École des Ponts
ParisTech

Deep Learning for Near-duplicated Patterns Discovery and Alignment in Artworks

Author: Xi Shen

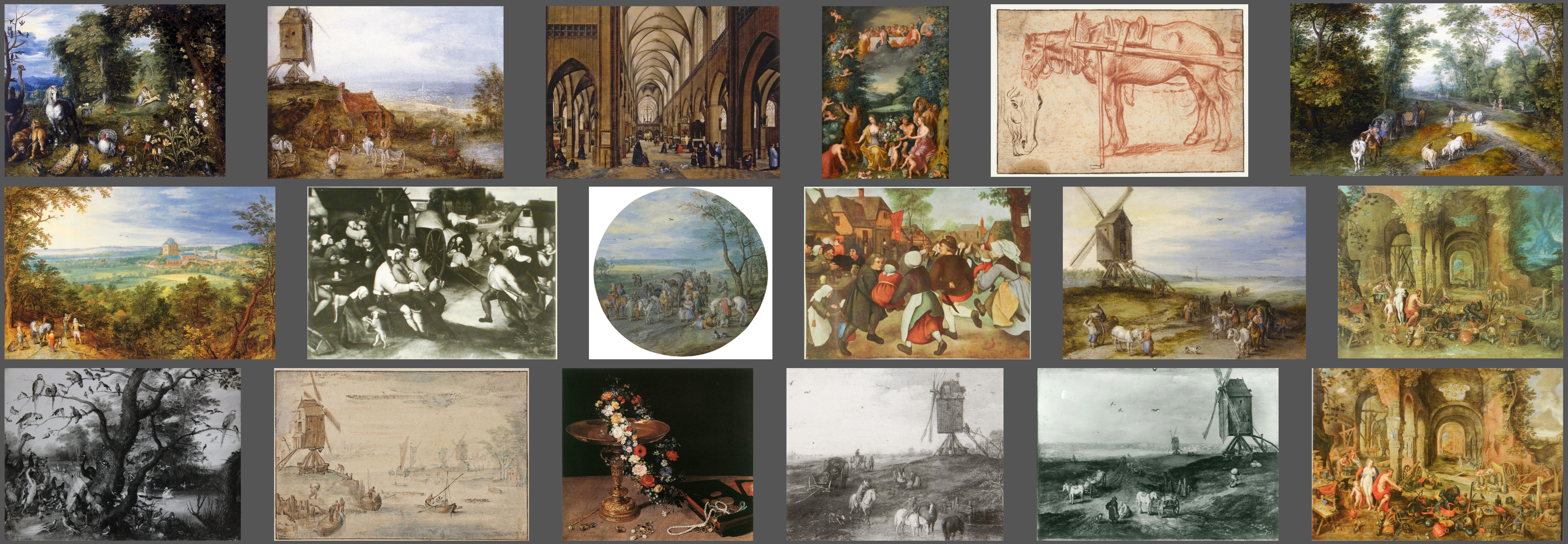
Thesis Advisor: Mathieu Aubry

Dec. 3, 2021

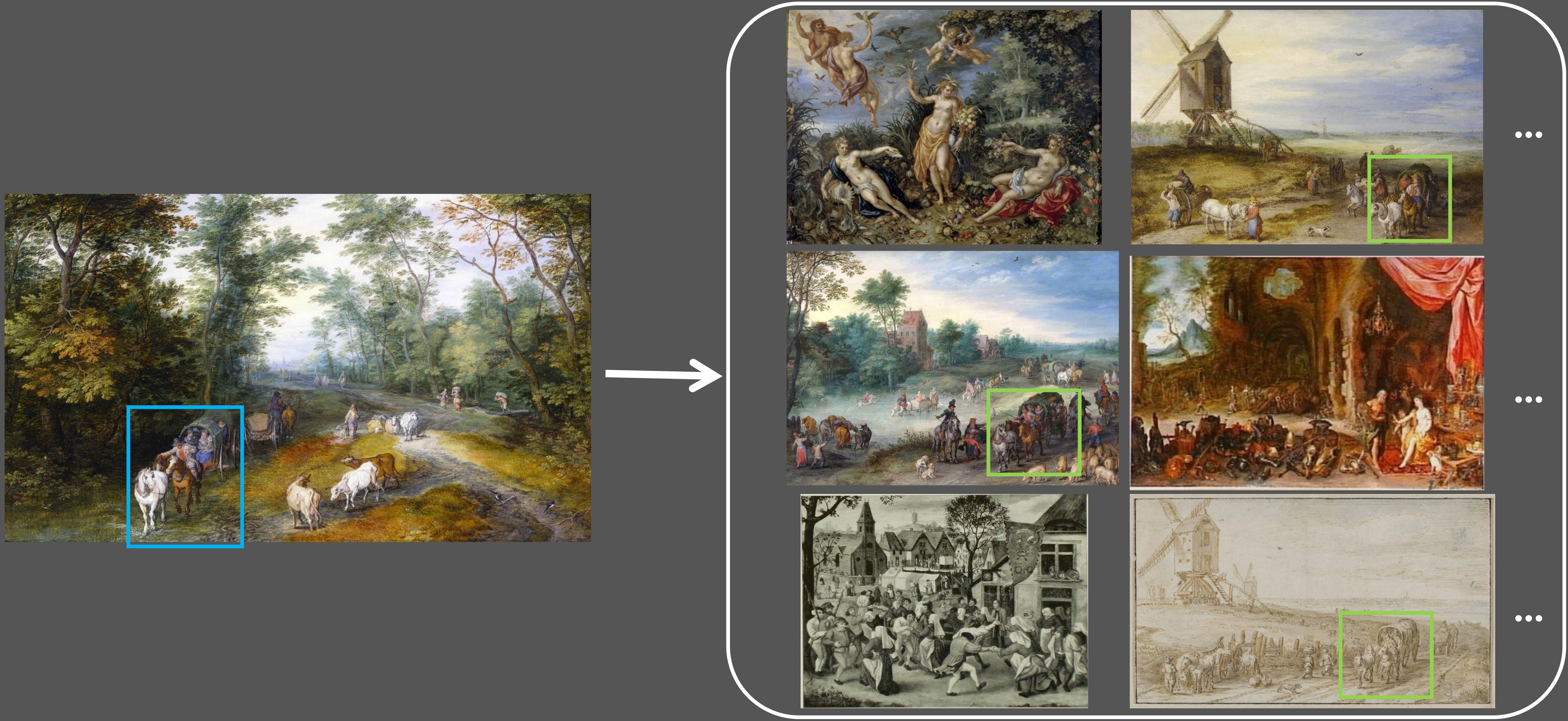
Motivation: art analysis on Brueghel

https://www.janbrueghel.net/

Several artworks in Brueghel (total 1 586 artworks):

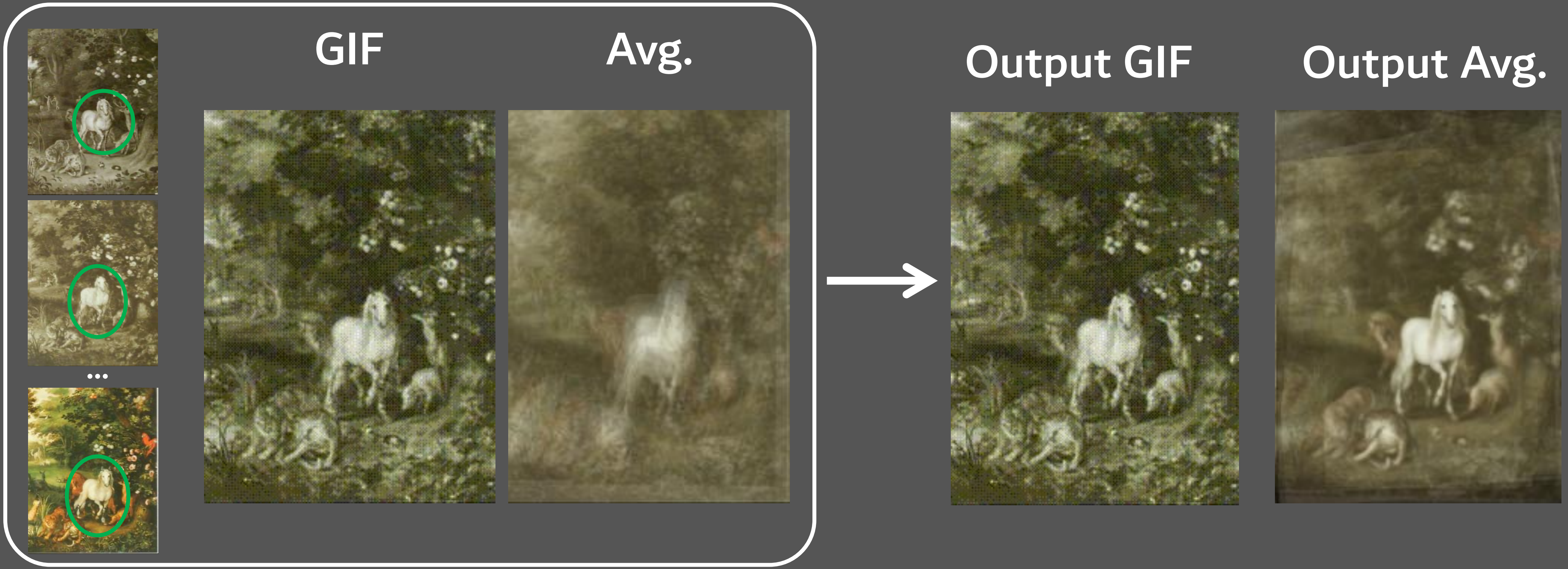


Motivation: One-shot detection

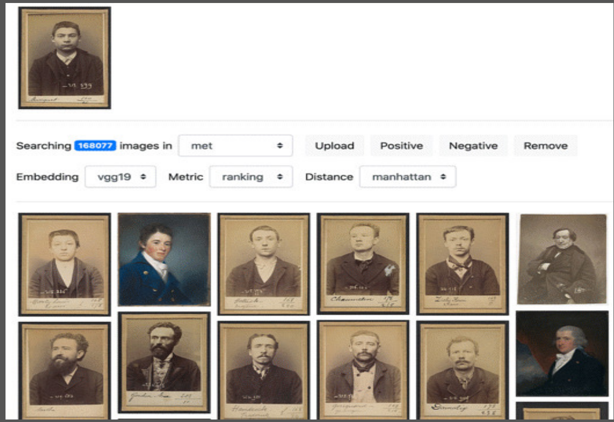


Motivation: Discovery

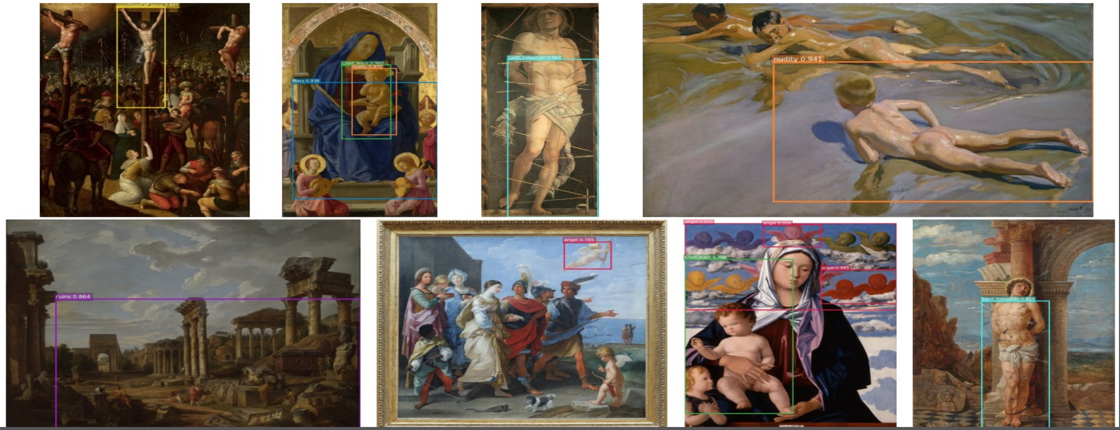
Motivation: Dense alignment



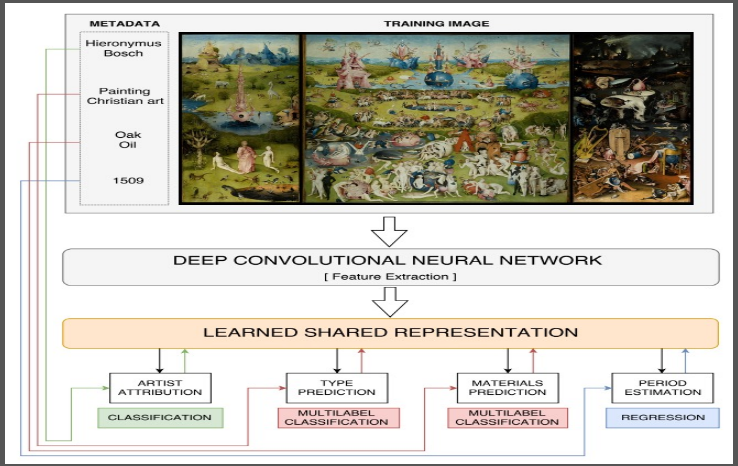
Related works: computer vision and art



Artwork retrieval. [Shrivastava et al. 2011; Crowley et al. 2015; Seguin et al. 2016 ...]



Object detection in artworks. [Ginasor et al. 2014; Crowley and Zisserman, 2014; Gonthier et al, 2018...]



Attributes prediction. [Karayev et al. 2014; Van Noord et al. 2015; Strezoski and Worring 2019 ...]

Creating artworks. [Gatys et al. 2015; Zhu et al. 2017; Elgammal et al. 2017...]

Challenges

1. Lack of supervision

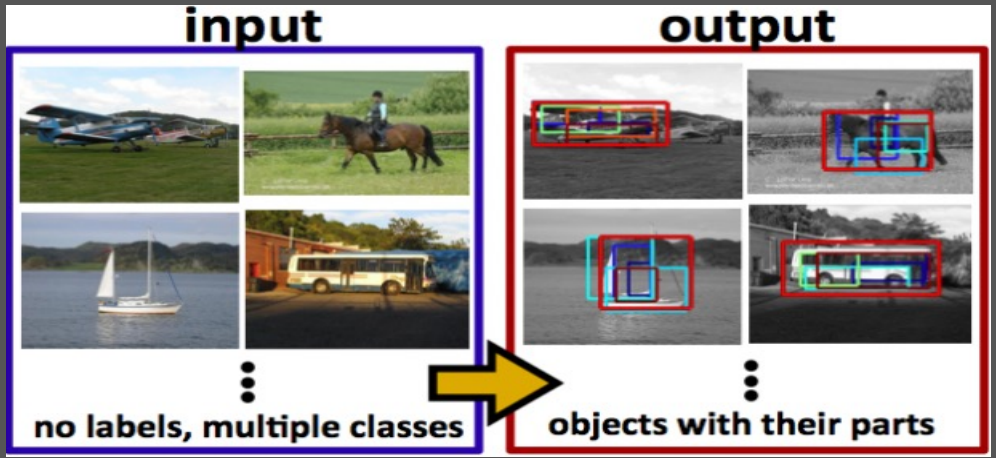
2. Diversity of depiction styles

3. Scalability

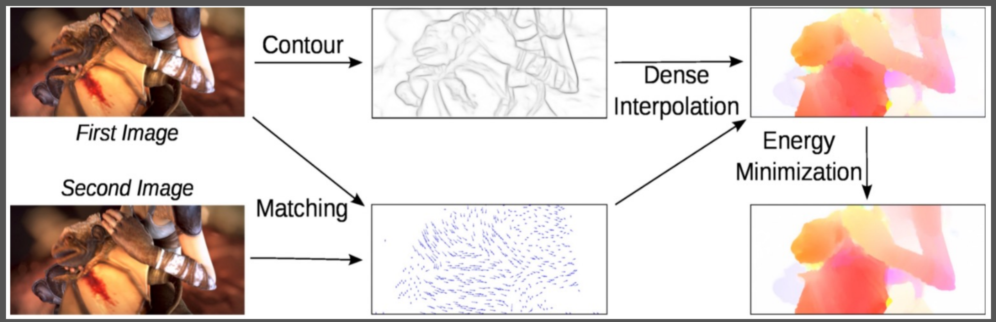
Related works: non-deep approaches

Image retrieval. [Sivic and Zisserman, 2003; Nister and Stewenius, 2006; Philbin et al. 2007...]

Object discovery. [Tang et al. 2014; Cho et al. 2015; Vo et al. 2019...]

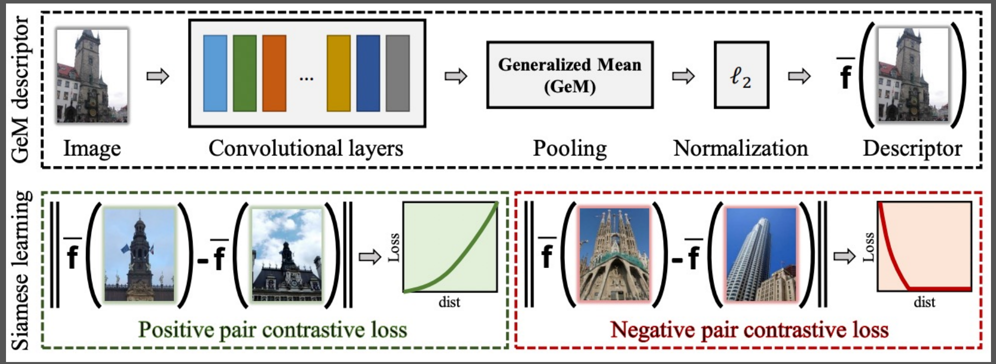


Optical flow. [Brox et al. 2009, Liu et al. 2010; Revaud et al. 2015...]

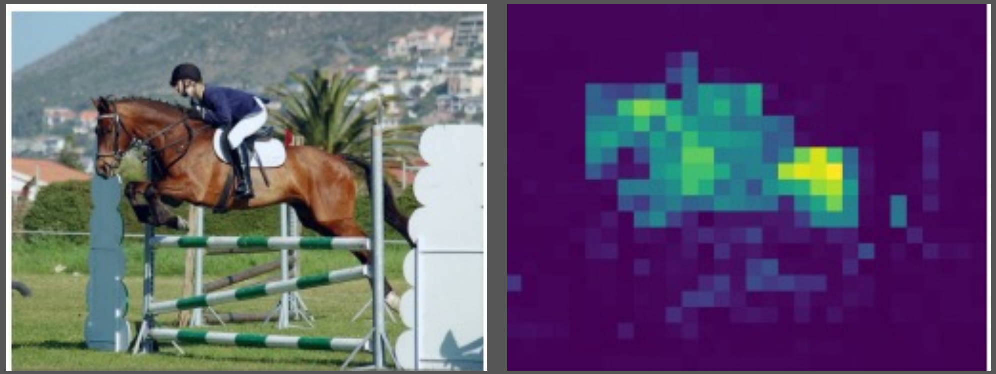


Related works: deep approaches

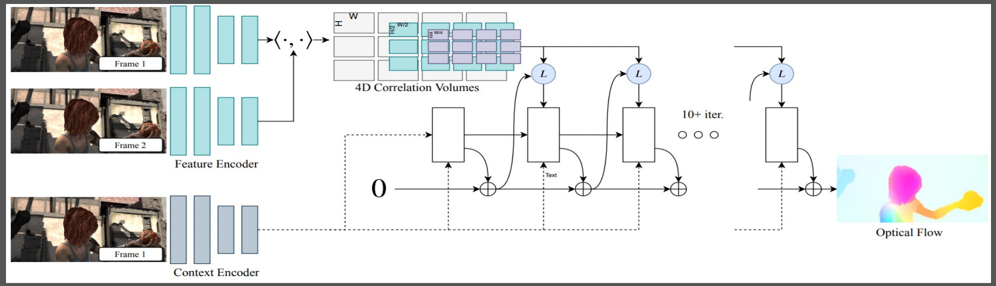
Image retrieval. [Babenko et al, 2014; Gordo et al, 2017; Radenovic et al. 2018...]



