# LSSCN Consortium

# Video based situation assessment for autonomous vehicles

Mahmud Mohammad, Ioannis Kaloskampis, Rossi Setchi, Yulia Hicks

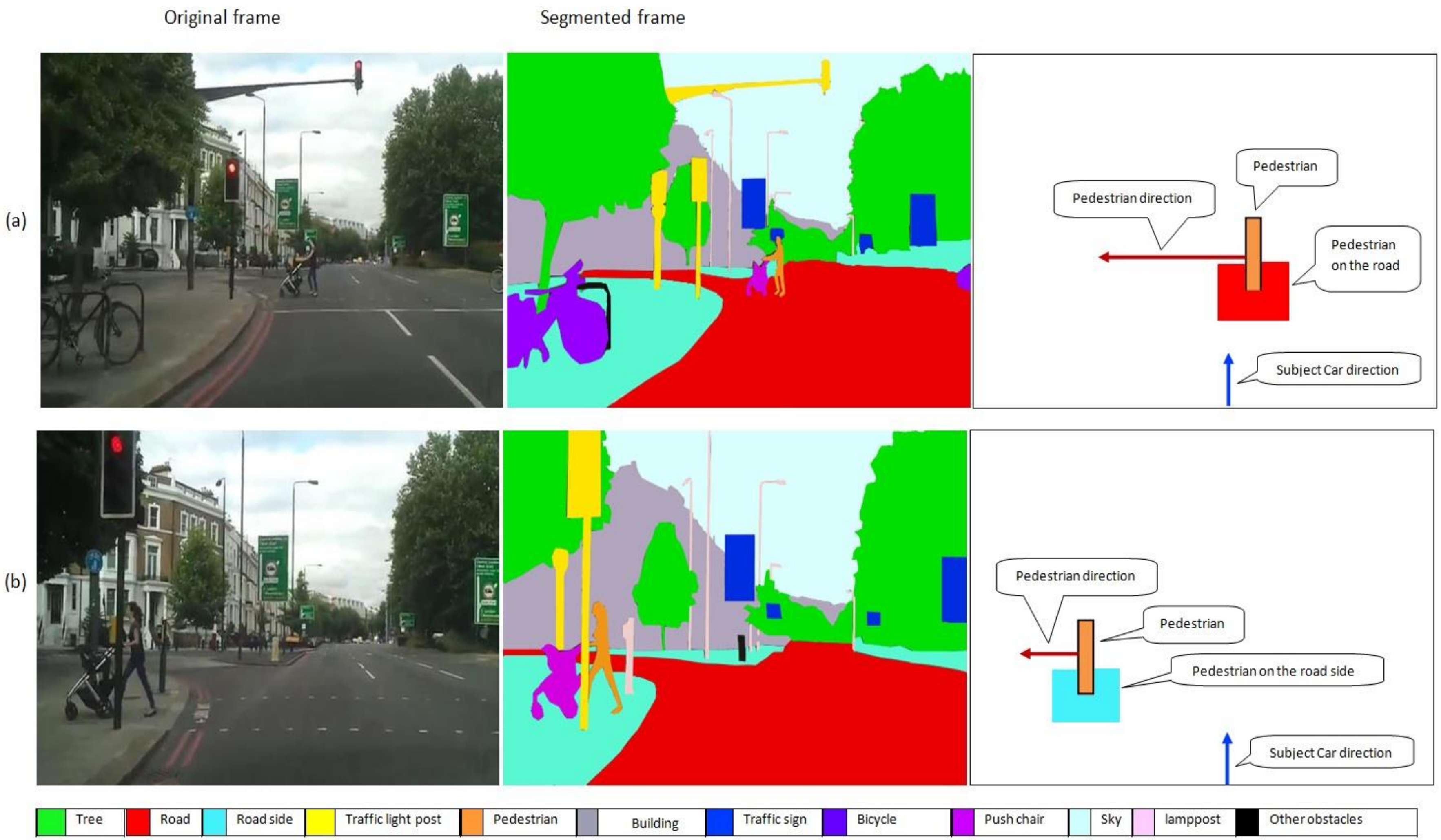
Cardiff University

# Situational assessment in road scenes

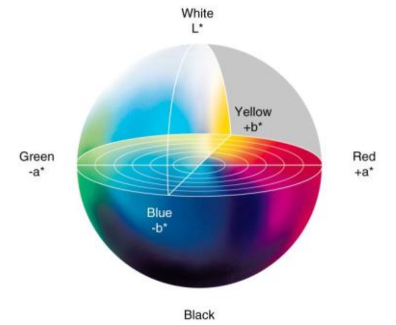
- Objects:
  - Type, location and behaviour (e.g. object speed, object motion direction).
- Road:
  - Road quality, road type, traffic signs
- Environment:
  - Environmental  conditions (e.g. weather and visibility conditions).

# Work carried out at Cardiff University

## Video segmentation, collision prediction, road type detection

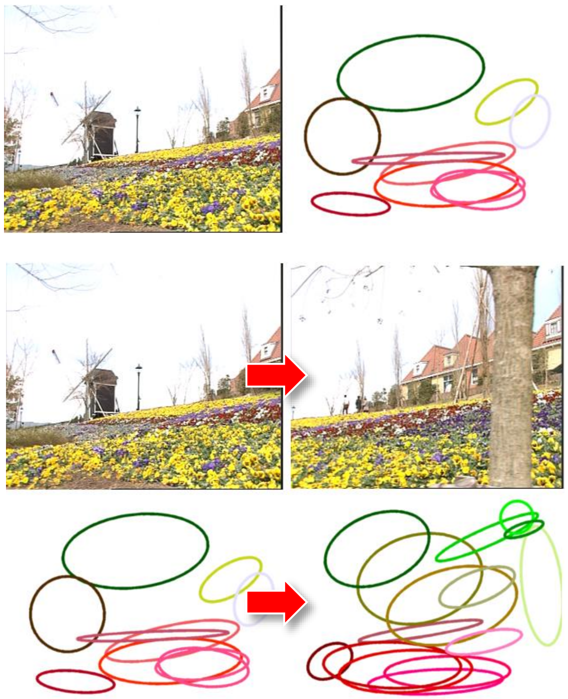


