ISSN 2661-2666 (Online) International Scientific Journal Monte (ISJM) DOI: 10.33807/monte.20232834 Volume 7, (No).2 (2023): April

# USE OF VIDEO ANALYTICS IN SECURITY SYSTEMS FOR CRIME PREVENTION AND EVIDENCE COLLECTION

**Prof.Asoc.Dr.Adrian Leka[1] Msc.Eraldi Ndoj[2]**
[1] leka-ad@live.com Luigj Gurakuqi University,Faculty of Law, Shkodër, Albania
[2] eraldindoj@gmail.com University of Tirana,Faculty of Law,Albania

## Entry

Video has always been an essential element for law enforcement agencies in maintaining and promoting public safety. From local police to elite specialized units, law enforcement agencies around the world rely on video surveillance to gather operational information and evidence needed to maintain order, protect citizens and enhance security, whether within a neighborhood or at state borders.

Although law enforcement agencies rely heavily on video surveillance, it is not always the most practical method. First, it is difficult to watch *live* video recordings. Any person charged with this task would be prone to distraction and error. Critical details can be overlooked if they do not pay full attention to the Video Management System ( *VMS* ) or recordings. Even if the person is fully focused, it is easy to miss important objects that appear, especially if there is a lot of activity in the area being recorded.

Apart *live video* viewing , video surveillance can also be used as supporting evidence in criminal investigations of incidents. Even this analysis does not escape from mistakes and lack of human attention, besides it takes a lot of time. Public and private spaces are monitored by multiple cameras, designed to cover every corner of the space. When records need to be checked, law enforcement agents must go through hours of records from several sources. Usually, it is impossible and ineffective to see all the records, so it is necessary to create a priority queue and set time limits for the completion of the work. Basically, it turns into a battle between the need to see all the evidence and the human resources needed for other investigations as well.

These needs of law enforcement agencies have given birth to Video *Content Analytics* (VCA). This paper will discuss how Video Content Analysis can be used to prevent crime and enable law enforcement and security agencies to overcome the challenges of video surveillance and utilize the full technological potential for productive video review and extracting predictive analytics from video content.

## What is artificial intelligence, how was it born and how did it develop?

Artificial intelligence (in English, *artificial intelligence* - AI) is a discipline that belongs to computer science and that studies the theoretical foundations, methodologies and techniques that allow the design of hardware systems and software systems that allow a computer to perform functions that, to the eye of an average observer , would seem to belong exclusively to human intelligence.

The above definition belongs to Professor Marko Somalviko [1], an eminent Italian engineer, specialized in the field of artificial intelligence, winner of the Joseph Engelberger International Prize for Robotics. With this brief description, it is possible to understand the essence of this discipline, which has been born and developed for years, but only recently has received wide recognition.

Artificial intelligence aims to mimic the human brain, which is based on neural networks or communities of closely interconnected neurons that change their configuration in response to external stimuli. In this sense, the brain has a learning function, and artificial models try to imitate this distinctive characteristic of it.

[1] Somalvico , Marco, *Intelligenza artificial* , in *Encyclopedia italiana di scienze letters and art* , Rome, Istituto vein Encyclopedia Italiana , 1991, Appendix V, pp. 735-738.

In the field of machine learning ( *Machine Learning* ), an artificial neural network (in English, *artificial neural network* - ANN) is a computational model composed of artificial *"neurons"* , inspired by a simplified model of a biological neural network .

The artificial neural network can be built from software programs, but also from dedicated hardware ( *Digital Signal Processing* or DSP). As early as 1943, scientists McSallow and Pitts [2]gave *"life"* to the first theoretical model of a simple artificial neuron. They describe an apparatus capable of receiving an *n number of* binary data [3]inputs to each of its elements, followed by a single output data for each. Such an apparatus is able to work with elementary Boolean functions [4]. In 1949, Heb [5], for the first time in history, raised the hypothesis of the possibility of instructing machines to learn in a manner similar to the way human intelligence learns.

This became a reality in 1958 by the hand of Frank Rosenblatt [6], the American psychologist who created the Perceptron, the first neural network based on the model of automated learning with a layer of input nodes (artificial neurons) and an output node. This model is *feedforward* , i.e. it is characterized by impulses that propagate in a single direction: forward. It deals with shape recognition, classifying them into two separate groups, and calculating simple functions. This, then, was a neural network with a limited scope.

With the birth of the *Multi Layer Perceptron* (MLP), this network became more complex. Inside it, between the input and output nodes, there is a hidden layer, where the information coming from the input layer is processed, which is then sent to the output node. It is a non-linear *feedforward* network : the input and output connections from each node are multiple. Thanks to this architecture, MLP can compute any function.

