

WikiArt Emotions: An Annotated Dataset of Emotions Evoked by Art

Saif M. Mohammad and Svetlana Kiritchenko

National Research Council Canada

{saif.mohammad,svetlana.kiritchenko}@nrc-cnrc.gc.ca

Abstract

Art is imaginative human creation meant to be appreciated, make people think, and evoke an emotional response. Here for the first time, we create a dataset of more than 4,000 pieces of art (mostly paintings) that has annotations for emotions evoked in the observer. The pieces of art are selected from WikiArt.org’s collection for four western styles (Renaissance Art, Post-Renaissance Art, Modern Art, and Contemporary Art). The art is annotated via crowdsourcing for one or more of twenty emotion categories (including neutral). In addition to emotions, the art is also annotated for whether it includes the depiction of a face and how much the observers like the art. The dataset, which we refer to as the *WikiArt Emotions Dataset*, can help answer several compelling questions, such as: what makes art evocative, how does art convey different emotions, what attributes of a painting make it well liked, what combinations of categories and emotions evoke strong emotional response, how much does the title of an art impact its emotional response, and what is the extent to which different categories of art evoke consistent emotions in people. We found that fear, happiness, love, and sadness were the dominant emotions that also obtained consistent annotations among the different annotators. We found that the title often impacts the affectual response to art. We show that pieces of art that depict faces draw more consistent emotional responses than those that do not. We also show, for each art category and emotion combination, the average agreements on the emotions evoked and the average art ratings. The WikiArt Emotions dataset also has applications in automatic image processing, as it can be used to develop systems that detect emotions evoked by art, and systems that can transform existing art (or even generate new art) that evokes the desired affectual response.

Keywords: art, images, emotions, image retrieval, emotion analysis, crowdsourcing, Renaissance art, modern art, image generation

1. Introduction

Art is imaginative human creation meant to be appreciated, make people think, and evoke an emotional response. Paintings are a popular form of art with a long and compelling history. (The paintings in the Chauvet-Pont-d’Arc Cave in southern France are about 32,000 years old.) Studies have also shown that using art to evoke an emotional response (and being creative, in general) are desired fitness attributes that have played a role in the natural selection of humans (Davies, 2012; Dutton, 2009; Miller, 2001; Aiken, 1998). Nonetheless, many of the mechanisms behind how and when paintings evoke emotions remain elusive. Several research questions remain unanswered, such as which emotions are commonly elicited by art?, why are some paintings more evocative than others?, why we like some paintings but not others?, and what is the relationship between the emotion an art evokes and how much we like it?

Museums across the world house hundreds of thousands of pieces of art and attract millions of visitors each year.¹ Yet, they display only a fraction of the art they own due to space constraints. Thus, a number of museums now have a substantial online presence. Availability of massive amounts of art online means that it is useful to have the ability to search for art with various attributes. Paintings are usually labeled with the title, the artist, and the style of painting, but they are not categorized for the emotions they evoke. Thus, automatically detecting the emotions evoked by art is of considerable importance. It can be used for organizing paintings by the emotions they evoke, for recommending paintings that accentuate or balance a particular mood, and for searching paintings of a certain style or genre that depict user-determined content in a user-determined affectual state (e.g., a Post-Renaissance painting showing angry peasants).

Paintings can be created in one of many styles such as realism, cubism, expressionism, minimalism, etc. They can belong to different genres such as still life, landscape, abstract, allegorical, figurative, etc. WikiArt.org displays 151,151 pieces of art (mostly paintings) corresponding to ten main art styles and 168 style categories.² We created the *WikiArt Emotions Dataset*, which includes emotion annotations for more than 4,000 pieces of art available on the WikiArt.org. The art in the WikiArt Emotions Dataset is from four western styles and twenty-two style categories (as listed in Table 2). The pieces of art are annotated for one or more of twenty emotion categories (as listed in Table 3). These emotion categories were chosen from the psychology literature on the theories of basic emotions (Ekman, 1992; Plutchik, 1980; Parrot, 2001) and on the theories of emotions elicited by art (Silvia, 2005; Silvia, 2009; Millis, 2001; Noy and Noy-Sharav, 2013). We obtained separate emotion annotations for when the observer sees only the image, sees only the title of the art, and sees both title and art together. In addition to emotions, the art is annotated for whether it includes the depiction of a face and the extent to which the observers liked the art—the *average art rating*.

We found that anticipation, fear, happiness, humility, love, optimism, sadness, surprise, and trust were frequently chosen as the emotions evoked by art. Fear, happiness, love, and sadness were also the emotions for which the annotators provided the most consistent labels. Other emotions were also found to be more frequent and consistently

²WikiArt displays both copyright protected and public domain art. The copyright protected art is displayed in accordance with the fair use principle: <https://www.wikiart.org/en/about>. They display historically significant artworks and provide low resolution copies that are unsuitable for commercial use. We do not distribute any of the art. We only provide URLs to the WikiArt.org pages, along with the crowdsourced annotations for these pieces of art.

¹https://en.wikipedia.org/wiki/List_of_most_visited_art_museums

annotated within paintings of a particular style. Examination of the image only, title only, and whole art (image and title together) annotations revealed that the title of the art markedly impacted the emotion evoked by a painting.

We show that paintings with faces (and to a lesser extent, paintings depicting a body but no face) elicited markedly more consistent emotion annotations than paintings that did not depict a face or body.

About 64% of all the annotated art was marked as liked (to some degree), 18% as disliked (to some degree), and 18% as neither liked nor disliked. We found that paintings evoking positive emotions were liked more, in general. We also found that paintings that bring to mind certain positive emotions such as love were liked much more than paintings that bring to mind other positive emotions such as humility. The difference was even more pronounced when comparing negative-emotion paintings; paintings bringing to mind regret, arrogance, and sadness were liked much more than paintings bringing to mind disgust, anger, or fear. Paintings evoking no emotion were the least liked paintings.

The WikiArt Emotions Dataset is made freely available for research on emotions in art as well as for developing automatic systems that can detect emotions evoked by art.³ We will also provide an interactive visualization that allows users to search for WikiArt paintings with desired attributes such as style, genre, emotion, and average art ratings.

2. Related Work

Automatically understanding the content of images and text is beneficial for several information extraction and information retrieval needs. Advances in vision and natural language processing have greatly improved the capabilities of automatic systems for understanding real-world images and text. Most of these automatic systems rely on supervised machine learning algorithms which require large amounts of human-labeled instances. Several computer vision datasets have been developed and made available. (See Appendix.) Resources such as ImageNet (Deng et al., 2009) and Microsoft Common Objects in Context (MS COCO) (Lin et al., 2014) are of particular interest for developing algorithms at the intersection of computer vision and natural language processing. ImageNet has thousands of images each for many WordNet noun concepts, whereas MS COCO has hundreds of thousands of images, many of which have English captions (descriptions).

Automatically detecting emotions has also gained considerable attention over recent years, especially from text (Mohammad, 2012b; Mohammad, 2012a; Zhu et al., 2014; Kiritchenko et al., 2014; Yang et al., 2007; Bollen et al., 2009; Wang et al., 2016; Mohammad and Bravo-Marquez, 2017; Mohammad et al., 2018) but also from images (Fasel and Luetin, 2003; De Silva et al., 1997; Zheng et al., 2010). However, image annotations for emotions have largely been limited to small datasets of facial expressions (Lucey et al., 2010; Susskind et al., 2007). Ours is the first dataset we know of that includes emotions annotations for thousands of pieces of art.

³WikiArt Emotions Project webpage:
<http://saifmohammad.com/WebPages/wikiartemotions.html>

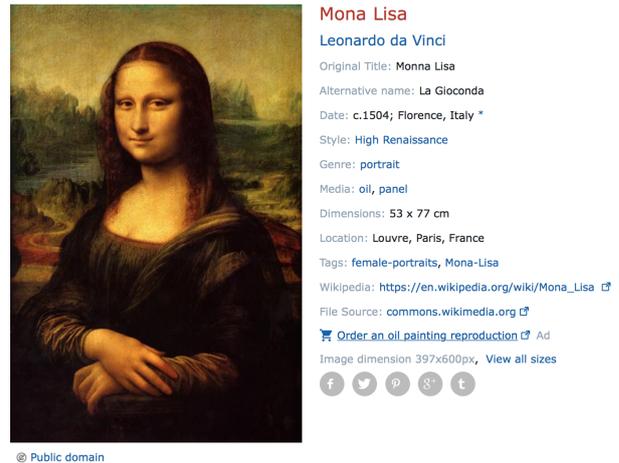


Figure 1: WikiArt.org’s page for the *Mona Lisa*. In the WikiArt Emotions Dataset, the *Mona Lisa* is labeled as evoking happiness, love, and trust; its average rating is 2.1 (in the range of -3 to 3).

