# Data Augmentation from Sketch

Debora Gil[1], Antonio Esteban-Lansaque[1], Sebastian Stefaniga[2], Mihail Gaianu[2], and Carles Sanchez[1]

[1] Comp. Vision Center and Comp. Science Dept. UAB. Barcelona.
{debora,csanchez}@cvc.uab.es
[2] West University of Timisoara, Comp. Science Dept, Romania

**Abstract.** State of the art machine learning methods need huge amounts of data with unambiguous annotations for their training. In the context of medical imaging this is, in general, a very difficult task due to limited access to clinical data, the time required for manual annotations and variability across experts. Simulated data could serve for data augmentation provided that its appearance was comparable to the actual appearance of intra-operative acquisitions. Generative Adversarial Networks (GANs) are a powerful tool for artistic style transfer, but lack a criteria for selecting epochs ensuring also preservation of intra-operative content.
We propose a multi-objective optimization strategy for a selection of cycleGAN epochs ensuring a mapping between virtual images and the intra-operative domain preserving anatomical content. Our approach has been applied to simulate intra-operative bronchoscopic videos and chest CT scans from virtual sketches generated using simple graphical primitives.

**Keywords:** data augmentation, cycleGANs, multi-objective optimization

## 1 Introduction

Medical imaging applications are challenging for machine learning methods due to the difficulty to generate reliable ground truth and the limited size of annotated databases. Data augmentation [1] has become a standard procedure to improve the training process. Data augmentation schemes increase the number of training samples by simple transformations (like translation, rotation, flip and scale) of the original dataset images. However, the diversity that can be gained from such modifications of the images is relatively small and introduces correlations in training data.

Virtual images obtained from computer simulations could be used to train classifiers and validate image processing methods if their appearance was comparable (in texture and color) to the actual appearance of intra-operative data. We consider that style transfer could be used to endow virtual simulated data with the content and texture of intra-operative scans using modern techniques for artistic style transfer.

Recent works [2, 3] have shown the power of Generative Adversarial Networks (GANs) for artistic style transfer. These methods generate stylized images comparing image feature representations extracted from pre-computed convolutional neural networks. A main challenge is the generation of pairs with similar content ensuring preservation of anatomical features. To avoid pairings, new approaches like [3] use GANs to transform images from one domain $A$ (like virtual simulations) into a domain $B$ (like interventional videos). The novelty of [3] is that a cyclic term is added in order to make the domain transfer bijective ($A \to B \to A$ and $B \to A \to B$). Although this endows maps with some content preservation, a main challenge [4] is the selection of the epochs most suitable for a given problem. Due to the oscillating behavior of adversarial training, most of the cases this is done manually.

Recently, several works using GANs for medical data augmentation have been proposed [5, 6, 7, 8]. These works lack for a criteria to chose the best trained epoch depending on the target task. As it is stated in [8], rigorous performance evaluation of GANs is an important research area, since is not clear how to quantitatively evaluate generative models[9]. In particular, when using statistical methods, like [3], aside probability distributions among pixels, their geometric relations related to anatomical content should also be taken into account. Such geometrical content should be considered in the definition of virtual data and it might not be easy to model in case of complex multi-organ scans. The most usual approach is to use anatomies already segmented, which might imply access to large amount of annotated data for generating new cases.

In this work, we propose a multi-objective optimization approach based on the Pareto front [10] to select the cycleGAN epoch achieving the best compromise between content preservation and style transfer. As a proof-of-concept, our multi-objective cycleGAN has been applied to simulate bronchoscopic videos and chest CT scans from synthetic models of the anatomy sketched using graphical primitives. As far as we know this is the first work generating several anatomies from simple geometrical models.

