Outlook: Deep Learning

Statistical Methods for Machine Learning

Christian Igel
Department of Computer Science

Outline

Outline

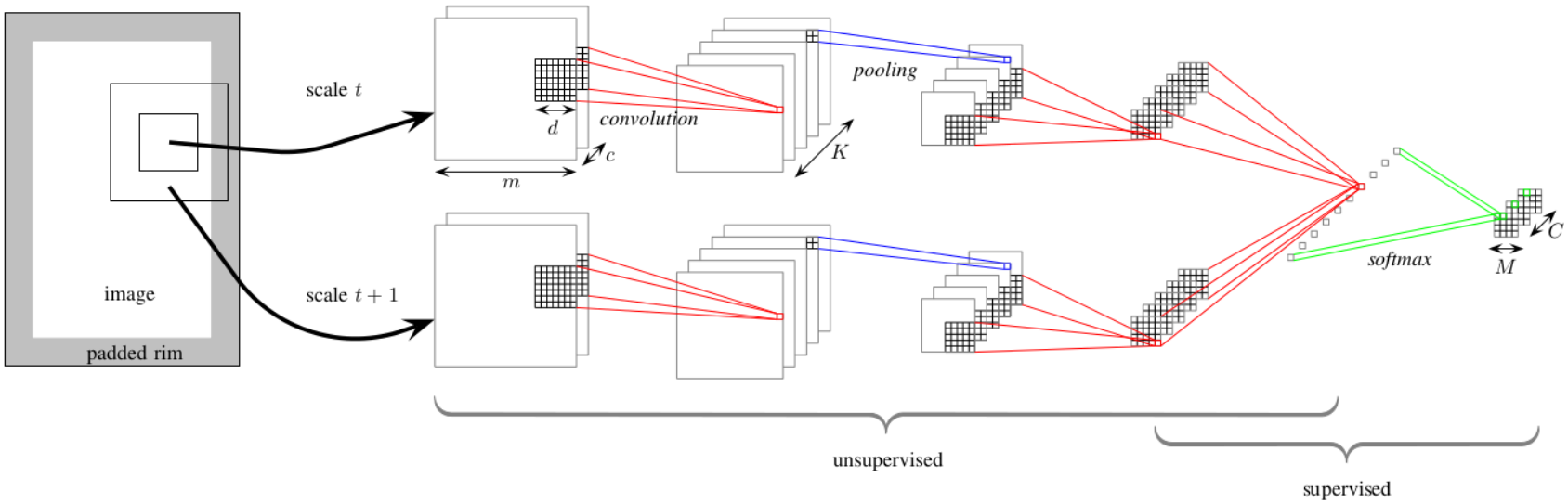
Deep learning

- Deep learning refers to machine learning architectures composed of multiple levels of non-linear transformations.
- The idea is to extract more and more abstract features from input data, to learn more abstract representations.
- The representations can be learned in a supervised and/or unsupervised manner.
- Example: Convolutional neural networks (CNNs) are popular deep learning architectures.

LeCun, Bottou, Bengio, Haffner. Gradient-based learning applied to document recognition. *Proceedings of the IEEE*, 1998

Convolutional neural network (CNN)

