

Introduction to Deep Learning

Chapter 4: CNN: Introduction

Bernd Bischl

Department of Statistics – LMU Munich

WS 2021/2022



CONVOLUTIONAL NEURAL NETWORKS

- Convolutional Neural Networks (CNN, or ConvNet) are a powerful family of neural networks that are inspired by biological process in which the connectivity pattern between neurons resembles the organization of the mammal visual cortex.

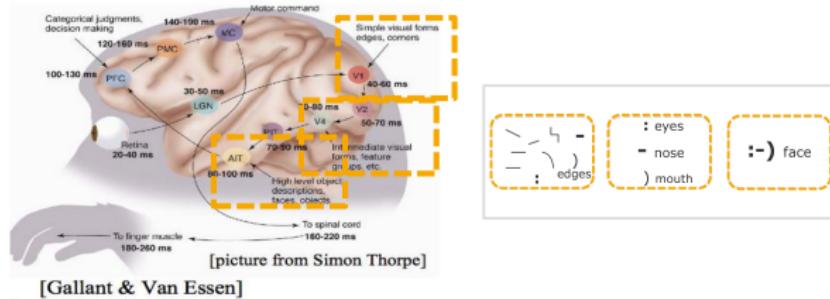


Figure: The ventral (recognition) pathway in the visual cortex has multiple stages: Retina - LGN - V1 - V2 - V4 - PIT - AIT, lots of intermediate representations

CONVOLUTIONAL NEURAL NETWORKS

- Since 2012, by the success of CNNs model in ImageNet competition, they are state-of-the-art in many fields such as computer vision.
- Common applications of CNN-based architectures in computer vision are:
 - Image classification.
 - Object detection /localization.
 - Semantic segmentation.
- Also widely applied in other domains such as natural language processing (NLP), Speech and even time-series data.
- Basic idea: a CNN automatically extracts visual, or, more generally, spatial features from an input such that it is able to make the optimal prediction based on those features.
- Therefore, it contains different building blocks...

CNNs - WHAT FOR?



Figure: All Tesla cars being produced now have full self-driving hardware and customers can purchase one today (source Tesla website) .Twenty-nine states from fifty states in the U.S. have enacted legislation related to autonomous vehicles. A convolutional neural network is used to map raw pixels from a single front-facing camera directly into steering commands. With minimum training data from humans, the system learns to drive in traffic on local roads with or without lane markings and on highways.

CNNs - WHAT FOR?

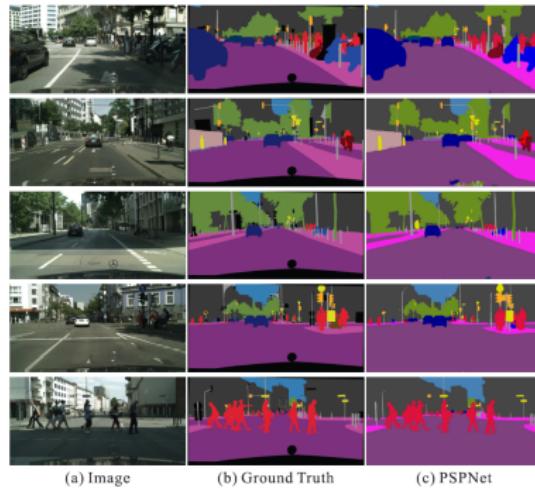


Figure: Given an input image, first CNN is used to get the feature map of the last convolutional layer, then a pyramid parsing module is applied to harvest different sub-region representations, followed by upsampling and concatenation layers to form the final feature representation, which carries both local and global context information. Finally, the representation is fed into a convolution layer to get the final per-pixel prediction. (Source: pyramid scene parsing network, by Zhao et. al, CVPR 2017)

CNNs - WHAT FOR?



Figure: Road segmentation (Mnih Volodymyr (2013)). Aerial images and possibly outdated map pixels are labeled.

CNNs - WHAT FOR?

CNN for personalized medicine: tracking, diagnosis and localization of Covid-19 patient

- CNN based model (RADLogists) for personalized Covid-19 detection: three CT scans from a single coronavirus patient diagnosed by RADLogists

