Generating High Quality Synthetic Skin Lesions for Boosting Automated Screening

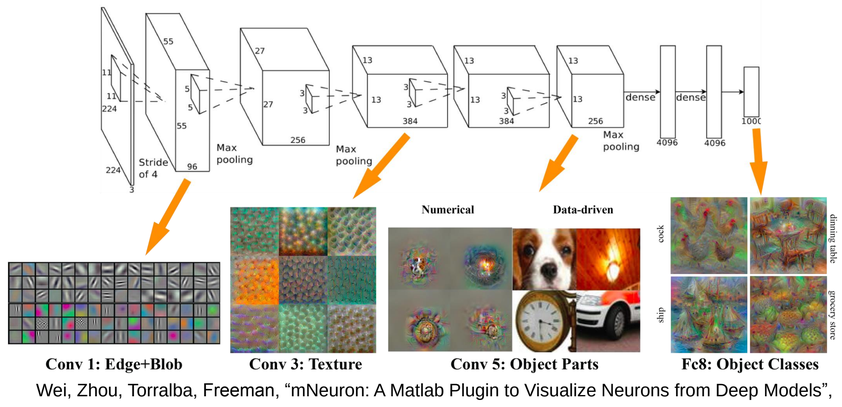
A. Bissoto, M. Fornaciali, E. Valle, S. Avila

University of Campinas, Campinas, Brazil

Melanoma incidence increases much faster then our ability to train new specialists. In that scenario, automatic triage is very attractive to optimize the time of those scarce professionals.

State of the art of automated skin lesion classification is dominated by the so called "deep learning" technique. Those are neural networks with many layers (sometimes more than 100), which have just recently become possible.

In the deep learning (below) the classification is made directly from the image, without the necessity of previous step of feature extraction: the technique itself learns to extract the relevant features; the model has many layers, which learn more and more abstract and specialized features



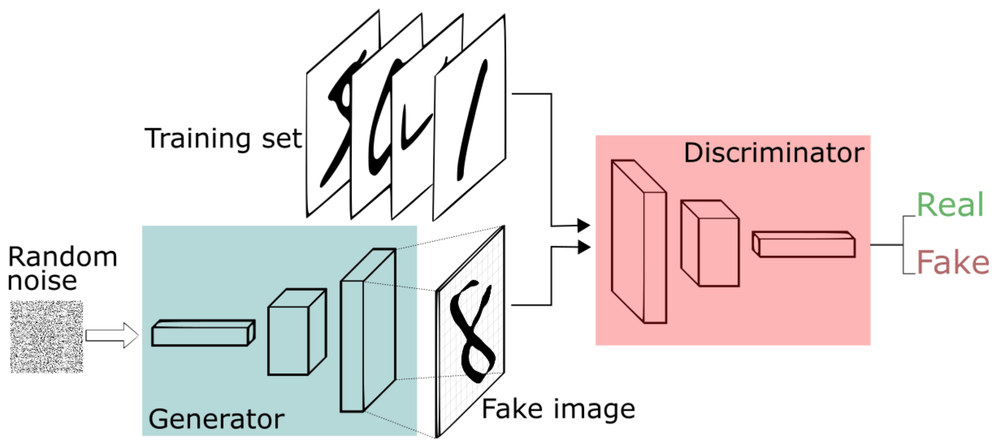
Wei, Zhou, Torralba, Freeman, “mNeuron: A Matlab Plugin to Visualize Neurons from Deep Models”, *vision03.csail.mit.edu/cnn_art/*

Deep learning is greedy for large amounts of annotated data, which is not common for medical applications. Medical imaging databases have at most a few thousand images, while deep neural networks are typically trained with hundreds of thousands.

Our main objective is to synthesize clinically-meaningful skin lesion images that are able to boost the predictive power of automatic triage systems.

We aim to generate synthetic lesions using the Generative Adversarial Network (GANs) approach.

