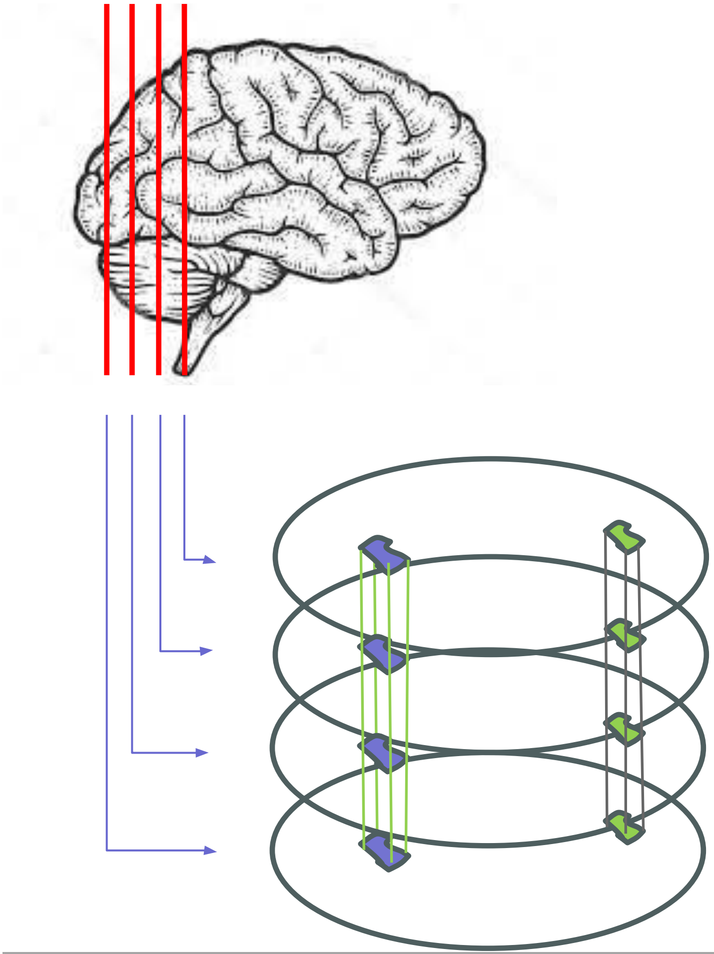
Heinrich Heine
Universität Düsseldorf

SuperBrain

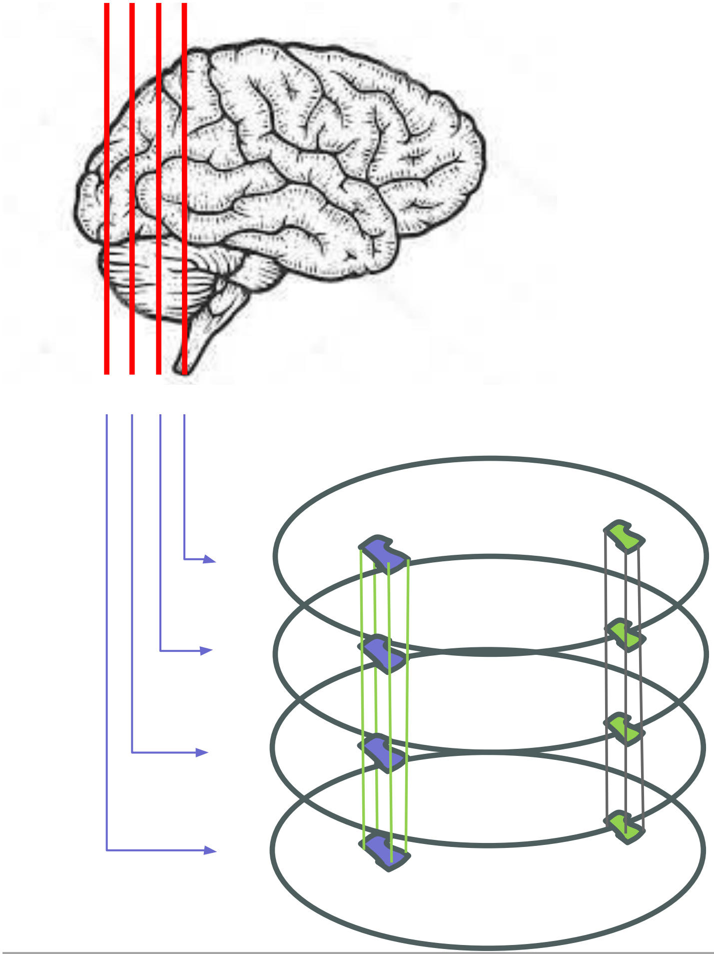
Image matching for successive scans of brain slices

Problem overview



- Reconstruct 3D model of multiple 2D brain-slice scans
- Two-step process: **Image matching** and image registration
- Highly non-linear through thin and flexible nature of brain slices

