Bachelor thesis

Colorization of panchromatic space images with Generative Adversarial Network

Patrycja Cieplicka
Automation and robotics
Academic year 2019/2020

Supervisor: dr inż. Andrzej Kordecki

1. Introduction

Process of colorization could be described as assigning colors to the gray-scaled images in order to get the reality seen with human eyes. In this thesis, a method of colorization of panchromatic space images, which could be used in such tasks as image segmentation, detection of roads or buildings, as well as tracking of objects, was presented. To colorize images one of the state of art artificial algorithm - generative adversarial network - was used. An important element of the work was to adjust proper parameters of training. To this end, the effect of label smoothing, the learning rate value, the number of epochs have been checked, and also the network architecture has been slightly modified. The solutions used in the work were based on two spaces of color representation: RGB and CIELab.

2. Generative Adversarial Network

Generative Adversarial Network (GAN) consists of two neural networks, which are competing with each other. In colorization model variation of GANs, Conditional GAN, where both generator and discriminator are conditioned on some data (panchromatic images), where used. The discriminator gets colored images from both generator and original data along with the gray-scaled input as the condition and tries to decide which pair contains the true colored image. On the other hand, generator, wants to produce colored images, which are the most similar to the reality, from panchromatic input - its goal is to fool the Discriminator. This two networks are trained together in min-max game. Generator architecture is based on U-net, encoder-decoder network with skip connections and discriminator is convolutional network.

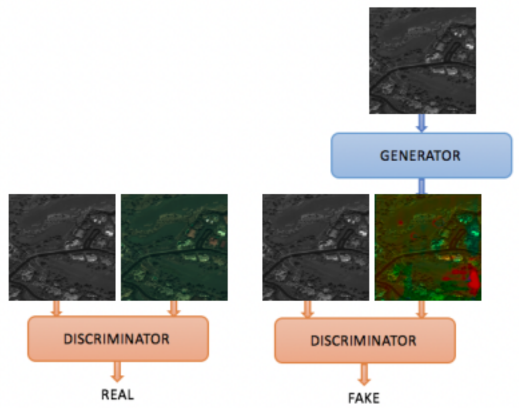

Figure: Conditional GANs

Final loss functions, which describe mathematically how GAN works, for discriminator (1) and generator (2) are presented below:

$$\max_{\theta_D} J^{(D)}(\theta_D, \theta_G) = \max_{\theta_D}(\mathbb{E}_z[\log(D(x))] + \mathbb{E}_z[\log(1 - D(G(z)))] \quad (1)$$

$$\max_{\theta_G} J^{(G)}(\theta_D, \theta_G) = \max_{\theta_G} \mathbb{E}_z[\log(D(G(z)))] \quad (2)$$

Colorization models

As part of the work, two different colorization models were created depending on the training pictures used. In first model generator predicts two layers *a* and *b* in Lab space from panchromatic image. Second model predicts three layers *L*, *a* and *b* in Lab color space or *R*, *G* and *B* in RGB color space.

3. Results

Below are presented three the best colorization for first model (predicts *a* and *b* layers). To measure the performance of the colorization Mean Squarred Error (MSE), Peak Signal-to-noise Ratio (PSNR) and color difference in Lab space ΔE_{ab} were calculated.

Table: Colorization results for three the best models

Label (false/true)	Epochs	MSE	PSNR	ΔE_{ab}
0; 0.9	2750	19.69	35.18	2.69
0; 0.95	3000	18.4	35.48	2.74
0; 0.95	1250	18.24	35.52	2.84

Figure: Colorization comparison. From left: first model (label 0.9, 2750 epochs), second model (label 0.95, 3000 epochs), third model (transposed convolution, label 0.95, 1250 epochs), original photos

4. Conclusions

- Colorization with Generative Adversarial Network produces satisfactory results, which can be used in such tasks as image segmentation, detection and tracking objects.
- The GAN learning process is unstable and changes rapidly depending on number of epochs, so more frequent saving of models would bring better results. Label smoothing and higher generator's learning rate than discriminator's have positive effect on results, however it is difficult to find appropriate parameters. Transposed convolution made GAN's training more stable.
- Proposed solution can be also used to colorize photos from Mars rovers and probes exploring other celestial bodies. Moreover, it can be used to colorize videos from cameras in robots.