Multi-class Skin Lesion Segmentation for Cutaneous T-cell Lymphomas on High-Resolution Clinical Images

Zihao Liu[1,2], Haihao Pan[3], Chen Gong[1], Zejia Fan[1], Yujie Wen[3], Tingting Jiang[1], Ruiqin Xiong[1], Hang Li[3], and Yang Wang[3]

[1] NELVT, Department of Computer Science, Peking University, Beijing, China
[2] Advanced Institute of Information Technology, Peking University, China
[3] Peking University First Hospital, Beijing, China
{lzh19961031,ttjiang}@pku.edu.cn

Abstract. Automated skin lesion segmentation is essential to assist doctors in diagnosis. Most methods focus on lesion segmentation of dermoscopy images, while a few focus on clinical images. Nearly all the existing methods tackle the binary segmentation problem as to distinguish lesion parts from normal skin parts, and are designed for diseases with localized solitary skin lesion. Besides, the chararcteristics of both the dermoscopy images and the clinical images are fourfold: (1) Only one skin lesion exists in the image. (2) The skin lesion mostly appears in the center of the image. (3) The backgrounds are similar between different images of same modality. (4) The resolution of images isn't high, with an average of about 1500×1200 in several popular datasets. In contrast, this paper focuses on a four-class segmentation task for Cutaneous T-cell lymphomas(CTCL), an extremely aggressive skin disease with three visually similar kinds of lesions. For the first time, we collect a new dataset, which only contains clinical images captured from different body areas of human. The main characteristics of these images differ from all the existing images in four aspects: (1) Multiple skin lesion parts exist in each image. (2) The skin lesion parts are widely scattered in different areas of the image. (3) The background of the images has a large variety. (4) All the images have high resolutions, with an average of 3255×2535. According to the characteristics and difficulties of CTCL, we design a new Multi Knowledge Learning Network (MKLN). The experimental results demonstrate the superiority of our method, which meet the clinical needs.

Keywords: Skin lesion segmentation · Clinical image · Neural network.

1 Introduction

Skin disease is one of the most common type of disease in the world, with nearly 5 million new cases estimated every year [18, 15]. In recent years, more and more computer-aided methods are devoted to segmenting the lesion areas in order to assist doctors in the diagnosis process. Normally, there're two modalities of imaging: clinical images and dermoscopy images. Clinical images are captured by standard digital camera and can provide a global representation in view, angle and lighting, as for a better understanding and observation of the disease. In contrast, dermoscopy images are obtained by a microscope using incident light and oil immersion to make subsurface structures of

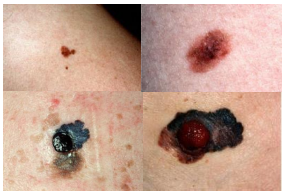

Fig. 1. Existing clinial images.

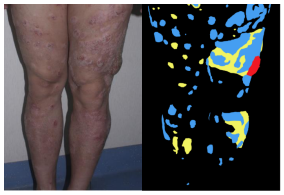

Fig. 2. CTCL clinical image.

Fig. 3. CTCL dataset.

a small skin region accessible, which permit a more detailed inspection of skin lesions. Most of the existing segmentation methods are designed for dermoscopy images [11, 4, 2, 21, 5, 20, 9]. In this work, we focus on clinical image segmentation.

There are also a few methods for clinical image segmentation [17, 1, 6, 10, 7]. The authors in [7] propose a texture-based method to learn the representative texture distributions, as for a better localization of lesion area. For deep-learning based methods, a FCN-based method is introduced in [17], with an extra critic module for better predictions of boundaries. To tackle the disturbing factors such as illumination variations from skin surface, the method in [1] proposes a zoom-out window to capture local and global information for accurate extraction of lesion regions. However, nearly all the above existing methods are designed for the binary segmentation task to distinguish between lesion area and skin area. Besides, the datasets mainly contain diseases with localized soliatary skin lesions, which means only one lesion area appeared and is usually located in the center of the images.

