Article in Press IJIMAI journal

# Aligning Figurative Paintings With Their Sources for Semantic Interpretation

Sinem Aslan[1,2], Luc Steels[3]

[1] Ege University, International Computer Institute, Bornova, Izmir (Turkey)
[2] Ca' Foscari University of Venice, DAIS & ECLT, Venice (Italy)
[3] Barcelona Supercomputing Center, Barcelona (Spain)

Received 14 April 2022 | Accepted 2 March 2023 | Early Access 17 April 2023

## Abstract

This paper reports steps in probing the artistic methods of figurative painters through computational algorithms. We explore a comparative method that investigates the relation between the source of a painting, typically a photograph or an earlier painting, and the painting itself. A first crucial step in this process is to find the source and to crop, standardize and align it to the painting so that a comparison becomes possible. The next step is to apply different low-level algorithms to construct difference maps for color, edges, texture, brightness, etc. From this basis, various subsequent operations become possible to detect and compare features of the image, such as facial action units and the emotions they signify. This paper demonstrates a pipeline we have built and tested using paintings by a renowned contemporary painter Luc Tuymans. We focus in this paper particularly on the alignment process, on edge difference maps, and on the utility of the comparative method for bringing out the semantic significance of a painting.

## Keywords

Artistic Methods, Computer Vision, Edge Detection, Figurative Art Analysis, Image Alignment.

DOI: 10.9781/ijimai.2023.04.004

## I. The Comparative Method

This paper reports on research into the artistic methods used by figurative painters using computational algorithms. One aspect of the artistic method concerns style. Human viewers quickly see whether a landscape or a face is painted in a romantic, impressionist, expressionist or cubist style. Much work in AI, with remarkable results, has been done on capturing an artist's style or period and generating new works in a similar style [1], [2]. Although style transfer is very interesting, it is not discussed in this paper because we are interested in another aspect of the artistic method, namely the *expression of meaning*.

Painters, particularly figurative painters, want to mean something with their work and they introduce *signifiers* that convey these meanings. They introduce focal points and centers of interest, pronounced edges, textural regions with less detail, brightness contrasts, unusual defigurations of objects, etc. [3]. Whether these potential signifiers become real signifiers is determined by later semantic processing that make use of the context and world knowledge to interpret the painting. For example, a portrait is typically not a photographic rendering of the depicted person's face. The painter selects, highlights, and transforms the source image (either a live model or a photograph) in order to express meanings at many levels. For example, he or she may want to convey the personality and affective state of the person, his or her moral attitudes, and the socio-cultural context in which the person lived. This is well illustrated by Francis Bacon's series of popes all inspired by the famous painting by Velasquez of Pope Innocent X.

* Corresponding author.

E-mail address: sinem.aslan@unive.it

There is already a vast and fast growing literature using computer vision algorithms for various functions related to art interpretation, such as classification [4], object detection [5], similarity retrieval [6], sentiment analysis [7], and generative art [8] to name just a few of the most active areas. Most of this work is based on the use of low level image analysis and image transformation although with the advent of deep learning and the availability of semantically annotated image datasets, there is now a general trend towards the detection and interpretation of features that contribute to meaning [9].

Using this prior art, there are two approaches to study the artistic methods that artists use to transform source images into paintings. One approach is to start from a photograph of a face or a real world situation like a still life, construct a semantic interpretation that includes various levels of meaning such as affect, character, perspective, moral implication, socio-cultural context, etc. and then convert this photograph plus its desired meanings into a painting by making informed choices about cropping, lighting, color and tone, brush strokes, edges, level of textural detail, contrast, etc. [10]. This approach is being pursued in the context of *non-photorealistic* or *artistic rendering* (*NPR*) [11], which has become more and more sophisticated lately to include parameters driven by semantic criteria such as emotional or personality analysis rather than random perturbations [12].

Another complementary approach, advocated in this paper, goes in the other direction. We call it the *comparative method*. It starts from the source photograph and background knowledge about the painter and from the catalog and studies how this painter actually transformed the source photograph and what possible meanings could have played a role. This approach is therefore a way to study a painter's artistic method, not just his style, as is done in research on style transfer, but what kind of meanings have been found to be important to express

Please cite this article in press as:
S. Aslan, L. Steels. Aligning Figurative Paintings With Their Sources for Semantic Interpretation, International Journal of Interactive Multimedia and Artificial Intelligence, (2023), http://dx.doi.org/10.9781/ijimai.2023.04.004

and what expression strategies have been employed. The benefit of a comparative method is not only increased understanding of a painting and the oeuvre and approach of a painter. This kind of analysis could also yield insights and methods that could feed into the NPR approach by shedding more light on how creative artists achieve non-photorealistic rendering. Normally this method can only be used if a source image (which could also be an earlier painting by the same or another painter) is available, however there have also been remarkable experiments, in relation to the work of Rembrandt, where a new photograph is made of an existing person that resembles a figure painted centuries ago [13] and then we can use the comparative method starting from this photograph.

This paper takes steps toward an application of the comparative method. It requires that we find first the source of the painting under investigation by interacting directly with the painter or by historical research. The massive number of images now available on the Internet and existing image search algorithms should be very helpful in this respect. The next step is to overlay the relevant parts of the source on the target painting so that they become visually comparable. Painters may isolate only a small area of a source. They may leave out details, for example to make the object on the painting less tied to its source and hence more universal. They may stretch represented objects, shift them with respect to each other and change their orientation. They may change the color choices of significant surfaces, add or remove edges, etc. Many of these actions are geared towards the creation of potential signifiers and influence the way a viewer reacts to the painting.

More concretely, we focus first in this paper, which builds further on the results reported in [14], on two concrete technical challenges necessary to make the comparative method applicable: (i) finding the geometric operations of translation, scaling and rotation which align the painting and its source, and (ii) computing edge difference maps. In a final section we go back to the bigger picture and demonstrate the utility of difference maps, specifically by extracting facial activation units and focusing on those where the edge difference maps have identified regions of interest.

## II. The Case Studies

Using paintings by Luc Tuymans, a contemporary Flemish painter, we have done a number of concrete case studies for both technical challenges. It's worth to note that working with a living artist makes it possible to validate our methodology and tells us whether the algorithms have yielded valuable results, not only for viewers, curators or art historians but also for those who create the artworks. These case studies have lead to an exhibition called 'Secrets'[1], Artificial Intelligence and Luc Tuymans' at the BOZAR cultural center in Brussels, which raised considerable impact in the community of artists, art curators and interested viewers.

Luc Tuymans is currently considered one of the most important contemporary painters [15]. His solo exhibitions took place at some of the most prestigious and influential art centres in the world, such as the MOMA in New York, the MCA Museum of Contemporary Art in Chicago, the Palazzo Grassi in Venice, the Städel Museum in Frankfurt, the National Art Museum of Beijing, BOZAR in Brussels, etc. We have been fortunate to be in direct and frequent contact with this painter and to have access to relevant parts of his digital archives. Luc Tuymans is very articulate in describing both his own artistic method and the methods used by other painters [16] and he almost always works on the basis of a photographic image, which he has supplied or validated for the case studies we have undertaken.

[1] https://readymag.com/u3083945729/secrets-guide/

We have examined quite a number of paintings from the 2019-2020 solo exhibition of Luc Tuymans in the Palazzo Grassi in Venice. In this paper we demonstrate our work using two oil paintings from this exhibition, shown in Fig. 1: *K.*, which depicts the face of a young woman looking energetically into the future, and *Secrets*, which depicts the face of an older man in a somber mood. *K.* is bigger than *Secrets* and they are quite different in terms of color usage and general emotional impact.

