TreEnhance: An Automatic Tree-Search Based Method for Low-Light Image Enhancement

Marco Cotogni, Claudio Cusano
Dept. of Dept. of Electrical, Computer and Biomedical Engineering
University of Pavia, Italy
Via Ferrata 1, 27100, Pavia, Italy.
mengo.cotogni01@universitadipavia.it
claudio.cusano@unipv.it

ABSTRACT

In this paper we present TreEnhance, an automatic method for low-light image enhancement capable of improving the quality of digital images. The method combines tree search theory, and in particular the Monte Carlo Tree Search (MCTS) algorithm, with deep reinforcement learning. Given as input a low-light image, TreEnhance produces as output its enhanced version together with the sequence of image editing operations used to obtain it. The method repeatedly alternates two main phases. In the generation phase a modified version of MCTS explores the space of image editing operations and selects the most promising sequence. In the optimization phase the parameters of a neural network, implementing the enhancement policy, are updated. After training, two different inference solutions are proposed for the enhancement of new images: one is based on MCTS and is more accurate but more time and memory consuming; the other directly applies the learned policy and is faster but slightly less precise. Unlike other methods from the state of the art, TreEnhance does not pose any constraint on the image resolution and can be used in a variety of scenarios with minimal tuning. We tested the method on two datasets: the Low-Light dataset and the Adobe Five-K dataset obtaining good results from both a qualitative and a quantitative point of view.

Keywords Low-Light Image Enhancement · Deep Reinforcement Learning · Automatic Image Retouching · Image processing · Tree Search

1 Introduction

Image enhancement is an automatic method for low-light image enhancement capable of improving the quality of digital images. The method combines tree search theory, and in particular the Monte Carlo Tree Search (MCTS) algorithm, with deep reinforcement learning. Given as input a low-light image, TreEnhance produces as output its enhanced version together with the sequence of image editing operations used to obtain it. The method repeatedly alternates two main phases. In the generation phase a modified version of MCTS explores the space of image editing operations and selects the most promising sequence. In the optimization phase the parameters of a neural network, implementing the enhancement policy, are updated. After training, two different inference solutions are proposed for the enhancement of new images: one is based on MCTS and is more accurate but more time and memory consuming; the other directly applies the learned policy and is faster but slightly less precise. Unlike other methods from the state of the art, TreEnhance does not pose any constraint on the image resolution and can be used in a variety of scenarios with minimal tuning. We tested the method on two datasets: the Low-Light dataset and the Adobe Five-K dataset obtaining good results from both a qualitative and a quantitative point of view.

Keywords Low-Light Image Enhancement · Deep Reinforcement Learning · Automatic Image Retouching · Image processing · Tree Search

1 Introduction

Image enhancement is a well-known image processing problem that consists in improving the appearance of a low quality image to obtain a high quality version without changing the content. Nowadays, Deep Convolutional Neural Networks are the methods of choice for many image processing tasks such as image enhancement, super resolution, image inpainting, etc. For this reason, neural networks are now part of many of the image processing tools used by professional photographers to post-process their pictures. However, the use of this kind of tools requires a certain level of experience and knowledge, making desirable to have both semi-automatic and fully automatic procedures.

Among neural models, image-to-image translation methods [1-5] have been shown to be very effective for image enhancement. However, they suffer of two major problems: the presence of artifacts their output images, and the limitation on the resolution of the input that is often imposed by the architecture of the neural network.

Moreover, most recent methods work as black boxes, without providing any insight about the transformation found by the neural network. For a professional usage, it is of paramount importance that the method provides an “explainable” solution that makes it possible for the user to verify it and possibly to adjust it tow his or her needs. Possible solutions to provide an explanation to the neural networks’ predictions were explored in the recent years [6-8]. In these works the authors proposed methods to analyze the predictions obtained by the model and obtain some explanation behind the networks’ reasoning.
On the basis of these considerations, we propose here TreEnhance, a low-light image enhancement method based on tree search and deep reinforcement learning. The method is able to obtain high quality output images with the application of a sequence of image editing operations. The sequence selected by the method is fully explainable, making it suitable not only for automatic enhancement, but also as a tool for a variety of interactive scenarios. For instance, it could be used for educational purposes to teach beginners how to edit their own images.

The method does not pose any constraint on the resolution of input images. In fact it can deal with both high- and low-resolution images without any modification. We tested our method on two different datasets tuning only the parameters of the image editing operations.

The main contributions presented in this paper are:

- To the best of our knowledge, this is the first work presenting the application of tree-search theory to image enhancement problems.
- Differently from several previous methods in the state of the art, TreEnhance acts as a fully explainable neural architecture for image enhancement. This method provides along with the resulting output image the sequence of enhancing operators selected by the algorithm to enhance the low quality image. The images provided as output are good from a qualitative and quantitative points of view without the introduction of any artifact and without constraints on the image resolution.
- Once the model has been trained, TreEnhance presents two different inference strategies for enhance low quality image based on the data, time an memory constraints.

The paper is organized as follows: we begin with the presentation of the state of the art in image enhancement in Section 2. In Section 3, we explain how TreEnhance works, what are the steps of the method and which neural architecture it uses. In Section 4, we describe the datasets used for the experiments and the results obtained. Finally we conclude the paper with Section 5 and Section 6 where we discuss the results obtained and possible future investigations on this topic.

2 Related Work

In recent years, many image enhancement methods have been proposed. Among these, one of the most effective approaches is that represented by image-to-image translation methods. In this family of methods, a neural network learns how to directly convert an image from one domain to another (i.e. low quality to high quality), while preserving its visual content. These methods learn by observing thousands of pairs of low-quality/high-quality images.

Ravirathinam et al. proposed a multi-context framework based on a modified version of the UNet architecture for low-light image enhancement [9]. In this work the authors combined a perceptual loss, a structural loss and a patch-wise euclidean loss to enhance a low-light input image.

Xu et al. presented a decomposition and enhancement method for low-light image enhancement working in the frequency domain. The architecture presented learns how to recover image objects low-frequency layers. Once these image objects have been recovered, the method is able to enhance high-frequency image details.