Object discovery. [Vo et al. 2020; Chen et al. 2020; Simeoni et al. 2021...]



Optical flow. [Dosovitskiy et al. 2015; Ilg et al. 2017; Teed and Deng 2020...]



Contributions in my thesis

1. Style-invariant feature from self-supervision

2. Co-segmentation from synthetic data

3. Dense image alignment from reconstruction

Contributions in my thesis

1. Style-invariant feature from self-supervision

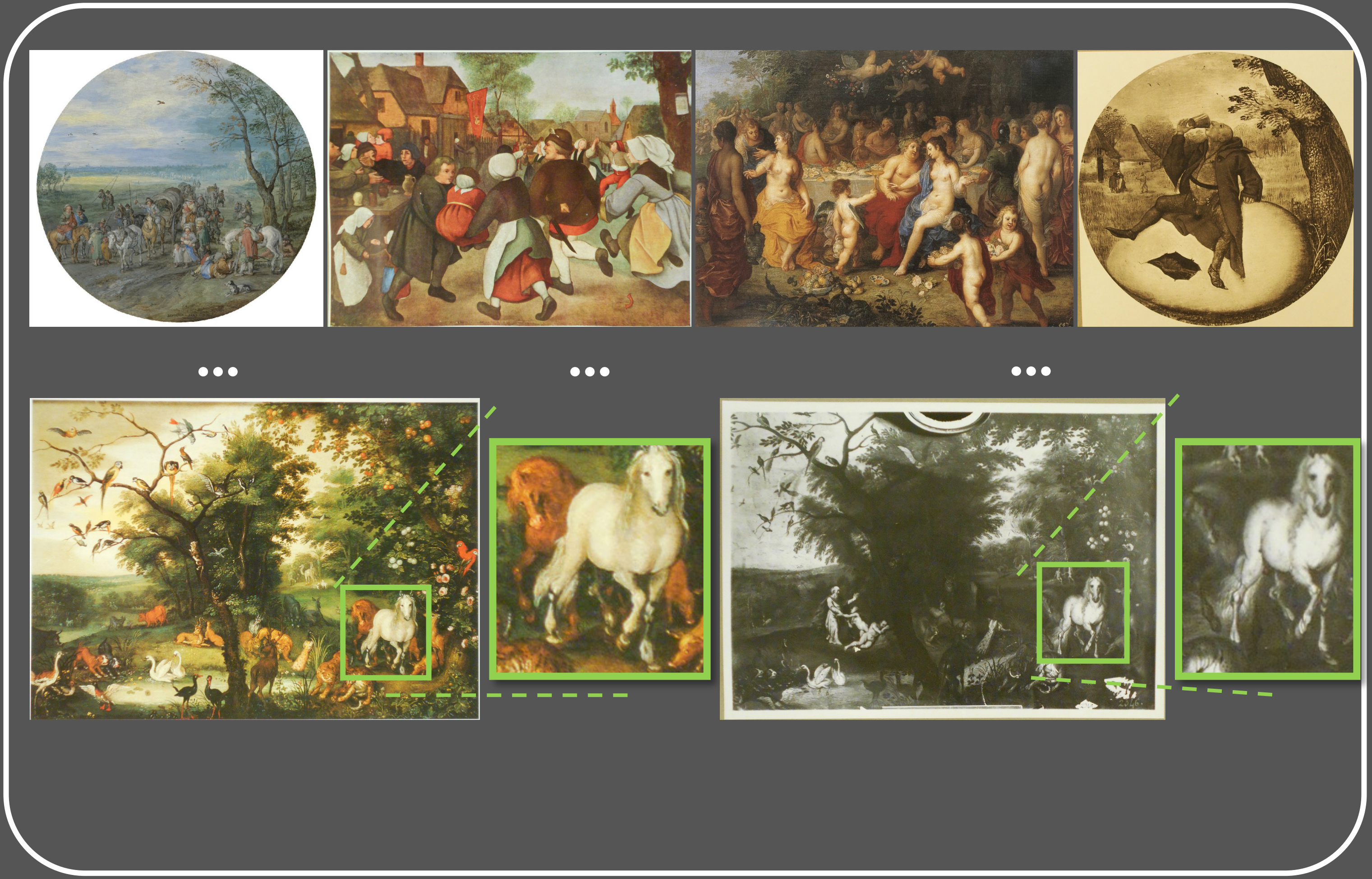
2. Co-segmentation from synthetic data

3. Dense image alignment from reconstruction

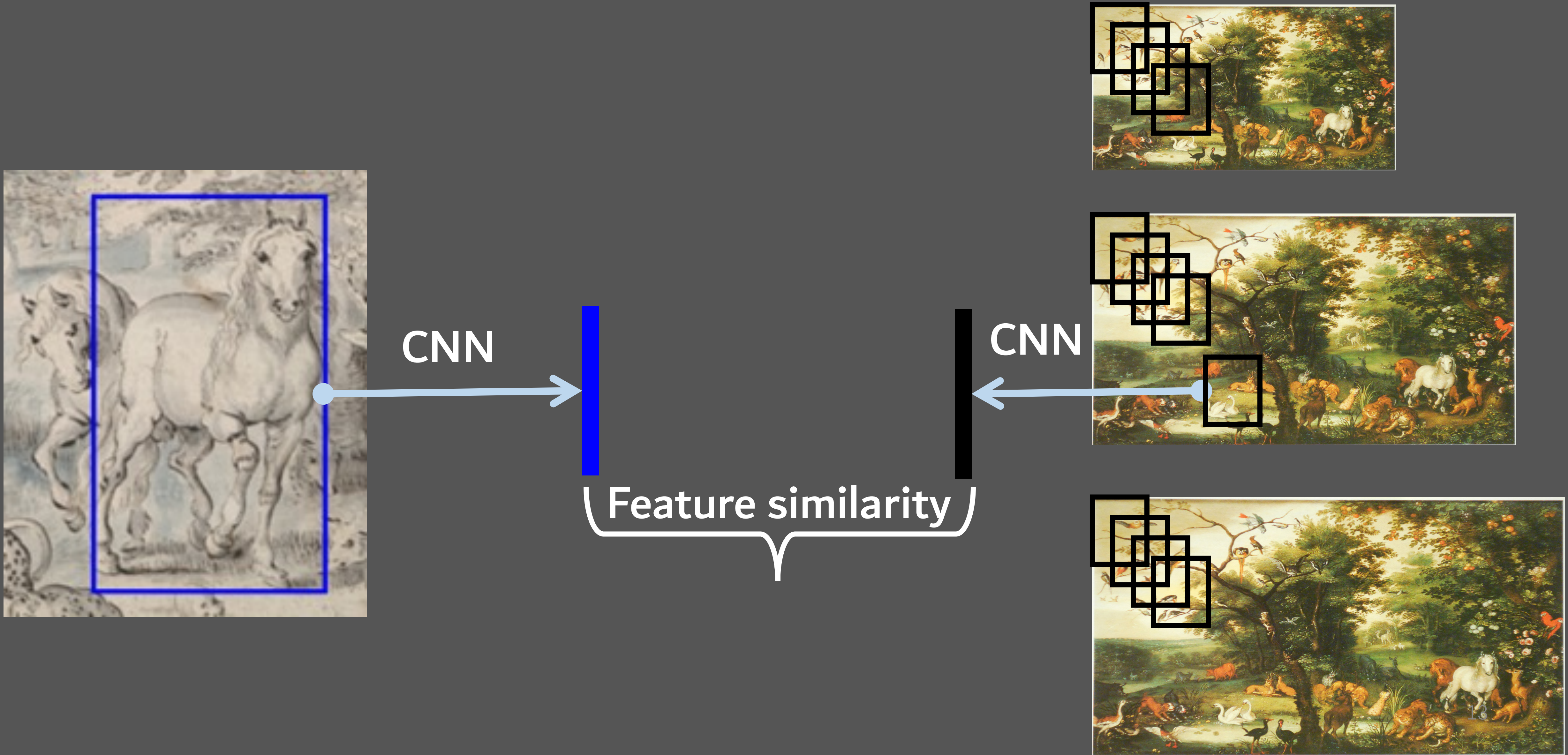
Task: one-shot detection

Query

Detect near-duplicated patterns



Solution: multi-scale feature matching



Problem: ImageNet feature results

Query

Top Matches

ImageNet features are not invariant to style

No training data available

Key idea: metric learning

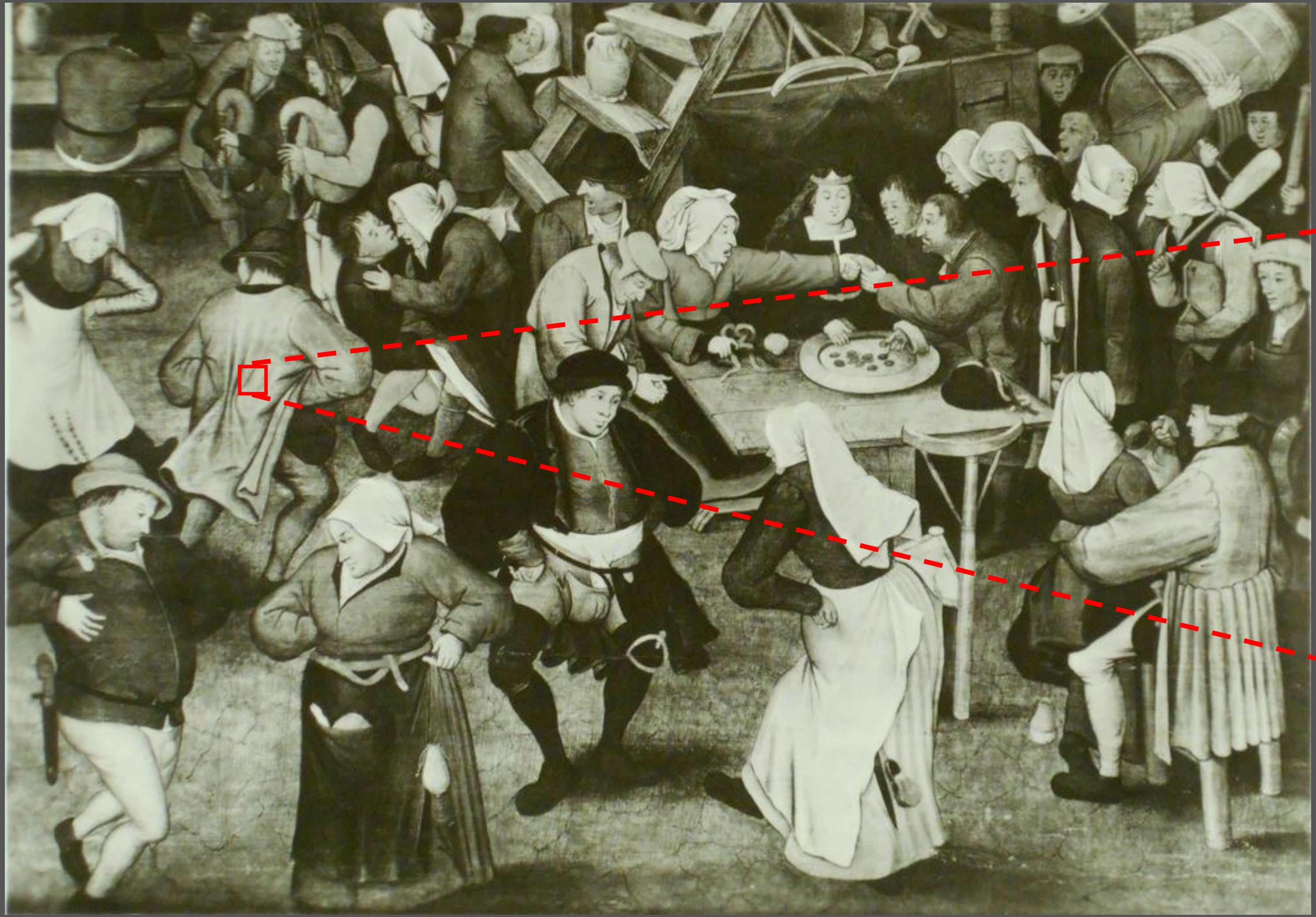
Positive (P_1 and P_2) and negative pair (P_1 and N_1) :

$$L(P_1, P_2, N_1) = \max(s(P_1, N_1), 1-\lambda) - \min(s(P_1, P_2), \lambda)$$

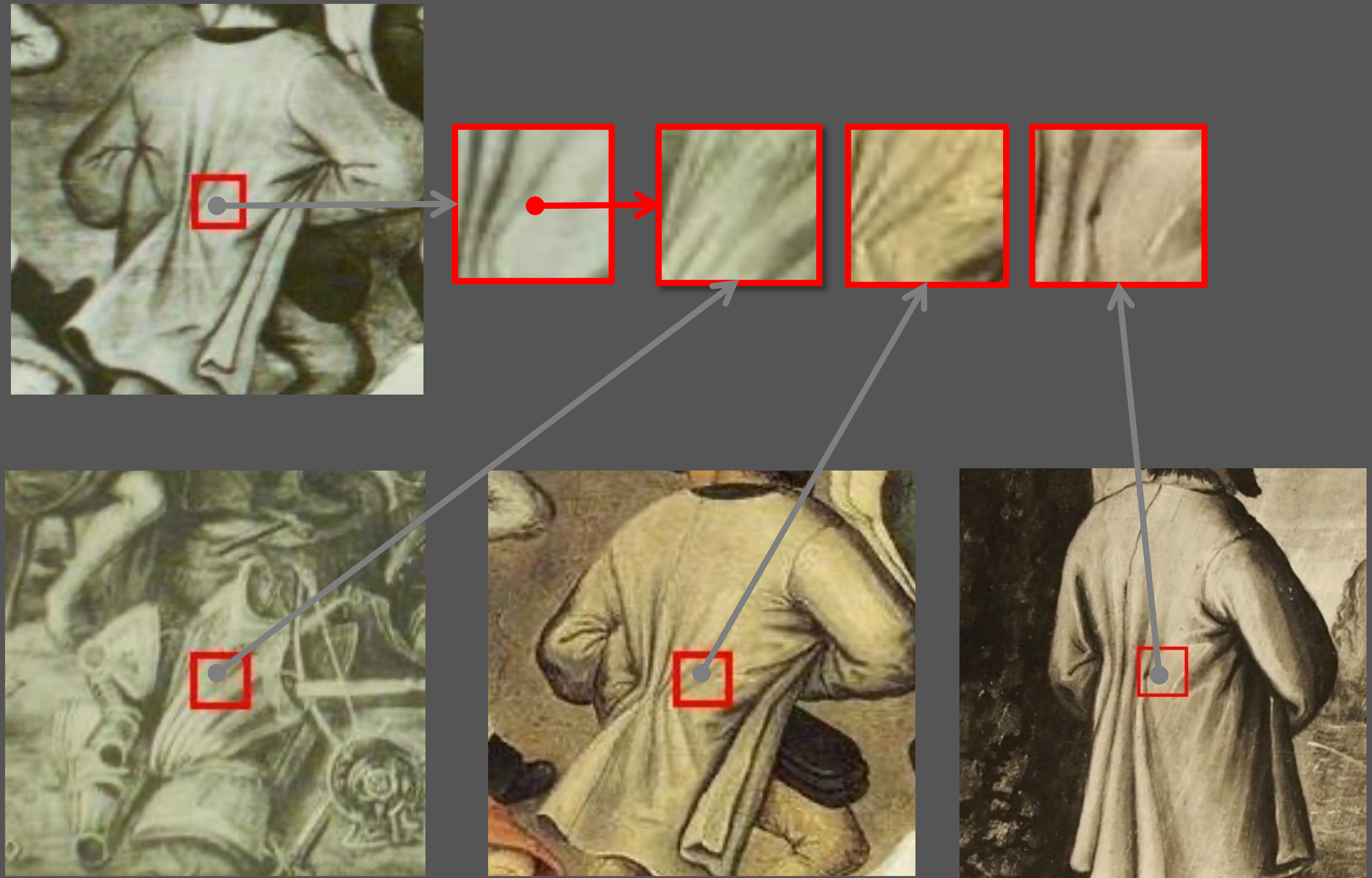
- s: cosine similarity
- λ: hyper-parameter in the triplet loss

Question: how to find positive / negative pairs?

