MelaGo: Gamifying Medical Image Annotation

Vassilis-Javed Khan[1], Pim Meijer[1], Michele Paludetti[1], Reka Magyari[1], Dominique van Berkum[1] and Veronika Cheplygina[2]

[1]Department of Industrial Design, Eindhoven University of Technology
[2]Department of Biomedical Engineering, Eindhoven University of Technology

1 Introduction

Machine learning (ML) algorithms for medical image analysis can be very effective, but typically need a lot of annotated images for good performance. In computer vision, crowdsourcing has been used to gather datasets with millions of images. Recently, a few studies have also shown promising results in medical imaging, for example [1].

Unfortunately standard crowdsourcing platforms were not built with medical images in mind, and lack interaction possibilities that would be desirable (i) to increase the ease of annotation, and (ii) to motivate the annotators, and as such, to increase quantity and/or quality. Thus the role of designing appropriate interfaces for crowds is of outmost importance for further expanding application of crowdsourcing annotations. For instance, designing gamification features has shown to positively affect crowd workers [2].

2 Methods

We explore design features of MelaGo: an app that supports easy annotation of skin lesion images (Fig. 1, left). The app targets annotation of visual attributes: asymmetry, border irregularity, and color variety, known to be used by experts [3]. Each user has their own "patient", visually represented by an avatar, whose physical state will increase when images are annotated, and decrease when the app is neglected.

After a short training phase, the users are presented with pairs of images from the ISIC 2017 challenge [4], and asked to indicate which image is more asymmetric and so forth. While ML algorithms for skin lesion classification cannot be trained with such inputs alone, they could potentially make training with existing datasets more robust.

In an initial test of the app, we examined whether gamification features affected the quantity of the annotations produced. Twenty participants (students from Eindhoven University of Technology) were recruited and offered a gift certificate for participating in the study. Each participant was assigned into a group which had the app with gamification features (progress bar of completing the current level, the "patient" represented by an avatar, unlockable voucher), or a control group.

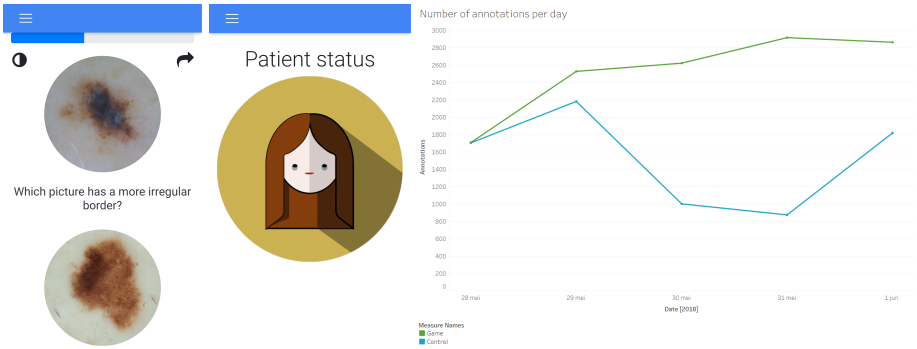

Fig. 1. Left: Screenshots from the MelaGo app, available at `http://www.melago.xyz`. Right: number of annotations per day by the groups with and without gamification.

3 Results and Discussion

We found that the gamification features positively affected the quantity annotations made. In Fig. 1 (right) we show the total number of annotations made by day by each of the two groups. While initially both groups seemed to be making similar progress, the interest of the control group then dropped. Although it increased again close to the end of the testing period, it still didn't match the continued interest of the gamification group.

The quality of the annotations is more difficult to assess since we only have ground truth whether the lesion is a melanoma or not. We did find that the gamification group spent more time per annotation, suggesting their annotations could be of higher quality. In follow-up work we plan to use the pairwise annotations with representation learning algorithms and compare classifiers trained on the representations, to assess the quality of each group.

References

1. Cheplygina, V., Perez-Rovira, A., Kuo, W., Tiddens, H., de Bruijne, M.: Early experiences with crowdsourcing airway annotations in chest CT. In: Large-scale Annotation of Biomedical data and Expert Label Synthesis (MICCAI LABELS). (2016) 209–218
2. Hamari, J., Koivisto, J., Sarsa, H.: Does gamification work? - A literature review of empirical studies on gamification. Proceedings of the Annual Hawaii International Conference on System Sciences (2014) 3025–3034
3. Abbasi, N.R., Shaw, H.M., Rigel, D.S., Friedman, R.J., McCarthy, W.H., Osman, I., Kopf, A.W., Polsky, D.: Early diagnosis of cutaneous melanoma: revisiting the abcd criteria. Jama **292**(22) (2004) 2771–2776
4. Codella, N.C., Gutman, D., Celebi, M.E., Helba, B., Marchetti, M.A., Dusza, S.W., Kalloo, A., Liopyris, K., Mishra, N., Kittler, H., et al.: 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 preprint arXiv:1710.05006 (2017)