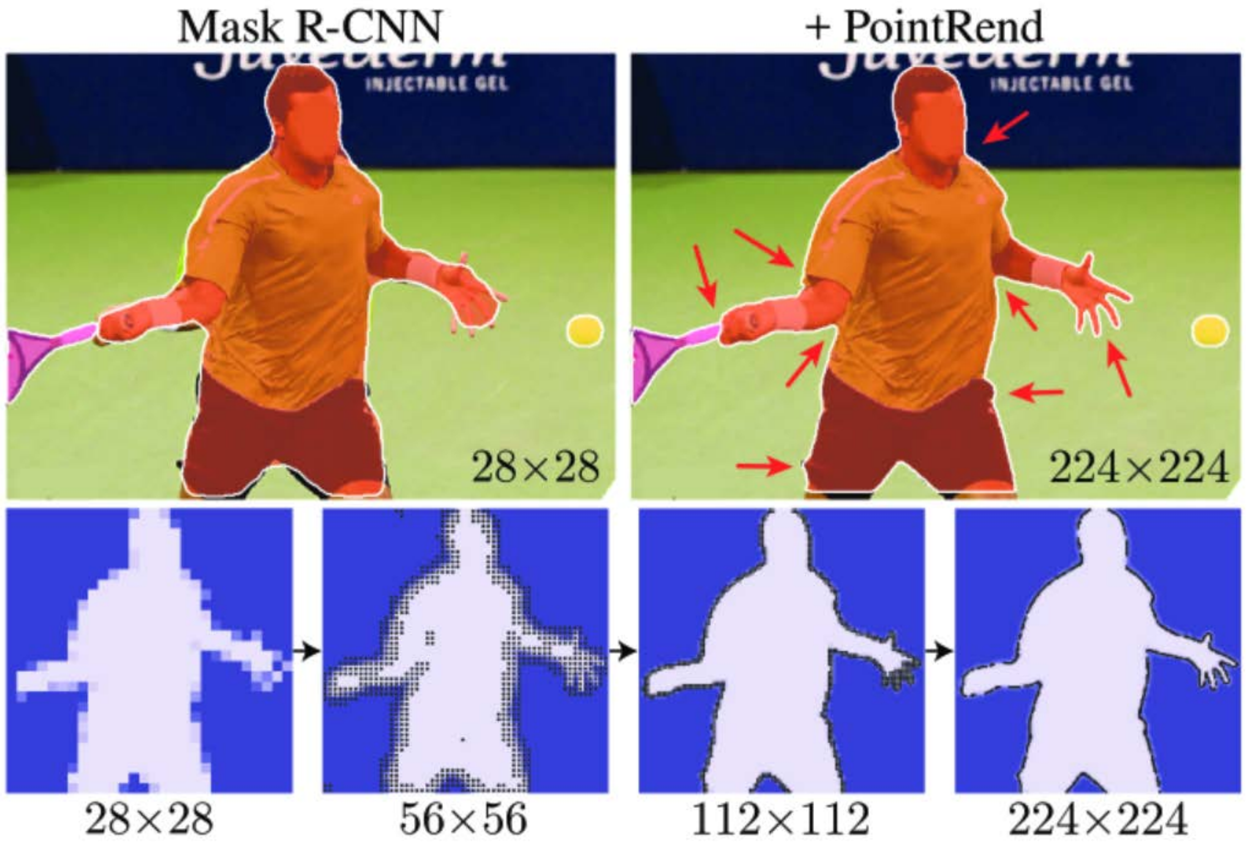
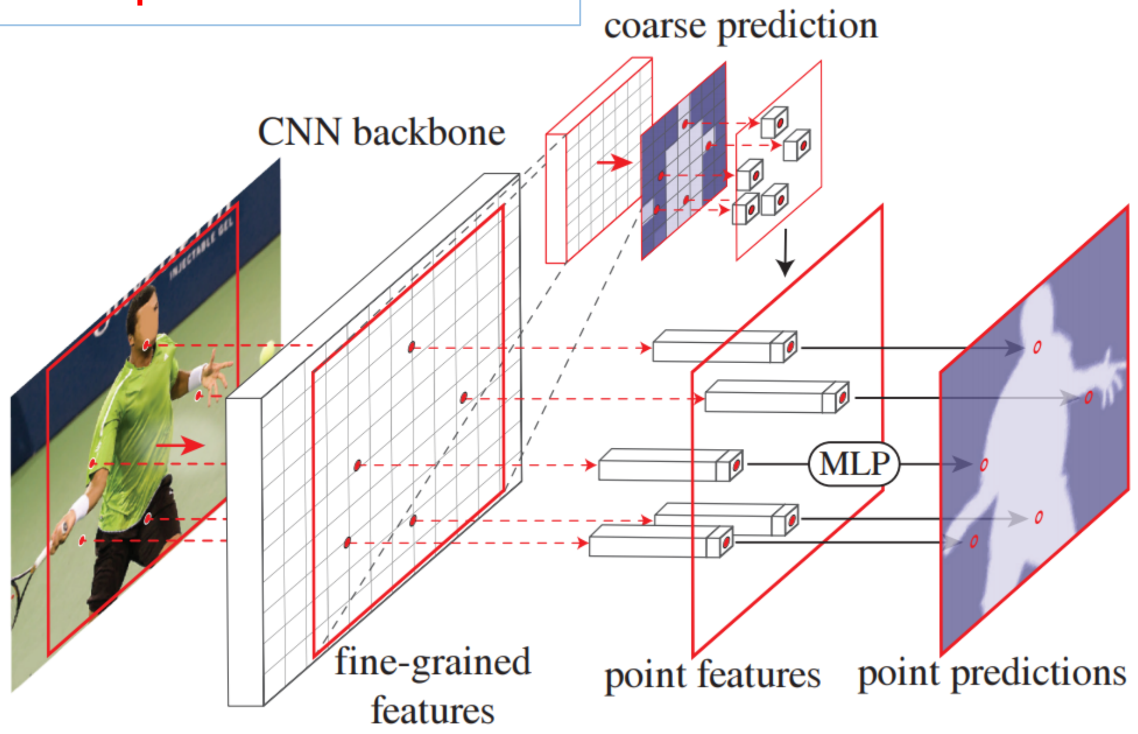
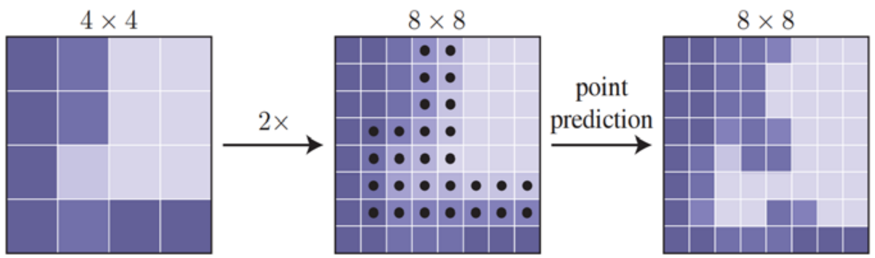
PointRend: Image Segmentation as Rendering
(Alexander Kirillov, Yuxin Wu, Kaiming He, and Ross Girshick)

