Journal Title
XX(X):1–13

DOI: 10.1177/ToBeAssigned
www.sagepub.com/
SAGE

Stantne in imaginum classibus animi motus? A study on the ability of Iconclass to decipher the visual language of emotions in classical imagery

Alessandro Adamou[1], Davide Picca[2], and Francesca Tancini[3]

Abstract

This paper explores the intersection of iconography, iconology, and sentic computing to enhance the analysis and understanding of artworks. We focus on the Greek mythology associated with the Trojan War, which lends itself to highly diverse visual interpretations over many centuries, and exploit the availability of imagery annotated with Iconclass, a widely used taxonomy of visual art themes, figures, and concepts. Our study consists of a quantitative analysis with non-expert annotators and a qualitative study with art historians. The results show that laypersons agree on emotions derived from the iconography in universally tragic events, while expert knowledge boosts consensus in cases where emotional resonance depends on context. Both studies converge on a need to integrate an iconological layer with the one for which Iconclass is usually employed, in order to enrich the cataloging, analysis, and understanding of artworks. Our work contributes to broader discussions about innovative and inclusive practices in art classification, highlighting the need for a holistic approach that accommodates both intellectual and affective dimensions in visual art as in text.

Keywords

iconography, iconology, Iconclass, RDF, SKOS, emotion mining, sentic computing, art history, digital hermeneutics

Introduction

In the realm of art history, iconography emerges as a key methodology for unraveling the complexities of visual art. It is the discipline that delves into the rich narrative of content and subject matter that artworks embody (Godzic (1981); Vago et al. (2021)). This nuanced approach enables art historians to identify and contextualize the subjects depicted within a work, thereby facilitating attribution to specific geographical, cultural, or chronological contexts. For instance, the recognition of specific motifs can be linked to the time and place of a particular event, such as the Saints Sebastian and Roch protecting against a particular outbreak of the plague. Iconography transcends the surface-level examination of style and structure, in that it focuses more on what is depicted, rather than how and why in that way. However, by detecting and analyzing themes and motifs, iconographical studies provide a foundation for a deeper understanding of the artworks' meanings, which in turn incorporates notions of style, structure, and historical context of its authorship. This latter discipline is often referred to as *iconology*.

The relationship between iconography and iconology is subject to ongoing debate in art historical circles. While iconography is concerned with the identification and contextualization of subjects and motifs, iconology seeks to uncover the deeper, often symbolic meanings embedded within these visual representations. The distinction between these two concepts is not always clear-cut, and some scholars argue that the dichotomy between iconography and iconology is artificial (Džalto (2023)). Nevertheless, probing the underlying narratives and ideologies that shape the visual language of artworks remains a vital aspect of art historical analysis. By exploring the intersections and tensions between iconography and iconology, art historians can gain a more nuanced understanding of the complex meanings and contexts that underlie the artworks they study. A dual analysis of subject and symbolic meanings allows for a further level of interpretation. An example is the case of Trojan priestess Cassandra standing with her back to a burning Troy in an 1898 painting by Evelyn De Morgan (1855-1919) now at the De Morgan Collection, Barnsley (inv. 22): the theme, chosen by this Pre-Raphaelite artist, has been likened to the struggles of British women in making themselves heard at the turn of the 20th Century, De Morgan herself having been a feminist and a suffragette (Coleman (2014)).

Background on Iconclass

What is certainly the standard semantic tool for digital iconography, at the disposal of scholars and practitioners

[1]Bibliotheca Hertziana - Max Planck Institute for Art History, Italy
[2]University of Lausanne, Switzerland
[3]Newcastle University, UK

Corresponding author:
Alessandro Adamou, Bibliotheca Hertziana - Max Planck Institute for Art History Via Gregoriana 28, 00187 Rome, Italy.
Email: alessandro.adamou@biblhertz.it

in the GLAM sector (Galleries, Libraries, Archives, and Museums) since over four decades, is Iconclass* (Becker et al. (1983)). This classification system is a taxonomy of visual art themes, figures, and concepts, covering a broad spectrum of themes and subjects across various cultures and historical periods, from non-representational art and abstract concepts to narrative subjects taken from literature, history, or religion. It is used to classify a large variety of images, from painting to photography. Iconclass is structured as a decimal classification system with ten divisions acting as access points to more than 28,000 subjects corresponding to just as many alphanumeric unique codes.

In this study, we focus the Greek mythology associated to the Trojan war. The Iconclass categories–or *notations*–of interest begin with `94C` to `94K`, or with `95`, where `9` is the class for 'Classical Mythology and Ancient History', and `94` and `95` are the two subclasses for 'Greek heroic legends'. While notation `94` is hierarchically organized along a phenomenological dimension, with subclasses covering specific books in Homer's poems and events therein, notation `95` is centered around heroes and heroines, with subclasses covering phases of their lives, or activities and relationships that shape their characters. For instance, `94F82` is the category that represents the death of Patroclus at the hands of Hector, whereas `95B(CASSANDRA)6` relates to the suffering and misfortune (in general) of Cassandra, the cursed seer of Troy. Different criteria govern other parts of the taxonomy: as another key topic of our study, emotions are contemplated as abstract concepts in notation `56` "Emotion", and as indicators of human types in notation `32A8` "passions, emotions, affections".

Problem statement and outline of the study

Being Iconclass an almost exquisitely iconographical tool, its widespread adoption begs the question about its ability to support, at least indirectly, queries of iconological nature as well. Is it able to inform studies on the possible interpretations of what is being depicted? Can a standard practice be devised to annotate a digital record of Evelyn De Morgan's "Cassandra" with notations pertaining to its iconography–the torment of Cassandra; the fall of Troy–as well as to its iconology–political enfranchisement of women and its subterms: feminism and suffragettes–and the latter according to what credited art-historical interpretation?

To lay the groundwork for such a study, with this paper we intend to explore the extent to which Iconclass can encode nuanced features related to emotions expressed through elements of the iconography itself, including but not limited to facial expressions, pose and gesture, and even objects such as drapes and their movement. The choice of Trojan War mythology is motivated by the fact that it is represented with a great variety of expressions in a vast range of visual arts, reaching up to Neoclassicism in the late 18th and early 19th centuries. This theme is also used across contexts more widely open to non-conventional interpretation than other widespread artistic themes, such as Christian sacred imagery.

Our work borrows from the theories of *sentic computing* formulated by Cambria and Hussain (2015), which bridge computational linguistics, semiotics, and affective computing. The set of emotional coordinates that we employ comes from the sentic computing model SenticNet, which we use in two separate studies. The first (quantitative) study, first described in Adamou and Picca (2024), has a team of non-experts independently annotate a corpus of 300 images from the photographic collections of the Bibliotheca Hertziana, which are already tagged with Iconclass notations of our interest. These annotations are compared to a purely linguistic rating, extracted from the Iconclass notations and their associated keywords in the Iconclass dataset. Upon an initial interpretation of the results, a second (qualitative) study is conducted with scholars in art history on a hand-picked set of 24 images tagged with the Iconclass notations upon which the raters of the first study disagreed the most.

In our setting, the Iconclass tagging of artworks has been performed at face value–i.e. based on factual data and not on interpretive work or original research. The results of our studies showed that laypersons are more likely to agree over the emotions expressed by the iconography, where the affective resonance of the episode is unambiguous, such as for events that are universally tragic. For events whose emotional resonance depends on the point of view chosen by the artist, expert knowledge of the episodes and of the contexts of the artwork's creation boosts consensus. This leads us to surmise that the expert annotation of digital images with an ulterior layer that makes an iconological usage of Iconclass, possibly in combination with other vocabularies, would indeed enhance the semantic expressivity of digital catalogues of visual art. This is further corroborated by recent advancements in the design of ontology systems for representing iconology.