In 1986, starting from Werbo's thesis [7]on how the learning parameters of MLP can be established and thanks to the contribution of David Rumelhart, Jeffrey Hinton and Ronald Williams [8], the famous Error-Back Propagation was elaborated [9].

With the backpropagation algorithm we enter the present, because it is still used today. EBP allows to improve successive stages of automated learning of a neural network, until they take it to the level of deep learning. We are talking, therefore, about *Deep Learning.*

## Deep Learning

*Deep Learning* is that field of research on machine learning ( *Machine Learning* ) and on artificial intelligence, which is based on different levels of representation, corresponding to hierarchies of characteristics of factors or concepts, where high-level concepts are defined on base of low level ones. In other words [10], *Deep Learning* refers to a set of techniques based on artificial neural networks organized in different layers, where each layer calculates the values for the next layer, so that the information is always processed in a more complete way.

[2]Warren S. McCulloch, Walter H. Pitt's and the article The their : *"A Logical Calculus of the Ideas Immanent in Nervous Activity".*

[3] Number beam IS A number The EXPRESS IN SySTEm binary number , that is IN his method _ EXPRESSION MATHEMATICAL THAT use only two symbols : 0 and 1. System binary number is used BY almost THE all computer AND based devices _ IN computers .

[4] Boolean algebra or The logic boolean IS A branch of algebra THAT USE ABOUT THE created STATEMENTS THE truth / of fake . Expressions boolean use AND, OR, XOR and NOT operators for THE compare values AND ABOUT THE LEARNED A RESULT THE really OR THE false .

[5]DO Hebb.

[6]Frank Rosenblatt.

[7]Paul John Werbos AND the article The his "Beyond Regression: New Tools for Prediction and Analysis in the Behavioral Science", 1974.

[8] David E. Rumelhart, Geoffrey E. Hinton, Ronald J. Williams.

[9]This is it A algorithm THAT use DESCENT gradual . IN A NETWORK artificial nerve and A function error , algorithm Calculate error rate _ IN relative to the weight of the network . calculation proceeds BACKWARD through network , starting BY error rate _ IN the last layers of going TO those THE the money .

[10] According to the definition of the Artificial Intelligence Observatory of the Polytechnic of Milan.

The learning algorithms used to train neural networks depend on the field of application for which the network is designed and the typology of the network itself ( *feedforward* or *feedback* ). In this sense, *Deep Learning* is a community of automated learning techniques, through which artificial neural networks are exposed to large amounts of data, so that they can learn to perform certain tasks through automatic learning algorithms.

*Deep Learning* systems allow, among other things, to transcribe speech into text, to individualize and interpret the interests of network users, etc. In the field of medicine, the most advanced clinical diagnostic software is based on *Deep Learning systems* . In the field of motoring, efforts for automated driving of cars have started since 1925. This field, more tangible for the general public, although still without complete success, widely uses Deep Learning in various technologies: *cameras* , [11]radars [12], external sensors [13], front laser scanner [14], GPS satellite system [15]. It is thought that fully automated cars will be in circulation within the year 2030. But what is important for this work, *Deep Learning systems* are precisely those that allow to analyze the content of images and videos, identifying the people and objects that appear. in them.

## Artificial intelligence and video surveillance

Now that we have a clearer understanding of the concept of *Deep Learning* , we can enter more concretely into our topic and see how artificial intelligence is revolutionizing the video surveillance sector. Artificial intelligence and Deep Learning in this field appear in the form of video analysis or video content analysis.

The term *"video analysis"* refers to the ability of a camera surveillance (or video surveillance) system to analyze the content of recorded scenes and perform certain actions following the recording of a certain event that occurs in the recorded scene. or what happens after the evolution of the recorded scene. Thus, the video surveillance system that is equipped with video analysis not only records according to the modes defined in the system configurations, but, analyzing the content of the scenes in real time, is able to alert the operator of the control center about the occurrence of a event that may be dangerous or, at least, of interest. At this point, we want to take the opportunity to explain that video analysis is not an exclusive function of cameras, but is applied to many other video devices with security functions. In 2021, video analysis is found in cameras, network video recorders and related software, despite differences in analysis typology related to the type of technology or the characteristics of the manufacturing brand.

We'll start with the differences between video analytics, intelligent video analytics, and the latest developments in the field of video analytics. Everyone has heard talk, at least once in their life, about motion detection (in English, *motion detection* ). It is about the class of algorithms that allow to detect the movement of an object within a video. These algorithms are applied almost exclusively in video surveillance systems and can have different levels of sensitivity in video stream analysis. These algorithms work on pixels.

