



XD IDEAS / PRINCIPLES / EMERGING TECHNOLOGY

What Is Computer Vision & How Does it Work?

An Introduction



Nick Babich

Jul 28, 2020



For many decades, people dreamed of creating machines with the characteristics of human intelligence, those that can think and act like humans. One of the most fascinating ideas was to give computers the ability to "see" and interpret the world around them. The fiction of yesterday has become the fact of today.

Thanks to advancements in artificial intelligence and computational power, computer vision technology has taken a huge leap toward integration in our daily lives. The computer vision market is expected to reach \$48.6 billion by 2022, making it an extremely promising UX technology.

In this article, we will review the concept of computer vision, discuss how this technology evolved, and share a few excellent examples where this technology can be applied in our lives.

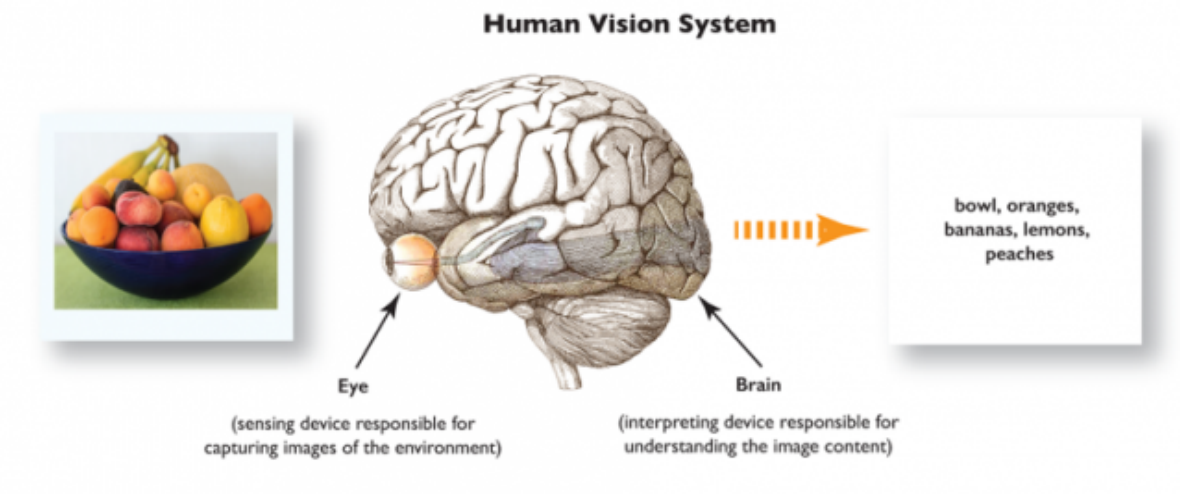
What is computer vision?

Analysis of visual data based on the human thought process!

Computer vision is the field of computer science that focuses on creating digital systems that can process, analyze, and make sense of visual data (images or videos) in the same way that humans do. The concept of computer vision is based on teaching computers to process an image at a pixel level and understand it. Technically, machines attempt to retrieve visual information, handle it, and interpret results through special software algorithms.

Based on pixel level!





Model to simulate human interpretation of images

Human vision and computer vision systems process visual data in a similar way. Image credit [manning](#).

Here are a few common tasks that computer vision systems can be used for:

- **Object classification.** The system parses visual content and classifies the object on a photo/video to the defined category. For example, the system can find a dog among all objects in the image.
- **Object identification.** The system parses visual content and identifies a particular object on a photo/video. For example, the system can find a specific dog among the dogs in the image.
- **Object tracking.** The system processes video finds the object (or objects) that match search criteria and track its movement.



How does computer vision work?

Computer vision technology tends to mimic the way the human brain works. But how does our brain solve visual object recognition? One of the popular [hypothesis](#) states that our brains rely on patterns to decode individual objects. This concept is used to create computer vision systems.

