FEATURE

MISTRAL®
... *Partners in Real Time*

Security & Surveillance
Role of Artificial Intelligence

By **Pawan Vashisht**
General Manager, Homeland Security BU
Mistral Solutions

Installation and use of CCTV cameras for the purpose of security & surveillance is a no-brainer. Cameras are considered a fundamental commodity for setting up any surveillance infrastructure, but at the same time, 24x7 monitoring of hundreds or thousands of video feeds by operators (which is humanely impossible) doesn't serve the purpose of providing proactive surveillance and quick response to breaches.

Software based video content analytics (VCA) provides a certain level of reprieve by raising real-time alerts for a few standard breaches like tripwire, left baggage, motion detection etc., but the in-accuracy and false-positives far outweighed the potential benefits, to an extent that, most of the operators disable these analytics to avoid the innumerable false alarms

Off late, with the use of artificial intelligence (AI) and deep learning techniques, VCA software are being trained to detect, identify and distinguish various objects in video by exposing it to a large number of tagged examples. This is called 'supervised learning.' For example, to ensure that the VCA software identifies a person riding a 2-wheeler without a helmet, the software will be trained with various kinds of sample images or videos having helmet-wearing riders on the bike. This way, the system will learn to distinguish between riders wearing a helmet and the ones without it. In addition to AI-based object classification, computer vision algorithms are also being used to extract other types of data such as absolute speed and size, direction, colour, path and area. This data can then be searched to concentrate the video analytics effort on relevant information.

Artificial intelligence (AI), machine learning (ML), deep learning (DNN) and convolutional neural networks (CNN) are all technologies that attempt to imitate human behaviour.

AI is an all-encompassing term, machine learning – a type of AI is the study of computer algorithms. Deep learning is a specific discipline of machine learning and CNN is a specific type of deep learning. Deep learning is a machine learning method that uses deep neural networks and has been around for long time. In last decade, with the availability of significant amount of data and increased computational power, experts have been able to take the theoretical ideas of deep learning and put them to practical use, specifically in the domain of computer vision.

AI in video content analytics

The objective of VCA software is to analyse the video stream, one frame at a time, and create a structured database of information out of the unstructured video data. The VCA engine accepts the raw video stream, which can be a live or recorded video from a camera and then converts the same to a format which it can understand and then process the same using computer vision & deep learning technology. As part of this processing, it performs the following critical tasks:

- **Object detection –** Where are the objects in the frame?
- **Object segmentation –** What pixels belong to the object?
- **Object tracking –** Where all the object moved in the frame?
- **Object classification –** What is the broad category of the object?
- **Object recognition –** Which objects are there in the frame?

In addition to the above operations, the various object attributes like time stamp, colour, size are also extracted and saved as part of the metadata. Deep learning classification & recognition algorithms are used here to ensure higher accuracy. This metadata is then processed to perform various kinds of analytics.

Face detection, recognition & alert

Accurate face detection and recognition is very critical to law enforcement agencies. It helps in identifying people-of-interest and is also helpful in post-incident investigations. Broadly, some of the benefits of facial recognition (FR) application are:

- Automatic attendance.
- Automatic recognition of authorized individuals or re-identification of unknown people.
- Automatic alert for blacklisted/ barred people or no-go zone breach.
- Customizable MIS reports (alerts/ movements/ area-access/ area-usage).

Precise face recognition rapidly pinpoints people of interest in real-time using digital images extracted from the video, external image sources and pre-defined watchlists. The accuracy of face recognition is dependent on many factors including camera location (placement), distance of cameras, resolution (pixels), video quality, lighting, quality of the face image, angle of camera, and type of camera.

Unique features such as the distance between the eyes, width of the nose and shape of the cheekbones, which are the specific identifiers of the face, are extracted and coded into a feature vector that represents a specific face. This feature vector is stored in the database and is used to compare it to the watchlist or as a reference when faces are searched for. With the advancement of AI based deep learning algorithms, FR systems can now be trained with DNN models having a large set of sample faces. In addition, the advancements in highly efficient GPU technology has ensured that the facial recognition can be done at a large scale and in real-time.

Traffic & road safety

AI technology has enabled VCA applications to detect traffic violations accurately and automatically. Prior to AI, the accuracy of these detections was very low, and lack of computational resources had made it economically unviable to implement at large scale. Availability of large set of video data has enabled the respective DNN models to be trained effectively, there-

by increasing the accuracy manifold. Here are some of the VCA uses cases for the traffic and road safety:

- No-helmet and triple riding detection.
- Wrong way driving or illegal turn detection.
- No-parking violation detection.
- License plate detection.
- Stop-line crossing detection.
- No-seatbelt or mobile usage detection.
- Over-speeding detection.