This paper aims to raise awareness in the digital humanities community of the need for a holistic approach to the understanding of visual art. We intend to follow up with a proposal that accommodates both the intellectual and affective dimensions, thereby enhancing the analytical capabilities of art history and iconography, and offering a complementary lens through which artworks can be classified and understood (Pavlopoulos et al. (2022); Picca and Richard (2023); Picca (2024); Picca and Pavlopoulos (2024)). Using a theme that has been repeatedly represented through history, this work also offers insights into how similar emotions are depicted across cultures and epochs. It also seeks to explore the feasibility and benefits of an integration with Iconclass, hypothesizing that a combined approach could significantly enrich the cataloging, analysis, and understanding of artworks. Through a balanced examination of Iconclass alongside an emerging potential emotional classification system, this study aims to contribute to broader discussions about innovative and inclusive practices in the classification of art.

After an overview on related work in digital iconography, iconology and sentic computing, we delve into the materials and methods employed. We then describe the quantitative analysis and discuss its results leading to the qualitative study. The concluding section offers a discussion touching upon potential strategies for acting upon these outcomes.

*Iconclass, `http://iconclass.org/`

Related Work

Recent advancements in the detection and analysis of emotions within iconographic artworks illustrate a promising intersection of art history and cognitive science and technology. These interdisciplinary efforts aim to enhance our understanding of how emotions are conveyed and perceived in artistic representations, leading to more nuanced interpretations and accessible art historical knowledge.

Emotion mining of visual art and digital hermeneutics

One groundbreaking approach involves the adaptation of machine learning techniques, specifically convolutional neural networks (CNNs), to interpret emotional content in artworks. Research by González-Martín et al. (2024) indicates that CNNs, traditionally trained on non-artistic images, can be effectively adapted to the artistic domain by addressing the cross-depiction problem. This involves algorithms such as QuickShift, which enhance the network's ability to generalize across different artistic styles, thereby improving accuracy in emotion detection. This methodology not only bridges the gap between digital image processing and art analysis, but also opens up possibilities for more robust cataloging and understanding of emotional expressions across diverse art forms.

Further enriching the domain are multimodal frameworks that integrate various types of data (e.g., visual, textual, and auditory) to analyze emotional content. Originally developed for conversational dynamics in videos, such frameworks can be adapted for art, considering how narratives within artworks contribute to emotional impact. This suggests a layered approach to emotion detection, where the interplay of different modalities can provide a deeper understanding of how artworks engage viewers emotionally. Hazarika et al. (2018) present such a multimodal emotion detection framework, which could potentially be applied to the analysis of emotional content in iconic artworks .

Moreover, studies examining cognitive responses to iconic versus realistic depictions reveal that iconic representations, such as those in cartoons or stylized graphics, communicate emotional information more effectively. Kendall et al. (2016) highlight the differences in neural processing that occur when viewers encounter iconic versus realistic images, suggesting that iconicity enhances emotional communication through visual art, potentially due to the simplified and exaggerated features that better capture and convey emotional states.

The influence of demographic factors on emotional responses is also critical. Research by Ko and Yu (2016) investigates gender differences in responses to iconic designs, using facial expression recognition software to analyze how different genders perceive and react emotionally to the same visual stimuli. Such studies highlight the need for considering a variety of viewer backgrounds when analyzing emotional responses to art, providing insights into how personal experiences and cultural contexts might influence emotional interpretation.

Lastly, we observe that the potential of generative deep learning models shows promise. Particularly, the BLIP model has been shown to offer remarkable performance for the automated generation of textual captions for artistic imagery (Zhu et al. (2024)): an aspect where it will be of interest to investigate the degree to which prompt engineering and model fine-tuning can be directed towards performing interpretive acts as part of the generative process.

Models for digital iconology

Complementing these computational and psychological approaches, ontological frameworks have been developed to model the complex layers of meaning in artworks. The ICON ontology, introduced by Sartini et al. (2023), provides a structured model for representing artistic interpretations of artworks' subject matter (i.e., iconographies) and meanings (i.e., symbols, iconological aspects). By conceptualizing authoritative knowledge and notions taken from Panofsky's levels of interpretation theory, ICON focuses on the granularity of interpretations, allowing descriptions from the pre-iconographical, iconographical, and iconological levels. The ontology is aligned with other ontologies from the domains of cultural descriptions (ArCo, CIDOC-CRM, VIR), semiotics (DOLCE), bibliometrics (CITO), and symbolism (Simulation Ontology), ensuring a robust and extensible schema. The ICON ontology is openly available and compliant with FAIR principles, facilitating its integration into various research and application contexts.

Building upon the foundational work of the ICON ontology (Sartini et al. 2023), several recent studies have further advanced the formal representation of iconographic and iconological knowledge in the context of cultural heritage.

Sartini (2024) introduced IICONGRAPH, a knowledge graph designed to refine and extend the iconographic and iconological statements present in existing cultural heritage knowledge graphs such as ArCo and Wikidata. By addressing persistent gaps in the representation of cultural symbolism, IICONGRAPH offers a more granular and semantically rich framework for modeling the visual and contextual elements of artworks. The ontology adheres to FAIR principles, ensuring its reusability and integration into broader semantic web applications.

Baroncini et al. (2021) proposed a data model that extends existing ontologies to better capture the nuances of iconological analysis. Their model addresses the need for representing the cultural and historical contexts of artworks, facilitating more comprehensive digital art history research. By analyzing eleven case studies, the authors identified gaps in current ontological representations and suggested new terms to enhance the expressiveness of semantic models in capturing iconological interpretations.

Gartner (2020) presented an ontology aimed at conceptualizing iconographical recognition of subjects in artworks, aligning with Panofsky's second level of interpretation. This ontology seeks to enhance the semantic annotation of visual art by providing structured representations of iconographic elements. The study emphasizes the importance of machine-readable ontologies in improving the identification and analysis of iconographic subjects within digital art history.

Additionally, the universality of emotional responses to sensory inputs extends beyond visual arts to other domains such as poetry, where Auracher et al. (2010) explore how sound iconicity in poetry can evoke specific emotions.

This research suggests parallels in visual art, where certain visual forms or styles may universally trigger emotional responses, enhancing our understanding of the cross-sensory dimensions of emotional perception in art.

These diverse approaches not only underscore the complexity of emotion detection in art but also highlight the potential for developing more sophisticated tools and methodologies, which can cater to the multifaceted nature of art perception and appreciation.

Materials and methods

The selection of classical iconography as a case study is motivated by several factors that render it an ideal subject for investigation. Firstly, as mentioned earlier, the classical iconography of Greek myths encompasses a vast array of diverse representations spanning centuries and cultures. This multiplicity of interpretations facilitates iconological analysis, but also allows scholars to gauge the influence of style and context on the interpretation, since these factors vary greatly across the centuries.

A second consideration is that the structure of Iconclass on classical mythology is rigid yet twofold, having one event-centric branch and another character-centric branch. While the latter occasionally touches upon the emotional payload of a character's vicissitudes, the former tends to focus on the literal depiction of events, with limited attention to emotional nuances. Consequently, these Iconclass categories often neglect to explicitly capture emotional cues, instead relying on contextual information to convey affective meaning. For instance, certain categories, such as `94F83` (Achilles' grief over Patroclus), explicitly encode emotional content, whereas others, like `94C11` (Paris and the golden apple), require the emotional resonance to be inferred from other factors, which are partly discussed here.

The Bibliotheca Hertziana Photographic Collection

The Photographic Collection of the Bibliotheca Hertziana, or Fotothek,[†] occupies a key position within the realm of art historical research. Established in Rome, this archive was deliberately designed to facilitate scholarly inquiry into the study of art history, with a particular emphasis on Italian and Mediterranean art. As a comprehensive repository of photographic reproductions, it encompasses a vast array of artworks, including ancient sculptures, frescoes, Renaissance paintings, and architectural monuments.