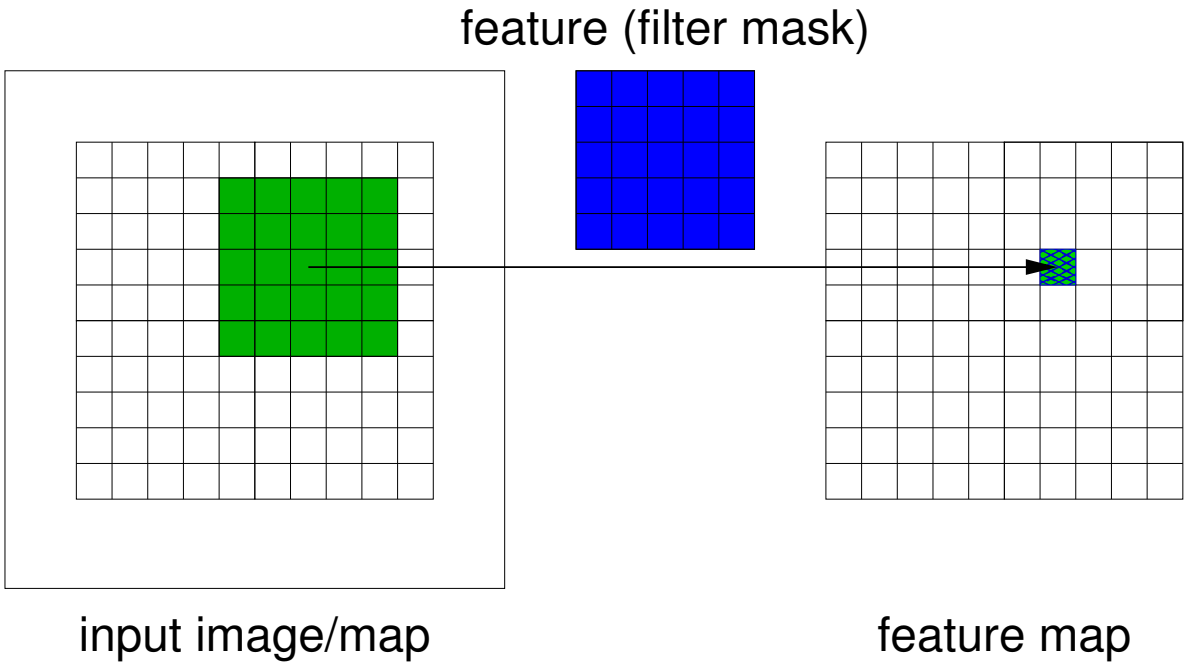
Example of a special CNN architecture we use:



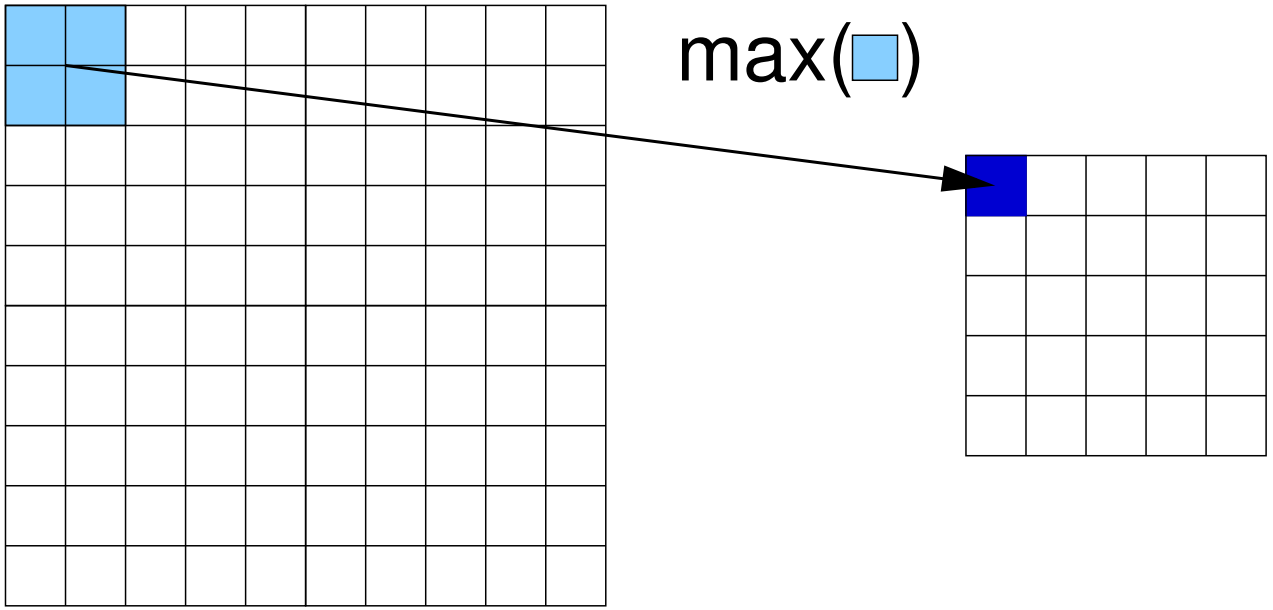
A canonical CNN consists of

- convolutional layers, each of which producing several *feature maps*; interleaved with
- pooling layers; and
- a standard neural network on top.

