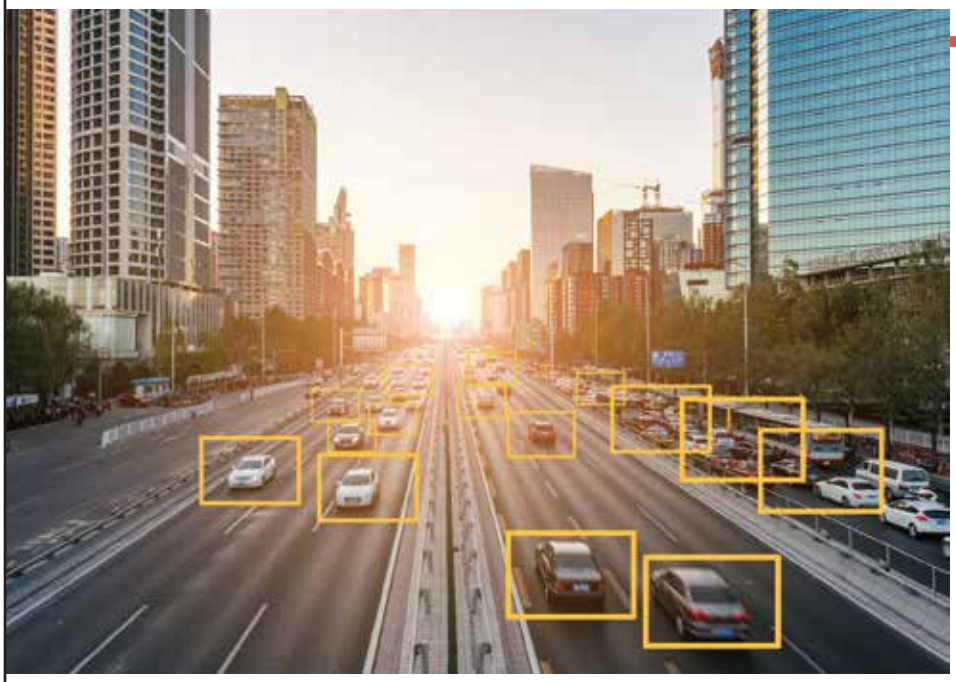
MACHINE LEARNING

UNDERSTANDING AI IN VIDEO SURVEILLANCE

Applying human intelligence to computer programs

By: Brian Carle

Many video surveillance professionals have come across the terms Artificial Intelligence (AI), Machine Learning (ML), and Deep Learning (DL). But what do those terms mean, and how do they affect Video Surveillance?



AI, Machine Learning and Deep Learning

AI is a term that loosely refers to applying human intelligence to computer programs or allowing programs to learn over time with the goal of producing better results as they learn. Machine Learning is a technique used to achieve a level of AI, and Deep Learning is an evolution of Machine Learning. In short, Deep Learning is an advanced, more sophisticated Machine Learning technique, and both are methods of achieving a level of AI.

Application in video surveillance. In video surveillance, video analytics uses Machine Learning and Deep Learning methods to identify objects, classify them, and determine their properties.

Whenever people receive new information, our brains attempt to compare the data to similar items in order to make sense of it. This comparative approach is the same concept that Machine and Deep Learning algorithms employ.

Machine and Deep Learning algorithms differ in how they are programmed to determine what constitutes a known object. Machine Learning requires more human intervention from a programmer to establish desired parameters in order to achieve the desired outcome. Deep Learning identifies object attributes independently and may consider characteristics the programmers would not.

Machine learning versus deep learning. What do Machine Learning and Deep Learning mean for Video Analytics? Both approaches describe programming methods where a system learns based on a data set. With Machine Learning, the attributes of the data a system looks for are usually preset, or corrected for, by human programmers. For instance, the system may be programmed to delineate an object that is taller than it is wide, with limbs moving in specified ways, and so on, and label this object a “person.”

Deep Learning is considered superior to Machine Learning, in part because the programmers may not recognize the most relevant criteria. Using the previous algorithm to identify a person, a seated and stationary person may not trigger an accurate detection.

With Deep Learning, the video analytic algorithms are fed an extensive data set representing an object. This step is called training, where the algorithm trains itself to recognize a type of object. For example, the system is fed thousands of images of people of varying genders, styles of clothing, ethnic backgrounds, images taken at different angles, and more.

The algorithm figures out attributes that are similar as well as dissimilar, and also determines how to weigh the relevance of those characteristics. After analyzing thousands of images, the algorithm may calculate the majority of images include a triangular-shaped object near the upper part of the image, with two darkened oval spots near its bottom, which we would think of as a nose on someone's face. In fact, the algorithm may have identified many other such characteristics we wouldn't think of.

Training the system is done by the developers of the software before it is used by a consumer. The process takes a substantial amount of computing power; much more than what is required to detect and classify objects when used in the field. The result is a file that is referenced by the system to determine if a detected object matches the classification.

Because the Deep Learning process uses the machine to determine object characteristics, it has led to analytics which can provide much more granular classification. For instance, older approaches may be able to detect a person, but Deep Learning based analytics can detect whether the person is a man, woman, or child. It may also be able to detect associated characteristics of an individual as well as vehicle type or make.

Learning over time. Typically, AI in video surveillance is trained at design time and, in some cases, does not get progressively "smarter" when used in the field. Deep Learning and Machine Learning do have this capability, however, and if used, can employ analytics which can learn over time.

Typical applications may include systems that determine what is normal in a scene. For instance, a school hallway experiences a rush of traffic about every 45 minutes between class periods. During that high traffic time, the traffic is dispersed and not concentrated in any particular area.

Furthermore, it is unusual for all the people to be moving at a very high speed. If the system detects an unusual concentration of objects, it could indicate a fight broke out. If all the people are running in the same direction outside of the usual inter-class period, it could indicate an emergency situation.

Smarter Systems, Better Results

Video surveillance systems produce huge volumes of data. Monitoring and filtering through such vast quantities of information makes the task of quickly identifying security incidents and finding evidence more difficult than ever.

Intelligent systems using Deep Learning can help us identify evidence much more promptly and analyze video in real-time to alert system operators of suspected events, providing better results for your security program.