Ignatov et al. presented an image-to-image translation pipeline for enhancing smartphone pictures [4]. In this work the ground truth images were obtained with a professional camera. In order to deal with misalignment problems due to the different resolutions of the cameras, the authors trained the neural model with aligned patches obtained from the original images.

Pix2Pix, Cycle-GAN and EnlightenGAN are methods based on an adversarial learning schema which involves a generator and a discriminator. In a image-to-image translation scenario, the generator learns how to produce images very similar to the training ones, the discriminator learns how to distinguish real training images from generated images. The optimal point is reached when the discriminator is not able to distinguish the images in the training from those produced by the generator [1–3]. These works show very good performances in several computer vision tasks including image enhancement.

Lv et al. proposed a method based on local features extractors for low light images [10]. A similar approach for underexposed images has been proposed instead by Wang et al. [11].

Zhang proposed a pipeline based on decomposition network and illumination adjustment to tune the exposure in low-light images [12]. The same approach has been further explored by other research groups in the following years [13–15].
Liu et al. presented a Retinex rule-based model for low-light image enhancement. The method firstly defines the neural architecture following the optimization process of Retinex-based models. Then, a two-level search strategy is presented to define suitable architectures for noise removal and illuminant estimation.\textsuperscript{[15]}

Differently from image-to-image translation methods, parametric methods do not model directly the mapping between low-quality and high-quality images but they learn instead the parameters of the color transformation to be applied to the low-quality image.

Bianco et al. presented two works: in the first one, a neural network learns the parameters of a color transformation. This transformation is applied to a filtered image to recover its original color curve. In the second one, the parameters of a color transformation are combined with a basis function to enhance the visual content of low-quality images\textsuperscript{[16,17]}. Another work from the same authors explores the use of a neural network to estimate the coefficients of splines, that are the used as color curves\textsuperscript{[18]}.

Zhang et al. proposed an Exposure Correction Network able to estimate the best S-shaped curve (a non-linear curve used to adjust the exposure of shadow-tone areas in images) to restore low-light images. This curve is then applied to the low-quality input image to enhance it\textsuperscript{[19]}.

Chai et al. presented an approach for parametric color enhancement. In this work, a convolutional neural network learns the parameters of a quadratic color transformation in a supervised learning scenario\textsuperscript{[20]}.

Kim et al. proposed a method based on representative color transformations. This method uses local and global enhancement modules to determine the most representative colors in input images and to estimate the transformation for these colors. Then, the model defines the enhanced colors by using the transformed colors. The criteria used to define the enhanced colors is based on the similarity between representative and input colors\textsuperscript{[21]}.

Guo et al. presented a zero-reference method for low-light image enhancement. In this work, a deep curve estimator network takes as input a low-quality image and it estimates a series of light-enhancement curves. These curves are then iteratively applied to the RGB input image obtaining the enhanced image. In order to train the estimator in a zero-reference scenario, four differentiable non-reference losses were involved\textsuperscript{[22]}.

Another kind of related approaches, are those based on reinforcement learning. In this family of methods an agent selects the enhancement operators to be applied to a low-quality image to obtain its enhanced version. During the training a numerical reward is used to improve the agent.

Park et al. presented a deep Q-learning based approach for image enhancement. The reward is modeled as the difference between the distance of the image from the ground truth before and after the application of the operator\textsuperscript{[23]}.

Yang et al. proposed a method based on Markov Decision Process for real-time exposure control. In this work, given the current frame to the agent, a fully convolutional neural network is trained by using the Gaussian policy gradient algorithm. The method is able to optimize the trade-off among convergence, minimal temporal oscillation and quality of the produced image\textsuperscript{[24]}.

Hu et al. proposed a white-box approach for image enhancement. This work is based on an actor-critic algorithm to enhance the content of a low-quality image\textsuperscript{[25]}.

Yu et al. presented a deep Q-learning based tool-chain for image restoration. In this work, a neural network learns what are the most suitable restoration operators that should be applied to a corrupted image to restore its content\textsuperscript{[26]}.

Furuta et al. presented an actor-critic approach for image processing based on a pixel-wise reward. They applied the method to several image processing tasks such as color enhancement, image restoration, etc.\textsuperscript{[27]}.

Yu et al. proposed a mixed approach based on GANs and deep reinforcement learning\textsuperscript{[28]}.

Finally, a new approach presented by Zhang et al. is based on a transformer architecture for enhancing low-quality images. This algorithm divides the images in patch, applies an embedding and passes the created vectors through several two-branches transformer modules obtaining the enhanced image\textsuperscript{[5]}.

3 Method

TreEnhance is a tree-based deep reinforcement learning method that finds sequences of image editing operations to improve the quality of digital photographs. It does so by emulating the behavior of a professional photo editor who selects the operations among those available in the software he or she uses. The selection follows a trial-and-error strategy in which multiple operations (such as gamma correction, contrast adjustment, etc.) are repeatedly tried, verified,
canceled, and combined. This way many different sequences of operations are evaluated and the most promising one is taken.

More precisely, given an image \( x \) the method selects the editing operations \( a_1, \ldots, a_n \) from a finite set \( \mathcal{A} \) of operations, and apply them to produce the enhanced image \( \hat{y} \):

\[
\hat{y} = a_n(a_{n-1}(\ldots a_1(x))). \tag{1}
\]

The sequence is terminated by a special STOP operation.

To select the operations, the method builds a search tree that explores a large number of candidate sequences. A convolutional neural network guides the growth of the tree by promoting the exploration of the operations that are more likely to contribute to the improvement of the image. The network also evaluates the images obtained during the process so that the best one can be selected.

The network is trained in a supervised way with the goal of obtaining an enhanced image \( \hat{y} \) as close as possible to a target image \( y \) manually edited by a human expert. The following sections present the method in details.

### 3.1 Training Procedure

A set of input/target image pairs is used to train the neural network. The training procedure alternates two phases: generation and optimization. In the generation phase a revised version of the Monte Carlo Tree Search (MCTS) algorithm is used to grow trees associated to a random selection of the available images pairs. In the optimization phase the images obtained in the generation phase are used to train the parameters of the neural network.

