

UCAN: A Learning-based Model to Enhance Poorly Exposed Images

Lucas R. V. Messias, Cristiano R. Steffens, Paulo L. J. Drews-Jr and Silvia S. C. Botelho

Centro de Ciências Computacionais - C3

Universidade Federal do Rio Grande - FURG

Rio Grande, Brazil

Email: l.r.v.m, cristianosteffens, paulodrews, silviacb@furg.br

Abstract—Image enhancement is a critical process in image-based systems. In these systems, image quality is a crucial factor to achieve a good performance. Scenes with a dynamic range above the capability of the camera or poor lighting are challenging conditions, which usually result in low contrast images, and, with that, we can have the underexposure and/or overexposure problem. In this work, our aim is to restore ill-exposed images. For this purpose, we present UCAN, a small and fast learning-based model capable to restore and enhance poorly exposed images. The obtained results are evaluated using image quality indicators which show that the proposed network is able to improve images damaged by real and simulated exposure. Qualitative and quantitative results show that the proposed model outperforms the existing models for this objective.

I. INTRODUCTION

Scenes with high contrast represent a challenge for image acquisition systems. Images acquired from conventional cameras, which operate in the visible light spectrum, are commonly affected by artifacts and distortions resulting from the excess or lack of light. The radiance of the scene near the limits of the acquisition system results in underexposure and overexposure. When the radiance passes through those limits (clipping), the information is lost [1].

Underexposure is a phenomenon of digital photography that occurs when the camera’s sensor can not correctly capture the scene details in the darkest areas. Underexposure occurs due to inadequate lighting, insufficient exposure time, or small opening of the iris lens limiting the amount of light that reaches the sensor. Overexposure, on the other end, occurs when the sensor receives too much light and becomes unable to distinguish the scene details in the brighter parts of the image. Overexposure usually leads to large saturated areas where the scene content information is lost.

To estimate the radiance of an inappropriately exposed image, we need to restore and enhance non-clipped pixels to improve clarity and color fidelity, as well as reconstruct the regions where the signal has been clipped. Therefore, Deep-Learning models can overcome the limitations of traditional image enhancement approaches by being able to learn objects, textures, and patterns from a wide set of training data. Neural networks can enhance image restoration outcomes with the help of semantics and contextual information learned from the instances it has been trained on.

Life sciences, machine vision, voice recognition, natural language processing, autonomous vehicles, and a wide set of difficult image-based tasks have benefited from the advancements in artificial neural networks [2]. Estimating an ill-exposed image’s irradiance requires preservation and enhancement of the pixels with details to improve clarity and color accuracy, as well as reconstruction techniques for clipped regions. Deep-Learning models provide a viable alternative to perform the de-clipping task.

We present an end-to-end convolutional neural network for image enhancement for ill-exposed sRGB images. Our network architecture is designed to broaden the receptive field. The model is significantly smaller than other state-of-the-art CNNs working on this problem domain in terms of trainable parameters. The **main contributions** of this paper are summarized as follows: i) We present UCAN, a fast and small exposure correction model which is able to enhance color texture from an ill-exposed image; ii) We introduce Receptive Field Expander Block, allowing the model to access a wide region in the neighborhood, which reduces the need for scaling layers in the network. iii) We design a custom content-based objective function to maximize restoration and reconstruction on almost clipped regions; and iv) We provide quantitative and qualitative results on both under and over-exposed images, outperforming recently proposed methods.

II. RELATED WORKS

For ill-exposed images, luminance and color correction incorporate elements from distinct areas of image processing such as contrast enhancement, signal reconstruction, noise suppression, tone mapping, and image completion. Thus, the literature includes histogram equalization [3], [4], dehaze-based contrast enhancement [5], Retinex based contrast enhancement [6], camera response based models [7], as well as exposure fusion based models [8].