3. WikiArt Emotions Dataset

We now describe how we created the WikiArt Emotions Dataset.

3.1. Compiling the Art

As of January 2018, WikiArt.org had 151,151 pieces of art (mostly paintings) corresponding to ten main art styles and 168 style categories. See Table 1 for details. The art is also independently classified into 54 genres. Portrait, landscape, genre painting, abstract and religious painting are the genres with the most items. The art can be in one of 183 different media. Oil, canvas, paper, watercolor, and panel are the most common media. For each piece of art, the website provides the title, the image, the name of the artist, the year in which the piece of art was created, the style, the genre, and the medium. Figure 1 shows the page WikiArt.org provides for the *Mona Lisa*. We collected the URLs and the meta-information for all of these pieces of art and store them in a simple, easy to process file format. The data is made freely available via our WikiArt Emotions project webpage for non-commercial research and art- or education-related purposes.⁴

For our human annotation work, we chose about 200 paintings each from twenty-two categories (4,105 paintings in total). The categories chosen were the most populous ones (categories with more than 1000 paintings) from four styles: Modern Art, Post-Renaissance Art, Renaissance Art, and Contemporary Art.⁵ For each chosen category, we selected up to 200 paintings displayed on WikiArt.org’s ‘Featured’ tab for that category. (WikiArt selects certain paintings from each category to feature more prominently on its website. These are particularly significant pieces of art.) Table 2 summarizes these details.

⁴<http://saifmohammad.com/WebPages/wikiartemotions.html>

⁵Art Nouveau (Modern), Symbolism, Naive Art (Primitivism), Conceptual Art, Mannerism (Late Renaissance), and Academicism also each had more than 1000 paintings, but we chose not to annotate paintings of these styles for now.

Style	#Categories	#Items
Modern Art	90	90576
Post-Renaissance Art	13	39497
Contemporary Art	33	8906
Renaissance Art	6	7511
Japanese Art	9	2909
Chinese Art	2	746
Medieval Art	7	636
Islamic Art	6	306
Native Art	1	50
Korean Art	1	14
Total	168	151,151

Table 1: The number of items in each art style on WikiArt.org. The styles are listed in decreasing order of the number of items.

Style	Style Category	# Items	
		Total	Annotated
<i>Contemporary Art</i>			
	Minimalism	2001	200
<i>Modern Art</i>			
	Impressionism	14862	200
	Expressionism	10629	200
	Post-Impressionism	7405	200
	Surrealism	6813	198
	Abstract Expressionism	4367	200
	Cubism	2963	200
	Pop Art	2004	200
	Abstract Art	1812	200
	Art Informel	1807	200
	Color Field Painting	1585	200
	Neo-Expressionism	1304	200
	Magic Realism	1289	153
	Lyrical Abstraction	1124	200
<i>Post-Renaissance Art</i>			
	Realism	13972	200
	Romanticism	10929	200
	Baroque	5498	200
	Neoclassicism	3450	197
	Rococo	2868	200
<i>Renaissance Art</i>			
	Northern Renaissance	2867	192
	High Renaissance	1465	104
	Early Renaissance	1405	119
Total		151,151	4,105

Table 2: The styles and categories whose items are annotated for emotions in the WikiArt Emotions Project. The total number of items in each category as well as the number of items chosen for annotations are also shown. Some items belong to more than one category. The styles are shown in reverse chronological order. The categories are shown in decreasing order of the number of total items.

3.2. Designing the Questionnaire

Art can be annotated for emotions in several different ways. We describe below some of the choices we made, and the motivations behind them.

What Emotion Question to Ask: Just as text, one can label art for emotions from many perspectives: what emotion is the painter trying to convey, what emotion is felt by the

observer (the person viewing the painting), what emotion is felt by the entities depicted in the painting, how does the observer feel towards entities depicted in the painting, etc. All of these are worthy annotations to pursue. However, in this work we focus on the emotions evoked in the observer (the annotator) by the painting. That decided, it is still worth explicitly articulating what it means for a painting to evoke an emotion, as here too, many different interpretations exist. Should one label a painting with sadness if it depicts an entity in an unhappy situation, but the observer does not feel sadness on seeing the painting? How should one label an art depicting and evoking many different emotions, for example, a scene of an angry mother elephant defending her calf from a predator? And so on. For this annotation project we chose to instruct annotators to label all emotions that the painting brings to mind. Our exact instructions in this regard are shown below:

Art is imaginative human creation meant to be appreciated and evoke an emotional response. We will show you pieces of art, mostly paintings, one at a time. Your task is to identify the emotions that the art evokes, that is, all emotions that the art brings to mind. For example:

- *the image of someone suffering, brings to mind sadness.*
- *the image of a mother elephant fighting a lion to protect its calf, may bring to mind the mother's fear of losing the calf, anger at the lion, and admiration of the mother's bravery.*
- *the image of a rich tyrant enjoying his feast may bring to mind both his conceit and your disgust for him.*
- *the symmetry of lines and shapes in non-representative art may bring a sense of calm, whereas looking at the juxtaposition of shapes in a different art may evoke a sense of conflict.*

Which Emotions Apply Frequently to Art: Humans are capable of recognizing hundreds of emotions and it is likely that all of them can be evoked from paintings. However, some emotions are more frequent than others and come more easily to mind. Further, different individual experiences may prime different people to easily recall different sets of emotions. Also, emotion boundaries are fuzzy and some emotion pairs are more similar than others. All of this means that an open-ended question asking annotators to enter the emotions evoked through a text box is sub-optimal. Thus we chose to provide a set of options (each corresponding to a closely related emotions set) and asked annotators to check all emotions that apply. We chose the options from these sources:

- The psychology literature on basic emotions (Ekman, 1992; Plutchik, 1980; Parrot, 2001).
- The psychology literature on emotions elicited by art (Silvia, 2005; Silvia, 2009; Millis, 2001; Noy and Noy-Sharav, 2013).
- Our own annotations of the WikiArt paintings in a small pilot effort.

Polarity	Emotion Category	Abbreviation
<i>Positive</i>	gratitude , thankfulness, or indebtedness	grat
	happiness , calmness, pleasure, or ecstasy	happ
	humility , modesty, unpretentiousness, or simplicity	humi
	love or affection	love
	optimism , hopefulness, or confidence	opti
<i>Negative</i>	trust , admiration, respect, dignity, or honor	trus
	anger , annoyance, or rage	ange
	arrogance , vanity, hubris, or conceit	arro
	disgust , dislike, indifference, or hate	disg
	fear , anxiety, vulnerability, or terror	fear
	pessimism , cynicism, or lack of confidence	peSSI
	regret , guilt, or remorse	regr
<i>Other or Mixed</i>	sadness , pensiveness, loneliness, or grief	sadn
	shame , humiliation, or disgrace	sham
	agreeableness , acceptance, submission, or compliance	agre
	anticipation , interest, curiosity, suspicion, or vigilance	anti
	disagreeableness , defiance, conflict, or strife	disa
	surprise , surrealism, amazement, or confusion	surp
	shyness , self-consciousness, reserve, or reticence	shyn
neutral	neut	

Table 3: The list of emotions provided to annotators to label the image, the title text, and the art (title and image).

We grouped similar emotions into a single option. The final result was 19 options of closely-related emotion sets and a final neutral option. The options were arranged in three sets ‘positive’, ‘negative’, and ‘mixed or other’ as shown in Table 3, to facilitate ease of annotation. A text box was also provided for the annotators to capture any additional emotions that were not part of the pre-defined set of the 19 options. Many extra emotions were entered by the annotators, including uncertainty, amusement, and jealousy. However, none of the proposed additional emotions was used more than 20 times overall, which indicates that the pre-defined set of the 19 emotions we provided covered the art emotion space well.