PointRend uses a subdivision strategy to adaptively select a non-uniform set of points at which to compute labels.

Only think about the important (incorrect) region.







ELF: Embedded Localization of Features in Pre-Trained CNN

Benbihi, Assia and Geist, Matthieu and Pradalier ICCV 2019

An Guoyuan
20184637

Content

1 Background

Common Requirements

- Problem 1:
  - Detect the same point *independently* in both images

No chance to match!

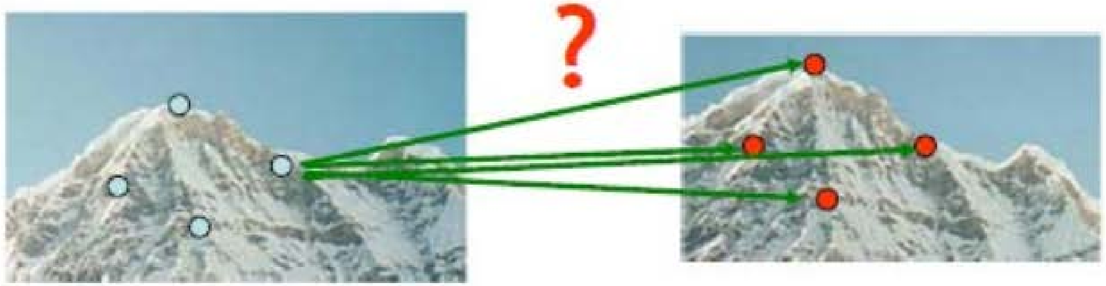
requirement1

We need a repeatable detector!

Slide credit: Darya Frolova, Denis Simakov

Common Requirements

- Problem 1:
  - Detect the same point *independently* in both images
- Problem 2:
  - For each point correctly recognize the corresponding one



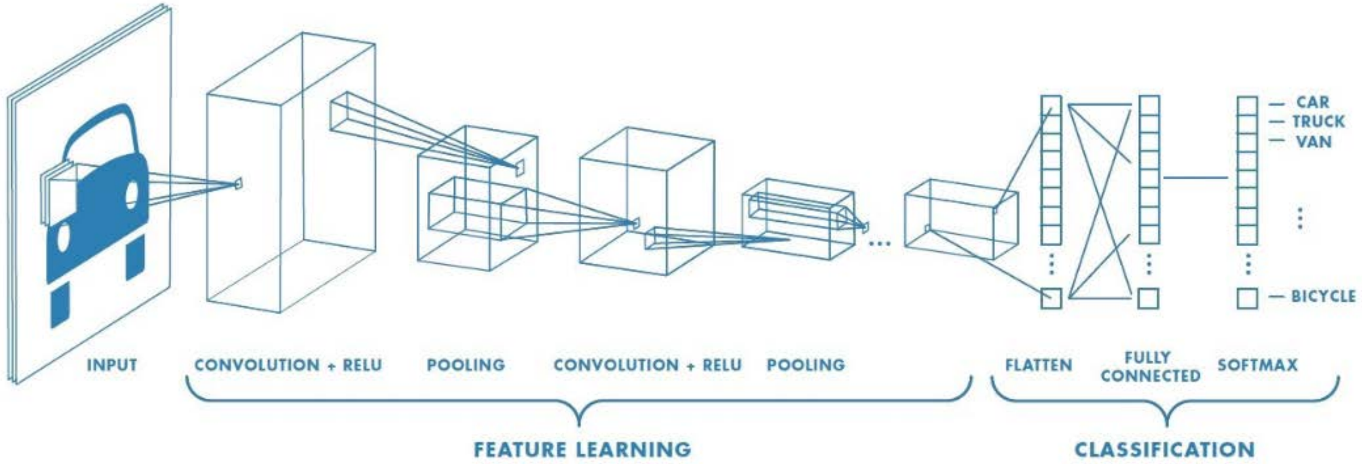
requirement2

We need a reliable and distinctive descriptor!

