



One of the most common uses of Computer vision is in autonomous cars

Automating tasks of visual recognition

Computer vision is one of the fields of artificial intelligence which was developed from digital image processing, and whose focus is enabling computer systems to retrieve information from images and to understand images in the same way humans do.

Filip Štetić, Machine Learning Engineer, Megatrend Poslovna Rješenja

There are various ways of using computer vision, those include:

- Image classification** – the system gets an image with an object as input and outputs which class that object belongs to. For example: input is an image with a cat on it and classes are “Dog” and “Cat”
- Object detection** – the system has to find the locations of objects in an image
- Image segmentation** – partitioning an image into multiple segments with similar characteristics

Object tracking – mostly in videos where a user is interested in object movement from frame to frame

The main concept used today are neural networks. It's an algorithm which mimics the human neural system using the idea of neuron activation and neuron connection using synapses to model mathematical functions.

DEVELOPMENT

Work on computer image processing started more than 60 years ago, in 1957, at the

First digital scan of an image
Credit: NIST



American National Institute for Standards and Technology, where a group of engineers, led by Russell Kirsch, succeeded in making the first ever digital scan of an image. The image had become so famous that the Life magazine included it in its

article about 100 images that changed the world.

Lawrence Roberts is generally considered the father of computer vision. In his doctorate thesis in 1963 at MIT, he presented a process for getting information about a 3D object from a 2D image. He is also interesting because he later went to work for DARPA and took part in developing the Internet.

A step towards current methods was made in 1979, when a Japanese computer scientist Kunihiko Fukushima developed an artificial network for pattern recognition, which was made from convolutional layers. They were revolutionary because they treated a part of the image as a whole and in that way they didn't ignore mutual dependence of neighboring pixels. He called it Neocognitron and it is

undoubtedly the origin of networks which are even currently dominating the world of automatic visual recognition.

Development of faster and more efficient image processing algorithms enabled the making of larger datasets. The ImageNet dataset, undoubtedly the most famous visual recognition dataset, was made in 2010 and it consists of more than a million images divided into 1000 classes of everyday objects (animals, types of balls, modes of transport, etc.). Today's precision for computer systems is more than 97%, while human precision is around 95%.

USES

Computer vision found its use in many industrial branches. Some include:

Autonomous vehicles

Currently a widespread topic are autonomous vehicles. Driving being one of the most complex actions people perform makes the task of autonomous driving complex and hard. The first successful attempt happened in 1994 in Paris when a team, led by German computer scientist Ernst Dickmann, drove in two Mercedes 500 SELs. They drove up to 130 km/h, changed lanes and reacted to cars around them, all autonomously.

Today the development is led by Tesla, their vehicles offer the *Full Self-Driving* package which enables autonomous driving on highways. Information about the environment is obtained from eight cameras mounted in a way which creates a 360 degrees' field of view and on distances of up to 250 meters.

Health care

90% of medical data is in the form of images so it's natural to use computer vision as a helping tool for doctors. Using it increases speed and accuracy of diagnosis, especially with conditions which need early detection, like tumors or cancers. It's also used in monitor-

ing the health and symptoms of patients during and after operations. For example, in the Winnie Palmer Hospital in Orlando they use computer vision to more accurately keep track of blood loss during childbirth.

Facial recognition

Facial recognition is an important feature in Meta (former Facebook) and Apple. Meta started with it in 2010, which brought up the talk about privacy and data security. They use it in recognizing users in images and generating alternative image captions for visually impaired and blind users. At the end of 2021 they decided to get rid of it because of growing societal concerns, and lack of clear rules. Apple uses it in organizing private images, grouping those which contain the same person. They also added facial recognition as a biometric authentication for unlocking a device in 2017.

Agriculture

Computer vision in agriculture helps in cheap, efficient and highly accurate automa-



Rear seat of the Mercedes where the computer was placed for easier access and monitoring. Credit: Reinhold Behringer

tion. Some of the tasks it solves are: monitoring for healthy crop growth, prevention and control of diseases, insects and weeds, automatic harvesting using machines or robots, and all of that in a fully automated farm. A farm like that was started in 2018 by the American company Iron Ox. Apart from sensors for monitoring they also use two robots: *Grover*, a flat driving robot that moves plant stands depending on the needs, and *Phil*, a robot that takes care of nutri-

ents and watering of plants.

CONCLUSION

Although the growth of computer vision is impressive and it already provides useful and applicable solutions, the development in the field is far from finished. There are many more problems which aren't solved, but given the size of investments and market size which was estimated to be worth 11.3 billion dollars in 2020., the trend of progress will continue.



Grover (under the table) and Phil (on the right). Credit: Iron Ox