

Computer Vision

**Conor Weldon**

**N00191746**

Computer Vision, CA2

DL836 BSc (Hons) in Creative Computing

Year 4, Term 2

Table of Contents

[Acknowledgments 3](#_Toc124629242)

[Understand computer vision 4](#_Toc124629243)

[Computer Vision models and capabilities 4](#_Toc124629244)

[Computer vision services in Microsoft Azure 7](#_Toc124629245)

# Acknowledgments

The author would like to thank the reviewer Tim Mc Nichols, Faculty of Computer Science, (Creative Computing), Institute of Art, Design & Technology, for very useful comments and suggestions.

He would also like to thank his colleagues in work who are and are not, part of the MTC team.  
  
Ireland, Dublin   
Conor P. Weldon January 2023

# Understand computer vision

Visual processing is the focus of the branch of AI known as computer vision. Let's examine a few of the potential applications of computer vision.

A fantastic illustration of the capabilities of computer vision is the Seeing AI app. The Seeing AI software, created for the blind and low vision users, uses artificial intelligence to describe surrounding people, text, and objects and to open up the visual world.

## Computer Vision models and capabilities

Machine learning models, which may be used to process sensory information from cameras, movies, or photos, are the foundation of the majority of computer vision solutions. Traditional computer vision tasks are outlined in the table below.

| **Task** | **Description** |
| --- | --- |
| Image classification | An image of a taxi with the label "Taxi"  In order to categorize photos premised on their properties, a machine learning model must be trained. For instance, you could employ an image classification model in some kind of a traffic surveillance system to categorize photographs depending on the kinds of vehicles they show, which including taxis, buses, bicycles, and so forth. |
| Object detection | An image of a street with buses, cars, and cyclists identified and highlighted with a bounding box  Machine learning models for object recognition are taught to categorize specific items inside an image and pinpoint their exact locations using bounding boxes. For instance, a traffic monitoring system may employ object detection to locate various automobile classifications. |
| Semantic segmentation | An image of a street with the pixels belonging to buses, cars, and cyclists identified  A sophisticated ML approach called semantic segmentation allows for the classification of distinct picture pixels in accordance with the item with which they correspond. To emphasize distinct cars using certain colours, a traffic surveillance system, for instance, can overlay traffic photos with "mask" layers. |
| Image analysis | An image of a person with a dog on a street and the caption "A person with a dog on a street"  To retrieve data from photos, such as "tags" that might assist classifying the picture or even provide meaningful comments that briefly describe the scenario depicted in the picture, developers can develop systems that incorporate both machine learning models with sophisticated image analysis techniques. |
| Face detection, analysis, and recognition | An image of multiple people on a city street with their faces highlighted  Finding a person or peoples faces in a picture is done using a particular kind of object detection called face detection. This may be used in conjunction with facial geometry analysis as well as segmentation methods to identify people according to their face traits. |
| Optical character recognition (OCR) | An image of a building with the sign "Toronto Dominion Bank", which is highlighted  To find and understand text in photographs, a technique called optical character recognition is utilized. OCR may be utilized to retrieve data from scanned documents like emails, bills, and forms as well as to scan text from images of things like business fronts or road signage. |

## Computer vision services in Microsoft Azure

Microsoft Azure provides the following cognitive services to help you create computer vision solutions:

| **Service** | **Capabilities** |
| --- | --- |
| **Computer Vision** | With the help of this service, you may analyse photographs and videos and retrieve text, "tags", objects, and summaries from them. |
| **Custom Vision** | Utilize this tool to educate personalized object recognition and picture classification models by utilizing your own photographs. |
| **Face** | You may create face detection and facial recognition applications using the Face service. |
| **Form Recognizer** | Utilize this service to get data out of scanned documents and invoices.. |