Article

Towards Automatic Image Enhancement with Genetic Programming and Machine Learning

João Correia [1], **Nereida Rodriguez-Fernandez** [2], **Leonardo Vieira** [1], **Juan Romero** [2,*] and **Penousal Machado** [1]

1   CISUC, Department of Informatics Engineering, University of Coimbra, 3030 Coimbra, Portugal; jncor@dei.uc.pt (J.C.); lmavieira@student.dei.uc.pt (L.V.); machado@dei.uc.pt (P.M.)
2   Department of Computer Science and Information Technologies, Faculty of Communication Science, University of A Coruña, 15071 A Coruña, Spain; nereida.rodriguezf@udc.es
*   Correspondence: jj@udc.es

Abstract: Image Enhancement (IE) is an image processing procedure in which the image's original information is improved, highlighting specific features to ease post-processing analyses by a human or machine. State-of-the-art image enhancement pipelines apply solutions to fixed and static constraints to solve specific issues in isolation. In this work, an IE system for image marketing is proposed, more precisely, real estate marketing, where the objective is to enhance the commercial appeal of the images, while maintaining a level of realism and similarity with the original image. This work proposes a generic image enhancement pipeline that combines state-of-the-art image processing filters, Machine Learning methods, and Evolutionary approaches, such as Genetic Programming (GP), to create a dynamic framework for Image Enhancement. The GP-based system is trained to optimize 4 metrics: Neural Image Assessment (NIMA) technical and BRISQUE, which evaluate the technical quality of the images; and NIMA aesthetics and PhotoILike, that evaluate the commercial attractiveness. It is shown that the GP model was able to find the best image quality enhancement (0.97 NIMA Aesthetics), while maintaining a high level of similarity with the original images (Structural Similarity Index Measure (SSIM) of 0.88). The framework has better performance according to the image quality metrics than the off-the-shelf image enhancement tool and the framework's isolated parts.

Keywords: genetic programming; image enhancement; image filters; computer vision

1. Introduction

IE is an essential and ever-demanding branch of image processing and computer vision that allows visual improvement to images by manipulating their attributes. It can be used to alter an image in several different ways, for instance, by highlighting a specific feature to ease post-processing analyses by a human or a machine or by increasing its human perceived aesthetic. Each image has a set of attributes, such as size, color space, contrast, brightness, saturation, distortions, artifacts, noise, format, etc., that define how we, and different computer software solutions, perceive it. All these features are not isolated but, rather, interact with each other. The image format and compression, for instance, is pivotal to every other aspect of the image, as different formats allow for distinct color and compression properties that can, in turn, account for distinct visual characteristics. These attributes are often not well balanced between each other nor optimized for the image context and may cause a diverse range of deformations in the image quality.

Moreover, IE is useful for different tasks and fields that benefit from pre-processing an image to improve human or machine perception. Such fields are medical imagery, space imagery, bio-metric, photography, and video editing, among others [1–3]. Aesthetic improvement via IE is also one of the most popular uses of these techniques. A great example of such use of IE is the post-processing that is applied to an image before introducing it in a context where image appeal is crucial, for instance, in advertisements. Today, many consumers choose their purchases via the Internet, with countless stimuli and options.

On this basis, it is not surprising that more and more decisions are made based on the aesthetic value of products [4] and the images in their advertisements. From a commercial perspective, it is interesting to note that advertisers spend around US$ 13 billion annually on banner ads [5]. Intending to improve images' click-through rate (CTR), advertisers rely on image filter applications to adjust them. In this context, we decided to focus on a sector that advertises high economic value products and that the majority of the population consumes at some point in their lives: real estate. Currently, a large part of the real estate business is on the Internet, real estate portals, and the agencies' websites. In most cases, the first step to selling a real estate property, even before a visit, is to show pictures of it. If those photographs are not attractive to the client, it is probable that the visit is lost and, consequently, the sale itself.

Recent advances in IE include image processing or Machine Learning approaches to solve specific problems [6]. Due to the vast set of applications and contexts, researchers rely on a set of image filters [7] or generative Machine Learning models, such as Generative Adversarial Networks (GAN) [8]. However, there is no simple solution that fits all IE problems. Specifically, real estate image marketing presents a complex IE scenario, where the overall aesthetics and technical aspects of the image must be adjusted to deliver a realistic, credible, and attractive result for customers. Real estate images are diverse in terms of content and the conditions in which they are taken (resolutions, perspectives, and lighting conditions, among others). Some photos are taken with professional-level cameras, while others are taken with inexpensive smartphones. There are also often indoor and outdoor photos, day and night, or there might even be a close-up of an ornament or a vase.

Considering all of this, there is space for an automatic tool that combines state-of-the-art approaches and adapts dynamically to different input conditions, a requirement for the real estate marketing problem.

The problem of IE can be formulated as a set of operations performed to the input image that can result in a better output from the perspective of qualitative or quantitative metrics that fulfill the objective. A core idea of the presented work is that GP can be used to leverage the application of a set of state-of-the-art IE operations to create automatically IE pipelines. We resort to ELAINE (EvoLutionAry Image eNhancEment) to handle the generation of the image pipelines [9].

Based on state of the art, Pix2Pix GANs, such as Enlighten-GAN [8], showed promising results on solving several generalized contrast and illumination problems in benchmark datasets. Thus, the framework combines these two approaches into a conditional framework, wherein a first step the GAN attempts to perform generic IE and the GP pipeline can be optimized to subsets of data. These approaches combined offer a more dynamic and adaptable solution, while providing suitable solutions to solve different types of problems.

Through this work, an IE system capable of enhancing the visual aesthetic quality of real estate images is developed. To accomplish this, we relied on different techniques, arranged in an end-to-end layout, to formulate a system that aims to process images with different characteristics properly. More precisely, we aim to combine image processing techniques by using Genetic Programming to create complete image pipelines combined with Machine Learning via GANs into a single framework for IE. This work is being performed while in contact with *PhotoILike*, a Spanish company dedicated to the use of Artificial Intelligence techniques for the aesthetic valuation of real estate images. They contribute by providing valuable tools and data acquired during the experimentation phase.

In summary, this work contributes to the field of study by: (i) performing different IE experiments with Classical image processing filters, Pix2Pix GANs, and a set of filters created with GP; (ii) performing different tests using state-of-the-art *no-reference* Image Quality Assessment (IQA) metrics of image technical quality and aesthetics; (iii) using ELAINE, a Genetic Programming approach, for the generation of image enhancement pipelines; and (iv) developing a system capable of improving the visual aesthetics of real estate images.

The remainder of the document is structured as follows. Section 2 reviews works related to image enhancement. Next, Section 3 describes the proposed approach to development of an adaptive IE system. Section 4 explains the experiments, as well as the setups, and shows and analyzes the results. Afterwards, a discussion on the results is presented in Section 5. Finally, Section 6 presents overall conclusions.

2. Related Works

There are multiple types of IE techniques with different purposes and characteristics. The research around this field revolves around Machine Learning models, computer vision pipelines by applying filters, or both. This section reviews the works related to IE techniques, Classical or using Machine Learning, which are more related to our approach.

IE is a sub-field of computer vision and image processing [10] that aims to develop ways of improving the perception of an image specific feature or general quality, by a person or a computer, in a specific context [6].

Although, by definition, IE can be performed manually, this work will focus on automatic IE achieved algorithmically or via Machine Learning. Automatic IE bears interesting challenges, specifically when trying to manipulate multiple aspects of the image at the same time, because individual features are not independent of each other. Figure 1 proposes a high-level taxonomy divided into three main layers, for IE techniques.
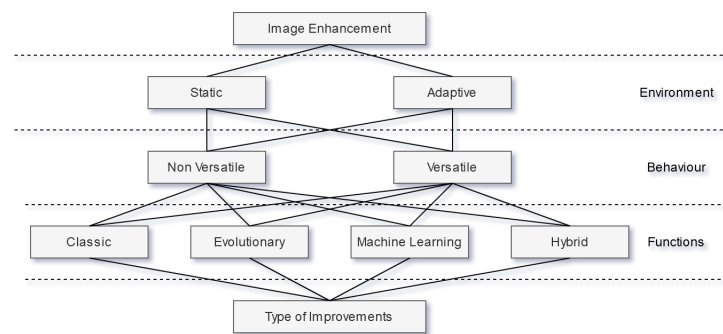

Figure 1. Proposed taxonomy for image enhancement techniques.