#### 3.1.1 Generation

Starting from the root node, which represents the input image, the method selects an image editing operation (an action) which results in a new image (i.e. a new node). Nodes and actions form a tree which is built by a sequence of iterations, each one adding a new node. A numerical score, called return, is assigned to each node. The goal is to find actions that are likely to lead to nodes associated to high returns. Each iteration of MCTS includes four major steps [29]:

1. **Selection**: a tree policy is applied to select a leaf node.
2. **Expansion**: new child nodes (one for each possible action) are attached to the selected leaf. They will all become new leaves of the tree.
3. **Evaluation**: the return is assigned to the node.
4. **Backup**: the return is propagated from the selected node up to the root of the tree.

In the original MCTS the evaluation step consists in the repeated application of a random rollout policy until a terminal node is found and the corresponding return is obtained.

Here, similarly to the method proposed by Silver et al. [30], the nodes chosen in the selection step are processed by a convolutional neural network. The network takes as input an image \( x \) and computes a policy \( \pi(x, \cdot) \) and a value \( v(x) \). The policy estimates the probability distribution of the best operation over all the available editing operations (\( \pi(x, a) \) is the probability estimate that \( a \) is the best operation for \( x \)). The value, instead, estimates the expected return \( r(\cdot) \) obtained by the descendants of \( x \).

The tree policy used in the selection step consists in the maximization of the Upper Confidence Bound (UCB) [31] which accounts for both the exploration of the less frequently visited nodes and the exploitation of those that obtained high returns on average:

\[
\text{UCB}(x) = \bar{r}(x) + c \cdot \pi(x_p, a) \sqrt{\frac{N(x_p)}{N(x) + 1}}. \tag{2}
\]

where \( \bar{r}(x) \) is the average return that has been propagated through \( x \) in previous iterations, \( a \in \mathcal{A} \) is the operation that produced \( x \) from its parent \( x_p \) (i.e. \( x = a(x_p) \)), and \( N(\cdot) \) denotes the number of times a given node has been visited in previous selection steps. The coefficient \( c \) balances the exploitation and the exploration terms.

The selection step starts from the root and descends the tree by choosing the nodes with the highest bound [2] until a leaf is reached. As a further mechanism to promote exploration we add a small amount of Dirichlet noise to the policy. When a terminal node is found, it is directly evaluated instead of being processed by the neural network. This happens in two cases: when a maximum depth is reached, and when the special STOP operation is selected. In the first case the search fails and the return is just zero. In the second case the return depends on the distance between the resulting image \( x \) and the target image \( y \):

\[
\bar{r}(x) = \exp(-\alpha \|x - y\|^2), \tag{3}
\]
where $\alpha$ is a tunable parameter. When a set number of MCTS iteration is reached, a policy $\rho$ for the root node is obtained.

$$
\rho(x, a) = \frac{N(a(x))}{N(x)}.
$$

This policy represents the probability distribution of the root’s child visit. From this probability distribution a random sample is extracted and the corresponding action is applied to the root image. This ensures a good degree of exploration by allowing the generation of multiple sequences of operations for the same image. From this new image (that becomes the new root of the tree) the complete procedure restarts and it is repeated until a terminal node is reached. Figure 1 depicts the generation phase of TreEnhance.

### 3.1.2 Optimization

The roots visited during the generation phase are used to build a training set. Each training element is a triplet combining (i) the image $x$, (ii) the return $r(x)$, and (iii) the probabilities of the stochastic policy $\rho(x, \cdot)$. The parameters of the neural architecture are updated by minimizing the loss function $L$:

$$
L = \lambda(r(x) - v(x))^2 - \sum_{a \in A} \rho(x, a) \log \pi(x, a),
$$

which combines the squared estimation error of the return with the cross entropy measuring the disagreement between the stochastic policy $\rho$ and the network policy $\pi$. The parameter $\lambda$ is a tunable hyperparameter used to balance the weight of the two terms in the loss.

The whole process is repeated multiple times, each time producing more training data in the generation phase and improving the neural network in the optimization phase.

### 3.2 Neural Architecture

The convolutional neural network used in this work is a modified version of the ResNet-18 [32] where the last fully connected classification layer is replaced by two feed forward heads. Both include two fully connected linear layers. The first head also includes the softmax function to produce the vector representing the policy $\pi(x, \cdot)$. The second one is followed by the application of the sigmoid activation function to compute the scalar value $v(x)$.

Figure 1: TreEnhance generation phase. To make the figure more clear only part of the tree is actually shown.
Figure 2: Neural network architecture. For the complete architecture of the ResNet-18, please refer to the original paper. The function $\pi(x, \cdot)$ is the policy over the editing operations and $v(x)$ is the estimated return.

We decided to consider the ResNet-18 as a backbone architecture since with a deeper model there is a high chance to overfit the (relatively small) datasets we used in the experiments. However, in principle there are no special limitations on the architecture of the neural network. The network is trained from scratch after random initialization.

The neural network accepts images of any spatial resolution. The only limitation is the amount of memory available in the system. During training an image needs to be stored for each node in the tree. Images can be downsampled to increase the number of trees that are grown in parallel and to speed-up training in general (in the experiments we resampled images to the resolution of $256 \times 256$ pixels).

Figure 2 shows the neural network architecture. In the first layer of the two heads, in order to avoid overfitting, dropout is applied with probability $p$, where $p$ is a tunable hyperparameter. For our experiments the value of this parameter was tuned to 0.6.

3.3 Inference

After training, TreEnhance can be used in two different ways: by building a tree (tree search) or by using directly the trained neural network (policy-based search). The first strategy is more accurate but it also slower. The second is faster and slightly less precise.

3.3.1 Tree Search

At inference time the method builds the tree as in the generation phase, but without having access to the target image for the computation of the return of terminal nodes. When the STOP action is selected or the maximum depth is reached, the return is estimated instead by $v(x)$. Once the tree has been built, a sequence of operations is obtained by descending the tree taking at each node the actions that maximize the policy $\rho(x, \cdot)$.

