Road User Detection

Performance & Operating Guide

PREPARED BY	Spectro-AI	SUBJECT	Road user detection algorithm
RUNS ON	DJI Matrice 4 Series (on-board)	MORE INFO	spectroai.ai
ALGORITHM	TrafficEye-SAI-V1	DOCUMENT	v1.0 · Public
DATE	07 June 2026	MORE INFO	spectroai.ai

Contents

OPERATIONAL SAFETY The algorithm is an operator-assistance tool and does not replace the pilot's responsibility to maintain safe flight or comply with applicable aviation regulations. All detections should be verified before operational decisions are made.

INTENDED USE This algorithm is designed to assist operators in locating road users during aerial observation. It can highlight areas requiring further investigation, but should not be used as a replacement for visual verification, operator assessment or any procedure your organisation applies.

OBSERVING PEOPLE: PRIVACY AND PROPORTIONALITY This algorithm detects people and vehicles from a visible-light camera, at heights where individuals, faces, clothing and number plates may be identifiable in the recorded footage. That makes any flight a form of personal data collection, and in most jurisdictions it is regulated as such. Before deploying it, confirm your legal basis, any notification or signage obligations, how long footage and detections are retained, who may access them, and whether the footage requires blurring or other minimisation. A detection is not an identification and must never be treated as one.

1 About this guide

This guide describes how the Spectro-AI road user detection algorithm performs when running on-board a DJI Matrice 4 Series aircraft, and how to fly and configure the camera to obtain the most reliable results. It is written for drone operators and field users rather than for specialists.

The algorithm runs directly on the aircraft and reports detections in real time during the flight through the DJI Pilot 2 application. It works from the visible-light (RGB) camera and recognises seven classes of road user, from pedestrians through to long vehicles.

Its purpose is to give an operator a faster read of a scene than watching a video feed unaided: marking where people and vehicles are in the picture so that attention can be directed to the right part of it. Scenes of this kind are typically busy, and the algorithm will often place tens of boxes in a single frame.

A DETECTION AID, NOT A DECISION The algorithm supports human judgement. It does not replace visual verification, operator assessment, or any procedure your organisation applies. Every detection is a prompt for a person to look: never a conclusion in itself.

Performance figures presented in this guide were established under validated conditions and should be used as operational guidance rather than guaranteed performance. Detection accuracy varies with flight height, camera angle, lighting, weather, and above all with how crowded the scene is.

2 Supported aircraft and camera configuration

This is a visible-light algorithm. It uses the RGB camera, not the thermal payload, and therefore runs on every aircraft in the Matrice 4 Series: the Matrice 4E, 4T, 4D and 4TD. On the thermal variants it still works from the visible camera and does not read the thermal feed.

Table 2.1 *Aircraft compatibility*

Aircraft	Status	Reason
DJI Matrice 4E	**Supported**	RGB camera fitted
DJI Matrice 4T	**Supported**	RGB camera fitted

Aircraft	Status	Reason
DJI Matrice 4D	**Supported**	RGB camera fitted (dock variant)
DJI Matrice 4TD	**Supported**	RGB camera fitted (dock variant)

DAYLIGHT ALGORITHM There is no thermal palette to set and no palette constraint of any kind, but there is a lighting constraint, and it is the equivalent one. The algorithm needs light. It works best in daylight and in good visibility; in poor light and darkness it may struggle or fail to function altogether, and it has no thermal capability to fall back on.

3 Detection classes

The algorithm reports seven classes covering the range of road users. Each detection appears on the live image as a box, a label and a confidence score, and every object in the image is boxed separately.

