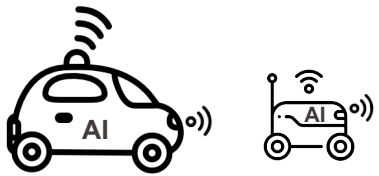
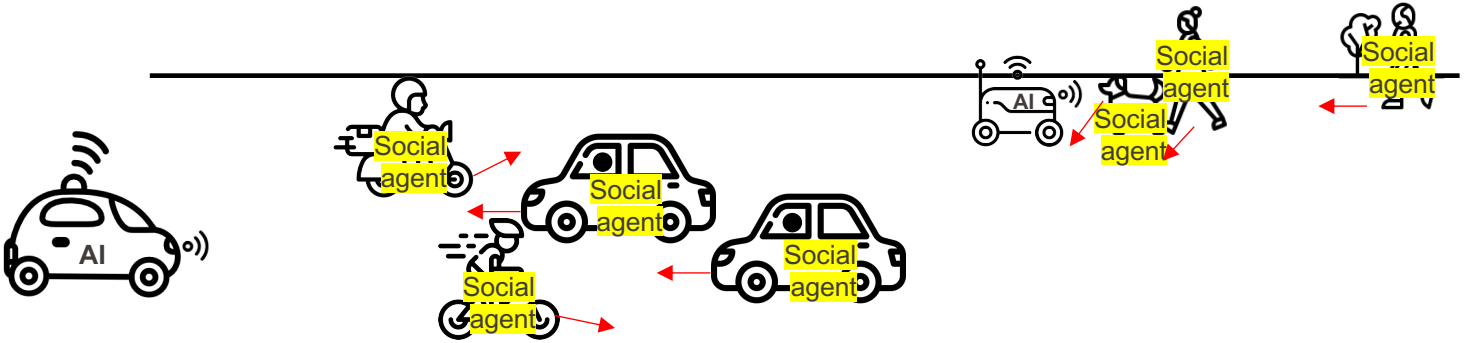


Crowd-Robot Interaction

Prof. Alexandre Alahi



*Social agent = any moving entity in the world (driver, pedestrian, cyclist…) interacting with other moving entities

"Humans subconsciously **forecast the future**…
Autonomous Vehicles must have the same **forecasting** capability to **harmlessly** and **effectively co-exist**",
Our lab goal.

