DD2424: Project Proposal → GANcedonia

Miquel Corominas Larsson miquell@kth.se
Marcel Catà Villà marcelcv@kth.se
Borja Rodríguez Gàlvez borjarg@kth.se

April 13, 2018

1 Project description

In this project, we propose doing object transfiguration applied to fruit images. For this purpose, we are going to use a variant of Generative Adversarial Networks (GANs) [1] called Cycle GANs [2].

The task of object transfiguration consists on translating the objects from one domain, X, to another, Y. Therefore, we want to train a function $F : X \to Y$ such that the transformation of an object from the first domain, $\hat{y} = F(x),\ x \in X$; is indistinguishable from an object on the second domain $y \in Y$.

Standard GANs [1] try to generate F in a stochastic fashion; and learn the best fit trying to minimize an adversarial loss function, $\mathcal{L}^F_{adv}$. The novelty provided by CycleGANs [2] is the fact they are *cycle consistent*; which means they make sure forward, $F : X \to Y$, and backward translations, $B : Y \to X$, are the inverse from one another and that both of them are bijections. They take care of that (i.e., $F(B(x)) \approx x$ and $B(F(y)) \approx y$) by introducing a cycle consistency loss, $\mathcal{L}_{con}$, along with the forward and backward adversarial losses, $\mathcal{L}^F_{adv}$ and $\mathcal{L}^B_{adv}$.

Hence, the task will be to minimize a loss function based on the above mentioned losses:

$$F^*, B^* = \arg\min_{F,G} \max_{D_x,D_y} (\mathcal{L}_{tot}) \tag{1}$$

where D_x and D_y are the adversarial discriminators and the total loss is defined as:

$$\mathcal{L}_{tot} = \mathcal{L}^F_{adv} + \mathcal{L}^G_{adv} + \lambda \mathcal{L}_{con} \tag{2}$$

with λ controlling the importance of the cyclic consistence.

We will therefore analyze the importance of the cyclic consistence and the adversarial losses and observe if other objective functions (e.g., [3]) lead to better results.

2 Data

The data source will be the Fruits 360 dataset [4]. This dataset consists of 38409 labeled fruit images of 60 different classes, from which we will only use a small selection of fruit types. The images have a normalized size of 100x100 pixels, and have been pre-processed (background has been removed). Moreover, if necessary, complementary images will be added to the data from the CIFAR-100 [5] and FIDS30 [6] datasets or/and some data augmentation techniques will be used.

3 Specifications

This project will be carried out in Python. Specifically, we will use PyTorch [7] as the neural net framework to do the implementation. Nevertheless, implementation of other frameworks will be studied along side, such as TensorFlow [8].

References

[1] I. Goodfellow, "Nips 2016 tutorial: Generative adversarial networks," *arXiv preprint arXiv:1701.00160*, 2016.

[2] J.-Y. Zhu, T. Park, P. Isola, and A. A. Efros, "Unpaired image-to-image translation using cycle-consistent adversarial networkss," in *Computer Vision (ICCV), 2017 IEEE International Conference on*, 2017.

[3] Z. Yi, H. Zhang, P. Tan, and M. Gong, "Dualgan: Unsupervised dual learning for image-to-image translation," *arXiv preprint*, 2017.

[4] H. Mureşan and M. Oltean, "Fruit recognition from images using deep learning," *arXiv preprint arXiv:1712.00580*, 2017.

[5] A. Krizhevsky and G. Hinton, "Learning multiple layers of features from tiny images," 2009.

[6] Škrjanec Marko, "Automatic fruit recognition using computer vision." (Mentor: Matej Kristan), Fakulteta za računalništvo in informatiko, Univerza v Ljubljani, 2013.

[7] A. Paszke, S. Gross, S. Chintala, G. Chanan, E. Yang, Z. DeVito, Z. Lin, A. Desmaison, L. Antiga, and A. Lerer, "Automatic differentiation in pytorch," in *NIPS-W*, 2017.

[8] M. Abadi, P. Barham, J. Chen, Z. Chen, A. Davis, J. Dean, M. Devin, S. Ghemawat, G. Irving, M. Isard, *et al.*, "Tensorflow: A system for large-scale machine learning.," in *OSDI*, vol. 16, pp. 265–283, 2016.