## 2 Multi-Objective cycleGAN

Given two domains $Virtual$, $V$, and $Real$, $R$, style transfer learns two (bijective) maps ($G_r$, $G_v$) from one domain onto the other one:

$$G_r : Virtual \to Real \quad G_v : Real \to Virtual \tag{1}$$

with the map composition $G_r(G_v)$ and $G_v(G_r)$ being the identity on each domain. Following [3], maps are given by auto-encoders trained to optimize:

$$\boldsymbol{\ell}(G_r, G_v, D_r, D_v) = \boldsymbol{\ell}_{\mathbf{GAN}}(G_v, G_r, D_r, D_v, V, R) + \lambda \boldsymbol{\ell}_{\mathbf{cyc}}(G_v, G_r) \tag{2}$$

The term $\boldsymbol{\ell}_{\mathbf{GAN}}$ measures how good are $G_v, G_r$ transferring images from one domain to the other one, while $\boldsymbol{\ell}_{\mathbf{cyc}}$ is a "cycle consistency loss" introduced to force bijective mappings.

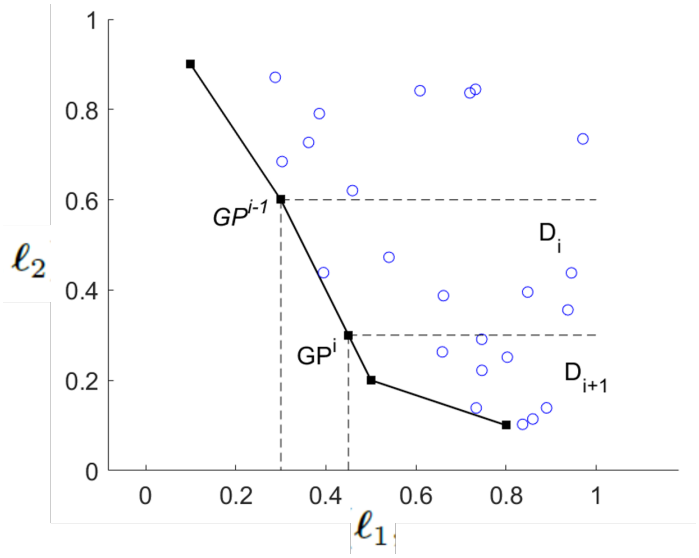


Fig. 1: Pareto front of cycle-GAN 2-objective optimization.

The minimization problem is solved by adversarial training as:

$$G_r^*, G_v^* = \min_{G_r, G_v} \left( \max_{D_r, D_v} \boldsymbol{\ell}(G_r, G_v, D_r, D_v) \right) \tag{3}$$

In this manner, $G_r^*$ and $G_v^*$ are optimized so that $G_r$, $G_v$ minimize (2) while the adversarial $D_r$, $D_v$ maximize it. These conditions (minimize while maximizing at the same time) might be in conflict considered into a single optimization process and might hinder the convergence of the training back-propagation.

We propose to consider separately the optimization of each of the terms in $\boldsymbol{\ell}$ and pose adversarial training as a multi-objective optimization [10] problem:

$$\begin{aligned} G_r^*, G_v^* = \min_{G_r, G_v} (\boldsymbol{\ell}_1, \boldsymbol{\ell}_2) = \min_{G_r, G_v} (\boldsymbol{\ell}_{\mathbf{cyc}}, \boldsymbol{\ell}_{\mathbf{GAN}}) = \\ \min_{G_r, G_v} (\boldsymbol{\ell}_{\mathbf{cyc}}(G_r, G_v), \max_{D_r, D_v} \boldsymbol{\ell}_{\mathbf{GAN}}(G_v, G_r, D_r, D_v)) \end{aligned} \tag{4}$$

The solution to (4) is computed from the Pareto front [10] consisting in the set of dominating configurations that outperform in any of the objectives without degrading at least one of the other ones. For this two-objective problem we compute an approximation to the Pareto front from back-propagation epochs as follows.