Forecasting is essential…

time t

Autonomous ✓

t+1

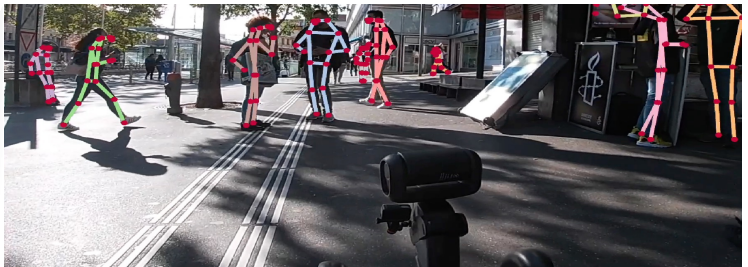
Autonomous ✓

t+2

Autonomous ✗

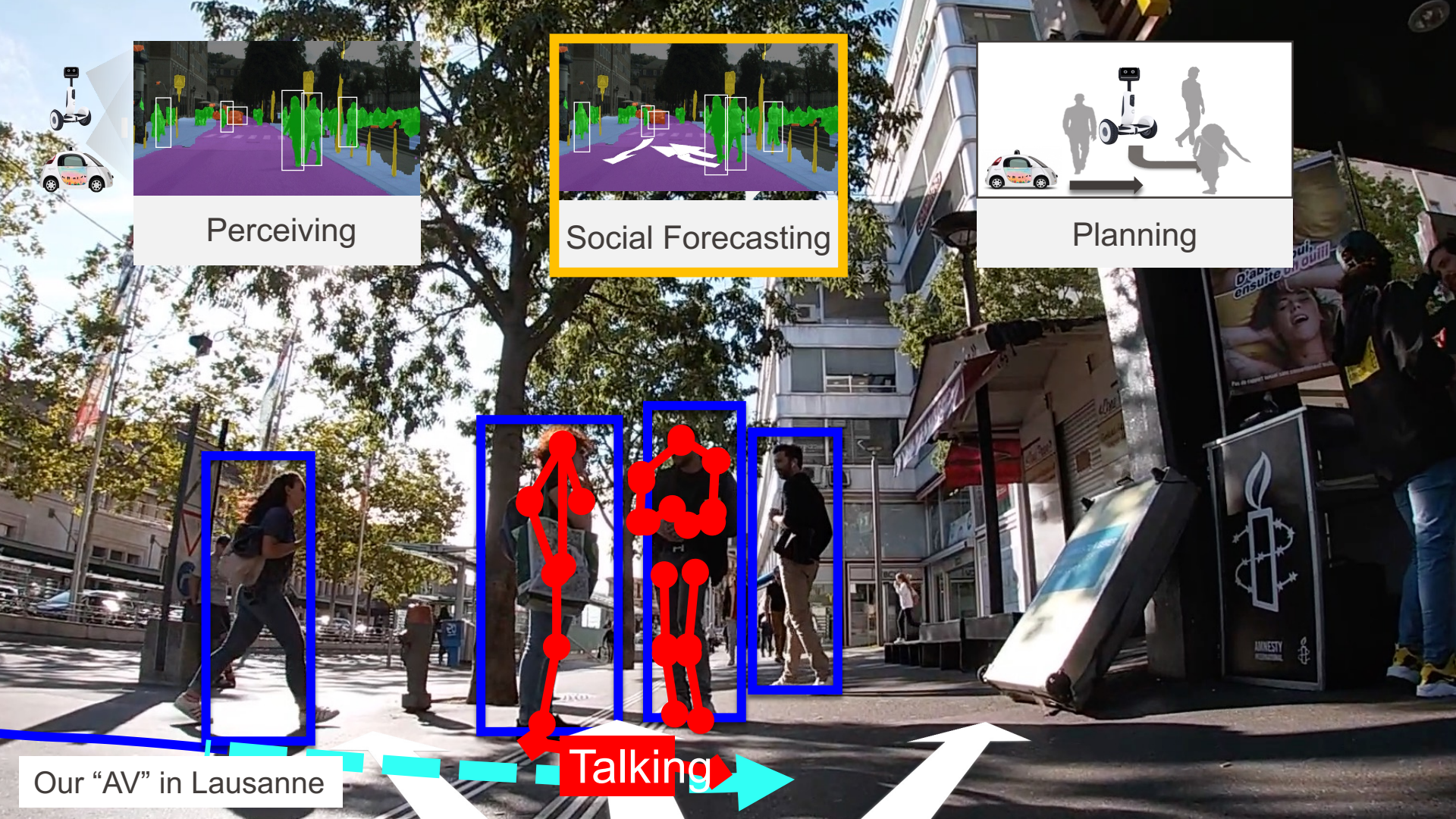
=> Our robot freezes in close human proximity

t+3

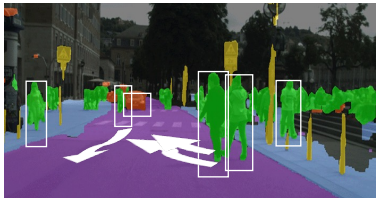


Autonomous ✗

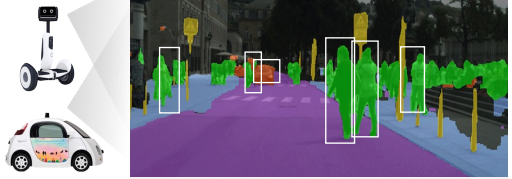
=> Our robot does not comply with social norms