Emotions evoked by the image alone, the title alone, and the art as a whole: We asked annotators to identify the emotions that the art evokes in three scenarios:

- Scenario I: we present only the image (no title), and ask the annotator to identify the emotions it evokes;
- Scenario II: we present only the title of the art (no image), and ask the annotator to identify the emotions it evokes;
- Scenario III: we present both the title and the image of the art, and ask the annotator to identify the emotions that the art as a whole evokes.

We instruct the annotators so that:

- when answering the question about the title (scenario II), they should not try to recollect what they answered earlier for the image that goes with it. Their response should be based solely on the title.
- When answering the question about the title–image combination (scenario III), they should not try to recollect what they answered earlier for the image alone or for the title alone. Their response should be based on what the art evokes.

To help the annotators focus only on the question at hand (and exclude influences from earlier responses), we show five instances in scenario I in a random order, followed by five instances in scenario II in a different random order, followed by five in scenario III in another random order.

Questions Asked: Scenario I question is shown below:

Q1. Examine the art above (the image). Which of the following describe the emotions it brings to mind? Select all that apply.

(Twenty options as shown in Table 3.)

The Questions for scenario II and III (Q2 and Q3), looked identical to Q1, except that they asked the annotator to examine the title and the art (image and title), respectively.

In Scenario III (image and title) we asked the following additional questions.

Q4. Which of the following best describes how you feel about the piece of art?

3: I like it a lot.

2: I like it.

1: I like it somewhat.

0: I neither like it nor dislike it.

–1: I dislike it somewhat.

–2: I dislike it.

–3: I dislike it a lot.

Q5. Which of the following is true about the image? Click all that apply.

– the image shows the face of at least one person or animal (select if there is any indication of a face anywhere in the image)

– the image shows the body of at least one person or animal (select if there is any indication of a body anywhere in the image)

– none of the above

Example instances were provided in advance with examples of suitable responses.

3.3. Crowdsourcing Annotations

We annotated all of our data by crowdsourcing. Links to the art and the annotation questionnaires were uploaded on the crowdsourcing platform, CrowdFlower.⁶ All annotators for our tasks had already agreed to the CrowdFlower terms of agreement. They chose to do our task among the hundreds available, based on interest and compensation provided. Respondents were free to annotate as many instances as they wished to. The annotation task was approved by the National Research Council Canada’s Institutional Review Board, which reviewed the proposed methods to ensure that they were ethical. Special attention was paid to obtaining informed consent and protecting participant anonymity.

About 2% of the instances were annotated internally beforehand (by the authors). These instances are referred to as gold instances. The gold instances are interspersed with other instances. If a crowd-worker answers a gold instance question incorrectly, they are immediately notified of the error. If the worker’s accuracy on the gold instances falls below 70%, they are refused further annotation, and all of their annotations are discarded. This serves as a mechanism to avoid malicious annotations. We mainly used the face-body question (Q5) as gold, but we also used, although sparingly, the emotion questions (Q1–Q3) as gold. Even though the emotional response to art is somewhat subjective, there are instances where some emotions clearly apply.

On the CrowdFlower task settings, we specified that we needed annotations from ten people for each instance. However, because of the way the gold instances are setup, they are annotated by more than ten people. The median number of annotations is still ten. In all, 308 people annotated between 20 and 1,525 pieces of art. A total of 41,985 sets of responses (for Q1–Q5) were obtained for the 4,105 pieces of art.

Annotation Aggregation and Machine Learning Datasets: For each item (image, title, or art), we will refer to the emotion that receives the majority of the votes from the annotators as the *predominant emotion*. In case of ties, all emotions with the majority vote are considered the predominant emotions. When aggregating the responses to obtain the full set of emotion labels for an item, we wanted to include not just the predominant emotion, but all others that apply, even if their presence is more subtle. Thus, we chose a somewhat generous aggregation criteria: if at least 40% of the responses (four out of ten people) indicate that a certain emotion applies, then that label is chosen. We will refer to this as *Ag4* dataset. 929 images, 1332 titles, and 823 paintings did not receive sufficient votes to be labeled with any emotion. These items were set aside. The rest of the items and their emotion labels can be used to train and test machine learning algorithms to predict the emotions evoked by art.

We also created two other versions of the labeled dataset by using an aggregation threshold of 30% and 50%, respectively. (If at least 30%/50% of the responses (three/five out of ten people) indicate that a certain emotion applies, then that label is chosen.) We will refer to them as

Ag3 and *Ag5* datasets. In our own future work, we will be working mainly with the *Ag4* version of the data, however, the other versions will also be made available for those interested in those variants.

Class Distribution: The ‘% votes’ rows of Table 4 show the percentage of times each emotion was selected by the annotators. The ‘*Ag3*’, ‘*Ag4*’, and ‘*Ag5*’ rows show the distribution of labels in the WikiArt Emotions dataset after aggregation. The numbers in each of the rows sum up to more than 100% because an item may be labeled with more than one emotion. Observe that anticipation, fear, happiness, humility, love, optimism, sadness, surprise, and trust get a high number of votes, whereas the rest get only a small percentage of votes. Observe also that as the aggregation threshold is increased (*Ag3* through *Ag5*), the percentage of tweets labeled with the less-frequent emotions reduces. For example, even though anger received 3.5% of the total art votes and 2% of the art pieces were labeled with anger when using *Ag3*, only 1% of tweets have anger as a label when using *Ag5*. Table 5 shows the percentage of times each emotion got the majority of votes, and was thus selected as the predominant emotion. Observe that some emotions such as humility and trust have markedly lower percentages as the predominant emotion than as one of the applicable emotions.

Tables 9 in the Appendix shows the proportions of the items in the WikiArt Emotions dataset corresponding to art with no face or body depicted, art with a body depicted but no face, and art with a face depicted broken down by art category. One can see that the vast majority of the Renaissance and Post-Renaissance art depict faces, with the lowest proportion corresponding to Romanticism (0.71) and the highest proportion corresponding to the Early and High Renaissance (1.00). Minimalism, Abstract Art, and Color Field Paintings have the lowest depiction rates of face or body (0 to 0.01). Table 10 in the Appendix shows these proportions broken down by emotion. We observe that certain emotions such as arrogance, shame, love, gratitude, and trust are depicted predominantly through faces (face-present proportions greater than 0.7). In contrast, emotions of anticipation, surprise, and disgust are predominantly evoked by paintings without any depictions of face or body (face-present proportions less than 0.4). The percentages for neutral indicate that a vast majority of the paintings that did not evoke any emotion did not depict a face or a body (neither face nor body proportion of 0.88).

4. Agreement

Emotion annotations of art are not expected to be highly consistent across people for a number of reasons, including: differences in human experience that impact how they perceive art, the subtle ways in which art can express affect, and fuzzy boundaries of affect categories. With the annotations on the WikiArts Emotion dataset, we can now determine the extent to which this agreement exists across different emotions, and how the agreements are impacted by attributes of the painting such as style, style category, and depictions of faces.

