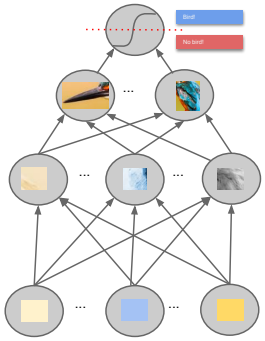
Deep Learning

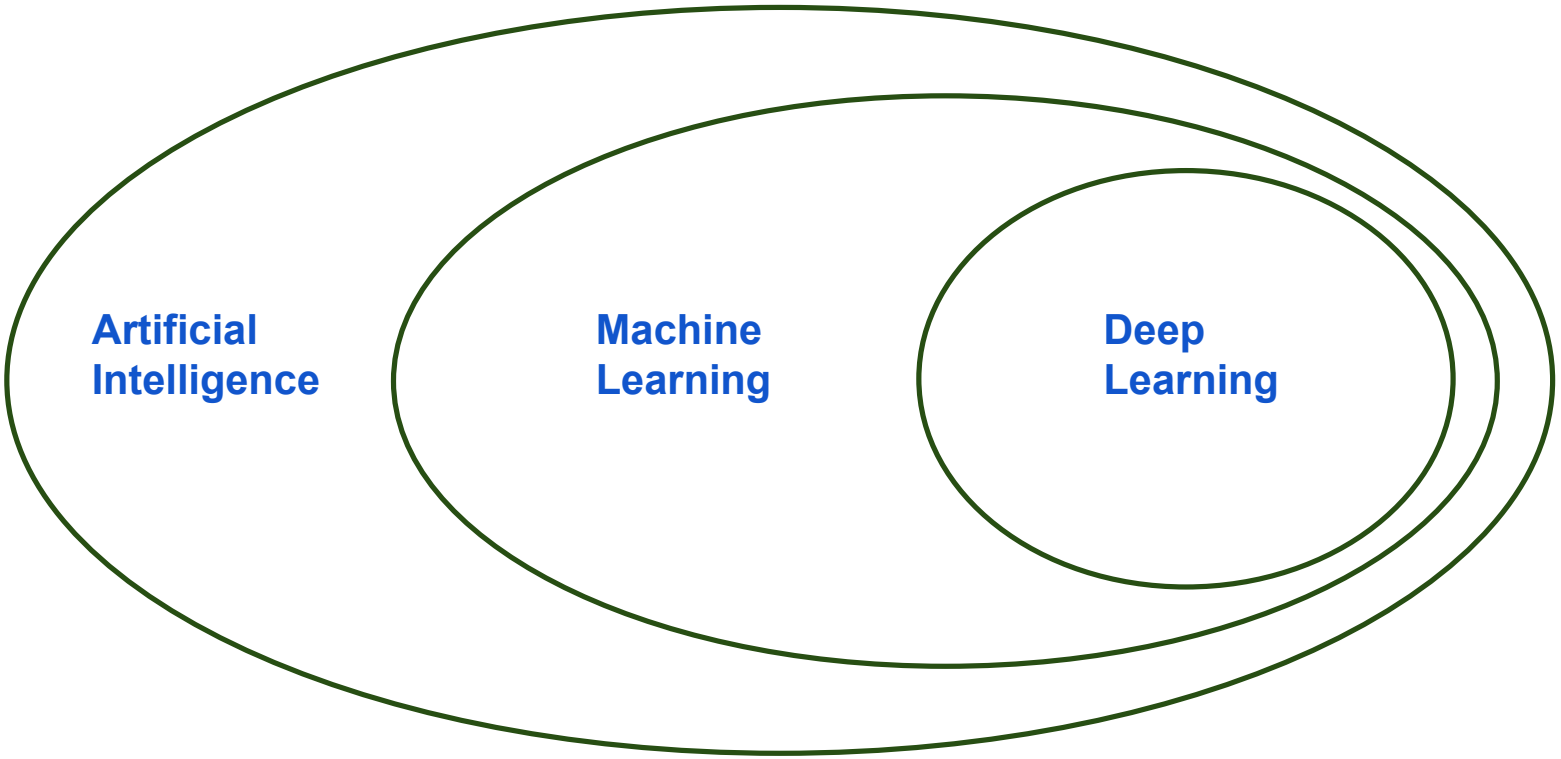
Introduction



Learning goals

- Relationship of DL and ML
- Concept of representation or feature learning
- Use-cases and data types for DL methods

WHAT IS DEEP LEARNING



- Deep learning is a subfield of ML based on artificial neural networks.