Conceived as an integral component of the Hertziana's research infrastructure, the Photographic Collection reflects the vision of its founder, Henriette Hertz, to create a visual resource that would provide scholars with unfettered access to Italian art, irrespective of their geographical location. Over the years, the collection has expanded to comprise over 1,500,000 images documenting monuments and artworks, serving as indispensable tools for researchers. The progressive adoption of digital technologies by the hosting institute has led to a continued scanning and digitization endeavor of the entire material, as well as making a substantial portion of it available via the IIIF protocols.[‡] Each digitized photograph is accompanied by metadata adhering to the Marburg MIDAS schema (Bove et al. (2001)), which includes information about the artwork's creator, date, and iconographic themes.

As MIDAS predates linked open data, many of the metadata associated to an image consist of plaintext values mostly in German, however, the schema mandates that the artwork's subject matter be indicated as Iconclass classification codes. Though only the labels of these codes in German are imported into the Hertziana's systems, the codes themselves are internationalized. At the Photographic Collection, the assignment of these codes is carried out either by resident staff, or by scholars in art history as part of dedicated digitization campaigns, still under the supervision of resident staff. The tagging policy is that there should be no iconological commitment. Therefore, scholars are asked to tag an image with only the Iconclass codes that can be derived from the work's title or metadata and, to a limited extent, from its iconography, but not on the basis of a scholar's interpretation or specific research. This partly explains why emotion-specific notations, like those on the branches `32A8` and `56`, are seldom used in the Hertziana Photographic Collection, and almost never in our dataset: only if a painting were titled e.g. "The rage of Achilles" could it be tagged with `56E2` "anger, rage", but not in cases where this emotion can be visually derived but is not explicit. Therefore, we cannot assume to rely on the explicit usage of these codes to derive the emotional content of a depiction.

Corpus construction

For our qualitative study, the one that required the largest workforce, we could avail ourselves of a cohort of twenty annotators. The initial iconographic corpus was therefore created with the intent of striking a balance between workload assigned to annotators and dataset size and variety.

In the Hertziana Photographic Collection, photographs are grouped by the work of art, or "Objekt", that they depict; therefore, photographs of the same sculpture from different angles, by different photographers, or color and black/white photographs of the same painting are listed in one single IIIF manifest. An Objekt can be broken into several parts, e.g. paintings belonging to the same pictorial cycle, sides to a coin, or details of one historical building. The IIIF manifests also reflect this parthood relation, however, Iconclass codes are assigned individually to the whole or to each part and no inheritance is implied nor displayed on the online catalog.

To build our corpora, we first retrieved the identifiers of all the Objekte that were tagged with at least one Iconclass notation belonging in the `94C` to `94K` range, or with one in the `95` group, which is about individual characters. Although this was originally performed through a local XQuery, it is now also possible through a REST search API.

From the images tagged with codes in the `95` group, we excluded characters not associated to Trojan War mythology, including but not limited to[§] the "Iliad" and "Odyssey". Since the quantitative study required the image corpus to be submitted to third parties for annotation, it had to comprise

[†] Fotothek, `https://foto.biblhertz.it/`.

[‡] Internet Image Interoperability Framework, `https://iiif.io/`.

[§] A typical outlier prominently figured in art history is the Laocoön, who is found in Virgil and in the *Posthomerica*, but not in Homer.

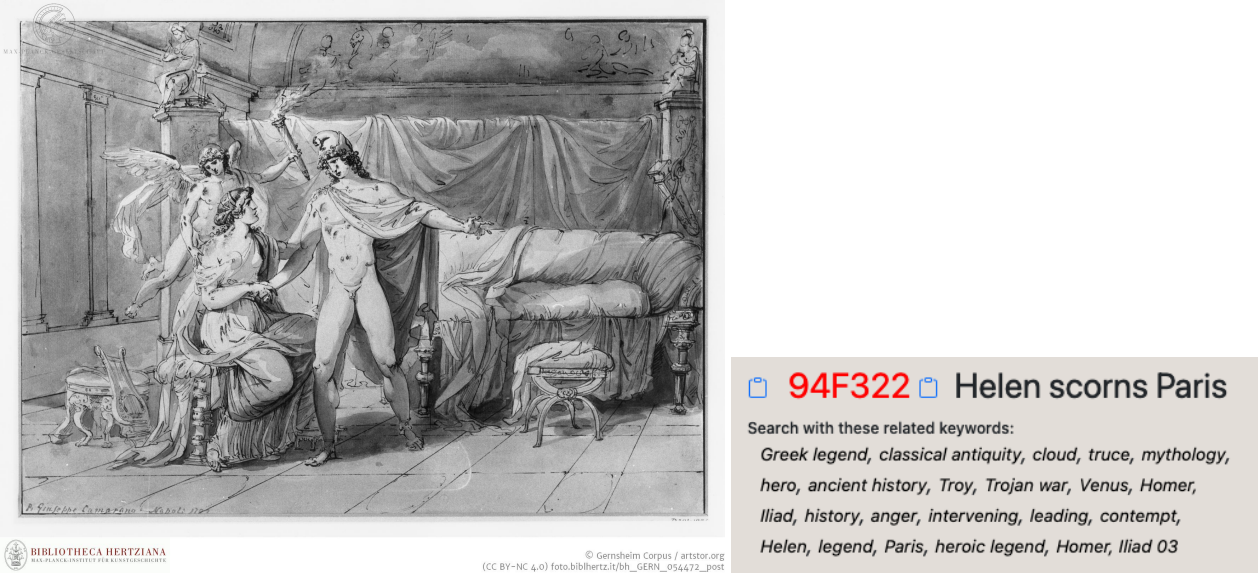

Figure 1. Left: Giuseppe Cammarano, "Paris and Helen, with Hymen", drawing, 1782 or 1792. Corpus Gernsheim (https://foto.biblhertz.it/obj/08123696), CC BY-NC 4.0. Right: detail of Iconclass 94F322, with the associated set of keywords, retrieved from https://iconclass.org/94F322.

only images that were licensed free of charge for public use ('freigegeben') by the rights owner. A harvester was written for fetching the actual images and ran from outside the network of the Max Planck Society, so that the retrieval of non-licensed images would fail. We filtered the resulting dataset to limit the number of different photos of the same work. Introducing a few negative examples, i.e. images with other Iconclasses in the 94-95 range, yielded a corpus of 300 images for the quantitative study.

A second corpus, themed around the most controversial results from the first study, was later built to be used for the qualitative study. Since this study only involved the authors and resident scholars, restricting to openly available images was no longer necessary and images with rights retained could also be used. This focused corpus consisted of 24 images, half of which in common with the first corpus.

Example: Cammarano's "Helen scorns Paris". Iconclass associates to each category code a bag of keywords available in multiple languages, which represent topics relevant to that category. For instance, Figure 1 shows a photograph of an 18th-century Italian drawing by the Neoclassical painter Giuseppe Cammarano (1766-1850), the original in pen and ink being at the Victoria and Albert Museum, London (inv. D201-1889). It shows the Trojan prince Paris, the beautiful Helen, and the winged figure of Hymen, the god of marriage ceremonies. The scene depicted is most possibly to be linked to the moment when, in the third book of the Iliad, Paris returns to his chambers after cowardly fleeing from a duel against Menelaus. Helen chides her husband as a coward, but Paris brings her to the bridal bed: "come, let us take our joy, couched together in love" (Il. 3, v. 441). In the Hertziana Photographic Collection, the image is annotated with Iconclass 94F322 ("Helen scorns Paris"). That category is decorated with keywords like 'Trojan war', 'Helen', 'truce', but also 'anger' and 'contempt'. Concerning the last two, we observe that, while 'contempt' is directly assigned to this category, 'anger' appears because it is inherited from notation 94, which is general for the tenth year of the Trojan war. We expect this keyword inheritance mechanism to deeply affect the outcome of our study.

