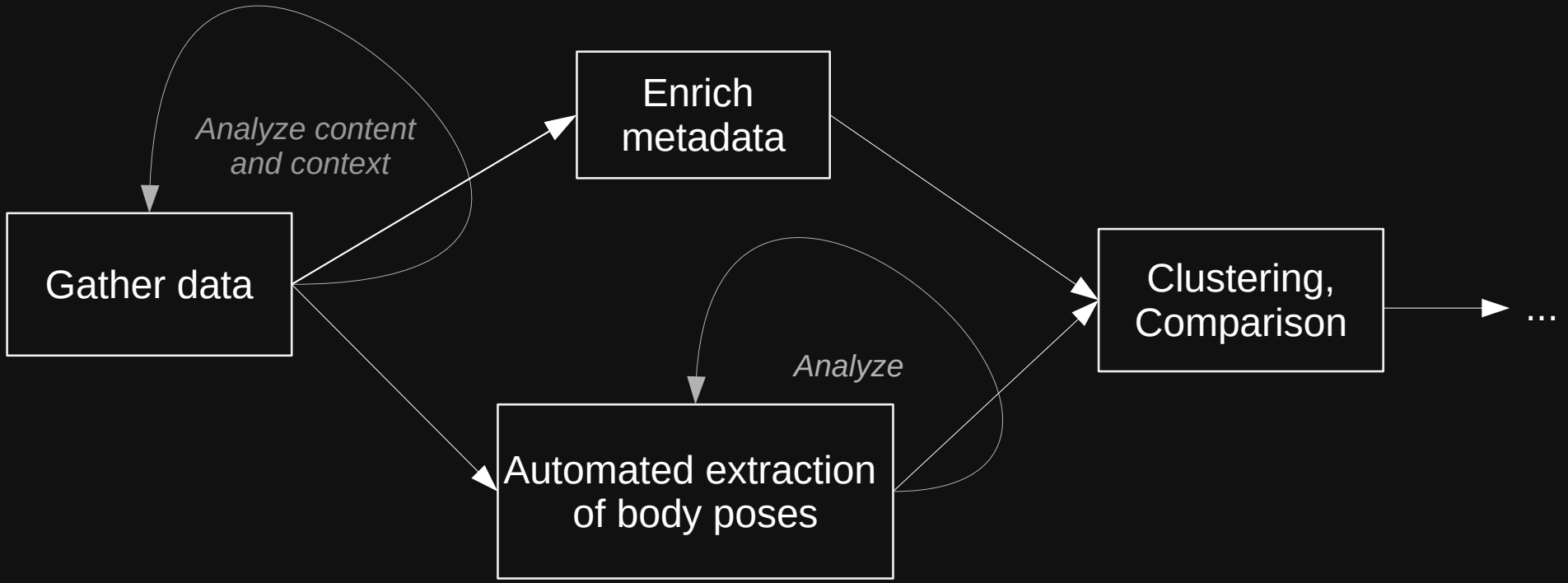
Missing annotations: towards automated gesture recognition on artworks

Valentine Bernasconi, PhD fellow
Digital Visual Studies
Max Planck Society / University of Zurich

Computational and Historical Analysis of Hands and Gestures in Early Modern Art

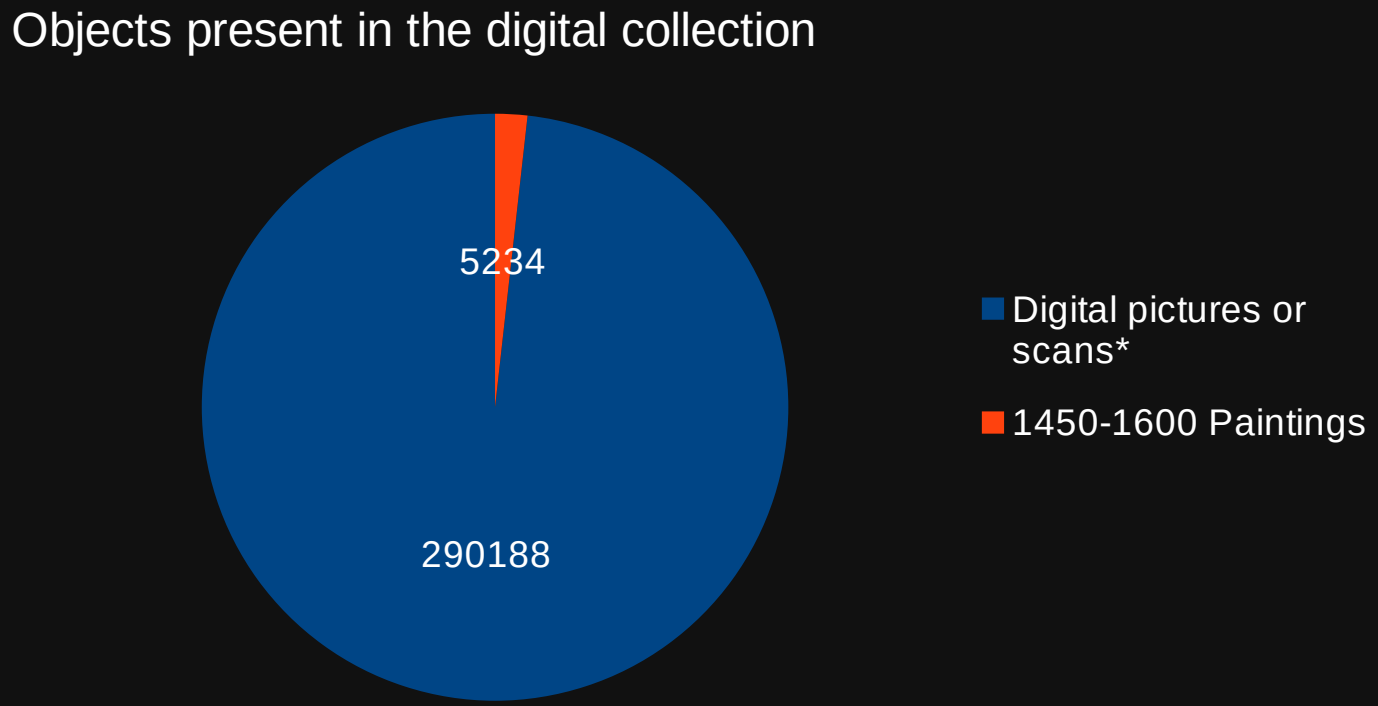
- Understanding most recurrent gestures in Early Modern art
- How gestures participate in the narrative system of paintings
- Who are the source painters of specific gestural patterns
- What are their evolutions