The most common example is the one found in most cameras that have a video analysis function and that allows, during system configuration, to create a line beyond which any movement generates an alarm (usually a fence or siege). In more modern versions, there is an opportunity to create entire areas of interest.

The simplicity of the operation of this algorithm places it outside the category of intelligent systems. The difficulty of using it lies in the fact that it generates a lot of false positive alarms, since it is activated by any movement, even by a change in lighting. This has been one of the factors that has influenced distrust towards video analysis. Today this functionality is advised to be used mainly to set the recording in motion. So the cameras start recording when a movement is detected in the defined area and stop when the state of stability returns. In this way, it is avoided to record when nothing happens, which saves the storage capacity of the recorder. This makes it even easier to search the records if needed later.

[11] To distinguish the obstacles provided and the lane lines.

[12] They allow detailed 360° views around the car and lane changes, as they detect cars approaching from behind.

[13] Identify objects when the car is moving at low speed.

[14] It allows more detailed scanning of the road ahead of the car and distinguishing potential hazards.

[15] It allows the car to understand where it is and what is around it.

Video analysis became moderately intelligent with the introduction of vertical algorithms, capable of identifying certain movement characteristics, such as falls (Slip&Fall Detection), direction of movement (Direction Detection) and the trajectory of objects and persons (Auto Tracker), or analyze objects that do not move, such as the detection of abandoned objects (Object Left Detection) or removed (Asset Detection).

The real intelligence came when the analysis was able to distinguish who or what it was analyzing (mainly, as we will see below, the human face, the human body, and machines). This was achieved through neural networks, which we talked about above. Intelligent algorithms stop working on pixels and work on objects.

Manufacturers of smart video surveillance devices work on the basis of large libraries of neural networks created by companies such as Google, Nvidia and Facebook to create algorithms that are able to recognize only objects that are important for security (cars, people, animals , etc.). The special algorithms are necessary because the data, with which the aforementioned networks have been supplied, have been general and do not specifically refer to security situations.

Initially, the cameras were designed to detect people and cars within the scene they were recording, without the need to create a security line or perimeter.

Then perimeter security evolved, with the combined use of thermal cameras [16]and PTZ cameras [17]to guard the perimeter of facilities. Thermal cameras distinguish people or cars by detecting infrared rays. The thermal cameras transmit this signal to the PTZ cameras, which follow the detected target in a chain, keeping it in the field of view.

Thermal cameras find wide use in the preservation of critical sites and equipment. Since last year they are being widely used to control the spread of Covid 19, measuring the temperature of people at the entrance of airports, government and commercial buildings, schools, hospitals, etc.

In objects of special importance and increased security requirements, the industry practice is to guard the perimeter with thermal cameras, which transmit signals to PTZ cameras, while the entire territory is also monitored with radars, and the perimeter with perimeter protection systems. Both radars and perimeter defense signal the PTZ cameras to track the target.

The radars monitor the entire territory in volume and can be programmed to give an alert for persons or motorized vehicles [18].

Perimeter protection is carried out through fiber installed in the fence, which is activated by touch, or small radars located at the endpoints, which create a virtual perimeter line, the passage of which activates the alarm, or hydraulic systems installed underground, which are activated by the passage of persons or motorized vehicles over them. These systems are also integrated with PTZ cameras for active target tracking.

All of this is combined with the artificial intelligence functions that are built into fixed cameras, rolling cameras and all other types of specialized cameras. Additionally, camera surveillance systems can be integrated with fire and burglar alarm systems, access/exit control, perimeter protection, radars, GPS trackers, and any other imaginable security or automation system and function. of buildings. This integration can be done based on VMS or on neutral platforms [19], which have no brand or functionality restrictions. It is precisely the latter, as we will see below, that are opening new horizons in the field of video analytics.

## Functions of video analytics nowadays

[16] A thermal imaging camera (infrared camera or thermal imaging camera) is a device that creates an image using infrared (IR) radiation, similar to a conventional camera that creates an image using visible light.

[17] PTZ stands for Pan-Tilt-Zoom (pan-tilt-zoom) and indicates that the camera is equipped with the function of moving in pursuit of the target or by command, as well as with the function of panoramic view and focus zoom.

[18] Radars can also be combined with GPS systems, so that authorized persons and vehicles equipped with GPS trackers can be recognized by the radar as *"friends"*.

[19] PSIM Platforms – Physical Security Information Management.

