Enhancing virtual bronchoscopy with intra-operative data using a multi-objective GAN

D.Gil[1],A.Esteban-Lansaque[1], A.Borras[1] and C.Sánchez[1]

1. Purpose

State of the art methods in computer vision 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. The particular field of intervention guiding has the extra difficulty of intra-operative recordings probably requiring the alteration of standard protocols.

Virtual endoscopy could be used to train classifiers and validate image processing methods if its appearance was comparable (in texture and color) to the actual appearance of intra-operative recordings. Modern techniques for artistic style transfer could be used to endow virtual endoscopic images with the content and texture of intra-operative videos provided that stylized images preserved the anatomical structure of, both, intra-operative images and simulated data.

This work addresses the generation of realistic endoscopic images using intra-operative data to augment the appearance of virtual endoscopy. In particular, we use cycleGAN [1] in a multi-objective optimization scheme to map virtual images into the intra-operative domain preserving their anatomical content.

2. Methods

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 [1], 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.

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.

We propose to consider separately the optimization of each of the terms in $\boldsymbol{\ell}$ and pose adversarial training as the following multi-objective optimization [2] problem:

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

The solution to (4) is computed from the Pareto front [2] consisting in the set of dominating configurations that outperform in any of the objectives without degrading at least one of the other ones.

The set of Pareto epochs achieve the best trade-off between the objective functions and, thus, are equivalent from the point of view of the cycleGAN. The epoch from the Pareto front best suited for augmentation of virtual endoscopic images is selected as the one that minimizes the L^2-difference:

$$\boldsymbol{\ell}_{\mathbf{Cont}} = \underset{v \in V}{mean} (||v - G_r(v)||_2) + \underset{r \in R}{mean} (||r - G_v(r)||_2) \tag{5}$$

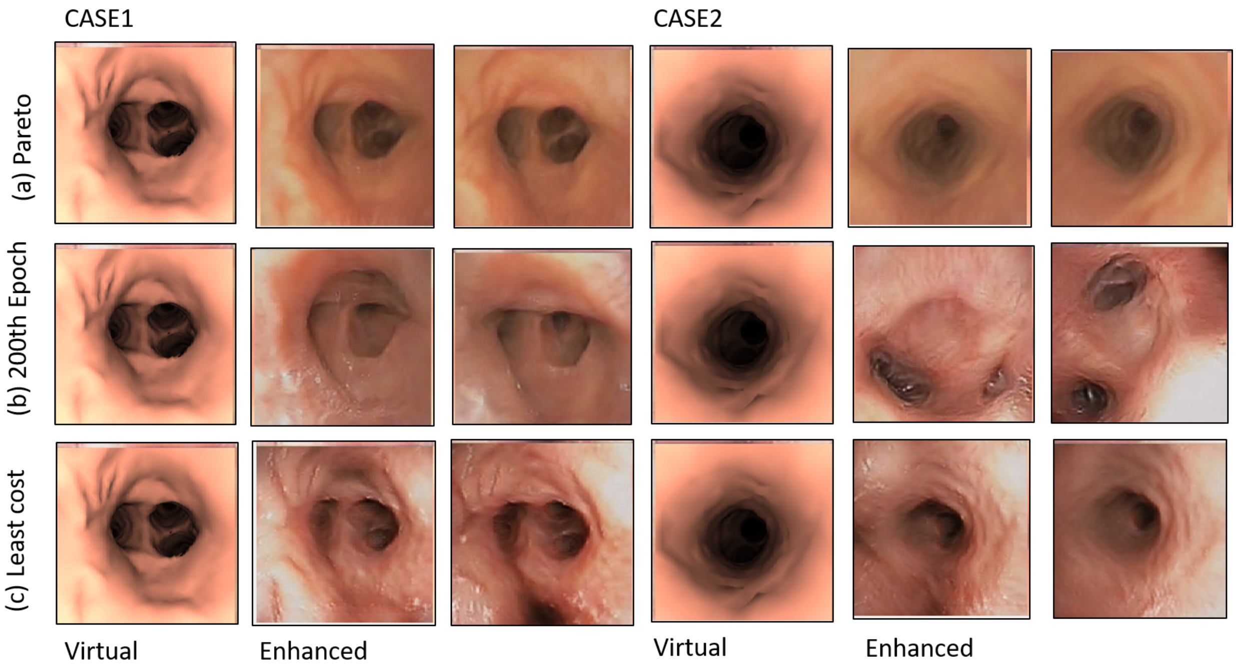

Figure 1: Virtual bronchoscopy enhanced using our method, (a), 200th epoch network, (b) and the least cost one, (c).

for $\underset{v\in V}{mean}$, $\underset{r\in R}{mean}$ denoting the average values on the training sets of virtual, V, and real, R, images.

3. Results

We trained our method 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.

After 200 epochs, our multi-objective approach selected epoch 50 as the one achieving the best compromise between intr-operative appearance and preservation of virtual anatomical content. To validate the visual quality of virtual bronchoscopies enhanced using this network, we used simulations generated from 4 additional CT scans acquired from a set of different patients. Simulations were also enhanced using the 200th epoch network and the network achieving the least value of the cost (2).

Figure 1 shows representative images of virtual bronchoscopies enhanced using our method (fig.1 (a)), the 200th epoch network (fig.1 (b)) and the least cost one (fig.1 (c)). For each case, we show two consecutive frames of the enhanced virtual sequence which should be very similar in appearance and content. Least cost images have sudden dark artifacts, while the 200th epoch yields highly unstable images that do not always match the original anatomy. The GAN trained using the proposed multi-objective strategy provides a stable appearance in images which are the most consistent with the original anatomical content of virtual images.

4. Conclusions

Visual inspection indicates that the proposed multi-objective approach to GANs improves the anatomical content of stylized virtual videos keeping an intra-operative appearance. Future quantitative analysis and observations will be conducted to analyze the performance of the multi-objective GAN in detail.

Acknowledgment

The research leading to these results has received funding from the European Unions 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.

References

[1] J.-Y. Zhu, T. Park, P. Isola, A. A. Efros, Unpaired image-to-image translation using cycle-consistent adversarial networks, arXiv preprint arXiv:1703.10593.

[2] K. Miettinen, Nonlinear multiobjective optimization, Vol. 12, Springer Science & Business Media, 2012.