Let $\boldsymbol{G}^k := (G_r^k, G_v^k)$ be the transformation maps at the $k$-th epoch and $\boldsymbol{GP} = (\boldsymbol{GP}^i)_{i=1}^{NP}$ be the set of epochs belonging to the Pareto front. Such Pareto maps can be iteratively computed from the values of the objective functions as:

$$\boldsymbol{GP}^i := \min_{\boldsymbol{G}^i \in \boldsymbol{D}_i} (\boldsymbol{\ell}_{\mathbf{cyc}}(\boldsymbol{G}^i)) \tag{5}$$

for $\boldsymbol{D}_1$ the set of maps for all epochs and $\boldsymbol{D}_i$, $i > 2$, the set of maps dominating $\boldsymbol{G}P^{i-1}$. In our case, $\boldsymbol{G}^i \in \boldsymbol{D}_i$ if it satisfies the following conditions:

1. $\boldsymbol{\ell}_{\mathbf{cyc}}(\boldsymbol{G}^i) > \boldsymbol{\ell}_{\mathbf{cyc}}(\boldsymbol{G}P^{i-1})$ 2. $\boldsymbol{\ell}_{\mathbf{GAN}}(\boldsymbol{G}^i) < \boldsymbol{\ell}_{\mathbf{GAN}}(\boldsymbol{G}P^{i-1})$

Fig.1 shows the Pareto front associated to the two-objective problem given by (4) with the dashed lines enclosing the region that includes the set of epochs that dominate a given $\boldsymbol{G}P^{i-1}$. The pairs of function values $(\boldsymbol{\ell}_{\mathbf{cyc}}(\boldsymbol{G}P^i), \boldsymbol{\ell}_{\mathbf{GAN}}(\boldsymbol{G}P^i))$, $\boldsymbol{G}P^i \in \boldsymbol{G}P$, given by the two objectives evaluated at the Pareto front are shown in solid squares joined with a solid line.

We note that, by construction, our approximation to the Pareto front contains the epoch achieving the minimum of (2). Due to adversarial training, this epoch might be an early one prone to have high $\boldsymbol{\ell}_{\mathbf{GAN}}$. In order to select Pareto epochs producing style images with higher real appearance, a clustering or threshold on Pareto epochs $\boldsymbol{\ell}_{\mathbf{GAN}}$ values could be considered as filtering post-processing.

## 3 Experiments

The proposed multi-objective cycleGAN has been applied to simulate intra-operative virtual bronchoscopies and chest CT scans. For each modality, a different cycleGAN was trained using intra-operative data and tested in virtual sketches of the anatomy observed in each modality. Each CycleGAN was trained for 200 epochs.

### 3.1 Intra-operative virtual bronchoscopies

For this experiment we trained a cycleGAN on a set of 5 ultrathin intra-operative recordings (defining the set of images of the *Real* domain, $R$) and 5 virtual bronchoscopies (defining the set of images of the $Virtual$ domain, $V$). Intra-operative videos were acquired during biopsy interventions at Hospital Bellvitge (Barcelona, Spain) using an Olympus Exera III HD Ultrathin videobronchoscope. Virtual bronchoscopies were generated using an own developed software from CT scans acquired for different patients with an Aquilion ONE (Toshiba Medical Systems, Otawara, Japan) using slice thickness and interval of 0.5 and 0.4 mm respectively. For each intra-operative video or virtual simulation, we randomly selected 500 images for training cycleGAN from scratch.

In order to validate our proposal, we generated 10 sequences using sketches of airways in bronchoscopic frames. Sketches of airways were produced using elliptical shapes of different sizes and eccentricity and considering several configurations in number and disposition in the image. For each configuration, a sequence of affine transformations was applied to simulate bronchoscopic navigation.