Problem overview

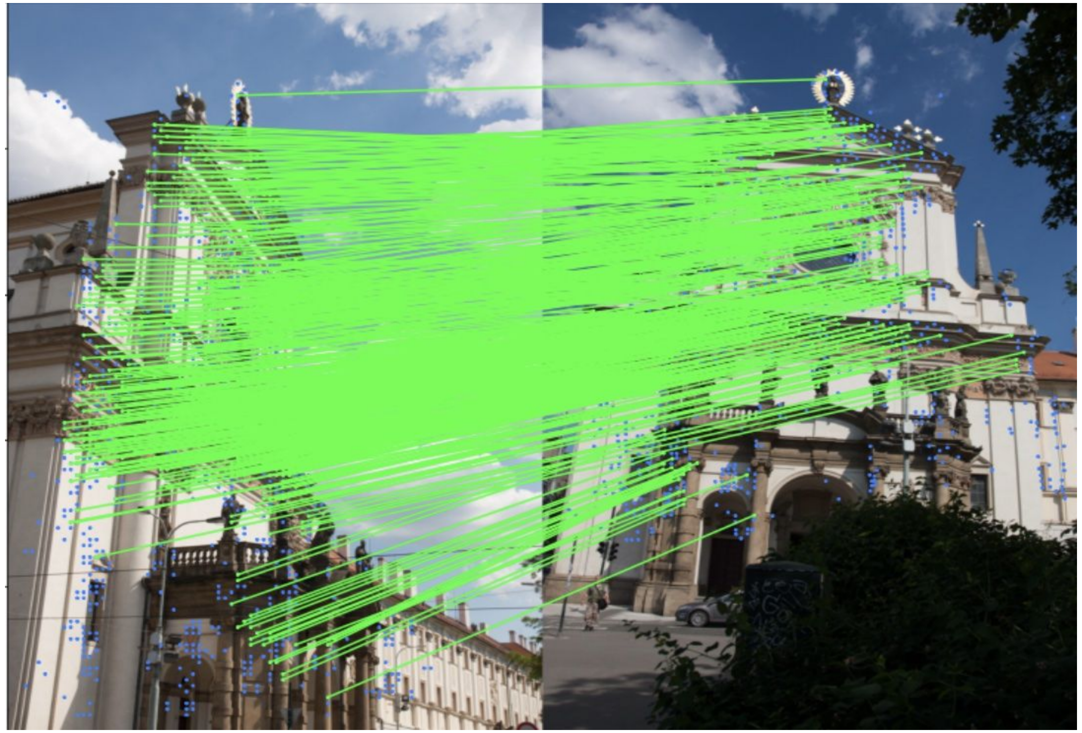


- Reconstruct 3D model of multiple 2D brain-slice scans
- Two-step process: **Image matching** and image registration
- Highly non-linear through thin and flexible nature of brain slices

→ This work thematizes the image matching step, hereinafter referred to as "Brain matching"

Classical image matching

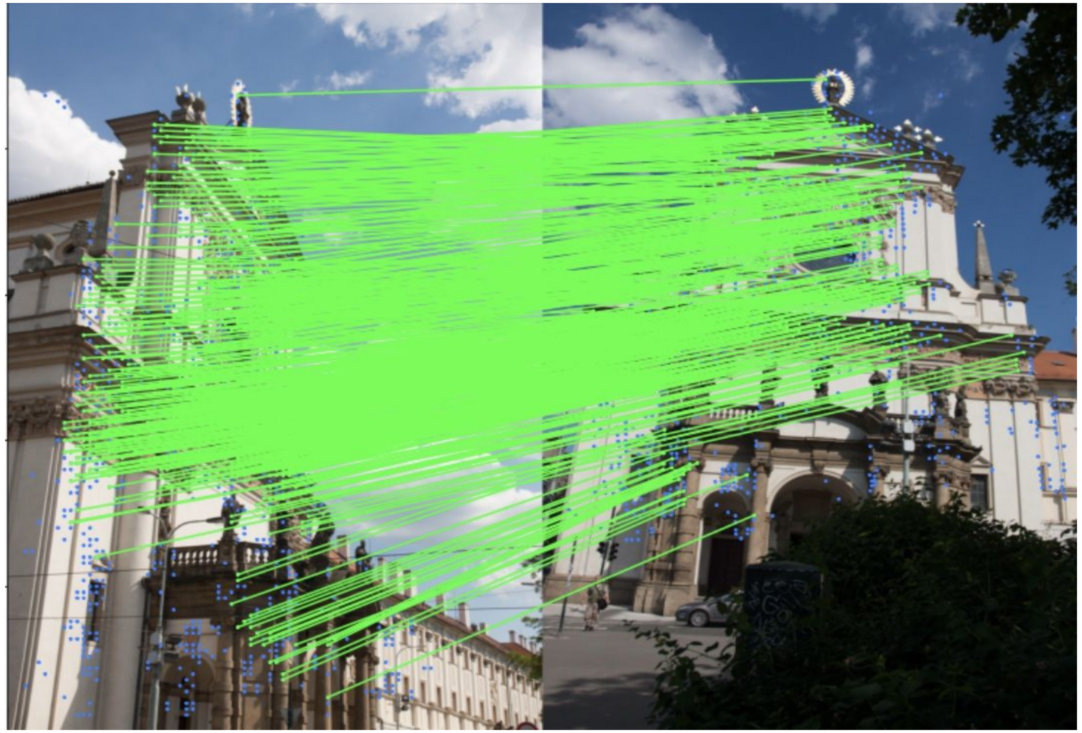
- “Image matching is a process of finding pixel and region correspondences between two images of the same scene.”
  ~ Kornia



https://kornia.readthedocs.io/

Classical image matching

- “Image matching is a process of finding pixel and region correspondences between two images of the same scene.”
  ~ Kornia