⁶<http://www.crowdfLOWER.com>

	agre	ange	anti	arro	disa	disg	fear	grat	happ	humi	love	opti	pess	regr	sadn	sham	shyn	surp	trus	neut
<i>Image (no Title)</i>																				
% votes	3.1	3.4	19.3	4.6	3.6	7.0	11.0	5.0	24.3	14.0	8.1	11.8	4.6	2.9	10.2	2.6	1.9	21.2	17.1	1.2
Ag3: % items label.	0.4	2.0	26.4	3.4	0.6	4.0	14.1	1.9	38.0	17.3	9.1	10.0	1.8	0.4	11.6	1.1	0.2	35.4	23.2	0.2
Ag4: % items label.	0.1	1.2	15.4	1.7	0.2	1.0	9.9	0.8	35.2	10.8	7.3	3.4	0.6	0.2	8.0	0.4	0.1	29.3	17.7	0.0
Ag5: % items label.	0.0	1.0	9.1	1.0	0.1	0.3	9.2	0.4	35.1	6.7	7.3	1.3	0.3	0.0	7.2	0.2	0.0	24.2	16.2	0.0
<i>Title (no Image)</i>																				
% votes	3.0	2.0	27.0	3.3	2.6	5.6	6.1	4.9	23.0	11.2	7.7	11.7	2.8	2.1	6.0	1.7	1.8	12.1	17.3	5.9
Ag3: % items label.	0.2	1.3	48.9	1.1	0.5	1.8	6.5	2.1	36.9	9.9	7.4	9.8	0.7	0.3	6.1	0.7	0.1	10.0	23.4	4.8
Ag4: % items label.	0.0	0.8	37.5	0.6	0.1	0.4	5.3	0.7	33.3	4.5	6.5	3.8	0.3	0.0	5.4	0.4	0.0	4.4	19.2	2.7
Ag5: % items label.	0.0	0.8	28.7	0.3	0.0	0.3	5.2	0.3	35.3	2.6	6.4	1.5	0.2	0.0	5.4	0.1	0.0	2.5	19.6	1.9
<i>Art (Image and Title)</i>																				
% votes	3.2	3.5	18.9	4.9	3.6	7.6	10.9	5.6	26.3	15.2	9.1	13.9	5.1	3.2	11.0	2.8	2.1	21.0	19.8	1.2
Ag3: % items label.	0.3	2.0	25.5	3.3	0.7	5.2	13.9	3.0	41.3	19.7	10.1	14.5	2.6	0.8	13.0	1.5	0.1	34.6	27.6	0.2
Ag4: % items label.	0.1	1.3	15.4	1.9	0.2	1.7	10.2	1.3	36.9	12.0	8.1	5.9	1.1	0.3	9.2	0.7	0.1	27.4	21.5	0.0
Ag5: % items label.	0.0	1.0	9.8	1.0	0.1	0.7	8.8	0.7	36.5	8.2	7.7	2.7	0.6	0.0	7.9	0.4	0.0	21.6	20.4	0.0

Table 4: **Applicable Emotion:** Percentage of votes for each emotion as being applicable and the percentage of items that were labeled with a given emotion (after aggregation of votes). Numbers greater than or equal to 10% are shown in bold.

	agre	ange	anti	arro	disa	disg	fear	grat	happ	humi	love	opti	pess	regr	sadn	sham	shyn	surp	trus	neut
<i>Image</i>	0.0	0.7	10.3	1.0	0.2	1.5	8.5	0.1	23.9	5.5	3.2	1.3	0.4	0.0	6.1	0.2	0.1	25.0	11.9	0.2
<i>Title</i>	0.1	0.4	34.9	0.4	0.2	1.1	4.2	0.2	23.8	3.1	3.1	2.3	0.2	0.0	3.8	0.2	0.0	4.4	14.5	3.2
<i>Art</i>	0.0	0.6	9.2	1.0	0.1	2.0	8.0	0.2	24.9	5.0	3.1	1.9	0.4	0.0	6.3	0.3	0.1	23.5	13.3	0.2

Table 5: **Predominant Emotion:** Percentage of items that were predominantly labeled with a given emotion. Numbers greater than or equal to 5% are shown in bold.

Figure 4 in the Appendix shows the Fleiss’ κ inter-rater agreement scores for each of the emotion classes across the different style categories. (Fleiss’ κ calculates the extent to which the observed agreement exceeds the one that would be expected by chance (Fleiss, 1971). However, note that correcting for chance remains controversial.⁷ Nonetheless, the relative variations in the values of κ are useful indicators of relative agreement.)

Observe that the κ scores range from close to 0 to 0.31 for the different emotion–category combinations. The close to 0 scores indicate that when considering all the paintings for some category–emotion pairs there is very little agreement beyond random chance. Nonetheless, there exist subsets of paintings, even for those category–emotion pairs, where agreement is higher. Scores closer to 0.31 indicate fair amounts of agreement for the sets as a whole. It should be noted that in general, agreement scores for art are lower than what one finds for text, which is expected. (Mohammad and Kiritchenko (2018) report Fleiss’ κ scores in the range of 0.32 to 0.47 for anger, fear, joy, and sadness conveyed by tweets, and lower scores for other emotions such as surprise, trust, and optimism.)

The κ scores are relatively higher for the Renaissance and Post-Renaissance art styles as compared to Modern Art and Contemporary Art. This is likely because, on average, Modern Art tends to be more abstract and non-representative. The κ scores are relatively high for basic emotions such as fear, happiness, sadness, anger, and

love, and lower for more complex emotions such as optimism, shame, guilt, and regret. Nonetheless, for certain category–emotion pairs the agreement is relatively high as compared to other categories and the same emotion. Examples include: Post-Renaissance categories with trust (especially, Romanticism–trust), Early Renaissance and Northern Renaissance with shame, Magic Realism with arrogance, Magic Realism with shame, Surrealism with surprise, Post-Impressionism with arrogance, and Early Renaissance with arrogance.

Figure 2 shows the agreement on the three partitions of the paintings corresponding to art with no face or body depicted, art with a body depicted but no face, and art with a face depicted. Observe that agreements are markedly higher for art with a body than with no body, and markedly higher again for art that depicts a face than art that only shows a body but no face. This is consistent with the hypothesis that human faces are an effective medium to convey emotions, and that depiction of even just a body without face is effective in conveying emotions and therefore elicits similar emotions in different observers.

Note that even though the κ scores shown here are lower than what one might find for other tasks such as part-of-speech tagging or named-entity recognition, these scores are closer to what one finds when annotating text for emotions (as indicated earlier). Further, the aggregation strategies of Ag4 and Ag5 described in the previous section, help filter out items with low inter-annotator agreement, and the remaining items can be used to train and test machine learning systems that detect emotions evoked by art.

⁷<http://www.john-uebersax.com/stat/kappa2.htm>
<http://www.agreestat.com/book3/bookexcerpts/chapter2.pdf>

Group	Emotion	Neither Face Nor Body Present	Body Present, No Face	Face Present	
Positive	gratitude	0.0100	0.0000	0.0400	
	happiness	0.1900	0.2000	0.2200	
	humility	0.0600	0.0600	0.0800	
	love	0.0300	0.1000	0.2500	
	optimism	0.0300	0.0400	0.0300	
	trust	0.0200	0.0600	0.1500	
Negative	anger	0.0200	0.0500	0.1400	
	arrogance	-0.0100	0.0200	0.0800	
	disgust	0.0200	0.0700	0.0700	
	fear	0.1100	0.1900	0.2400	
	pessimism	0.0300	0.0600	0.0600	
	regret	-0.0100	0.0200	0.0400	
	sadness	0.0700	0.1100	0.2300	
	shame	0.0100	0.1000	0.0800	
	Other	agreeableness	0.0000	0.0100	0.0000
		anticipation	0.0200	0.0300	0.0100
disagreeableness		0.0000	0.0100	0.0300	
neutral		0.0100	0.0000	0.0000	
shyness		0.0000	0.0200	0.0100	
surprise		0.0600	0.1100	0.0700	

Figure 2: Annotator agreement (Fleiss’ κ) for art pieces that show the face of a person or an animal (2,068 items), art pieces that show the body (and no face) of a person or an animal (227 items), and art pieces that show neither a face nor a body (1,810 items).

5. Emotions that Tend to Occur Together

Since we allow annotators to mark multiple emotions as being associated with an item, it is worth examining which emotions tend to be frequently voted for together (often evoked together by art). For every pair of emotions, i and j , we calculated the proportion of times an item received votes for both emotions i and j from an observer, out of all the votes for emotion i across all items. (See Figure 5 in the Appendix for the co-occurrence proportions.) The following pairs of emotions have scores greater than 0.4 indicating that when the first emotion is present, there is a greater than 40% chance that the second is also present. Emotion pairs of this kind include: gratitude–trust, love–happiness, pessimism–sadness, regret–sadness, shame–sadness, and surprise–anticipation. It is interesting to note that pessimism, regret, and shame have high co-occurrence with both fear and sadness. This suggests that these are complex emotions that include some elements of fear and sadness (two basic emotions) within them. Note also that for many emotion pairs, the association is markedly stronger in one direction than in the other direction. For example, pessimism is often indicative of sadness, but sadness is not often indicative of pessimism. As expected, highly contrasting emotions such as happiness and disgust have very low co-occurrence scores.

6. Emotions Evoked from the Image, the Title, and the Art

Titles of paintings impact how the observer views the art. They guide the observer by highlighting some aspect of the art.⁸ With our annotations, we wanted to quantify the impact titles have on the emotional response elicited by the art. Thus, as indicated earlier, we asked annotators to provide the emotions evoked by the image alone, the title alone,

⁸Titles are of different types such as sentimental, factual, abstract, and mysterious.