# Video segmentation – method overview




- Low level feature extraction (colour: L*a*b* space, pixel coordinates)
- Scene modelling using an evolving GMM

$$p(\mathbf{x}|\theta) = \sum_{i=1}^{K} a_i \mathcal{G}(\mathbf{x}|\mathbf{m}_i, \mathbf{C}_i)$$



# Video segmentation – method overview

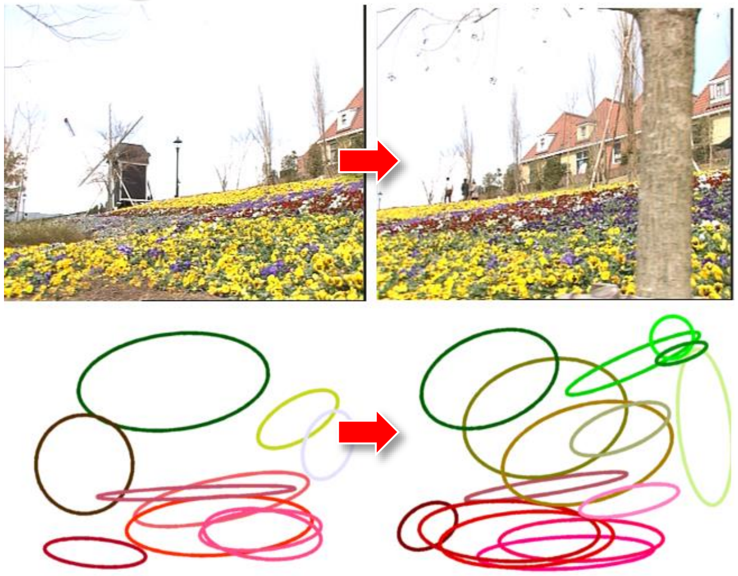
- GMM parameters change over time according to changes in video
  → evolving model
- The number of components K is estimated automatically for each frame.
- Achieved by merging the evolving GMM $\theta_{evolving}$ with a temporary GMM $\theta_{temp}$ trained on the current frame
- Merging is achieved using modified EM