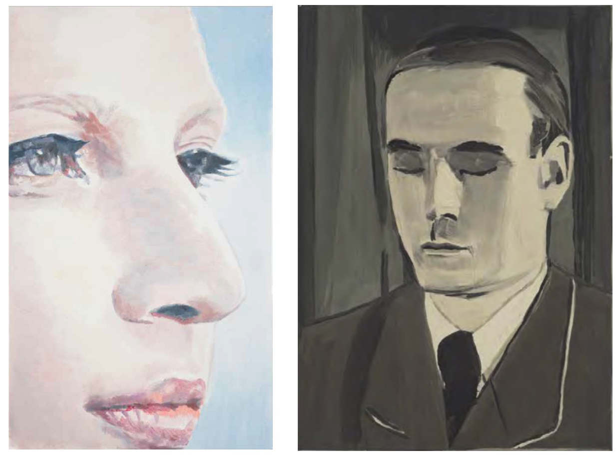

Fig. 1. Two of the oil paintings by Luc Tuymans used in our case studies. Left: *K.* (2017) oil on canvas 135 x 80,2 cm. Private collection. Right: *Secrets* (1990) oil on canvas 52 x 37cm. Private collection.

The overall workflow we used for this paper is illustrated for the painting *K.* in Fig. 2. From left to right, there is the identification of the original, the alignment process, edge detection and construction of comparing edge maps, and their use in further pattern recognition and semantic interpretation. Each of these steps is discussed in detail in the body of this paper, both for the painting *K.* and for *Secrets*.

## III. Sources

Due to direct interactions with the painter, we had access to the originals he used. But it is also interesting to try and find the sources of these originals and their wider context using the Internet. We used reverse image search as offered by several commercial search providers, namely Google (American), TinEye (American), Bing (American), Yandex (Russian) and Baidu (Chinese), using the paintings themselves as the key. These search engines indeed provide a large number of images that are visually related to the painting, with interesting cultural differences between the search engines, undoubtly based on the image repositories used to train the reverse image search algorithms. However, none of them yielded the original photograph nor the context in which it was taken. We hypothesize that this difficulty is related to the well known domain adaptation problem: Two images that humans see as quite similar are nevertheless not detected to be so due to differences in illumination, pose, image quality, texture, etc., because they cause a distribution change or domain shift between the domains derived from the respective images [17]. Solving domain adaptation is currently a frontier area in computer vision and we may expect that search engines will get better if new results are incorporated.

On the other hand, when we provide more information to search engines, we can retrieve the originals and their context. More information means that we supply the title or text from the catalog or we supply the source image that the painter originally provided to us. For the painting K. we show the results in Fig. 3. The painting is in fact a close-up of a face which is itself part of a larger scene

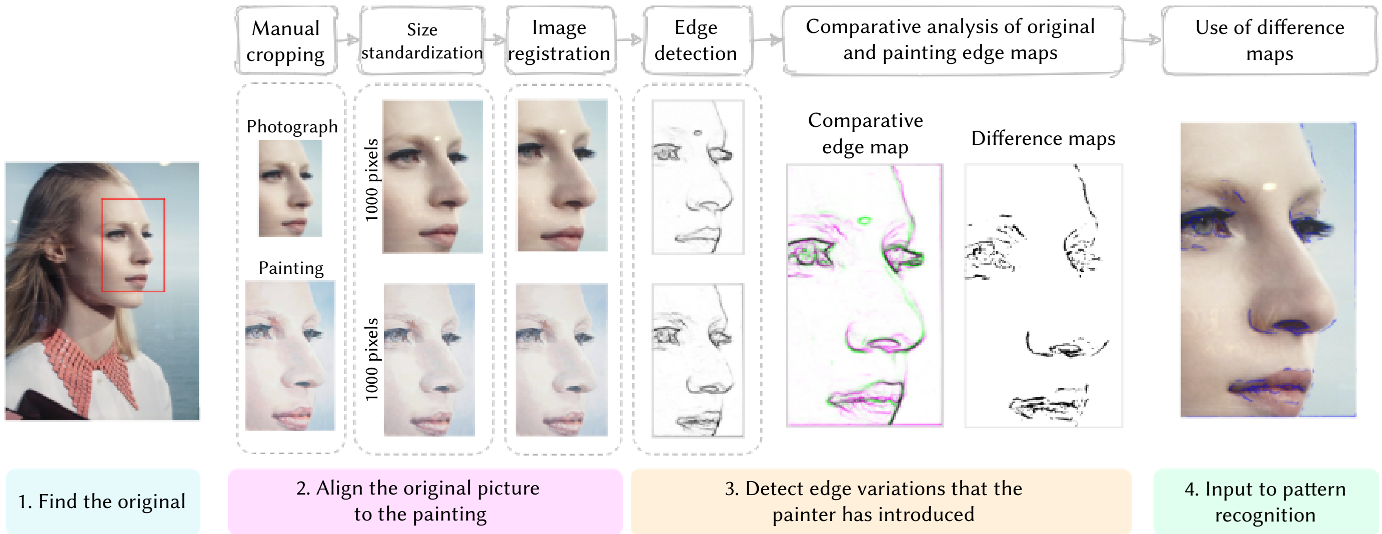


Fig. 2. Work flow discussed in this paper, going from identification and alignment to the construction of edge difference maps. It is illustrated for the painting *K.*.

coming from a Dior commercial. Knowing this context suggests that the direct inspiration is a fashion model from an advertising campaign. We see an objectification [18] of the human body, more concretely in this case of the human face, which is typical for advertising imagery in fashion or cosmetics. It can be said that this objectification is present in the photograph, and even more so in the painting considering the followings: with an excessive focus on the face, the context was almost completely eliminated, the details that normally make the faces look alive were softened, and the letter K. was chosen for the title of the painting instead of a real name for the woman depicted. Marc Donnadieu, curator of the Palazzo Grassi exhibition points in the catalogue to additional features of the face: 'smiles discreetly', 'defiant', 'expressive gaze', which are evoked through signifiers such as subtle changes in the lips, eyes and eyebrows, and a change in the nose.


Fig. 3. From left to right: painting, artist supplied source, original image from the Dior Autumn-Winter 2015 campaign photographed by Willy Vanderperre; clothes designed by Raf Simons and the fashion model is Julia Nobis.

For the painting *Secrets*, the original source is shown in Fig. 4. It is in fact a famous photograph of the Nazi architect and Third Reich minister of armament Albert Speer. The photograph was taken by Walter Frenz, the chief cameraman of Leni Riefenstahl. Seeing this source makes us realize at once that the secrets mentioned in the title have to do with denying knowledge and responsability for the atrocities of the war. The painter has used again the technique of zooming in on a segment of the original image and on removing iconic signifiers in the source (such as the Nazi insigna) in order to make the portrait more timeless and convey expressive and emotional meanings. We see also that the face has become more rectangular, almost looking like a mask. The eyes are closed signifying the hiding of secrets, a grey shadow hangs over the face, the nose is sharper, and the lips are tight.

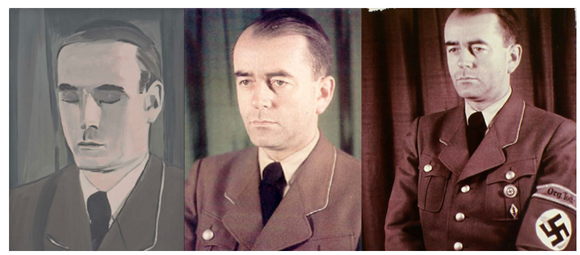

Fig. 4. From left to right: painting, cropped original, original. The cropped original is a photograph of Albert Speer provided by the artist. The original has been found through Google search using 'Albert Speer' as the key. It did not appear through reverse image search, neither with the painting as key nor with the cropped image supplied by the painter as key.

## IV. Aligning the Source Image to the Painting

We first focus on the geometric transformation process. We need algorithms that compute how the source was transformed to obtain the target painting, in other words the transformations that allow the source to be aligned as much as possible with the painting. Subsequent comparative visual processing rely entirely on whether such an alignment could be established. The relevant technique from computer vision for this purpose is called *image registration* now also often called *image alignment*. Image alignment is a well-studied problem in image processing and ready-made algorithms are available of all common computer vision platforms such as Matlab or OpenCV. We used the image registration algorithms available on Matlab[2].