Quantitative study: Emotion tagging in classical iconography

The primary goal of this experiment is to capture a range of emotional interpretations that viewers associate with specific scenes from classical mythology. The study was designed to explore how different viewers perceive and respond to classical themes and to the artist's nuanced interpretation of them. It also aims to assess the potential for integrating emotional responses into traditional iconographic cataloging systems, in the event that the latter are found to be lacking on that front. Lastly, it is meant to stress how much the correct identification of a specific set of emotions can help in defining the exact textual source linked to an image and how much the artist–or another influential actor, such as the patron commissioning the artwork–decides to part from it and why. In the longer run, we intend to devise a way to add a layer of emotional data to the existing Iconclass codes used in the archive, potentially enriching the academic and educational value of the Bibliotheca Hertziana holdings.

Having built the iconographic corpus, we proceed to generate two comparable "ratings" of it. On the one hand, the categories and keywords of Iconclass itself represent the source of the first rating: we intend to obtain a dataset that represents how Iconclass assesses the emotions associated to a scene depicted in art, based only on what event is known as being portrayed, and on what emotions are associated by it by art historians and classicists, who presumably contributed the terms of the keyword cloud associated to each Iconclass. On the other hand, we gather the independent responses of non-experts, who are informed as little as possible to minimize bias, and who represent likely consumers of visual art: this dataset is based on what emotions they perceive as elicited from an observation of the artwork itself. These two

ratings are then encoded as matrices whereupon it is possible to conduct various types of analytics, most importantly, detecting patterns of agreement between them.

Extracting affective terminology from Iconclass

Iconclass is also available in the Resource Description Framework (RDF) format. In this form, categories are related to one another using SKOS, e.g. `skos:broader` or `skos:related`, while the associated keywords that help convey aspects of the category semantics are represented in the Dublin Core schema as `dc:subject` predicates. The objects of these predicates are plaintext literals in multiple languages, not aligned with any vocabulary, dictionary, or authority file. An example for the episode of Polyxena's sacrifice is given in Figure 2. Note that, by construction, an Iconclass cannot have multiple parents, therefore it will not be the case that this class, or its related one of Iphigenia's sacrifice, will also share a common ancestor like e.g. `12B2122` "human sacrifice: non-Christian religions".

As demonstrated on the righthand side of Figure 1, it is also implied from the human-readable description of the Iconclass code, that it also inherits the keywords of the one(s) that is subsumes. This explains why 'anger' appears in the figure, even though it is not an intrinsic term of that specific Iconclass code. Therefore, emotion-laden terms appearing high up in the notation hierarchy are likely to sway the sentiment associated to specific episodes.

As a controlled vocabulary of emotions, offering the rating dimensions of this experiment, we adopt the SenticNet vocabulary. SenticNet was built from 100,000 concepts automatically extracted using a blend of symbolic and sub-symbolic AI techniques. Each concept includes a multiword expression, weights for four affective dimensions (temper, introspection, sensitivity, attitude), primary and secondary mood labels, a polarity score, and semantically related concepts. Particularly, it features a vocabulary of 24 terms, such as "anger", "delight" or "responsiveness", that denote emotions (Susanto et al. (2020)). For instance, the aforementioned word "truce" is associated to "contentment" as the SenticNet primary emotion and to "serenity" as secondary emotion.

The strategy for extracting the Iconclass affective rating for our image corpus was to build a knowledge graph. First, we extract subject-related metadata from the IIIF manifests of the images in the Hertziana Photographic Collection and represent them as RDF. Then, we crawl iconclass.org for the RDF data of the matching Iconclass codes, traversing the SKOS hierarchy and storing the results.

```
PREFIX sentic: <http://example.org/SenticNet#'>
SELECT DISTINCT ?subject WHERE {
  <https://iconclass.org/94H243>
    skos:broader*/sentic:hasSenticNet/dc:subject
      ?subject .
}
```

Listing 1: SPARQL query to retrieve the keywords associated to an Iconclass notation in the knowledge graph.

SenticNet extraction for each image was performed through querying this knowledge graph. Through a simple SPARQL query, we traverse the hierarchies of all the categories that an image is tagged with at the Hertziana, collecting all its `dc:subject` keywords in the process (see Listing 1 for an example). We then look up each keyword on the SenticNet dataset (available as a Python dictionary) and, for every match, take both its primary and secondary emotion. Therefore, for every image i we end up with a vector $F_i = \{n_{ij}\}$, where n_{ij} is an integer that counts the occurrences of emotion j on the Iconclass categories with which i is annotated in the Photographic Collection. The matrix $\mathcal{F} = \{F_i\}$ of size $n \times m$, where n is the number of images and m are the 24 SenticNet emotions, is thus the dataset of the emotional ratings of the image corpus according to Iconclass itself.

User study: image annotation with SenticNet

The next step is to obtain a counter-rating of our iconographic corpus from its viewers. Our starting hypothesis is that even who is not an expert in the classical subject matter is able to detect nuances in the emotions evoked from the depicted scene, and that the emotions detected by non-experts approximate those that experts would surmise from their knowledge of the episode. Whether these nuances are the result of the artist's own interpretation of the scene, or of shortcomings in the Iconclass scheme, is an iconological research question in and of itself, which the results should help us explore. The qualitative study carried upon these outcomes partly sets the postulates for such an investigation into a restricted set of works.

The counter-rating, to be compared against the one emerging from Iconclass, was obtained through a user study.[¶] Each participant was asked to annotate a subset of the image corpus with emotional tags, so that each of the 300 images would be annotated by exactly four participants. The annotation process required them to select one or more regions of each image, and to tag each region with exactly one of the 24 emotions from SenticNet.[‖] One region could only be tagged with multiple emotions by replicating the region itself. This task was performed independently, and the participants were encouraged to consider both the emotional tone of the scene and the emotional responses they believed the artwork was intended to evoke in an audience. They were not informed on what artwork they were looking at, what scene it depicted, or what Iconclass terms the Hertziana Photographic Collection had it annotated with. They were also not told how the emotion labels relate to one another in the SenticNet hourglass model, and it was in fact left up to them to assess the affective intensity of each label. Finally, they were given complete freedom to choose the shape and size of the regions–whether it highlighted a face, an entire body, another body part or, possibly, even a garment or other inanimate object.

These measures were all taken in the interest of minimizing annotator bias, as well as allowing them to express the emotions conveyed in one depicted scene at a fine granularity, should they choose to do so. This placed

[¶] The user base consisted of a cohort of 20 undergraduate students enrolled in the Computer Science for the Humanities course at the University of Lausanne.

[‖] For the sake of future studies, an additional option for "emotionlessness" was given to the users in case they still wanted to highlight a region of interest without assigning it a SenticNet term.

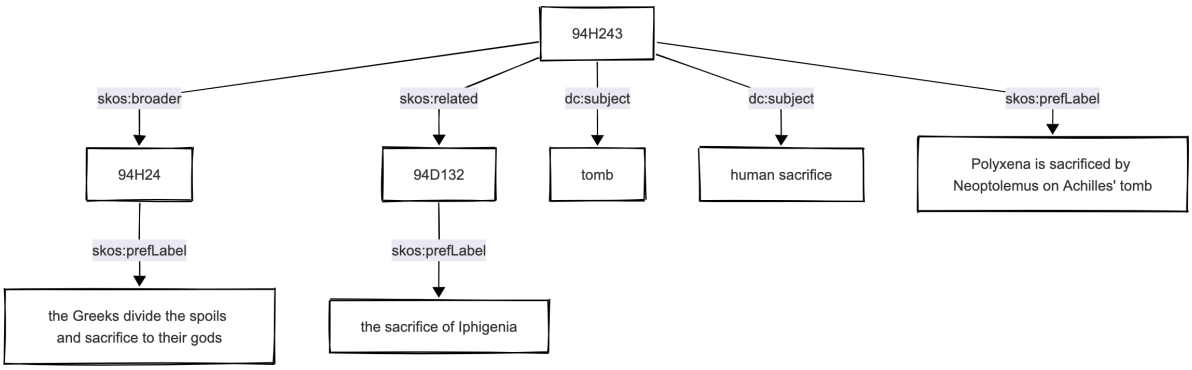

Figure 2. RDF representation of category `94H243` (from https://iconclass.org/94H243.rdf).

the users on a more extreme position than the one of typical visitors of a gallery or museum, who are guided, at a bare minimum, by the caption next to an artwork. However, the scenes in classical mythology are such that the artwork titles are often repeated descriptions of the scenes themselves, which likely influences the viewer's perception of emotions. While we obviously could not rule out that the study participants might have had a foundational understanding of the subjects depicted, we needed to counterbalance this possibility with minimal information in order to obtain ratings as unprejudiced as only possible.

