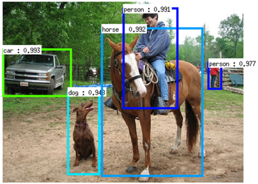
Object Detection using the Scattering Transform

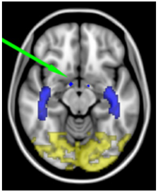
Marius Hobbhahn

February 18, 2019

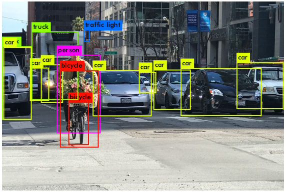
Object Detection



(A)



(B)



(C)

- Automatic annotation of everyday photos
- Brain disease recognition
- Traffic scene detection

Motivation

Problems

1. Loads of data necessary for training
2. Capability to generalize unclear for different circumstances (i.e. invariances w.r.t. many things)

Possible solutions

1. Filters that generalize quickly
2. Theoretical bounds for some invariances (i.e. translation, location)

Scattering Transform

Basic Idea

Static image filter that has certain theoretical guarantees with respect to invariances (i.e. location, scale, rotation).

$$\psi(u) = C_1(e^{iu.\xi} - C_2)e^{\frac{-|u|^2}{2\sigma^2}} \tag{1}$$

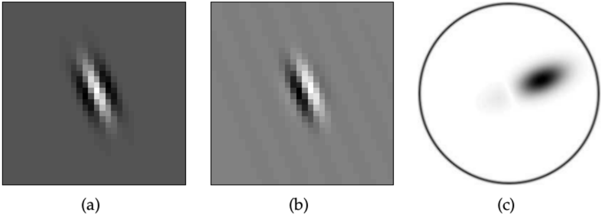


Figure 1: Complex morlet wavelet. a) Real part of ψ. b) Imaginary part of ψ. c) Fourier modulus $|\hat{\psi}|$.

Visualization of the filter bank

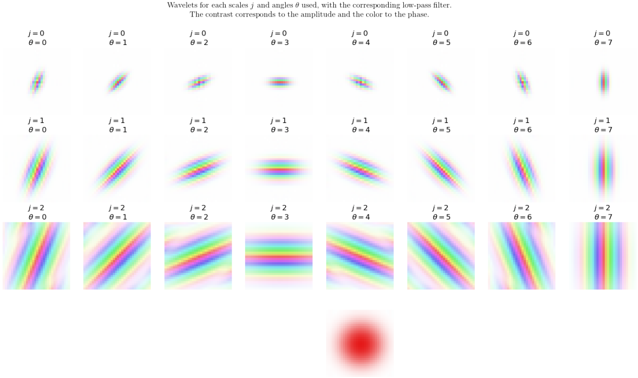


Figure 2: Visualization of the filter bank

Scattering Networks

Basic Idea

Apply the scattering transform multiple times to get higher order scattering coefficients.

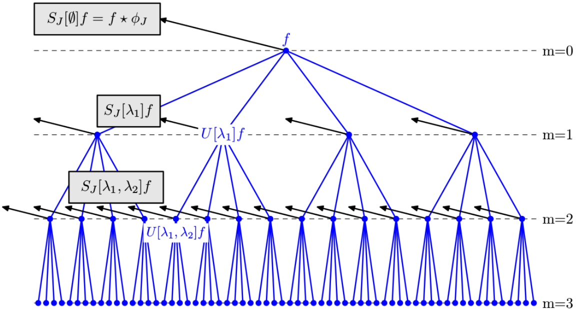


Figure 3

Visualization of the scattering coefficients

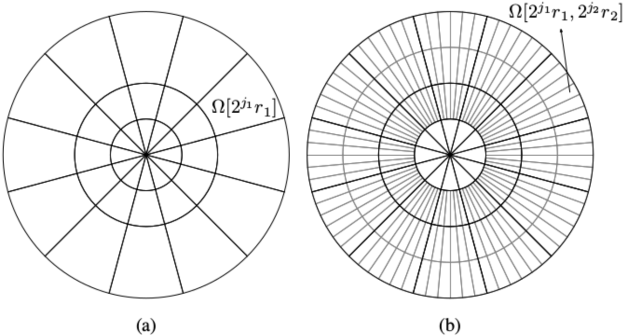


Figure 4: Visualization of the scattering coefficients for m=1 (left) and m=2 (right)