Variable number of GMM components

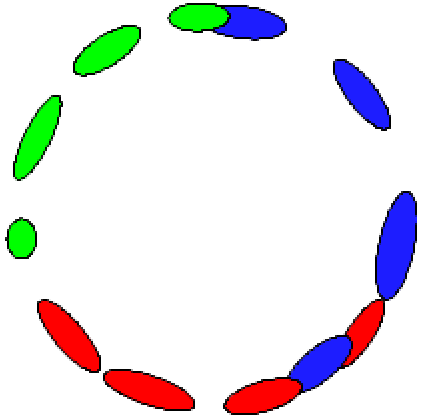
[Charron and Hicks, ICIP 2010]

[Kaloskampis and Hicks, ISCCSP 2014]

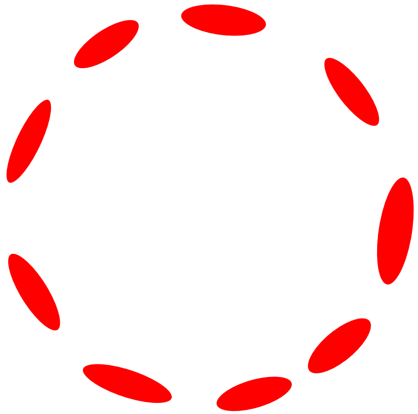
# Video segmentation – method overview



After Concatenation



Merging

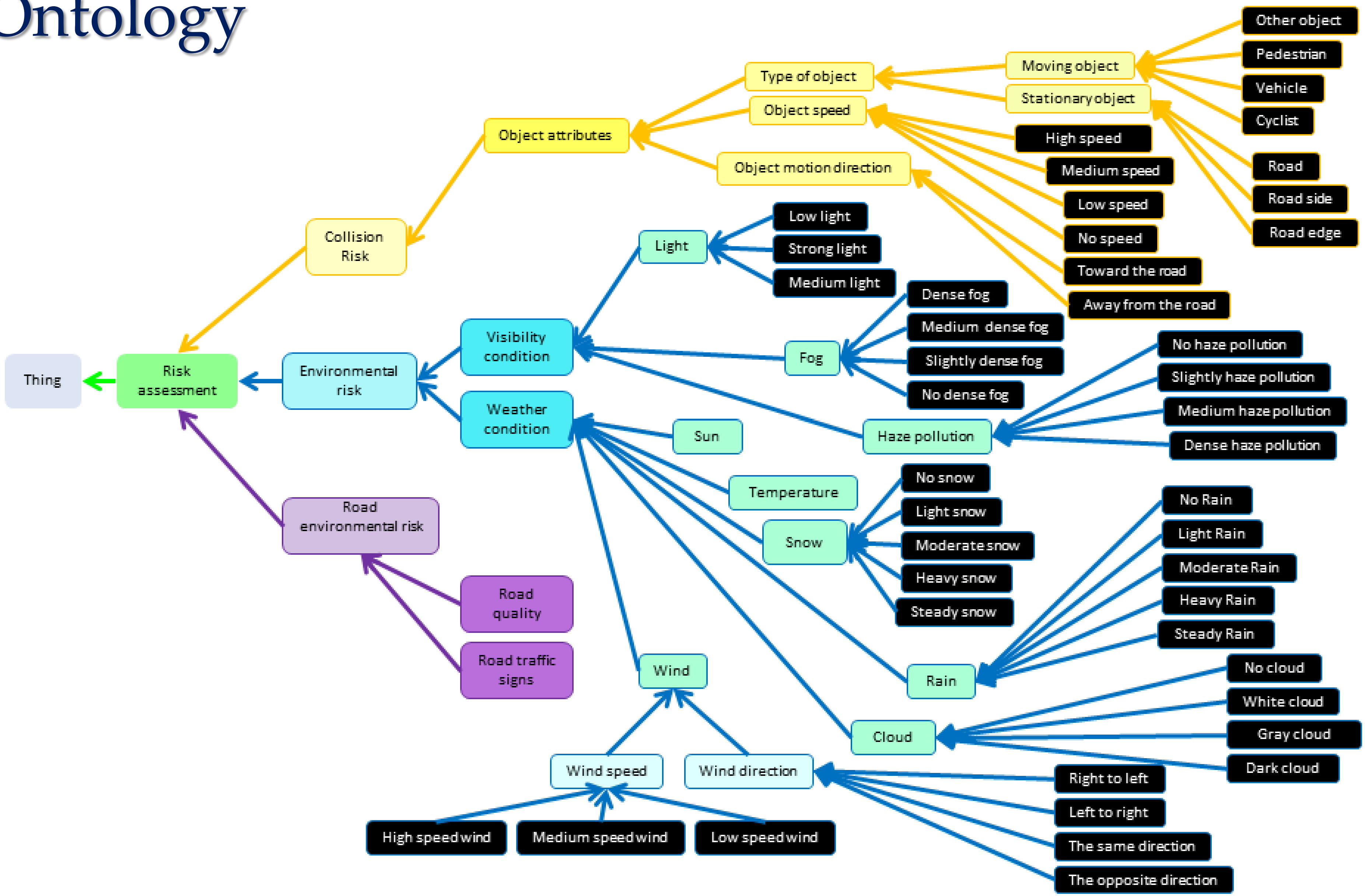


- All data used for GMM estimation (images, data points) are discarded
- By merging GMMs overlapping (redundant) Gaussians disappear
- Result: Efficient memory handling
- The frame is segmented with the merged GMM

# Collision prediction– method overview

- Detect key environment entities (road, buildings, etc) using video segmentation method developed at Cardiff.
- Detect and track pedestrians using Viola & Jones object detector and Kalman filter for tracking.
- Monitor the interactions of pedestrians with environment entities (relative positions and motion direction).
- Use an ontology which represents all key entities and their relationships to assess the risk of collision

# Ontology


Thing
Risk assessment
Collision Risk
Object attributes
Type of object
Object speed
Object motion direction
Moving object
Stationary object
Other object
Pedestrian
Vehicle
Cyclist
Road
Road side
Road edge
High speed
Medium speed
Low speed
No speed
