

# Quantitative Analysis of Visual Representation of Sign Elements in COVID-19 Context

Empirical Studies of the Arts  
1–21

© The Author(s) 2022


Article reuse guidelines:

sagepub.com/journals-permissions

DOI: 10.1177/02762374221104059

journals.sagepub.com/home/art



María Jesús Cano-Martínez<sup>1</sup>,  
Miguel Carrasco<sup>2</sup>, Joaquín Sandoval<sup>2</sup>,  
and César González-Martín<sup>3</sup> 

## Abstract

Visual representation as a means of communication uses elements to build a narrative. We propose using computer analysis to perform a quantitative analysis of the elements used in the visual creations that have been produced in reference to the epidemic, using 927 images compiled from The Covid Art Museum's Instagram account. This process has been carried out with techniques based on deep learning to detect objects contained in each study image. The research reveals the elements that are repeated in images to create narratives and the relations of association that are established in the sample. The predominant discourses in the sample do not show concern for the effects of illness. On the contrary, the impact and effects of confinement, through the prominent presence of elements such as human figures, windows, and buildings, are the most expressed experiences in the creations analyzed.

## Keywords

COVID-19, visual representation, computer analysis, the COVID art museum, instagram

<sup>1</sup>Jaume I University Faculty of Health Sciences, Castello de la Plana, Comunitat Valenciana, Spain

<sup>2</sup>Universidad Adolfo Ibáñez, Av. Diagonal Las Torres 2700, Of. 320. Edificio Talleres, Santiago, 7941169, Chile

<sup>3</sup>University of Cordoba, San Alberto Magno, Cordoba, 14071, Spain

## Corresponding Author:

César González-Martín, University of Cordoba, San Alberto Magno, Cordoba, 14071, Spain.

Email: cesar.gonzalez@uco.es

## Introduction

«(...) the best or most healthy way to respond to illness is through narrative» (Woods, 2012, p. 116).

We live surrounded by images. Thus, the representational baggage that things contain is not in fact overcome through discourse, but also through the images involved in every construct. In this sense, we have forgotten that «every image has at least two components: the reality it is reproducing and the meaning of that represented reality» (Aparici, 2009, p. 46). Beyond the merely mundane, human beings have developed a sophisticated way of creating images, which far from presenting reality, represent it, assigning it a story, a function. Thus, the image acts as a testimony to events, as liberator of expressions or emotions, or as a merely persuasive capable of inciting action and/or consumption.

Jovchelovitch and Routledge (2019) shows us in his research how representation is the basis of every knowledge system and the origin of understanding. Representation allows us to develop and implement the social constructionist approach (Hall, 1997) by favoring our understanding of the world and the relationships among members of a community.

More specific studies on the function of the image, in addition to the ontological role it plays (Banks & Morphy, 1999; Chaplin, 1996; Grady, 1996; Ruby, 1996), seek to materialize its effects at the social (Harper, 2012), or individual/psychological (Balçetis, 2010; Reavey, 2011) levels. Thus, on the sociological level, Grady (1996) points out the following items that images contribute to social studies and construction: «The first is Seeing, or the study of role of sight and vision in the construction of social organization and meaning. The second is Iconic Communication, or the study of how spontaneous and deliberate construction of images and imagery communicate information and can be used to manage relationships in society. The third area I have termed Doing Sociology Visually and is concerned with how techniques of producing and decoding images can be used to empirically investigate social organization, cultural meaning and psychological processes» (p. 10).

These aspects, which are aimed at the influence images have over social constructs, are reinforced in the compendium of qualitative image studies by Reavey (2011), who argues that the visual is an integral part of the way culture operates and especially now, thanks to the mass media.

As a complement to this influence of the image under a sociological approach, at the psychological level visual imagery is appreciated for its «(...) impact on how people experience the world they live in» (Reavey, 2011, XVI). This is why it is hard to ignore the complexity of the visual elements that mediate between personal experience and context. Therefore, when we take seriously how people's experiences are created and the contexts in which they emerge, it becomes hard to ignore the rich complex of visual media through which experiences are made reality.

Thus, when faced with a global event that has altered and modified our way of life, limiting our freedom and our individuality, as has been the case with the COVID-19

pandemic, this research aims to detect which denotative elements people use to express their experiences in this context through their visual creations and to detect whether common patterns and ideas emerge among the creators.

In comparison to other pandemics, COVID-19 has been widely documented and shared thanks to the extensive use of social networks such as Instagram, Whatsapp, Tik-Tok, etc., using different typologies of digital creations, the subjective experiences of this event have been recounted. A new disease without visual baggage or background like we have been able to see with cancer, HIV, or mental disease in the history of contemporary art, and which has nevertheless produced a huge number of images around it (Garden, 2010, 2015).

As a result, online repositories and initiatives have emerged to collect visual creations around the COVID-19 pandemic, such as Covid Art Museum (CAM)<sup>1</sup>, the Alameda Art Center in Santiago<sup>2</sup> or the Reina Sofia Museum in Spain,<sup>3</sup> as well as virtual exhibitions such as the MURAL proposal by the Spanish National Museum of Anthropology.<sup>4</sup>

All these sources of information make it possible to acquire and analyze large quantities of images to learn about individual experiences and the most recurrent discourses about this world event.

### *The Power of the Image: Effects of Representation*

«Representation is an essential part of the process by which meaning is produced and exchanged between members of a culture. It does involve the use of language, of signs and images which stand for or represent things» (Hall, 1997, p. 15).

Representation entails a system that combines the codes that we are capable of interpreting through learning and perception on the one hand, and the construction of signs on the other, which—in our case—visually represent concepts through «(...) ways of organizing, clustering, arranging and classifying concepts, and of establishing complex relations between them» (Hall, 1997, p. 17). A language.

Thus, visual representation is a process of social, cultural and ideological mediation (Desai, 2000; Knochel, 2013), selection, classification and decontextualization of elements with information from reality, with which a meaning is constructed through the structuring of these representation elements. This process fosters the transmission of an informative and/or emotional message based on individual or collective experience, which will ultimately be broadcast to the world in a language that must be shared by a community.

We speak of construction because a representation cannot reflect reality, as we all have different experiences and interpretations of 'reality'. In the case of the pandemic, we have all shared the same framework, but our experiences and interpretations of events, of reality, show subjective variations in their representation/description, meaning that the image created as artistic practice «(...) has come to be replaced by presentation, by a form of presenting reality more than representing it, while art,

especially contemporary art, flees the mimesis to present a new reality» (Cano, 2018, p. 104). Thus, when it comes to the pandemic we cannot speak of representations, but rather of presentations and dynamic constructs.

Some research assumes «(...) that images are not surfaces of universal messages whereby a common sense interpretation becomes apparent to all who look on. Instead, images serve as filters of how we see and are seen in the world, tinting our perceptions about our peers, building a bricolage of our understanding and assumptions about the other, and looking back at us to inform our self-concept» (Knochel, 2013, p. 17). However, in the pandemic context we are in, can the message of a subjective representation of experience be universal?

