Near InfraRed Imagery Colorization

Patricia L. Suarez, Angel D. Sappa, Boris X. Vintimilla and Riad I. Hammoud

Escuela Superior Politécnica del Litoral, ESPOL
FIEC-CIDIS
Campus Gustavo Galindo, 09-01-5863, Guayaquil, Ecuador

Computer Vision Center
Autonomous University of Barcelona
08193, Bellaterra, Barcelona, Spain

BAE Systems FAST LabsBarcelona
Burlington, MA 01803, USA

1. Introduction

- In many vision applications, including surveillance photo interpretation by imagery analysts and driving scene understanding by drivers looking at backup-aid cameras, RGB video sensors are preferred since depicted images are similar to the human visual perception system
- RGB image have limitations related with lighting conditions and object surface color
- The limitations mentioned above can be easily overcome using Near Infrared (NIR) imagery
- NIR image colorization shares some particularities with color correction/transfer [1]

RGB image

NIR image

2. Motivation and Objective

- Image acquisition devices have largely expanded in recent years, mainly due to the decrease in price of electronics together with the increase in computational power.
- In spite of the large amount of possibilities, when the information needs to be provided to a final user, the classical RGB representation is preferred, because help user understanding.

RESEARCH QUESTION

How to represent in the RGB color space an infrared image?

3. Proposed Method

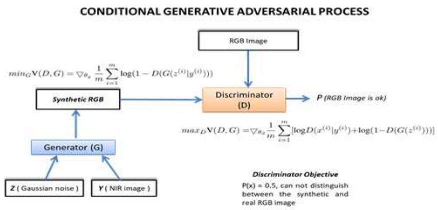
Generative Adversarial Networks

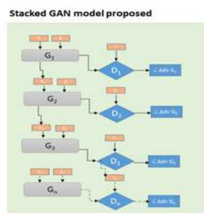
- Generative Adversarial Network (GAN) is a framework presented on [2], where two models are trained simultaneously
  - Generative model (G) that captures the data distribution
  - Discriminative model (D) that estimate the probability that a sample came from the training data rather than G
- Parameters for G and D are obtained from: $$\min_G \max_D V(D,G) = \mathbb{E}_{x\sim p\ \mathrm{data}(x)}[logD(x)] + \mathbb{E}_{z}\sim_{p\ \mathrm{data}(z)}[log(1-D(G(z)))]$$
- GANs can be extended to a conditional model if both the generator and discriminator are conditioned on some extra information: $$\min_G \max_D V(D,G) = \mathbb{E}_{x\sim p\ \mathrm{data}(x)}[logD(x|y)] + \mathbb{E}_{z}\sim_{p\ \mathrm{data}(z)}[log(1-D(G(z|y)))]$$
- The proposed approach is based on the usage of a Conditional Generative Adversarial Network (C-GAN) [3]

Stacked Conditional GAN model

Based on [5] a novel infrared imagery coloring estimation using Stacked C-GAN is proposed

- The model will receive as input a near infrared patch (NIR) fused with Gaussian noise to ensure more diversity of colors
- *l1* regularization term is added on each layer of the model to prevent overfitting





Multi-term loss function:

$$\mathcal{L}_{Final} = \mathcal{L}_{Adversarial} + \mathcal{L}_{Intensity} + \mathcal{L}_{SSIM}$$

- Adversarial loss (minimize the cross-entropy):

$$\mathcal{L}_{Adversarial} = -\sum_i logD(G_w(I_{z|y}),(I_{x|y})$$

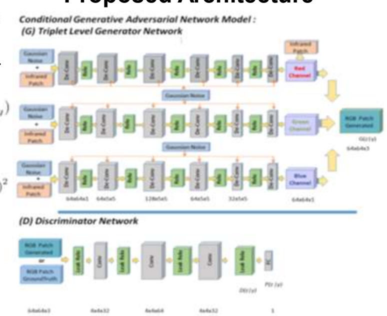
- Intensity loss:

$$\mathcal{L}_{Intensity} = \frac{1}{NM}\sum_{i=1}^{N}\sum_{j=1}^{M}(NDVIe_{i,j} - NDVIg_{i,j})^2$$

