Introduction to computer vision and real-time video analytics



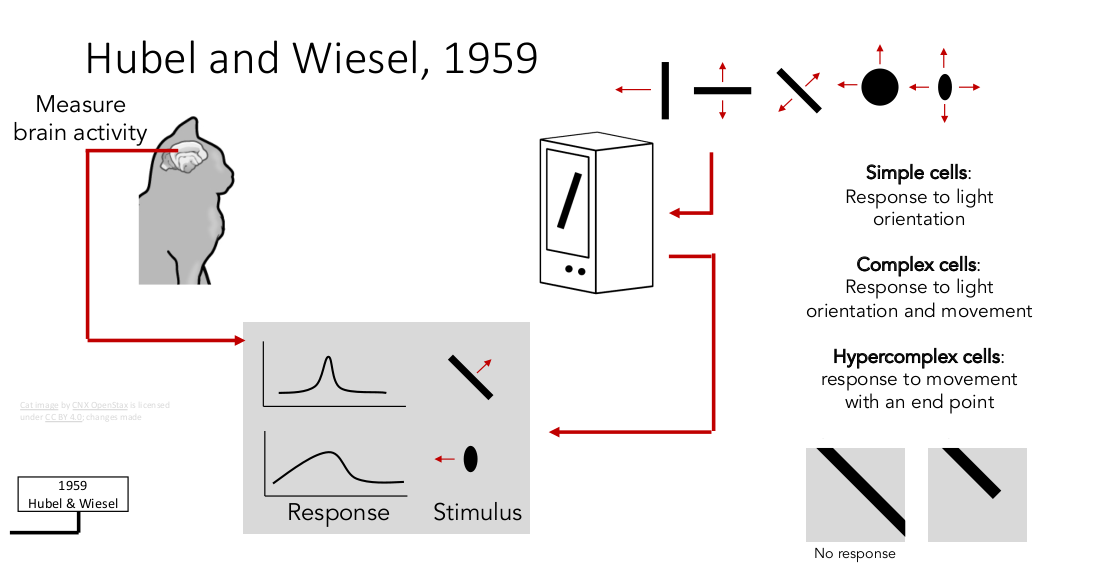
Consider every single computer with a digital camera around you to be an eye looking on, storing information, identifying and processing images similar to the human eye, and slowly understanding the visual world. In the simplest terms, this is Computer Vision(CV). CV employs a combination of Artificial Intelligence and Digital signal processing to capture and analyze visual scenes.

Understanding the industry relevance of computer vision is easy, in a single word, Automation. We are at the point where Robots have human-like recognition capabilities. This feature is automated for detecting data patterns and interacting with efficiency.

The efficacy of computer vision is probed in different fields of astrophysics, robotics, VDI(virtual desktop interfaces), cancer treatment, data processing, surveillance to name a few.

### **The history behind Computer Vision**

A popular misconception is that computer vision is a new technology but history tells us differently. As early as the 1960s, the technology behind CV was in development. It started with two neurophysiologists(David Hubel and Torsten Wiesel )  who wrote a paper( “Receptive fields of single neurons in the cat’s striate cortex”), where they used visual stimuli to elicit a response from a cat to understand if it shaped its cortical architecture.



In the experiment, an anesthetized cat had electrodes placed on the primary visual cortex area of the brain.

The aim was to study neuronal activity in the region while displaying various images. At first, all their efforts seem to be in vain. But in a few months, they noticed that the feline had a neuron response when it reacted to a line produced while changing slides in the projector.

The research established that simple and complex neurons are present in the primary visual cortex and that oriented edges are first detected at the start of visual processing. This is the core principle behind Deep learning which is at the core of Computer Vision.

The precursor of the modern computer vision was discovered in 1963 by Lawrence Roberts when he published a Ph.D. Thesis on “Machine perception of three-dimensional solids”, in which he described the visual world in dimensional objects of 2D and 3D.

His finding was that the process of 2D to 3D construction and the following display of 3D to 2D is the right starting point for future research into computer-aided 3D systems.