The platform employed for recording annotations in both our studies is *LabelStudio*[**], a data labelling tool that is available both on the cloud and as an open source local package. It was chosen, in its cloud-based variant, due to its flexibility in allowing multi-user rating, its automatic calculation of simple rater metrics, and its ability to also annotate text, which will prove useful in future developments of this work.

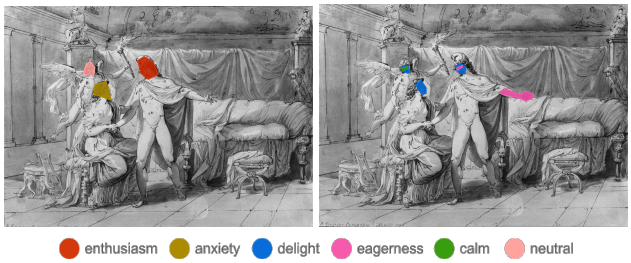

Figure 3. Two different annotations of the drawing "Paris and Helen, with Hymen".

Figure 3 shows two different ways in which the photograph of the drawing from Figure 1 was annotated. The participants interpreted Helen's facial expression very differently, noting 'anxiety' and 'delight' respectively–none of which matches the repulsion expected from the Iconclass category name, "Helen scorns Paris", nor from its subject keyword "contempt". They also used different but agreeable tags for the expressions of Paris and Hymen, yet one of them reinforced Paris' 'eagerness' by also annotating his tense arm with it. It is perhaps also worth noting that the textual episode illustrated by the image in question is particularly dense emotion-wise: Helen's disgust for her cowardly husband, with a sting of regret for having abandoned the valiant Menelaus, clashes against Paris' acceptance of his own defeat, immediately overtaken by sexual desire (hence the presence of Hymen as the god presiding marriage rites), Paris himself being a multi-faceted character acting with bravery or cowardice depending on the moment.

In the Hertziana Photographic Collection, no other items seem to be related to the same textual episode as the photograph of Cammarano's drawing, so that it is not possible to compare and contrast different annotations for the same event. It could then be of some use to bring into the analysis yet another scene from the Trojan War where both Paris and Helen are included, as, for example, the abduction of Helen, which is assigned Iconclass `94C133`. In the Hertziana Photographic Collection, this moment is documented by a few images (Figure 4), not always annotated with that Iconclass: the photograph of an engraving by German artist Johann Gerhard Huck (1759-1811) after a drawing by Raphael (1483-1520) now in Chatsworth House (inv. 903A); and the photograph of an engraving traced by French artist Louis Desplaces (1682-1739) after the painting by Guido Reni (1575-1642) now at the Musée du Louvre (inv. 539). Both prints are physically at the Museum Kunstpalast, Düsseldorf. The scene of the abduction is not included in the Iliad save a couple of passing references (Il. 3, v. 174, Il. 6, v. 348). It has long been questioned whether Helen willingly followed Paris, cajoled by Aphrodite, Eros, and possibly Peitho, goddess of persuasion, or if she was taken against her will. While Guido Reni, with his gallant procession of knights and courtesans, seems to opt for the first choice, with the added figures of Eros with his bow and arrow and the dark-skinned page boy with a little monkey on a leash, the one possibly after Raphael certainly looks much more similar to a violent abduction. That is one more reason for us to be interested in emotion labelling.

The result of the annotation process is the matrix $\mathcal{A}$ which, like $\mathcal{F}$, is of size $n \times m$, so that it could be used for the subsequent analysis of the emotional responses. These were further compared to the traditional thematic classifications provided by Iconclass codes.

Intermediate results

The great freedom granted to the human annotators in the selection of regions, aimed at facilitating the variety of perceptions, affects the number of times that an annotated

**LabelStudio, https://labelstud.io/.

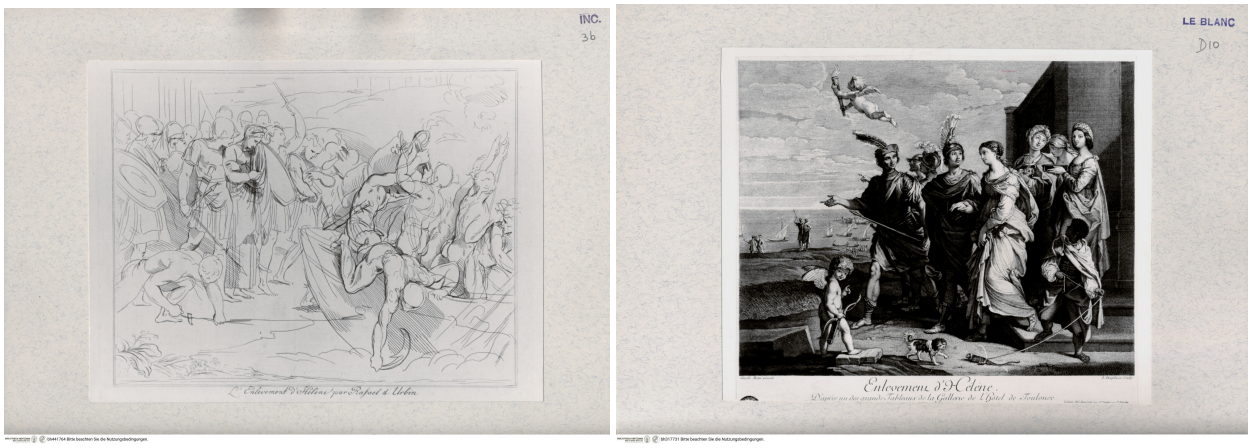

Figure 4. Two depictions of the abduction of Helen in the Hertziana Photographic Collection. Left: Johann Gerhard Huck, "Die Entführung der Helena" (print of a drawing by Raphael), 1781, Museum Kunstpalast, Düsseldorf, Public Domain Mark 1.0 (https://foto.biblhertz.it/obj/08079584?part=136). Right: Louis Desplaces, "L'enlevement d'Helène" (print of a painting by Guido Reni), Museum Kunstpalast, Düsseldorf (https://foto.biblhertz.it/obj/08039532), Public Domain Mark 1.0. Neither image is explicitly annotated with the Iconclass code of the episode (94C133).

value repeats itself. We use the frequency of these repetitions to reinforce the overall perception of the artwork's emotional resonance, therefore great agreement between the four annotators per image is not to be expected. Since our objective is to gauge how much artwork viewers "agree" with Iconclass on the emotional dimension, we consider matrix $\mathcal{F}$ (the labelling extracted from Iconclass) as representative of one rater, and matrix $\mathcal{A}$ (the labelling performed by all the participants as one) as another rater. We therefore proceed to calculate various forms of agreement between these two.

Having scrutinized several measures of rater reliability– i.e. agreement–, we adopt Krippendorff's Alpha as the reliability measure of our choice for this study. This is a very flexible mechanism, as it is able to accommodate categorical rating scales, as is the case of SenticNet's emotions, and missing data, since some users felt at times unable to choose a label. Unlike e.g. Cohen's and Fleiss' Kappas, however, the Alpha has the desirable feature of supporting cases where the total ratings do not amount for the same number for each item (Krippendorff (2013)), which is again our case due to allowing users to annotate multiple regions with the same emotion. This phenomenon was also likely in matrix $\mathcal{F}$, since multiple Iconclass keywords could have the same primary or secondary emotion in SenticNet, therefore one label could appear an indefinite number of times per image.