DEEP LEARNING AND NEURAL NETWORKS

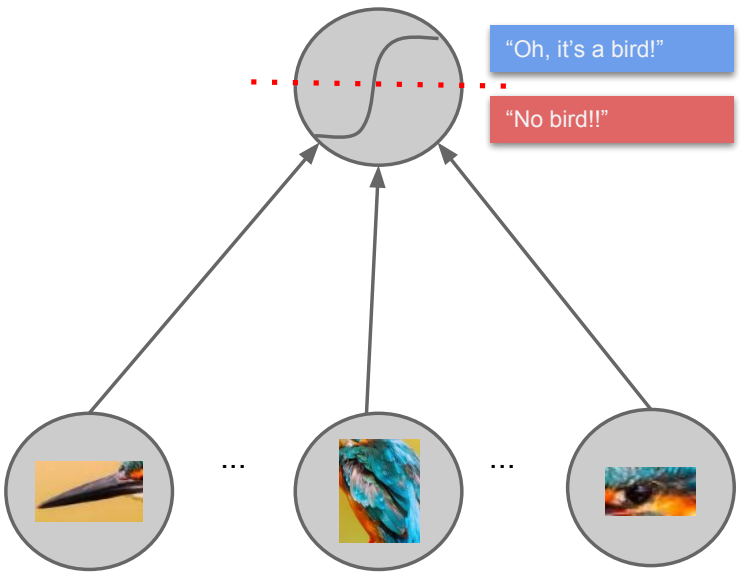
- Deep learning itself is not *new*:
  - Neural networks have been around since the 70s.
  - *Deep* neural networks, i.e., networks with multiple hidden layers, are not much younger.

- Why everybody is talking about deep learning now:
  1. Specialized, powerful hardware allows training of huge neural networks to push the state-of-the-art on difficult problems.
  2. Large amount of data is available.
  3. Special network architectures for image/text data.
  4. Better optimization and regularization strategies.

IMAGE CLASSIFICATION WITH NEURAL NETWORKS

"Machine learning algorithms, inspired by the brain, based on learning multiple levels of representation/abstraction."

Y. Bengio



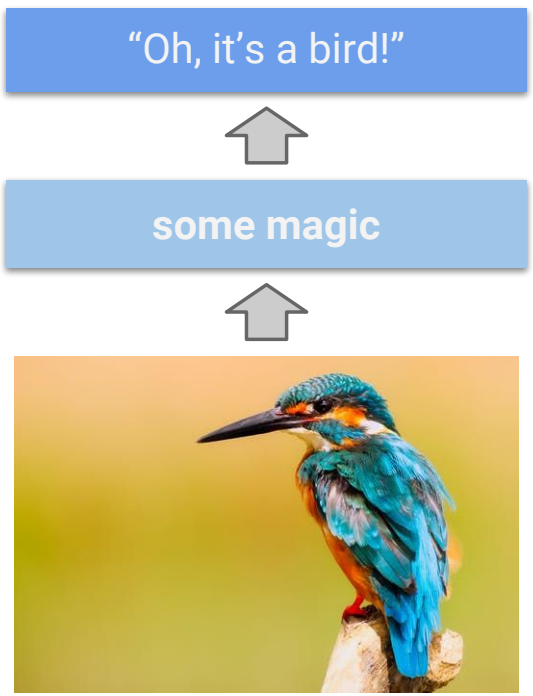


IMAGE CLASSIFICATION WITH NEURAL NETWORKS

"Machine learning algorithms, inspired by the brain, based on learning multiple levels of representation/abstraction."

