Vi DiMENSiONS

Powered by patent pending Abnormality Detection technology

WHITE PAPER – A Technology Overview

Executive Summary

Some statistics have shown that 99% of surveillance video being generated and recorded is never watched by anyone and produces no value whatsoever. This is because there is simply too much data to humanly process. This results in missed detections of abnormal events that could be happening in scene. Current methods have tried to automate this by using rule-based approaches in video analytics (e.g. send an alert when a person crosses a virtual trip wire for perimeter intrusion detection). However, these methods are not sophisticated enough to detect more complex human behaviours like rioting, fighting or even unknown threats. Most abnormal behaviours are too varied to be predictable beforehand and therefore video analytic rules are not effectively deployed to detect such events.

Vi Dimensions uses an unsupervised Machine Learning technique that is able to analyse vast amounts of surveillance video autonomously which does not require users to pre-define the rules for event detection. It is able to automatically identify patterns (i.e. motion, trajectories) in a scene and find deviations and abnormalities to warn of potential security and safety threats. It is able to do both real-time alerts and forensic search or video data mining of archived videos for abnormal event detection. The product is also able to combine with rule-based methods to deliver an even more powerful system in reducing false alarms.

The Need for Better Analytics

There is an exponential increase in the number of surveillance cameras being used for city-wide deployment but many recent events have indicated an ineffective surveillance monitoring in pre-empting or preventing such security incidents.

The hard truth is that out of all the thousands of cameras deployed in a city, only a small fraction (less than 10%) are actually looking at perimeters, entrances or exits where an analytic rule can effectively be applied. If we look at this 10% of cameras, we can see that out of the all the possible behaviours or events that can be detected, perhaps another 10% of these events can be detected by applying simple analytic rules. Putting it together means that we are only detecting 1% of all possible events, essentially missing out on the 99% of the events that could be of interest to us.

The inherent problem with rule-based analytics is that each rule specified is to detect a specific behaviour (e.g. loitering) and have to make good that claim with 90% accuracy. Much time and effort is then put into configuring this one rule and fine-tuning it to achieve the desired accuracy. If this particular event (i.e. loitering) does not occur, then the rule is basically sitting there doing nothing and produces no value. Worse, if it is inappropriately applied, it produces false alarms instead, adding to frustration of the users.

The rise in terrorist attacks shows the vulnerability associated with public places. They are no longer confined to high value targets like power plants or nuclear facilities. Attacks can now take place easily leaving little time for forewarning of danger even with the proliferation of CCTV. All these point to the fact that there is now a greater need to discover unusual and deviant behaviour and events as early as possible. Many of the security incidents are those that we are not looking out for and impossible to apply a rule beforehand. Therefore, we need better video analytics to surface more abnormal behaviours or events for scrutiny by the security staff.

The questions we need to ask ourselves are:

1) Can we come up with a better way of automating the discovery of abnormal behaviour?
2) Can we do it fast and autonomously applying it across thousands of cameras?

Why rule-based systems alone are simply not adequate

Current video analytic systems define rules in order to specify the detection of a particular behaviour or event. For example, in order to detect a “perimeter intrusion” event, a virtual trip wire has to be specified in the region of interest. The rule is set such that if this line is crossed by a human or object, the “perimeter intrusion” event is deemed to have occurred and this violation will lead to an alert. And so it is with the many other analytic rules set to detect ‘Loitering’, ‘Stopped Vehicles’, ‘Abandoned Objects’ etc. Even for people or vehicle counting, they depend on setting a line to be crossed in order for the count to be registered.

However, rule-based video analytic systems have the following shortcomings:

1) **Only for detecting simple known behaviours** - In order for a rule to be drawn, the CCTV operator needs to have some prior experience as to what kind of rules to apply for the different scenarios. This means that he/she needs to know beforehand the kind of behaviour/event they wish to detect. In many instances, this may be limited to simple behaviours that we can predict or wish to prevent against. However, what about other

unusual or unwanted abnormal behaviours and unknown threats which we might not even have thought about?