First general steps



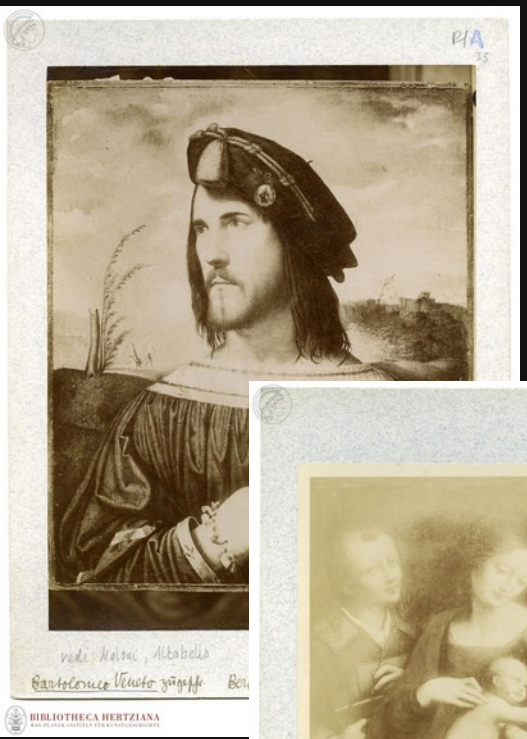
The dataset

- Fototeca, collection of the Bibliotheca Hertziana



*Objects present as per 11.2020

Visual content

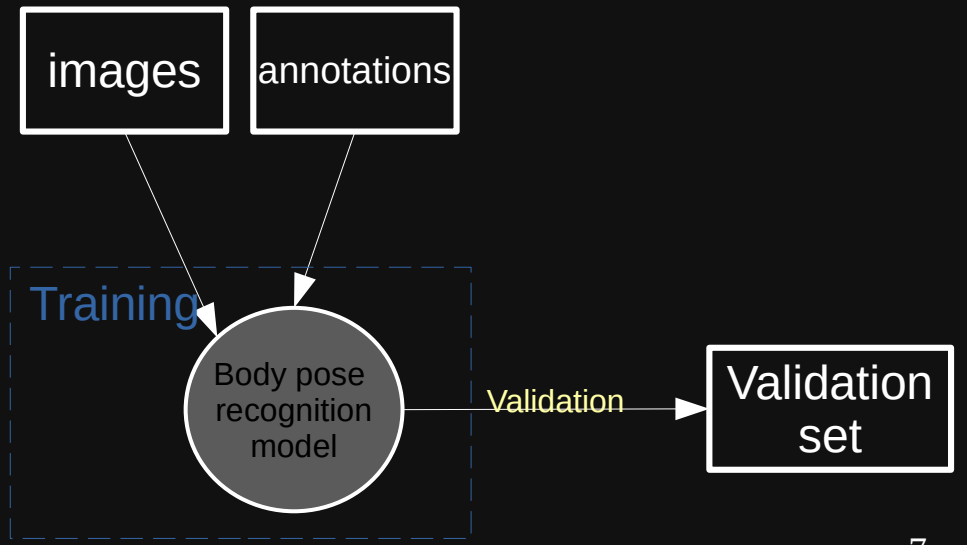


Automated recognition of body poses

- How many bodies
- Where they are positioned on the picture
- Description of the body poses

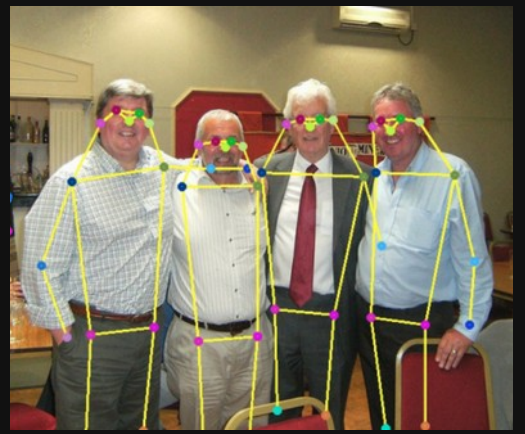
Computer vision

- Human Pose Estimation
- Existing ML models: openpose, Detectron2...
- Big training datasets



Microsoft COCO dataset

- 250’000 images of people with keypoints
- 17 body keypoints
- 68 face keypoints
- 42 hands keypoints



Microsoft COCO dataset

Problem

- Models trained on real images do not perform well on artworks
- Many differences of image content