Y. Bengio

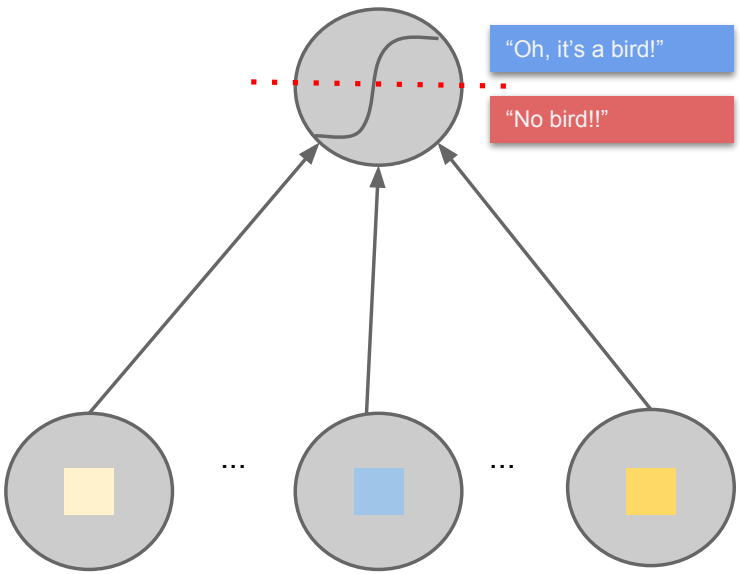
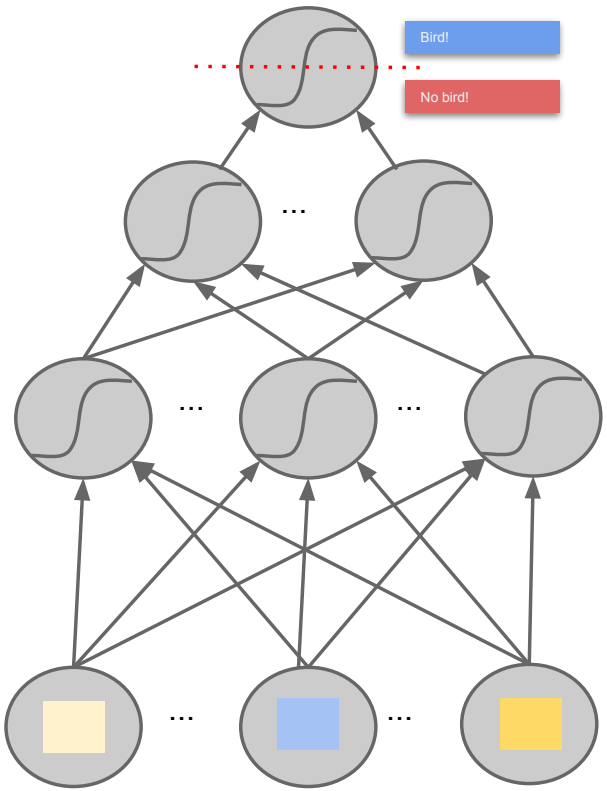


IMAGE CLASSIFICATION WITH NEURAL NETWORKS

"Machine learning algorithms, inspired by the brain, based on learning multiple levels of representation/abstraction."

