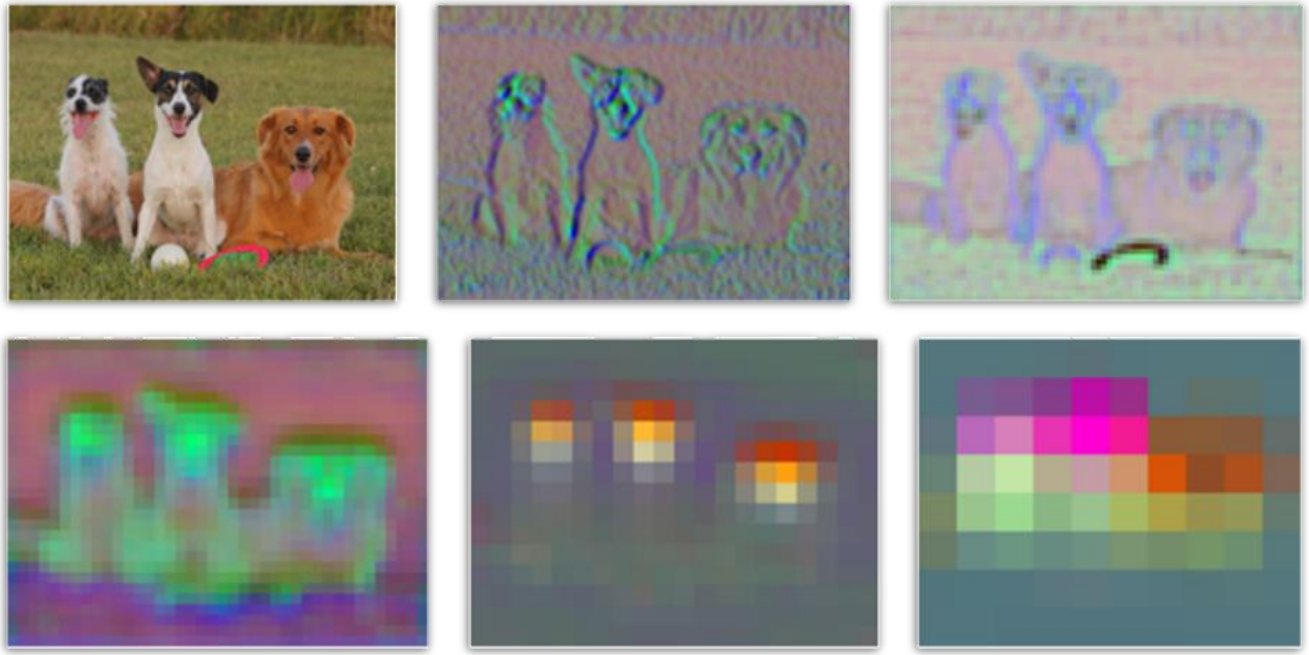
Modelling 2

STATISTICAL DATA MODELLING



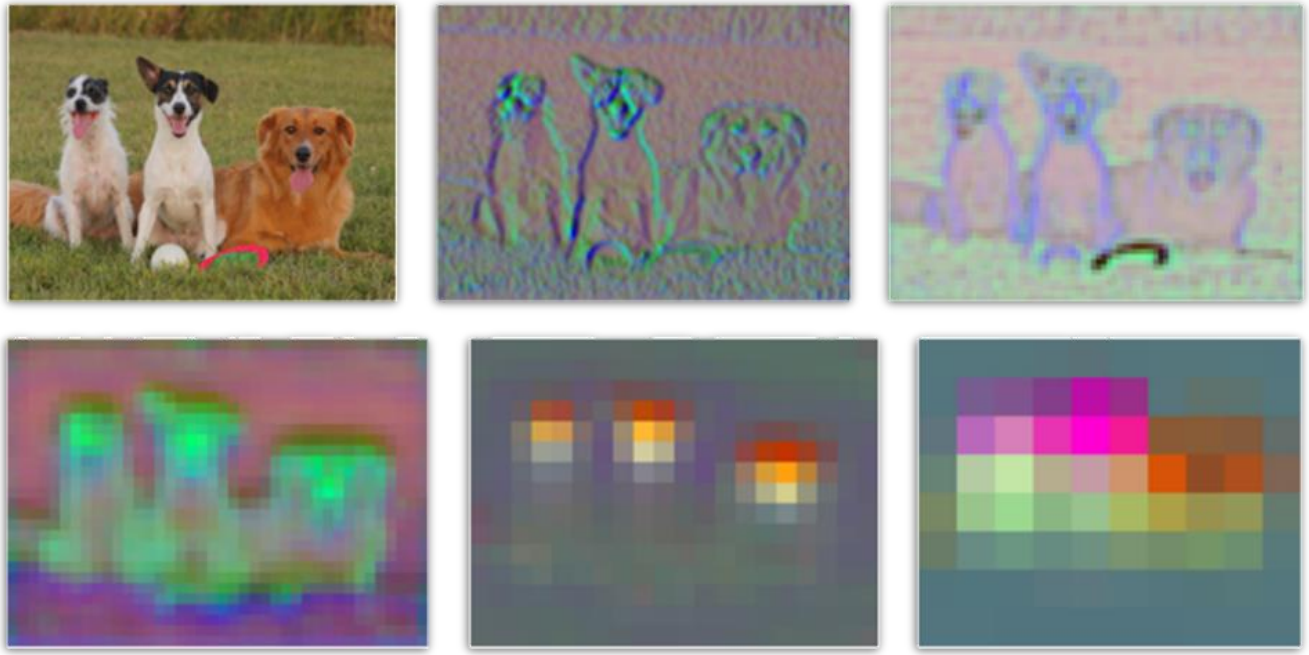
Chapter 10+ Intro

Overview of Part II

Michael Wand · Institut für Informatik, JGU Mainz · michael.wand@uni-mainz.de

Modelling 2