Krippendorff's Alpha, or α-agreement, is normalized between -1 and 1. A value greater than zero denotes inter-rater agreement; one close to zero means that the rating is neither more nor less reliable than a random choice of ratings; one lower than zero denotes disagreement.

Our first attempt was to consider the SenticNet emotions as the items being rated, thus calculating the Alpha between the *inverted* $\mathcal{F}$ and $\mathcal{A}$, to detect whether certain categories of emotion tend to gather more agreement than others. Two calculations were made: one that considers a value of zero (i.e. that emotion was never assigned to a certain item) as an effectively relevant zero, and one that considers it a missing rating. This was done because considering zero-values causes the agreements to approach randomness, but can help make subtleties emerge, which would not stand out by considering them as missing values. Some emotions were either never chosen by any human rater, or never extracted from the IconClass keywords, which is shown as the lack of a value in the last column. The results are shown in Table 1.

Table 1. α-agreement on emotions, rated considering zeroes both as relevant values and as missing values.

Emotion	α (zeroes as 0)	α (zeroes as missing)
acceptance	-0.02	-0.13
anger	-0.12	**1**
annoyance	-0.13	—
anxiety	-0.15	-0.17
bliss	0.02	—
calm	-0.2	0
contentment	-0.07	—
delight	-0.01	-0.22
disgust	-0.01	0.1
dislike	-0.12	-0.24
eagerness	-0.15	0
ecstasy	-0.03	-0.07
enthusiasm	-0.02	—
fear	-0.2	0
grief	0.03	-0.08
joy	-0.08	—
loathing	-0.06	—
melancholy	-0.11	—
pleasantness	-0.04	0
rage	-0.04	0
responsiveness	-0.17	—
sadness	-0.13	0
serenity	-0.1	-0.07
terror	-0.02	0.25

From the generally low absolute values in the α-agreement, evaluating by single emotion does not seem very telling. What is still striking, yet somewhat to be expected out of knowledge of the domain, is that the annotators agreed with Iconclass on emotions typically associated to the tragedy of the Greek *epos*, as indicated by the perfect agreement on *anger* and the 25% agreement on *terror*.

Sentiments that show disagreement in the 13-24 percent range appear either to be more subtle on the negativity front (e.g. *anxiety*, *dislike*), or to embody positiveness (e.g. *delight* or *acceptance*). We add that this disagreement is largely in favor of the human annotators, who used these terms more often due to perceiving nuances of positive feelings in the depicted scenes, which are not considered characteristic of the corresponding episodes or characters in the Trojan war. This would confirm the first statement of our starting hypothesis.

To confirm or disprove the above assumption, we can look at the Iconclass codes themselves. Recall that codes starting with `94` denote events, and those starting with `95` denote characters and their personal lives. Therefore, if agreement is primarily found over a universally tragic event, or a character primarily known for their dire fate, then we are closer to proving that Iconclass has a gap in representing classical iconography in their affective variety and subtleties.

We aggregate the ratings in $\mathcal{F}$ and $\mathcal{A}$ by Iconclass code and obtain two matrices of size 173×24, where 173 is the number of Iconclass categories in the range being considered, with which our corpus is annotated at the Hertziana Photographic Collection and the labels are again 24. A value in each cell represents the absolute number of occurrences of an emotion for an Iconclass code, considering all the images annotated with it. We then calculate the α-agreement between these matrices.

Table 2 shows the categories for which the highest agreement or disagreement was found. Indeed, the human annotators are shown to mostly agree with Iconclass over episodes–such as the mourning of a loved one or the retrieval of their dead body–whose tragic nature is always represented beyond a doubt. Examples include Thetis mourning Achilles, Orestes killing his mother, the dire fate of Cassandra, and the blinding of Polyphemus, this last one quite unsurprisingly, considering the raters were not made aware of the heroic standing of Ulysses in the episode, By contrast, the highest disagreement in the event classes is found where multiple engagements are at play, such as in the Judgment of Paris, or the human sacrifices of Iphigenia and Polyxena. In the latter case, the terror and anxiety of the sacrificial victims seem to be overshadowed, in the scene depictions, by the sense of satisfaction of the others attending the propitiatory event, in stark contrast with the focus on the victim that is established by Iconclass.

A similar argument applies to character classes: the lack of agreement over complex characters such as Achilles or Paris, appearing through their most general codes, owes to the intricacy of their stories, whereas characters like Orestes and Cassandra, whose artistic focus tends to be on their most grievous vicissitudes, enjoy a more faithful emotional representation in Iconclass.

These observations were formulated by the members of the project team who are neither classicists nor art historians. To give better framing to their assumptions, a dialogue was sought with scholars in art history and classical studies, which called for a second, qualitative study for this paper.

Qualitative study: Expert input

On the basis of the quantitative study, and of the non-expert conjectures formulated in the interpretation of their results, we integrated this work with an additional study that relied on expert input. The research questions that warranted this study shared the common goal of assessing whether or not there could be iconological motivations beneath the greatest disagreements that were detected by the quantitative study. If the annotators could not agree upon the emotions perceived from the artwork, was it due to the inability to always rely on standard factors, like facial expressions or gestures? Bearing in mind that agreement was calculated over Iconclass codes, rather than on individual images, are the depicted episodes of such intrinsic complexity that no two depictions can be guaranteed to carry the same emotional payload? Should it be implied that affective content cannot be appraised in and of itself, but requires profound knowledge of the depicted scene, the author's history, and stylistic underpinnings?

Setup

For this study, we singled out the Iconclass notations more frequently appearing in the Hertziana Photographic Collection, whose α-agreement was the lowest. We concentrated on episodes of the Iliad books–i.e. notations starting with `94`–that are featured in at least ten images, namely the human sacrifices of Iphigenia (`94D132`) and Polyxena (`94H243`) and the cycle of Helen and Paris (various notations in the `94C−94F` range). The latter also incorporates another disagreed-upon notation in the `95`-branch, namely `95A(PARIS)`, i.e. the life and story of the character of Paris in general.

A new, focused image corpus was built, comprising 24 images from the Hertziana Photographic Collection tagged with the relevant Iconclass notations. Ten images depicted the sacrifices of Iphigenia or Polyxena equally split, and fourteen showed various scenes from the cycle of Helen and Paris. The corpus was submitted to two scholars in art history, who were asked to perform a similar annotation procedure as in the quantitative study. The key differences were that: (a) information such as the work shown in the picture, its author and period, and the Iconclass tags and their descriptions, was disclosed to them; (b) they were also asked to argument on their own annotations, where possible, based on their knowledge of the source material and subject matter. The experts were asked to prioritise, if at all possible, the emotions that they perceived by looking at the artwork, over those that were the result of speculating over their knowledge and understanding of the episode.

Results and discussion

A quantitative evaluation was also performed on the annotations from the experts, who were either knowledgeable about, or informed of, the classical contexts. This was not done to compare the agreement of the experts with the users in the previous study since, unlike the former, none of the latter had rated the entire corpus given to them, but to single out pitfalls and cases warranting further investigation.

Because treating the SenticNet terms as distinct categories would not take into account their natural relations, we moved away from the category-based agreement computation of

Table 2. The most agreed or disagreed Iconclass codes, rated considering zeroes as relevant values.

