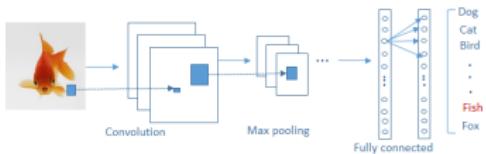


# Deep Learning

## CNN: Introduction

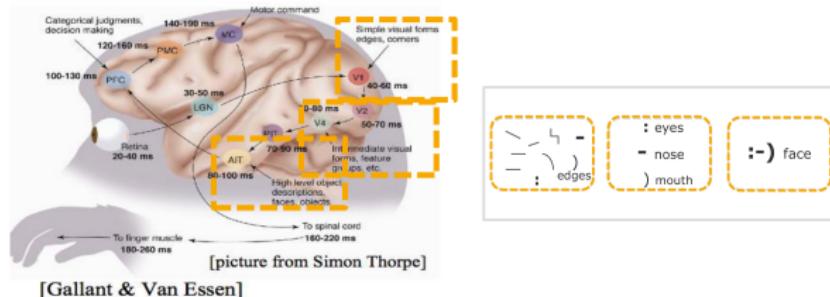


### Learning goals

- What are CNNs?
- When to apply CNNs?
- A glimpse into CNN architectures

# CONVOLUTIONAL NEURAL NETWORKS

- Convolutional Neural Networks (CNN, or ConvNet) are a powerful family of neural networks that are inspired by biological processes 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 etc., which consist of lots of intermediate representations.

# CONVOLUTIONAL NEURAL NETWORKS

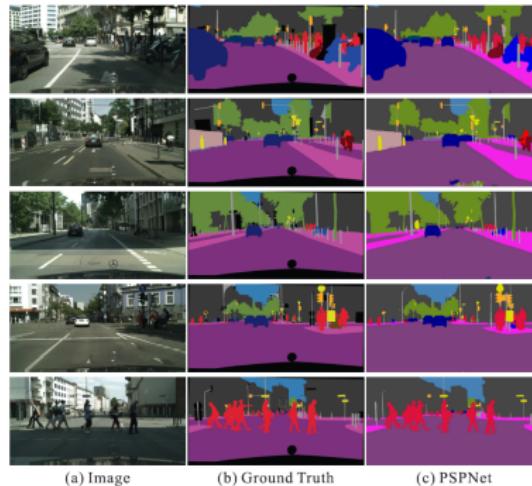
- Since 2012, given their success in the ILSVRC competition, CNNs are popular in many fields.
- Common applications of CNN-based architectures in computer vision are:
  - Image classification.
  - Object detection / localization.
  - Semantic segmentation.
- CNNs are widely applied in other domains such as natural language processing (NLP), audio, and time-series data.
- Basic idea: a CNN automatically extracts visual, or, more generally, spatial features from an input data such that it is able to make the optimal prediction based on the extracted features.
- It contains different building blocks and components.

# CNNs - WHAT FOR?



**Figure:** All Tesla cars being produced now have full self-driving hardware. A convolutional neural network is used to map raw pixels from a single front-facing camera directly into steering commands. The system learns to drive in traffic, on local roads, with or without lane markings as well as on highways (The Tesla Team, 2016).

# CNNs - WHAT FOR?



**Figure:** Given an input image, a CNN is first 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 (Zhao et al., 2017).

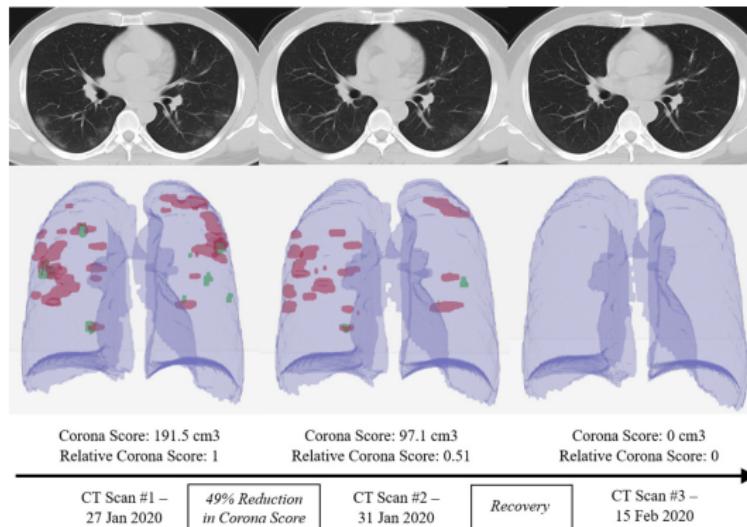
# CNNs - WHAT FOR?