2) **Not possible to specify all rule combinations –** You may specify a rule in one location but what if the violation happens in another location in the scene? It is not possible to figure out all the possible combinations of rules or even think of all the possible behaviours which could take place.
3) **Limited coverage** - Since rules are effective to detect only simple behaviours that are known beforehand, then they are mostly limited to protection of perimeters, exits or entrances where a simple rule can be specified. In a city-wide surveillance deployment of thousands of cameras, these usually represent a small fraction of the number of cameras many of which are not looking at such areas. Some statistics have shown that less than 10% of the deployed cameras have effective video analytics.
4) **Inappropriate use leads to high false alarm rates –** Since the only approach over the past 10 years have been the use of rule-based analytics, users have no choice but to try to apply them to all kinds of scenarios in order to automate surveillance monitoring. However, this inappropriate use of rules has often led to high false alarm rates.
5) **Some scenarios are difficult to specify any rules –** Complex scenarios with large movements of crowd means that it is often not possible to specify any rules beforehand.
6) **Time consuming in rule configurations –** Many hours are needed to first setup and configure the rule and later test it for accuracy in detecting the specified event. This is highly inefficient and not scalable for thousands of cameras.
7) **A new rule/template has to be defined for each new behaviour to be detected** – Rule-based methods require the algorithms to first be trained to recognize a specific event or behaviour. This means a parametric template is often used. This means that every time a new behaviour is to be detected, a new template has to be developed just for the new behaviour and often cannot be used for others.

Discovery versus Detection

Rule-based systems start with the approach of 'detecting' specific behaviours. In doing so, they miss detecting the majority of other behaviours not known beforehand. We propose the alternative approach to first 'discover' as many deviant behavioural patterns and activities as possible. In this way, we are able to 'surface' unusual events of interest that otherwise would have gone completely unnoticed even with human operators monitoring 24x7.

After the first level of discovery and surfacing of these abnormal events, we can then parse them through several decision layers for:

1) Follow-up action
2) Application of rule-based systems for more targeted detection
3) Fine-tune the sensitivity for surfacing other Abnormal Events.

Revolutionizing City Surveillance – Doing It Smarter

The key differentiating factor in our approach is that we discover unusual and abnormal behaviour and events by looking for deviations from the normal activity patterns. By 'abnormal event', we define this as any event that is unusual, rarely occurring and identified as a deviation from the normal patterns and activities of a particular scene. Unlike a rule-based system, we are not specifically looking for any behaviour in particular. For example, we are not looking for a fight event. However, if the fight event represents a deviation from the normal activity patterns of a scene, then it will be alerted as an abnormality by our system. This approach leads us to find more events that are of interest to us.

Our product, ARVAS, is an unsupervised Machine Learning system that does not require human intervention to automatically discover dominant motion patterns. This also means it does not require a human to specify the rules for event detection. Our Abnormality Detection Algorithm is based on a unique and novel approach. We have adapted it for surveillance videos where multiple motion patterns are occurring simultaneously and it can effectively infer their various patterns and starting times. Since the system is autonomous, it provides the means to automatically analyse hours of video easily. Our proposed system consists of three main components, namely,

a) The automatic abnormality detection (AAD) engine,
b) The rules engine, and
c) The AAD and rules fusion module.

AAD is a completely automatic, unsupervised algorithm to learn frequently occurring activity patterns in the scene. The functionality of AAD is the automatic detection of abnormal activities by looking for data out of the ordinary. With the set of frequently occurring activity patterns recovered by the unsupervised algorithm, the detection of abnormal activity will correspondingly be automatic. This detection is performed by matching the observed activity against the activity patterns recovered. In contrast to event or object centric methods described in the previous section, this method works automatically without requiring any human input.

Another key distinguishing feature with respect to other solutions available in the market is the combination of the rules and AAD engine. It is to be noted that the functionality of AAD and the rules engine are mutually orthogonal. Our combination of both the AAD and rules engine complements each other, reaping the benefit and the advantages mutually.

Novelty of Our Technology

1) The type of technology used is state of the art Machine Learning and Video Analytics (Computer Vision).
2) The system can implement both rule-based and statistical approaches in a common platform harnessing strengths from both types of method.
3) Unlike current systems, it is completely autonomous and does not require users to set any rules beforehand in order to discover an abnormal event.
4) It allows intelligent video analytics to be enabled on more percentage of cameras deployed. For example, usually only 10% of cameras are using video analytics and these are usually deployed on cameras looking at entrances and exits of buildings. With our system, potentially up to 100% of the cameras can be used. This enhances the capability of Sensor (camera) Networks.