### *Visual Semiotics: Elements for a Language of Representation*

The subjective world created by visual representation is perceived with material elements «(...) whose 'shape' includes all the meaning of the work» (Rettberg, 2018, p. 433), which constitutes a visual message. Following Visual Literacy (VL) theory, which defines itself as an interdisciplinary, multidisciplinary and multidimensional discipline for the development of skills to interpret and create visual messages through the use and manipulation of cultural elements (Rettberg, 2018), it is akin to a verbal and literary language, and as a language, the visual requires skills, competencies and learning to properly convey and perceive the message: «(...) VL involves cognitive functions such as critical viewing and thinking, imaging, visualizing, inferring as well as constructing meaning; but also communicating as well as evoking feelings and attitudes» (Cohn, 2020, p. 267).

As a language, visual representation is composed of graphic elements such as shapes, lines, colors, textures, etc.—letters in the linguistic context—primitive components that make up more complex elements (e.g. a face) - the word in traditional language. This «(...) involves an interaction between three primary structures, similar to the parallel architecture of linguistic systems (...): A meaning expressed by a modality (here: visual-graphic marks), which is organized using combinatorial grammatical structures» (Jackendoff cited in Cohn, 2020, p. 267).

These complex forms fulfill a function and become significant, depending on the cultural context; they contain meaning (concept) in themselves and the sign is established by combining them. The combination of signs in an image determines the message through a spatial structure of relationships. In the case of visual language, their correct perception and understanding will be determined by the degree of iconicity.

The visual features of images can be divided into general features (color, shape, texture, etc.) of domain-specific features like human faces (Sethi, Coman, & Stan, 2001). This division is similar in semantics, which establishes two levels: low-level description representation, which are the formal characteristics of the image: color, light, textures, etc., dependent on the concept that one seeks to transmit (Bloehdorn et al., 2005); and the highest level, which is the meaning (the concept) (Liu, Zhang, Lu, & Ma, 2007), established by the semantic relations between elements.

Both, the low-level and high-level features of the images, have been extensively analyzed (Joshi et al., 2011; Liu, Lughofer, Zeng, & Li, 2018; Sowmyayani & Rani, 2022; Tadi Bani & Fekri-Ershad, 2019; Zhang, He, & Lu, 2020, 2022; Zhao et al., 2021), and artistic creation has been a field of study for this (Achlioptas, Ovsjanikov, Haydarov, Elhoseiny, & Guibas, 2021; Fekete et al., 2022; Machajdik & Hanbury, 2010; Stamatopoulou & Cupchik, 2017; Yanulevskaya et al., 2012; Zhang et al., 2011; Zhao et al., 2014).

In our research, we analyze signs at the denotative - descriptive - level contents and representations in the images developed by users on the Internet to identify the signs, and combinations used by users in their visual representations based on the pandemic. Thus, like written or spoken language, we can discover the most-used words and phrases and reach a more connotative level of the expressions of shared feelings during the pandemic on the part of a group of people at the global level and compiled on the social network Instagram.

«Connotations are common associations connected to a sign, not private associations that only one individual might have, but associations and references that are shared by larger cultures or groups» (Rettberg, 2018, p. 433).

### *Health Humanities. Incorporation of a Visual Narrative as Discursive Element of the Disease*

Health humanities is an interdisciplinary field of study that seeks to incorporate a humanistic perspective to health and healthcare. Humanities (communications, cultural studies and linguistics, history, literature, philosophy, theology and religious studies and arts) are indispensable to human healthcare as they teach us what it means to be human. They teach us about the human condition, about human suffering and healing, and about human well-being and flourishing.

In a systematic review, the authors Costa, Kangasjarvi, and Charise (2020) provide us with the following data on the benefits of applying the humanities to the field of health: «(...) proponents of 'medical humanities' have argued that by encouraging attention to less overtly clinical aspects of healthcare, humanities-engaged clinical education helps build social and relational skills, often by fostering empathy and moral development; compassion; self-reflection; interpersonal skills; perspective-taking; openness to otherness; critical reflection; and tolerance for ambiguity.» (p. 1204).

In reference to representation, and specifically social representation, Moscovici (1984) argues there are three fundamental types: emancipated, controversial and hegemonic. These types of representation are also part of the metaphorical developments of different social constructs. In the case we are concerned with, it could be the social representation of the COVID-19 pandemic. Art can be situated in each of these, exercising different functions in the construction of a narrative of events, in addition to the dissemination and empowerment of some discourse over others. An example of this, regarding the second type described by Moscovici (hegemonic social representations),

was presented by Alfonso Baya (2015) in his thesis on the imagery of AIDS and contributions from Art, where he points out the main hegemonic representations: «The visual elements that the pictorial images have offered of disease over the course of history have been allegories of the disease in itself, such as JF Bocanegra's 'Allegory of the Plague' in the 1660s. The figure of the doctor visiting the sick, such as "Doctor's Visit" by J Steen (1625–79), or the terminally ill in bed. Other representations are directly inclined toward the representation of death, either as an allegory or as a human still life without action, movement or life» (p. 71). This leads us to wonder whether the narrative that is being created of the pandemic is putting forward certain hegemonic constructs of the same, above the display of the diversity of existing social realities; whether it is oriented toward the dissemination of images more aimed at controversy and the confrontation of opinions on the disease and its handling in different aspects, or whether it is fulfilling an emancipating function instead, where art is presented as a solution to the circumstances that the pandemic places us in.

### *Computer Analysis of Visual Representation*

From a behavioral perspective, studies indicate that visual perception of art occurs in a differentiated way among groups of artists and non-artists (Solso, 2000; Zeki, 2003). Human vision is not just perceived in terms of the visual aspect of the work of art, but rather there is an understanding that takes the artist beyond the work and even beyond the esthetic terms that might be represented in it (Ramachandran & Hirstein, 1999). This means that artists not only paint through their eyes, but rather through an understanding and interpretation of what is happening in their brains (Goguen & Myin, 1999; Quiroga Méndez, 2016; Zeki, 2003).

Thus, the question emerges as to whether it is possible for a computer to capture the nature of a painting. That is, what we experience in human form as the capacity for imagination and connection with art. The first works aimed at understanding art from the field of computer sciences were developed over two decades ago, thanks to the initial work of Taylor, Micolich, and Jonas (1999) to authenticate the works of art by Jackson Pollock. For their part, Widjaja, Leow, and Wu (2003) proposed the combination of multiple classifiers of four artists through skin pattern analysis. Li and Wang (2004) developed an algorithm for the automatic identification of Chinese ink paintings through wavelet analysis and MHMM (Multiresolution Hidden Markov Models) tools. Johnson et al. (2008) proposed an algorithm that uses a set of wavelet filters on different scales to extract patterns from Van Gogh's and other paintings (not Van Gogh). Shamir, Macura, Orlov, Eckley, and Goldberg (2010) proposed an automatic recognition system that differentiates the works of three types of art school: impressionism, abstract expressionism, and surrealism. For this, they used a set of algorithms to extract characteristics based on textures, statistical moments, Gabor filters and the analysis of edges, among others. As the literature describes, all the techniques prior to 2012 use the manual extraction of low-level characteristics according to chromatic, geometric, or statistical descriptors, which in

general perform poorly as they do not extract the underlying information from the styles contained in the images.