Alignment aims to spatially align multiple images of the same scene. Image alignment is widely used in a variety of application fields [19] for: (1) multi-view analysis, where images from different viewpoints of the same scene are aligned for a larger (either 2D or 3D) representation of the scene; (2) multi-temporal analysis, where images of the same scene taken at different times are aligned to detect changes over that time period; or (3) multi-modal analysis, where images of the same scene are acquired by different types of sensors and aligned for fusing information from different sources to obtain a more comprehensive representation of the scene [19].

For successful registration some pre-processing of the image data was required. If a painting encompasses only a smaller portion of a picture (as is the case here), the relevant region in the picture has to

[2] https://mathworks.com/discovery/image-registration.html

be cropped to exclude the uninterested region and automatic image alignment has to be performed on the cropped portion. Performing cropping before image alignment has been recommended as a useful operation in the literature. For example, following a feature-based image registration approach [20] demonstrated that cropping the overlapping area and thus restricting the search area generally improves the results of the feature matching algorithms. [21] showed that cropping images by discarding the unnecessary background aids the alignment algorithm because it ensures registration of the foreground rather than registration of the background which improves the registration performance.

Moreover, as mentioned in [22], [23] the cropping operation reduces registration time and prevents memory issues. While cropping can be done manually [20]–[22], [24]–[26], there have been a number of experiments to automate it [23], [27]. Here we do not investigate automated cropping but use hand-selected possible candidates for *K.* and *Secrets* as shown in Fig. 5. The manual selection does not need to be very precise but without it we do not get usable results, as shown in Fig. 8.

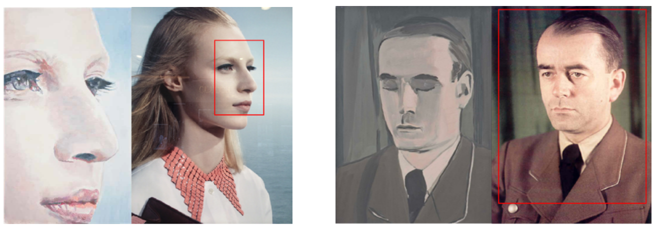

Fig. 5. Painting and selected source image for cropping *K.* (left) and *Secrets* (right).

In the terminology of image alignment algorithms, there are two images given as input. The one that remains unchanged is called the *reference* or *fixed* image, while the other, that is transformed to align with the reference image, is called the *sensed* or *moving* image [19], [28]. In this work, we consider the photograph, which is the original source image, as the *moving image* and try to align it to the painting by successively applying geometrical transformations on it. Thus, the *painting* is considered as the *fixed image* in image alignment terminology. This obviously reflects the *perspective of the painter* as the painter transforms the original photograph into a painting. Thus, in this paper, we discuss the painter's perspective.

[19] summarise the main steps of the majority of alignment methods as follows: 1) Either use *feature detection* to find salient elements such as lines, keypoints, or regions in both images, which can be done by well-known algorithms such as MSER, Harris, SURF, SIFT, etc., or use image pixels densely sampled on a regular grid as features [29]; 2) *Match* these features between the two images by means of similarity or correlation of the local neighborhoods of the features; and 3) *Estimate a transform model* to find the mapping function that transforms the moving image so that it overlays as well as possible with the reference one. We now show the result of applying these steps for our case study.

## A. Feature Detection and Matching

For these phases of the image alignment process, feature- based and area-based methods can both be used [19] and we have tried known algorithms for each method in order to see which approach is the most appropriate in the present context.

*Feature-based approaches* aim to match detected salient structures in both reference and moving images. We used the well known SURF algorithm on the gray-scaled original photograph and the painting [30]. We then *l*2-normalized the feature vectors to obtain the unit vectors and matched the features in the original photograph to the nearest neighbors in the features of the painting by computing the pair-wise distances (sum of squared differences) between feature vectors in the photograph and the painting. We used the default highest value for the match threshold $T$ ($0 < T \leq 100$) of the software platform[3], which is $T = 100$, i.e., two feature vectors match when the distance between them is less than or equal to 100. We performed a forward and backward matching between the photograph and the painting, and kept the best matches of the feature vectors. Results are shown in Fig. 6. For *K.*, 47 and 197 features were detected in the photograph and the painting, respectively and only six of them matched, while for *Secrets* 134 and 249 features were detected in the photograph and the painting, respectively, and 17 of them matched.

For a successful image alignment, the number of correctly matched features between the reference (fixed) and moving (sensed) image should be sufficiently high regardless of the geometrical or photometrical changes in the images [19]. This is not the case here. We observe in Fig. 6 that only two of the matches can be accepted as correct in the case of *K.*. On the other hand, while the number of matched points is higher, compared to *K.*, for *Secrets* there are still an unreasonable number of mismatches and they significantly affect alignment accuracy. We obtained similar results for other feature-based alignment algorithms, specifically the well known SIFT [31] and ORB [32] algorithms.

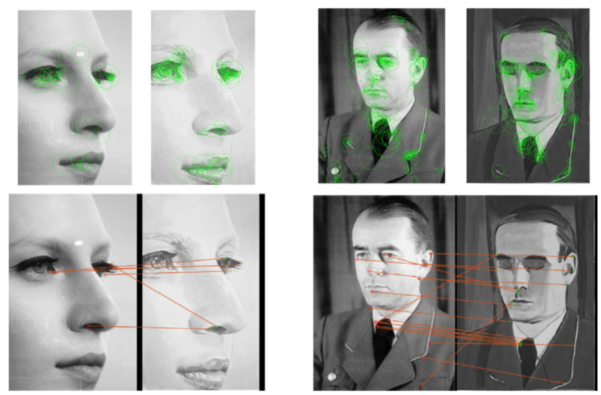

Fig. 6. Top: Detected SURF keypoints in the cropped photograph and the painting of *K.* (number of detected keypoints for the original source and painting images is 47 and 197, respectively), and *Secrets* (number of detected keyponts for the original source and painting images is 134 and 249, respectively); Bottom: Matched SURF features for *K.* and *Secrets* (number of matched features for K. and Secrets is 6 and 17, respectively).

Once there is a feature correspondence, the mapping function, also called transform model, can be estimated. It transforms the moving image so that it overlays as well as possible with the reference image, which requires finding a transformation function and estimating its parameters. A transform model hence characterizes the geometrical deformation to which the moving image has been subjected by the painter. For the present study we restricted the possible transforms to be shape-preserving so there could only be *rotation*, *translation*, and *isotropic scaling*.

Following feature-based matching, the M-estimator SAmple Consensus (MSAC) algorithm [33] was used to estimate the transform model parameters. The MSAC algorithm is a variant of the Random Sample Consensus (RANSAC) algorithm [33], [34], which is known to be more robust than RANSAC [35]. The quality of model estimation is evaluated using the sum of distances between all points to the estimated model differently from the RANSAC, which uses the number of inliers, i.e., correctly matched points, as the quality measure. Application of the transform to the moving image and overlay on the reference image is shown in Fig. 7. The results are not good at

[3] https://mathworks.com/help/vision/ref/matchfeatures.html

all, undoubtly because the feature-based approach used here (and for similar methods i.c. SIFT and ORB) does not work well for finding correspondences between photographs and paintings in the case of Luc Tuymans, simply because the operations done by the painter are too numerous - even though human observers immediately see that the same image contents are present. Perhaps a better result would be obtained if the painting is first transformed to look similar to a natural scene before alignment is attempted [36].

