Applying Genetic Algorithm and Image Quality Assessment for reproducible processing of low-light images

Olivier Parisot, Thomas Tamisier
Luxembourg Institute of Science and Technology (LIST)
5 Avenue des Hauts-Fourneaux
4362 Esch-sur-Alzette - Luxembourg
olivier.parisot@list.lu

Keywords:
Low-light images, Image Quality Assessment, Genetic algorithms.

Abstract:
Reproducible images preprocessing is fundamental in computer vision, whether to fairly compare process algorithms or to prepare new images corpus. In this paper, we propose an approach based on genetic algorithm combined to Image Quality Assessment methods to obtain a reproducible sequence of transformations for improving low-light images. Preliminary tests have been performed on state-of-the-art benchmarks.

1 INTRODUCTION

Images captured in poor lighting conditions often exhibit characteristics such as low brightness, low contrast, narrow gray scale, color distortion, and high noise - making them difficult for the human eyes to view details (Wang et al., 2020). Improvement of the quality of such images is a popular research area in computer vision.

In general, applying appropriate transformation to improve given image requires powerfull tools and strong expertise (Chaudhary et al., 2018). For instance, a regular user of dedicated software like Gimp or Photoshop process images by incrementally creating/modifying/merging layers until the result is satisfying. In order to automate as much as possible this workflow, two elements are essential. On the one hand, it is important to use specific metrics to guide the process: in this regard, Image Quality Assessment aims at estimating the quality of an image in a way that corresponds to a human subjective scoring of the same image (Zhai and Min, 2020). On the other hand, new techniques are constantly proposed in the literature to enhance images (Parekh et al., 2021); nevertheles, most of them are based on Deep Learning techniques that produce effective results – the effective transformation is then difficult to interpret or reproduce by another method (Buhrmester et al., 2021). This is particularly the case for low-light images, as shown by a recent survey presenting the recent works (Li et al., 2021).

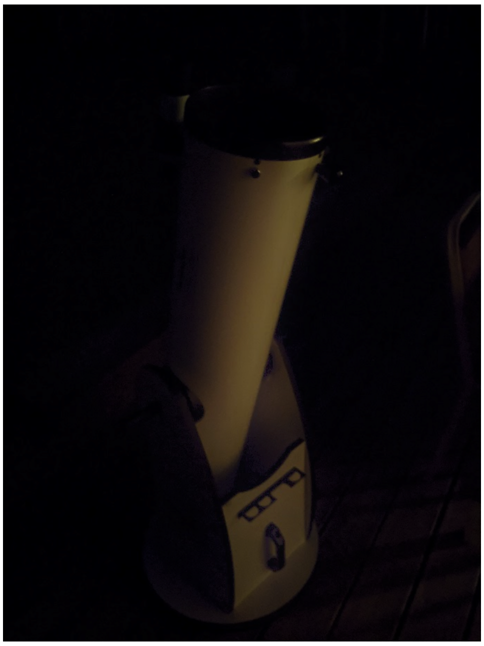


Figure 1: High resolution photography of a telescope captured by the author during the night time by using a smartphone. The picture was not processed by an additional software – only a minimal treatment was applied by the smartphone firmware.

However, in the context of academic research or industrial innovation, it is increasingly required to guarantee the reproducibility of experiments by keeping trace of the transformations performed on the images (Berg, 2018). As an example, a recent paper has shown that an important proportion of research works lacks of transparency regarding image handling and it may compromise the interpretation of the leading results (Miura and Nørrelykke, 2021). We can make an analogy with Machine Learning: data preprocessing should be transparent in order to lead to meaningful and trustable predictive models (Zelaya, 2019).

In this paper, we propose an approach based on a genetic algorithm to obtain a reproducible improvement of low-light images quality by relying on transformations monitored by Image Quality Assessment methods.

The rest of this article is organized as follows. Firstly, related works about Image Quality Assessment and image quality improvement of low-light images are briefly presented (Section 2). Then, our approach to improve the quality of low-light images is described (Section 3). Finally, a concrete implementation is detailed (Section 4), the results of preliminary experiments are discussed (Section 5) and we conclude by opening some perspectives (Section 6).

2 RELATED WORKS

2.1 Image Quality Assesment

Numerous Image Quality Assessment approaches were developed in recent years and an exhaustive list was already compiled (Zhai and Min, 2020). They are is widely used in benchmarks to compare the efficiency of image processing algorithms (Li et al., 2018).