3.3.2 Policy-based Search

The second approach consists in the repeated selection of the operation that maximizes the probability $\pi(x, \cdot)$ computed by the neural network. The procedure terminates when the STOP action is selected or the maximum number of operations is reached.

4 Experimental Results

To verify the quality of the solutions found by TreEnhance we evaluated it on two datasets widely used to assess the performance of image enhancement methods. More in detail, we used the LOw-light (LOL) dataset and the Adobe five-K dataset. For both datasets global editing operations for tone adjustment and contrast enhancement are required. In the case of LOL spatial operators (such as deblurring, edge enhancement, etc.) can also be useful.

4.1 LOw-Light Dataset

The LOL dataset [33] is composed of 500 pairs of low-light/normal-light images obtained from real scenes. The training set includes 485 images and the test set 15. One of the main problems of real-scene datasets is the alignment of the image pairs. For this reason the authors of the dataset removed all the pairs that have a mean squared error greater than a fixed threshold (0.1). The enhancement of the images in this dataset requires mostly the application of global
transformations, white balancing algorithms and spatial filters. The set of operations \( A \) we selected for this experiment includes:

- **Brightness adjustment**: a constant \( \delta \) is added to the values of all pixels:
  \[
  (a(x))_{ijc} = \text{clip}(x_{ijc} + \delta).
  \]
  The clip function limits the adjusted values to the \([0, 1]\) range (\( \text{clip}(x) = \min\{\max\{x, 0\}, 1\} \)). We considered eight variants of this operator: with \( \delta \in \{-0.1, +0.1\} \), and applied to all the color channels or to a single color channel at a time \( c \in \{R, G, B\} \).

- **Contrast adjustment**: the pixels values are stretched by a factor \( \beta \) around their mean:
  \[
  (a(x))_{ijc} = \text{clip}(\mu_c + \beta \times (x_{ijc} - \mu_c)),
  \]
  where \( \mu_c \) is the average channel value. We considered the variants with \( \beta \in \{0.8, 2.0\} \) and applied to all the channels, or just one.

- **Gamma Correction**: we considered two values \( \gamma \in \{0.6, 1.1\} \), and applied them to all or to one of the channels:
  \[
  (a(x))_{ijc} = (x_{ijc})^\gamma.
  \]

- **Saturation adjustment**: after the conversion to the HSV color space, the \( S \) channel is scaled by a factor \( s \in \{0.5, 2.0\} \) and then the pixel is converted back in the RGB space.

- **Hue rotation**: the hue channel of the HSV color space is shifted by \( h \) degrees (\( h \in \{-18.0, +18.0\} \)).

- **Gray world**: the gray world white balancing algorithm is applied to the image. This algorithm makes the assumption that the images, on average, are gray and the value of the illuminant (in each channel) could be modeled as a constant factor multiplied to the channel. For this reason, in order to remove the variation due to the illuminant, the algorithm normalizes each channel of the image with the average value of the pixels [34].

- **Max RGB**: the max RGB white balancing algorithm is applied. This algorithm relies on the fact that the brightest pixel in each channel should be considered as white. Each channel is divided by the highest value [35].

- **Median Filter**: replaces the value of a pixel with the median of the values of its neighbours. The dimension \( k \) of the filter’s window was tune to 3 obtaining a \( 3 \times 3 \) filter [36].

- **Gaussian Blur Filter**: blurs the input image with a Gaussian filter with standard deviation \( \sigma \) [36]. The parameter \( \sigma \) was set to 2.

- **Sharpening Filter**: increases the sharpness of intensity transitions in an image. The sharpening filter that we considered is the unsharp masking, defined by the kernel \( F \):
  \[
  F = \begin{bmatrix}
  -0.125 & -0.125 & -0.125 \\
  -0.125 & 2 & -0.125 \\
  -0.125 & -0.125 & -0.125 
  \end{bmatrix}.
  \]

- **Edge Enhancement Filter**: highlights the edges in the image.
  \[
  F = \begin{bmatrix}
  -0.5 & -0.5 & -0.5 \\
  -0.5 & 5 & -0.5 \\
  -0.5 & -0.5 & -0.5 
  \end{bmatrix}.
  \]

- **Detail Filter**: the main purpose of this filter is to enhance the details of the subjects/objects present in the image.
  \[
  F = \begin{bmatrix}
  0 & -0.17 & 0 \\
  -0.17 & 1.67 & -0.17 \\
  0 & -0.17 & 0 
  \end{bmatrix}.
  \]

- **Smoothing Filter**: replaces the value of the pixels with the average of their neighbors in the filter window [36].
  \[
  F = \begin{bmatrix}
  0.077 & 0.077 & 0.077 \\
  0.077 & 0.385 & 0.077 \\
  0.077 & 0.077 & 0.077 
  \end{bmatrix}.
  \]

- **STOP**: terminates the enhancement procedure.
Table 1: Comparison of state-of-the-art enhancement methods on the LoL dataset.

| Method       | PSNR | SSIM | ∆E |
|--------------|------|------|----|
| RRDNet [14]  | 11.36| 0.54 | 32.63|
| DSLR [41]    | 15.23| 0.68 | 23.47|
| ExCNet [19]  | 15.80| 0.67 | 21.47|
| Zero-DCE [22]| 16.15| 0.70 | 21.99|
| RetinexNet [33]| 17.16| 0.72 | 18.31|
| MBLLEN [10]  | 17.38| 0.75 | 20.48|
| DeepUPE [11] | 17.87| 0.73 | 20.56|
| KinD [12]    | 18.27| 0.83 | 15.02|
| KinD++ [12]  | 18.75| 0.82 | 15.13|
| EnlightenGAN [3]| 18.82| 0.78 | 15.06|
| DPED [4]     | 19.71| 0.81 | 14.81|
| HDRNet [42]  | 20.14| 0.83 | 14.26|
| RetinexDec [13]| 20.23| **0.84**| 13.10|
| RUAS [15]    | 18.22| 0.72 | -   |
| **TreEnhance**|**21.96**|0.81|**10.82**|

The total number of editing operations is 38 (eight brightness adjustments, eight contrast adjustments, eight gamma corrections, two saturation adjustments, two hue rotations, two white balancing algorithms, seven spatial filters and one stop). The parameters $\delta, \beta, \gamma, s, h, \sigma, k$ have been chosen to make it possible to enhance the images within a reasonable amount of steps. In the case in which, to save memory, TreEnhance is applied to resampled images, the parameters of the spatial operators need to be adjusted when finally applied to the original high resolution images.