Though neural networks (Zhang, 2000) have been present for more than 60 years, it was not until the methodology proposed by Krizhevsky, Sutskever, and Hinton (2017) that it marked a paradigm shift by implementing a convolutional neural network (CNN) with a more complex and operational architecture, which allowed automatic extraction of patterns from images without requiring the work of manual characteristic extraction. This last tool has been a major revolution, as CNNs have made it possible to extract mid-level properties present in images due to the internal process of the network itself. In the years that followed, exhaustive analyses of different sets of data have shown the great flexibility and power of extracting patterns from images, which can even be transferred even to other domains (Oquab, Bottou, Laptev, & Sivic, 2014; Razavian, Azizpour, Sullivan, & Carlsson, 2014; Wang & Deng, 2018; Westlake, Cai, & Hall, 2016).

Currently extensive databases have been developed that are composed of thousands or millions of images of digital photographs with their corresponding objects classified (i.e., SUN, PascalVOC, ImageNet, Google Image Database), which have allowed major progress to be made in identifying, recognizing, and classifying images and their objects (Russakovsky et al., 2015). This does not normally occur in artistic paintings where different styles are present and these change over time, in addition to the limited universe of artistic images compared to digital images like photographs of everyday objects (Crowley & Zisserman, 2014; Garcia & Vogiatzis, 2019; Hong & Kim, 2019; Ren, He, Girshick, & Sun, 2017). This leads to substantial differences in the statistics associated with the objects represented in the images. However, thanks to the development of CNN in digital photographs, it has been possible to transfer this knowledge to the area of art and paintings, in this way overcoming the adaptation of domain (Kulis, Saenko, & Darrell, 2011; Wang & Deng, 2018). An example of this is found in Karayev et al. (2014) work, which is aimed at differentiating the different esthetics not just of art, but also of digital images where different photographic techniques, compositions styles, moods, genders, and types of scenes can be distinguished. The same is the case with the work by Crowley and Zisserman (2015), who managed to detect and align objects in paintings based on their learning of the same using natural images (i.e., not paintings). In the same way, Westlake et al. (2016) makes progress in the study of artistic images, especially with the creation of a database of specialized in people with over 43 styles of painting. He emphasizes the use of learning transfer from the ImageNet network, modifying the output network. The study reveals the complexity that an identification algorithm faces in detecting a given study subject. In the case of detecting people, this can be done with multiple art styles and even so, the algorithm can identify their spatial location within the image (see Figure 1). This is a complex task given the countless types of configuration, form, pose, size, and abstraction that exist in the artists' interpretations.

In recent years, identification tools based on Deep Learning architectures (Sandoval, Pirogova, & Lech, 2019; Tan, Chan, Aguirre, & Tanaka, 2016) have evolved and significantly improved accuracy, speed and performance on issues related to identification



**Figure 1.** Semantic identification in different styles of painting on which the algorithm identifies the spatial location. Green box on image (Westlake et al., 2016).

of regions, known as region multi-state detectors. This category includes the following frameworks: SPP-net (spatial pyramid pooling) (He, Zhang, Ren, & Sun, 2014), R-FCN (region-based fully convolutional network) (region-based fully convolutional network) (Dai, Li, He, & Sun, 2016), FPN (Feature pyramid networks) (Lin et al., 2017), Fast R-CNN (Ren et al., 2017), Faster R-CNN (Ren et al., 2017), Mask R-CNN (He, Gkioxari, Dollar, & Girshick, 2017) and R-CNN (Zhao, Zheng, Xu, & Wu, 2019). Despite the wide variety of identification and classification tools, one of the architectures that has performed the best, especially with abstract images, has been the “You Only Look Once” (YOLO) network (Westlake et al., 2016). YOLO, proposed by Redmon, Divvala, Girshick, and Farhadi (2016), is a simplified network compared to previous architectures, requiring only reducing the size of the entire image and processing it directly over the network. Other networks usually require a preliminary process that consists in scrolling over the image with a sliding window and then using an adjustment process to find objects of interest along with adjusting the boundaries of each object found. In addition, YOLO processes the image by dividing it into a rectangular grid and then searching for and grouping quadrants together as the likelihood of finding an object increases. This process works globally, thus increasing the general performance in object detection. This means that the probabilities of the objects found in the image are simultaneously predicted and adjusted. For YOLO to operate, there is a training stage that the developers themselves have provided for, though it can be trained for a specific set of images (Burić, Pobar, & Ivašić-Kos, 2019; Dutta & Biswas, 2019; Ju, Wang, & Chang, 2019). YOLOv4 is currently the most accurate and fastest network among current identification tools (Bochkovskiy, Wang, & Liao, 2020; Sumit, Watada, Roy, & Rambli, 2020) and was used as the key object identification tool in this research.

## Materials and Methods

This article analyzes visual representations arising from the COVID-19 pandemic, using the Instagram profile entitled The Covid Art Museum (CAM), considered the first museum dedicated to this health emergency<sup>5</sup>, collected art produced during the pandemic, to give visibility to confined artists and promote culture when 95% of the world’s museums were closed<sup>6</sup>.



This online source of information produced by Internet users allows for the acquisition of an important representation of different expressions and discourses. Its availability on the network facilitates its access, extraction, and use for its processing and analysis through the methodology proposed in this research.

The period for the mass download of images ranges from the date that the CAM profile was created (19 March 2020) to 4 August 2020, with a total of 927 images obtained, and no selection criteria were applied (Figure 2).

### Computer Analysis

Current techniques based on Deep Learning algorithms generally have a limited number of objects that the network has been trained for. These networks generally work with objects or representations from the physical world and not necessarily abstract objects like those commonly found in art. Though there is a growing body of research related to the identification of elements and/or representations in artistic expressions, it is always done based on a limited and defined set.

The images used in our study represent a mix of physical objects and/or abstract representations. Analyzing this type of images is complex, especially due to the process of abstraction that the artist defines in his or her composition. The methodology proposed



**Figure 2.** Sample images from Instagram’s COVID Art museum account (image 1: Hojin Kang, @hojinkangdotcom; image 2: @Bix.rex.1; Image 3: @marciorodriguezphoto; Image 4: @rebordacao, Rebordação Embroidery; Image 5: Orane Tasky @oranetasky; Image 6: Paz Peralta, Title: The Fridge, @notpauz).

seeks to find the relationship between the objects in each image to find a general ontology that brings together the concepts contained in each image and how they are related to the current context of the COVID-19 pandemic. To facilitate this search, there was a need to redefine the classes as a function of the objects they contain, thus facilitating the networks learning process.

The analysis proposal is made up of four phases, which we detail below (Figure 3). The first phase consists in preparing the data. In this phase each of the objects to be detected are manually identified in the following categories:

Categories = { dog, person, cat, TV, car, outside, plants, books, window inside, bird, closed curtains, bicycle, plate, banana, ketchup, cup, hanging clothes, house, rubber gloves, garlic, vase, mask, shopping cart, groceries, toilet paper, abstract mask, handwashing, faucet, virus, nudity, kiss, phone, flower, fruits, wine, bathtub, bread, toilet paper pile, sanitizer, snacks, game controller, empty toilet, paper roll, soap, laptop, knife, bed, balcony, disinfecting wipes, bleach, soap bar, pen, syringe, motorcycle, thermometer, crowd, city landscape, street, buildings, bathroom, supermarket, soap dispenser, dish soap, videochat }

The process of choosing the categories was inductive. It began with the compilation of objects contained in the images and associating them with more general concepts. We defined 64 categories for our study, broad enough to detect a large proportion of the elements contained in the images.