Perceiving
Social Forecasting
Planning
Talking
Our “AV” in Lausanne

Alexandre Alahi, EPFL

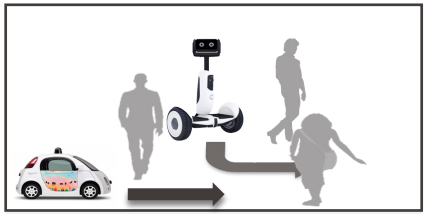


Social Forecasting



Perceiving

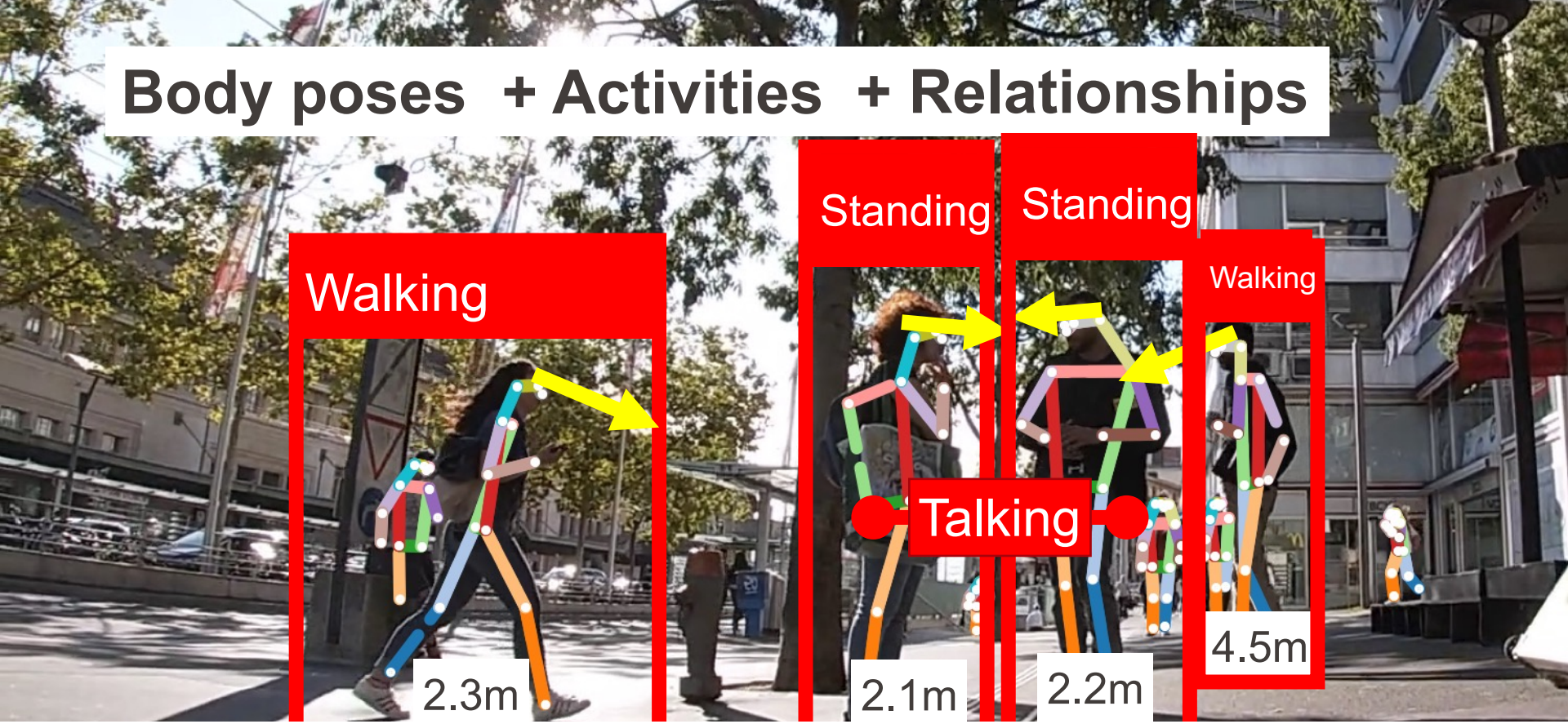
Socially-aware AI



Planning

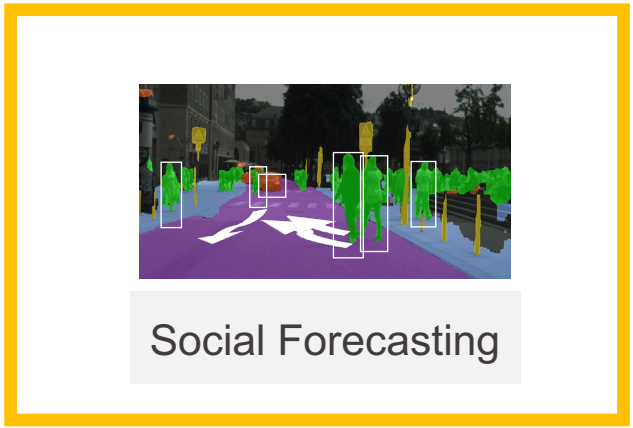
Object detection

Body poses + Activities + Relationships

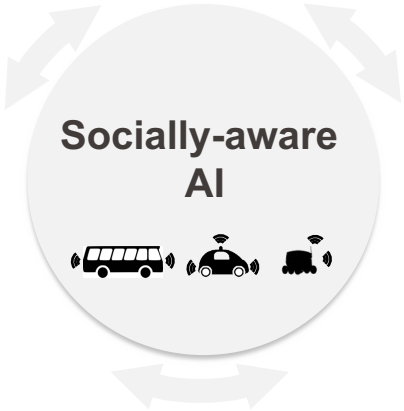


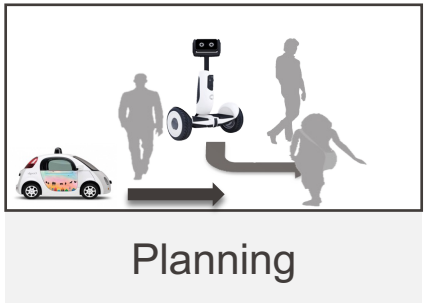
[1] PifPaf, **CVPR'19;**
[2] Keypoints communities, **ICCV'21;**
[4] Convolutional Relational Machine, **CVPR'19;**
[6] Monoloco, **ICCV'19;**

on-line demo: vitademo.epfl.ch/movements
[3] OpenPifPaf**,** open-source library, **IEEE ITS'21**
[5] Detecting 32 Pedestrian Attributes, **IEEE ITS'21**
[7] MonStereo, **ICRA'21**