We can distinguish two main types of techniques: Full-reference (FR) and Reduced-reference (RR) methods are based on a referential of images (raw/distorted) while No-reference (NR) and Blind methods intend to estimate single image quality (Liu et al., 2019). In this paper, we prefer to focus on NR and Blind approaches because because in practice it is very often difficult to obtain both raw and corrected images. Among them, we can mention:

- Classical methods like BRISQUE (Blind/Referenceless Image Spatial Quality Evaluator): a score between 0 and 100 is produced (0 for good quality image, 100 for poor quality) (Mittal et al., 2012).
- Recent Deep Learning methods like NIMA (Neural Image Assessment) – a set of Convolutional Neural Networks to estimate the aesthetic and technical quality of images: a score between 0 and 10 is produced (0 for poor quality, 10 for good quality) (Talebi and Milanfar, 2018).
- Dedicated techniques for low-light images like NLIEE (No-reference Low-light Image Enhancement Evaluation) (Zhang et al., 2021): the leading quality score represents various aspects like light, color comparison, noise and structure.

2.2 Genetic algorithm for images processing

Nature Inspired Optimization is a family of problem-solving approaches derived from natural processes. Among them, the most popular include genetic algorithms and particle swarm optimization (Li et al., 2020). These approaches are increasly applied in image processing for various tasks such as blur and noise reduction, restoration and segmentation (Dhal et al., 2019; Ramson et al., 2019). In particular, (Parisot and Tamisier, 2021) process images with a Nature Inspired Optimization Algorithm.

To the best of our knowledge, there are no much contributions about the reproducible transformations of low-light images by applying genetic algorithm guided by Image Quality Assessment techniques.

3 APPROACH

The cornerstone of our approach is defined as follows:

- An initial low-light image.
- A sequence of specific transformations applied on the initial image (examples: brighten, enhance, dehaze, adjust histogram, deblur, total variation denoise, etc.).
- A quality score evaluated by using a method S. This step is critical and drives the algorithm (quality serves here as the *fitness* of the solution, in the terminology used for evolutionary algorithms).

For a given low-light input image (I), by considering a quality evaluation method (S) and a maximum count of epochs (E), the following genetic algorithm computes the transformations sequences leading to an image with a better quality:

- A population is generated with P images: each image is a clone of the initial image I on which a random transformation has been applied or not. In fact, to ensure that the algorithm does not lead to a lower-quality image, it is important to keep at least one unmodified clone of the initial image in the population: at worst, it will remain the best solution.
- During E epochs:
  - The current best image or an other randomly selected image is cloned, and then a random transformation is applied: the newly created image is evaluated with S and added into the population.
  - An other image is randomly selected in the population and is stacked with the initial image (with a random weight): the newly created element is evaluated with S and added into the population.
  - According to the evaluation with S of the images present in the population, the worst images are selected and then removed from the population (to always keep P images in the population).
- The final result is the image of the consolidated population having the best quality estimation. The algorithm output is then an sequence of transformations that leads to an amelioration of the Image Quality Assessment.

The quality score resulting from the method (S) is evaluated by using both a Image Quality Assessment method and a brightness estimation. The quality estimator will be able to evaluate the global quality of the image while an explicit estimation of brightness may help to give a better score to brighter images as they tend to exhibit more details. As a result, we propose a quality score method (S) defined as follows:

- the quality score is the result of a selected Image Quality Assessment method.
- if the brightness of the image being evaluated is lower than that of the reference image, then a malus is applied to the score.
- Conversely, if the brightness of the image being evaluated is higher than that of the reference image, then a bonus is applied to the score.

To prevent the image from deviating too much from the original one, we have added a test comparing the similarity between the produced image and the initial image: if the similarity is too low (i.e. lower than a predefined threshold T), then the image score is strongly penalised and the last transformation is therefore not retained. The test is based here on the Structural Similarity Index (SSIM): in practice, the value is close to 1 when the two images are similar while the value is close to 0 when the images are really different.

4 PROTOTYPE

The algorithm has been implemented into a Python prototype. Various well-known open-source packages have been integrated. Images loading and transformations are realized with various dedicated packages like *openCV* [1] and *scikit-images* [2]. BRISQUE score is computed through the *image-quality* package [3] and NIMA scores are provided by a Tensorflow implementation [4].

By using these packages, these image transformations can be applied:

- Blurring and deblurring.
- Denoising/restoration: total variation, non local means, wavelets, bilateral, Noise2Noise (Lehtinen et al., 2018).
- Contrast adjustment / Histogram optimization by using CLAHE (Contrast Limited Adaptive histogram equalization, (Zuiderveld, 1994)).
- Background estimation and processing (Guo and Wan, 2018).
- Dehazing via Deep Learning methods like Cycle-Dehaze (Engin et al., 2018).
- Morphological transformations (like erode and dilate) (Sreedhar and Panlal, 2012).

