ICPR 2020: International Conference on Pattern Recognition

Thermal Image Enhancement using Generative Adversarial Network for Pedestrian Detection

Mohamed Amine Marnissi, Hajer Fradi, Anis Sahbani and Najoua Essoukri Ben Amara

Emails: medamine.marnissi@eniso.u-sousse.tn, hajer.fradi@issatso.rnu.tn, anis.sahbani@enovarobotics.com and najoua.benamara@eniso.rnu.tn

Introduction

Motivation:

- Infrared cameras often suffer from low-contrast, low-resolution and blurred details.
- These problems could limit the feasibility of different infrared imaging applications.

Objective: To enhance the visual quality of thermal images in order to improve the pedestrian detection performance.

State-of-the-art:

- Traditional methods based on Histogram Equalization (HE) and Contrast Limited Adaptive(CLAHE)
- Deep Learning methods (VDSR, SRCNN, TEN, CDN-MRF)
- Adversarial Learning methods(SRGAN, DCGAN)

Proposed Thermal Image Enhancement Architecture

- New Thermal Enhancement-GAN (TE-GAN) architecture basically inspired from EnlightenGAN [1] and DnCNN[2].
- The proposed architecture is composed of two modules with a post-processing step to cover different limitations of thermal images.
- Training TE-GAN architecture according to an overall loss function that combines perceptual, content, global and local losses