STATISTICAL DATA MODELLING

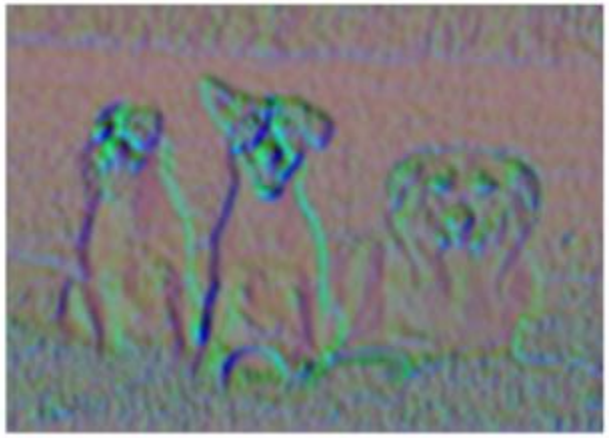
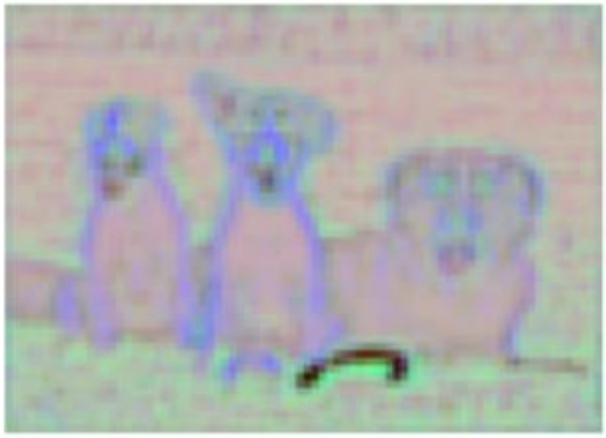
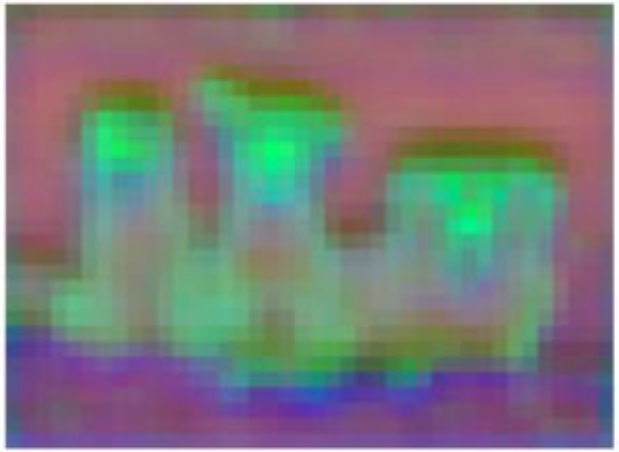
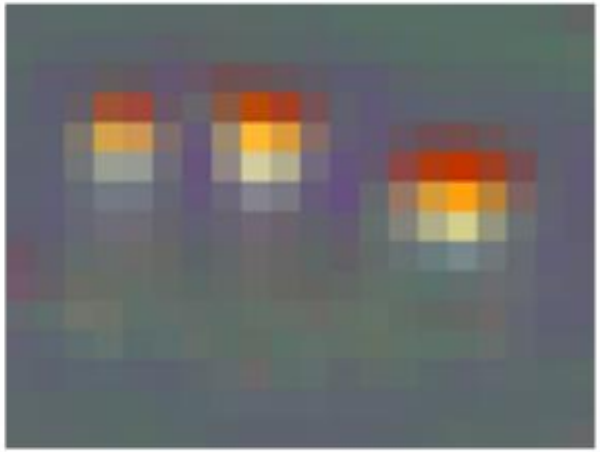
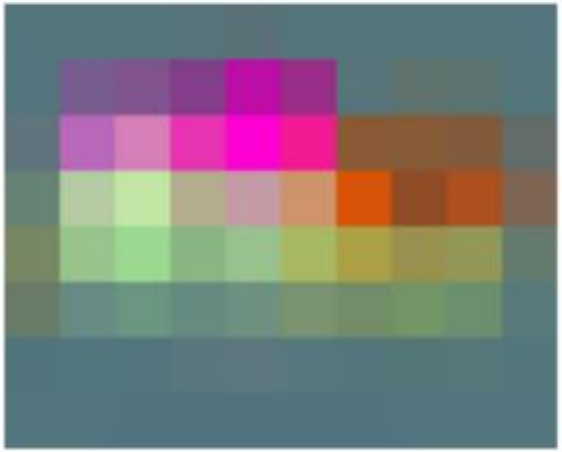


Chapter 10+ Intro

