Deep Learning in a nutshell

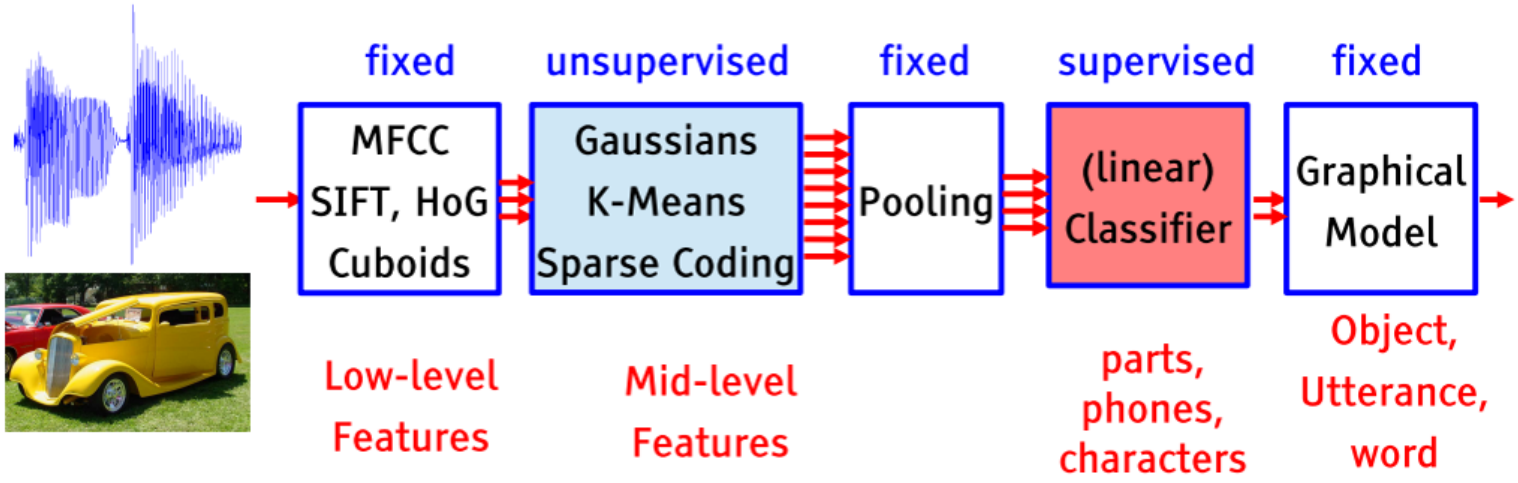
Elad Hoffer

Technion Israel Institute Of Technology,
Computer-Science department

Traditional machine learning

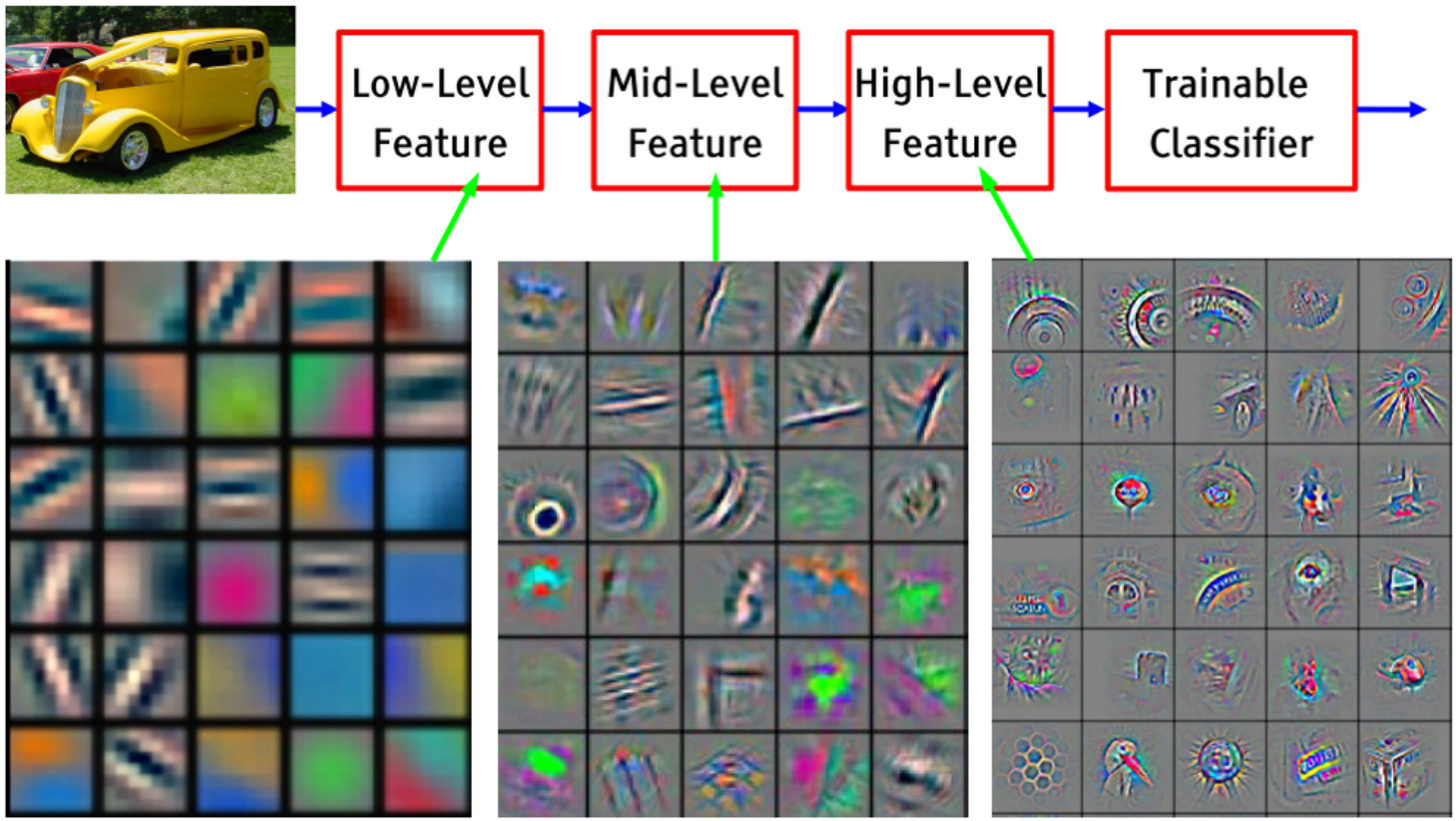
Traditional machine learning is about:

- Feature engineering
- Creating models and representations
- Optimizing and predicting



Feature learning

Can we create a model to learn and extract hierarchical feature representations from data?



Feature visualization of convolutional net trained on ImageNet from [Zeiler & Fergus 2013]

