



Learning from Small Data: Adapting Pretrained Diffusion Models for 17th Century Painting

Katarina Mohar, France Stele Institute of Art History, ZRC SAZU

AI and Artworks: Object Detection, Image Classification and Iconographic Analysis

Vienna, 16th April 2026

Joannes Almenak (c. 1640 – after 1684)

All images: © National Gallery of Slovenia

Preserved paintings:



The Card Players I



The Peddler

Copies after lost originals:



Shepherd and Shepherdess



The Card Players II



Boy with a Turkey



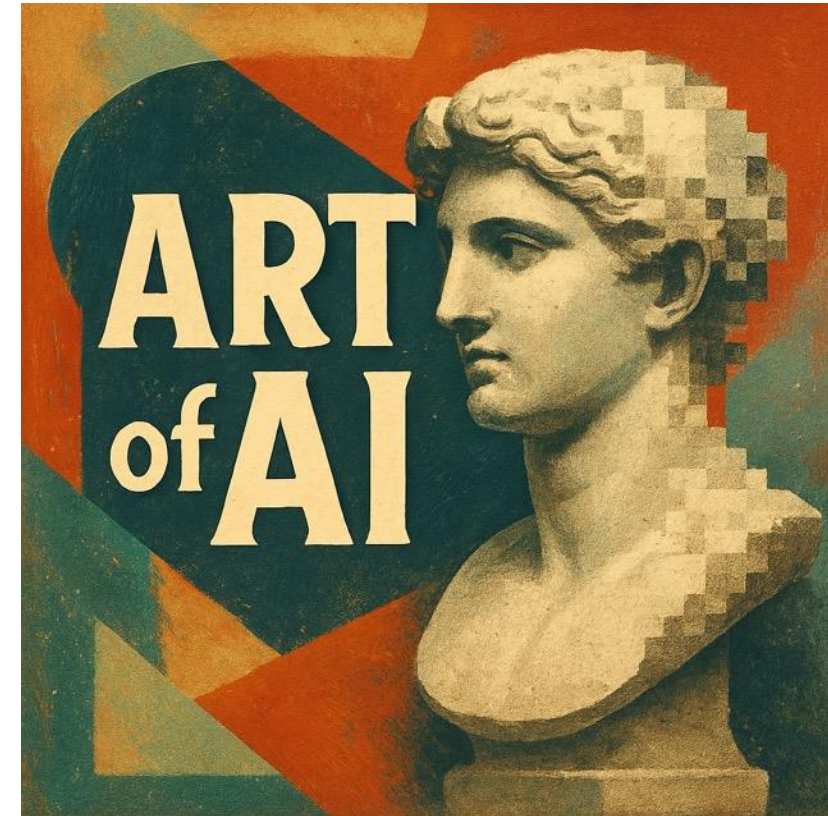
A Peasant Family

ARTofAI Project

Tools of the Future for Research of the Past. The Use of Generative Image Models for Humanities

@ ZRC SAZU, France Stele Institute for Art History, 2023–2026

- Interdisciplinary team: art historians, computer scientist, philosopher
- Main question: Can diffusion models, fine-tuned on very small datasets, support the study of damaged, fragmentary, or lost artworks?
- Approach: AI as a tool for controlled speculation, not a replacement for expertise
- Focus: style learning, recurrent visual patterns, speculative visual hypotheses



Training Data: Images

- High-quality, rights-cleared images
- Segmented into fragments: faces, hands, objects, details.
- Corpus: 4 preserved paintings → 99 fragments



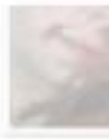
29.png
1,419 × 1,414



30.png
1,705 × 1,705



31.png
2,520 × 2,520



32.png
966 ×



35.png
696 × 696



36.png
4,336 × 4,336



37.png
1,243 × 1,243



38.png
774 × 774



41.png



42.png



43.png



44.png



Training Data: Captions

- We tested four caption types:
 - Simple – short, human-written labels
 - Expert – art-historical descriptions
 - VLM – AI-generated sentences
 - CLIP Tags – AI-generated keywords)
- Initial hypothesis: richer captions might affect what the model learns



Almenak: *Boy with a Turkey*, fragment/detail © NG SI

Simple: “turkey head eyes neck”

Expert: “A turkey, 17th century, genre painting, highly realistic, low-life genre, resembling Giacomo Ceruti, resembling Michiel Sweerts”

VLM: “The image is a close-up of a turkey's head and neck ... red feathers, which appear to be roosters ... realistic style ...”

CLIP: “a painting of a turkey ... Rembrandt ... anthropomorphized chicken ... the bird is wearing a bowtie ... walter white, 18th century ...”

