intelexvision

iSentry Camera Capabilities & Best Practice Manual

Version: 2025R2.12.09.2025

Copyright, trademarks, and disclaimer

Trademarks

iSentry is a trademark or registered trademarks of IntelexVision. All other trademarks mentioned in this guide are the property of their respective holders. This product may make use of third-party software for which specific terms and conditions may apply.

All names of people andorganizations usedinthe examples inthis textare fictitious. Any resemblance to any actual organization or person, living or dead, is purely coincidental and unintended.

Disclaimer

This document is intended for general Information purposes only and due care has been taken in its preparation. Any risk arising from the use of this information rests with the recipient, and nothing herein should be construed as constituting any kind of warranty.

IntelexVision may make changes without prior notice.

This product may make use of third-party software for which specific terms and conditions may apply.

Version control

Version:	Date:	Author:	Content:
1.0	23/12/2023	All	ALL
1.0.1	15/01/2024	W.Peters	
1.0.2	22/01/2024	W.Peters, T.Cao, V.Alonso	Revision
1.1	01/08/2025	W.Peters, T.Cao, M.Zapotoczny	Adding Aurora, ANPR, FR
1.2	01/09/2025	W.Peters	Aurora Revision
1.3	12/09/2025	M.Zapotoczny	Weapon, Gun Pose Detection

Introduction

iSentry is designed to analyse thousands of CCTV cameras in real-time and give that information to controllers immediately in the form of video alerts that are enriched through multiple data points to block out the noise of live video, from there iSentry can add instant visual insight to the huge amounts of footage generated by a large-scale CCTV deployment so that operators can focus on what matters.

This provides the ability to identify potential threats to security, operational challenges and opportunities, health and safety violations and regulatory compliance and can provide actionable insights, all in one place.

Using the latest Artificial Intelligence (AI) and Machine Learning techniques, iSentry shows operators what matters. We allow operators to focus on interesting events and respond in real-time to both precursor activities and incidents using our state-of-the-art analytic process flow, while excluding inconsequential events with our Deep Learning and rules engines.

Foreword

This Document is intended for general information purposes only, all due care has been taken in its preparation, any risk arising from the use of this information rests with the recipient, and nothing herein should be construed as constituting any kind of warranty.

All Names of People and organizations used in this document's examples are fictitious. Any resemblance to any actual organization or person, living or deceased, is purely coincidental and unintended.

All logos and software mentioned in this document are copywrites of their respective owners. This product may make use of third-party software for which specific terms and conditions apply.

Key topics covered include:

- Recommended camera specifications for face recognition
- Camera placement, mounting, and lighting advice
- The importance of video clarity and how to achieve it
- Example scenarios of good and bad camera setups
- Common pitfalls and how to avoid them

Scope

This document will provide insight into the world of camera configuration and best practice for the iSentry system.

System Access

You should have:

- Access to the premises and relevant mounting locations
- Necessary permissions and tools for camera installation and adjustment

Audience

This guide is intended for:

- System Integrators: Responsible for designing and installing camera infrastructure.
- Security Technicians: Tasked with the physical setup, adjustment, and maintenance of cameras.
- IT/Facility Managers: Overseeing system deployments and ongoing performance.

Abbreviations, Acronyms and Definitions

UB	Unusual Behaviour
LO	Left Object
OM	Object Motion
TREX	Threat Detection and Extraction
DL	Deep Learning

Pre-requisites

Technical Knowledge

- Basic Windows Proficiency. Ability to navigate Windows applications, log into systems, and manage basic settings.
- Data Entry Skills. Proficiency in entering data into various Windows-based software applications.
- An advanced understanding of video surveillance systems and camera installation
- An advanced understanding of key camera parameters (resolution, compression, GOP, shutter speed, WDR aperture, etc.)
- Basic skills in using installation tools and safety equipment
- Basic skills in using photo editing software

System Access

- Administrator Privileges. Access to a user account with the right privileges to be able to perform the intended activities.

- Network Permissions. Ensure necessary permissions are granted and configure network settings, including firewall exceptions.

- Access to the premises and relevant mounting locations
- Necessary permissions and tools for camera installation and adjustment

Preparatory Tasks

Familiarize yourself with this guide and any additional documentation.

What you should do:

- Carefully review this guide and your project requirements before beginning
- Confirm all required hardware and accessories are on site
- Plan for optimal lighting and infrastructure conditions at the installation site

Video Streams

General

Generally, iSentry can process a wide variety of video stream configurations, but for the best balance between performance and quality the following guidelines are provided:

- The single most important factor for iSentry video analytic performance is image quality, which is strongly related to camera's sensor quality. The better the quality of the image the better the iSentry analytic will perform, this is far more important than many other camera's specifications such as resolution or frame rate.

- Encoding: h.264 or h.265 is highly recommended, but other codecs are also supported.

- Resolution: 640 x 480 or 1280 x 720 are commonly used and recommended and will produce excellent results with reasonable resource utilization.

  1980 x 1080 will have a small analytic advantage, but at the cost of higher resource utilization. Many other resolutions are also supported.

- Framerates: Any framerate over 5 FPS is supported, but +- 10 - 12 FPS is recommended for general purposes. Here is a particular exception: if fast moving objects must be accurately tracked, higher frame rates may be required. In general, it is not recommended to exceed 25FPS.

- Keyframes Interval for h264/h265 format: Any value under 30 frames between keyframes is supported, with lower values being generally better (generally lower values results in higher bandwidth requirement). It is important to note that leaving this setting to manufacturer default often results in extremely high values (150+) in an attempt to minimize bandwidth utilization. High frame gap between keyframe values may lead to a deterioration in tracking and video quality and will often result in "jumps" in the video, causing false alerts. Therefore, it is very important to set this value correctly when any camera is added to the analytic.

- Bitrates: Both Variable and Constant bitrate encoding is supported, but constant bitrate is preferred to avoid unwanted video degradation due to camera firmware decision. Maximum bitrate values depend on stream settings but as a guideline the following can be offered:
  - 640 x 480 @10FPS , Maximum bitrate: ~1 Mbps
  - 1280 x 720 @ 10FPS , Maximum bitrate: ~2 Mbps
  - 1920 x 1080 @ 10FPS , Maximum bitrate: ~4 Mbps

NOTE: Avoid proprietary manufacturer compression techniques, or any non-standard bandwidth reducing mechanisms. iSentry must process the original, unprocessed camera stream to be able to function optimally/correctly.

Networking requirement

Inbound:

Bandwidth estimation can be done using the formulae below:

- Maximum bitrate X Number of cameras = Total bandwidth throughput required

NOTE: Actual bandwidth will vary.

Example:
50 Cameras @ 1080p with Maximum bitrate of 4Mbps:

- 4Mbps x 50 = 200Mbps

NOTE: This estimation assumes that NO other streams are enabled on the camera.

Outbound:

Outbound traffic will be a function primarily of alert volume and system architecture. There are a multitude of architecture options such as, centralized, centralized with distributed processing, distributed with centralized Deep Learning, etc. For an accurate assessment of outbound networking please consult an Intelex Vision sales or support consultant.

As a worst-case scenario from a video processing server point of view, 6MB per generated alert can be used to calculate outbound traffic, using peak alert volume as the yard stick.

For example, if we assume peak alert volume to be 6 alerts per minute / 1 alert per 10 seconds / 0.1 alert per second. This would equate to 0.1 * 6 = 0.6MBps or ~5Mbps as the bare minimum required to manage this peak volume number. Depending on the exact architecture being deployed, a small amount of extra bandwidth may be required for housekeeping and background tasks.

NOTE: It is considered best practice to not exceed 50% of the stated maximum throughput of the network card on a server/PC. For example, exceeding 500Mbps for a connection through a 1Gbps NIC is not recommended, and in this case a faster NIC is recommended.

Camera Considerations

Assumptions

All the data in this document (except where specifically stated otherwise) assumes the following camera and stream configuration:

- Resolution: 1280 x 720
- Encoding: h.264
- Bitrate control: Constant bitrate
- Maximum Bitrate: 2048Kbps
- Framerate: 12FPS
- Profile: High Profile
- Frames between keyframes: 12
- Frames between keyframes mode: Custom

Quality and clarity

Video quality is of paramount importance to iSentry. Video quality is hard to define completely in technical terms but in general a video is said to have a good quality when it has good contrast, targets are of sufficient size and is free of the following:

- compression noise
- sensor noise
- optical distortion
- motion blur

General

As a rule of thumb, using lower resolutions like 640 x 480 video with a high-quality sensor will yield better detection results than high resolution (e.g 4K) video generated with low quality sensor.

High resolution is less important than sensor quality, and often low-cost high-resolution cameras will include a lot of visual “noise” in the stream, reducing detection quality, while greatly increasing processing requirement: a resolution of 1280 x 720 is recommended as an excellent middle ground for general purpose use, if the sensor quality is good. For shorter distances a resolution of 640 x 480(360) with a good quality camera is generally excellent and will yield excellent results balanced with valuable resource efficiency.

It should also be noted that Thermal cameras are typically low resolution, and multi kilometer detection ranges are common. Many thermal camera manufacturers use software upscaling internally to make the output stream look bigger than what's been natively captured. Where possible, it is recommended to use the native camera resolution for processing.

Note: Stream configuration is also key to quality.

Detection Distances (Typical):

Distances based on ideal mounting positions, angles, and lighting.

NOTE: Typical lens focal length vs typical distance range for a visible light camera with a good quality sensor at a resolution of 1280 x 720.

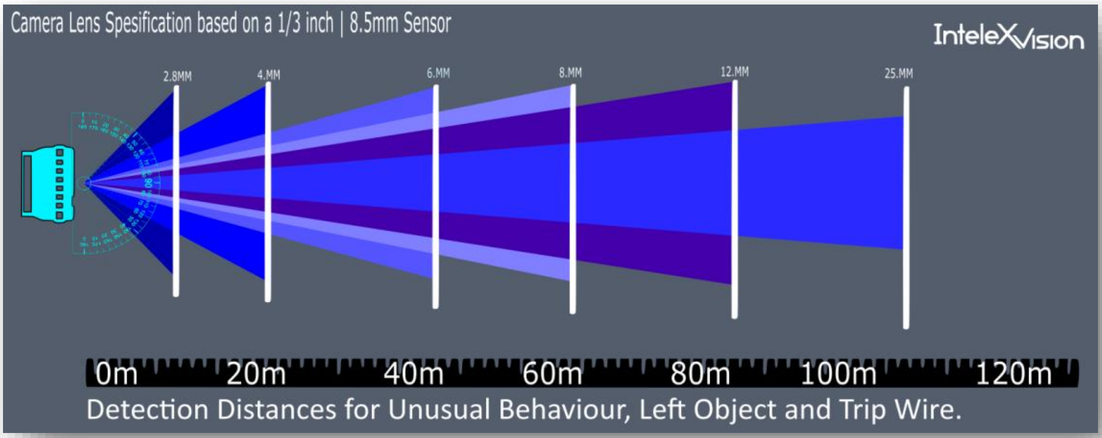
These guidelines are subject to camera and lens specifications and configuration and is meant as a rough guideline only!

Unusual Behaviour detection (UB), as well as all of the core analytics TREX Lite (Object Motion (OM) (Short distance), Left Object detection (LOD)):

- 2.8mm Cameras = 2 - 20m
- 6mm Cameras = 10 - 35m
- 12mm Cameras = 20 - 50m

Video Tripwire (VTW):

- 2.8 - 4mm Cameras = 2 - 20m
- 6mm Cameras = 10 – 35m
- 12mm Cameras = 20 – 50m



TREX Standard:

- 6mm Cameras – 40m – 150m
- 12mm Cameras – 80 – 200m
- 35mm Cameras – 100 – 300+m

TREX Premium and Elite:

- 6mm Cameras – 80m – 150m
- 12mm Cameras – 100 – 250m
- 35mm Cameras – 200 – 700+m
- 105mm Cameras – 250 – 1500+m

Camera Positioning (Angle and Height)

General Positioning guidelines:

Most Cameras operate well at 2.5 meters or higher above the ground, it is however recommended to place commercial cameras 3 meter or higher above ground for tampering reasons. In the case of Facial Extraction or recognition, this height is generally reduced to around 1.6 – 2m.

Please note that camera mounting height will impact detected object size.

It is important to note that low mounted and low angled cameras (less than 35 Degrees below horizontal), will be more vulnerable to the occlusion problem, where one object occludes another in the scene when multiple tracked objects are present. This problem is reduced by higher mounting positions and angles.

Deep learning:

The Deep learning system relies on various neural networks which are trained at "typical" CCTV angles, which in general are between 20 and 45 Degrees below horizontal. This implies that at extreme angles (e.g. top-down or above horizontal) the ability of the Deep Learning system to function accurately will be impaired.

Typically, at angles between 0 and 50 Degrees below horizontal, the performance of the neural networks is still very high, but at angles over 50 degrees below horizontal, the performance will start to be impacted negatively.

For cases where high angles are required (between 65 and 75 degrees below horizontal), a custom "Top-down network" has been developed specifically for person detection. This allows for the use of nearly vertically mounted cameras, or very high mounted cameras for accurate classification of people. Classification of other objects, such as trucks or buses, is possible but lower accuracy due to the top-down viewpoint.

iSentry Analytics

All the iSentry system analytics will function at any angle, but optimal angles would differ from use-case to use-case.All the iSentry system analytics will function at any angle, but optimal angles would differ from use-case to use-case.

Facial Extraction, and Facial recognition (SAFR)

Typically for facial extraction and/or recognition, cameras should be mounted at a low height, and near horizontally, with excellent illumination to be able to capture the face of subjects as clearly as possible. These cameras would also typically be used for very short distance monitoring.

General Short to Medium Distance Monitoring

In general, tilt angles around 35 Degrees below horizontal for short distance detections are recommended, but depending on detection distance required, as well as mounting height, this can vary.

General Entry and Exit Counting (Non-DL)

For general object counting, a wide-angle camera can be mounted in a top-down position (vertically), directly above a door, or entrance/exit hallway etc.
This type of positioning helps avoid occlusion and improves accuracy.
Since Deep Learning is not used, any object (within size restrictions) will be counted.
With a good position, where the entrance/exit zone is clearly defined and visible, with good lighting, while avoiding visual interference such as shadows and light wash etc. a very high degree of accuracy can be achieved. This counting mechanism is also extremely efficient, as it requires no deep learning processing.

Deep Learning Assisted Entry and Exit Counting

Deep Learning assisted counting mitigates for less-than-ideal camera positioning, by combining object tracking with classification. Here only classified objects are counted, and data is stored in the database for each object counted. With this method of counting, sub-optimal mounting positions can be mitigated to some extent, but the system will be sensitive to angle, lighting, occlusions and any other visual interference since the process relies on both tracking and classification.

Medium to Long Distance Monitoring

Generally medium to long distance detections are needed in the case of wide area surveillance and intrusion detection use-cases. Camera angles are typically less than 35 Degrees below horizontal, and mounting height will typically be higher, for longer distances.
With very high mounted cameras, care should be taken with regards to camera stability. TREX is extremely “tolerant” of slight camera movement, but severe “whipping” or movement due to wind or poor mounting would impair operation.

Analytic Specific Best Practices

It should be reiterated that object sizes and distances quoted in this document assume high quality video. As stated before, the quality of the video analysed is paramount to the quality of the outcome of this analysis, which is why stating pixel density is largely avoided since high resolution cameras may have high pixel density but this is no indication of image quality. For this reason, object sizes are used as a percentage of the scene size, as an indication of distance, rather than using a metric such as pixel density.

Unusual Behaviour

Unusual Behaviour detection is based on pixel level processing of objects in a CCTV camera scene. The manner in which objects behave in a scene will determine expectations for future behaviour, and when expectations are not met, alerts are generated. For optimal functioning of this system, objects must be clearly visible and easily distinguished from the background and each other, and occlusion should be avoided if possible. Ideally objects of interest should be no smaller than 15% of the scene size and no larger than 70% of the scene size. This typically translates to up to 40 metres or more depending on camera specification. Extremely short distance, long distance, quiet or very densely crowded scenes are examples of scenes not well suited for Unusual Behaviour analysis. Despite the above best practices, Unusual Behaviour detection is able to function in any scene provided with greater or lesser effectiveness.

Short distance object tracking

Advanced left/removed object detection is based on pixel level analysis of the scene. This analytic requires stable feed of some minutes (> 5 min) to learn the scene and the objects within it, after which it will alert on any objects removed or placed in this scene according to configuration. This analytic is extremely effective in identifying all manner of objects, even doors left open but effectiveness will depend on video quality much more so than pixel density generally objects larger than 3% and smaller than 80% of the scene size can be effectively detected when placed or removed. Although this analytic is typically used for shorter distances it can be very effective at distances up to or exceeding 100 metres depending on camera specification. Currently left or removed objects are treated as the same, no current method to distinguish the two types of events apart.