Example of the scattering coefficients

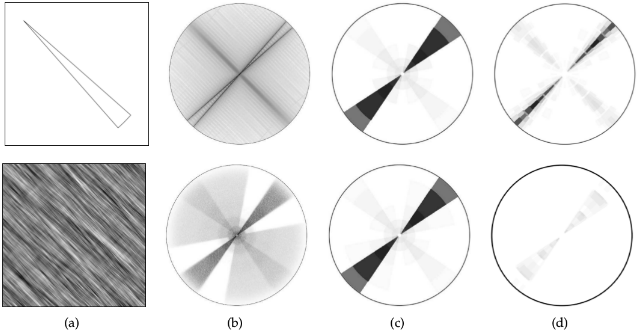


Figure 5: Scattering display of two images having the same first order scattering coefficients. a) Image, b) Fourrier modulus c) Coefficients with m=1, d) Coefficients with m=2

Hybrid scattering networks

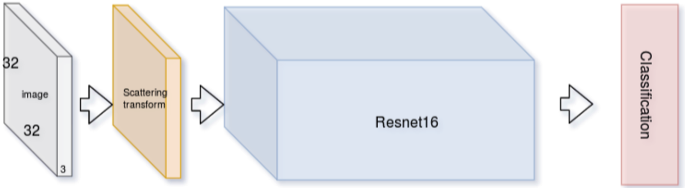


Figure 6: Architecture

Method	100	500	1000	Full
WRN 16-8	34.7 $\pm$ 0.8	46.5 $\pm$1.4	60.0 $\pm$1.8	**95.7**
VGG 16 [58]	25.5 $\pm$2.7	46.2 $\pm$ 2.6	56 $\pm$ 1.0	92.6
Scat + WRN	**38.9 $\pm$ 1.2**	**54.7$\pm$0.6**	**62.0$\pm$1.1**	93.1

Figure 7: Results of 2017 paper

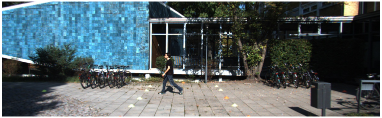
Datasets - Kitti

KITTI

- Content: street and traffic scenes
- number of samples: 7480

(A)

(B)



(C)

(D)

Datasets - VOC

VOC

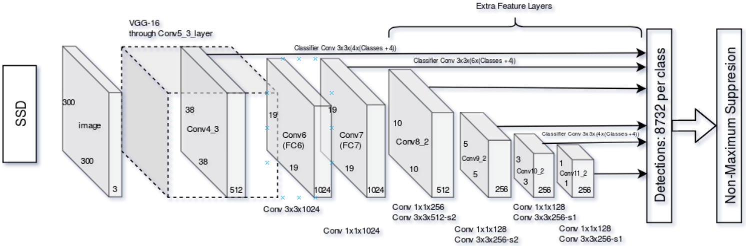
- miscellaneous (singular) objects
- VOC2007 and VOC2012
- number of samples: 9963 + 17125

(e)

(f)

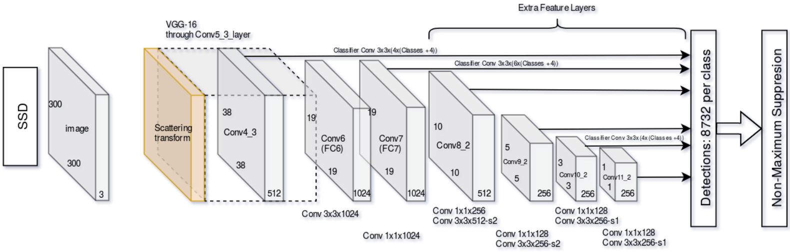
(g)