Table 3.1 *Detection classes*

Group	Classes	Size, and how easily they are found
Large vehicles	**Long-vehicle**	Buses, trucks, vans and similar. The easiest class: large, distinctive and detected from the greatest heights.
Cars	**Car**	Detected reliably through most of the height band. In the validation imagery these carry among the highest confidence scores.
Three-wheeled vehicles	**Tricycle · Awning-tricycle**	Light three-wheeled vehicles, with and without a canopy. Common in some regions and rare in others; where they are absent from a scene, these classes simply do not appear.
Two-wheeled vehicles	**Motorcycle · Bicycle**	Small and narrow, and often close to the rider. Harder than cars, and frequently confused with one another and with the person riding them.
People	**Person**	The smallest class and the hardest. Pedestrians are small from height, frequently overlap one another in crowds, and are often partly hidden by vehicles or street furniture.

THE PRINCIPLE BEHIND ALL SEVEN Larger objects are easier to detect than smaller ones, and isolated objects are easier than crowded ones. Long vehicles and cars sit at the easy end; motorcycles, bicycles and pedestrians at the hard end. Everything in Section 5 follows from this.

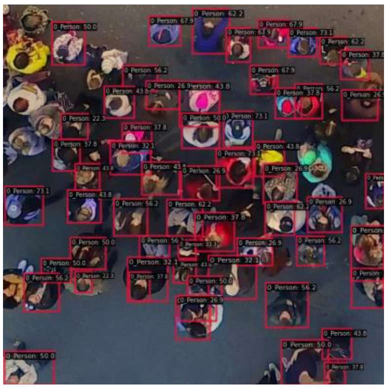
Crowding is the defining difficulty

This algorithm is unusual in the Spectro-AI range for the sheer number of objects it is asked to find at once. In the validation imagery, single images commonly contain thirty to forty detections, packed close together and frequently overlapping.

That density is the source of most of the errors an operator will see. Where road users stand or move close together, neighbouring objects merge into a single box, a partly hidden object is missed entirely, and small classes are lost against the larger ones beside them. It is also why confidence scores in this algorithm sit lower than in the rest of the range.

> **DO NOT TREAT DETECTION COUNTS AS SURVEY FIGURES** It is tempting to read the number of boxes on screen as a count of road users. Do not. Crowded scenes are exactly where objects merge into one another and where partly hidden ones are missed, so the count will generally be an undercount, and the busier the scene, the greater the shortfall. The output is a map of where road users are, not a traffic survey figure, and it should not be used for counting, flow measurement or any statistic without independent verification.

Table 3.2 *Example detections*

Detection (with bounding boxes)	What it shows
	**A busy scene at mixed scale** Around forty detections in a single frame, ranging from large vehicles filling much of the picture down to pedestrians only a few pixels across. This is the algorithm working at full stretch, and the condition most scenes will present.
	**Pedestrians in a group** Numerous Person detections at close spacing, with scores ranging widely: some in the seventies, others below thirty. Groups like this are where boxes merge and where the count becomes unreliable.
	**Vehicles and two-wheelers** A quieter scene with cars and motorcycles detected. Note the contrast in confidence: the car is scored well above the motorcycles beside it, which is the size effect in plain view.

Detection (with bounding boxes)	What it shows
	**Mixed road users** People and vehicles detected together across a range of distances. Scores fall away steadily as objects recede: the practical reason for the height guidance in Section 5.

4 Detection performance

Validation was performed on an independent dataset that was held out from training and evaluated separately. The results below describe performance under the validated conditions and should be regarded as estimates of expected performance rather than guarantees.

0.537	0.317
mAP@0.50	mAP@0.50–0.95

What these numbers mean

- **mAP@0.50 (0.537).** The headline measure of whether road users are found and correctly located. This is the lowest figure in the Spectro-AI on-board range, and it deserves to be stated plainly rather than buried: expect this algorithm to miss more than the others do.
- **mAP@0.50–0.95 (0.317).** A stricter measure that also rewards how precisely each box is drawn. In dense scenes, where boxes overlap and objects touch, drawing a tight box around the right object is hard: this figure reflects that difficulty as much as anything else.

