Visual people tracking with deep learning detection and feature tracking

Author: Marcos Pieras Sagardoy

Tutor: José María Cañas Plaza

Universidad Rey Juan Carlos

July 25, 2017

Index

Introduction

Introduction

Definició

Object tracking: Estimate the target state over time from image sequences

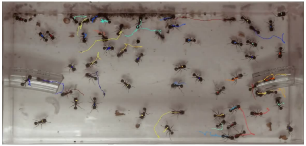
It is challenging field due to:

- Variations beacause of geometric changes.
- Variations due to photometric factors.
- Occlusions.
- Similar objects in the scene.

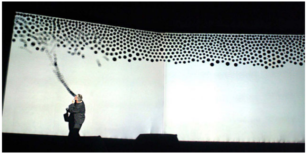
Applications

(a) Surveillance

(b) Sports

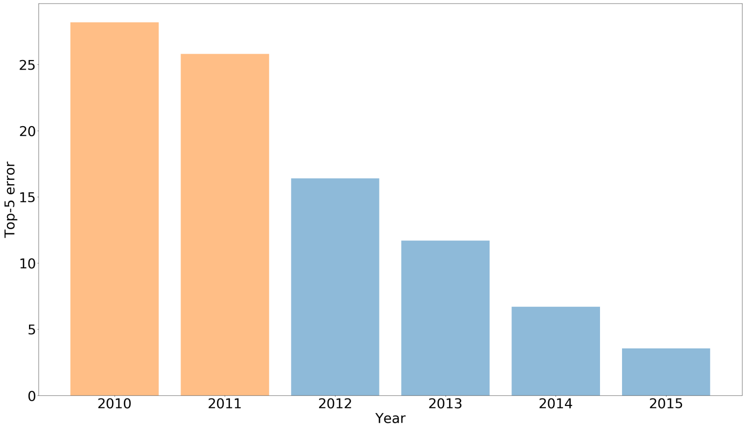


(c) Science



(d) Art

Why deep learning techniques ?



Objetives

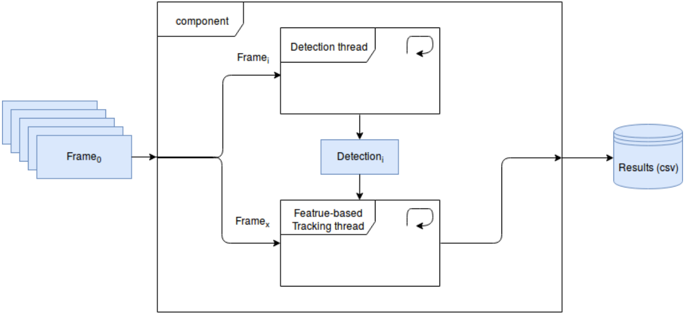
Develop and characterize an algorithm of multiple people tracking, combining:

- Neural networks
- Feature-based tracking

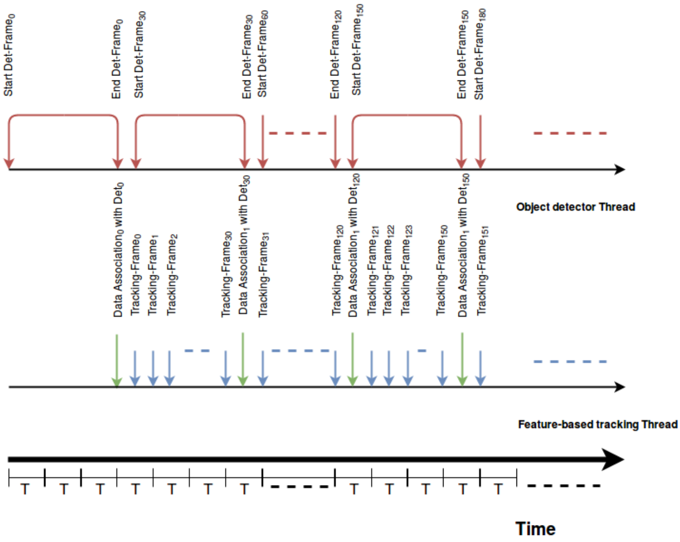
Testing of the component on an international databases, the Multiple object tracking challenge.