All of the previous methods depend entirely on the signal that is represented in the input image. Deep Learning based image processing has gained a lot of attention in recent years. Its applications include super-resolution [9], de-hazing [10], inpainting [11], as well as general image enhancement [12], [13], RAW low light image enhancement [14], and sRGB ill-exposure correction [15], [16].

III. METHODOLOGY

A. UCAN Model

U-Nets [17], [18] have become popular for image-to-image translation tasks. However, the amount of memory required to store the part results in the intermediate layers is a major feature affecting all models inspired by U-Net. Due to extensive usage of skip connections, where the early layers of the network (encoder) have to be stored for later use in the decoder layers, U-Nets require a lot of memory resources.

We propose UCAN, a new architecture based on ideas presented in [17], [19], and [13]. The proposed approach reduces the network memory requirements and increases the performance of image exposure adjustment. The network relies on dilated convolutions and learned downscaling and upscaling layers. Therefore, we substantially boost the performance in the restoration task and raise the size of each batch during the training process. Fig. 1 presents an overview of the network design.

Our Receptive Field Expander Block (RFEB), shown in Fig. 1, contains four dilated convolutional layers of 3×3 parallel with dilation values varying from 2^0 up to 2^3 . Therefore, each RFEB will cover 19 features in the input space utilizing just nine trainable weights per channel. Dilated convolutions favor context aggregation for each pixel by enabling the model to reach a wide neighborhood area, helping us to improve the scaling layers inside the network.

All hidden convolutional layers in the CNN model are followed by an Exponential Linear Unit (ELU). Through experiments, we have found it significantly speeds up the learning process and leads to overall higher image quality. ELUs' ability to result in negative values allows them to push the mean unit activation closer to zero, speeding up learning because they bring the gradient closer to the unit's natural gradient. We also use a ReLU non-linear activation function on the output layer to avoid negative outputs.

B. Loss Function

A custom error function is used to emphasize regions of the image closer to the sensor limit, and therefore more likely to suffer the adverse effects of inadequate exposure. This objective function combines, in a weighted combination, structural dissimilarity (DSSIM) and absolute mean error between pixel-by-pixel. DSSIM is based on SSIM [20], a similarity index calculated on various 3×3 window of an image.

Although DSSIM provides a good assessment of the similarity between two images, the index is unable to assess the pixel values at the exact position. Therefore, the objective function is complemented with the absolute error (AE) between pixel values of the model output and the input image. In order to favor the restoration of critical regions, greater weight is attributed in regions more prone to saturation or underexposure. For that, we use a matrix of weights W computed from the reference image b . Assuming that the images are represented in the interval $[0; 1]$, values close to the limits have a weight greater than values in the center of the scale.