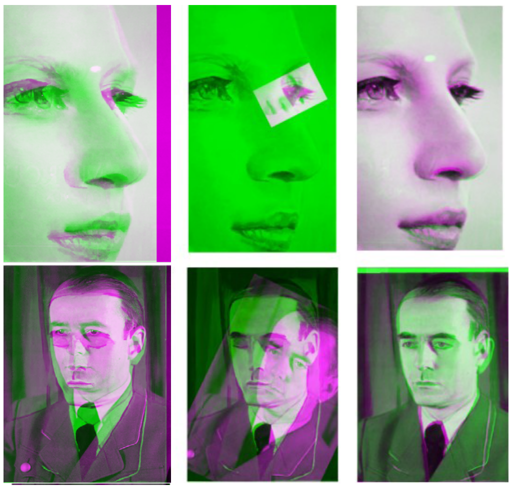

Fig. 7. Overlay of photograph on painting for *K.* (top) and *Secrets* (bottom): Gray regions in the composite image show where the two images have the same intensities. Magenta and green regions show where the intensities are different. Left: With the painting and the picture overlayed without image alignment; Middle: After image alignment using a feature-based (SURF) approach; Right: After image alignment using an area-based approach.

*Area-based approaches* do not attempt to detect salient regions, but use windows of predefined sizes or even entire images to estimate the correspondence. [19]. In this work, we used image pixel regions and one-plus-one evolutionary optimizer [37] for matching them, which is implemented using the Matlab Registration Estimator App[4] using its default parameter settings.

To speed up the process, this algorithm builds an image pyramid (both for the reference and moving input images) that has $N = 3$ levels where at each pyramid level the input image resolution is decreased by a factor of 2 in both image dimensions. Then, a coarse-to-fine hierarchical strategy is used to apply the alignment method, i.e., optimization starts at the coarsest level of the pyramid and continues at the finer levels until either convergence or $MaximumIterations = 100$ is reached. While going up to the finer resolutions, *estimates of feature correspondence* and *transform model parameters* are improved gradually [19]. The Mattes mutual information metric [38] is used to measure the similarity between reference and moving images in every optimization step. It was shown by [19] that this metric provides a more accurate alignment than the Mean Squares metric when the moving and reference images are from different modalities, as in our case.

One-plus-one evolutionary optimizer [37] refines the estimation of the parameters for the specified (similarity) transform model iteratively. A set of variations, called the *children*, of a given matrix of transformation parameters, called *parent*, is initially created using aggressive perturbations. If a child's parameters bring a better alignment, it becomes the new parent on the next iteration, otherwise, the parent stays the same and new children are computed with less aggressive changes to the parent's matrix. At every iteration of the optimization, the moving image is resampled by bilinear interpolation based on the transformation model estimated in that step and the similarity to be optimized between the reference and the transformed moving image is computed. For example, the obtained transform model for *Secrets* includes a translation with $t_x = -37.12$ and $t_y = 28.21$, a scaling by a factor $s_x = s_y = 1.09$ and a slight rotation by $\theta = -0.515$ degrees.

Finally, using the estimated transform model (from the feature-based or area-based approach), the moving image is resampled by bilinear interpolation and thus the images are said to be registered. Image alignment results obtained by feature-based and area-based approaches are shown in Fig. 7. It is seen that the accuracy of the alignment results with the area-based approach provides us almost a perfect alignment and we can build further on that base. For the feature-based approach we see that the original *K.* image has been shrunk and overlayed on the left eye and the original *Secrets* image has been rotated, both of which are not usable for further analysis.

It is of course possible to try many other methods for image alignment. We just mention two other promising approaches: based on using mutual information (MI) as a metric for alignment, as used by [39] for example, or using point cloud representations as originally developed for the reconstruction of 3D objects from multiple sources (laser scanner, digital cameras), as illustrated in [40]. However the present result is adequate for the next step in the pipeline.

Fig. 8. Results of alignment using the same area-based approach as in Fig. 7 but now without cropping, both for *K.* (left) and *Secrets* (right). The algorithm establishes a transform model, which is however totally unsatisfactory for further comparison.

## V. Edge Detection and Difference Maps

Having obtained an adequate alignment of the original picture with the painting, we can start to investigate the micro-transformations that the painter has introduced and their function [14]. These variations happen for different visual aspects, e.g., color, contrast, orientation, edges, contours, etc.. Thus, first these aspects are needed to be extracted from the painting and the aligned original picture, so that a comparison is possible. In the present work, we only look at edges, i.e., we explore which additional edges or edge variations the painter has introduced with respect to the original source photograph. Edges are certainly not the only vehicle that painters use to achieve visual effects and express meanings but it is a very important one. In our project we have also been investigating other types of features, specifically as related to the focal point [41] and color [42] and of course we are considering other features as well.

Our methodology consists of two steps. First, we detect the edges in both the source image and the painting. Then, for a comparative analysis, we construct a *difference map* that show all the edges for the source picture (magenta) and the painting (green) simultaneously, a *similarity map* that shows only the shared edges, and a *difference map* that shows which edges do not exist in the source image.

[4] https://mathworks.com/help/images/register-images-using-the-registration-estimator-app.html

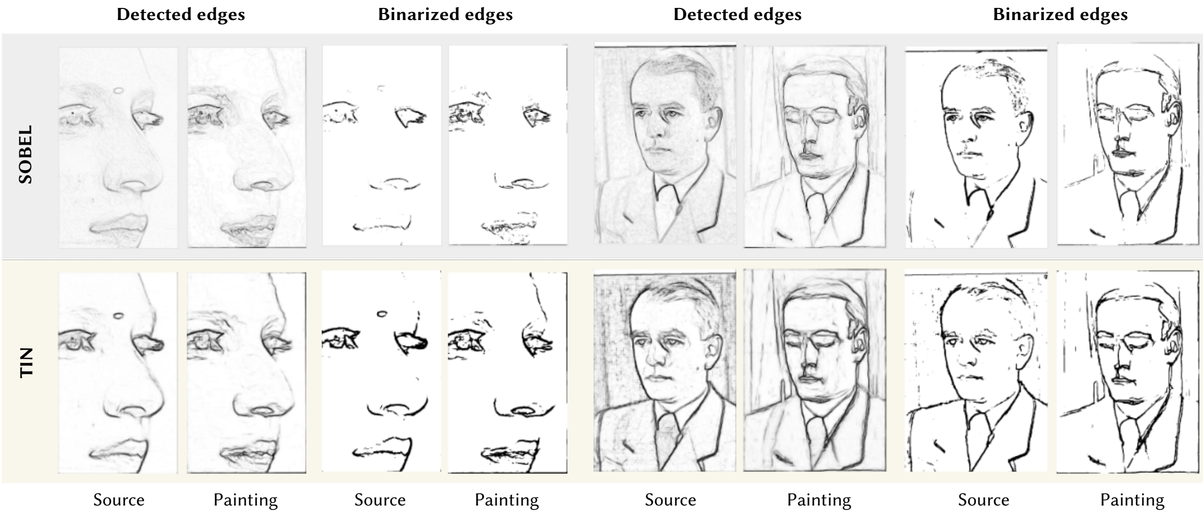


Fig. 9. Comparison of two edge detection methods, i.e., SOBEL [44] (1st row) and TIN [43] (2nd row), taking the painter's perspective (going from photograph to painting), with Non-Maximum Suppression (NMS) post-processing. We see that the TIN method provides clearer edges showing the variations introduced by the painter in a clearer way.

## A. Edge Detection

We wanted to compare the results with traditional edge detection methods, more specifically the *Sobel Isotropic* 3 × 3 *gradient operator* (SOBEL), and a deep neural network known as the *Traditional Inspired Network* (TIN) [43]. To achieve better-located edges, as mentioned in [43], we applied a post-processing operation, namely *Non-Maximum Suppression* (NMS), for both edge-detection methods. Specifically, we computed three edge maps for each input image at different scales, namely 1.5×, 1×, 0.5×, resized the resulting edge maps to the original image size, i.e. $x = 1000$, and averaged them to obtain the final edge map. More details about the two methods are as follows:

1. *The SOBEL edge detection* method [44], which has been widely used in image processing since the late 1960, is based on the derivation of a computationally efficient gradient operator. The gray-scaled input image $I$ is convolved with the following 3 × 3 kernels, to obtain the gradients for horizontal and vertical intensity changes, that are $G_x$ and $G_y$, respectively [44] as in Eq. (1).

$$G_x = \begin{bmatrix} -1 & 0 & +1 \\ -2 & 0 & +2 \\ -1 & 0 & +1 \end{bmatrix} * I,\ G_y = \begin{bmatrix} +1 & +2 & +1 \\ 0 & 0 & 0 \\ -1 & -2 & -1 \end{bmatrix} * I \ (1) \quad (1)$$

   Then, the resulting gradient approximations are combined with $G = \sqrt{G_x^2 + G_y^2}$ to have the gradient magnitude at each point in the image; this is the result displayed for the Sobel Method in the subsequent images in this paper.

