Segmentation of Skin Lesions by Superpixel Classification with Graph-Context CNN

Ángel Víctor Juanco Müller[1,2], João F.C. Mota[2][0000−0001−7263−8255], and Corné Hoogendoorn[1][0000−0002−4914−9936]

[1] Canon Medical Research Europe
[2] Heriot Watt University aj112@hw.ac.uk

1 Introduction

Accurate delineation of skin lesions is prone to inter-expert variability [4], which calls for automatic segmentation methods. In this paper, we process superpixel graphs with a novel Graph Convolution layer to segment skin lesions.

2 Methods & Data

First, the *Simple Linear Iterative Clustering* (SLIC) algorithm [2] divides an input image into n superpixels of approximately 250 pixels each. We then form a Region Adjacency Graph (RAG) over these partitions, which is the input of the proposed network (figure 1). Input features are the output of feeding 50×50 image patches, centred at the superpixel centroids, to a convolution layer.

Positional Encoding. The distances from superpixels to the central one, normalised to the interval $(0, 1)$, are element-wise added to the input features.

Graph-Context CNN (gc-CNN). We build on the Graph Convolution Network [9], where $\widetilde{\mathbf{A}} = \widehat{\mathbf{D}}^{-1/2} \cdot \widehat{\mathbf{A}} \cdot \widehat{\mathbf{D}}^{-1/2}$ guides the propagation of information. $\widehat{\mathbf{A}}$ is the adjacency matrix with self loops and $\widehat{\mathbf{D}}$ its degree matrix. At layer l, we convolve a tensor $\mathcal{X}^l$ of n feature maps, each with c channels, height h and width w, with a learnable kernel $\mathbf{K}_\Theta$. Fold and unfold operations are used to enable the propagation of features through the RAG. Putting it all together, the gc-CNN layer can be written as:

$$\mathcal{X}^{l+1} = \texttt{fold}_{f \to c \times h \times w}(\widetilde{\mathbf{A}} \cdot (\texttt{unfold}_{c \times h \times w \to f}(\mathcal{X}^l * \mathbf{K}_\Theta)) + \mathbf{b}_\Theta), \tag{1}$$

where $*$ denotes the 2-D convolution operation, and $\mathbf{b}_\Theta$ is a learnable bias vector. Sequences of gc-CNN, the ReLU non-linearity and layer normalization (LN) [3] are combined with skip connections [7] in the feature extraction block.

Feature Extraction Block. This block (figure 1) comprises a residual block and a transition block. The latter halves the spatial dimension, increases the channel dimension and updates $\widehat{\mathbf{A}}$ by re-utilizing the feature maps corresponding to neighbouring nodes. The mean value of their element-wise product, mapped to the range $(-1, 1)$ with the Tanh function and weighted by a parameter $\beta = 0.001$, are added to the entries of $\widehat{\mathbf{A}}$.

MLP block. A Multi Layer Perceptron (MLP) with layers of 2048, 1024, 512, 128, 32, 16 and 2 neurons, respectively, with ReLU activation and LN [3], predicts the lesion probability for the superpixels.

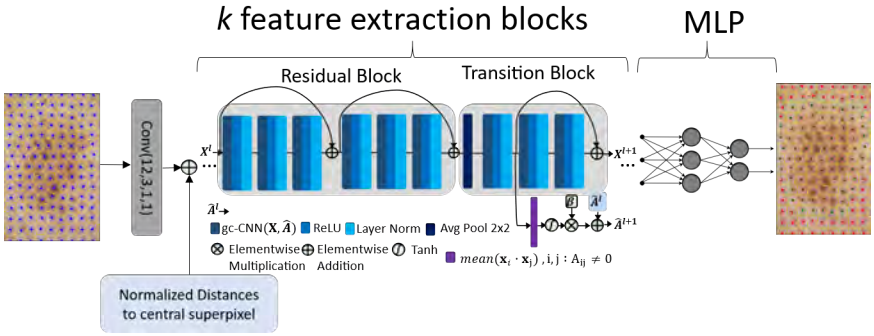

Fig. 1. Overall network, comprising k feature extraction blocks and a MLP. The feature extraction block is formed by a residual block and a transition block.

