Deep Learning Thermal Tracker (DLTT) Technical Note

Information applies to VCAserver 2.4.0.

Document version v0.1

Introduction 3
Applications 3
Limitations 3
General Guidelines 4
Camera Position and Angle 4
Obstructions 4
Detection Range 4
Frequently Asked Questions 5
Does the Deep Learning Thermal Tracker (DLTT) improve during operation? 5
Is calibration still required? 5
What image settings should I use on the camera? 5
Setup 6
General 6

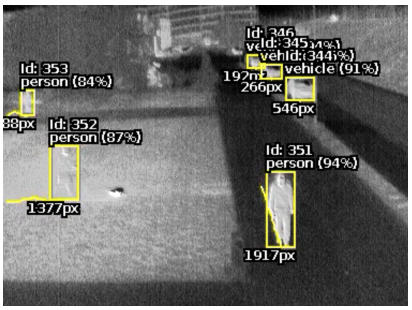
Introduction

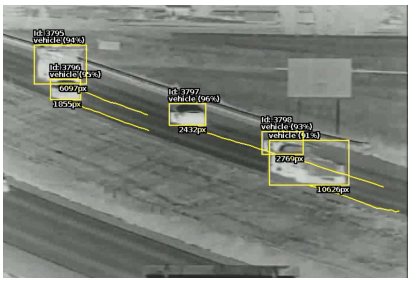
This technical note details the best practices for using the Deep Learning Thermal Tracker (DLTT) along with some answers to common questions about when to use it.

Applications

The Deep Learning Thermal Tracker (DLTT) is designed to detect, track and classify a variety of objects in a scene using a pre-trained model.

- Security
  - Perimeter protection
  - Areas where access is restricted
  - Remote site security
  - Sterile zone
  - Object detection (i.e, Vehicle, person)





Limitations

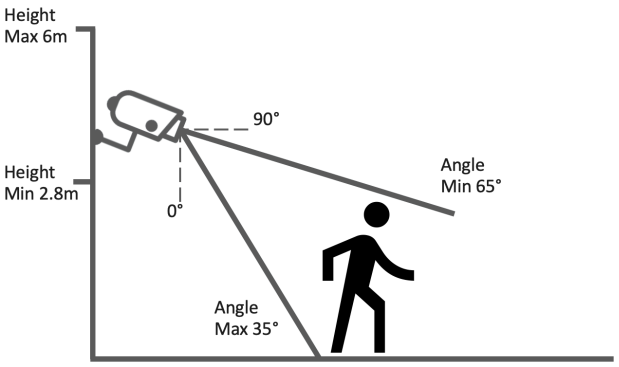
As with many algorithms, there are cases where performance will be suboptimal. Examples include:

- Fisheye cameras
- RGB cameras
- Small pixel size objects (objects in the distance)
- Extremely crowded and dense scenes
- Thermal cameras with low contrast variation

General Guidelines

The following guidelines should be used to provide optimal tracking and classification.

Camera Position and Angle



- Minimum camera height of 2.8m from the ground plane and no higher than 6m. However, higher zoom lenses will increase these figures.
- Cameras should be mounted on a stable surface to reduce the effects of vibration or other environmental factors.
- The camera view tilt angle (pitch) should be within ~35° from the horizontal.
- The camera view roll angle should be within +/- 10° from the horizontal.
- For high tracking accuracy, position the camera so objects are present for at least 2 seconds.

Obstructions

Camera scenes should be clear of environmental factors that can reduce object detection and tracking.

- Avoid locations that include foliage and large objects that can interfere with tracking.
- Be aware of static heat sources that could interfere with the thermal camera

Class	Min Pixels on Target	Height %
person	300px	5%
vehicle	1220px	6.2%

Detection Range

Physical detection range (in metres) will vary based on a camera's field of view, which is defined by the camera lens and focal point.