Toward the road
Away from the road
Environmental risk
Visibility condition
Weather condition
Light
Low light
Strong light
Medium light
Fog
Dense fog
Medium dense fog
Slightly dense fog
No dense fog
Haze pollution
No haze pollution
Slightly haze pollution
Medium haze pollution
Dense haze pollution
Sun
Temperature
Snow
No snow
Light snow
Moderate snow
Heavy snow
Steady snow
Rain
No Rain
Light Rain
Moderate Rain
Heavy Rain
Steady Rain
Cloud
No cloud
White cloud
Gray cloud
Dark cloud
Wind
Wind speed
Wind direction
High speed wind
Medium speed wind
Low speed wind
Right to left
Left to right
The same direction
The opposite direction
Road environmental risk
Road quality
Road traffic signs

# Ontology rules and implementation

- The ontology's inference rules are formed in the semantic web rule language (SWRL).

- Sample rules are given here:

Pedestrian (?p) ^ Road (?r) ^ hasHighSpeed (?p; ?s) ^ objectOnTheRoad (?p; ?r)  highRisk (?p; ?a)

Pedestrian (?p) ^ Road (?r) ^ hasHighSpeed (?p; ?s) ^ objectOnTheRoadEdge (?p; ?re) ^ hasAwayFromThe (?p; ?r)  mediumRisk (?p; ?a)

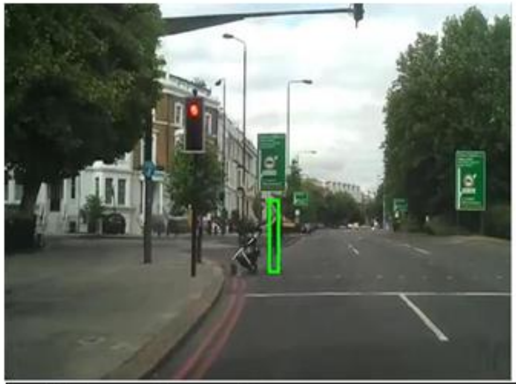
Pedestrian (?p) ^ Road (?r) ^ hasHighSpeed (?p; ?s) ^ objectOnTheRoadS ide (?p; ?rs) ^ hasAwayFromThe (?p; ?r)  lowRisk (?p; ?a)