2. *Deep neural networks* were originally designed for high-level computer vision tasks, e.g., object or scene recognition. Since edge detection is a simpler task, a lightweight deep learning network with reduced computational complexity can provide high-quality edges [43]. Motivated by this fact, in this work, we used a lightweight deep neural network architecture named *Traditional Method Inspired Network* (*TIN*) [43] where state-of-the-art accuracy performances were reported on the BSDS500 test set.

The TIN framework is composed of three modules: *Feature Extractor*, *Enrichment*, and *Summarizer*, which roughly correspond to gradient, low pass filter, and pixel connection in the traditional edge detection schemes [43]. In particular:

- *Feature Extractor* is formed by the 3 × 3 convolutional neural network layer which is designed to simulate the gradient operators (such as Sobel operator).
- *Enrichment* aims to remove the noise or tiny/isolated edge candidates by using multi scale filtering dilated convolutions, and
- *Summarizer* produces the final edges by fusing the outputs of the previous layer.

Two architectures, TIN1 and TIN2, were proposed by [43]. TIN1 is composed of the aforementioned three modules and TIN2 is a stack of two TIN1s, where output of the first module of the first TIN1 is downsampled by max-pooling in half and given as input to the first module of the second TIN1. The pre-training of the TIN method was performed on three datasets, i.e., BSDS500 (natural images), PASCAL VOC (natural images), and NYUDv2 (indoor images) [43]. Since higher performances were reported by TIN2 compared to TIN1 in [43], we employed the TIN2 architecture using the code published by the authors5 and their pretrained model on the aforementioned datasets.

Once we computed the grayscale edge maps with either SOBEL or TIN, we binarized them so that the significant edges be highlighted more. In the binarization operation we used a global threshold computed using Otsu's method, which chooses the threshold to minimize the intra-class variance and accordingly maximizes the inter-class variance of the thresholded black and white pixels [45]. Fig. 9 shows edge maps computed by each method and the outcome after thresholding. After thresholding it is observed that a significant amount of edges detected by the SOBEL method were removed, while the higher quantity of edges detected by TIN method were preserved. We proceed for further analysis with the edges detected by the TIN method, since TIN preserved the significant edges better. Of course, it is possible to try still other edge detection algorithms, e.g., [46], and the Canny algorithm [47] would be a prime candidate, but given the results with TIN we continue with this solution.

### 1. Comparative Analysis of Detected Edges

Finally, we compute the edge maps shown in Fig. 10. At the comparative edge map on the left side, detected grayscale edges by TIN are shown in magenta for the source picture, green for the painting, and black when the edges overlap. The equal edge map in the middle shows the overlapping edges in the binarized image, which can be considered as the locations where the painter has not introduced any modifications. We are most interested in the difference map on the right. It shows the edges introduced or emphasized by the painter and not found in the source image.

[5] https://github.com/jannctu/TIN


Fig. 10. Comparative analysis of the detected edges by TIN on the source image and painting. In the comparative grayscale edge map (left), the painting edges are in magenta and the source picture edges are in green. The similarity edge map (middle) shows the overlapped binarized edges. The difference edge map (right) shows the edge modifications that were introduced by the painter which do not exist in the source picture.

## VI. Step Towards Semantic Interpretation

The edge difference maps are just one of the many difference maps we can make but instead of considering other difference maps, we turn to the topic of semantic interpretation which is the ultimate goal of our work. It is important to point out that probing for the presence and meaning of signifiers requires more than the alignment and low level feature analysis discussed so far. We need to apply pattern recognition, such as (i.e., speaking for portrait paintings) facial expression recognition [48], [49] or recognition of attributes such as glasses, lipstick, hat, gender, hair color and hair shape, eyebrow shape, nose shape, lip shape, race, face shape, existence and shape of moustache and/or beard, etc. [50], as well as knowledge level processing based on common sense, world knowledge and historical knowledge.

We discuss in this section first what we can potentially learn from the edge difference maps and next how edge difference maps can be used in cooperation with other pattern recognition algorithms, more specifically face detection, facial behavior analysis and emotion recognition to probe semantic interpretation. A discussion on the role of knowledge level processing is beyond the scope of this paper.

### *A. Semantic Interpretation Using Edge Difference Maps*

Fig. 11 shows the edge difference map overlayed on the painting to illustrate that the following regions have been slightly altered [14]: (a) bottom part of the lips area, (b) the pupil and the area around the right eye, (c) the nose area and the curve at the right wing of the nose, (d) the left eye, especially the corner with the nose, and (e) the region above the left eye. These are therefore centers of interest and should be focal areas for subsequent pattern recognition and interpretation algorithms. The changes between the source and the painting are very subtle, nevertheless they give an overall change in the facial expression, as suggested by Marc Donnadieu in the exhibition catalogue: *K.* ‘smiles discreetly’, ‘is defiant’, and has an ‘expressive gaze’. The eyes are more open towards the world on the painting, also because they are in a lighter blue. The right mouth corner emphasizes a faint smile.

Continuing the interpretation of *K.*, we observe that in addition to deleting all context and objectifying the woman in the image, there is a strong cropping of the original image (Fig. 5) which also causes a strong focus on the eye gaze and on the main components of the face: the eyes, the nose and the lips. This zooming and focusing is so

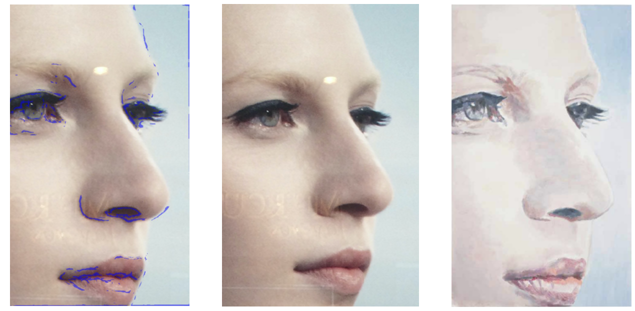

Fig. 11. Left: Original source image of *K.*. Middle: Edge difference map overlayed on aligned original source. Right: Painting *K.*.

strong that state-of-the-art image algorithms have great difficulty. For example, YOLOv3 [51] does not recognize that *K.* represents a face, but instead labels the eye area as a bird (Fig. 12) and Mask R-CNN, another common pixel labelling algorithm [52], does not do much better. Also state of the art facial behavioral analysis algorithms have difficulty to recognize the facial components. This is shown in Fig. 12 for the application of OpenFace [53]. For the painting *K.* (middle of Fig. 12), OpenFace 2.0 has difficulty to detect the facial landmarks, and recognize the facial action units and head orientation on the cropped source and it does not recognize any of them on the painting.

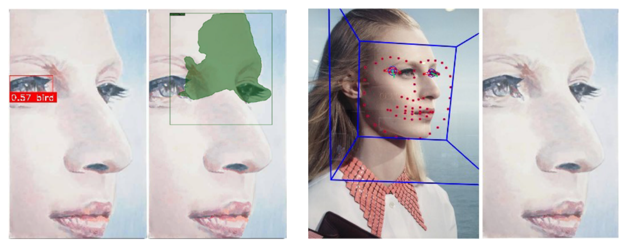


Fig. 12. Left: Segmentation and labeling of *K.* YOLO segments the eye and labels this segment as a bird with 0.57 % certainty and Mask R-CNN labels an area at the forehead as person with 0.74 % certainty. Right: Application of OpenFace to analyze the facial activation units and head orientation. Wrong results are obtained for the cropped image and no results at all for the painting.