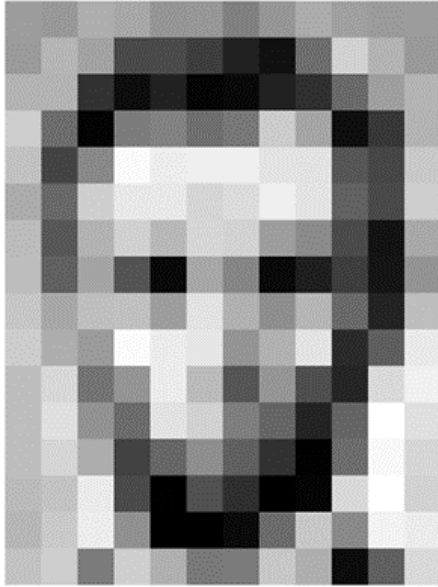
Concept of computer vision systems!

Computer vision algorithms that we use today are based on pattern recognition. We train computers on a massive amount of visual data—computers process images, label objects on them, and find patterns in those objects. For example, if we send a million images of flowers, the computer will analyze them, identify patterns that are similar to all flowers and, at the end of this process, will create a model “flower.” As a result, the computer will be able to accurately detect whether a particular image is a flower every time we send them pictures.

Golan Levin, in his article [Image Processing and Computer Vision](#), provides technical details about the process that machines follow in interpreting images. In short, machines interpret images as a series of pixels, each with their own set of color values. For example, below is a picture of Abraham Lincoln. Each pixel's brightness in this image is represented by a single 8-bit number, ranging from 0 (black) to 255 (white). These numbers are what software sees when you input an image. This data is provided as an input to the computer vision algorithm that will be responsible for further analysis and decision making.

Use of RGB for pixel values and serves as input.





157	153	174	168	150	152	129	151	172	161	155	156
155	182	163	74	75	62	33	17	110	210	180	154
180	180	50	14	34	6	10	33	48	106	159	181
206	109	5	124	131	111	120	204	166	15	56	180
194	68	137	251	237	239	239	228	227	87	71	201
172	105	207	233	233	214	220	239	228	98	74	206
188	88	179	209	185	215	211	158	139	75	20	169
189	97	165	84	10	168	134	11	31	62	22	148
199	168	191	193	158	227	178	143	182	106	36	190
205	174	155	252	236	231	149	178	228	43	95	234
190	216	116	149	236	187	86	150	79	38	218	241
190	224	147	108	227	210	127	102	36	101	255	224
190	214	173	66	103	143	96	50	2	109	249	215
187	196	235	75	1	81	47	0	6	217	255	211
183	202	237	145	0	0	12	108	200	138	243	236
195	206	123	207	177	121	123	200	175	13	96	218

157	153	174	168	150	152	129	151	172	161	155	156
155	182	163	74	75	62	33	17	110	210	180	154
180	180	50	14	34	6	10	33	48	106	159	181
206	109	5	124	131	111	120	204	166	15	56	180
194	68	137	251	237	239	239	228	227	87	71	201
172	105	207	233	233	214	220	239	228	98	74	206
188	88	179	209	185	215	211	158	139	75	20	169
189	97	165	84	10	168	134	11	31	62	22	148
199	168	191	193	158	227	178	143	182	106	36	190
205	174	155	252	236	231	149	178	228	43	95	234
190	216	116	149	236	187	86	150	79	38	218	241
190	224	147	108	227	210	127	102	36	101	255	224
190	214	173	66	103	143	96	50	2	109	249	215
187	196	235	75	1	81	47	0	6	217	255	211
183	202	237	145	0	0	12	108	200	138	243	236
195	206	123	207	177	121	123	200	175	13	96	218

Color values of individual pixels are converted into a simple array of numbers used as input for a computer vision algorithm. Image credit [openframeworks](#).

Storage of colors in an array serves as an input for a machine learning model.

The evolution of computer vision

Computer vision is not a new technology; the first experiments with computer vision started in the [1950s](#), and back then, it was used to interpret typewritten and handwritten text. At that time, computer vision analysis procedures were relatively simple but required a lot of work from human operators who had to provide data samples for analysis manually. As you probably guess, it was hard to provide a lot of data when doing it manually. Plus, the computational power wasn't good enough, so the error margin for this analysis was pretty high.