Slide credit: Darya Frolova, Denis Simakov

- Repeatable key points detector → Harris
- Reliable descriptor → LoG; SIFT

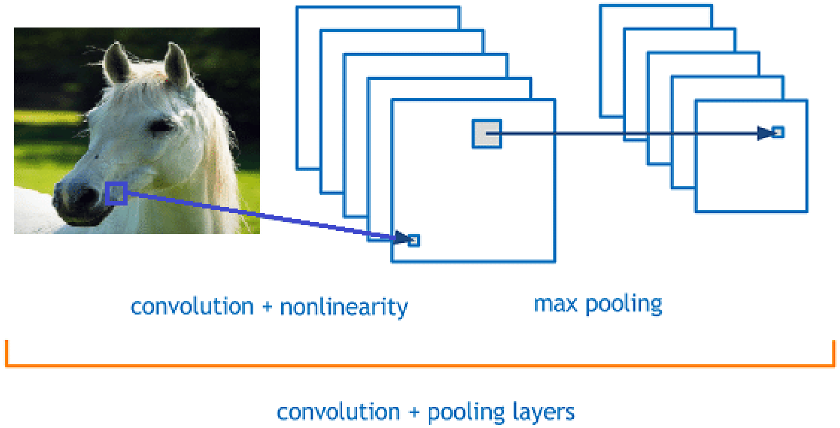
A milestone: Convolutional Neural Network

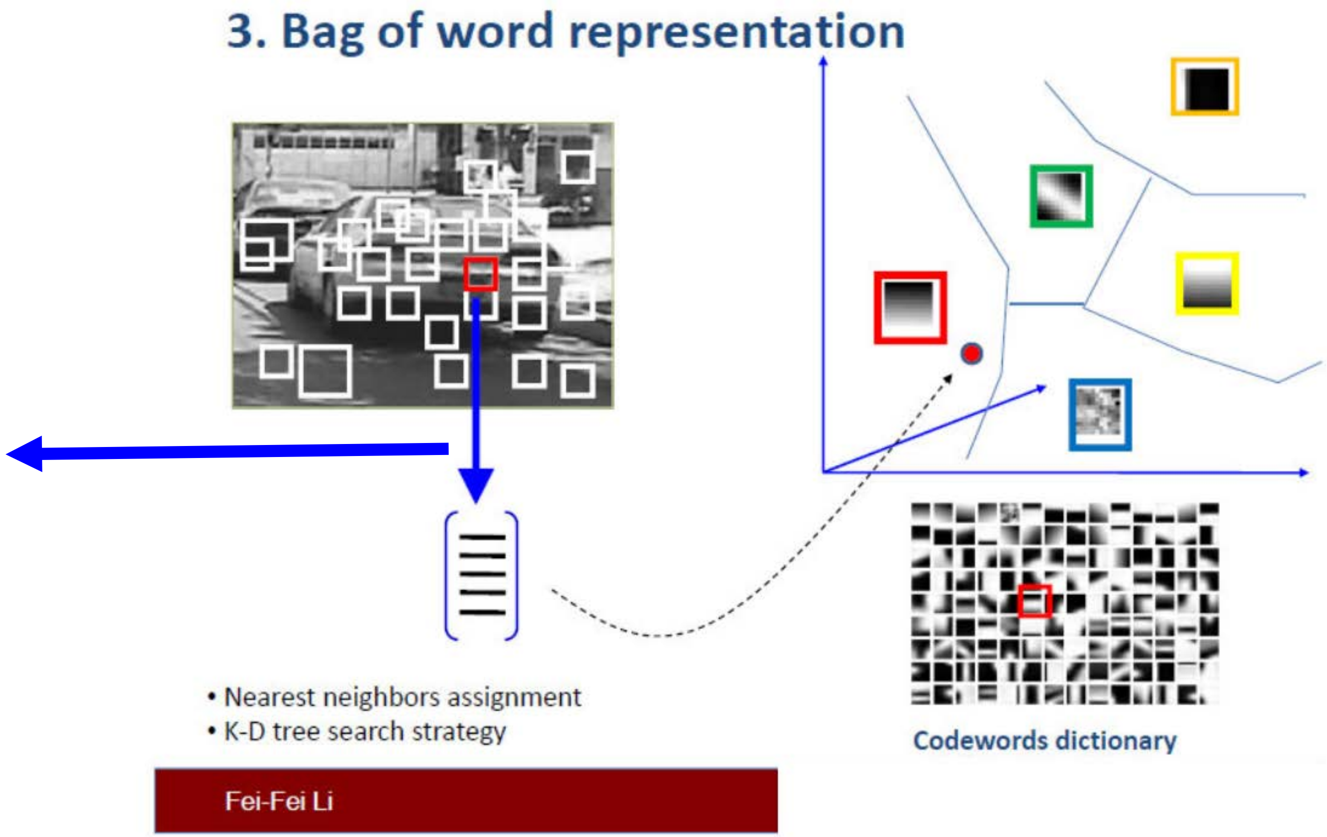


How to use CNN in image search?

Reference:Hands-on Machine Learning with Scikit-Learn, Keras, and TensorFlow, 2nd ed., O'Reilly, 2019
https ://towardsdatascience.com/a-comprehensive-guide-to-convolutional-neural-networks-the-eli5-way-3bd2b1164a53

