Seeing the Differences in Artistry among Art Fields by using Multi-task Learning

Ryo Sato, Fumihiko Sakaue and Jun Sato
Nagoya Institute of Technology, Japan
{r-sato@cv., sakaue@, junsato@}nitech.ac.jp

Keywords: Multi-Task Learning, Relevance of Art Fields, Analysis of Artistry.

Abstract: In this paper, we propose a method for analyzing the relevance of artistry among multiple art fields by using deep neural networks. Artistry is thought to exist in various man-made objects, such as paintings, sculptures, architectures, and gardens. However, we are not sure if the artistry or the human aesthetic sensitivities in these different art fields is the same or different. Therefore, we in this paper propose a method for analyzing the relevance of artistry among multiple art fields by using deep neural networks. In particular, we show that by using the multi-task learning, the relevance of multiple art fields can be analyzed efficiently.

1 INTRODUCTION

In recent years, various new methods have been developed in deep learning, and their applications are advancing in various fields such as image recognition and generation. Especially in recent years, it has become possible to generate highly accurate images that deceive the human eye by using GAN (Goodfellow et al., 2014). While the technology for generating realistic images has evolved significantly, the technology for realizing the artistic creativity of human beings is still developing.

As a deep learning network that simulates human artistry, Elgammal et al. proposed a method for generating artistic paintings by using a GAN-based network called CAN (Elgammal et al., 2017). They considered the relationship between the knowledge of past works and new artistry, and based on that, they tried to minimize the deviation from the distribution of artistry while maximizing the deviation from the existing style for generating creative artistic works. As a result, they succeeded in creating a work that does not follow the traditional style of artistry.

While research on the generation of artistic works is progressing, the analysis of artistry, i.e. artistic creativity of human beings, has not yet progressed much. Many methods have been proposed for classifying artistic paintings according to artist, style and genre (Tan et al., 2016; Anwer et al., 2016; Bianco et al., 2019), but these methods simply classify artistic works based on their similarity and cannot measure the artistic creativity that exists in these artistic works.

Figure 1: Various art fields.

The concept of artistry is vague and its definition is not clear even among experts (The Metaphysics Research Lab, 1995). Fig. 2 shows Picasso's "The Weeping Woman" (Picasso,) and Van Gogh's "Sunflowers" (van Gogh,). Both are famous works of art and are highly evaluated by experts and the general public, but it is not clear where in these works we feel the artistry. Also it is not clear why the works made by ordinary people are not artistic. While it can be confirmed that the artistry is ambiguous, it is certain that there is something that fascinates many people in art works such as "The Weeping Woman" and "Sunflower". That is, it seems that these images have sufficient information for determining the presence or absence of artistry. Therefore, in this research, we analyze what the artistry of these works of art is by using deep neural network.

It is known that paintings are not the only things that have artistry, but various other man-made objects such as sculptures, architecture, and gardens are considered to have artistry as shown in Fig. 1. Therefore, by clarifying the artistry common to these multiple fields, it is considered that the artistry can be grasped more objectively. However, it is not clear whether

Sato, R., Sakaue, F. and Sato, J.
Seeing the Differences in Artistry among Art Fields by using Multi-task Learning.
DOI: 10.5220/0010902300003124
In *Proceedings of the 17th International Joint Conference on Computer Vision, Imaging and Computer Graphics Theory and Applications (VISIGRAPP 2022) - Volume 5: VISAPP*, pages 609-616
ISBN: 978-989-758-555-5; ISSN: 2184-4321

”The Weeping Woman” by Picasso

”Sunflower” by Gogh

Figure 2: Work of art in painting.

Figure 3: Output images of CAN.

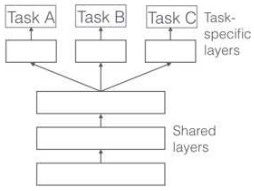
the artistry in these different fields is the same or different. If the artistry is common among the multiple fields, the network that can identify the artistry of one field is considered to have the ability to identify the artistry of other fields as well.

Thus, in this research, we train classifiers that distinguish between the presence and absence of artistry in multiple art fields using multi-task learning (Caruana, 1997). By analyzing the improvement in artistry identification in the multi-task learning, we clarify the commonalities of artistry among different fields. We also analyze the characteristics of the shared layer in the network after the training and show the relationships of artistry among different fields.