Figure 2 plots the two cost functions for all epochs with the 35 Pareto epochs in black crosses. The epochs selected for these experiment are shown in squares, green ones for epochs on the Pareto front, red ones, otherwise. We have selected 2

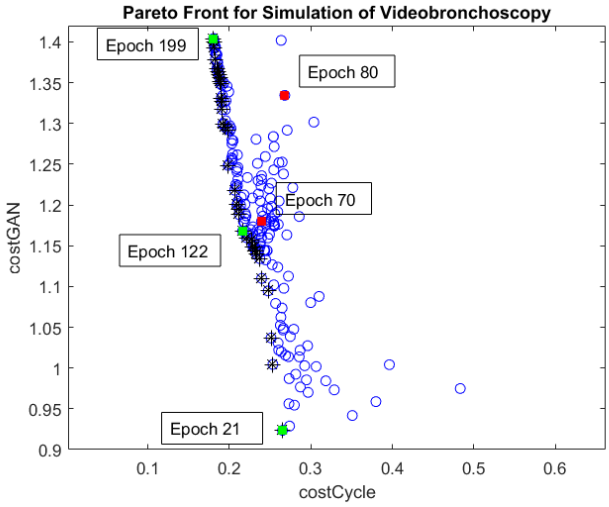


Fig. 2: Pareto front for Intra-operative virtual bronchoscopies.

epochs not belonging to the Pareto front, one (epoch 80) far from it, another one (epoch 70) close to it. The epochs selected on the Pareto front are the boundary ones (199, 21, which is the one achieving the minimum value of (2)) and the epoch (122) closest to epoch 70.

Figure 3 shows a mosaic of representative sketches and their augmentation using the 5 epochs. For each sketch, we show consecutive frames to check the stability of the augmented images. Images obtained from Pareto epochs have a green frame, red otherwise. As it can be seen, the appearance of images generated with Pareto front epochs are more stable and keep better texture. Also brightness, color and the shape of holes are more realistic. Epoch 21 artifacts on image boundary and lack of texture can be attributed to an early stage of cycleGAN training. It is worth noticing that images generated with epochs not belonging to the Pareto front (like epoch 80) present an unstable anatomy with new holes that were not in the sketches.

### 3.2 Chest CT scans

For this experiment we trained a cycleGAN on a set of 11 chest CT volumes (defining the *Real* domain) and their segmentation (defining the $Virtual$ domain) of the main lung structures. CT scans were acquired in inspiration with an Aquilion ONE (Toshiba Medical Systems, Otawara, Japan) using slice thickness and interval of 0.5 and 0.4 mm respectively. Segmentations included the body, lungs, pulmonary vessels and airways and were computed using an own-designed software. Each volume (original CT scan and its segmentation) was uniformly sampled in 250 short axis planes covering the whole volume of lungs for training cycleGAN from scratch.

