Rediscovering ICA properties in neural networks applied to mixtures of signals problems

INSTITUTE FOR OPTICAL SYSTEMS

HTWG Hochschule Konstanz University of Applied Sciences

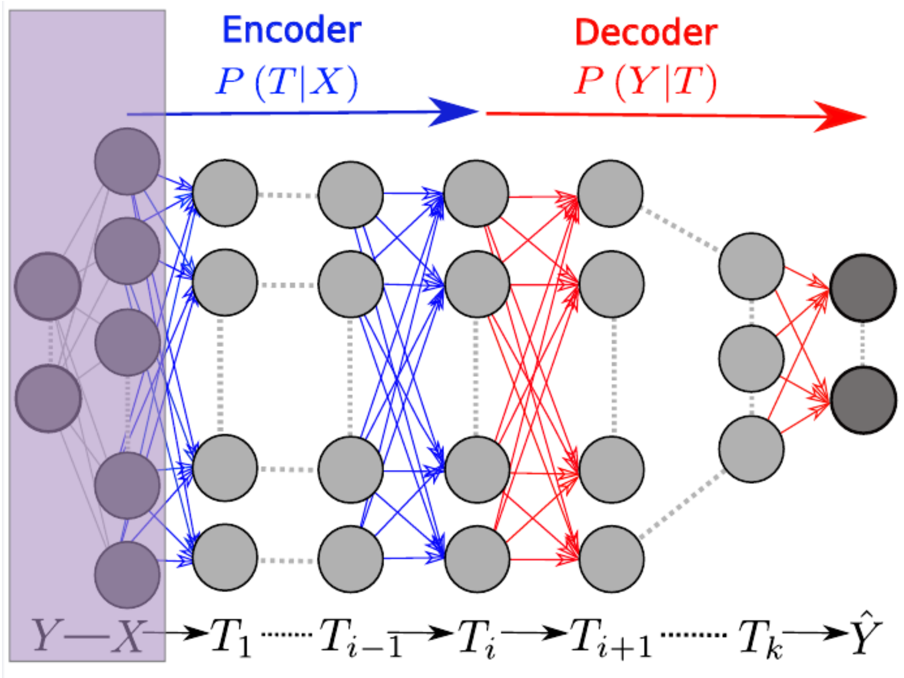
Matthias Hermann,
Michael Grunwald, Thomas Gnädig and Matthias O. Franz

HTWG Konstanz
Institute for Optical Systems

Deep Learning Day 2018

14 September 2018

INSTITUTE FOR OPTICAL SYSTEMS



Introduction

Neural networks are black boxes

Tishby and Zaslavsky (2015) describe neural networks as Markov chains of successive representations of the input layer and suggested studying them with information theory.

Analyzing ICA properties

Relevance

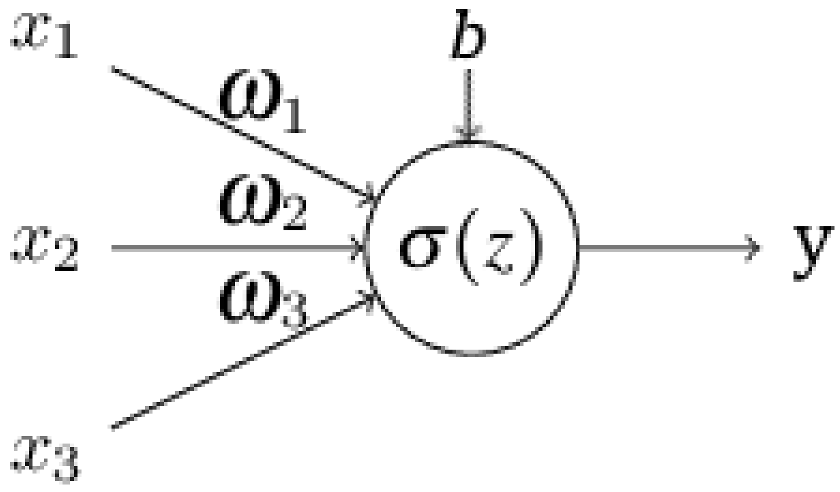
- Knowledge about statistical properties of neural networks enables new **initialization**, visualization, regularization or **optimization** techniques.
- Special focus is on initialization or pre-training in cases with few data labels which is useful in order to train **unsupervised** or enable **transfer learning**.

Hypothesis

In order to perform non-linear classification or regression neural networks must **implicitly** conduct an **Independent Component Analysis** (ICA).