The first layer from the top, Environment, divides the context where a technique is inserted in two categories, Static and Adaptive. Static means that, after the deployment of such technique, there are no changes to itself or surrounding components that affect said technique performance. Adaptive, on the other hand, implies the existence of some change to the technique or surrounding components to improve its performance over time. The second layer, Behavior, divides the way a technique reacts to different inputs in two categories, versatile and non-versatile. Versatile means that the technique is capable of properly handling input images with different characteristics, whereas non-versatile techniques are specialized on a single type of image. The third layer, Functions, divides the technical foundation of each technique in 4 categories, Classic, Evolutionary, Machine Learning, and Hybrid. Classic refers to techniques based on mathematical and statistical evaluation and manipulation of the input image in order to improve it. Evolutionary and Machine Learning refer to techniques based on Evolutionary computation methods, which are inspired by biological evolution mechanisms, and Machine Learning methods, respectively. The Hybrid category encompasses techniques that take advantage of different technical foundations that belong in different categories. Finally, all these layers converge on the type of improvement trying to be achieved. Common approaches seek to improve images aesthetically by manipulating their brightness and contrast. However, other manipulations are also possible, such as image super-resolution, automatic selective crop, and artifact removal. Next, we will discuss what these mean and present some research examples for each one.

As we are interested in enhancing the perceived aesthetics for humans, we should focus on aspects of the image that most appeal to the human eye. Perceptual IE is a sub-category of IE that includes models that take in consideration the *"Principles of Human Visual System (HVS)"* to better enhance images [11]. A few of the most important principles that have played important roles in existing methods are *contrast sensitivity function* that maps how the human eye reacts to different levels of contrast in different situations; *multi-scale and multi-orientation decomposition*, which explains how the human eye adjusts objects at various scales, orientations and distances; and visual masking, that refers to the phenomenon that occurs when an image appears to have lower contrast or brightness when surrounded by other stronger stimulus called the mask [11].

Concerning Classical IE techniques, Zhuo et al., in Reference [12], proposed a noise reduction pipeline that starts with a hybrid camera system that takes a near-infrared flash photo (N) at the same time as the normal photo (V) is taken. In 2016, Talebi et al. proposed [13] a novel way of improving an image detail and contrast by expanding on Laplacian operators of edge-aware filter kernels in order to develop a robust method capable of enhancing the details of an image without compromising its overall quality by boosting noise and artifacts, which is a major concern when working with edge-aware filters. Moreover, Wong et al. proposed in the same year another pipeline approach that tries to bridge the problem where approaches that are only based on intensity enhancement may produce artifacts in "over-enhanced" regions, and lack enrichment on color-based features [14]. An experimental phase is conducted with various natural scene images showing that this simple and efficient pipeline is a suitable choice to enhance color images. For their part, Wencheng et al. recently proposed an IE pipeline that aims to improve the overall brightness and contrast of low-illumination images [7]. The enhancement function takes inspiration from the Weber–Fechner law that suggests that the brightness perceived by the human eye follows approximately a logarithmic function. So, different illumination values would require different adjustments. The result showed significant improvement in brightness and contrast when compared to other Classical approaches, while preserving the details.

Lastly, to finish with the classic techniques, it is worth mentioning the work of Afifi et al., who recently presented, in Reference [15], a novel way of performing white-balance within a camera image processing pipeline. They propose a pipeline that renders multiple *"tiny versions"* of the original image, each with a different white-balance setting. Afifi et al. also recently presented a few other research studies concerning image white-balance and illumination, but by using more advanced computational techniques, such as deep learning [16].

Regarding IE approaches using Machine Learning, a few examples are reviewed below. Zhang et al. proposes a very deep architecture to perform image denoising [17]. As opposed to the presented work by Zhuo et al., where a near-IR camera setup was proposed to tackle the denoising problem, this work presents a novel, very deep Convolutional Neural Network (CNN) architecture that manipulates the final image. Experiments with this model showed that it could remove Gaussian noise with an unknown noise level, which is not possible in traditional discriminative models trained for specific levels. Deng et al. introduced an adversarial learning model called *EnhanceGAN* in Reference [18]. *EnhanceGAN* only requires weak supervision and is based on the GAN framework. However, the generator G does not generate images itself but, rather, learns an operation of transformation. This means that the work is extensible to other types of operators besides the ones presented. Chen et al. took an interesting approach on extremely low light and low signal-to-noise IE [19]. Instead of manipulating the resulting image using Classical techniques, as in the work of Wencheng et al. mentioned above, the proposed model is an end-to-end CNN architecture that receives as input the raw sensor data of the camera, replacing the traditional pipeline, which rarely performs well under the short-exposure scenario presented. Jiang et al. recently introduced another GAN-based model dubbed *Enlighten-GAN*, an unsupervised GAN that can be trained without image pairs [8].

Kang et al. presented an interesting take on IE by introducing a framework that explores the value of the personalization factor in automatic IE [20], where both Classical techniques and Machine Learning models are used, making it an example of a hybrid approach to the problem. During an initial training phase, a database is built by presenting the user to a set of image examples in a simple interface that allows him to select the enhancement he likes the most in a set of enhanced versions of the original image. This database stores the original image features and the enhancement parameters that the user chose. When a new input is received, a pre-processing phase occurs, where white balance and contrast stretch is performed to make slight improvements to bad photos or photos that the system is not trained to handle. The system then searches for the most similar images in the database, and the parameters prompted by the user are used for the enhancement. The adjustments used in this work were *"white balancing via changes in temperature and tint, and contrast manipulation via changes in power and S-curves"*.

In summary, to the best of our knowledge, most of the approaches are primarily non-modular pipelines with fixed parameterization and applied to solve specific issues with the input image. Based on the review, we moved to implement a set of image filters in our approach to provide flexibility to the pipelines that we aim to evolve.

3. The Proposed Approach

The objective of this work is the development of an adaptive IE system, and its application on real estate pictures, as part of the collaboration with the Spanish company *PhotoILike*. We had access to datasets addressed in Section 4.2 and to an objective image qualification software for real estate images. This section will discuss how we approached the development of a solution to the real estate problem.

Looking at the information shown previously in Section 2, the Evolutionary- and Machine Learning-based approaches can produce models that better adapt to different kinds of inputs. Our research also shows that global enhancement, meaning the simultaneous enhancement of multiple features of an image, is a challenging goal to achieve, as the solution space of an IE solution is vast. Common approaches to automatic IE focus on individual and more straightforward tasks, such as color balancing, in order to prevent over-enhancement where an image becomes too far-off from the original one, losing important context-sensitive information.

3.1. Proposed Pipeline Architecture

Having this in mind, possible architectures were explored from all three methods, Classical, Machine Learning, and Evolutionary, while maintaining the main focus on an end-to-end layout. To undertake the adaptive aspect of our desired solution, i.e., the capability of performing differently according to the input, a pipeline composed by both an *online* and *offline* section, was idealized. An overview of the proposed pipeline architecture is demonstrated in Figure 2, and further detail of both *online* and *offline* modules is given in the next sections.

In order to have a better understanding of our results, we performed statistical work to infer how our solution stands against one other solution that tries to achieve the same objective. It is important to note that this architecture is very modular as it allows for the replacement of almost any component. This facilitates any change to the pipeline to improve it or tailor it to a distinct context.
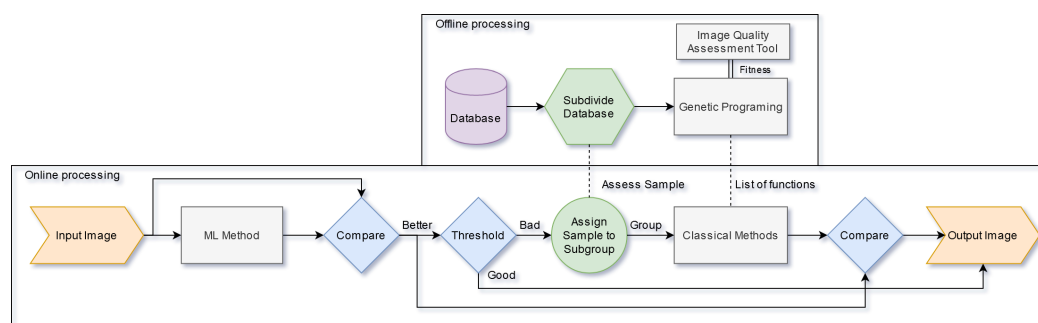
Figure 2. Schematic of the proposed end-to-end IE pipeline architecture. The pipeline is divided in two main sections, *online* and *offline*, containing the components computed in real-time over the input image, and the components independent from the input sample, respectively.

3.2. Offline Pipeline Section

The offline section is responsible for all the processing that is not possible to be executed without heavily compromising the pipeline's time efficiency. Thus, the offline operations can be continuously performed and updated, consequently altering the pipeline performance.