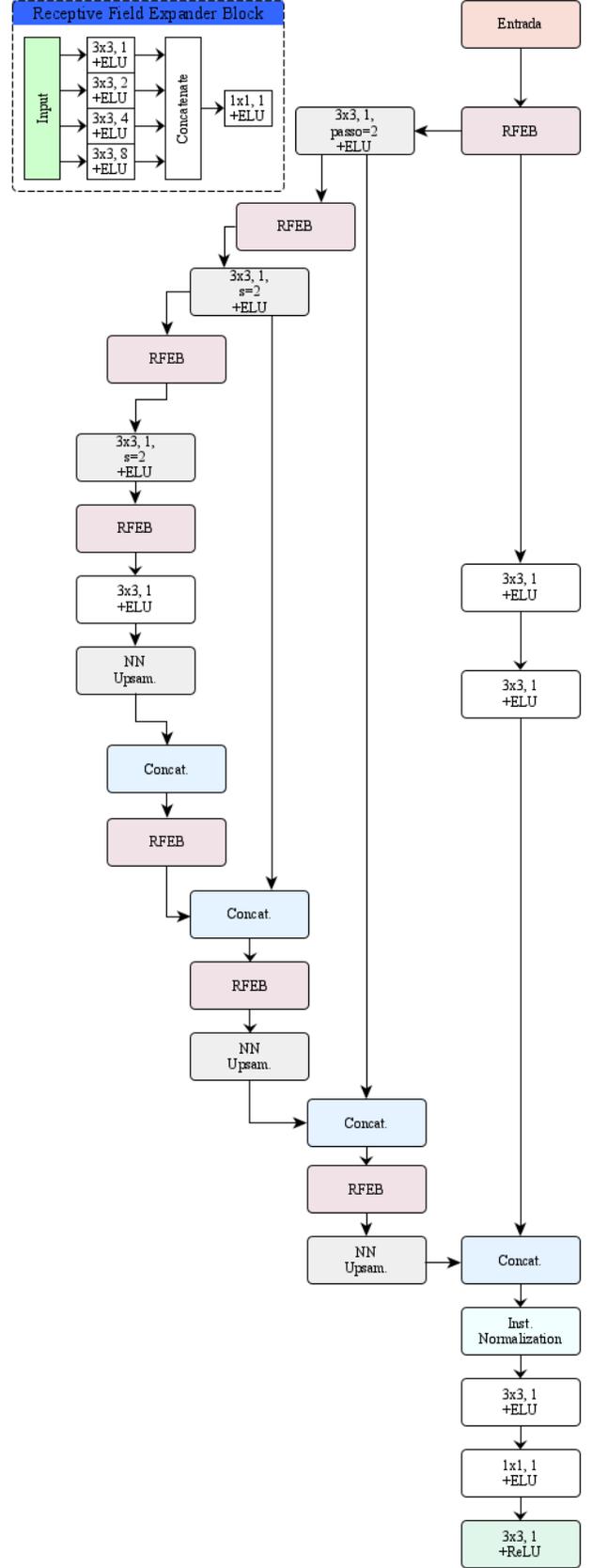


Fig. 1: UCAN architecture overview.

$$AE(\hat{I}, I^*) = |I^* - \hat{I}|. \quad (1)$$

$$W = |I^* - 0,5|. \quad (2)$$

In the end, the objective function is given by

$$\mathcal{L}(\hat{I}, I^*) = \lambda W \circ AE(\hat{I}, I^*) + (1 - \lambda)DSSIM(\hat{I}, I^*), \quad (3)$$

where \hat{I} and I^* represents the enhanced image and the target image. The empirical constant $\lambda = 0.2$ is used to compensate for the difference in scale between the two error functions.

C. Datasets

We used four sets of images with the ill-exposure distortion. Half of them are simulated making use of gamma transformation. Gamma power transformation is a nonlinear operation used to encode and decode luminance values in image systems [22] given by the equation:

$$I_{deg} = I^\gamma. \quad (4)$$

To generate the simulated datasets, we use gamma values between [2, 8] for underexposure and between $[\frac{1}{2}, \frac{1}{8}]$ for overexposure.

Simulated: FiveK-based and HDR+ Burst. The MIT-Adobe FiveK Dataset [23] contains 5,000 photographs shot with SLR cameras from a variety of various photographers. The HDR+ Burst Photography Dataset, initially presented by Hasinoff et al. [24], comprises sequences of images in different exposures by smartphone cameras.

Real: A6300 and Cai Multi-Exposure Datasets. Proposed by Steffens *et al.* [25], the A6300 dataset is composed of sets of 4 images for each scene: an appropriately exposed image using a single photograph, an underexposed image, an overexposed image, and a composition of the previous ones using the Tone Mapping method of Mertens et al. [26]. The Cai dataset, presented in [15], it consists of 589 image sets with separate exposure settings for each scene and a tone-mapped composition using the [26] method.

D. Training Phase

The proposed model is adjusted and tested on four different sets of images. In all cases, 70% of the dataset is used for training and the remainder for testing. The samples used for each stage are selected at random. The Adam optimizer [27] is used with the standard hyper-parameters. Weights are updated in mini-batches of 8 images with varying resolution. All data used for training is paired.

