Convolutional Neural Networks II

COSC 410: Applied Machine Learning
Fall 2025
Prof. Forrest Davis

November 4, 2025

Warm-up

1. Discuss with your neighbor the highlight of your family weekend.
2. Apply Mean Pooling (like Max Pooling, which we have seen, but now the mean) with a kernel of size 2×3 and a stride of one to the following matrix:

$$\begin{bmatrix} 5 & 8 & 4 & 3 \\ 6 & 6 & 1 & 2 \\ 5 & 5 & 5 & 7 \end{bmatrix}$$

Logistics

- Codelet 5 is due Nov 14
- Midterm Exam II is now Nov. 19

Learning Objectives

- Describe the graphical representation of a CNN
- Describe three core tasks, including their input-output, their loss, and ways to evaluate their output extrinsically
- Understand some additional design decisions in real world settings

Summary: We conclude our discussion of convolutional neural networks with some real world applications and ways to interpret models.

Another Way to Motivate Deeper CNNs

LAST CLASS, WE TACITLY motivated deeper CNNs (i.e., CNNs with many layers) by pointing out early layers learn basic features while later layers learn high order features (like ears, eyes, etc). Under that motivation, we want a deeper CNN because it allows the model to learn more complex relationship in the data. Here's another way to show this. Work through the following points with a small group.

Question

1. Draw a simple fully connected one layer feed-forward neural neural network mapping 5 inputs to 5 hidden nodes. Ignore the bias and label the weights.
2. The *receptive* field for every hidden node in the network you drew is 5, as it takes into account five nodes in the input. Draw a modified version of your network with a receptive field of 3. Label the weights and ignore the bias.
3. Finally, draw a modified version of the network you just drew. This time add another hidden layer with five nodes and a receptive field of 3 to the layer below.
4. How can we relabel the weights for the figure you drew for question 2 so it represents applying a kernel of size 1×3 to an image that has 5 pixels?
5. What is the receptive field for the middle node, in terms of the input nodes, in the final layer for the node you drew in response to question 3?

Application One: Image Classification

THE CORE TASK OF IMAGE CLASSIFICATION is to label images with the content of the image.

- **Dataset Example:** ImageNet

Practice Problems

Work with a small group to apply what you've learned in class to give an approach you would take to operationalizing this task.

1. What are the inputs to the task and what are the outputs?
2. What loss function would you use to train the model?
3. What measure(s) would you use to evaluate the ability of your system?

Application Two: Object Detection

THE CORE TASK OF OBJECT DETECTION is to identify the region of an image that corresponds to a label.

- **Dataset Example:** Open Images Dataset

Practice Problems

Work with a small group to apply what you've learned in class to give an approach you would take to operationalizing this task.

1. What are the inputs to the task and what are the outputs?
2. What loss function would you use to train the model?
3. What measure(s) would you use to evaluate the ability of your system?

Application Three: Semantic Segmentation

THE CORE TASK OF SEMANTIC SEGMENTATION is to identify the region of an image that corresponds to a label.

- **Dataset Example:** COCO: Common Objects in Context

Practice Problems

Work with a small group to apply what you've learned in class to give an approach you would take to operationalizing this task.

1. What are the inputs to the task and what are the outputs?
2. What loss function would you use to train the model?
3. What measure(s) would you use to evaluate the ability of your system?

Before Next Class

- Reading and pre-class quiz
- Review/Work on Codelet 5