WHY THE FIGURES ARE LOWER HERE This is a harder problem than the rest of the range, not a weaker piece of engineering. The algorithm is asked to separate seven classes at once, in scenes containing tens of objects packed close together, where road users routinely overlap, hide one another behind vehicles and street furniture, and appear at wildly different sizes within the same frame. Crowding and occlusion are the dominant causes of the shortfall. Used with that expectation, the algorithm is useful; used as though it were exhaustive, it will disappoint.

Broken down by how large objects appear in the image, the pattern is consistent with everything above.

Table 4.1 *How reliability changes with apparent size*

How it appears in frame	Reliability	What this means in the air
Large	**Best**	Long vehicles and cars at working height. Found consistently and scored highest.
Medium	**Moderate**	Two- and three-wheeled vehicles, and pedestrians at closer range. Generally found, with regular misses in crowded scenes.

How it appears in frame	Reliability	What this means in the air
Small	**Weak**	Distant pedestrians and two-wheelers. Reliability here is roughly half that on large objects. Expect substantial misses, and fly lower if these matter.

Per-class figures were not measured for this build, so these results should not be read as applying equally to all seven classes: the class table in Section 3 sets out where the differences lie. Precision and recall at a single fixed confidence threshold were likewise not measured.

> **CONFIDENCE-THRESHOLD SLIDER** Start at approximately 0.30. This is lower than most of the Spectro-AI range and deliberately so: in the validation imagery, correct detections are routinely scored in the twenties and thirties, so a threshold set to make the display look tidy will discard a large share of genuine road users. Lower it further when the priority is to miss as little as possible and you are prepared to check more false alarms. Raise it when the screen is so busy that the boxes are no longer readable, but understand that you are trading away real detections when you do.

Common causes of missed detections

- **Crowding and occlusion.** The dominant cause. Road users standing or moving close together merge into single boxes, and anything behind a vehicle, a structure or another person is not seen at all.
- **Small and distant objects.** Pedestrians and two-wheelers lose detail quickly with height. This is the second biggest cause and the one the operator has most control over.
- **Poor light.** Dusk, night, deep shadow and heavy overcast all reduce contrast. There is no thermal fallback.
- **Unusual viewing angles.** Views outside the validated range present shapes the algorithm has not been trained on.
- **Motion blur.** Fast-moving traffic combined with aircraft movement softens the image and costs detections.

5 Operating recommendations

Flight height and camera angle

Table 5.1 *Operating envelope*

Parameter	Setting	Notes
Camera	**Visible (RGB)**	The algorithm reads the visible feed. It does not use the thermal camera, even where one is fitted.
Flight height	**10 – 100 m above ground**	A wide band. Lower gives larger, clearer road users and better results on people and two-wheelers; higher covers more of a junction or corridor at once.
Camera angle	**Nadir to approximately −30°**	Straight down through to a distinctly angled view. Both are represented in the validated imagery.
Preferred angle	**Near-nadir for counting and coverage**	A top-down view gives the most even coverage and the least overlap between

Parameter	Setting	Notes
		objects: which matters more here than in any other algorithm in the range. Use angled views to see along a street or under an obstruction.
Lighting	**Good daylight**	Night and low light are not validated and there is no thermal fallback.
Confidence threshold	**0.30 to start**	Adjustable in DJI Pilot 2. Tune to the scene and to how much checking you are prepared to do.
Zoom	**[Insert validated zoom settings]**	Not recorded in the validation imagery. Optical zoom may be used to enlarge distant road users.

HOW TO USE THIS Decide before the flight which road users matter. If long vehicles and cars are the objective, work the upper end of the height band and cover far more ground per pass. If pedestrians, cyclists or motorcyclists matter, come down: they are the classes that fail first, and height is the lever that recovers them. Prefer a near-nadir view wherever the scene is busy, because overlap between objects is this algorithm's principal weakness and looking straight down minimises it.