The four datasets are interspersed during the training step so that the model can learn the problem worked on more quickly and effectively. The idea is to use the simulated datasets at the beginning of the training, bringing to the model the basic characteristics of degradation, and the real ones at the end, refining the model to what it will adjust to challenges in real miss-exposed images.

Model	PSNR	MAE	SSIM
UCAN	2.070.E+01	9.282E-02	8.885E-01
U-net [17]	1.913E+01	1.119E-01	8.445E-01
Can24 [13]	1.663E+01	1.640E-01	8.152E-01
DHE [4]	1.483E+01	1.709E-01	7.56E-01
Ying [7]	1.408E+01	2.241E-01	7.498E-01
Fu [21]	1.366E+01	2.310E-01	7.393E-01
Input	1.409E+01	2.441E-01	6.749E-01

TABLE I: Quantitative results of related methods.

Training for underexposed and overexposed images is carried out separately, resulting in a specific model for restoring underexposed images and a specific model for overexposed images. Training is terminated once 300 batches of images are processed without making improvements larger than 10^{-5} . An identical criterion is applied to the U-Net [17] and CAN [13] architectures used in the comparison.

IV. RESULTS

In this section, we compare the enhancement methods from the literature with the UCAN model on the test data. In the quantitative evaluation, we use three image quality measurements: Peak Signal-to-Noise ratio (PSNR), Mean Absolute Error (MAE), and Structural Similarity (SSIM) [20]. With UCAN model, three non-learning-based models [4], [7], [21] and two learning-based models [13], [17], trained in the same conditions, were tested. Table I shows that the proposed model outperforms the enhancement methods with a big advantage. In all the metrics, the learning-based models are better in this task than the conventional ones, showing that the road to resolve the ill-exposure degradation is with neural networks.

Qualitatively, Fig. 2 shows a pair of ill-exposed images applied to the enhancement methods. We note a major change in the brightness of the components, restoring texture and re-coloring. In wide regions where all three channels are clipped, we find that our model is still unable to restore the surface smoothness, resulting in images that retain the JPEG compression block artifacts. Overall, a subjective assessment demonstrates the robustness and efficacy of the experimental approach for difficult situations of ill-exposure.

V. CONCLUSION

In this work, we propose UCAN, a CNN-based image enhancement model designed to optimize signal restoration and feature restoration of poorly exposed sRGB pictures. Ucan combines an architecture inspired by U-Net with a Receptive Field Expander Block that enables proper global luminance correction to be achieved. A custom-built objective function is presented that emphasizes restoring the color on this kind of problem and preserves the gradients of the input image. In terms of brightness improvement, contrast enhancement, and edge reconstruction, quantitative and qualitative assessment utilizing four datasets has shown our model substantially improved over the results of conventional image processing methods. As future work, we plan to optimize the smoothness of recovered regions (de-blocking), the synthesis of texture,