In this work, we aim at a new four-class segmentation task for a supremely aggressive skin disease called Cutaneous T-cell lymphomas (CTCL), on clinical images with totally different characteristics from existing ones. CTCL is a severe group of extranodal non-Hodgkin lymphomas [19, 8], which is a skin cancer with an increasing incidence rate in recent years. Survival declines dramatically as the disease progresses [16]. The unique manifestation of CTCL is three types of lesions on the skin: patches, plaques and tumors, which are increasing in severity and change at the disease progress [14, 3]. Thus, for auxiliary diagnosis, four-class segmentation is needed. Meanwhile, an important clinical application called "mSWAT Score Evaluation" is needed for this disease, which is implemented by the modified Severity Weighted Assessment Tool to calculate the percentage of the area of each skin lesion type over the area of whole body to track the clinical response for the treatments [13]. However, in clinical work, due to irregular skin lesions, it's difficult for doctors to make accurate and repeatable estimates during calculations. Thus, involving computers to automatically calculate the area will be very beneficial to improve the accuracy. For a better area evaluation result, segmenting all the lesion parts correctly first by the computer is important. So this four-class segmentation task is significant and the fundation of the automated mSWAT Score Evaluation.

However, there're several difficulties and uniqueness of this task, comparing to previous tasks. Firstly, the visual difference within different lesion parts or between one lesion part and a normal skin part could all be very similar. Next, multiple and connected skin lesions could appear throughout the body. Thus, for this disease, clinical, instead of dermoscopy images, are needed to captured from different parts of body sur-

Table 1. Comparasion between the existing skin lesion segmentation task and our task.

Attribute	Previous problem	Our problem
Number of classes	2 (Only lesion and normal skin)	4 (Three visually similar lesions)
Disease	With localized solitary skin lesion	With generalized skin lesion
Image Modality	Mostly dermoscopy, few clinical	Clinical
Characteristics of images	(1) Only one skin lesion part (2) Lesion part appears in the center (3) Similar background (4) Low to medium resolution	(1) Multiple skin lesion parts (2) Lesion parts are widely scattered (3) Large variety of background (4) High resolution

face area. Besides, high-resolution images are needed for accurate estimation of lesion areas. Fig. 1 contains examples of existing clinical images. Fig. 2 contains an example of CTCL image and an annotation map of all the lesion areas. In the map, yellow, blue and red denote patch, plaque and tumors, respectively. As we can see, the CTCL image is different from existing clinical images no matter from the characteristics and number of lesion types. Besides, the three types of lesions could concurrently appear in a single lesion area and are visually similar. Table 1 is the comparasion between our task and the previous task. To sum up, the main uniqueness and challenges for this task are: (1) It is a four-class segmentation task and the visual difference between different classes could be very similar, no matter between different lesion types within a lesion area or between lesion parts and normal skin parts. (2) The clinical images all have very high resolution and have distributed lesion parts existing in different parts, making the task more difficult. (3) The clinical images suffer from a large variety of backgrounds since they're captured in different body areas. However, to the best of our knowledge, there's no any previous work to solve the segmentation task for CTCL on clinical images. Therefore, this task is challenging yet has a great clinical application and research value.

In order to solve the above problems, for the first time, we collect a dataset, which contains 57 clinical images for CTCL, with full annotation for each lesion area and the corresponding lesion type in every image. Then, we design a novel two-branch Multi Knowledge Learning Network (MKLN), with a Lesion Area Learning Module for each branch and a Feature Co-Learning Module. One branch resizes the image to obtain global information, called "global branch"; the other branch crops the image to obtain detailed local information, called "local branch". The Lesion Area Learning Module is equipped by each branch to tackle the two challenges: large background variety and visual similarity between lesion part and normal skin part. It utilizes multi-level feature to perform binary segmentation and provide an attention map, as for a better learning of lesion edge knowledge. The Feature Co-Learning Module is designed to deal with the other challenges, including high resolution and visual similarity between different lesion types within a lesion area. It takes the output features of the two branches, learns an attention map for each branch, and enables them to mutually facilitate and learn knowledge from each other. Our method effectively addresses the difficulties of CTCL and achieves good experimental results on our dataset, meeting clinical needs.

