

Synopsis : Image super-resolution using hierarchical VAEs

Students : Paolo Zifferero, s200149

Prasad Jagtap, s200109

Paul Connetable, s151404

1 Motivation and background

Increasing an image resolution or reconstructing a high-resolution image from a coarse resolution image is a very active area of research. This is called image *super-resolution*, and can be used to improve the image quality of movies or pictures, as well as store them in coarser resolution, and therefore use little space. This challenge was carried out at the beginning by *bicubic interpolation* or *sparse-coding-based* methods. Quickly, Deep Learning models proved their efficiency at solving this ill-posed problem. At first, CNNs were proposed to tackle this problem [1], by letting them learn the features of the image, and reconstruct them in higher resolution. Later, the use of VAEs was proposed to improve quality of the reconstructed images [2]. GANs have also been used to improve the perceptual quality of the images, to make them look more realistic [3]. The latest state of the art *super-resolution neural network* [4] is a *deep hierarchical VAE*. The more recent models like the first ones cited above have proven to be much more adequate to the task, and have also been able to generate new images, separate an image's content & style, and use them separately as in google's deep-dream.

2 Milestones

Our first step will be to choose an image data set and prepare it for super-resolution by applying a Gaussian filter and subsampling images. We are leaning towards using the CIFAR data-set. Next, we will create a CNN similar to the one described by *Dong et al.*[1], and test it on a the selected data-set. Later, we will adapt the already existing state-of-the-art BIVA model to our requirements, and perform image SR on samples taken from the data-set.

References

- [1] C. Dong, C. C. Loy, K. He, and X. Tang, "Image super-resolution using deep convolutional networks," *IEEE transactions on pattern analysis and machine intelligence*, vol. 38, no. 2, pp. 295–307, 2015.
- [2] C. K. Sønderby, T. Raiko, L. Maaløe, S. K. Sønderby, and O. Winther, "Ladder variational autoencoders," in *Advances in neural information processing systems*, pp. 3738–3746, 2016.

- [3] C. Ledig, L. Theis, F. Huszár, J. Caballero, A. Cunningham, A. Acosta, A. Aitken, A. Tejani, J. Totz, Z. Wang, *et al.*, “Photo-realistic single image super-resolution using a generative adversarial network,” in *Proceedings of the IEEE conference on computer vision and pattern recognition*, pp. 4681–4690, 2017.
- [4] A. Vahdat and J. Kautz, “Nvae: A deep hierarchical variational autoencoder,” *arXiv preprint arXiv:2007.03898*, 2020.