Moreover, the brightness was evaluated by a method proposed by (Rex Finley, 2006).

The prototype was tested on a computing infrastructure with the following hardware configuration: 40 cores and 128 GB RAM (Intel(R) Xeon(R) Silver 4210 CPU @ 2.20GHz) and NVIDIA Tesla V100-PCIE-32GB. The CUDA [5]

[1] https://pypi.org/project/opencv-python/
[2] https://pypi.org/project/scikit-image/
[3] https://pypi.org/project/image-quality/
[4] https://github.com/idealo/image-quality-assessment
[5] https://developer.nvidia.com/cuda-zone

Table 1: Experiments on the VIP-LowLight benchmark: the (average,min,max) values are listed for each metric (before the algorithm execution).

	Raw images
BRISQUE	(22.8029, 21.0121, 26.0876)
NIMA-aesthetic	(4.3338, 4.0052, 4.6494)
NIMA-technical	(4.3953, 4.1569, 4.9886)
Noise variance	(5.7839, 1.701, 10.696)

Table 2: Experiments on the VIP-LowLight benchmark: the (average,min,max) values are listed for each metric (after the algorithm execution).

	Processed images
BRISQUE	(2.1749, 0.3492, 13.7365)
NIMA-aesthetic	(4.7018, 3.9337, 5.8975)
NIMA-technical	(4.7208, 4.2394, 5.0744)
Noise variance	(11.855, 2.337, 71.425)

et NUMBA [6] frameworks have been used to optimize the usage of the hardware (CPUs and GPUs).

5 FIRST EXPERIMENTS

The prototype was executed on low-light benchmarks, i.e. with images coming fron the LOL dataset (Wei et al., 2018) and the VIP-lowLight dataset (Chung and Wong, 2016), as shown in Figure 2 and in Table 1.

The presented method can thus be seamlessly inserted into any image processing workflow; not only is it possible to reproduce the image processing sequence, but it also allows to modify it afterwards if needed (for manual adjustments according to the specificities of the images – such as additional denoising).

Moreover, the first experiments show that the results obtained on the benchmarks are globally satisfactory. Table 1 and Table 2 have been obtained with the following hyperparameters: BRISQUE as targetted Image Quality Assessment score combined with brightness control, an initial population of 20 images, 50 maximum epochs and 0.25 as minimum similarity. According to significant runs, this setting offers the best tradeoff between quality improvement and execution time. BRISQUE score has been computed afterwards to check the quality of the algorithm inputs / outputs and Noise Variance (Immerkaer, 1996) has been esimated to highlight the level of noises in the benchmark.

[6] http://numba.pydata.org/

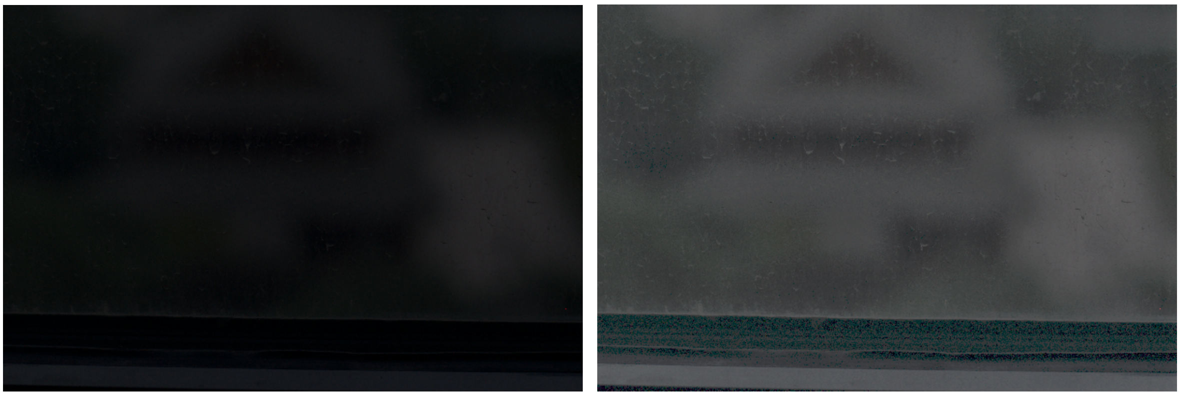
Finally, a word on performances: the time needed for the experiments was reasonnable on the infrastructure described above (from a few seconds to several dozen seconds per image – depending of the images shapes). During our preliminary tests, we have ran the algorithm on small (Figure 2) and large images (Figure 1) – and the computation time was not the same: the image transformation operations obviously took more time on high resolution images. In practice, a tradeoff between algorithm efficiency and execution time is required, and it may be controlled by the genetic algorithm settings (epochs count, population size, etc.). An other *trick* consists in using minimzed version of raw images during the genetic algorithm execution (let say by reducing the size by a factor of let two): once the sequence is calculated, it can be further applied to the original image. The quality evaluation will be less precise, but it will greatly accelerate the execution of the algorithm.