The parameters of the training procedure described in Section 3 have been selected on the basis of preliminary experiments, also taking into account the cost in terms of computational resources. In each generation phase we randomly selected 100 training images and for each one we applied 10 000 MCTS iterations to grow the corresponding tree (we set to ten the maximum tree depth). The resolution of images was $600 \times 400$ pixels and standard data augmentation techniques (cropping, flipping and resizing) were used to improve the diversity of the training set. Ten $(x, r(x), p(x, \cdot))$ triplets were sampled from each tree and added to the training set. The parameter $\alpha$ in Equation (3) was set to 0.05 and the exploration parameter $c$ in Equation (2) was set to 10.

In each optimization phase the parameters of the neural network have been updated by minimizing the average loss with 160 iterations of the ADAMW [37] optimizer, with learning rate $10^{-3}$ and weight decay $10^{-2}$.

At the end of the training procedure the method has been applied to the test set. For each test image a tree of 1000 nodes has been grown and a final editing sequence was selected with the procedures described in Section 3.3. Table 1 reports the result of TreEnhance on the LOL test set with the Tree Search inference procedure.

We compared the results of our method, with several state of the art architectures in terms of PSNR [38], SSIM [39], $\Delta E$ [40]. TreEnhance was able to outperform all the other methods considered in terms of PSNR and $\Delta E$. Figure [3] presents the results of the application of TreEnhance on the LOL test set. From a qualitative and a quantitative point of view the results obtained are very satisfactory.

4.2 Adobe Five-K Dataset

The second dataset we considered is the Adobe Five-K dataset [43]. It includes 5000 high-resolution images in the device RAW format and in the Adobe RGB space after the work of five different photo editing experts (Expert A, B, C, D, E). We followed the procedure presented by Hu et al. [25] to split the images in a training (4000 images) and a test set (1000 images). As target images we took those edited by Expert C, since it is the most consistent of the five.

For this second dataset, the set of enhancing operators $\mathcal{A}$ was reduced by removing the spatial filters and the white balancing algorithms, exploiting the knowledge that the experts used only those operations that can be obtained with the “color curves” tool of photo editing applications. For the Five-K dataset, the difference between input and ground truth is smaller than that of the LOL dataset. For this reason we changed the parameters of the operators obtaining more fine-grained actions.
Figure 3: Results of the application of TreEnhance on the LOL test set. From left to right: low-light images, TreEnhance output and ground truth images.

- Brightness adjustment: $\delta \in \{-0.05, +0.05\}$;
- Contrast adjustment: $\beta \in \{0.894, 1.414\}$;
- Gamma Correction: $\gamma \in \{0.775, 1.05\}$;
- Saturation adjustment: $s \in \{0.707, 1.414\}$;
- Hue rotation: $h \in \{-9.0, +9.0\}$.

The total number of operations is 29: the STOP action, two saturation adjustments, two hue rotations, brightness, contrast and gamma correction applied either to a single channel or to three channels simultaneously.

The training procedure adopted for this second dataset was the same explained in Section 3. For each generation we considered 100 images from the training set, we downsampled them to $256 \times 256$ pixels and we applied data augmentation as for the LOL dataset. We executed 10 000 MCTS iterations (the maximum depth was set to ten). The value of the parameters $c$ in Equation (2) was set to 4: a value smaller than that used for the LOL dataset to account for the reduced number of operations that requires a smaller degree of exploration. The value of the parameter $\alpha$ in Equation (2) is 0.05, the same used for the LOL dataset. After each generation phase, we trained the network with mini batch gradient descent using the ADAMW optimizer for 70 iterations.

We compared our method (with the Tree Search inference procedure) with several other methods from the state of the art: results are reported in Table 2.

Since the images provided in the Five-K are RAW format images, before training the algorithm, a preprocessing was needed. In order to obtain a fair comparison with the other state of the art methods, we applied these algorithms, using the code provided by the authors, to our preprocessed data. From the comparison we can see that TreEnhance is competitive on this dataset performing very well from a quantitative point of view. Moreover, it reached the best results in terms of LPIPS [44], a metric that measures the perceptual similarity between two images. It also obtained a very good SSIM value. In terms of the pixel-level metrics PSNR and $\Delta E$, other methods obtained better results. This fact depends on the granularity of the set of editing operations considered. To obtain more accurate results at the pixel level, we would need to include more fine-grained operations, which would make significantly longer the sequences found by
Table 2: Comparison of state-of-the-art enhancement methods on the Adobe five-k dataset.

| Method              | LPIPS↓ | PSNR↑ | ΔE↓ | SSIM↑ |
|---------------------|--------|-------|-----|-------|
| Exposure [25]       | 0.16   | 18.74 | 13.57 | 0.81  |
| CycleGan [2]        | 0.16   | 19.38 | 12.95 | 0.78  |
| DCE-Net [5]         | 0.13   | 20.47 | 11.77 | 0.85  |
| DCE-Net-Pooling [5] | 0.15   | 20.33 | 12.13 | 0.85  |
| DaR [23]            | 0.10   | 20.91 | 13.05 | 0.86  |
| Unfiltering [16]    | 0.09   | 21.67 | 10.22 | 0.88  |
| Pix2Pix [1]         | 0.09   | 23.05 | 8.84  | 0.86  |
| HDRNet [42]         | 0.08   | 22.31 | 9.53  | 0.89  |
| Star-DCE [5]        | 0.08   | 23.55 | 8.45  | 0.89  |
| Parametric [17]     | 0.07   | 23.20 | 8.52  | 0.90  |
| TreEnhance          | 0.06   | 21.24 | 11.25 | 0.89  |

Moreover, when the sequences of operators becomes to long, the level of explainability of the model is reduced. In Figure 4, we compared the images obtained using TreEnhance and other state-of-the-art methods to enhance a sample from the Five-K test set.