Through the use of new technology, computer vision will be a very big part of the future.

Today, we do not have any shortage of computer power. Cloud computing, paired with robust algorithms, can help us solve even the most complex problems. But not just the new hardware paired with sophisticated algorithms (we will review them in the next section) is driving computer vision technology forward; the impressive amount of publicly available visual data that we generate every day is responsible for the recent process of

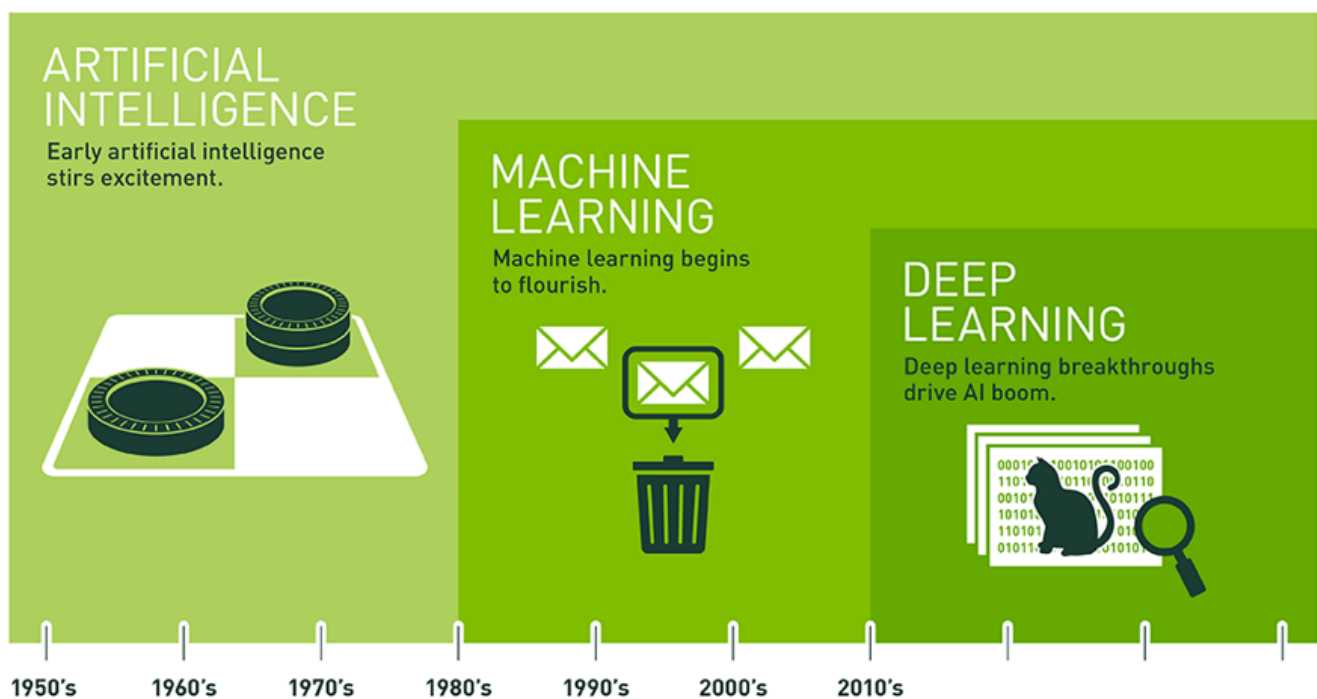


technology. According to [Forbes](#), users share online more than three billion images daily, and this data is used to train the computer vision systems.

Deep learning revolution

Connects to my interest to do deep learning in ISM!

To understand the recent process of computer vision technology, we need to dive into algorithms this technique relies on. Modern computer vision relies on deep learning, a specific subset of machine learning, which uses algorithms to glean insights from data. Machine learning, on the other hand, relies on artificial intelligence, which acts as a foundation for both technologies (check [AI design best practices](#) to learn more about design for AI).