6 CONCLUSION

This paper presented an approach based on a genetic algorithm to improve the quality of a given low-light images from a reproducible sequence of transformations. A prototype based on Image Quality Assessment methods was implemented and tested on various state-of-the-art low-light images databases.

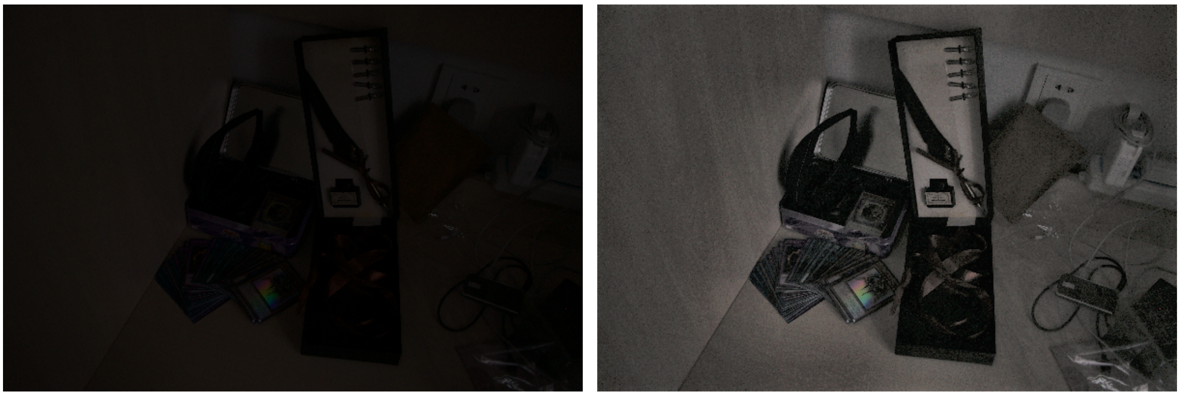
Thanks to academic and operational partners, we will set-up real-world use-cases to validate the approach. In parallel, we will improve the prototype by automatically generating the Python source code to transform the image as provided by Automated Machine Learning platforms for predictive models. Finally, we will work to improve execution performance by distributing calculations via frameworks like Spark because the Map/Reduce concept may drastically speed-up genetic algorithms execution.

Acknowledgments: This work was carried our during the MILAN project (MachIne Learning for AstroNomy) – funded by the Luxembourg National Research Fund. The tests were realized on the LIST Artificial Intelligence and Data Analytics platform (LIST AIDA). Special thanks to Raynald Jadoul and Jean-François Merche for their support.



(a) The original raw image. (b) The processed image.

Figure 2: Raw image coming from the LOL dataset (2a), and the second one was processed with our algorithm (2b). The original raw image has the following characteristics: BRISQUE=25.0677, noise variance=2.779.The following sequence has been computed: *adjust gamma (sigma=1), sum with 0.7x the original image, sum with 0.2x the original image, increase contrast, CLAHE (clipLimit=1).* The processed image has the following characteristics: BRISQUE=6.2736, noise variance=5.54.



(a) The original raw image. (b) The processed image.

Figure 3: An other raw image coming from the LOL dataset (3a), and the second one was processed with our algorithm (3b). The original raw image has the following characteristics: BRISQUE=21.8989, noise-variance=2.798 The following sequence has been computed: *blur (sigma=0.05), enhance (factor=1.05), CLAHE (clipLimit=2), sum with 0.3x the original image, enhance (factor=0.95), CLAHE (clipLimit=1).* The processed image has the following characteristics: BRISQUE=20.1891, noise-variance=6.523

REFERENCES

Berg, J. (2018). Progress on reproducibility.

Buhrmester, V., Münch, D., and Arens, M. (2021). Analysis of explainers of black box deep neural networks for computer vision: A survey. *Machine Learning and Knowledge Extraction*, 3(4):966–989.

Chaudhary, P., Shaw, K., and Mallick, P. K. (2018). A survey on image enhancement techniques using aesthetic community. In *International Conference on Intelligent Computing and Applications*, pages 585–596. Springer.