- Structural loss (computed from [4]):

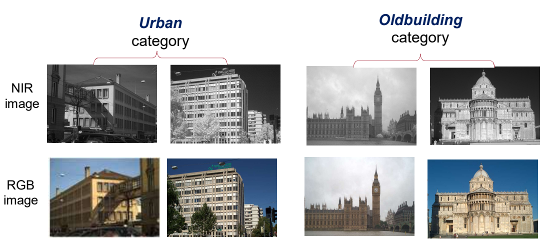
$$\mathcal{L}_{SSIM} = \frac{1}{NM}\sum_{p=1}^{P} 1 - SSIM(p)$$

Proposed Architecture

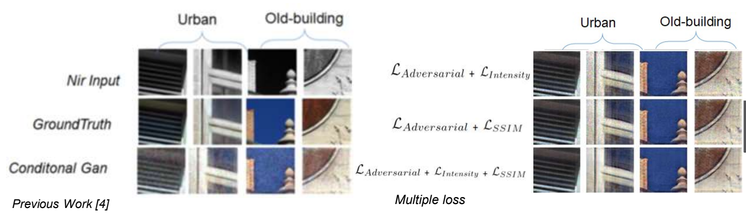


4. Experimental Results

The proposed approach has been evaluated with a large data set [5] (country and field categories have been used), some illustrative and quantitative results are provided below:



GAN with different loss functions:



ANGULAR ERRORS (AE), MEAN SQUARED ERRORS (MSE) AND STRUCTURAL SIMILARITIES (SSIM) OBTAINED WITH THE PROPOSED PROPOSED STACKED CONDITIONAL GAN ARCHITECTURE BY USING DIFFERENT LOSS FUNCTIONS (SSIM VALUES, THE BIGGER THE BETTER).

Training	AE		MSE		SSIM	
	Urban	*Old-Building*	*Urban*	*Old-Building*	*Urban*	*Old-Building*
Conditional GAN from [4]	5.77	5.96	18.91	18.25	0.84	0.86
Proposed Stacked Conditional GAN with $\mathcal{L}_{Adversarial} + \mathcal{L}_{Intensity}$	5.43	5.21	18.74	18.11	0.86	0.89
Proposed Stacked Conditional GAN with $\mathcal{L}_{Adversarial} + \mathcal{L}_{SSIM}$	5.32	4.97	18.53	18.02	0.90	0.91
Proposed Stacked Conditional GAN with $\mathcal{L}_{final}$	5.04	4.78	17.63	17.34	0.90	0.91

5. Conclusions

- A novel Stacked Conditional Generative Adversarial Network model has been proposed for NIR image colorization.
- Different loss functions have been evaluated to help the learning model to produce good results.
- Other strategies will be considered for improving results like to use Cycled consistent GAN's or Variational Autoencoder with adversarial learning process

6. References

1. M. Oliveira, A. D. Sappa, and V. Santos, "Unsupervised local color correction for coarsely registered images," in Computer Vision and Pattern Recognition (CVPR). IEEE, June 2011.
2. I. Goodfellow, J. Pouget-Abadie, M. Mirza, B. Xu, D. Warde, S. Ozair, A. Courville, and Y. Bengio, "Generative adversarial nets," in Advances in neural information processing systems, 2014.
3. M. Mirza and S. Osindero, "Conditional generative adversarial nets," preprint arXiv:141.178 2014.
4. P. L. Suarez, A. D. Sappa, and B. X. Vintimilla, "Colorizing infrared images through a triplet conditional dcgan architecture," in 19th International Conference on Image Analysis and Processing, 2017
5. X. Huang, Y. Li, O. Poursaeed, J. Hopcroft, and S. Belongie, "Stacked generative networks" preprint arXiv:1612.04357, 2016.