Social Forecasting


Perceiving

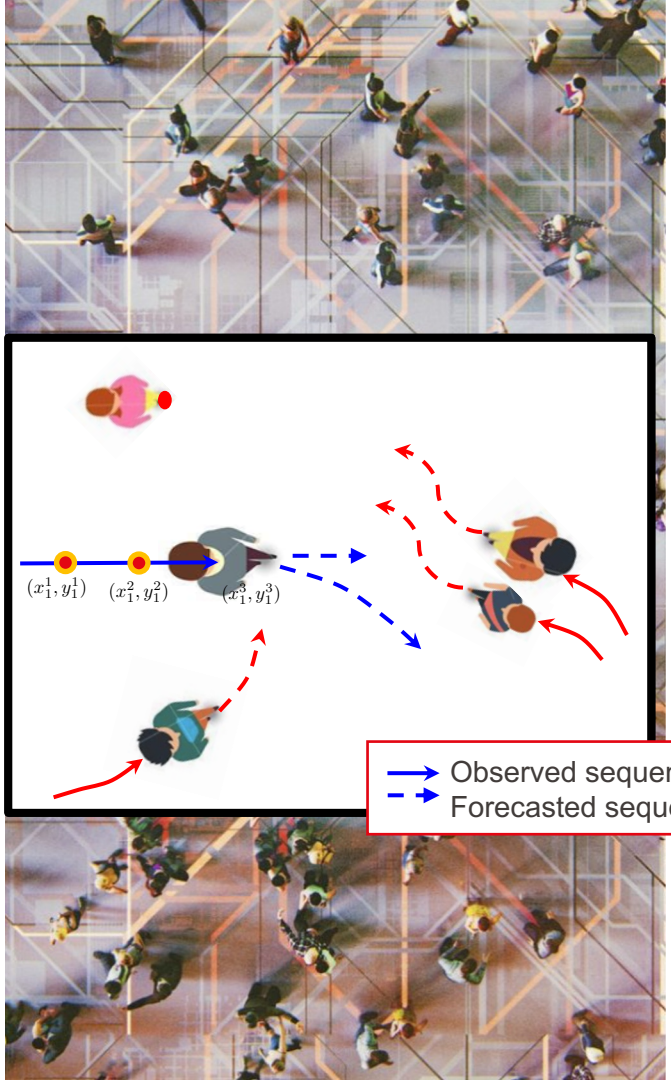

Socially-aware
AI


Planning

Alexandre Alahi, EPFL

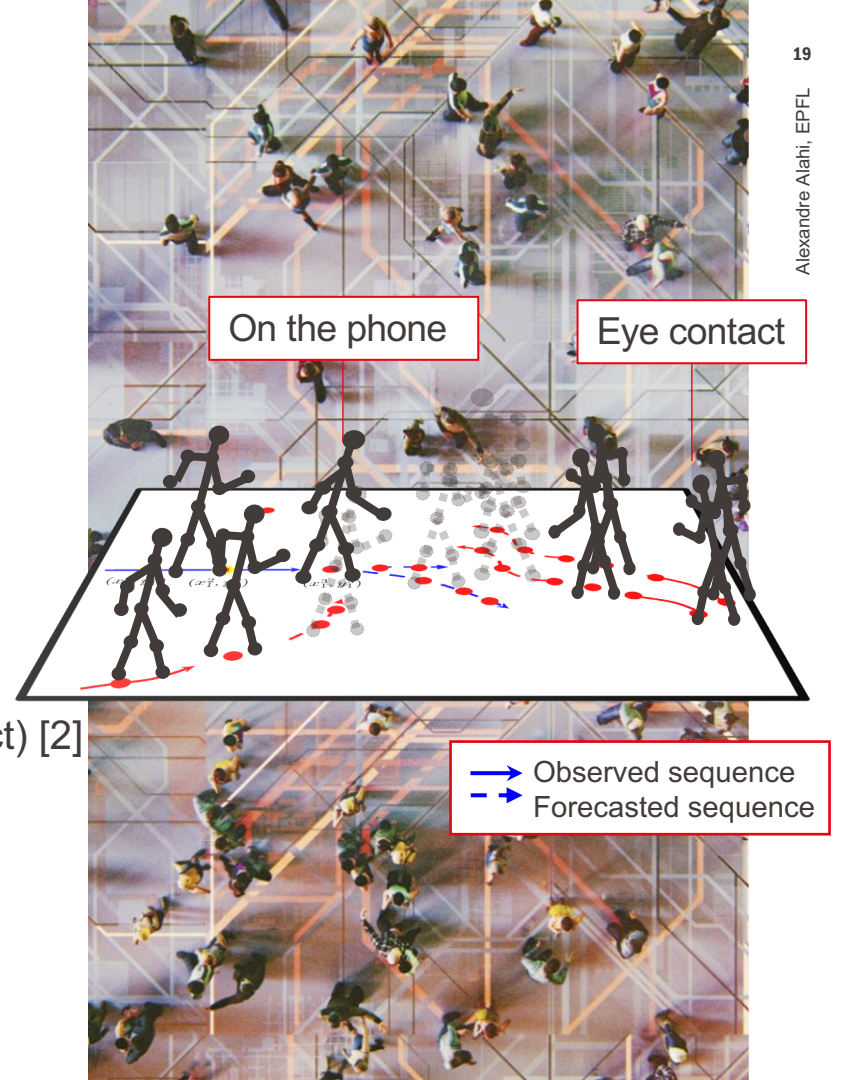
Social Forecasting
(w/ pedestrians)