A rectangular region is identified for each object and associated with one of the categories listed above. In some cases, due to the large number of objects contained in the same image, we have manually identified 60% of the objects contained in it. That is, if an image contains more versions of the same object, then only 60% of these are tagged.

In the second phase (training), we used a YOLOv4 convolutional neural network. The data was divided into three groups in this phase: training (70%), validation (20%), and test (10%). In this way, all the previously identified objects are trained in the network to be automatically classified in subsequent steps. As results, we put

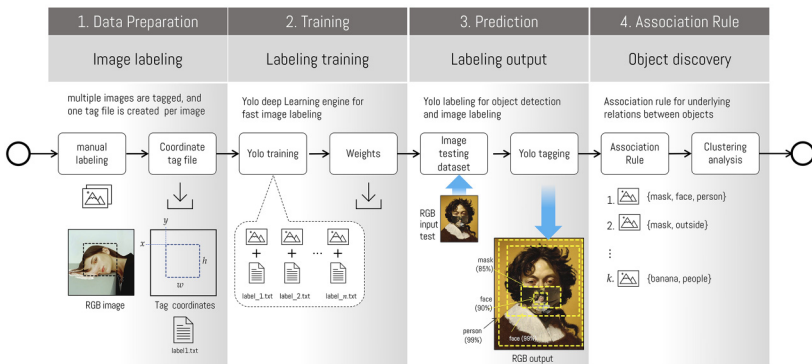


Figure 3. Process of preparing and identifying objects and their relationships.

forward a set of network training weights specific to the set of images of the problem. In the third phase, the network is assessed on the same set of training images. Though we know that normally we seek to find objects in images that have not previously been trained, it is important to note that not all the objects were tagged in many of the training set's images. In this way we increased the number of objects detected in the training images. In contrast, we did not use YOLO as a prediction tool, but rather as a complementary tool for the object identification process.

Lastly, the fourth phase uses all the objects detected in each image as a knowledge base. To find a relationship among these and their relationship with different images, we used an association rule algorithm (Agrawal, Mannila, Srikant, Toivonen, & Verkamo, 1996) that allows the ontology in the images to be defined. In this way, in addition to finding the frequency of the objects' appearance in each image, the rules of association allow us to determine the underlying relationship in objects that do not initially seem to be related but which are, through the search of patterns that group the relationship together. To facilitate the visual analysis, we used the graphic representation of the rules of association, in which each node represents an object from the set of study images and the arrow indicates the relationship between antecedent items (LHS) and consequent items (RHS). For example, if an object {buildings, street} repeatedly appears in an image, then it implies a certain level of probability that the object {outside} will also be present in the image. In this way, the algorithm can build a relationship according to two indicators: Support and confidence, where support determines the percentage of relationships that appear in the dataset and Confidence is the frequency with which consequent items appear along with antecedent ones.

## Results

The sample analysis reveals that the most common elements used to construct narratives are people, an outside view, an interior scene, face masks, windows, vegetable elements and buildings (see Figure 4).

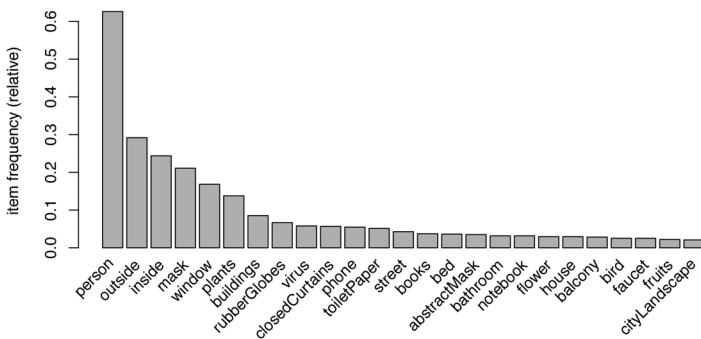
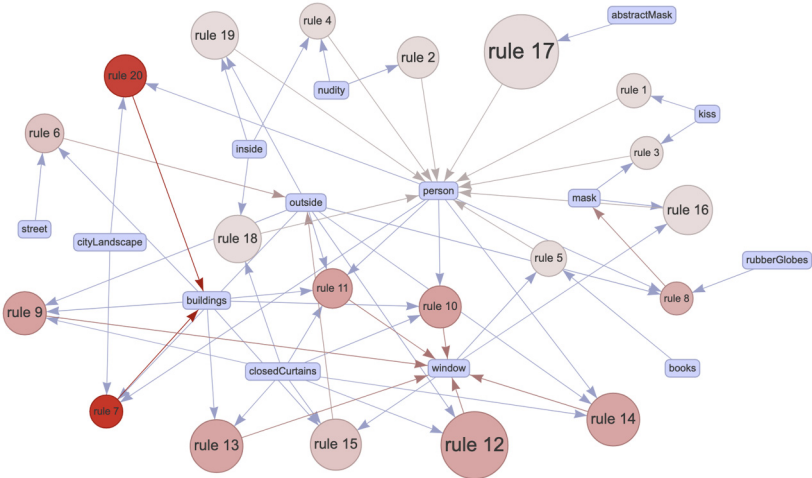


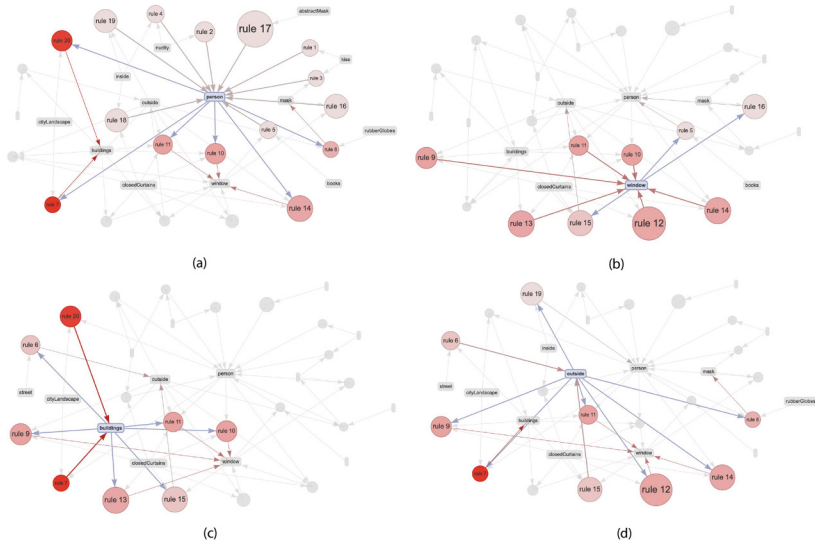
Figure 4. Main frequencies of objects in the set of images.



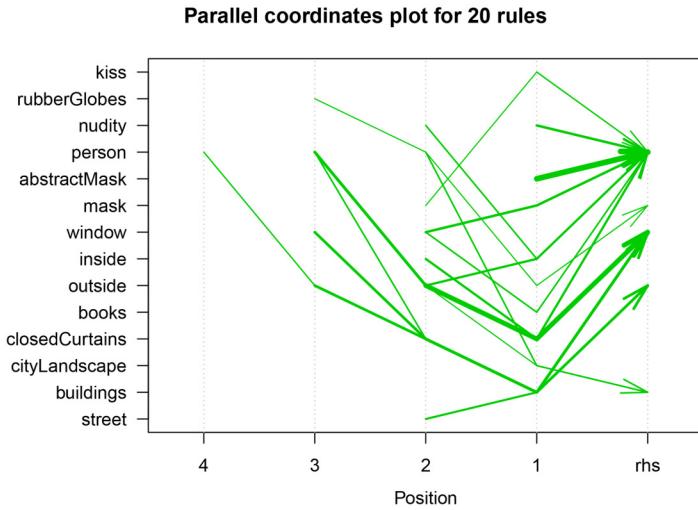
**Figure 5.** A priori relationship between the objects found within the same image.