Pedestrian (?p) ^ Road (?r) ^ hasNoSpeed (?p; ?s) ^ objectOnTheRoadS ide (?p; ?rs)  noRisk (?p; ?a)

**p** pedestrian, **r** road, **re** road edge
**rs** road side, **a** assessment.

- Tools: Protégé resource, Pellet reasoner, SPARQL query

# System output



Video one: Frame 16
Object: Pedestrian
Object's location: On the road
Object's speed: Low speed
Direction: 90 degrees w.r.t. the car
Risk assessment: High risk

(a)

Video two: Frame 66
Object: Pedestrian
Object's location: On the road side
Object's speed: Medium speed
Direction: 90 degrees w.r.t. the car
Risk assessment: Low risk

(b)

Video two: Frame 21
Object: Pedestrian
Object's location: On the road
Object's speed: High speed
Direction: 270 degrees w.r.t. the car
Risk assessment: High risk

(c)

Video three: Frame 16
Object: Pedestrian
Object's location: On the road
Object's speed: High speed
Direction: 270 degrees w.r.t. the car
Risk assessment: High risk

(d)

Video four: Frame 60
Object: Pedestrian
Object's location: On the road side
Object's speed: High speed
Direction: 180 degrees w.r.t. the car
Risk assessment: No risk

(e)

Video five: Frame 161
Object: Pedestrian
Object's location: On the road
Object's speed: Medium speed
Direction: 90 degrees w.r.t. the car
Risk assessment: High risk

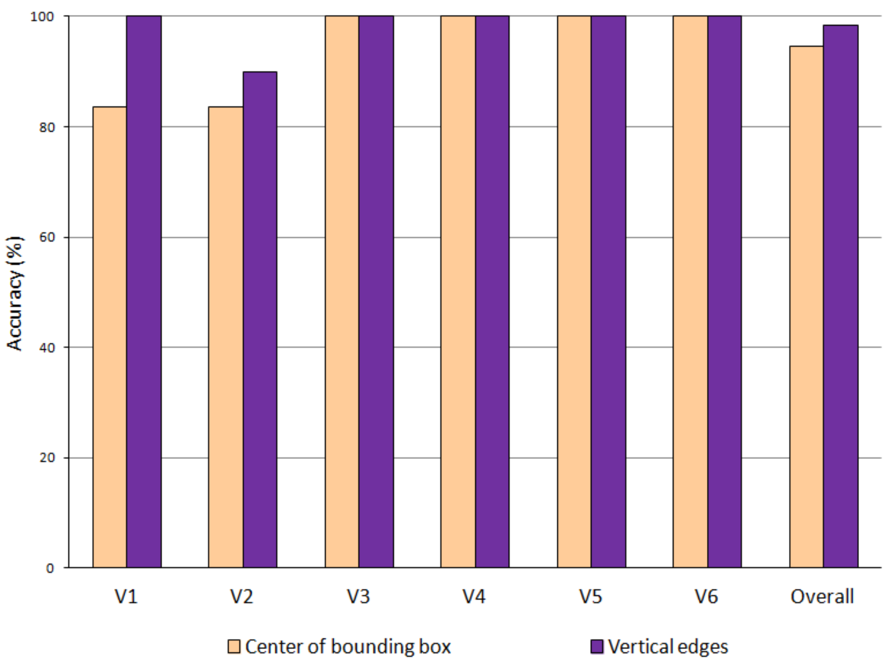
(f)

# Experimental evaluation - Dataset

- We created a dataset comprising six videos featuring pedestrian behaviour in road scenes with various degrees of risk.
- Frame rate: 25 fps to 30 fps.
- Resolution: 640 * 480 (resized).
- All videos captured from right-hand drive vehicles and correspond to the driver's perspective.
- Ground truth for the dataset was provided by two independent observers.

