CANVAS: Compositional Analysis of Visual Art Structure

What is composition in art?
- The arrangement of different visual elements in relation to one another.
- Controls human perception of aesthetics and meaning
- "Rule of thirds", "golden ratio", "symmetry"

How is composition currently studied?
- Formal visual analysis by artists and art historians.
- Patterns in composition reveal the beliefs and values behind artistic movements or eras.
- Works are studied & interpreted one by one.

A scalable framework enables large-scale analysis of composition, revealing patterns across centuries of art.

After two preliminary stages involving experimentation with exclusive classification, CANVAS transitioned to a different framework, where **each principle is analyzed independently**.

CANVAS focused on 3 principles. 1,000 Impressionist paintings were flagged as *present* or *absent* for each principle.

Steelyard Composition: One dominant mass balanced by a smaller, separate counter-mass.

Horizontal Thirds: 3 bands spanning the painting's width.

Radiating Lines: 3+ lines radiating from a central point.

(a) Steelyard Composition   (b) Horizontal Thirds   (c) Radiating Lines

Figure 1: *Examples paintings exhibiting each attribute.* Images from *Impressionist Landscapes Paintings* dataset by Rob Gonsalves, 2021, Kaggle (https://www.kaggle.com/datasets/robgonsalves/impressionistlandscapespaintings). CC BY-SA 4.0.

These definitions enable models to detect compositional features without being constrained by mutually exclusive categories, providing the basis for a scalable framework.

- For the initial CNNs, 1,000 impressionist landscape paintings are manually labeled as "present" or "absent" for all three attributes.
- To test for the presence of Steelyard Composition using object detection, a larger dataset of 3,000 paintings are manually labeled with either "present" or "absent".

Convolutional Neural Networks
- CANVAS trains an individual CNN for each binary attribute using 150×150 RGB images as input.
- Each network uses two convolutional layers with ReLU and max-pooling, followed by a fully connected layer and a sigmoid output to predict attribute presence.
- Models are optimized with Adam and binary cross-entropy loss, with class weighting to handle dataset imbalance.

Bryant, D. (2024). Machine Learning Image Classifier for Identifying Composition in Landscape Paintings.

Lee, B., Seo, M. K., Kim, D., Shin, I. S., Schich, M., Jeong, H., & Han, S. K. (2020). Dissecting landscape art history with information theory. *Proceedings of the National Academy of Sciences*, 117(43), 26580-26590.

Lee, J. T., Kim, H. U., Lee, C., & Kim, C. S. (2018). Photographic composition classification and dominant geometric element detection for outdoor scenes. *Journal of Visual Communication and Image Representation*, 55, 91-105.

Liu, D., Puri, R., Kamath, N., & Bhattacharya, S. (2019). Modeling image composition for visual aesthetic assessment. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops* (pp. 0-0).

Liu, S., Zeng, Z., Ren, T., Li, F., Zhang, H., Yang, J., ... & Zhang, L. (2024, September). Grounding dino: Marrying dino with grounded pre-training for open-set object detection. In *European conference on computer vision* (pp. 38-55). Cham: Springer Nature Switzerland.

Madhu, P., Marquart, T., Kosti, R., Bell, P., Maier, A., & Christlein, V. (2020, August). Understanding compositional structures in art historical images using pose and gaze priors: Towards scene understanding in digital art history. In *European Conference on Computer Vision* (pp. 109-125). Cham: Springer International Publishing.

Payne, E. A. (1957). *Composition of outdoor painting.* EP Payne.

Vandor, A., Van Vollenhoven, M., Weiss, G., & Spanakis, G. (2020). Can we detect harmony in artistic compositions? A machine learning approach. *arXiv preprint arXiv:2012.05633.*

Flagging Steelyard using Object Detection

To replicate how humans perceive and recognize different visual structures in art, CANVAS uses Grounding DINO, a vision–language open-set object detector, to generate bounding boxes around semantic objects in the painting.

(a) With Steelyard   (b) Without Steelyard

Figure 2: *Examples of object detections.* Images from *Impressionist Landscapes Paintings* dataset by Rob Gonsalves, 2021, Kaggle (https://www.kaggle.com/datasets/robgonsalves/impressionistlandscapespaintings). CC BY-SA 4.0.

5 features are calculated (object area ratios, center of visual mass, top-bottom and left-right symmetry, dominant object position, spatial spread).

These features are then fed into a neural network classifier with three fully connected layers (256, 128, and 64 neurons) followed by batch normalization and dropout to predict the probability that a painting exhibits Steelyard Composition.

Results without Object Detection

Label / Model Prediction	Steelyard	No Steelyard
Steelyard	29.5%	25.0%
No Steelyard	22.0%	23.5%

Table 1: Steelyard Prediction w/o Object Detection

Label / Model Detection	Hor. Thirds	No Hor. Thirds
Hor. Thirds	28.5%	23.0%
No Hor. Thirds	23.0%	25.5%

Table 2: Horizontal Thirds Prediction w/o Object Detection

Label / Model Prediction	Radiating Lines	No Radiating Lines
Radiating Lines	32%	35.5%
No Radiating Lines	19.5%	13.0%

Table 3: Radiating Lines Prediction w/o Object Detection

Results with Object Detection

Label / Model Prediction	Steelyard	No Steelyard
Steelyard	59.0%	18.0%
No Steelyard	9.0%	14.0%

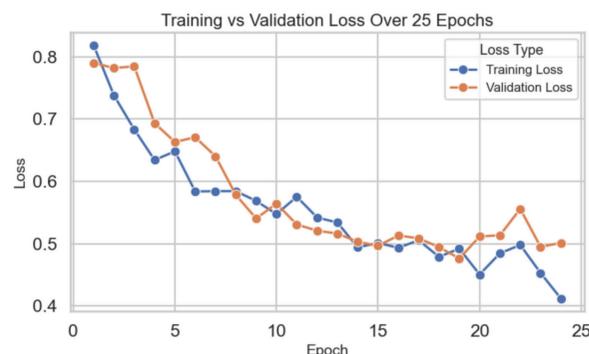Table 4: Steelyard Prediction Using Object Detection

Figure 3: *Training and Validation Loss over 25 Epochs.* Graph created by the student researcher using Seaborn, 2025.
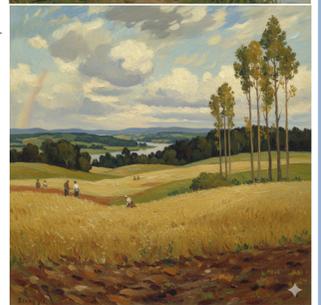

Figure 4: *Paintings generated using 9 inputted images strongly exhibiting each attribute.* Images created in November 2025 using Gemini 3.

- Composition is defined by the arrangement and **relationships between meaningful elements**, rather than color or contrast alone.
- CANVAS provides a scalable, quantitative framework for studying composition.
- CANVAS **interrogates the extent to which machines can understand and interpret abstract concepts like art on a human level.**

Ella Lu