



Deep Learning for Medical Applications in Deficiency of Labels

Christoph Baur¹ <c.baur@tum.de>

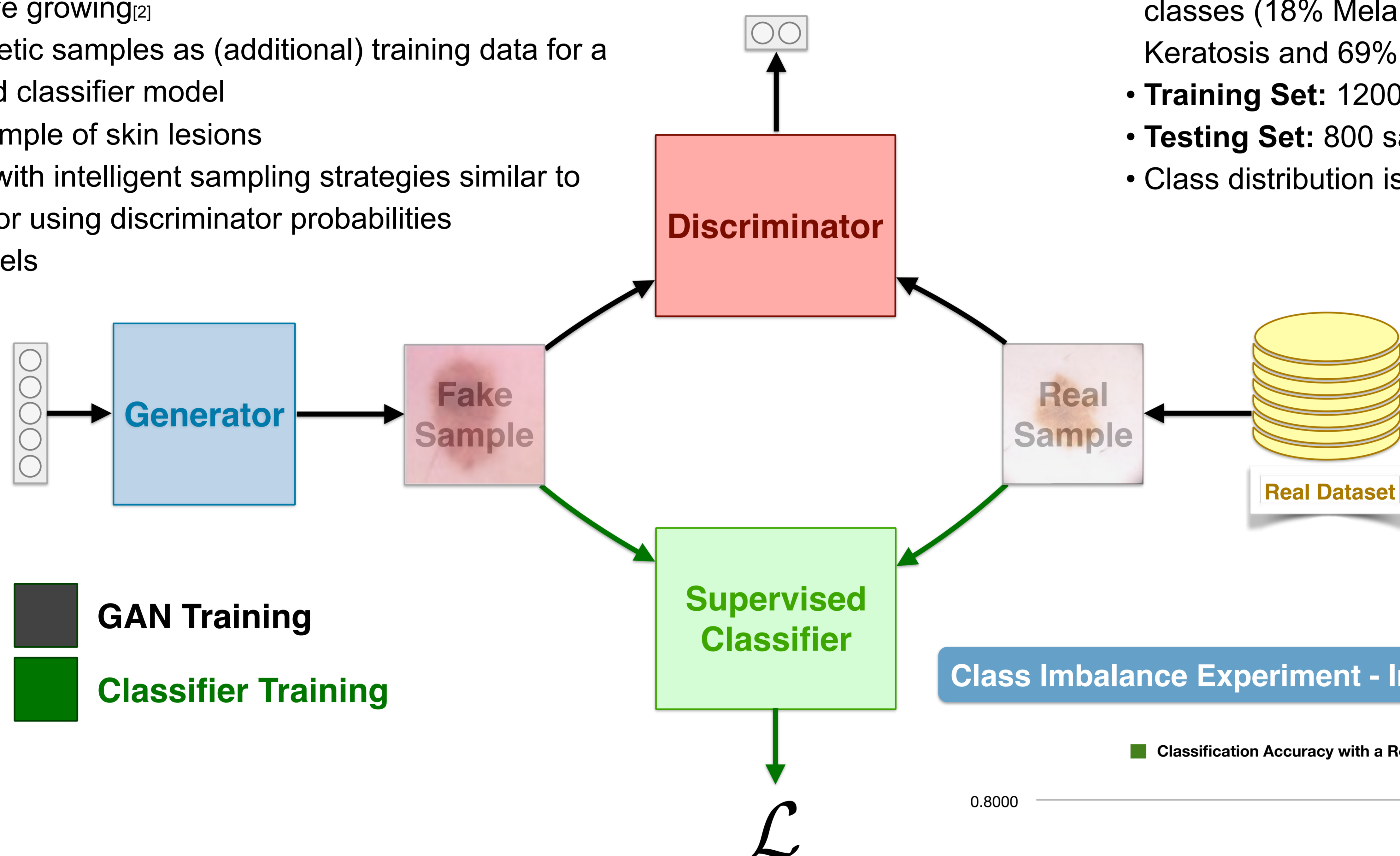
1. Computer Aided Medical Procedures (CAMP), Technische Universität München, Munich, Germany

Overview and Motivation

- Supervised Machine Learning in the Medical Image Analysis Field is generally impeded by a lack of labeled data
- Generative Adversarial Networks^[1] (GANs) have shown the capability to synthesize images with unprecedented levels of realism
- Recent work has been devoted to the synthesis of high resolution images and has shown outstanding results^[2]
- We aim to leverage synthetic samples generated with GANs to tackle the problems of data scarcity and class imbalance