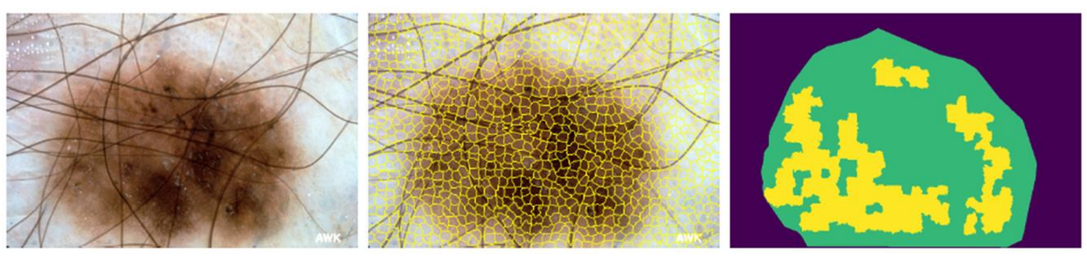
GANs model the real image appearance by forcing the synthesized samples to be really similar from real images



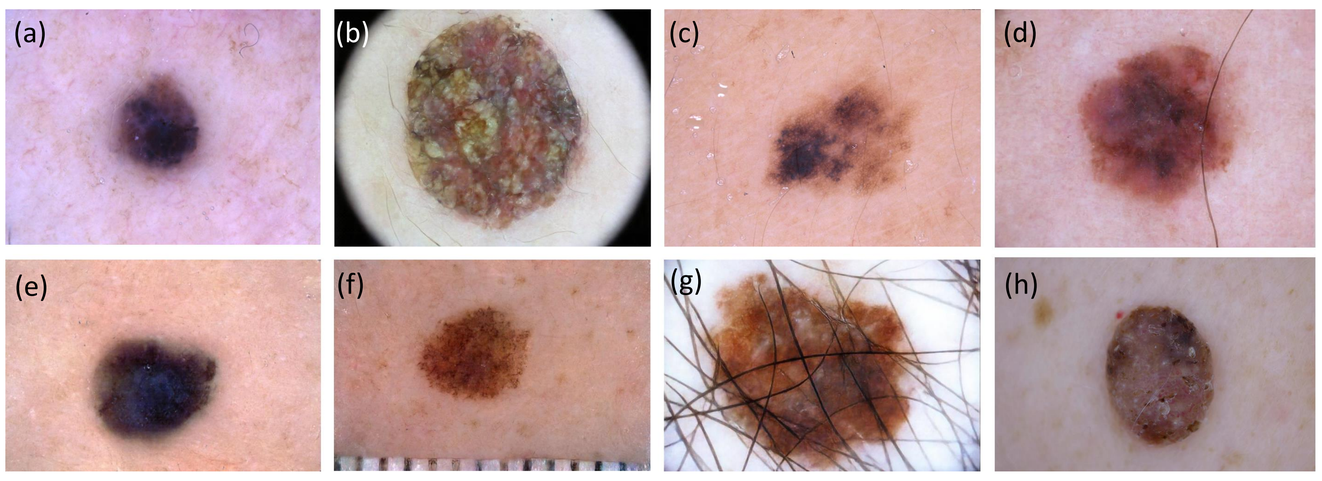
Reproduced from https://deeplearning4j.org/generative-adversarial-network © DL4J.org

Instead of following the traditional approach of generating images from random noise, we teach the network the malignancy markers — pigment network, negative network, streaks, milia-like cysts, and globules — while incorporating the specificities of a lesion border.

We feed such information directly to the network, using a **semantic map** and an **instance map**. Semantic maps are blobs that show the presence and the location of the 5 malignancy markers within the same lesions' segmentation masks. Instance maps take information from superpixels, which group similar pixels creating visually meaningful blobs, limiting each unit regarding their meaning.