https://kornia.readthedocs.io/

→ notion of interest point detection is semantically ill-defined

Classical image matching vs. Brain matching

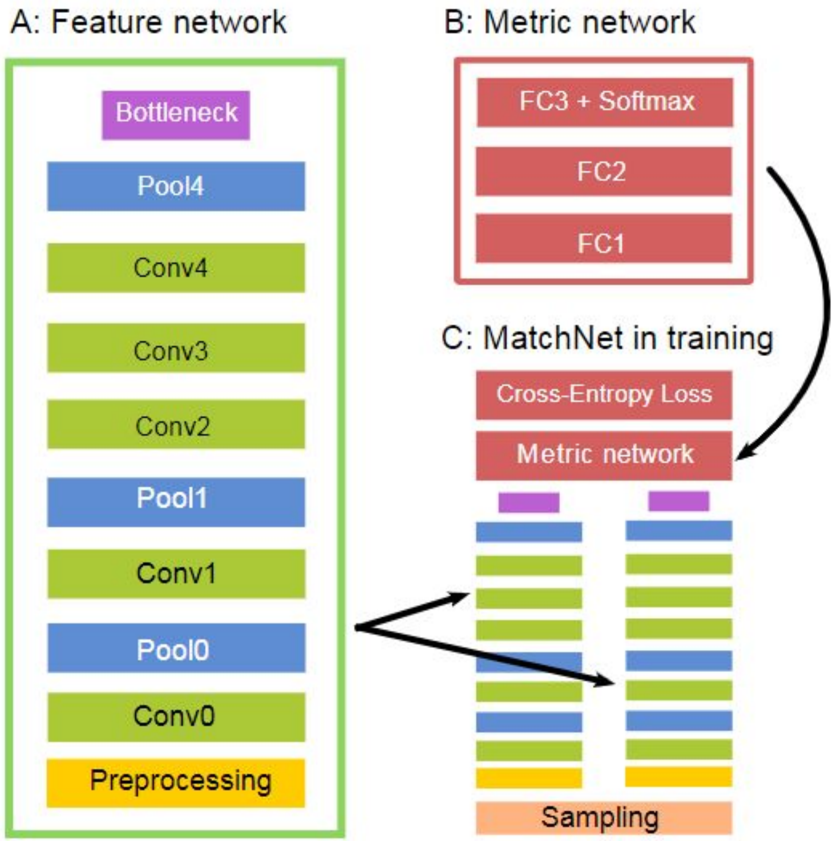
- Classical image matching: Natural images of same object from different perspectives
- Brain matching: Medical images of different objects from same perspective

Classical image matching vs. Brain matching

- Classical image matching: Natural images of same object from different perspectives
- Brain matching: Medical images of different objects from same perspective

Some image matching approaches make use of augmentations. This is suitable for natural images, because augmentations are similar to image changes in the real world. But: Inapplicable for brain matching!

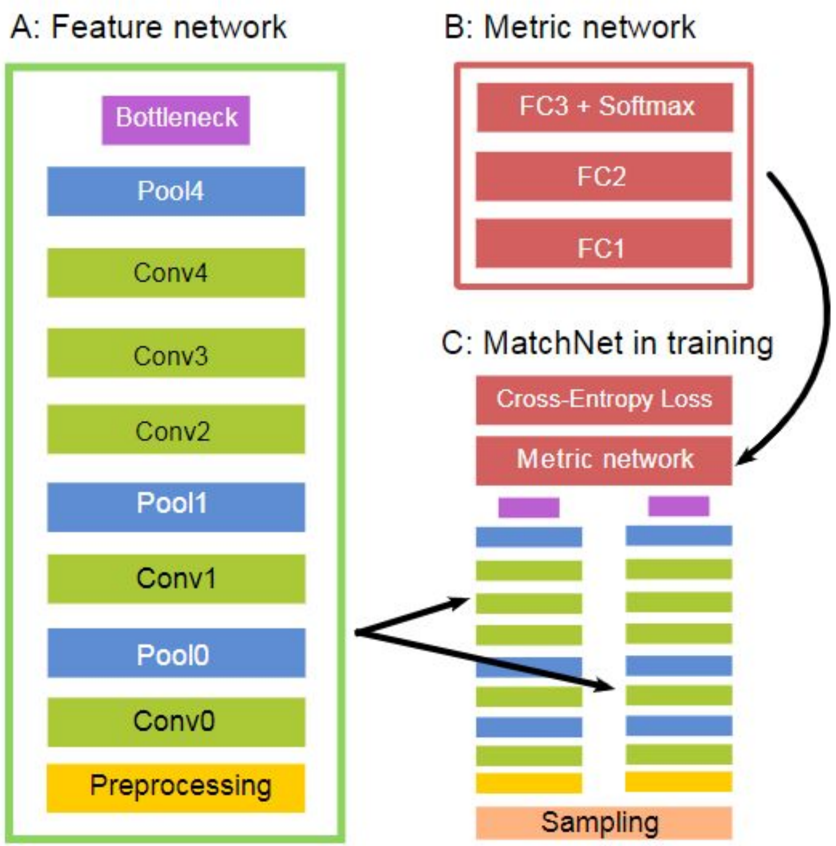
MatchNet (2015)



MatchNet: Unifying Feature and Metric Learning for Patch-Based Matching

- Extract positive and negative pairs of patches (UBC patch dataset)
- Train a binary CNN classifier with cross-entropy

MatchNet (2015)