Methodology

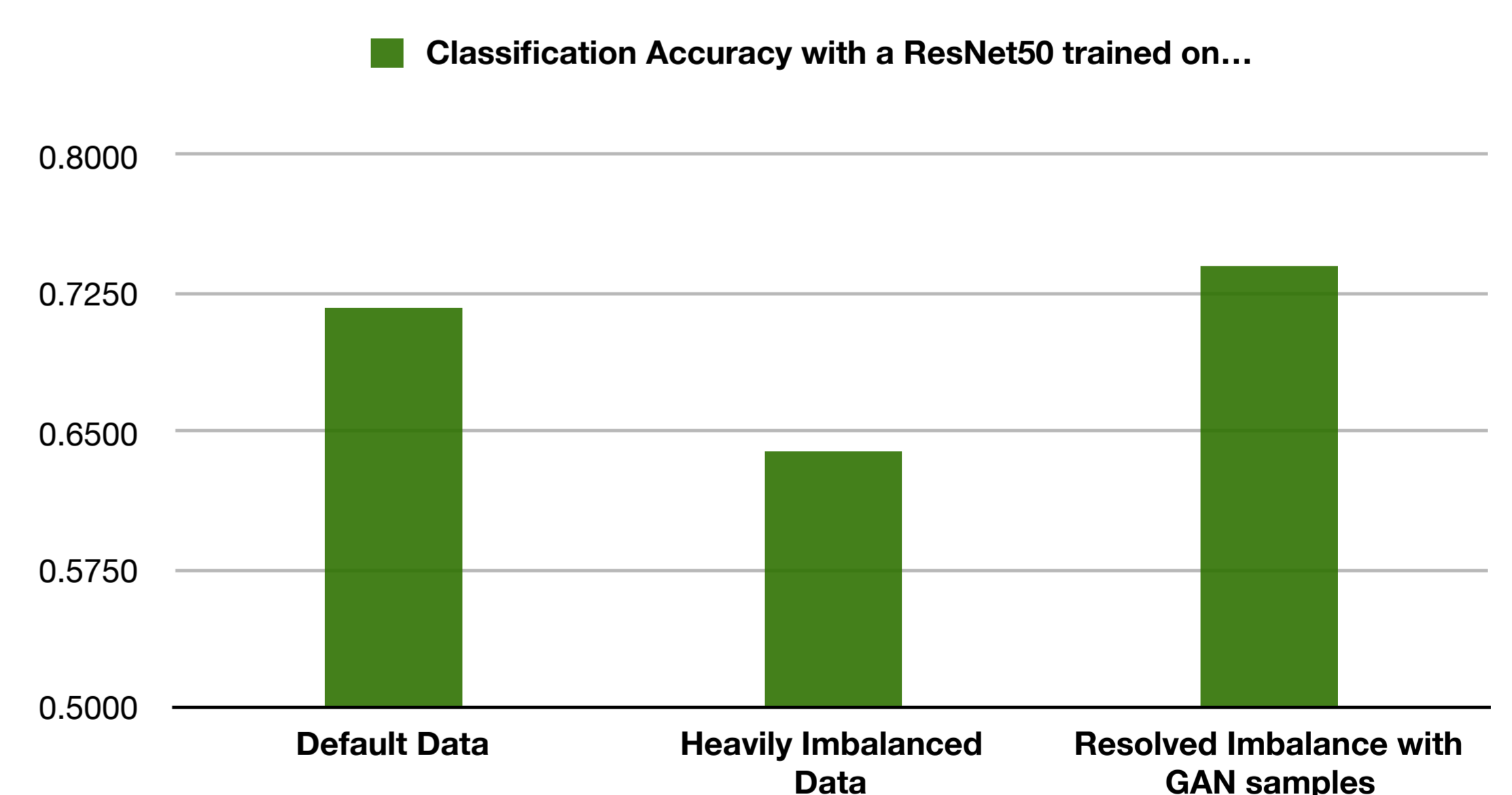
- Train GANs for realistic, high resolution image synthesis (256x256px and above) using progressive growing^[2]
- Use synthetic samples as (additional) training data for a supervised classifier model
- At the example of skin lesions
- Combine with intelligent sampling strategies similar to CASED^[4] or using discriminator probabilities as soft labels