**Table 1.** Result of the a priori Algorithm for the First 20 Rules of Association Ordered According to Confidence: LHS (Antecedent), RHS (Consequent).

ID	LHS (Antecedent)	RHS (Consequent)	Support	Confidence	Lift
1	{kiss}	{person}	0.0120	1.0000	1.597
2	{nudity}	{person}	0.0175	1.0000	1.597
3	{kiss,mask}	{person}	0.0109	1.0000	1.597
4	{inside,nudity}	{person}	0.0131	1.0000	1.597
5	{books>window}	{person}	0.0131	1.0000	1.597
6	{buildings,street}	{outside}	0.0153	1.0000	3.427
7	{cityLandscape,outside,person}	{buildings}	0.0109	1.0000	11.731
8	{outside,person,rubberGloves}	{mask}	0.0109	1.0000	4.741
9	{buildings,closedCurtains,outside}	{window}	0.0186	1.0000	5.942
10	{buildings,closedCurtains,person}	{window}	0.0131	1.0000	5.942
11	{buildings,closedCurtains,outside,person}	{window}	0.0120	1.0000	5.942
12	{closedCurtains,outside}	{window}	0.0284	0.9630	5.722
13	{buildings,closedCurtains}	{window}	0.0197	0.9474	5.629
14	{closedCurtains,outside,person}	{window}	0.0197	0.9474	5.629
15	{buildings,closedCurtains>window}	{outside}	0.0186	0.9444	3.237
16	{mask>window}	{person}	0.0175	0.9412	1.503
17	{abstractMask}	{person}	0.0328	0.9375	1.497
18	{closedCurtains,inside}	{person}	0.0164	0.9375	1.497
19	{inside,outside}	{person}	0.0164	0.9375	1.497
20	{cityLandscape,person}	{buildings}	0.0131	0.9231	10.828



**Figure 6.** Four main antecedents with the largest number of rules of association between antecedents and consequents.



**Figure 7.** Parallel representation graph (paracord) identifying the main relationships that connect with RHS.

However, frequency-based analysis only allows identifying the objects that appear most frequently in the set of images, but it does not identify the combinations between the elements represented, which constitute the decision-making to build the message, like the sum of words in a sentence.

This limitation has been resolved with a priori analysis, which allows finding the relationships between multiple objects, as shown in Table 1, ordered according to confidence.

Figure 5 below graphically shows the map of association relationships between objects in the same image from sample established in Table 1, where the radius of the circle represents the support of the measure and the intensity of color.

The four elements represented with the most antecedent and consequent rules of association in the sample are: a) Person, b) Window, c) Building, d) Outside. (See Figures 6 and 7).

## **Discussion and Conclusions**

The study of the denotative elements that make up the images produced in this context, obtained from The Covid Art Museum account on the social network Instagram, has helped us to understand how we associate a series of objects and images on a concrete reality, establishing a visual representation of the same.

Connecting with the main idea of how visual representation is based on the selection, classification and decontextualization of elements to forge a meaning through the way they are structured, computer analysis has revealed the elements most used and their associations.

We have been able to see that, through quantification in terms of the frequency of the objects contained in the images, the isolated elements most commonly used by creators for their visual representations were people, an outside view, an interior scene, face masks, windows, vegetable elements and buildings. In this regard, it is worth highlighting the high frequency with which the object “face mask” appears, how a foreign element in our lives appears in representations associated with the pandemic and has incorporated itself into collective imagery.

However, the frequency study does not allow one to know the use of the elements and their relationships in creating holistic discourses on the pandemic. For this, a methodology of graphic representation of rules of association was applied, to automatically establish the underlying relationships between objects detected in the data. This is a more statistically valid study and allows relationships to be obtained that would not normally be detected by an observer or critic.

Those rules show us that the use of people, windows, buildings, and outside scenes are the elements that have related the most with others to construct the visual representation of subjective experiences in the context of the COVID-19 pandemic.

Thus, despite the subjective discourses and creations, the representations share certain patterns and ideas when it comes to choosing the denotative elements, where the human figure focuses the discourse and the window element is detected as the

link with confinement, the restriction most emphasized by creators as a symbol of the limitation of movement, in addition to representations of buildings, the habitat where the experience has taken place, and the representation of outdoor spaces as a representation of the desire to regain freedom. In this way, one can see that the construction of visual narratives among the images, in terms of associations, is produced by means of these four denotative elements (people, buildings, windows and outdoor scenes) and this has generated visual discourses, not so much on the disease, but its impact on the collateral damage that has been caused and the changes in our everyday dynamics, beyond the physical, economic, social damage, etc.

Thus, no significant uses of symbolic or iconographic elements traditionally associated with disease were detected (e.g., a skull as a reference to death, the prostrate and weak individual, the clock as a reference to the time that has passed, etc.), iconographic features that have even been analyzed in relation to previous pandemics in the works of Ana María Díaz Perez (1991) or, more recently, that of Carla Joan Dias de Angelo (2018). We also could not find images of saturated hospitals, deaths, rising poverty or increased social tension, or they are not representative in the sample.

Based on the results, we can conclude that there are a series of shared patterns and images that are repeated, each one originating in different authors, thus revealing a shared collective imagery, to a certain extent inherent to a globalized society that has experienced a common pandemic, and the use of social networks as a global channel of expression, which have constituted a visual imagery but with limited use of resources and codes in creations.

This research shows how social networks have become a channel for expressing and communicating our everyday lives, becoming a source of research to understand contemporary phenomena, realizing the potential of interdisciplinary research in the digital humanities to achieve new knowledge.

## Limitations and Future Research

Recent research by Elena Semino (2020) on the metaphors have been generated at the narrative level around the pandemic establishes that: «Metaphors are crucial tools for communication and thinking, and can be particularly useful in public health communication. For example, the ‘second wave’ metaphor suggests that there is renewed danger and threat from the virus, and may therefore encourage compliance with measures aimed at reducing transmission» (p.5).

This metaphor is a representative example of how the pandemic has been described at the narrative level and how through the metaphor it is possible to build a concrete reality. In this sense, it is interesting to note how at the visual level, from a freer and more personal position like artistic discourse, metaphors have been built in a very different way from the way done in verbal discourse. Thus, ‘waves’ or ‘battlefields’—narrative metaphors described in Semino (2020)—give way to other narratives, which comprise a field to be explored in future research.

The methodology proposed in our research has allowed relations between the objects that underly individual representations to be empirically established. Based on the aforementioned conclusions, there is significant room for work to extract the vast amount of information contained in images, such as the color, shape, spatial location of elements, among others, and in that way study the process that leads artists to define the elements contained in a work.

The results obtained will be compared with other artworks from the history of art that deal with other pandemics or diseases to detect similarities and differences in the visual elements, and thus understand the strategies used in each historical moment.