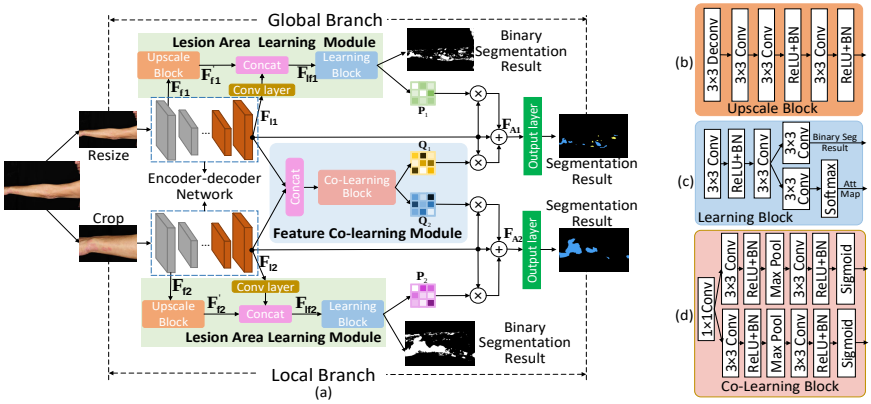

Fig. 4. The overall architecture of the MKLN is illustrated in (a). (b),(c) are the architecture of the Upscale Block and the Learning Block in Lesion Area Learning Module, (d) is the architecture of the Co-learning Block in Feature Co-learning Module.

Our contributions are two-fold: (1) The first paper to focus on the segmentation task for CTCL and a new dataset with clinical images is collected; (2) A new method is designed based on the characteristics of CTCL and achieves state-of-the art results, meeting clinical needs.

2 Dataset

We collaborate with the hospital and collect 57 clinical images, which can be called "CTCL dataset". The images are captured by the Nikon D800 Camera. Three experienced dermatologists participated in annotating. Each lesion area and the corresponding lesion type are annotated in every image. Several examples of our dataset are showed in Fig. 3. The collected clinical images are sampled from patients with different sex and wide age range, which is from 30 to 60 years old. The images include totally 12 human body regions, containing head, neck, anterior trunk, arms, forearms, hands, posterior trunk, buttocks, thighs, legs, feet and groin. Thus the background of the images suffer from a large variety. Resolution of the images is very high, with an average of 3255×2535. The largest resolution is 6000×2921. More examples of the CTCL dataset can be found in supplementary material.

3 Method

The whole architecture of MKLN is illustrated in Fig. 4, which is a two-branch network and is primarily based on an encoder-decoder network [12] with two other modules. The global branch takes the resized image as the input and the local branch takes the randomly cropped image, which has the same size as resized image, as the input. The first module is Lesion Area Learning Module, which is equipped by each branch. It takes the

feature from the first encoding block and the last decoding block as the input, generates an attention map containing the lesion edge feature and obtains a binary segmentation result which distinguish between lesion areas and non-lesion areas. To enable the network to take the advantage of both the global and local branch, the Feature Co-learning Module takes the output feature of last decoding block from each branch and generates an attention map for each branch, containing the learned knowledge from the other branch.

Lesion Area Learning Module As stated in Sec. 1, the background of CTCL clinical images has a large diversity since the images are captured from different body areas. Besides, the visual difference between lesion areas and normal skin areas is similar, for example between patches and normal skin in Fig. 2. Meanwhile, the lesion parts could also look similar to the background. Thus, adding additional constraint and supervision between lesion parts and other parts is really appealing and compelling for a better result. Therefore, we designed this module as an extra regularization for each branch.