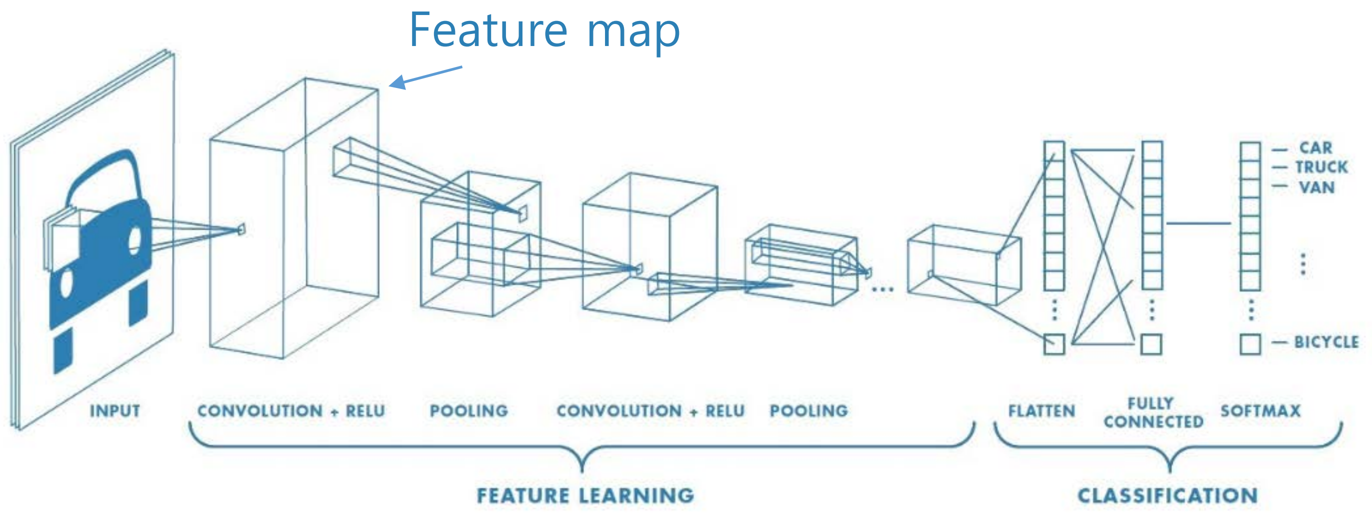
We often use CNN as a descriptor





Q: how could CNN be helpful for the detector?

ELF: Embedded Localization of Features in Pre-Trained CNN

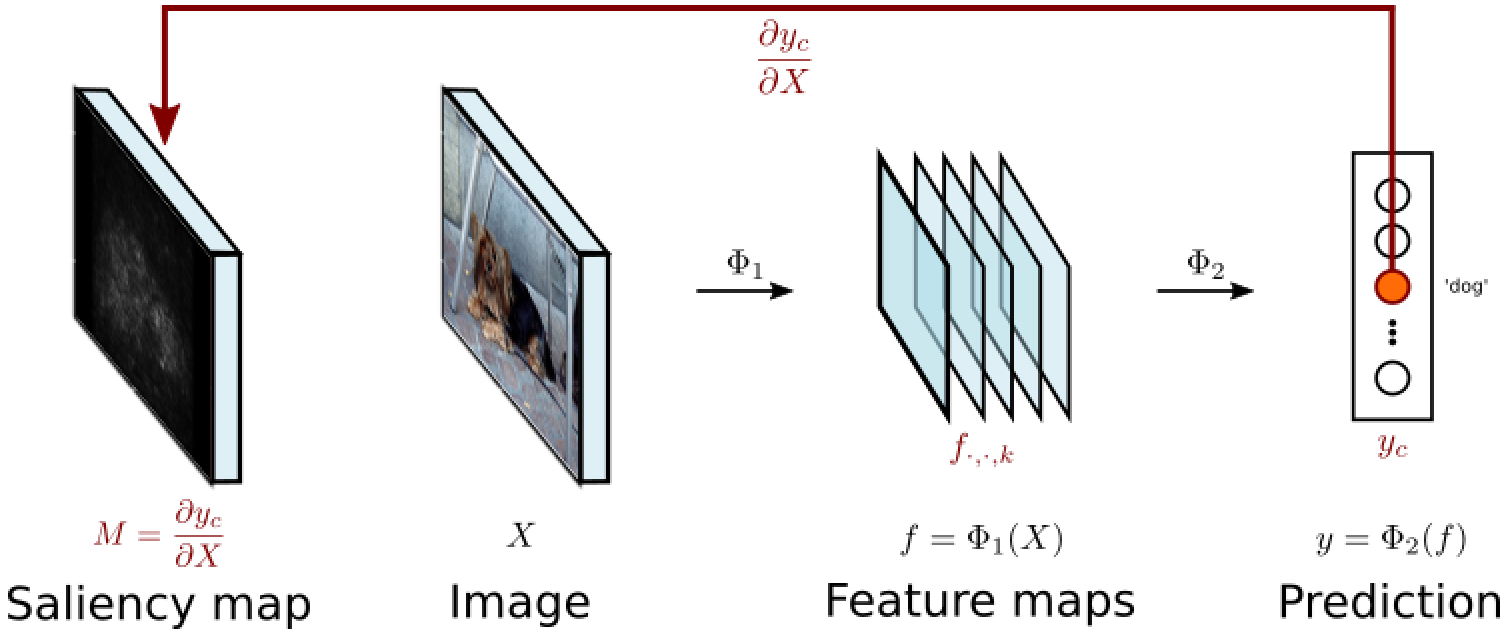


2 Method (part 1) – Saliency Maps

- During backpropagation, the gradient is helpful for detecting keypoints.

Saliency map

- During backpropagation, the gradient is helpful for detecting keypoints.



http://research.sualab.com/assets/images/interpretable-machine-learning-overview-2/saliency-map-with-gradient-concept.png

Saliency map from feature map.