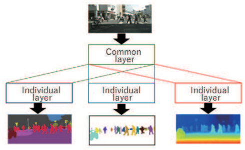
2 RELATED WORK

Many methods have been proposed for classifying artistic paintings according to their artist, style and genre (Tan et al., 2016; Anwer et al., 2016; Bianco et al., 2019). However, these methods simply classify artistic works based on their similarity, and they cannot estimate the artistic creativity that exists in these artistic works.

Although the analysis of human creativity is difficult and has not yet progressed much, the study on generating creative works by using deep neural network has started. Elgammal et al. (Elgammal et al., 2017) have developed Creative Adversarial Networks (CAN), which outputs artistic paintings by using generative adversarial network (GAN) (Goodfellow et al., 2014; Isola et al., 2017). In their method, the relationship between the past works and new artistry is analyzed, and the deviation from the distribution of artistry is minimized while the deviation from the existing style is maximized for generating creative artistic works. As a result, they succeeded in creating a work that does not follow the traditional style of artistry as shown in Fig. 3 and making us think that the work is artistic.



(a) basic network structure



(b) example of multi-task learning (Kendall et al., 2018)

Figure 4: Multi-task learning.

While networks that perform single tasks are being developed for solving various problems, networks that efficiently perform multiple tasks are also being developed. This is called multi-task learning (Caruana, 1997). The multi-task learning learns multiple related tasks at the same time, and the common features of these tasks can be obtained in a single network. As a result, the multi-task learning can improve the accuracy of individual task execution (Kendall et al., 2018; Liu et al., 2019).

Traditionally, multi-task learning networks in computer vision follow a simple outline which consists of a global feature extractor made of convolutional layers shared by all tasks followed by an individual output branch for each task as shown in Fig. 4 (a). For example, when we want to train the segmentation, object detection, and depth estimation, the network structure of the multi-task learning can be designed as shown in Fig. 4 (b) (Kendall et al., 2018). More recently, various forms of more complex network structures have been proposed. Since the objective of this research is to clarify the existence of artistry common to all art fields, we use a network structure for multi-task learning suitable for this objective.

3 ARTISTRY ANALYSIS IN MULTIPLE ART FIELDS

In this research, we use multi-task learning for analyzing the relationships among multiple different art fields. As shown in Fig. 1, artistry may exists in

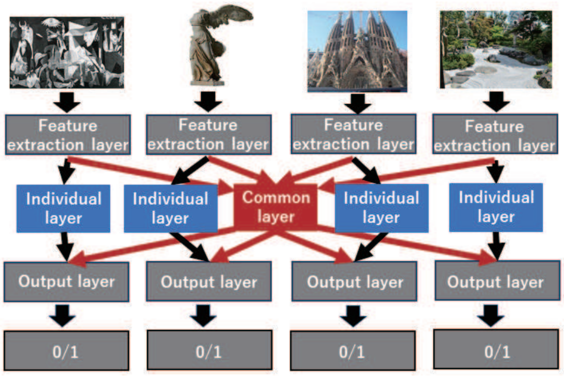


Figure 5: Network structure of multi-task learning in this research.

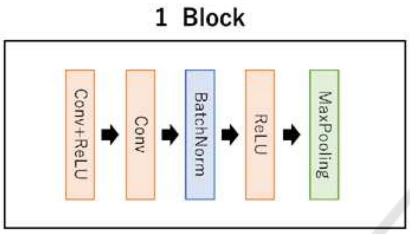


Figure 6: Basis block used in this research.

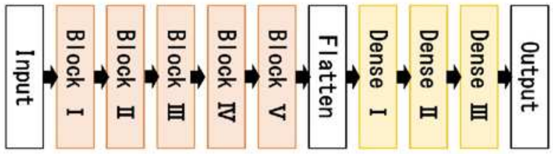


Figure 7: Network structure for single-task learning.

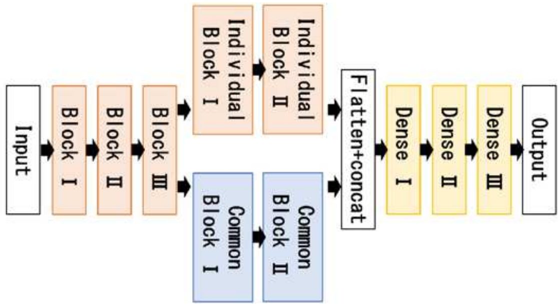


Figure 8: Network structure for multi-task learning.

many man-made objects, such as paintings, sculptures, architecture, and gardens. However, it is unclear whether the artistry of those different art fields is measured on the same scale or on different scales. In this research, we analyze the commonality of artistry among these art fields by learning the relationships of artistry among these art fields using multi-task learning.