For a better distinction between skin lesions and non-skin lesions, which is a binary segmentation task, edge and texture information are required. The low-level feature contains rich edge and texture information, meanwhile the high-level feature contains rich semantic information, which are all helpful for the binary segmentation. Thus, the output features of the first encoding block and the last decoding block are taken as input, denoted as F_f and F_l respectively. F_f goes through an upscale block and the upscaled feature $F_f^{'}$ will be obtained. Meanwhile, F_l will go through two 3×3 conv layers and concatenated with $F_f^{'}$, denoted as F_{lf}. A learning block is applied to F_{lf}, which will output the binary segmentation result and an attention map P with rich edge and boundary features. The architecture of the upscale block and learning block is illustrated in Fig. 4(b) and Fig. 4(c) respectively. The prediction result of the binary segmentation is supervised through binary cross-entropy loss, as illustrated in Eqn. (1):

$$L_{bce} = -\sum_{i}(y_i log(\hat{y_i}) + (1 - y_i) log(1 - \hat{y_i})) \tag{1}$$

where y_i denotes the groundtruth of each pixel and $\hat{y_i}$ is the predicted results. Meanwhile, an element-wise multiplication between P and F_l is applied. The result will be multiplied by a learnable factor α and then applied a matrix sum calculation to F_l.

With the extra supervision, the learned lesion area knowledge is better able to guide the extraction of discriminative features in high-level layers to learn a better edge and texture representation and knowledge for each lesion area.

Feature Co-learning Module The global branch resizes the image, and captures more global information and comprehensive semantic information and generates a better overall segmentation result. But the loss of detailed local and texture information leads to inaccurate classification between different classes of lesions, especially when different lesion types appear in a same lesion part. In contrast, the local branch crops the image, and captures more detailed local information, including boundary and texture information. Within each lesion area, the classification between different lesion types performs better. But due to the lack of global semantic information, the performance is also limited. However, since the resolution of images in our dataset is too high, with

an average of 3255 $\times$ 2535, neither of the single branch performs good. Therefore, we propose to combine the information of two branches and design this module as a connection between the two branches.

The module takes F_{l1} and F_{l2} as the input, which contains high-level semantic information, and concatenates them together firstly. Then, the concatenated features will go through a co-learning block, the architecture of which is illustrated in Fig. 4(d). Finally, an attention map Q for each branch will be generated. For the global branch, the generated attention map is expected to contain more detailed local information from the local branch for complementary. Similarly for the local branch, the generated attention map is expected to contain more global semantic knowledge for facilitation. Finally, an element-wise multiplication between Q and F_l is performed. The result will be multiplied by a learnable factor β and then applied a summation to F_l. The final output feature of these two modules is obtained by Eqn. (2):

$$F_A = \alpha \cdot (P \cdot F_l) + \beta \cdot (Q \cdot F_l) + F_l \tag{2}$$

where "$\cdot$" means element-wise multiplication.

After that, the output feature of each branch passes a 3 $\times$ 3 convolutional output layer to obtain the four-class segmentation result, which is supervised by cross-entropy loss L_{ce}, as illustrated in Eqn. (3):

$$L_{ce} = -\sum_{i}\sum_{j}(y_{ij} log(\hat{y}_{ij})) \tag{3}$$

where y_{ij} denotes whether pixel i belongs to label j and $\hat{y}_{ij}$ means the classification score of class j for pixel i.

This module automatically mines and exchanges the knowledge that each branch has learned, enabling them to facilitate each other. To the best of our knowledge, we are the first to take advantage of both of the two strategies and enable them to facilitate each other in the skin lesion segmentation task. To be mentioned, our method is totally different from [1] in the perspective and definition of "global" and "local", which results to a different design of the network architecture. Method in [1] just uses a slightly bigger window with the same center of local patch as global structure which still ignores the general information of the full image, and ignores to utilize the two branches to facilitate each other. But we successfully capture global information of the whole image, and directly exchange the information and knowledge from two branches to benefit each other.

A hybrid loss function is used as the final loss of each branch, which is weighted sum of two cross-entropy loss. It's illustrated in Eqn. (4) :

$$L = L_{bce} + \lambda L_{ce} \tag{4}$$

where λ is a hyper-parameter to linearly combine two losses in order to balance them.