Y. Bengio



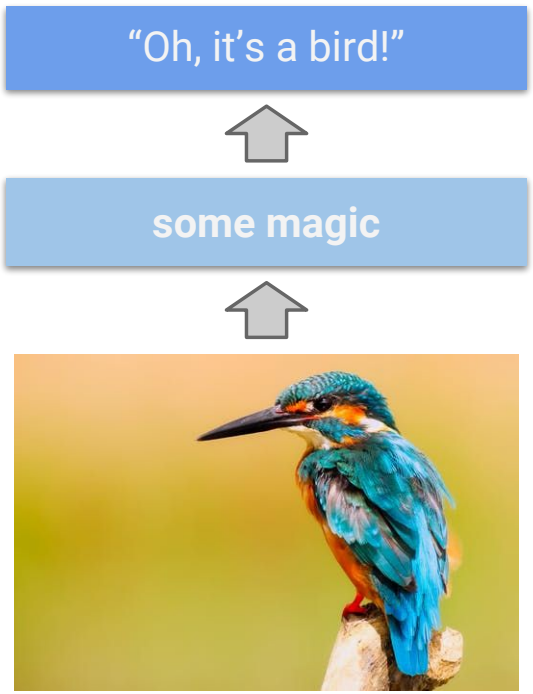
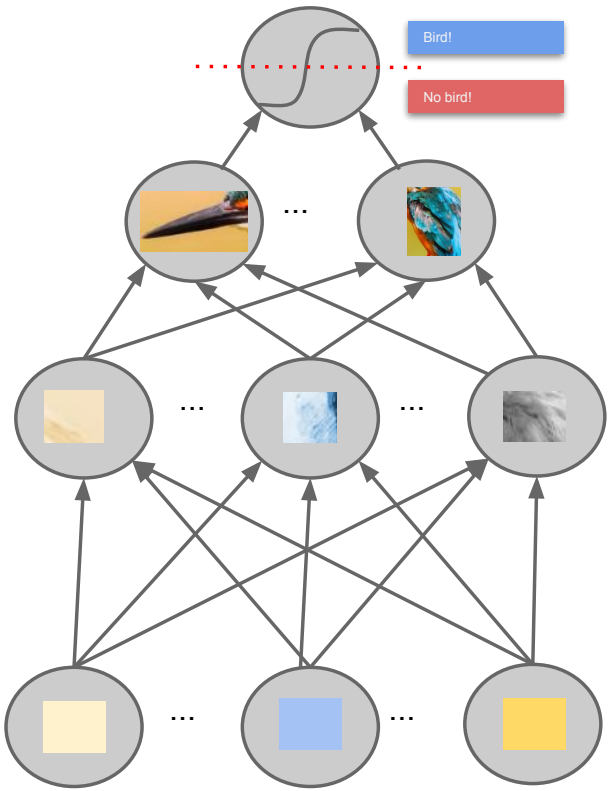
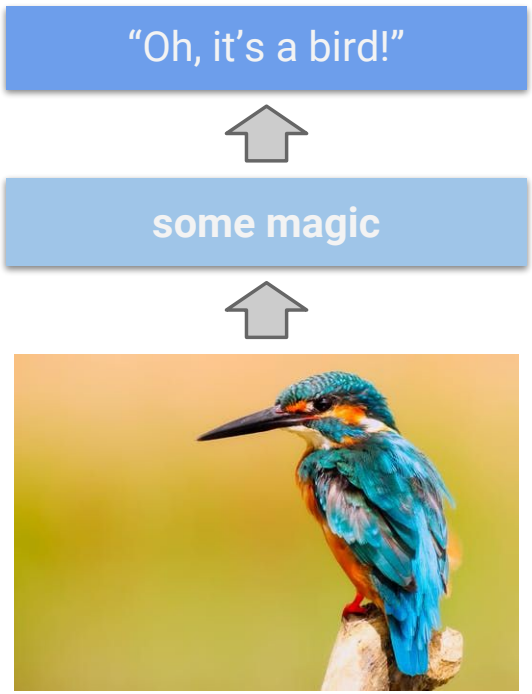


IMAGE CLASSIFICATION WITH NEURAL NETWORKS

"Machine learning algorithms, inspired by the brain, based on learning multiple levels of representation/abstraction."

Y. Bengio





POSSIBLE USE-CASES

Deep learning can be extremely valuable if the data has these properties:

- It is high dimensional.
- Each single feature itself is not very informative but only a combination of them might be.
- There is a large amount of training data.

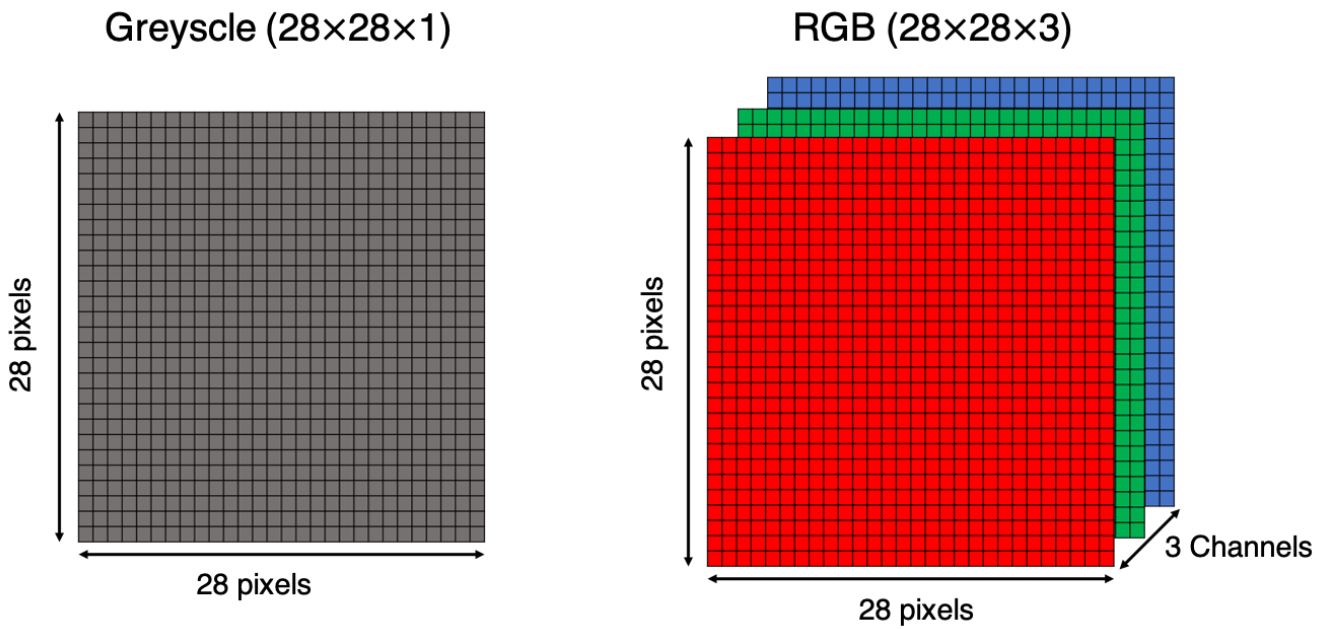
This implies that for tabular data, deep learning is rarely the correct model choice.