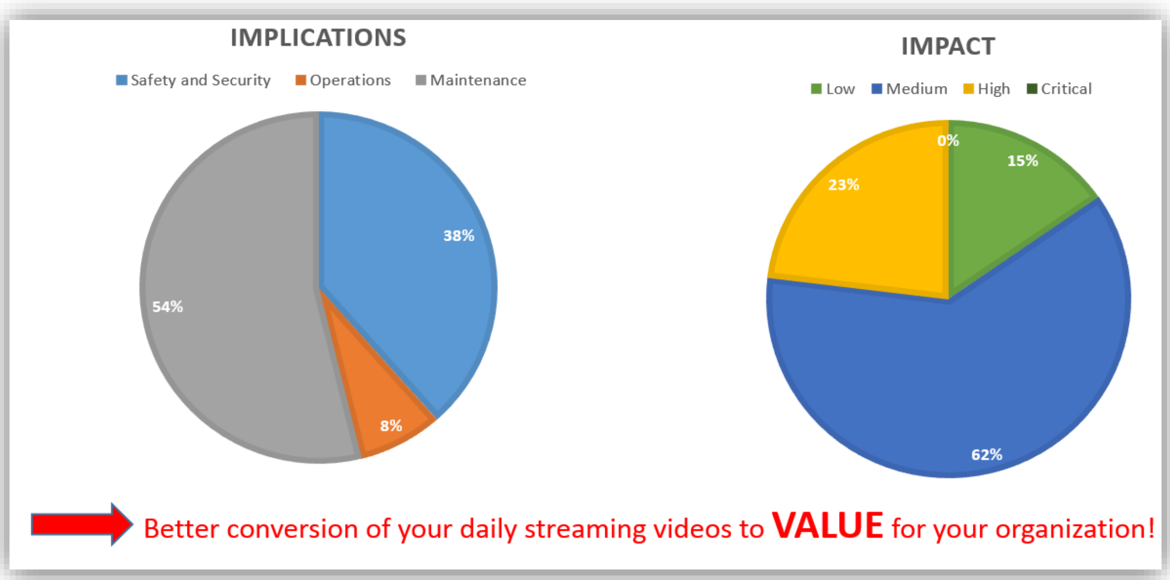
5) It enables users to find abnormal events that they are previously unaware of and not even looking for.

Key User Benefits:

- **Massive manpower savings** – Human CCTV operators are not needed to constantly eyeball and monitor the thousands of cameras for events.
- **Improved flexibility and capability** – Users need not define rules as the system can self-learn and detect abnormalities.
- **Improved efficiency** – Users can be more proactive and not reactive since the system can notify the user ahead of time to potential threats based on abnormal events and occurrences he was not actively looking for.
- **Increased surveillance coverage** - Previously only about 10% of cameras deployed can use rule-based analytics. Our technology can increase this to potentially 100% coverage.

Conversion of live streaming video to VALUE!

Beyond security and surveillance, we have shown that the discovery of unusual patterns or behaviours can lead to a better conversion of the streaming video to value for an organisation. For example, we have shown that events with safety, operational, maintenance implications can be discovered as well. Often this is directly from the discovery of unusual events or behaviours which, CCTV operators watching the constantly streaming live video, were not able to pick out and not even aware were happening.



