start the training with a batch size of 4, and the proposed architecture is optimized by Adam optimizer. The learning rate of the algorithm is set to 1e-4 (0.0001). We trained all the models for 5 epochs with a lower learning rate so that a more generalized model can be built.

4. Results And Discussion

We have implemented five Architectures of skin lesion segmentation using the ISIC2017 data set, and we have done the step of data augmentation, and we got the results shown in Table 1. We notice in Table 1 that the deeplabv3plus method shows the best results in F1 with an average of 89.82%, Jaccard with an average of 83.7% and Recall with an average of 91.51%, while the mobilenetv2_unet method showed good results in Precision with an average of 93.49%. The Resnet50_unet method showed the highest rate of Accuracy up to 95.76%. The results showed a significant improvement after using the data augmentation step.

Table 1 Results of Skin Lesion Segmentation on ISIC 2017 Before and After Pre-processing

MODEL	Accuracy		F1		Jaccard		Recall		Precision	
	Before	After	Before	After	Before	After	Before	After	Before	After
Deeplabv3+	0.9512	0.9531	0.8341	0.8982	0.7500	0.8370	0.7982	0.9151	0.9273	0.9082
Inception-ResNet-v2-unet	0.9546	0.9548	0.8395	0.8909	0.7641	0.8302	0.8209	0.8920	0.9041	0.9164
Mobilenetv2_unet	0.9465	0.9329	0.8315	0.8639	0.7467	0.7959	0.8158	0.8373	0.9078	0.9349
Resnet50_unet	0.9576	0.9496	0.8593	0.8568	0.7813	0.7943	0.8539	0.8460	0.9066	0.9043
Vgg19_unet	0.9449	0.9454	0.8192	0.8775	0.7256	0.8098	0.8990	0.8699	0.8091	0.9224
U-Net	0.9341	0.9199	0.7966	0.8062	0.7011	0.7253	0.8456	0.8426	0.8357	0.8530

Table 2 shows the time consumed for training the segmentation on the original data set, after Pre-processing, and the total Params for each model.

Table 2 show the time consumed for training the segmentation and Total Params for each mode

MODEL	Time Training Per Epoch		Total Params
	Before	After	
Deeplabv3+	0:53:29	4:56:16	17,830,209
Inception-ResNet-v2-unet	0:33:16	3:35:45	36,793,089
Mobilenetv2_unet	0:29:07	3:03:28	11,753,809
Resnet50_unet	0:34:56	3:28:05	20,676,545
Vgg19_unet	0:43:12	4:46:25	31,172,033
U-Net	0:50:54	4:52:04	31,055,297

5. Conclusion

Melanoma is a dangerous and metastatic cancer that may be fatal if not diagnosed early, and it has a high ability to invade other tissues and organs. Initially, careful segmenting of melanoma can help treat this life-threatening problem. But, manual segmenting needs good-trained experts. So, automatic skin segmentation is considered an enthusiastic study at present. In this paper, we investigate the applicability of deep learning approaches to the segmentation of skin lesions by evaluating five architectures: Deeplabv3plus, Inception-ResNet-v2-unet, mobilenetv2_unet, Resnet50_unet, vgg19_unet. All methods were trained on the ISIC2017 dataset. The methods were trained on the original dataset, and then the dataset was pre-processed for use in training the five methods. We used different types of original image transformations, such as grid distortion, horizontal flip, and vertical flip, which increased the number of images in the training set. Our experiments show that the five structures performed well and closely in the results of skin lesion image segmentation. Overall, the Deeplabv3+ architecture showed significant results compared to the rest of the architecture in F1 as high as 89%, Jaccard as high as 83% and Recall as high as 91%. This is because the Deeplabv3+ method designed a simple decoder to merge the low and high-level features, corresponding to structure details and semantic information, to extract more information from the input image. Using Atreus and depthwise separable convolution to computer faster while maintaining the quite similar precision. ResNet50 model used as the backbone here for better performance (both accuracy and speed). The number of parameters is not very large. In addition, we observed that the pre-processing steps had a good effect on the results.

References

1. “American Cancer Society,” (January 12, 2021) Available: https://www.cancer.org/cancer/melanoma-skin-cancer/about/key-statistics.html
2. Masood A, Al-Jumaily AA(2013) “Computer aided diagnostic support system for skin cancer: a review of techniques and algorithms.” Int J Biomed Imaging 2013
3. C. Barata, M. Emre Celebi, and J. S. Marques, "Improving dermoscopy image classification using color constancy," IEEE Journal of Biomedical and Health Informatics, vol. 19, pp. 1146-1152, 2015.
4. Adegun A, Viriri S (2019) “An enhanced deep learning framework for skin lesions segmentation.” In: International conference on computational collective intelligence, pp 414-425. Springer, Cham
5. Nasir M, Khan MA, Sharif M, Lali IU, Saba T, Iqbal T (2018) An improved strategy for skin lesion detection and classification using uniform segmentation and feature selection based approach. Microsc Res Tech 18(6):528–543
6. Ali, H., Elmogy, M., ALdaidamony, E., Atwan, A. MRI BRAIN IMAGE SEGMENTATION BASED ON CASCADED FRACTIONAL-ORDER DARWINIAN PARTICLE SWARM OPTIMIZATION AND MEAN SHIFT. International Journal of Intelligent Computing and Information Sciences, 2015; 15(1): 71-83. doi: 10.21608/ijicis.2015.10912
7. Aouf, M. Application of K-MTh Algorithm for Accurate Lunge Cancer Detection. International Journal of Intelligent Computing and Information Sciences, 2019; 19(2): 1-20. doi: 10.21608/ijicis.2019.71026
8. Laban, N., Abdellatif, B., Moushier, H., shedeed, H., Tolba, M. ENHANCED PIXEL BASED URBAN AREA CLASSIFICATION OF SATELLITE IMAGES USING CONVOLUTIONAL NEURAL NETWORK. *International Journal of Intelligent Computing and Information Sciences*, 2021; (): 1-16. doi: 10.21608/ijicis.2021.79070.1099

