MO434 - Deep Learning
Convolutional Neural Networks

Alexandre Xavier Falcão

Institute of Computing - UNICAMP

afalcao@ic.unicamp.br

Agenda

- Convolutional neural network (CNN).
- Building a CNN model for image classification.
- Classical modules to enhance CNN models.
- Using and fine-tuning pretrained models.
- How to explain CNNs using information visualization.

Convolutional Neural Network (CNN)

- CNNs emerged from the study of the visual cortex of cat brains and have been used for image classification since 1980s.

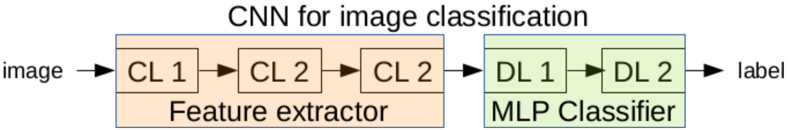
Convolutional Neural Network (CNN)

- CNNs emerged from the study of the visual cortex of cat brains and have been used for image classification since 1980s.
- David Hubel and Torsten Wiesel (Nobel Prize in Medicine, 1981) showed that
  - some neurons react to small receptive fields (adjacency relations),
  - others with a same receptive field react to lines of different orientations (filters of a kernel bank), and
  - neurons with larger receptive fields react to more complex patterns (small adjacency relations in deeper layers correspond to larger regions in the input image).

Convolutional Neural Network (CNN)

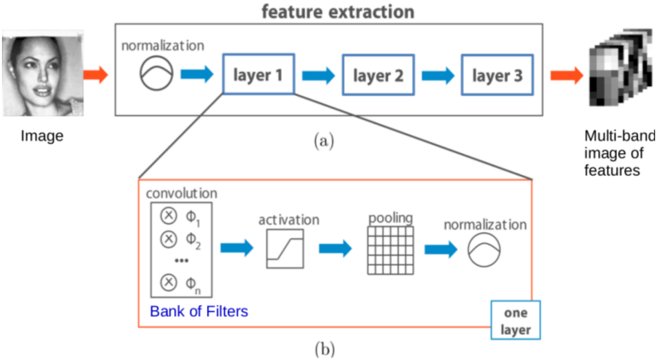
- CNNs emerged from the study of the visual cortex of cat brains and have been used for image classification since 1980s.
- David Hubel and Torsten Wiesel (Nobel Prize in Medicine, 1981) showed that
  - some neurons react to small receptive fields (adjacency relations),
  - others with a same receptive field react to lines of different orientations (filters of a kernel bank), and
  - neurons with larger receptive fields react to more complex patterns (small adjacency relations in deeper layers correspond to larger regions in the input image).
- From lower- to higher-level image patterns can be extracted by stacking more convolutional layers – a feature extractor.

- Yann LeCun proposed LeNet-5 (1998) by adding pooling and normalization to convolutional layers.
- The ImageNet dataset (Fei-Fei Li, 2009) accelerated the development and success of several other CNNs, starting with AlexNet (Alex Krizhevsky, 2012).
- Such convolutional layers (CL) have much less parameters than dense layers (DL), which finally makes training by backpropagation feasible for deep neural networks.



Convolutional neural network for image classification

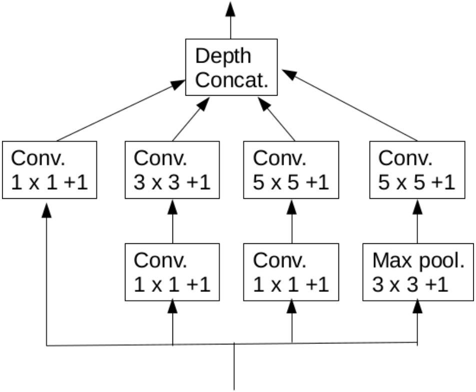
A convolutional layer may contain convolution, activation, pooling, normalization, skip connections, and subnetworks (modules/layer blocks) – e.g., LeNet-5, AlexNet, and VGGNet.



By flattening or by global average pooling, its output becomes a global feature vector for the input of the first dense layer.