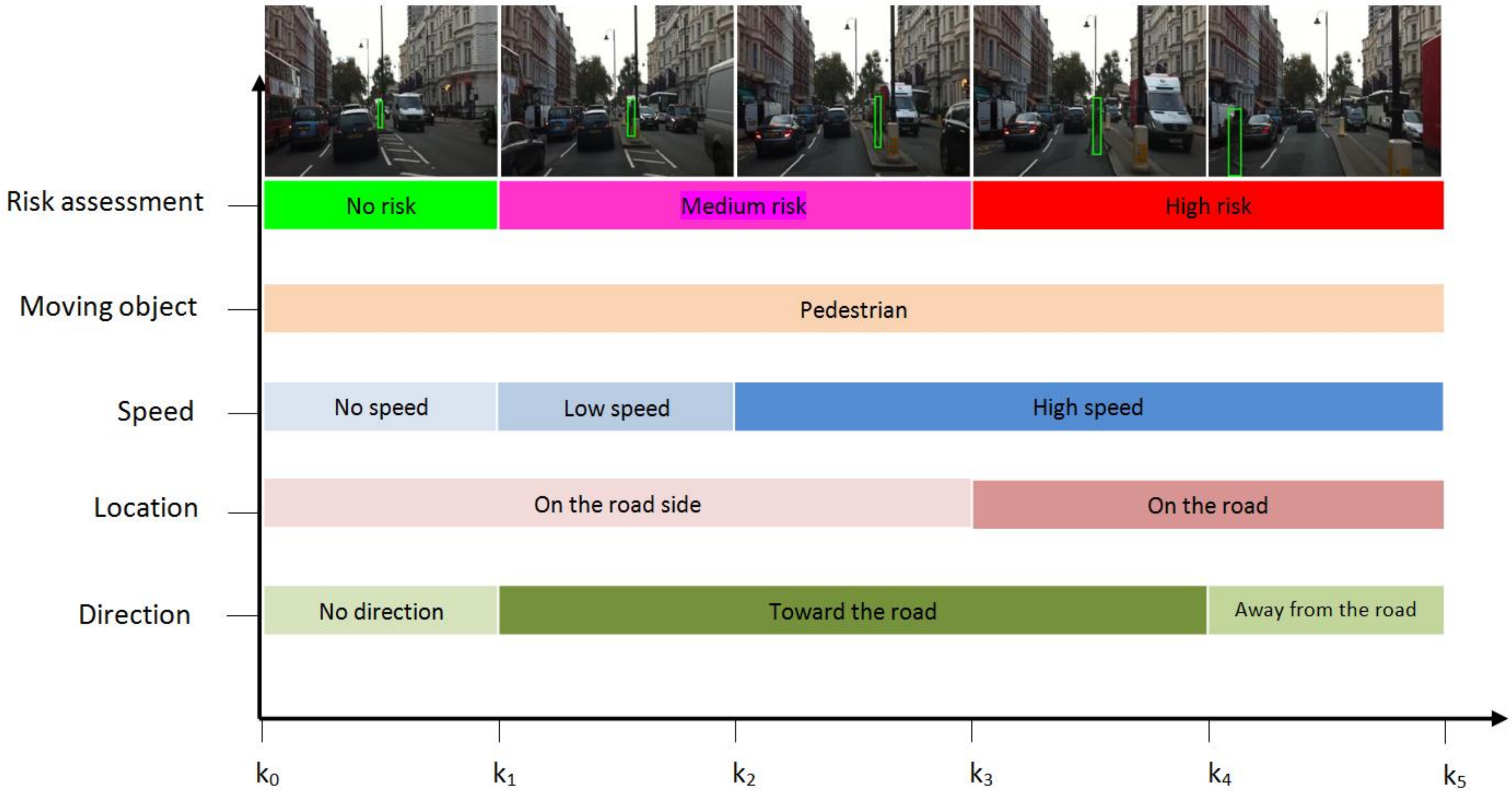
# Experimental evaluation - Results

- Implementation details:
  - EvoGMM (Kaloskampis & Hicks 2014) to detect key scene entities (road, pavement etc.)
  - VJ classifier for pedestrian detection
  - Kalman filter based tracker
- Results in terms of accuracy in assessment of the degree of risk in a given scene



# Experimental evaluation - Results

- Event based representation of the results



# Research in road type detection

- Vision based road-type classification:
  - Specify the road type based on the video content of the scene.
  - Important step towards road scene understanding.
- Previous work in road type detection:
  - Mioulet et al., 2013;
  - Tang and Breckon, 2011
- Unlike previous work, our approach takes into account all scene regions.