Video Tripwire

The the iSentry video tripwire analytic is designed primarily for short distance tracking and is part of the iSentry core analytics package. It is designed to alert on any object crossing it's virtual line, in short to medium distance applications. Typically, this analytic would be used for short distance intrusion detection as well as general surveillance applications. Assuming good quality video objects as small as 5% of the scene size can easily be tracked, depending on illumination and if enough contrast exists in the video to clearly distinguish the object from the background. This analytic would typically be used for applications up to 60 metres or more depending on camera specification.

Defence

The iSentry Defence analytic is a specialised derivative of the TREX analytic. Defence is particularly designed for fast moving objects crossing over fence lines. A practical example of an application for defence would be an environment where objects may be thrown over the fence surrounding the facility. Defence is designed to avoid false positives while being able to track any object (even small fast-moving objects) crossing a virtual fence. This analytic was not designed for long distance but is very effective for short and medium distance. With good quality video with a high frame rate objects as small as 4% of the image size could be tracked crossing a virtual fence. Depending on the camera used this may translate to anything up to 40 metres or more detection distance.

TREX

Several subcategories exist under the umbrella of the TREX analytic, namely TREX Standard, TREX Premium and TREX Elite.

- TREX Standard: This is the typical analytic used for medium distance intrusion and wide area surveillance applications. Correctly configured with the appropriate quality video, moving objects as small as 3% of the scene size can be successfully tracked. Depending on camera specification this analytic is typically used at distances up to 300 metres or more.
- TREX Premium: This analytic is typically used for long distance intrusion detection and wide area surveillance applications. Using appropriate cameras, up to multi kilometre detections are possible. When using good quality thermal cameras for example, moving objects 1% of the scene size (or even less) are successfully tracked. 640 x 480 assuming ideal conditions and mounting.

- TREX Elite: This analytic is very similar to TREX Premium in its detection capabilities. What distinguishes TREX Elite from TREX Premium is the fact that it is designed to monitor moving PTZ cameras. This is achieved by monitoring for scene changes (Camera patrol movement) and avoiding alerting when this happens. After movement has been detected the system will allow for stabilisation of the camera and then start monitoring the scene once more.

As with all analytics, following best practices for mounting positions, distance and size will increase system performance, and therefore practical utility.

Deep Learning

Assumptions

Resolution:

1280 x 720 resolution is assumed.

Deep learning Configuration:

Minimum threshold: 25%

Mode: Accuracy

Camera Sensor and Lens Considerations

For Deep Learning, the camera used must have a good quality sensor and lens. The relative size of the target objects to the overall view of the entire frame or scene is the most important factor affecting the performance of iSentry DL modules.

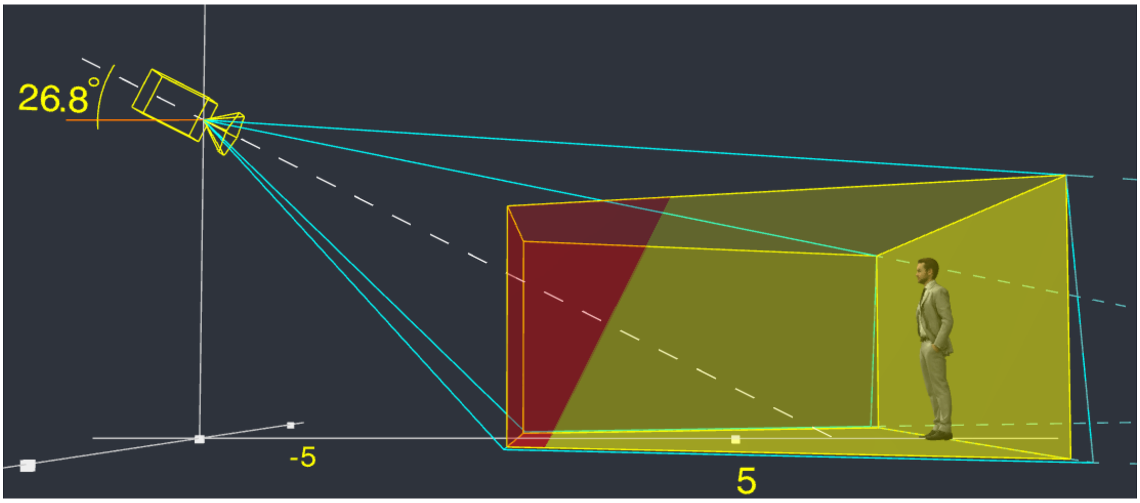
iSentry Detection Modules

The iSentry system employs deep learning to enrich the data for pre-existing alerts triggered by its real time core analytics. Each of the many neural networks employed will function optimally under specific circumstances based on the use case each one was designed to meet. The following are some of the neural networks available along with best practise information:

Action Recognition

The action recognition module is designed to detect specific actions carried out by people (such as running, falling, sitting down, lying down). The camera mounting position, lighting and video quality are all critically important factors in determining the quality of outcome of the detections of this module. Typically, this module would be used on cameras mounted at around a 35-degree angle, ideally where subjects pass through the scene from side to side as opposed to towards or away from the camera. The people tracked by this module should ideally be no larger than 70% of the scene size and no less than 40% of the scene size, as it is designed for close distance detections.

This Module requires one of the General Object Detection modules to be enabled. Ideally mounting height should be 1.8 to 4m and mounting angle should be from 20 to 45 Degrees below horizontal. This model may be incompatible with analytics that required more 'top-down' view.



Fighting Detection

What the analytics actually "see"

Fight detection is action recognition: models look for group/individual aggressive behaviours over short video chunks - e.g., fights, assault, abuse, vandalism. They do not "predict", only classify what's visible.

AI is probabilistic: precision/recall trade-offs are inherent; design must manage occlusion, motion blur and lighting to keep signals clean.

False Positives & Contextual Considerations

- Similar actions: Some behaviours (dancing, hugging, handshakes, celebrations, crowd surges) may resemble fighting and be flagged.

Figure 1. The following example illustrates a false positive. The act of embracing and expressing affection in a demonstrative fashion was categorised as a form of conflict.

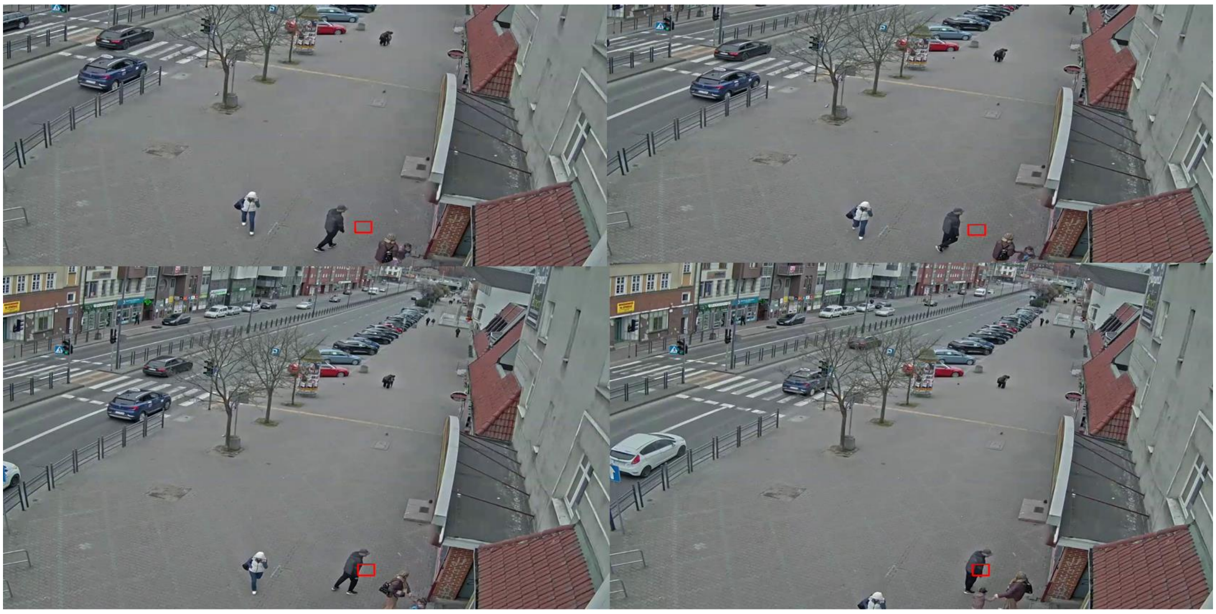


Figure 2. . The following example illustrates a false positive. The act of playing with a child was categorised as a form of conflict

- Vandalism: Aggressive vandalism (kicking doors, smashing windows, striking objects) may be detected as fighting if the motion pattern resembles assault. Treat such alerts as high priority, but verify context before escalation.
- Occlusion: In dense or obstructed scenes, single-camera analytics may confuse crowding with pushing/shoving. Use multiple camera angles for confirmation.

Camera Position & Angle

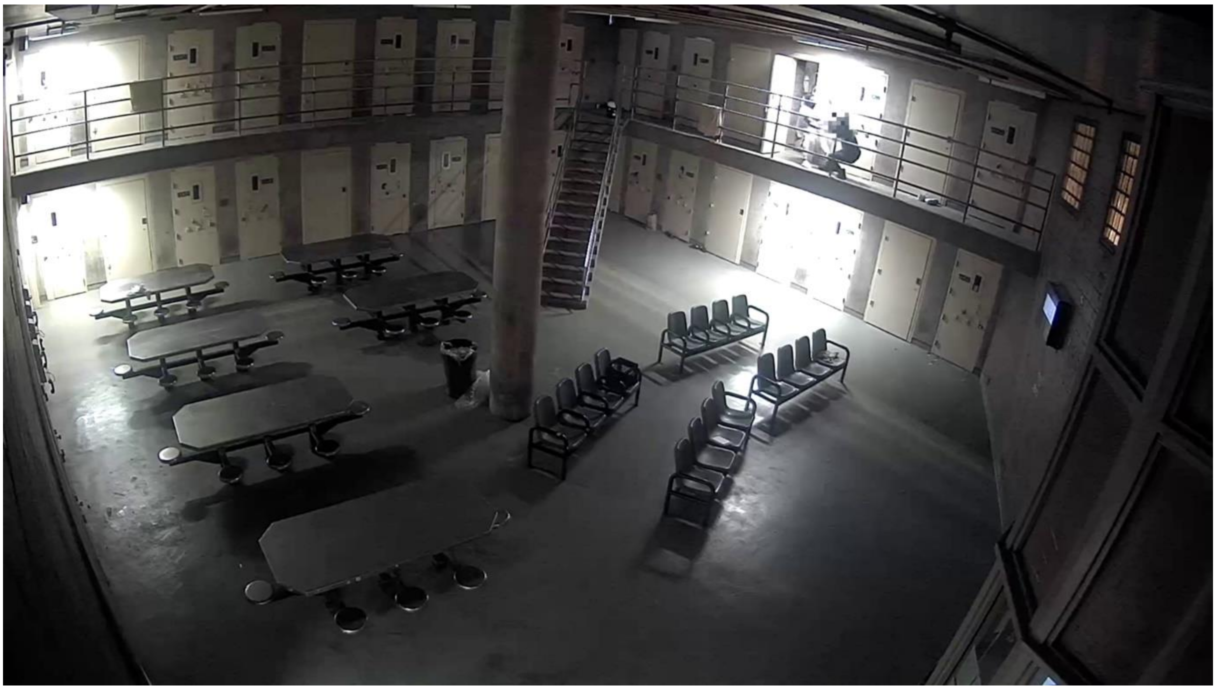
- Mounting height: 2.8–4 m
- Tilt: 20–45° below horizontal; avoid >50° (deep-learning performance drops).
- Roll: keep within ±10° to preserve body geometry.
- Occlusion risk: higher mounting reduces crowd overlap; avoid placing cameras directly against walls.
- Subject scale: people should occupy ~20–50% of frame height for action recognition. Avoid wide-angle and busy scenes
- Avoid top-down view. The model is not designed to detect action using top-down view
- 

Figure 3. An example of a false negative – the fight was not classified because the size of the objects is insufficient, with a very wide scene. In such cases, it is best to use the Spotlight function in critical areas and increase the sensitivity of UB detector.

Coverage & Framing

- Keep people in view ≥2 seconds – action classifiers need temporal context.
- Design for side-to-side movement across the FoV (better body articulation) rather than towards/away.
- Define an "effective area" to cut edge-of-frame false positives (door thresholds, escalator, frame exits).
- Avoid fisheye/dewarp skeleton/pose models degrade with distortion.

Critical areas: in high-risk locations (bars, concourses, prison yards), use two or more cameras with overlapping views. This reduces false positives and confirms fights even if one angle suffers occlusion.

Image Quality

- Resolution: 720p is sufficient for close/mid-range fights.
- Frame rate: 12–15 fps minimum, with maximum 1s GOP,
- Shutter speed: 1/100–1/250 s to minimise blur from punches or falling
- Bitrate: constant or capped VBR; avoid aggressive compression or non-standard codecs. Use "high" profile to reduce compression artefacts that obscure limbs/hands.

Lighting

- Minimum ~20 lux recommended.
- IR/night mode: works, but confidence drops; validate in POC.
- If the camera is facing into very strong light sources (sunset, stage lights, car headlights, flash strobes, etc.), people in the scene often appear as dark silhouettes. When that happens:
    - Body parts blur together → skeleton/action models can't separate arms from torso.
    - Movements lose detail → a punch, shove, or kick may look like a simple shadow flicker.
    - Detection confidence collapses → the system may either miss a real fight or generate false positives.

So in practice, when setting up cameras for fight detection:

- Don't aim them directly into bright light sources (avoid filming *towards* windows, doors, floodlights).
- Ensure even front or side lighting on the subjects.
- If unavoidable (e.g. entrances), compensate with additional lighting or re-position cameras to reduce the silhouette effect.

Analytics Setup

- Requires General Object Detection enabled
- Fight detection is CPU/GPU intensive – use alert-driven mode (trigger via TREX or UB) rather than continuous if resources are limited.
- Use continuous detection mode in critical areas to lower the number of false negatives.

Spotlight Optimisation

- Use iSentry Spotlight to create virtual zoomed-in "fight-zones" (bar areas, queue lines) from a single camera, increasing pixel density and reducing GPU load.

False Positives & Contextual Considerations

- Similar actions: Some behaviours (dancing, hugging, handshakes, celebrations, crowd surges) may resemble fighting and be flagged.
- Vandalism: Aggressive vandalism (kicking doors, smashing windows, striking objects) may be detected as fighting if the motion pattern resembles assault. Treat such alerts as high priority, but verify context before escalation.
- Occlusion: In dense or obstructed scenes, single-camera analytics may confuse pushing/shoving with benign crowding. Use multiple camera angles for confirmation.

Disclaimer for Specs

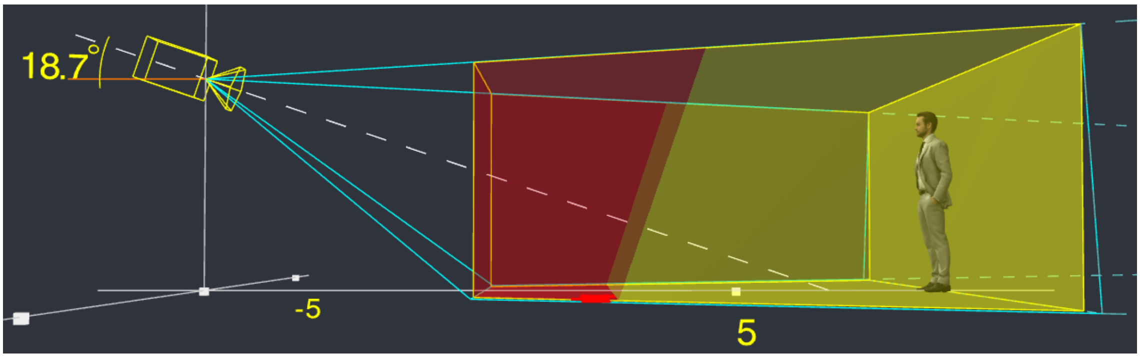
Violence detection analytics identify visible aggressive actions but do not guarantee recognition of every incident. Dancing, celebrations, vandalism and other non-violent actions may sometimes be classified as fighting. Multi-camera setups are recommended in high-risk areas to reduce occlusion errors. Performance depends on image quality, subject size, occlusion, lighting, and stream configuration. Aurora Verify can reduce false positives but may not eliminate them because it's not designed to work on "action" based alerts. Always validate performance in site-specific POCs.

Pose Verification

Pose Verification works in tandem with Action Recognition. When enabled this module will enhance the quality of the action recognition module but with added processing overhead. This module is also required for Hands Raised modules.