ConvNets

Deep learning is about learning *hierarchical* high-level representation of data

- This is usually done by end-to-end training of *Artificial neural networks*

Neural networks

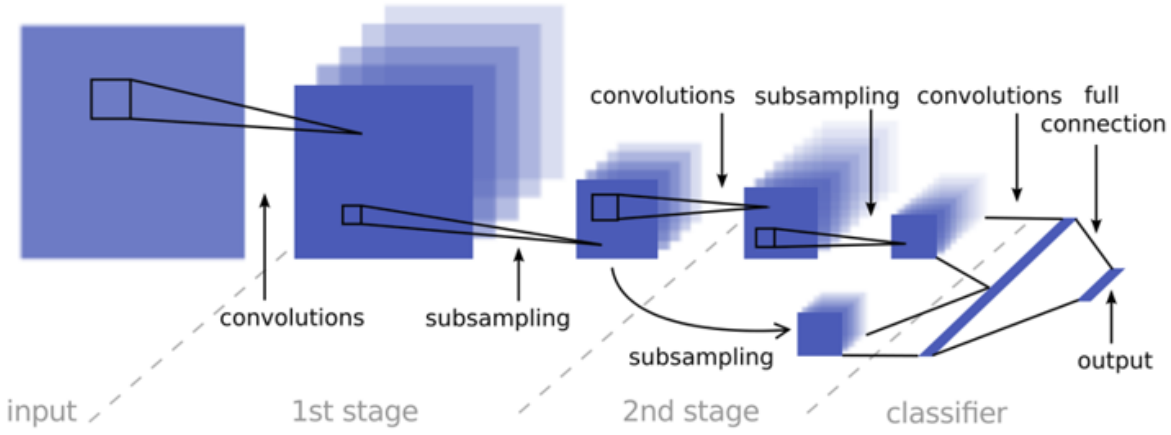
This is not a new approach (as early as 1940s), but we finally have enough data and computational resources to make it shine.

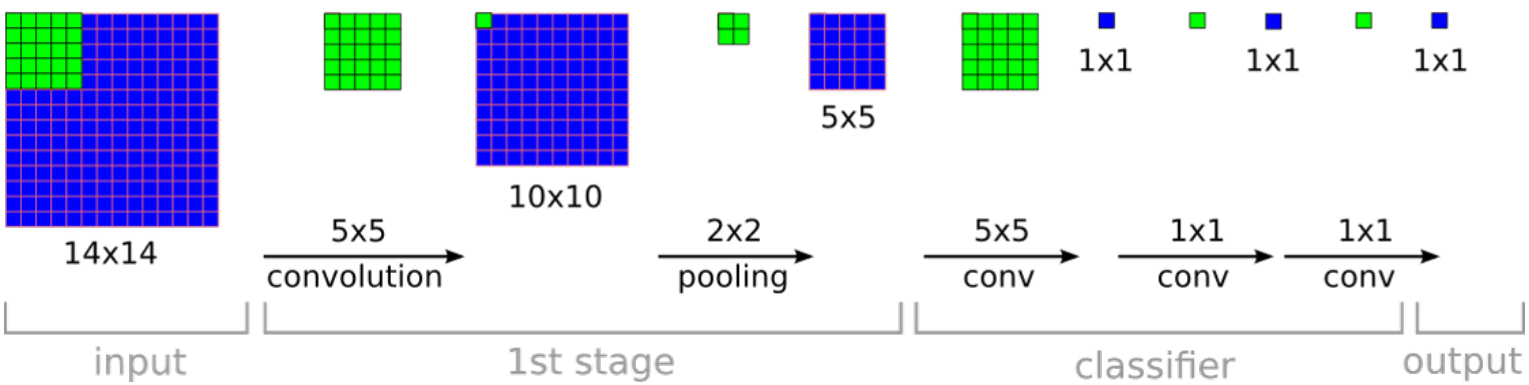
Currently state-of-the-art approach in several domains, most noteworthy:

- Vision - object-classification, detection etc.
- Speech recognition

ConvNets

Convolutional network (ConvNet/CNN), is a type of deep network used for vision related tasks