Main differences

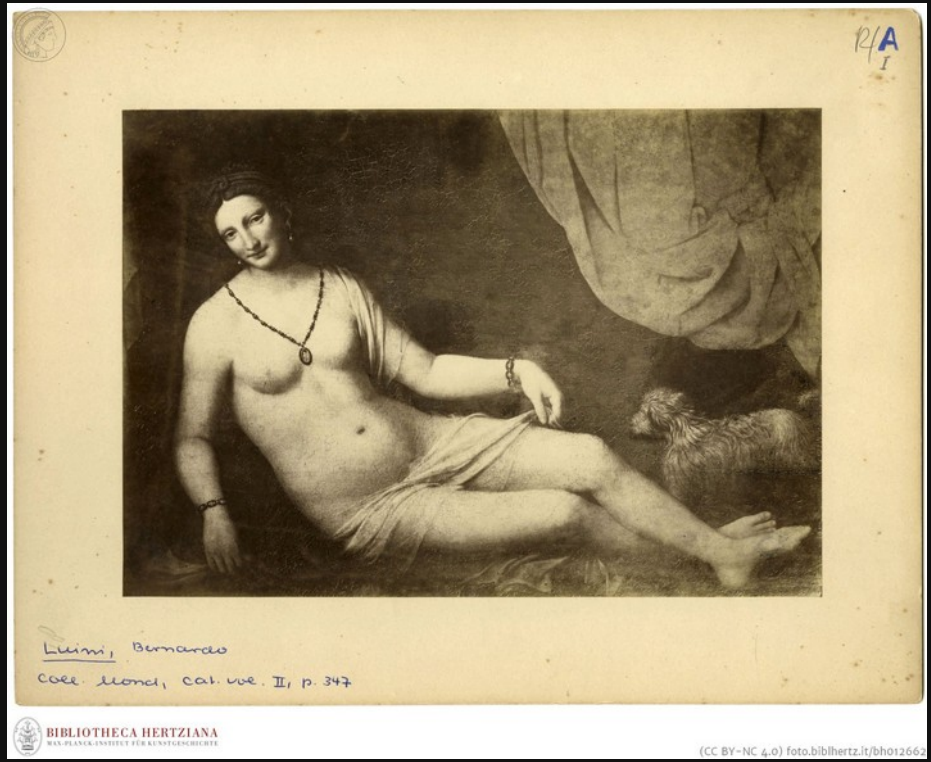
- Layout
- Content
- Shape

Main differences

- Layout
- Conteant
- Shape

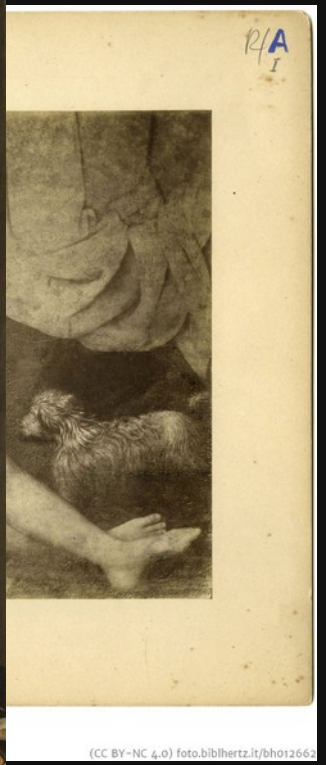
Main differences

- Layout
- Content
- Shape



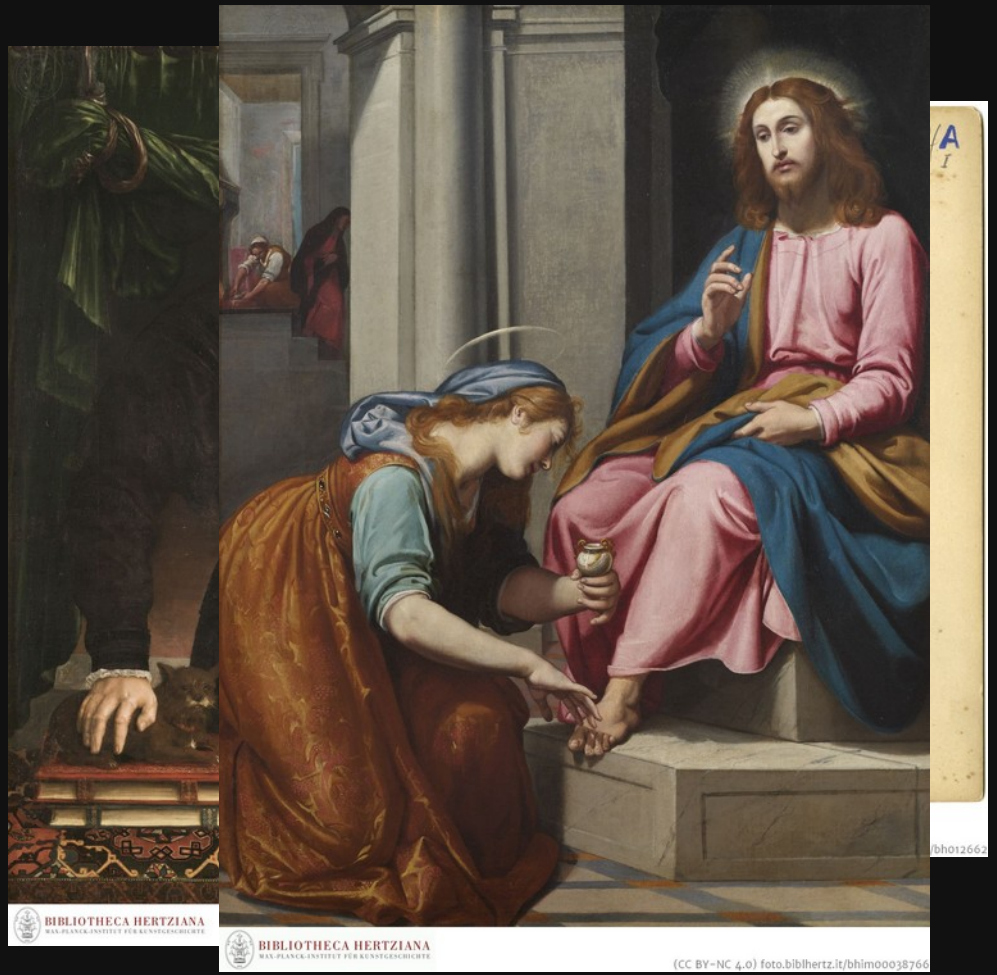
Main differences

- Layout
- Content
- Shape



