Master 2 Internship - Université de Bordeaux

Image restoration with semantic priors

Context The internship will be realized within the ANR PostProdLEAP project, that gathers researchers in applied mathematics (IMB, MAP5) and Computer Science (LaBRI). The project aims at providing tools for archive post-production, dedicated to image and video restoration, super-resolution and colorization. The project includes interactions with the artists and historians from the company Composite Films, world leader in restoration and colorization of films and archive footage. Through this collaboration, the objective is to give the artists control on the final results provided by the developed methods, while making their tasks less time consuming and tedious than with current professional software.

When:	5 to 6 months Internship starting between January and April 2020
Where:	Institut de Mathématiques de Bordeaux, Talence, France
Salary:	≈540 €/month
Expected skills:	Computer science or applied mathematics Image processing and analysis, machine learning, computer vision, Matlab, python.
Perspective:	A PhD grant is proposed at the end of the internship
Application:	Send by email, **before December 31th 2019**, a CV and a statement of interest to nicolas.papadakis@math.u-bordeaux.fr and andres.almansa@parisdescartes.fr

Objectives and methodology The objective of the internship and is to propose new editing tools inspired by recent patch-based and deep learning approaches. The project focuses on archival film restoration. In contrast to other image restoration areas, an explicit degradation model is not available in general for old degraded films. Therefore we need to rely more heavily on learning-based methods and on semantic information that can be obtained from the degraded film such as textured background, buildings, objects, animals, dress, human faces, and on historical documents that can link those semantic categories to the desired restored appearance. The methods will be trained, tested and validated on datasets of video archives with different resolutions and degrees of quality, created specially during the project from movies restored by Composite Films. Meanwhile, standard annotated databases (faces with CelebA or natural images with Imagenet) will be considered.

In the literature, related methods rely on end-to-end architectures [1] for joint restoration and colorization. This kind of approach gives no possibility for the user to monitor the output. During the project, the candidate will rather study generative models such as Variational Auto-Encoders (VAEs) [6, 4], Generative Adveserial Networks (GANs) [5] or their Wasserstein generalizations (WGANs) [2]. These methods are designed to generate new samples from a given discrete dataset. To that end, generative models learn a mapping from a (low dimensional) continuous latent space to the data space, in order to approximate the data distribution. Then, tuning the latent variable offers an efficient way to control the output of the model. This is essential to develop graphical user interfaces including track bars that will be used by artists. To propose adaptive and reliable generative models, several limitations have to be addressed.

Control of the latent space. Having a structured and disentangled [10] latent space allows for controlling explicitly some characteristics of the generated images. Incorporating semantic information in latent spaces [9] is a main objective to propose tunable models that can be used by restorers and colorists.

Resizeable generative models. In order to restore or generate images of variable size, it is necessary to define "resizeable" generative networks (*i.e.* that can be trained and tested on images of different sizes). While end-to-end models with convolutional architectures can do so, neural-networks-based generative models are a priori not resizeable since they are based on latent spaces of fixed dimension. Optimal-transport and patch-based modeling [7] has the potential to circumvent this issue. Posterior patch aggregation can nevertheless destroy the patch distribution of the produced image and is not satisfying up to now. These solutions might be combined with recent advances in specially tailored latent-spaces that allow them to be resized [3, 12, 11].

Quality of generative models. Statistical tests [8] have shown that VAEs have higher recall performances whereas GANs architectures reach better precision. Such analysis will be taken into account to provide models with theoretical guarantees, and training methods that combine the benefits of both GANs and VAEs.

References

[1] J. Antic. Deoldify project, 2018. `https://github.com/jantic/DeOldify`.

[2] M. Arjovsky, S. Chintala, and L. Bottou. Wasserstein GAN, 2017.

[3] U. Bergmann, N. Jetchev, and R. Vollgraf. Learning texture manifolds with the periodic spatial GAN. In *(ICML) International Conference on Machine Learning*, volume 1, pages 722–730, 2017.

[4] B. Dai and D. Wipf. Diagnosing and Enhancing VAE Models. *ICLR*, pages 1–42, 2019.

[5] I. J. Goodfellow, J. Pouget-Abadie, M. Mirza, B. Xu, D. Warde-Farley, S. Ozair, A. Courville, and Y. Bengio. Generative Adversarial Networks. *(NeurIPS) Advances in Neural Information Processing Systems 27*, pages 2672–2680, 2014.

[6] D. P. Kingma and M. Welling. Auto-Encoding Variational Bayes. In *(ICLR) International Conference on Learning Representations*, number Ml, pages 1–14, 2013.

[7] A. Leclaire and J. Rabin. A Fast Multi-layer Approximation to Semi-discrete Optimal Transport. In *(SSVM) International Conference on Scale-Space and Variational Methods in Computer Vision*, volume 11603 LNCS, pages 341–353, 2019. HAL preprint.

[8] M. S. M. Sajjadi, O. Bachem, M. Lucic, O. Bousquet, and S. Gelly. Assessing Generative Models via Precision and Recall. In *(NeurIPS) Neural Information Processing Systems*, 2018.

[9] Y. Shen, J. Gu, X. Tang, and B. Zhou. Interpreting the Latent Space of GANs for Semantic Face Editing, 2019.

[10] Z. Shu, M. Sahasrabudhe, R. Alp Güler, D. Samaras, N. Paragios, I. Kokkinos, A. Guler, D. Samaras, N. Paragios, and I. Kokkinos. Deforming autoencoders: Unsupervised disentangling of shape and appearance. In *(ECCV) European Conference on Computer Vision*, volume 11214 LNCS, pages 664–680, 2018.

[11] N. Yu, C. Barnes, E. Shechtman, S. Amirghodsi, and M. Lukac. Texture Mixer: A Network for Controllable Synthesis and Interpolation of Texture. In *(CVPR) Conference on Computer Vision and Pattern Recognition*, 2019.

[12] Y. Zhou, Z. Zhu, X. Bai, D. Lischinski, D. Cohen-Or, and H. Huang. Non-stationary texture synthesis by adversarial expansion. *ACM Transactions on Graphics*, 37(4):1–13, 2018.