Chung, A. G. and Wong, A. (2016). Noise suppression and contrast enhancement via bayesian residual transform (brt) in low-light conditions. *Journal of Computational Vision and Imaging Systems*, 2(1).

Dhal, K. G., Ray, S., Das, A., and Das, S. (2019). A survey on nature-inspired optimization algorithms and their application in image enhancement domain. *Archives of Computational Methods in Engineering*, 26(5):1607–1638.

Engin, D., Genç, A., and Ekenel, H. K. (2018). Cycle-dehaze: Enhanced cyclegan for single image dehazing. In *The IEEE Conference on Computer Vision and Pattern Recognition (CVPR) Workshops*.

Guo, Y. and Wan, Y. (2018). Image enhancement algorithm based on background enhancement coefficient. In *2018 10th International Conference on Communications, Circuits and Systems (ICCCAS)*, pages 413–417.

Immerkaer, J. (1996). Fast noise variance estimation. *Computer vision and image understanding*, 64(2):300–302.

Lehtinen, J., Munkberg, J., Hasselgren, J., Laine, S., Karras, T., Aittala, M., and Aila, T. (2018). Noise2noise: Learning image restoration without clean data. *arXiv preprint arXiv:1803.04189*.

Li, B., Ren, W., Fu, D., Tao, D., Feng, D., Zeng, W., and Wang, Z. (2018). Benchmarking single-image dehazing and beyond. *IEEE Trans. on Image Processing*, 28(1).

Li, C., Guo, C., Han, L.-H., Jiang, J., Cheng, M.-M., Gu, J., and Loy, C. C. (2021). Low-light image and video enhancement using deep learning: a survey. *IEEE Transactions on Pattern Analysis & Machine Intelligence*, (01):1–1.

Li, H., Liu, X., Huang, Z., Zeng, C., Zou, P., Chu, Z., and Yi, J. (2020). Newly emerging nature-inspired optimization-algorithm review, unified framework, evaluation, and behavioural parameter optimization. *IEEE Access*, 8:72620–72649.

Liu, Y.-H., Yang, K.-F., and Yan, H.-M. (2019). No-reference image quality assessment method based on visual parameters. *Journal of Electronic Science and Tech.*, 17(2).

Mittal, A., Moorthy, A. K., and Bovik, A. C. (2012). No-reference image quality assessment in the spatial domain. *IEEE Trans. on image processing*, 21(12):4695–4708.

Miura, K. and Nørrelykke, S. F. (2021). Reproducible image handling and analysis. *The EMBO journal*, 40(3):e105889.

Parekh, J., Turakhia, P., Bhinderwala, H., and Dhage, S. N. (2021). A survey of image enhancement and object detection methods. *Advances in Computer, Communication and Computational Sciences*, pages 1035–1047.

Parisot, O. and Tamisier, T. (2021). Reproducible improvement of images quality through nature inspired optimisation. In *International Conference on Cooperative Design, Visualization and Engineering*, pages 335–341. Springer.

Ramson, S. J., Raju, K. L., Vishnu, S., and Anagnostopoulos, T. (2019). Nature inspired optimization techniques for image processing—a short review. *NIO Techniques for Image Processing Applications*, pages 113–145.

Rex Finley, D. (2006). Hsp color model — alternative to hsv (hsb) and hsl.

Sreedhar, K. and Panlal, B. (2012). Enhancement of images using morphological transformation. *arXiv preprint arXiv:1203.2514*.

Talebi, H. and Milanfar, P. (2018). Nima: Neural image assessment. *IEEE Trans. on Image Processing*, 27(8):3998–4011.

Wang, W., Wu, X., Yuan, X., and Gao, Z. (2020). An experiment-based review of low-light image enhancement methods. *IEEE Access*, 8:87884–87917.

Wei, C., Wang, W., Yang, W., and Liu, J. (2018). Deep retinex decomposition for low-light enhancement. *arXiv preprint arXiv:1808.04560*.

Zelaya, C. V. G. (2019). Towards explaining the effects of data preprocessing on machine learning. In *2019 IEEE 35th ICDE*, pages 2086–2090. IEEE.

Zhai, G. and Min, X. (2020). Perceptual image quality assessment: a survey. *Science China Information Sciences*, 63:1–52.

Zhang, Z., Sun, W., Min, X., Zhu, W., Wang, T., Lu, W., and Zhai, G. (2021). A no-reference evaluation metric for low-light image enhancement. In *2021 IEEE International Conference on Multimedia and Expo (ICME)*, pages 1–6. IEEE.

Zuiderveld, K. (1994). *Contrast Limited Adaptive Histogram Equalization*, page 474–485. Academic Press Professional, Inc., USA.