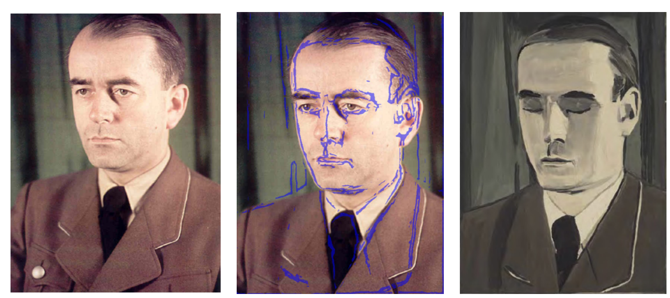

Fig. 13. Left: Original source image of *Secrets*. Middle: Edge difference map overlayed on aligned original source. Right: The painting itself.

The painting *Secrets* uses the same artistic method as *K.*: selecting a small portion of the original and cropping the image to focus on the face only, although the cropping is not so drastic so that semantic labeling (see left of Fig. 14) and algorithmic facial behavioral analysis (Right of Fig. 14) is now feasible with the current state-of-the-art in computer vision. All the insigna that point in a direct way to nazism have been removed so that a more universal image and a focus on the inner state of hiding secrets becomes the main topic. From the perspective of edge detection (see bottom series in Fig. 10) we see many more changes in the edges compared to *K.*, mostly in the following regions: (a) eye-lids region, (b) nose and lip region, (c) vertical regions on the border of the face, (d) intense regions under the chin, (e) left and right line on the jacket (left and right) (Fig. 13).

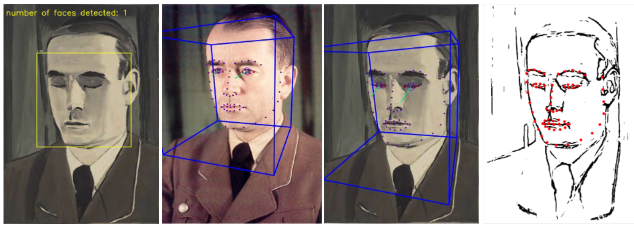


Fig. 14. Left: Semantic labeling of *Secrets*. A single face is recognized. Middle: Application of OpenFace on *Secrets*, both the source and the painting. Right: Projection of the activation units detected in the painting on the edge difference map.

It is not too difficult for humans to interpret these signifiers given common world knowledge. For example the left and right line of the jacket is a sign that the portrayed person is wearing a military uniform. The eye-lids are closed which is an ambiguous signifier that could point to sleeping, meditation, self-reflection, but also "to ignore something bad and pretend it is not happening" (Cambridge English Dictionary) in other words denial and hiding secrets. The nose is sharper, there is a moustache-like area under the nose, the eyebrows are more pronounced, the lips are more tight. These signifiers suggest an authoritarian attitude and one of hiding secrets. For example, tight lipped is defined in the Cambridge English dictionary as: "Someone who is tight-lipped is pressing their lips together to avoid showing anger, or is refusing to speak about something."

### *B. Semantic Interpretation Using Facial Behavioral Analysis*

Automatically detecting these interpretations (and we have just given a few examples) requires many more pattern recognition algorithms and the use of semantic web resources (thesauri, dictionaries, knowledge graphs, distributional semantics), but we can still illustrate further the comparative methodology by focusing on the expression of emotion using behavioral face analysis. Given the difficulties encountered with *K.* we probe *Secrets* only.

We have used the existing OpenFace 2.0 [53] to first reconstruct the facial action units and from there the emotions using the emotion facial action coding system (EMFACS) [54].

As with the edge difference maps, we are keenly interested in the differences between facial actions on the painting and the source picture. The painter can either *adopt* the facial actions of the source, and therefore their emotional expression, *amplify* them to express the emotion more strongly, or *diminish* them to weaken the emotional expression. The edge differences obtained in the previous section identify the *regions of interest* on which the interpretation of facial emotion recognition should focus (see right most image in Fig. 14).

OpenFace automatically detects facial landmarks, which are areas that are under the control of facial muscles (called facial action units) and are therefore available to express emotions. They include control of eyebrows (inner, outer brow), lips (upper lip raiser, lip corner puller, lip corner depressor), blinking or closing of the eyes, etc. The predicted intensity of all action units for Secrets are summarized in Fig. 15, both for the painting and for the picture. We can see very clearly that a number of action units have a much higher activation intensity, most notably the eyebrows which are thicker and more raised on the painting, the lip stretcher and parting which both make the lips look more tight, and the blink (which is actually the closing of the eyes).

Mapping AUs onto a number of emotion categories is still an active research area [55]. Several previous studies on facial emotion recognition have proposed to use computational algorithms, such as ANNs [56], [57], and SVMs [58], [59]. In this work, we follow the approach suggested by classic psychological studies [54], [60], that claim combined movements of the facial muscles are associated

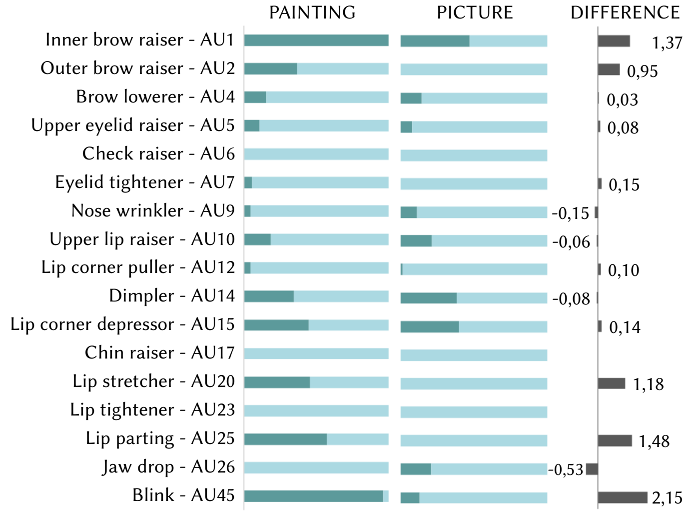


Fig. 15. Activation levels for different facial muscles, which are the basic features for the facial expression of emotion, on the painting Secrets and its original source image.

with one of the seven basic emotions [54], [57], [61] (see Table I). For example, sadness is calculated from the combination of action unit 1 (inner brow raiser), 4 (brow lowerer) and 15 (lip corner depressor). Thus, based on the predicted action-unit intensities by OpenFace 2.0, it is possible to characterize the presence of particular emotions, namely happiness, sadness, surprise, fear, anger, disgust, and contempt [54]. It can be seen in Table 1 that the strongest emotions are fear and sadness and to some extent surprise. They have been amplified in the painting implying that the painter has wanted to emphasize them. Happiness and anger are not expressed, neither in the picture or the painting, and disgust and contempt have roughly equal low levels in the picture and the painting.

Consequently, an important finding in the last section can be noted that strong cropping on the source picture and painting prevented YOLOv3, Mask R-CNN and OpenFace from recognizing faces and facial landmarks, as in *K.*. However, when cropping was not so drastic as in *Secrets*, they performed well and allowed us to conduct the comparative approach on the source picture and the painting in terms of facial expression and emotion recognition.

TABLE I. Characterization of Presence of Emotions on the Painting *Secrets* and Its Source Picture by Emotion-related Facial Actions [54]

| Emotion | Action units | Painting | Picture | Difference |
|---|---|---|---|---|
| Happiness | AU6+AU12 | 0,13 | 0,03 | 0.1 |
| Sadness | AU1+AU4+AU15 | 4,14 | 2,6 | 1,54 |
| Surprise | AU1+AU2+AU5+AU26 | 3,81 | 1,94 | 1.87 |
| Fear | AU1+AU2+AU4+AU5+AU7+AU20+AU26 | 5,54 | 2,31 | 3.23 |
| Anger | AU4+AU5+AU7+AU23 | 0,83 | 0,57 | 0.26 |
| Disgust | AU9+AU15+AU17 | 1,29 | 1,3 | -0.01 |
| Contempt | AUR12+AUR14 | 1,03 | 1,01 | 0.02 |

## VII. Conclusions

This paper reports on our ongoing research into artistic methods using AI algorithms in which we focus on how painters create signifiers to express meaning. We explored here a comparative method in which we compare the original source with the painting and applied this approach to the works of the contemporary Flemish painter Luc Tuymans. We investigated possible methods for aligning a painting and its source and used edge detection and the construction of comparative edge maps, to detect centers of interest. We found that an area-based alignment process gives by far better results compared to feature-based alignment and that the TIN edge detection method followed by the construction of aggregated edge maps gives useful candidates for further interpretation.