Regarding the YOLO tool, which distinguishes at least 80 categories natively, our research has proposed detecting 64 categories other than those already defined by YOLO, taking the network structure and training weights as the basis. We will consider reducing this number in future improvements, grouping similar concepts together, thus facilitating the search for shared patterns among the different artistic representations and, in particular, expanding the characteristics extracted from the images that allow us to find other new relations on the object categories.

By using digital creations extracted from the Internet, a high level of geographical representation is assumed, but this is unknown. From a more *netnographic* approach, the demographics of the images and/or the users could be used to analyze the level of globalization in the use of visual codes in the representations of the pandemic, or, on the contrary, whether there are significant differences according to the country of origin.

Differences in creative strategies may be produced by cultural context, but in the context of the COVID-19 pandemic, perhaps the health policies applied in each country could have influenced the visual representation of this event. This may be an important area of research within the health humanities.

### **Declaration of Conflicting Interests**

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

### **Funding**

The author(s) received no financial support for the research, authorship, and/or publication of this article.

### **ORCID iD**

César González-Martín  <https://orcid.org/0000-0002-9017-3196>

### **Notes**

1. See more in <https://www.instagram.com/covidartmuseum/>
2. See more in <https://www.centroartealameda.tv/>
3. See more in <https://www.lavanguardia.com/cultura/20200416/48541348579/exposicion-virtual-museo-refina-sofia-pandemia-fotos-clemente-bernad.html>



4. See more in <http://www.culturaydeporte.gob.es/mnantropologia/exposvirtuales/mural.html>
5. See more in: <https://www.lavanguardia.com/magazine/experiencias/20200402/48247199504/covid-19-coronavirus-instagram-arte-barcelona-museo-covid.html>
6. See more in: <https://en.unesco.org/news/response-chiles-museums-cultural-impact-covid-19>

## References

- Achlioptas, P., Ovsjanikov, M., Haydarov, K., Elhoseiny, M., & Guibas, L. (2021). ArtEmis: Affective Language for Visual Art. *ArXiv:2101.07396 [Cs]*. <http://arxiv.org/abs/2101.07396>.
- Agrawal, R., Mannila, H., Srikant, R., Toivonen, H., & Verkamo, A. I. (1996). Fast discovery of association rules. *Advances in knowledge discovery and data mining* (pp. 307–328). Menlo Park, California, USA: American Association for Artificial Intelligence.
- Aparici, R. (2009). *La imagen: Análisis y representación de la realidad*. <http://site.ebrary.com/id/11322999>.
- Balçetis, E. (Ed.). (2010). *Social psychology of visual perception*. USA: Psychology Press.
- Banks, M., & Morphy, H. (Eds.). (1999). *Rethinking visual anthropology (1st publ. In paperback)*. UK: Yale Univ. Press.
- Baya, A. (2015). *El imaginario del sida en la cultura visual: Construcción, significación y cuestionamiento a través de estrategias artísticas* [Doctoral dissertation, Universidad de Granada]. <https://digibug.ugr.es/handle/10481/49417>.
- Bloehdorn, S., Petridis, K., Saathoff, C., Simou, N., Tzouvaras, V., ..., M. G. Strintzis (2005). Semantic annotation of images and videos for multimedia analysis. In A. Gómez-Pérez, & J. Euzenat (Eds.), *The semantic web: Research and applications* (Vol. 3532, pp. 592–607). Berlin, Heidelberg: Springer. [https://doi.org/10.1007/11431053\\_40](https://doi.org/10.1007/11431053_40).
- Bochkovskiy, A., Wang, C.-Y., & Liao, H.-Y. M. (2020). YOLOv4: Optimal Speed and Accuracy of Object Detection. *ArXiv:2004.10934 [Cs, Eess]*. <http://arxiv.org/abs/2004.10934>.
- Burić, M., Pobar, M., & Ivašić-Kos, M. (2019). Adapting YOLO Network for Ball and Player Detection: *Proceedings of the 8th International Conference on Pattern Recognition Applications and Methods*, 845–851. <https://doi.org/10.5220/0007582008450851>.
- Cano-Martínez, M. (2018). *Escondido tras la piel. Representaciones y afrontamientos del dolor y el sufrimiento desde el Arte de acción* [Doctoral dissertation, Universidad de Granada]. <https://digibug.ugr.es/handle/10481/52552>.
- Carla, J. D. A. d. S. (2018). Muerte y representación en la Edad Media: consideraciones sobre la imagen, la iconografía de la muerte y la influencia de la Peste Negra en el surgimiento de los temas macabros. *De Medio Aevo*, 12(1), 239–258.
- Chaplin, E. M. (1996). *Sociology and visual representation (reprint)*. London: Routledge.
- Cohn, N. (2020). Visual narrative comprehension: Universal or not? *Psychonomic Bulletin & Review*, 27(2), 266–285. <https://doi.org/10.3758/s13423-019-01670-1>.
- Costa, M., Kangasjarvi, E., & Charise, A. (2020). Beyond empathy: A qualitative exploration of arts and humanities in pre-professional (baccalaureate) health education. *Advances in Health Sciences Education*, 25(5), 1203–1226. <https://doi.org/10.1007/s10459-020-09964-z>.
- Crowley, E., & Zisserman, A. (2014). The state of the art: Object retrieval in paintings using discriminative regions. *Proceedings of the British Machine Vision Conference, 2014*, 38.1–38.12. <https://doi.org/10.5244/C.28.38>.
- Crowley, E., & Zisserman, A. (2015). In search of art. In L. Agapito, M. M. Bronstein, & C. Rother (Eds.), *Computer vision—ECCV 2014 workshops* (Vol. 8925, pp. 54–70). Cham: Springer International Publishing. [https://doi.org/10.1007/978-3-319-16178-5\\_4](https://doi.org/10.1007/978-3-319-16178-5_4).