Training. We label a superpixel as *lesion* if more than a 60% of its pixels have that value in the ground truth[3]. The ADAM optimizer [8] minimizes the Cross Entropy function for 100 epochs with learning rate $\lambda = 5 \cdot 10^{-3}$ and batch size 1. We use Pytorch Geometric [5] on an 11GB NVIDIA GeForce GTX 1080 Ti.

Datasets. We train on the ISIC-2017 [6] training split; 6 crops are extracted and flipped for each image. We run inference on two dermatoscopic datasets, (ISIC-2017 [6] and PH^2 [10]), and a non-dermatoscopic dataset (DermQuest [1]) which has less controlled illumination conditions.

Post-processing. We keep only the connected component closest to the image centre, which reduces false positives caused by illumination gradients in DermQuest. We also apply binary opening and hole filling operations, which is useful on ISIC-2017 although it decreases performance on PH^2 (see figure 2).

3 Results & Conclusion

Results. Table 1 reports Sensitivity (SEN), Specificity (SPE), Accuracy (ACC) and F1 score (F1) for the different categories in the PH^2 dataset, comparing against the results reported in [11], where a colour distance criterium was used to merge superpixels. Figure 2 shows the F1 histograms for all three datasets.

Conclusion. We explore a segmentation technique that combines superpixel graphs with Deep Learning. Our results for melanoma lesions surpass those in [11], and our model generalises well to unseen data. Future work will seek to improve performance and apply the method to other medical domains.

[3] 60% yields an acceptable F1 score, around 90%, compared to the original masks.

Table 1. Our results on PH^2, in terms of mean and standard deviation, after post-processing, for models with various k, and those reported in [11]. NN stands for Normal Nevi, AN for Atypical Nevi, M for Melanoma. The highest figures are in bold.

Ref	Class	SEN	SPE	ACC	F1
Patiño *et al.* [11]	NN (80 images)	92.12	**96.42**	**95.24±6.37**	**92.92**
ours ($k = 1$)		97.27±5.66	91.01±6.08	92.88±3.50	82.46±9.46
ours ($k = 2$)		94.21±13.87	86.86±16.42	89.52±12.52	79.72±16.17
ours ($k = 3$)		**97.39±7.84**	86.29±15.54	89.23±12.28	79.23±16.36
Patiño *et al.* [11]	AN (80 images)	92.25	**93.54**	**93.14±8.41**	**91.12**
ours ($k = 1$)		96.87±4.70	90.31±8.69	92.98±4.20	86.65±7.18
ours ($k = 2$)		96.39±7.06	85.76±15.96	89.78±10.77	83.54±12.16
ours ($k = 3$)		**97.12±4.45**	87.73±13.89	91.38±9.60	85.32±11.76
Patiño *et al.* [11]	M (40 images)	**86.45**	68.70	75.19±22.16	77.79
ours ($k = 1$)		84.07±23.70	83.19 ±17.07	81.27±20.09	83.31±18.08
ours ($k = 2$)		89.26 ± 16.29	80.71 ±14.00	83.76±13.18	85.80±12.36
ours ($k = 3$)		86.32 ±20.54	**86.18 ±10.39**	**84.04±17.61**	**86.15±14.76**
Patiño *et al.* [11]	All (200 images)	91.04	**89.73**	90.39±14.19	**89.18**
ours ($k = 1$)		94.47±12.69	89.16±10.60	**90.60 ±10.70**	84.31±11.19
ours ($k = 2$)		94.09±12.52	85.19±15.94	88.47±12.23	82.47±14.14
ours ($k = 3$)		**95.07±11.67**	86.85±14.00	89.05 ±12.90	83.05±14.69