- **Input**: several sequences of states
- **Output**: forecast the future states, *e.g.*, next 5 seconds



Social Forecasting
(w/ pedestrians)

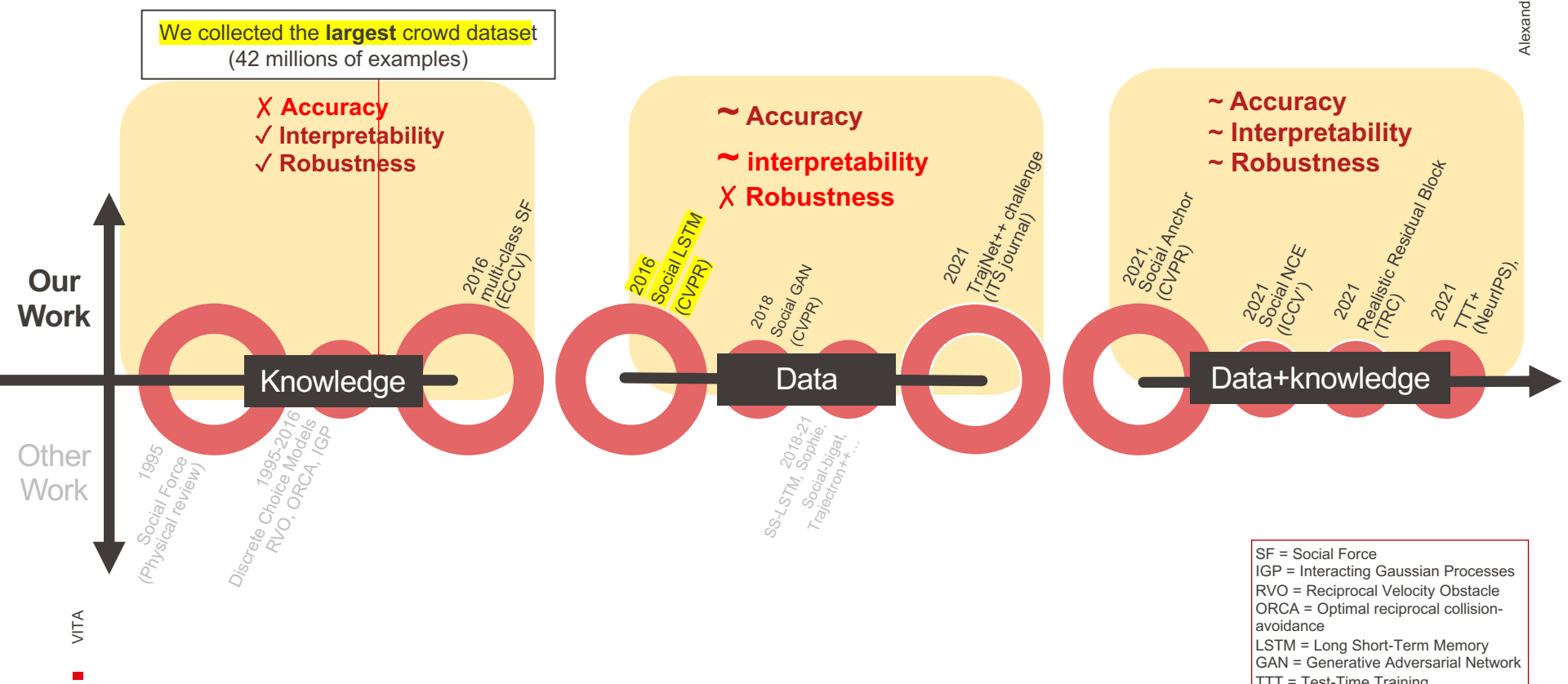
- **Input**: several sequences of states
- **Output**: forecast the future states, *e.g.*, next 5 seconds
- **State**:
  - (x^t,y^t) coordinates in time
  - Body pose [1]
  - Attributes (*e.g.*, on the phone, eye contact) [2]
- Challenge 1: agent-agent interactions
- Challenge 2: disentangle physics from social



