Development of an intelligent system to detect and predict pedestrians' intentions in urban environments
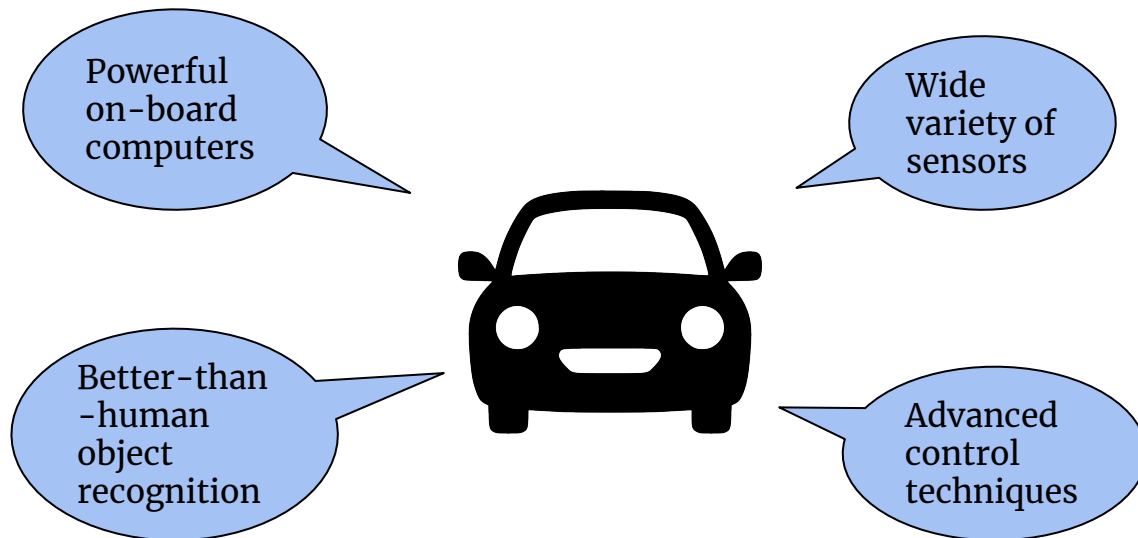
Alberto Barragán Moreno

Outline

- The problem
- State of the Art
- Proposed solution
- Results
- Conclusions
- Future development
- Live demonstration

The problem

Autonomous-driving technology is progressing really fast...

Powerful on-board computers

Wide variety of sensors

Better-than-human object recognition

Advanced control techniques

Waymo's experimental vehicles are considered level 4. Tesla claims they will reach level 5 by 2020.

The problem

But... what about interactions?

- Vehicle-to-Vehicle (V2V) and Vehicle-to-Infrastructure (V2I) are "easy" to solve. Their behavior is simple and predictable, and high-speed technologies (5G, WiFi) allow for almost real-time communications.

- Vehicle-to-Pedestrian (V2P) is more challenging... humans act in quite strange ways and that makes it difficult for machines to anticipate their actions.

State of the Art

- **Academic research**

  Joint Attention in Autonomous
  Driving (JAAD) dataset

  NUSCENES by APTIV
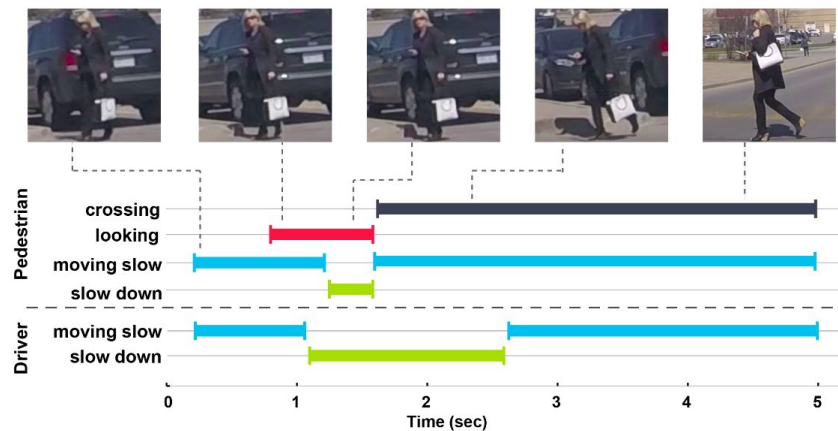
- **Private companies**
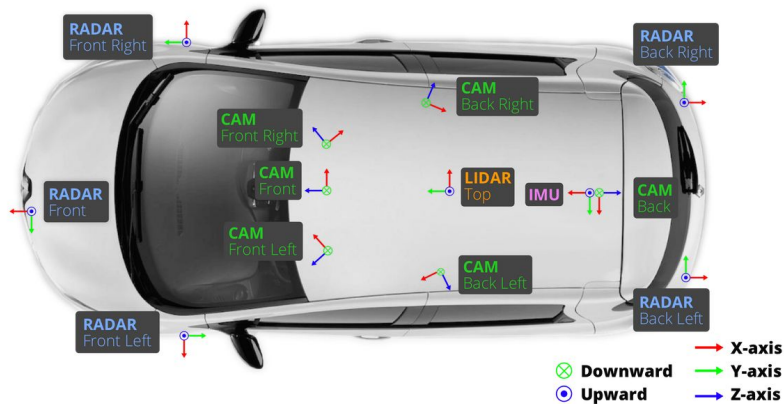
  PERCEPTIVE AUTOMATA