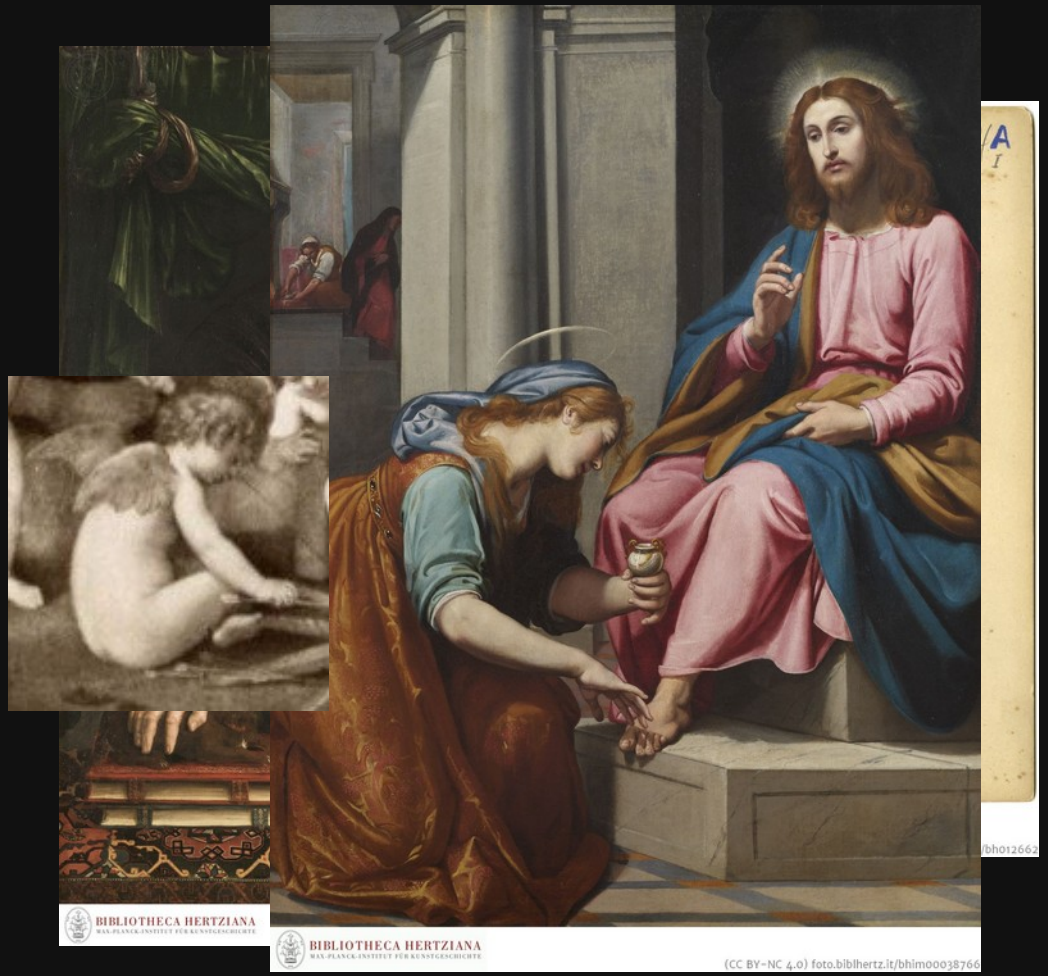
Main differences

- Layout
- Content
- Shape



Main differences

- Layout
- Content
- Shape

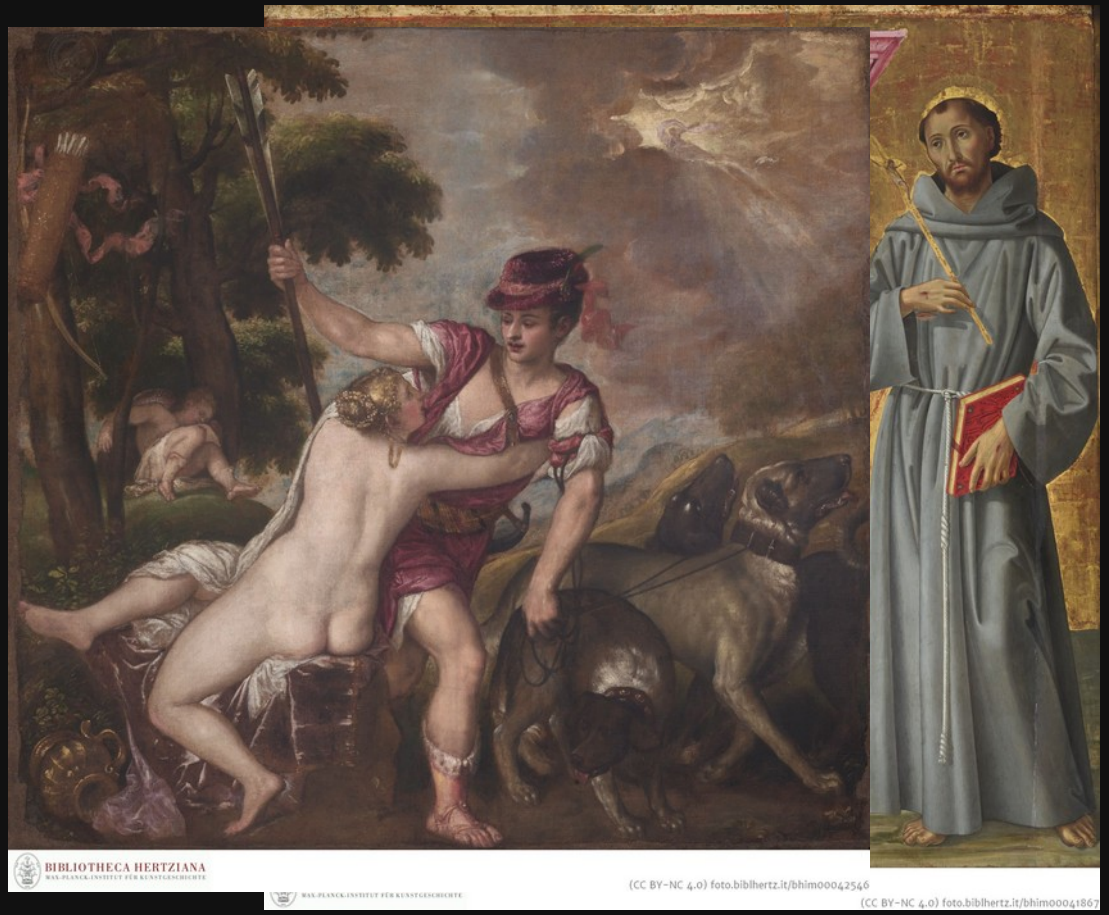


Main differences

- Layout
- Content
- Shapes

Main differences

- Layout
- Content
- Shapes

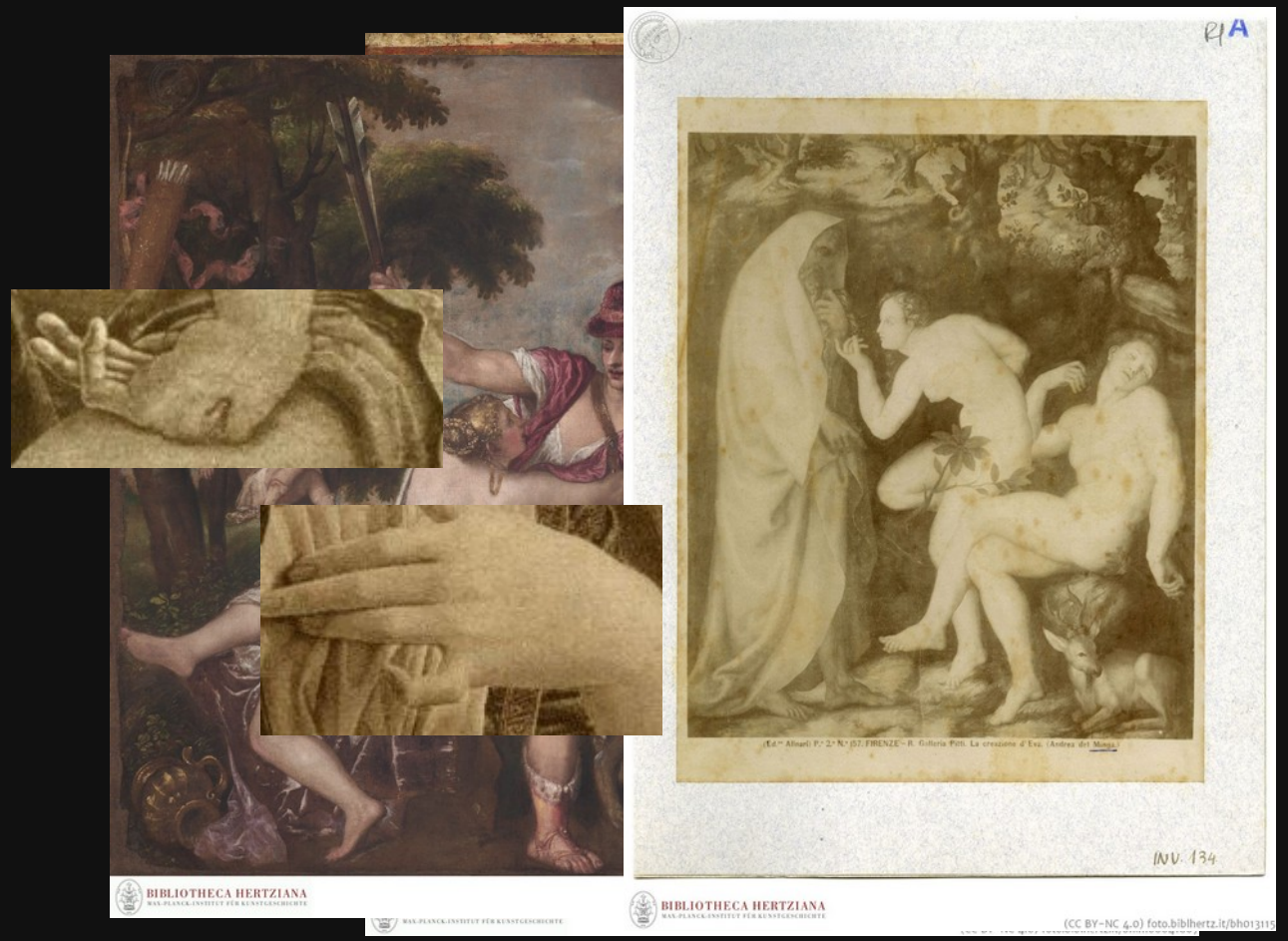