	% Match
a. image–art	53.32
b. title–art	30.80
c. image–title	27.20

Table 6: The percentage of annotations that have exactly the same emotion sets selected for image and art, title and art, or image and title by the same annotator.

and the art as whole (image and title). From these annotations, we calculated: a. the percentage of times a piece of art (image and title) was annotated with the same set of emotions as just the image; b. the percentage of times a piece of art (image and title) was annotated with the same set of emotions as just the title; and c. the percentage of times the image was annotated with the same set of emotions as just the title. Here, two sets of emotion labels are considered different if any one of the emotions in one set is not present in the other set. Table 6 shows the results.

Observe that the title often conveys a different set of emotions than the image alone or the art as a whole. In contrast, the art and image often convey the same sets of emotions, but there is a large percentage of instances where they differ. This shows that the title of an art plays a substantial role in the emotions evoked by the art.

7. What Makes Art Well Liked?

Art is judged in many ways: by how engaging, thought-provoking, or evocative it is, by the amount of expertise needed to create the art, by how easy it is to understand what is being communicated, by how pleasing the shapes and colours are, etc. Further, one may find the painting very engaging, but not want it in their home. Rather than asking people to judge all of these facets, we asked our annotators to simply rate the extent to which they liked or disliked the painting overall (Question 5). Table 7 shows the distribution of annotations that the art pieces received. Observe that the majority of the art is well liked, with the ratings of 2 (like it) and 1 (somewhat like it) being the most common. Around 18% of the pieces are marked as disliked (to varying degrees) and another 18% of the pieces are marked as being neither liked nor disliked.

Table 8 gives the average art ratings for the different art styles. Table 11 in the Appendix gives a breakdown of the art ratings by style category. Observe that Post-Renaissance and Renaissance pieces are liked the most (especially Realism, Rococo, Neoclassicism, and High Renaissance). Even though Modern Art overall received lower average rating score, Impressionism is the most liked category among all twenty-two considered in this work. Minimalism (Contemporary Art) and Art Informel (Modern Art) received the lowest ratings.

Table 12 in the Appendix gives a breakdown of the art ratings by emotion. We observe that art which evokes no emotion (neutral) and art which evokes disgust receive the lowest average scores. In contrast, art that evokes positive emotions such as love, gratitude, happiness, humility, optimism, and trust obtain some of the highest average scores. It is interesting to note that pieces of art that evoke negative

Like Description	Rating	% Annotations
like it a lot	3	17.41
like it	2	24.20
like it somewhat	1	22.01
neither like it nor dislike it	0	17.93
dislike it somewhat	-1	8.12
dislike it	-2	6.31
dislike it a lot	-3	4.02

Table 7: Distribution of art ratings.

Art Category	Ave. Rating
<i>Contemporary Art</i>	-0.07
<i>Modern Art</i>	0.67
<i>Post-Renaissance Art</i>	1.52
<i>Renaissance Art</i>	1.29
Average	0.91

Table 8: Average art ratings per art category.

emotions such as sadness, arrogance, and regret received markedly higher average scores than surprise and other negative emotions such as pessimism, shame, fear, and anger. Figure 3 in the Appendix shows the full breakdown of average art ratings for each category–emotion pair. Romanticism, Neoclassicism, and Impressionism paintings evoking love as well as Impressionism paintings evoking optimism received the highest scores (2.15–2.20).

8. Future Work and Applications

This paper examines attributes of a painting such as its style and content (face, body, none) and the emotions it evokes. We are currently analyzing the role of the features of the observer such as gender, age, and personality, on the emotions they perceive in art. We are also interested in conducting further annotations amongst the sets of paintings that evoke happiness, love, fear, and sadness, to determine the *intensities* of emotions they evoke. This will allow for a ranking of paintings by joy intensity, fear intensity, etc. We also want to determine whether paintings that evoke intense amounts of an emotion are also the ones that are, on average, liked more. We will also annotate the paintings that depict faces and bodies to determine whether the left or right side of the face or body is shown more prominently in the art. These annotations will help test the hypothesis that art that depicts the left side of a person’s face or body is on average found to be more appealing (left-cheek bias) (Powell and Schirillo, 2011; Blackburn and Schirillo, 2012).

The WikiArt Emotions dataset has many applications in automatic image and text processing, including those listed below:

- To train and test machine learning algorithms that can predict the emotions evoked by art. It will be interesting to determine the accuracies of unimodal (image- or text-only based) systems as well as multi-modal (text- and image-based) systems to detect the emotions. We will conduct experiments to determine the extent to which different modalities (text and image) are useful in detect-

ing emotion intensity, and under what circumstances they provide complementary information.

- To conduct experiments to determine what characteristics of images make them particularly evocative.
- To develop deep learning algorithms for art generation; for instance, to create systems that can transform a given piece of art (especially abstract paintings) to alter the affective reaction it evokes (for example, transforming a painting to make it evoke more sadness or more conflict).

We are currently developing an interactive visualization that allows users to search for WikiArt.org paintings with desired attributes such as style, genre, emotion, and average art ratings.

9. Conclusions

We created the *WikiArt Emotions Dataset*, which includes emotion annotations for more than 4,000 pieces of art from four western styles (Modern Art, Post-Renaissance Art, Renaissance Art, and Contemporary Art) and 22 style categories. The art is annotated for one or more of twenty emotion categories (including neutral). We also obtained separate emotion annotations for when the observer sees only the image and sees only the title of the art. We found that fear, happiness, love, and sadness were the dominant emotions that also obtained consistent annotations among the different annotators. Other emotions were also found to be more frequent and consistently annotated within paintings of particular style categories. Examination of the image only, title only, and whole art annotations revealed that the title of the art markedly impacted the emotions evoked by a painting.

The WikiArt Emotions dataset also has annotations for whether the painting includes the depiction of a face, a body, or neither. We found that paintings with faces (and to a lesser extent, paintings depicting a body but no face) elicited markedly more consistent emotion annotations. Finally, the dataset includes ratings given by people corresponding to the extent to which they liked or disliked the art. About 64% of the art was marked as liked (to some degree), 18% as disliked (to some degree), and 18% as neither liked nor disliked. We found that paintings evoking positive emotions were liked more, in general. We also found that paintings evoking certain positive emotions such as love were liked much more than paintings evoking other positive emotions such as humility. The difference was even more pronounced when comparing paintings evoking negative emotions; paintings evoking regret, arrogance, and sadness were liked much more than paintings evoking disgust, anger, or fear. Paintings evoking no emotion and disgust, were some of the least liked paintings.

The WikiArt Emotions Dataset is made freely available for educational purposes and to facilitate research in emotions, art, human psychology, and automatic image analysis/generation.

10. Acknowledgements

We thank Peter Turney for suggesting that we examine emotions evoked by art.