WTH is going on?

Michael Wand · Institut für Informatik, JGU Mainz · michael.wand@uni-mainz.de

Dogs

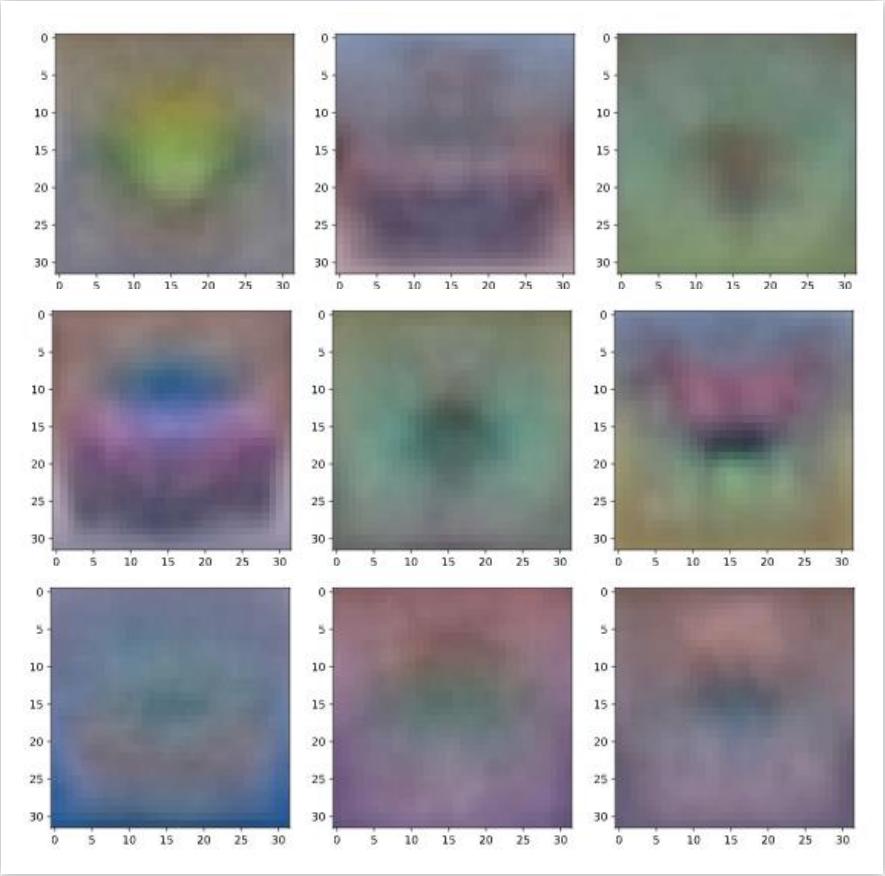
PCA-Embedding of activation patterns of a VGG network
(input & output after pooling layer 1,2,3,4,5)