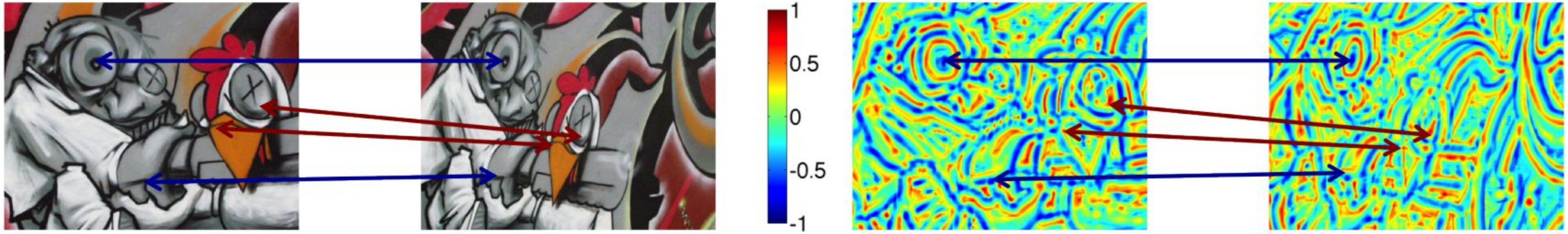


MatchNet: Unifying Feature and Metric Learning for Patch-Based Matching

- Extract positive and negative pairs of patches (UBC patch dataset)
- Train a binary CNN classifier with cross-entropy

→ finds patch-correspondences, not pixel-to-pixel

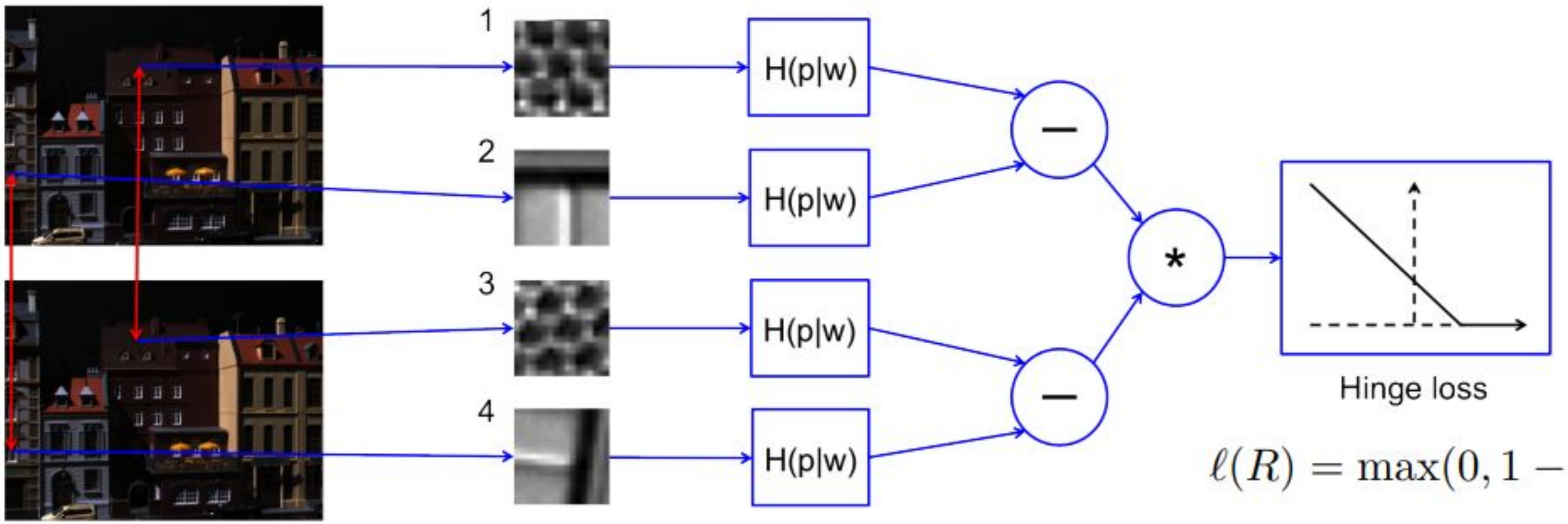
QuadNetworks (2016)



Quad-Networks: unsupervised learning to rank for interest point detection

- Unsupervised rank learning: Map patches of original and transformed image to a single real-valued response
- Corresponding patches from images should have the same position in their sorted intra-image ranking
- *If one point is higher in the ranking than another one, it should still be higher after a transformation*

QuadNetworks (2016)

