Integrating clinically-relevant features into skin lesion classification

Veltmeijer, E.; Karaoglu, S.; Gevers, T.

Publication date
2019
Document Version
Final published version
Published in
BNAIC/BENELEARN 2019 : proceedings of the Reference AI & ML Conference for Belgium, Netherlands & Luxemburg
License
CC BY

Link to publication

Citation for published version (APA):
Veltmeijer, E., Karaoglu, S., & Gevers, T. (2019). Integrating clinically-relevant features into skin lesion classification. In K. Beuls, B. Bogaerts, G. Bontempi, P. Geurts, N. Harley, B. Lebichot, T. Lenaerts, G. Louppe, & P. Van Eecke (Eds.), *BNAIC/BENELEARN 2019 : proceedings of the Reference AI & ML Conference for Belgium, Netherlands & Luxemburg: proceedings of the 31st Benelux Conference on Artificial Intelligence (BNAIC 2019) and the 28th Belgian Dutch Conference on Machine Learning (Benelearn 2019), Brussels, Belgium, November 6-8, 2019* [127] (CEUR Workshop Proceedings; Vol. 2491). CEUR-WS. http://ceur-ws.org/Vol-2491/abstract127.pdf

Integrating clinically-relevant features into skin lesion classification

Emmeke A. Veltmeijer[1,2], Sezer Karaoglu[2], and Theo Gevers[1,2]

[1] University of Amsterdam, Amsterdam, The Netherlands
[2] 3DUniversum, Amsterdam, The Netherlands

Keywords: melanoma· skin lesion classification· multi-task learning· explainable AI.

1 Introduction

Skin cancer is one of the most common types of cancer, melanoma being the most deadly form of skin cancer [3]. The treatment of melanoma benefits tremendously from early detection, survival rates quadrupling with an early treatment as compared to a late detection and treatment [3,1]. Thus, society has a lot to gain from improving the (early) detection of skin cancer, and melanoma specifically. Computer-aided diagnosis systems could help with this, by improving the early detection accuracy and lowering the threshold for patients to get a first assessment. However, such a diagnosing model usually outputs nothing but a classification, making it difficult for both doctor and patient to understand and accept the outcome. To increase both the performance and explainability of the model, we propose the integration of clinically-relevant features into the task of skin lesion classification in a multi-task framework.

2 Methods

The integration of these clinically-relevant features is done by building a multi-task neural network, where relevant features serve as auxiliary (guiding) output next to the main classification output. This multi-task framework guides the network during training as well as serving to increase the model's explainability during testing. The latter is done by indicating the presence or absence of these relevant features, thereby 'explaining' the main classification output. Several data augmentation techniques, including GANs, are analyzed for the purpose of truthfully increasing the number of images for different skin lesion classes, thereby reducing overfitting.
When an expert performs a first assessment of a skin lesion, it focuses on certain features of the lesion that could be a sign of malignancy (melanoma). A common used method is the ABCDE-rule [4]. We annotate the dataset of the ISIC Challenge 2017 [2] with the presence of the first three of these features (**A**symmetry,

irregular **B**orders, multiple **C**olours) in each skin lesion image, since these can be derived from 2D images with no extra information about scale or temporal change. This information is then provided to the network as auxiliary output to train on, next to the main ground truth classification (classifying malignant melanoma and two benign skin lesion types). Next to this multi-task network using clinically-relevant image features, we experiment with another multi-task network using segmentation masks of the images as auxiliary output, as these provide valuable and detailed information about some of these clinically-relevant features (the asymmetry and border shape).

3 Results

We find that the combination of classic data augmentation and GAN data augmentation works best for the classification performance, the latter generating realistic results that follow the diagnostic criteria for each class. Using ABC-features as auxiliary output, the classification performance does not improve. Instead, for a specific benign skin lesion class, nevus, performance decreases. We find that the (malignant) melanoma class does significantly benefit from the use of segmentation masks as auxiliary output.

4 Discussion

The class that is negatively influenced by the addition of ABC-features as auxiliary output, nevus, is possibly not able to retrieve useful information from them: the other two classes are annotated to either have the highest presence of ABC-features (the malignant melanoma class), or the lowest presence (a benign class). The nevus class falls in between, and we suggest it does therefore not benefit from the added features. Instead, these might 'distract' the network from learning the right features to classify nevus images.
The finding that the (malignant) melanoma class benefits from the use of segmentation masks as auxiliary output is according to our hypothesis: this class can be distinguished from the other two by its asymmetrical shape and irregular border (the other two typically having symmetrical shapes and regular borders), information that is present in the segmentation masks.

References

1. American Cancer Society. Survival rates for melanoma skin cancer, https://www.cancer.org/cancer/melanoma-skin-cancer/detection-diagnosis-staging/survival-rates-for-melanoma-skin-cancer-by-stage.html. Accessed: on 14 June 2019
2. Codella, N. C., Gutman, D., Celebi, M. E., ... and Halpern, A.: Skin lesion analysis toward melanoma detection. In: 15th ISBI, 168–172. IEEE (2018)
3. Siegel, R. L., Miller, K. D., and Jemal, A.: Cancer statistics, 2019. CA: a cancer journal for clinicians **69**(1), 7–34 (2019)
4. Tsao, H., Olazagasti, J. M., ... & Begolka, W. S.: Early detection of melanoma: reviewing the ABCDEs. Journal of the Am. Acad. of Derm. **72**(4), 717–723 (2015)