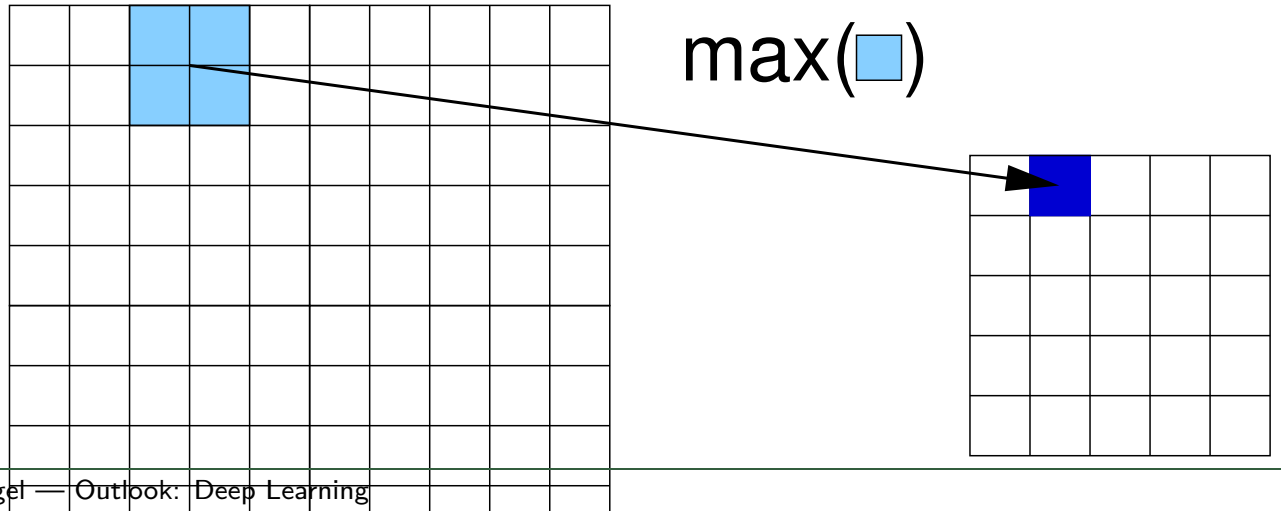
Convolutional layer



Pooling layer



input map



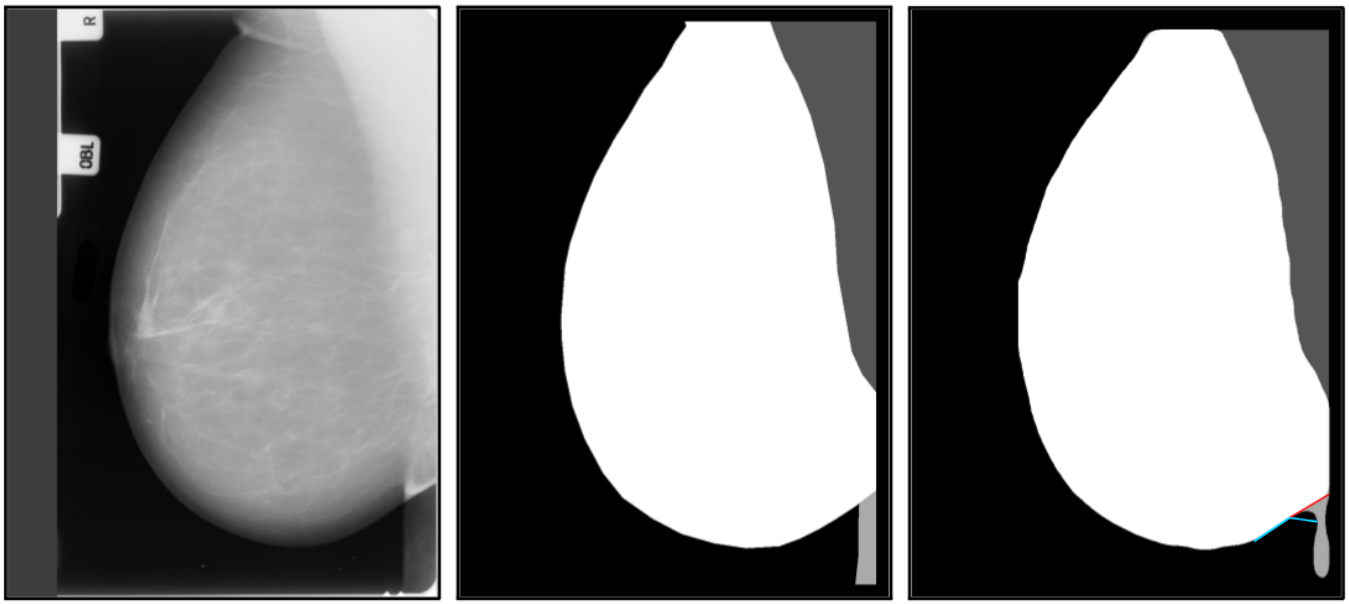
Breast cancer

- Breast cancer is the most frequent cause of death among women
- Screening programs halve the risk of death[1] and reduce mortality by 28-36 %[2]
- Still: 33 % of cancers are missed,[3] 70 % of referrals are false positives,[4] and 25 % of cancers could have been detected earlier[5]
- Goal: More accurate image-based biomarkers allowing personalized breast cancer screening

