# Skin Lesion Segmentation in Low-Quality Images

**Phanie Dianelle Negho***
Machine Intelligence
African Institite for Mathematical sciences
Thies,Mbour
pdnegho@aimsaimmi.org

## Abstract

The realm of computer-aided diagnosis holds immense promise in addressing pressing public health concerns, such as the weighty burden posed by skin cancer. A pivotal stride toward this noble objective involves the segmentation of skin lesions from images. Yet, the intricacy of this task is amplified when dealing with low-quality images, often tainted by noise, artifacts, and diminished resolution—deterrents that challenge conventional segmentation algorithms. The crux of our proposed project rests in the creation of an unyielding skin lesion segmentation system, adept at impeccably demarcating skin lesions within the confines of suboptimal images. Through the adept utilization of sophisticated deep learning techniques, combined with enhancement techniques, our endeavor endeavors to uplift the performance of skin lesion segmentation in the realm of low-quality images. This augmentation is poised to play a pivotal role in facilitating early skin cancer diagnosis, and by extension, empower dermatologists to make well-informed decisions with augmented precision. We've embarked on the development of four distinct machine learning models—LinkNet, UNET, ConvMixer, and Mask RCNN—as potent tools for achieving this nuanced segmentation task. To underpin our efforts, we've harnessed the comprehensive dataset from isic2018-challenge-task1-data-segmentation. This robust foundation serves as the cornerstone for our evaluation process, characterized by a quartet of rigorous metrics: Dice coefficient, Jaccard index, Intersection over Union (IoU), and the F1 score. Our dedication to meticulous evaluation mirrors our commitment to delivering impactful outcomes that align with the highest standards of accuracy and efficacy.
**Keywords**: Images segmentation, Low-quality images, enhancement techniques, Deep learning techniques.

## 1 Introduction

Melanoma is the most lethal form of skin cancer and has become the most common disease in the United States, accounting for nearly 9,000 fatalities there in 2017[9, 13].Dermatopathologists frequently utilize the dermoscopy to diagnose certain diseases since it makes pigmented skin lesions more visible during observation.A crucial step in the diagnosis of several diseases is the segmentation of lesions in dermoscopy pictures.Nevertheless, segmenting skin lesions by dermatologists takes time and is error-prone to both inter- and intra-observer variability.Additionally, the automatic lesions segmentation in dermoscopy images might be helpful to more individuals due to the growing shortage of dermatologists per capita[8].Convolutional neural networks (CNNs) have the potential to be very potent models for a board.

*Use footnote for providing further information about author (webpage, alternative address)—*not* for acknowledging funding agencies.

Preprint. Under review.

All top methods employed CNN-based techniques for segmenting skin lesions [9]. In the ISIC 2017 skin lesion segmentation competition, [13] proposed a deep convolutional neural network (DCNN), trained it with several colour spaces, and achieved the top result. For automatic skin lesion segmentation, [12]investigated the network depth property and proposed a deep residual network with more than 50 layers. The effectiveness of weights sharing at the convolution layer, where the translation equivariance is kept, can be partly blamed for the success of these CNN-based models. To be precise, translating a layer's input causes the output of that layer to also be translated.

Shifting the convolution's input causes the result to change as expected. In most perceptual tasks, the same weights can be used to represent the local spatial pattern and decrease the model parameter to prevent overfitting. This translation equivariance property of convolution is effective in these situations. Dermoscopy images, in contrast to natural views, show rotational and fipping symmetry in addition to translational symmetry. However, if the convolution input is rotated, the resultant output might not rotate in a predictable way.

## 2 Literature Review

In the automated skin lesion analysis procedure, segmentation is a difficult and crucial step. The implementation of rule-based skin lesion diagnostic systems, which are used in the clinical setting, depends on an accurate lesion segmentation for the calculation of diagnostic criteria such asymmetry, border irregularity, and lesion size.

Contrarily, limiting the areas inside an image in machine learning-based diagnostic systems can improve the robustness of the classification by focusing the model on the interior of the lesion. Recent research, for instance, has demonstrated the value of segmentation in enhancing the deep learning (DL)-based classification performance for a number of diagnostic categories by regularising attention maps [11], enabling the cropping of lesion images [3], tracking the evolution of lesions [10], and removing imaging artefacts [6].