  INTUO

  HUMANISING AUTONOMY

State of the Art

- **Publicly available datasets**
  - JAAD:
    - focused on pedestrian behaviors
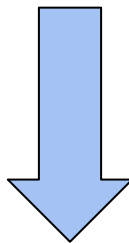   - vague vehicle monitoring

State of the Art

- **Publicly available datasets**
  - nuScenes:
    - focused on multi-sensor perception
    - just a few human behaviors are considered

Proposed solution

Lack of a big-scale dataset

(no unified methodology, nor specific tools)

Establish new, all-round procedures and methodology

Create a new suite of integrated tools

Proposed solution

Combine data from the main sources of information we use as humans

Objectives

1. To analyze the main factors influencing the behavior of pedestrians in urban environments, especially in crossing/not-crossing scenarios

2. To establish an unified methodology covering all the way from image acquisition routines to feature analysis

3. To develop a suite of integrated tools to facilitate and accelerate the successive processes in the proposed methodology

4. To test and verify both the methodology an the tools suite by performing a small-scale analysis of pedestrian behavior in a series of urban scenarios

The Pedestrian Awareness Dataset (PAD) suite

All-in-one video processing, subject detection and attributes management solution

any2avi

clip2frames

clip2vid

frames2clip

hevc2avi

vid2clip

BBox Creator

BBox Identification

BBox Remover

BBox Tracker

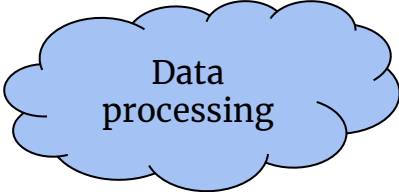Attributes Manager

Data Fusion

YOLO
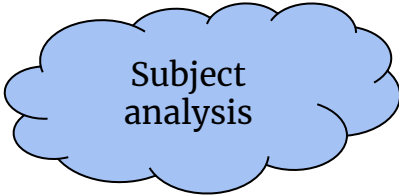
BORIS

The PAD Methodology

Filming planning and execution

- Location scouting
- Route tracing
- Hardware setup

Data processing

- Formatting
- Video cutting
- Frame extraction

Subject analysis

- Detection
- Identification
- Tracking
- Behavior and context annotation
- Pedestrian attributes

Results

Clips	Characteristics	Session
001-033	1164x874 @ 25fps	10/04 (W. a.)
034-050	1920x1080 @ 60fps, 120º	18/05 (S. m.)
051-075	1920x1080 @ 60fps, 120º	14/06 (F. e.)

Data available	Scenes
Video	75
Behavioral and attributes annotations	70
Bounding boxes and subject tracking	28
Video with tracking and behaviors	28
Vehicle parameters	0

