

L'image augmentée : entre histoire de l'art et computer vision

Databases : between image analysis and generation

MARIA GIULIA DONDERO, ADRIEN DELIEGE ET ENZO D'ARMENIO

MAY 23, 2024

L'image augmentée : entre histoire de l'art et computer vision

Databases : between image analysis and generation

Experiments : augmenting Susanna and the Elders

ADRIEN DELIEGE

MAY 23, 2024

L'image augmentée : entre histoire de l'art et computer vision

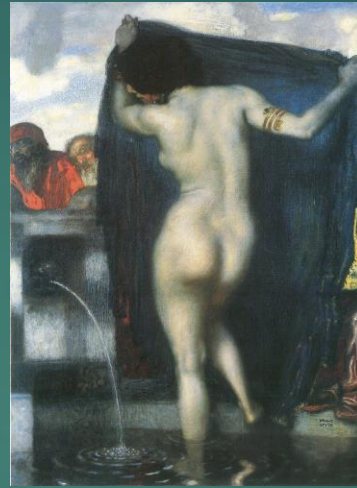
Databases : between image analysis and generation

Experiments : augmenting Susanna and the Elders

ADRIEN DELIEGE

MAY 23, 2024

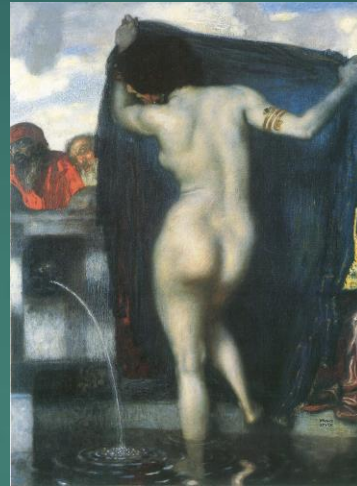
Susanna and the Elders



What comes here ?

+ dozens more!
A whole database of Susannas...

Susanna and the Elders



Not augmenting a given image with extra analyses

Not augmenting with pixels around an image

**Augmenting the theme
*Susanna and the Elders***

+ dozens more!
A whole database of Susannas...

Using DALL·E 3 (OpenAI)



- To add more details
- For safety reasons, to prevent:
 - violent, adult content
 - biases, hurting ethnicities
 - hurting public figures
 - using style of living artist

Exp. 1: General culture of the machine



“A painting of Susanna and the Elders”

“A classical painting showcasing an old religious story from the Bible. Depict a scene with three individuals - a young Middle-Eastern woman named Susanna, dressed in beautiful, richly-colored fabrics, and two elderly men, one of them of Caucasian descent and the other of South Asian descent, looking intensely at her. The scene takes place in a lush garden, with hints of Greek architecture in the surroundings.”