By using multi-task learning, it is possible to learn common features between tasks and task-specific features at the same time. Therefore, in this research, we construct a classifier that identifies the magnitude of artistry of each work in each art field by single-task learning and multi-task learning respectively. In the multi-task learning, it is possible to learn the common part in the artistry more efficiently. Therefore, if there exists a commonality of artistry in these art fields, it is expected that the accuracy of identifying the artistry will be improved by using the multi-task learning compared to the single-task learning.

For example, if there is something in common between the artistry of painting and the artistry of sculpture, the score of the multi-task learning will be better than that of the single-task learning in the artistry identification task for painting and sculpture. That is, the depth of the relationship of artistry in each art field can be measured by observing the degree of performance improvement of multi-task learning with respect to single-task learning.

In this research, we take up four fields of art, painting, architecture, sculpture, and garden, and analyze the relationship of artistry in these four fields by using multi-task learning.

4 MULTI-TASK LEARNING FOR ARTISTRY IDENTIFICATION

The network structure of multi-task learning used in this research is as shown in Fig. 5. As shown in Fig. 5, the network takes images of painting, architecture, sculpture, and garden as inputs and the magnitude of their artistry as outputs. In this research, the magnitude of artistry is considered to be a binary value with or without artistry, and so the output of the network is set to 0 or 1.

Since the appearance of art works in each field is very different, we first input the image to the individual feature extraction layers that extract the unique features in each art field, as shown in Fig. 5. Since it is considered that the artistry that has become sufficiently abstract has features that are common among multiple fields, the output from the individual feature extraction layers is passed to the shared common layers. Since the final identification of artistry may differ in each field, the output from the common layers is passed to the individual output layers, and finally the presence or absence of artistry is output as 0 or 1. In addition, since some abstracted artistry is common and some are unique, the individual feature extraction layers exist in parallel with the common feature extraction layers as shown in Fig. 5.

Table 1: Criteria for judging artistic or non-artistic.

	artistic work	non-artistic
Pictures	Works of famous painters	Amateur work
Architecture	Works of famous architects	Private houses, apartments, buildings
Sculptures	Works of famous sculptors	Amateur work
Gardens	Famous garden	Private house garden

Figure 9: Examples of non-artistic paintings.

Figure 10: Examples of artistic painting.

Figure 11: Examples of non-artistic architecture.

Figure 12: Examples of artistic architecture.

Table 2: Accuracy of artistry identification.

	painting	architecture	sculpture	garden
single-task learning	0.6834	0.8353	0.7280	0.8825
multi-task learning	**0.7128**	**0.8518**	**0.7482**	**0.9295**

For training the proposed network, we use supervised learning with the ground truth of artistry of each work. Let y^t_c be the output obtained by inputting the image x of class c in task t to the multi-task network f as follows:

$$y^t_c = f(x) \tag{1}$$

When the ground truth of y^t_c in that image is $\hat{y}^t_c$, the loss function of each task can be defined by using the multi-class cross entropy as follows:

$$Loss^t = -\sum_c y^t_c \log \hat{y_c}^t \tag{2}$$

When we have T tasks in the multi-task learning, the multi-task learning can be realized by solving the minimization problem that minimizes the following multi-task loss:

$$Loss_{multi} = \sum_{t=1}^{T} \lambda_t Loss^t \tag{3}$$

where, λ_t $(t = 1, \cdots, T)$ are hyperparameters.

We next explain the detail of our network. In this research, we use the combination of Convolution, Batch Normalization, ReLU, and MaxPooling as a basis block of CNN (Fukushima and Miyake, 1982) as shown in Fig. 6, and the network is constructed by combining the basis blocks.

We want to evaluate the effectiveness of multi-task learning when there is a commonality in artistry among multiple art fields, so we build a network

Figure 13: Examples of non-artistic sculpture.

Figure 14: Examples of artistic sculpture.

Figure 15: Examples of non-artistic garden.