Iconclass	Alpha	Name
94C1131	-0.32	the Judgment of Paris (without Mercury)
94D132	-0.29	the sacrifice of Iphigenia
94F31	-0.27	Paris and Menelaus duelling
94G533	0.17	Thetis mourning Achilles
94H151	-0.27	the Trojan horse
94H243	-0.27	Polyxena is sacrificed by Neoptolemus on Achilles' tomb
94H2452	0.20	Hecuba finds her dead son Polydorus on the sea-shore
94I134	0.21	Polyphemus is blinded with a pointed stake
94I6	-0.28	the end of the suitors
95A(ACHILLES)	-0.28	(story of) Achilles
95A(DIOMEDES)4	0.35	Diomedes - aggressive, unfriendly activities and relationships
95A(LAOMEDON)31	-0.29	Neptune builds the walls of Troy as Apollo tends Laomedon's flock
95A(ORESTES)312	0.18	Orestes kills his mother Clytaemnestra, and Aegisthus
95A(PARIS)	-0.36	(story of) Paris (Alexander)
95A(ULYSSES)	-0.35	(story of) Ulysses
95B(CASSANDRA)6	0.24	(story of) Cassandra - suffering, misfortune
95B(CIRCE)	-0.33	Circe
95B(LAODAMIA)21	-0.27	Laodamia and Protesilaus
95B(LAODAMIA)6	-0.27	suffering, misfortune of Laodamia

Figure 5. Two depictions of human sacrifice in the focused image corpus. Left: Domenico Corvi, "Neoptolemus sacrifices Polyxena on the tomb of Achilles", 1791-1800, Museo Civico di Viterbo, CC BY-NC-SA 4.0. Right: Pietro Berrettini, "The Sacrifice of Iphigenia" (drawing based on Domenichino's fresco in Villa Giustiniani-Odescalchi, Bassano Romano), CNE.

the qualitative study and organized the labels in a way that reflects their placement in the hourglass model by Susanto et al. (2020). This model considers each term in SenticNet as one of six possible gradations of four key dimensions, namely *attitude*, *introspection*, *sensitivity*, and *temper*. For instance, "bliss" and "rage" are at opposite extremes of *temper*, while "anger" and "rage" are neighbors. We assigned a signed normalized value of module in $(0, 1]$ to each term identified by a dimension (e.g. 1 to "bliss", -0.66 to "anger", -1 to "rage"). A value of 0 is reserved for when the user explicitly annotates a region with "emotionlessness", or no annotations along that dimension of terms were used.

Having discrete yet numeric ratings instead of categorical ones, and a guaranteed rating for each annotator, one can resort to simpler measures than α-agreement to assess their reliability: we chose the Intraclass Correlation Coefficient, or ICC (Koo and Li (2016)), particularly ICC3 (Single fixed raters), which fits our case where each subject has been measured by all–i.e. both–raters. Since the experts knew the Iconclass codes and descriptions of each image, and since said descriptions often matched the artwork titles, the agreement between the experts and Iconclass was not computed, as it would likely have been biased. The computed ICCs for each dimension and either class of scenes, as well as for the entire corpus rated as whole, are shown in Table 3.

In all cases but one, a positive ICC was computed, denoting agreeement, in most cases exceeding the 0.5 threshold. It follows that an understanding or knowledge of the mythical episode being depicted likely guides the perception of the emotional resonance in the artwork. This rules in favor of a static association of affective terms to Iconclass notations, as is currently the case, though not necessarily propagating them from higher-level notations. However, it is reasonable to presume that the gap in

Table 3. Expert inter-rater agreement, measured as intraclass correlations coefficient (ICC3) within each SenticNet class.

Iconclass	ICC3 (Single fixed raters)			
	attitude	introsp.	sensitiv.	temper
sacrifices	.542	.622	.264	-.235
Helen & Paris	.080	.627	.547	.622
combined	.301	.695	.540	.566

agreement can be filled by complementing iconographical annotations with iconological ones, whether or not the terms still come from Iconclass itself.

To investigate the few cases of disagreement or near-zero values (i.e. akin to random ratings), the comments left by the expert annotators were scrutinized. It emerged that the terms along the *temper* dimensions, which range from "bliss" to "rage", were very rarely used in the depiction of sacrifices, and the cause for disagreement is that the raters picked positive terms from different dimensions. For instance, a depiction of Iphigenia's sacrifice after a fresco by Domenichino (Figure 5, right) shows the variant where Iphigenia is saved by Artemis, right as she is about to replace the sacrificial victim with a deer. Artemis was tagged with "eagerness" (*sensitivity*) by one rater and with "calmness" (*temper*) by another. For other characters, the scholars have commented that the SenticNet vocabulary was not sufficiently rich, or not aligned with their desiderata, to convey what they felt worth expressing: one remarked a lack of terms for "curiosity" or "desperation", while another had to fall back to "terror" to rate the priest's expression, where they would have used "surprise"[††] had it been available. One also commented that it was hard to distinguish between joy, bliss, and enthusiasm, which are, unbeknownst to the scholars themselves, placed on three different dimensions in the SenticNet hourglass model.

The art historians argued that emotions cannot be rationally defined while keeping the scholar's knowledge of the episode represented separate from what is actually shown. By their own admission, by knowing what was going on in a scene, they could be anticipating emotions that were not (or not yet) shown in the picture.

There were two added complications: one was the style and intended audience: for instance, Rococo or Neoclassical paintings have more restrained emotions, or none at all in the case of negative emotions, so the annotation will be 'flawed' from the outset. The other was that, when the work is a preparatory sketch or the image is degraded for copyright reasons, facial expressions are often illegible or not clearly visible, so the marking has to fall back to other factors such as gestures, the movements of drapes, or the scholar's knowledge of the episode.

The materials and tools employed for both analyses, sans the images, are available at https://github.com/unil-ish/Hertziana_IconClass_Public.

Conclusion

These experiments highlighted the potential for incorporating emotional tagging into the classification and analysis of artworks. While the second study with scholars in art history confirmed that an understanding of the scene and characters, whether from personal knowledge or from the classification scheme, helps place the artwork on emotional coordinates, the first study showed that there are nuances that cannot be entrusted to a genuinely iconographical scheme.

The nature of Iconclass as a tool for iconography would partly explain the reluctance of curators to use it for categorizing visual art from an iconological angle, even though parts of the thesaurus, such as the underutilized branch 56, would lend themselves to interpretive acts. While the efficacy of an "enhanced Iconclass" would be debatable in that sense, it could be argued that a layered approach where Iconclass can be applied, possibly along with dedicated schemas, could be taken. Recent formal ontological approaches like Baroncini et al. (2021) achieve a remarkable degree of expressivity and are open to further investigation into their usability in the area of curatorial processes in art history. We leave this to the endeavors of the emerging digital hermeneutics discipline .

Iconclass has, in recent years, become a SKOS thesaurus and is now, for all intents and purposes, a resource of the semantic web. For this reason, and due to its widespread employment, it was the scheme of choice for this study, though it remains of interest to evaluate the iconological potential of bespoke yet reasonably well-known alternatives like the *Warburg Classification System* and Garnier's *Thesaurus Iconographique*.

Among the possible steps ahead, it would be useful to collect and analyze the results of emotion mining performed on different images linked through common iconography. Several research questions would arise, such as:

- How much and in which respects does the representation of emotions change, when the images are produced by the same artist or artistic movement, in the same historical moment? Or by different artists in different moments?
- Focusing on the same iconography, which are the specific variant/invariant emotions at stake? Is it possible to isolate an emotive specificity connected to the iconography itself, or to the characters involved? Is it possible to isolate specific emotional nuances internal to all the images with the same iconography? Would it be possible to integrate that invariant emotional dimension with the Iconclass framework?
- Which is the emotive trajectory of the variant ones? How much do they change, if checked against the different context for which they were commissioned and designed to compare, contrast and analyse the variant and invariant emotions at stake?

And, more generally:

- Are the emotions annotated on specific areas of the images/figures (i.e. heads, as the first results of this analysis mostly seem to suggest), or are they also linked to wider proxemics, to limbs or gestures, or to setting and locations?

††"Surprise" was in fact present in an earlier version of the SenticNet hourglass model, but has since been deprecated.

- Do the variant or invariant emotions appear to be linked to specific artistic styles or techniques? Do they respond to the iconography alone, or are they influenced by a particularly emotive style (Expressionism, Romanticism and the like) or by a particularly gestural praxis?