$\mathbf{I}$ Image $D_I = H_I \cdot W_I \cdot C_I$

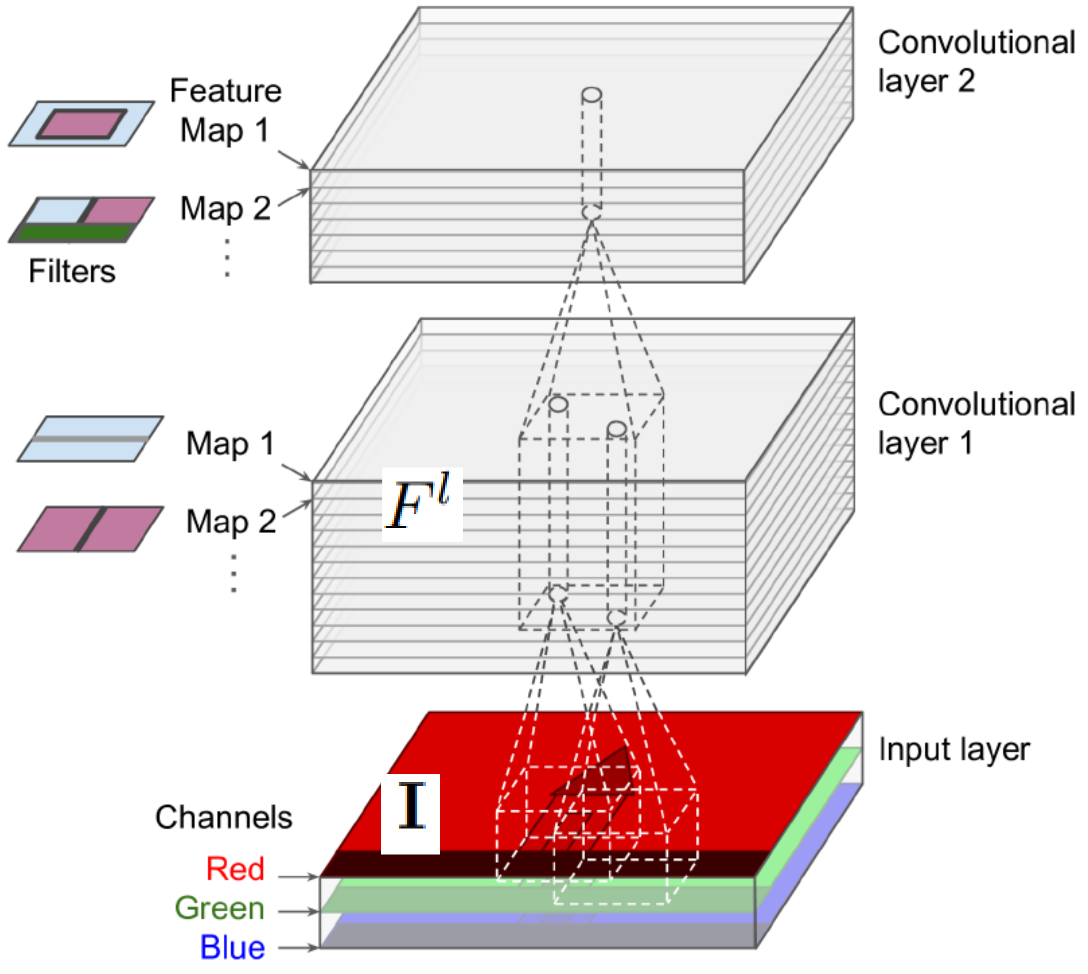
F^l Feature Map $D_F = H_l \cdot W_l \cdot C_l$

Saliency Map $S^l(\mathbf{I}) = \left| {}^t F^l(\mathbf{I}) \cdot \nabla_I F^l \right|$

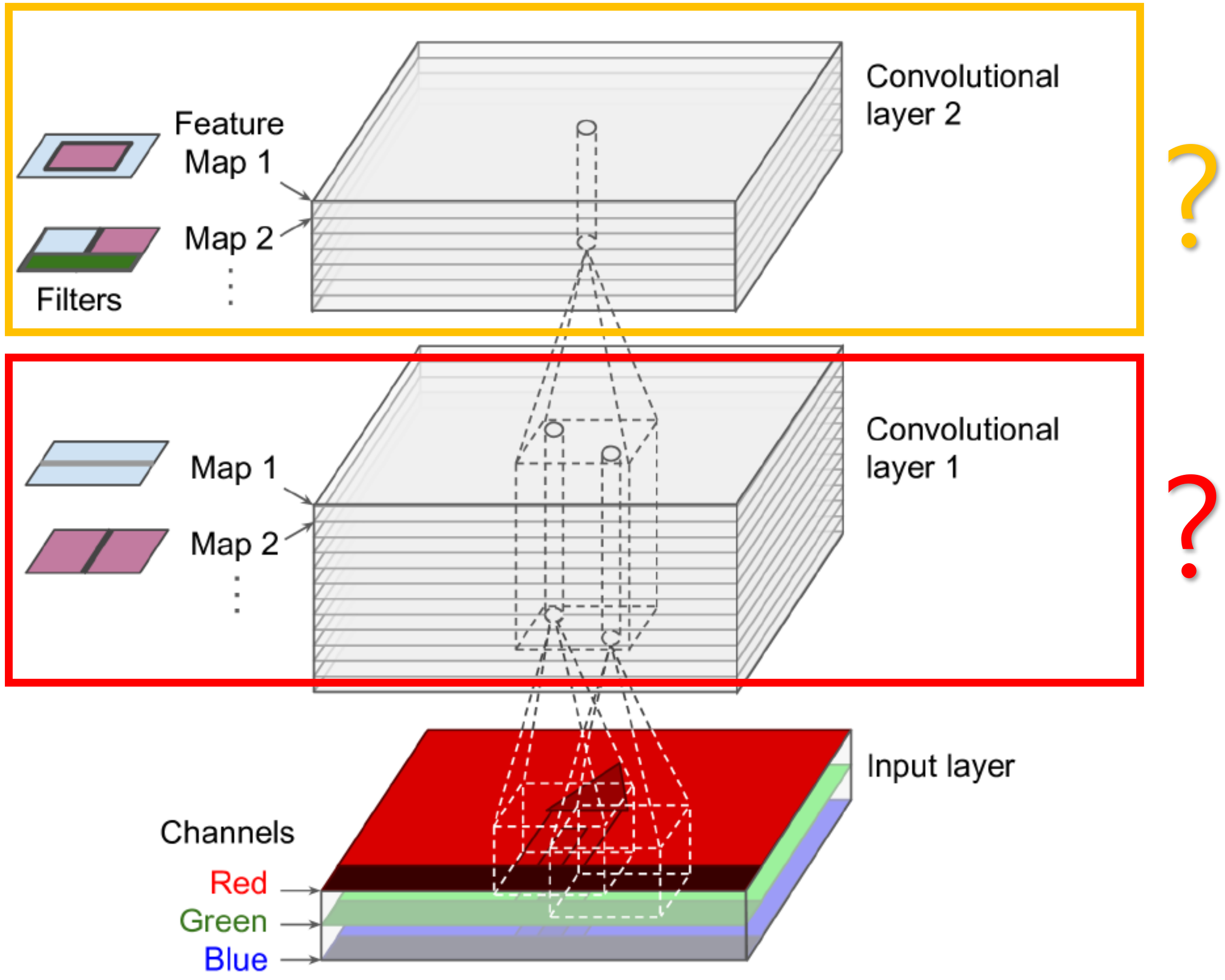
Apply the correlation $\nabla_I F^l$ to the features $F^l(\mathbf{I})$ specifically and generate a visualization in image space $S^l(\mathbf{I})$.

