ArtHist.net

Can Large Language Models Speak Art? A Call for Scholarly Collaboration

Gianmarco Spinaci

Recent advances in generative multimodal Large Language Models (LLMs) have attracted significant attention; however, their actual capabilities in the Art History domain remain underexplored. While these systems process both images and texts, there is a lack of systematic, empirical evidence regarding their knowledge of art-historical subjects, visual interpretation, and contextual reasoning. As part of my PhD research, this project aims to address this gap by investigating the current capabilities and limitations of multimodal LLMs in Art History.

The primary objective of this research is to examine how general-purpose multimodal LLMs respond to art historical questions, with particular focus on bias, hallucination, and domain-specific knowledge. By evaluating model behavior across diverse topics, the project aims to identify areas of success and failure and to assess the implications for their application in scholarly research.

To support this investigation, I invite members of the ArtHist.net community—including art historians, researchers, and educators—to contribute multiple-choice questions, ideally accompanied by images. These questions should assess art historical knowledge, visual analysis, iconography, attribution, chronology, or contextual understanding. All contributions will be fully attributed, and the collected material will constitute an open, scholarly dataset for evaluating multimodal LLMs and critically examining claims of artificial general intelligence within Art History.

If you are interested in contributing questions or discussing this project further, I would welcome your response.

Gianmarco Spinaci (gianmarco.spinaci2 [at] unibo.it)
PhD Student @ Cultural Heritage in the Digital Ecosystem, University of Bologna
Digital Humanities Research Associate @ Villa I Tatti, The Harvard University Center for Italian Renaissance Studies