In the scenario presented, any Evolutionary computation component had to be performed offline, considering the significant time consumption inherent to such techniques and the requirements we must meet on the final solution. The objective is to evolve the conditional set of implemented images filters to be applied to each input sample, utilizing the score of IQA tools to guide evolution. Throughout the development, we gave special attention to GP techniques as the Evolutionary component since these techniques are designed to evolve programs as solutions, which is similar to our required solution. The implemented Classical functions were simple functions that altered the image in a specific way. We implemented functions that manipulated image color balance, contrast, saturation, brightness, noise, and sharpness. GP is usually utilized to optimize specific problems, meaning, in the IE context, the improvement of a single image at a time. Nonetheless, a general solution capable of showing improvement on a diverse set of images was sought. To achieve this, in each evaluation, each individual was evaluated on multiple different image groups from the database. The input is assigned to one of the groups during the online processing, and the corresponding set of functions is applied. More details on this will be presented in Section 3.4.

3.3. Online Pipeline Section

On the other hand, the online section is responsible for performing all the operations that are applied to the specific input image. It is allocated to quicker tasks, such as using trained Machine Learning models and performing image processing Classical filters to the input image. We also propose the introduction of a *comparison phase* that guarantees the best output possible out of the multiple alterations computed during the pipeline processing. The comparison phase introduces some overhead in the pipeline. Moreover, we also propose the introduction of a *threshold phase* that filters images in which quality can presumably be even more enhanced, from those considered to have such quality that is attempting to improve them even further may degrade their quality.

Regarding the Machine Learning component, we were inclined to explore adversarial and convolutional networks as they have continuously indicated to perform well on image manipulation tasks. We focused on a recent architecture that performs image-to-image translation [21]. This component would help to subtly improve the overall image quality; thus, it is placed at the beginning of the pipeline as a way to provide a *head start* for the following methods. More details on this will be presented in Section 3.5.

Pseudo-Code in Algorithm 1 sums up the pipeline's online decision process, where *IQA* represents a IQA function that returns a score based on the input image's quality; *ML*

represents a Machine Learning component in the pipeline; and *EC* represents the set of Classical functions applied to an image that is returned by the Evolutionary component.

Algorithm 1: Overview of the pipeline's online decision process

---

    Input: Input Image
    Output: Output Enhanced Image
    I = Input Image
    if *IQA(I) > IQA(ML(I))* **then**
        if *IQA(I) > Quality Threshold* **then**
          | **return** I
        end
        if *IQA(I) > IQA(EC(I))* **then**
          | **return** I
        else
          | **return** EC(I)
        end
    else
        I = ML(I)
        if *IQA(I) > Quality Threshold* **then**
          | **return** I
        end
        if *IQA(I) > IQA(EC(I))* **then**
          | **return** I
        else
          | **return** EC(I)
        end
    end

3.4. Genetic Programming Engine—ELAINE

The analysis of the related work and preliminary work with the implemented filters showed that individually the filters can perform well under the right conditions but may struggle with versatility, i.e., input images under different conditions may require adjustments on parameters. Moreover, it is clear that applying different filters sequentially can produce unique results and that slight adjustments in the order of the filters may cause significant changes to the output. Furthermore, some images may require the application of one or two filters, depending on different conditions. The need for more filters led us to conclude that, for automatic IE, we need to find a generic pipeline for the application of image filters suitable to different input images. Contemplating these insights, a way to automatically generate pipelines that compute image filters applied to the input images was developed. We opted for GP using a tree representation since the problem inherently can be viewed as a program, a succession of steps and decisions of what filters should be applied on the input image and by what order. GP provides us with a representation suitable for exploring solutions in a structured and flexible way, with variation operators well-defined and adaptable to our problem. Thus, the system developed was later named ELAINE (EvoLutionAry Image eNhancEment), a GP engine that allows the evolution of image enhancement pipelines based on pre-defined image filters.

Making use of GP requires the definition of the primitives and terminals that will be available to the population during the evolution. The scenario presented wants to evolve a sequence of filter functions that generally receive at least an image and a numeric value as input, i.e., the filter's parameters. We defined a primitive set containing all the seven Classical functions previously implemented, a terminal set containing the input image, and *ephemeral constants* ranging from -1 to 1. Each function then mapped the defined range in order to adapt it to the desired magnitude. Each function parameter range was empirically defined so that the function provided acceptable results. Additionally, we

introduced an "*if-then-else*" function that, depending on the boolean value of a condition, returns the output of the "*then* tree" or the "*else* tree", allowing the same program solution to behave differently according to the input characteristics.

Since one of the primitives is the whole image, it required values that could be used for comparison to make conditionals. To make this possible, a set of "conditional functions" were introduced. Thus, we added to the primitive set five functions that extracted relevant features from the image. The features are image-related features that capture characteristics of the perceived quality of the image: noise, contrast, saturation, brightness, and sharpness. To extract the noise, we used the work proposed in Reference [22] to estimate the images Gaussian noise quickly. For contrast, the *RMS* contrast [23] was calculated, meaning standard deviation of pixel intensities. For saturation, we averaged the pixels' intensity in the *S* channel of the *HSV* color system. For brightness, the *HSP* color system was used [24], as it grants a brightness value closer to the real human perception when compared to the Luminance (L) channel of the *HSL* or the value (V) channel from the *HSV*. We then averaged the perceived brightness (P) channel to obtain a final value. Finally, for sharpness, we applied a Laplacian filter, calculated the variance of the output, and used that as a sharpness score, as proposed in Reference [25]. All functions were adapted to expect an image as input and an ephemeral constant, which serves as a threshold for that condition. Figure 3 shows a graphical example of a possible individual. The entire implementation was performed using *DEAP* [26] for Python as the base Evolutionary engine.
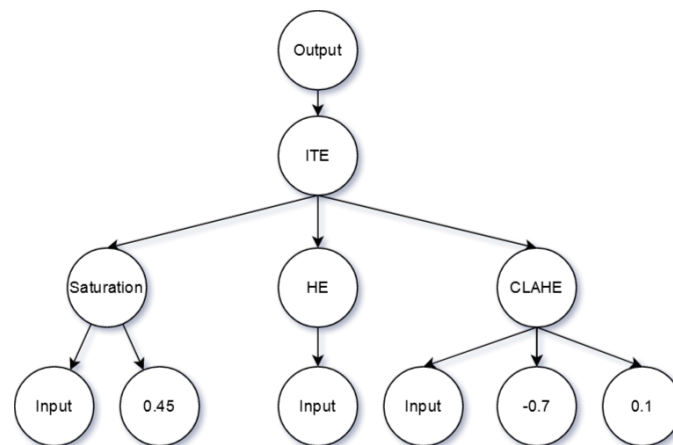

Figure 3. Graphical example of an individual. The numbers represent the ephemeral constants, the *ITE* node represents the *if-then-else* primitive, and the Saturation node represents the conditional function.

Besides the conditional functions and terminals, the function set of ELAINE is also composed of a set of 7 previously reviewed image enhancement filters. The filters focus on five main aspects of IE approaches: contrast adjustment, brightness adjustment, color balance, noise removal, and edge enhancement (also referred to as *sharpening*). The contrast in image processing is the range of intensity values available to an image. The contrast stretching is a point operation method that tries to improve the image contrast by linearly increasing the difference between the maximum intensity value and the minimum intensity value in an image, increasing the contrast level. The work of Bazeille et al. [27] discusses and reports such a filter for IE. Histogram Equalization (HE) is another method that tries to improve an image's quality by manipulating the contrast. It does this by spreading the most common intensity values by the less common ones, increasing the global contrast of an image. This method is highly used, and there are multiple iterations and discussions about its results [28]. Contrast Limited Adaptive Histogram Equalization (CLAHE) is yet another contrast enhancement method and adaptable for different use cases [29], which consists of an iteration of the Adaptive Histogram Equalization (AHE) technique that is an improved version of the regular histogram equalization. Thus, CLAHE improves upon the AHE by clipping the maximum intensity values of each region and redistributing the clipped values

uniformly throughout the histogram before applying the equalization. Gamma Correction (CB) tries to accommodate the fact that the HVS perceives brightness in a non-linear way. This is performed by scaling each pixel brightness from [0–255] to [0–1] and applying an expression to map the original values. Non-local Means Denoising (NLMD) [30], as the name implies, tries to reduce the existing noise in an image. It replaces the value of each pixel in each channel with the average of similar pixels. Unsharp Masking (UM) is an IE technique that sharpens the edges of an image [18]. It does that by subtracting a blurred version from the original image to create a mask. This mask is then applied to the original image, enhancing edges and details. Simplest Color Balance (SCB) was proposed by Limare et al. in Reference [31]. The algorithm tries to remove incorrect color cast by scaling each channel histogram to the complete *0–255* range via affine transform.