Object Detection

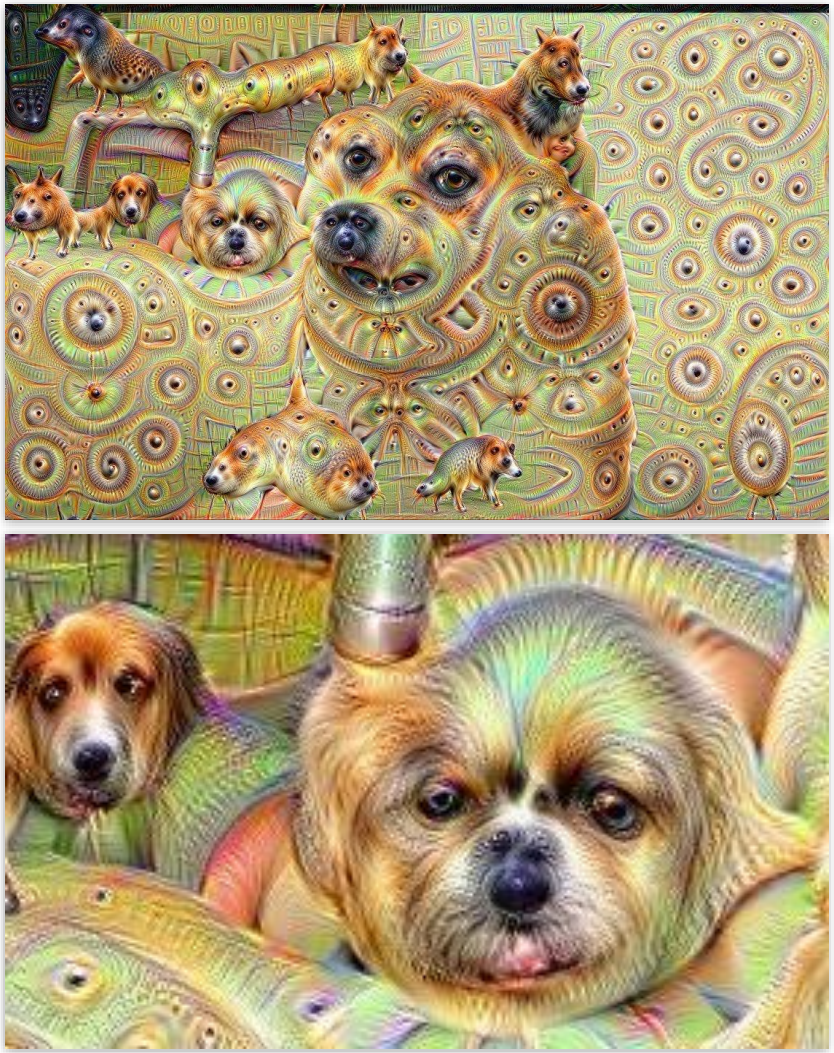
PASCAL-VOC 2007

- “Small” data set
  - 20 categories, 12 000 training images
- “Deformable Parts Model”: mAP ~35%
  - SOTA in 2008
  - PASCAL VOC "Lifetime Achievement" Prize in 2010
  - Cascade Eff-B7 NAS-FPN: mAP ~89% (SOTA 2021)
- DPMv5 (2012): Worst categories
  - “Potted plants”: 11%, “Dog” 23%, “Cat” 27%