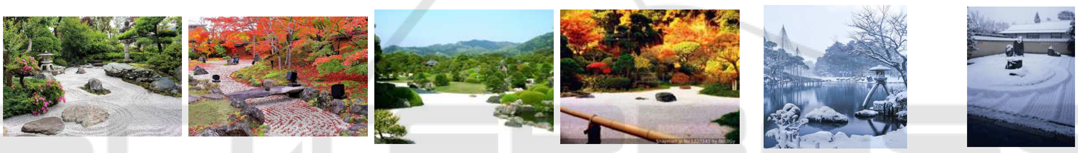


Figure 16: Examples of artistic garden.

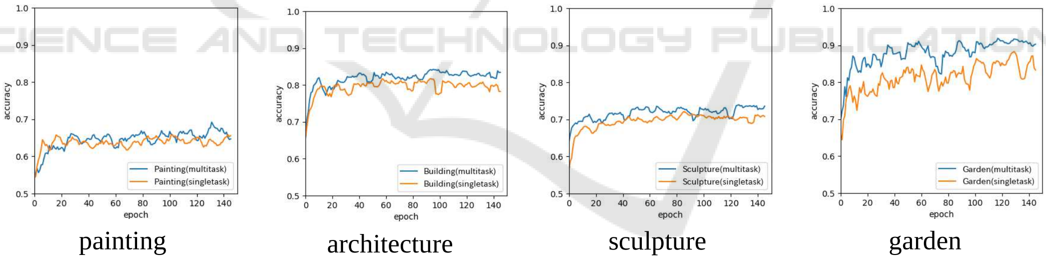


Figure 17: Changes in accuracy of artistry identification.

for multi-task learning and a network for single-task learning respectively for comparison.

The network structure for single-task learning is as shown in Fig. 7, and the network structure of multi-task learning is as shown in Fig. 8.

As shown in Fig. 8, the multi-task learning network has individual blocks for each art field and common blocks common to all art fields. In this network, the feature extraction is performed first for each art field as in the case of single-task learning, and then it is divided into individual blocks and a shared common blocks. The output of these blocks is input to the individual dense network as in the case of single-task learning.

5 DATASET

In this research, we use images from four art fields of painting, architecture, sculpture, and garden as artistic works. Since the ground truth value of artistry, i.e. artistic or non-artistic, is very difficult to fix, we set the ground truth value of artistry of each work whether it is made by famous artist or not (Sachant et al., 2016).

Table 1 shows the criteria for judging artistic or non-artistic in each art field. In the case of paintings and sculptures, those created by famous artists are considered as artistic works, and those created by amateurs such as students are considered as non-artistic.

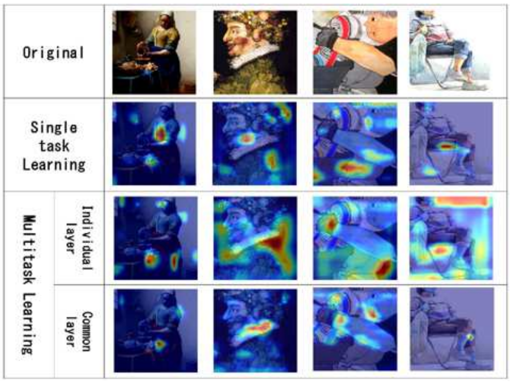


Figure 18: Painting

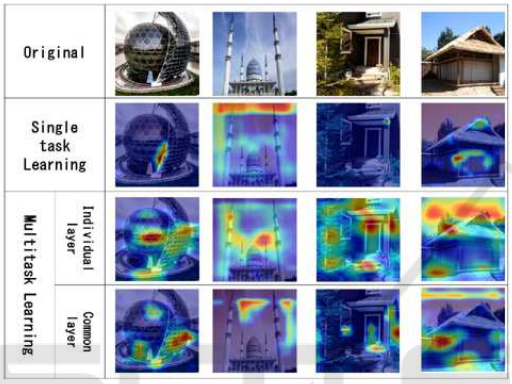


Figure 19: Architecture.

In the case of architectures, the works of famous architects are considered as artistic works, and ordinary private houses, apartments, and buildings are considered as non-artistic. In the case of gardens, the famous gardens are considered as artistic, and the gardens of ordinary private houses are considered as non-artistic.

The examples of artistic works and non-artistic works collected based on these criterion in painting, architecture, sculpture, and garden are shown in Figs. 9 to 16. We have collected 1000 images for artistic work and 1000 images for non-artistic work in each art field. We applied data augmentation to these images, and obtained 20000 images for artistic work and 20000 images for non-artistic work in each art field. Of these, 30000 were used for training and 10000 were used for testing in each art field.