Hands Raised

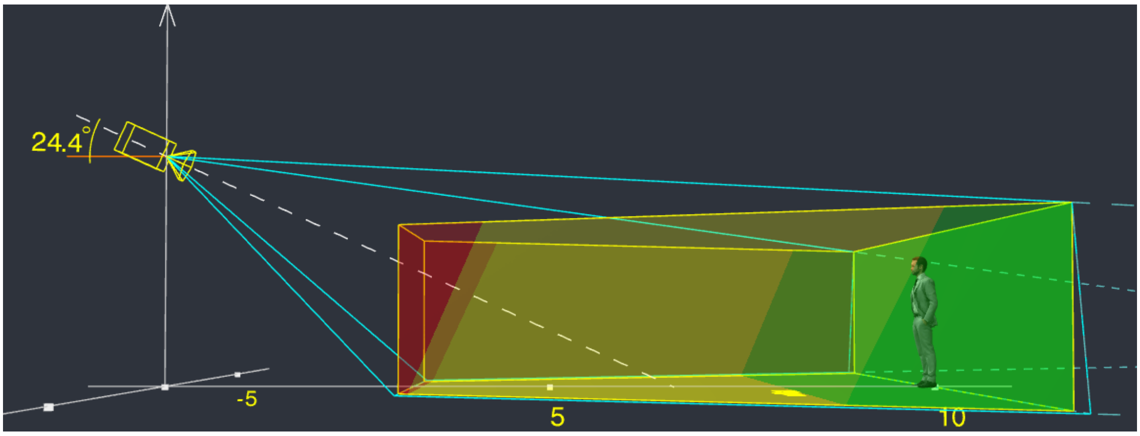
Two modules exist, Single Hand Raised and Both Hands Raised. Both of these modules require the pose verification module to be enabled. These modules require near horizontally mounted cameras and are only effective at close distance. People being analysed by this module should be no larger than 85% of the scene and no smaller than 60% of the scene. This Module requires one of the General Object Detection modules to be enabled. Ideally mounting height should be 1.8 to 2.5m and mounting angle should be from 0 to 35 Degrees below horizontal. Shallower angles will improve hands raised detection and is therefore recommended for this use-case.



Colour Estimation

This module is fully dependent on the light quality in the scene as well as the quality of the sensor in the camera. Objects analysed should be no larger than 70% of the scene size and smaller than 30% of the scene size.

When analysing object colour only the prevailing colour of the detected object or the background within the detected object bounding box is considered and care should be taken with objects placed against a brightly coloured background as this may affect accuracy of detection. Camera/sensor's white balance settings and external lighting will have a major impact on the accuracy of the detection also, and full spectrum lights or sunlight is ideal. This Module requires one of the General Object Detection modules to be enabled. Ideally mounting height should be 1.8 to 4m and mounting angle should be from 20 to 45 Degrees below horizontal.



Fire and Smoke Detection

1. What Firefly detects

- Smoke (including early, low-contrast plumes).
- Open flame (small to large).
- Sparks / radiance bursts (e.g., grinding, short arcing)—as radiance anomalies.

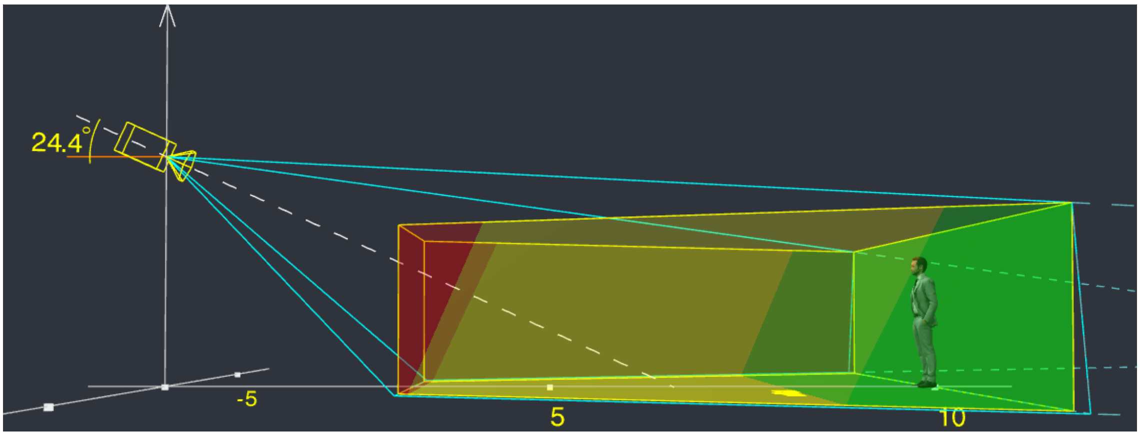
2. Camera & Lens recommendations

- Use visible-light colour cameras (our strong recommendation for smoke/fire).
- Preferred sensors:
  - 2 MP on 1/2.8" (e.g., Sony IMX290/IMX291) or better.
  - Cameras with 1/1.2" sensors and enhanced colour modes are excellent where light is limited.
- IR mode: supported, but expect more false positives/negatives in some scenes (e.g., shimmering IR back-light, monotone scenes). Prioritise colour whenever feasible.
- Lens: avoid extreme ultra-wide fields where possible; strong edge distortion and sky-dominance increase noise (competitor docs advise keeping HFoV <100° to limit distortion)

Why colour over IR? Subtle grey/blue/brown smoke tints, flame cores, and spark radiance cues are more separable in colour channels; IR can collapse these differences and exaggerate flicker.

3. Pixel density & scene size

- Firefly can detect and classify smoke and fire without any hard minimum "target area" but for best results we <u>recommend</u> keeping the smoke or fire clearly visible and at least 25% size of the scene and no larger than 70% of the scene.
- Design intent: Aim for clear visibility of early smoke/flame in the live frame (colour separation and texture), not a numeric PPM/px quota. If you must choose, zoom slightly tighter on the risk zone rather than covering excessive sky/bright backgrounds
- Ideally mounting height should be 1.8 to 4m and mounting angle should be from 20 to 45 Degrees below horizontal.
- Mounting angle: Prefer flatter angles with unobstructed view of the ignition plane. Avoid pointing into bright skies or high-contrast windows
- Wind & outdoors: Cross-wind can push smoke towards/away from the lens and create unusual apparent motion. Wind can alter plume direction and speed and even prevent detection in extreme cases—plan coverage accordingly (e.g., two angles on open yards)



Example of distance to detect flame or smoke by opening angle & flame width (to keep 25% size use Firefly Spotlight function to digital zoom in important areas)

Typical lens	Opening angle	0.3 m flame	0.5 m flame	1.0 m flame
~6 mm	60°	8 m	12 m	25 m
~12 mm	40°	12 m	22 m	40 m
~30 mm	15°	35 m	50 m	100 m

Disclaimer – indicative distances only:
The flame/smoke detection ranges shown are reference values, not guaranteed limits or real distances. Real-world performance varies with smoke transparency/opacity, smoke colour and growth rate, plume speed and direction (wind/stack effect), lighting and back-light, background contrast, camera resolution/bitrate/compression, lens FOV and focus, installation height/angle, scene occlusion, exposure/WDR/white balance, and IR vs colour mode. Detector sensitivity/verifier settings (e.g., Aurora Verify correlation) also affect results. Always validate on site with commissioning tests across day/night lighting states and use conservative design margins for life-safety applications. For best performance smoke and fire should occupy 20% of the image.

Figure 4. An example of excessive distance with an overly wide viewing angle of the camera.

Figure 5 Zoomed-in camera region using Firefly Spotlight Virtual Camera..

4. Critical infrastructure pairing (high-risk sites)

- Use dual sensing for early, reliable detection.
  For waste/garbage halls, battery storage, recycling lines, and warehouses with combustibles, deploy visible-light colour cameras (primary) alongside thermal cameras (supporting). Thermal helps with dust, fog, low contrast, and partial occlusion, improving early recall while colour maintains smoke/flame/sparks visibility.
- Alarm correlation in iSentryMMS (AND within a time window).

To reduce nuisance alarms, correlate Firefly (visual) events with a thermal alarm using Conditional Trigger plus a timer:

1. Create Rule A: Firefly – smoke/flame/sparks detected.
2. Create Rule B: Thermal – hot spot / over-temperature.
3. Create Rule C (Conditional): IF Rule A AND Rule B occur within X seconds → raise alarm.
   a. Typical X: 5–15 s indoors, 10–30 s in large or outdoor areas.
   b. Optionally include severity/temperature threshold from thermal to avoid minor heat sources.

- Important safety caveat (smoke vs heat).
  Thermal cameras do not "see" smoke. Relying solely on dual verification (visual AND thermal) can increase false negatives—e.g., cold smoke without a visible hot spot.
- For life-safety/mission-critical zones, use a fail-safe logic:
  - Primary action (high priority): raise alarm on Firefly smoke/flame/sparks OR thermal hot spot (whichever comes first).
  - Secondary action (escalation/filtering): use the AND-within-window correlation to prioritise/route alarms (e.g., higher severity, voice call-out), while still keeping the single-source alarm visible to operators.
- This approach preserves high recall while still reducing false positives when both modalities agree.

- Good practice for thermal pairing.
  - Set sensible temperature thresholds and rates of rise; avoid triggers from sun-heated surfaces or transient reflections.
  - Calibrate for emissivity and avoid aiming through glass or on highly reflective surfaces.
  - Mask known heat sources (machinery housings, exhaust flues) and use schedules for hot-work periods.
- Use Aurora Verify for context.
  You can pass candidate alarms to Aurora Verify to distinguish legitimate hot work (e.g., a blowtorch on a bench) from abnormal incidents (e.g., a vehicle catching fire, unattended smoke). This reduces false positives without suppressing early smoke-only events. But please be aware that this can introduce delay in alarm triggering. Also please keep in mind that this can also reducing false positives at cost of some false negatives due to hallucinations.

5. Aurora Verify — cutting false positives & adding context

Use Aurora Verify in two ways:

1. Human-in-the-loop verification to suppress nuisance alarms (lights, reflections, compression artefacts) before escalation.
2. Context classification (natural-language prompts) to distinguish legitimate hot work vs danger:
    - "blowtorch on workbench" → normal operation, non-incident
    - "smoke from engine bay, vehicle stationary, no workers nearby" → escalate
      This addresses the gap typical smoke/fire engines have: they are visual-phenomenon detectors, not use-case reasoners — Aurora Verify adds the missing scene understanding layer.

6. Configuration tips (Firefly)

- Active areas: Draw effective zones over genuine risk regions (reduces noise—an approach also recommended by competitor planning guides).
- Sensitivity vs verification time:
    - Critical life/fire safety → bias recall: higher sensitivity, shorter verification; then route via Aurora Verify to filter nuisance.
    - Operational monitoring → moderate sensitivity, moderate verification.
- Radiance mode: Enable where cutting/grinding occurs to catch radiance bursts; combine with work-hours schedules to reduce noise outside shifts.
- Multi-angle design outdoors: Cover from at least two bearings where wind or occlusion is likely.

7. Commissioning & acceptance

- Dry runs with small controlled smoke (e.g., smoke pellet at safe distance) and shielded test flame where permissible by site policy.
- Capture before/after tweaks: exposure, gain, WDR, colour mode; confirm latency and notification path.
- For sites with dynamic lighting (day, dusk, IR), test each state

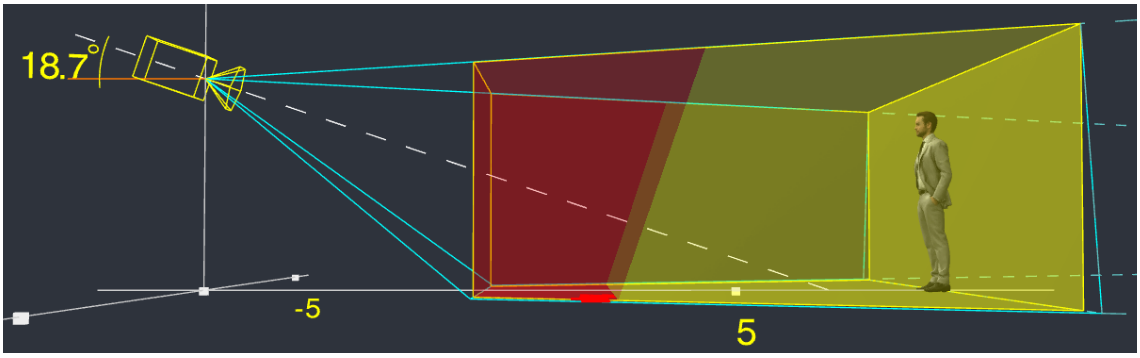
8. Operational stance on safety (important)

If detecting smoke or fire is mission-critical, always prioritise reducing false negatives, even at the cost of more false positives. That means higher sensitivity, shorter verification and trained operator validation.

Face Detection

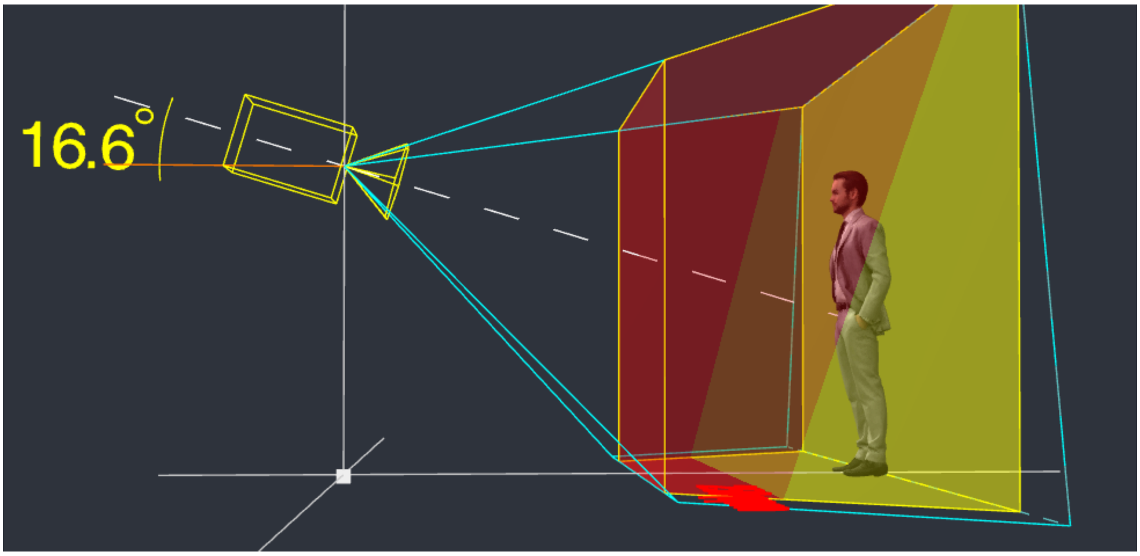
Face detection is a module used to just detect faces of persons in a scene.

This module is typically used to enhance standard object detection by additionally detecting faces. It can also be useful for counting and detecting people looking in the direction of the camera. The face that needs to be detected must be visible to the camera, and at least 8% of the scene size and no more than 40% of the scene size. For smaller faces, the quality of the video will be paramount. This module is typically for shorter distances (although still effective at medium distance), and for low mounted cameras, with an angle of less than 35 degrees below horizontal.



Gender Estimation

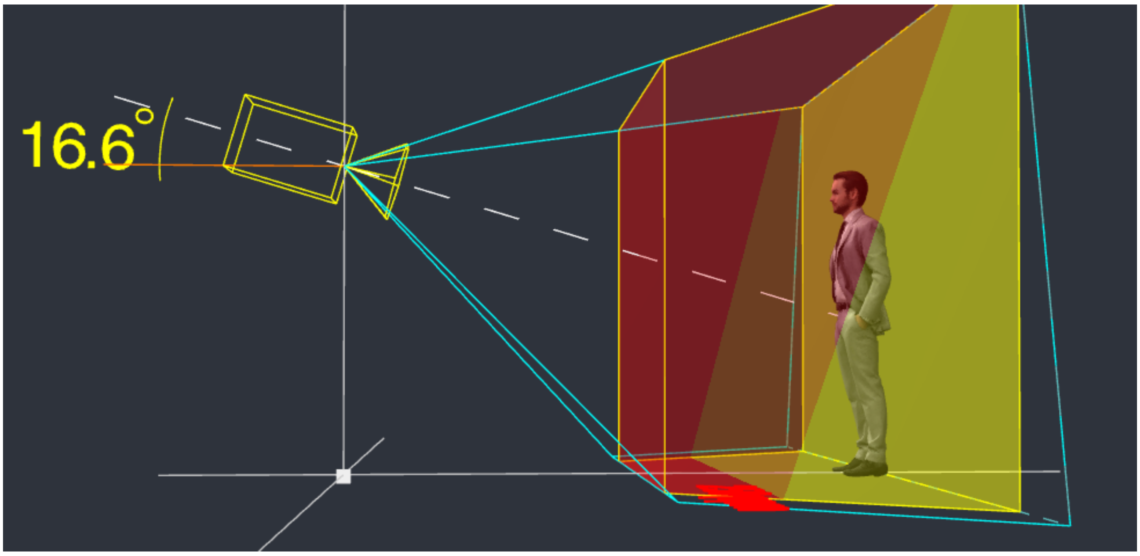
This module relies on the Face detection module and will add an estimation of the age of the face detected. This module depends on having a really good view of the face detected, and is therefore highly dependent on video quality, visibility of the face and angle of view (Horizontal or nearly horizontal angle of view preferred). The detected face must be between 20% and 40% of the scene size, and this module is designed for short distance only.