[1] Otto et al. *Cancer Epidemiol Biomarkers Prev* 21, 2012
[2] Broeders et al. *J Med Screen* 19, 2012
[3] Karssemeijer et al. *Radiology* 227, 2003
[4] Yankaskas et al. *Am J Roentgenol* 177, 2001
[5] Timmers et al. *Eur J Public Health*, 2012

Breast segmentation

Breast segmentation is the first step in the analysis.



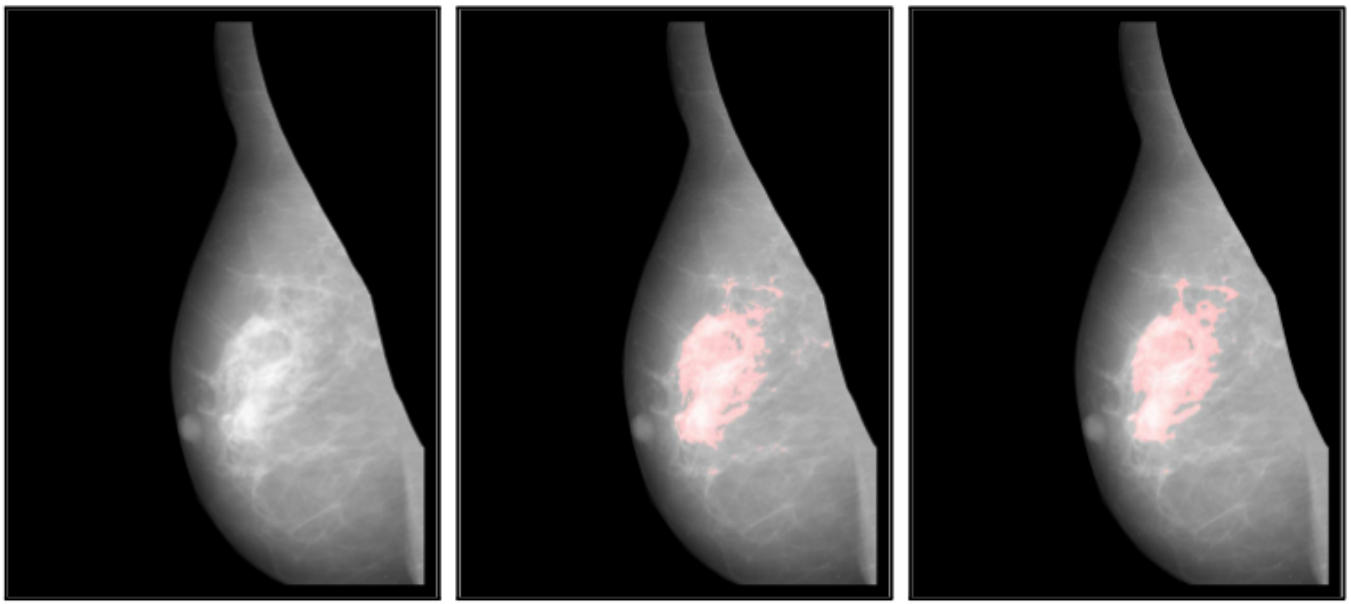
input manual CNN

Dataset	Background	Pectoral muscle	Breast tissue
Nijmegen ($N = 495$)	0.99 ± 0.01	0.95 ± 0.08	0.98 ± 0.01
Mayo ($N = 717$)	0.99 ± 0.01	0.93 ± 0.11	0.97 ± 0.02

Dice's coefficients

Breast density scoring

Breast density is related to breast cancer risk, the risk of missing breast cancer, and the risk of false positive referral.