- Dai, J., Li, Y., He, K., & Sun, J. (2016). R-FCN: Object detection via region-based fully convolutional networks. *Proceedings of the 30th International Conference on Neural Information Processing Systems*, Barcelona, Spain, 2016.
- Desai, D. (2000). Imaging difference: The politics of representation in multicultural art education. *Studies in Art Education*, 41(2), 114. <https://doi.org/10.2307/1320658>.
- Dutta, A., & Biswas, S. (2019). CNN based Extraction of Panels/Characters from Bengali Comic Book Page Images. *ArXiv:1910.09233 [Cs]*. <http://arxiv.org/abs/1910.09233/objidref>>.
- Fekete, A., Pelowski, M., Specker, E., Brieber, D., Rosenberg, R., & Leder, H. (2022). The Vienna Art Picture System (VAPS): A data set of 999 paintings and subjective ratings for art and aesthetics research. *Psychology of Aesthetics, Creativity, and the Arts*. No Pagination Specified-No Pagination Specified. <https://doi.org/10.1037/aca0000460>
- Garcia, N., & Vogiatzis, G. (2019). How to read paintings: Semantic art understanding with multi-modal retrieval. In L. Leal-Taixé, & S. Roth (Eds.), *Computer vision – ECCV 2018 workshops* (Vol. 11130, pp. 676–691). Cham: Springer International Publishing. [https://doi.org/10.1007/978-3-030-11012-3\\_52](https://doi.org/10.1007/978-3-030-11012-3_52).
- Garden, R. (2010). Disability and narrative: New directions for medicine and the medical humanities. *Medical Humanities*, 36(2), 70–74. <https://doi.org/10.1136/jmh.2010.004143>.
- Garden, R. (2015). Who speaks for whom? Health humanities and the ethics of representation. *Medical Humanities*, 41(2), 77–80. <https://doi.org/10.1136/medhum-2014-010642>.
- Goguen, J., & Myin, E. (1999). *Art and the brain*. UK: Imprint Academic.
- Grady, J. (1996). The scope of visual sociology. *Visual Sociology*, 11(2), 10–24. <https://doi.org/10.1080/14725869608583762>.
- Hall, S. (1997). *Representation and the media*. <https://www.mediaed.org/transcripts/Stuart-Hall-Representation-and-the-Media-Transcript.pdf>.
- Harper, D. A. (2012). *Visual sociology*. London: Routledge.
- He, K., Gkioxari, G., Dollár, P., & Girshick, R. (2017). Mask R-CNN. 2017 *IEEE International Conference on Computer Vision (ICCV)*, 2980–2988. <https://doi.org/10.1109/ICCV.2017.322>.
- He, K., Zhang, X., Ren, S., & Sun, J. (2014). Spatial Pyramid Pooling in Deep Convolutional Networks for Visual Recognition. *ArXiv:1406.4729 [Cs]*, 8691, 346–361. [https://doi.org/10.1007/978-3-319-10578-9\\_23](https://doi.org/10.1007/978-3-319-10578-9_23).
- Hong, Y., & Kim, J. (2019). Art painting detection and identification based on deep learning and image local features. *Multimedia Tools and Applications*, 78(6), 6513–6528. <https://doi.org/10.1007/s11042-018-6387-5>.
- Johnson, C., Hendriks, E., Bereznoy, I., Brevdo, E., Hughes, S., & Daubechies, I., ... Wang J. (2008). Image processing for artist identification. *IEEE Signal Processing Magazine*, 25(4), 37–48. <https://doi.org/10.1109/MSP.2008.923513>.
- Joshi, D., Datta, R., Fedorovskaya, E., Luong, Q.-T., Wang, J. Z., Li, J., & Luo, J. (2011). Aesthetics and emotions in images. *IEEE Signal Processing Magazine*, 28(5), 94–115. <https://doi.org/10.1109/MSP.2011.941851>.
- Jovchelovitch, S. & Routledge. (2019). *Knowledge in context: Representations, community and culture*. London: Routledge Taylor & Francis Group.
- Ju, L., Wang, H., & Chang. (2019). The application of improved YOLO V3 in multi-scale target detection. *Applied Sciences*, 9(18), 3775. <https://doi.org/10.3390/app9183775>.
- Karayev, S., Trentacoste, M., Han, H., Agarwala, A., Darrell, T., Hertzmann, A., & Winnemoeller, H. (2014). Recognizing Image Style. *ArXiv:1311.3715 [Cs]*. <http://arxiv.org/abs/1311.3715>.

- Knochel (2013). Assembling visuality: Social Media, everyday imaging, and critical thinking in digital visual culture. *Visual Arts Research*, 39(2), 13. <https://doi.org/10.5406/visuartsrese.39.2.0013>.
- Krizhevsky, A., Sutskever, I., & Hinton, G. E. (2017). Imagenet classification with deep convolutional neural networks. *Communications of the ACM*, 60(6), 84–90. <https://doi.org/10.1145/3065386>.
- Kulis, B., Saenko, K., & Darrell, T. (2011). What you saw is not what you get: Domain adaptation using asymmetric kernel transforms. *CVPR, 2011*, 1785–1792. <https://doi.org/10.1109/CVPR.2011.5995702>.
- Li, J., & Wang, J. Z. (2004). Studying digital imagery of ancient paintings by mixtures of stochastic models. *IEEE Transactions on Image Processing*, 13(3), 340–353. <https://doi.org/10.1109/TIP.2003.821349>.
- Lin, T.-Y., Dollar, P., Girshick, R., He, K., Hariharan, B., & Belongie, S. (2017). Feature Pyramid Networks for Object Detection. 2017 *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 936–944. <https://doi.org/10.1109/CVPR.2017.106>.
- Liu, J., Lughofer, E., Zeng, X., & Li, Z. (2018). The power of visual texture in aesthetic perception: An exploration of the predictability of perceived aesthetic emotions. *Computational Intelligence and Neuroscience*, 2018, 1812980. <https://doi.org/10.1155/2018/1812980>.
- Liu, Y., Zhang, D., Lu, G., & Ma, W.-Y. (2007). A survey of content-based image retrieval with high-level semantics. *Pattern Recognition*, 40(1), 262–282. <https://doi.org/10.1016/j.patcog.2006.04.045>.
- Machajdik, J., & Hanbury, A. (2010). Affective image classification using features inspired by psychology and art theory. *Proceedings of the 18th ACM International Conference on Multimedia*, Firenze, Italy, 2010.
- Moscovici, S. (1984). The phenomenon of social representations. In R. Farr, & S. Moscovici (Eds.), *Social representations* (pp. 3–69). Cambridge: Cambridge University Press.
- Oquab, M., Bottou, L., Laptev, I., & Sivic, J. (2014). Learning and transferring mid-level image representations using convolutional neural networks. 2014 *IEEE Conference on Computer Vision and Pattern Recognition*, 1717–1724. <https://doi.org/10.1109/CVPR.2014.222>.
- Perez, A. M. D. (1991). *Iconografía de los santos protectores de epidemias y enfermedades en Canarias* [Doctoral dissertation]. Universidad de la Laguna.
- Quiroga Méndez, M. D. P. (2016). Dar un contenido al tiempo, una interpretación arquetipal del arte en la postmodernidad. *Arte, Individuo y Sociedad*, 28(3), 567–582. [https://doi.org/10.5209/rev\\_ARIS.2016.v28.n3.51614](https://doi.org/10.5209/rev_ARIS.2016.v28.n3.51614).
- Ramachandran, V., & Hirstein, W. (1999). The science of art: A neurological theory of aesthetic experience. *Journal of Consciousness Studies*, 6, 15–51.
- Razavian, A. S., Azizpour, H., Sullivan, J., & Carlsson, S. (2014). CNN Features off-the-shelf: An Astounding Baseline for Recognition. *ArXiv:1403.6382 [Cs]*. <http://arxiv.org/abs/1403.6382>.
- Reavey, P. (Ed.). (2011). *Visual methods in psychology: Using and interpreting images in qualitative research*. NY: Routledge.
- Redmon, J., Divvala, S., Girshick, R., & Farhadi, A. (2016). You only Look once: Unified, real-time object detection. 2016 *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 779–788. <https://doi.org/10.1109/CVPR.2016.91>.
- Ren, S., He, K., Girshick, R., & Sun, J. (2017). Faster R-CNN: Towards real-time object detection with region proposal networks. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 39(6), 1137–1149. <https://doi.org/10.1109/TPAMI.2016.2577031>.