All these analytics can be implemented on the pre-installed CCTV cameras, provided their resolution is at least 1080p. A lot of smart city projects in India are already in different stages of deployment for these AI based analytics to detect traffic violations intelligently. There are still certain challenges like the poor video quality, night-time detection and vehicles like car or bikes gone undetected since bigger vehicles like bus or truck hide them from the camera view. Hence these analytic systems can not detect 100% of the violations, but it can make sure that only the 100% genuine violations are reported.

Object tracking

During CCTV based post-incident investigations, it is very important to understand the entry/ exit routes of an object-of-interest at the site of the incident. This is very helpful in identifying the vehicles, in case of hit-and-run accidents or to identify a person, who may have left a suspicious package at the incident site. The object-of-interest could be a car, bike, truck, auto-rickshaw, mini-van, man or a woman. Each of these objects can be further defined by various attributes like colour, size, direction-of-move, speed, gender, body attire etc.

Using computer vision algorithms, once the object in a frame is detected and segmented, it can then be matched against a set of defined categories which could be a car, bike, truck, man/woman with a cap, jacket or backpack etc. The VCA software can be trained to identify these categories by using DNN models. A bigger training-set would result in better detection rate. Once the object-of-interest is detected and matched, the object segmentation defines the pixels used by the object and the movement of those pixels across the video frames can be tracked from multiple CCTV cameras , thereby giving the entry/ exit route of the object.

Even though it looks simple in theory, this detection and tracking also has its own set of challenges like distinguishing similar looking same-coloured cars or a very low number of pixels defining an object in the frame. Broadly, each investigation incident is different and may need certain custom changes in the algorithm to increase the accuracy, which may not be feasible in a commercial off-the-shelf product.

Video forensics

AI based deep learning technology can also help in solving crimes, if captured on CCTV cameras. Sometime ago, a high-profile killing during the night was captured on a camera, wherein the assailants were on a bike and investigation team wanted to know the colour, model and any distinguishing marks on the bike since the faces were covered by helmets. The grayscale IR night video feed was not of much help. To identify the colour of bike, the team replaced the existing cameras with 4K resolution cameras and captured the available background in colour and then performed the colour-to-grayscale degra-

dation of this newly captured video feed, along with grayscale to colour upgradation of the night-time video. This process gave them an idea about the colour palette of the background, which was then matched with the upgraded grayscale image to finally identify the colour of bike. Machine learning techniques were used for grayscale to colour conversion and to match the two video backgrounds, which finally helped in deducing the probable colour of the bike.

AI based machine learning algorithms can help in many other forensic activities as well. These could include:

- Vehicle model detection.
- 3D face reconstruction.
- Video enhancement by image super-resolution.
- Video de-hazing and noise reduction.
- License plate de-hazing.
- Predictive image searching.

Again, the success-percentage of these techniques really depends on the quality of video frames provided as input.

Pre-emptive policing

A combination of the above-mentioned analytics can be integrated onto a single system to enable law-enforcement agencies to take pre-emptive actions. A facial recognition system can work seamlessly in wild (uncontrolled outdoor) environment on standard CCTV cameras allowing the police a great deal of flexibility in tracking and recognizing known criminals from a pre-defined data base. In case a known criminal or fugitive is identified in a zone of interest, immediate alerts can be generated for on-field law enforcement teams thereby enabling them to take quick action.

The system can also perform crowd behaviour analysis in areas of interest. In case of any calamity or fire or accident or mishap, which generally results in crowd formation or sudden change in behaviour, alerts can be generated enabling pre-emptive police action.

Analytic data about the traffic behaviour & violations apart from past accidents can be correlated to provide analytical feedback so that necessary changes can be made in order to reduce incidents.

Conclusion

Artificial intelligence is the next evolution in video analytics. Owing to the advent of high-performance GPU hardware, deep learning based AI techniques are being widely adopted by various VCA software OEMs. This improves the detection accuracy without increasing the hardware cost exponentially. For end-users, it greatly reduces the workload of security staff and brings significant benefits by detecting unusual incidents and solving a lot of video forensic problems. Moreover, it enables them to use the massive amount of CCTV video data generated for system training purpose instead of getting overwritten over a period. In future, the quality of detection will continue to improve thereby improving adoption of AI in security and surveillance. ■