Main differences

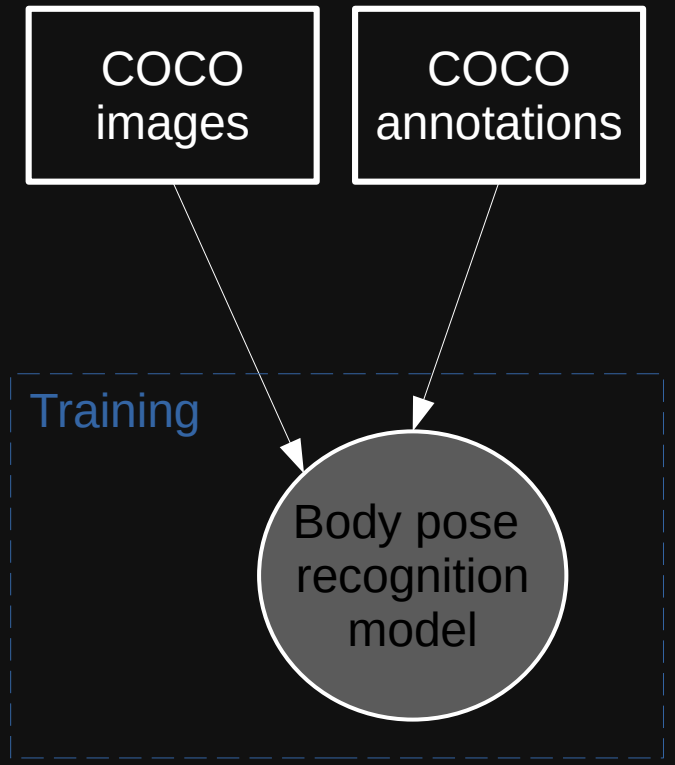
- Layout
- Content
- Shapes

Main differences

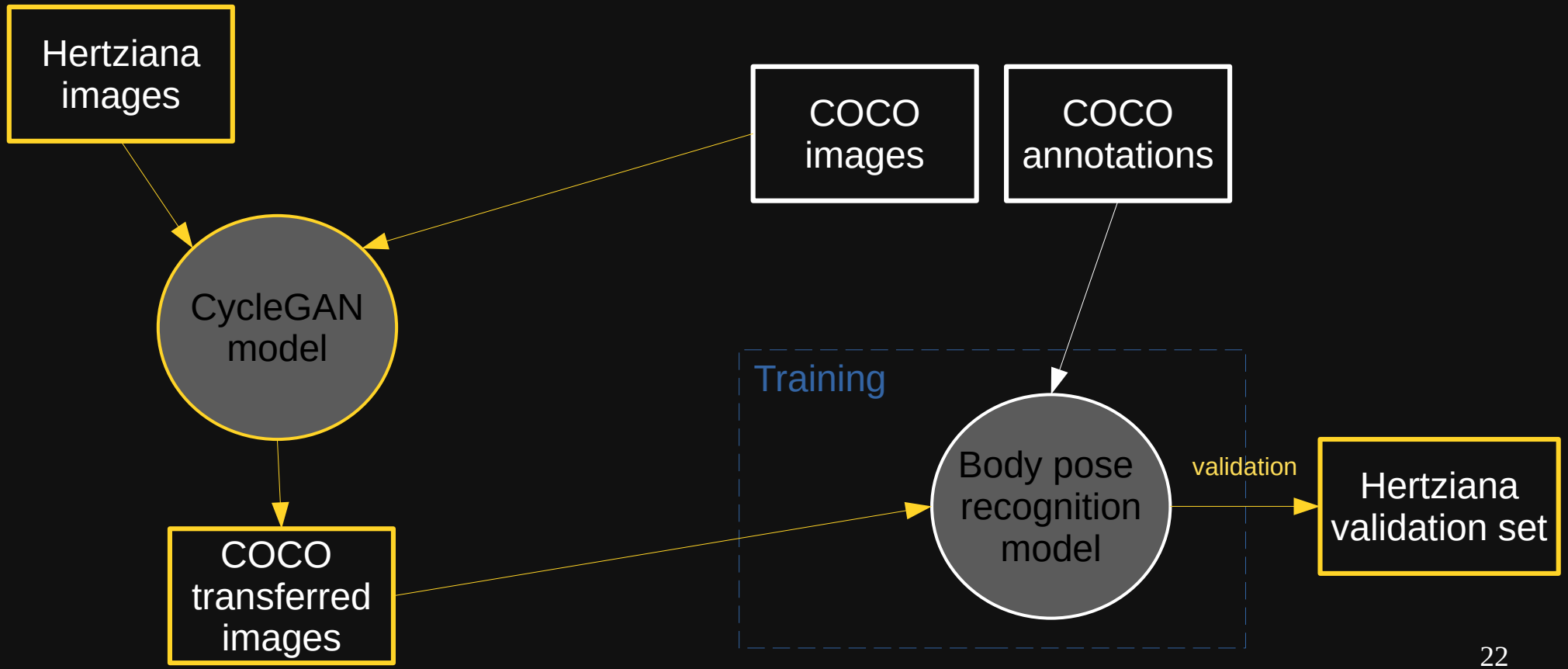
- Layout
- Content
- Shapes



Domain adaptation

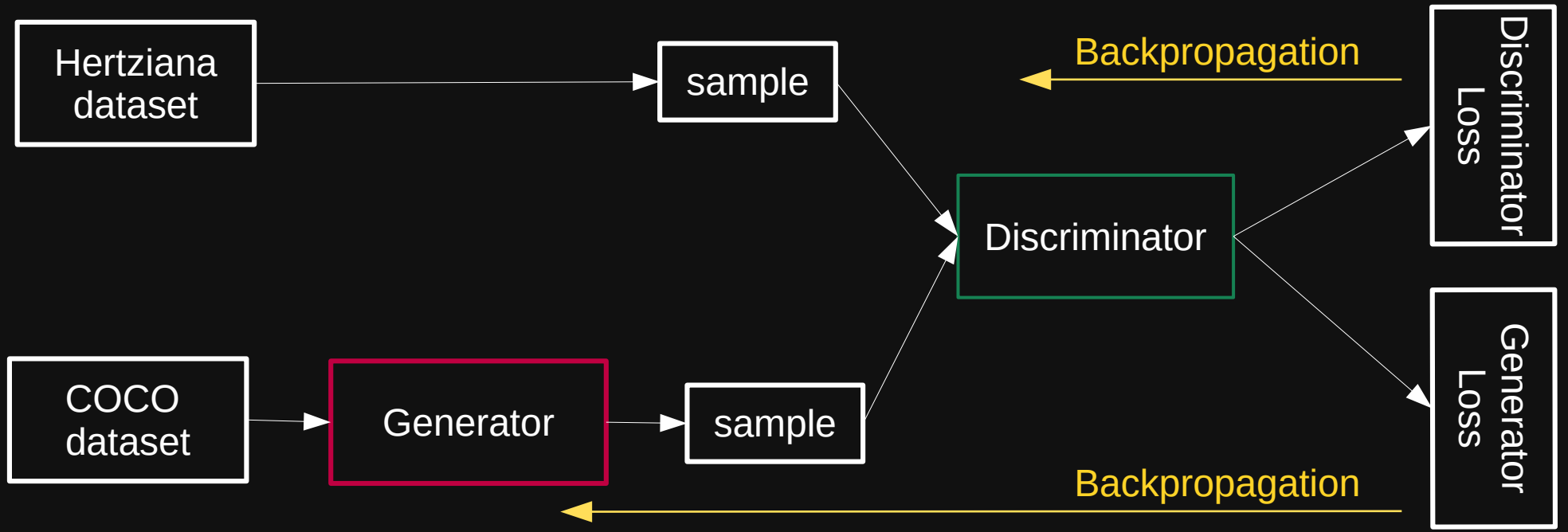


Domain adaptation



Domain adaptation

- Generative Adversarial Networks