Software implementation

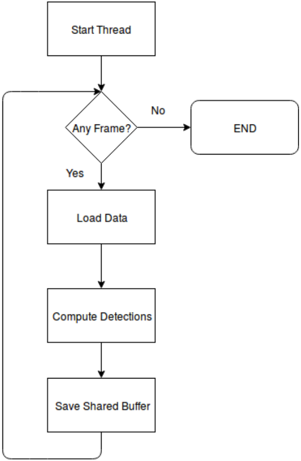
System Overview: Diagram



System Overview: Temporal diagram

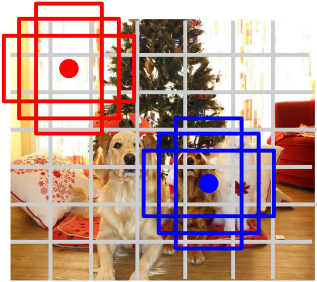


Object detector thread: Overview



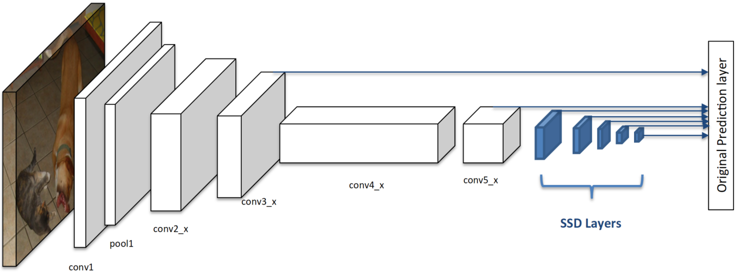
Single Shot Multibox Detector [SSD] I

(e) Input image.

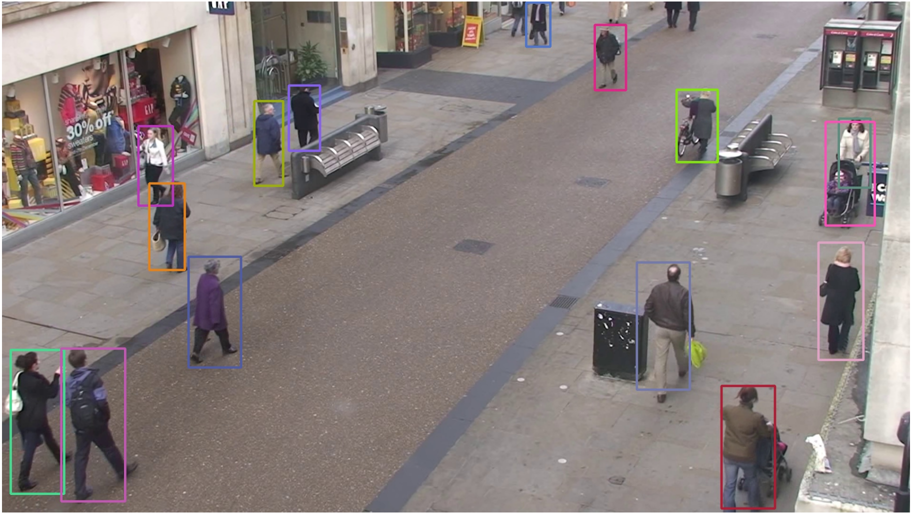


(f) Divided image.