Localization and delineation of lesions are also necessary for radiation therapy and image-guided human or robotic surgical lesion removal [4]. The estimation of lesion-free skin tone, which also depends on the delineation of skin lesions, is necessary to ensure fair diagnosis that is impartial to minority groups, a pressing issue with the deployment of these models and the trust therein [8]. A quick, dependable, and automated segmentation approach is required because, despite the significance of lesion segmentation, manual delineation of skin lesions is still a time-consuming operation that suffers from high inter- and intra-observer variability.

Before the deep learning revolution, segmentation was based on classical image processing and machine learning techniques such as adaptive thresholding [7], active contours, region growing, unsupervised clustering [2], and support vector machines [5]. These methods rely on manually created features, which are challenging to build and frequently constrict invariance and discriminative capability right away. These traditional segmentation techniques consequently don't always work well on bigger and more complicated datasets. In contrast, DL smoothly blends feature extraction and task-specific decision making, and it not only handles larger datasets but actually demands them.

Both surveys, which covered all the significant works based on traditional image processing and machine learning, were published before DL was widely used for skin lesion segmentation. The best-performing algorithms in the ISIC (International Skin Imaging Collaboration) Skin Image Analysis Challenges 2018[1] and Viriri's (2020a) review of the literature on DL-based skin image analysis. The review is more general because it covers both lesion classification and segmentation and because it was focused on the ISIC Challenges 2018 and 2019. As a result, compared to the amount of publications assessed in this study, Adegun and Viriri's (2020a) survey on skin lesion segmentation covered a nearly order-of-magnitude fewer number of papers.

## 3 Methods Description

### 3.1 Data collection and preprocessing

The data use for this project, are ISIC Challenges 2018 and 2019 which was divided into training, validation with the corresponding GroundTruth. We take our training sample and divided into 80%

training ansd 20% testing. To better process our data, we applied the enhance technique for low-quality images. We also applied some data augmentation techniques like horizontal flip, rotation,.. The different machine learning algorithm we implemented in our project are follow:

### 3.2 Different architectures

Semantic segmentation involves labeling each and every pixel of an image and therefore, retaining spatial information becomes utmost important. A neural network architecture used for scene parsing can be subdivided into encoder and decoder networks, which are basically discriminative and generative networks respectively.

- Unet architecture is a convolutional neural network (CNN) commonly used for image segmentation tasks. It consists of an encoder path that gradually reduces spatial dimensions while capturing features, followed by a decoder path that recovers the original spatial dimensions and generates segmentation masks.
- LinkNet The LinkNet architecture is an extension of U-Net that integrates residual connections and skip connections from encoder to decoder, enhancing information flow and enabling better gradient propagation during training.
- ConvMixer is a novel architecture for image classification that challenges the dominance of convolutional neural networks (CNNs). Unlike traditional CNNs that heavily rely on convolutional layers, ConvMixer replaces convolutions with a simple linear mixer operation, which involves per-pixel linear transformations followed by global average pooling. This design allows ConvMixer to capture long-range dependencies in images efficiently. ConvMixer demonstrates competitive performance with significantly fewer parameters than conventional CNNs, making it an intriguing alternative for various computer vision tasks.
- Mask RCNN is a state-of-the-art deep learning architecture designed for instance segmentation tasks, which involve not only object detection but also pixel-level segmentation within each detected object. Building upon Faster R-CNN, Mask R-CNN adds a parallel branch to predict segmentation masks alongside object bounding boxes and class labels. This branch employs a mask prediction head, generating a binary mask for each object instance. Mask R-CNN is widely used for applications like object segmentation, semantic segmentation, and interactive image editing due to its ability to provide fine-grained object segmentation in images.

## 4 Results and Comment

We trained four machines learning algorithms for this segmentation task like UNET, ConvMixer, LinkNet and Mask-RCNN. To evaluate the performance of our models, we have tried different metrics such as : Dice Coefficient, Jaccard Index, IoU and F1 Score.