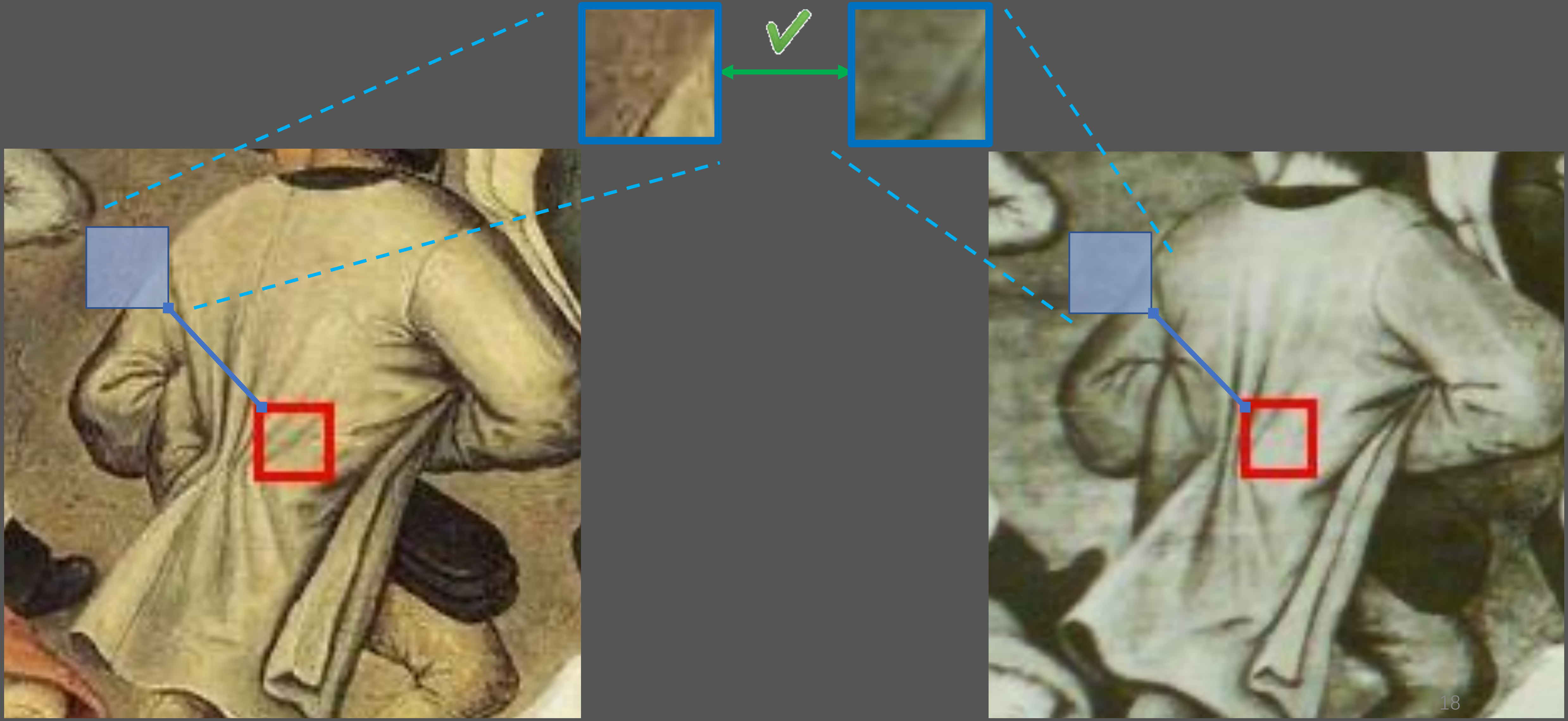
Positive pairs: query patch sampling

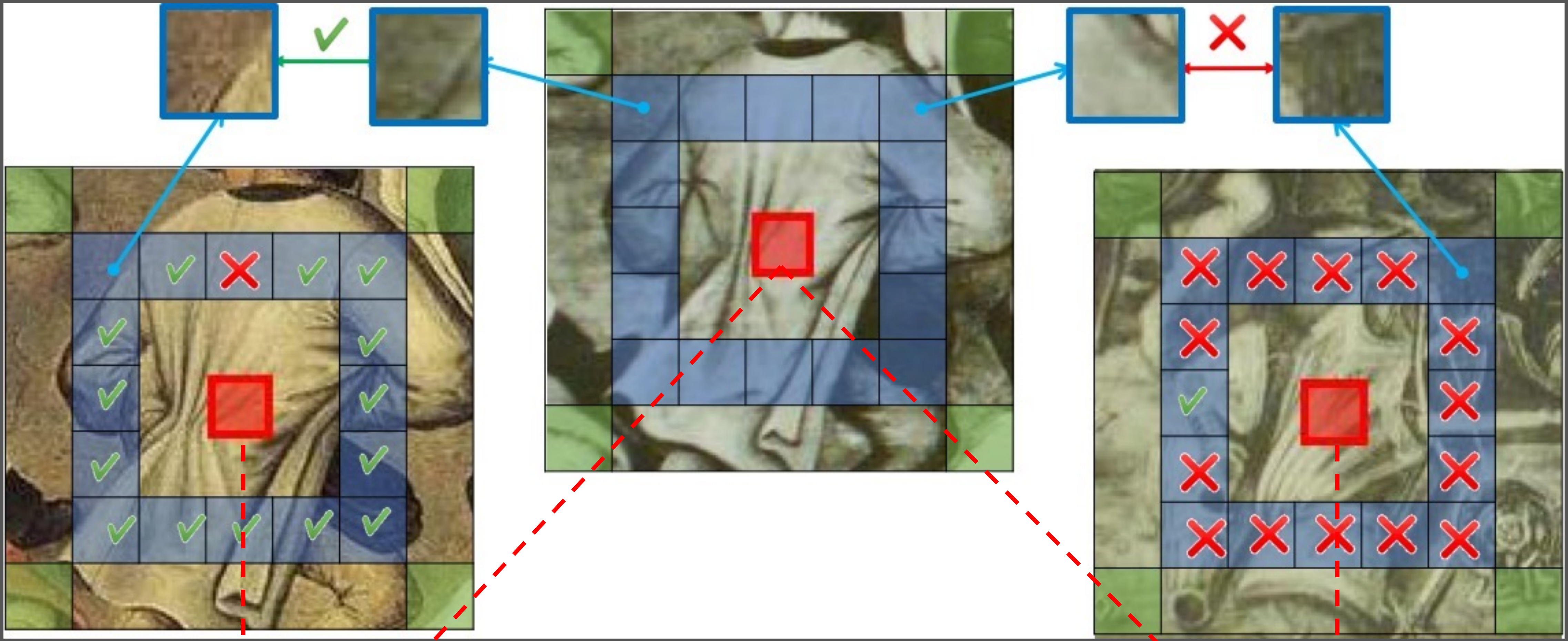


Positive pairs: candidates via matching



Positive pairs: validation from consistency

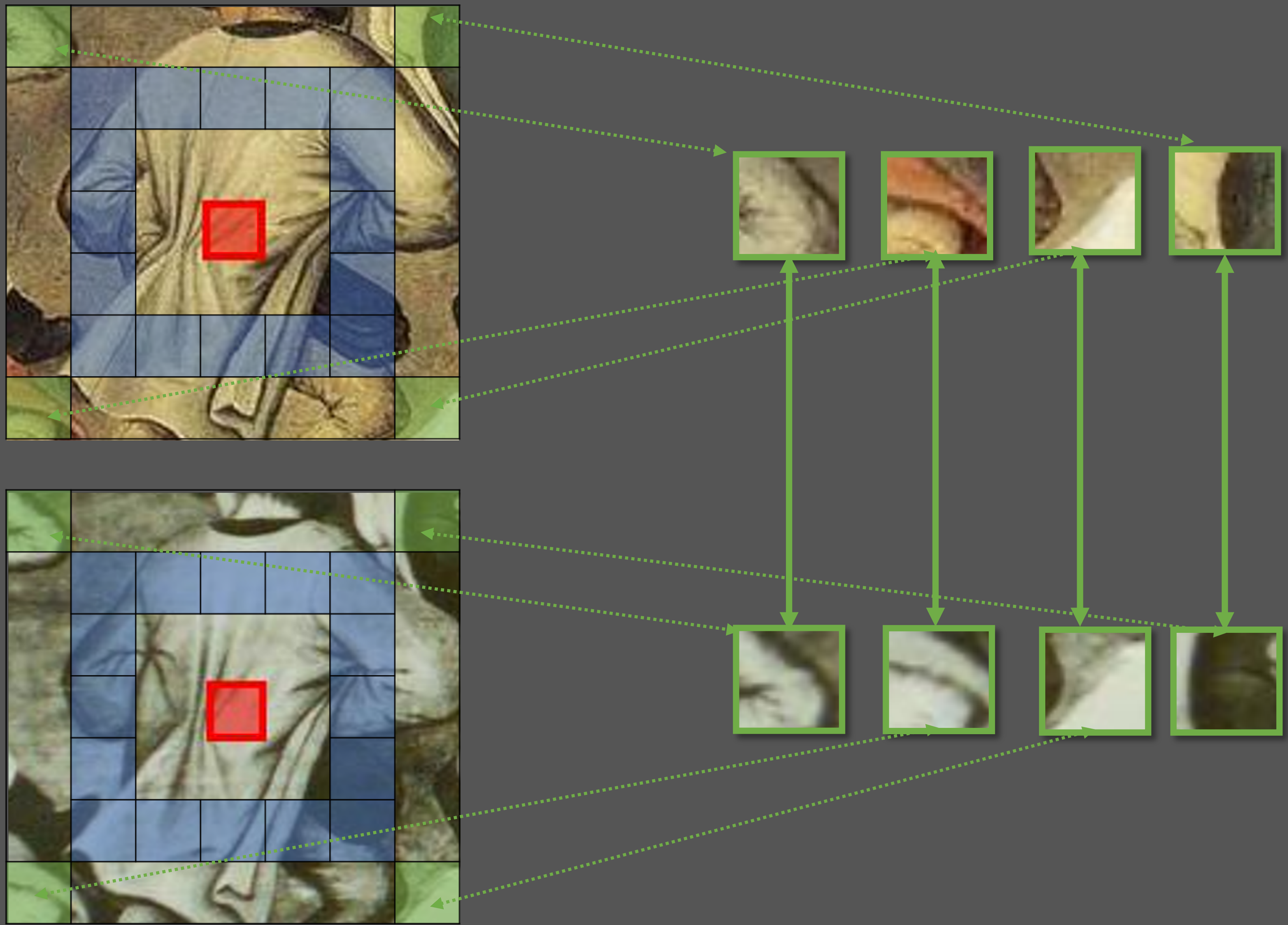




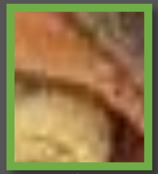
Score ( , ) = 15. / 16

Score ( , ) = 1. / 16

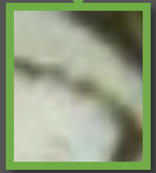
Positive pairs: hard positive mining



Negative pairs



Positive pair

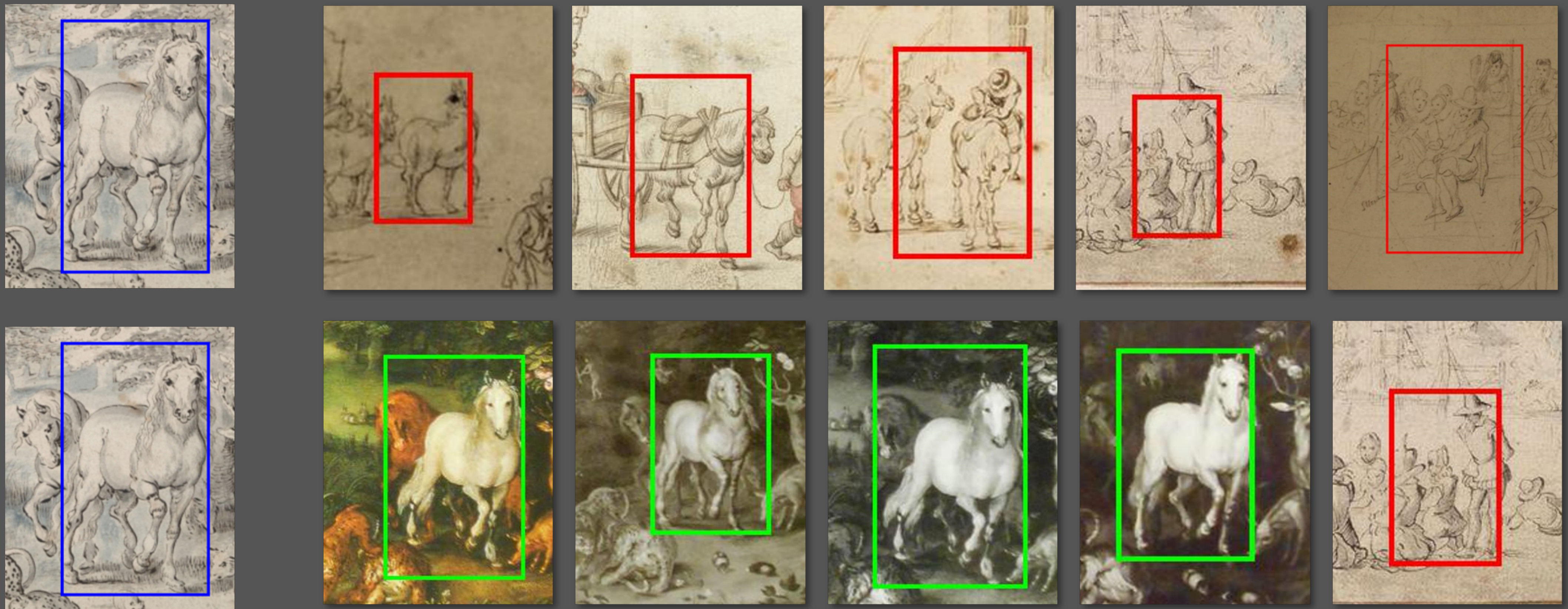


Top-k matches
as
negative pairs

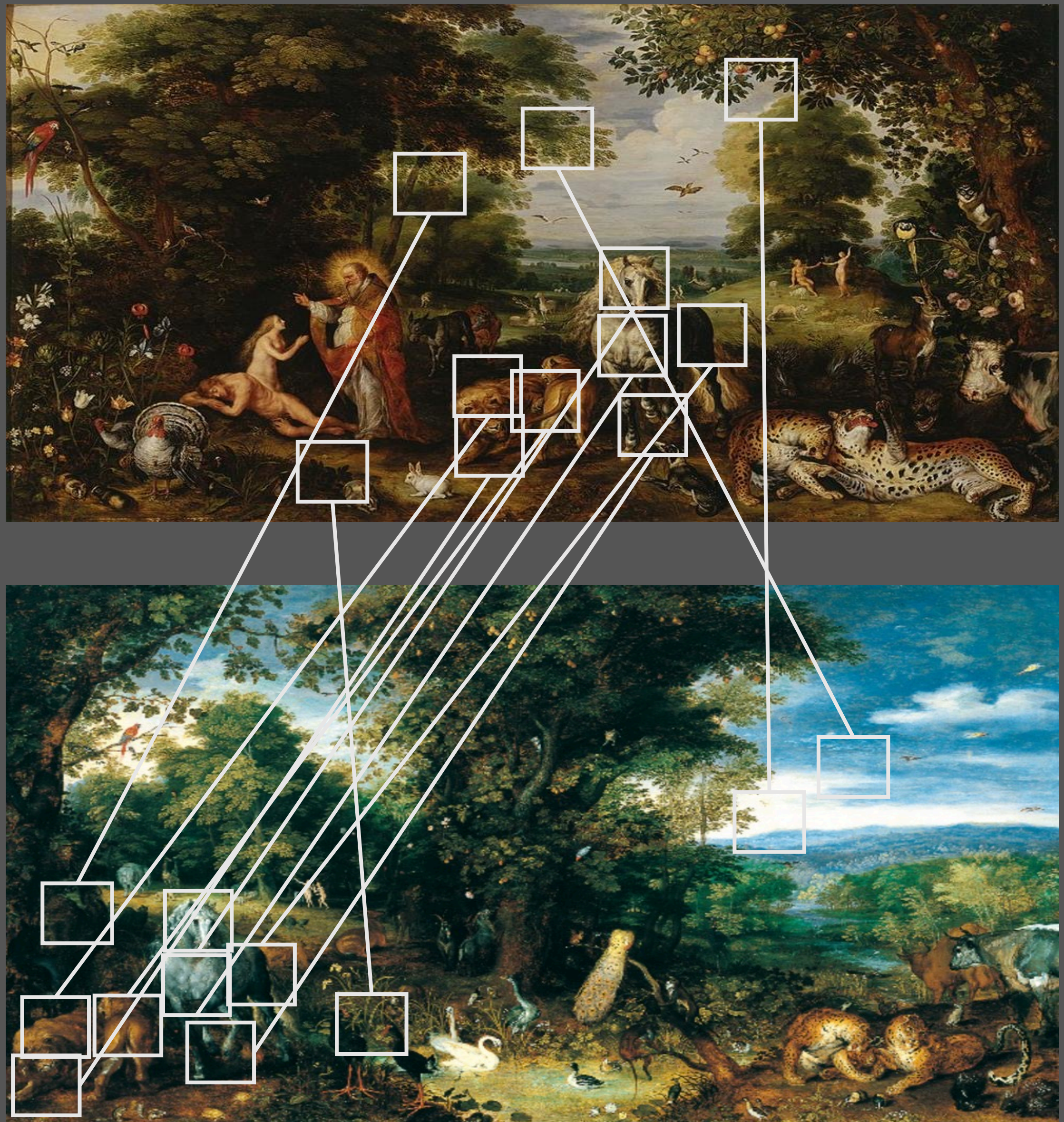
Visual results

ImageNet Feature

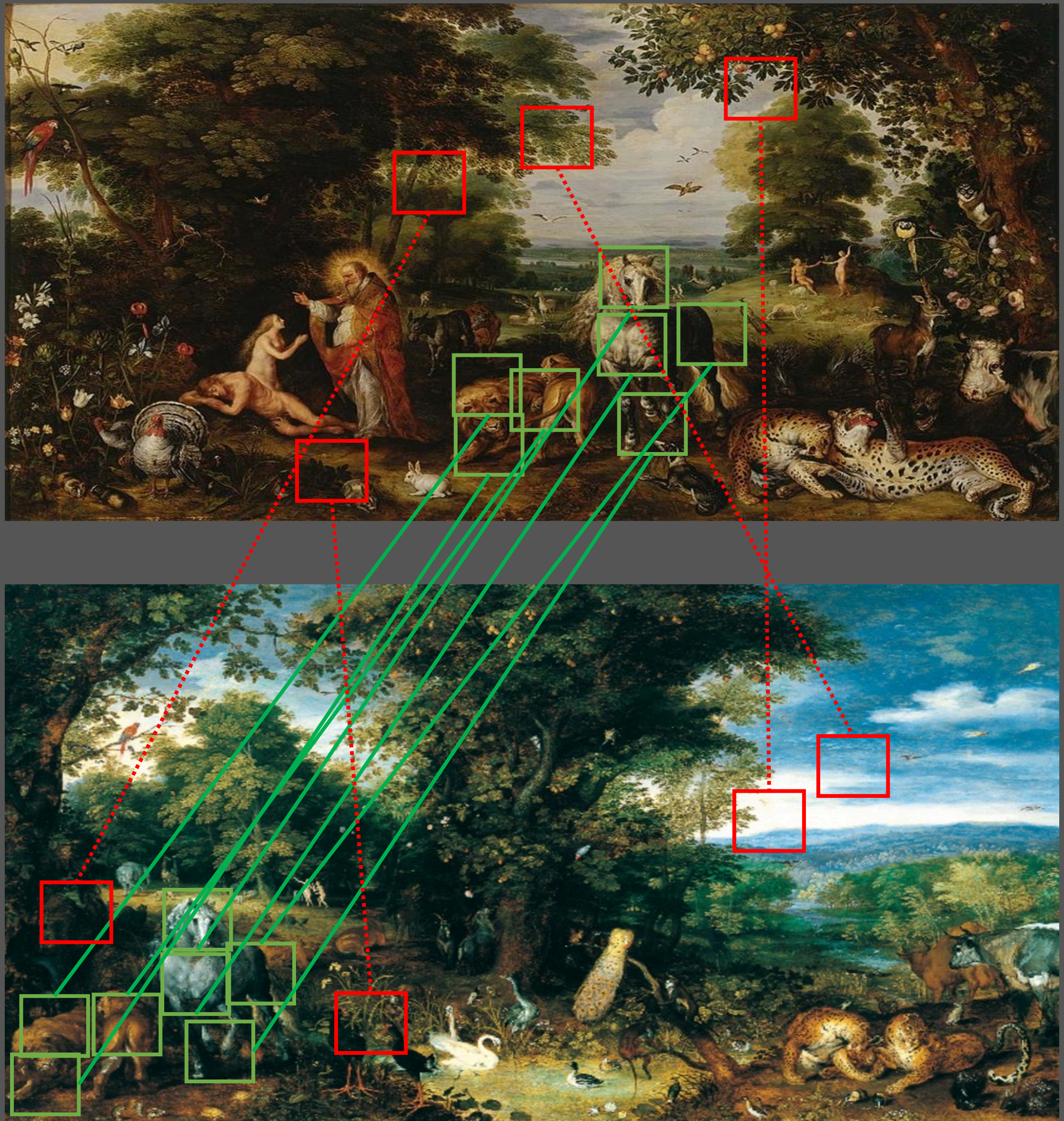
Our Feature



Discovery score



Discovery score



Discovery score

$$S(I_1, I_2) = \sum_{i \in \mathcal{I}} e^{-\frac{\|x_1^i - \mathcal{T}(x_2^i)\|^2}{2\sigma^2}} \boxed{s(f_1^i, f_2^i)}$$

- $\mathcal{I}$: inlier set
- $\mathtt{I}_1, \mathtt{I}_2$: pair of regions
- $\mathcal{T}$: geometric transformation
- $\mathtt{s}$: similarity metric
- $\mathtt{f}_1^{\mathtt{i}}, \mathtt{f}_2^{\mathtt{i}}$: features of i_{th} correspondence
- $\mathtt{x}_1^{\mathtt{i}}, \mathtt{x}_2^{\mathtt{i}}$: positions of i_{th} correspondence

Discovery score

$$S(I_1, I_2) = \sum_{i \in \mathcal{I}} \boxed{e^{-\frac{\|x_1^i - \mathcal{T}(x_2^i)\|^2}{2\sigma^2}}} s(f_1^i, f_2^i)$$

- $\mathcal{I}$: inlier set
- $\mathtt{I}_1, \mathtt{I}_2$: pair of regions
- $\mathcal{T}$: geometric transformation
- $\mathtt{s}$: similarity metric
- $\mathtt{f}_1^{\mathtt{i}}, \mathtt{f}_2^{\mathtt{i}}$: features of i_{th} correspondence
- $\mathtt{x}_1^{\mathtt{i}}, \mathtt{x}_2^{\mathtt{i}}$: positions of i_{th} correspondence

Discovery score

$$S(I_1, I_2) = \sum_{i \in \mathcal{I}} e^{-\frac{\|x_1^i - \mathcal{T}(x_2^i)\|^2}{2\sigma^2}} \boxed{s}(f_1^i, f_2^i)$$

- $\mathcal{I}$: inlier set
- $\mathtt{I}_1, \mathtt{I}_2$: pair of regions
- $\mathcal{T}$: geometric transformation
- $\boxed{\mathtt{s}\text{: similarity metric}}$
- $\mathtt{f}_1^{\mathtt{i}}, \mathtt{f}_2^{\mathtt{i}}$: features of i_{th} correspondence
- $\mathtt{x}_1^{\mathtt{i}}, \mathtt{x}_2^{\mathtt{i}}$: positions of i_{th} correspondence

Discovery

Discovery Score:

$$S(I_1, I_2) = \sum_{i \in \mathcal{I}} e^{-\frac{\|x_1^i - \boxed{\mathcal{T}}(x_2^i)\|^2}{2\sigma^2}} s(f_1^i, f_2^i)$$

- $\mathcal{I}$: inlier set
- $\mathtt{I}_1, \mathtt{I}_2$: pair of regions
- $\mathcal{T}$: geometric transformation
- $\mathtt{s}$: similarity metric
- $\mathtt{f}_1^{\mathtt{i}}, \mathtt{f}_2^{\mathtt{i}}$: features of i_{th} correspondence
- $\mathtt{x}_1^{\mathtt{i}}, \mathtt{x}_2^{\mathtt{i}}$: positions of i_{th} correspondence

ImageNet
Feature

Our
Feature

Discovery
+
our feature

Brueghel dataset

Results: one-shot detection on Brueghel

Feature \ Method	Cosine similarity	Discovery score
ImageNet pre-taining	58.0	54.8
Context Prediction [11]	58.8	64.29
Ours (trained on Brueghel)	**75.3**	**76.4**

Table 1: Experimental results on Brueghel, IoU > 0.3.