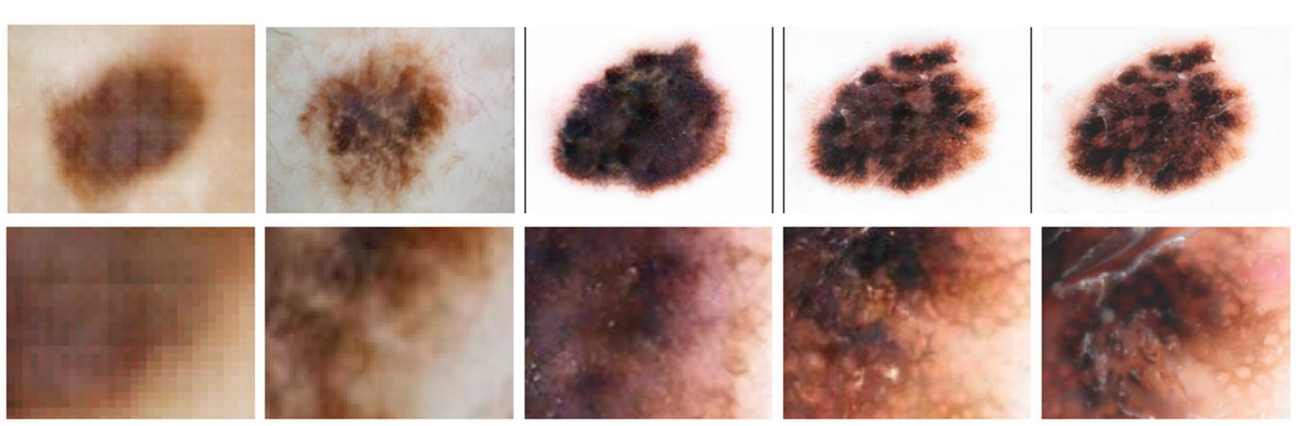


Real image | Instance map (Superpixels) | Semantic label map



Synthetic skin lesion images can be a solution to the scarcity of medical images for training deep learning models for automated triage. Among the 8 images above, 3 are real skin lesions and 5 are synthetic. Can you guess which ones are real? See the answer at the bottom right of the poster.

We compared four approaches for generating synthetic lesions: (a) a traditional GAN; (b) the progressive GAN: The model is progressively fed with images of increasing resolution along training; (c) pix2pixHD GAN: The model receive as inputs (c) only semantic maps; and (d) both semantic and instance maps. This model architecture is specialized to work with high-resolution samples (requirement for skin lesion images). For quality comparison, (e) is the real image. In the figure below, the first row are the full images while the second row are the zoom-in to focus on the details.



(a) (b) (c) (d) (e)

We evaluate the impact of inflating the training set of a classification network with synthetic images. Comparing the area under the ROC curve (AUC) of both experiments, we confirm the synthetic images contain features that characterize a lesion as malignant or benign. Even more, the results suggest the synthetic images contain features that are beyond the boundaries of the real images, which improves the classification network by an average of 1.3 percentage point and keeps the network more stable.

Training data	AUC (%)
Real images	83.4 ± 0.9
Real + synthetic images	84.7 ± 0.5

We demonstrate the applicability of GANs to generate high-definition, visually-appealing, and clinically-meaningful synthetic skin lesion images. We also show that incorporating them to the training dataset of automatic skin lesion classifiers improves the predictive power of the method, not only by the size of the extended dataset, but mainly due to the similarity of the synthetic lesions with real skin lesion.

This work is part of a long-term project to improve automatic melanoma screening and triage.
Know our publications: *sites.google.com/site/robustmelanomascreening*

An idea? A question? A collaboration opportunity? Stay in touch:
sandra@ic.unicamp.br · alceubissoto@gmail.com · michelfornaciali@gmail.com
recodbr.wordpress.com · facebook.com/recodbr · twitter.com/recodbr

A. Bissoto is funded by a CNPq grant. M. Fornaciali and E. Valle are partially funded by Google Research Awards for Latin America. E. Valle is also partially funded by a CNPq PQ-2 grant (311905/2017-0) and Universal grant (424958/2016-3). We gratefully acknowledge NVIDIA for the donation of GPUs, Microsoft Azure for the GPU-powered cloud platform, and CCES/Unicamp (Center for Computational Engineering & Sciences) for the GPUs used in this work. The RECOD Lab is funded from FAPESP, CAPES, and CNPq.

The real ones: (a), (c) and (h)