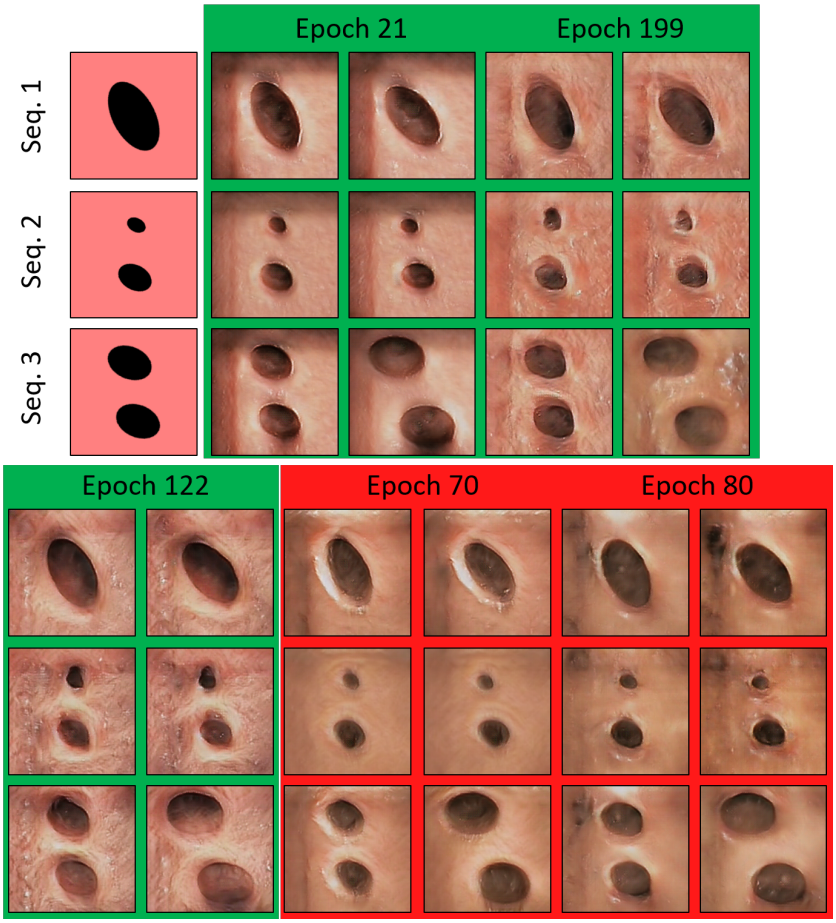


Fig. 3: Generated intra-operative virtual bronchoscopies for 3 representative sketch sequences: images from Pareto front epochs in green frame and images from non-Pareto epochs in red one.

A total number of 10 anatomical sketches were generated using 3D graphical primitives for each lung structure. Body was modelled with a cylinder, lungs with ellipsoids and, both, vessels and airways were modelled as tubular structures with several branching levels. The thickness of the tubes was set depending on the level to account for different lumen sizes. To endow sketches with more realistic geometry, the cardiac and abdominal cavities were also modelled using ellipsoids. The ranges for primitives sizes were statistically learned from segmented CTs. For each sketched anatomy 250 short axis planes were sampled for testing.

In this case, the Pareto front had 7 epochs. As in the previous experiment, we selected the Pareto boundary epochs (196 and the least cost 29), two epochs not belonging to the Pareto front (epoch 92 far from it and epoch 53 close to it)

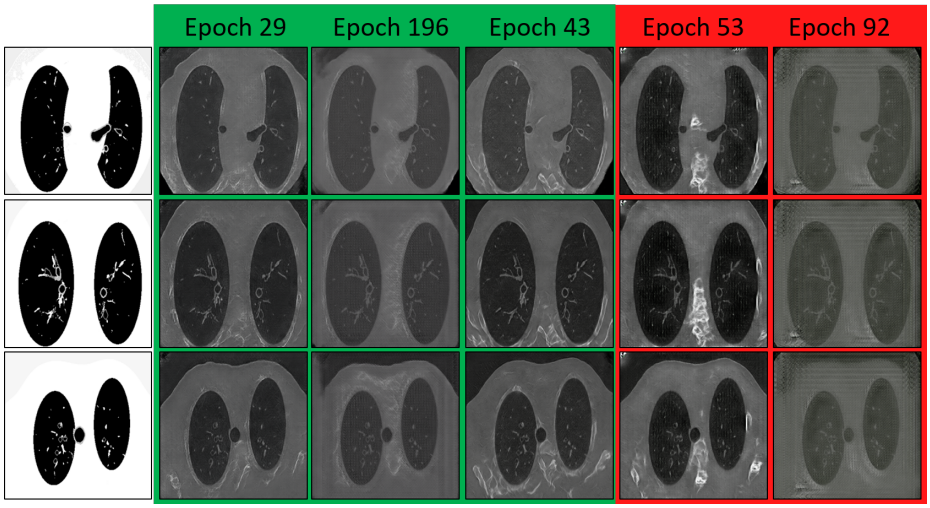


Fig. 4: Generated CT virtual scans for 3 representative (distal, mid, basal) sketches: images from Pareto front epochs in green frame and images from non-Pareto epochs in red one.

and the Pareto epoch (43) closest to epoch 53. Figure 4 shows a mosaic of representative sketches and their augmentation using these 5 epochs. Image frame colors are like in fig.3. As before, images generated with epochs not belonging to the Pareto front present different artifacts, like a stripe pattern (epoch 92), lack of texture in lungs (epoch 53) and bright structures in the body (epoch 53). Meanwhile, images generated with Pareto epochs keep a stable appearance similar to the one of a chest CT scan.