Currently, all major manufacturers of cameras, video recorders and VMS offer video analytics functions, which are based on artificial intelligence and *Deep Learning* . The cameras themselves have become more capable of distinguishing certain categories of objects.

On the other hand, the ability of video surveillance systems to analyze the content of the recording has evolved beyond the simple recognition of people or cars, as a general category. Today, video surveillance systems can be designed with many artificial intelligence functions, using *Deep Learning algorithms* built into recorders or VMS. These functions can be used during *live recording* to signal via alarms, but also to check the recordings at a later time, if this is necessary.

Video analytics uses metadata, which is attribute attribution information extracted from objects of interest that can be used for data collection. Currently, three main metadata are used in the security industry: human face, human body and machine metadata. Face metadata includes sex, age, glasses, masks, expressions, beard, etc. Body metadata includes T-shirts, pants, clothing colors, hair, backpacks, etc. Car metadata includes license plate, color, make, model, etc.

The metadata recognition function can be used to determine whether objects are present in the image or video. If objects are present, their position and dimensions are recorded, characteristics and model are extracted. The model is compared with those registered in the database and the system notifies if there is a match.

Below we list the functionalities that can currently be realized through video analysis:

- Classification of categories: person, two-wheeled vehicles, other vehicles, animals.
- Classification of persons: man, woman, child.
- Classification of two-wheeled vehicles: bicycles, motorcycles.
- Classification of vehicles: car, truck, van, minibus, bus, train, plane, boat.
- Vehicle changes: lights on, lights off.
- Person attributes: bottom clothing (length, color), top clothing (sleeves, color), hats, masks, shoulder bag, backpack.
- Color recognition.
- Identification of persons and means with similar attributes.
- Face recognition.
- Recognition of car license plates.
- Direction of movement.
- Passing on a certain route.
- Movement speed.
- Proximity.
- Dimensions.
- Area.
- Time of standing still.
- Movement of objects.
- Recreating the movements of a person or machine.
- Monitoring of personnel (movement or lack of movement).
- Smoke and fire detection (thermal cameras).
- Loud sound detection.
- Fisheye Dewarping (using Fisheye cameras [20]as panoramic and PTZ cameras).
- Heatmap (people/vehicle traffic map).
- Crowd monitoring.
- Counting people.

[20] *Fisheye* lenses are wide-angle lenses that produce a strong visual distortion that creates a wide panoramic or hemispherical image. These lenses achieve very wide viewing angles.

- Checking the filling of shelves [21].

## The role of video analytics in crime prevention

When the system is designed in such a way, the above functions can be performed during *live video recordings* . This is the essential role video surveillance plays in crime prevention.

The advancement of technology is trying to expand the role of video analytics in crime prevention. The most advanced software not only performs object recognition based on metadata, but also analyzes and suggests the degree of risk of an object or action. This is accomplished through the simultaneous analysis of dimensional proportions, perspective and movements. For this purpose, certain manufacturers have performed *"custom training"* of the video analysis algorithm. These trainings mean creating specific situations related to video surveillance and exposing the algorithm to them hundreds of thousands of times, until the algorithm is able to generalize. Using these learning techniques, behavior analysis modules have been created, which distinguish potentially dangerous situations by detecting certain positions that the human body takes.

So far we described how video analytics is used to prevent crime in the context of a particular security system installed in a concrete facility. However, the same logic is used to create intelligence and help prevent crime on a larger scale. Recently developed software platforms enable the analysis of large amounts of records, obtained from many sources simultaneously. Moreover, these platforms have managed to overcome one of the main difficulties of applying video analytics, which is brand exclusivity in hardware-software compatibility. More and more brand-neutral software solutions are being implemented, that is, they allow the connection to the same network of cameras of different brands and their joint analysis, based on the same criteria [22].

These categories of software are being successfully used at the level of communities or cities or in certain industries. They enable traffic control, pedestrian behavior analysis, crowd management, etc., becoming useful aids in law enforcement, urban planning, public transportation planning, etc.

## The role of video analytics in evidence gathering

What was mentioned above about the functionalities of video surveillance systems equipped with video analytics serve to prevent crime, as a special event, and criminality, as a phenomenon that affects society in general. When it comes to analyzing recordings to discover evidence of crime, we are faced with the same question: how do we analyze hours and days of recordings from several sources to find evidence that can identify the perpetrators or be used against them in court?

As a rule, any camera surveillance system that has artificial intelligence functionalities can also offer these functionalities in *playback* , i.e. on saved recordings. Of course, in the most common cases, this requires that, at a minimum, the camera, recorder, or VMS have video analytics capabilities. The search on the saved records is done with those criteria that the system has configured for the preventive function. Thus, for example, a system that has facial recognition functionality can be used to find a suspect's face based on a pattern provided by the investigator, or a system that has license plate recognition functionality can be used to search for a assigned license plate number.