The paper is of interest because it reports on how a variety of computer vision and pattern recognition (e.g., YOLO, Mask R-CNN, OpenFace algorithms) fare with respect to the analysis of paintings. As was already known from earlier work, techniques that work for photographs of real world images do not carry over very well to paintings. Nevertheless for some paintings they already give adequate results for the application of the comparative method.

We are aware that this paper is only a small step in the fascinating but highly complex process through which painters create meaning and viewers reconstruct meanings. The final section nevertheless gave an idea in which direction we are going. We have linked the outcome of the edge analysis with the results of pattern recognition through OpenFace and showed how the artist amplifies certain emotions compatible with the way he wants to frame the depicted character.

## Acknowledgments

Experiments and preparation of the paper was partially funded by the EU Pathfinder project MUHAI through the Venice International University and by the EU Humane-AI.net coordination project. Additional funding for the interaction with Luc Tuymans came from the EU STARTS project through a scientist in residence grant to LS.

## References

[1] L. Gatys, A. Ecker, M. Bethge, “A neural algorithm of artistic style,” in *16th Annual Meeting of the Vision Sciences Society (VSS 2016)*, 2016, p. 326, Scholar One, Inc.

[2] Y. Jing, Y. Yang, Z. Feng, J. Ye, Y. Yu, M. Song, “Neural style transfer: A review,” *IEEE transactions on visualization and computer graphics*, vol. 26, no. 11, pp. 3365–3385, 2019.

[3] D. Hockney, M. Gayford, *A history of pictures. From the cave to the computer screen.* London, UK: Thames and Hudson, 2006.

[4] S. Liu, J. Yang, S. S. Agaian, C. Yuan, “Novel features for art movement classification of portrait paintings,” *Image and Vision Computing*, vol. 108, p. 104121, 2021.

[5] N. Gonthier, Y. Gousseau, S. Ladjal, O. Bonfait, “Weakly supervised object detection in artworks,” in *Proceedings of the European Conference on Computer Vision (ECCV) Workshops*, 2018, pp. 0–0.

[6] B. Seguin, C. Striolo, F. Kaplan, *et al.*, “Visual link retrieval in a database of paintings,” in *European conference on computer vision*, 2016, pp. 753–767, Springer.

[7] Y. Lin, “Sentiment analysis of painting based on deep learning,” in *International Conference on Application of Intelligent Systems in Multi-modal Information Analytics*, 2020, pp. 651–655, Springer.

[8] D. Foster, *Generative deep learning: teaching machines to paint, write, compose, and play.* O’Reilly Media, 2019.

[9] E. Cetinic, “Iconographic image captioning for artworks,” in *ICPR International Workshops and Challenges*, 2021, pp. 502–516.

[10] S. DiPaola, “Painterly rendered portraits from photographs using a knowledge-based approach,” in *Human Vision and Electronic Imaging XII*, vol. 6492, 2007, p. 649203, International Society for Optics and Photonics.

[11] A. A. Gooch, J. Long, L. Ji, A. Estey, B. S. Gooch, “Viewing progress in non-photorealistic rendering through heinlein’s lens,” in *Proceedings of the 8th International Symposium on Non-Photorealistic Animation and Rendering*, NPAR ’10, New York, NY, USA, 2010, pp. 165–171, Association for Computing Machinery.

[12] O. N. Yalçın, N. Abukhodair, S. DiPaola, “Empathic ai painter: A computational creativity system with embodied conversational interaction,” in *Proceedings of the NeurIPS 2019 Competition and Demonstration Track*, vol. 123, 2020, pp. 131–141.

[13] S. DiPaola, C. Riebe, J. Enns, “Rembrandt’s textural agency: A shared perspective in visual art and science,” *Leonardo*, vol. 43(2), pp. 145–151, 2020.

[14] S. Aslan, L. Steels, “Identifying centres of interest in paintings using alignment and edge detection,” in *ICPR International Workshops and Challenges. ICPR 2021. Lecture Notes in Computer Science, vol 12663*, 2021, pp. 589–603, Springer, Cham.

[15] N. S. U. Loock, J-V. Aliaga, *Luc Tuymans.* London: Phaidon, 1996.

[16] L. Tuymans, *The Image Revisited. In conversation with G. Boehm, T. Clark and H. De Wolf.* Brussels: Ludion, 2018.

[17] M. Wang, W. Deng, “Deep visual domain adaptation: A survey,” *Neurocomputing*, vol. 312, pp. 135–153, 2018.

[18] B. Fredrickson, T.-A. Roberts, “Objectification theory. toward understanding women’s lived experiences and mental health risks,” *Psychology of women quarterly*, vol. 21, pp. 173–206, 1997.

[19] B. Zitova, J. Flusser, “Image registration methods: a survey,” *Image and vision computing*, vol. 21, no. 11, pp. 977–1000, 2003.

[20] F. A. Onyango, “Multi-resolution automated image registration,” Master’s thesis, University of Twente, 2017.

[21] K. P. Wilkie, “Mutual information based methods to localize image registration,” Master’s thesis, University of Waterloo, 2005.

[22] S. Wognum, S. Heethuis, T. Rosario, M. Hoogeman, Bel, “Validation of deformable image registration algorithms on ct images of ex vivo porcine bladders with fiducial markers,” *Medical physics*, vol. 41, no. 7, p. 071916, 2014.

[23] M. Koenig, S. Kohle, H.-O. Peitgen, “Automatic cropping of breast regions for registration in mr mammography,” in *Medical Imaging 2005: Image Processing*, vol. 5747, 2005, pp. 1563–1570.

[24] E. Zacharaki, G. Matsopoulos, P. Asvestas, K. Nikita, K. Grondahl, H. Grondahl, “A digital subtraction radiography scheme based on automatic multiresolution registration,” *Dentomaxillofacial radiology*, vol. 33, no. 6, pp. 379–390, 2004.

[25] P. L. Chow, D. B. Stout, E. Komisopoulou, A. F. Chatziioannou, “A method of image registration for small animal, multi-modality imaging,” *Physics in Medicine & Biology*, vol. 51, no. 2, p. 379, 2006.

[26] Y. Sun, M.-P. Jolly, J. Moura, “Integrated registration of dynamic renal perfusion mr images,” in *International Conference on Image Processing, ICIP’04.*, vol. 3, 2004, pp. 1923–1926, IEEE.

[27] Y. Guo, Y. Liu, T. Georgiou, M. Lew, “A review of semantic segmentation using deep neural networks,” *International Journal of Multimedia Information Retrieval*, vol. 7, pp. 87–93, 2017.

[28] A. A. Goshtasby, *2-D and 3-D image registration: for medical, remote sensing, and industrial applications.* John Wiley & Sons, 2005.

[29] T. Tuytelaars, “Dense interest points,” in *2010 IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, 2010, pp. 2281–2288, IEEE.

[30] H. Bay, A. Ess, T. Tuytelaars, L. Van Gool, “Speeded- up robust features (surf),” *Computer vision and image understanding*, vol. 110, no. 3, pp. 346–359, 2008.

[31] D. Lowe, “Object recognition from local scale-invariant features,” in *Proceedings of the 7th IEEE International Conference on Computer Vision*, vol. 2, 1999, pp. 1150–1157 vol.2.

[32] E. Rublee, V. Rabaud, K. Konolige, G. R. Bradski, “ORB: an efficient alternative to SIFT or SURF,” in *IEEE International Conference on Computer Vision, ICCV*, 2011, pp. 2564–2571.