Dreams
of a Classifier



Linear SVM



Google Inception CNN

SVN: LibLinear on CIFAR-10
Deep Dream: [https://artofericwayne.com/2015/07/08/google-deep-dream-getting-too-good]

Style-GAN
[Karras et al. 2018]

PCA
[Turk & Pentland 1987]

Tero Karras, Samuli Laine, Timo Aila:
A Style-Based Generator Architecture for Generative Adversarial Networks, 2018

[StyleGan result Wikipedia user OwlsMcGee, https://en.wikipedia.org/wiki/File:Woman_1.jpg]

[PCA results courtesy of D. Schwarz, D. Klaus, A. Rübe]

Open Questions

(also: to remain open after the lecture)

Mysteries about Deep-NNs

Why does it work so well?

- Bias-Variance-Trade-Off, NFL-Theorem:
  The network must encode strong prior knowledge
- What is the inductive bias of deep learning?

Why does it generalize robustly?

- Limited tendency towards overfitting
- Architectures do not need to much fine-tuning

Why can we optimize Deep-NNs?

- Highly non-convex objective, but solutions are good

Mysteries about Deep-NNs

Why

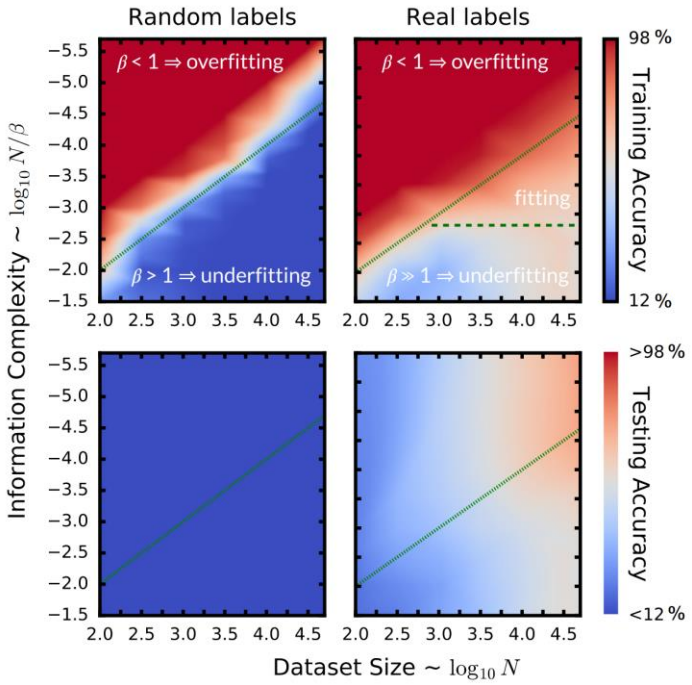
- Bi
  Th
- Wh

Why

- Li
- Ar

Why

- Highly non-convex objective, but solutions are good



Alessandro Achille, Stefano Soatto
Emergence of Invariance and Disentanglement in Deep Representations
Journal of Machine Learning Research 18 (2018) 1-34
(Figure 1, CC-BY 4.0)

Mysteries about Deep-NNs

Why does it work so well?

- Bias-Variance-Trade-Off, NFL-Theorem:
  The network must encode strong prior knowledge
- What is the inductive bias of deep learning?

Why does it generalize robustly?

- Limited tendency towards overfitting
- Architectures do not need to much fine-tuning

Why can we optimize Deep-NNs?

- Highly non-convex objective, but solutions are good

Part II – Understanding (Deep) Learning

Methods for analyzing

- Complex systems
- Deep neural networks

(Planned) Topics

- Manifolds & the curse of dimensionality
- Symmetry & group theory
- Physical dynamics & self-organization
- Information theory & bottlenecks
- (Initialization and training)
- (Variational approximations)