4 Experiments

4.1 Implementation Details

The proposed method is evaluated on the newly introduced clinical dataset. Five-fold cross-validation is used for our experiments. Each image is resized and randomly cropped

to 608 $\times$ 608. At the training phase, the batch size is set to 2 on two NVIDIA GTX 1080Ti GPUs. At the testing phase, after each branch gets the four-class segmentation result for the whole image, the average probabilistic score of the two branches is calculated as the final segmentation result. The learning rate is initialized to 0.001 and decayed by 0.1 every 20 epochs. In Eqn. (4), λ is set to 1. We think both losses are equally important. The learnable factor α and β are trained together with the network. Four evaluation metrics are used in the experiments, including Accuracy(ACC), Dice Similarity Coefficient(DSC), Sensitivity(SE) and Specificity(SP).

4.2 Ablation Study

We conduct the ablation study on CTCL dataset to investigate the impact of the proposed two modules of our method on the performance. The specific performance is listed in Table 2. In the table, "**R**" and "**C**" mean the resize and crop strategy for the experimentes respectively. "**L**" denotes Lesion Area Learning Module and "**F**" represents Feature Co-learning Module. We first conduct experiments using baseline model [12] for each branch, denoted as **"MKLN(R)"** and **"MKLN(C)"** in the first two columns. Then we add **Lesion Area Learning Module** to each branch, denoted as **"MKLN(RL)"** and **"MKLN(CL)"**. The performance of all the four metrics improves. Compared to the baseline, **MKLN(RL)** yields an DSC of 0.695 and **MKLN(CL)** yields an DSC of 0.636, which are 4.1% and 1.7% higher than the baseline respectively. The Accuracy is improved by 3.1% and 2.1% respectively, which all verify the effectiveness and necessity for the Lesion Area Learning Module.

After that, **Feature Co-learning Module** is adopted to the network, the performance of two branch is denoted as "**MKLN(RLF)**" and "**MKLN(CLF)**". The performance of **MKLN(RLF)** is better than **MKLN(R)** and **MKLN(RL)**. The DSC is further increased to 0.769, with a gain of 11.7% comparing to **MKLN(R)**, and 7.4%, comparing to **MKLN(RL)**. Compared to **MKLN(C)** and **MKLN(CL)**, the DSC improves 5.3%and 3.6% respectively. Besides, results of the other three metrics all have a growth. The improvement of adopting Feature Co-learning Module is larger than the Lesion Area Learning Module, which further proves the effectiveness and advancement of this module, since no other methods have been utilized the knowledge from two branchs. Finally, a weighted sum is applied to combine the probabilistic score of **MKLN(RLF)** and **MKLN(CLF)** for a final segmentation result, which is denoted as **"MKLN(Full)"**. The performance of the full model on all the four metrics is the best, further demonstrates the effectiveness of introducing two branch and combining both of their knowledge. Besides, global branch has a better performance than local branch, we surmise it happens because the cropped image loses too much global semantic information, due to the high resolution of the whole image. To sum up, the encouraging results show the advantage of our method.

4.3 Comparison with other Methods

Since no any other related method are designed for this problem, we choose the most related work [17] and [1] to compare the results. Both [17] and [1] have similarities to our method in perspective of solving the segmentation task. As introduced in Sec. 1,

Table 2. Four-class segmentation performance by baseline methods, our method and other methods for comparasion on the CTCL dataset.

Method	MKLN (R)	MKLN (C)	MKLN (RL)	MKLN (CL)	MKLN (RLF)	MKLN (CLF)	MKLN (Full)	Izadi et al.[17]	Jafari et al.[1]
ACC	0.643	0.656	0.678	0.667	0.715	0.683	**0.725**	0.663	0.652
DSC	0.652	0.619	0.695	0.636	0.769	0.672	**0.801**	0.716	0.709
SE	0.624	0.701	0.721	0.761	0.881	0.778	**0.884**	0.844	0.825
SP	0.593	0.598	0.645	0.640	0.736	0.707	**0.776**	0.680	0.673

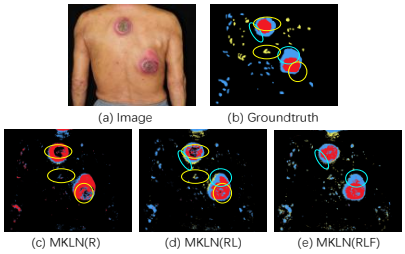

Fig. 5. Visualization results of the two modules.