[33] P. H. Torr, A. Zisserman, “Mlesac: A new robust estimator with application to estimating image geometry,” *Computer vision and image understanding*, vol. 78, no. 1, pp. 138–156, 2000.

[34] R. Hartley, A. Zisserman, “Multiple view geometry in computer vision. cambridge university press, isbn: 0521540518,” 2004.

[35] S. Choi, T. Kim, W. Yu, “Performance evaluation of ransac family,” *Journal of Computer Vision*, vol. 24, no. 3, pp. 271–300, 1997.

[36] M. Tomei, L. Baraldi, M. Cornia, R. Cucchiara, “What was monet seeing while painting? translating artworks to photo-realistic images,” in *Proceedings of the European Conference on Computer Vision (ECCV) Workshops*, 2018, pp. 0–0.

[37] M. Styner, C. Brechbuhler, G. Szckely, G. Gerig, “Parametric estimate of intensity inhomogeneities applied to mri,” *IEEE Trans. on medical*

*imaging*, vol. 19, no. 3, pp. 153–165, 2000.

[38] S. Rahunathan, D. Stredney, P. Schmalbrock, B. D. Clymer, "Image registration using rigid registration and maximization of mutual information," in *13th Annu. Med. Meets Virtual Reality Conf*, 2005.

[39] A. Dame, E. Marchand, "Second-order optimization of mutual information for real-time image registration," *IEEE Transactions on Image Processing*, vol. 21, no. 9, pp. 4190–4203, 2012, doi: 10.1109/TIP.2012.2199124.

[40] N. Crombez, G. Caron, E. Mouaddib, "3d point cloud model colorization by dense registration of digital images.," *International Archives of the Photogrammetry, Remote Sensing & Spatial Information Sciences*, 2015.

[41] L. Steels, B. Wahle, "Perceiving the focal point of a painting with ai. case studies on works of luc tuymans.," in *12th International Conference on Agents and Artificial Intelligence*, 2020.

[42] S. Aslan, L. Steels, "Finding signifiers and their possible interpretations using colour," in *Conference on Colors and Cultures. Couleurs et Cultures. Universit'e de Haute-Alsace, Mulhouse (France)*, 2021.

[43] K. Wibisono, H.-M. Hang, "Traditional method inspired deep neural network for edge detection," in *IEEE Int. Conf. on Image Processing (ICIP)*, 2020, pp. 678–682.

[44] I. Sobel, G. Feldman, "A 3x3 isotropic gradient operator for image processing," *A talk at the Stanford Artificial Project*, pp. 271–272, 1968.

[45] N. Otsu, "A threshold selection method from gray- level histograms," *IEEE transactions on systems, man, and cybernetics*, vol. 9, no. 1, pp. 62–66, 1979.

[46] M. González-Hidalgo, S. Massanet, A. Mir, D. Ruiz- Aguilera, "A new edge detector based on uninorms," in *International Conference on Information Processing and Management of Uncertainty in Knowledge-Based Systems*, 2014, pp. 184–193, Springer.

[47] J. Canny, "A computational approach to edge detection," *IEEE Transactions on pattern analysis and machine intelligence*, no. 6, pp. 679–698, 1986.

[48] T. Baltrusaitis, P. Robinson, L.-P. Morency, "Openface: An open source facial behavior analysis toolkit.," in *WACV*, 2016, pp. 1–10, IEEE Computer Society.

[49] J. Hong, H. J. Lee, Y. Kim, Y. M. Ro, "Face tells detailed expression: Generating comprehensive facial expression sentence through facial action units," in *International Conference on Multimedia Modeling*, 2020, pp. 100–111, Springer.

[50] S. Bekhet, H. Alahmer, "A robust deep learning approach for glasses detection in non-standard facial images," *IET Biometrics*, vol. 10, no. 1, pp. 74–86, 2021.

[51] J. Redmon, A. Farhadi, "Yolov3: An incremental improvement," *arXiv preprint arXiv:1804.02767*, 2018.

[52] K. He, G. Gkioxari, P. Dollár, R. Girshick, "Mask r- cnn," in *Proceedings of the IEEE international conference on computer vision*, 2017, pp. 2961–2969.

[53] T. Baltrusaitis, A. Zadeh, Y. C. Lim, L.-P. Morency, "Openface 2.0: Facial behavior analysis toolkit," in *13th IEEE International Conference on Automatic Face & Gesture Recognition (FG 2018)*, 2018, pp. 59–66, IEEE.

[54] W. V. Friesen, P. Ekman, *et al.*, "Emfacs-7: Emotional facial action coding system," *Unpublished manuscript, University of California at San Francisco*, vol. 2, no. 36, p. 1, 1983.

[55] S. Velusamy, H. Kannan, B. Anand, A. Sharma, Navathe, "A method to infer emotions from facial action units," in *2011 IEEE international conference on acoustics, speech and signal processing (ICASSP)*, 2011, pp. 2028–2031, IEEE.

[56] B. Fasel, F. Monay, D. Gatica-Perez, "Latent semantic analysis of facial action codes for automatic facial expression recognition," in *Proceedings of the 6th ACM SIGMM international workshop on Multimedia information retrieval*, 2004, pp. 181–188.

[57] M. F. Valstar, M. Pantic, "Biologically vs. logic inspired encoding of facial actions and emotions in video," in *2006 IEEE International Conference on Multimedia and Expo*, 2006, pp. 325–328, IEEE.

[58] M. S. Bartlett, G. Littlewort, C. Lainscsek, I. Fasel, J. Movellan, "Machine learning methods for fully automatic recognition of facial expressions and facial actions," in *2004 IEEE International Conference on Systems, Man and Cybernetics (IEEE Cat. No. 04CH37583)*, vol. 1, 2004, pp. 592–597, IEEE.

[59] L. Yao, Y. Wan, H. Ni, B. Xu, "Action unit classification for facial expression recognition using active learning and svm," *Multimedia Tools and Applications*, vol. 80, no. 16, pp. 24287–24301, 2021.

[60] P. Ekman, W. V. Friesen, J. C. Hager, *Facial Action Coding System: Facial action coding system: the manual: on CD- ROM.* Research Nexus, 2002.

[61] M. Berkane, K. Belhouchette, H. Belhadef, "Emotion recognition approach using multilayer perceptron network and motion estimation," *International Journal of Synthetic Emotions (IJSE)*, vol. 10, no. 1, pp. 38–53, 2019.

Sinem Aslan

Sinem Aslan received her Ph.D. in Computer Science from Ege University (Turkey), in collaboration with the Electrical and Electronics Engineering Department of Boğaziçi University (Turkey), in 2016. Following her doctorate, she held postdoctoral researcher positions at IVL of the University of Milano-Bicocca, ECLT and DAIS of Ca' Foscari University of Venice (Italy), and Ege University (Turkey). She is currently an Assistant Professor at the Department of Environmental Sciences, Informatics and Statistics at Ca' Foscari University of Venice (Italy). Her recent research has focused on computer vision and machine learning, applied to fine arts and cultural heritage analysis.

Luc Steels

Luc Steels studied linguistics at the University of Antwerp (Belgium) and computer science at the Massachusetts Institute of Technology (USA). His main research field is Artificial Intelligence covering a wide range of intelligent abilities, including vision, robotic behavior, conceptual representations and language. In 1983 he became a professor of computer science at the University of Brussels (VUB). He has been co-founder and chairman (from 1990 until 1995) of the VUB Computer Science Department (Faculty of Sciences). He founded the Sony Computer Science Laboratory in Paris in 1996 and became its first director. After that he became ICREA research professor at the Institute for Evolutionary Biology (CSIC,UPF). Steels has participated in dozens of large-scale European projects and more than 30 PhD theses have been granted under his direction. He has produced over 200 articles and edited 15 books directly related to his research. During the past decade he has focused on theories for the origins and evolution of language using computer simulations and robotic experiments to discover and test them.