Deep learning fits inside machine learning, a subset of artificial intelligence. Image credit [Nvidia](#).

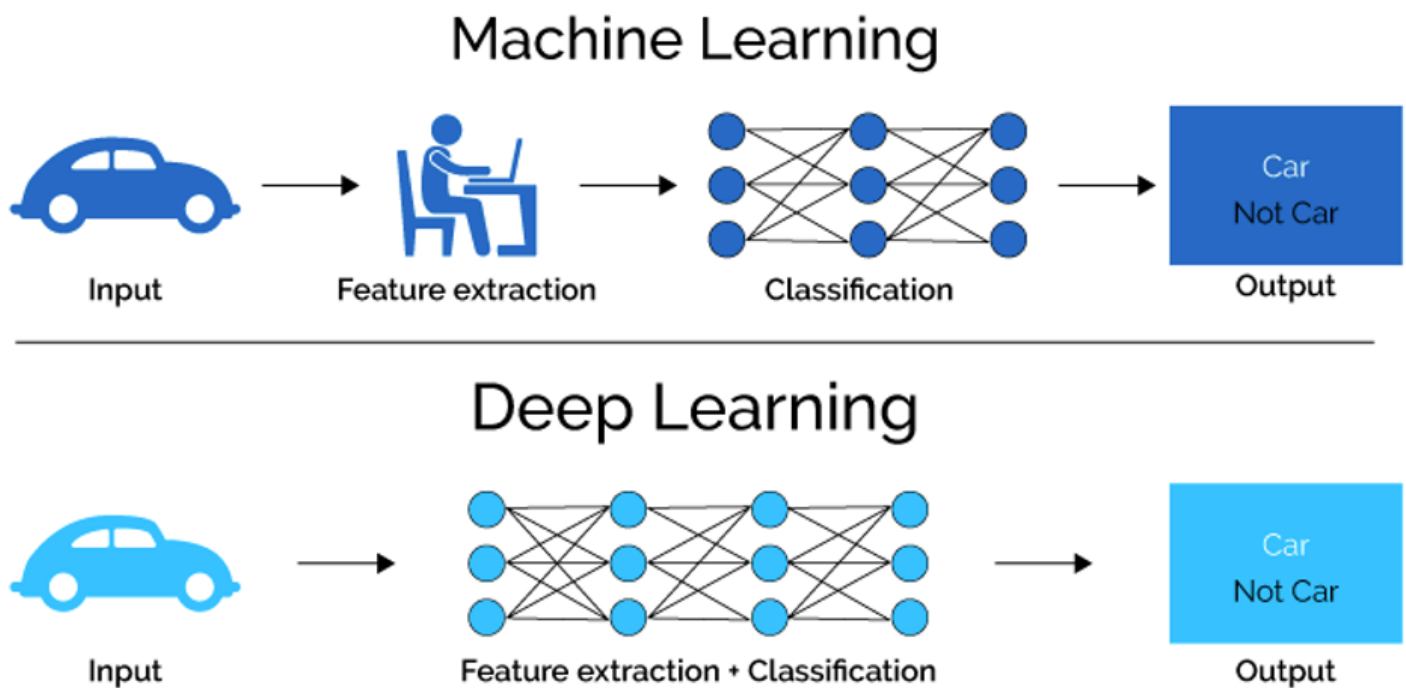
Deep learning represents a more effective way to do computer vision—it uses a specific algorithm called a neural network. The neural networks are



used to extract patterns from provided data samples. The algorithms are inspired by the human understanding of how brains function, in particular, the interconnections between the neurons in the cerebral cortex.

How a neural network works!

At the core level of a neural network is the perceptron, the mathematical representation of a biological neuron. Similar to biologic neurons in the cerebral cortex, it's possible to have several layers of interconnected perceptrons. Input values (raw data) get passed through the network created by perceptrons and end up in the output layer, which is a prediction, or a highly educated guess about a certain object. For example, at the end of the analysis, the machine can classify an object with X% confidence.



Machine learning uses algorithms to parse data while deep learning relies on layers of artificial neural networks (ANN). Image credit [Quora](#).

