

Computer vision

1. The process

00:00 - 00:13

In the previous lesson we mentioned that deep learning is especially well-suited for working with images and text. In this lesson we'll focus on images and how they are used in computer vision applications.

2. Computer vision

00:13 - 00:44

So, what is computer vision? The goal of computer vision is to help computers see and understand the content of digital images. Computer vision is necessary to enable, for example, self-driving cars. Manufacturers such as Tesla, BMW, Volvo, and Audi use multiple cameras to acquire images from the environment so that their self-driving cars can detect objects, lane markings, and traffic signs to safely drive.

3. Image data

00:44 - 01:08

To understand how it works, we first need to find out what image data looks like. An image is made up of pixels. These pixels contain information about color and intensity. On the slide, you can see a grayscale pixelated image. Each pixel's intensity can be represented by a number between 0 and 255.

4. Image data

01:08 - 01:44

But what about colored images? A colored image is generally stored in the RGB system. RGB stands for Red, Green, and Blue. Each image can be thought of as being represented by three rasters, one for each color channel. This means that you need three times the amount of data to store a color image compared to a grayscale one. So, digital images can actually be seen as a bunch of numbers. Just like before, these numbers can be used as features for your machine learning model.

5. Face recognition

01:44 - 01:57

Imagine you want to build a system to recognize people from pictures. The first step is to get some pictures, in this case from your lovely instructors, and use these as input.

6. Face recognition

01:57 - 02:02

The intensities of each pixel can be passed into a neural network.

7. Face recognition

02:02 - 02:07

Its job is to figure out the identity of the person on the picture.

8. Face recognition

02:07 - 02:36

Just like before, the neurons in the middle will compute various values by themselves. Typically, when feeding a neural network images, neurons in the earlier stages will learn to detect edges later on parts of objects, like eyes and noses for example, the final neurons will learn to detect shapes of faces. In the end, the network will put all of this together to output the identity of the person in the image.

9. Training the neural network

02:36 - 02:59

Don't forget, part of the magic of neural networks is that you don't really need to worry about what it is doing in the middle. All you need to do is give it a lot of images of faces, the features, as well as the correct identity, the labels, and during training the learning algorithm will figure out by itself what each of these neurons in the middle should be computing.

10. Applications

02:59 - 03:40

Many popular computer vision applications involve recognizing things in images. Examples include facial recognition, as you saw in the example, self-driving vehicles, automatic detection of tumors in CT scans, and many more. Not only can computer vision applications understand images, but we're also at the point, where they can create realistic images. For example, deep fake is a software that is used to depict people in fake videos they did not actually appear in. By understanding what makes up a human face, deep fake can generate new faces.

11. Let's practice!

03:40 - 03:46

Let's move on to some exercises and see how familiar you are with computer vision!