Age Estimation

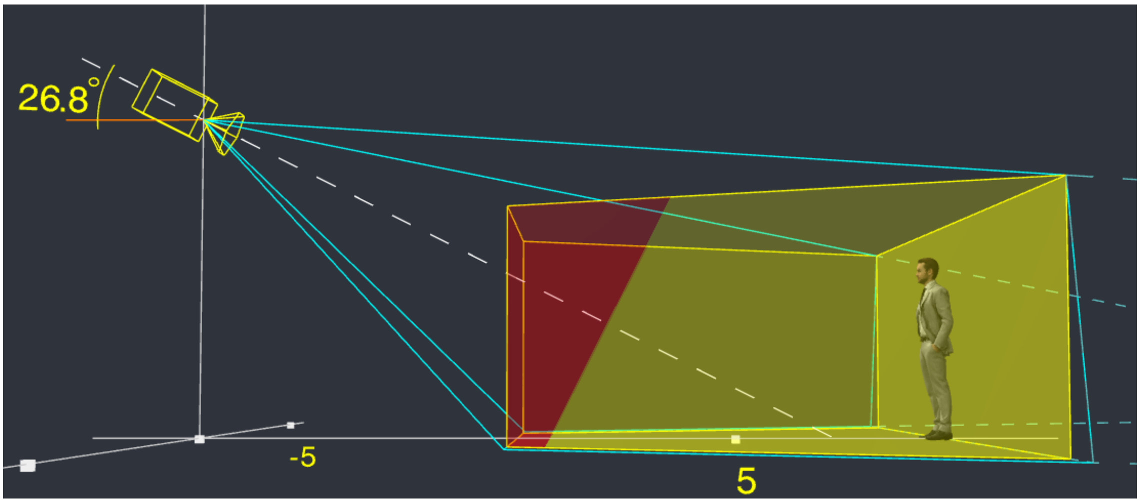
This module relies on the Face detection module and will add an estimation of the age of the face detected.

This module depends on having a really good view of the face detected, and is therefore highly dependent on video quality, visibility of the face and angle of view (Horizontal or nearly horizontal angle of view preferred). The detected face must be between 20% and 40% of the scene size, and this module is designed for short distance only.



Helmet Detection

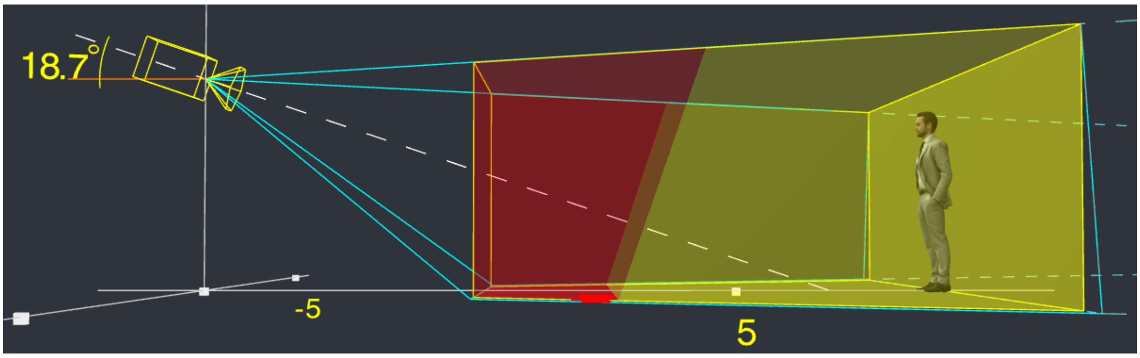
This module a specifically built to detect motorcycle and other helmets. Helmets must be clearly visible and at least 15 % of the scene size. Ideally mounting height should be 1.8 to 4m and mounting angle should be from 20 to 45 Degrees below horizontal.



Face Mask Compliance Detection

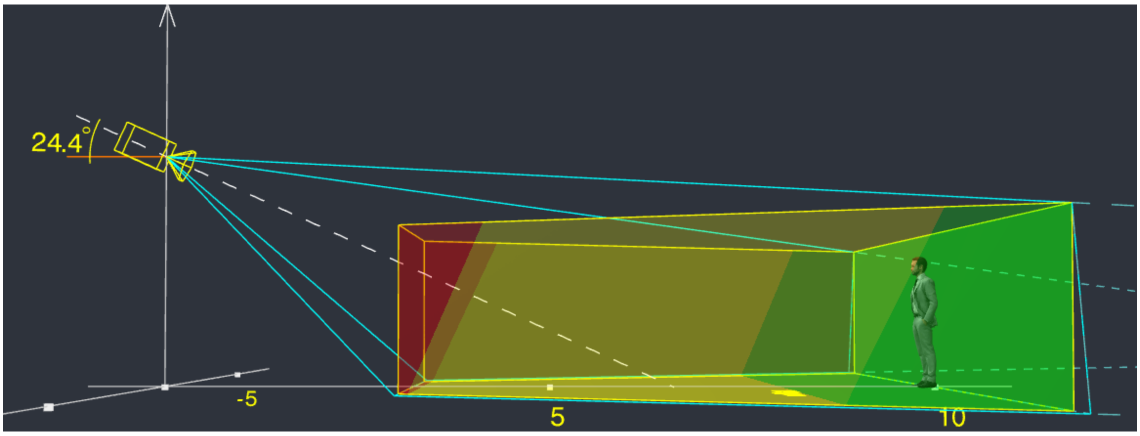
The mask compliance module will detect whether a person's nose and mouth is fully covered, partially covered or not covered at all. Ideally mounting height should be 1.8

to 4m and mounting angle should be from 20 to 45 Degrees below horizontal. The detected face must be at least 15% of the scene size, and this module is designed for short distance only.



PPE Compliance Detection

This module was designed to detect high visibility jackets, hard hats, safety goggles, ear defenders as well as face shields. This module was specifically trained for common PPE equipment and may require additional training for specific use cases. The equipment that must be detected must be clearly visible. The individual wearing the safety equipment must be at least 25% of the scene size and no larger than 80% of the scene size. Ideally mounting height should be 1.8 to 4m and mounting angle should be from 20 to 45 Degrees below horizontal.



Weapon Detection and Gun Pose Detection

The weapon detection module can detect weapons such as handguns and rifles.Gun Pose Detection module can detect when a person is trying to use a gun (shooting pose). It is important to note that hidden or partially occluded weapons will not be detected. For successful detections of weapons these must be clearly visible and unambiguous in appearance.

Camera Requirements for Optimal Gun Detection

To maximise the accuracy and performance of the iSentry system, cameras and infrastructure must be designed and configured to the following technical criteria. This guidance is descriptive so installers, designers, and security leads can make clear trade-offs and document residual risk.

1) Safety and Operating Principle (Read First)

- Human-in-the-loop is mandatory. Gun/weapon alerts are advisory and must be verified by a trained operator before any action. Do not arm/disarm locks, trigger sirens, or initiate lockdowns purely from an AI label.

- Residual risk exists. Missed detections (FN) and false alerts (FP) will occur. Select and sign off a risk posture per site.

- Aurora GenAI is advisory only. It can add context but may be wrong (LLM "hallucination"). Never auto-downgrade an alert because attire "looks like" a uniform; treat such hints as supplementary, not decisive.

2) The MINIMUM Pixel Density and Resolution (PPF/PPM and %-of-frame)

The detection distance and accuracy are based on the camera resolution, angle and zoom level. Simply put, if a human eye can identify the gun, then the system is also able to detect it. Effective detection depends on object size in pixels as well as scene complexity, camera angle, lighting condition and codec used and validate with %-of-frame checks so the spec holds across fields of view and resolutions..

Unified targets (well-lit ~300 lux; increase by ~20% for ~299-100 and by 40% for scenes ~99-50 lux)

- Handguns and rifles: ≈ 190 PPM (≈ 58 PPF).

- Apparent size checks:

  - Handgun (typical ~0.19 m × 0.13 m): ≥ ~3% of frame width or ≥ ~3.5% of frame height (whichever axis the gun spans most).
  - Rifle (typical ~0.90 m × 0.25 m): ≥ ~13–14% of frame width or ≥ ~6–7% of frame height.

Use both: meet 190 PPM and the %-of-frame criterion that corresponds to the firearm's orientation. This prevents the "wide-angle looks big in megapixels but tiny on screen" trap.

Practical notes

- Keep the field of view tight so hands/forearm + object are consistently large enough.

- Use Spotlight/virtual ROIs to crop key lanes so small pistols appear larger to the analytic without moving the physical camera.
- Quick conversion: PPF = PPM × 0.3048 → 190 PPM ≈ 58 PPF.
- If the site chooses FN-Minimised, exceed 190 PPM or tighten ROIs.

Resolution & sensor

• HD (720p) streams are highly recommended for strong performance at practical compute cost.

• Higher resolutions (4 MP/4K) massively increase compute, can increase analysis latency, and often underperform if the FOV stays wide (object remains small). If you need more detail, use optical zoom / tighten FOV rather than simply raising resolution.

• Larger sensors help with noise and edges when available; resolution uplift alone is not a substitute for correct geometry.

• Prefer larger sensors for lower noise and sharper edges (minimum 1/2.8" for 1Mpx/2Mpx, 1/2" for higher resolution)

• Higher resolutions capture more detail but demand better lighting to avoid motion blur and noise.

3) Risk Posture & Tuning (must be chosen and signed off)

Choose one per site (or per camera group) and document it on the commissioning form:

FN-Minimised (Safety-max):

• Goal: reduce missed weapons even if FP rate increases.

• Settings: lower detection thresholds; tighter ROIs/Spotlights; broader alerting; faster operator escalation.

Balanced:

• Goal: practical balance of FP workload vs timely capture.

• Settings: platform defaults with local tuning and ROIs for critical zones.

FP-Minimised (Workload-min):

• Goal: reduce nuisance alerts at the cost of potentially higher FN.

• Settings: higher thresholds; more aggressive masking; ROI focus on choke points only.

• Not recommended for high-risk areas (front-of-house, entrances, tills, security screening).

Important: For high-risk environments (public-facing venues, schools, transport hubs), prefer FN-Minimised or Balanced. Document the rationale and approvals.

4) Illumination and Dynamic Range

Lighting quality is as important as pixels. Aim for even, flicker-free light with good subject contrast.

• Design baseline: target ~300 lux along the approach to detection lines.

- Minimum operational: keep ≥ 50 lux; below this, increase PPF targets by ~20% and expect more misses.

- Backlight management: enable WDR/BLC where needed, but test for artefacts (double borders/ghosting) that can create false shapes. Adjust until edges look clean on a moving hand holding a dark object.

- Low light: if colour collapses, allow monochrome/IR mode for contrast—but prefer adding white light to preserve edges and prevent blur.

- Avoid flares and bloom: keep sun, headlights and strong luminaires out of the lens; diffuse or reduce IR if you see halos.

- Reflective surfaces: minimise mirrors, glossy floors and glass dominating the scene; reflections confuse edges and silhouettes.

Lighting tips & tricks

- Even illumination: aim for ≥ 150 lux across the detection area to reduce shadows and colour shifts.

- Avoid direct light sources in the lens.

- Avoid drastic lighting changes: e.g., skylights, big south-facing windows without shades.

- Beware indirect light: IR reflections from walls, floors or glass doors can lower contrast.

5) Motion, Shutter and Frame Rate

Fast hand motion (draw/aim) is where blur kills detection. Balance exposure with frame rate and add light rather than letting the shutter run long.

- Analytics rate: maintain ≥ 12.5–15 fps of analysable frames.

- Shutter guideline: start around 1/30 s; if you see smearing on hands/forearms, add light and keep the shutter short rather than over-using noise reduction.

- Auto-exposure (AE): allowed, but cap the minimum frame rate so AE cannot drop below the analytics rate.

- Noise/sharpness: keep temporal/spatial NR and sharpening near defaults; aggressive NR creates trails and false edges.

- Visibility window: ensure the firearm (or the posture suggesting one) is clearly visible for > ~1 second to allow stable classification.

5a) Indicative Detection Performance vs Time-in-View

Values below are planning ranges, not guarantees. Replace with the site's own acceptance test results (see "How to generate your site's table" below).

Definitions

- Continuous visibility time: a contiguous sequence of frames in which hands/forearm and the suspected weapon occupy ≥ 10% of the frame (or meet the design PPF/PPM), are < 50% occluded, and are not smeared by motion blur; the scene meets ≥ 15 fps capture and shutter ≈ 1/30 s (or faster) with compliant lighting.

• Pd (probability of detection): share of test runs where an alert is produced while the object remains visible (no later than 0.5 s after the window ends).

• Pfa (false alarm rate): share of non-weapon runs producing a weapon alert.

Indicative planning ranges (replace with your measured data)

Continuous visibility window	Typical Pd range (compliant scene)	Common failure modes	Notes
0.3–0.5 s	30–60%	Hand smear, low fps, occlusion spikes	Very short reveals; add light / shorten shutter; consider tighter ROIs. Increase detector sensitivity (TREX)
0.5–1.0 s	60–90%	Backlight, IR bloom, reflective floors, VBR artefacts	Ensure ≥ 15 fps, GOP ≈ FPS,
1.0–2.0 s	80–96%	Minor occlusions, backlight, motion blur and compression artifacts	This is the design target window; tune thresholds/ROIs
2.0–3.0 s	90–98%	Complex/busy scene, wrong detector	Use Spotlight/ROIs so the object is larger to AI. Increase detector sensitivity, increase alert frames, lower time between alert frames
≥ 3.0 s	95–99%	Deep shadow, severe WDR artefacts	If lower than this, re-check PPF/PPM, lux, shutter and NR settings.

How to generate your site's table

1. Fix the scene to design criteria (PPF/PPM, lux, fps/shutter, WDR tuned, GOP≈FPS, TCP).

2. Prepare scenarios (safe props only): near/mid distance; day/night; mild backlight; partial occlusion.

3. Run ≥ 30 trials per window (0.5 s, 1 s, 2 s, 3 s). Start timing when the weapon first meets the definition above; count a detection if the alert fires within that window or ≤ 0.5 s after.

4. Compute Pd and Pfa per window, report sample size n, and (ideally) a 95% CI.

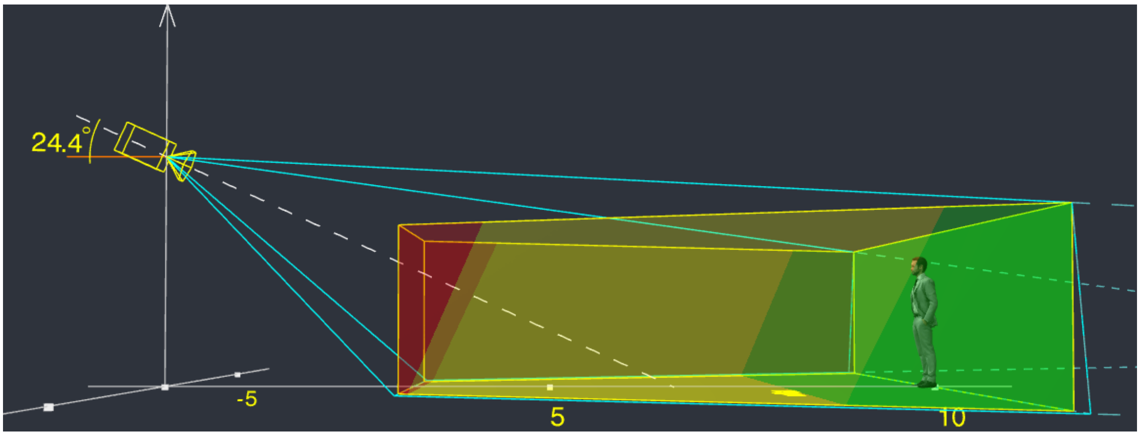
5. Publish your site table:

Window	Pd% (n=__)	Pfa% (n=__)	Conditions (PPF/PPM, lux, fps/shutter)
0.5–1.0 s	__%	__ %	e.g., 50 PPF, 250 lux, 15 fps / 1/33 s
1.0–2.0 s	__%	__%	…
2.0–3.0 s	__%	__%	…
≥ 3.0 s	__%	__%	…

6) Camera Positioning and Field of View

Mounting geometry decides how the model "sees" hands and objects.

• Angle & height: mount 1.8–4 m high with the view ~20–45° below horizontal. Avoid bird's-eye (top-down > 50–60°) — it distorts shapes and reduces weapon classification reliability.



• Near-perpendicular perspective: prefer oblique, human-eye-like views across the area of interest.

• Occlusion: plan lines of sight so hands/forearms are not blocked by pillars, shelving, signage or foliage. Hidden or heavily occluded items are not detected.