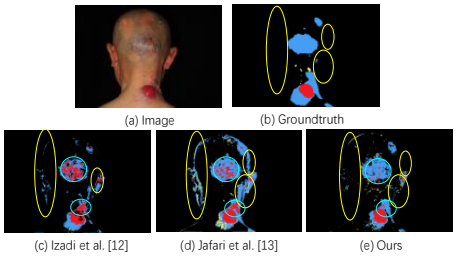

Fig. 6. Visualization results of three methods.

they are both deep-learning based method on clinical images. Besides, [17] produces a more accurate results on lesion boundary by introducing a critic network, and [1] also utilizes global and local information. The source codes for these two methods are not publicly available, so we implemented them based on our best understanding and obtained the results by same five-fold cross-validation setting.

The results of our method, [17] and [1] are shown in the last three columns of Table 2. The performance on ACC, DSC, SE and SP is evaluated. Our method achieve the best result in all the metrics, with an DSC of **0.801**, which exceeds [17] by 8.5% and [1] by 9.2%. Our ACC, SE and SP all perform the best.

4.4 Visualization Results

In addition to quantitative results, qualitative segmentation results are also provided. Fig. 5 demonstrates the effectiveness of the two modules. The comparison illustrated by yellow circles proves that with the lesion area learning module, the details and lesion boundaries and edges are clearer. The cyan circles areas demonstrates some misclassified lesion parts are corrected classified, resulting to a better global segmentation result. The results comparation between our method and the other two methods are shown in Fig. 6. The encircled areas by yellow circles have a better boundary and edge classification result. The areas encircled by cyan circle achieve a better classification result of lesion parts. So that our results are more close to the groundtruth. Fig. 7 visualizes the attention maps of the two modules. Compared to the activation of original feature map, the attention map of **L** successfully focuses on the edges of lesion parts, and the attention map of **F** further focuses on the detailed lesion parts.

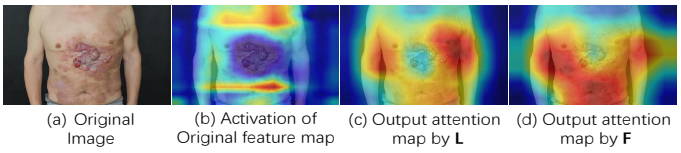

Fig. 7. Visualization of attention maps of **L** and **F**.

5 Conclusion

In this paper, we focus on a new four-class segmentation problem for Cutaneous T-cell lymphomas(CTCL), an extremely severe skin disease with three types of visually similar and decentralized skin lesions. For the first time, we collect a new dataset which contains clinical images with high resolutions from different body areas. Then we propose a novel Multi Knowledge-Learning Network(MKLN), including a Lesion Area Learning Module and a Feature Co-Learning Module to address this problem, which achieves very good performance and meets the clinical needs.

Acknowledgement. This work was partially supported by the Natural Science Foundation of China under contracts 61572042, 61772041, 81922058 and National Key R&D Program of China 2019YFC0840700. We also acknowledge the Clinical Medicine Plus X-Young Scholars Project, and High-Performance Computing Platform of Peking University for providing computational resources.

References

1. Jafari M, H and Karimi, N and Nasr-Esfahani, E and et al: Skin lesion segmentation in clinical images using deep learning. In: International conference on pattern recognition. pp. 337-342 (2016)
2. Diego, P., Jonathan, A., John, B.: Automatic skin lesion segmentation on dermoscopic images by the means of superpixel merging. In: International Conference on Medical Image Computing and Computer-Assisted Intervention. pp. 728–736 (2018)
3. Girardi, M., Heald P, W., Wilson L, D.: The pathogenesis of mycosis fungoides. New England Journal of Medicine **350**(19), 1978–1988 (2004)
4. Huan, W., Guotai, W., Ze, S., Shaoting, Z.: Automated segmentation of skin lesion based on pyramid attention network. In: International Workshop on Machine Learning in Medical Imaging. pp. 435–443 (2019)
5. Huan, W., Guotai, W., Zhihan, X., Wenhui, L., Shaoting, Z.: Star shape prior in fully convolutional networks for skin lesion segmentation. In: International Workshop on Machine Learning in Medical Imaging. pp. 611–619 (2019)
6. Idir, F., Malika, B.: Multi-scale contrast based skin lesion segmentation in digital images. Optik **185**, 794–811 (2019)
7. Jeffrey, G., Alexander, W., David A., C.: A segmentation of skin lesions from digital images using joint statistical texture distinctiveness. Pattern Recognition **61**(4), 1220–1230 (2014)
8. Kaveri, K., Michael, X., Martin, W.: Changing incidence trends of cutaneous t-cell lymphoma. JAMA dermatology **149**(11), 1295–1299 (2013)