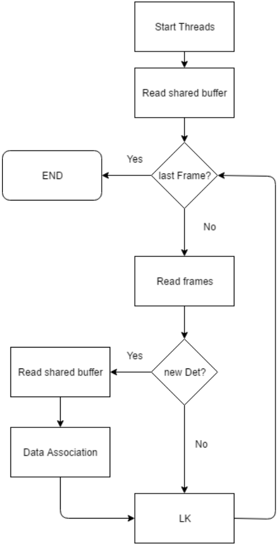
Single Shot Multibox Detector [SSD] II



Result Object detector thread



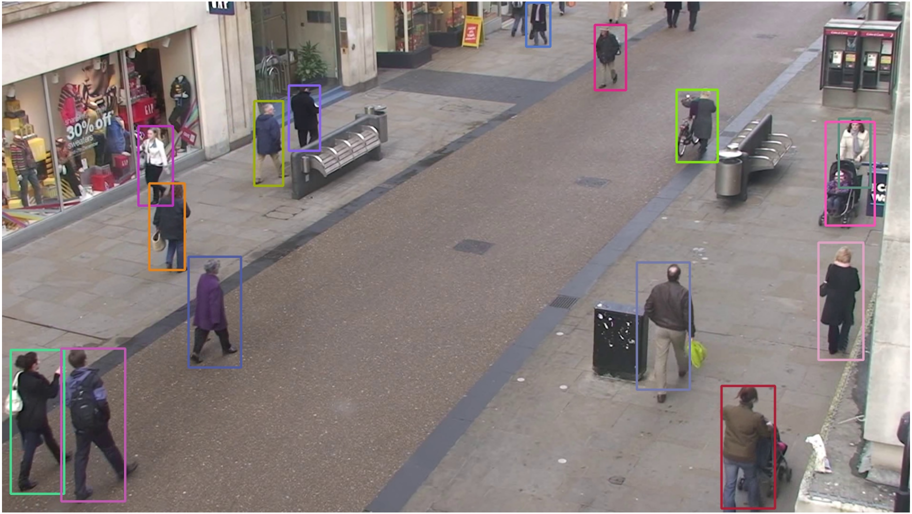
Feature-based tracking thread: Overview



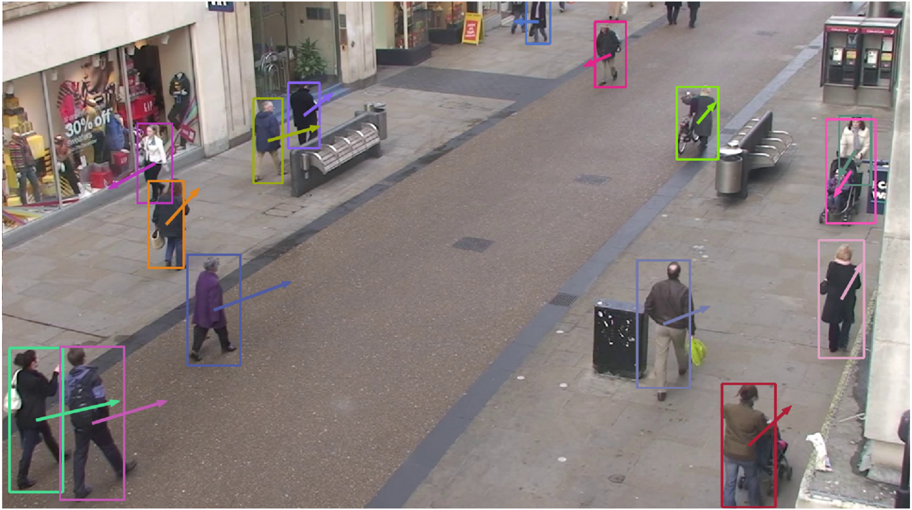
Feature-based tracking overview

(g) Detection. (h) Points. (i) Matching. (j) Displacement.