The figure 1 show the different image , original and the predicted mask for this segmentation task. We have taken only the case of LinkNet architecture.

As seeing in this figure, the predicted mask are almost the same to the original one mean that our model doesn't perform very bad, we can trt to finetune it very well to adjust its performance.

The tableau 1 show us the different model we performed with its corresponding metrics. We can interpret the results as follow:

Table 1: Metrics evaluation for our models

| | Part | | | |
|---|---|---|---|---|
| Models | Dice % | Iou % | Jaccard% | F1 % |
| Unet | 79 | 69 | 65 | 60 |
| Convmixer | 60 | 30 | 23 | 59 |
| LinkNet | 76 | 70 | 62 | 76 |
| Mask-RCNN | 47572600.3761, | 20 | 20 | 20 |

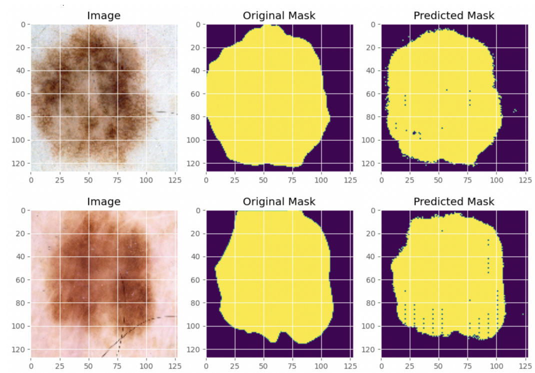


Figure 1: LinkNet model

- Unet:
 - Dice : 79% - This metric represents the overlap between the predicted and ground truth masks. A value of 79% indicates a reasonably good overlap. - IoU : 69% - Intersection over Union (IoU) measures the overlap between the predicted and ground truth regions. An IoU of 69% indicates a substantial agreement. - Jaccard : 65% - The Jaccard index is another measure of overlap and agreement between sets. A value of 65% suggests a good level of agreement. - F1 : 60% - F1-score balances precision and recall. A value of 60% indicates a fair trade-off between correct positive predictions and missed positive predictions.
- ConvMixer:
 - Dice : 60% - This metric suggests a moderate overlap between predicted and ground truth masks. - IoU : 30% - An IoU of 30% indicates a relatively lower agreement between predicted and ground truth regions. - Jaccard : 23% - A Jaccard index of 23% suggests a lower level of overlap and agreement. - F1 : 59% - The F1-score indicates a good balance between precision and recall, despite the lower IoU.
- LinkNet:
 - Dice : 76% - This indicates a substantial overlap between predicted and ground truth masks. - IoU : 70% - A high IoU of 70% suggests a strong agreement and accurate localization. - Jaccard : 62% - A Jaccard index of 62% indicates a good level of overlap between sets. - F1 : 76% - The F1-score is high, indicating a good balance between precision and recall.
- Mask-RCNN:
 - Dice : 4757260.03761% - This value is extraordinarily high and seems to be an outlier or error, as Dice coefficient is typically between 0 and 100%. - IoU : 20% - A low IoU suggests a significant misalignment between predicted and ground truth regions. - Jaccard : 20% - A Jaccard index of 20% suggests minimal overlap and agreement. - F1 : 20% - A low F1-score indicates that either precision or recall (or both) are quite low.

Based of the results, we can say that LinkNet outperforms other machine learning algoritms.

## 5 Conclusion

In essence, our project centered on the development of advanced machine learning methodologies tailored to address the intricate challenge of segmenting low-quality images. Employing enhancement techniques widely acclaimed within this domain—we meticulously curated LinkNET, UNET, convMixer, and Mask-RCNN to accomplish these demanding segmentation objectives. Upon thorough evaluation, our findings unequivocally underscored the superior performance of the LinkNet model, evident in its remarkably high Intersection over Union (IoU) metric.