[1] Our PifPaf, CVPR'19
[2] Our 32 attributes detector, ITS transactions'21

Alexandre Alahi, EPFL

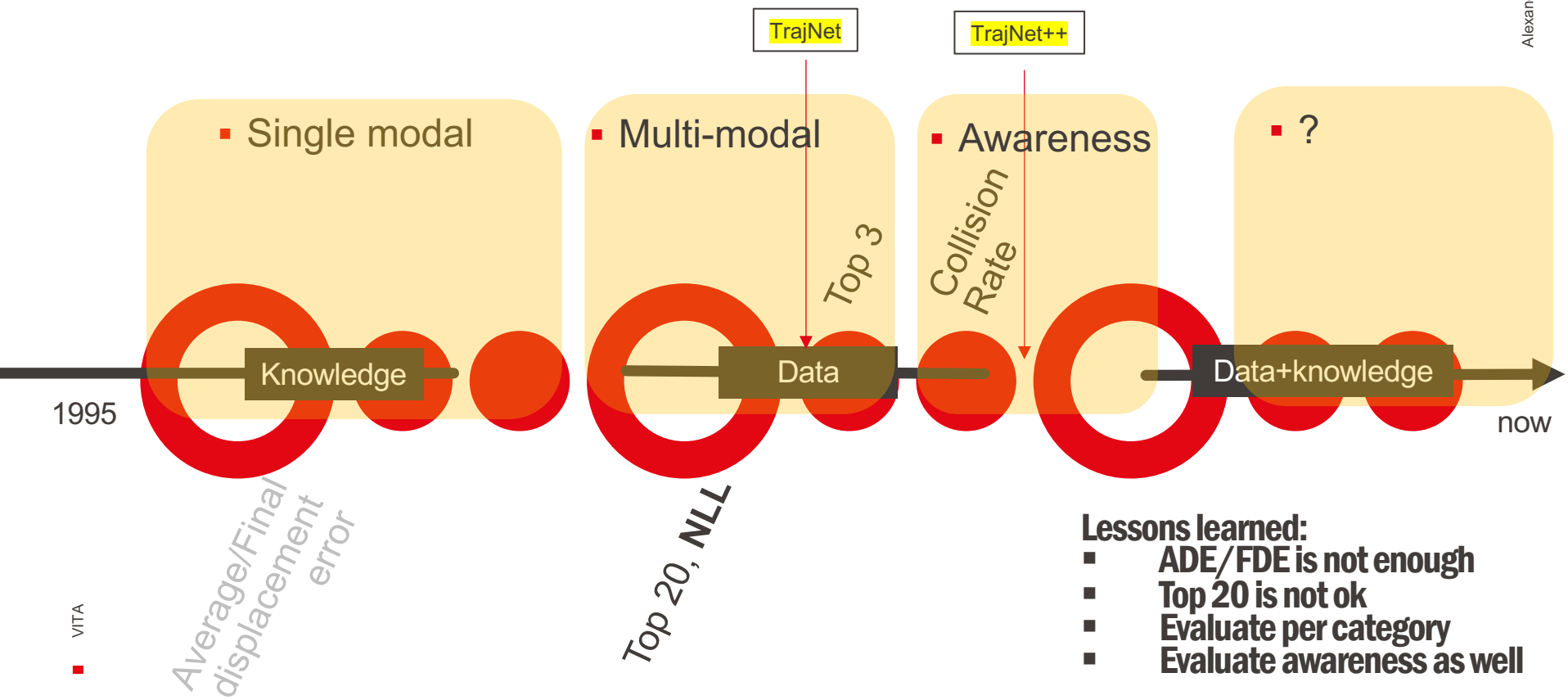
Learning paradigms



SF = Social Force
IGP = Interacting Gaussian Processes
RVO = Reciprocal Velocity Obstacle
ORCA = Optimal reciprocal collision-avoidance
LSTM = Long Short-Term Memory
GAN = Generative Adversarial Network
TTT = Test-Time Training

Alexandre Alahi, EPFL

Accuracy over the years



Lessons learned:

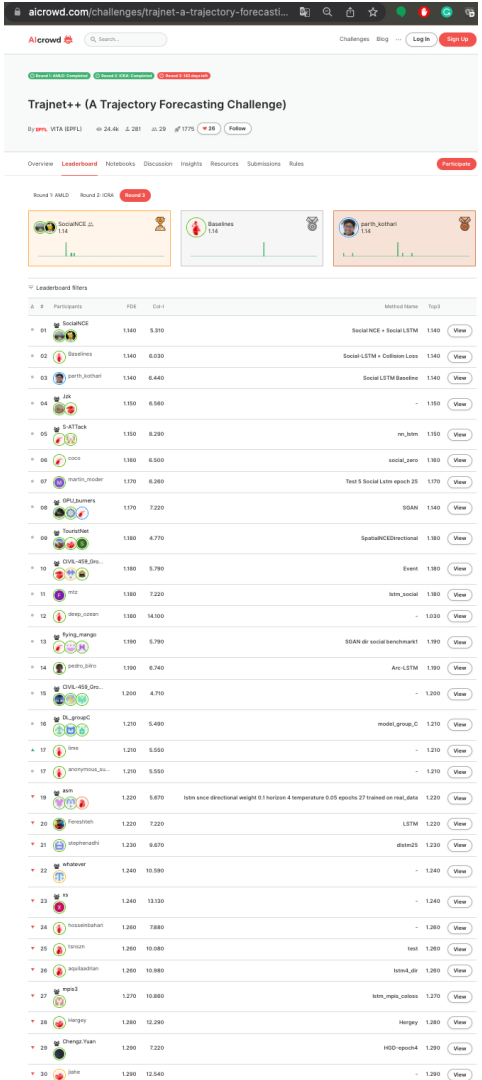
- ADE/FDE is not enough
- Top 20 is not ok
- Evaluate per category
- Evaluate awareness as well