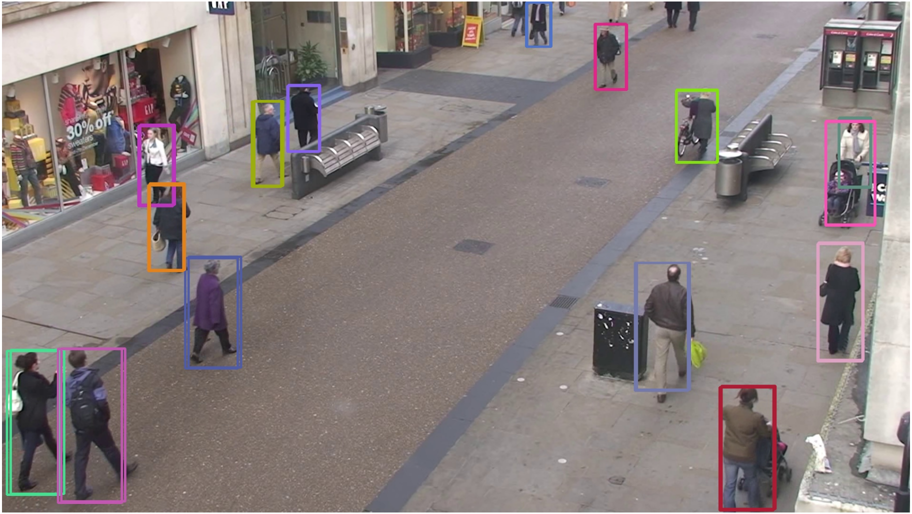
Blobs



Displacement



Upload estimation

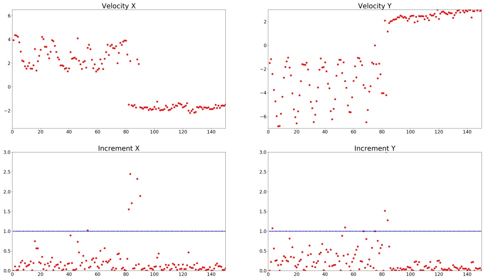


Tracking failure

(k) Trajectory

Tracking failure

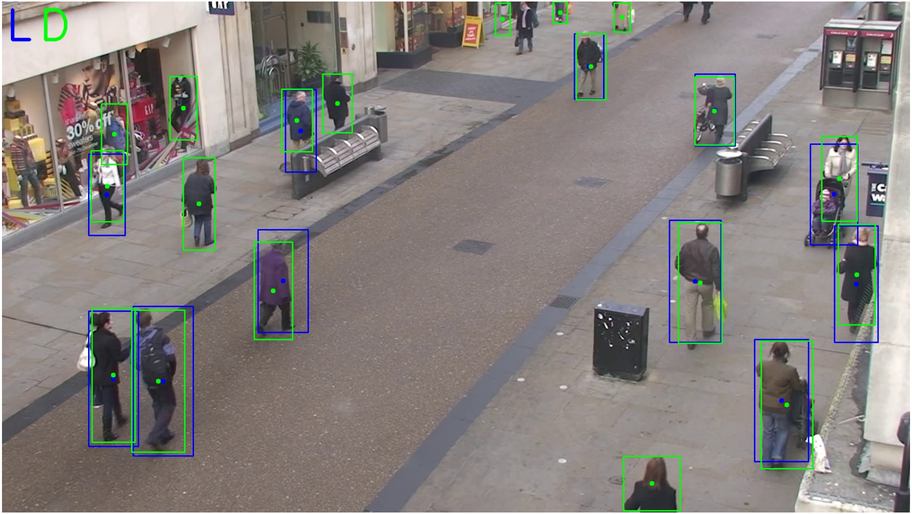
(a) Trajectory



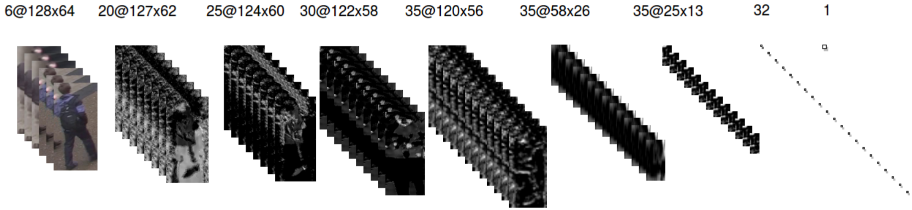
Data Association I

- **Situation 1**, blob with nearby detection, replaced by blob detection.
- **Situation 2**, blob without nearby detection, blob continues.
- **Situation 3**, blob detection, new or past ?

Data Association II

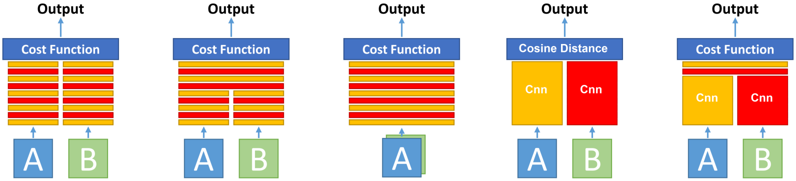


Data Association: Siamese architecture



Experiments

Reidentification architectures



(a) Cost Function

(b) In-Network

(c) Joint data input

(d) Feature extractor with cosine distance

(e) Feature network fine-tuned

Data augmentation

(f) Original (g) Random brightness (h) Random crop (i) Vertical flip (j) Gaussian blur (k) Random shadow

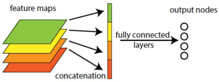
(l) Zoom in (m) Transla- (n) Zoom out (o) Gaussian (p) Opposite

Parameters training network I

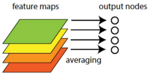
- **Loss**: Binary cross entropy.
- **Optimizer**: Adam with exponential decay.
- **Activation**: ReLu.
- **Initialization**: He. initialization. Biases with the value of 0.1.
- **Regularization**: Dropout in the fully connected layer.

Parameters training network II

- **Preprocessing**: Centre the data and normalized.
- **Final layers**: Flatten, Global Average Pooling and Spatial pyramid pooling.



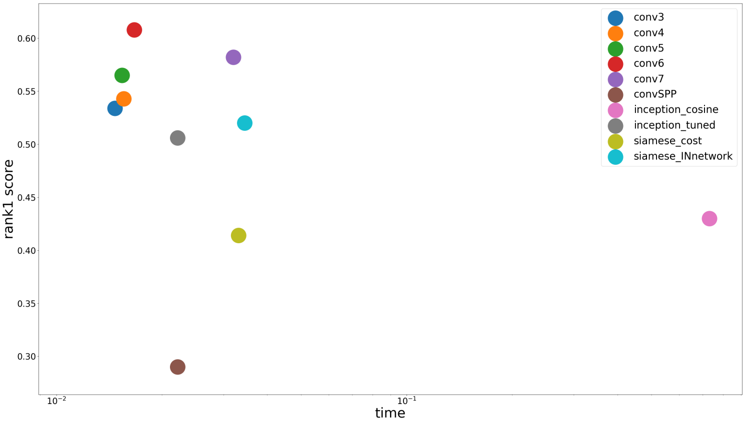
(q) Flatten



(r) Global Average Pooling

- **Output**: Sigmoid activation.