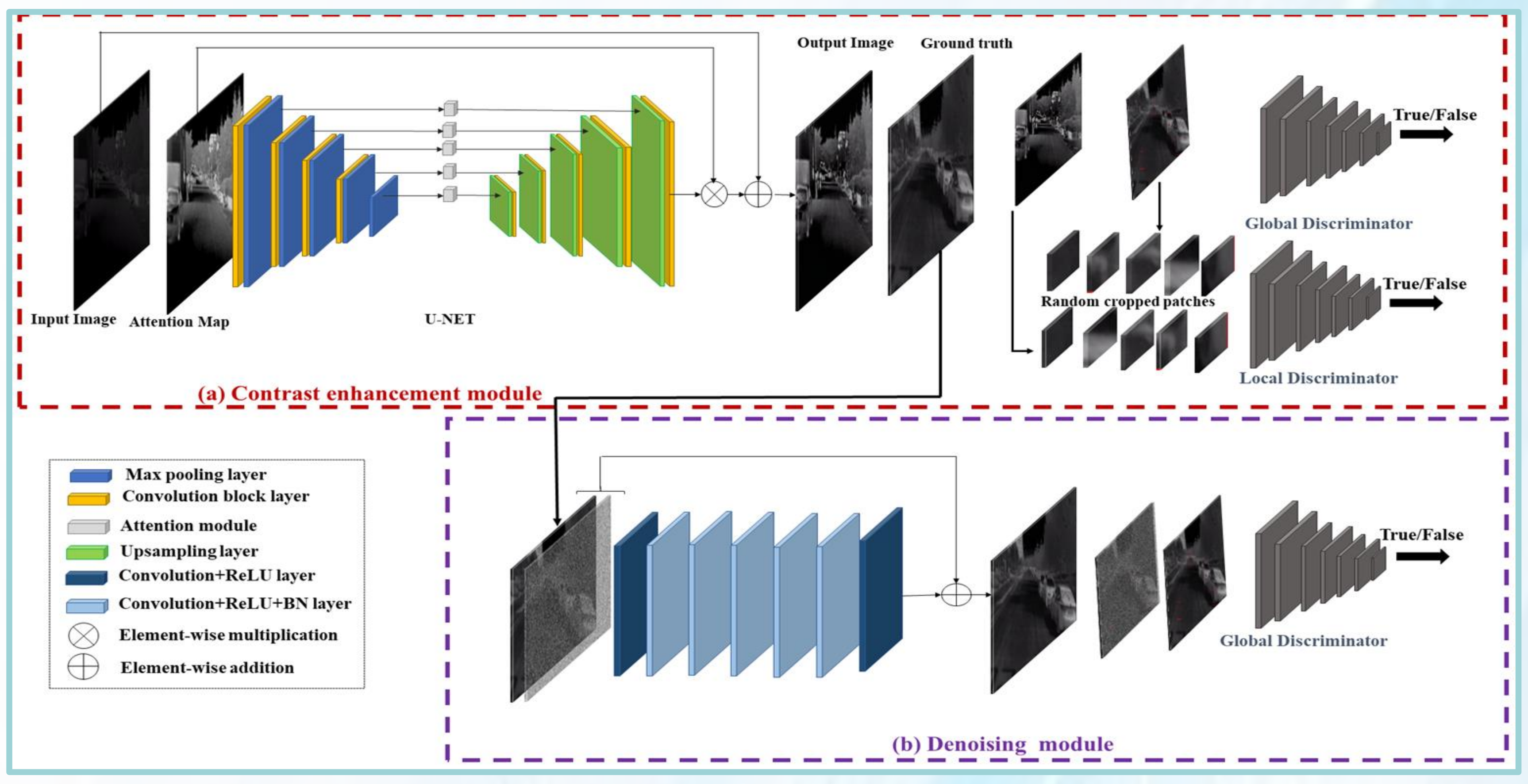

Fig.1 :The proposed TE-GAN architecture composed of two modules:(a) contrast enhancement and (b) denoising.

1.Contrast Enhancement Module

To improve the contrast using U-NET generator, global and local discriminators.

❑ **Perceptual Loss**

To compute the distance between the output image and the ground-truth based on high-level representations extracted from VGG pre-trained model.

2. Denoising Module

To remove the noise level by a CNN generator and a global discriminator.

❑ **Content Loss**

To minimize the low-level content errors between the noisy image and the denoised generated image using Pixel-wise MSE,.

3. Post-processing Module

To highlight the edges and decrease the visual blur effects by means of a convolutional edge enhancement filter

Dataset

KAIST dataset [3].

Total number of frames(RGB and thermal): 95k
Training set : 7601 thermal images
Testing set : 2252 thermal images

Fig.2: Examples of thermal images from KAIST datatset.

Experiments & Results

1.Visual quality evaluation in terms of:

- *The Peak Signal to Noise Ratio (PSNR)*
- *The Structural Similarity Index Metric (SSIM)*

	HE	CLAHE	TE-GAN
PSNR	7.81	11.92	13.92
SSIM	0.34	0.37	0.50

Tab.1: Comparison of the TE-GAN architecture to HE and CLAHE.

Fig.3: Qualitative results of our proposed architecture TE-GAN compared to other enhancement methods.

2. Detection results with and without enhancement in terms:

- *The mean Average Precision (mAP)*
- *The Log Average Miss Rate (LAMR)*

Testing conditions	Metric	Without enhancement	With enhancement
Day	mAP	0.61	0.63
	LAMR	0.41	0.40
Night	mAP	0.66	0.73
	LAMR	0.26	0.20
All	mAP	0.62	0.65
	LAMR	0.45	0.43

Tab.2: Comparison of the detection performance of YOLOv3 detector with and without TE-GAN enhancement.

Fig.4:Some results of pedestrian detection using YOLOv3 on thermal images from KAIST dataset with and without TE-GAN enhancement.

Conclusion & Future work

- The effectiveness of the proposed TE-GAN architecture is proven by obtaining better quantitative and qualitative results compared to the original thermal images and to other existing enhancement methods.
- The detection results using YOLOv3 detector are improved by means of TE-GAN architecture with a significant margin.
- Extension of the proposed TE-GAN architecture to incorporate a super-resolution module.
- Extend the proposed approach to other applications such as people tracking and activity recognition.

References

[1] Jiang, Yifan, et al. "Enlightengan: Deep light enhancement without paired supervision." arXiv preprint arXiv:1906.06972 (2019).
[2] Zhang, Kai, et al. "Beyond a gaussian denoiser: Residual learning of deep cnn for image denoising." IEEE Transactions on Image Processing 26.7 (2017): 3142-3155.
[3] Hwang, Soonmin, et al. "Multispectral pedestrian detection: Benchmark dataset and baseline." Proceedings of the IEEE conference on computer vision and pattern recognition. 2015.

This work has been supported by the DGVR research fund from the Tunisian Ministry of Higher Education and Scientific Research that is gratefully acknowledged.