Results: one-shot detection on Brueghel

Feature \ Method	Cosine similarity	Discovery score
ImageNet pre-taining	58.0	54.8
Context Prediction [11]	58.8	64.29
Ours (trained on Brueghel)	**75.3**	**76.4**

Table 1: Experimental results on Brueghel, IoU > 0.3.

Results: one-shot detection on Brueghel

Feature \ Method	Cosine similarity	Discovery score
ImageNet pre-taining	58.0	54.8
Context Prediction [11]	58.8	64.29
Ours (trained on Brueghel)	**75.3**	**76.4**

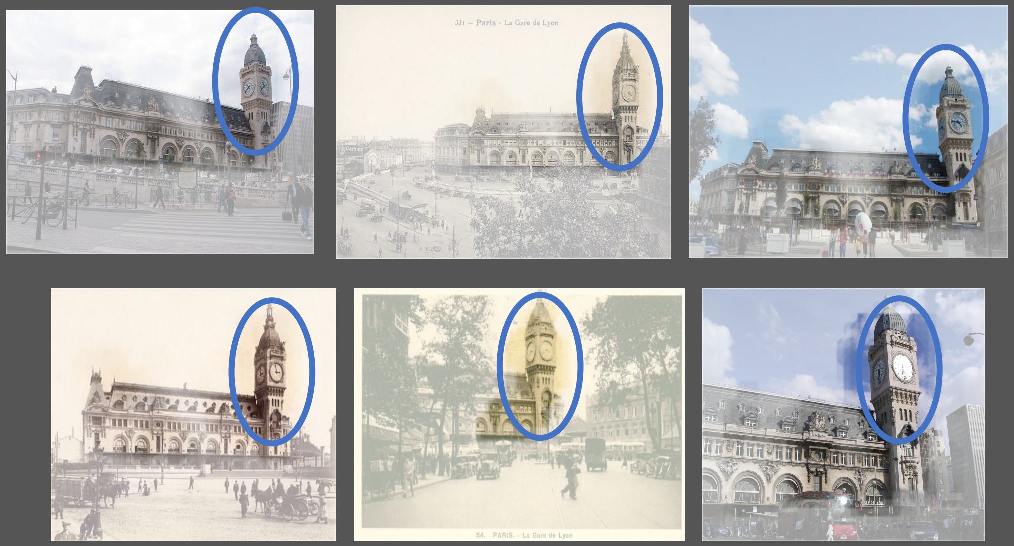
Table 1: Experimental results on Brueghel, IoU > 0.3.

Results of discovery

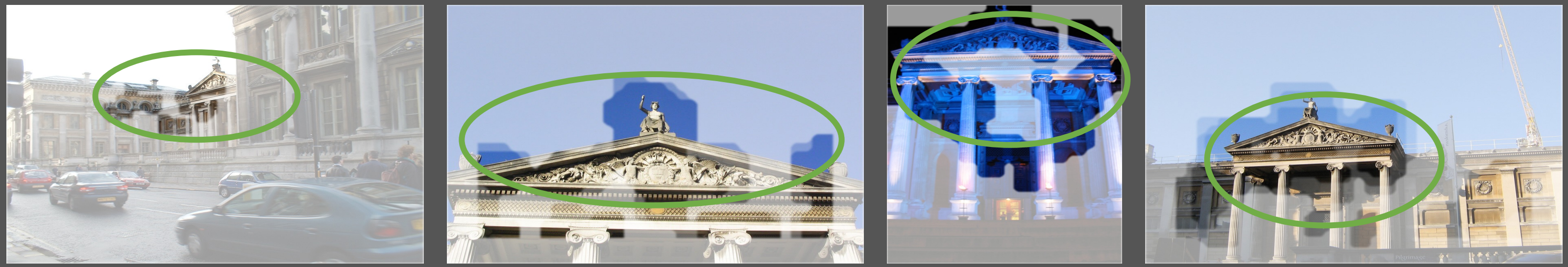
Other results: geo-localization

Method	LTLL (%)	Oxford (%)
B. Fernando et al.[16]	56.1	-
F. Radenović et al.[35]	-	87.8
ResNet18 max-pool, image level	59.8	14.0
ResNet18 + discovery	80.9	85.0
Ours (trained LTLL + discovery)	**88.5**	83.6
Ours (trained Oxford + discovery)	85.6	**85.7**

Table 2: Classification accuracy on LTLL and retrieval mAP on Oxford5K



Discovered group in LTLL



Discovered group in Oxford

Summary

- Annotations on the Brueghel to evaluate one-shot detection

- Self-supervised training strategy to learn style-invariant features

- Multi-scale feature matching to discover repeated patterns

Contributions in my thesis

1. Style-invariant feature from self-supervision

2. Co-segmentation from synthetic data

3. Dense image alignment from reconstruction

Task: co-segmentation in a pair of images

Input

predicted masks

Problem: no training data available

Key idea: synthetic pairs with duplicated patterns

Source and selected segment

Background

Key idea: direct copy-paste

Direct
Copy-paste

Key idea: our blending

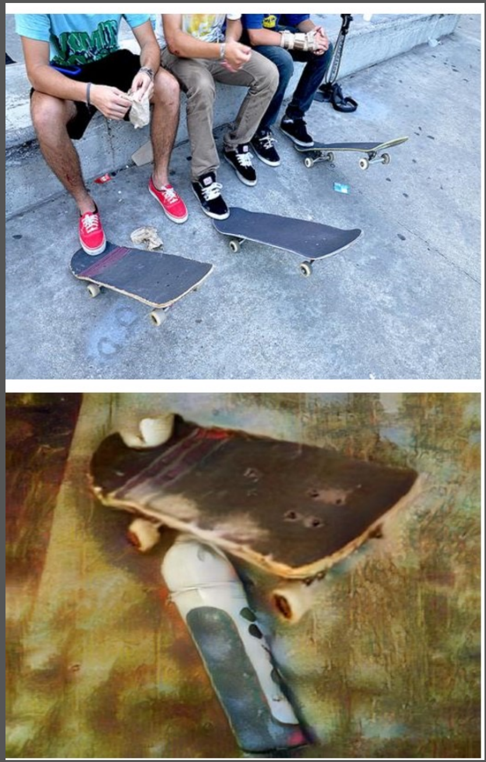
Our blending

Poisson blending [Pérez et al. 2003]

\+

Style transfer [Huang and Belongie 2017]

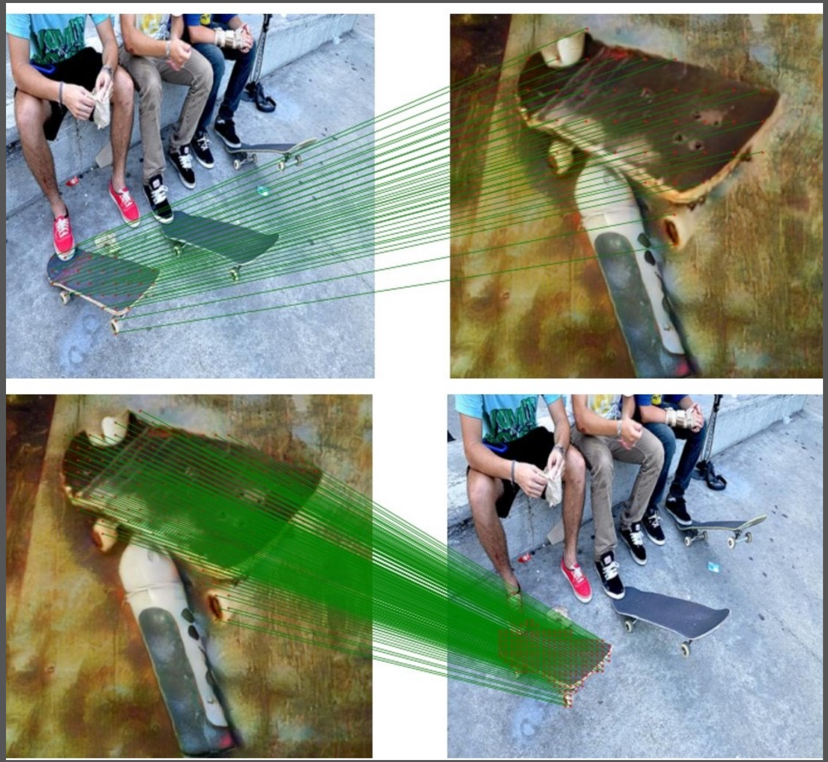
Annotations



Generated images

Annotations

Masks



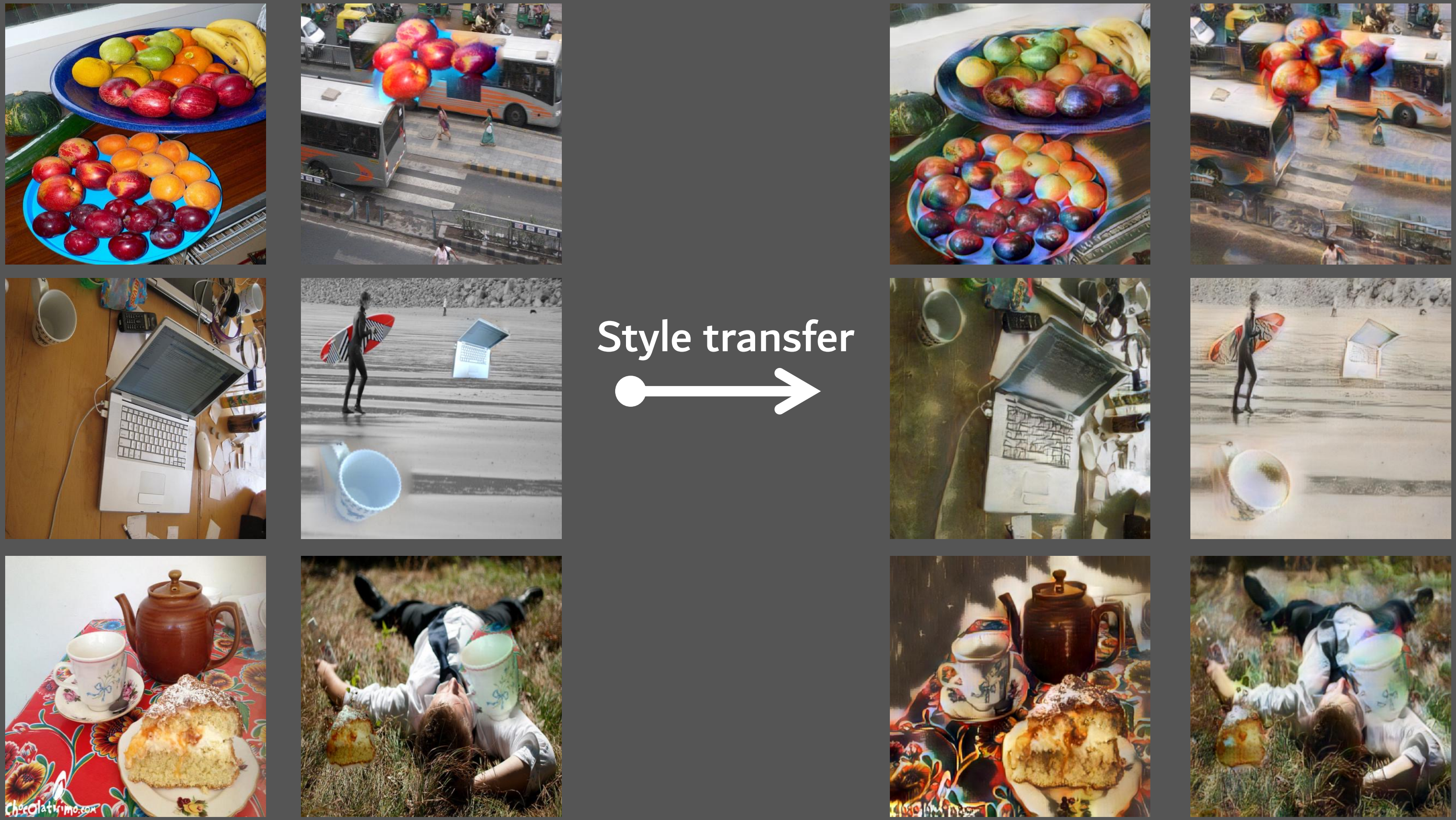
Correspondences

Training data

One object
COCO Seg.

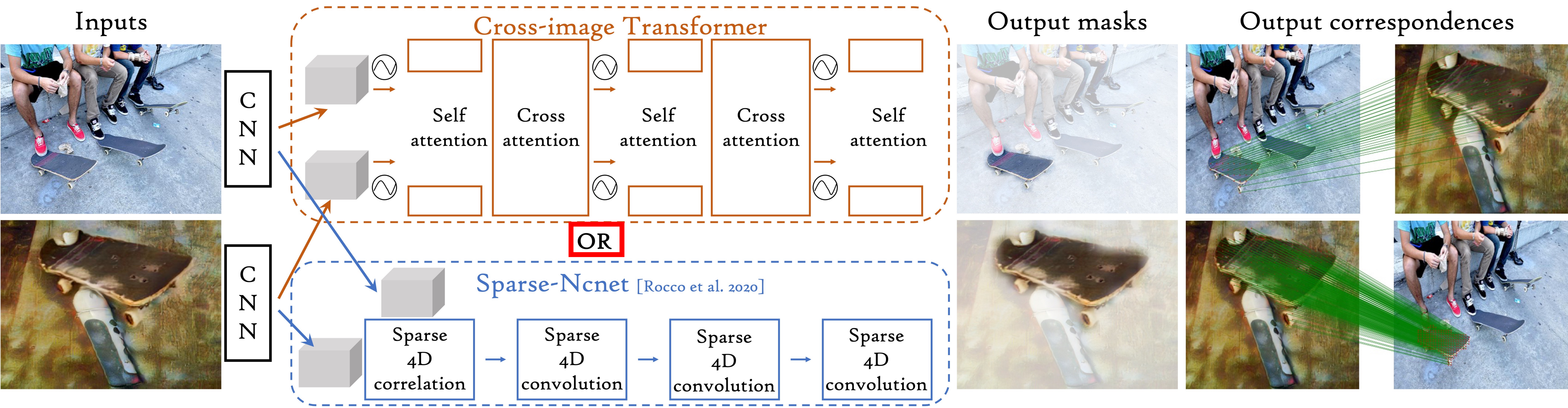
Two objects
COCO Seg.

Two objects
Unsup. Seg.