Simple Single Shot MultiBox Detector (SSD)



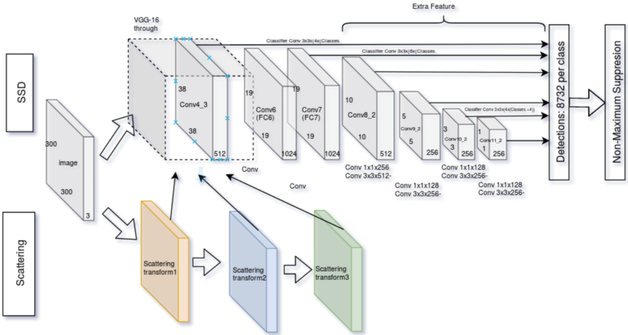
Sequential Scattering SSD

- Scattering is applied before data is piped through SSD



Continuous Fusion Scattering SSD

- Data is piped through scattering and standard SSD and continuously merged at different stages



Furthers plans and additional ideas

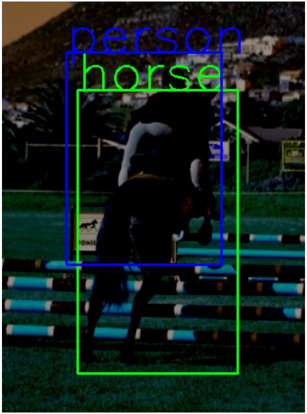
- try to achieve same baseline as other SSD implementations on kitti did: $\sim$ 50%
- use other base-architecture: faster RCNN, masked RCNN, ...
- integrate further datasets: COCO, FDDB, ...

Status Quo in pictures - Kitti

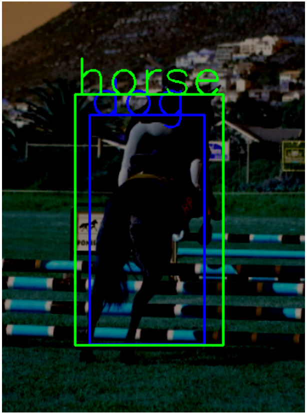
(h)

(i)

Status Quo in pictures - VOC + scattering



(j)



(k)

Status Quo in numbers

dataset	train accuracy	test accuracy
VOC	0.76	0.61
Kitti	0.82	0.12
kitti1000x300	-	0.05
Kitti small	0.95	0.0
Kitti small1000x300	0.95	0.0
Scattering Hybrid VOC	-	0.06

Questions and Suggestions

- Any questions?
- Any suggestions, tips, ... are very welcome

References

[BM12], [SM13], [OM14], [OBZ17], [ACC+17]

- Tameem Adel, Taco Cohen, Matthan Caan, Max Welling, On behalf of the AGEhIV study group Initiative, and the Alzheimer's Disease Neuroimaging.
  3d scattering transforms for disease classification in neuroimaging.
  NeuroImage: Clinical, 14:506–517, 2017.
  Exported from https://app.dimensions.ai on 2018/10/21.

- Joan Bruna and Stéphane Mallat.
  Invariant scattering convolution networks.
  CoRR, abs/1203.1513, 2012.

- Edouard Oyallon, Eugene Belilovsky, and Sergey Zagoruyko.
  Scaling the scattering transform: Deep hybrid networks.
  CoRR, abs/1703.08961, 2017.

- Edouard Oyallon and Stéphane Mallat.
  Deep roto-translation scattering for object classification.
  CoRR, abs/1412.8659, 2014.

- Laurent Sifre and Stephane Mallat.
  Rotation, scaling and deformation invariant scattering for texture discrimination.
  In *Proceedings of the 2013 IEEE Conference on Computer Vision and Pattern Recognition*, CVPR '13, pages 1233–1240, Washington, DC, USA, 2013. IEEE Computer Society.