# Overview of the proposed method

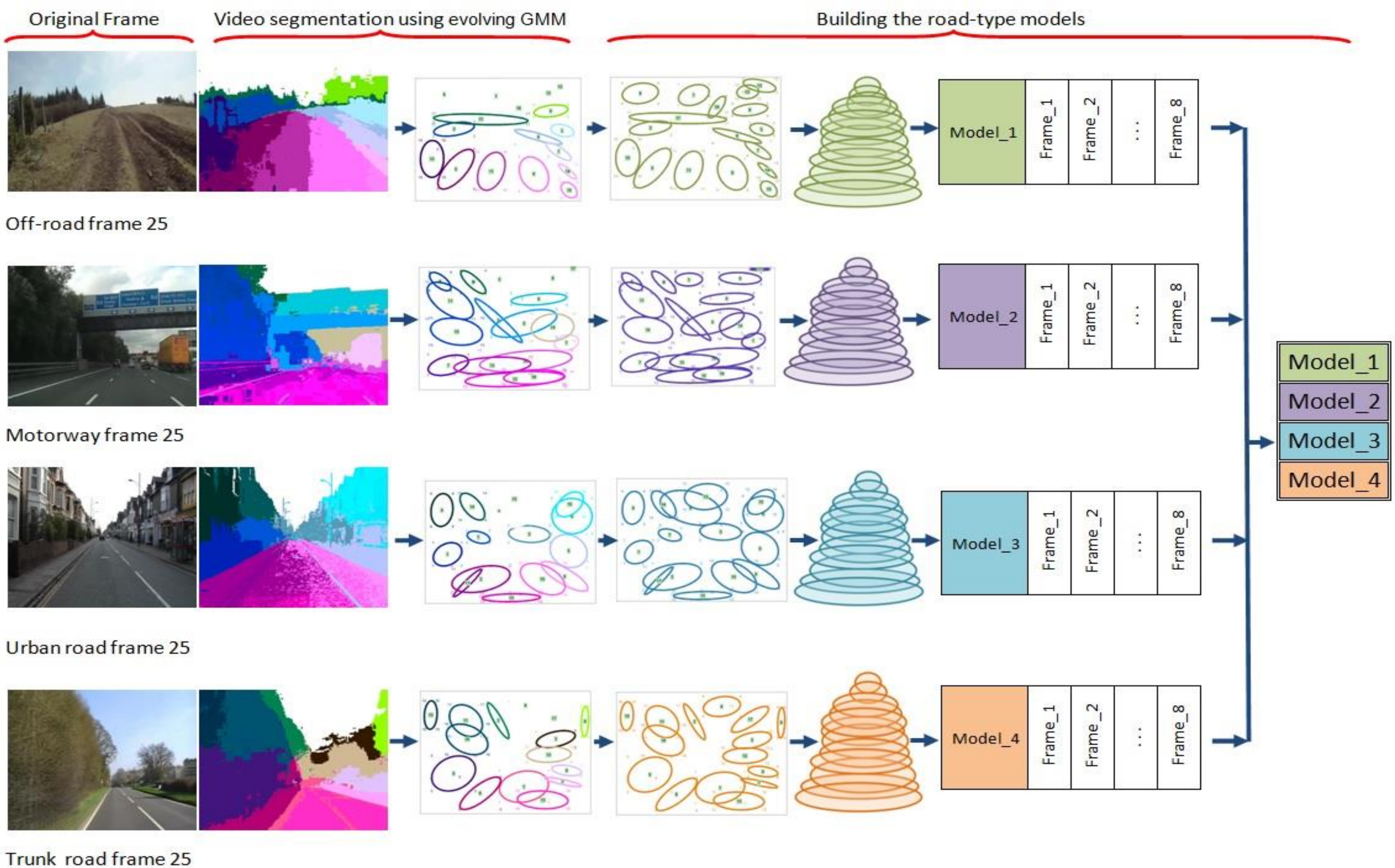
- We propose a method to classify road types from data obtained from a monocular camera.
- We consider a four-class problem (motorway, off-road, trunk road, and urban road).
- Our method consists of two stages:

1. Building a statistical model for each road type offline using video segmentation.
2. Online classification of new video frames.