Convolutional neural network for image classification

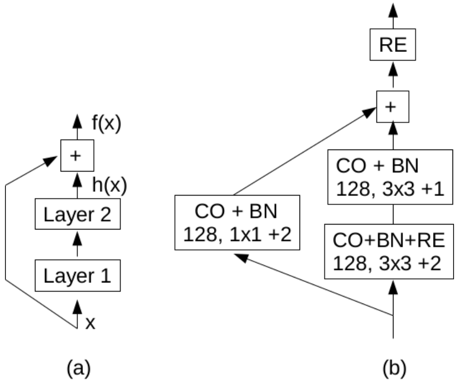
GoogLeNet (Christian Szegedy et al., 2015) incorporated inception modules to CNNs.



A depth concatenation layer concatenates the output of multiple convolutional layers $w \times w + s$, with $w \times w$ adjacencies, padding $w/2$, and stride s.

Convolutional neural network for image classification

ResNet (Kaiming He et al., 2015) incorporated residual modules, whose goal is to force a residual output $f(x) = h(x) - x$ for an input x that passes through a subnetwork of output $h(x)$.

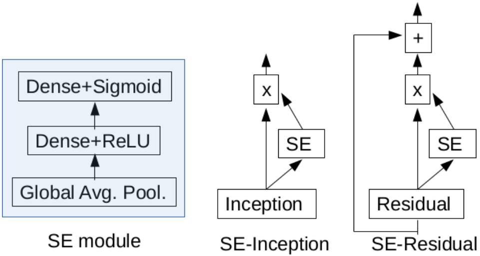


CO+BN+RE is convolution followed by batch normalization and ReLU.

Convolutional neural network for image classification

- Xception net (François Chollet, 2016) added depthwise separable modules, which use a distinct spatial filter per channel followed by a convolution with 1×1 filters.

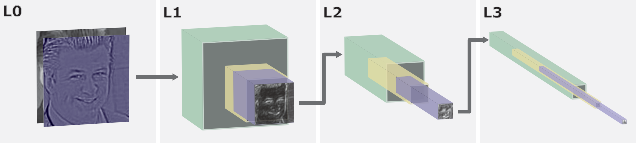
- Xception net (François Chollet, 2016) added depthwise separable modules, which use a distinct spatial filter per channel followed by a convolution with 1×1 filters.
- SENet (Jie Hu et al., 2018) added squeeze-and-excitation module to recalibrate feature maps from another module.



An SE module outputs the importance in $[0, 1]$ of each channel at the output of a given module (inception/residual).

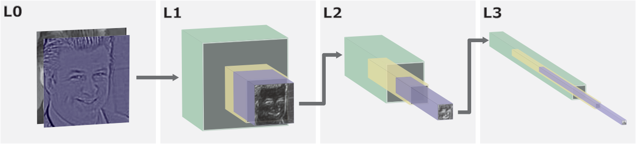
Convolutional neural network for image classification

Deeper layers usually reduce spatial resolution (strides > 1) and increase the number of filters (number of output channels).



Convolutional neural network for image classification

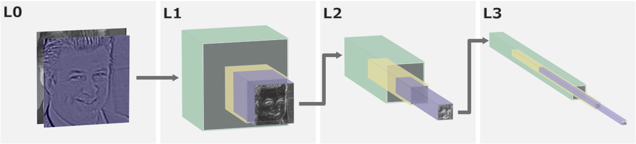
Deeper layers usually reduce spatial resolution (strides > 1) and increase the number of filters (number of output channels).



- Class separability: convolutional layers aim to map images of distinct classes onto different subspaces, such that the classes can be separated by one or two dense layers.

Convolutional neural network for image classification

Deeper layers usually reduce spatial resolution (strides > 1) and increase the number of filters (number of output channels).

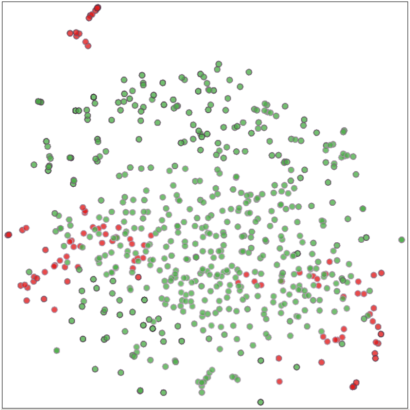


- Class separability: convolutional layers aim to map images of distinct classes onto different subspaces, such that the classes can be separated by one or two dense layers.