**Figure:** Road segmentation: Aerial images and possibly outdated map pixels are segmented (Mnih & Hinton, 2010).

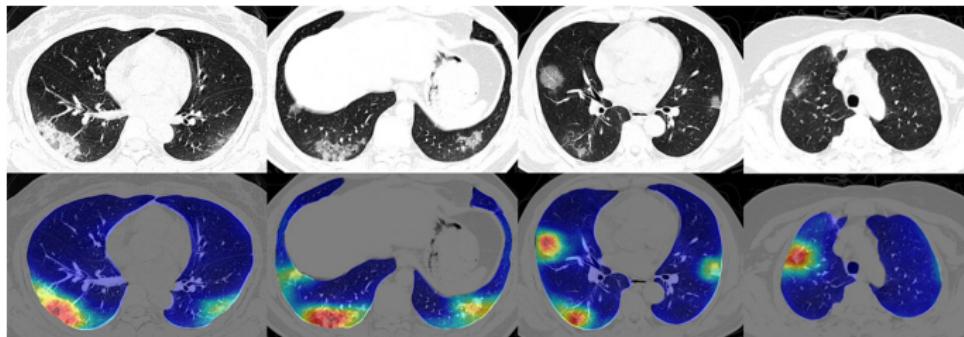
# CNNs - WHAT FOR?

CNN for personalized medicine e.g tracking, diagnosis and localization of Covid-19 patients.



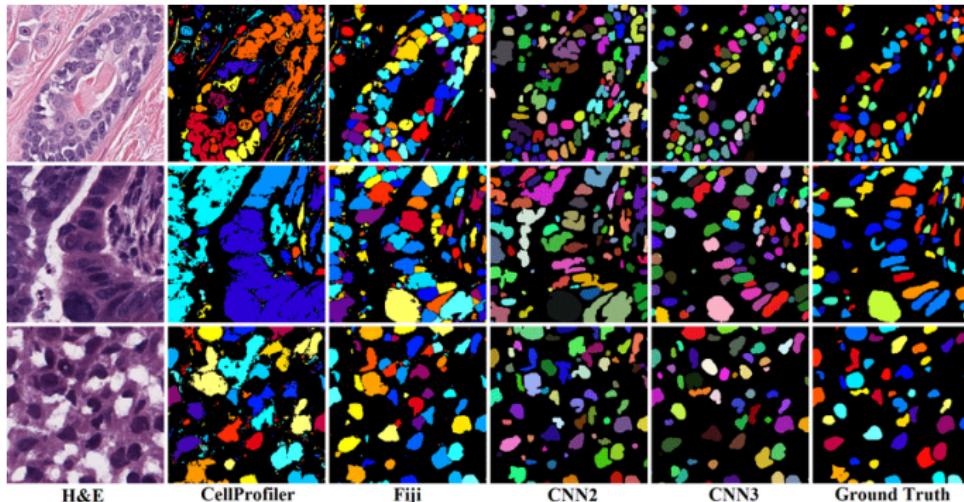
**Figure:** CNN based method (RadLogics algorithm) for personalized Covid-19 detection: three CT scans from a single Corona virus patient (Scudellari, 2020).

# CNNs - WHAT FOR?



**Figure:** Four COVID-19 lung CT scans at the top with corresponding colored maps showing Corona virus abnormalities at the bottom using the RadLogics algorithm (Scudellari, 2020).