## 4 Conclusions

In this paper, we propose a generative algorithm based on cycleGANs to produce synthetic medical data from simple graphical primitives roughly approximating the anatomical structure of organs. In particular, we present a novel multi-objective analysis of cycleGAN training epochs for the selection of epochs achieving the best compromise between generation and discrimination.

Models have been tested in synthetic videobronchoscopy sequences and chest CT scans computed using simple graphical primitives. Visual inspection of this proof-of-concept indicates that networks might admit a training on limited amount of data thanks to a numerical analysis of the back-propagation scheme. Results also indicate that the epochs selected by our multi-objective approach have a good generalization power in testing on data with geometrical structure different from the one of training data.

These results encourage further investigation of the performance of using sketches (easy to generate without large gathering of medical data) for training

networks and augment sets using more realistic anatomies computed from few medical scans.

**Acknowledgments**

The research leading to these results has received funding from the European Union?s Horizon 2020 research and innovation programme under the Marie Sk?odowska-Curie grant agreement No 712949 (TECNIOspring PLUS) and from the Agency for Business Competitiveness of the Government of Catalonia. Work supported by Spanish Projects FIS-G64384969, Generalitat de Catalunya, 2017-SGR-1624 and CERCA-Programme. D.Gil is a Serra Hunter fellow. The Titan X Pascal used for this research was donated by the NVIDIA Corporation.

## References

[1] Alex Krizhevsky, Ilya Sutskever, and Geoffrey E Hinton. Imagenet classification with deep convolutional neural networks. In *Advances in neural information processing systems*, pages 1097–1105, 2012.

[2] Justin Johnson, Alexandre Alahi, and Li Fei-Fei. Perceptual losses for real-time style transfer and super-resolution. In *European Conference on Computer Vision*, pages 694–711. Springer, 2016.

[3] Jun-Yan Zhu, Taesung Park, Phillip Isola, and Alexei A Efros. Unpaired image-to-image translation using cycle-consistent adversarial networks. *arXiv preprint arXiv:1703.10593*, 2017.

[4] Alec Radford, Luke Metz, and Soumith Chintala. Unsupervised representation learning with deep convolutional generative adversarial networks. *arXiv preprint arXiv:1511.06434*, 2015.

[5] Maayan Frid-Adar, Eyal Klang, Michal Amitai, Jacob Goldberger, and Hayit Greenspan. Synthetic data augmentation using gan for improved liver lesion classification. In *2018 IEEE 15th International Symposium on Biomedical Imaging (ISBI 2018)*, pages 289–293. IEEE, 2018.

[6] Hoo-Chang Shin, Neil A Tenenholtz, Jameson K Rogers, Christopher G Schwarz, Matthew L Senjem, Jeffrey L Gunter, Katherine P Andriole, and Mark Michalski. Medical image synthesis for data augmentation and anonymization using generative adversarial networks. In *International Workshop on Simulation and Synthesis in Medical Imaging*, pages 1–11. Springer, 2018.

[7] Maayan Frid-Adar, Idit Diamant, Eyal Klang, Michal Amitai, Jacob Goldberger, and Hayit Greenspan. Gan-based synthetic medical image augmentation for increased cnn performance in liver lesion classification. *Neurocomputing*, 321:321–331, 2018.

[8] Francesco Calimeri, Aldo Marzullo, Claudio Stamile, and Giorgio Terracina. Biomedical data augmentation using generative adversarial neural networks. In *International conference on artificial neural networks*, pages 626–634. Springer, 2017.

[9] Ian Goodfellow, Yoshua Bengio, and Aaron Courville. *Deep learning*. MIT press, 2016.

[10] Kaisa Miettinen. *Nonlinear multiobjective optimization*, volume 12. Springer Science & Business Media, 2012.