Angle, perspective and crowding

- **Nadir separates objects; oblique stacks them.** Looking straight down, two road users standing a metre apart appear a metre apart. Viewed at an angle, they may overlap in the picture entirely. In a crowded scene this difference decides whether both are detected or only one.
- **Use the angled view deliberately, not by default.** An oblique view is useful for seeing along a street, under a tree line or beneath a canopy, and for showing the faces of vehicles. It is the wrong choice for surveying a busy junction.
- **Objects shrink up the image.** In an angled view, road users near the top of the picture are much further away and much smaller than those at the bottom. Expect detections to thin out towards the top of the image.
- **Move rather than zoom where you can.** Optical zoom genuinely resolves more detail. Digital zoom only enlarges what has already been captured. Flying lower is better than magnifying from height.

Lighting and weather

- **Best conditions.** Bright, even daylight with the sun high enough that shadows are short. Overcast but bright is often better than harsh low sun.
- **Shadow and glare.** Long shadows from buildings and vehicles hide road users and create dark shapes that can be flagged in error. Low sun straight into the lens washes out the picture entirely.
- **Wet roads.** Standing water and wet asphalt reflect sky and headlights, obscuring road users and producing false detections.
- **Rain, mist and spray.** All reduce contrast and put droplets on the lens. Check the lens after any wet flight.

- **Night.** Not validated. This algorithm has no thermal capability and should not be relied upon after dark.

Lens cleanliness

Dirt, water droplets or condensation on the camera lens soften the image and may degrade detection performance, and a pedestrian only a few pixels across is exactly what a soft image loses first. Inspect the lens before each flight and clean it as required, and allow time for condensation to clear when a cold aircraft is brought into warmer, more humid air.

6 Limitations

Each limitation below is paired with what the operator should actually do about it.

Table 6.1 *Limitations and operator response*

Limitation	What to do
Crowding and overlap	The principal limitation. In busy scenes, neighbouring road users merge into single boxes and hidden ones are missed. Fly near-nadir to minimise overlap, come lower, and never read the number of boxes as a count of people or vehicles.
Detection counts are not survey figures	Counts will generally understate what is present, and the busier the scene the greater the shortfall. Do not use the output for counting, flow measurement or any published statistic without independent verification.
Objects hidden behind others	Pedestrians behind vehicles, road users under trees, canopies or bridges, and anything inside a vehicle cannot be seen. Work the scene from more than one position rather than assuming one pass has covered it.
Small and distant road users	Reliability on small objects is roughly half that on large ones. Descend or use optical zoom rather than lowering the threshold: no confidence setting recovers detail the camera never captured.
Similar classes are confused	Motorcycles, bicycles and their riders are readily mistaken for one another, and the three-wheeled classes for each other. The label may be wrong while the detection is still correct: confirm what an object is before acting on the class name.
Poor light and night	Detection may be reduced in low light, deep shadow or heavy overcast, and night operation is not validated. Fly in daylight; there is no thermal fallback.
Unusual viewing angles	Detection may be reduced outside the range in Table 5.1: in particular from a near-level view. Keep the camera between nadir and approximately −30°.
Regional variation in vehicle types	The three-wheeled classes are common in some regions and rare in others, and local vehicle types not represented in training may be classed incorrectly or missed. Run a short familiarisation flight on unfamiliar ground before relying on the output.
It reports a class, not an identity	A Person detection means a person-shaped object was seen. It says nothing about who. Never treat a detection as an identification, and follow your obligations on personal data.

Limitation	What to do
An absence of detections proves nothing	It is not evidence that a road or area is clear. Use your own eyes on the feed and follow your normal procedure regardless of what the algorithm has reported.

7 Operator workflow

Before flight

- **Camera set to visible (RGB).** Confirm the visible camera is producing a live image. The algorithm does not read a thermal feed.
- **Lighting assessed.** Good daylight, and note where the sun is. Long shadows and low sun into the lens both cost detections.
- **Lens inspected and clean.** Dirt and condensation cost small targets first.
- **Condensation allowed to clear.** If the aircraft has come from a cold vehicle into warm, humid air, give the optics time to settle.
- **Scene density considered.** A busy junction will produce far more misses than a quiet street. Set expectations before the flight, not after it.
- **Confidence threshold set.** 0.30 to start; lower if the priority is to miss as little as possible.
- **Height and angle planned.** Within 10–100 m, nadir to approximately −30°. Choose according to what you are looking for: lower for pedestrians and two-wheelers.
- **Permissions confirmed.** Where the flight will observe people, confirm your legal basis and any notification obligations have been met.

