

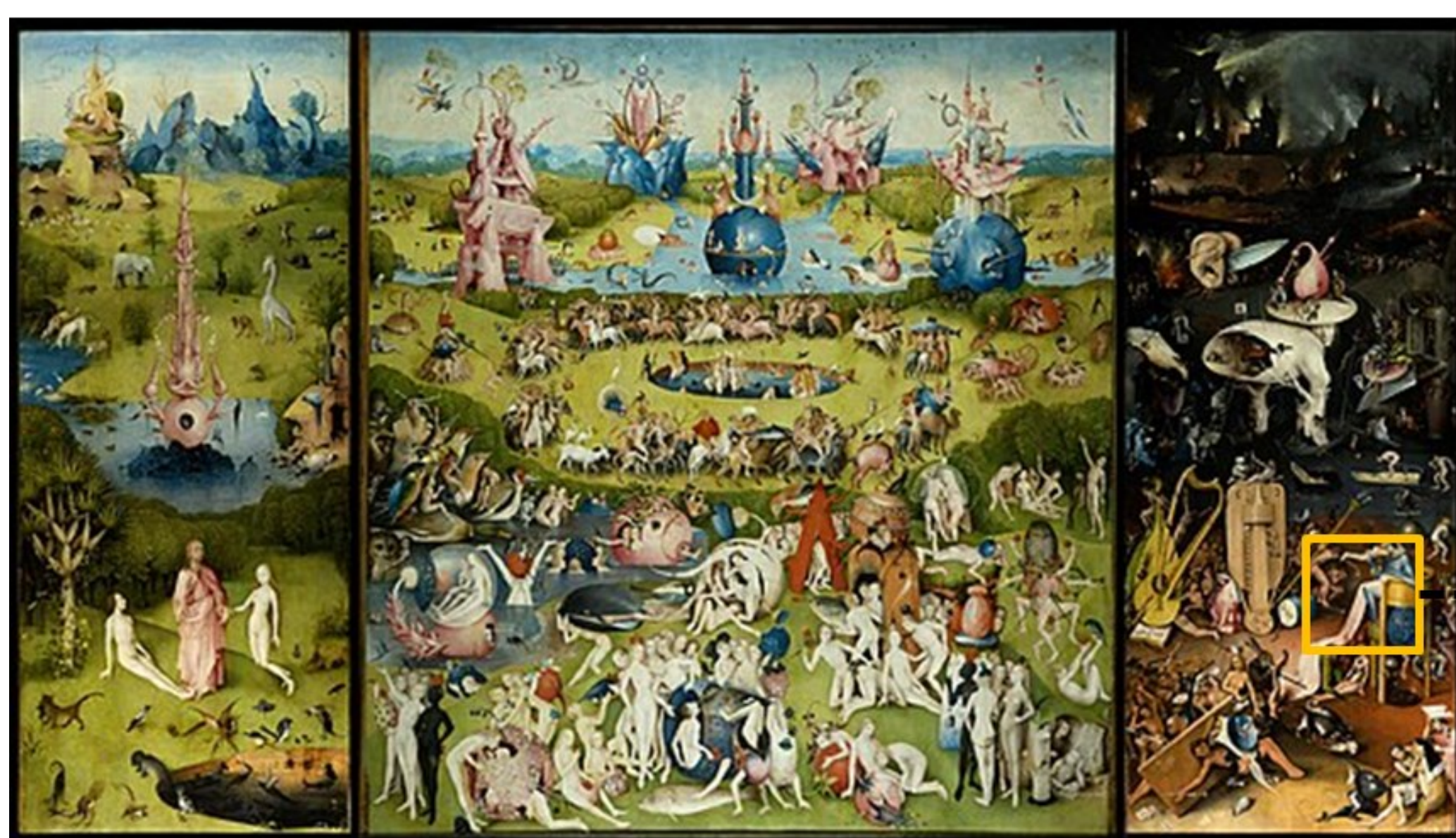
Captioning Bosch: A Twitter Bot

Cornelia Ferner

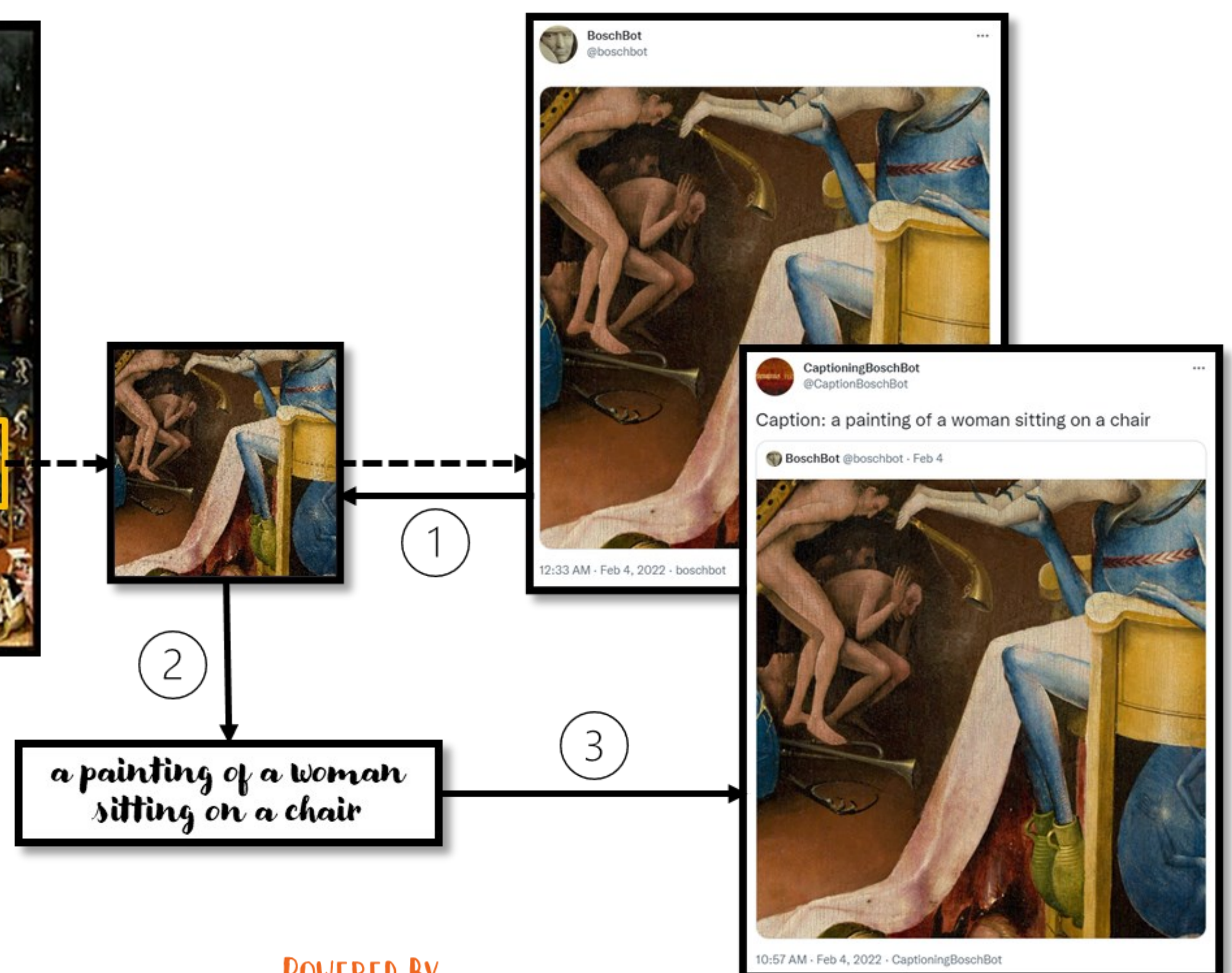
Salzburg University of Applied Sciences

The artworks by Dutch painter Hieronymus Bosch are well known for their incredible wealth of details. A popular Twitter bot regularly posts small segments of the digitized paintings on Twitter, thus relieving their density and making them more accessible. CaptioningBoschBot, the Twitter bot presented in this demo, reverses the creative process of the artist: It uses the out-of-context painting segments as input for an encoder-

decoder model to generate captions that interpret the painted objects. As the model was only trained on realistic, photographic images, curious interpretations of the otherworldly details can be observed. The generated captions are again posted on Twitter to encourage discussions about Bosch's masterpieces and the AI technology in general.



The Garden of Earthly Delights, Hieronymus Bosch, 1490-1510



METHOD

① Accessing Tweets

Retrieve Tweets from BoschBot and access posts containing media to extract the URL of the image segment. Load and preprocess the image.

② Generating Captions

A Vision Transformer as encoder and a GPT-2 network as decoder are used to generate a caption for the image segment.

POWERED BY
Hugging Face

③ Posting Retweets

The original post is retweeted by CaptionBoschBot, combining both the image and the generated caption.

CHALLENGES

During training, the model has never seen art before. While the Vision Transformer was pretrained on ImageNet, the image captioning pipeline was trained on the MS COCO captions dataset containing photorealistic images.

There are three main challenges that demonstrate the weaknesses of a model unaware of art:

- Ⓐ Fixation on cracking or overemphasis
- Ⓑ Lost focus or high-level descriptions
- Ⓒ Comprehension gap or lack of context

Ⓐ

A PAINTING OF A
WOMAN AND A MAN
IN A FIELD.

Ⓑ

A GROUP OF BIRDS
SITTING ON TOP OF
EACH OTHER.

Ⓒ

A WOMAN HOLDING A
SURFBOARD.

CONCLUSION

CaptioningBoschBot illustrates that creativity is not inherent in an AI model per default. It lacks the means to recognize art and the vocabulary to describe it.

Without retraining the network on images of artwork and their descriptions, it is impossible to generate anything else than objective descriptions.

This, however leads to curious interpretations and challenges our way of perceiving the original painting.