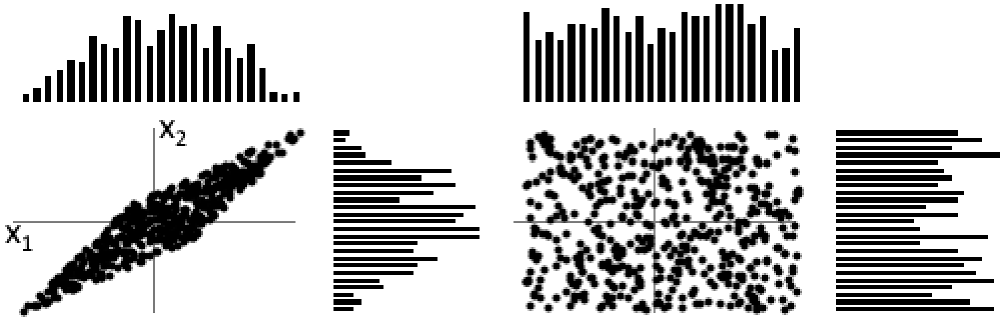
Methods

Neural Networks



Independent Component Analysis

Non-linear technique for decorrelating independent signals



Criterion

- Maximize Entropy:

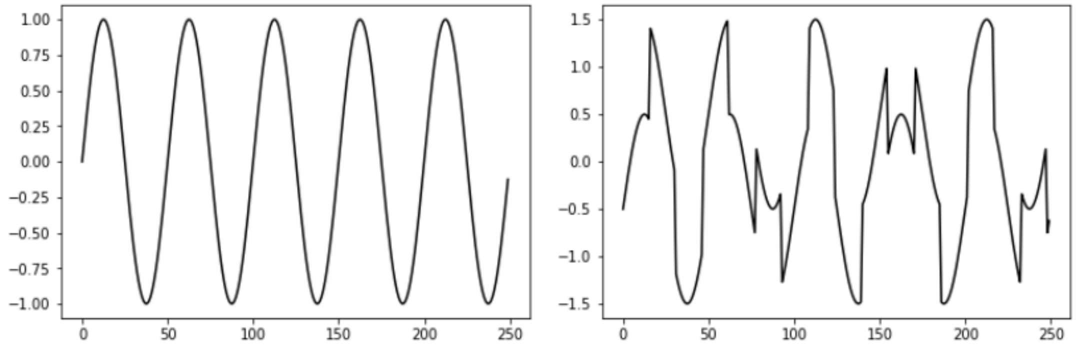
$$H(X) = -\sum_{x\,e\,X} p(x) log p(x)$$

- Maximize Kurtosis:

$$Kurz[X] = E\left[(\frac{X-\mu}{\sigma})^4\right] - 3$$

Results “*Binary mixed signal classification*”

