LESION ATTRIBUTES SEGMENTATION FOR MELANOMA DETECTION WITH MULTI-TASK U-NET

Eric Z. Chen[1], Xu Dong[2], Xiaoxiao Li[3], Hongda Jiang[4], Ruichen Rong[5] and Junyan Wu[6]

[1]Dana-Farber Cancer Institute, Boston, MA, USA [2]Virginia Tech, Blacksburg, VA, USA
[3]Yale University, New Haven, CT, USA [4]East China University of Science and Technology, Shanghai, China
[5]UT Southwestern Medical Center, Dallas, TX, USA [6]Cleerly Inc., New York City, New York, USA

Abstract

- Melanoma is the most deadly form of skin cancer worldwide. It is crucial to identify the specific lesion patterns for accurate diagnosis of melanoma. However, the common lesion patterns are not consistently present and cause **sparse label problems** in the data.
- we propose a **multi-task U-Net model** to automatically detect lesion attributes of melanoma. The network includes two tasks, one is the classification task to classify if the lesion attributes present, and the other is the segmentation task to segment the attributes in the images.
- Our multi-task U-Net model achieves a Jaccard index of 0.433 on official test data of ISIC 2018 Challenges task 2, which is the **best single-model result** and ranks the 5th place on the final leaderboard.

Methods

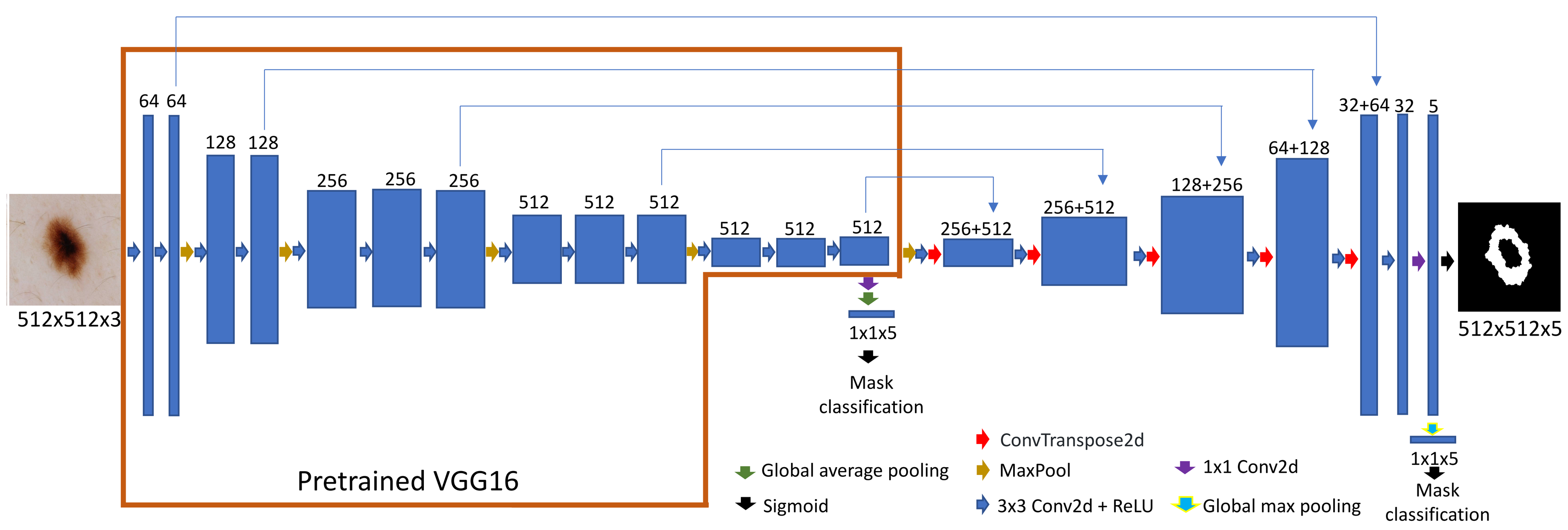

Figure 1. Network structure using U-Net architecture. The encoder part of the U-Net is replaced with a pretrained VGG16 network. We also add two classification heads to the network, where one is added to the middle layer with a 1 × 1 convolutional layer followed by a global average pooling layer and the other is added to the last layer with a global max pooling layer.

Loss function:
- Pixel-wise segmentation loss: a combination of binary cross entropy loss L and Jaccard index J: loss1= L - J
- Classification loss: binary cross entropy loss L using the middle layer (loss2) and using the last layer (loss3)
- Final loss: loss = loss1 + 0.5 x loss2 + 0.5 x loss3

Data:
- We downloaded the task 2 dataset from ISIC 2018 Challenges website.
- For each skin image, this task is aimed at detecting five lesion attributes (pigment network, negative network, streaks, milia-like cysts, and globules).
- The dataset consists of 2,594 images and 12,970 ground truth masks (five masks for each skin image) as training data. There are 100 images as validation data and 1,000 images as test data but no masks were provided.

