

# Art Report Generation and Evaluation Using LLM and Computer Vision for Visually Impaired Individuals

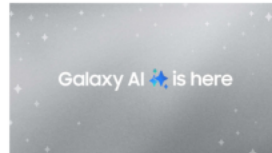
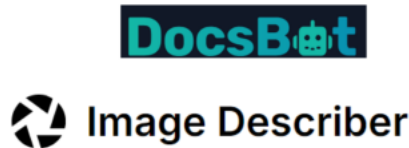
Gonzalez y Gonzalez, Sarai  
saraiglez@yglez@gmail.com - BME, Hungary



## Introduction

Individuals with visual impairments often face significant barriers in accessing and appreciating visual art. While some AI tools exist to generate image descriptions, they frequently fall short in conveying the nuanced details and emotional depth inherent in artworks. This limitation hinders the ability of visually impaired individuals to fully experience and interpret art, leading to a gap in cultural accessibility.

Some AI tools to generate image descriptions:



## Objective

Develop an AI system that analyzes images of cultural artifacts or historical sites paintings to generate descriptive reports, making them accessible to individuals with visual impairments.

## Research questions:

1. How can AI-generated descriptive reports of cultural artifacts improve accessibility for individuals with visual impairments?
2. Would it be beneficial or not for visually impaired people at the end?

## Solution

Develop an advanced AI-driven system specifically designed to analyze art images and generate rich, descriptive narratives that capture both the technical and emotional aspects of artworks. This system would integrate computer vision techniques to assess visual elements such as color, texture, and composition, with Large Language Models (LLMs) to produce detailed and contextually relevant descriptions. By focusing on the unique characteristics of art, the system aims to provide visually impaired users with a more immersive and informative experience.



Fig. 1 Collection of Mr. and Mrs. Paul Mellon, "Flower Beds in Holland, c. 1883, Vincent Van Gogh" Dutch Paintings of the Nineteenth Century, NGA Online Editions.

## Methodology

- Using WikiArt dataset try to use CNN to excel in identifying intricate visual features, patterns, and styles within paintings.
- Integrate a LLM to generate coherent and contextually rich narratives.
- User-Centric Design and Testing with visually impaired individuals to understand their needs and preferences.
- Conduct usability testing to gather feedback on the clarity, usefulness, and emotional resonance of the generated descriptions.

## Expected Outcome


 <p>GT: Pointillism</p>	 <p>the art style of the painting is...</p> <ul style="list-style-type: none"><li>Pointillism (48.53%)</li><li>Fauvism (37.45%)</li><li>Color Field Painting (3.52%)</li><li>Post Impressionism (2.54%)</li><li>Impressionism (2.16%)</li><li>Synthetic Cubism (1.79%)</li><li>Art Nouveau Modern (1.06%)</li><li>Naive Art Primitivism (0.83%)</li><li>Analytical Cubism (0.5%)</li><li>Romanticism (0.46%)</li><li>Contemporary Realism (0.23%)</li></ul>	 <p>USER: To which art style does this painting belong?</p> <p>ASSISTANT: This painting belongs to the Pointillist style, which is characterized by the use of small, distinct dots or brushstrokes to create a visual effect. The painting features a beach scene with sailboats, people, and trees, all depicted in a vibrant and detailed manner. The Pointillist style was popular in the late 19th and early 20th centuries, and it is evident in this painting that the artist has used this technique to create a lively and engaging scene.</p>	 <p>The art style of this painting is Cubism.</p> <p>The art style of this painting is Realism.</p> <p>The art style of this painting is Impressionism.</p>	 <p>To which art style does this painting belong?</p> <p>The painting depicted belongs to the art style known as Pointillism. Pointillism is a technique of painting in which small, distinct dots of color are applied in patterns to form an image. It was developed by Georges Seurat and Paul Signac in the late 19th century. This approach to painting uses the science of optics to create vibrant, luminous images.</p>
---	--	--	---	--

Fig. 2 Strafforello et al. (2024) Example of art style classification with VLMs on the paintings.



Prompt: Please write a paragraph of formal analysis for this painting.