9. Ukaoha, K., Igodan, E. Architecture Optimization Model for the Deep Neural Network. *International Journal of Intelligent Computing and Information Sciences*, 2019; 19(2): 1-16. doi: 10.21608/ijicis.2019.96101
10. Nazih, W., Hifny, Y., S. Elkilani, W., Abdelkader, T. Fast Detection of Distributed Denial of Service Attacks in VoIP Networks Using Convolutional Neural Networks. International Journal of Intelligent Computing and Information Sciences, 2020; 20(2): 125-138. doi: 10.21608/ijicis.2021.51555.1046
11. Z. E. Diame, M. N. Al-Berry, M. A. . -M. Salem and M. Roushdy, "Deep Learning Architiectures For Aided Melanoma Skin Disease Recognition: A Review," 2021 International Mobile, Intelligent, and Ubiquitous Computing Conference (MIUCC), 2021, pp. 324-329, doi: 10.1109/MIUCC52538.2021.9447615.
12. Wang, Yujie, Simon Sun, Jahow Yu and Limin Yu. "Skin Lesion Segmentation Using Atrous Convolution via DeepLab v3." ArXiv abs/1807.08891 (2018): n. pag.
13. Azad R., Asadi-Aghbolaghi M., Fathy M., Escalera S. (2020) Attention Deeplabv3+: Multi-level Context Attention Mechanism for Skin Lesion Segmentation. In: Bartoli A., Fusiello A. (eds) Computer Vision – ECCV 2020 Workshops. ECCV 2020. Lecture Notes in Computer Science, vol 12535. Springer, Cham. https://doi.org/10.1007/978-3-030-66415-2_16
14. Khouloud, S., Ahlem, M., Fadel, T. et al. W-net and inception residual network for skin lesion segmentation and classification. Appl Intell (2021). https://doi.org/10.1007/s10489-021-02652-4
15. Guha S.R., Rafizul Haque S.M. (2020) Performance Comparison of Machine Learning-Based Classification of Skin Diseases from Skin Lesion Images. In: Bindhu V., Chen J., Tavares J. (eds) International Conference on Communication, Computing and Electronics Systems. Lecture Notes in Electrical Engineering, vol 637. Springer, Singapore. https://doi.org/10.1007/978-981-15-2612-1_2
16. Adi Wibowo, Satriawan Rasyid Purnama, Panji Wisnu Wirawan, Hanif Rasyidi, Lightweight encoder-decoder model for automatic skin lesion segmentation, Informatics in Medicine Unlocked, Volume 25, 2021.
17. Andrade C., Teixeira L.F., Vasconcelos M.J.M., Rosado L. (2020) Deep Learning Models for Segmentation of Mobile-Acquired Dermatological Images. In: Campilho A., Karray F., Wang Z. (eds) Image Analysis and Recognition. ICIAR 2020. Lecture Notes in Computer Science, vol 12132. Springer, Cham. https://doi.org/10.1007/978-3-030-50516-5_20
18. Zafar K, Gilani SO, Waris A, Ahmed A, Jamil M, Khan MN, Sohail Kashif A. Skin Lesion Segmentation from Dermoscopic Images Using Convolutional Neural Network. Sensors (Basel). 2020 Mar 13;20(6):1601. doi: 10.3390/s20061601. PMID: 32183041; PMCID: PMC7147706.
19. Jha, Debesh, M. Riegler, Dag Johansen, Paal Halvorsen and Haavard D. Johansen. "DoubleU-Net: A Deep Convolutional Neural Network for Medical Image Segmentation." 2020 IEEE 33rd International Symposium on Computer-Based Medical Systems (CBMS) (2020): 558-564.
20. Jasil, S.P.G., Ulagamuthalvi, V. Deep learning architecture using transfer learning for classification of skin lesions. J Ambient Intell Human Comput (2021). https://doi.org/10.1007/s12652-021-03062-7
21. 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: 2018 IEEE 15th ISBI 2018. pp. 168–172. IEEE (2018)
22. Chen, Liang-Chieh, Yukun Zhu, George Papandreou, Florian Schroff and Hartwig Adam. "Encoder-Decoder with Atrous Separable Convolution for Semantic Image Segmentation." ArXiv abs/1802.02611 (2018): n. pag.

23. He, Kaiming, X. Zhang, Shaoqing Ren and Jian Sun. “Deep Residual Learning for Image Recognition.” 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR) (2016): 770-778.
24. Szegedy, Christian, Sergey Ioffe, Vincent Vanhoucke and Alexander Amir Alemi. “Inception-v4, Inception-ResNet and the Impact of Residual Connections on Learning.” AAAI (2017).
25. Ronneberger, Olaf, Philipp Fischer and Thomas Brox. “U-Net: Convolutional Networks for Biomedical Image Segmentation.” MICCAI (2015).
26. Sandler, Mark, Andrew G. Howard, Menglong Zhu, Andrey Zhmoginov and Liang-Chieh Chen. “MobileNetV2: Inverted Residuals and Linear Bottlenecks.” 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition (2018): 4510-4520.
27. Simonyan, Karen and Andrew Zisserman. “Very Deep Convolutional Networks for Large-Scale Image Recognition.” CoRR abs/1409.1556 (2015): n. pag.
28. Kingma, Diederik P. and Jimmy Ba. “Adam: A Method for Stochastic Optimization.” CoRR abs/1412.6980 (2015): n. pag.