Computer Vision Primer

Author: Colin Cumming

What is Computer Vision?

Computer vision is a field of artificial intelligence that trains computers to interpret and understand the visual world [1]. Essentially, you are trying to enable machines to react to images in a controlled fashion, much like humans do. Computer vision has already been implemented into many industries and will continue to as accuracy and technology advances. It is being used for facial recognition in security systems, analyzing banking information, detecting disease in plant life.

Types of Computer Vision

Computer Vision is not as simple as one way of interpreting images. Depending on your desired outcome, certain techniques may need to be applied. In more advanced applications such as self driving cars, multiple techniques may be used at different times (edge detection, image classification, etc.) to accomplish the goal of the entire system. Some examples of computer vision are:

Image segmentation → creating regions within an image based on specific criteria (colour, type of object, texture).

Object detection → identifying an object in an image. Could also be multiple objects in an image.

Facial recognition → finding and identifying human faces in an image. Edge detection – find boundaries/edges in an image. Unlike image segmentation, edge detection looks at the edges which are creating regions, image segmentation is finding distinct regions in an image.

Image classification → grouping images given a specific criterion.

Feature matching → matching similarities in an image to potentially locate or classify objects between multiple images.

Important Libraries

With the growth of computer vision many tools and libraries have been implemented into popular development platforms like Python and C++. Let's cover some important libraries with prebuilt tools and techniques.

OpenCV is an open-source computer vision library. It is available in Python and C++. It is the most popular computer vision library which means there is lots of educational information about there for using its many modules. It has prebuilt modules for many of the techniques listed in the above section which makes it great for quickly developing models and testing.

TensorFlow is a machine learning library that specializes in deep learning algorithms. This library can be useful when implementing classification and object detection which requires deep learning and less structured data.

Numpy is a data science library. This may seem out of place when mentioning computer vision libraries. However, OpenCV can handle images as Numpy arrays and much of your data preprocessing and

reading can be done in Numpy. Treating images as Numpy arrays with each array value pertaining to pixel values is important to implementing any computer vision techniques.

Conclusion

Hopefully by now you have learned the basics of what computer vision is and how you might implement it into your own development projects. Computer vision is an integral part of many industries and is an important skill to learn as an AI developer. Below are some further resources to explore if you are interested in learning more.

https://www.youtube.com/watch?v=OcycT1Jwsns

https://www.youtube.com/playlist?list=PLf7L7Kg8_FNxHATtLwDceyh72QQL9pvpQ