The generic software, which we talked about above, offers the most video analytics capabilities compared to the VMS of the camera manufacturers.

Video analytics on stored recordings saves investigators time, human resources and technological capacity. Recordings can be analyzed at their source and using the installed software, without having to be transferred to any other device. Thus, investigators can only use that part of the records that contains valid evidence.

[21] This list is comprehensive. Not all brands of security systems offer all of these analytics functionalities. Some of them are in cameras, some in recorders and most in VMS.

[22] See, for example , Milesight , Axxonsoft , Eocortex , etc. _

Lately, generic software solutions are evolving towards enabling video analytics even for surveillance system recordings that do not have artificial intelligence functionality themselves. Software solutions have been developed and are being successfully used in developed countries that allow video recordings from different surveillance systems to be taken and analyzed jointly, based on criteria set by the investigator [23]. The limitations in use that these solutions have do not come from the lack of technical ability to perform video analysis, but from the difficulty of integrating different brands of cameras, recorders and VMS, due to copyrights and the costs that come from purchase of these rights.

Currently, software solutions of this category have not yet been developed to include all popular models of cameras, recorders or VMS. Furthermore, since these solutions were originally developed in the USA, they leave out some major brands of security equipment, which are not allowed to be used by the US government [24]. For this reason, the effectiveness of the use of these software by law enforcement agencies outside the US is very limited.

## Conclusion

Despite all these advances and advantages of artificial intelligence in video surveillance systems, in our country we come across very few, if not any security systems that use artificial intelligence and video analytics.

The main reasons are financial constraints and lack of knowledge from end customers and installers/integrators themselves. Financial constraints are not simply a lack of funds, but an inability to analyze costs, needs and benefits in detail. In our country, this is also observed to a large extent among government clients, who do not invest in advanced technology, nor do they make long-term budget plans for this purpose.

## REFERENCES

Kardas K, Cicekli NK. SVAS: surveillance video analysis system. Expert Syst Appl. 2017.

Tzelepis C, Galanopoulos D, Mezaris V, Patras I. Learning to detect video events from zero or very few video examples. Image Vis Comput. 2016 (ISSN 0262-8856).

Guraya FF, Cheikh FA. Neural networks based visual attention model for video surveillance. Neurocomputing. 2015 (ISSN 0925-2312).

Pathak AR, Pandey M, Rautaray S. Application of deep learning for object detection. Proceedings Comput Sci. 2018 (ISSN 1877-0509).

Tsakanikas V, Dagiuklas T. Video surveillance systems-current status and future trends. Comput Electr Eng.

Wang Y, Zhang D, Liu Y, Dai B, Lee LH. Enhancing transportation systems via deep learning: a survey. Transport Res Part C Emerg Technol. 2018. (ISSN 0968-090X).

Huang H, Xu Y, Huang Y, Yang Q, Zhou Z. Pedestrian tracking by learning deep features. J Vis Commun Image Represent. (ISSN 1047-3203).

Pang S, del Coz JJ, Yu Z, Luaces O, Díez J. Deep learning to frame objects for visual target tracking. Eng Appl Artif Intell. (ISSN 0952-1976).

[23] See, for example , Briefcam HOW solution LEADER software IN this direction .

[24] Including Dahua and Hikvision that ARE two MARKS MORE People THE cAMERA IN THE all the world .

Hassan MM, Uddin MZ, Mohamed A, Almogren A. A robust human activity recognition system using smartphone sensors and deep learning. Future Gener Comput Syst. (ISSN 0167-739X).

Mammadli R, Wolf F, Jannesari A. The art of getting deep neural networks in shape. ACM Trans Archit Code Optim. 2019.

Fan Z, Song X, Xia T, Jiang R, Shibasaki R, Sakuramachi R. Online Deep Ensemble Learning for Predicting Citywide Human Mobility. Proc ACM Interact Mob Wearable Ubiquitous Technol. 2018.

Arceda VM, Fabián KF, Laura PL, Tito JR, Cáceres JG. Fast face detection in violent video scenes. Electron Notes Theor Comput Sci. 2016.

Tay L, Jebb AT, Woo SE. Video capture of human behaviors: towards a Big Data approach. Curr Opin Behav Sci. 2017 (ISSN 2352-1546).

Cermeño E, Pérez A, Sigüenza JA. Intelligent video surveillance beyond robust background modeling. Expert Syst Appl. 2018.