Along with the Evolutionary engine, we have created a website that allows a user to test the evolved filters with ELAINE (http://elaine.dei.uc.pt), as shown in Figure 4. The interaction goes as follows: upload an image; select a filter from the preset evolved filters; submit to the system; and see the generated image which may be downloaded. Note that the filters are examples of outputs for the particular sample image. Since the pipelines have conditionals based on image properties, different images with the same filter can have different effects.
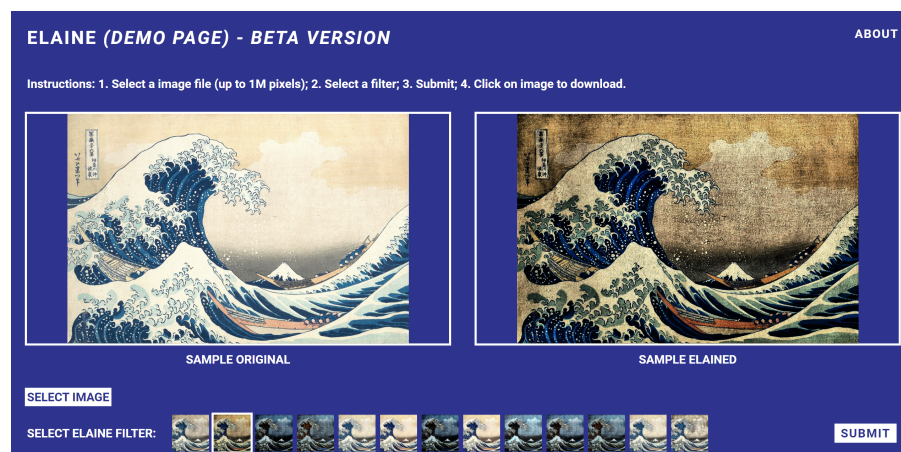

Figure 4. ELAINE's filter tester website.

3.5. Machine Learning Model—Pix2Pix GAN

As reviewed in Section 2, GANs are, most of the time, the "go-to" architecture for current state-of-the-art image processing due to their ability on delivering automatically relevant results for most problems. Therefore, the possibility of applying a similar approach to our problem was explored. Briefly defined, GANs are a generative Machine Learning model proposed by Goodfellow et al., in 2014 [32], taking advantage of the concept of adversarial learning to create generative models capable of creating samples from the same distribution of the training data. This work focused on experimenting with an iteration of this architecture proposed by Isola et al., in Reference [21], referred to as *Pix2Pix GAN*, and is introduced as a *"general-purpose solution to image-to-image translation problems"*. One main change that is provided over the typical GAN is that, instead of learning to map noise input to an output that goes accordingly to the training dataset, it learns to map an input image to a variance of that input image based on the training dataset. Furthermore, *Pix2Pix* suggests the training of the generator not only by adversarial loss but also by direct comparison with the target image using *L1* loss. Both loss metrics in conjunction encourage the generator to produce images in the domain of the target set that are plausible translations for the original input image. Because of this, *Pix2Pix* GAN is deemed as a good solution to be applied to the IE field. This hypothesis was explored by implementing and testing a *Pix2Pix GAN* according to the original paper specifications. *Pix2pix* GAN is a supervised Machine Learning model, and, as such, it requires labeled datasets during training. As we will see

in the experiments, to include the Pix2Pix GAN, it was necessary to train on a set of images suitable to our problem, labeled with enhanced images and original ones.

4. Experiments in Real Estate Image Marketing

According to a study carried out by the Spanish company *PhotoILike*, real estate agencies invest hundreds of euros per month in advertising their listings on portals. In addition, they spend a great deal of time creating these ads, especially in the visual part. *PhotoILike* [33] developed a system that scores and sorts real estate images according to their commercial impact in order to achieve a higher conversion rate. At this point, experiments with IE were considered to increase the value of the system so that it would not only score and sort the images but could also increase the commercial impact of the images to improve the conversion of the ads.

4.1. Image Quality Assessment

Although our work has its focus on IE, it is essential to understand how an image's quality can be measured. He et al. proposes the definition of image quality in three levels: fidelity, perception, and aesthetics [34]. Fidelity is how well the image preserved the original information. Perception is how well the image is perceived according to every part of the HVS. Lastly, the image's aesthetic is the most subjective level because it varies from person to person. It is also the most difficult to measure objectively because *"aesthetics is too nonrepresentational to be characterized using mathematical models"* [34]. Since we are dealing with image quality measurements, which most often is derived from a subjective appreciation, we must have a deterministic and automated way of qualifying an image's quality. To tackle this problem, we made use of 4 distinct *no-reference* IQA tools, where *no-reference* means that the evaluation does not depend on a target nor reference image to evaluate the quality of an input image.

PhotoILike (*PHIL*) is an IQA service provided by the Spanish company with the same name. This service is based on a proprietary AI system [33] that predicts the commercial attractiveness of each image, based on human evaluations. More than a million human evaluations of real estate images were employed to build a dataset for this AI. At the time of this experiment, only authentic real estate images, obtained from real estate portals, were employed in this dataset. This service receives an image and returns a value from 1 to 10, where 1 means the worst attractiveness, and 10 the best attractiveness.

Blind Image Spatial Quality Evaluator [35] (*BRISQUE*) is a *no-reference* IQA tool, proposed by Mittal et al., used in image enhancement contexts [36,37]. As opposed to the previous methods, *BRISQUE* is based on a set of Classical feature extraction procedures that computes a collection of 36 features per image. This tool originally outputs a value between 0 and 100, where 0 represents the best quality, and 100 the worst. However, for the outputs to be in concordance with the previously presented methods, the output was mapped to a 1 to 10 range, where 1 means is the worst score, and 10 the best.

NIMA [38] is a *no-reference* IQA tool based on a deep CNN, proposed by Talebi and Milanfar. The paper highlights how the same architecture, trained with different datasets, leads to state-of-the-art performance predicting technical and aesthetic scores. As the paper states, NIMA technical judgment considers noise, blur, and compression artefacts, among other image features. On the other hand, the NIMA aesthetic evaluation aims to quantify the semantic level characteristics associated with images' emotions and beauty. Both provided models predict the final score as an average of a distribution of scores between 1 and 10, where 1 means the worst score, and 10 the best. Both models, use photographs with high user rating from the *dpchallenge* (https://www.dpchallenge.com/, accessed on 28 September 2021) website.

4.2. Dataset

To assess the quality of this system and to train the GAN, large quantities of images are required. To tackle this problem, we used the dataset provided by the Spanish company

PhotoILike. To examine our results unbiased, we separated in an early stage 10% of the 12,090 images from the rest of the dataset. Thus, 1209 images represent the test set. Figure 5 shows the type of images that appear, that range from landscapes, house fronts, kitchens, bathrooms, and garages. The rest served as a training set to our approaches, i.e., the dataset used by the Evolutionary approach to be selected along the evaluation's Evolutionary process. To have a baseline for our measurements, we examined the original images of this test set. Figure 6 represents four histograms, one histogram for each metric. It is noticeable that the *BRISQUE* tool is much more propitious to giving the maximum possible score than any other metric. In contrast, we see that the *NIMA* aesthetic model only ranges between medium scores, mostly between the 4 and 6 score range. *PHIL* evaluation shows that the test subset has a more spread-out range of scores, ranging from scores of 1 to 9.

Upon conducting the initial experiments with ELAINE [9], we realized that we must track how the image changes after the image enhancement process. Thus, we also resort to two image similarity metrics to track the change between the input and the output: the SSIM and Mean Squared Error (MSE). SSIM extracts 3 key features from an image: Luminance, Contrast, and Structure, with the objective of capturing the degradation of structural information between two images in order to calculate their similarity. MSE is a common metric to quantify the difference in the values of each of the corresponding pixels between the sample and the reference images.

