Course-4 Week1 Summary: Convolutional Neural Networks

Computer Vision

Computer vision is one of the areas that's been advancing rapidly thanks to deep learning. Deep learning computer vision is now helping self-driving cars figure out where the other cars and pedestrians around so as to avoid them. Making face recognition work much better than ever before, so that perhaps some of you will soon, or perhaps already, be able to unlock a phone, unlock even a door using just your face. And there are many apps that show you pictures of food, or pictures of a hotel, or just fun pictures of scenery. And some of the companies that build those apps are using deep learning to help show you the most attractive, the most beautiful, or the most relevant pictures. For example, when I was working on speech recognition, I sometimes actually took inspiration from ideas from computer vision and borrowed them into the speech literature.

So, if you're building a self-driving car, maybe you don't just need to figure out that there are other cars in this image. But instead, you need to figure out the position of the other cars in this picture, so that your car can avoid them. In object detection, usually, we have to not just figure out that these other objects say cars and picture, but also draw boxes around them. We have some other way of recognizing where in the picture are these objects. They can be multiple cars in the same picture, or at least every one of them within a certain distance of your car.

Let's see how you can implement this and illustrate convolutions, using the example of Edge Detection.

Edge Detection Example

The convolution operation is one of the fundamental building blocks of a convolutional neural network.

Using edge detection as the motivating example in this video, you will see how the convolution operation. Given a picture like that for a computer to figure out what are the objects in this picture, the first thing you might do is maybe detect vertical edges in this image. The convolution operation gives you a convenient way to specify how to find these vertical edges in an image.

How to take this and use it as one of the basic building blocks of a Convolution Neural Network.

Padding

In order to build deep neural networks one modification to the basic convolutional operation that you need to really use is padding.

Let's talk about how you can implement Strided convolutions.

Strided Convolutions

It is another piece of the basic building block of convolutions as used in Convolutional Neural Networks.

Convolutions Over Volume

You've seen how convolutions over 2D images works.

Now, let's see how you can implement convolutions over, not just 2D images, but over three dimensional

In order to detect edges or some other feature in this image, you can vault this, not with a three by three filter, as we have previously, but now with also with a 3D filter, that's going to be three by three by three.

So the filter itself will also have three layers corresponding to the red, green, and blue channels.

So to give these things some names, this first six here, that's the height of the image, that's the width, and this three is the number of channels. And your filter also similarly has a height, a width, and the number of channels. The number of channels in your image must match the number of channels in your filter, so these two numbers have to be equal.

One Layer of a Convolutional Network

So now that you know how to implement convolutions over volumes, you now are ready to implement one layer of the convolutional neural network.

Simple Convolutional Network Example

you saw the building blocks of a single layer, of a single convolution layer in the ConvNet.

Now let's go through a concrete example of a deep convolutional neural network. And this will give you some practice with the notation that we introduced toward the end of the last video as well.

Let's say you have an image, and you want to do image classification, or image recognition. Where you want to take as input an image, x, and decide is this a cat or not, 0 or 1, so it's a classification problem. A ConvNet you could use for this task. So this would be a pretty typical example of a ConvNet.

What's the stride? What's the padding and how many filters are used?

And both later this week as well as next week, we'll give some suggestions and some guidelines on how to make these choices. But for now, maybe one thing to take away from this is that as you go deeper in a neural network, typically you start off with larger images, 39 by 39.

Pooling Layers

ConvNets often also use pooling layers to reduce the size of the representation, to speed the computation, as well as make some of the features that detects a bit more robust.

CNN Example

You now know pretty much all the building blocks of building a full convolutional neural network. A more complex example of a ConvNet.

Why convolutions? this final video for this week, let's talk a bit about why convolutions are so useful when you include them in your neural networks.

And then finally, let's briefly talk about how to put this all together and how you could train a convolution neural network when you have a label training

I think there are two main advantages of convolutional layers over just using fully connected layers.

And the advantages are parameter sharing and sparsity of connections.