Deep Learning - Foundations and Concepts

Chapter 10. Convolutional Networks

nonlineark@github

March 6, 2025

Outline

Computer vision

- Computer vision was one of the first fields to be transformed by the deep learning revolution, predominantly thanks to the CNN architecture.
- Recently alternative architectures based on transformers have become competitive with convolutional networks in some applications.
- Some applications for machine learning in computer vision: Classification, detection, segmentation, caption generation, synthesis, inpainting, style transfer, super-resolution, depth prediction, scene reconstruction.

Image data

- The structure of an image:
  - An image comprises a rectangular array of pixels.
  - Each pixel has either a grey-scale intensity or a triplet of red, green and blue channels each with its own intensity value.
- Challenges of applying neural networks to image data:
  - Images generally have a high dimensionality.
  - Image data is highly structured.
- Local correlations can be used to encode strong inductive biases into a neural network, leading to models with far fewer parameters and with much better generalization accuracy.