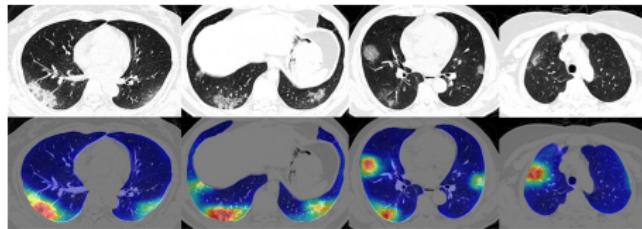
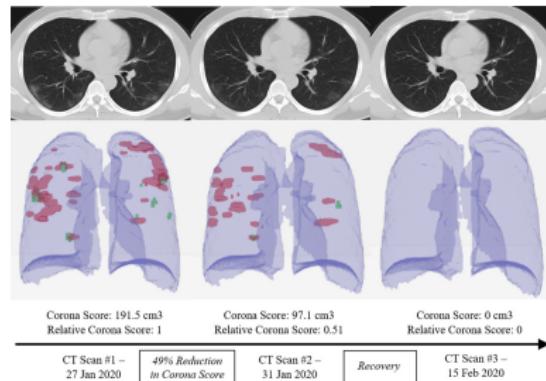


Figure: Four COVID-19 lung CT scans (top) with corresponding colored maps showing coronavirus abnormalities (bottom). (source: IEEE Spectrum)

CNNs - WHAT FOR?

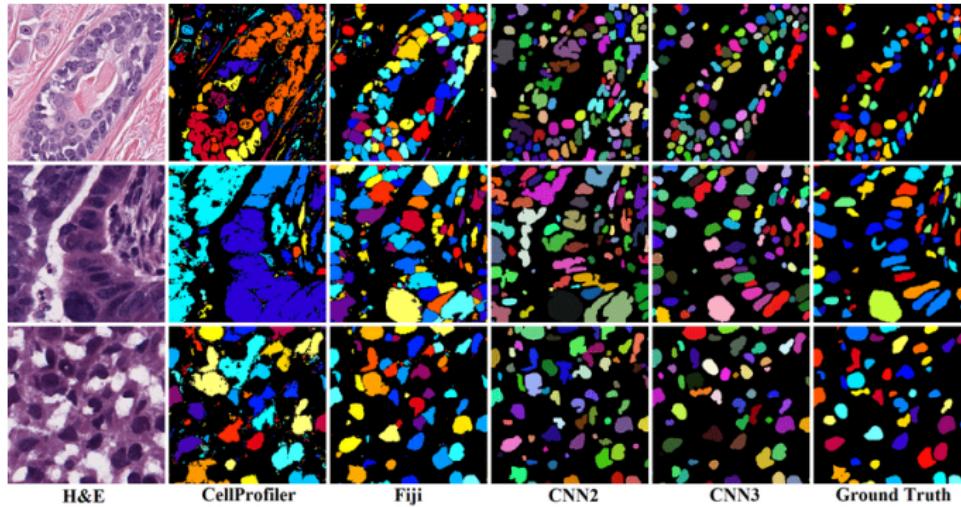


Figure: Nuclear segmentation in digital microscopic tissue images can enable extraction of high-quality features for nuclear morphometrics and other analysis in computational pathology. (source: Kummar et. al. IEEE Transaction Medical Imaging)

CNNs - WHAT FOR?



Figure: Colorful Image Colorization (Zhang et al. (2016)). Given a grayscale photograph as input (top row), this network attacks the problem of hallucinating a plausible color version of the photograph (bottom row, i.e. the prediction of the network). Realizing this task manually consumes many hours of time.

CNNs - WHAT FOR?