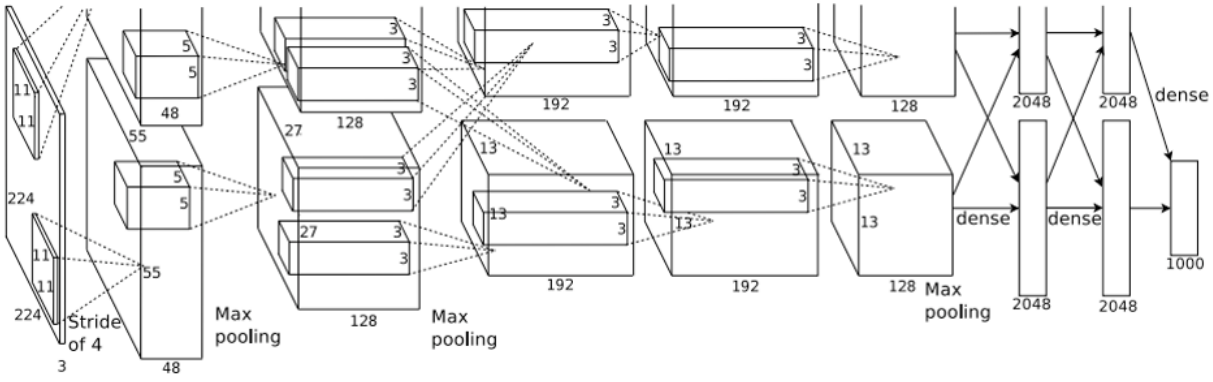




ConvNets

In 2012 Convolutional Networks were reintroduced and popularized (invented in the 90s)

- AlexNet was submitted to the ImageNet ILSVRC challenge in 2012 and significantly outperformed the second runner-up (top 5 error of 16% vs 26% error).
- The Network was deeper, bigger, than previous networks - 60M parameters, 8 trainable layers.