$\nabla_I F^l$

- The correlation between the feature space and image space
- For every node in feature map, calculate the gradient to all the pixels in the image.



Q: how to find the best feature map?

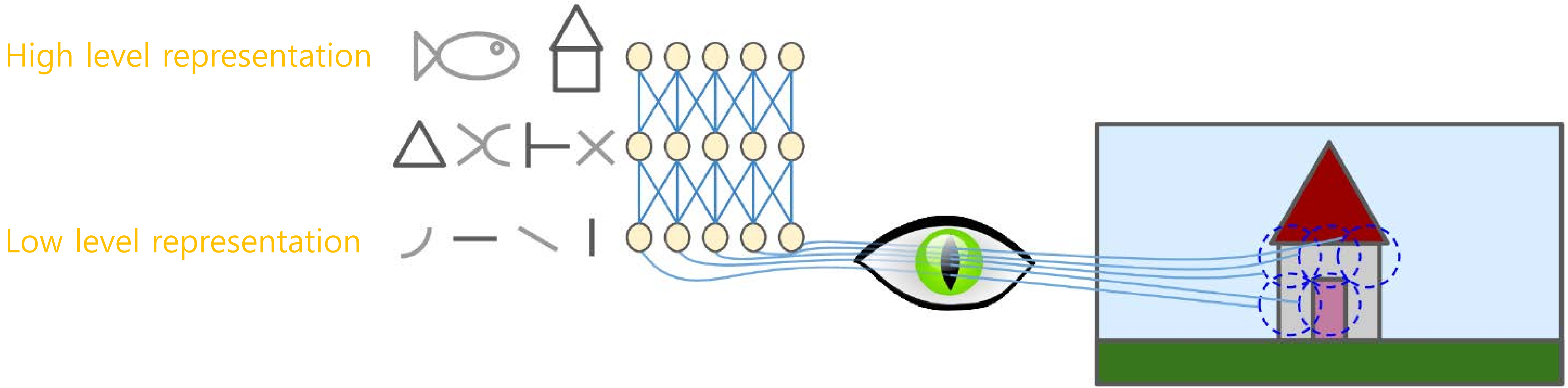


3 Method (part 2) – Feature Map Selection

- High level representation
- High resolution localization information

High level representation

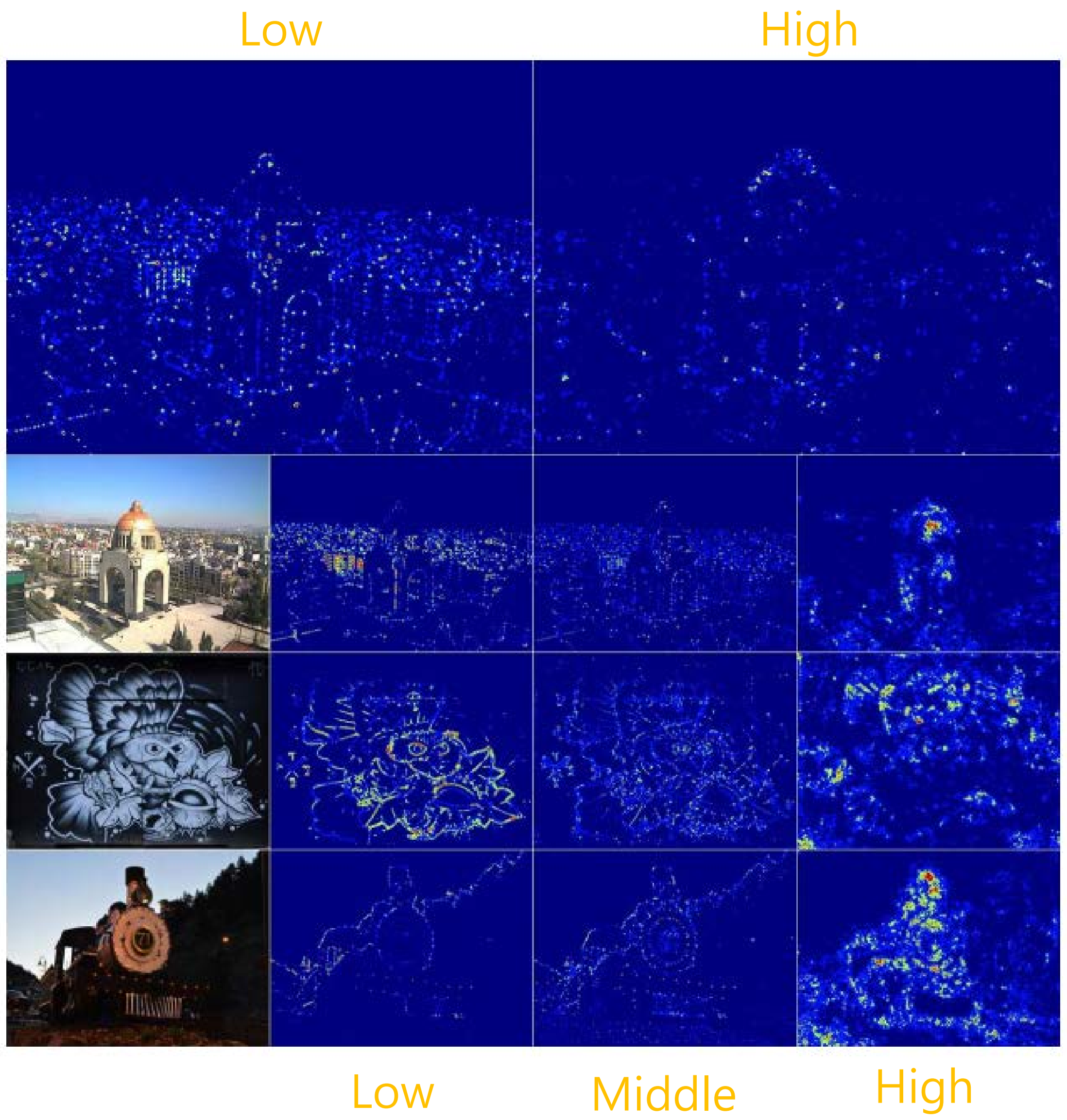
- To represent an image, the higher, the better.