- Without extensive tuning, models like random forests or gradient boosting will outperform deep learning most of the time.
- One exception is data with categorical features with many levels.

POSSIBLE USE-CASE: IMAGES

- **High Dimensional**: A color image with 255×255 (3 Colors) pixels already has 195075 features.
- **Informative**: A single pixel is not meaningful in itself.
- **Training Data**: Depending on applications huge amounts of data are available.

Architecture: **C**onvolutional **N**eural **N**etworks (CNN)



POSSIBLE USE-CASE: IMAGES

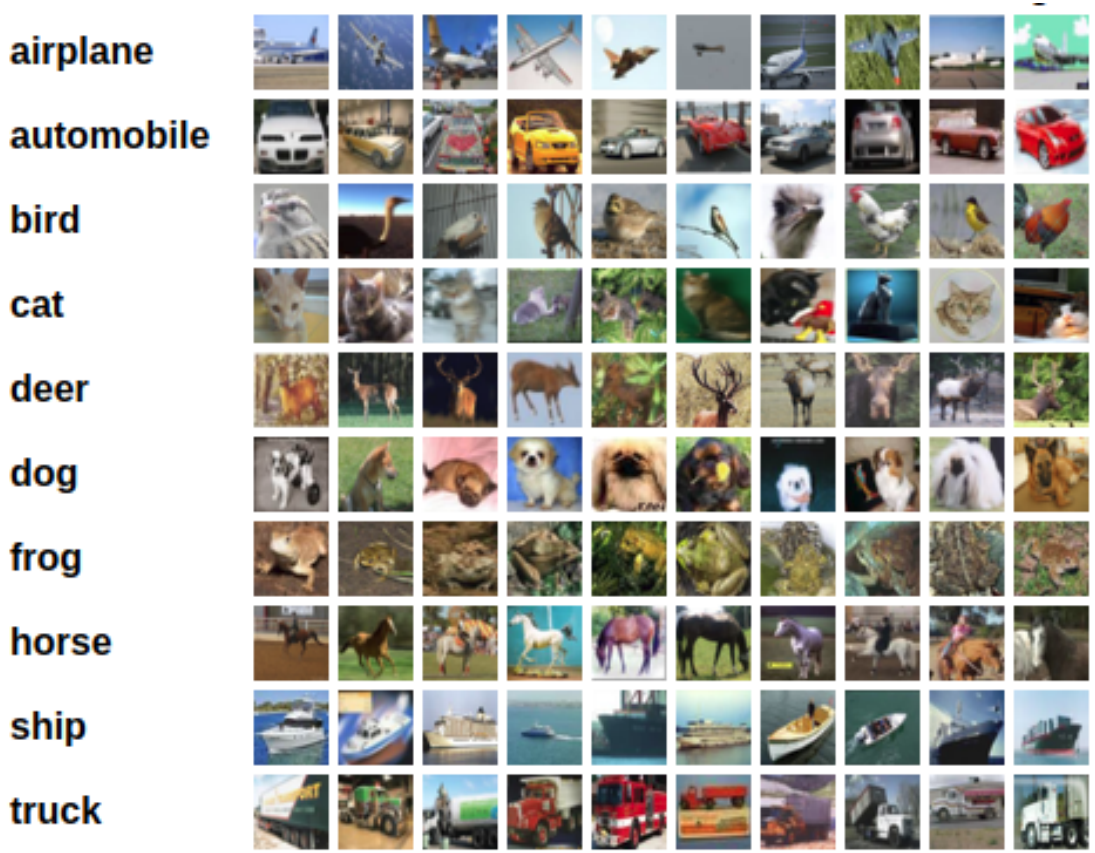


Figure: Example for image classification (Krizhevsky, 2009)

Image classification tries to predict a single label for each image.
CIFAR-10 is a well-known dataset used for image classification. It consists of 60,000 32x32 color images containing one of 10 object classes, with 6000 images per class.

