FROM MOTION TO UNDERSTANDING: HOW DEEP LEARNING ENHANCES AUTOMATIC INCIDENT DETECTION

Emmanuel MARCHAND, Maciej SEKULA, Christophe COLOMBEL

Citilog, FR

DOI 10.3217/978-3-99161-087-8-038 (CC BY-SA 4.0)

ABSTRACT

Automatic Incident Detection (AID) through video analytics is a cornerstone of modern traffic management. Historically, AID systems relied on pixel-based algorithms, detecting incidents by analysing changes in pixel intensity and movement over time. While foundational, these traditional methods often struggle with environmental noise and static scenes, leading to missed incidents and high false alarm frequencies.

This presentation by Citilog explores the paradigm shift brought by Deep Learning (DL) in AID. Unlike pixel-based methods that track abstract motion, DL models are trained for true object recognition and scene understanding. We will demonstrate how this transition fundamentally improves detection quality – simultaneously increasing true detection rates and drastically decreasing false alarms – through three key examples:

- **Stopped Vehicle Detection:** Traditional pixel-based systems frequently trigger false alarms from moving shadows or sudden lighting changes. Deep Learning accurately distinguishes between a cast shadow and an actual stationary vehicle, ensuring high reliability.

- **Pedestrian Detection:** Legacy analytics typically require tracking a small group of pixels over a long trajectory to confidently identify a pedestrian. In contrast, Deep Learning can recognize a human form instantly within a single frame, enabling much faster and more accurate detection.

- **Smoke Detection:** Pixel-based methods are largely limited to detecting a general loss of visibility by comparing the current scene against historical reference images. This requires time to establish a baseline and often cannot confirm if the visibility drop is actually smoke. Deep Learning, however, can instantly recognize the unique visual structure and patterns of smoke within a single frame, resulting in significantly faster and more precise alerts.

Deep Learning participates to define a new standard for AID accuracy, paving the way for safer, more efficient, and highly responsive road networks.

1. INTRODUCTION – THE EVOLUTION OF VIDEO ANALYTICS

For decades, Automatic Incident Detection (AID) has served as the critical "eyes" of Traffic Control Centres. However, those eyes have traditionally been limited by their reliance on **pixel-based background subtraction**. This method functions by establishing a "background" image and identifying moving objects as "foreground" based on changes in pixel intensity and contrast. While revolutionary at its inception, pixel-based analytics are fundamentally reactive: they detect *change*, but they do not *understand* the context of that change.

The inherent weakness of pixel-based systems lies in their extreme sensitivity to environmental noise. In complex infrastructures like tunnels or high-traffic corridors, variables such as shifting shadows, headlight reflections, and swaying equipment all trigger pixel changes that the system misinterprets as incidents. This leads to "false alarm fatigue," where the sheer volume of nuisance alerts can cause operators to miss real emergencies.

The Deep Learning Shift

Citilog's transition to Deep Learning (DL) represents a move from simple motion detection to **Advanced Computer Vision**. Instead of calculating pixel differentials, DL models (Convolutional Neural Networks) are trained on thousands of annotated images to recognize the "features" of objects—the specific geometry of a vehicle, the silhouette of a pedestrian, or the unique texture of smoke. This paper explores how this shift from tracking movement to identifying objects has delivered a significant improvement in detection quality, simultaneously increasing safety and operational efficiency across three critical use cases.

2. CASE STUDY 1 – STOPPED VEHICLE DETECTION

Stopped Vehicle Detection is the most frequent requirement for AID, yet it is where pixel-based systems are most prone to error due to the dynamic lighting environments of modern tunnels.

The Challenge: Luminance Adaptation and "Ghost" Incidents

Tunnels are quite stable environments; but the interior lighting is constantly adapted to match outdoor ambient light levels to ensure driver safety. As the sun moves, banks of lights are turned on, off, or dimmed in stages. To a pixel-based system, these **intensity variations** or **outdoor light entering the tunnel** are indistinguishable from physical movement. When a light bank dims, the algorithm interprets the sudden darkening as a large, stationary object. Furthermore, these lighting shifts create "new" shadows from infrastructure. Because the pixel-based logic only sees a change against a stored reference, these daily luminance adjustments may trigger "ghost" alarms, forcing operators to manually verify many non-events.

The Deep Learning Advantage

By utilizing Deep Learning, new algorithms perform **True Object Classification**. The system identifies a vehicle based on its structural features—wheels, windshields, and body contours—rather than just detecting a change in light and dark pixel clusters.

- **Feature-Based Recognition:** A shadow created by a lighting transition lacks a chassis or a roofline; the DL model ignores it as environmental noise.
- **Operational Efficiency:** This ensures that operators only receive alerts for actual physical hazards, neutralizing the impact of daily lighting cycles.