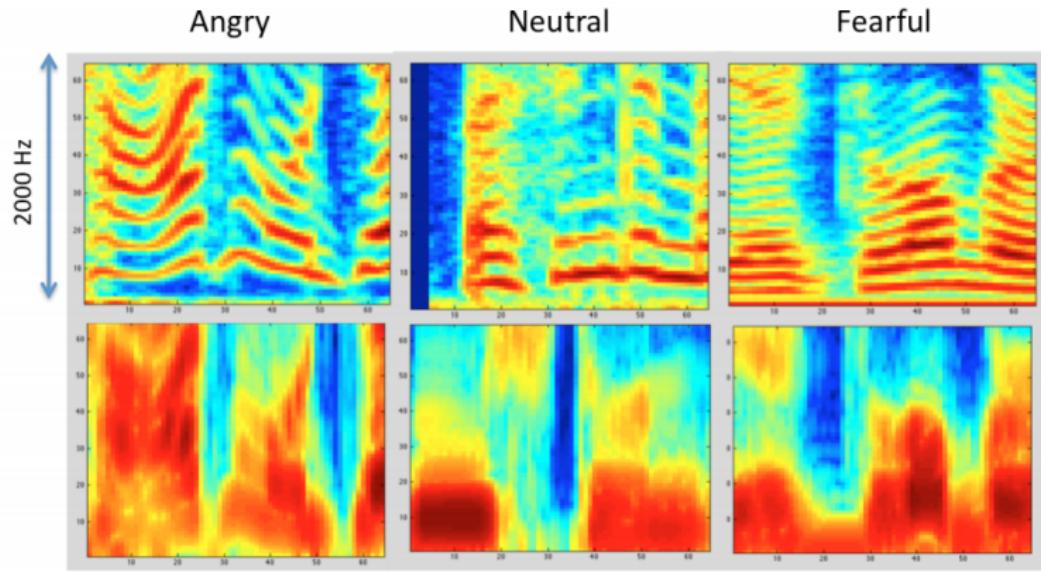
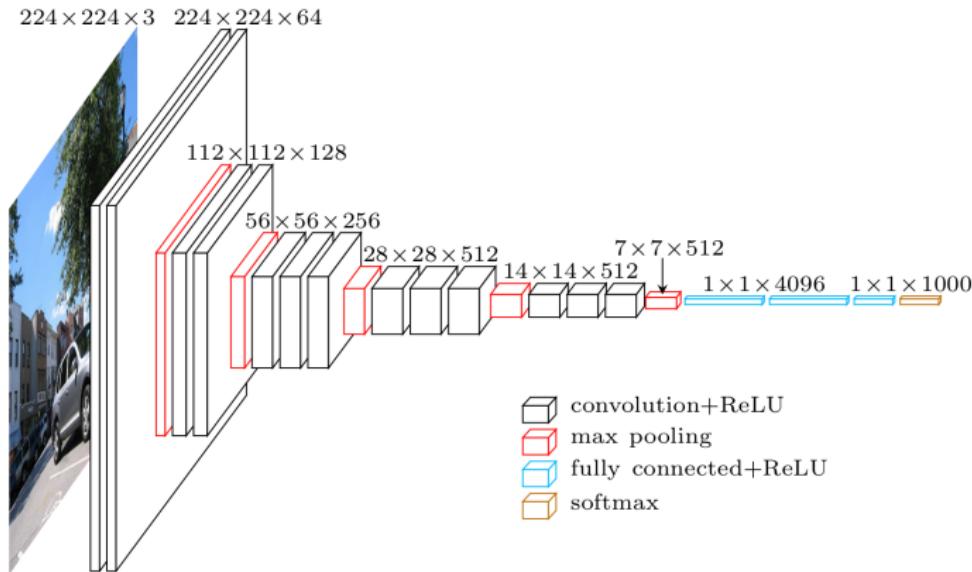


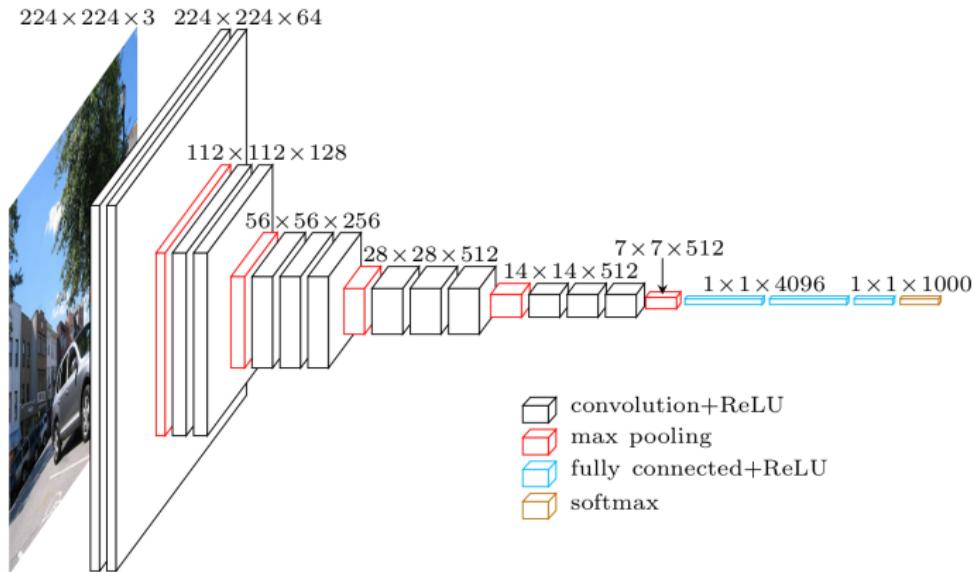
Figure: Speech recognition (Anand & Verma (2015)). Convolutional neural network to extract features from audio data in order to classify emotions.

CNNs - A FIRST GLIMPSE



- Input layer: contains the input (-image) as data matrices.
- **Convolutions:** extract feature maps from a previous layer.
- **Pooling:** reduces the dimensionality of any input and filter robust features.

CNNs - A FIRST GLIMPSE



- Fully connected layer: standard layer that connects feature map elements with the output neurons.
- Softmax: squashes output values to probability scores.