Alexandre Alahi, EPFL

Trajnet++

- Open-source library (> 15 models)
  - https://github.com/vita-epfl/trajnetplusplusdata
- Data+evaluation protocols
- Challenge on Aicrowd
  - https://www.aicrowd.com/challenges/trajnet-a-trajectory-forecasting-challenge

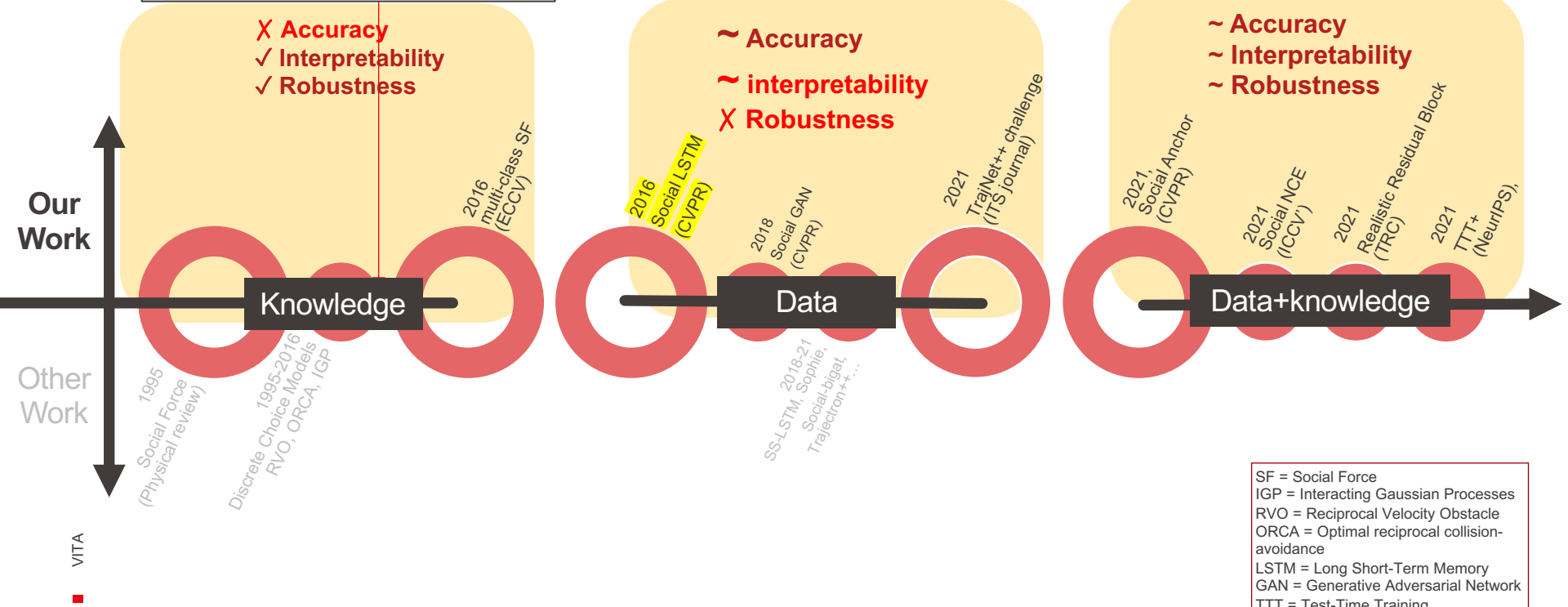
GitHub



Alexandre Alahi, EPFL

What about Robustness?

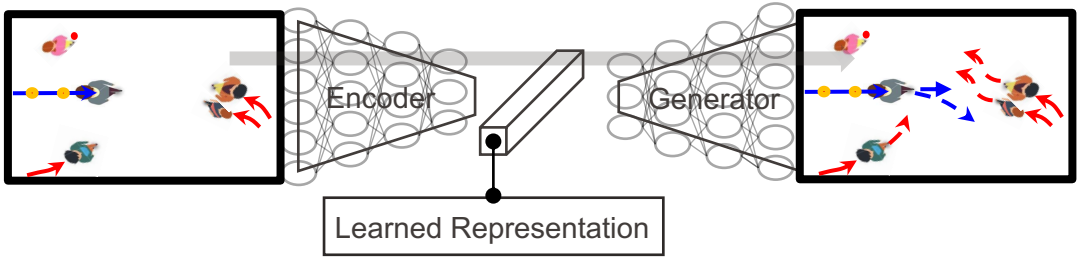
We collected the **largest** crowd dataset
(42 millions of examples)



SF = Social Force
IGP = Interacting Gaussian Processes
RVO = Reciprocal Velocity Obstacle
ORCA = Optimal reciprocal collision-avoidance
LSTM = Long Short-Term Memory
GAN = Generative Adversarial Network
TTT = Test-Time Training