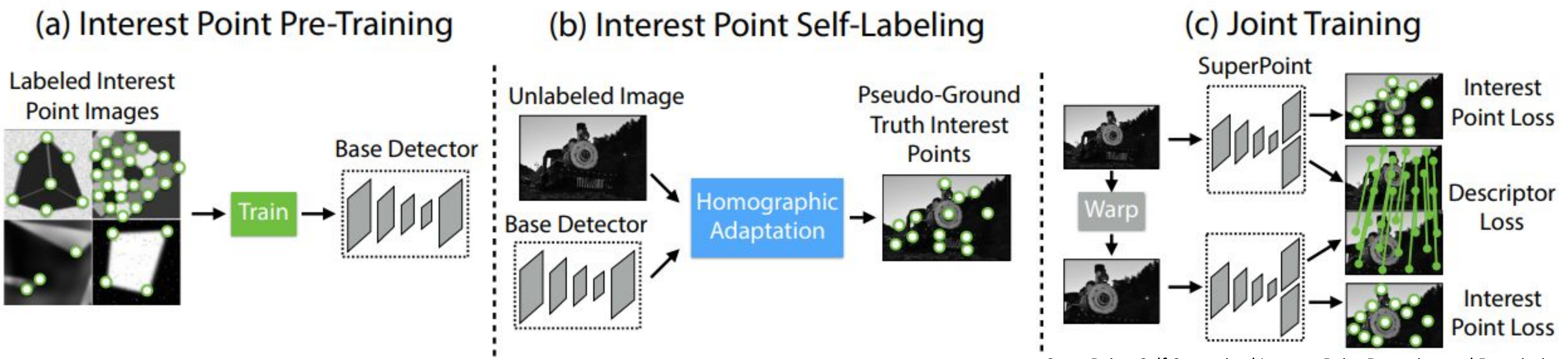


$$\ell(R) = \max(0, 1 - R)$$

Quad-Networks: unsupervised learning to rank for interest point detection

→ dependent on augmentations

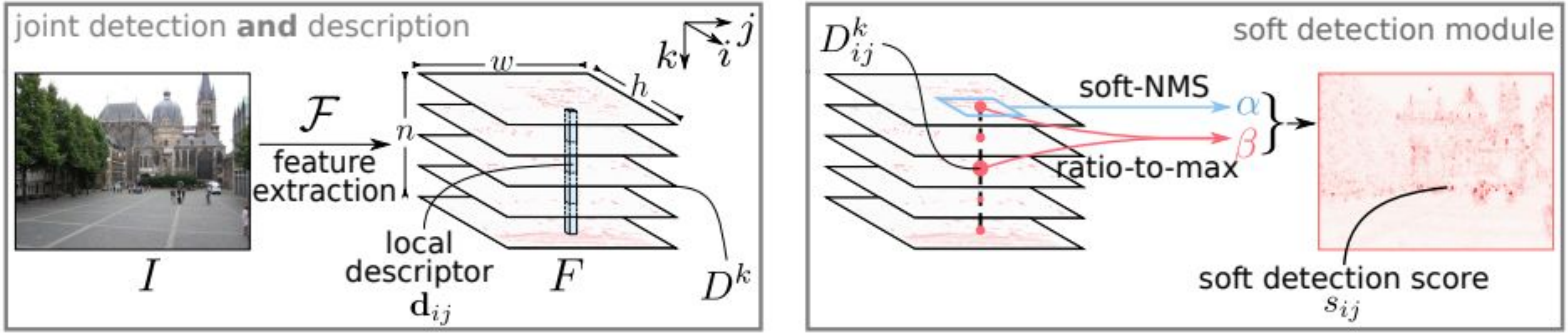
SuperPoint (2017)



SuperPoint: Self-Supervised Interest Point Detection and Description

- Self-supervised
- Model works on full-sized images
- Computes keypoints and descriptors in one forward pass

D2Net (2019)



D2-Net: A Trainable CNN for Joint Description and Detection of Local Features

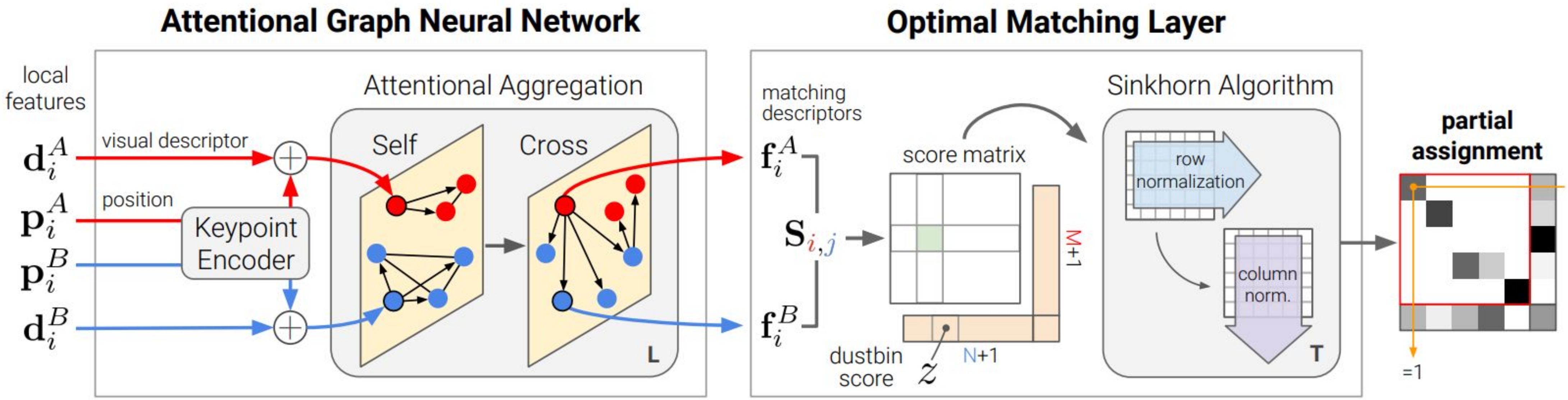
- Detect-and-describe instead of detect-then-describe
- Performs better under extreme appearance changes than previous approaches

Approaches so far…

… used a naive way of assigning matches:

- By finding mutual nearest neighbors, i.e. match descriptors from 2 images with min. euclidean distance
- Ignoring the assignment-structure of keypoints