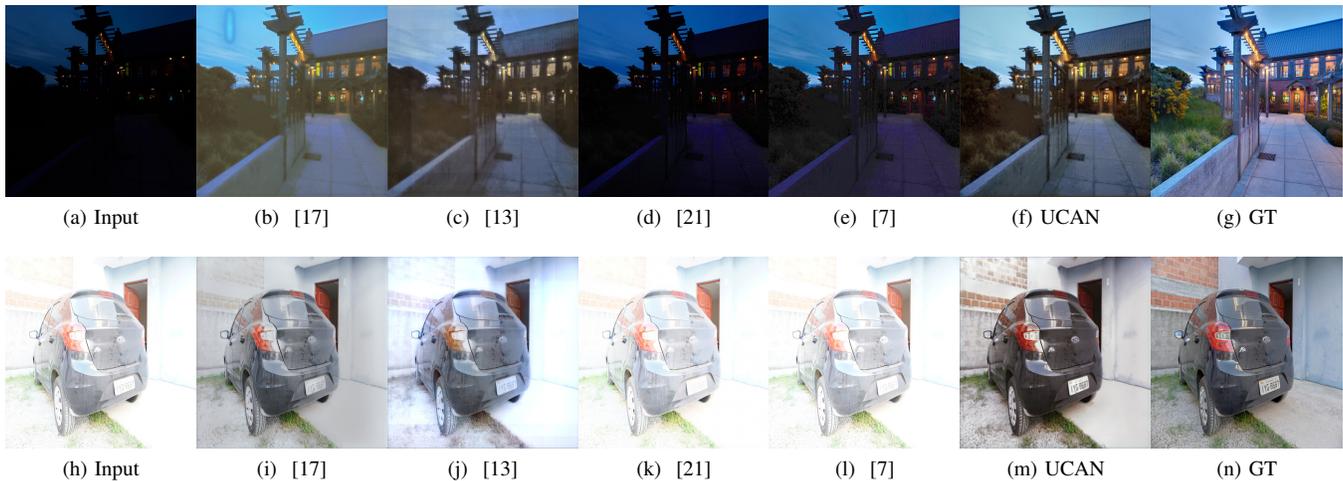


Fig. 2: Qualitative results of related methods, where [a-g] represents underexposure and [h-n] represents overexposure.

and the completion of broad clipping utilizing semantic characteristics.