Source Blended Source Blended

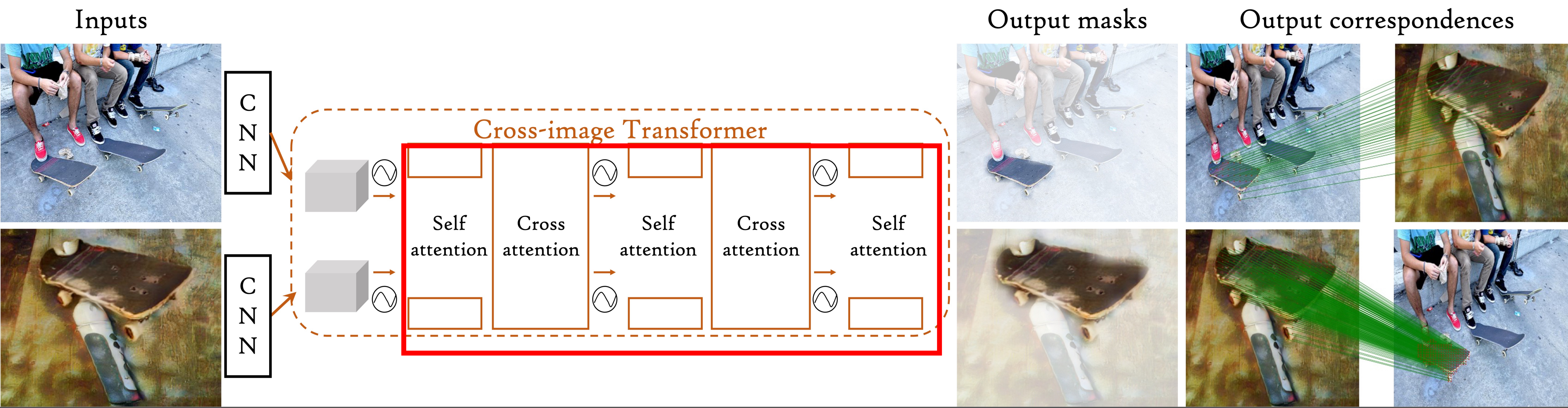
Key idea: learning co-segmentation



Objective function:

$$\mathcal{L}^s_{sup} = \underbrace{CE(\mathbf{M}^s_{gt}, \mathbf{M}^s)}_{\mathcal{L}_{mask}} + \underbrace{CE(\mathbf{M}^s_{gt}, \mathbf{M}^t(\mathbf{C}^{s\to t}))}_{\mathcal{L}_{tmask}} + \underbrace{\eta \frac{1}{\sum_{i,j} \mathbf{M}^s_{gt}(i,j)} \sum_{i,j} \mathbf{M}^s_{gt}(i,j) \|\mathbf{C}^{s\to t}(i,j) - \mathbf{C}^{s\to t}_{gt}(i,j)\|}_{\mathcal{L}_{corr}}$$

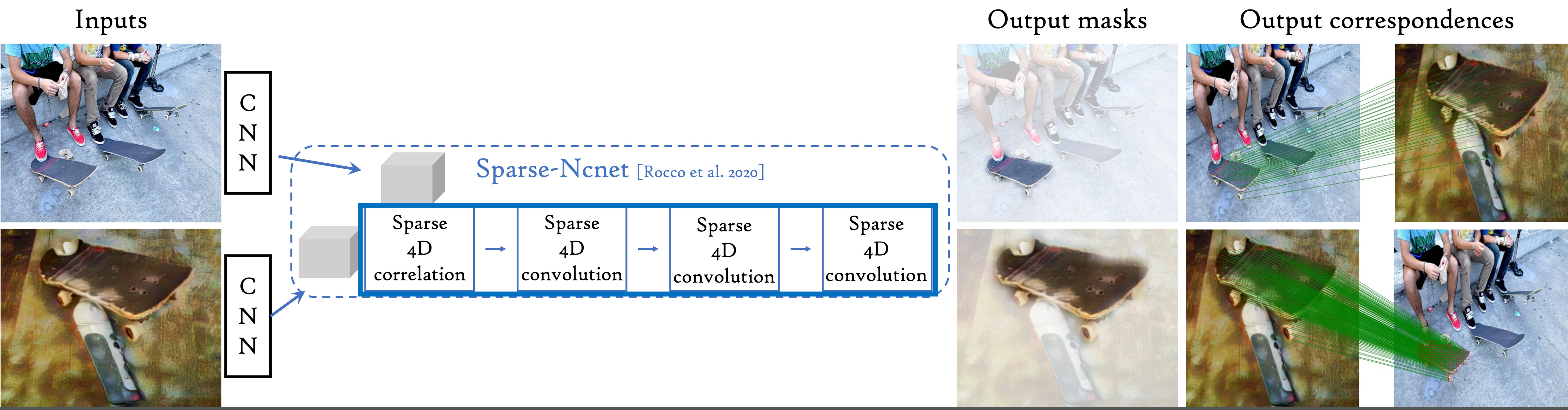
Key idea: learning co-segmentation



Objective function:

$$\mathcal{L}^s_{sup} = \underbrace{CE(\mathbf{M}^s_{gt}, \mathbf{M}^s)}_{\mathcal{L}_{mask}} + \underbrace{CE(\mathbf{M}^s_{gt}, \mathbf{M}^t(\mathbf{C}^{s\to t}))}_{\mathcal{L}_{tmask}} + \underbrace{\eta \frac{1}{\sum_{i,j} \mathbf{M}^s_{gt}(i,j)} \sum_{i,j} \mathbf{M}^s_{gt}(i,j) \| \mathbf{C}^{s\to t}(i,j) - \mathbf{C}^{s\to t}_{gt}(i,j) \|}_{\mathcal{L}_{corr}}$$

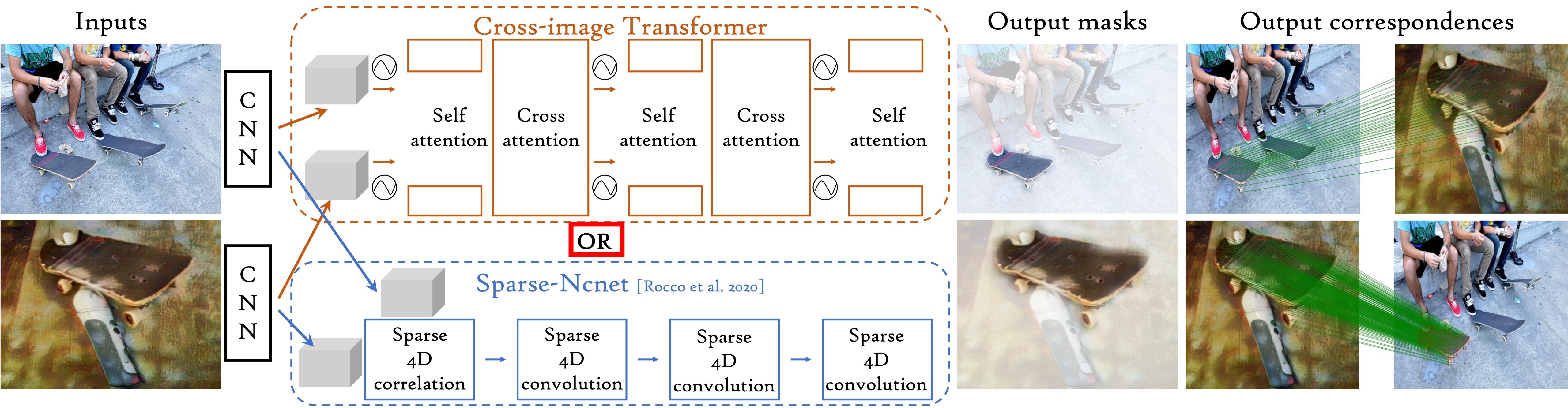
Key idea: learning co-segmentation



Objective function:

$$\mathcal{L}^s_{sup} = \underbrace{CE(\mathbf{M}^s_{gt}, \mathbf{M}^s)}_{\mathcal{L}_{mask}} + \underbrace{CE(\mathbf{M}^s_{gt}, \mathbf{M}^t(\mathbf{C}^{s\to t}))}_{\mathcal{L}_{tmask}} + \underbrace{\eta \frac{1}{\sum_{i,j} \mathbf{M}^s_{gt}(i,j)} \sum_{i,j} \mathbf{M}^s_{gt}(i,j) \| \mathbf{C}^{s\to t}(i,j) - \mathbf{C}^{s\to t}_{gt}(i,j) \|}_{\mathcal{L}_{corr}}$$

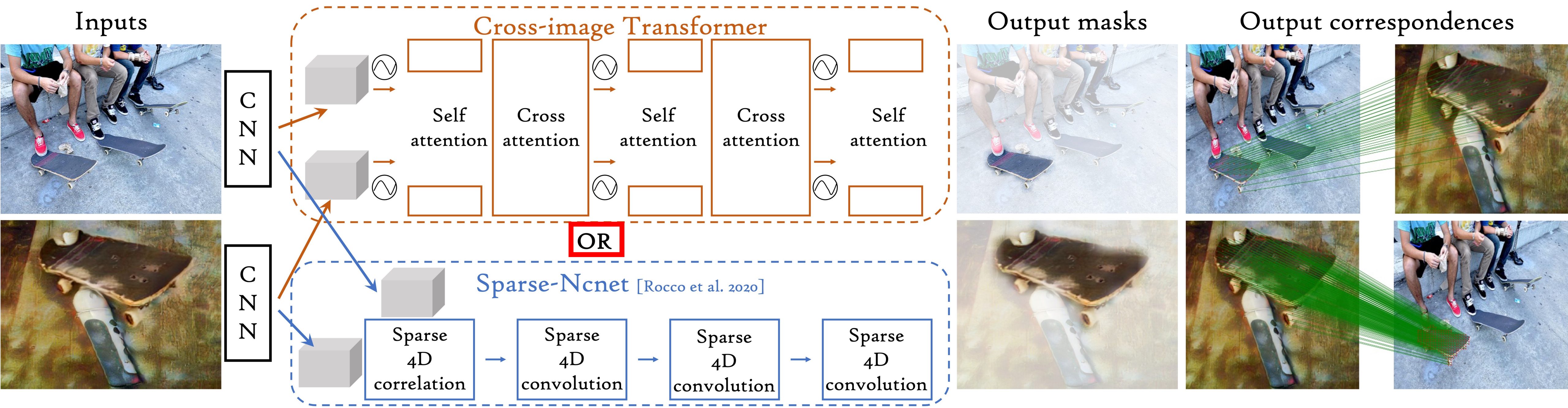
Key idea: learning co-segmentation



Objective function:

$$\mathcal{L}^s_{sup} = \boxed{\underbrace{CE(\mathbf{M}^s_{gt}, \mathbf{M}^s)}_{\mathcal{L}_{mask}}} + \underbrace{CE(\mathbf{M}^s_{gt}, \mathbf{M}^t(\mathbf{C}^{s\to t}))}_{\mathcal{L}_{tmask}} + \underbrace{\eta \frac{1}{\sum_{i,j}\mathbf{M}^s_{gt}(i,j)} \sum_{i,j} \mathbf{M}^s_{gt}(i,j)\|\mathbf{C}^{s\to t}(i,j) - \mathbf{C}^{s\to t}_{gt}(i,j)\|}_{\mathcal{L}_{corr}}$$

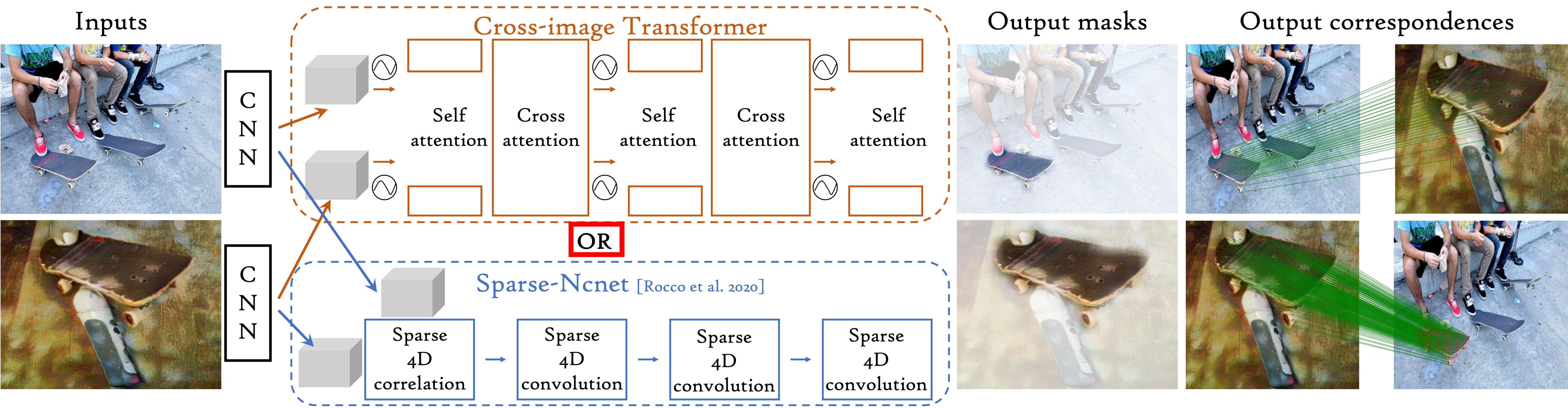
Key idea: learning co-segmentation



Objective function:

$$\mathcal{L}^s_{sup} = \underbrace{CE(\mathbf{M}^s_{gt}, \mathbf{M}^s)}_{\mathcal{L}_{mask}} + \underbrace{CE(\mathbf{M}^s_{gt}, \mathbf{M}^t(\mathbf{C}^{s\to t}))}_{\mathcal{L}_{tmask}} + \boxed{\underbrace{\eta \frac{1}{\sum_{i,j} \mathbf{M}^s_{gt}(i,j)} \sum_{i,j} \mathbf{M}^s_{gt}(i,j) \|\mathbf{C}^{s\to t}(i,j) - \mathbf{C}^{s\to t}_{gt}(i,j)\|}_{\mathcal{L}_{corr}}}$$

Key idea: learning co-segmentation



Objective function:

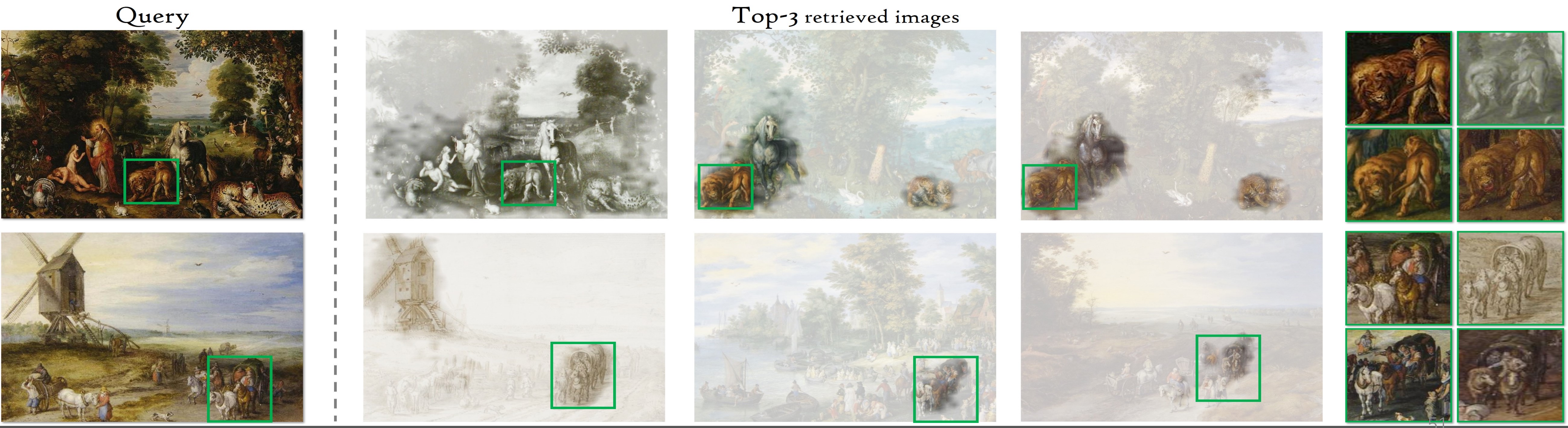
$$\mathcal{L}^s_{sup} = \underbrace{CE(\mathbf{M}^s_{gt}, \mathbf{M}^s)}_{\mathcal{L}_{mask}} + \boxed{\underbrace{CE(\mathbf{M}^s_{gt}, \mathbf{M}^t(\mathbf{C}^{s\to t}))}_{\mathcal{L}_{tmask}}} + \underbrace{\eta \frac{1}{\sum_{i,j} \mathbf{M}^s_{gt}(i,j)} \sum_{i,j} \mathbf{M}^s_{gt}(i,j) \|\mathbf{C}^{s\to t}(i,j) - \mathbf{C}^{s\to t}_{gt}(i,j)\|}_{\mathcal{L}_{corr}}$$

Experiments: one-shot detection on Brueghel

Score between a pair of images

$$\mathcal{S}(\mathbf{I}^s, \mathbf{I}^t) = \sum_{i,j} \underbrace{\mathbf{M}^s_{joint}(i,j)}_{\text{Mask}} \boxed{\underbrace{cos(\mathbf{F}^s(i,j), \mathbf{F}^t(\mathbf{C}^{s \to t}(i,j)))}_{\text{Feat. similarity}}}$$

$$\mathbf{M}^s_{joint}(i,j) = \mathbf{M}^t(\mathbf{C}^{s \to t}(i,j))\mathbf{M}^s(i,j)$$



Experiments: one-shot detection on Brueghel

Feat. + Methods	mAP	
	Retrieval	Det.(IoU > 0.3)
Shen et al. [53] + cos [53]	75.5	75.3
Shen et al. [53] + discovery [53]	76.6	76.4
MocoV2 [8] + cos [53]	79.0	78.7
MocoV2 [8] + discovery [53]	80.8	79.6
Ours + Unsupervised segments		
Transformer	81.8	79.4
Sparse-Ncnet	82.8	73.4
Ours + COCO segments [35]		
Transformer	**84.4**	**81.8**
Sparse-Ncnet	83.3	73.7

Table 1. Art detail retrieval and detection on Brueghel [53]. For detection, we employ ArtMiner (Brueghel [53] + cos [53]) as a post-processing and reports results with IoU > 0.3 [53]

Experiments: one-shot detection on Brueghel

Feat. + Methods	mAP Retrieval	mAP Det.(IoU > 0.3)
Shen et al. [53] + cos [53]	75.5	75.3
Shen et al. [53] + discovery [53]	76.6	76.4
MocoV2 [8] + cos [53]	79.0	78.7
MocoV2 [8] + discovery [53]	80.8	79.6
Ours + Unsupervised segments		
Transformer	81.8	79.4
Sparse-Ncnet	82.8	73.4
Ours + COCO segments [35]		
Transformer	**84.4**	**81.8**
Sparse-Ncnet	83.3	73.7

Table 1. Art detail retrieval and detection on Brueghel [53]. For detection, we employ ArtMiner (Brueghel [53] + cos [53]) as a post-processing and reports results with IoU > 0.3 [53]

Discovery on Brueghel: Correspondences graph

Correspondences $\mathcal{V} = \{v_1, v_2, \ldots, v_i, \ldots, v_j, \ldots\}$

Graph: $\mathcal{G} = (\mathcal{V}, \mathcal{E})$

Mask Pred.