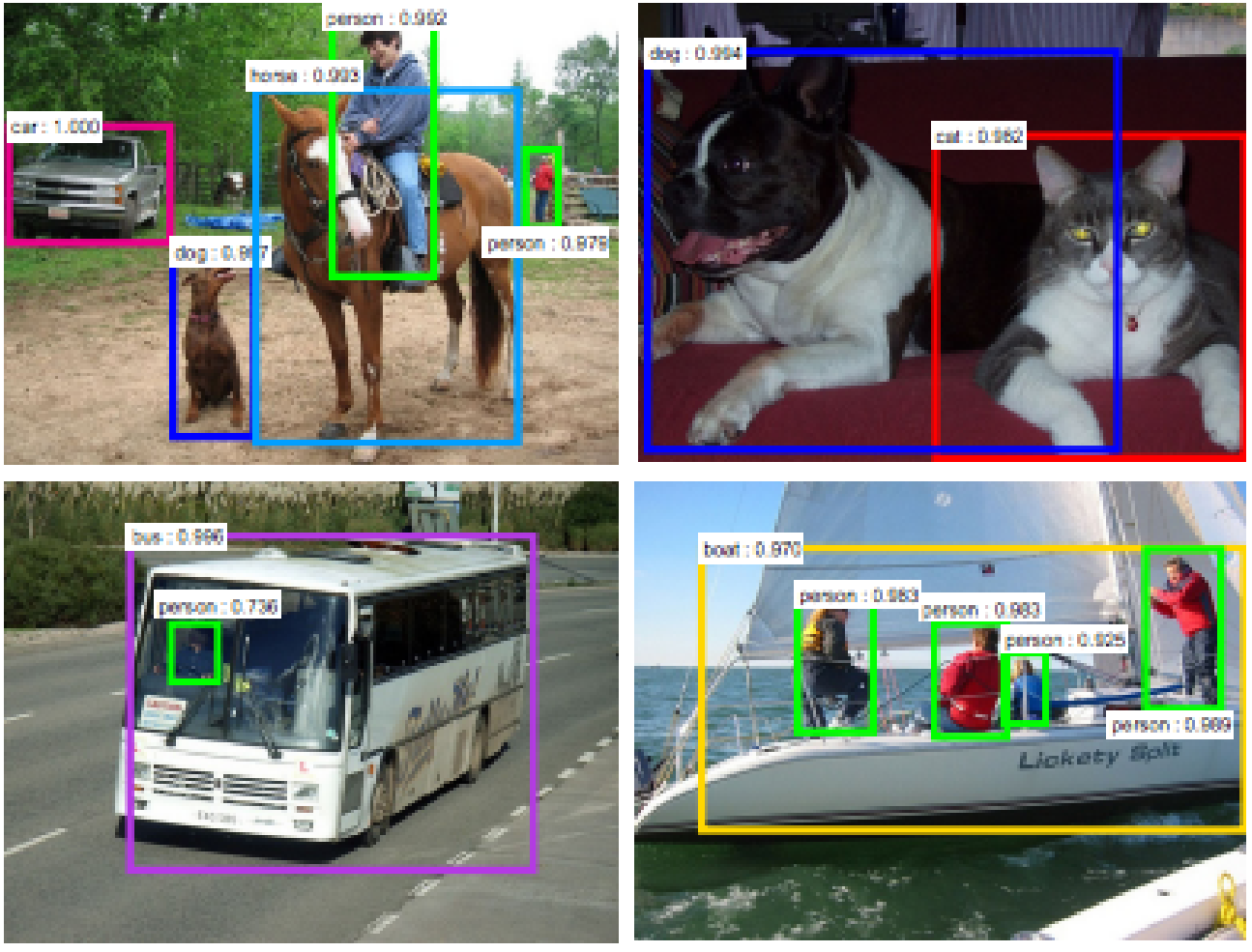


ConvNets

- We now use much deeper, more complex networks, and CNNs are the de-facto standard for visual recognition
- Deep learning models also work very well for many other tasks...

Localization + Detection with CNNs

Localize and *detect* objects in an image.