Domain adaptation

- Cycle GAN

Real image

Transferred image

Domain adaptation

- Cycle GAN

Real image

Transferred image

Domain adaptation

- Cycle GAN

Real image

Transferred image

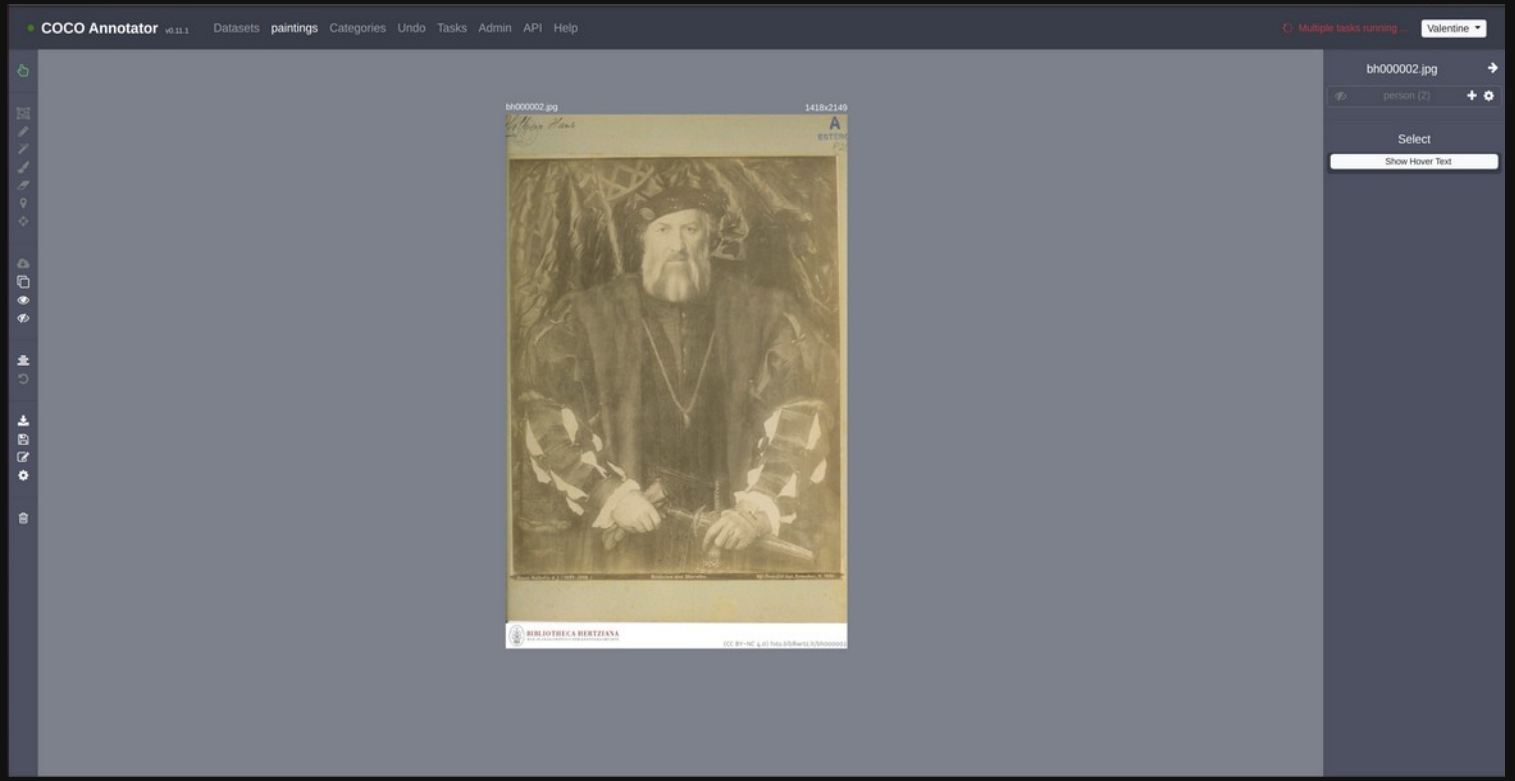
Domain adaptation

- Cycle GAN

Real image

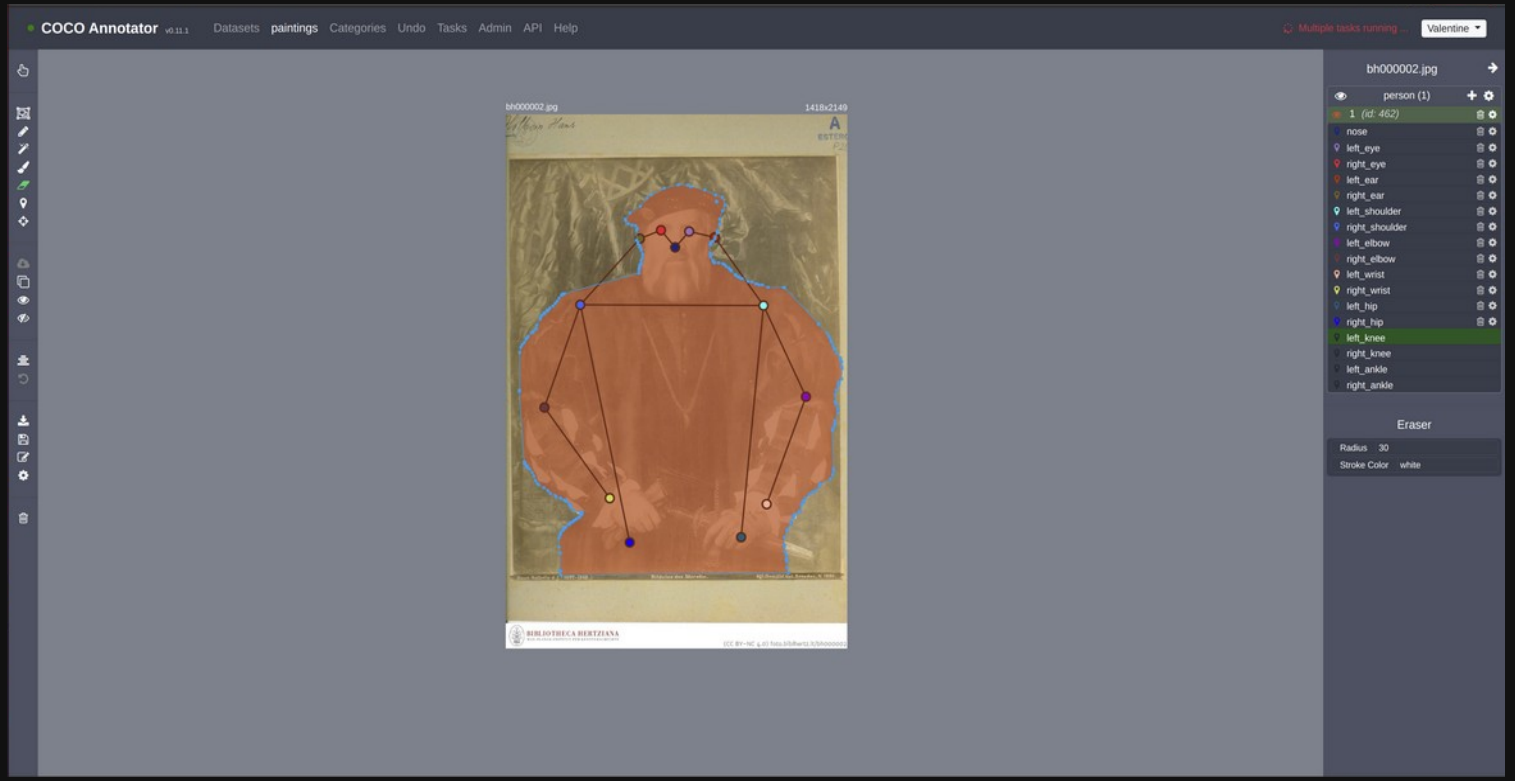
Transferred image

