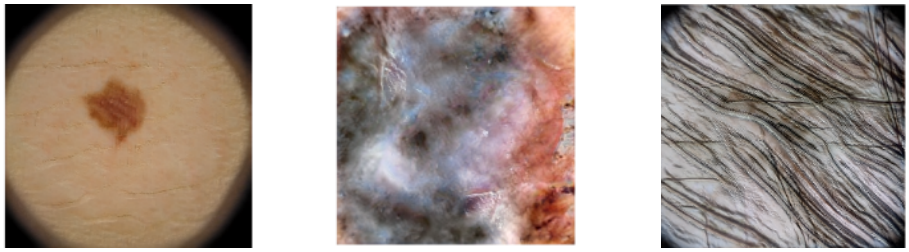
Skin Disease Diagnosis with Adversarial Training and Diffusion Models

Supervisors: Filipe Silva (DETI-IEETA, fmsilva@ua.pt), Pétia Georgieva (DETI-IEETA, petia@ua.pt)

INTRODUCTION

Skin lesions are a growing health concern requiring early and accurate diagnosis for successful treatment. Dermatologists rely on visual examination and biopsies, but subjectivity in visual assessment and limitations in access to specialists can create difficulties for early detection. Deep learning has achieved remarkable success in medical image analysis, with convolutional neural networks (CNNs) demonstrating promising results in skin lesion classification [1,2]. However, a significant challenge in training these models lies in the limited availability of high-quality, labeled medical data. **Generative models** can address this problem by creating synthetic images that augment the real dataset, enhancing the model's ability to generalize predictions. While Generative Adversarial Networks (GANs) have been a popular choice for generating images [3], **diffusion models** [4] offer advantages in the context of skin lesion diagnosis. First, by learning the underlying distribution of real data, these models often provide more realistic and diverse images compared to GANs, who suffer from artifacts and inconsistencies (see Fig. 1). Second, GANs involve a complex training process where a generator and a discriminator compete with each other, making them unstable and prone to collapse. Conversely, diffusion models follow a well-defined training procedure with a clear objective, leading to more stable and reliable training.



(a) (b) (c)

Fig. 1: Examples of synthetic melanoma images generated by a StyleGAN2-Ada model with unrealistic details [1]: (a) checkerboard pattern, (b) mode collapse and (c) lack of skin lesion.

OBJECTIVES

This dissertation proposes a structured approach to investigate the potential of adversarial training with diffusion models in order to address data scarcity and improving the generalizability of deep learning classifiers. In the critical task of skin lesion classification, where subtle details can be crucial, the higher fidelity of diffusion-generated images can be advantageous [5,6]. Additionally, diffusion models allow for more control over the generation process by guiding the model towards specific image features. This can be particularly beneficial in skin lesion diagnosis, where focusing on generating synthetic images that mimic challenging or rare lesion types can improve the classifier's ability to handle those cases.

Bearing this in mind, the research will focus on the following specific objectives:

1. Train a deep CNN classifier for skin lesion classification using a dataset of dermoscopy images.
2. Develop a diffusion model to generate synthetic skin lesion images that are similar to real data.
3. Implement an adversarial (iterative) training process where the diffusion model is guided to generate synthetic data that targets challenging cases identified on a held-out test set.
4. Evaluate the impact of the generated synthetic data on the classifier's performance, comparing its accuracy, sensitivity, and specificity on the test set before and after adversarial training.

WORK PLAN

The main tasks to be carried out can be summarized in the following points:

1. **Literature Review**
   - Review research on skin lesion diagnosis using deep learning, encompassing performance metrics, data augmentation techniques, synthetic data generation, and diffusion models.
   - Become familiar with publicly available datasets for research on skin lesion diagnosis, such as the ISIC Archive (International Skin Imaging Collaboration Archive, https://www.isic-archive.com/), the ISIC melanoma dataset (https://www.kaggle.com/datasets/andrewmvd/isic-2019), the HAM10000 dataset (https://paperswithcode.com/dataset/ham10000-1), among others.

2. **Deep Classifier Deployment**
   - Choose a pre-trained CNN architecture suitable for image classification and train the classifier on the real dataset of skin lesion images.
   - Evaluate the classifier's performance on the separate test set using metrics like accuracy, sensitivity, and specificity.

3. **Diffusion Model**
   - Choose a diffusion model architecture for generating synthetic skin lesions and train the model on the real dataset of skin lesion images.
   - Tune hyperparameters of the diffusion model for optimal performance and assess the quality and realism of the generated synthetic images.

4. **Adversarial Training**
   - Develop an iterative training loop: (i) identify misclassified images from the test set; (ii) use these challenging cases to guide the diffusion model to generate synthetic data that replicates their features; and (iii) re-train the classifier on the real and newly generated synthetic data.
   - Compare the performance metrics before and after adversarial training to assess the effectiveness of the synthetic data in improving classification, particularly for challenging cases.

5. **Dissertation and Documentation.** The final stage involves writing the master dissertation and additional documentation, such as code repositories and reports, to ensure reproducibility and facilitate future research.

REFERENCES

[1] Choy, S. P., Kim, B. J., Paolino, A., Tan, W. R., Lim, S. M. L., Seo, J., ... & Mahil, S. K. (2023). Systematic review of deep learning image analyses for the diagnosis and monitoring of skin disease. *NPJ Digital Medicine*, *6*(1), 180.

[2] Venkatesh, K. P., Raza, M. M., Nickel, G., Wang, S., & Kvedar, J. C. (2024). Deep learning models across the range of skin disease. *npj Digital Medicine*, *7*(1), 32.

[3] Silva, G. (2021). Data augmentation and deep classification with generative adversarial networks. Master's dissertation, University of Aveiro. https://ria.ua.pt/handle/10773/32283

[4] Kazerouni, A., Aghdam, E. K., Heidari, M., Azad, R., Fayyaz, M., Hacihaliloglu, I., & Merhof, D. (2023). Diffusion models in medical imaging: A comprehensive survey. *Medical Image Analysis*, 102846.

[5] Kincso, B., Solis, S., Cirone, K., Kawahara, J., Slade, D., Abid, L., ... & Fazekas, I. (2024). Diffusion-Based Data Augmentation for Skin Disease Classification: Impact Across Original Medical Datasets to Fully Synthetic Images. In *Deep Generative Models: Third MICCAI Workshop, DGM4MICCAI 2023, Held in Conjunction with MICCAI 2023, Vancouver, BC, Canada, October 8, 2023, Proceedings* (Vol. 14533, p. 99). Springer Nature.

[6] Sagers, L. W., Diao, J. A., Groh, M., Rajpurkar, P., Adamson, A. S., & Manrai, A. K. (2022). Improving dermatology classifiers across populations using images generated by large diffusion models. *arXiv preprint arXiv:2211.13352*.