Where we can apply computer vision technology

Some people think that computer vision is something from the distant [future of design](#). Not true. Computer vision is already integrated into ma

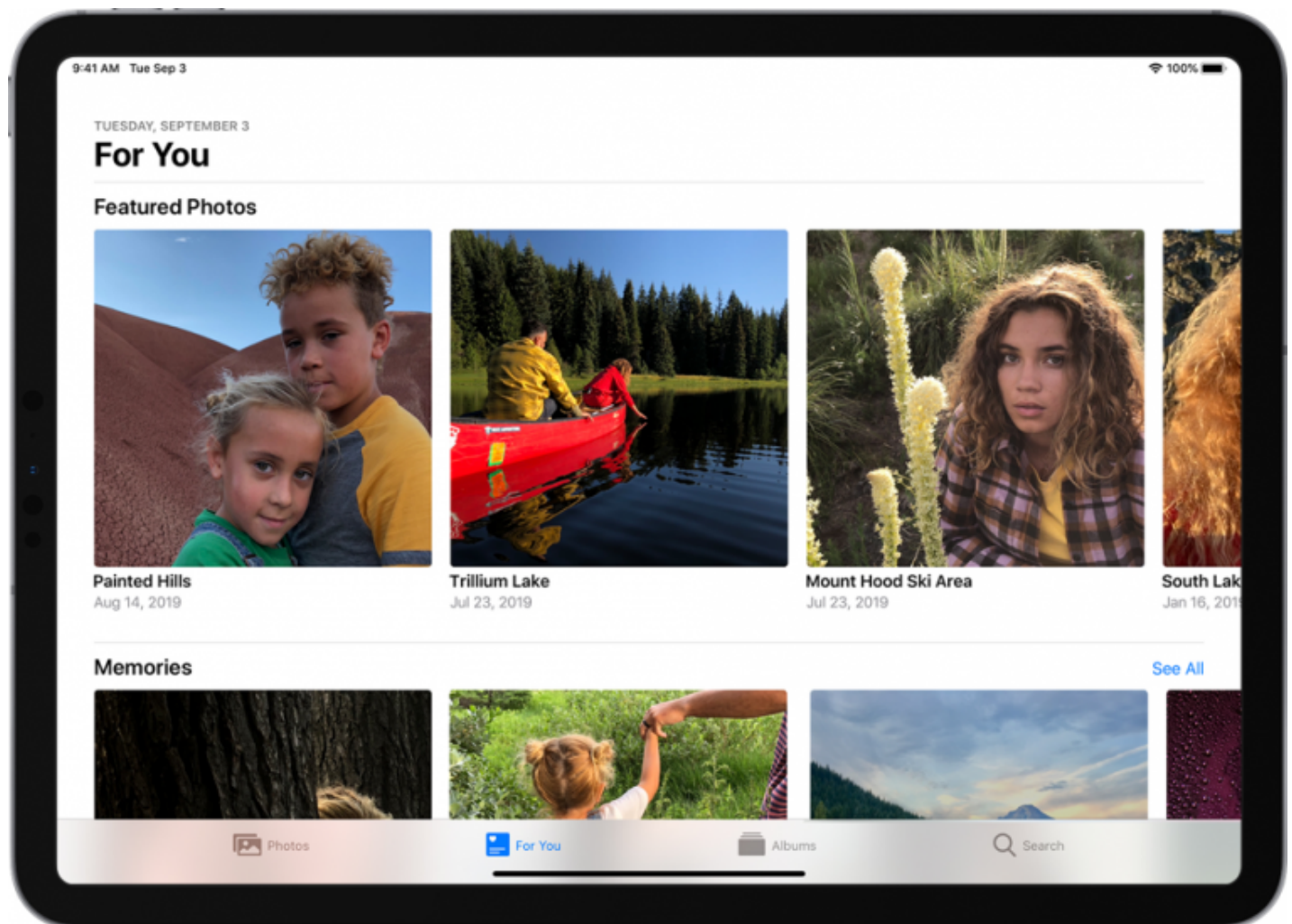


areas of our life. Below are just a few notable examples of how we use this technology today:

Content organization

Applications of computer vision!