LoRA vs. DreamBooth: Expert Evaluation

- 7 expert evaluators, paired comparison study
- LoRA preferred over DreamBooth in:
 - 73.8% of coherence judgments
 - 69.0% of composition judgments
 - 66.7% of style judgments

LoRA



DreamBooth



LoRA: Visible Stylistic Shift

- Outputs often shifted away from the base model
- Shared tendencies across caption types
- Closer to Almenak's palette, light, and surface treatment
- Caption type had little visible effect on style



Almenak: *The Card Players I*

Base model

Simple

VLM

Expert

CLIP





Base model



CLIP Captions



Simple Captions



Prompt: A group of men playing cards in a smoky tavern interior, candlelit and dim, in a Flemish Baroque genre painting with visible brushwork, chiaroscuro lighting, and earthy palette

Captioning: Weak Effect on Style, Stronger on Content

- No major stylistic difference between caption sets
- Differences appeared more clearly in semantic behaviour
- Simple / VLM often showed more literal reuse of training fragments
- Expert / CLIP sometimes allowed more flexible recombination
- Captioning seemed to shape content more than style



Almenak: *The Peddler*



Expert Captions



Prompt: A woman in a 20th Century kitchen

Simple Captions



Training Data



Almenak: *The Card Players I* (detail)

Expert Captions



Simple Captions



Training Data



Almenak: *The Peddler*
(detail)

Prompt: An elderly peddler showing trinkets to a peasant woman outside a cottage, depicted in the style of a 17th-century oil painting, with expressive gestures, ochre tones, and soft painterly edges

What the Experiment Clarified

- Even very small corpora can shift a pretrained diffusion model
- LoRA worked better than DreamBooth
- Captioning shaped semantic orientation more than style
- Composition remained the major unresolved issue
- With small datasets, structure matters as much as quantity

The Dataset

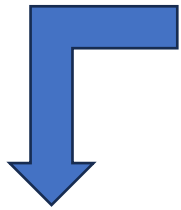
Image Corpus:

- 99 fragments cropped from 4 preserved works



Captions created in 4 distinct styles:

- Simple (Human)
- Expert (Art Historian)
- VLM (AI-generated full sentence descriptions)
- CLIP Tags (AI-generated keywords/textual tags)



Simple: turkey, head, eyes, neck

Expert: A turkey, 17th century, genre painting, highly realistic, low-life genre, resembling Giacomo Ceruti, resembling Michiel Sweerts

VLM: The image is a close-up of a turkey's head and neck. The turkey is facing towards the left side of the image, with its head turned slightly to the right. The head of the turkey is covered in red feathers, which appear to be roosters. The feathers are arranged in a rooster-like pattern, with some overlapping each other. The beak is slightly open, and the eyes are looking directly at the viewer. The background is a dark blue, and there is a white curtain visible in the background. The painting is done in a realistic style, with loose brushstrokes and vibrant colors.

CLIP: a painting of a turkey, high detail baroque oil painting, detailed baroque oil painting, (rembrandt), anthropomorphized chicken, the bird is wearing a bowtie, giuseppe arcimboldo walter white, 18th century oil painting, 18th century oil painting, coxcomb, with the beak of an eagle, flemish painting



Simple: boy turkey sitting face gestures background

Expert: Painting, oil on canvas, depicting a boy of approximately 15 years of age, sitting in a dark courtyard and feeding a turkey. A boy has a sad and numb expression on his face. Realistic portrayal of working boy, and a portrayal of poverty. In the back of the painting two hens are depicted. In the style of Italian genre painting. Inspired by Flemish and Dutch Caravaggisti. Genre scene from the second half of the 17th century, low life genre. Brown and earthy colour palette.

CLIP: a painting of a man holding a chicken, inspired by David Teniers the Younger, inspired by David Teniers the Elder, inspired by David Teniers III, inspired by Pieter Lastman, inspired by Pieter Claesz, inspired by Jan Steen, dutch golden age painting, inspired by Friedrich von Amerling, by David Teniers III, inspired by Willem van Aelst

VLM: The image is a painting of a young man sitting on a chair with a rooster in his lap. The man is wearing a black jacket, white pants, and black boots. He has a serious expression on his face and is looking directly at the viewer. The rooster is a light brown color with a red comb and is resting its head on the armrest of the chair. In front of the man, there is a wooden bowl filled with black pebbles. Behind the man is a brick wall with a window and a wooden bench. On the left side of the image, there are two chickens sitting on the bench. The painting is done in a realistic style with loose brushstrokes and vibrant colors.