9. Lei, S., Jianzhe, L., Z. Jane, W., Haoqian, W.: Dense-residual attention network for skin lesion segmentation. In: International Workshop on Machine Learning in Medical Imaging. pp. 319–327 (2019)
10. Maciel, Z., Eliezer, F., Jacob, S.: A simple weighted thresholding method for the segmentation of pigmented skin lesions in macroscopic images. Pattern Recognition **64**, 92–104 (2017)
11. Md. Mostafa Kamal, S., Hatem A., R., Farhan, A., Syeda Furruka, B., Adel, S., Vivek Kumar, S., Forhad U. H., C., Saddam, A., Santiago, R., Petia, R., Puig, D.: Slsdeep: Skin lesion segmentation based on dilated residual and pyramid pooling networks. In: International Conference on Medical Image Computing and Computer-Assisted Intervention. pp. 21–29 (2018)
12. Olaf, R., Philipp, F., Thomas, B.: U-net: Convolutional networks for biomedical image segmentation. In: International Conference on Medical Image Computing and Computer-Assisted Intervention. pp. 234–241 (2015)
13. Olsen, E., Whittaker, S., Kim, Y., et al: Clinical end points and response criteria in mycosis fungoides and sézary syndrome: a consensus statement of the international society for cutaneous lymphomas, the united states cutaneous lymphoma consortium, and the cutaneous lymphoma task force of the european organisation for research and treatment of cancer. Journal of clinical oncology **29**(18), 2598 (2011)
14. Pulitzer, M.: Cutaneous t-cell lymphoma. Clinics in Laboratory Medicine **37**(3), 527–546 (2017)
15. Rogers, H.W., Weinstock, M.A., Feldman, S.R., Coldiron, B.M.: Incidence estimate of nonmelanoma skin cancer (keratinocyte carcinomas) in the US population, 2012. JAMA dermatology **151**(10), 1081–1086 (2015)
16. Ryan A., W.: Cutaneous t-cell lymphoma: 2017 update on diagnosis, risk-stratification, and management. American journal of hematology **92**(10), 1085–1102 (2017)
17. Saeed, I., Zahra, M., Jeremy, K., Ghassan, H.: Generative adversarial networks to segment skin lesions. In: IEEE 15th International Symposium on Biomedical Imaging. pp. 881–884 (2018)
18. Siegel, R.L., Miller, K.D., Jemal, A.: Cancer statistics, 2015. CA: a cancer journal for clinicians **65**(1), 5–29 (2015)
19. Willemze, R., Jaffe E, S., Burg, G, e.a.: Who-eortc classification for cutaneous lymphomas. Blood **105**(10), 3768–3785 (2005)
20. Xiaomeng, L., Lequan, Y., Chi-Wing, F., Pheng-Ann, H.: Deeply supervised rotation equivariant network for lesion segmentation in dermoscopy images. In: OR 2.0 Context-Aware Operating Theaters, Computer Assisted Robotic Endoscopy, Clinical Image-Based Procedures, and Skin Image Analysis. pp. 235–243 (2018)
21. Zahra, M., Ghassan, H.: Star shape prior in fully convolutional networks for skin lesion segmentation. In: International Conference on Medical Image Computing and Computer-Assisted Intervention. pp. 737–745 (2018)