POSSIBLE USE-CASE: IMAGES

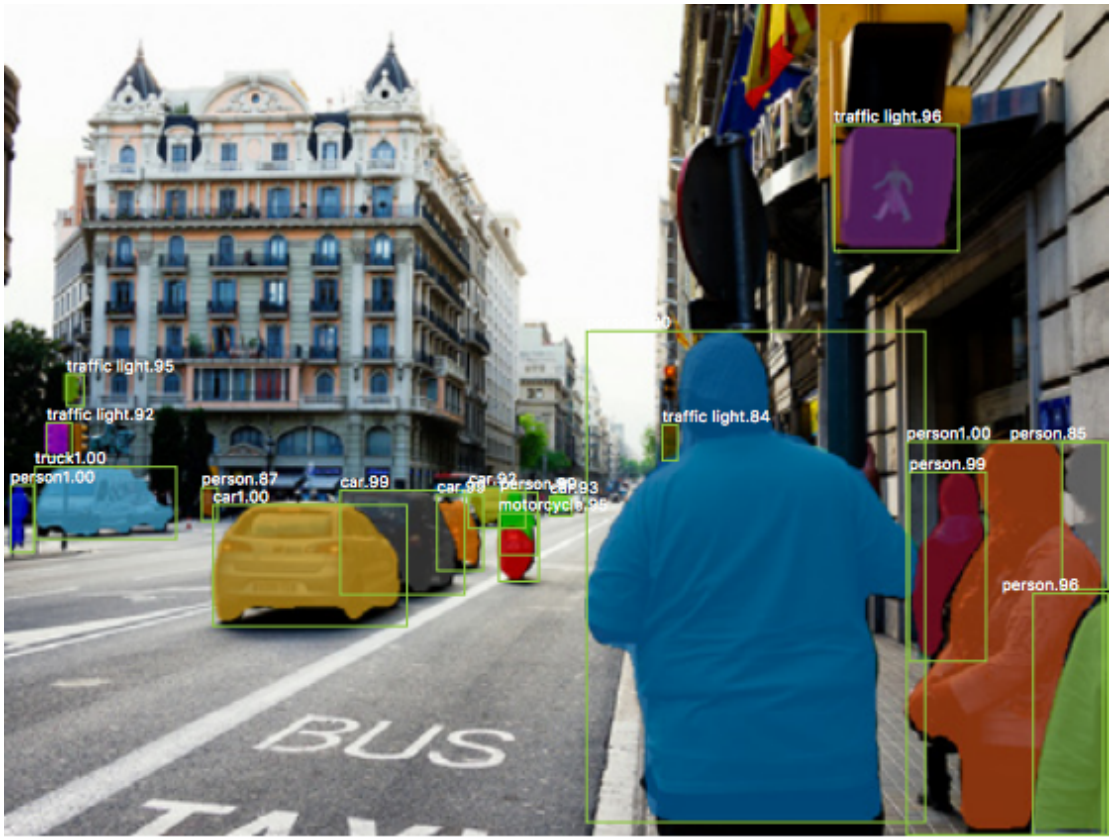


Figure: Example for object detection (He et al., 2017)

Object Detection Mask R-CNN is a general framework for instance segmentation, that efficiently detects objects in an image while simultaneously generating a high-quality segmentation mask for each instance.

