OBJECT DETECTION USING THE SCATTERING TRANSFORM

Marius Hobbhahn

April 30, 2019

MOTIVATION

PROBLEMS

1. Loads of data necessary for training
2. Capability to generalize unclear for different circumstances (i.e. invariances w.r.t. translation, deformations, scale(?), rotation(?)

POSSIBLE SOLUTIONS

1. Filters that generalize quickly
2. Theoretical bounds for some invariances (translation) or Lipschitz continuity w.r.t deformations

SCATTERING TRANSFORM

BASIC IDEA

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

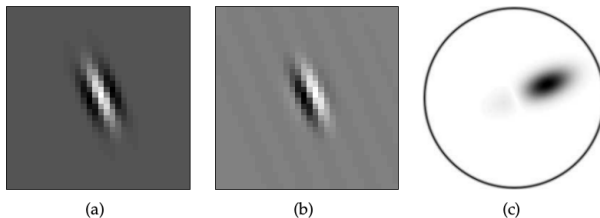$$\psi(u) = C_1(e^{iu.\xi} - C_2)e^{\frac{-|u|^2}{2\sigma^2}} \tag{1}$$

(a)   (b)   (c)

FIGURE 1: Complex morlet wavelet. a) Real part of ψ. b) Imaginary part of ψ. c) Fourier modulus $|\hat{\psi}|$.
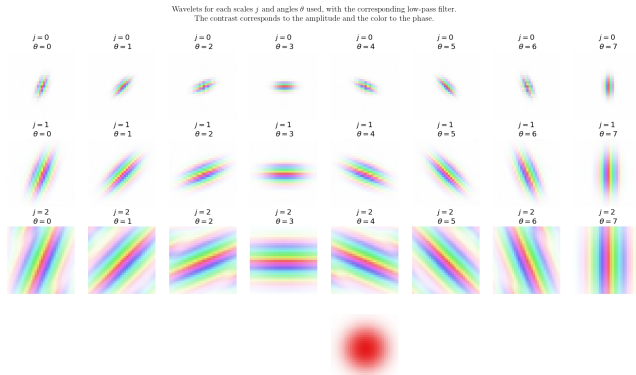
VISUALIZATION OF THE FILTER BANK

FIGURE 2: Visualization of the filter bank

SCATTERING NETWORKS

BASIC IDEA

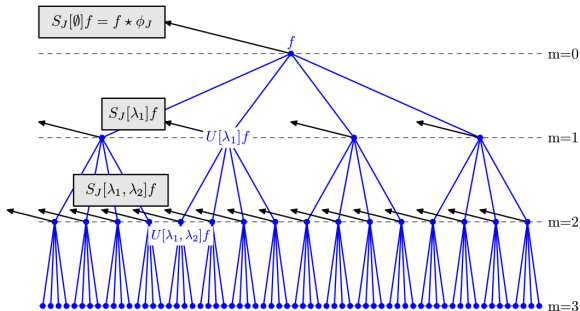Apply the scattering transform multiple times to get higher order scattering coefficients.

FIGURE 3

EXAMPLE

$$i \cdot (1 + JL) \tag{2}$$

$$i \cdot (1 + JL + \frac{1}{2}J(J-1)L^2) \tag{3}$$

- ▶ Let $J = 2, L = 8, N, M = 32, 32$ for a RGB image.
- ▶ number of outputs of the scattering network for $m = 1$:

$$3 \cdot (1 + 2 * 8) = 51$$

- ▶ number of outputs of the scattering network for $m = 2$:

$$3 \cdot (1 + 16 + 0.5 * 2 * 1 * 64) = 243$$

- ▶ all outputs of size 8x8
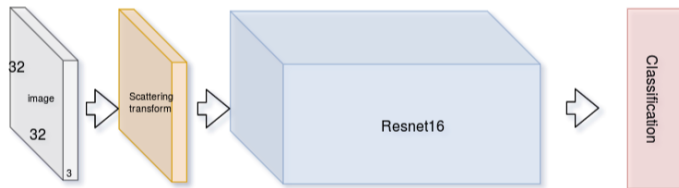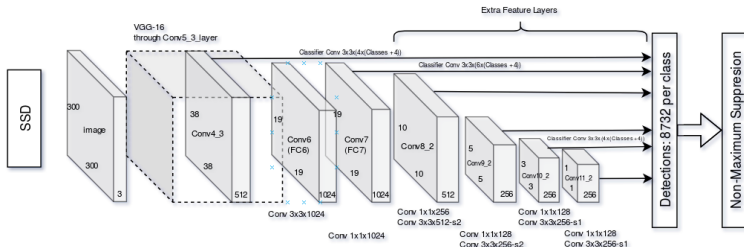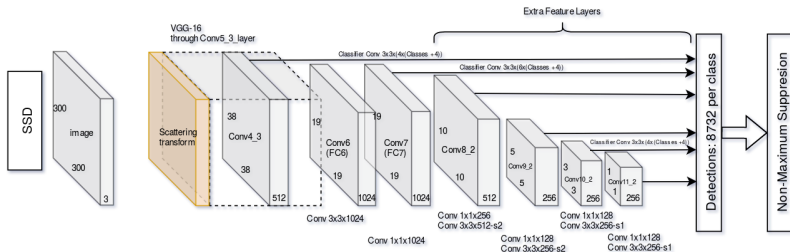
HYBRID SCATTERING NETWORKS

FIGURE 4: Architecture

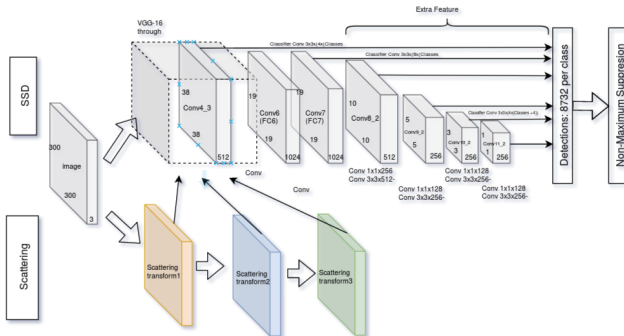# SIMPLE SINGLE SHOT MULTIBOX DETECTOR (SSD)

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

STATUS QUO

- ▶ experimented with different hyperparameters and features for baseline SSD (pretrained, batch norm, augmentations, etc.) → still very bad results for kitti
- ▶ implemented sequential and parallel scattering → very high variance within the results
- ▶ created toy datasets to test invariances

TOY DATASETS

IDEA

Test if the promised invariants or continuities hold in empirical experiments

- ▶ Transformation dataset
- ▶ Scale dataset
- ▶ Rotation dataset
- ▶ Deformation dataset

TRANSFORMATION DATASET

(A)　　　　　(B)　　　　　(C)　　　　　(D)　　　　　(E)

(F)　　　　　(G)　　　　　(H)　　　　　(I)　　　　　(J)

SCALE DATASET

(K) (L) (M) (N) (O)

(P) (Q) (R) (S) (T)

Rotation dataset

(u) (v) (w) (x) (y)

(z) (aa) (ab) (ac) (ad)

DEFORMATION DATASET

(AE)   (AF)   (AG)   (AH)   (AI)

(AJ)   (AK)   (AL)   (AM)   (AN)

FURTHERS PLANS AND ADDITIONAL IDEAS

- ▶ do baseline and scattering experiments on the currently existing datasets
- ▶ test pretrained and batch norm for scattering nets
- ▶ use other base-architecture: faster RCNN, masked RCNN, ... (trying to get the detectorch framework to run)

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.