Model training:
- We first resized all the skin images and corresponding masks to 512 × 512. Then for each skin image, we concatenated the five corresponding grayscale masks into a five-channel mask.
- Both skin images and mask images were divided by 255 to scale the values into [0,1]
- We first froze all the encoder layers and only trained the decoder layers for 50 epochs. Then we trained the whole network for another 250 epochs. During training, we augmented the skin images and masks by random flipping, rotation, scaling as well as randomly adjusting brightness and saturation.

Results

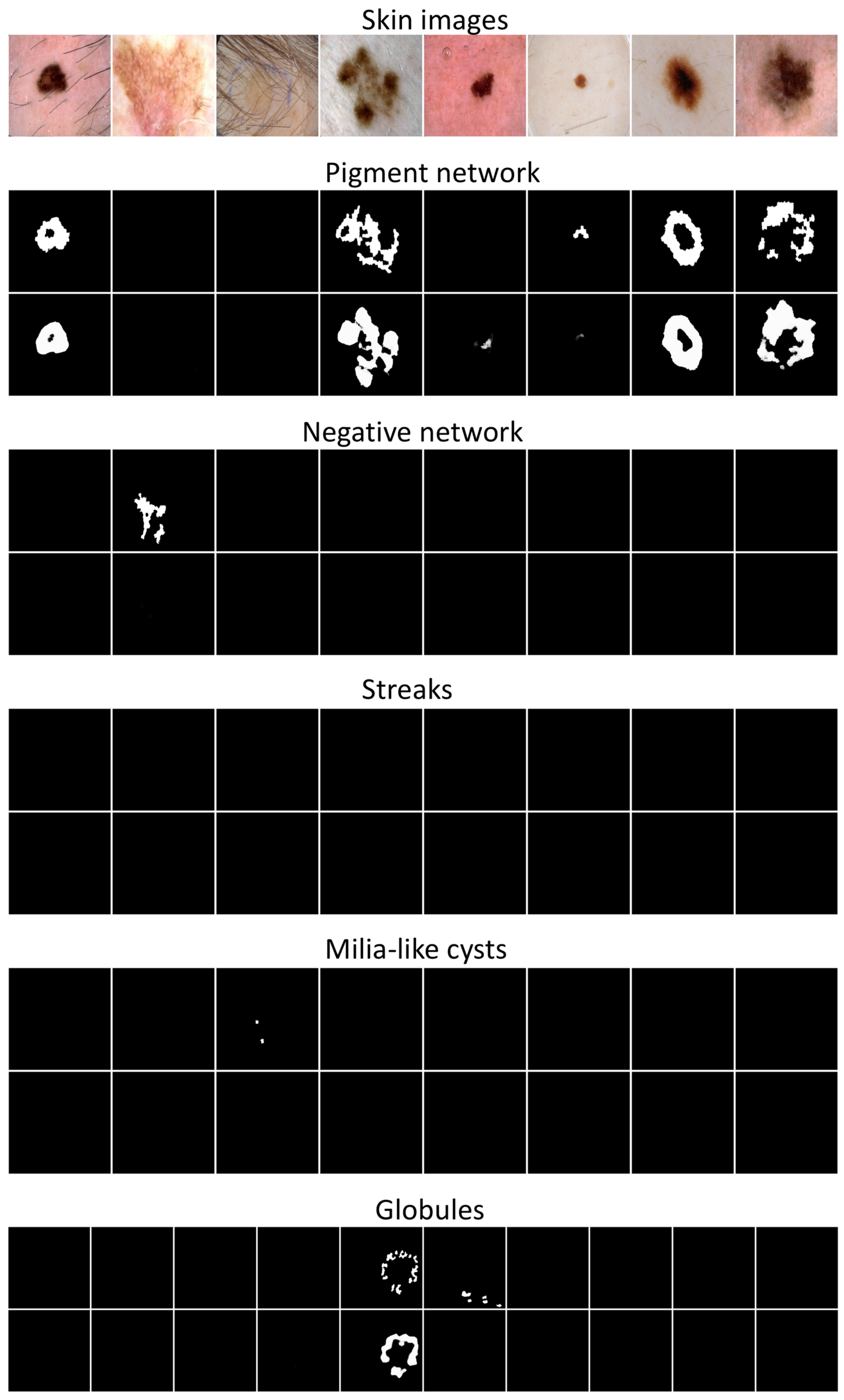

Figure 2. Examples of skin images and corresponding lesion attributes. For each attribute, the top row is the ground truth masks and the bottom row is the predicted masks from multi- task U-Net model. Note that many of the ground truth masks are empty, indicating no corresponding lesion attributes in the skin images.

Attributes	N	Percentage
Pigment network	1522	58.7%
Negative network	189	7.3%
Streaks	100	2.9%
Milia-like cysts	681	26.3%
Globules	602	23.2%
Total skin images	2594	100%

Table 1. Summary of non-empty masks in the training data.

Models	Jaccard Index
Multi-task U-Net	0.477
U-Net	0.262
Mask R-CNN	0.233

Table 2. Comparison of different models.

References

- ISIC 2018: Skin Lesion Analysis Towards Melanoma Detection https://challenge2018.isic-archive.com/
- Shvets et. al. Automatic instrument segmentation in robot-assisted surgery using deep learning. arXiv 2018