• Crowds: overlapping bodies raise miss rates; use multiple angles or narrow ROIs for choke points.

• Vehicles: through-glass and in-vehicle views are harder; treat as reduced-confidence scenarios.

- Stream rotation: avoid rotated/corridor streams for gun detection unless explicitly supported in your camera hardware and software.

---

7) Video Compression, Transport and Stability

- Codec: H.264 (Main/High) or H.265; prefer CBR for stability.
- FPS: ≥ 15 fps (25/30 if bandwidth allows).
- GOP: set GOP ≈ FPS (e.g., 15 fps → 15-frame GOP). Keep keyframe interval ≤ 2× FPS (stay under ~30).
- Transport: prefer TCP on unreliable links to avoid loss/jitter; avoid aggressive "smart codecs" that alter shapes/frame cadence.
- Clocking: keep cameras and servers NTP-synchronised so events align with video.
- Network headroom: design bandwidth so bitrate does not cliff under motion; dropped frames translate into missed detections.

8) System Health Monitoring

- Continuously monitor connectivity, FPS and bitrate per camera.
- Alert on FPS drops, bitrate starvation, RTSP errors and keyframe gaps.
- Review a sample of alerts weekly for FP sources and tuning opportunities.
- Re-test after any change in lighting, angle, compression or rules.

9) Maintenance and Re-calibration

- Lens care: clean lenses routinely; dust, grease and moisture degrade edges and contrast.
- Physical alignment: verify mounts and FOV after maintenance or accidental bumps; small shifts can invalidate ROIs and trigger assumptions.
- Environment drift: as layouts, shelving, signage or foliage change, update ROIs/masks and re-verify PPF/PPM at the detection line.

10) Tips for Minimising False Positives (Aurora GenAI — Advisory Only)

Aurora can contextualise alerts (e.g., "object appears to be a firearm" vs "phone/tool", or "this looks like a screen/poster"). Use it as secondary input only.

- Quality pre-conditions: good light, clean edges, subject big enough in frame.
- Latency: depending on mode and load, Aurora may add ~0.5 to a few seconds. Configure parallel routing so the base alarm reaches operators immediately; let Aurora's text arrive afterwards.
- Never auto-downgrade or suppress an alarm purely because Aurora suggests "uniform" or "authorised"; verify via human judgement or secondary means (radio check, second camera).

11) Commissioning (Plain-English Walk-through)

1. Frame the shot so the planned detection line hits the PPF/PPM targets (190PPM adjusted for light).

2. Measure light at those lines (aim ~300 lux; keep ≥ 50 lux). Add white light where needed.

3. Tune exposure: AE on with minimum 1/30s, min FPS ≥ 12.5–15; keep NR/sharpness moderate. be careful with aggressive WDR.

4. Verify stream hygiene: H.264/H.265 CBR; FPS ≥ 15; GOP ≈ FPS; monitor stream jitter, it should be <100

5. Dry-run with safe props or customer test clips; prove alerts reach the SOC and that operators can verify quickly.

6. Document risk posture (FN-Minimised / Balanced / FP-Minimised) and get sign-off.

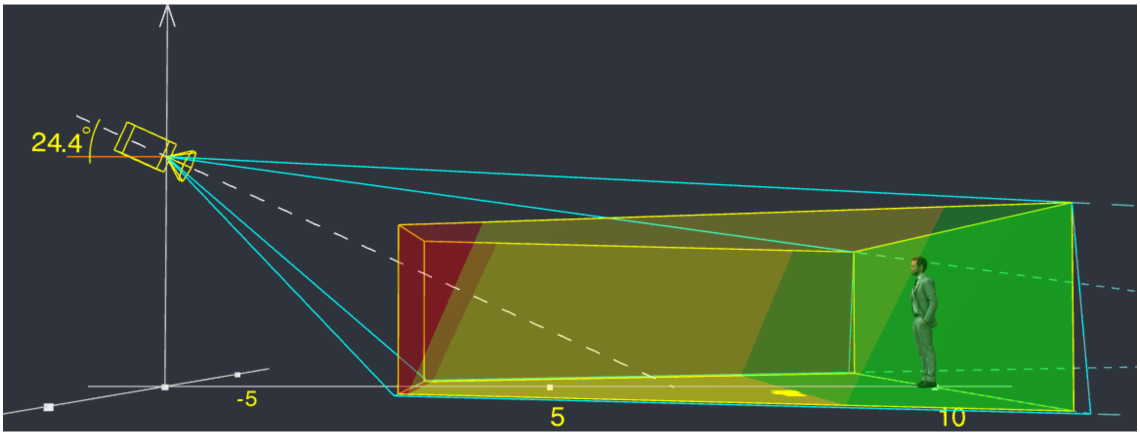
7. Record limitations: concealed/occluded items, steep top-down angles, glass/glare scenes and dense crowds reduce reliability.

12) What the System Will Not Do

- Detect concealed, bagged or heavily occluded weapons.
- Perform reliably from near-vertical (bird's-eye) angles.
- Provide legal proof of authorisation from attire; uniforms/insignia can be occluded, counterfeit or ambiguous.
- Operate as a certified life-safety device without a human in the loop.

Wheelchair Detection

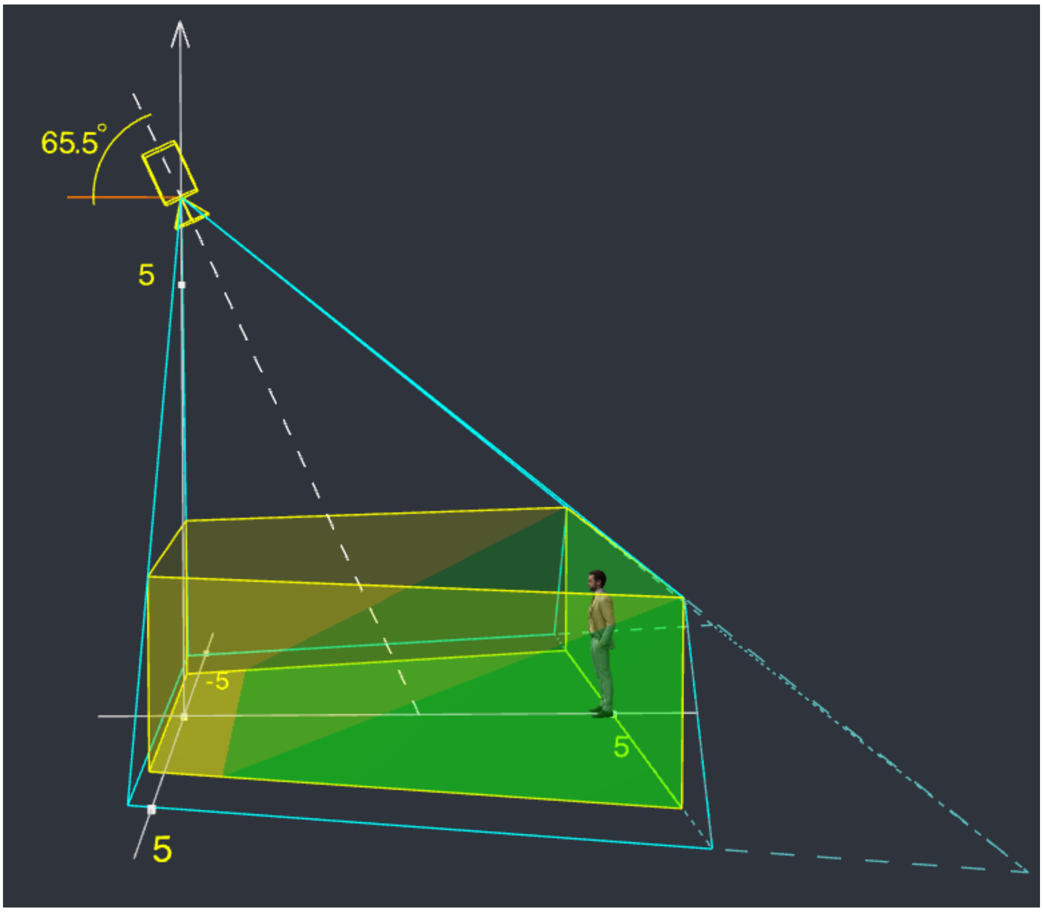
This module will detect wheelchairs as long as they are clearly visible and no smaller than 30% of the scene size and no larger than 70% of the scene size. Ideally mounting height should be 1.8 to 4m and mounting angle should be from 20 to 45 Degrees below horizontal.



Top-down Object Detection

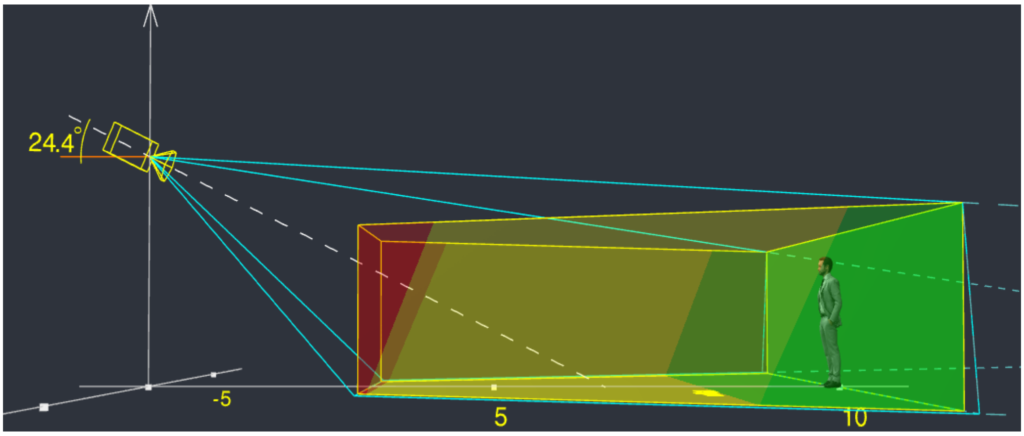
The top-down module has been specifically built for high, near vertically mounted (Not vertically mounted) cameras. Ideal mounting angles would be between 65 degrees to 75 degrees below horizontal. This module is specifically designed for person detection and persons as small as 5% of the scene size can be successfully detected, however the accuracy of this module is highly dependent on video quality, illumination, contrast, mounting position as well as angle.

When applying this module for vertically mounted or top-down cameras, quality of classifications will reduce dramatically, similarly using this module with a low mounted camera with an angle less than 60 degrees below horizontal will result in inferior quality classifications with an increase in false positive detections.



General Object Detection and Classification

General object detection includes a multitude of objects that can be detected. This module is well suited to most use cases and should typically be the module of choice. Because of the variety of objects available for detection size vary greatly depending on the object type and how easily the object is recognised. Focusing on person detection, a person should be at least 5% of the scene size and no larger than 70% of the scene size. Ideally mounting height should be 1.8 to 4m and mounting angle should be from 20 to 45 Degrees below



Performance Considerations

Camera and Image Notes

The images below were taken with an UniView IPC262ER9-X10DU CCTV camera with 12mm lens at 1280 x 720.

DL Settings:

- Minimum threshold: 25%
- Mode: Accuracy

Below performance results assume a similar camera, lens and configuration.

General Object Detection and Classification

In accuracy mode, there is a probability of 90% that people can be detected when they have pixel density of about 25 pixels/meter. An average adult at this density will have size of 20x45 pixels (width x height). At lower density, 18 pixel/meter, the detection probability drops to about 70%.

	Person	Car
25 Pixels/m (1.5% x 6%)	90%	
18 Pixels/m	70%	

Close to medium distance:

Far distance:

ANPR Camera Guidelines

General Requirements and Recommendations

Any video camera is designed for video surveillance, however, it serves various purposes. It is crucial to understand the conditions under which information will be transmitted, as well as the quality of that transmission. In this specific case, the video camera should only be focused on one task - the recognition of vehicle license plates. Therefore, when designing an LPR (License Plate Recognition) system, we recommend selecting video cameras with the following technical specifications.

Video Camera Technical Requirements

The following technical requirements are important for video cameras used in license plate recognition:

1. Sensitivity: Sensitivity should be at least 0.01 lux.

2. Matrix Size: The matrix size should be no less than 1/2.8 for HD and FullHD camera or 1/1.8 for 4Mpx.
3. Image Resolution: FullHD (2Mpx) Increasing the camera resolution does not improve the license plate recognition function, as it may reduce the camera's photosensitivity.
4. Lens Selection: Choose a lens with the highest aperture, preferably not less than F/1.4. Mind the depth of field.
5. Day/Night Mode: Color cameras must have a full Day/Night mode with a removable IR cut filter.
6. Electronic Shutter Speed: The camera should allow setting a fixed electronic shutter speed according to the following guidelines:
   - 1/500 for speeds up to 40 km/h
   - 1/1000 for speeds up to 80 km/h
   - 1/2000 for speeds of 160 km/h and above

Refer to the table below for the recommended maximum shutter time corresponding to different camera angles and car speeds, using 1 ms (1/1000 s) as the unit of measurement:

Camera Angle	30 km/h	50 km/h	80 km/h	110 km/h	130 km/h
5 °	1/50	1/80	1/125	1/200	1/200
10 °	1/100	1/160	1/250	1/400	1/500
15 °	1/160	1/250	1/400	1/500	1/640

20 °	1/200	1/320	1/500	1/800	1/1 000
25 °	1/250	1/400	1/640	1/1 000	1/1 000
30 °	1/320	1/500	1/800	1/1 000	1/1 250

For details, see the Maximum Shutter Time section of this manual.

Pixel Density

In video surveillance, a parameter known as pixel density is commonly used. Pixel density is typically expressed as pixels per meter. It is important not to confuse the pixel density in video surveillance with the pixel density of monitors, which is measured in pixels per inch (PPI). The advantage of using the term 'pixel density' is that it encompasses several factors, including sensor size, pixel count, lens focal length, and distance to the observed object.

When designing an LPR system, it is crucial to calculate the required pixel density to ensure optimal image quality for reliable license plate recognition. The license plate must be captured with enough pixels to clearly distinguish the letters and numbers. The best recognition results are typically achieved within a specific range of license plate sizes in the image, measured in pixels, which is determined through practical experimentation.

For a European standard license plate, it is recommended to have a minimum of 75 pixels to adequately capture the letters with full contrast. Most LPR software requires a width of 100-150 pixels across the license plate for optimal performance:

To ensure the effective performance of the LPR plugin, it is necessary for the recognition area to have a minimum height or width of 200 pixels. Additionally, the aspect ratio should be within the range of 0.6 to 2, maintaining the proportional relationship between the dimensions. These criteria help optimize the functionality and accuracy of the LPR plugin.

Examples of license plates for different countries are shown in the picture below:

(g)

(h)

a) India
b) USA (California)
c) Australia
d) Pakistan
e) England
f) Italy
g) Japan
h) Iran

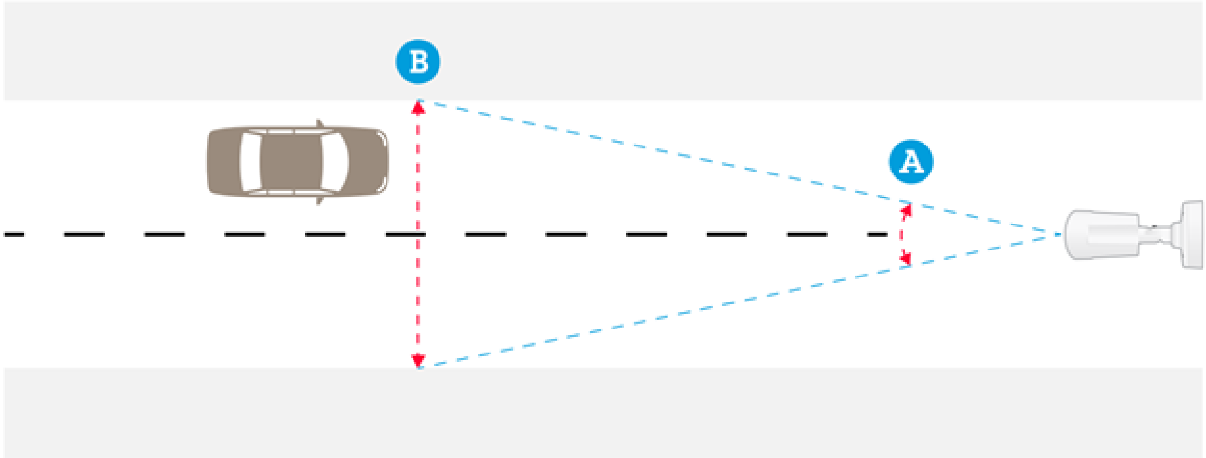
In most cases, a resolution of up to 1280x720 pixels, which corresponds to the HD standard of 1MP (720P), is sufficient. Using higher-resolution cameras does not provide significant benefits.