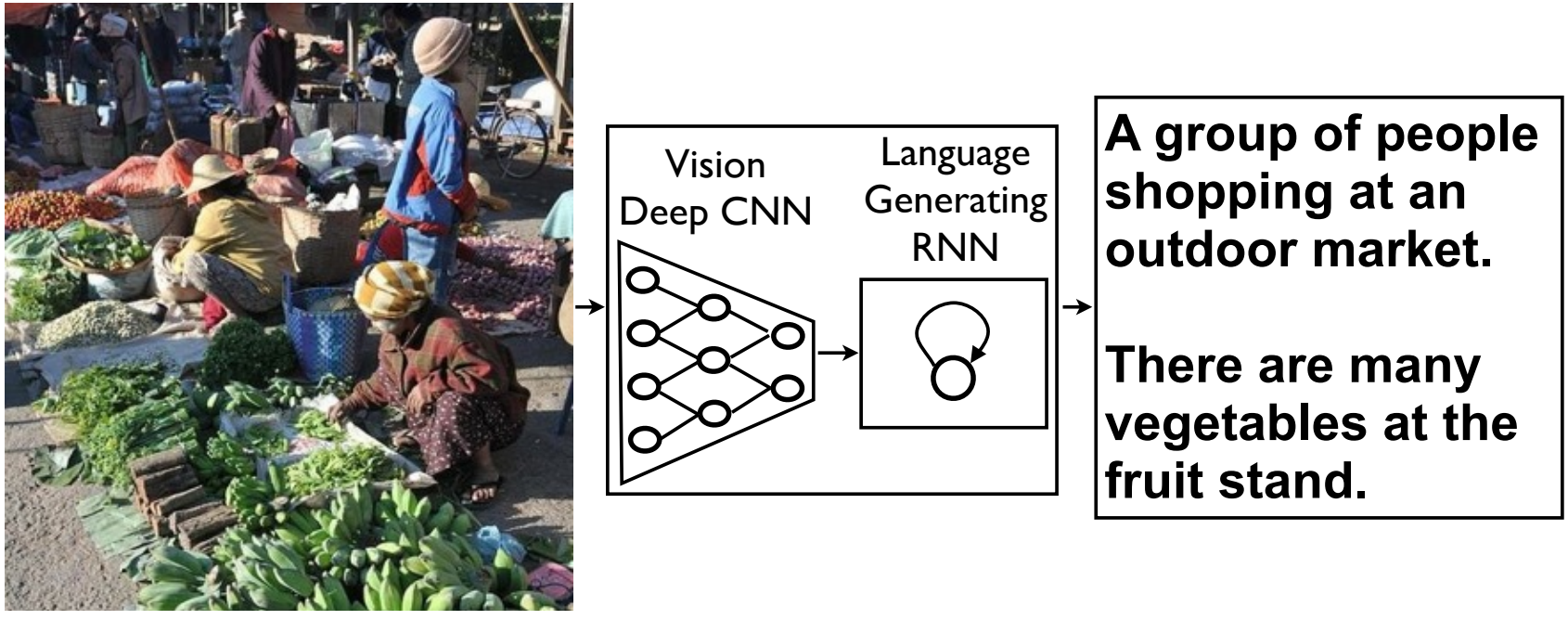


Segmentation using CNNs

Image segmentation (*DeepMask* Pinheiro 15′)