Creation of a validation set



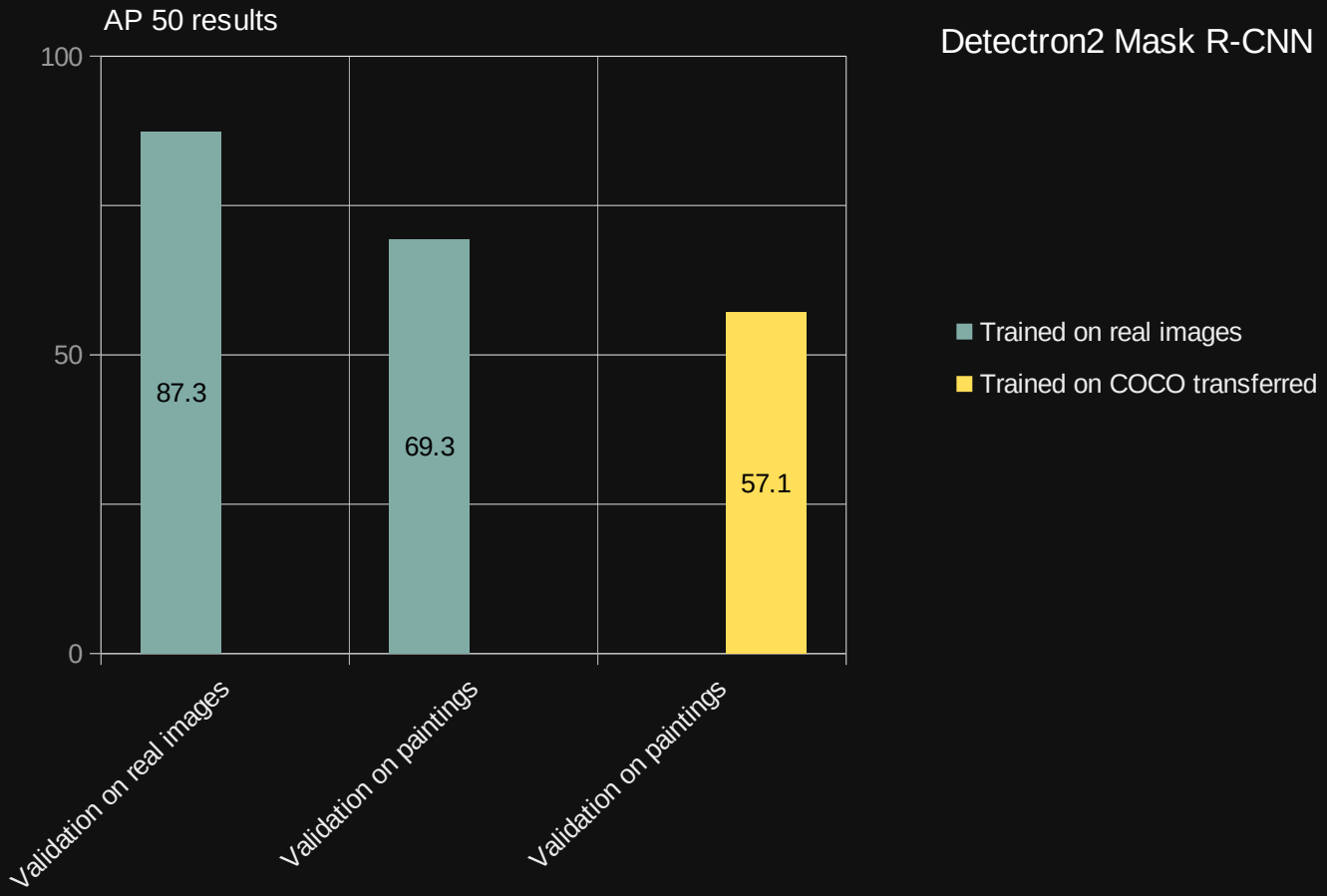
Justin Brooks, *COCO Annotator*, https://github.com/jsbroks/coco-annotator/, 2019

Creation of a validation set



Justin Brooks, *COCO Annotator*, https://github.com/jsbroks/coco-annotator/, 2019

First results comparison



What else?