REFERENCES

- [1] S. W. Hasinoff, "Saturation (imaging)," in *Computer Vision*. Springer, 2014, pp. 699–701.
- [2] Y. LeCun, Y. Bengio, and G. Hinton, "Deep learning," *nature*, vol. 521, no. 7553, pp. 436–444, 2015.
- [3] H. Ibrahim and N. S. P. Kong, "Brightness preserving dynamic histogram equalization for image contrast enhancement," *IEEE Transactions on Consumer Electronics*, vol. 53, no. 4, pp. 1752–1758, 2007.
- [4] M. Abdullah-Al-Wadud, M. H. Kabir, M. A. A. Dewan, and O. Chae, "A dynamic histogram equalization for image contrast enhancement," *IEEE Transactions on Consumer Electronics*, vol. 53, no. 2, pp. 593–600, May 2007.
- [5] X. Dong, G. Wang, Y. Pang, W. Li, J. Wen, W. Meng, and Y. Lu, "Fast efficient algorithm for enhancement of low lighting video," in *2011 IEEE International Conference on Multimedia and Expo*. IEEE, 2011, pp. 1–6.
- [6] A. B. Petro, C. Sbert, and J.-M. Morel, "Multiscale retinex," *Image Processing On Line*, pp. 71–88, 2014.
- [7] Z. Ying, G. Li, Y. Ren, R. Wang, and W. Wang, "A new low-light image enhancement algorithm using camera response model," in *Computer Vision Workshop (ICCVW), 2017 IEEE International Conference on*. IEEE, 2017, pp. 3015–3022.
- [8] —, "A new image contrast enhancement algorithm using exposure fusion framework," in *International Conference on Computer Analysis of Images and Patterns*. Springer, 2017, pp. 36–46.
- [9] T.-C. Wang, M.-Y. Liu, J.-Y. Zhu, A. Tao, J. Kautz, and B. Catanzaro, "High-resolution image synthesis and semantic manipulation with conditional gans," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 2018, pp. 8798–8807.
- [10] L. T. Gonçalves, J. F. de Oliveira Gaya, P. J. L. D. Junior, and S. S. da Costa Botelho, "Guidednet: Single image dehazing using an end-to-end convolutional neural network," in *2018 31st SIBGRAPI Conference on Graphics, Patterns and Images (SIBGRAPI)*. IEEE, 2018, pp. 79–86.
- [11] D. Pathak, P. Krahenbuhl, J. Donahue, T. Darrell, and A. A. Efros, "Context encoders: Feature learning by inpainting," in *The IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, June 2016.
- [12] A. Ignatov, N. Kobyshev, R. Timofte, K. Vanhoey, and L. Van Gool, "Wespe: Weakly supervised photo enhancer for digital cameras," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition Workshops*, 2018, pp. 691–700.
- [13] Q. Chen, J. Xu, and V. Koltun, "Fast image processing with fully-convolutional networks," in *IEEE International Conference on Computer Vision*, vol. 9, 2017, pp. 2516–2525.
- [14] C. Chen, Q. Chen, J. Xu, and V. Koltun, "Learning to see in the dark," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2018.
- [15] J. Cai, S. Gu, and L. Zhang, "Learning a deep single image contrast enhancer from multi-exposure images," *IEEE Transactions on Image Processing*, vol. 27, no. 4, pp. 2049–2062, 2018.
- [16] C. Li, J. Guo, F. Porikli, and Y. Pang, "Lightnet: a convolutional neural network for weakly illuminated image enhancement," *Pattern Recognition Letters*, vol. 104, pp. 15–22, 2018.
- [17] O. Ronneberger, P. Fischer, and T. Brox, "U-net: Convolutional networks for biomedical image segmentation," in *International Conference on Medical image computing and computer-assisted intervention*. Springer, 2015, pp. 234–241.
- [18] Z. Zhou, M. M. R. Siddiquee, N. Tajbakhsh, and J. Liang, "Unet++: A nested u-net architecture for medical image segmentation," in *Deep Learning in Medical Image Analysis and Multimodal Learning for Clinical Decision Support*. Springer, 2018, pp. 3–11.
- [19] C. Steffens, P. Drews-Jr, and S. Botelho, "Deep learning based exposure correction for image exposure correction with application in computer vision for robotics," in *Latin American Robotic Symposium and Brazilian Symposium on Robotics (LARS/SBR)*. IEEE, 2018, pp. 194–200.
- [20] Z. Wang, A. C. Bovik, H. R. Sheikh, and E. P. Simoncelli, "Image quality assessment: from error visibility to structural similarity," *IEEE transactions on image processing*, vol. 13, no. 4, pp. 600–612, 2004.
- [21] X. Fu, Y. Liao, D. Zeng, Y. Huang, X.-P. Zhang, and X. Ding, "A probabilistic method for image enhancement with simultaneous illumination and reflectance estimation," *IEEE Transactions on Image Processing*, vol. 24, no. 12, pp. 4965–4977, 2015.
- [22] R. Szeliski, *Computer vision: algorithms and applications*. Springer Science & Business Media, 2010.
- [23] V. Bychkovsky, S. Paris, E. Chan, and F. Durand, "Learning photographic global tonal adjustment with a database of input / output image pairs," in *The Twenty-Fourth IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2011.
- [24] S. W. Hasinoff, D. Sharlet, R. Geiss, A. Adams, J. T. Barron, F. Kainz, J. Chen, and M. Levoy, "Burst photography for high dynamic range and low-light imaging on mobile cameras," *ACM Transactions on Graphics (TOG)*, vol. 35, no. 6, p. 192, 2016.
- [25] C. R. Steffens, L. R. V. Messias, P. J. L. Drews-Jr, and S. S. d. C. Botelho, "Cnn based image restoration," *Journal of Intelligent & Robotic Systems*, vol. 99, no. 3, pp. 609–627, Sep 2020. [Online]. Available: <https://doi.org/10.1007/s10846-019-01124-9>
- [26] T. Mertens, J. Kautz, and F. Van Reeth, "Exposure fusion," in *Computer Graphics and Applications, 2007. PG'07. 15th Pacific Conference on*. IEEE, 2007, pp. 382–390.
- [27] D. P. Kingma and J. Ba, "Adam: A method for stochastic optimization," *arXiv preprint arXiv:1412.6980*, 2014.