Caption generation

Another trend is coupling image and language understanding

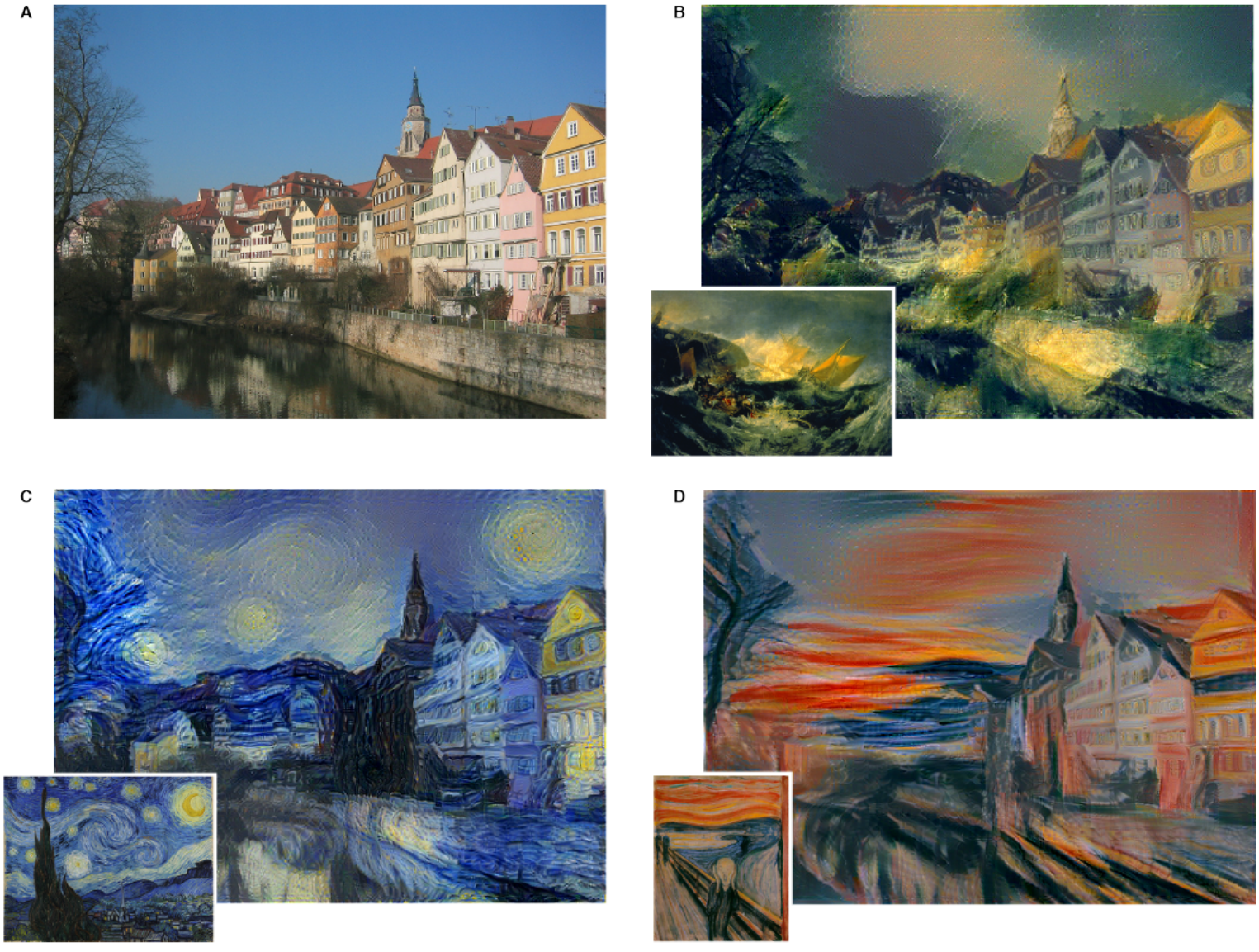


Generative adversarial networks

Even generating images (Radford et al. 15′):

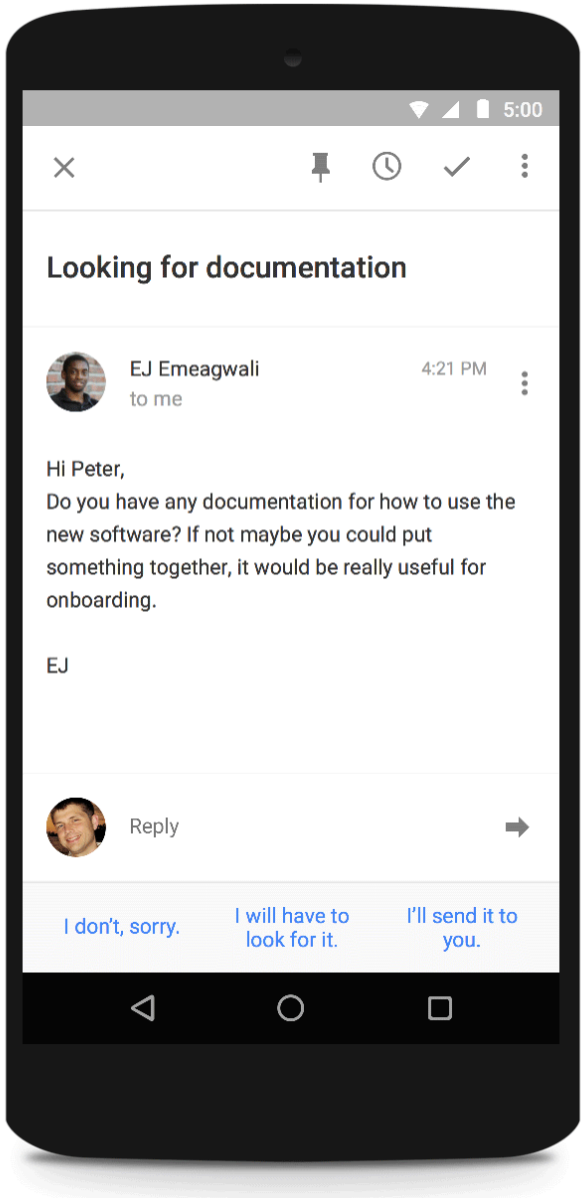
Generative adversarial networks

Style transfer (Gatys 15')



Sequence to sequence

Or sequences of text (recurrent networks):



My research - comparative learning

Taking inspiration from human learning, we often find that humans tend to relate more easily to comparative measures than to explicit discrimination.

- E.g: Instead of answering the question:

What is this flower? / What are the key properties of it?

My research - comparative learning

Taking inspiration from human learning, we often find that humans tend to relate more easily to comparative measures than to explicit discrimination.

- E.g: Instead of answering the question:

What is this flower? / What are the key properties of it?

- It is much easier for humans to answer

Which of these is more similar? / What are the shared properties?