- Need for solutions to handle artistic specifics:
  - Context
  - Body morphology
  - Textures
  - Visual features
  - ...

Pose Annotations Project for Artworks (PAPA)

- Creation of an annotated dataset for artworks
- Use standard COCO annotation format
- Participatory annotation platform:
  - Segmentation
  - Keypoints annotation
  - Validation

Future questions

- Should we create more categories than *person*?
- To what kind of artwork could the annotations be extended?
- How do we annotate more abstract content?

References

- Cao Z. and al., *OpenPose: Realtime Multi-Person 2D Pose Estimation using Part Affinity Fields*, IEEE Transactions on Pattern Analysis and Machine Intelligence, 2019
- Güler R. and al., *Densepose: Dense human pose estimation in the wild*, Proceedings of the IEEE Conference on Computer Vision and Pattern Recogntion, pages 7297-7308, 2018
- Supreeth N. and al., *Contextual Attention for Hand Detection in the Wild*, International Conference on Computer Vision (ICCV), 2019
- Antoniou, Antreas, Amos Storkey, and Harrison Edwards. "Data Augmentation Generative Adversarial Networks." *ArXiv:1711.04340 [Cs, Stat]*, March 21, 2018. http://arxiv.org/abs/1711.04340.
- Badea, Mihai, Corneliu Florea, Laura Florea, and Constantin Vertan. "Can We Teach Computers to Understand Art? Domain Adaptation for Enhancing Deep Networks Capacity to De-Abstract Art." *ArXiv:1712.03727 [Cs]*, December 11, 2017. http://arxiv.org/abs/1712.03727.
- Bousmalis, Konstantinos, Nathan Silberman, David Dohan, Dumitru Erhan, and Dilip Krishnan. "Unsupervised Pixel-Level Domain Adaptation with Generative Adversarial Networks." *ArXiv:1612.05424 [Cs]*, August 23, 2017. http://arxiv.org/abs/1612.05424.
- Farahani, Abolfazl, Sahar Voghoei, Khaled Rasheed, and Hamid R. Arabnia. "A Brief Review of Domain Adaptation." *ArXiv:2010.03978 [Cs]*, October 7, 2020. http://arxiv.org/abs/2010.03978.
- Gastineau, Eric Luc Adrien. "Domain Adaptation via Data Augmentation." Thesis, Georgia Institute of Technology, 2020. https://smartech.gatech.edu/handle/1853/62862.
- Huang, Xun, and Serge Belongie. "Arbitrary Style Transfer in Real-Time with Adaptive Instance Normalization." *ArXiv:1703.06868 [Cs]*, July 30, 2017. http://arxiv.org/abs/1703.06868.
- Rojtberg, Pavel, Thomas Pöllabauer, and Arjan Kuijper. "Style-Transfer GANs for Bridging the Domain Gap in Synthetic Pose Estimator Training." *ArXiv:2004.13681 [Cs]*, April 28, 2020. http://arxiv.org/abs/2004.13681.
- L. Impett and S. Süsstrunk, "Pose and Pathosformel in Aby Warburg's Bilderatlas,", 2016, Accessed: Jun. 03, 2021. [Online

References

- "Styled Transfer Learning for Art Historical Images › Digital Humanities and Social Sciences." Accessed December 7, 2020. https://www.izdigital.fau.eu/2020/06/15/styled-transfer-learning-for-art-historical-images/.
- Thomas, Christopher, and Adriana Kovashka. "Artistic Object Recognition by Unsupervised Style Adaptation." In *Computer Vision – ACCV 2018*, edited by C. V. Jawahar, Hongdong Li, Greg Mori, and Konrad Schindler, 11363:460–76. Lecture Notes in Computer Science. Cham: Springer International Publishing, 2019. https://doi.org/10.1007/978-3-030-20893-6_29.
- R. Girshick, "Fast R-CNN," arXiv:1504.08083 [cs], Sep. 2015, Accessed: Jun. 03, 2021. [Online]. Available: http://arxiv.org/abs/1504.08083
- "Fototeca | Bibliotheca Hertziana – Istituto Max Planck per la storia dell'arte." https://www.biblhertz.it/it/photographic-collection (accessed Jun. 03, 2021).
- "COCO - Common Objects in Context." https://cocodataset.org/#format-data (accessed Jun. 03, 2021).
- T.-Y. Lin et al., "Microsoft COCO: Common Objects in Context," arXiv:1405.0312 [cs], Feb. 2015, Accessed: Jun. 03, 2021. [Online]. Available: http://arxiv.org/abs/1405.0312
- "MPII Human Pose Database." http://human-pose.mpi-inf.mpg.de/ (accessed 03, 2021).
- F. Milani and P. Fraternali, "A Data Set and a Convolutional Model for Iconography Classification in Paintings," arXiv:2010.11697 [cs], Nov. 2020, Accessed: Jun. 01, 2021. [Online]. Available: http://arxiv.org/abs/2010.11697
- V. Marsocci and L. Lastilla, "POSE-ID-on—A Novel Framework for Artwork Pose Clustering," ISPRS International Journal of Geo-Information, vol. 10, no. 4, Art. no. 4, Apr. 2021, doi: 10.3390/ijgi10040257.
- P. Madhu et al., Enhancing Human Pose Estimation in Ancient Vase Paintings via Perceptually-grounded Style Transfer Learning. 2020
- "Wild Mont-Blanc : un nouveau programme de science participative à succès ! | Sciences en relief," May 27, 2020. https://blog.creamontblanc.org?p=2879 (accessed Jun. 03, 2021).
- G. Raschia and B. Hervy, "Recherche de consensus : les enseignements de la plate-forme RECITAL de transcription participative," Lille, France, Oct. 2018. Accessed: Jun. 03, 2021. [Online]. Available: https://hal.archives-ouvertes.fr/hal-02536772