To determine the actual pixel density, follow these steps:

1. Save a frame from the video camera
2. Open the saved image in Paint (or any image editing software)
3. Use a block or rectangle tool to highlight the vehicle number
4. Check if the width of the number (in pixels) matches the desired value you have set
5. If the width does not match, adjust the camera lens accordingly

From a video camera's perspective, the number of pixels on a license plate is determined by the image sensor's resolution and the field of view. A variable focal length lens provides the flexibility to adjust the field of view by zooming in or out.

The width of the scene visible in the image (B) depends on the chosen field of view (A):



The number of pixels across the width of the license plate depends on the camera's resolution and the width of the scene. In this example, a camera with a resolution of 1080x1920 pixels:

1. When zoomed in on one lane (4 meters wide), the license plate occupies 250 pixels
2. When zoomed out to cover almost two lanes (6.5 meters wide), the license plate occupies 154 pixels

The table below illustrates the recommended horizontal field of view to effectively cover one, two, and three lanes at various coverage distances, taking into account the sensor format, focal length of the camera lens, and potential lens distortion:

Capture Distance:	5 m (~ 16 ft)	10 m (~ 33 ft)	30 m (~ 98 ft)	16 m (~ 164 ft)	80 m (~ 262 ft)
1 lane < 4 m (~ 13 ft)	33º – 44º	17º – 23º	6º – 8º	3º – 6º	2º – 3º
2 lanes < 8 m (~ 25 ft)	62º – 77º	33º – 44º	11º – 15º	7º – 9º	4º – 6º
3 lanes < 12 m (~ 39 ft)	84º – 100º	48º – 62º	17º – 23º	10º – 14º	6º – 9º

Note that the values provided in the table are defined based on European standard license plates. It's important to consider that smaller license plates may require a higher resolution to ensure sufficient clarity and accurate recognition.

The minimum resolution is indicated in the table below:

Minimum Resolution	
1 lane, width < 4 m (~ 13 ft)	1 MP (HD, 720p)
2 lanes width < 8 m (~ 25 ft)	2 MP (Full HD, 1080p)
3 lanes width < 12 m (~ 39 ft)	5 MP (5MP Lite,1920p)

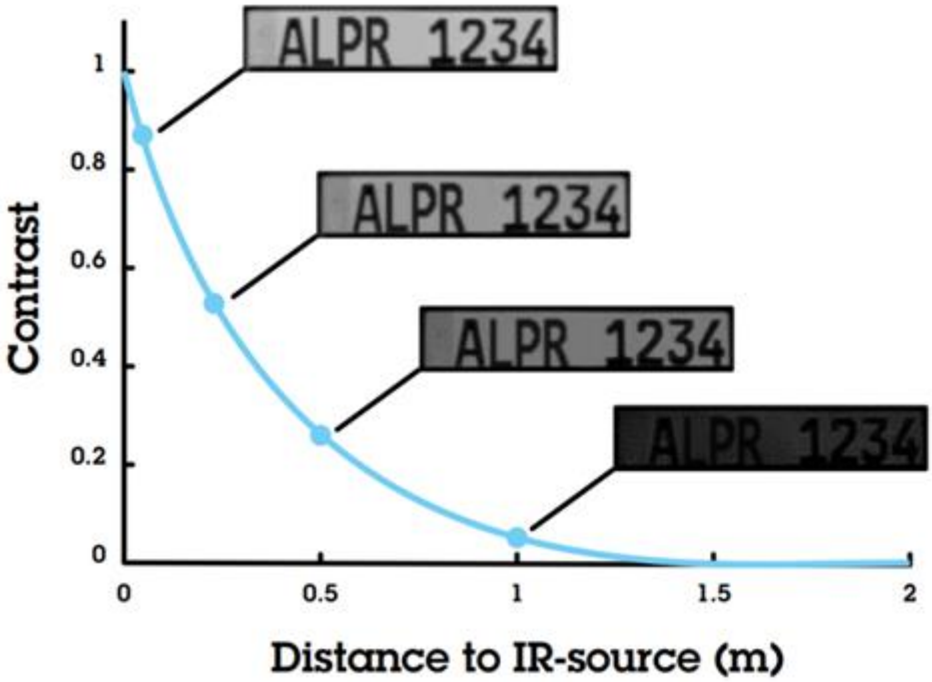
Using high-resolution video cameras has the disadvantage of increasing the processing time for LPR software to analyze each image. This extended processing time can pose a higher risk of missing certain numbers, especially in situations of heavy traffic.

IR Light

Artificial lighting is essential for capturing license plates at night. Infrared (IR) light is commonly employed due to its invisible nature and non-blinding effect on drivers. Most license plates have infrared-reflective properties, allowing IR light to enhance visibility and contrast even in dark or cloudy weather conditions. IR light can originate from LEDs integrated into the camera or external IR sources.

It is important to note that the intensity of IR light diminishes quadratically as the distance between the light source and the object increases. This relationship applies to reflective objects like license plates. Consequently, to maintain consistent visibility, a fourfold increase in IR power is required for each doubling of the distance between the light source and the object.

The brightness and contrast of a license plate within a camera image decrease rapidly as the IR source is moved farther away from the camera (perpendicular to the road):

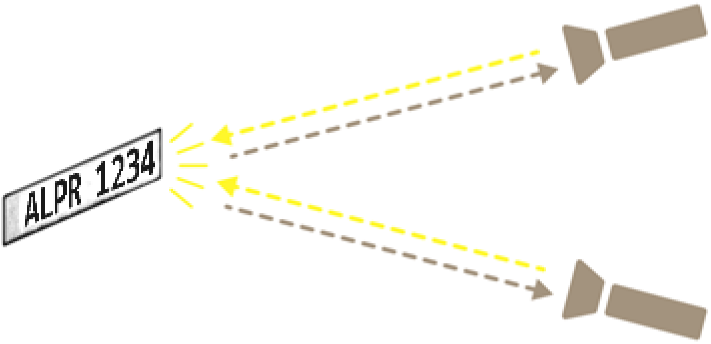


The acceptable quality of night license plate recognition is shown in the picture below:

External Sources of Infrared Radiation

If the IR range is insufficient for the built-in LEDs, or if the camera does not have built-in IR LEDs, infrared sources external to the camera can be used. The light cone of the IR source must match the camera's field of view at the appropriate zoom level. License plates are made of reflective material, which means they reflect light right back where it came from, no matter what angle the light hits the license plate. When using an external IR source, the reflected IR light will return to the source.

For this reason, external IR sources must be placed close to the camera so that the reflected light can actually enter the camera:



The following pictures illustrate the contrast of a license plate at a distance of 10 m from a passing vehicle, considering the distance between the camera and an external source of infrared radiation:

1. The IR source is positioned 5 m away from the video camera
2. The IR source is positioned on the same optical axis as the video camera

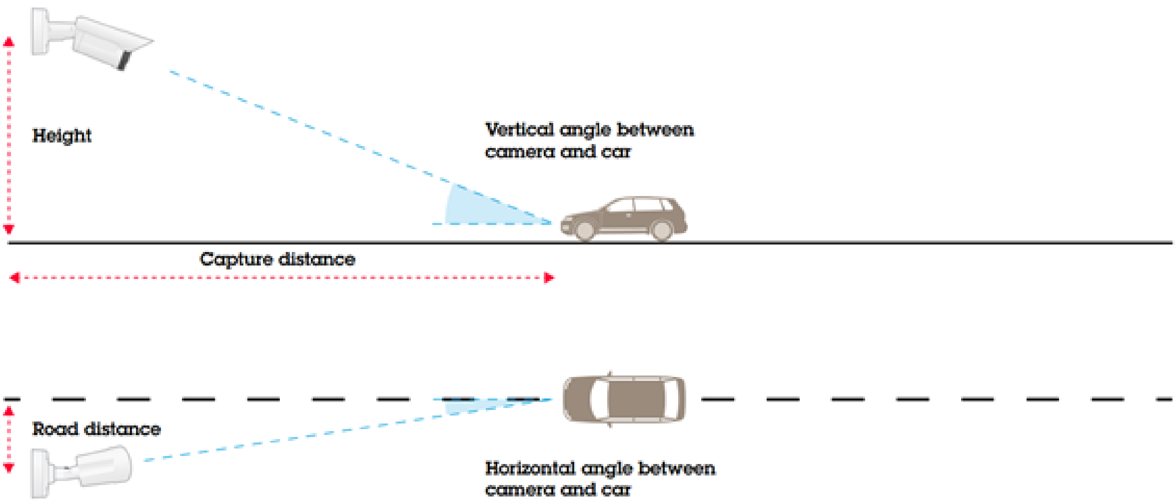
To ensure optimal illumination, it is crucial to place the IR source parallel to the camera. This arrangement ensures that the light falls on the section of the road captured within the camera's field of view, enhancing license plate visibility.

Video Camera Position

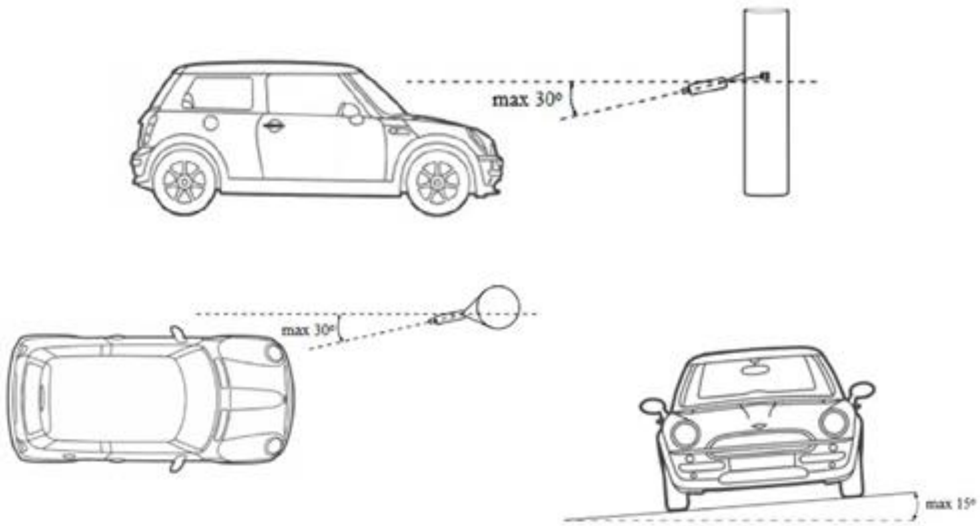
The main parameters of the video camera position are:

1. Height of installation above the track surface.
2. Vertical angle between the optical axis of the video camera and the average vehicle speed vector.
3. Horizontal angle between the optical axis of the video camera and the average vehicle speed vector.
4. Angle of the video camera image raster to the track surface.

The mounting height and distance from the center of the road determine the angle between the camera and the direction of vehicle motion:



Tilting cameras have specific requirements that should be taken into consideration:



1. Proper angle adjustment: The camera should be tilted at an appropriate angle to capture the desired field of view and maximize the visibility of license plates.
2. Avoid extreme tilting: Excessive tilting can distort the perspective and make it challenging to accurately capture license plates. It is important to find the right balance to ensure optimal image quality.
3. Consider height and distance: The height and distance of the camera from the target area play a crucial role in determining the required tilt. These factors affect the camera's perspective and coverage area.
4. Account for vehicle speed: The speed of the vehicles passing through the camera's field of view should also be considered when setting the tilt angle. Faster-moving vehicles may require a different tilt angle compared to slower ones to ensure clear and accurate license plate capture.
5. Test and adjust: It is recommended to test the camera's tilt angle in real-world conditions and make necessary adjustments to achieve the best results.

It is advisable to minimize the angle between the camera and the vehicle's direction of movement for optimal results. Ideally, the camera should be positioned directly above the vehicles at a moderate height. However, to prevent the camera from being dazzled by strong lights, it is recommended to place it above the headlights. The recommended suspension height for the camera is between 1.5 and 2 meters above the road surface.

To determine the total angle between the camera and the vehicle's direction of movement, refer to the tables provided in the tables below. It is recommended to keep the total angle below 30° for optimal performance.

Camera angles at road distance 0:

Capture Distance: Height:	5 m (~ 16 ft)	10 m (~ 33 ft)	30 m (~ 98 ft)	50 m (~ 164 ft)	80 m (~ 262 ft)
1.5 m (~ 5 ft)	17º	8.5º	2.9º	1.7º	1.1º

3 m (~ 10 ft)	31º	17º	5.7º	3.4º	2.1º
5 m (~ 16 ft)	45º	27º	9.5º	5.7º	3.6º
7 m (~ 23 ft)	54º	35º	13º	8.0º	5.0º
10 m (~ 33 ft)	63º	45º	18º	11º	7.1º

Camera angles at road distance 2m (~ 7 ft):

Capture Distance: Height:	5 m (~ 16 ft)	10 m (~ 33 ft)	30 m (~ 98 ft)	50 m (~ 164 ft)	80 m (~ 262 ft)
1.5 m (~ 5 ft)	27º	14º	4.8º	2.9º	1.8º
3 m (~ 10 ft)	36º	20º	6.9º	4.1º	2.6º
5 m (~ 16 ft)	47º	28º	10º	6.1º	3.9º
7 m (~ 23 ft)	56º	36º	14º	8.3º	5.2º
10 m (~ 33 ft)	64º	46º	19º	12º	7.3º

The red text indicates that the angle is too large for LPR.

Examples of correct camera installation:

Consider the following recommendations for proper video camera installation:

- Ensure Stability: Install the camera on a stationary structure that remains stable and free from swinging or vibration.
- Minimize Vibration: Check the reaction of roadside posts to heavy vehicles passing by and choose a location with minimal vibration to ensure optimal license plate recognition.
- Avoid bright light sources: Position the camera in such a way that it is not directly facing bright light sources like street lamps. These sources can interfere with the camera's auto-exposure function and cause glare and reflections, hindering the clarity of captured images.
- Minimize changing light sources: It is preferable to avoid having light sources within the camera's field of view that slowly change their position. Such sources can introduce variations in lighting conditions, which may affect the quality of license plate recognition.

- Eliminate interference: Ensure that there are no static or dynamic obstructions within the camera's observation cone. Trees, large signage, utility poles, or other objects can obstruct the view and interfere with license plate recognition. It is important to have a clear line of sight to the desired area of surveillance.

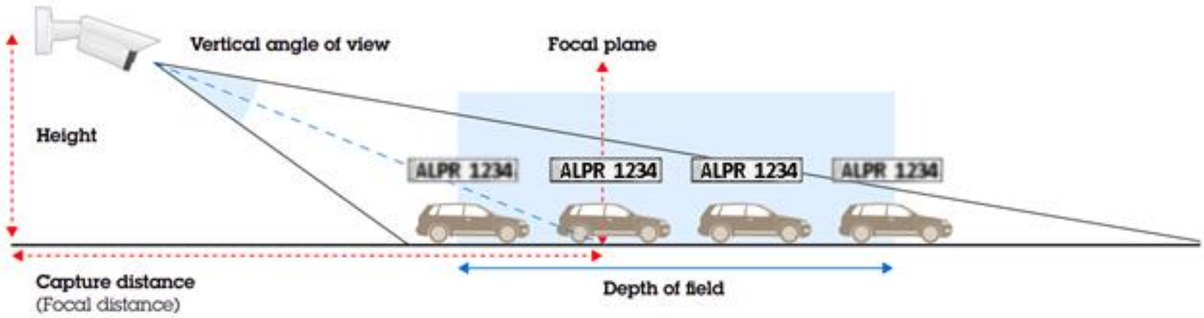
Camera aligment

The camera should be aimed at the road in a way that positions the corresponding lanes in the center of the image. The field of view should be adjusted to cover the desired number of lanes, without exceeding that range. Additionally, the camera's rotation angle should be adjusted to ensure that the license plate appears parallel to the edges of the image:

The allowable maximum swing angle deviation is illustrated in the picture below:

Depth of Field

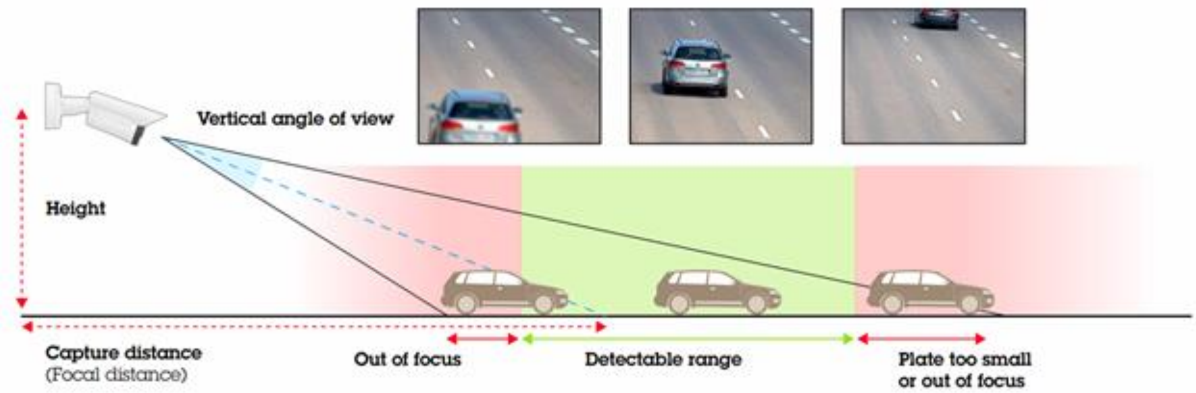
The camera must be properly focused to ensure clear and legible license plate images. However, it is important for the image to remain clear not only at a specific distance but also within a range of distances around the focal plane, as depicted in the picture below:



The extent of this range is referred to as the depth of field (DOF). Typically, the DOF is greater for longer capture distances. To increase the DOF, one can reduce the aperture size. However, this adjustment is primarily necessary for short capture distances, specifically those below 10 meters, where the DOF tends to be shallow. It is important to exercise caution when reducing the aperture, as it can limit the camera's performance in low-light conditions.