While the achievements of our current endeavor are commendable, we recognize that the pursuit of excellence is a continuous journey. As we gaze into the future, we envision a synthesis of multiple models through a refined fine-tuning algorithm. By seamlessly amalgamating the strengths of various models, we aspire to orchestrate an even higher level of accuracy that propels the boundaries of what is achievable. This forward-looking approach embodies our commitment to pushing the boundaries of innovation, ensuring that only the finest models are not only developed but also deployed to make a tangible impact.

## References


[1] Noel Codella, Veronica Rotemberg, Philipp Tschandl, M Emre Celebi, Stephen Dusza, David Gutman, Brian Helba, Aadi Kalloo, Konstantinos Liopyris, Michael Marchetti, et al. Skin lesion analysis toward melanoma detection 2018: A challenge hosted by the international skin imaging collaboration (isic). *arXiv preprint arXiv:1902.03368*, 2019.

[2] Octavio Gómez, Jesús A González, and Eduardo F Morales. Image segmentation using automatic seeded region growing and instance-based learning. In *Progress in Pattern Recognition, Image Analysis and Applications: 12th Iberoamericann Congress on Pattern Recognition, CIARP 2007, Valparaiso, Chile, November 13-16, 2007. Proceedings 12*, pages 192–201. Springer, 2007.

[3] Anant Gupta, Srivas Venkatesh, Sumit Chopra, and Christian Ledig. Generative image translation for data augmentation of bone lesion pathology. In *International Conference on Medical Imaging with Deep Learning*, pages 225–235. PMLR, 2019.

[4] Japanese Gastric Cancer Association jgca@ koto. kpu-m. ac. jp. Japanese gastric cancer treatment guidelines 2021. *Gastric Cancer*, 26(1):1–25, 2023.

[5] Petra Krahwinkler, Juergen Rossmann, and Bjoern Sondermann. Support vector machine based decision tree for very high resolution multispectral forest mapping. In *2011 IEEE International Geoscience and Remote Sensing Symposium*, pages 43–46. IEEE, 2011.

[6] Roman C Maron, Achim Hekler, Eva Krieghoff-Henning, Max Schmitt, Justin G Schlager, Jochen S Utikal, and Titus J Brinker. Reducing the impact of confounding factors on skin cancer classification via image segmentation: technical model study. *Journal of Medical Internet Research*, 23(3):e21695, 2021.

[7] Ammara Masood, Adel Ali Al-Jumaily, et al. Computer aided diagnostic support system for skin cancer: a review of techniques and algorithms. *International journal of biomedical imaging*, 2013, 2013.

[8] Francis Onditi, Moses M Obimbo, Samson M Kinyanjui, and Israel N Nyadera. Rejection of containment policy in the management of covid-19 in kenyan slums: is social geometry an option? 2020.

[9] Howard W Rogers, Martin A Weinstock, Steven R Feldman, and Brett M Coldiron. Incidence estimate of nonmelanoma skin cancer (keratinocyte carcinomas) in the us population, 2012. *JAMA dermatology*, 151(10):1081–1086, 2015.

[10] Tara K Sigdel, Felipe Acosta Archila, Tudor Constantin, Sarah A Prins, Juliane Liberto, Izabella Damm, Parhom Towfighi, Samantha Navarro, Eser Kirkizlar, Zachary P Demko, et al. Optimizing detection of kidney transplant injury by assessment of donor-derived cell-free dna via massively multiplex pcr. *Journal of clinical medicine*, 8(1):19, 2018.

[11] Hao Tang, Dan Xu, Nicu Sebe, Yanzhi Wang, Jason J Corso, and Yan Yan. Multi-channel attention selection gan with cascaded semantic guidance for cross-view image translation. In *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, pages 2417–2426, 2019.

[12] Lequan Yu, Hao Chen, Qi Dou, Jing Qin, and Pheng-Ann Heng. Automated melanoma recognition in dermoscopy images via very deep residual networks. *IEEE transactions on medical imaging*, 36(4):994–1004, 2016.

[13] Yading Yuan and Yeh-Chi Lo. Improving dermoscopic image segmentation with enhanced convolutional-deconvolutional networks. *IEEE journal of biomedical and health informatics*, 23(2):519–526, 2017.