Figure 5. Sample of images from the test dataset of real estate image marketing.
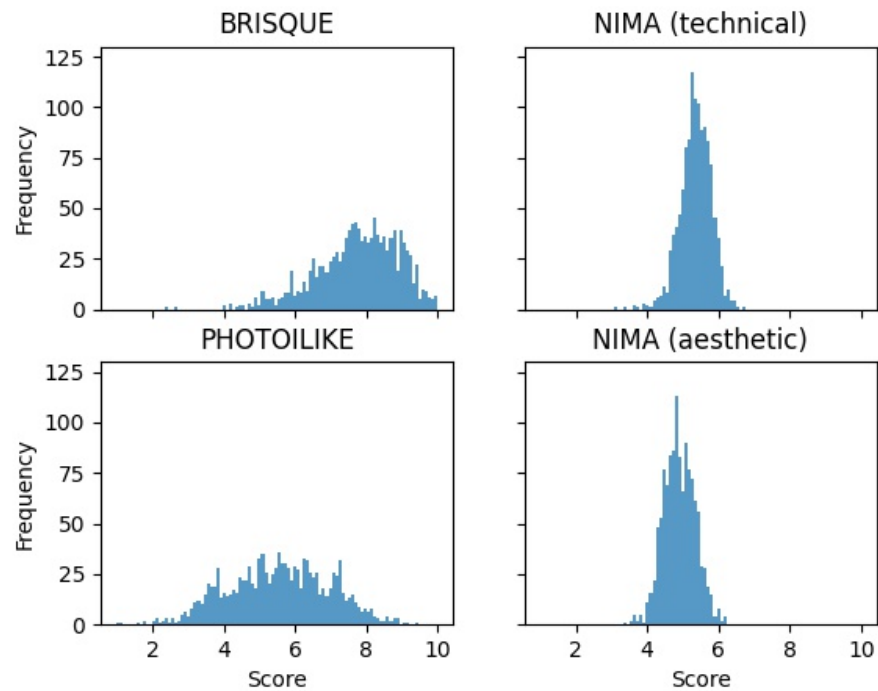
Figure 6. Graphical representation of the test dataset scores, computed by all four *no-reference* IQA tools used during the experiments.

4.3. Experimental Setup

The experimental setup is comprised of the configuration of the ELAINE—the GP engine, the ML approach by training the Pix2Pix GAN, and setting up thresholds for the overall pipeline approach. The entire GP configuration used is presented in Table 1.

The decisions about operators and probabilities were based on previous works, that employed GP to general purpose image generation [39], generation of photorealistic face images [40], and the previous work with images filters [9]. The maximum depth is set to 10 to avoid the creation of really complex filters. We plan to test different parameters in future experiments. However, the evolution of fitness in the experiments carried out shows that the GP system works correctly.

Table 1. Summary of the GP configuration used during the experiments.

Cross-over	One-point
Mutation	Sub-Tree mutation—adds a tree with depth between 0 and 2
Selection	Tournament Size 3
Tree Generation	Ramped half-and-half
Population Size	80
Number of Generations	150
Crossover probability	75%
Mutation probability	5%
Elite size	1
Max Depth	10

As mentioned in Section 3, it is not the objective of these experiments to produce a solution that improves upon a specific image. Instead, we endeavor in searching for a solution as generalist as possible. To achieve this, it is necessary to perform a fundamental change to the typical GP evaluation. Each solution will be evaluated on a set of 10 randomly selected images from the training subset. The average fitness of all images will be consid-

ered the individual's fitness. In addition to that, a new set of 10 images is selected at each generation to prevent overfitting further a specific group of images. For this reason, it is expected a significant variation in fitness from one generation to another. In all performed experiments, the individual with the overall highest fitness was selected as the test subject for validation.

4.4. Experimental Results

This section presents the results from the different approaches carried out to tackle the problem of real estate image marketing. Note that part of the validation of ELAINE, the GP engine, was already completed in a previous work [9]. A conclusion from the previous work was that making the objective only about aesthetics could sometimes yield overly transformed images. However, from the fitness function perspective, the results were correct, and optimizing the function, from a more practical perspective, the transformations to the input image turned them into something unrealistic and, in some cases, not suitable for this real estate image marketing problem.

The experimentation started by evaluating the original dataset with the 4 *no-reference* IQA methods in Table 2. After that, to have a criterion for all the future experiments, we measured the improvement achieved by *One-Click*, the external IE software, on the test dataset using all available metrics. In addition to this, we also conducted the same experience on a list of Classical functions, that we will name CL, manually arranged by us based on expertise and personal tastes, with the default parameterization. This list was as follows: *Contrast Balance, CLAHE, Unsharp Masking, Non-local means denoising, and Contrast Stretching*. These two experiments will be an essential "cornerstone" for the rest of the result analysis. The results are presented in Tables 2 and 3, and in Figure 7. For better readability, *NIMA* models are abbreviated to respective initials, and *PhotoILike* is abbreviated to *PHIL*. While interpreting the results, we will focus more on SSIM, rather than *MSE*, since SSIM gives a much more valid metric of similarity. Nonetheless, *MSE* will still be presented, rounded to units. All of the results are presented in Table 3 and show the metrics improvement over the original images scores. Negative improvement means degradation of quality according to the respective metric. An output example of both methods is presented in Figure 8.

Table 2. Average (μ) and standard deviation (σ) of the original images from the test dataset using 4 *no-reference* metrics. All of the metrics values range from 1 to 10, where 1 means the lowest quality, and 10 highest quality.

	Fit.	*NIMA A.*	*NIMA T.*	*BRISQUE*	*PHIL*
μ	Original	4.91	5.36	7.77	5.56
σ	Original	0.45	0.43	1.15	1.41

In the GAN model, the positive class is the enhanced version of an input image that is considered as a target. Note that, following the original implementation, each image was resized to 256 by 256. The setup employed is that of the authors of Enlighten-GAN [8], while using our dataset. The model uses an Adam optimizer with a batch size of 32 trained 100 epochs with a learning rate of 0.0001 and another 100 epochs with linear learning rate decay until 0. Table 3 (rows GAN A and GAN B) shows the results for both trained models.

We selected the model that showed the best performance in our metrics, GAN A, and performed all the subsequent tests using it. The first test we ran using this model was to apply the GAN over the original image. After the GAN, we applied the Classical functions introduced in Section 3.4. Table 3 shows these results (GAN + CL), and Figure 9 showcases a couple of examples of the combination of GAN and Classical functions.
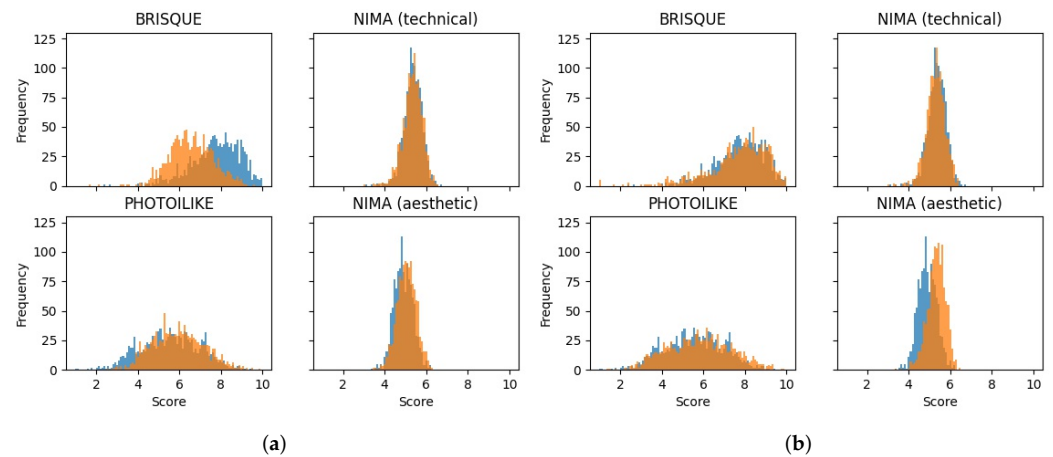
<div align="center">(a)　　　　　　　　　　　　　　　(b)</div>

Figure 7. (**a**) Graphical comparison between the original test dataset (blue) and the results from the *One-Click* (orange), computed by all four *no-reference* IQA tools used during the experiments; (**b**) graphical comparison between the original test dataset (blue) and the results from the list of Classical functions manually selected (orange), computed by all four *no-reference* IQA tools used during the experiments.

<div align="center">(**a**)　　　　　　　　　　　　　(**b**)　　　　　　　　　　　　　(**c**)</div>

Figure 8. (**a**) Original image as input; (**b**) *One-Click* output; (**c**) Classical List output.

Regarding the GP system, the pipeline can be constructed to optimize different metrics. In this work, three configurations, that differ only in the fitness function, were selected that use the following formulations:

- **GP.A**—(N.A.) NIMA Aesthetic model: purely based on aesthetics.
- **GP.B**—(AVG(N.T. + N.A.)): a Average of both NIMA technical and aesthetic models, which takes into account image aesthetics, while balancing the technical part.
- **GP.C**—($N.A.^2 * SSIM$): combination of NIMA Aesthetic model and SSIM similarity metric.

For each selected method, we performed 10 runs due to the stochastic nature of the GP engine. Then, the best solution from each group of 10 experiments was selected. Table 3 contains the results of the 3 chosen solutions. Figure 10 presents examples for each solution.