Detectable Range

The detectable range refers to the range of distances along the road where the license plate is visible and readable in the camera's image. Ideally, the detectable range encompasses the full field of view of the camera, but this may not always be achievable. The depth of field of the camera can impose limitations on the detectable range, and vehicles that are located far away may appear too small to be distinctly recognized:



Weather conditions, including snow, rain, and fog, can significantly impair visibility over long capture ranges, thereby restricting the detection range.

During daylight hours and favorable weather conditions, the detection range expands as the capture range increases. When dealing with high-speed vehicles, it is necessary to employ a longer capture distance to allow sufficient time for reading the license plate before the vehicle moves out of view.

Recommended Capture Distance

Recommended minimum capture distances for various vehicle speeds are provided in the reference table. These distances are based on an assumed detection time of 0.2 seconds, allowing the license plate recognition (LPR) software to analyze five frames per second. It's important to note that the number of frames analyzed per second may vary depending on the LPR software, processor capabilities, and image resolution. The table below serves as a general guideline:

Vehicle Speed	Recommended Minimum Capture Distance
30 km/h (~ 19 mph)	7 m (~ 23 ft)
50 km/h (~ 31 mph)	11 m (~ 36 ft)
80 km/h (~ 50 mph)	24 m (~ 79 ft)
100 km/h (~ 62 mph)	27 m (~ 89 ft)
130 km/h (~ 81 mph)	30 m (~ 98 ft)

When operating in low-light conditions, the maximum achievable capture distance is often constrained by the limited reach of infrared (IR) illumination. To extend the IR range, the utilization of more powerful external IR sources can be considered.

Video Camera Settings

Before commencing the camera installation process, it is crucial to thoroughly review the accompanying documentation provided by the selected manufacturer for the specific model. Pay close attention to the following technical details:

- Any IP camera necessitates the following components: a power cable, a network cable, a standard port connection to a router or switch, connectors, a bracket, and a power supply. It is essential to select the appropriate accessories that are compatible with your chosen model.
- If the camera is equipped with pan/tilt mechanisms, additional mechanical adjustments must be performed after installation and initial setup. Failure to do so may result in the remote pan/tilt functionalities not functioning properly.

Camera settings play a vital role in achieving optimal performance. Cameras specifically designed for capturing license plates are typically preconfigured with appropriate settings, requiring minimal adjustments. However, for other camera models, it may be necessary to modify the following settings.

Maximum Shutter Time

The maximum shutter time depends on both the camera settings and the speed of the vehicle. If the camera's shutter time is too long, it can result in motion blur when capturing moving vehicles. For example, a car moving at high speed can appear blurred in an image taken with an exposure time of 1/30 second, as demonstrated in the picture below:

A car approaching the camera will not traverse the image horizontally but rather appear larger as it gets closer. This effect is typically insignificant. However, when there is an angle between the camera and the direction of travel, the car will move laterally in the image at a speed determined by the angle. This lateral movement can cause motion blur when using a standard shutter time of approximately 1/30 second, thus necessitating a limitation on the maximum shutter time.

Recommended maximum shutter times based on the camera angle and vehicle speed (1 ms = 1/1000 s):

Camera Angle:	30 km/h (~ 19 mph)	50 km/h (~ 31 mph)	80 km/h (~ 50 mph)	110 km/h (~ 68 mph)	130 km/h (~ 81 mph)
5º	19.3 ms	11.6 ms	7.2 ms	5.3 ms	4.5 ms
10º	9.7 ms	5.8 ms	3.6 ms	2.6 ms	2.2 ms
15º	6.5 ms	3.9 ms	2.4 ms	1.8 ms	1.5 ms
20º	4.9 ms	2.9 ms	1.8 ms	1.3 ms	1.1 ms
25º	4.0 ms	2.4 ms	1.5 ms	1.1 ms	0.9 ms
30º	3.4 ms	2.0 ms	1.3 ms	0.9 ms	0.8 ms

Maximum Grain

To prevent license plate overexposure, it is advisable to limit the maximum gain of the camera. This can be achieved by adjusting the available IR intensity, considering the distance to the vehicles, and taking into account the camera's sensitivity. By finding the right balance, the license plate can be captured clearly without being excessively bright, as demonstrated in the picture below:

WDR

Wide dynamic range (WDR) techniques are employed to enhance the dynamic range of an image. This technology proves particularly valuable in revealing details that may be concealed in shadows or preventing excessive glare from overwhelming the camera.

However, it is important to note that WDR can introduce motion artifacts in images of moving vehicles, depending on the implementation in a specific camera. Unless explicitly specified in the camera specifications, we advise disabling WDR when capturing license plates.

To ensure optimal recognition, it is recommended to examine a freeze-frame of the license plate, ensuring clear and well-defined borders. This guarantees that the recognition module will have no difficulty identifying the license plate accurately.

Tone Mapping or Local Contrast

The adjustment of tone mapping plays a crucial role in enhancing detail and contrast in the darker areas of an image. For cameras equipped with an ARTPEC-6 chip or earlier, this setting can be modified through the local contrast option. However, ARTPEC-7 cameras offer improved functionality and a dedicated tone-mapping feature.

Utilizing tone mapping or adjusting the local contrast can effectively enhance the visibility of license plate letters and numbers. Nevertheless, it's important to exercise caution as setting it too high may result in increased glare and reflections from headlights, ultimately producing a gray and noisy image.

Balancing the local contrast setting on an ARTPEC-6 camera involves finding the right tradeoff between good contrast in the license plate and potential drawbacks such as increased noise and glare in the overall image:

Local contrast 0% | Local contrast 20% | Local contrast 50%

To achieve optimal results, we recommend keeping the local contrast below 50% on cameras equipped with the ARTPEC-6 chip. Additionally, disabling the color mode is advised, as a black-and-white image offers greater contrast and is more accurately interpreted by the recognition module.

Face Recognition

Introduction

Successful face recognition relies not only on the right software and hardware, but—critically—on the clarity and quality of your video stream. A high-quality, clean video feed is essential for reliable identification. Even the most advanced recognition algorithms can only work with what they see: if the video is blurry, poorly lit, or has compression artefacts, accuracy drops dramatically.
This guide provides essential instructions on selecting, installing, and configuring cameras for face recognition systems. Proper installation is critical—cutting corners may save time up front, but leads to expensive troubleshooting and unreliable performance.

Please read and follow these guidelines carefully.

Note: The quality of the reference photo used for ID Cards is just as important. Even with perfect installation, poor-quality reference images will lead to failed matches.

Reference Photo Requirements

Before installation, ensure your ID Card reference photos meet these standards:

- The face should occupy no more than 70–80% of the image.
- Minimum resolution: 150×150 pixels.
- Photo quality should match official documents: sharp focus, natural colour, good lighting, uniform background, no compression artefacts, and a neutral expression

Camera Selection: Key Specifications

When choosing a camera for face recognition, prioritise these features:

Setting	Minimum Requirement	Notes
Resolution	FullHD (2MPix), 1920x1080, 1080p	Higher resolutions *do not* guarantee better recognition—higher resolution can reduce low-light performance and slow processing.
Sensitivity	At least 0.06 lux (F1.6, IRE30, 1/30s)	Higher sensitivity is better, even in well-lit areas. Higher sensitivity allows for a faster shutter speed and more clean detection.
Lens Tilt	≤ 20° (vertical/horizontal)	Avoid severe angles—faces must be as frontal as possible.
Aperture	Minimum F1.6	Wider apertures gather more light.
Matrix Size	≥ 1/2.8"	Larger sensors perform better in varied lighting.
Shutter Speed	1/100 sec or faster	Slow shutter = motion blur (e.g. 1/25s causes ghosting on moving faces).
Frame Rate	Minimum 5 fps (12 fps recommended)	Higher frame rates improve the capture of moving faces.

Setting	Minimum Requirement	Notes
Focal length	8mm+	Wide-angle lenses introduce distortion and can affect recognition performance

Installation Best Practices

Before installing a camera for face recognition, several critical criteria must be addressed to achieve the best possible identification results:

- Installation Location: Select a spot that offers an unobstructed, clear view of the area where faces will be captured. Avoid locations with obstructions, reflective surfaces, or direct sunlight in the field of view.
- Lighting: Good, even lighting is essential. Avoid extremes—overexposed (too bright) or underexposed (too dark) scenes drastically reduce recognition accuracy.
- Camera Angle: Position the camera so faces are as frontal as possible. Avoid steep angles that distort or obscure facial features.
- Mount Height: The camera should be installed to capture faces at approximately eye level—this maximises recognition accuracy and consistency.
- Tilt Angle: Maintain a tilt (vertical and horizontal) of no more than 20°. Excessive tilt degrades alignment and facial detail.
- Lens Quality: Use cameras with high-quality lenses. Cheap or low-grade lenses introduce blur and distortion that can ruin image clarity. Example of how distortion affects face shape.

- Image Quality: Configure the camera for maximum clarity and sharpness, using high resolution and minimal compression. Clean, sharp video is non-negotiable for face recognition.

Camera Installation Requirements

Surveillance cameras are often installed high, near ceilings, for practical reasons: to avoid tampering, accidental damage, and to provide a wide field of view. However, for face recognition, eye-level installation is strongly recommended wherever possible. If higher mounting is required, strictly observe optimal angle and distance guidelines.

Installation trade-offs:

- Higher mounting: Protects cameras, but can result in top-down views and poor facial visibility.
- Eye-level mounting: Best for face capture, but may require extra protection (housing, vandal-proof domes).

Recommended camera tilt for face recognition

Area Color	Tilt Angle	Recognition Quality
Red	50º+	Low
Yellow	15º - 45º'	Average
Green	0º - 15º	High

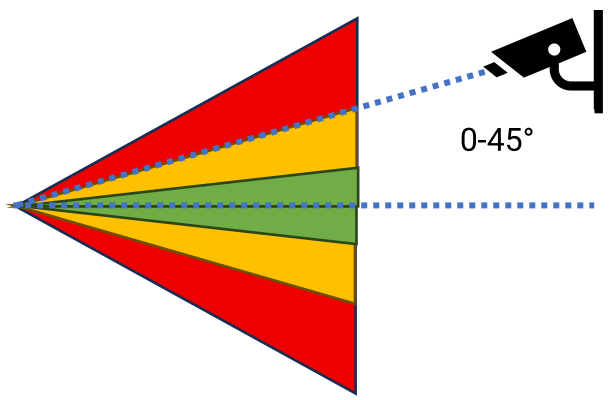


Figure 6. Recommended camera tilt angle for best results

Recommended installation heights:

Height (m)	Height (ft)	Detection Zone Range (m)	Detection Zone Range (ft)
2.0	6.5	4–8	13–26
2.25	7.4	5–8	15–26
2.5	8.2	6–8	19–26
2.75	9.0	7–8	23–26
3.0	9.8	8–10	26–33

For recognition of children, mount the camera at 1.4 metres (4.6 ft) above the floor.

Example of Full HD (1080p) camera, 15° tilt angle, 21° HFOV, 2.25m installation height:

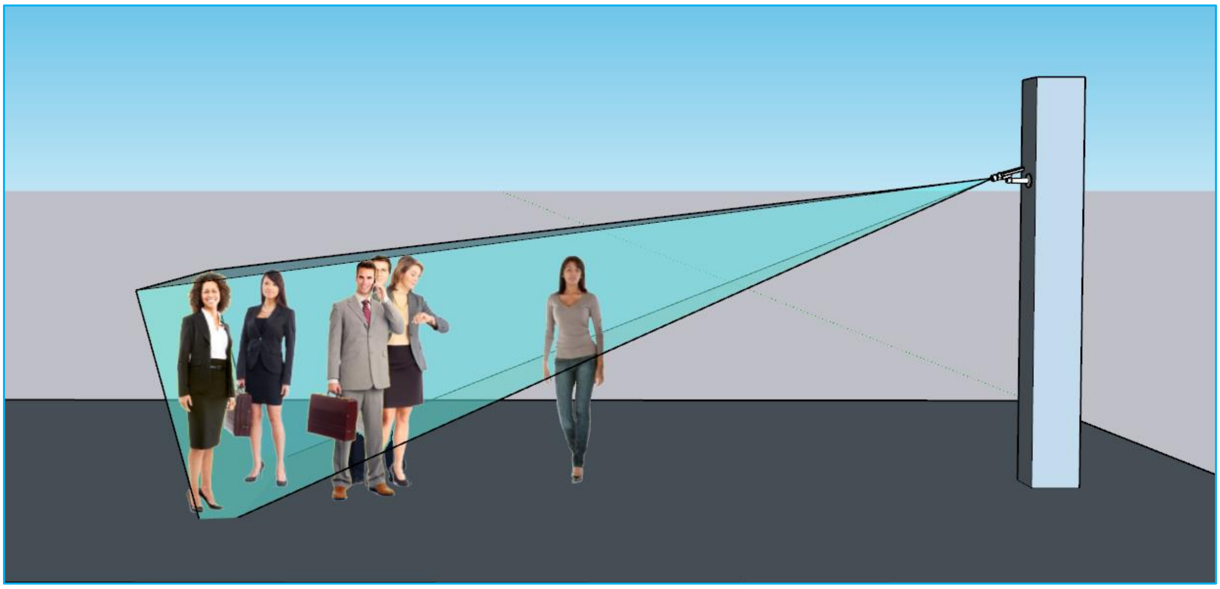


Image Quality Requirements

Image quality is absolutely critical for face recognition performance. A clean, sharp, and well-lit video feed is essential.

Key parameters:

- Sensor Sensitivity: Higher sensitivity improves clarity, especially in low light.
- Resolution: Higher is generally better, but only if lens and lighting are sufficient.
- Shutter Speed: Fast shutter speeds (≤ 1/100 sec) prevent motion blur and ensure sharp frames.

How to check if your image is good enough

Interocular Distance (IOD) and Pixels per Meter

- IOD (Interocular Distance) = distance between the centres of a person's pupils (in pixels, as measured in an image)
- Pixels per Meter (ppm) = how many pixels in your image correspond to 1 metre in the real world, at the position of the face

To calculate IOD:

1. Save a video frame and open it in a graphics editor.
2. Draw a line between the person's pupils.
3. Measure the distance (in pixels)—this is the interocular distance (IOD).

- Minimum IOD for face detection[1]: 10 pixels.
- Minimum IOD for face recognition[2]: 21 pixels (recommended: 40+ pixels).
- If IOD < 10 pixels, adjust camera placement or settings.

Formula for Pixel per meter (PPM)

$$\textit{Pixel per meter}\ [PPM] = \frac{IOD\ [pixels]}{IOD\ [meters]}$$

- Example:
  If IOD (pixels) = 21
  Real-life IOD = 63 mm = 0.063 m (average adult human IOD)

Pixels per meter≈333

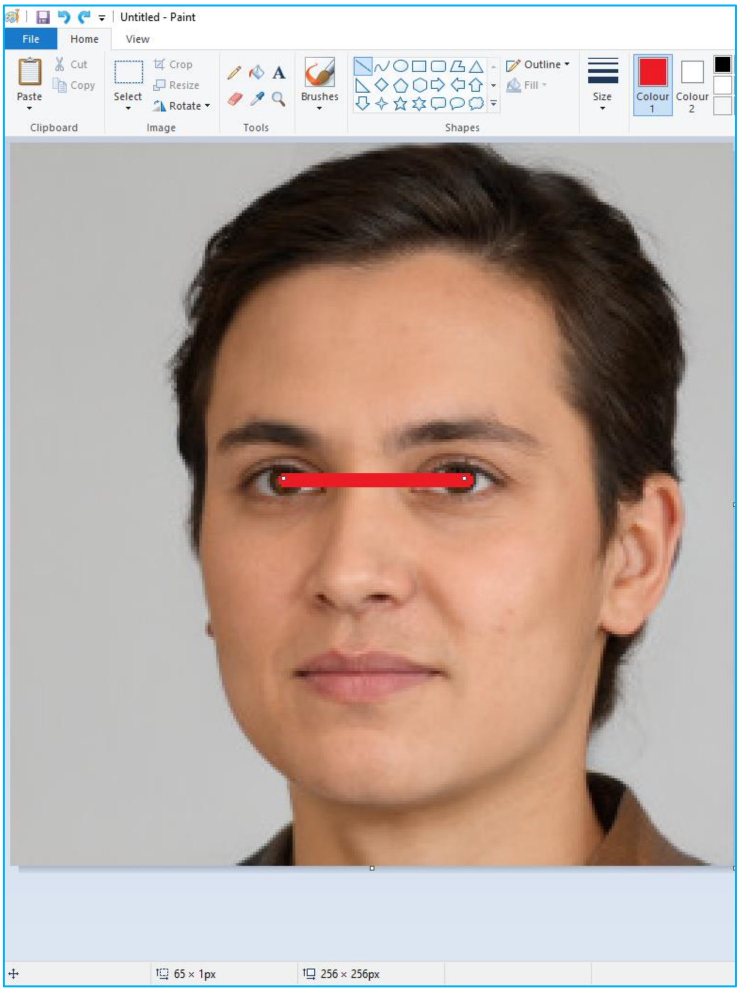
It is therefore recommended that a minimum of 333 pixels per metre of the scene be provided for a face recognition solution.