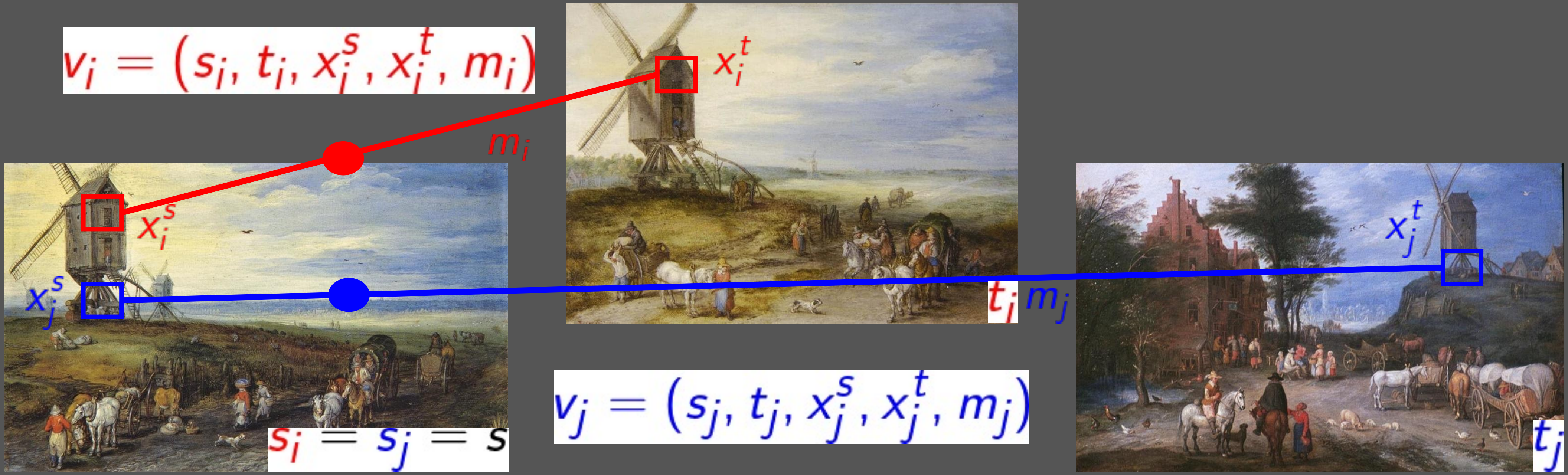
$$v_i = (\underbrace{s_i, t_i}_{\text{Indexes of source and targe images}}, \overbrace{x_i^s, x_i^t}^{\text{Corr. Pred.}}, m_i)$$

Discovery on Brueghel: Correspondences graph

3-cycle Consistency $\mathcal{E} = \{e_{i,j}\}$

Graph: $\mathcal{G} = (\mathcal{V}, \mathcal{E})$

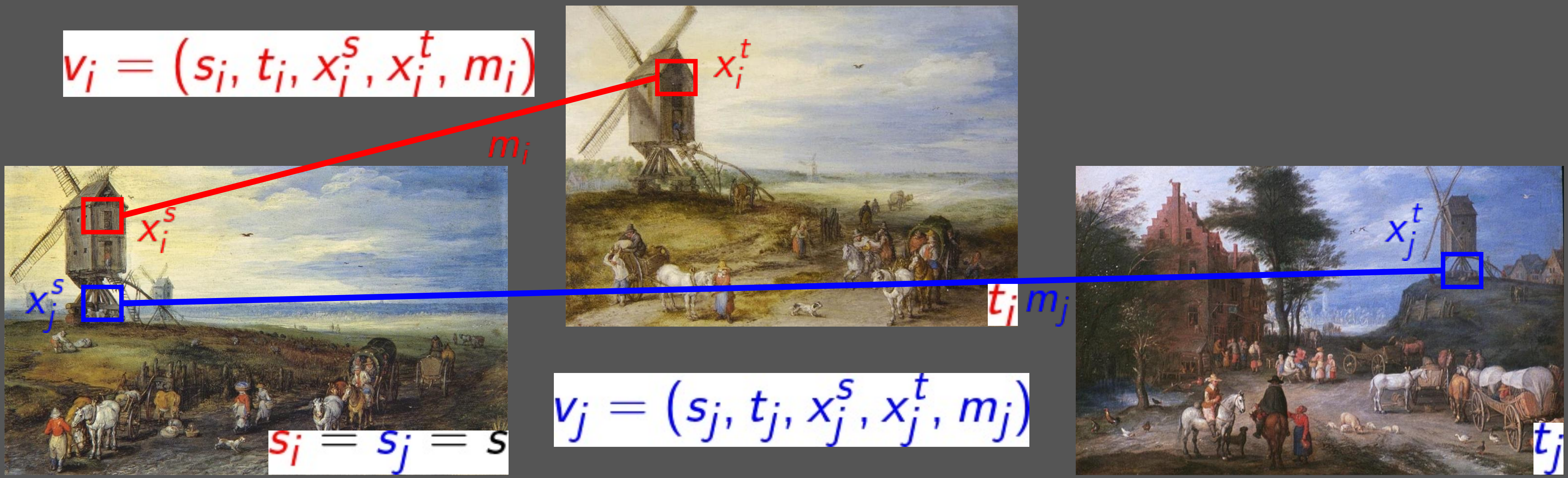
$v_i = (s_i, t_i, x_i^s, x_i^t, m_i)$



$v_j = (s_j, t_j, x_j^s, x_j^t, m_j)$

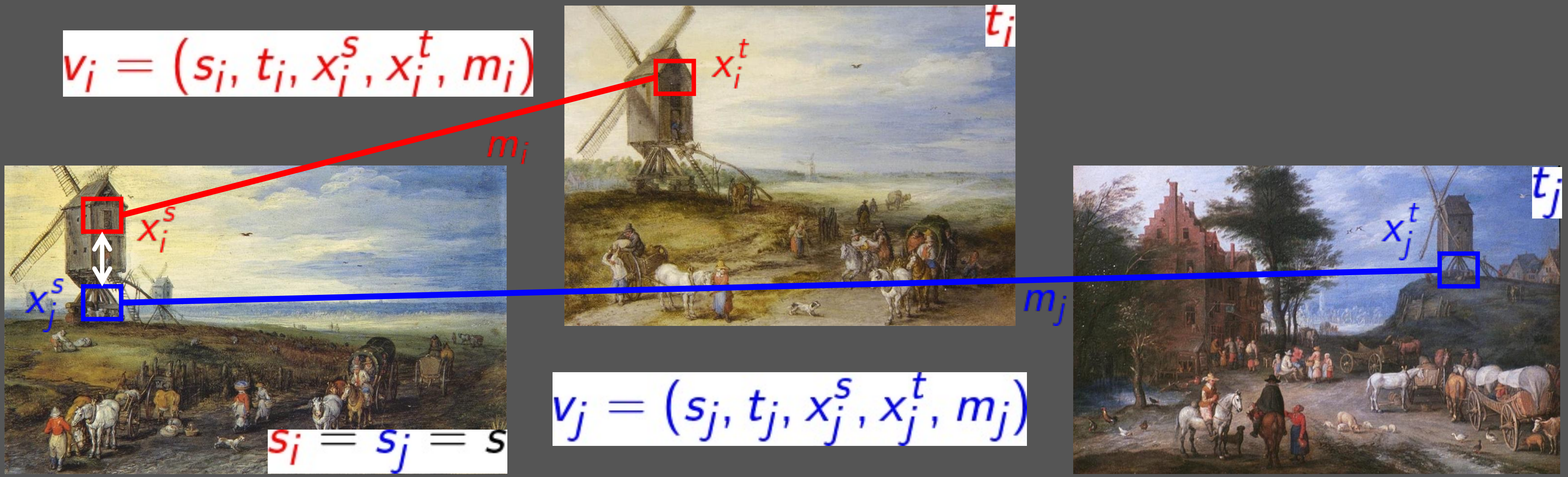
Discovery on Brueghel: Correspondences graph

$$\frac{1}{2}\boxed{m_i m_j}\exp\left(\frac{||x_i^s - x_j^s||}{\sigma}\right)\left[\exp\left(\frac{||x_i^t - C^{t_j \to t_i}(x_j^t)||}{\sigma}\right) + \exp\left(\frac{||x_j^t - C^{t_i \to t_j}(x_i^t)||}{\sigma}\right)\right]$$



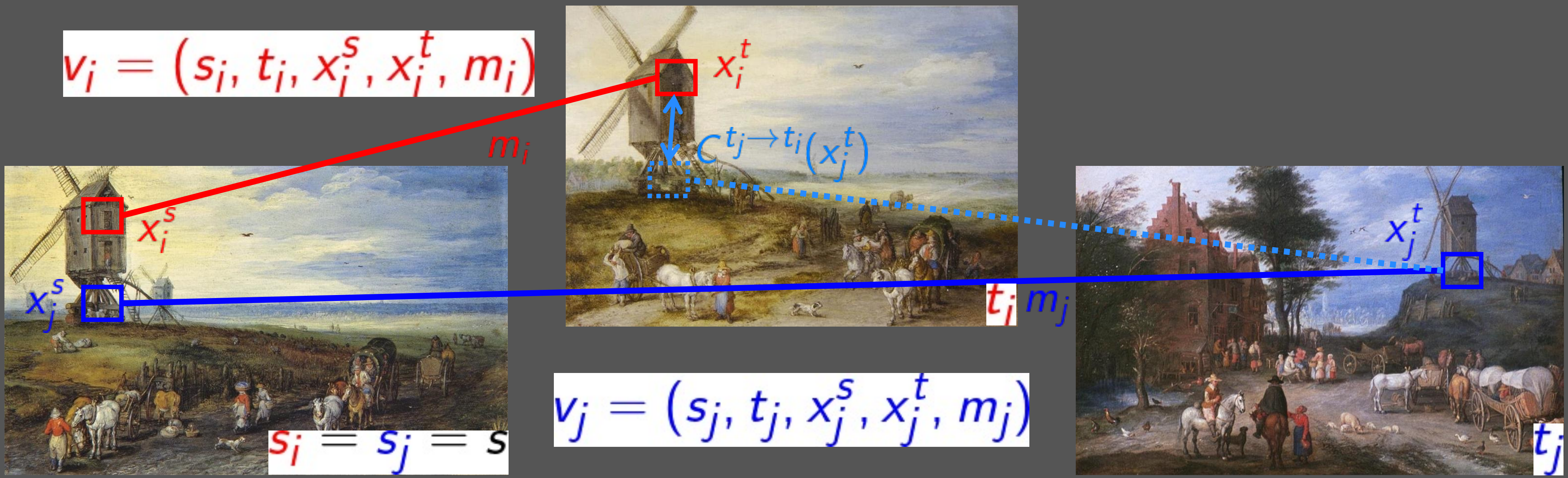
Discovery on Brueghel: Correspondences graph

$$\frac{1}{2} m_i m_j \boxed{\exp\left(\frac{||x_i^s - x_j^s||}{\sigma}\right)}\left[\exp\left(\frac{||x_i^t - C^{t_j \to t_i}(x_j^t)||}{\sigma}\right) + \exp\left(\frac{||x_j^t - C^{t_i \to t_j}(x_i^t)||}{\sigma}\right)\right]$$



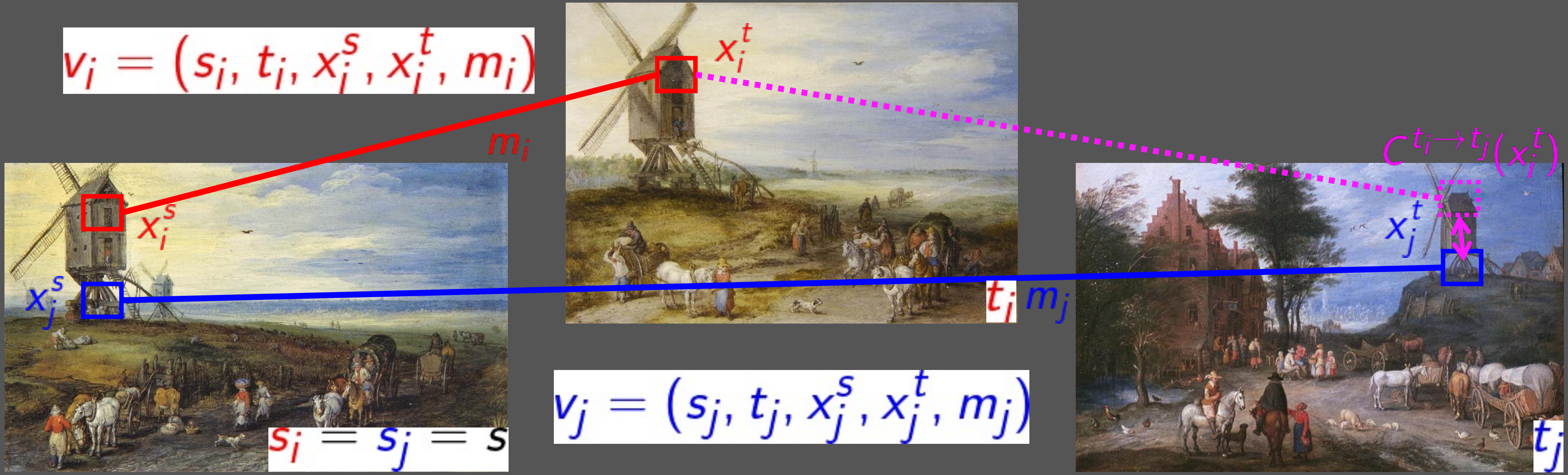
Discovery on Brueghel: Correspondences graph

$$\frac{1}{2} m_i m_j \exp\left(\frac{\|x_i^s - x_j^s\|}{\sigma}\right)\left[\boxed{\exp\left(\frac{\|x_i^t - C^{t_j \to t_i}(x_j^t)\|}{\sigma}\right)} + \exp\left(\frac{\|x_j^t - C^{t_i \to t_j}(x_i^t)\|}{\sigma}\right)\right]$$

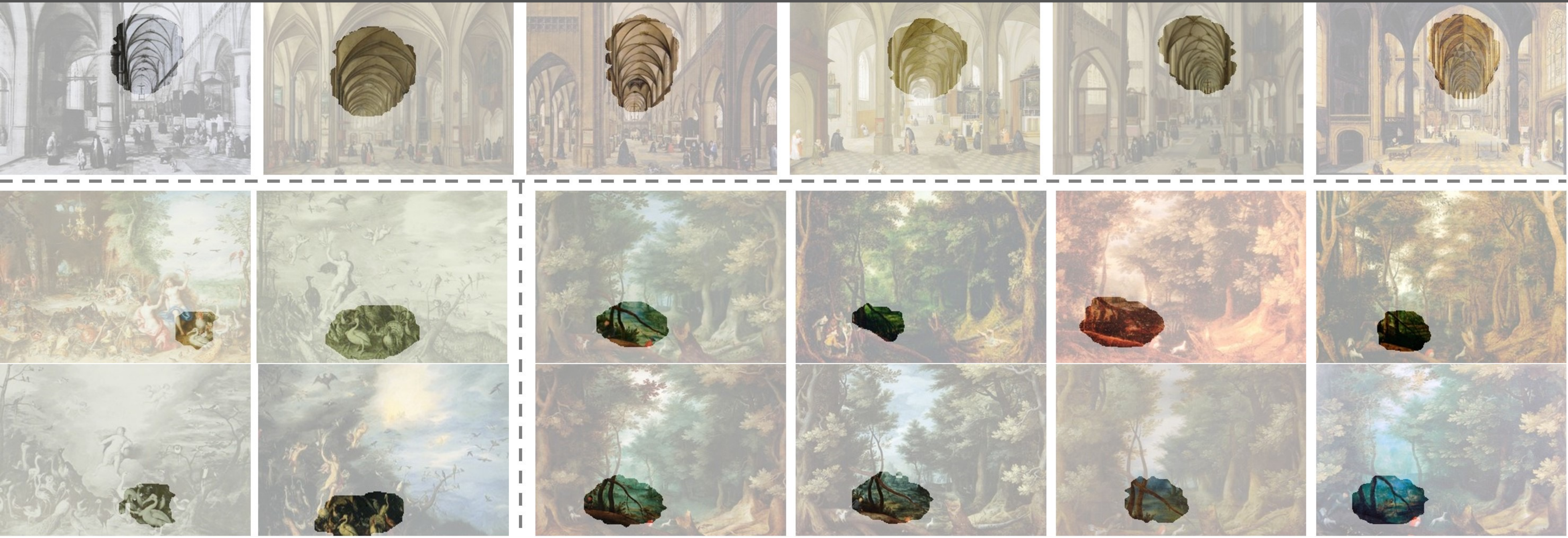


Discovery on Brueghel: Correspondences graph

$$\frac{1}{2} m_i m_j \exp\left(\frac{||x_i^s - x_j^s||}{\sigma}\right)\left[\exp\left(\frac{||x_i^t - C^{t_j \to t_i}(x_j^t)||}{\sigma}\right) + \boxed{\exp\left(\frac{||x_j^t - C^{t_i \to t_j}(x_i^t)||}{\sigma}\right)}\right]$$



Experiments: discovery on Brueghel [Shen et al. 2019]

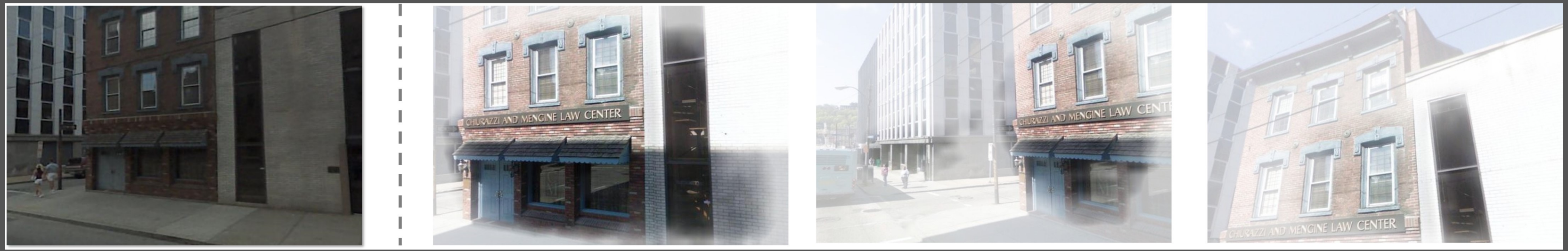


Other results

Discovery on the dataset of [Rubinstein et al. 2013]

Place recognition Tokyo24/7 [Torii et al. 2015]

Place recognition Pitts30K [Torri et al. 2013]