11. Bibliographical References

- Aiken, N. E. (1998). *The biological origins of art*. Praeger Publishers/Greenwood Publishing Group.
- Blackburn, K. and Schirillo, J. (2012). Emotive hemispheric differences measured in real-life portraits using pupil diameter and subjective aesthetic preferences. *Experimental brain research*, 219(4):447–455.
- Bollen, J., Mao, H., and Pepe, A. (2009). Modeling public mood and emotion: Twitter sentiment and socioeconomic phenomena. In *Proceedings of the Fifth International Conference on Weblogs and Social Media*, pages 450–453.
- Davies, S. (2012). *The artful species: aesthetics, art, and evolution*. OUP Oxford.
- De Silva, L. C., Miyasato, T., and Nakatsu, R. (1997). Facial emotion recognition using multi-modal information. In *Proceedings of the International Conference on Information, Communications and Signal Processing*, volume 1, pages 397–401.
- Deng, J., Dong, W., Socher, R., Li, L.-J., Li, K., and Fei-Fei, L. (2009). Imagenet: A large-scale hierarchical image database. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pages 248–255.
- Dutton, D. (2009). *The art instinct: Beauty, pleasure, & human evolution*. Oxford University Press, USA.
- Ekman, P. (1992). An argument for basic emotions. *Cognition and Emotion*, 6(3):169–200.
- Fasel, B. and Luetttin, J. (2003). Automatic facial expression analysis: a survey. *Pattern recognition*, 36(1):259–275.
- Fleiss, J. L. (1971). Measuring nominal scale agreement among many raters. *Psychological Bulletin*, 76(5):378–382.
- Kiritchenko, S., Zhu, X., and Mohammad, S. M. (2014). Sentiment analysis of short informal texts. *Journal of Artificial Intelligence Research*, 50:723–762.
- Lin, T., Maire, M., Belongie, S. J., Bourdev, L. D., Girshick, R. B., Hays, J., Perona, P., Ramanan, D., Dollár, P., and Zitnick, C. L. (2014). Microsoft COCO: common objects in context. *CoRR*, abs/1405.0312.
- Lucey, P., Cohn, J. F., Kanade, T., Saragih, J., Ambadar, Z., and Matthews, I. (2010). The extended Cohn-Kanade dataset (ck+): A complete dataset for action unit and emotion-specified expression. In *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition Workshops (CVPRW)*, pages 94–101.
- Miller, G. F. (2001). Aesthetic fitness: How sexual selection shaped artistic virtuosity as a fitness indicator and aesthetic preferences as mate choice criteria. *Bulletin of Psychology and the Arts*, 2(1):20–25.
- Millis, K. (2001). Making meaning brings pleasure: The influence of titles on aesthetic experiences. *Emotion*, 1(3):320.
- Mohammad, S. M. and Bravo-Marquez, F. (2017). WASSA-2017 shared task on emotion intensity. In *Proceedings of the Workshop on Computational Approaches to Subjectivity, Sentiment and Social Media Analysis (WASSA)*, Copenhagen, Denmark.
- Mohammad, S. M. and Kiritchenko, S. (2018). Understanding emotions: A dataset of tweets to study interactions between affect categories. In *Proceedings of the 11th Edition of the Language Resources and Evaluation Conference (LREC-2018)*, Miyazaki, Japan.
- Mohammad, S. M., Bravo-Marquez, F., Salameh, M., and Kiritchenko, S. (2018). Semeval-2018 Task 1: Affect in tweets. In *Proceedings of International Workshop on Semantic Evaluation (SemEval-2018)*, New Orleans, LA, USA.
- Mohammad, S. (2012a). #Emotional Tweets. In *Proceedings of the First Joint Conference on Lexical and Computational Semantics (*SEM)*, pages 246–255, Montréal, Canada.
- Mohammad, S. M. (2012b). From once upon a time to happily ever after: Tracking emotions in mail and books. *Decision Support Systems*, 53(4):730–741.
- Noy, P. and Noy-Sharav, D. (2013). Art and emotions. *International Journal of Applied Psychoanalytic Studies*, 10(2):100–107.
- Parrot, W. (2001). *Emotions in Social Psychology*. Psychology Press.
- Plutchik, R. (1980). A general psychoevolutionary theory of emotion. *Emotion: Theory, research, and experience*, 1(3):3–33.
- Powell, W. R. and Schirillo, J. A. (2011). Hemispheric laterality measured in Rembrandt’s portraits using pupil diameter and aesthetic verbal judgements. *Cognition & emotion*, 25(5):868–885.
- Silvia, P. J. (2005). Emotional responses to art: From collocation and arousal to cognition and emotion. *Review of general psychology*, 9(4):342.
- Silvia, P. J. (2009). Looking past pleasure: anger, confusion, disgust, pride, surprise, and other unusual aesthetic emotions. *Psychology of Aesthetics, Creativity, and the Arts*, 3(1):48.
- Susskind, J., Littlewort, G., Bartlett, M., Movellan, J., and Anderson, A. (2007). Human and computer recognition of facial expressions of emotion. *Neuropsychologia*, 45(1):152–162.
- Wang, B., Liakata, M., Zubiaga, A., Procter, R., and Jensen, E. (2016). SMILE: Twitter emotion classification using domain adaptation. In *Proceedings of the CEUR Workshop*, volume 1619, pages 15–21.
- Yang, C., Lin, K. H.-Y., and Chen, H.-H. (2007). Building emotion lexicon from weblog corpora. In *Proceedings of the 45th Annual Meeting of the ACL*, pages 133–136.
- Zheng, W., Tang, H., Lin, Z., and Huang, T. S. (2010). Emotion recognition from arbitrary view facial images. In *Proceedings of the European Conference on Computer Vision*, pages 490–503.
- Zhu, X., Guo, H., Mohammad, S., and Kiritchenko, S. (2014). An empirical study on the effect of negation words on sentiment. In *Proceedings of the 52nd Annual Meeting of the Association for Computational Linguistics*, pages 304–313, Baltimore, Maryland, June.

Art Category	Neither face nor body	Body, no face	Face pres.
<i>Contemporary Art</i>			
Minimalism	0.99	0.00	0.01
<i>Modern Art</i>			
Abstract Art	0.97	0.01	0.01
Abstract Expressionism	0.92	0.04	0.04
Art Informel	0.95	0.01	0.04
Color Field Painting	0.99	0.01	0.01
Cubism	0.49	0.07	0.45
Expressionism	0.19	0.10	0.71
Impressionism	0.22	0.09	0.69
Lyrical Abstraction	1.00	0.00	0.01
Magic Realism	0.38	0.10	0.52
Neo-Expressionism	0.33	0.13	0.55
Pop Art	0.48	0.07	0.44
Post-Impressionism	0.38	0.08	0.54
Surrealism	0.49	0.10	0.41
<i>Post-Renaissance Art</i>			
Baroque	0.12	0.06	0.82
Neoclassicism	0.03	0.04	0.94
Realism	0.17	0.10	0.74
Rococo	0.03	0.06	0.92
Romanticism	0.19	0.11	0.71
<i>Renaissance Art</i>			
Early Renaissance	0.00	0.00	1.00
High Renaissance	0.00	0.00	1.00
Northern Renaissance	0.02	0.04	0.94

Table 9: Proportions of items in the WikiArt Emotions dataset corresponding to art with no face or body depicted, art with a body depicted but no face, and art with a face depicted **broken down by art category**.

Appendix

Tables 9 and 10 show the proportions of items in the WikiArt Emotions dataset corresponding to art with no face or body depicted, art with a body depicted but no face, and art with a face depicted broken down by category and emotion, respectively.

Table 11 gives a breakdown of the art ratings by art category. Table 12 gives a breakdown of the art ratings by emotion. Figure 3 shows the full breakdown of average art ratings for each category–emotion pair.

Figure 4 shows the Fleiss’ κ inter-rater agreement scores for each of the emotion classes across the different style categories.

Figure 5 shows, for every pair of emotions, i and j , the proportion of times an item received votes for both emotions i and j from an observer, out of all the votes for emotion i .

Some freely available computer vision datasets:

- <http://www.computervisiononline.com/datasets>
- <http://www.cvpapers.com/datasets.html>
- <http://riemenschneider.hayko.at/vision/dataset/>
- http://clickdamage.com/sourcecode/cv_datasets.php
- <http://cocodataset.org/>
- <http://www.image-net.org>

Emotion	Neither face nor body	Body, no face	Face pres.
<i>Positive</i>			
gratitude	0.21	0.05	0.74
happiness	0.30	0.07	0.62
humility	0.29	0.07	0.64
love	0.17	0.04	0.80
optimism	0.38	0.06	0.56
trust	0.19	0.04	0.76
<i>Negative</i>			
anger	0.40	0.04	0.56
arrogance	0.20	0.03	0.77
disgust	0.57	0.04	0.38
fear	0.41	0.06	0.53
pessimism	0.44	0.06	0.50
regret	0.32	0.06	0.62
sadness	0.31	0.06	0.62
shame	0.22	0.08	0.71
<i>Other or Mixed</i>			
agreeableness	0.29	0.05	0.65
anticipation	0.60	0.05	0.35
disagreeableness	0.49	0.05	0.46
shyness	0.40	0.06	0.54
surprise	0.71	0.04	0.25
neutral	0.88	0.02	0.10

Table 10: Proportions of items in the WikiArt Emotions dataset corresponding to art with no face or body depicted, art with a body depicted but no face, and art with a face depicted **broken down by emotion**.

