15 Further CNNs

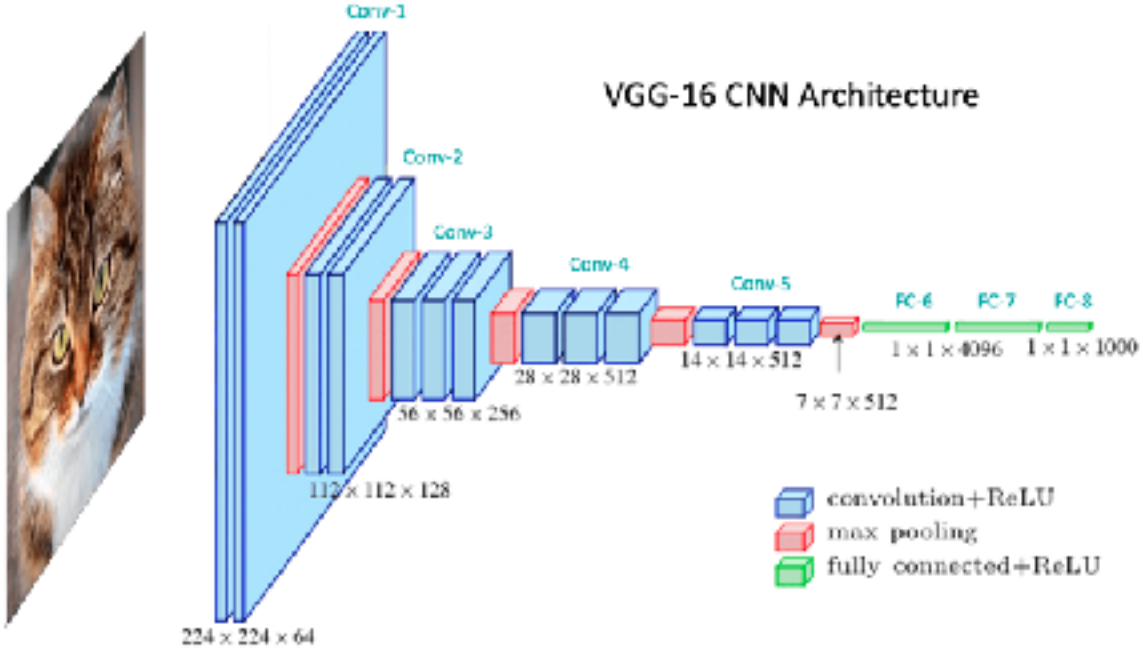
Time to grow a backbone

15.1
Backbones for CNNs.

What are Backbones?

- In the context of Convolutional Neural Networks a "backbone" refers to the base part of the network that is responsible for extracting features from the input data
- Selecting the right Backbone for the task is crucial as it has an affect on multiple attributes of the model

Example of a popular Backbone



Importance of Backbones

- **Feature Extraction**: The primary role of the backbone in a CNN is to extract meaningful features from the input data (images, videos, etc.)
- This involves capturing various aspects such as edges, textures, and patterns that are crucial for understanding the content of the data

Importance of Backbones

- **Hierarchical Processing**: CNN backbones typically consist of multiple layers, where each layer builds upon the features extracted by the previous layers
- This hierarchical structure allows the network to learn complex and abstract representations of the data

Importance of Backbones

- **Transfer Learning**: A well-trained backbone can be used across different tasks and datasets
- This is known as transfer learning, where a backbone trained on a large and diverse dataset (like ImageNet) is reused for other tasks, significantly reducing the training time and data requirements

Importance of Backbones

- **Efficiency and Performance**: The architecture of the backbone affects the efficiency and performance of the CNN
- A well-designed backbone can strike a balance between computational efficiency and the ability to capture relevant features, which is vital for deploying models in resource-constrained environments (like mobile devices)

Importance of Backbones

- **Adaptability**: Backbones can be adapted to different scales and complexities of tasks
- For instance, lighter backbones like MobileNet are used for applications where speed and low memory footprint are crucial, while more complex backbones like ResNet are used for tasks where accuracy is more important

15.2
Getting a Backbone.

Backbone Architectures for CNNs

- There are several popular architectures used in Convolutional Neural Networks
- Each has its unique characteristics and applications

VGG (Visual Geometry Group)

- VGG models, particularly VGG16 and VGG19, are known for their deep architectures with consecutive convolution layers.
- They are straightforward but computationally intensive, often used as a baseline for comparison in many computer vision tasks

ResNet (Residual Networks)

- ResNet architectures, especially ResNet-50, ResNet-101, and ResNet-152, introduced the concept of residual learning with shortcut connections to alleviate the vanishing gradient problem in very deep networks
- They are highly influential and widely used for a variety of tasks due to their excellent performance and efficiency

MobileNet

- Designed for mobile and edge devices, MobileNet architectures use depthwise separable convolutions to reduce the model size and computational complexity while maintaining high performance
- They are ideal for applications where resources are limited

15.3
Selecting the **right Backbone**.

How do you choose the right backbone for the job?

- Akin to choosing the right tool for the job it is important to choose the right backbone for the task at hand
- Choosing between different backbone architectures requires consideration of various factors
- Let’s decide between two architectures by asking some key questions

Deciding between ResNet-50 vs VGG16

- What is the Complexity of the Task?
  - Is the task relatively simple (e.g., basic image classification) or complex (e.g., detailed object detection or segmentation)? ResNet-50, with its deeper architecture, is generally better suited for more complex tasks

Deciding between ResNet-50 vs VGG

- How Large and Varied is the Dataset?
  - Deeper networks like ResNet-50 require more data to train effectively without overfitting. VGG16 might be sufficient for smaller datasets

Deciding between ResNet-50 vs VGG

- What are the Computational Constraints?
  - Do you have limitations in terms of computational resources? VGG16 is more computationally intensive, especially in terms of memory due to its fully connected layers, whereas ResNet-50 is more efficient

Deciding between ResNet-50 vs VGG

- What is the Required Inference Time?
  - How fast does the model need to make predictions? If inference speed is a crucial factor, especially in real-time applications, the more efficient architecture (ResNet-50) might be preferable

Deciding between ResNet-50 vs VGG

- What is the Training Infrastructure Available?
  - Do you have access to powerful GPUs or are you working with limited resources? Training deeper networks like ResNet-50 typically requires better hardware

Deciding between ResNet-50 vs VGG

- What is the Nature of the Input Data?
  - Are the input images high-resolution, or do they contain fine-grained details? Deeper networks like ResNet-50 might be better at capturing details in high-resolution images

15.3
Selecting the **right Backbone**.

END.