Summary

- Learning co-segmentation from synthetic pairs

- Discovering patterns using the correspondence graph

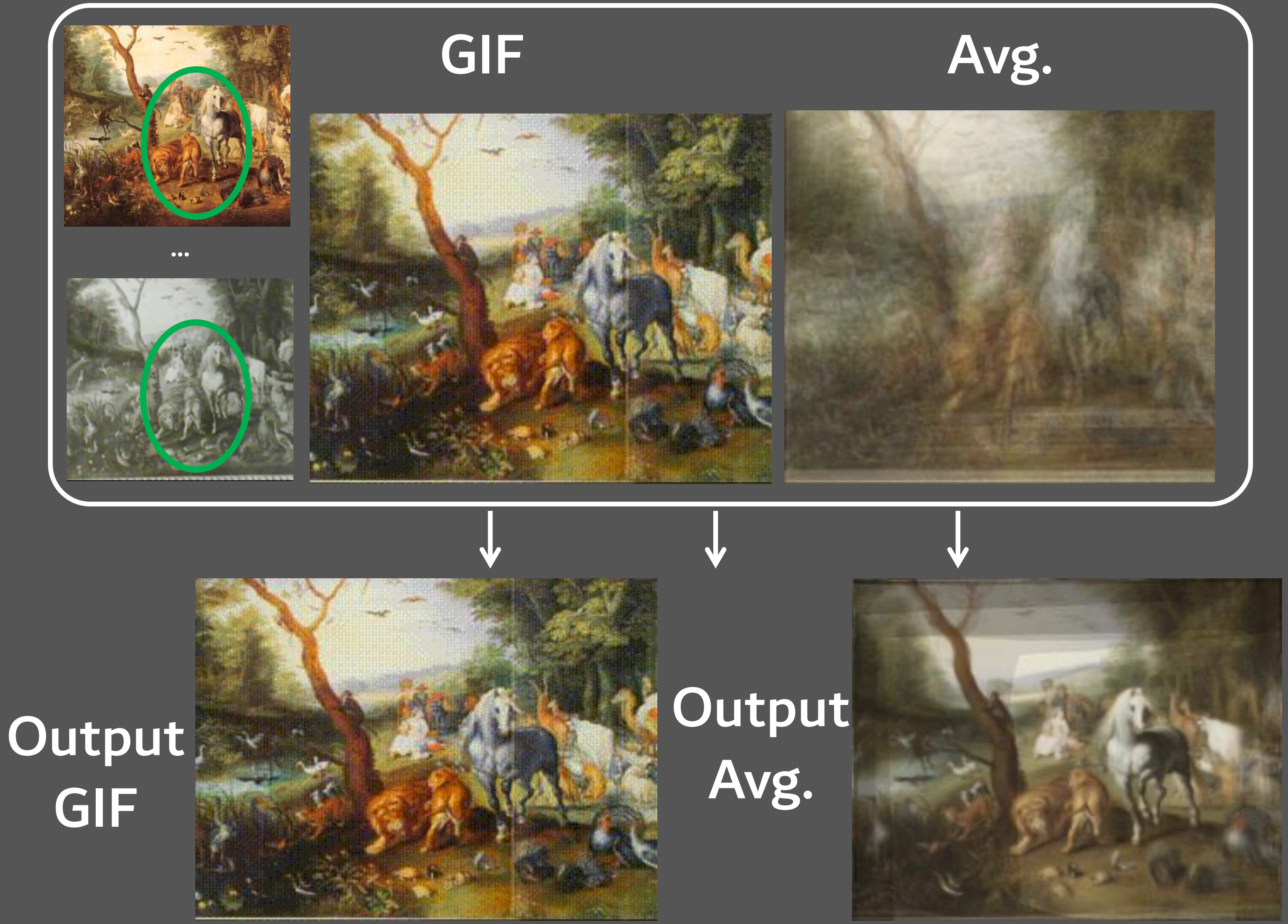
Contributions in my thesis

1. Style-invariant feature from self-supervision

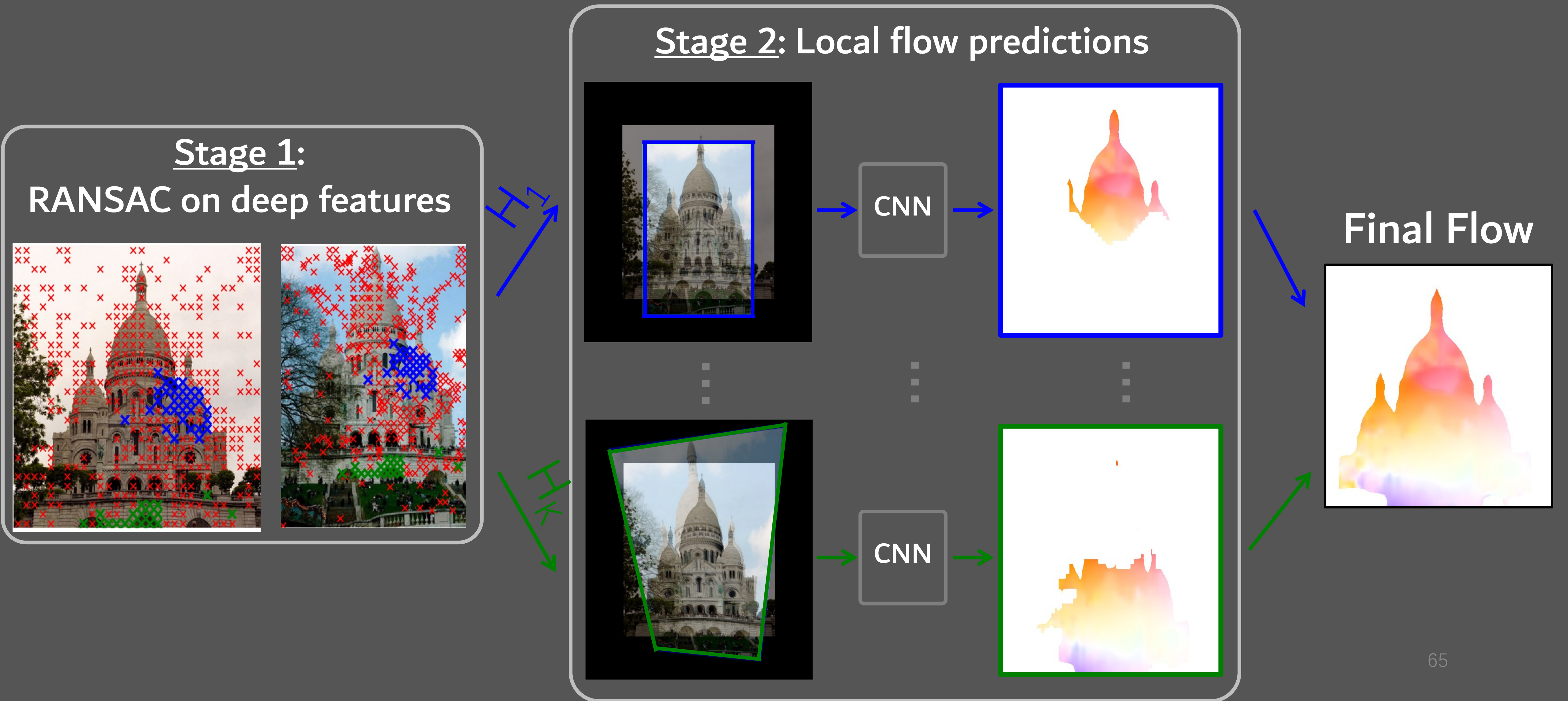
2. Co-segmentation from synthetic data

3. Dense image alignment from reconstruction

Problem: generic image alignment



Key idea: an unsupervised two-stage method

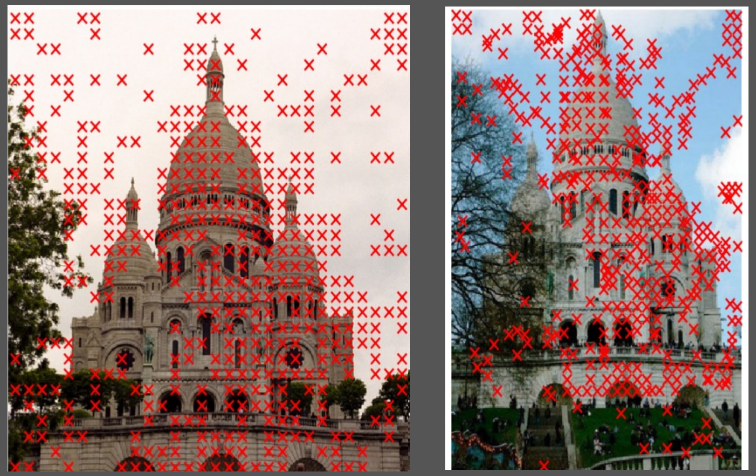


Key idea: an unsupervised two-stage method

Stage 1:
RANSAC on deep features

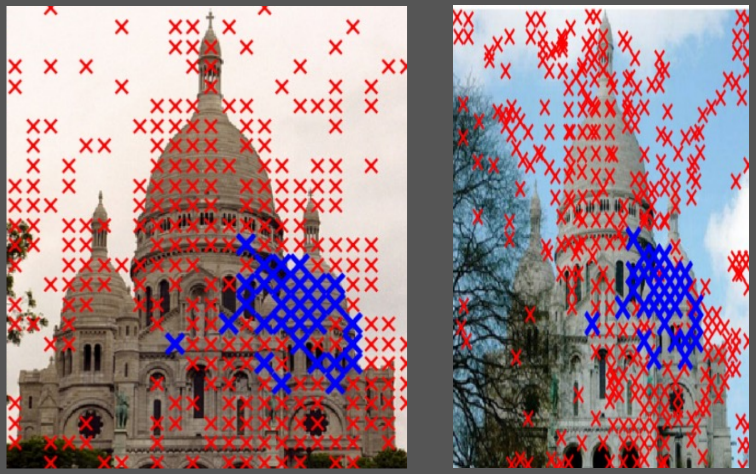
Key idea: an unsupervised two-stage method

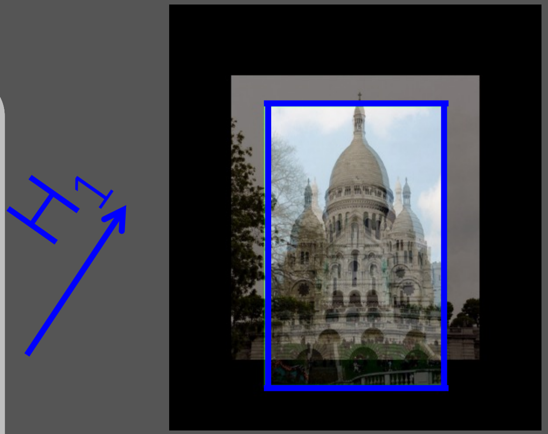
Stage 1:
RANSAC on deep features



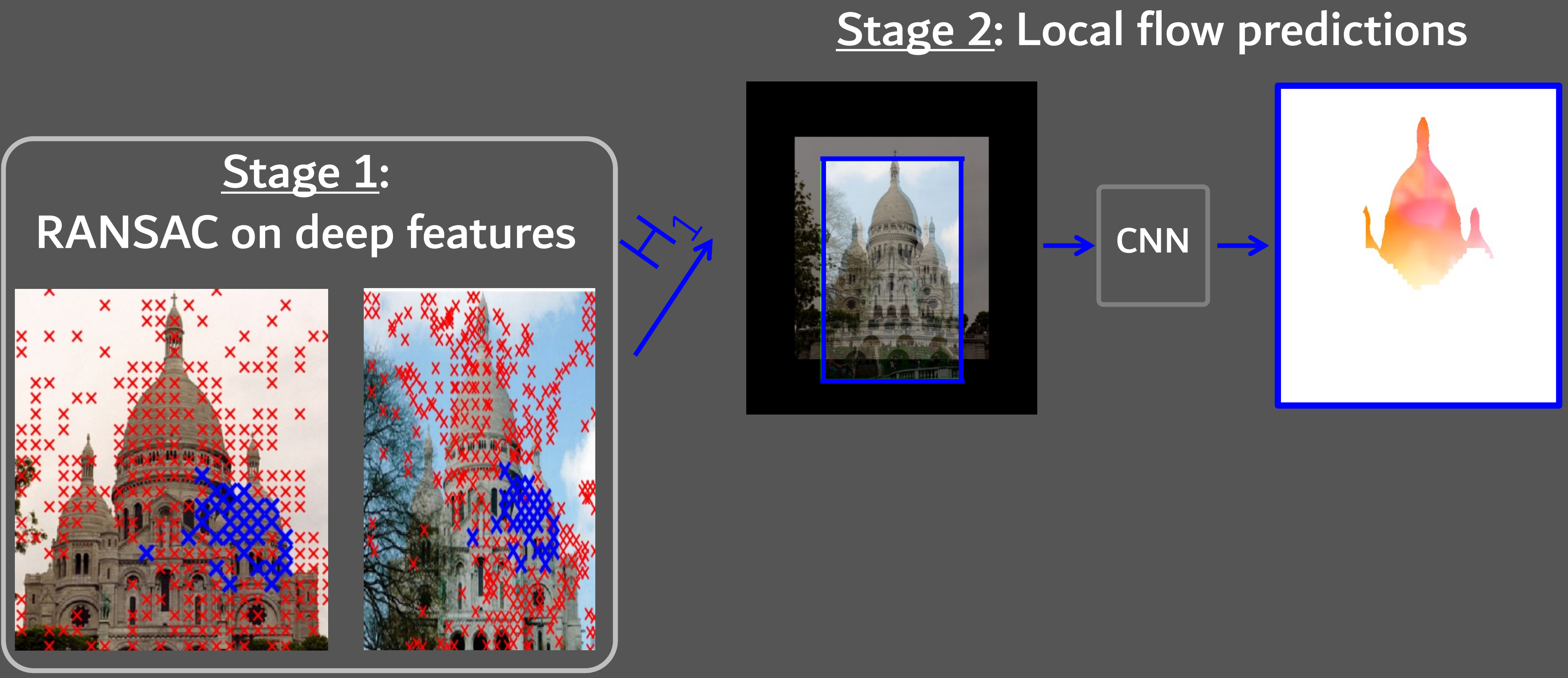
Key idea: an unsupervised two-stage method

Stage 1:
RANSAC on deep features





Key idea: an unsupervised two-stage method

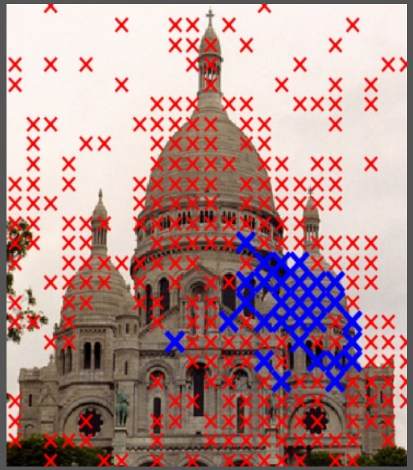


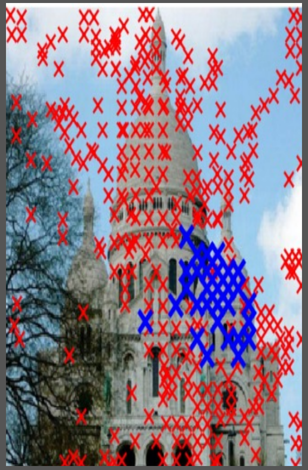
Key idea: an unsupervised two-stage method

SSIM + mask + cycle-consistency loss

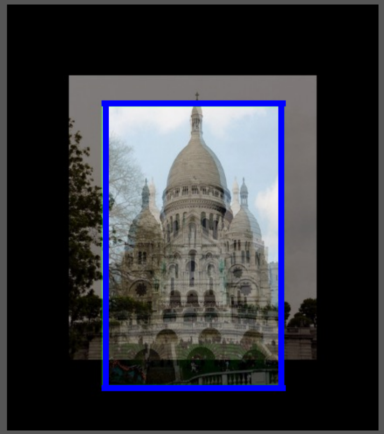
Stage 2: Local flow predictions

Stage 1:
RANSAC on deep features

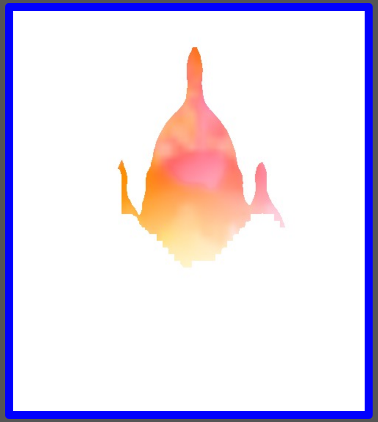




H^{-1}



CNN



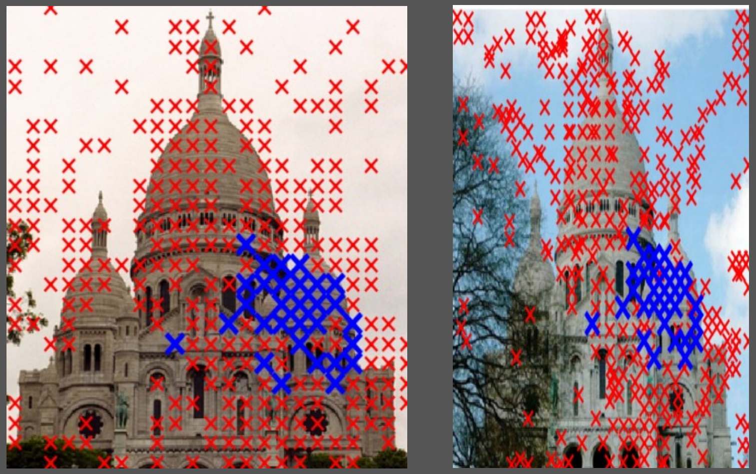
$$\underbrace{\mathcal{L}_m(I_s, I_t)}_{\text{Mask loss}} = \sum_{(x,y)\in I_t} \left|\underbrace{M_t^{cycle}(x,y)}_{\text{Confidence at (x,y)}} - 1\right|$$

Key idea: an unsupervised two-stage method

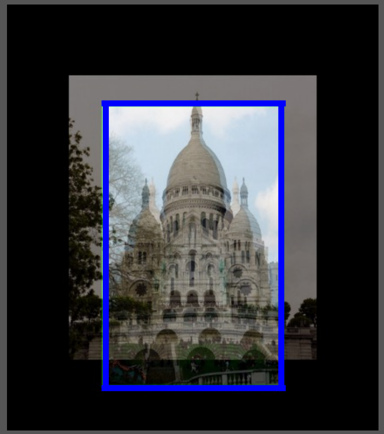
SSIM + mask + cycle-consistency loss

Stage 2: Local flow predictions

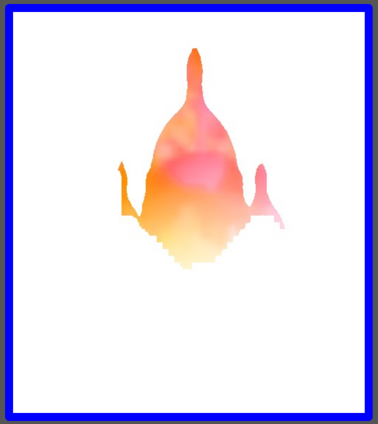
Stage 1:
RANSAC on deep features



H^{-1}



→ CNN →

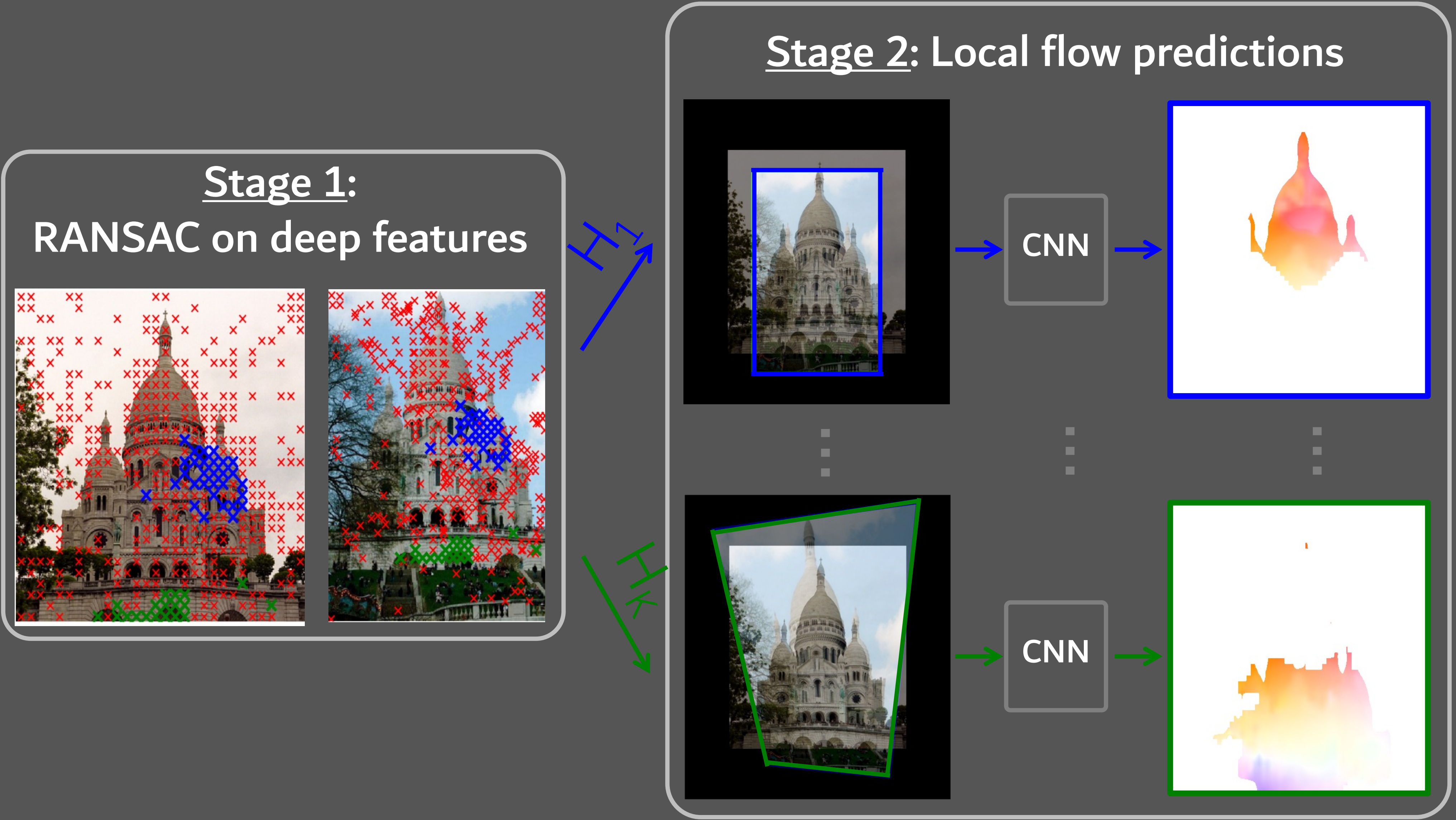


$$\mathcal{L}_{rec}^{SSIM}(I_s, I_t) = \sum_{(x,y)\in I_t} \underbrace{M_t^{cycle}(x,y)}\left(1 - SSIM\left(I_s(x',y'), I_t(x,y)\right)\right)$$