Figure 1: Stopped vehicle in a tunnel, Source: *Citilog CT-IM-AID_Functions_specifications-RevL.pdf*

Stopped vehicle	**Detection rate**	**False alarm (FA) frequency**
Pixel-based analytics	Indoor > 97 %; Outdoor > 95 %	Indoor: < 0,2 FA/camera/day Outdoor: < **1 FA/camera/day**
Deep Learning	Indoor > 99 %; Outdoor > 97 %	Indoor: < 0.02–0.1 FA/cam/day; Outdoor: < **0.1–0.25 FA/cam/day**

Sources: audits done for a tunnel for Service Public de Wallonie, Citilog CT-IM-AID_Functions_specifications-RevL.pdf

3. CASE STUDY 2 – PEDESTRIAN DETECTION

Detecting a pedestrian on a highway or in a tunnel is a high-stakes task where speed of detection is as critical as accuracy.

The Limitation of Trajectories

Pixel-based analytics struggle with pedestrians because a human being represents a very small number of pixels compared to a vehicle. To avoid false alarms from birds or debris, traditional systems usually require a **"validation trajectory."** The system must track the movement of those pixels over a significant distance and time (often several seconds) to confirm the "blob" moves like a human. In a fast-moving traffic environment, these lost seconds are a safety risk.

The Deep Learning Advantage

Deep Learning enables **Instantaneous Detection**. Because the model has been trained on the human form (the "skeleton" and posture), it can identify a pedestrian from a single frame.

- **Speed:** Detection occurs the moment a person enters the field of view, even if they are standing still or only partially visible behind a barrier.
- **Reliability:** By recognizing the object itself rather than its path, the system is far less likely to be fooled by non-human movement.

Figure 2: Pedestrian detected when entering the FoV, Source: www.citilog.com

Pedestrian detection	Detection rate	False alarm (FA) frequency
Pixel-based analytics	97% Only on **sidewalks**	< 0,1 FA/camera/day
Deep Learning	99% On **any traffic lane or sidewalk**	< 0,05 FA/camera/day

Sources: Citilog CT-IM-AID_Functions_specifications-RevL.pdf, indoor

4. CASE STUDY 3 – SMOKE DETECTION

Smoke detection is perhaps the most technically demanding AID task due to the semi-transparent, non-rigid, and shifting nature of the "object."

The Problem with View Obstruction

Traditional smoke detection relies on **Visibility Analysis**, comparing the live feed against a "historical reference image." This method assumes that any significant "loss of detail" or "graying" of the image is smoke. A major flaw occurs when a **large vehicle, such as a truck, stops directly in front of the camera**. The truck's flat side "hides" the background details of the tunnel. The pixel-based system sees that the reference (the tunnel wall or road) has disappeared and incorrectly triggers a "Loss of Visibility" or "Smoke" alarm, despite there being no fire.

The Deep Learning Advantage

Deep Learning treats smoke as a **distinct visual texture and pattern** rather than a mere lack of visibility. It looks for the characteristic turbulent flow and opacity gradients unique to smoke.

- **Contextual Intelligence:** When a large truck stops and covers the Field of View, the DL model identifies the object as a "Truck." It understands that the background is obstructed by a solid object, not obscured by an airborne particulate.
- **No Reference Required:** Because the model recognizes smoke "on sight," it does not need a historical baseline. It can detect smoke the second it appears, even if the camera has just been turned on.

- **Superior Accuracy:** By differentiating between a solid obstruction (a truck) and a fluid one (smoke), Citilog provides a much higher level of confidence for fire response teams

Figure 3: Smoke in a tunnel, Source: www.citilog.com

Smoke detection	Detection rate	False alarm (FA) frequency
Pixel-based analytics	95% Any type of loss of visibility, not specific	< 0,1 FA/camera/day
Deep Learning	95% Specific for **smoke**	< 0,05 FA/camera/day

Sources: Citilog CT-IM-AID_Functions_specifications-RevL.pdf, indoor

5. CONCLUSION – A NEW STANDARD FOR ROAD SAFETY

The transition from pixel-based analytics to Deep Learning is a fundamental re-engineering of how we monitor road safety. By moving from **Motion to Understanding**, using Deep Learning has addressed the two greatest pain points of traffic control centre operators: the "missed event" and the "nuisance alarm."

Summary of Improvements:

- **Reliable stopped vehicle detection:** The system is no longer fooled by necessary daily **luminance adaptations** or "ghost" shadows created by tunnel lighting cycles.
- **Instantaneous Pedestrian Alerts:** Detection is based on **visual form** rather than movement trajectories, allowing for immediate intervention.
- **Intelligent Smoke Detection:** The system distinguishes between **physical obstructions** (like a stopped truck) and actual smoke, ensuring fire alarms are only triggered by true threats.

As Deep Learning models continue to evolve and edge computing power increases, the gap between traditional analytics and DL-based systems will only widen. This technology represents a significant step forward, ensuring that traffic managers can move away from managing software filters and back to their primary mission: responding to incidents and saving lives in a complex, ever-changing environment.