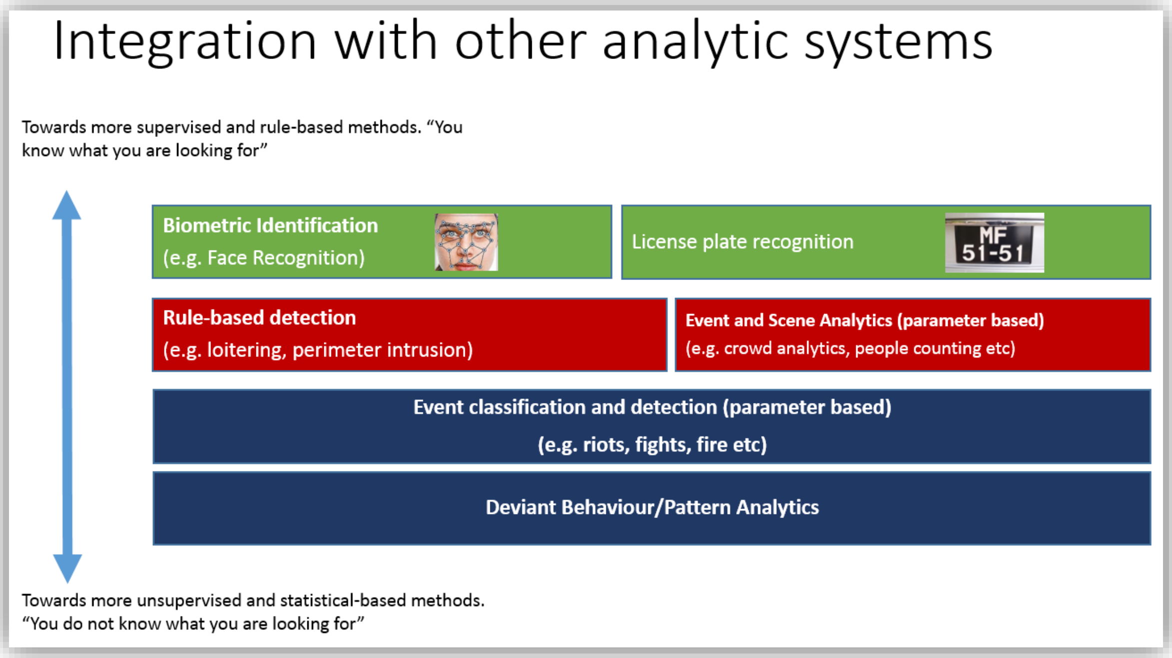
Towards An Integrated Approach

The rule-based approach however, is not redundant but has serious limitations which curb its effectiveness. We propose an integrated approach combining unsupervised Machine Learning with rules. In this way, one can help reduce the false alarm of the other. This also enhances the discovery of abnormal unknown events while the other can be used for detection of known events. Our method allows the combination in the following ways:

1) **Chaining of rule with AAD** – There can be a scenario where using a rule alone in a scenario could lead to high false alarms. Example. In this case, we can de-activate the rule first and only trigger it when an abnormality is discovered.
2) **Suppress false alarms from rules** – E.g. The passing headlights of cars causes reflection from a wet road surface is usually difficult to be differentiated by a normal rule-based system and causes many false alarms. The AAD can be used to recognize that these are normally occurring patterns in a scene and hence suppress such false alarms. The rule can then take care of the positive detection.
3) **Rule and AAD used in tandem** - AAD and rule can be used for different parts of a camera's field of view leading to better more targeted event detections.
4) **Multi-modal confirmation from both AAD and rule** – We can specify an alert only when it fulfils both requirements of AAD and rule violation.

Stacked Analytic Approach

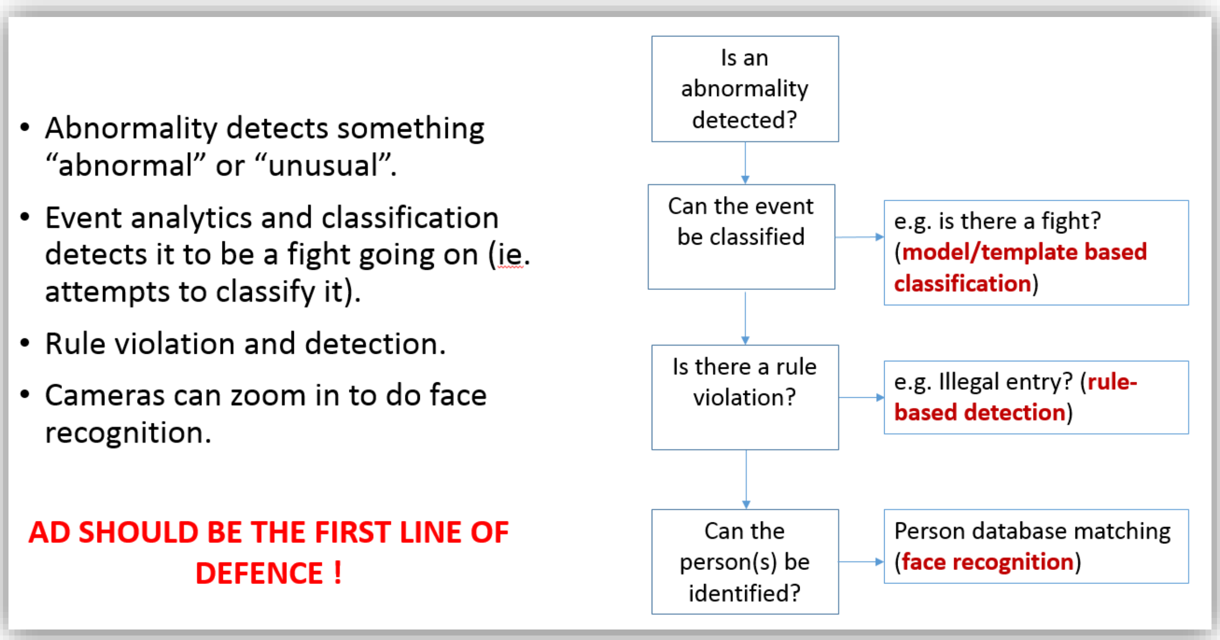
Different analytic systems can be stacked one on top of the other. If you know what you are looking for, you will probably be using an approach on the upper layers. For example, if you have a face or license plate to match against, it means you already know what you are looking for and a rule-based approach will be best suited for your needs. And so with the other layers from top to bottom. A face recognition or license plate recognition system will probably be on the top most layer. Next comes rule-based system which specify a rule or behaviour you are looking for. The bottom most layer will be our abnormality detection system which surfaces or discovers unusual activities which deviate from the normal patterns of the scene.