POSSIBLE USE-CASE: IMAGES

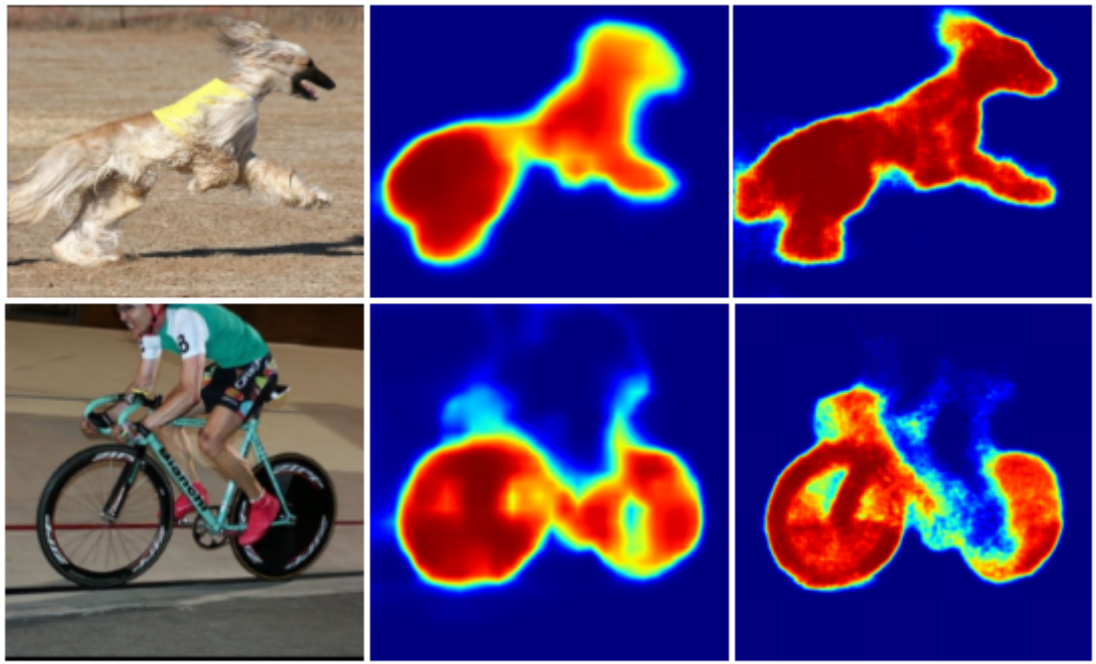


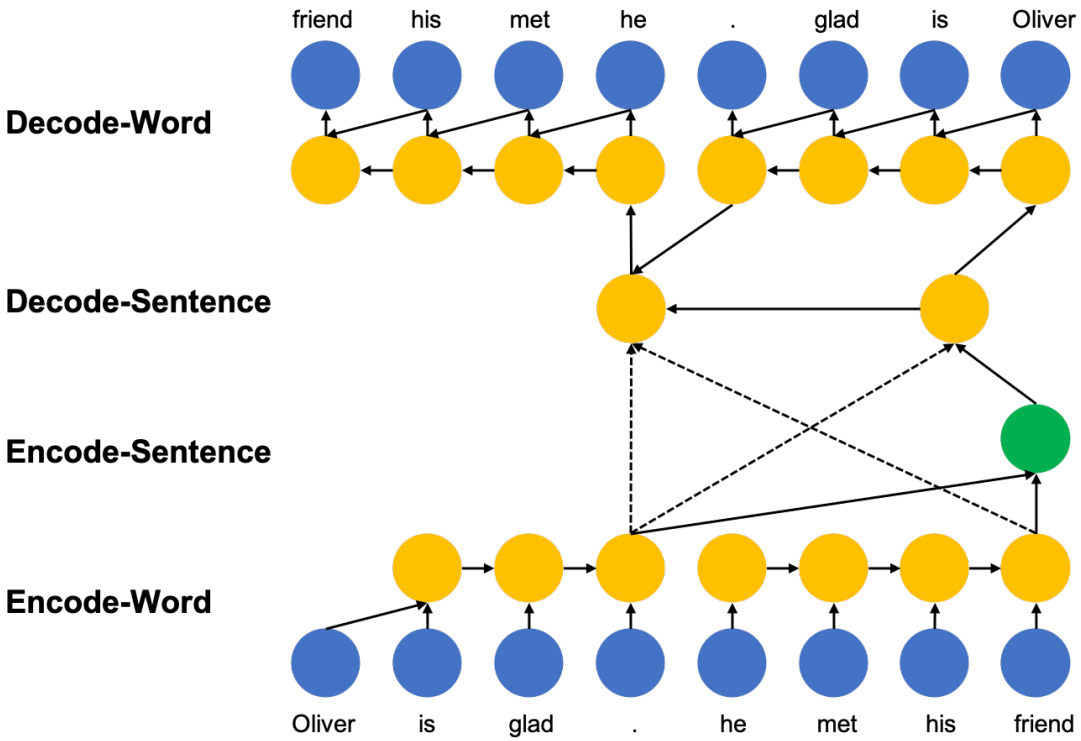
Figure: Example for image segmentation (Noh et al., 2015))

Image segmentation partitions the image into (multiple) segments.