Comparison reidentification methods



Validation experiments

Results by sequence

	GT	MT	PT	ML	FP	FN	IDs	MOTA	MOTP	FPS
02	54	0	13	41	2181	15526	113	0.1	67.1	9.02
04	83	0	41	42	5495	33980	290	16.6	71.1	12.3
05	125	3	43	79	28571	4713	109	-12.2	67.8	17.94
09	25	1	19	5	932	3225	71	19.7	62	10.52
10	54	0	4	50	404	11647	81	1.5	68.4	14.23
11	69	0	16	53	948	7366	72	8.6	71.4	17.49
13	107	0	9	98	1315	10743	32	-5.6	67.1	20.5
Global	517	3	127	387	18896	78999	618	10.8	70.3	15.85

Table: Results algorithm by sequences.

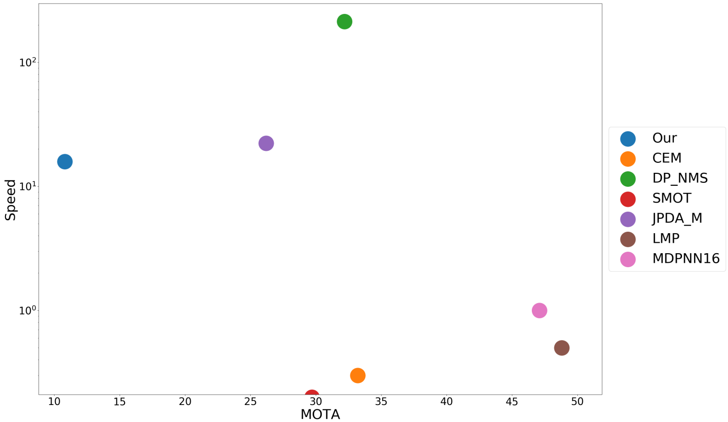
Weaknesses

(s) Hight texture blob

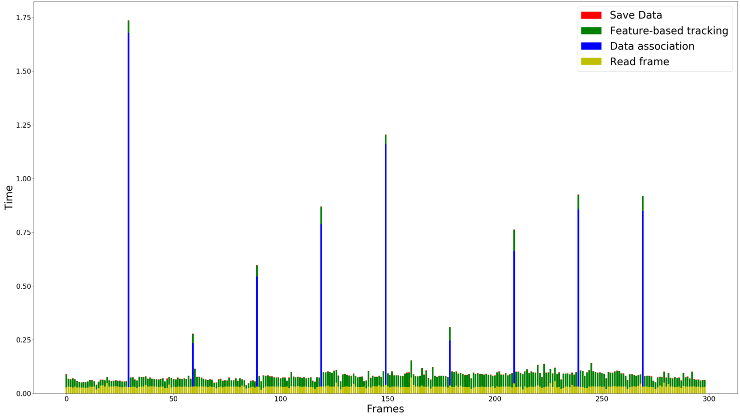
(t) Low texture blob

(u) Far away blob

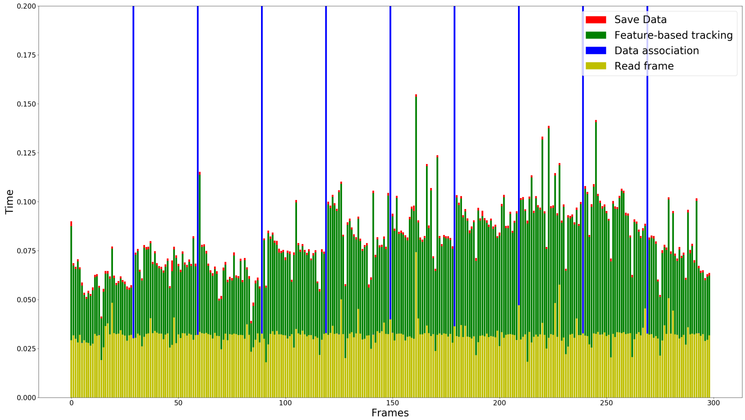
Comparison with SOTA



Timing performance I



Timing performance II



Conclusions

Conclusions

- Object detector using deep learning.
- Development of a Featured-based tracking module.
- Merging detections and feature-based tracking, adding a reidentification module.
- Testing of the component on an international databases.

Thank you for your attention !