# CNNs - WHAT FOR?



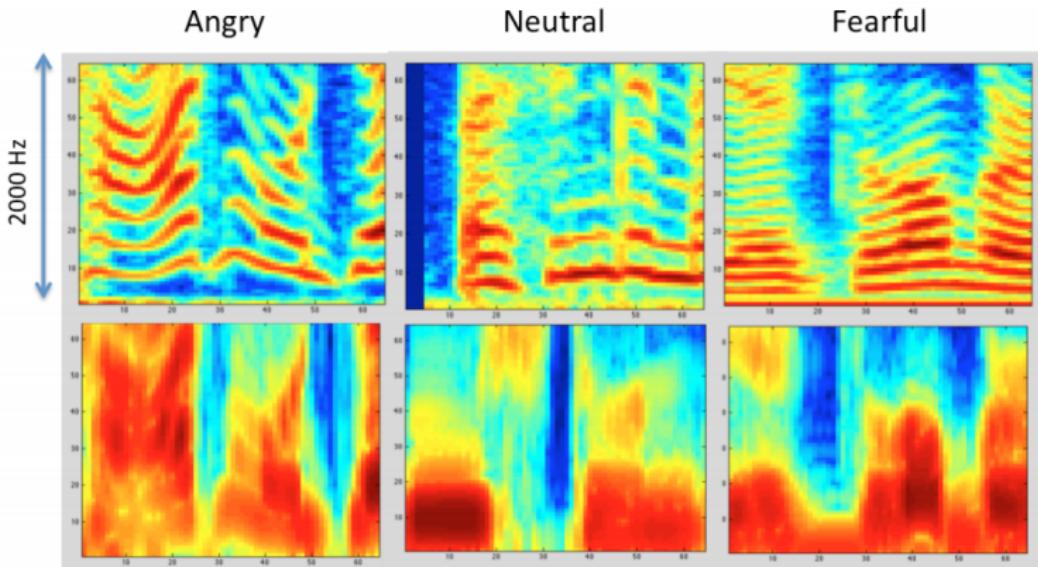
**Figure:** Various analyses in computational pathology are possible. For example, nuclear segmentation in digital microscopic tissue images enable extraction of high-quality features for nuclear morphometrics (Kumar et al., 2017).

# CNNs - WHAT FOR?



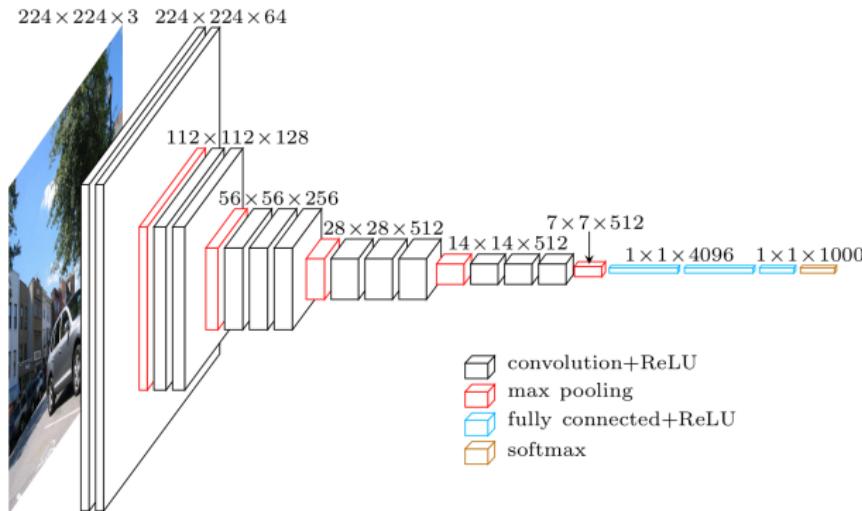
**Figure:** Image Colorization: Given a grayscale photo as the input (top row), this network solves the problem of hallucinating a plausible color version of the photo (bottom row, i.e. the prediction of the network) (Zhang et al., 2016).

# CNNs - WHAT FOR?



**Figure:** Speech recognition: Convolutional neural network is used to learn features from the audio data in order to classify emotions (Anand, 2015).