LLaVA-1.5-7B

The image features a woman sitting on a chair, wearing a white dress and holding a needle in her hand. She appears to be engaged in a sewing activity, possibly mending her clothing. The scene is set in a garden, with a potted plant located near the woman. The woman's attire, including the white dress, and the garden setting create a sense of tranquility and simplicity. The focus on the woman's hands and the needle she holds emphasizes her concentration on her task, while the potted plant adds a touch of nature to the scene. Overall, the image captures a moment of quiet contemplation and creativity in a serene outdoor setting.

ShareGPT4V-7B

The image is a captivating impressionist painting that captures a moment of tranquility in a garden. The central figure is a young woman, elegantly dressed in a white dress, seated on a bench. She is engrossed in her hands, perhaps lost in thought or simply enjoying the serenity of her surroundings. The garden around her is a riot of colors, with red flowers adding a vibrant contrast to the lush greenery. The painting is executed in an impressionist style, characterized by loose brushstrokes that capture the interplay of light and color. The artist's use of color and light brings a sense of depth and realism to the scene, while the focus on the woman's contemplative state adds a layer of emotional depth to the composition. The overall effect is a harmonious blend of color, light, and emotion, a hallmark of the impressionist genre.

GPT-4V

This painting depicts a young woman absorbed in her needlework, seated amidst a lush, verdant garden setting. The artist has employed a vibrant color palette, with the warm tones of the woman's dress and the surrounding foliage creating a harmonious and inviting atmosphere. The woman's concentration is evident in her downcast gaze and the delicate movements of her hands, suggesting a sense of tranquility and introspection. The interplay of light and shadow across the scene further enhances the painting's sense of depth and atmosphere, drawing the viewer into the intimate, contemplative moment being captured. Overall, the work conveys a serene, introspective mood, highlighting the beauty and quiet simplicity of the everyday.

GalleryGPT (Ours)

This exquisite piece of artistry is a testament to the mastery of capturing human emotion and intimate moments. The visual characteristics of the painting are striking, with the use of soft, muted colors that create a sense of warmth and tranquility. The central figure, a young woman, is depicted in a domestic setting, engrossed in the task of sewing. The artist's use of light and shadow is particularly noteworthy, with the light source seemingly emanating from the left, casting gentle shadows that add depth and dimension to the scene. The woman's facial expression and body language convey a sense of concentration and introspection, while her surroundings, including the ornate chair and the patterned fabric, add a layer of visual interest. The artist's brushwork is delicate and precise, particularly in the rendering of the woman's hands and face, which are imbued with a sense of realism and detail. The overall composition is balanced and harmonious, with the figure and her surroundings occupying the majority of the canvas, leaving little space for distraction. This, coupled with the artist's use of perspective, draws the viewer's attention towards the figure, making her the focal point of the painting.

Fig. 3 Bin et al. (2024) An example for qualitative comparison of formal analysis generation by several powerful LLMs.

## References

- Strafforello et al. (2024) Have Large Vision-Language Models Mastered Art History? <https://doi.org/10.48550/arXiv.2409.03521>
- Bin, et al. (2024) GalleryGPT: Analyzing Paintings with Large Multimodal Models. <https://doi.org/10.48550/arXiv.2408.00491>
- Liang et al. (2024) A Comprehensive Survey and Guide to Multimodal Large Language Models in Vision-Language Tasks. <https://doi.org/10.48550/arXiv.2411.06284>
- Stork, D. G. (2024) Computer Vision, ML, and AI in the Study of Fine Art. Computer Vision, ML, and AI in the Study of Fine Art – Communications of the ACM. [https://cacm.acm.org/research/computer-vision-ml-and-ai-in-the-study-of-fine-art/?utm\\_source=chatgpt.com](https://cacm.acm.org/research/computer-vision-ml-and-ai-in-the-study-of-fine-art/?utm_source=chatgpt.com)
- Acknowledgements. The HCAIM (the Human-Centred AI Master's Programme) Project is Co-Financed by the Connecting Europe Facility of the European Union Under Grant N° CEF-TC-2020-1 Digital Skills 2020- U-IA-0068. This Poster Was Created As Part of the Blended IP Event Organized Under Erasmus + Programme of the European Union.