In this way, the user can better decide when to use which method. Simply put, if you know what you are looking for, you will use a higher level analytics (e.g. face recognition). If you do not know what you are looking for or do not have a specific behaviour in mind, you will then use AAD to filter out or surface these events.

Sequential Analytic Approach – AAD as First Line of Defence

The different types of analytics can also be used sequentially as the below diagram illustrates. First, we use AAD to see if an abnormal event has occurred. Then we check if this event can be classified. Next, we ask if this there is a rule violation. In some scenarios, an unusual event may have occurred but if there is no rule violation, we do not want an alert. Finally, we can check if the person violating the rule can be identified if we have a face recognition system.



Thus, we see that AAD is used as the first ‘skin’ layer or first line of defence for surfacing any unusual activity to be cross-checked with other analytic systems for classification or identification.

The Vision for Smarter City Surveillance

Abnormal events can be contextual. For example an abnormal event occurring in a cluster of cameras could signify an event that spreads over a geographical location which leads to a higher level of alert. As in the case of the recent terrorist attacks, it can be seen that people exiting buildings in different manners (e.g. rushing out of doors, climbing out of windows, descending on poles etc) signify many forms of the same crowd dispersal and escape event. To AAD, this is commutative and unlike rule-based approach where it has to first recognize these different type of escape behaviours, we can simply show an increase in abnormal behaviours coming from multiple cameras and send an amplified alert to the operators.

Imagine a command centre with thousands of cameras streaming video, finding abnormalities in just 10% of these cameras can already convert to better value than using existing rule-based approaches where most of the time, there are no such positive detections. Even finding a significant abnormal event in one camera of high security threat will be worth the while and in order to do this, we need 100% of deployed cameras to have AAD.

Purpose Built for Big Video Data Analytics

Many of the current rule-based system are not deployable on the cloud. Precisely because there is no need for such a massive scale video analytics and because of the limited deployment for rule-based analytics. However, if we imagine that for every camera, it is possible to find abnormalities, then we will need to harness the compute resources of the cloud to enable such large scale analytics. Finding abnormalities in this case is more akin to Big Data Analytics than Video Analytics because of the similarity in scale.

Large scale analytics require it to be engineered from the start to be scalable to the cloud. This requires modularity in coding and ensuring its agnostic nature on different operating system and platforms. Our AAD system is highly customizable and built from the start with large scalability in mind.

About Vi Dimensions

Vi Dimensions was founded in 2015 with the simple idea that video analytics can be done in a much better and efficient way with the ultimate goal to revolutionize safe city surveillance harnessing thousands of cameras.

The company uses its patented algorithms and proprietary unsupervised Machine Learning techniques to derive meaningful information and actionable insights from live streaming video data. This translates to immediate value to the customer not only in terms of security and surveillance but also improves the organisation's safety, operational and maintenance aspects.

Our advanced and innovative system analyses vast amounts of real-time streaming (or archived) data autonomously for abnormal behavior and events. It does not require human intervention to automatically discover dominant motion patterns which means that unlike conventional systems, it does not require a human to specify rules necessary for detection.

Contact

Vi Dimensions Pte Ltd
16 Ayer Rajah Crescent
Tempco Technominium
#05-04
Singapore 139965
Tel: +65 65702331
www.vidimensions.com
enquiries@vidimensions.com