WikiArt Emotions Project homepage:

<http://saifmohammad.com/WebPages/wikiartemotions.html>

Art Category	Ave. Rating
<i>Contemporary Art</i>	
Minimalism	-0.07
Average	-0.07
<i>Modern Art</i>	
Abstract Art	0.29
Abstract Expressionism	0.20
Art Informel	0.04
Color Field Painting	0.22
Cubism	0.75
Expressionism	0.98
Impressionism	1.69
Lyrical Abstraction	0.47
Magic Realism	1.29
Neo-Expressionism	0.39
Pop Art	0.48
Post-Impressionism	1.43
Surrealism	0.45
Average	0.67
<i>Post-Renaissance Art</i>	
Baroque	1.39
Neoclassicism	1.56
Realism	1.58
Rococo	1.58
Romanticism	1.49
Average	1.52
<i>Renaissance Art</i>	
Early Renaissance	1.20
High Renaissance	1.50
Northern Renaissance	1.18
Average	1.29
Average (all categories)	0.91

Table 11: Average art ratings per art category.

Emotion	Ave. Rating
<i>Positive</i>	
gratitude	1.87
happiness	1.79
humility	1.62
love	1.95
optimism	1.72
trust	1.76
Average	1.79
<i>Negative</i>	
anger	0.41
arrogance	0.80
disgust	-0.38
fear	0.27
pessimism	0.39
regret	0.89
sadness	0.79
shame	0.48
Average	0.46
<i>Other or Mixed</i>	
agreeableness	1.60
anticipation	0.99
disagreeableness	0.60
shyness	1.10
surprise	0.49
neutral	-0.43
Average	0.73
Average (all emotions)	0.94

Table 12: Average art ratings per emotion.

Art Style	Art Category	Agree..	Anger	Antici..	Arroga..	Disagr..	Disgust	Fear	Gratitu..	Happi..	Humili..	Love	Optimi..	Pessi..	Regret	Sadne..	Shame	Shyne..	Surpri..	Trust	Neutral
Contemporary Art	Minimalism	1.19	-0.22	0.43	0.23	-0.24	-0.76	-0.18	1.50	1.65	0.48	1.30	1.07	-0.34	0.41	-0.15	-0.54	-0.13	-0.06	1.52	-0.66
	Abstract Art	1.62	-0.22	0.61	0.52	0.15	-0.70	-0.23	1.74	1.65	1.14	1.91	1.33	-0.26	0.23	-0.03	0.02	0.81	0.28	1.54	-0.21
Modern Art	Abstract Expressionism	1.46	0.21	0.64	0.75	0.04	-0.79	-0.22	1.88	1.55	0.92	1.70	1.36	0.05	0.85	0.33	0.08	0.71	0.28	1.62	-0.41
	Art Informel	1.70	-0.30	0.56	0.16	-0.08	-0.75	-0.34	1.84	1.55	0.84	1.50	1.34	-0.35	0.00	0.22	-0.40	0.40	0.13	1.46	-0.32
Renaissance Art	Color Field Painting	1.19	-0.52	0.48	0.33	-0.34	-0.88	-0.67	1.44	1.53	0.77	1.51	1.27	-0.59	0.31	-0.13	-1.17	0.03	0.08	1.61	-0.47
	Cubism	1.86	0.34	0.93	0.55	0.52	-0.26	0.31	1.60	1.61	1.32	1.80	1.55	0.25	0.68	0.68	0.25	0.94	0.52	1.65	-0.23
Post Renaissance Art	Expressionism	1.34	0.67	1.20	0.65	1.10	-0.14	0.31	1.84	1.61	1.56	1.81	1.71	0.69	1.03	0.70	0.40	1.00	0.82	1.67	0.00
	Impressionism	1.76	1.51	1.69	1.16	1.66	0.82	1.09	1.94	1.99	1.96	2.15	2.15	1.01	1.48	1.20	1.01	1.57	1.61	1.96	1.20
Renaissance Art	Lyrical Abstraction	1.38	0.18	0.85	1.08	0.44	-0.76	-0.14	1.98	1.75	1.26	1.78	1.51	0.18	0.66	0.33	0.83	0.85	0.39	1.82	-0.44
	Magic Realism	1.92	0.92	1.44	1.20	1.05	-0.16	0.67	2.01	1.92	1.74	2.07	2.04	0.74	1.50	1.04	0.59	1.65	1.15	1.95	0.30
Post Renaissance Art	Neo-Expressio..	1.54	-0.22	0.86	0.13	0.04	-0.70	-0.27	1.43	1.60	1.09	1.77	1.39	0.11	0.30	0.42	-0.01	0.86	0.45	1.59	-0.59
	Pop Art	1.34	0.19	0.79	0.32	0.52	-0.54	0.07	1.62	1.59	1.07	1.62	1.39	0.31	0.56	0.50	0.01	0.68	0.34	1.43	-0.82
Renaissance Art	Post-Impressi..	1.56	1.04	1.40	0.93	1.54	0.59	0.53	1.94	1.84	1.71	2.01	1.90	0.78	1.14	1.09	0.87	1.49	1.19	1.82	-0.13
	Surrealism	1.54	-0.09	0.98	0.43	0.10	-0.78	-0.21	1.74	1.66	1.33	1.86	1.52	0.13	0.45	0.50	0.08	0.37	0.41	1.71	-0.48
Post Renaissance Art	Baroque	1.71	1.26	1.68	1.02	1.35	0.78	1.01	1.92	1.77	1.84	2.03	1.81	1.05	1.41	1.15	0.93	1.89	1.57	1.80	-1.00
	Neoclassicism	1.91	1.13	1.80	1.08	1.59	0.85	1.10	2.09	1.94	2.04	2.16	1.97	1.28	1.52	1.21	1.12	1.69	1.77	1.84	0.50
Renaissance Art	Realism	1.77	0.64	1.67	1.14	1.53	0.63	0.96	2.09	1.93	1.88	2.12	2.03	0.94	1.34	1.28	1.16	1.79	1.59	1.92	0.88
	Rococo	1.79	0.87	1.60	1.01	1.14	0.52	0.84	2.02	1.96	1.90	2.10	1.99	0.99	1.44	1.17	1.06	1.88	1.52	1.82	0.00
Renaissance Art	Romanticism	1.73	1.02	1.71	1.23	1.16	0.70	1.01	2.08	1.88	1.89	2.20	1.99	0.94	1.31	1.24	0.88	1.70	1.57	1.82	0.50
	Early Renaissance High	1.56	0.67	1.41	0.63	1.01	0.28	0.65	1.70	1.62	1.63	1.81	1.61	0.73	0.82	0.69	0.42	1.26	1.64	1.65	0.00
Renaissance Art	Renaissance Northern	1.52	1.07	1.69	1.11	1.48	0.21	0.81	1.87	1.81	1.88	1.99	1.94	0.95	1.24	1.02	0.31	1.45	1.54	1.84	2.00
	Renaissance	1.40	0.62	1.31	0.71	0.76	0.01	0.70	1.75	1.67	1.69	1.81	1.75	0.61	0.64	0.90	0.61	1.56	1.47	1.66	0.22

Figure 3: A breakdown of average art ratings for each category–emotion pair in the WikiArt Emotions dataset. The scores range from -3 (most disliked) to 3 (most liked). Positive scores are shown in shades of green; darker shades for scores closer to 3. Negative scores are shown in orange; darker shades for scores closer to -3.