- Neuron specialization: the goal of hidden dense layers is to reduce dimensionality and specialize neurons as the faces of one hyperpolyhedron per class (see Fundamentals of DNNs - Part I).

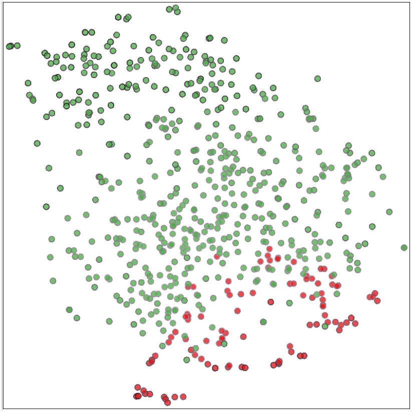
Class separability

Deeper convolutional layers should increase class separability.



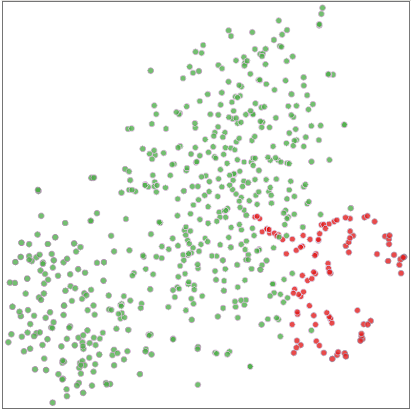
Feature projections (t-SNE) after layers 10, 11, 12, and 13 for larvae of helminth and impurities using VGG-16.

Deeper convolutional layers should increase class separability.



Feature projections (t-SNE) after layers 10, 11, 12, and 13 for larvae of helminth and impurities using VGG-16.

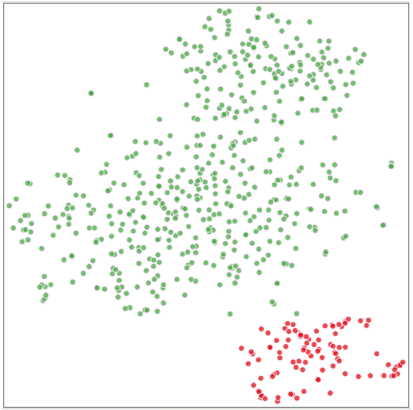
Deeper convolutional layers should increase class separability.



Feature projections (t-SNE) after layers 10, 11, 12, and 13 for larvae of helminth and impurities using VGG-16.

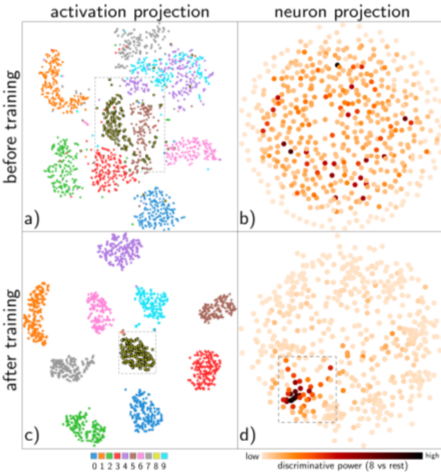
Class separability

Deeper convolutional layers should increase class separability.



Feature projections (t-SNE) after layers 10, 11, 12, and 13 for larvae of helminth and impurities using VGG-16.

Neuron specialization



Neuron projections (MDS, right) colored by their discriminative power for class 8 versus the others in a digit dataset.

Building CNNs with PyTorch

How to

- specify operations such as convolution, activation, pooling, batch normalization and skip connections,
- build a simple CNN for image classification,
- train and test our model, and
- save the model for future use (Building a simple CNN).

An image classification problem

Let's build, train, and test a more complex CNN for image classification. We will

- build models from scratch (Building models from scratch) and
- use pretrained models to extract features for image classification (Using pre-trained models).

- How to visualize activations and the importance of image regions in a model decision ▸ (Explaining the model).

Explaining CNNs with information visualization

- How to visualize activations and the importance of image regions in a model decision (Explaining the model).

- The next lecture will introduce CNN applications to image analysis and your project in this course.