6 EXPERIMENTS

We next show the experimental results of identifying artistry. In this experiment, both the multi-task learning network and the single-task learning network were trained with 150 epochs using images on artistic and non-artistic work shown in section 5, and their accuracy of artistry identification was compared using test images.

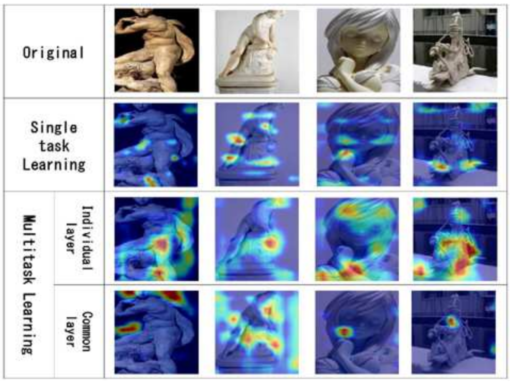


Figure 20: Sculpture.

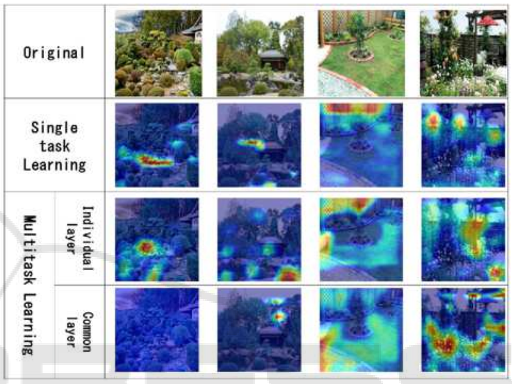


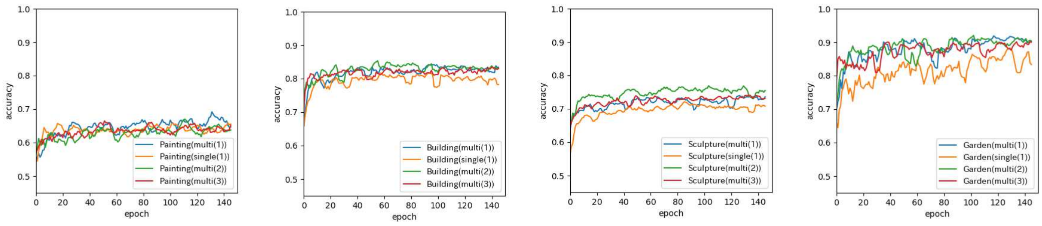
Figure 21: Garden.

Table 3: Number of parameters in each learning method.

	parameter
single-task learning (1)	9,700,722
multi-task learning (1)	26,699,634
multi-task learning (2)	9,184,626
multi-task learning (3)	4,694,898

6.1 Accuracy of Artistry Identification

Fig. 17 shows the changes in accuracy of artistry identification for the test data in multi-task learning and single-task learning respectively, and table 2 shows the results of the maximum identification rate for each field in each learning method. From these results, we find that the accuracy of artistry identification is higher in multi-task learning in all art fields. The results indicate that the artistry of these four art fields is related to each other. Especially in garden, the multi-task learning can improve the identification rate drastically compared to the single-task learning, indicating that it is more relevant to other fields than painting, architecture and sculpture.



painting architecture sculpture garden

Figure 22: Comparison of identification rate.

Table 4: Accuracy of artistry identification.

	painting	architecture	sculpture	garden
single-task learning (1)	0.6834	0.8353	0.7280	0.8825
multi-task learning (1)	**0.7128**	0.8518	0.7482	0.9295
multi-task learning (2)	0.6957	**0.8650**	**0.7820**	**0.9314**
multi-task learning (3)	0.7086	0.8482	0.7490	0.9167

6.2 Visualization of Attention in Artistry Identification

Since we were able to learn a classifier that discriminates the presence or absence of artistry, we next visualize the attention in the artistry identification task by using GradCAM (Selvaraju et al., 2017). In particular, we visualize which part of the work the artistry identifier focuses on to judge the presence or absence of artistry. We also compare whether there is a difference in the place of interest between single-task learning and multi-task learning.

In the artistry identification networks, the output of the final convolution layer is considered to represent the most important feature in identification, so the final convolution layer of block 5 is used for single-task learning, and the final individual layer and the final common layer are used for multi-task learning to visualize attention by GradCAM.

