✓ 100 XP

Introduction

3 minutes

Suppose you have images that contain text. The images can include road signs, scanned documents that are in an image file format such as JPEG or PNG, or a picture of a white board that was used during a meeting. Perhaps the text is printed, typewritten, or even includes handwriting.

The ability for computer systems to process written and printed text is an area of artificial intelligence (AI) where *computer vision* intersects with *natural language processing*. Vision capabilities are needed to "read" the text, and then natural language processing capabilities make sense of it.

The foundation of processing text in images is *optical character recognition* (OCR), in which a model can be trained to recognize individual shapes as letters, numerals, punctuation, or other elements of text. Much of the early work on implementing this kind of capability was performed by postal services to support automatic sorting of mail based on postal codes. Since then, the state-of-the-art for reading text has moved on, and it's now possible to build models that can detect printed or handwritten text in an image and read it line-by-line and word-by-word.

In this module, we'll focus on the use of OCR technologies to detect text in images and convert it into a text-based data format, which can then be stored, printed, or used as the input for further processing or analysis.

Uses of OCR

The ability to recognize printed and handwritten text in images, is beneficial in scenarios such as note taking, digitizing medical records or historical documents, scanning checks for bank deposits, and more.

Next unit: Get started with Vision Studio on Azure

Continue >