A lot of research was done on Neural Networks. It involved producing the theory of CNN(Convolutional neural network) which is the basis for image classification, NLP(natural language processing), image, and video recognition to name a few.

It is composed of convoluted(hidden) layers of neural network forming a class of deep neural networks. The input is images and the process is based on predictive arrangements of neurons concerning height, width, and depth. The output is based on a class score.

In 2001 when the Object(face) detection framework theory was introduced by Paul Viola and Michael Jones which used a binary classifier to identify facial features.

Five years after the introduction of the Viola/Jones algorithm, Fujitsu the Japanese camera maker released a real-time face detection feature in their new devices which was the first foray into modern computer vision.

### **Now**

Computer vision has gone from being theory on paper to real-life run time application on our hands. We can identify objects and faces, recognize them, and even track them successfully through a combination CV with other technologies.

Modern CV is used in tandem with other technologies.

Computer vision along with Deep learning(part of the Machine learning paradigm) is used to trace the pattern in data collection. Artificial systems are now trained to mimic human reactions in images and even interpret them.

It is a complex interdisciplinary field with a far-reaching impact and relevance throughout all technologies. It is based on Artificial Intelligence and is now applied to Machine learning, and further down to Deep learning.

### **Industry Use Case**

There are many use cases for Computer vision in different fields of agriculture, surveillance data processing, manufacturing, entertainment, healthcare, and transportation.

In Agriculture, Precision cropping uses a technology where aerial imagery is combined with CV to weed out good crops and bad crops by ultra analyses of the available visual data. AirSurf is a recently developed software that combines the two to improve cropping. It was used successfully in lettuce farming with exceptional results.

In Healthcare, during the era of the pandemic CV was successfully employed to weed out and detect anomalies present in chest X rays to confirm the diagnosis of pneumonia. This was based on an experiment conducted by medical students at Cranfield University.

In Manufacturing, RPA(Robotic Process Automation) and Machine vision are utilized in unmanned aerial vehicles (UAVs) to increase the field of vision and also used to go through and classify photographs to filter out abusive images. IoT is already automated and used in the manufacturing sector and along with computer vision it can eliminate human intervention in production lines.

### **Real-time video analytics in Computer Vision**

Why is there a significance of using video in Computer Vision?

Videos are important as they convey more information, they pack in a lot more information in comparison with a mere picture. Tracking an object/person is easily done in a video but impossible through a picture. There are obvious cons to using video analytics with Computer Vision. They are data storage and data processing. But we are at the point where processing power and storage can meet the threshold for this tech to work.

Video content curated in real-time together with computer vision provides the perfect platform for advanced surveillance. Currently, Amazon employs deep learning and CV together in the real-time sorting of vegetables and fruits. It is based on software produced by Anolytics.

PowerAI Vision is a software suite that uses deep learning models that are trained by Automatic machine learning methods for the detection and classification of objects and images. It was recently used by IBM for its in-house CV software to produce IBM Intelligent Video Analytics (IVA). It has a cross-industry appeal in the analytics in asset management, retail, worker safety compliance to name a few.



The high-resolution thermal cameras along with deep learning were used to capture infrared images of passengers in Taiwan during the covid 19 pandemic and these high-resolution images are run over computer vision algorithms to detect possibly infected passengers.

The self-driving cars are being tested by different major international conglomerates such as Google and Tesla. They are an expert use of real-time video analytics with multiple tools like ultrasonic sensors and LiDAR along with Computer vision to operate autonomous vehicles. This technology helps in successfully identifying road signs, obstacles, incoming traffic to safely orient the vehicle accordingly.

Using video surveillance to determine patterns of human behavior, crowd analysis, object recognition, and human movement could be used for security or in improving performance at workplaces.

### **Future**

The question that remains unanswered is that ‘Have we successfully tapped the full potential of Computer vision?’. The answer would be a resounding ‘No’, there are so many untapped and unexplored benefits and applications we are yet to inspect.