# Video segmentation

- The statistical road model is based on the evolving GMM algorithm for video segmentation (Kaloskampis & Hicks, 2014), selected for the following reasons:
  - The algorithm maintains the temporal coherence of the segments between different frames.
  - Online rather than a batch method, hence does not require all frames at once.

# Building the road type model

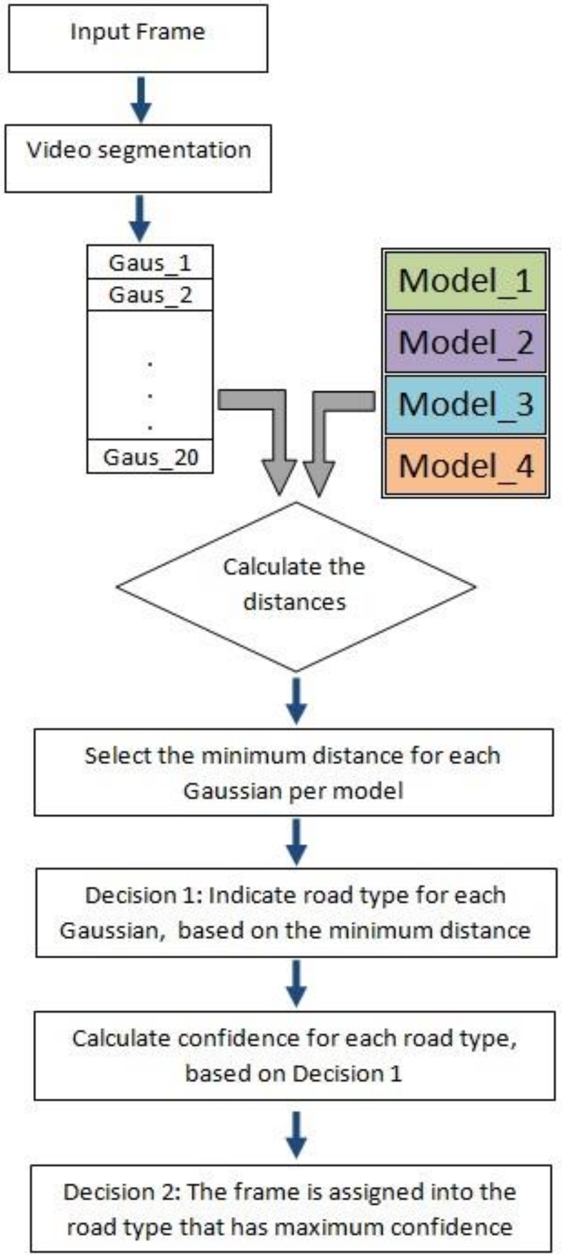


# Classification

- The second stage is the online classification of new video frames, in two steps.

  1. A GMM is created for the new frame. We estimate the proximity between the new frame's Gaussians and the models obtained from first stage with the Bhattacharyya distance .

  2. The road type confidence score is calculated based on the size of the segment corresponding to each classified Gaussian.



# Experimental results

- We built the model for each road type using 8 videos of 25 frames each.
- We resized the resolution of all video frames to 640 * 480; the frame rate is between 25~30 fps.
- All videos were captured from right-hand drive vehicles and correspond to the driver's perspective.
- For testing, we used 800 video frames illustrating each road type (not used when building our models).
- We benchmark our method against the state of the art.

# Experimental results

- Results are presented in terms of % classification accuracy.
- Our method achieves higher accuracy for each road type individually and higher overall accuracy.
- The difference between the two methods is more evident in the classification of the off-road environment.