Computer vision systems already help us organize our content. Apple Photos is an excellent example. The app has access to our photo collections, and it automatically adds tags to photos and allows us to browse a more structured collection of photographs. What makes Apple Photos great is that the app creates a curated view of your best moments for you.



In the For You section of Photos for iOS, you can see featured content that the app created so you can view your favorite moments. Image credit [Apple](#).



Facial recognition

Facial recognition technology is used to match photos of people's faces to their identities. This technology is integrated into major products that we use every day. For example, Facebook is using computer vision to identify people in photos.

Another application of using image data to recognize faces!

Facial recognition is a crucial technology for biometric authentication. Many mobile devices available on the market today allow users to unlock devices by showing their faces. A front face camera is used for facial recognition; mobile devices process this image and, based on analysis, can tell whether the person who is holding a device is authorized on this device. The beauty of this technology is that it works really fast.

Augmented reality

Object detection to detect physical space (a growing application of computer vision)

Computer vision is a core element of augmented reality apps. This technology helps AR apps to detect physical objects (both surfaces and individual objects within a given physical space) in real-time and use this information to place virtual objects within the physical environment.



The Ikea Place app uses AR to help users understand whether the furniture they want to buy will fit into their interior. Image credit [Wired](#).

Self-driving cars

Application in real world
currently (Tesla)

Computer vision enables cars to make sense of their surroundings. A smart vehicle has a few cameras that capture videos from different angles and send videos as an input signal to the computer vision software. The system processes the video in real-time and detects objects like road marking, objects near the car (such as pedestrians or other cars), traffic lights, etc. One of the most notable examples of applications of this technology is autopilot in Tesla cars.





The Tesla video demonstrates how autopilot works in Tesla Model 3. Video credit [YouTube](#).

Health

Diagnosing diseases with image recognition

Image information is a key element for diagnosis in medicine because it accounts for **90 percent of all medical data**. Many diagnoses in health are based on image processing—X-rays, MRI, and mammography, just to name a few. And image segmentation proved its effectiveness during medical scans analysis. For example, computer vision algorithms can detect **diabetic retinopathy**, the fastest-growing cause of blindness. Computer vision can process pictures of the back of the eye (see below) and rate them for disease presence and severity.



Computer vision algorithms can be used to process retinal fundus photographs to screen for diabetic retinopathy. Image credit [ai.googleblog](#).

Reliance on detection by computer vision algorithms

Cancer detection is another notable example. Accuracy in diagnosing different forms of cancer is vital. According to [Google](#), computer vision tools assist in detecting cancer metastasis with much higher precision than human doctors. Below you can see a closeup of a lymph node biopsy. The tissue contains a breast cancer metastasis as well as areas that look similar to the tumor but are benign. The computer vision algorithm successfully identifies the tumor region (bright green) and is not confused by the normal areas that look like tumors.



Applying computer vision technology during a lymph node biopsy can help detect the tumor region. Image credit [Google](#).

Agriculture

Many agricultural organizations employ computer vision to monitor the harvest and solve the **common agricultural problems** such as weeds emergence or nutrient deficiency. Computer vision systems process images from satellites, drones, or planes, and attempt to detect the problems in the early phase, which helps to avoid unnecessary financial losses.



Conclusion

Will be big in the future and thus I am very eager to learn about the back-end of computer vision now!

Computer vision is a popular topic in [articles about new technology](#). A different approach to using data is what makes this technology different.

Tremendous amounts of data that we create daily, which some people think as a curse of our generation, are actually used for our benefit—the data can teach computers to see and understand objects. This technology also demonstrates an important step that our civilization makes toward creating artificial intelligence that will be as sophisticated as humans.



Emerging Technology

Words by
Nick Babich



Nick Babich is UX architect and writer. Nick has spent the last 10 years working in the software industry with a specialized focus on research and development. He counts advertising, psychology, and cinema among his myriad interests.