The output result of GradCAM is shown in Fig. 18 to Fig. 21. From these figures, we find that in paintings, the artistry is identified mainly by focusing on the person. In the case of architecture, the artistry identifier focuses on a single limited place in single-task learning, but in multi-task learning, the points of interest are distributed throughout. In the case of sculpture, it can be confirmed that the points of interest are sparsely distributed throughout both in single-task learning and multi-task learning. The range of attention is larger in multi-task learning as in the case of architectures. Finally, in the case of garden, we find that the artistry identifier was paying attention to plants and stones. From these results, it is considered that the multi-task learning pays more attention to the entire work than the single-task learning, and this property of multi-task learning makes it possible to judge the presence or absence of artistry more accurately.

6.3 Number of Parameters

We next evaluate the change in accuracy when the number of parameters is changed.

In the experiment in section 6.1, the number of parameters for multi-task learning was larger than the number of parameters for single-task learning. Therefore, we next evaluate the accuracy of artistry identification in the case where the number of parameters for multi-task learning is equal to or less than the number of parameters for single-task learning.

The number of parameters for each learning method is as shown in table 3. Single-task learning (1) and multi-task learning (1) are the networks used in the previous experiment. The multi-task learning (2) uses a network with the same number of parameters as single-task learning (1), and the multi-task learning (3) uses a network with fewer parameters than single-task learning (1). The accuracy of artistry identification by these networks are shown in table 4 and Fig. 22.

As shown in table 4 and the graph in Fig. 22, we find that the artistry identification rate is higher in multi-task learning (2) than in other cases and the number of parameters should not be too large in the multi-task learning.

7 CONCLUSION

In this paper, we proposed a method for analyzing the relevance of artistry among multiple art fields by using deep neural networks. In particular, we showed that by using the multi-task learning, the relevance of multiple art fields can be analyzed quantitatively. In our experimental results, we showed that the accuracy of artistry identification becomes higher in the multi-task learning. The results show that the artistry in different art fields is related to each other.

REFERENCES

Anwer, R. M., Khan, F. S., van de Weijer, J., and Laaksonen, J. (2016). Combining holistic and part-based deep representations for computational painting categorization. In *Proc. International Conference on Multimedia Retrieval*, pages 339–342.

Bianco, S., Mazzini, D., Napoletano, P., and Schettini, R. (2019). Multitask painting categorization by deep multibranch neural network. *Expert Systems with Applications*, 135:90–101.

Caruana, R. (1997). Multi-task learning. *Machine Learning*, 28:41–75.

Elgammal, A., Liu, B., Elhoseiny, M., and Mazzone, M. (2017). Can: Creatice adversarial networks generating ”art” by learning about styles and deviationg from style norms. *arXiv: 1706.07068*.

Fukushima, K. and Miyake, S. (1982). Neocognitron: A new algorithm for pattern recognition tolerant of deformations and shifts in position. *Pattern Recognition*, 15(6):455–469.

Goodfellow, I., Pouget-Abadie, J., Mirza, M., Xu, B., Warde-Farley, D., Ozair, S., Courville, A., and Bengio, Y. (2014). Generative adversarial nets. In *Advances in neural information processing systems*, pages 2672–2680.

Isola, P., Zhu, J.-Y., Zhou, T., and Efros, A. A. (2017). Image-to-image translation with conditional adversarial networks. In *The IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 1125–1134.

Kendall, A., Gal, Y., and Cipolla, R. (2018). Multi-task learning using uncertainty to weigh losses for scene geometry and semantics. In *Proc. IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*.

Liu, S., Johns, E., and Davison, A. J. (2019). End-to-end multi-task learning with attention. In *Proc. IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 1871–1880.

Picasso, P. Weeping woman.

Sachant, P. J., Blood, P., LeMieux, J., and Tekippe, R., editors (2016). *Introduction to Art:Design, Context, and Meaning*. University of North Georgia.

Selvaraju, R. R., Cogswell, M., Das, A., Vedantam, R., Parikh, D., and Batra, D. (2017). Grad-cam: Visual explanations from deep networks via gradient-based localization. In *ICCV*, pages 618–626.

Tan, W. R., Chan, C. S., Aguirre, H. E., and Tanaka, K. (2016). Ceci n'est pas une pipe: A deep convolutional network for fine-art paintings classification. In *Proc. International Conference on Image Processing*.

The Metaphysics Research Lab (1995). Stanford encyclopedia of philosophy.

van Gogh, V. W. Still life - vase with fifteen sunflowers.