POSSIBLE USE-CASE: TEXT

- **High Dimensional**: Each word can be a single feature (300000 words in the German language).
- **Informative**: A single word does not provide much context.
- **Training Data**: Huge amounts of text data available.

Architecture: **R**ecurrent **N**eural **N**etworks (RNN)



POSSIBLE USE-CASE: TEXT CLASSIFICATION

Positive

Great job! Your customer support is fantastic.

Neutral

Not bad, but it should be improved in the future.

Negative

The worst customer service I have ever seen.

Sentiment Analysis is the application of natural language processing to systematically identify the emotional and subjective information in texts.

POSSIBLE USE-CASE: TEXT

Machine Translation (e.g. google translate) Neural machine translation exploits neural networks to predict the likelihood of a sequence of words, typically modeling entire sentences in a single integrated model.

APPLICATIONS OF DEEP LEARNING: SPEECH

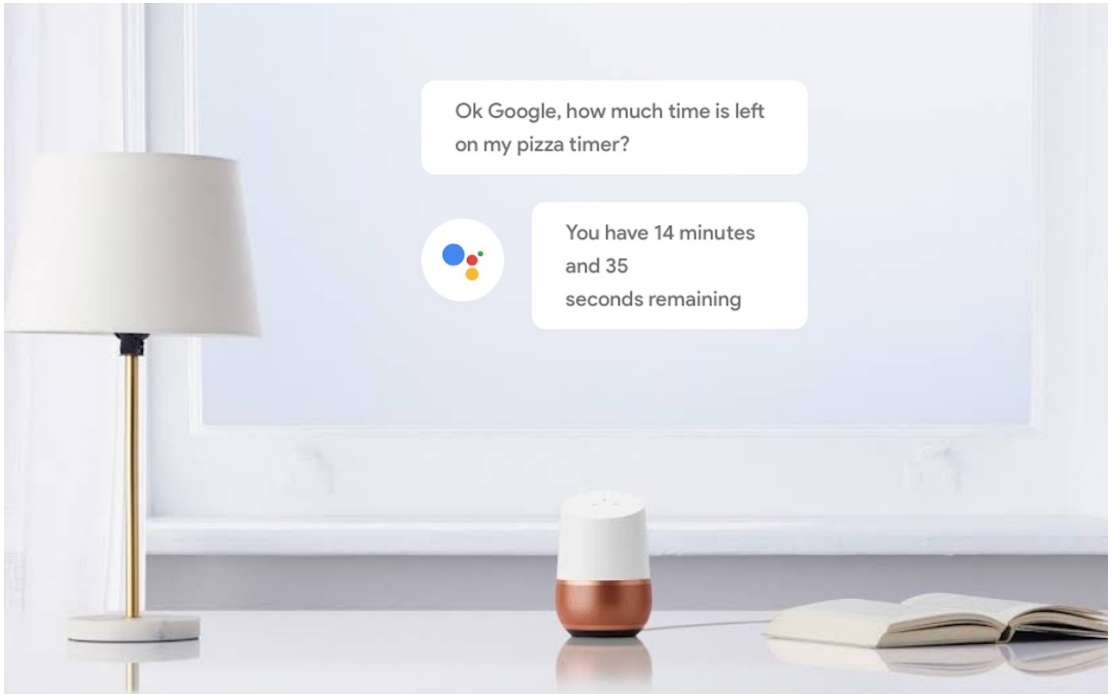


Figure: Example for speech recognition and generation (Google)

Speech Recognition and Generation (e.g. google assistant) Neural network extracts features from audio data for downstream tasks, e.g., to classify emotions in speech.

REFERENCES

- Krizhevsky, A. (2009). *Learning Multiple Layers of Features from Tiny Images.*
- He, K., Gkioxari, G., Dollár, P., & Girshick, R. (2017). *Mask R-CNN*.
- Noh, H., Hong, S., & Han, B. (2015). *Learning Deconvolution Network for Semantic Segmentation*.
- Google. (n.d.). *Smart speaker mit google assistant*. Google. `https://assistant.google.com/intl/de_de/platforms/speakers/`