In the experiments performed, the average image enhancement is higher when the image has a lower original score. This demonstrates that it is much easier to improve upon a bad image than it is for a good one. This is a very rational conclusion since better images have, overall, "less room for improvement", making the enhancement operation harder and, most importantly, in some cases, degrading the quality of the images.

Figure 9. Examples of outputs from the GAN model. On the left, we have the input and original image; on the right, we have the output of after GAN + CL.

In order to test all the methods together, we will use the framework discussed (see Algorithm 1). The framework employs two components, called *compare* component and the *threshold* component.

The *threshold* component is the *Quality Threshold*, referred to in Algorithm 1, that was designed to avoid the over-enhancement of the input image. It limits the transformation of those images that cannot be improved further without degrading the original quality. The NIMA T metric was selected due to the fact that it responds negatively to over-enhancement, based on the previous results. So, while using the *threshold* component, any follow-up

transformation is performed only if the resulting image has its NIMA T value higher than 5.36.

On the other side, we added a *compare* component to the pipeline, where a transformation will be performed only if the resulting *PhotoILike* metric is higher than the original *PhotoILike* metric. We selected *PhotoILike* as it is the closest metric to our use case scenario.

Table 4 presents the results for each GP solution *GP.A*, *GP.B*, and *GP.C*, respectively, using the *compare* component and the *threshold* component, both separately and in conjunction. For simplicity of presentation, we only show the average score of each metric. For comparison purposes, we highlighted in green every cell that scored better than *One-Click*. The performance of the *PhotoILike* metric is critical to us, as this tool was made exclusively for our use case scenario.
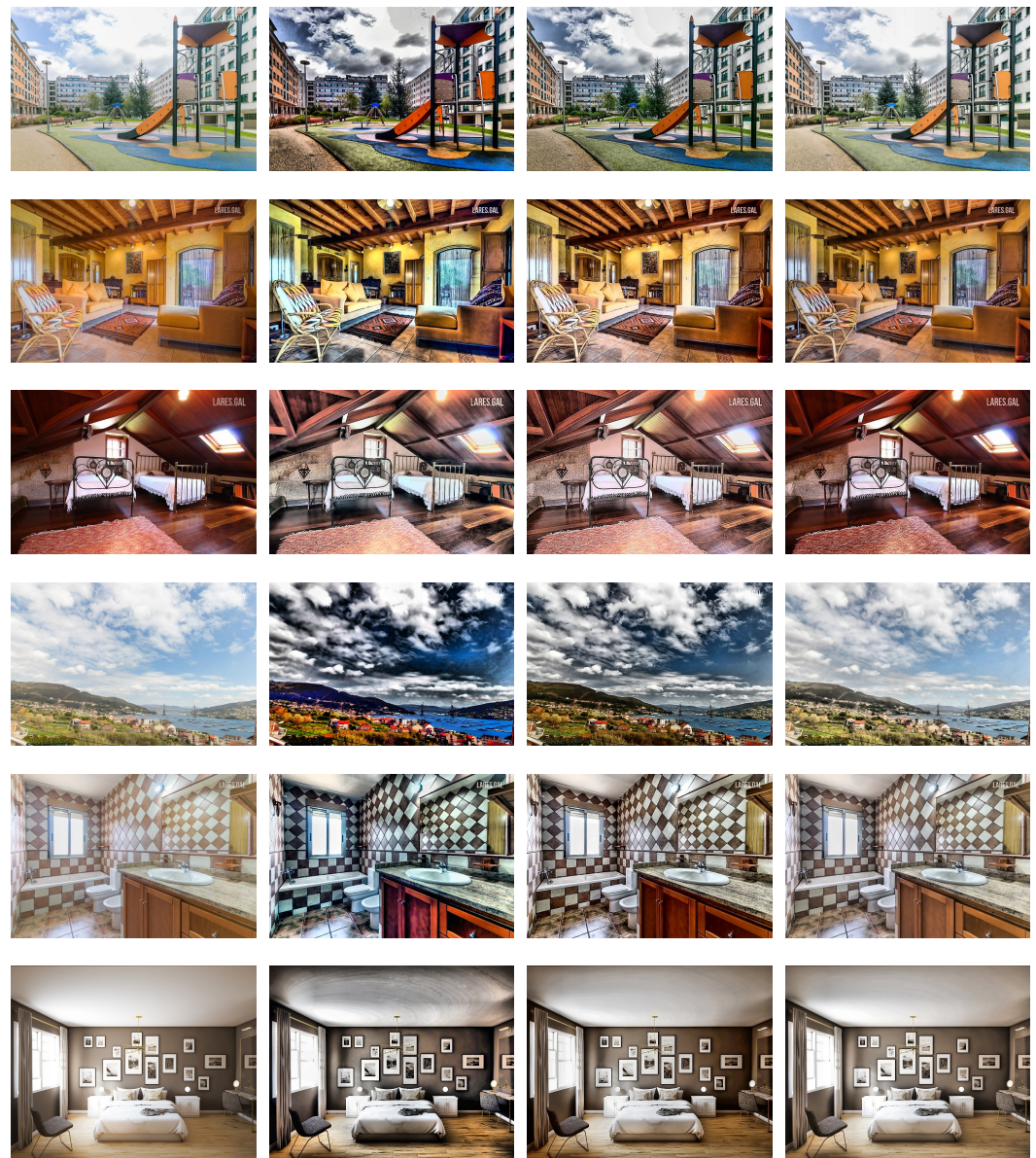
Figure 10. From left to right: Original image; output using solution from GP-A; output using solution from GP-B; output using solution from GP-C.

Overall, we can observe that the GP approach can provide suitable solutions to be included in the pipeline. Results also indicate that the approach surpasses the Pix2Pix GAN trained for the context of this problem. Every evolved pipeline increases the enhancement

of an image more than the One-Click tool, the manually arranged Classical List, and the GAN. It is a testimony to the power of Genetic Programming applied to this particular real-world problem.

The results attained with the automatic image enhancement pipelines solve most pitfalls of only using a system. With thresholds, an input image can be enhanced by multiple models.

Table 3. Average (μ) and standard deviation (σ) of the improvement and similarity using all of the setups and configurations on the test dataset. Highlighted are the cells that presented the best average score in each metric. All of the *no-reference* metrics values range from 1 to 10, where 1 means the lowest quality, and 10 highest quality. The improvement is calculated by subtracting the original score from the the resulting one. SSIM goes from 0 to 1; higher is better. In MSE, lower is better.

	Fit.	NIMA A.	NIMA T.	BRISQUE	PHIL	SSIM	MSE
μ	One-Click	0.16	−0.02	−1.19	0.39	0.80	5250
σ	One-Click	0.25	0.23	0.69	0.69	0.07	4175
μ	Classical List (CL)	0.43	−0.06	−0.08	0.25	0.66	3805
σ	Classical List (CL)	0.31	0.26	1.86	0.64	0.09	2407
μ	GAN A	0.08	−0.04	−0.22	0.15	0.92	677
σ	GAN A	0.16	0.16	0.67	0.30	0.05	317
μ	GAN B	−0.14	−0.14	−0.37	−0.05	0.90	941
σ	GAN B	0.17	0.18	0.92	0.66	0.05	538
μ	GAN + CL	0.51	-0.10	0.13	0.38	0.61	4577
σ	GAN + CL	0.36	0.27	1.52	0.69	0.08	2368
μ	GP.A	0.97	−0.13	0.63	0.44	0.58	8703
σ	GP.A	0.40	0.35	1.43	0.85	0.08	4120
μ	GP.B	0.88	−0.21	0.20	0.51	0.69	5574
σ	GP.B	0.41	0.30	1.51	0.79	0.11	3738
μ	GP.C	0.58	−0.17	−0.53	0.37	0.88	1349
σ	GP.C	0.30	0.24	1.03	0.69	0.04	591

Table 4. The values represent the average improvement of each pipeline configuration, using all of the GP solutions. All of the highlighted cells represent values in which the performance was better than *One-Click*. All of the *no-reference* metrics values range from 1 to 10, where 1 means the lowest quality, and 10 highest quality. The improvement is calculated by subtracting the original score from the the resulting one.