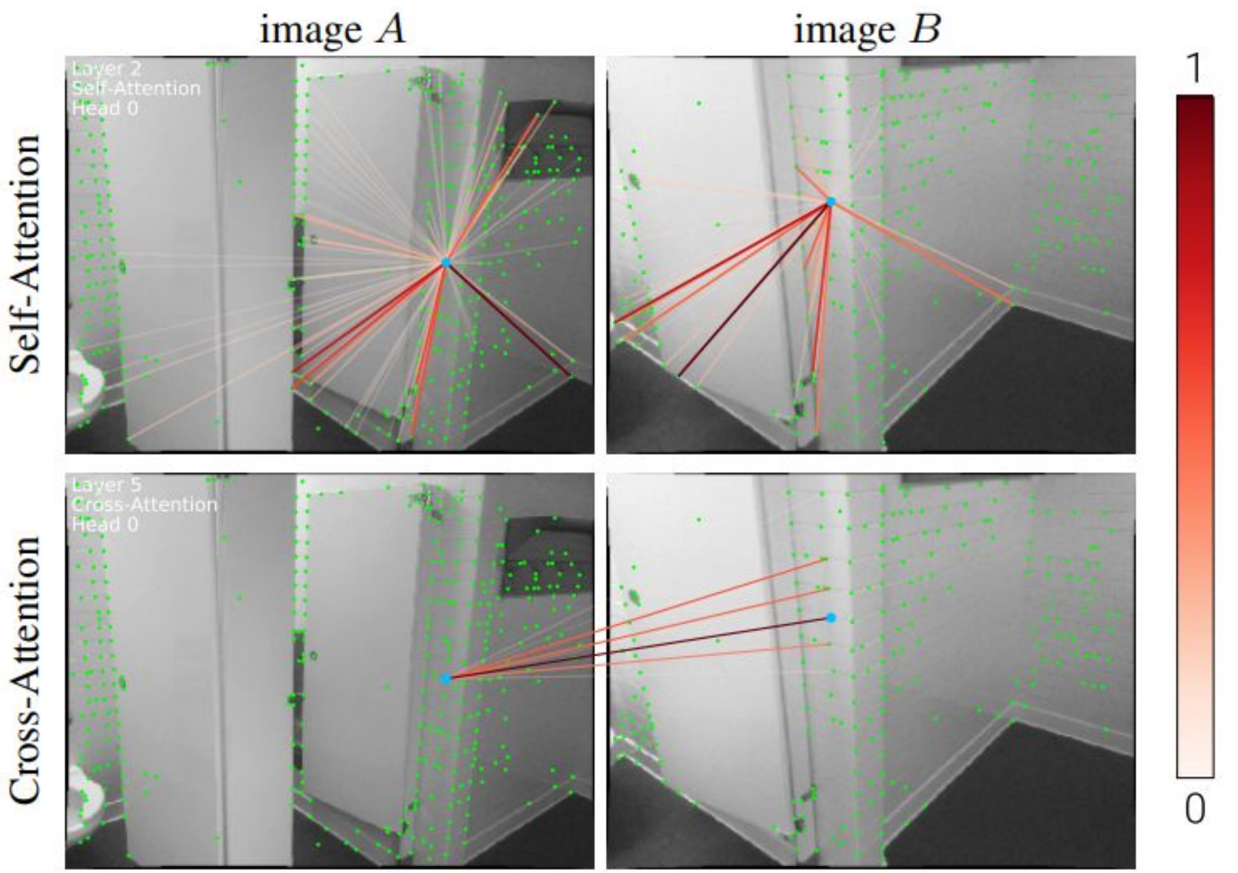
SuperGlue (2019)



SuperGlue: Learning Feature Matching with Graph Neural Networks

- SuperGlue solves this limitation by learning the assignment structure (GIVEN local features, e.g. from SuperPoint)
- Self-attention (intra-image) and cross-attention (inter-image) allow the model to consider its spatial and visual relationship with other keypoints

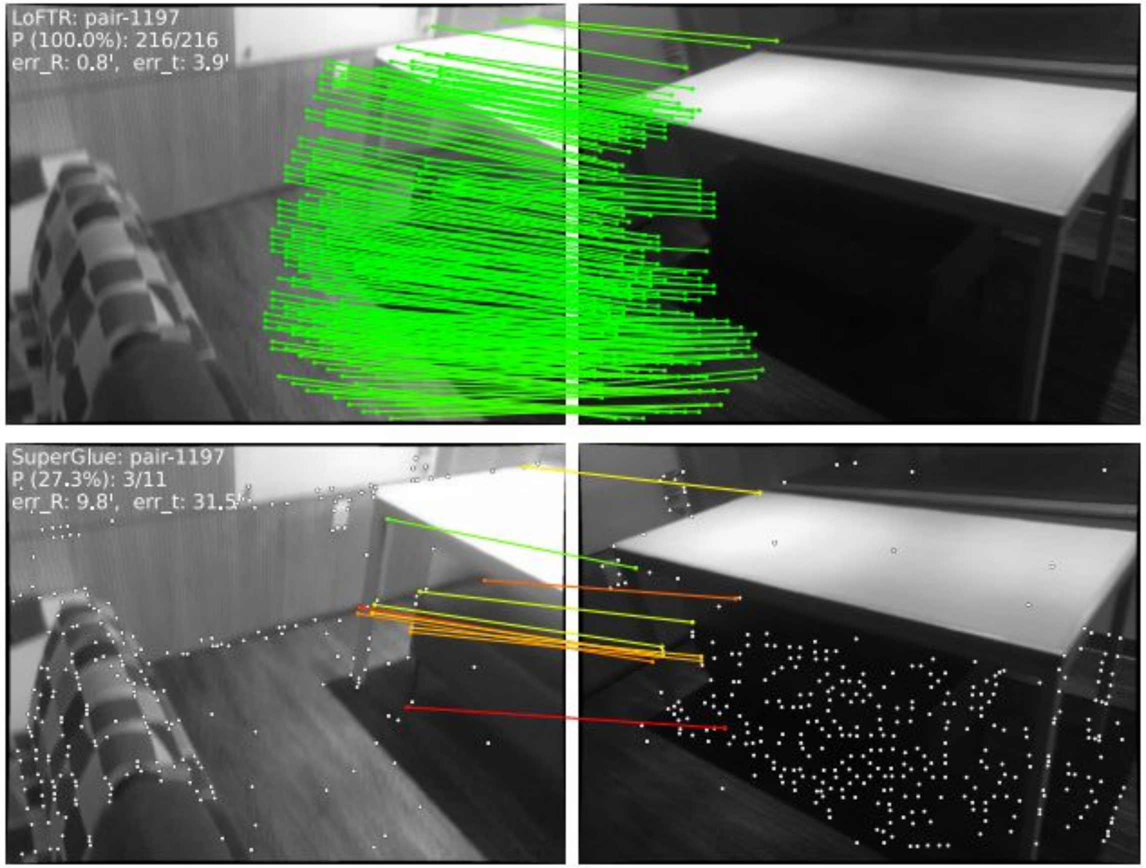
SuperGlue (2019)



