Low-light image enhancement based on U-Net and Haar wavelet

Elissavet Batziou1,2, Konstantinos Ioannidis1, Ioannis Patras2, Stefanos Vrochidis1, Ioannis Kompatsiaris1

[1]Information Technologies Institute, Centre for Research and Technology Hellas, Thessaloniki, Greece

{batziou.el,kioannid,stefanos,ikom}@iti.gr

[2]School of Electronic Engineering and Computer Science, Queen Mary University of London, Mile End Road London E1 4NS

{e.batziou,ipatras}@qmul.ac.uk

CENTRE FOR RESEARCH & TECHNOLOGY - HELLAS
Information Technologies Institute

Abstract

Low-light image enhancement has been very popular for improving the visual quality of image representations, while low-light images often require advanced techniques to improve the perception of information for a human viewer. One of the main objectives in increasing the lighting conditions is to retain patterns, texture, and style with minimal deviations from the considered image. To this direction, we propose a low-light image enhancement method with Haar wavelet-based pooling to preserve texture regions and increase their quality

Introduction

Image acquisition in low-lighting conditions is challenging to be improved when low brightness, low contrast, a narrow gray range, color distortion and severe noise exists, and no other sensor is involved as for example multiple cameras of different specifications.

Challenges:

- maximize the information involved from the input image on the available patterns
- extract the colour information that is hidden behind the low values of luminance.

Contribution of our work:

- Photorealistic style transfer for under-exposed image enhancement. Feature aggregation combines features of different scales to reach an improved image reconstruction through photorealistic style transfer to achieve a natural stylization effect.
- A modified U-Net based architecture is proposed that involves dense blocks and wavelet pooling layers.
- Encoding using low frequency (LL) component and unpooling using high frequency components (LH, HL, HH). Haar wavelet decomposes the original signal into channels that capture different components, which leads to better image enhancement.

Methodology

A modified U-Net-based network is combined with wavelet transformations and Adaptive Instance Normalization (AdaIN). The proposed network is comprised by the following three sub-components: encoder, enhancer and decoder.

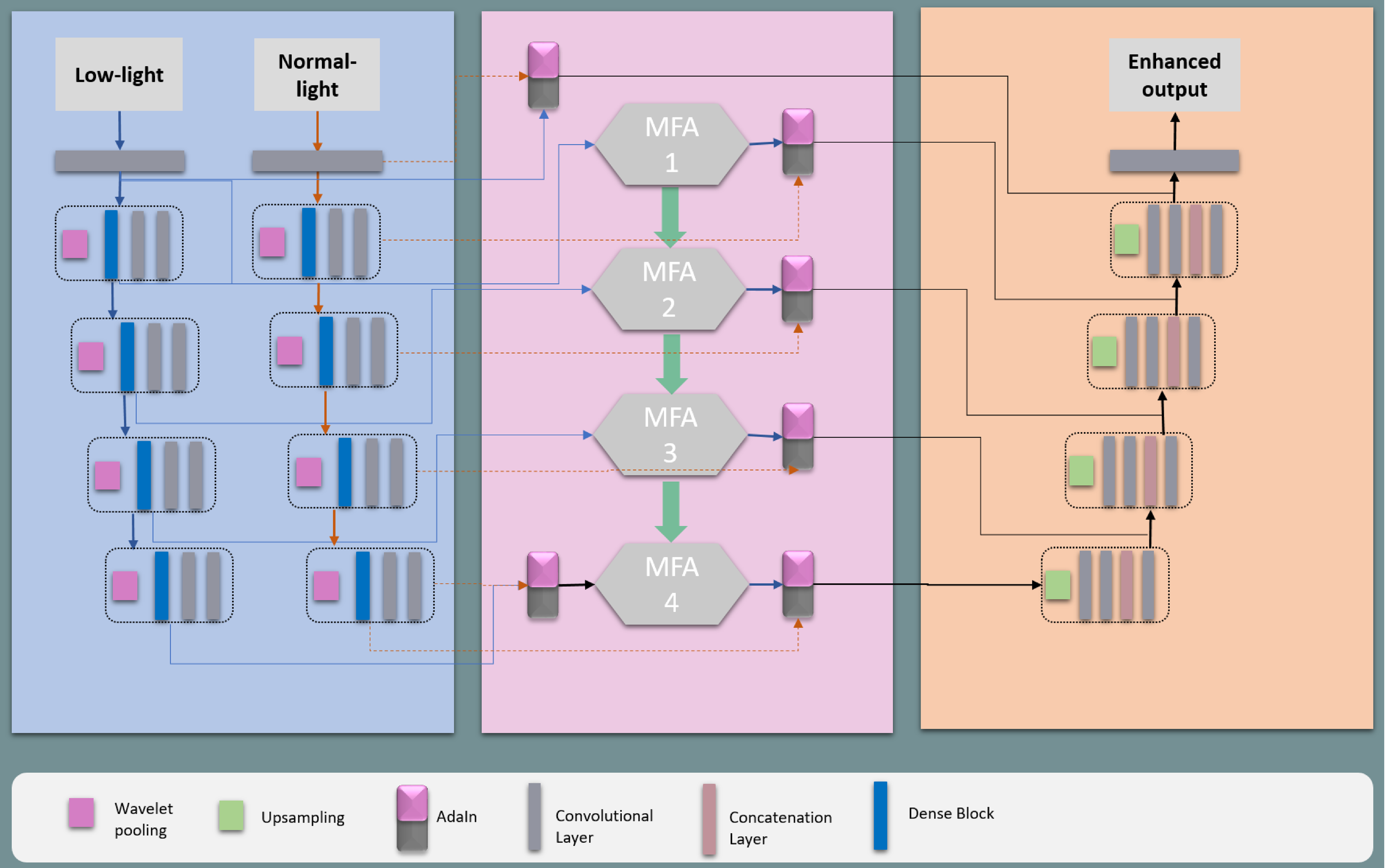
Encoder: has convolutional layers and downsampling dense blocks (blue box). Downsampling dense blocks contain:

- a Haar wavelet-based pooling layer
- a dense block (three densely connected conv. layers)
- two convolutional layers

Decoder: mirrors the encoder structure (yellow box) with four upsampling blocks. Each block consists of:

- an upsampling layer
- three convolutional layers
- a concatenation operation that receives the corresponding encoder feature coming from the enhancer module.

Enhancer: passes multi-scale features from the encoder to the decoder (purple box). Feature aggregation allows the network to incorporate spatial information from various scales and keep the detailed information of the input image.



Datasets

LOL dataset has 500 image pairs:

- 485 pairs randomly selected for training
- 15 pairs for testing.

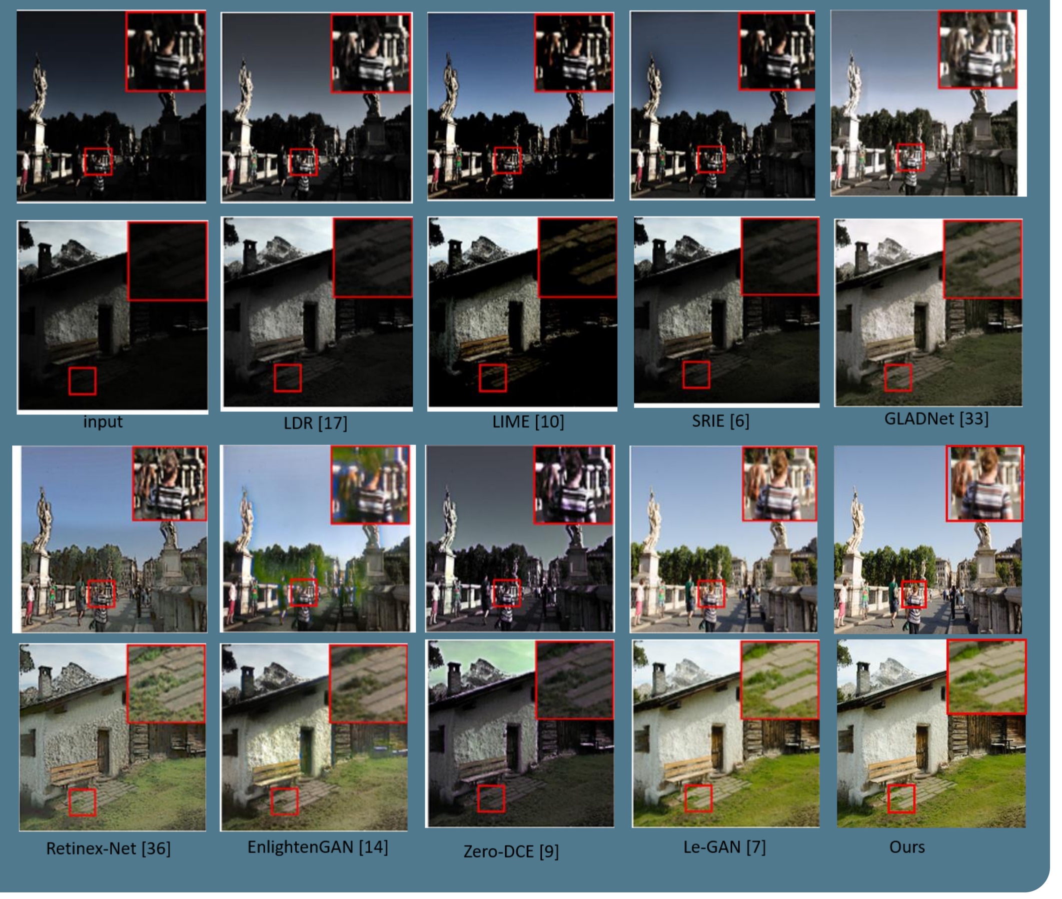
SYN dataset has 1,000 image pairs

- 950 pairs randomly selected for training
- 50 pairs for testing

Results

- Enhanced output image
- Measured performance
  - Peak Signal-to-NoiseRatio (PSNR) scores correspond to the average value of the complete test set.
  - Stractural Similarity Index (SSIM) across the average value of the complete test set
- Computational cost

Methods	LOL		SYN		Times	
	PSNR	SSIM	PSNR	SSIM	training	testing
LIME	15.484	0.634	14.318	0.554	-	-
Robust Retinex	13.876	0.669	16.873	0.714	-	-
JED	13.685	0.657	-	-	-	-
HE	14.800	0.386	-	-	-	-
AMSR	11.737	0.304	-	-	-	-
CMR	12.327	0.472	-	-	-	-
LightenNet	10.301	0.401	-	-	-	-
Zero-DCE	10.770	0.426	15.600	0.796	-	-
FSP	19.347	0.751			-	-
LDR	15.484	0.634	13.187	0.591	-	-
SRIE	17.440	0.649	14.487	0.639	-	-
GLADNet	20.314	0.739	16.761	0.797	-	-
Retinex-Net	17.780	0.425	16.286	0.779	-	-
EnlightenGAN	18.850	0.736	16.073	**0.827**	-	-
Le-GAN	**22.449**	**0.886**	**24.014**	**0.899**	25.6h	8.43ms