Setup	Components	NIMA A.	NIMA T.	BRISQUE	PHIL	SSIM	MSE
GP.A	Nothing	1.06	−0.24	0.62	0.53	0.54	9611
	Both	0.44	−0.11	0.08	0.47	0.79	3912
	Compare	0.72	−0.17	0.30	0.70	0.68	6221
	Threshold	0.62	−0.15	0.18	0.35	0.70	5869
GP.B	Nothing	0.92	−0.31	0.50	0.50	0.65	6119
	Both	0.40	−0.15	0.03	0.46	0.83	26
	Compare	0.67	−0.23	0.21	0.69	0.74	4145
	Threshold	0.54	−0.19	0.10	0.33	0.76	3806
GP.C	Nothing	0.69	−0.27	−0.36	0.43	0.82	23
	Both	0.31	−0.14	−0.32	0.41	0.90	1135
	Compare	0.49	−0.20	−0.37	0.60	0.87	16
	Threshold	0.41	−0.17	−0.34	0.28	0.86	16

5. Discussion

In this study, experiments have been carried out using four different metrics. Two of the metrics analyze the images from an aesthetic point of view (NIMA A and PHIL) and the other two from a quality point of view (NIMA T and BRISQUE). Table 2 and Figure 6 show how the two metrics with the highest standard deviation are BRISQUE and PHIL. The only metric specially trained for the real estate domain is PHIL. For its part, BRISQUE was trained from images of natural scenes.

The first filters applied in this work for IE are the One-Click filter and a set of Classical filters. In both cases, the metrics that take into account the aesthetic value (NIMA A and PHIL) increase the mean value of the images (see Table 3 and Figure 7). However, with the averages based on image quality (NIMA T and BRISQUE), the mean value decreases after using these filters, especially in the case of BRISQUE.

This trend is maintained when trying to improve the image with GAN trained on a One-Click basis (Table 3). GAN A increases the mean value of the images for the aesthetics metrics and decreases it for the quality metrics, but with lower levels of variation. Note that the GAN generates images much closer to the original image (SSIM of 0.90) and a BRISQUE value that, although negative, is closer to 0.

Table 3 shows the results of combining GAN A with the Classical filters (GAN + CL), showing that, for the first time, the value of the quality metric BRISQUE is positive. The value of the other quality metric, NIMA T, is negative but close to 0. However, it is important to note that the SSIM is the lowest so far, indicating that the new images are quite far from the original. This very low SSIM (0.61) can be explained by the combination of the GAN with the Classical filters since the Classical filters independently already have a very low SSIM (0.66). This degree of difference with the original image, however, is not so clearly represented in the examples shown in Figure 9.

The two aesthetics-based metrics (NIMA A and PHIL) provide higher values when extreme processing is applied to the images. This behavior may be due to the fact that the metrics were created based on real sets of images, which by nature do not usually include extreme processing. It is necessary to incorporate human-evaluated extreme examples in the training of the aesthetics metrics, in order to provide a reference of this type of images and, thus, avoid over-evaluation and consequent over-processing. Currently, examples with these extreme examples are incorporated in the PHIL metric updates.

Concerning the experiments performed with GP, the choice of fitness completely conditions the behavior of the system. In the first case, GP.A, the fitness only tries to improve the result of an aesthetics metric: NIMA A. The result is an interesting filter from an artistic point of view, providing images with higher average scores in the two aesthetic metrics (NIMA A and PHIL) and in the quality metric BRISQUE. However, the resulting image presents a low similarity to the original, which manifests an unrealistic appearance (as can be seen in Figure 10). In fact, the lack of realism means that the image cannot be used in the context of real estate marketing, although it can be explored in a more artistic context.

The other two fitness used seek to maintain a balance between image enhancement and similarity to the original. The fundamental difference is in the metric each uses to try to obtain a realistic result, with GP.B using the NIMA T metric, and GP.C using the SSIM. The latter mechanism is more robust, providing more realistic images (with an SSIM of 0.88), therefore being more appropriate for application in real estate problems.

6. Conclusions

This work proposes an automatic image enhancement approach that takes an input image and generates an output suitable for a real-world application, specifically for real estate image marketing, where aesthetics and verisimilitude are of paramount importance.

In the context of real estate advertisements, overly-processed images are not well received by consumers, as they appear unreliable. For the consumer, it is important that the image faithfully represents the actual property.

This, coupled with the intrinsic subjectivity of aesthetic image evaluation, makes image enhancement a difficult task to solve. Moreover, a single method is often not the best solution for a variety of cases; thus, a dynamic and automatic solution has been created to solve the problem.

We presented the results of using Classical image processing filters, Pix2Pix GANs, and a set of filters created using GP.

We performed tests using state-of-the-art *no-reference* IQA metrics: two metrics that analyzes the image quality (NIMA T. and BRISQUE), and two metrics that focus on image aesthetics (NIMA A and PHIL). The last one, qualifies an input image based on its appeal for online real estate image marketing. We present some analysis of the four metrics, including some examples of the use of this metrics with real estate images. The BRISQUE and PHIL IQA metrics present more spread-out range of scores. We also made use and one *full-reference* method, SSIM, and MSE, as measures of similarity.

As Table 3 shows, the results of the GANs follow the same trend as those obtained with the Classical and One-Click filters: the value of the aesthetic metrics (NIMA A and PHIL) is higher than that of the quality metrics (NIMA T and BRISQUE), although, with the GANs, the SSIM value increases, with a result closer to the original image.

The best results for both quality and aesthetics metrics were achieved using GP. The GP system was designed using parameters that provided good results in previous work by the authors of References [9,39,40].

Three different fitness functions were designed for the GP system. The first (GP.A) uses a metric based on aesthetic value and provides images that improve this metric by 0.97. In fact, this enhancement is the largest achieved in all experiments in this study. However, the processing of these images is too strong, creating very interesting variations on the original images that may present artistic interest but are not suitable for presentation in real estate advertisements.

The other two fitness functions are based on the sum of the values of two metrics, one being an aesthetics metric, and the other a quality or similarity metric. This results in more bounded processing suitable for use in real estate imaging. GP.B generates the highest increase in the PHIL metric, and GP.C results in images with high similarity to the original, presenting the highest SSIM of all experiments (SSIM = 0.88).

On the other hand, we argue the importance of a threshold in IE solutions, as the results repeatedly showed that every image has an inherent maximum quality that was attempting to improve any further and will most of the time cause quality degradation. This came along with the conclusion that images in which the original quality is lower are generally easier to improve upon than those that demonstrate higher quality right away.

We also developed a pipeline framework composed of Machine Learning methods, image processing filters, and an Evolutionary approach to create dynamic image enhancement pipelines capable of addressing this problem.

IE is an extensive field; thus, it is virtually impossible to expand upon every detail of this work. The core of our solution relies on seven different elementary Classical functions. For future work, this set of functions can be altered in order to obtain different results. The parameters range for each function is also manually defined; thus, it could also be changed to allow the generated solutions to have more or less impact on the image. All this implies the possibility of creating a similar system where enhancement intensity is a controllable parameter.

Another line of development was also detected in the validation of the GP system solutions. The obtained solutions are tested on 10 individuals, and the average improvement of each is considered. This makes sense because improvements can be obtained on a high number of images. However, it may be interesting to evolve filters that are very useful for some images, even if they are poorly suited for others. To incorporate these filters that are interesting only in specific cases, it would be necessary to test the GP solutions on more images and only take into account the percentage of images where the individual generates a higher improvement.

Finally, the GP system is incorporated in the framework but as an offline system. Although this system can be activated periodically, the fact that it is offline makes the framework less dynamic to the introduction of new image types, which require new filters generated by the GP system. In the future, we will test new forms of integration that allow the training of the system to be included in the online cycle.