With a completely different hermeneutical approach, it would be useful to ask how much, if at all, emotions and their consciousness change if the context, textual sources, and/or artistic personalities are recognizable or known to the annotator/s (i.e. if the latter are in the art history field). And again, with a yet different approach, would emotion mining allow the perception of a wider history of emotions themselves? How are they visualized and visually conceptualized in the different historical moments and contexts, by different artists, following the specifications of different patrons, designed for different uses and contexts?

In the direction of studies in digital hermeneutics, it comes natural, in this context, to relate iconographical and iconological studies to the interpretation of text, which connects to our earlier work on the Iliad Pavlopoulos et al. (2022); Picca and Pavlopoulos (2024) and on the profiling of literary characters Egloff et al. (2019). Because Iconclass codes in Classics are organized by episode, we envisage that a *trait d'union* between linguistic and iconographical studies lies in detecting the events shared by the poems and artworks: this branch of our study is currently underway and will be the topic of subsequent publications.

Acknowledgements

The authors wish to extend their thanks to Dr. Chiara Capulli of the Lise Meitner research group at the Bibliotheca Hertziana - Max Planck Institute for Art History, for her invaluable contribution in the appraisal of the visual material for our qualitative study.

References

A. Adamou, D. Picca, Assessing the expressivity of iconclass to embody emotional features in classical iconography, in: A. Bikakis, R. Ferrario, S. Jean, B. Markhoff, A. Mosca, M. N. Asmundo (Eds.), Proceedings of the fourth edition of the International Workshop on Semantic Web and Ontology Design for Cultural Heritage, Tours, France, October 30-31, 2024, volume 3809 of *CEUR Workshop Proceedings*, CEUR-WS.org, 2024. URL: https://ceur-ws.org/Vol-3809/paper4.pdf.

J. Auracher, S. Albers, Y. Zhai, G. Gareeva, T. Stavniychuk, P Is for Happiness, N Is for Sadness: Universals in Sound Iconicity to Detect Emotions in Poetry, Discourse Processes 48 (2010) 1–25. doi:10.1080/01638531003674894.

S. Baroncini, M. Daquino, F. Tomasi, Modelling art interpretation and meaning: A data model for describing iconology and iconography, arXiv preprint arXiv:2106.12967 (2021).

J. Becker, H. V. D. Waal, L. D. Couprie, R. Fuchs, E. Tholén, Iconclass: an iconographic classification system, Art Libraries Journal 8 (1983) 32 – 49. doi:10.1017/S0307472200003436.

J. Bove, L. Heusinger, A. Kailus, Marburger Informations-, Dokumentations- und Administrations-System (MIDAS), Digitales Informationssystem für Kunst- und Sozialgeschichte (DISKUS), 2001. URL: http://archiv.ub.uni-heidelberg.de/artdok/3770/1/Bove_Heusinger_Kailus_MIDAS_Handbuch_2001.pdf.

E. Cambria, A. Hussain, Sentic computing, Cogn. Comput. 7 (2015) 183–185. doi:10.1007/S12559-015-9325-0.

E. Coleman, Teaching history in 100 objects – Cassandra, 2014. URL: https://www.demorgan.org.uk/teaching-history-in-100-objects-cassandra/.

D. Džalto, Iconography and iconology, in: Oxford Research Encyclopedia of Religion, Oxford University Press, 2023. doi:10.1093/acrefore/9780199340378.013.914.

M. Egloff, D. Picca, A. Adamou, Extraction of character profiles from the Gutenberg Archive, in: E. Garoufallou, F. Fallucchi, E. W. De Luca (Eds.), Metadata and Semantic Research - 13th International Conference, MTSR 2019, volume 1057 of *Communications in Computer and Information Science*, Springer, 2019, pp. 367–372. doi:10.1007/978-3-030-36599-8_32.

R. Gartner, Towards an ontology-based iconography, Digital Scholarship in the Humanities 35 (2020) 43–53.

W. Godzic, Iconographic-iconological method in film research, Artibus et Historiae 2 (1981) 151–157. doi:10.2307/1483108.

C. González-Martín, M. Carrasco, T. G. W. Wielandt, Detection of emotions in artworks using a convolutional neural network trained on non-artistic images: a methodology to reduce the cross-depiction problem, Empirical Studies of the Arts 42 (2024) 38–64. doi:10.1177/02762374231163481.

D. Hazarika, S. Poria, R. Mihalcea, E. Cambria, R. Zimmermann, ICON: Interactive Conversational Memory Network for Multimodal Emotion Detection, in: E. Riloff, D. Chiang, J. Hockenmaier, J. Tsujii (Eds.), Proceedings of the 2018 Conference on Empirical Methods in Natural Language Processing, Association for Computational Linguistics, Brussels, Belgium, 2018, pp. 2594–2604. doi:10.18653/v1/D18-1280.

L. N. Kendall, Q. Raffaelli, A. Kingstone, R. M. Todd, Iconic faces are not real faces: enhanced emotion detection and altered neural processing as faces become more iconic, Cognitive Research: Principles and Implications 1 (2016) 19. doi:10.1186/s41235-016-0021-8.

C.-H. Ko, C.-Y. Yu, Gender differences in emotional responses to iconic design, in: 2016 5th IIAI International Congress on Advanced Applied Informatics (IIAI-AAI), 2016, pp. 785–790. URL: https://ieeexplore.ieee.org/document/7557718. doi:10.1109/IIAI-AAI.2016.35.

T. Koo, M. Li, A guideline of selecting and reporting intraclass correlation coefficients for reliability research, Journal of Chiropractic Medicine 15 (2016). doi:10.1016/j.jcm.2016.02.012.

K. Krippendorff, Content Analysis: An Introduction to Its Methodology, SAGE Publications, Inc., Thousand Oaks, California, 2013, pp. 221–250. doi:10.4135/9781071878781.

J. Pavlopoulos, A. Xenos, D. Picca, Sentiment analysis of Homeric text: The 1st Book of Iliad, in: Proceedings of the Thirteenth Language Resources and Evaluation Conference, 2022, pp. 7071–7077. URL: https://aclanthology.org/2022.lrec-1.765/.

D. Picca, C. Richard, Unveiling emotional landscapes in Plautus and Terentius comedies: A computational approach for qualitative analysis, in: Proceedings of the Ancient Language

Processing Workshop, 2023, pp. 88–95. URL: https://aclanthology.org/2023.alp-1.10/.

D. Picca, Fluid Dynamics-Inspired Emotional Analysis in Shakespearean Tragedies: A Novel Computational Linguistics Methodology, in: Proceedings of the 2nd Workshop on Mathematical Natural Language Processing@ LREC-COLING 2024, 2024, pp. 11–18. URL: https://aclanthology.org/2024.mathnlp-1.2/.

D. Picca, J. Pavlopoulos, The Iliad: a French dataset for emotion recognition in classical literature, in: Proceedings of the fourteenth language resources and evaluation conference (LREC-COLING 2024), 2024.

B. Sartini, IICONGRAPH: Improved iconographic and iconological statements in knowledge graphs, arXiv preprint arXiv:2402.00048 (2024).

B. Sartini, S. Baroncini, M. van Erp, F. Tomasi, A. Gangemi, ICON: An ontology for comprehensive artistic interpretations, Journal on Computing and Cultural Heritage 16 (2023) 1–38.

Y. Susanto, A. G. Livingstone, B. C. Ng, E. Cambria, The hourglass model revisited, IEEE Intelligent Systems 35 (2020) 96–102. doi:10.1109/MIS.2020.2992799.

N. O. P. Vago, F. Milani, P. Fraternali, R. da Silva Torres, Comparing CAM algorithms for the identification of salient image features in iconography artwork analysis, Journal of Imaging 7 (2021). doi:10.3390/jimaging7070106.

D. Zhu, J. Chen, K. Haydarov, X. Shen, W. Zhang, M. Elhoseiny, Chatgpt asks, BLIP-2 answers: Automatic questioning towards enriched visual descriptions, Trans. Mach. Learn. Res. 2024 (2024). URL: https://openreview.net/forum?id=1LoVwFkZNo.