- Rettberg, J. W. (2018). Self-representation in social media. In J. Burgess, A. Marwick, & T. Poell (Eds.), *The SAGE handbook of social media* (pp. 429–443). London: Sage Publications Ltd. <https://doi.org/10.4135/9781473984066.n24>.
- Ruby, J. (1996). Visual anthropology. In D. Levinson, & M. Ember (Eds.), *Encyclopedia of cultural anthropology* (Vol. 4, pp. 1345–1351). New York: Henry Holt and Company.
- Russakovsky, O., Deng, J., Su, H., Krause, J., Satheesh, S., ... Fei-Fei, L. (2015). Imagenet large scale visual recognition challenge. *International Journal of Computer Vision*, 115(3), 211–252. <https://doi.org/10.1007/s11263-015-0816-y>.
- Sandoval, C., Pirogova, E., & Lech, M. (2019). Two-Stage deep learning approach to the classification of fine-art paintings. *IEEE Access*, 7, 41770–41781. <https://doi.org/10.1109/ACCESS.2019.2907986>.
- Semino, E. (2020). COVID-19: A forest fire rather than a wave? *Metode Science Studies Journal*, 11(5). <https://doi.org/10.7203/metode.11.19336>.
- Sethi, I. K., Coman, I. L., & Stan, D. (2001). Mining association rules between low-level image features and high-level concepts. Proc. SPIE 4384. *Data Mining and Knowledge Discovery: Theory, Tools, and Technology III* (27 March 2001).
- Shamir, L., Macura, T., Orlov, N., Eckley, D. M., & Goldberg, G. (2010). Impressionism, expressionism, surrealism: Automated recognition of painters and schools of art. *ACM Transactions on Applied Perception*, 21(2). <https://doi.org/10.1145/1670671.1670672>.
- Solso, R. L. (2000). The cognitive neuroscience of art: A preliminary fMRI observation. *Journal of Consciousness Studies*, 7(8-9), 75–86. <https://www.ingentaconnect.com/content/imp/jcs/2000/00000007/f0020008/1043>.
- Sowmyayani, S., & Rani, P. A. J. (2022). Salient object based visual sentiment analysis by combining deep features and handcrafted features. *Multimedia Tools and Applications*, 81(6), 7941–7955. <https://doi.org/10.1007/s11042-022-11982-5>.
- Stamatopoulou, D., & Cupchik, G. C. (2017). The feeling of the form: Style as dynamic ‘textured’ expression. *Art & Perception*, 5(3), 262–298. <https://doi.org/10.1163/22134913-00002066>.
- Sumit, S. S., Watada, J., Roy, A., & Rambli, D. (2020). In object detection deep learning methods, YOLO shows supremum to mask R-CNN. *Journal of Physics: Conference Series*, 1529, 042086. <https://doi.org/10.1088/1742-6596/1529/4/042086>.
- Tadi Bani, N., & Fekri-Ershad, S. (2019). Content-based image retrieval based on combination of texture and colour information extracted in spatial and frequency domains. *The Electronic Library*, 37(4), 650–666. <https://doi.org/10.1108/EL-03-2019-0067>.
- Tan, W. R., Chan, C. S., Aguirre, H. E., & Tanaka, K. (2016). Ceci n’est pas une pipe: A deep convolutional network for fine-art paintings classification. *2016 IEEE International Conference on Image Processing (ICIP)*, 3703–3707. <https://doi.org/10.1109/ICIP.2016.7533051>.
- Taylor, R. P., Micolich, A. P., & Jonas, D. (1999). Fractal analysis of Pollock’s drip paintings. *Nature*, 399(6735), 422–422. <https://doi.org/10.1038/20833>.
- Wang, M., & Deng, W. (2018). Deep Visual Domain Adaptation: A Survey. *ArXiv:1802.03601 [Cs]*. <http://arxiv.org/abs/1802.03601>.
- Westlake, N., Cai, H., & Hall, P. (2016). Detecting people in artwork with CNNs. In G. Hua, & H. Jégou (Eds.), *Computer vision – ECCV 2016 workshops* (Vol. 9913, pp. 825–841). Cham: Springer International Publishing. [https://doi.org/10.1007/978-3-319-46604-0\\_57](https://doi.org/10.1007/978-3-319-46604-0_57).

- Widjaja, I., Leow, W. K., & Wu, F.-C. (2003). Identifying painters from color profiles of skin patches in painting images. *Proceedings 2003 International Conference on Image Processing (Cat. No.03CH37429)*, 1, 1-845–1-848. <https://doi.org/10.1109/ICIP.2003.1247095>.
- Woods, A. (2012). Beyond the wounded storyteller: Rethinking narrativity, illness and embodied self-experience. Chapter 6. In H. Carel, & R. Cooper (Eds.), *Health, illness and disease: Philosophical essays* (pp. 113–128). Utrecht: Acumen Publishing.
- Yanulevskaya, V., Uijlings, J., Bruni, E., Sartori, A., Zamboni, E., ... Sebe, N. (2012). In the eye of the beholder: Employing statistical analysis and eye tracking for analyzing abstract paintings. *Proceedings of the 20th ACM International Conference on Multimedia*, 349–358. <https://doi.org/10.1145/2393347.2393399>.
- Zeki, S. (2003). *Inner vision: An exploration of art and the brain (reprinted (with corr.))*. NY: Oxford Univ. Press.
- Zhang, G. P. (2000). Neural networks for classification: A survey. *IEEE Transactions on Systems, Man, and Cybernetics, Part C (Applications and Reviews)*, 30(4), 451–462. <https://doi.org/10.1109/5326.897072>.
- Zhang, H., Augilius, E., Honkela, T., Laaksonen, J., Gamper, H., & Alene, H. (2011). Analyzing emotional semantics of abstract art using low-level image features. In J. Gama, E. Bradley, & J. Hollmen (Eds.), *Advances in intelligent data analysis X: Ida 2011* (Vol. 7014, pp. 413). Berlin, Heidelberg: Springer-Verlag Berlin. [https://doi.org/10.1007/978-3-642-24800-9\\_38](https://doi.org/10.1007/978-3-642-24800-9_38).
- Zhang, J., Liu, X., Chen, M., Ye, Q., & Wang, Z. (2022). Image sentiment classification via multi-level sentiment region correlation analysis. *Neurocomputing*, 469, 221–233. <https://doi.org/10.1016/j.neucom.2021.10.062>.
- Zhang, W., He, X., & Lu, W. (2020). Exploring discriminative representations for image emotion recognition with CNNs. *Ieee Transactions on Multimedia*, 22(2), 515–523. <https://doi.org/10.1109/TMM.2019.2928998>.
- Zhao, S., Gao, Y., Jiang, X., Yao, H., Chua, T.-S., & Sun, X. (2014). Exploring principles-of-art features for image emotion recognition. *Proceedings of the 22nd ACM International Conference on Multimedia*, 47–56. <https://doi.org/10.1145/2647868.2654930>.
- Zhao, S., Yao, X., Yang, J., Jia, G., Ding, G., ... Keutzer, K. (2021). Affective Image Content Analysis: Two Decades Review and New Perspectives. *ArXiv:2106.16125 [Cs]*. <http://arxiv.org/abs/2106.16125>.
- Zhao, Z.-Q., Zheng, P., Xu, S.-T., & Wu, X. (2019). Object detection with deep learning: A review. *IEEE Transactions on Neural Networks and Learning Systems*, 30(11), 3212–3232. <https://doi.org/10.1109/TNNLS.2018.2876865>.