To better understand the differences among the images produced by the methods in the comparison, we organized a subjective evaluation test. We asked ten participants (all with some experience in the field) to select the best looking among seven versions of the same image. This selection process was repeated for 100 different images from the Five-K test set. The compared methods were TreEnhance, Pix2Pix, CycleGAN, Star-DCE, DaR, Exposure and HDRNet. The participants used a system that shows the seven alternatives in random order that, without any reference to the methods. The order used to present the 100 images was also different for each user.

From the results shown in Figure 5, TreEnhance was the most selected method with the 23.63% of the users preferences followed by HDRNet (22.53%) and Star-DCE (20.54%). The other methods showed much lower preferences: Pix2Pix 15.95%, Exposure 7.48%, DaR 5.58% and CycleGAN 4.29%.

These results confirm the ability of TreEnhance of producing high quality images from a perceptive point of view even if the pixel-level metrics are not the best one. This is confirmed by the best value of the perceptual LPIPS metric. An informal interview after the test pointed out that for most users there were three methods that were clearly superior than the others. These methods, were very stable in the choice of the tonality and their output was free from any artifacts, a very negative element for the others. The users considered the best three methods as substantially equivalent.

From the performed user study, TreEnhance is very competitive respect several methods for image enhancement. Moreover is able to produce images similar to modern state-of-the-art methods.

### 5 Discussion

The results obtained by TreEnhance in the two different tasks are satisfactory from both a quantitative and a qualitative point of view. The high quality output images are obtained with the application of basic image editing operations, making the process easy to explain. During the inference phase TreEnhance evaluates several candidate sequences of operators but at the end only the most promising is selected and applied to the input image.

In this section we discuss the memory and time requirements of TreEnhance, and we analyze more in detail its qualities.

#### 5.1 Memory and Time Requirements

The generation phase of the training procedure is the most demanding step of TreEnhance in terms of memory. This depends on the number of MCTS steps, which also corresponds to the number of nodes in the trees. With respect to other state-of-the-art methods, TreEnhance does not impose any limitations on the resolution of the input images. However, to save memory (and therefore to allow the growth of larger trees) it is possible to work with resampled images. Once the sequence of editing operations has been found on a low-resolution version of the input image, it can be applied to the original high-resolution version. In this case the parameters of spatial operators must be suitably adjusted. Figure 6 shows some examples of the application of TreEnhance on high resolution images.

We observed that a good degree of exploration should be ensured in the initial phases of the training procedure. In fact, at the beginning it tends to look like a breadth-first search where many nodes at the same level are explored until
Figure 4: Examples of test images in the Five-K dataset processed by methods in the state of the art. The left most column shows the input image and the right most the ground truth.
Figure 5: Results of the user study on the 100 images from the Five-K test set.

Figure 6: Examples of the application of TreEnhance to high resolution test images from the Five-k and LOL datasets.
M. Cotogni and C. Cusano

Table 3: Comparison of TreEnhance inference strategies on the two dataset considered.

| Search strategy | Five-K PSNR | ∆E | SSIM | LOL PSNR | ∆E | SSIM |
|-----------------|-------------|-----|------|-----------|-----|------|
| Tree            | 21.24       | 11.25 | 0.89 | 21.96     | 10.82 | 0.81 |
| Policy-based    | 20.06       | 12.44 | 0.84 | 21.86     | 10.99 | 0.80 |
| Guided          | 25.54       | 7.18  | 0.92 | 24.34     | 8.44  | 0.84 |

Figure 7: Accuracy obtained by the TreEnhance on the LOL test set varying the number of MCTS search steps (left) and the exploration/exploitation coefficient (right).

As explained in Section 3, only a random subset of the training set is involved in each generation. This subset is split in batches (four in our experiments) in order to make use of all the available memory. To speed up the generation phase, the trees corresponding to each one of the images in the batch grow in parallel. In our experiments we used an Intel® Core™ i7-8700 CPU @ 3.20GHz with 48 GB of memory. For the evaluation step of the generation phase and for the optimization phase we used a NVIDIA RTX™ 3080Ti graphic card with 12 GB of memory.

5.2 Tree-based and Policy-based Enhancement

As explained in Section 3, TreEnhance can be used in the inference phase in two different ways: by building a tree for each input image or by directly using the neural network as a policy. In Table 3 we compared the two different inference strategies on the two test sets. Tree Search enhancement is the most accurate strategy since it considers a wide range of alternatives. Policy-based Search strategy acts greedy by taking each time the most promising operation. For this reason it is faster than the tree-based approach and uses less memory but is slightly less accurate.

The additional complexity of the Tree Search strategy can be modulated by limiting the number of MCTS steps during the inference phase. Figure 7(left) shows that the accuracy of the method increases with the number of steps. Another parameter that can be adjusted during the search is the coefficient $c$ in Equation 2. From Figure 7(right) it seems that during inference a large value of $c$ is beneficial.

5.3 Guided Search

As a further experiment we verified if TreEnhance can be used as a tool for “reverse engineering” the work of a professional photo editor. To do so, we modified TreEnhance by using the ground truth during the inference phase (Equation 3). This way the MCTS search will retrieve the sequence of operations that can be used to obtain a retouched version of the image from the input. We called this third strategy “Guided Search”.

We reported the results obtained in Table 3. The accuracy obtained in this scenario is very high confirming the effectiveness of MCTS in finding good sequences of image editing operations. This experiment suggests that this strategy could be included in an educational software, like a puzzle to be solved, to teach beginners how to enhance a low-quality image and what kind operators they should use to obtain a given result.
5.4 Analysis of the Sequences

The sequences produced by the three inference strategies are correlated. In fact, as it is shown in Figure 8, all the strategies favor global enhancing operators. In particular, the most selected operators are those that increase the brightness, the gamma, and the contrast. This depends on the fact that the original LOL test images are low light ones and need to be brightened. Operations that work on single channels are more prone to be selected when the current image is close to the target one.