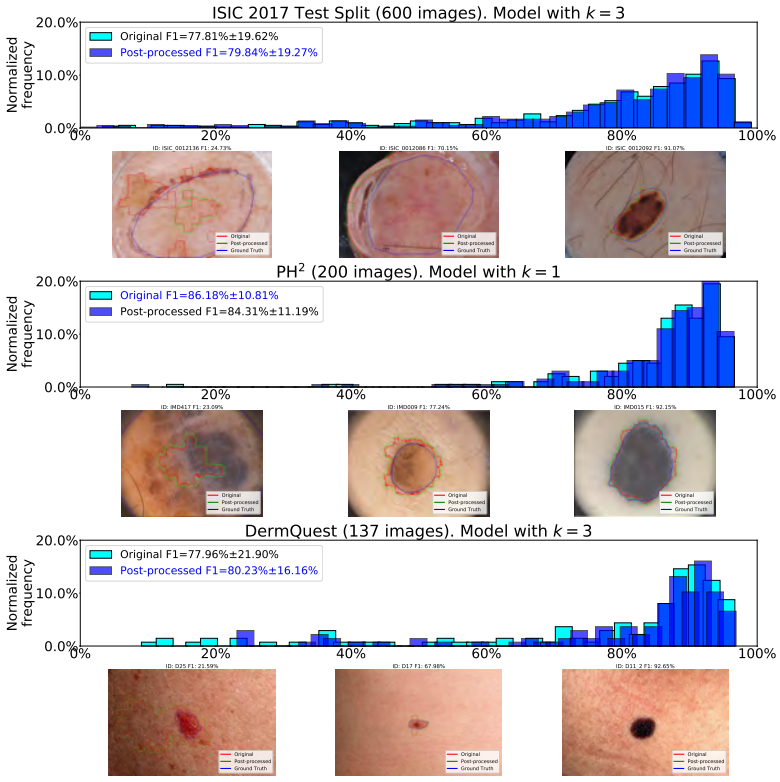

Fig. 2. F1 histograms and sample outputs for the ISIC-2017 test set (top), PH^2 (middle) and the DermQuest (bottom).

References

1. University of Waterloo VIPlab, https://uwaterloo.ca/vision-image-processing-lab/research-demos/skin-cancer-detection
2. Achanta, R., Süsstrunk, A.S., Smith, K., Lucchi, A., Fua, P.S.: SLIC Superpixels Compared to State-of-the-Art Superpixel Methods. IEEE Transactions on Pattern Analysis and Machine Intelligence **34**(11), 2274–2282 (2012)
3. Ba, J.L., Kiros, J.R., Hinton, G.E.: Layer Normalization. In: NIPS 2016. vol. 1050, p. 21 (2016), http://arxiv.org/abs/1607.06450
4. Bogo, F., Peruch, F., Fortina, A.B.., Peserico, E.: Where's the Lesion?: Variability in Human and Automated Segmentation of Dermoscopy Images of Melanocytic Skin Lesions. In: Dermoscopy image analysis, pp. 67–96 (2015)
5. Fey, M., Lenssen, J.E.: Fast graph representation learning with pytorch geometric. In: ICLR 2019. pp. 1–9. No. 1 (2019)
6. Halpern, A., Helba, B., Gutman, D., Codella, N., Celebi, M.E.: 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). arXiv: 1710.05006 https://arxiv.org/abs/1710.05006
7. He, K., Zhang, X., Ren, S., Sun, J.: Deep residual learning for image recognition. In: 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). pp. 770–778 (2016). https://doi.org/10.1109/CVPR.2016.90
8. Kingma, D.P., Ba, J.L.: Adam: A method for stochastic optimization. In: ICLR 2015. pp. 1–15 (2015)
9. Kipf, T.N., Welling, M.: Semi-supervised classification with graph convolutional networks. In: ICLR. pp. 1–14 (2017)
10. Mendonça, T.F., Ferreira, P.M., Marçal, A., Catarina Barata, J., Joana Rocha, J.: Ph2: A public database for the analysis of dermoscopic images. Dermoscopy image analysis. (2015)
11. Patiño, D., Avendaño, J., Branch, J.W.: Automatic skin lesion segmentation on dermoscopic images by the means of superpixel merging. In: MICCAI 2018. vol. 1, pp. 728–736 (2018)