Exp. 1: General culture of the machine

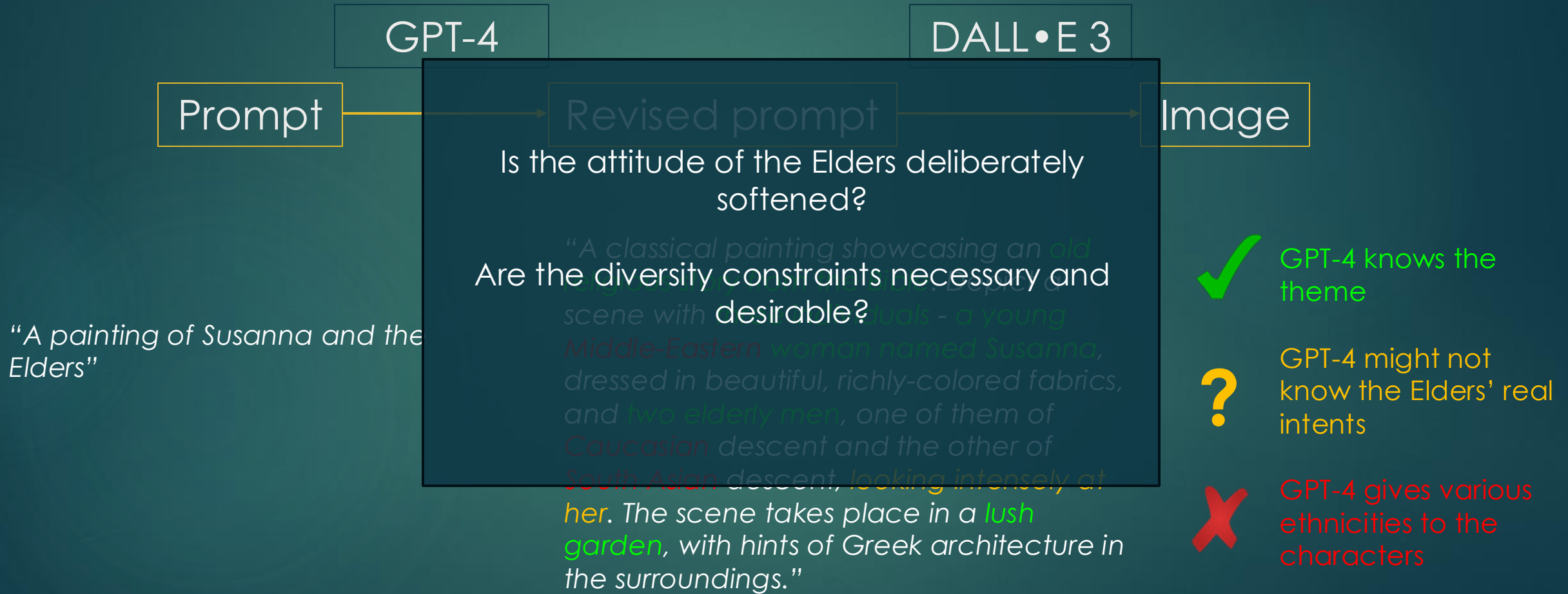


“A painting of Susanna and the Elders”

“A classical painting showcasing an *old religious story from the Bible*. Depict a scene with *three individuals - a young Middle-Eastern woman named Susanna*, dressed in beautiful, richly-colored fabrics, and *two elderly men*, one of them of *Caucasian* descent and the other of *South Asian* descent, *looking intensely at her*. The scene takes place in a *lush garden*, with hints of Greek architecture in the surroundings.”

- ✓ GPT-4 knows the theme
- ? GPT-4 might not know the Elders' real intents
- ✗ GPT-4 gives various ethnicities to the characters

Exp. 1: General culture of the machine



Exp. 1: General culture of the machine



“A painting of Susanna and the Elders”

“A classical painting showcasing an *old religious story from the Bible*. Depict a scene with *three individuals - a young Middle-Eastern woman named Susanna*, dressed in beautiful, richly-colored fabrics, and *two elderly men*, one of them of *Caucasian* descent and the other of *South Asian* descent, *looking intensely at her*. The scene takes place in a *lush garden*, with hints of Greek architecture in the surroundings.”



Exp. 1: General culture of the machine

*“A classical painting showcasing an old religious story from the Bible. Depict a scene with **three** individuals - a young **Middle-Eastern** woman named Susanna, dressed in beautiful, richly-colored fabrics, and two elderly men, one of them of Caucasian descent and the other of **South Asian** descent, looking intensely at her. The scene takes place in a lush garden, with hints of Greek architecture in the surroundings.”*

Difficulties to represent

- A given number of instances
- Ethnicities
- Hands





Exp. 1: General culture of the machine

OK to some extent, but can we really qualify these images as augmentations of *Susanna and the Elders*?

Exp. 2: Wikipedia's prompt (Tintoretto)



What really matters to augment a theme?