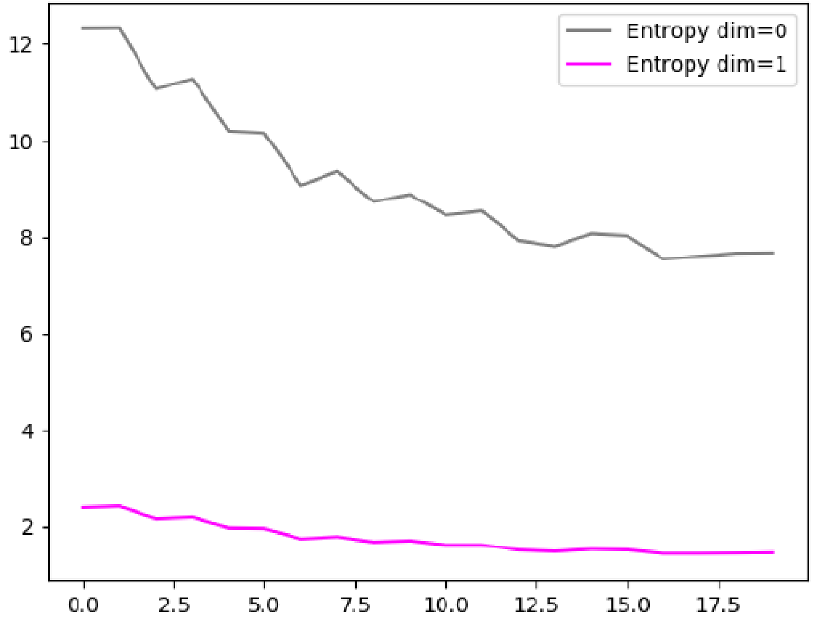
Training data

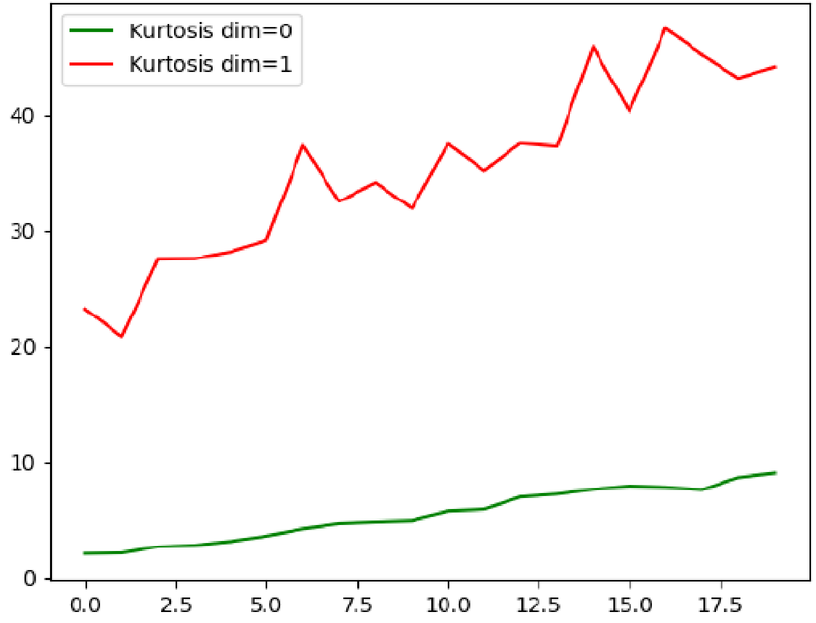


Generalization power

	sin	shark	rect	pyramid	rrect	random
sin	1,00	1,00	1,00	1,00	1,00	0,50
shark	1,00	1,00	1,00	1,00	1,00	0,50
rect	1,00	1,00	1,00	1,00	1,00	0,50
pyramid	1,00	1,00	1,00	1,00	1,00	0,50
rrect	0,99	0,99	0,99	0,99	0,99	0,49
random	0,50	0,50	0,50	0,50	0,50	0,50

Results “*Entropy and Kurtosis of a MNISTS classifier*”





Entropy decreases during training indicating non uniform data distribution
→ Entropy not well suited for measuring continuous random variables (binning)

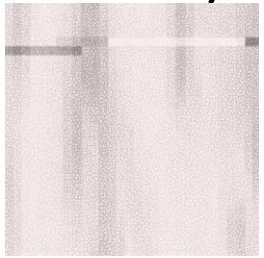
Kurtosis i.e. non-Gaussianity increases during training indicating ICA criterion
→ Comparison in non-linear space difficult

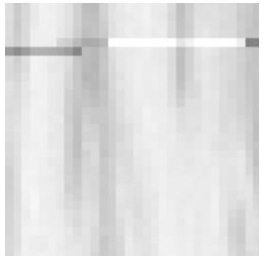
IOS
INSTITUTE FOR OPTICAL SYSTEMS

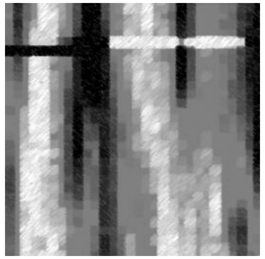
Research plan

Future work

- Adding negentropy as a more robust measure for non-Gaussianity
- Using natural images with added distortions instead of mixtures of signals
- Formulate ICA as decision problem
- Analyze mutual information between input and output







… why do things look as they do ?
K. Koffka, Principles of Gestalt Psychology, 1935

Thank you for your attention!

Matthias Hermann

Institute for Optical Systems
HTWG Konstanz

matthias.hermann@htwg-konstanz.de

www.ios.htwg-konstanz.de

IOS
INSTITUTE FOR OPTICAL SYSTEMS

References

- Schwartz-Ziv, Ravid and Naftali Tishby (2017). Opening the black box of Deep Neural Networks via Information. ICRI-CI.
- Tishby, Naftali and Noga Zaslavsky (2015). Deep learning and the information bottleneck principle. IEEE Information Theor Workshop.

IOS
INSTITUTE FOR OPTICAL SYSTEMS