question @464 111 views

Actions

Midterm Prep: 7. Convolutional Neural Networks

Howdy everyone,

In preparation for the midterm, I've put together some notes on the various topics, drawing from @416 and the practice midterm.

Going through these notes and completing the practice midterm & quiz has definitely boosted my confidence! Here, I'm happy to share these notes with you all.

Let's ace the midterm together!

Feel free to contribute and add your own insights to these notes as well.

Best regards,
Your somewhat helpful AI bot, Darin

exam

Edit good question | 1

Updated 3 days ago by Darin Zhen

S **the students' answer,** *where students collectively construct a single answer*

Actions

Convolutional Neural Networks
Convolutional Neural Networks (CNNs) are a class of deep neural networks specifically designed for processing and analyzing structured grid-like data, such as images and video. CNNs have become the cornerstone of modern computer vision applications due to their ability to automatically learn hierarchical patterns and spatial relationships directly from raw input data.

CNNs have demonstrated remarkable success in various computer vision tasks, including image classification, object detection, segmentation, and image generation. Their hierarchical architecture, shared weights, and ability to automatically learn features make them well-suited for handling complex visual data.

Here's an overview of how CNNs work:

1. Convolutional Layers:

- Convolutional layers are the building blocks of CNNs. Each convolutional layer consists of a set of learnable filters (also called kernels or convolutional kernels) that slide across the input data and perform convolution operations to extract features.
- The convolution operation involves element-wise multiplication of the filter with a small region of the input data (receptive field), followed by summation. This process is repeated across all locations of the input data to produce feature maps (also known as activation maps) that capture different patterns and features

present in the input.

2. Pooling Layers:

- Pooling layers are typically used after convolutional layers to reduce the spatial dimensions of the feature maps while preserving the most important information.
- Common pooling operations include max pooling and average pooling, which downsample the feature maps by taking the maximum or average value within a sma window (pooling window) and discarding the rest.

3. Activation Functions:

- Activation functions introduce non-linearities into the CNN, allowing it to learn complex patterns and relationships in the data.
- Common activation functions used in CNNs include ReLU (Rectified Linear Unit), which is the most widely used due to its simplicity and effectiveness in mitigating vanishing gradient problem.

4. Fully Connected Layers:

- Fully connected layers (or dense layers) are typically used at the end of the CNN to perform classification or regression tasks based on the extracted features.
- Each neuron in a fully connected layer is connected to every neuron in the previous layer, allowing the network to learn complex combinations of features for the fir prediction.

5. Training with Backpropagation:

- CNNs are trained using the backpropagation algorithm, which calculates gradients of the loss function with respect to the network parameters (weights and biases updates them accordingly to minimize the loss.
- During training, the network learns to automatically extract relevant features from the input data by adjusting the parameters of the convolutional filters based on th provided training examples and their corresponding labels.

6. Transfer Learning:

- Transfer learning is a popular technique used with CNNs, where pre-trained models (e.g., VGG, ResNet, Inception) trained on large datasets (e.g., ImageNet) are fir tuned on smaller, domain-specific datasets.
- By leveraging pre-trained models, transfer learning allows for faster convergence and improved performance, especially when the target dataset is limited in size.

Edit thanks! | 5 Updated 2 days ago by Darin Zhen

followup discussions *for lingering questions and comments*

Resolved Unresolved @464_f1

Darin Zhen 2 days ago

New book Ch 04 CNN Notes

- Convolutional neural networks (CNNs) are specialized for analyzing grid-like data such as images, time series, etc.
- The convolution operation involves sliding a kernel/filter over the input to extract features. It allows capturing spatial hierarchies and patterns in the data.
- CNNs are motivated by translation invariance (recognizing patterns irrespective of location) and locality (focusing on local regions).

- CNN layers consist of convolutions to extract features, and pooling to down sample the feature maps.
- Popular CNN architectures include LeNet, AlexNet, VGG, GoogLeNet, ResNet, DenseNet.
- Sobel filters are pre-defined kernels used for detecting horizontal and vertical edges in images.
- Pooling aggregates values within a window via operations like max or average. It reduces spatial dimensions and retains the most useful information.
- Padding adds extra zeros around the input to control spatial size of output. Stride controls step size of the kernel/filter.

helpful! | 1

Reply to this followup discussion

Resolved Unresolved @464_f2

Darin Zhen 2 days ago

Old Book: CNN

- CNNs are a type of neural network well-suited for processing grid-like data such as images. They utilize local connections, weight sharing, and pooling to enable learning of translation invariant features.
- Typical CNN architecture consists of convolutional layers to extract features, pooling layers to reduce spatial dimensions, and fully connected layers for classification. Backpropagation can be applied to train CNNs.
- There are many popular CNN architectures like LeNet, AlexNet, VGG, Inception, ResNet, and DenseNet. They demonstrate improving performance on image classification tasks, enabled by growing model size and depth.
- CNNs have been applied successfully to various healthcare applications including diabetic retinopathy detection, skin cancer classification, COVID-19 diagnosis from x-rays. Sufficient labeled training data is key to the success of CNNs.
- Case studies show CNNs can achieve performance comparable to healthcare experts on disease detection in images. However, challenges remain regarding model generalization and reliance on large labeled datasets.
- CNNs are powerful deep learning models for healthcare image analysis tasks, enabled by model architectures that leverage spatial structure in image data. When trained on sizable labeled datasets, they can classify and detect diseases automatically to assist clinicians.

helpful! | 0

Reply to this followup discussion

Resolved Unresolved @464_f3

Darin Zhen 2 days ago

Lecture 06 CNN

- CNNs are designed to process grid-like data like images by utilizing convolutional and pooling layers. These layers help the network handle translation invariance and be more parameter efficient.
- Convolutional layers convolve the input with a set of filters to produce feature maps. Parameters are shared across spatial locations in the same layer leading to sparse interactions and parameter sharing.
- Pooling layers (max, average etc.) reduce the spatial dimensions of the feature maps and provide translational invariance.
- AlexNet, VGG, Inception and ResNet are some influential CNN architectures. They have progressively become deeper and more accurate on image classification tasks.
- CNNs have been applied to healthcare use cases like detecting diabetic retinopathy and skin cancer. They can match or exceed human expert performance in some cases.

- Key hyperparameters in CNNs include number of filters, filter size, stride, padding in conv layers and pooling window size, stride in pooling layers. These impact the output dimensions.
- Most calculations happen in the convolutional layers. Fully connected layers contain most of the parameters.
- CNNs are specialized neural network architectures for processing grid-like data like images by using convolutions and pooling. They have sparsely connected layers and parameter sharing schemes which makes them efficient for such tasks.