SuperGlue: Learning Feature Matching with Graph Neural Networks

Limitations of CNNs (LoFTR 2021)

- CNNs have limited receptive fields which may not distinguish indistinctive regions
- Global context / larger receptive fields are crucial for distinguishing points based on their relative position

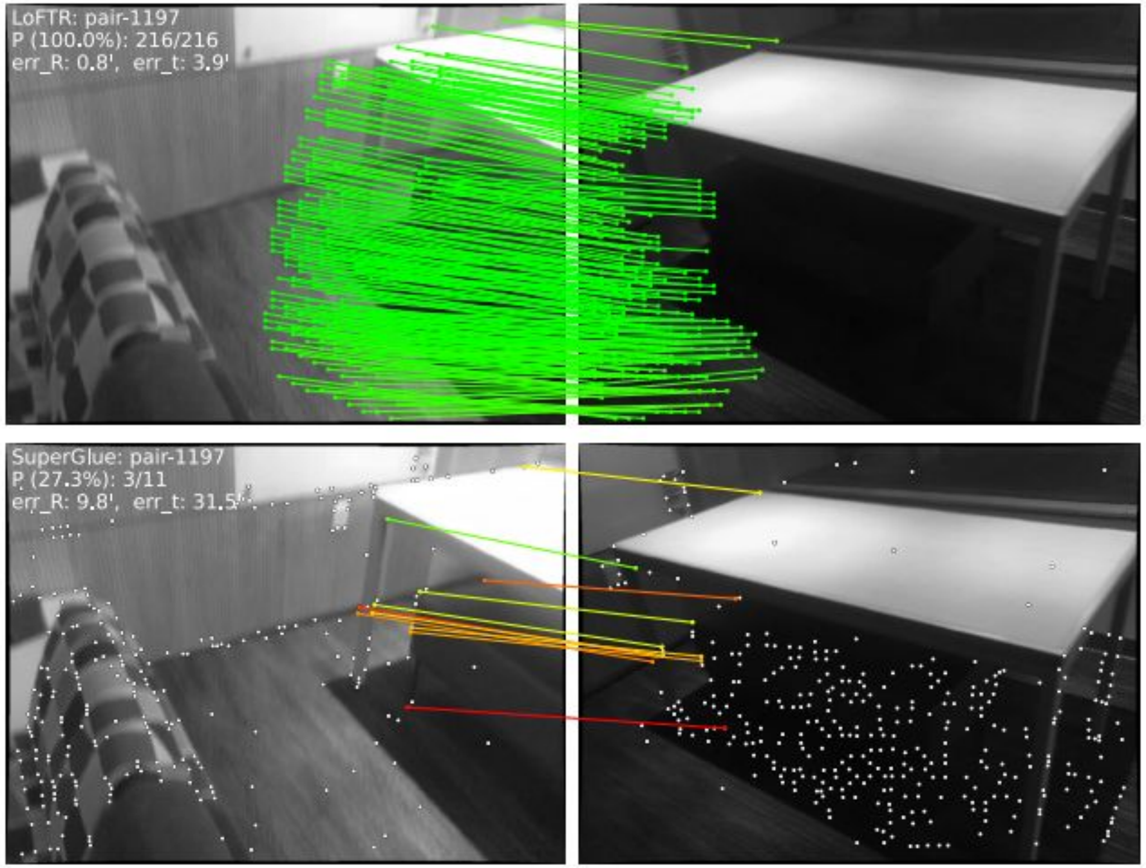


LoFTR: Detector-Free Local Feature Matching with Transformers

Limitations of CNNs (LoFTR 2021)