Results
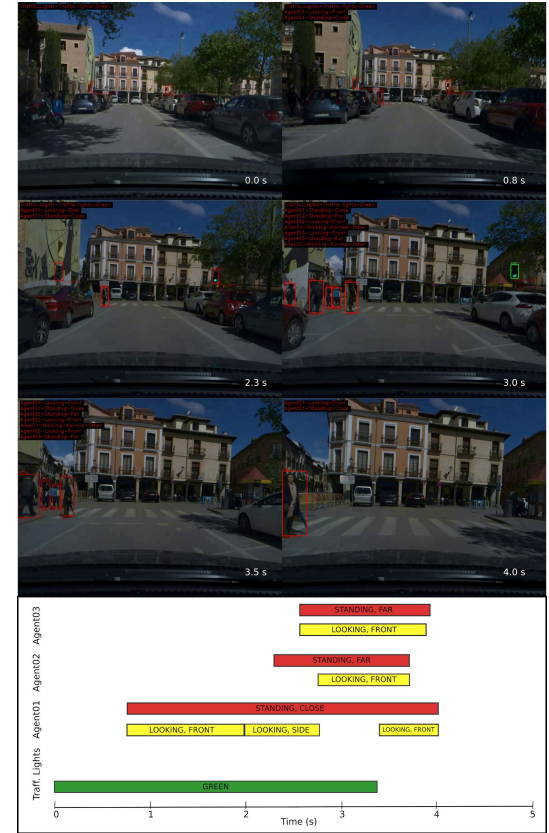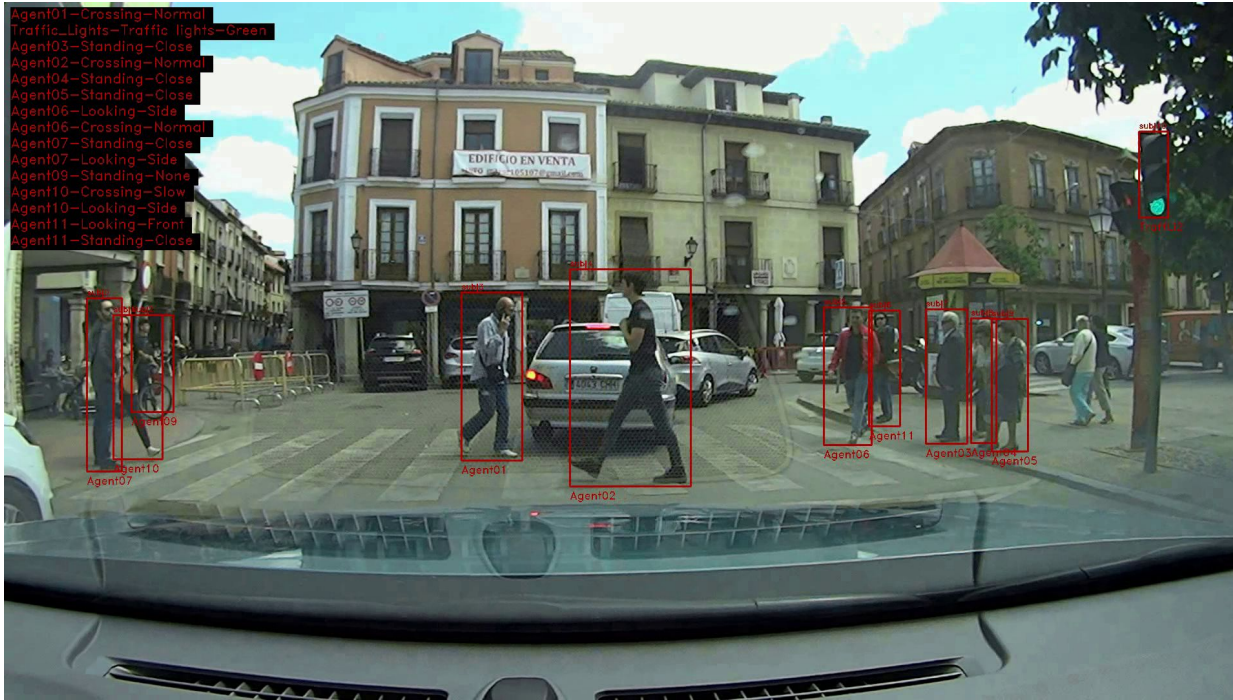
Results

Gathered data is not enough for a relevant behavioral study... but some useful observations can be made:

- Both pedestrians and cars obey traffic rules (kind of)

- Time-to-Cross vs. Time-to-Collision is one of the most important influencers in the decision-making process

- Presence of pedestrians and belonging to a group also have a big impact on attention

- Driver and vehicle behavior also affect pedestrians' actions, but no data could be obtained

Conclusions

- The suite covers most of the tasks needed to perform the behavioral analysis

- Detection with YOLO has to be improved, as it constitutes the bottleneck of the project

- BORIS is not the optimal tool for annotation, but it is more than enough

- Methodology has been correctly validated, with minor changes between the first version and the one used in the end

- The lack of vehicle data makes it impossible to perform a comprehensive analysis

- No proprietary software nor services have been necessary

Further development

- Multi-sensor setup, following nuScene's approach, to collect more inner and outer data

- More diverse scenes: nighttime, adverse weather, urban architectures, small towns

- Pedestrians' traffic lights status

- Monitoring of other vehicles

- Model the attention of a pedestrian due to peripheral vision rather than direct looking

- Port the methodology to static urban elements

Thank you for your attention

Questions time and demonstration