Style	Category	Agree..	Anger	Antici..	Arroga..	Disagr..	Disgust	Fear	Gratitu..	Happi..	Humili..	Love	Optimi..	Pessi..	Regret	Sadne..	Shame	Shyne..	Surpri..	Trust	Neutral
Contemporary Art	Minimalism	-0.01	0.01	0.00	0.01	-0.01	-0.02	0.07	-0.01	0.02	0.00	0.08	0.01	0.01	-0.01	0.07	0.01	0.01	-0.01	0.01	-0.01
	Abstract Art	0.01	0.04	0.01	-0.02	0.00	0.01	0.08	0.00	0.09	0.00	0.00	0.01	0.01	-0.01	0.02	0.04	-0.01	0.00	-0.01	0.00
	Abstract Expressionism	0.01	0.02	0.02	-0.02	-0.02	0.02	0.10	-0.01	0.06	0.03	0.03	0.03	0.04	0.00	0.04	0.00	-0.01	0.01	-0.01	0.00
	Art Informel	0.00	0.03	0.00	-0.01	0.00	-0.01	0.07	-0.01	0.07	-0.02	0.03	0.03	0.04	0.00	0.03	0.01	0.00	-0.01	-0.01	0.00
Modern Art	Color Field Painting	0.01	0.02	-0.01	-0.01	0.00	0.00	0.03	0.02	0.07	0.00	0.00	0.02	0.04	-0.02	0.07	0.03	0.00	0.01	0.00	0.00
	Cubism	-0.01	0.03	0.02	0.06	0.01	0.01	0.11	0.00	0.11	0.05	0.11	0.01	0.04	0.03	0.09	0.07	0.01	0.07	0.05	0.02
	Expressionism	0.00	0.04	0.00	0.04	-0.01	0.04	0.20	0.03	0.18	0.06	0.18	0.03	0.10	0.07	0.22	0.06	0.02	0.01	0.04	0.00
	Impressionism	0.00	0.08	-0.01	0.06	0.00	0.02	0.16	0.02	0.21	0.01	0.19	0.03	0.06	0.00	0.19	0.02	-0.01	-0.02	0.05	0.00
Post Renaissance Art	Lyrical Abstraction	-0.01	0.02	-0.01	-0.02	-0.02	0.00	0.11	0.01	0.09	0.02	0.03	0.01	0.02	-0.01	0.06	-0.01	0.00	0.02	-0.01	0.01
	Magic Realism	0.00	0.07	0.04	0.14	0.00	0.05	0.14	0.02	0.13	0.06	0.13	0.06	0.01	0.02	0.11	0.15	0.00	0.06	0.07	0.04
	Neo-Expressionism	-0.01	0.10	0.02	0.04	0.01	0.03	0.13	0.00	0.20	0.03	0.18	0.04	0.05	0.01	0.11	0.05	0.00	0.03	0.05	0.00
	Pop Art	0.00	0.06	0.00	0.03	0.01	0.02	0.11	-0.01	0.16	0.04	0.20	0.02	0.02	0.01	0.11	0.06	0.00	0.06	0.09	0.01
Renaissance Art	Post-Impressionism	-0.01	0.10	0.00	0.11	0.01	0.05	0.05	-0.01	0.13	0.03	0.18	0.02	0.03	0.03	0.20	0.05	0.05	0.03	0.05	0.00
	Surrealism	0.02	0.11	0.02	0.01	0.00	0.03	0.19	0.00	0.16	0.05	0.19	0.03	0.02	0.00	0.09	0.03	0.02	0.11	-0.01	-0.01
	Baroque	-0.01	0.20	0.02	0.10	0.05	0.09	0.29	0.04	0.23	0.06	0.22	0.03	0.01	0.02	0.16	0.12	0.00	0.01	0.12	0.00
	Neoclassicism	-0.01	0.10	0.00	0.10	0.02	0.04	0.23	0.01	0.20	0.10	0.26	0.00	0.03	0.04	0.26	0.09	0.00	-0.01	0.13	0.00
Renaissance Art	Realism	0.03	0.06	0.00	0.09	0.04	0.10	0.24	0.01	0.22	0.08	0.20	0.02	0.09	0.03	0.25	0.06	0.00	-0.02	0.09	0.00
	Rococo	0.00	0.09	-0.01	0.07	0.03	0.08	0.24	0.00	0.21	0.08	0.27	0.02	0.07	0.02	0.28	0.04	0.00	0.00	0.14	0.00
	Romanticism	0.00	0.20	-0.02	0.09	0.02	0.08	0.31	0.01	0.23	0.07	0.13	0.04	0.05	0.04	0.23	0.09	-0.01	-0.02	0.19	0.05
	Early Renaissance	-0.01	0.17	-0.01	0.14	0.07	0.00	0.18	0.04	0.12	0.03	0.28	0.00	0.03	0.05	0.30	0.12	-0.01	-0.02	0.05	0.00
Renaissance Art	High Renaissance	0.01	0.10	-0.02	0.07	0.02	0.09	0.23	0.06	0.13	0.04	0.27	0.02	0.00	0.14	0.05	0.05	0.01	-0.02	0.04	0.00
	Northern Renaissance	-0.02	0.10	-0.02	0.05	0.07	0.09	0.22	0.04	0.14	0.05	0.31	0.02	0.04	0.04	0.26	0.11	0.00	0.02	0.10	0.00

Figure 4: Annotator agreement (Fleiss' κ) per emotion and art category. The number of items in each category is shown in Table 2.

Emotion	Agreee..	Anger	Antici..	Arroga..	Disagr..	Disgust	Fear	Gratitu..	Happi..	Humili..	Love	Optimi..	Pessi..	Regret	Sadne..	Shame	Shyne..	Surpri..	Trust
agreeableness	1.00	0.11	0.35	0.17	0.12	0.12	0.14	0.23	0.30	0.25	0.18	0.21	0.12	0.12	0.11	0.09	0.11	0.21	0.38
anger	0.17	1.00	0.24	0.26	0.24	0.28	0.35	0.19	0.18	0.19	0.16	0.15	0.22	0.20	0.26	0.20	0.16	0.22	0.22
anticipation	0.04	0.02	1.00	0.04	0.03	0.03	0.05	0.05	0.10	0.06	0.05	0.06	0.03	0.03	0.04	0.02	0.02	0.25	0.15
arrogance	0.16	0.16	0.26	1.00	0.17	0.15	0.18	0.19	0.20	0.18	0.17	0.14	0.15	0.14	0.15	0.10	0.12	0.21	0.33
disagreeableness	0.15	0.18	0.32	0.22	1.00	0.29	0.31	0.27	0.26	0.31	0.21	0.14	0.25	0.22	0.22	0.16	0.19	0.34	0.24
disgust	0.07	0.11	0.14	0.09	0.14	1.00	0.16	0.12	0.12	0.14	0.09	0.06	0.13	0.11	0.12	0.09	0.09	0.14	0.10
fear	0.07	0.12	0.20	0.10	0.14	0.14	1.00	0.12	0.12	0.14	0.09	0.06	0.17	0.15	0.27	0.11	0.09	0.16	0.12
gratitude	0.15	0.08	0.27	0.13	0.14	0.13	0.14	1.00	0.34	0.34	0.23	0.21	0.14	0.14	0.14	0.07	0.10	0.20	0.43
happiness	0.04	0.02	0.11	0.03	0.03	0.03	0.03	0.08	1.00	0.15	0.16	0.18	0.03	0.03	0.04	0.02	0.02	0.07	0.13
humility	0.07	0.03	0.14	0.05	0.07	0.06	0.07	0.15	0.31	1.00	0.13	0.15	0.07	0.08	0.09	0.04	0.05	0.09	0.20
love	0.07	0.04	0.17	0.07	0.07	0.06	0.07	0.15	0.49	0.19	1.00	0.19	0.06	0.07	0.08	0.04	0.06	0.14	0.24
optimism	0.06	0.03	0.13	0.04	0.03	0.03	0.03	0.09	0.35	0.15	0.12	1.00	0.03	0.04	0.04	0.03	0.03	0.06	0.20
pessimism	0.13	0.15	0.28	0.17	0.23	0.24	0.35	0.24	0.25	0.27	0.17	0.13	1.00	0.25	0.42	0.17	0.15	0.26	0.21
regret	0.17	0.18	0.32	0.22	0.26	0.27	0.39	0.31	0.33	0.39	0.25	0.19	0.33	1.00	0.49	0.23	0.18	0.29	0.27
sadness	0.06	0.09	0.17	0.08	0.10	0.11	0.27	0.11	0.13	0.16	0.10	0.07	0.20	0.18	1.00	0.14	0.08	0.14	0.14
shame	0.17	0.23	0.24	0.20	0.24	0.27	0.39	0.20	0.20	0.25	0.18	0.17	0.28	0.29	0.48	1.00	0.22	0.24	0.20
shyness	0.19	0.17	0.28	0.21	0.25	0.24	0.27	0.26	0.26	0.31	0.23	0.18	0.23	0.21	0.25	0.20	1.00	0.24	0.25
surprise	0.06	0.04	0.53	0.06	0.08	0.06	0.08	0.08	0.13	0.09	0.09	0.06	0.06	0.06	0.07	0.04	0.04	1.00	0.24
trust	0.07	0.03	0.23	0.07	0.04	0.03	0.04	0.12	0.17	0.13	0.11	0.14	0.04	0.04	0.05	0.02	0.03	0.17	1.00

Figure 5: The proportion of votes for each pair of emotions. The number in cell (i, j) shows the proportion of items annotators labeled with both emotions i and j out of all the items annotators labeled with emotion i .