High resolution localization information

- To find a accurate location, the lower, the better.

Low level saliency maps activate pixels more accurately.



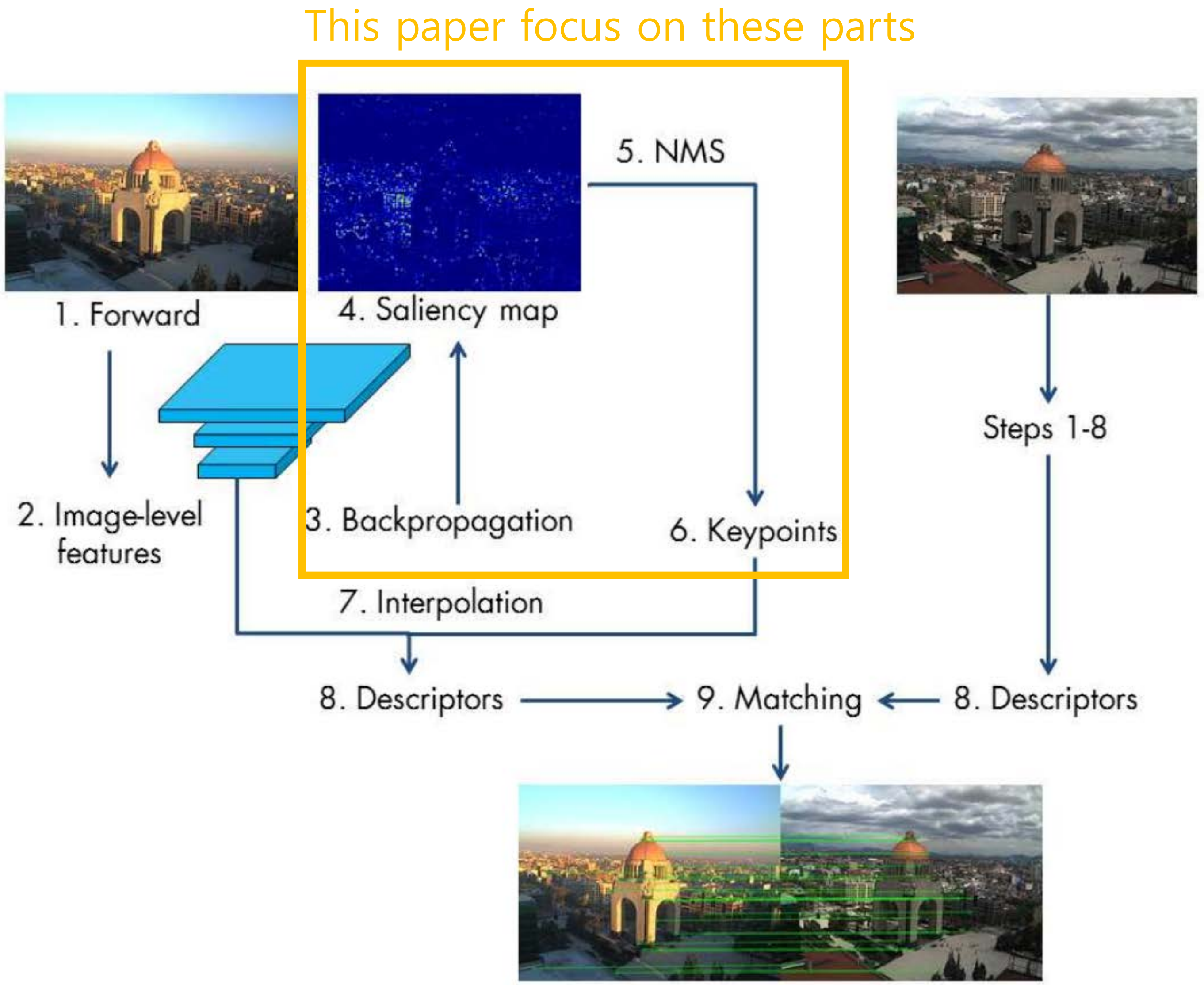
Summary for feature map selection

- High level representation → the higher, the better
- High resolution localization information → the lower, the better