Tree-based and policy-based strategies tend to over-select the gamma correction action all over the tree channels with respect to the guided search (action 6). Guided search, in fact, tends to have a more balanced distribution of the operations over the three channels with respect to the other strategies, resulting in higher quality images.

5.5 Artifacts

One of the main features of the generative models for image enhancement, like those based on GANs, is the ability to accurately model color distributions. However, as a drawback they often introduce visible artifacts in their outputs. Figure 9 shows a comparison among two generative methods (Pix2Pix and CycleGAN) and two reinforcement learning methods (TreEnhance and Exposure) on two Five-K test images. In both cases generative methods produced a very good color distribution, but with the presence of visible noise and artifacts. In fact, the value of perceptual metrics such as LPIPS tends to be relatively bad for this kind of methods. Reinforcement learning methods do not introduce any artifacts, but at the cost, sometimes, of less saturated colors.

6 Conclusions

In this paper we presented TreEnhance, a light-weight fully automatic and explainable low-light image enhancement method. TreEnhance shows that it is possible to obtain very good performances in two image enhancement tasks. It finds solutions that are fully explainable, since they can be decomposed as a sequence of simple image processing operations.

The method can be used for other enhancement tasks just by changing the set of image editing operations. Thanks to the support for multiple search strategies, TreEnhance is suitable for a variety of scenarios with different requirements in terms of constraints on memory and processing time, and with different levels of interaction with the user.

TreEnhance demonstrates how tree-search theory, and in particular MCTS, can be a very useful tool for image processing and we plan to investigate its application to other tasks like image retargeting, super resolution and image retrieval.
Figure 9: Comparison of generative and reinforcement learning methods on two Five-K test images.

References

[1] Phillip Isola, Jun-Yan Zhu, Tinghui Zhou, and Alexei A Efros. Image-to-image translation with conditional adversarial networks. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 1125–1134, 2017.

[2] Jun-Yan Zhu, Taesung Park, Phillip Isola, and Alexei A Efros. Unpaired image-to-image translation using cycle-consistent adversarial networks. In *Proceedings of the IEEE international conference on computer vision*, pages 2223–2232, 2017.

[3] Yifan Jiang, Xinyu Gong, Ding Liu, Yu Cheng, Chen Fang, Xiaohui Shen, Jianchao Yang, Pan Zhou, and Zhangyang Wang. Enlightengan: Deep light enhancement without paired supervision. *IEEE Transactions on Image Processing*, 30:2340–2349, 2021.

[4] Andrey Ignatov, Nikolay Kobyshev, Radu Timofte, Kenneth Vanhoey, and Luc Van Gool. Dslr-quality photos on mobile devices with deep convolutional networks. In *Proceedings of the IEEE International Conference on Computer Vision*, pages 3277–3285, 2017.

[5] Zhaoyang Zhang, Yitong Jiang, Jun Jiang, Xiaogang Wang, Ping Luo, and Jinwei Gu. Star: A structure-aware lightweight transformer for real-time image enhancement. In *Proceedings of the IEEE/CVF International
Conference on Computer Vision, pages 4106–4115, 2021.

[6] Scott M Lundberg and Su-In Lee. A unified approach to interpreting model predictions. Advances in neural information processing systems, 30, 2017.

[7] Ramprasaath R Selvaraju, Michael Cogswell, Abhishek Das, Ramakrishna Vedantam, Devi Parikh, and Dhruv Batra. Grad-cam: Visual explanations from deep networks via gradient-based localization. In Proceedings of the IEEE international conference on computer vision, pages 618–626, 2017.

[8] Marco Tulio Ribeiro, Sameer Singh, and Carlos Guestrin. "why should i trust you?" explaining the predictions of any classifier. In Proceedings of the 22nd ACM SIGKDD international conference on knowledge discovery and data mining, pages 1135–1144, 2016.

[9] Praveen Ravirathinam, Divyam Goel, and J Jennifer Ranjani. C-lienet: A multi-context low-light image enhancement network. IEEE Access, 9:31053–31064, 2021.

[10] Feifan Lv, Feng Lu, Jianhua Wu, and Chongsoon Lim. MBLEEN: Low-light image/video enhancement using CNNs. In BMVC, page 220, 2018.

[11] Ruixing Wang, Qing Zhang, Chi-Wing Fu, Xiaoyong Shen, Wei-Shi Zheng, and Jiaya Jia. Underexposed photo enhancement using deep illumination estimation. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, pages 6849–6857, 2019.

[12] Yonghua Zhang, Jiawan Zhang, and Xiaojie Guo. Kindling the darkness: A practical low-light image enhancer. In Proceedings of the 27th ACM international conference on multimedia, pages 1632–1640, 2019.

[13] Xiaojian Lv, Yujing Sun, Jun Zhang, Feng Jiang, and Shengping Zhang. Low-light image enhancement via deep retinex decomposition and bilateral learning. Signal Processing: Image Communication, 99:116466, 2021.

[14] Anqi Zhu, Lin Zhang, Ying Shen, Yong Ma, Shengjie Zhao, and Yicong Zhou. Zero-shot restoration of underexposed images via robust retinex decomposition. In 2020 IEEE International Conference on Multimedia and Expo (ICME), pages 1–6. IEEE, 2020.

[15] Risheng Liu, Long Ma, Jiaao Zhang, Xin Fan, and Zhongxuan Luo. Retinex-inspired unrolling with cooperative prior architecture search for low-light image enhancement. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, pages 10561–10570, 2021.

[16] Simone Bianco, Claudio Cusano, Flavio Piccoli, and Raimondo Schettini. Artistic photo filter removal using convolutional neural networks. Journal of Electronic Imaging, 27(1):011004, 2017.

[17] Simone Bianco, Claudio Cusano, Flavio Piccoli, and Raimondo Schettini. Learning parametric functions for color image enhancement. In International Workshop on Computational Color Imaging, pages 209–220. Springer, 2019.

[18] Simone Bianco, Claudio Cusano, Flavio Piccoli, and Raimondo Schettini. Personalized image enhancement using neural spline color transforms. IEEE Transactions on Image Processing, 29:6223–6236, 2020.