We want:

- More control on the composition
- More **focus on important elements**, in particular the intrusive behavior of the elders



Exp. 2: Wikipedia's prompt (Tintoretto)



~~"The painting shows a naked Susanna~~
A painting showing a woman sitting in a garden beside a pool, leaning against a tree, and facing a dense rose trellis, against which leans a mirror. ~~Susanna~~ She gazes into the mirror. Around her on the ground are several bright objects including a white silk shawl, pearls, jewelry, a comb, and silverware, which together form a kind of still life. Two elderly men peer around either end of the rose trellis to watch her. ~~Susanna~~ She is sitting absorbed by her reflection in the mirror and unaware of the intruders."

"A *detailed* painting of a **white, middle-aged** woman sitting by a *calm* pool in an *enchanted* garden. She is leaning against a *robust* tree while fixing her gaze into a mirror that rests against a *lush* rose trellis. In the foreground, there lies an assortment of *vibrant* objects such as a silken *white* shawl, *gleaming* pearls, *shiny* jewelry, a *finely carved* comb, and *clean* silverware, forming an *aesthetically pleasing* still life. Two elderly men of **Middle-Eastern** and **Hispanic** ethnicity lurk around either end of the rose trellis, spying on the woman who remains engrossed in her mirror's reflection."

Added ethnicities and physical characteristics

Beautifying adjectives

Exp. 2: Wikipedia's prompt (Tintoretto)



~~"The painting shows a naked Susanna. A painting showing a woman sitting in a garden beside a pool, leaning against a tree, and facing a dense rose trellis, against which leans a mirror. ~~Susanna~~ She gazes into the mirror. Around her on the ground are several bright objects including a white silk shawl, pearls, jewelry, a comb, and silverware, which together form a kind of still life. Two elderly men peer around either end of the rose trellis to watch her. ~~Susanna~~ She is sitting absorbed by her reflection in the mirror and unaware of the intruders."~~

"A *detailed* painting of a **white, middle-aged** woman sitting by a **calm** pool in an **enchanted** garden. She is leaning against a **robust** tree while fixing her gaze into a mirror that rests against a **lush** rose trellis. In the foreground, there lies an assortment of **vibrant** objects such as a silken **white** shawl, **gleaming** pearls, **shiny** jewelry, a **finely carved** comb, and **clean** silverware, forming an **aesthetically pleasing** still life. Two elderly men of **Middle-Eastern** and **Hispanic** ethnicity lurk around either end of the rose trellis, spying on the woman who remains engrossed in her mirror's reflection."



Exp. 2: Wikipedia's prompt (Tintoretto)

"A detailed painting of a white, middle-aged woman sitting by a calm pool in an enchanting garden. She is **leaning against a robust tree** while fixing her gaze into a mirror that rests against a lush rose trellis. In the foreground, there lies an assortment of vibrant objects such as a silken white shawl, gleaming pearls, shiny jewelry, a finely carved comb, and clean silverware, forming an aesthetically pleasing still life. **Two elderly men of Middle-Eastern and Hispanic ethnicity** lurk around **either end** of the rose trellis, spying on the woman who remains engrossed in her mirror's reflection."

Difficulties to represent

- A given number of instances
- Ethnicities
- Hands
- Some compositional aspects



Exp. 2: Wikipedia's prompt (Tintoretto)



OK to some extent again.
Are these better
augmentations?



Simple memorisation +
adaptation or real creation
capabilities?

Exp. 3: Compositional variations: genders, ages

Susanna → man, Elders → old women



Susanna → old woman, Elders → young men



Exp. 3: Compositional variations: ambiance, place

Atmosphere of death



At the beach



Conclusions

Augmenting *Susanna and the Elders* is somehow doable with tools like DALL•E 3.

However, the images created may sometimes barely be considered as faithful to the theme.

The real difficulty lies within the generation of the intents of the elders. This difficulty is for now technology-related.

Other difficulties may arise, but are more aesthetic than fundamental.