It can consequently be concluded that a camera with a resolution of 1920x1080 is capable of monitoring a scene that is 5.8 metres wide, or alternatively, five individuals in a single instance.

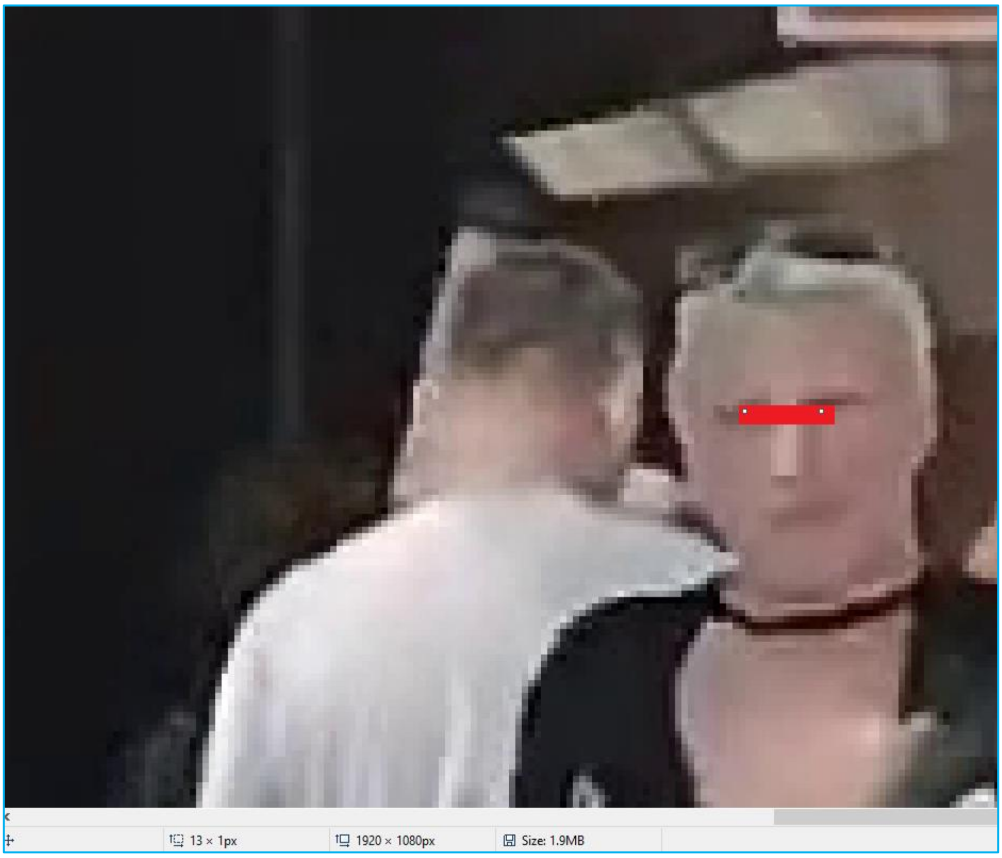
In instances where a wider area requires coverage or the capture of additional subjects, the installation of additional cameras is recommended

[1] **Face Detection** is the process of simply finding and locating human faces within an image or video frame. It answers the question: *"Is there a face in this picture, and where is it?"* The system draws a box around the detected face but does not attempt to identify who it belongs to.

[2] **Face Recognition** goes a step further. After a face is detected, recognition analyses the facial features and compares them to a database of known faces (such as a list of employees or registered visitors). It answers: *"Whose face is this?"* Face recognition is used for identification, verification, or access control.



Example of incorrect angle and IOD



Shutter Speed Adjustment

Blurry images = poor recognition. To keep faces sharp:

- Switch from auto to manual shutter mode.
- Set shutter speed to 1/100 sec or faster.

Compression Settings

- Minimise video compression. Set compression to as low as possible (ideally 0–10%). Excessive compression creates artefacts that destroy critical face details.
- Avoid advanced compression features like zip streaming or H.265+, AI H.265, SmartH.264 etc for face recognition applications.
- When using H.264/H.265, always test to ensure compression isn't hurting image quality.
- More bandwidth is better: higher bitrate = cleaner image.
- Recommended I-Frame interval is 1-2 per second or more for more dynamic scenes.

Lighting Requirements

Even, front-facing light is ideal. Lighting below 150 lux makes recognition unreliable. Best practices:

- Use overhead, diffuse lighting for consistent results—even when people move.
- Avoid harsh shadows and backlighting.

- In challenging environments (e.g., parking, entrances), add supplementary lights as needed.
- Check camera performance at different times of day—natural light changes can ruin recognition if not accounted for.

Camera Positioning Requirements

Install cameras where people naturally slow down (gates, turnstiles, doors). The system requires 0.2–1 second of clear face view for recognition.

In corridors, ensure faces are visible for at least 2–3 seconds in the frame, with good lighting.

Minimise visual noise: Keep backgrounds simple and free from:

- Info displays
- Structural supports
- Trees, plants, or moving vehicles
- Large groups or crowding

Aurora

What is Aurora?

Prompt Structure

Good prompts are specific, focused, and grounded in observable features. Avoid ambiguity or multiple questions. It is highly recommended to use English for prompt creation.

Goal	Good Prompt	Why It Works	Bad Prompt	Why It Fails
PPE Compliance	"Is the person wearing a helmet?"	Simple yes/no	"Is the person dressed properly?"	Too vague
Object Presence	"Is the person holding a bag?"	Clear binary	"What is the person holding?"	Open-ended, higher hallucination
Vehicle Check	"Is the tarp on the trailer closed?"	Yes/no, visual context	"Is the truck safe?"	Ambiguous, subjective
Scene Context	"Is anyone lying on the floor?"	Binary, precise	"Is the person lying down or sitting on floor?"	Multiple options = unclear result
Intrusion Detection	"Is there a person in the area?"	Defined action	"Is this an intrusion?"	Needs policy/context to interpret
Workplace Safety	"Is the worker using a safety harness?"	Binary compliance check	"Is the worker behaving safely?"	Behaviour is subjective

Object Detection	"Is there a pallet on the floor?"	One object, one location	"Is there anything unusual?"	Too abstract

🔁 Best Practice

- Rephrase descriptive questions into confirmatory ones.
  ❌ *"What is the person doing?"*
  ✅ *"Is the person lying on the ground?"*
- Avoid lists or compound prompts. Break them up:
  ❌ *"Is the person wearing a helmet and gloves?"*
  ✅ *"Is the person wearing a helmet?"* then ask again for gloves.
- Anchor the question to scene elements.
  ✅ *"Is the person on the left holding a red object?"*

Prompt-Review Checklist (operators must tick all before deployment)

1. Written in English, ≤ 15 words, starts with Is / Are.
2. Targets one object or condition – no lists, no commas.
3. Avoids protected traits (race, religion, age, gender, disability, nationality).
4. Tested in the Aurora Tester APP utility and returns the expected YES/NO.

Aurora Tester app – lets you paste a single snapshot, type your question, and preview Aurora's answer before going live. Always validate every new question and every new camera scene.

Crafting Prompts with GenAI LLM (ChatGPT, GROK, META) agent

Aurora accepts English questions only.
If you are unsure how to phrase a concise YES/NO question, let some GenAI agent draft it for you:

Step-by-step

1. Open ChatGPT and paste the template below.
2. Paste *once* if you have just one item, or replace [LIST] with multiple items to get a separate question for each.
3. Copy the single-line output verbatim into Aurora's prompt field.

Template:

You are building ONE atomic YES/NO question for an AI that analyses a single CCTV image.

CONSTRAINTS

– Output must be ≤15 words, start with "Is" or "Are", no commas.

– Verify ONE visual item or condition only.

– Use plain visual nouns, not compound labels .

– Do not add commentary, lists, or extra text — return ONE sentence only.

CONTEXT

A snapshot is attached so you can understand the scene.

TASK

Create the question(s) to verify: [LIST YOUR ITEM(S) HERE].

Return each question on its own line.

Return the question now.

Example:

Prompt:

You are building ONE atomic YES/NO question for an AI that analyses a single CCTV image. CONSTRAINTS – Output must be ≤15 words, start with “Is” or “Are”, no commas. – Verify ONE visual item or condition only. – Use plain visual nouns, not compound labels . – Do not add commentary, lists, or extra text — return ONE sentence only. CONTEXT A snapshot is attached so you can understand the scene. TASK Create the question(s) to verify: facemask, apron, gloves, person standing Return each question on its own line. Return the question now.

Answer:

Is the person wearing a facemask?
Is the person wearing an apron?
Is the person wearing gloves?
Is the person standing upright?

One-Frame Limitation & Action-Based Alarms

Aurora operates on a single still image — either the full frame, cropped region, or bounding box. It does not evaluate motion or time-series data. This imposes important limitations for action-based scenarios.

❗ Avoid using Aurora alone for:

- Fight detection
- Car crashes
- Acts of vandalism
- Any dynamic physical interaction
- Any action that needs multiple frames to understand:
  For example:
    - Running vs walking vs standing still;
    - Direction of movement;
    - Moving or placing an object;
    - How fast an object is moving(slow vs fast);
    - ...

🧠 A still frame may miss the nuance or context (e.g. someone falling vs lying down vs stretching), leading to:

- Hallucinated descriptions (false positives)
- Missed events (false negatives)
- Confusing or vague analysis

For such cases, combine Aurora with TREX, Unusual Behaviour or other frame-sequenced analytics, where these analytics can solve the problem.

Text, Numbers, Signs

Aurora can understand text, gauges, and printed signs in context, but it is not an OCR engine.

Do not use it for:

- Number plate recognition (use LPR module instead)
- Passport/ID reading
- Serial numbers

However, it can answer:

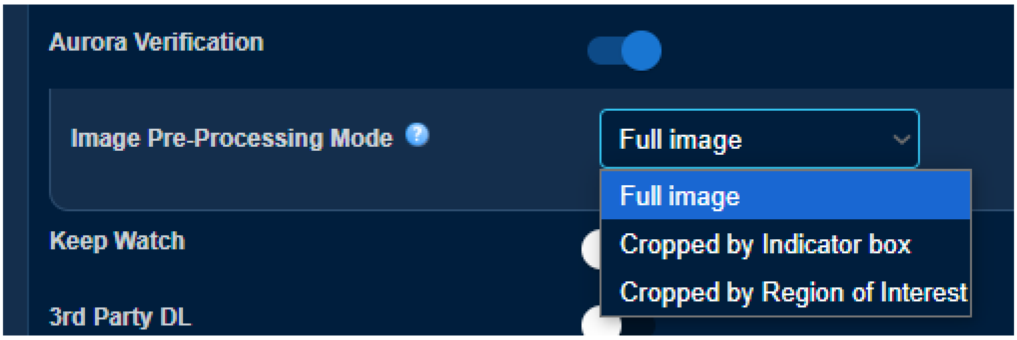
✅ "Is the gauge needle pointing to 120?"
✅ "Is there any text on the side of the truck?"
✅ "Can you read the logo or name on the uniform?"

For these tasks:

- Use Region of Interest or Full Image Mode.
- Ensure at least 12–16 GB of VRAM is available for full-frame parsing with text.

Aurora Image Pre-Processing Mode

Aurora supports three modes of operation depending on the alert context and performance needs



Cropped by Indicator box (Fastest)

- Aurora focuses only on the object's bounding box (e.g., a detected person).
- ✅ Best for: Simple classification (helmet, carrying object).
- ❌ Avoid when: Context matters (e.g., are they interacting with others?).

Region of Interest (Balanced)

- Aurora analyses the region surrounding the triggered alarm (e.g., TW line, Left Object zone).
- ✅ Best for: Scene-specific checks like direction, entry/exit, object position.
- ❌ Avoid when: Subject extends beyond margin or action occurs in another area.

Full Image Mode (Slowest but Richest)

- Analyses the entire image frame, including all objects and interactions.
- ✅ Best for: Complex scenes like vehicle loading, crowd dynamics, fire/smoke detection.

- ⚠ Requires significantly more GPU time and memory — not recommended for rapid high-volume alert environments. No focus on cause of the alert or alert object, which could cause vague or non-specific responses.

Choosing the Right Zoom & Processing Mode

Use Case Example	Recommended Mode	Why
"Is the person holding a knife?"	Cropped by Indicator Box	Precise focus on hands/arms
"Is the trailer door open?"	Full Image	Entire truck context needed
"Is anyone inside the restricted zone?"	Cropped by ROI	Alert zone context is important
"Is the person wearing high-vis clothing?"	Cropped by Indicator Box	Clear bounding box suffices
"Are there pallets left on the dock?"	Full Image	Scene-wide object placement

Tip: Use *iSentry Spotlight* to create virtual cameras or zoomed-in views to optimise scene clarity for Aurora, especially when verifying small items like phones, tools, or labels

System Performance Considerations

- The larger the scene analysed, the more GPU and RAM Aurora consumes.
- For high-volume deployments (e.g., 7500+ alerts/day), use bounding box/margin modes for speed.
- For detailed audit scenarios or complex scenes, schedule full-frame Aurora checks during off-peak hours.

Camera Requirements & Image Quality

Aurora's output quality depends on what it can see.

To ensure reliable interpretation:

- Use cameras that provide sufficient Pixels Per Metre (PPM) for the intended classification task.
- Avoid extremely high-angle overhead views in general, particularly when looking at face orientation, objects in hands, PPE, etc.
- For text or fine detail (e.g., printed labels), ensure the image is:
    - High image quality;
    - sharp
    - well-lit
    - zoomed-in if needed (consider Spotlight)

If the object is too small, occluded, or poorly illuminated — Aurora will fail or hallucinate.

Object-type examples	Suggested pixel-per-metre (ppm) at target area	Details
Knife, small hand-tool, safety glove	≈ 120 ppm	Gives ~24 pixels across a 20 cm object—enough edge detail for Aurora to confirm presence/absence.
Wheel-chock, fire extinguisher, hi-vis vest	50 – 70 ppm	Medium-sized items remain clearly outlined.
Pallet, trailer door, car body	25 – 35 ppm	Scene context is more important than fine detail.

Because the *absolute* number of pixels Aurora can dedicate to the object drops rapidly as the object shrinks in the frame. If a knife spans only 8 pixels, edge information is lost and the model will either hallucinate or fail to detect it. A car, meanwhile, fills hundreds of pixels even at very low ppm, so the model still sees enough detail.

Prohibited and Restricted Use

The following uses of Aurora are strictly prohibited and must never occur under any circumstances. Failure to comply may result in legal, regulatory, or contractual breaches.

You must not use Aurora for:

- Serving as the sole source of legal, security, or safety evidence without independent human validation.
- Making autonomous decisions or taking direct actions without a qualified human in the loop (e.g., automatic escalation, sanctions, or enforcement based solely on Aurora's output).
- Behavioural or psychological inference (e.g., “Is this person stealing?”, “Is this person angry or agitated?”).
- Operation in low-light, heavily obstructed, or visually degraded scenes where accurate interpretation is compromised.
- Profiling, identifying, or making judgements based on race, gender, age, religion, disability, nationality, or any other protected characteristic.
- Any activity that exploits the vulnerabilities of specific groups due to their age, social, physical, or psychological characteristics, in a manner that could cause or is likely to cause physical or psychological harm.

Important:

- A qualified human operator must always review and verify Aurora’s outputs before taking any security, safety, legal, or compliance action.
- Direct automation of critical decisions or policies (such as access denial, reporting, or enforcement) based solely on Aurora alerts is strictly prohibited.

Ethical & Bias Disclaimer

Aurora’s outputs are generated based on probabilistic models trained on diverse but imperfect datasets. As such:

- Aurora may reflect inherent biases from the training data.
- It should not be used as sole evidence for disciplinary or legal actions.
- Users must verify conclusions through multiple sources or human review.
- Discriminatory or profiling prompts are strictly prohibited.

Intelexvision assumes no liability for improper or discriminatory use of Aurora results.