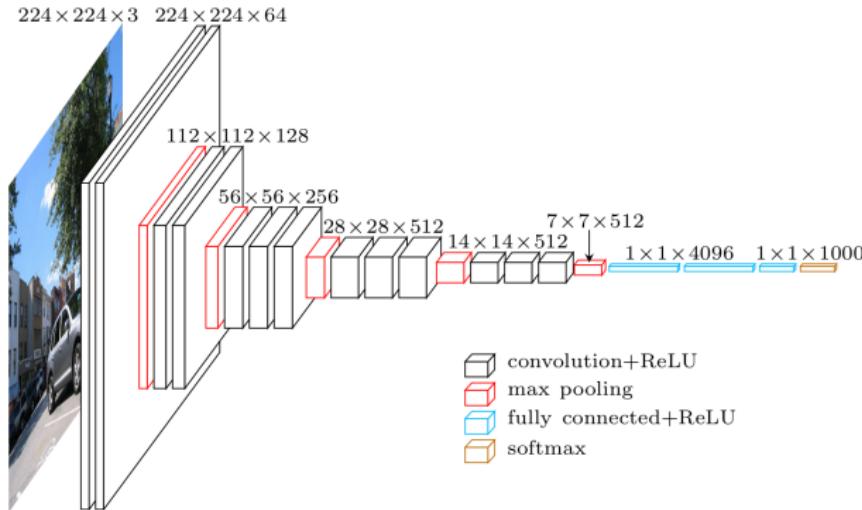
# CNNs - A FIRST GLIMPSE



**Figure:** Example architecture (Simonyan & Zisserman, 2015)

- **Input layer** takes input data (e.g. image, audio).
- **Convolution layers** extract feature maps from the previous layers.
- **Pooling layers** reduce the dimensionality of feature maps and filter meaningful features.

# CNNs - A FIRST GLIMPSE



**Figure:** Example architecture (Simonyan & Zisserman, 2015)

- **Fully connected layers** connect feature map elements to the output neurons.
- **Softmax** converts output values to probability scores.

# REFERENCES

-  The Tesla Team. (2016, October 19). *All Tesla cars being produced now have full self-driving hardware: Tesla Switzerland*. Tesla.  
[https://www.tesla.com/de\\_ch/blog/all-tesla-cars-being-produced-now-have-full-self-driving-hardware](https://www.tesla.com/de_ch/blog/all-tesla-cars-being-produced-now-have-full-self-driving-hardware)
-  Mnih, V., & Hinton, G. E. (2010). Learning to Detect Roads in High-Resolution Aerial Images. In K. Daniilidis, P. Maragos, & N. Paragios (Eds.), *Computer Vision – ECCV 2010* (pp. 210–223). Springer Berlin Heidelberg.
-  Zhao, H., Shi, J., Qi, X., Wang, X., & Jia, J. (2017, July). *Pyramid Scene Parsing Network*. Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR).
-  Scudellari, M. (2020, March 31). *Hospitals deploy AI tools to detect COVID-19 on chest scans*. IEEE Spectrum. <https://spectrum.ieee.org/hospitals-deploy-ai-tools-detect-covid19-chest-scans>
-  Kumar, N., Verma, R., Sharma, S., Bhargava, S., Vahadane, A., & Sethi, A. (03 2017). A Dataset and a Technique for Generalized Nuclear Segmentation for Computational Pathology. *IEEE Transactions on Medical Imaging*, 36, 1–1.  
doi:10.1109/TMI.2017.2677499

# REFERENCES

-  Zhang, R., Isola, P., & Efros, A. A. (2016). Colorful Image Colorization. In B. Leibe, J. Matas, N. Sebe, & M. Welling (Eds.), *Computer Vision – ECCV 2016* (pp. 649–666). Cham: Springer International Publishing.
-  Anand, N. (2015). *Convolted Feelings Convolutional and recurrent nets for detecting emotion from audio data*. Retrieved from <https://api.semanticscholar.org/CorpusID:209374156>
-  Simonyan, K., & Zisserman, A. (2015). Very Deep Convolutional Networks for Large-Scale Image Recognition.