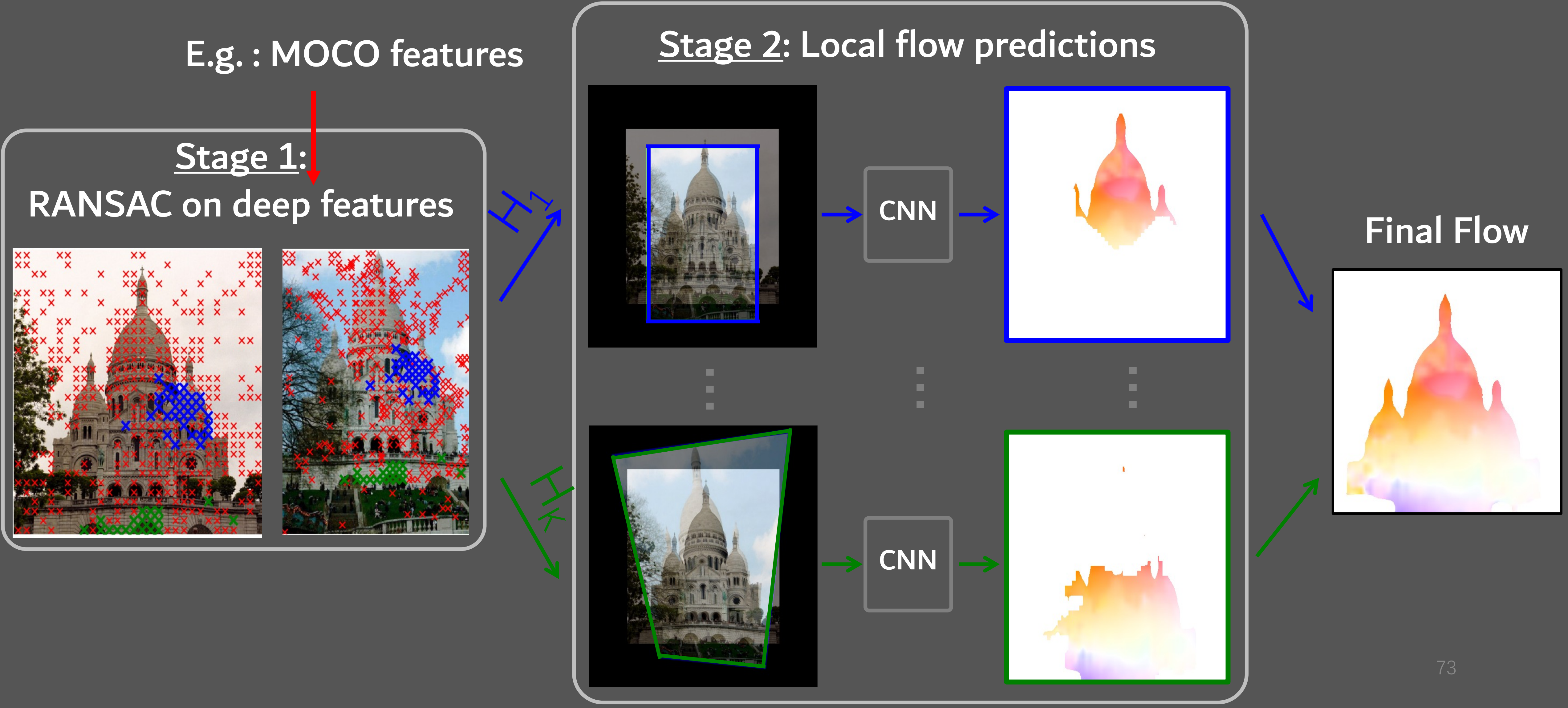
Confident regions, with $(x, y) = \mathbf{F}_{s\to t}(x', y')$

$$\mathcal{L}_c(I_s, I_t) = \sum_{(x,y)\in I_t} \overbrace{M_t^{cycle}(x,y)}\|(x',y'), \mathbf{F}_{t\to s}(x,y)\|_2$$

Key idea: an unsupervised two-stage method



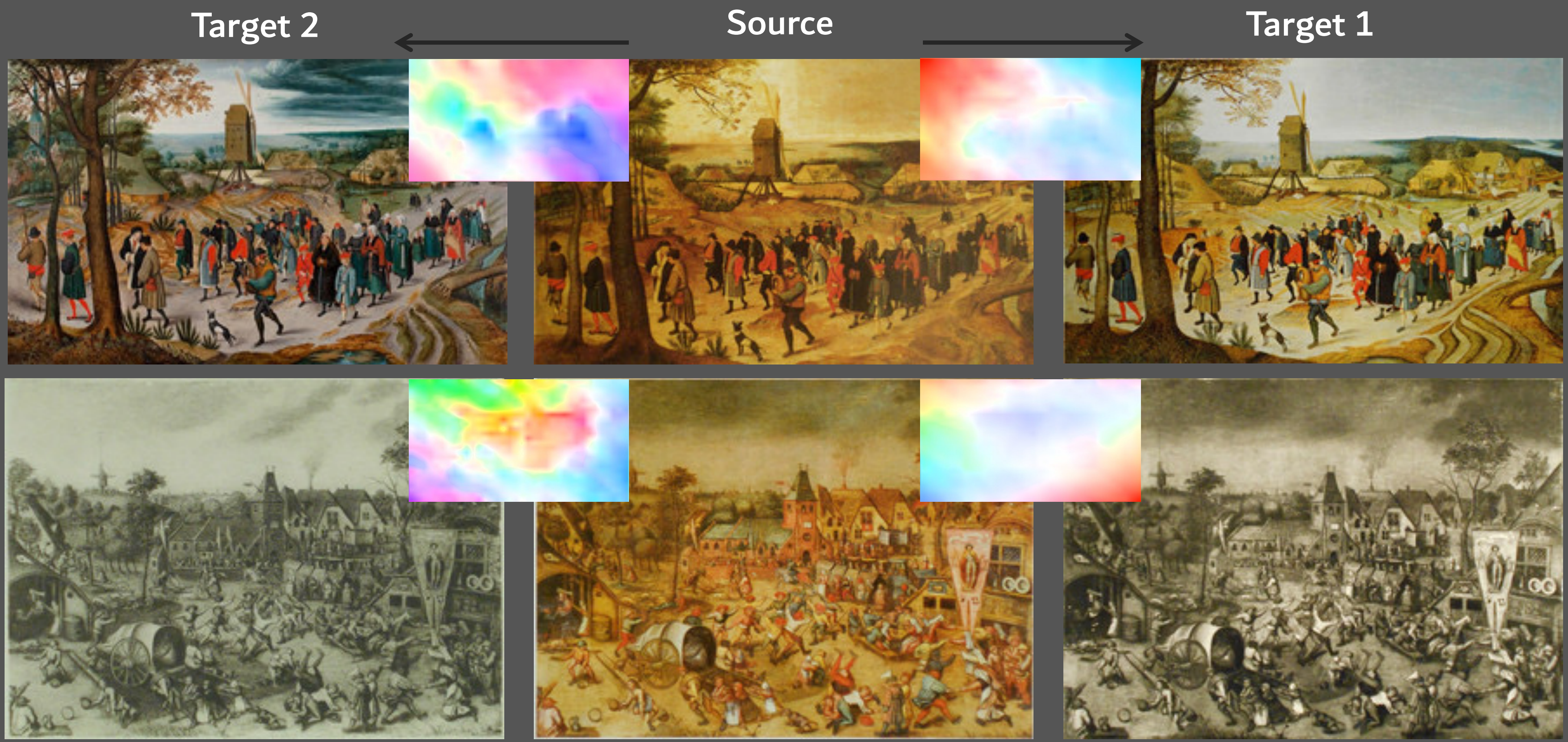
Key idea: an unsupervised two-stage method



Experiments: artwork alignment

Inputs W/o alignment Coarse alignment Fine alignment Top: Coarse flow
Bottom: Fine flow

Experiments: artwork analysis



More visual results can be found in the project page : http://imagine.enpc.fr/~shenx/SegSwap/

Experiments: aligning a group of art details

Discovered patterns in [Shen et al. 2019]

Our fine alignment

Experiments: aligning a group of Internet images

...

...

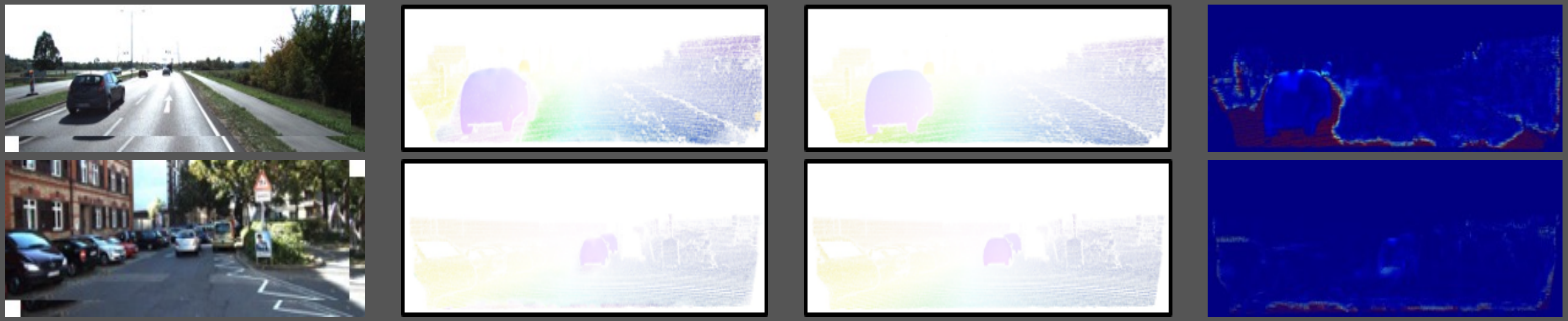
...

Inputs

Our fine alignment

Other results

Optical flow on
KITTI [Morit and Geiger, 2015]
and
Hpatches [Vassileios et al. 2017]



Input | Pred. flow | G.T. flow | Error map

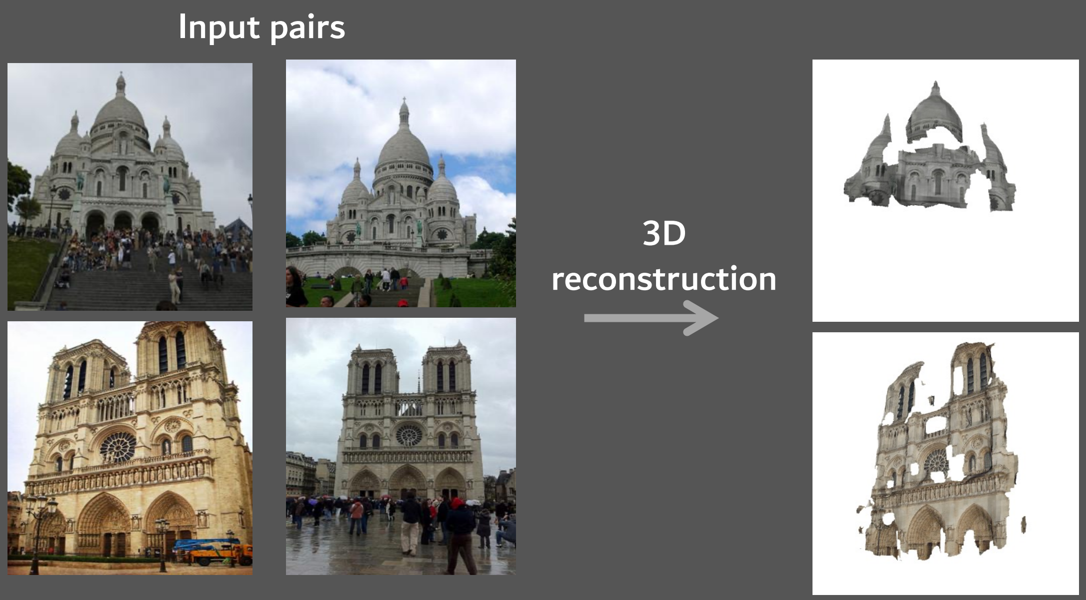
Sparse correspondences on
RobotCar [Will et al. 2017, Mans et al. 2019]
and
MegaDepth [Vassileios et al. 2017]

Inputs | Avg. coarse align. | Avg. fine align. | Top: coarse flow
Bottom: fine flow

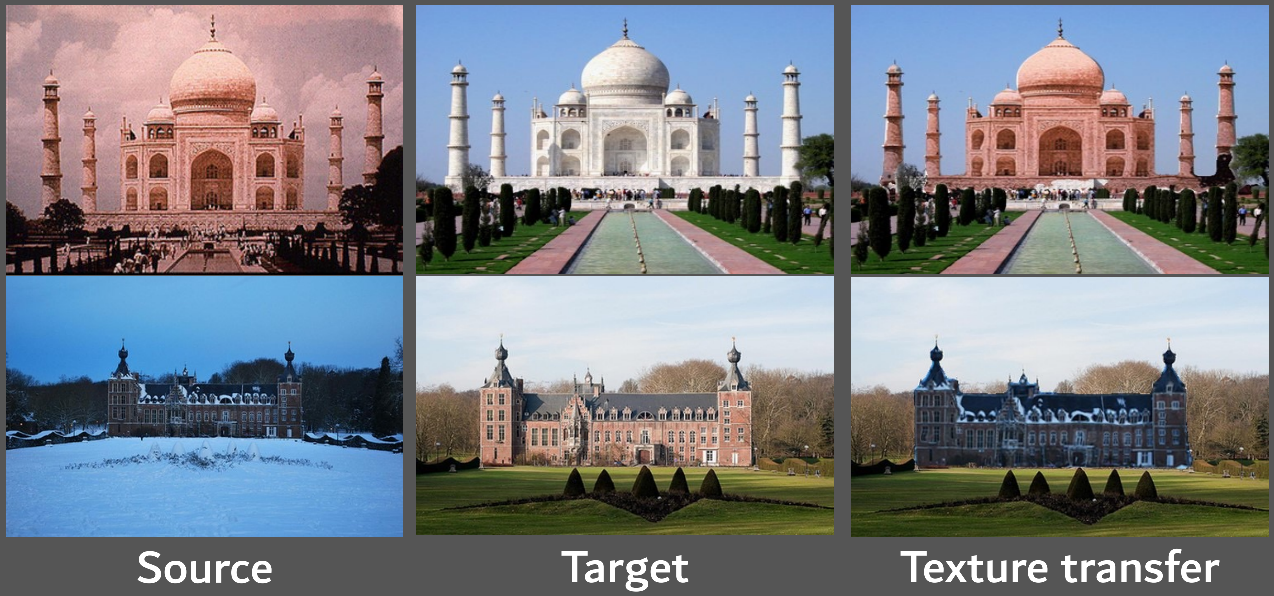
And two-view geometric estimation,
3D reconstruction and texture
transfer…

Other results

Two-view geometric estimation on YFCC100M [Thomee et al. 2016; Zhang et al. 2019] and Aachen day-night [Sattle et al. 2018]



Texture transfer on LTLL [Fernando et al. 2015]



Summary

- Unsupervised two-stage method for dense image alignment

- Superior performances on artworks alignment, optical flow, sparse correspondences and 3D reconstruction

Conclusion

Three unsupervised deep learning methods for near-duplicated patterns discovery and alignment in artworks:

- Style-invariant feature from self-supervision
- Learning co-segmentation from synthetic data
- Dense image alignment from reconstruction

Publications

Discovering Visual Patterns in Art Collections with Spatially-consistent Feature Learning.
Xi Shen, Alexei A. Efros, Mathieu Aubry,
CVPR 2019; Project page and code: http://imagine.enpc.fr/~shenx/ArtMiner/

Spatially-consistent Feature Matching and Learning for Art Collections and Watermark Recognition.
Xi Shen, Robin Champenois, Shiry Ginosar, Ilaria Pastrolin, Morgane Rousselot, Oumayma Bounou, Tom Monnier, Spyros Gidaris, François Bougard, Pierre-Guillaume Raverdy, Marie-Françoise Limon, Christine Bénèvent, Marc Smith, Olivier Poncet, K. Bender, Joyeux-Prunel Béatrice, Elizabeth Honig, Alexei A. Efros, Mathieu Aubry
IJCV Minor Revision, 2021; Project page and code: http://imagine.enpc.fr/~shenx/HisImgAna/

Learning Co-segmentation by Segment Swapping for Retrieval and Discovery.
Xi Shen, Alexei A. Efros, Armand Joulin, Mathieu Aubry,
In submission; Project page and code: http://imagine.enpc.fr/~shenx/SegSwap/

RANSAC-Flow: Generic Two-stage Image Alignment.
Xi Shen, François Darmon, Alexei A. Efros, Mathieu Aubry,
ECCV 2020; Project page and code: http://imagine.enpc.fr/~shenx/RANSAC-Flow/

Additional publications on historical data analysis

Large-Scale Historical Watermark Recognition: dataset and a new consistency-based approach.
Xi Shen, Ilaria Pastrolin, Oumayma Bounou, Spyros Gidaris, Marc Smith, Olivier Poncet, Mathieu Aubry
ICPR, 2021, Project page and code: http://imagine.enpc.fr/~shenx/Watermark/

A Web Application for Watermark Recognition.
Oumayma Bounou, Tom Monnier, Ilaria Pastrolin, **Xi Shen**, Christine Bénèvent, Marie-Françoise Limon-Bonnet, François Bougard, Mathieu Aubry, Marc Smith, Olivier Poncet, Pierre-Guillaume Raverdy
Journal of Data Mining and Digital Humanities, 2021 , Web application: https://filigranes.inria.fr/#/filigrane-search

Image Collation: Matching illustrations in manuscripts.
Ryad Kaoua, **Xi Shen**, Alexandra Durr, Stavros Lazaris, David Picard, Mathieu Aubry
ICDAR, 2021, Project page and code: http://imagine.enpc.fr/~shenx/ImageCollation/

Other publications

Few –shot learning

Empirical Bayes Transductive Meta-Learning with Synthetic Gradients.
Shell Xu Hu, Pablo G Moreno, Yang Xiao, **Xi Shen**, Guillaume Obozinski, Neil D Lawrence, Andreas Damianou
ICLR, 2020
Code: https://github.com/hushell/sib_meta_learn

Re-ranking for image retrieval and transductive few-shot classification.
Xi Shen, Yang Xiao, Shell Hu, Othman Sbai, Mathieu Aubry
NeurIPS, 2021
Project page and code: http://imagine.enpc.fr/~shenx/SSR/

Weakly supervised learning

Marginalized Average Attentional Network for Weakly-Supervised Learning.
Yuan Yuan, Yueming Lyu, **Xi Shen**, Ivor W Tsang, Dit-Yan Yeung
ICLR, 2019
Code: https://github.com/yyuanad/MAAN

Future works

- An advanced annotation system incorporating unsupervised / weakly supervised techniques, interaction with users

- End-to-end multi-object multi-image discovery

Thanks to everybody I interacted with

Thanks for your attention!
Questions ?