[19] Lin Zhang, Lijun Zhang, Xiao Liu, Ying Shen, Shaoming Zhang, and Shengjie Zhao. Zero-shot restoration of back-lit images using deep internal learning. In Proceedings of the 27th ACM International Conference on Multimedia, pages 1623–1631, 2019.

[20] Yoav Chai, Raja Giryes, and Lior Wolf. Supervised and unsupervised learning of parameterized color enhancement. In Proceedings of the IEEE/CVF Winter Conference on Applications of Computer Vision, pages 992–1000, 2020.

[21] Hanul Kim, Su-Min Choi, Chang-Su Kim, and Yeong Jun Koh. Representative color transform for image enhancement. In Proceedings of the IEEE/CVF International Conference on Computer Vision, pages 4459–4468, 2021.

[22] Chunle Guo, Chongyi Li, Jichang Guo, Chen Change Loy, Junhui Hou, Sam Kwong, and Runmin Cong. Zero-reference deep curve estimation for low-light image enhancement. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, pages 1780–1789, 2020.

[23] Jongchan Park, Joon-Young Lee, Donggeun Yoo, and In So Kweon. Distort-and-recover: Color enhancement using deep reinforcement learning. In Proceedings of the IEEE conference on computer vision and pattern recognition, pages 5928–5936, 2018.

[24] Huan Yang, Baoyuan Wang, Noranart Vesdapunt, Minyi Guo, and Sing Bing Kang. Personalized exposure control using adaptive metering and reinforcement learning. IEEE transactions on visualization and computer graphics, 25(10):2953–2968, 2018.

[25] Yuanming Hu, Hao He, Chenxi Xu, Baoyuan Wang, and Stephen Lin. Exposure: A white-box photo post-processing framework. ACM Transactions on Graphics (TOG), 37(2):1–17, 2018.
[26] Ke Yu, Chao Dong, Liang Lin, and Chen Change Loy. Crafting a toolchain for image restoration by deep reinforcement learning. In Proceedings of the IEEE conference on computer vision and pattern recognition, pages 2443–2452, 2018.

[27] Ryoosuke Furuta, Naoto Inoue, and Toshikiko Yamasaki. PixelRL: Fully convolutional network with reinforcement learning for image processing. IEEE Transactions on Multimedia, 22(7):1704–1719, 2019.

[28] Runsheng Yu, Wenyu Liu, Yasen Zhang, Zhi Qu, Deli Zhao, and Bo Zhang. Deepeexposure: Learning to expose photos with asynchronously reinforced adversarial learning. In Proceedings of the 32nd International Conference on Neural Information Processing Systems, pages 2153–2163, 2018.

[29] Cameron B Browne, Edward Powley, Daniel Whitehouse, Simon M Lucas, Peter I Cowling, Philipp Rohlfshagen, Stephen Tavener, Diego Perez, Spyridon Samothrakis, and Simon Colton. A survey of monte carlo tree search methods. IEEE Transactions on Computational Intelligence and AI in games, 4(1):1–43, 2012.

[30] David Silver, Thomas Hubert, Julian Schrittwieser, Ioannis Antonoglou, Matthew Lai, Arthur Guez, Marc Lanctot, Laurent Sifre, Dharshan Kumaran, Thore Graepel, et al. Mastering chess and shogi by self-play with a general reinforcement learning algorithm. arXiv preprint arXiv:1712.01815, 2017.

[31] Sébastien Bubeck and Nicolo Cesa-Bianchi. Regret analysis of stochastic and nonstochastic multi-armed bandit problems. arXiv preprint arXiv:1204.5721, 2012.

[32] Kaiming He, Xiangyu Zhang, Shaoqing Ren, and Jian Sun. Deep residual learning for image recognition. In Proceedings of the IEEE conference on computer vision and pattern recognition, pages 770–778, 2016.

[33] Runsheng Yu, Wenyu Liu, Yasen Zhang, Zhi Qu, Deli Zhao, and Bo Zhang. Deepeexposure: Learning to expose photos with asynchronously reinforced adversarial learning. In Proceedings of the 32nd International Conference on Neural Information Processing Systems, pages 2153–2163, 2018.

[34] Edmund Y Lam. Combining gray world and retinex theory for automatic white balance in digital photography. In Proceedings of the Ninth International Symposium on Consumer Electronics, 2005.(ISCE 2005), pages 134–139. IEEE, 2005.

[35] Edwin H Land. The retinex theory of color vision. Scientific american, 237(6):108–129, 1977.

[36] Rafael C. Gonzalez and Richard E. Woods. Digital Image Processing (3rd Edition). Prentice-Hall, Inc., USA, 2006.

[37] Ilya Loshchilov and Frank Hutter. Decoupled weight decay regularization. arXiv preprint arXiv:1711.05101, 2017.

[38] Onur Kele¸s, M. Akın Yılmaz, A. Murat Tekalp, Cansu Korkmaz, and Zafer Dogan. On the computation of psnr for a set of images or video, 2021.

[39] Alain Hore and Djemel Ziou. Image quality metrics: Psnr vs. ssim. In 2010 20th international conference on pattern recognition, pages 2366–2369. IEEE, 2010.

[40] T.M. Goodman. Colour Design: Theories and Applications. Elsevier, 2012.

[41] Seokjae Lim and Wonjun Kim. Dslr: Deep stacked laplacian restorer for low-light image enhancement. IEEE Transactions on Multimedia, 2020.

[42] Michaël Gharbi, Jiawen Chen, Jonathan T Barron, Samuel W Hasinoff, and Frédo Durand. Deep bilateral learning for real-time image enhancement. ACM Transactions on Graphics (TOG), 36(4):1–12, 2017.

[43] Vladimir Bychkovsky, Sylvain Paris, Eric Chan, and Frédo 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, 2011.

[44] Richard Zhang, Phillip Isola, Alexei A Efros, Eli Shechtman, and Oliver Wang. The unreasonable effectiveness of deep features as a perceptual metric. In CVPR, 2018.