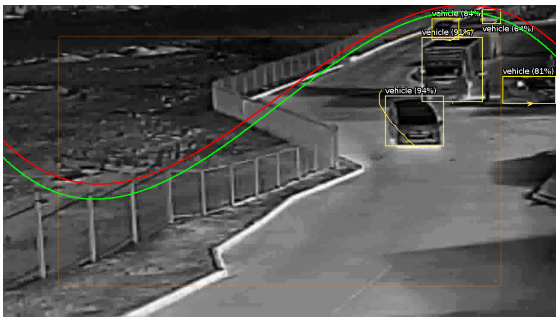
The ***pixels on target*** value (at 480p resolution) indicates object size based on the number of pixels in the bounding box of a tracked object. This number is relative to the input resolution, VCAserver can report this number under Burnt-in Annotation settings. For example, a tracked person in a 480p RTSP stream will have a higher pixels on target value than the same person in a 240p RTSP stream from the same camera.

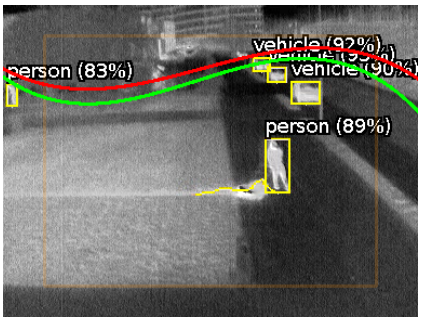
The **height %** is a relative measure of object size based on the approximate height of an object relative to the field of view. It will be the same regardless of input resolution.

The below images illustrate expected max detection range for the Deep Learning Thermal Tracker (DLTT) (red: min pixels on target, yellow: average min pixels on target):

Deep Learning Thermal Tracker

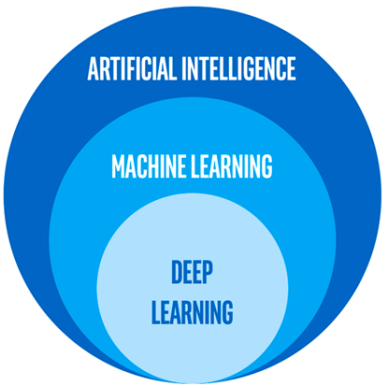
Technical Note





Frequently Asked Questions

Does the Deep Learning Thermal Tracker (DLTT) improve during operation?



The Deep Learning Thermal Tracker (DLTT) detects and classifies objects using a pre-trained model. This is created using large sets of reference data; it does not learn while in use.

Is calibration still required?

Because the Deep Learning Thermal Tracker (DLTT) detects and classifies objects that are in the scene, no calibration is currently required. If estimated speed and height metadata are required, then calibration will be required to be done. However, it will not impact the accuracy or range of detection.

What image settings should I use on the camera?

The image quality is an essential component to achieving accurate detection and tracking. The optimal resolution is 640x480 or D1 (720x480). Defining a lower resolution will reduce detection and tracking accuracy.

The optimal frame rate is 15 fps, providing a higher framerate will only result in an increase in resource usage without providing an increase in detection and tracking.

For sterile zone applications where the object speed will be minimal, tracking at 6fps is also supported. A lower framerate will result in an error.

The image quality and bitrate are important to ensure a good quality image is provided for video analytics. Where possible, the image quality and bitrate should be set to a high or max setting, with bitrate defined near or at the maximum available limit. This ensures that a good image can be presented for the tracker to detect and track accurately.

Setup

The Deep Learning Thermal Tracker (DLTT) is active on the channel as soon as it is selected. On first use, the tracker optimises to the available GPU(s). Depending on the GPU, this process can take up to 45 minutes to complete. During this initialisation process, a message is displayed on the channel preview page.

The tracker assesses every frame being delivered regardless of activity. This means that the GPU loading will be constant.

General

- The Deep Learning Thermal Tracker (DLTT) is controlled through licensing. Ensure the correct license has been activated on the system.
- Licenses are assigned to channels. Ensure the correct license is assigned to the channel.
- Ensure you have selected the correct tracker for your application.
- Ensure a good quality image is being delivered to the analytic engine.
- Any object that can be classified will be identified and tracked, use *object filters* to control which objects will trigger an event.
- Check that the prerequisites are installed and working correctly. The Deep Learning Thermal Tracker (DLTT) requires an Nvidia graphics card and CUDA libraries or a compatible accelerator.

Note: More information on the exact requirements can be found in our support portal, in the manual and through the hardware requirements document. (https://vcatechnology.com/downloads/)