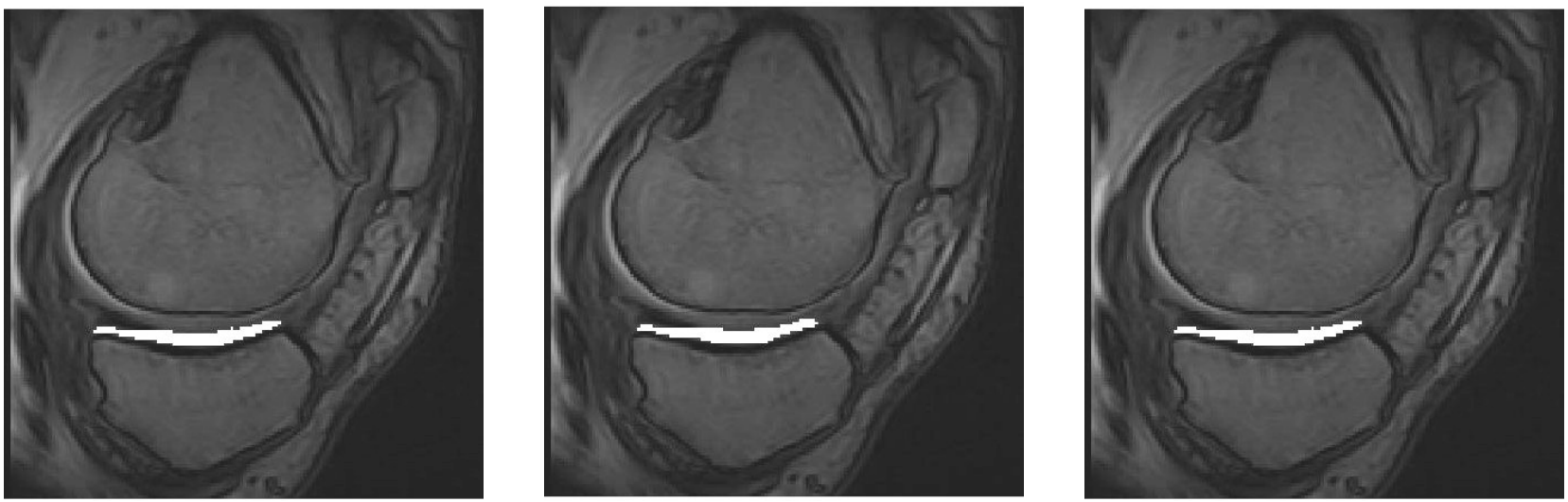
input manual CNN

Current work: Better biomarkers using CNN density scores and CNN texture scores

Petersen, Nielsen, Diao, Karssemeijer, Lillholm. Breast Tissue Segmentation and Mammographic Risk Scoring Using Deep Learning. *Digital Mammography / IWDM*, 2014

Knee cartilage segmentation

- Cartilage segmentation in knee MRI is the method of choice for quantifying cartilage deterioration.
- Cartilage deterioration implies osteoarthritis, one of the major reasons for work disability in the western world.



input manual CNN

Prasoon, Petersen, Igel, Lauze, Dam, Nielsen. Deep Feature Learning for Knee Cartilage Segmentation Using a Triplanar Convolutional Neural Network. *Medical Image Computing and Computer Assisted Intervention (MICCAI)*, LNCS 8150, 2013

Outline

Summary

- ⊕ CNNs are well suited for image and sound processing; they have won several competitions recently.
- ⊕ CNNs exploit "spatial" structure in the input; in particular, they incorporate translation invariance.
- ⊕ Layers can be (pre-)trained sequentially.
- ⊕ Unsupervised and supervised learning can be combined.
- ⊕ Training CNNs typically scales linearly in training set size.
- ⊕ CNNs profit from massively parallel computing (e.g., GPUs)
- ⊕/⊖ Deep networks can represent highly non-linear models.
- ⊖ Tuning the architecture of a deep learning system can be very difficult.
- ⊖ Deep learning is badly understood theoretically.

The revival of deep neural networks is partly due to the availability of fast (parallel) hardware and larger training data sets (which makes overfitting less of a problem).

Some deep learning reviews

Bengio. Learning Deep Architectures for AI. *Foundations and Trends in Machine Learning* 2(1): 1–127, 2009

Deng, Yu. Deep Learning: Methods and Applications. *Foundations and Trends in Signal Processing* 7(3-4):197-387, 2013

Schmidhuber. Deep Learning in Neural Networks: An Overview. *Neural Networks* 61:85–117, 2015