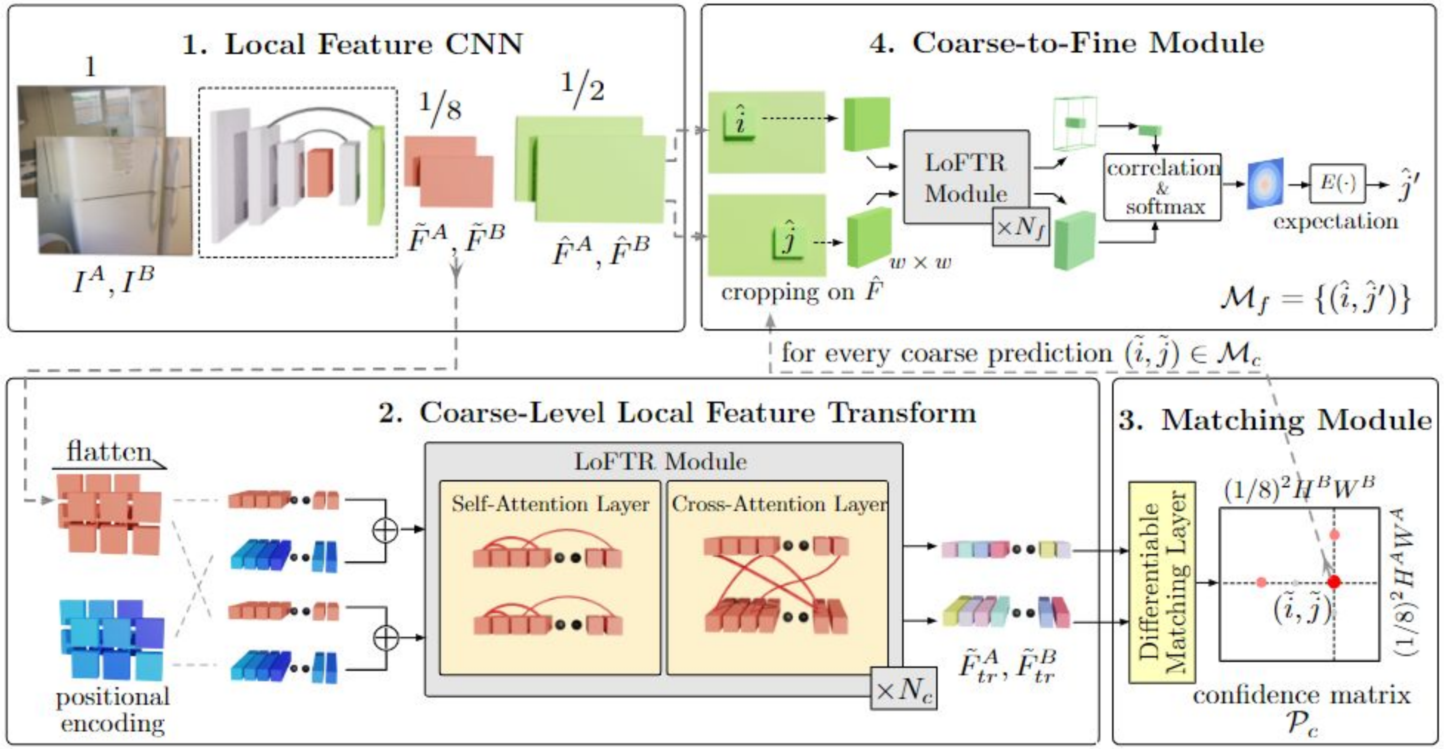
- CNNs have limited receptive fields which may not distinguish indistinctive regions
- Global context / larger receptive fields are crucial for distinguishing points based on their relative position

→ **TRANSFORMER**



LoFTR: Detector-Free Local Feature Matching with Transformers

LoFTR (2021)



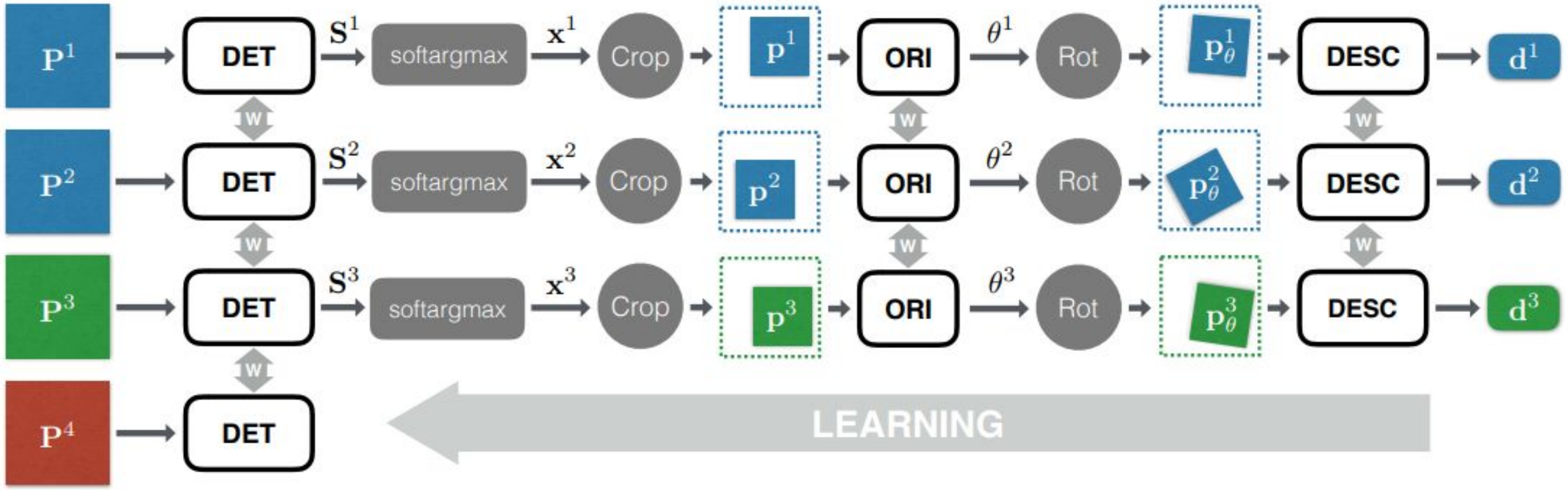
LoFTR: Detector-Free Local Feature Matching with Transformers

- Detector-free approach
- First calculate coarse matches, then refine them

Further reading…

- ASpanFormer (2022)
- QuadTree (2022)
- MatchFormer (2022)
- DKM (2022)
- LightGlue (2023)
- CasMTR (2023)
- and many more…

LIFT (2016)



- Train descriptor, orientation estimator and detector separately on feature points produced by SfM
- P^1 and P^2 are corresponding patches of the same physical point, P^3 of a different physical point and P^4 does not contain any feature point

SIFT vs. LIFT