During flight

- **Hold the planned height.** Climbing shrinks pedestrians and two-wheelers below what the algorithm can use.
- **Keep the camera within the validated angle range.** Near-nadir wherever the scene is busy; angled only when you need to see along a street or under an obstruction.
- **Fly at a steady, moderate speed.** Motion blur costs detections, and moving traffic supplies enough of it already.
- **Work in overlapping passes.** Objects hidden on one pass may be visible from another bearing.
- **Watch the scores as well as the boxes.** Consistently low scores across a pass usually mean too much height, poor light, or a scene too crowded to separate.
- **Re-tune between scenes.** A threshold that suited a quiet street will flood the screen at a busy junction, and vice versa.

After a detection

- **Verify on the live image.** Zoom in and look at what is actually inside the box before treating it as anything.
- **Zoom in before judging.** A box a few pixels across tells you very little until you look at it properly.
- **Do not rely on the label alone.** The class may be wrong while the detection is still real, and a warm mark is not an identification.
- **Search around the detection.** Where one road user has been found, look for others nearby that may be smaller, partly hidden or merged into the same box.
- **Follow your organisation's procedure.** The algorithm informs that procedure; it does not replace it and does not authorise any action on its own.

- **Record the conditions.** Note height, angle, light and how busy the scene was alongside anything you keep. It explains the result later.

> **BUILD YOUR OWN RANGE CARD** No validated detection distances are published. On an early flight over ground you know, note the heights at which pedestrians stop being detected reliably, and separately the heights at which cars and long vehicles do. The gap between those two numbers is large, and knowing it for your own aircraft is what lets you choose a survey height with confidence. Repeat the exercise on a busy scene and a quiet one: the difference will be instructive.

8 Deployment

The algorithm is licensed to the individual aircraft serial number and delivered through DJI Pilot 2. Once activated it runs entirely on-board: all processing happens on the aircraft, and no network connection is required during operation.

Table 8.1 *Deployment summary*

Item	Detail
Where it runs	On-board the aircraft, in real time, from the visible camera feed.
Network connection	Not required in flight. Imagery is not sent off the aircraft for processing, which also means detections are not transmitted anywhere by the algorithm itself.
Licensing	Tied to the aircraft serial number.
Version identification	The algorithm name and version are shown in the DJI Pilot 2 interface. Note that the algorithm version and the version of this guide are numbered separately: record the algorithm version shown on the aircraft alongside any results you keep, so that findings can be matched to the build that produced them.
Firmware compatibility	Keep aircraft firmware and DJI Pilot 2 at compatible versions. Updating one and not the other is the most common cause of an algorithm failing to load. [Insert minimum firmware / DJI Pilot 2 version.]
If it does not load	Confirm the licence is active for that airframe, check firmware and Pilot 2 versions, and restart the aircraft and controller. If it still does not appear, fly the task using the live camera feed directly and report the fault: the algorithm is an aid, and its absence should not stop the sortie.
If detections look wrong	Check height and angle against Table 5.1 first, then the light, then the lens, then the threshold. Most reported problems are a height problem or a crowding problem rather than a fault, and if the scene is dense, some level of missed detection is expected behaviour rather than a malfunction.

> **VALIDATED VERSION** The validated performance described in this guide applies to the released algorithm version identified on the cover of this document, TrafficEye-SAI-V1. Later algorithm updates, aircraft firmware changes and different camera configurations may perform differently, and the figures here should not be assumed to carry over to them. Check the algorithm version shown in DJI Pilot 2 against the cover of this guide, and record it alongside any results you keep.

Questions, or results in conditions not covered here? Visit spectroai.ai.