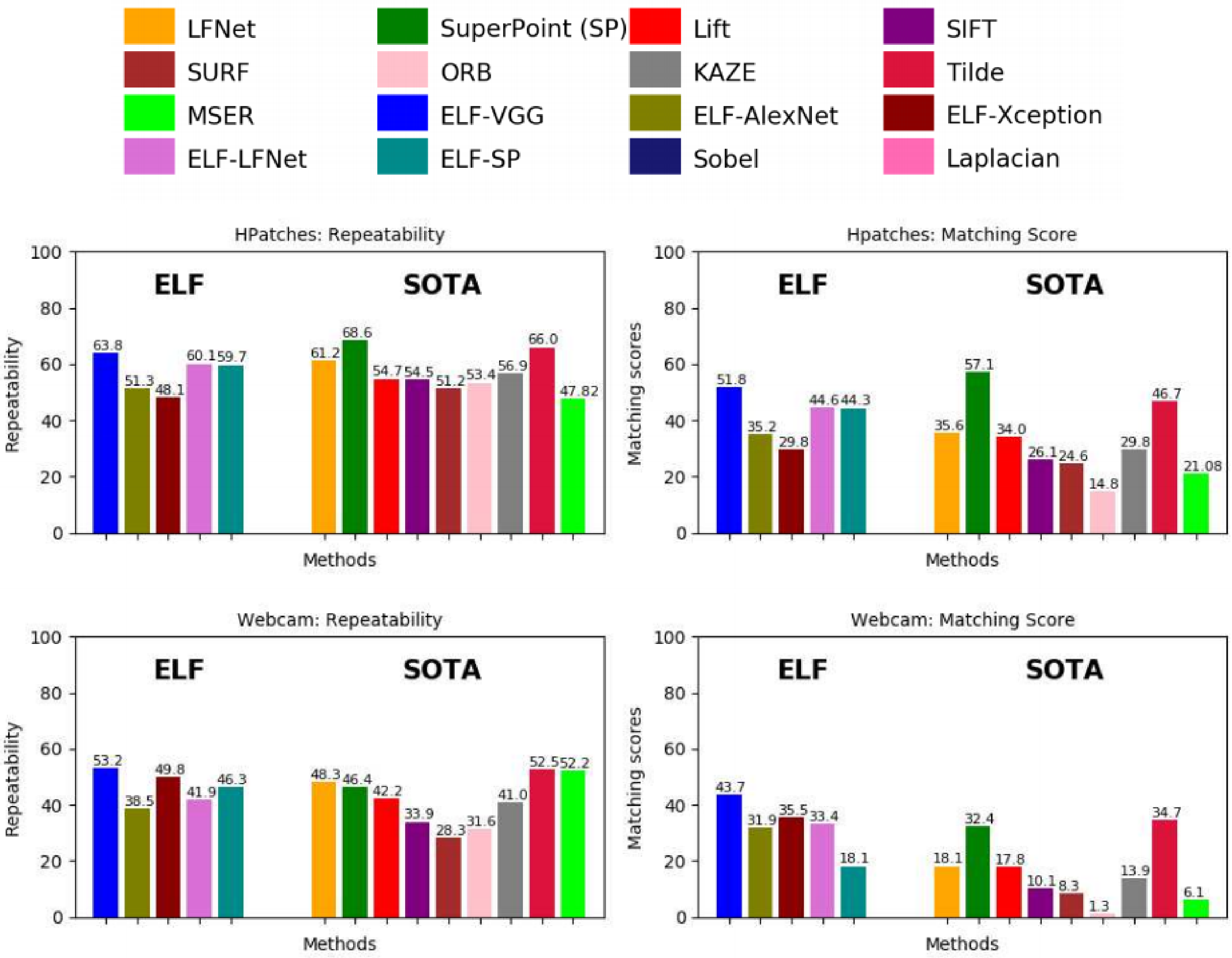
Solution: Visually observe the highest level which provides accurate localization.

4 Review and Result

Review


This paper focus on these parts
1. Forward
2. Image-level features
3. Backpropagation
4. Saliency map
5. NMS
6. Keypoints
7. Interpolation
8. Descriptors
9. Matching
8. Descriptors
Steps 1-8

Result



LFNet
SuperPoint (SP)
Lift
SIFT
SURF
ORB
KAZE
Tilde
MSER
ELF-VGG
ELF-AlexNet
ELF-Xception
ELF-LFNet
ELF-SP
Sobel
Laplacian
HPatches: Repeatability
Repeatability
ELF
SOTA
63.8
51.3
48.1
60.1
59.7
61.2
68.6
54.7
54.5
51.2
53.4
56.9
66.0
47.82
Methods
Hpatches: Matching Score
Matching scores
51.8
35.2
29.8
44.6
44.3
35.6
57.1
34.0
26.1
24.6
14.8
29.8
46.7
21.08
Webcam: Repeatability
53.2
38.5
49.8
41.9
46.3
48.3
46.4
42.2
33.9
28.3
31.6
41.0
52.5
52.2
Webcam: Matching Score
43.7
31.9
35.5
33.4
18.1
18.1
32.4
17.8
10.1
8.3
1.3
13.9
34.7
6.1
0
20
40
60
80
100

5 Discussion

Discussion

- Main Contribution:
  - Feature map based saliency map
  - Only use pre-trained CNN

- New directions
  - Harris on feature map.
  - Selecting the best feature map: SIFT-LoG