References

1. Rana, S. A Review of Medical Image Enhancement Techniques for Image Processing. *Int. J. Curr. Eng. Technol.* **2011**, *5*, 1282–1286. [CrossRef]
2. Pandey, P.; Dewangan, K.K.; Dewangan, D.K. Satellite image enhancement techniques—A comparative study. In Proceedings of the 2017 International Conference on Energy, Communication, Data Analytics and Soft Computing (ICECDS), Chennai, India, 1–2 August 2017; pp. 597–602. [CrossRef]
3. Schuch, P.; Schulz, S.; Busch, C. Survey on the impact of fingerprint image enhancement. *IET Biom.* **2018**, *7*, 102–115. [CrossRef]
4. Bloch, P.H.; Brunel, F.F.; Arnold, T.J. Individual differences in the centrality of visual product aesthetics: Concept and measurement. *J. Consum. Res.* **2003**, *29*, 551–565. [CrossRef]
5. Chae, Y.; Nakazawa, M.; Stenger, B. Enhancing product images for click-through rate improvement. In Proceedings of the 2018 25th IEEE International Conference on Image Processing (ICIP), Athens, Greece, 7–10 October 2018; pp. 1428–1432.
6. Maini, R.; Aggarwal, H. A Comprehensive Review of Image Enhancement Techniques. *arXiv* **2010**, arXiv:1003.4053.
7. Wang, W.; Chen, Z.; Yuan, X.; Wu, X. Adaptive image enhancement method for correcting low-illumination images. *Inf. Sci.* **2019**, *496*, 25–41. [CrossRef]
8. Jiang, Y.; Gong, X.; Liu, D.; Cheng, Y.; Fang, C.; Shen, X.; Yang, J.; Zhou, P.; Wang, Z. EnlightenGAN: Deep Light Enhancement without Paired Supervision. *arXiv* **2019**, arXiv:1906.06972.
9. Correia, J.; Vieira, L.; Rodriguez-Fernandez, N.; Romero, J.; Machado, P. Evolving Image Enhancement Pipelines. In *Artificial Intelligence in Music, Sound, Art and Design, Proceedings of the 10th International Conference, EvoMUSART 2021, Held as Part of EvoStar 2021, Virtual Event, 7–9 April 2021*; Lecture Notes in Computer Science; Romero, J., Martins, T., Rodríguez-Fernández, N., Eds.; Springer: Berlin/Heidelberg, Germany, 2021; Volume 12693, pp. 82–97. [CrossRef]
10. Umbaugh, S.E. *Computer Vision and Image Processing: A Practical Approach Using Cviptools with Cdrom*, 1st ed.; Prentice Hall PTR: Hoboken, NJ, USA, 1997.
11. Ghadiyaram, D.; Goodall, T.; Choi, L.K.; Bovik, A.C. Perceptual Image Enhancement. In *Encyclopedia of Image Processing*; Laplante, P.A., Ed.; CRC Press: Boca Raton, FL, USA, 2018.
12. Zhuo, S.; Zhang, X.; Miao, X.; Sim, T. Enhancing low light images using near infrared flash images. In Proceedings of the International Conference on Image Processing, ICIP, Hong Kong, China, 26–29 September 2010; pp. 2537–2540. [CrossRef]
13. Talebi, H.; Milanfar, P. Fast Multi-Layer Laplacian Enhancement. *IEEE Trans. Comput. Imaging* **2016**. [CrossRef]
14. Wong, C.Y.; Jiang, G.; Rahman, M.A.; Liu, S.; Lin, S.C.F.; Kwok, N.; Shi, H.; Yu, Y.H.; Wu, T. Histogram equalization and optimal profile compression based approach for colour image enhancement. *J. Vis. Commun. Image Represent.* **2016**, *38*, 802–813. [CrossRef]
15. Afifi, M.; Punnappurath, A.; Abdelhamed, A.; Karaimer, H.C.; Abuolaim, A.; Brown, M.S. Color Temperature Tuning: Allowing Accurate Post-Capture White-Balance Editing. In *Color Imaging Conference (CIC)*; Society for Imaging Science and Technology: Springfield, VA, USA, 2019.
16. Afifi, M.; Brown, M.S. Deep White-Balance Editing. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Seattle, WA, USA, 14–19 June 2020.
17. Zhang, K.; Zuo, W.; Chen, Y.; Meng, D.; Zhang, L. Beyond a Gaussian denoiser: Residual learning of deep CNN for image denoising. *IEEE Trans. Image Process.* **2017**, *26*, 3142–3155. [CrossRef]
18. Deng, Y.; Loy, C.C.; Tang, X. Aesthetic-driven image enhancement by adversarial learning. In Proceedings of the 2018 ACM Multimedia Conference (MM 2018), Seoul, Korea, 22–26 October 2018; pp. 870–878. [CrossRef]

19. Chen, C.; Chen, Q.; Xu, J.; Koltun, V. Learning to See in the Dark. In Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition, Salt Lake City, UT, USA, 18–22 June 2018; pp. 3291–3300. [CrossRef]

20. Kang, S.B.; Kapoor, A.; Lischinski, D. Personalization of image enhancement. In Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition, San Francisco, CA, USA, 13–18 June 2010; pp. 1799–1806. [CrossRef]

21. Isola, P.; Zhu, J.; Zhou, T.; Efros, A.A. Image-to-Image Translation with Conditional Adversarial Networks. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Las Vegas, NV, USA, 27–30 June 2016; pp. 1125–1134.

22. Immerkær, J. Fast Noise Variance Estimation. *Comput. Vis. Image Underst.* **1996**, *64*, 300–302. [CrossRef]

23. Peli, E. Contrast in complex images. *J. Opt. Soc. Am. A* **1990**, *7*, 2032–2040. [CrossRef]

24. Rex Finley, D. HSP Color Model—Alternative to HSV (HSB) and HSL. 2006. Available online: http://alienryderflex.com/hsp.html (accessed on 20 January 2022).

25. Pech-Pacheco, J.L.; Cristobal, G.; Chamorro-Martinez, J.; Fernandez-Valdivia, J. Diatom autofocusing in brightfield microscopy: A comparative study. In Proceedings of the 15th International Conference on Pattern Recognition, ICPR-2000, Barcelona, Spain, 3–7 September 2000; Volume 3, pp. 314–317.

26. Fortin, F.A.; De Rainville, F.M.; Gardner, M.A.; Parizeau, M.; Gagné, C. DEAP: Evolutionary Algorithms Made Easy. *J. Mach. Learn. Res.* **2012**, *13*, 2171–2175.

27. Bazeille, S.; Quidu, I.; Jaulin, L.; Malkasse, J.P. Automatic underwater image pre-processing. In Proceedings of the CMM'06, Brest, France, 16–19 October 2006.

28. Xie, Y.; Ning, L.; Wang, M.; Li, C. Image Enhancement Based on Histogram Equalization. *J. Phys. Conf. Ser.* **2019**, *1314*, 012161. [CrossRef]

29. Chang, Y.; Jung, C.; Ke, P.; Song, H.; Hwang, J. Automatic Contrast-Limited Adaptive Histogram Equalization with Dual Gamma Correction. *IEEE Access* **2018**, *6*, 11782–11792. [CrossRef]

30. Buades, A.; Coll, B.; Morel, J.M. Non-Local Means Denoising. *Image Process. Online* **2011**, *1*, 208–212. [CrossRef]

31. Limare, N.; Lisani, J.L.; Morel, J.M.; Petro, A.B.; Sbert, C. Simplest Color Balance. *Image Process. Online* **2011**, *1*, 297–315. [CrossRef]

32. Goodfellow, I.; Pouget-Abadie, J.; Mirza, M.; Xu, B.; Warde-Farley, D.; Ozair, S.; Courville, A.; Bengio, Y. Generative Adversarial Nets. In *Advances in Neural Information Processing Systems 27*; Ghahramani, Z., Welling, M., Cortes, C., Lawrence, N.D., Weinberger, K.Q., Eds.; Curran Associates, Inc.: Red Hook, NY, USA, 2014; pp. 2672–2680.

33. Rodriguez-Fernandez, N.; Alvarez-Gonzalez, S.; Santos, I.; Torrente-Patiño, A.; Carballal, A.; Romero, J. Validation of an Aesthetic Assessment System for Commercial Tasks. *Entropy* **2022**, *24*, 103. [CrossRef]

34. He, L.; Gao, F.; Hou, W.; Hao, L. Objective image quality assessment: A survey. *Int. J. Comput. Math.* **2014**, *91*, 2374–2388. [CrossRef]

35. Mittal, A.; Moorthy, A.K.; Bovik, A.C. No-Reference Image Quality Assessment in the Spatial Domain. *IEEE Trans. Image Process.* **2012**, *21*, 4695–4708. [CrossRef]

36. Lim, J.; Heo, M.; Lee, C.; Kim, C.S. Contrast enhancement of noisy low-light images based on structure-texture-noise decomposition. *J. Vis. Commun. Image Represent.* **2017**, *45*, 107–121. [CrossRef]

37. Wang, G.; Li, L.; Li, Q.; Gu, K.; Lu, Z.; Qian, J. Perceptual evaluation of single-image super-resolution reconstruction. In Proceedings of the 2017 IEEE International Conference on Image Processing (ICIP), Beijing, China, 17–20 September 2017; pp. 3145–3149.

38. Esfandarani, H.T.; Milanfar, P. NIMA: Neural Image Assessment. *IEEE Trans. Image Process.* **2017**, *27*, 3998–4011.

39. Machado, P.; Cardoso, A. All the Truth About NEvAr. *Appl. Intell.* **2002**, *16*, 101–118. [CrossRef]

40. Correia, J.; Martins, T.; Machado, P. Evolutionary Data Augmentation in Deep Face Detection. In Proceedings of the 2019 Genetic and Evolutionary Computation Conference (GECCO 2019), Prague, Czech Republic, 13–17 July 2019.