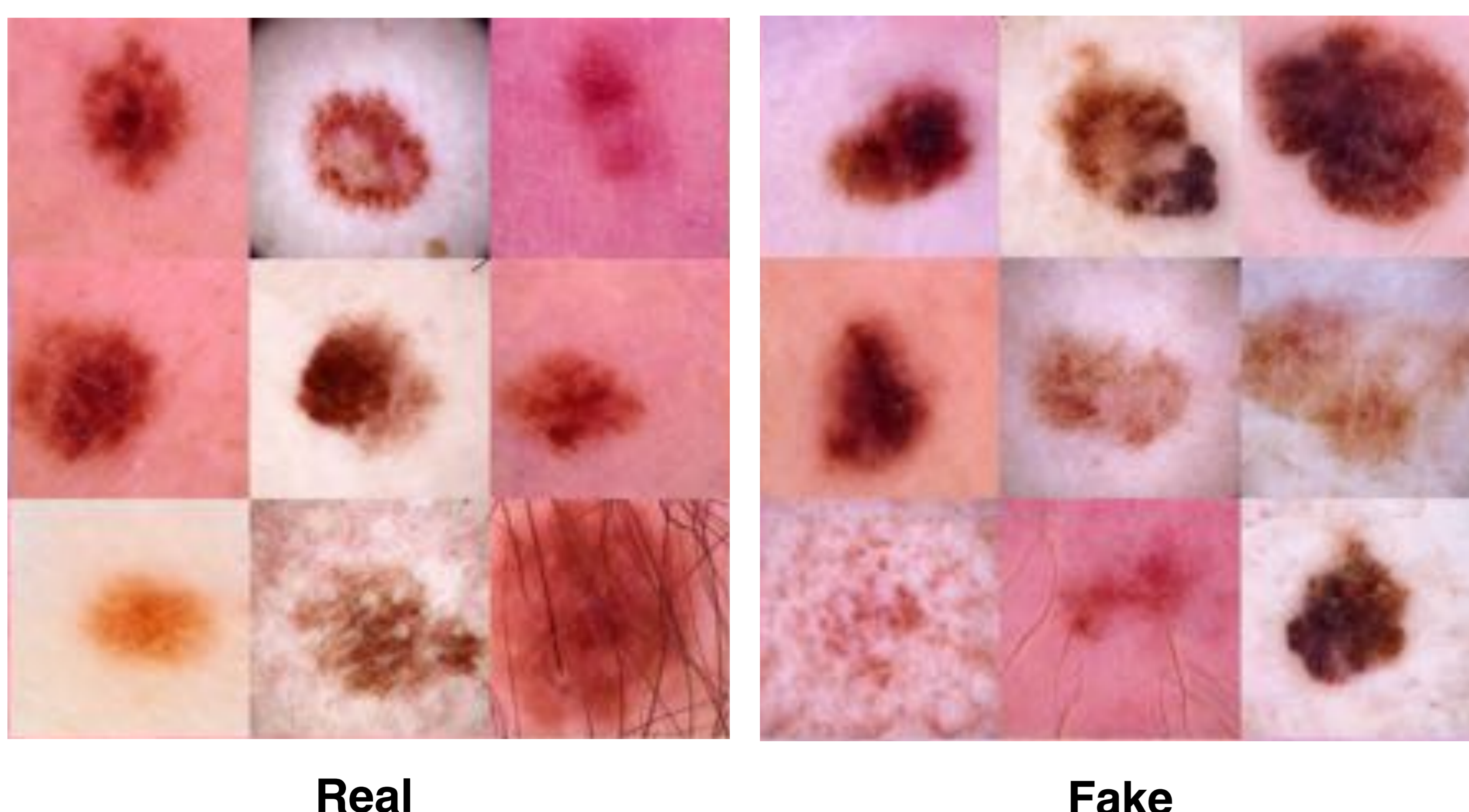
Dataset

- **ISIC2017** Dataset^[3]: 2000 images with 3 classes (18% Melanoma, 13% Seborrheic Keratosis and 69% Benign Nevi)
- **Training Set**: 1200 samples
- **Testing Set**: 800 samples
- Class distribution is preserved in each set

Class Imbalance Experiment - Initial Results



Real Data vs Synthesis Results



Conclusion & Open Questions

- Synthesis of visually appealing lesions works well
- Initial results show that synthetic samples can be used to augment datasets for resolving class imbalance
- **But:** How to evaluate realism of generated samples quantitatively and qualitatively?
- **But:** Realistic synthesis of granular structures such as hair
- **But:** How to reliably investigate memorization of the GAN training dataset?

References

- [1] Goodfellow, I., Pouget-Abadie, J., Mirza, M., Xu, B., Warde-Farley, D., Ozair, S., Courville, A. and Bengio, Y., 2014. Generative adversarial nets. In Advances in neural information processing systems (pp. 2672-2680).
- [2] Karras, T., Aila, T., Laine, S. and Lehtinen, J., 2017. Progressive growing of gans for improved quality, stability, and variation. arXiv preprint arXiv:1710.10196.
- [3] https://challenge.kitware.com/#challenge/ISIC_2017%3A_Skin_Lesion_Analysis_Towards_Melanoma_Detection
- [4] Jesson, A., Guizard, N., Ghalehjegh, S.H., Goblot, D., Soudan, F. and Chapados, N., 2017, September. CASED: Curriculum Adaptive Sampling for Extreme Data Imbalance. In International Conference on Medical Image Computing and Computer-Assisted Intervention (pp. 639-646). Springer, Cham.