Ours	**21.266**	**0.784**	**19.212**	0.716	**10.7h**	**5.27ms**



Results (cont.)

Method	Graphical card	memory
Le-GAN	3 NVIDIA 3090ti GPUs	3*24GB
Ours	1 NVIDIA GeForce RTX 2060 SUPER	8GB

Qualitative evaluation:

- there are some bright regions of the results generated by the compared methods, which are over- exposed in the enhanced outputs.
- the proposed approach performs well in all datasets with nearly no artifacts and generates quite realistic normal-light images.

Quantitative evaluation:

- the highest PSNR values for each image test correspond to the output of the proposed framework and the method of Le-GAN
- the proposed approach is significantly more efficient, without the need of the additional computational resources
- the training time of Le-GAN method is 2.5 times higher the training time of the proposed approach
- Le-GAN uses 3 NVIDIA 3090ti GPUs contrary to our method which uses the NVIDIA GeForce RTX 2060 SUPER.

Conclusion and Future work

- A photorealistic style transfer approach is adapted and modified to low-light image enhancement.
- The method is based on a U-net architecture using dense blocks and three main modules, namely encoder, enhancer and decoder.
- The encoder and enhancer include a Haar wavelet based pooling layer.
- Experiments in two datasets compared the proposed method with recent works in image enhancement, showing that the proposed approach achieves SoA with less computational resources in training and testing processes.
- Future modifications will focus on evaluating further the impact of the wavelet incorporation and their deployment in other levels of the architecture.

Acknowledgements

This work was partially supported by the European Commission under contracts H2020-952133 xR4DRAMA and H2020-958161 ASHVIN.