Alexandre Alahi, EPFL

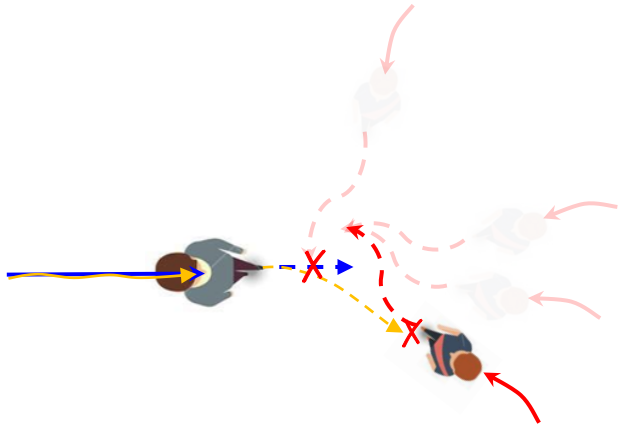
New evaluation protocol



Outcome

✓ **New evaluation** based on realistic adverserial examples [1]

✓ **Robust training**

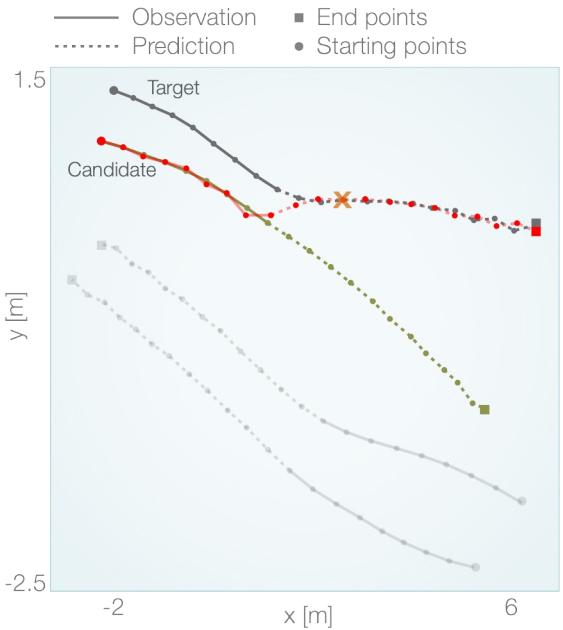


- Observed sequence
- Forecasted sequence by [2]
- Perturbed observation by < 7 cm
- Forecasted sequence leading to collision
- ✗ **Collision**

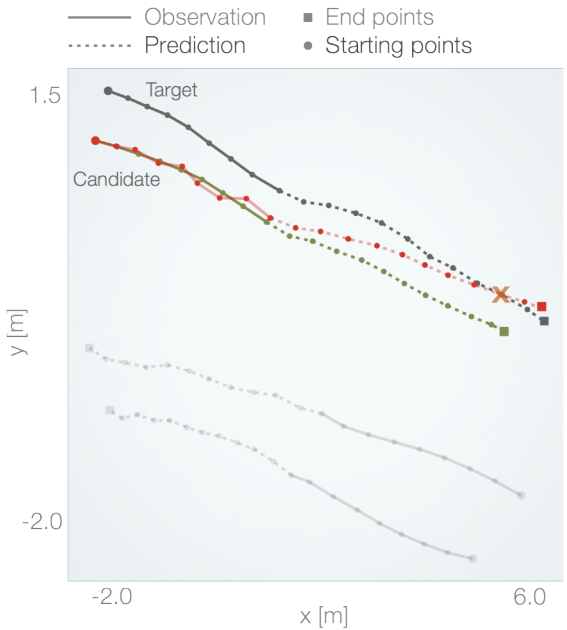
[1] S-attack library, **TRC'22**, https://s-attack.github.io/

[2] Ynet, ICCV'21, Top ranked model in Trajnet++ public challenge

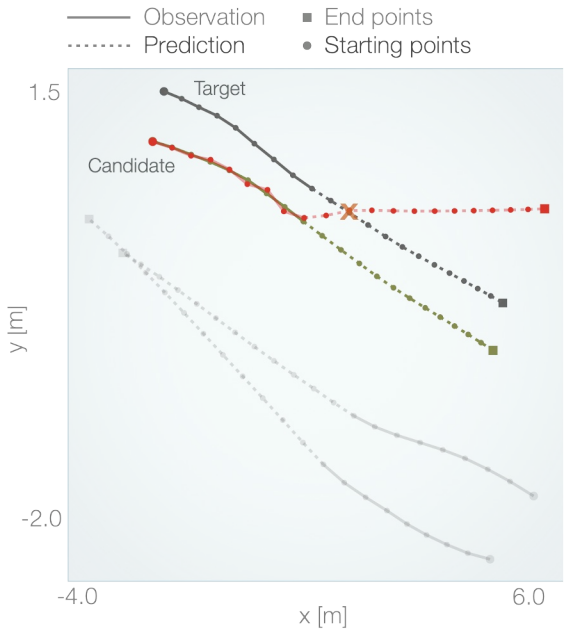
Qualitative results



D-Pool[ITS'2021]



S-STGCNN[CVPR'20]

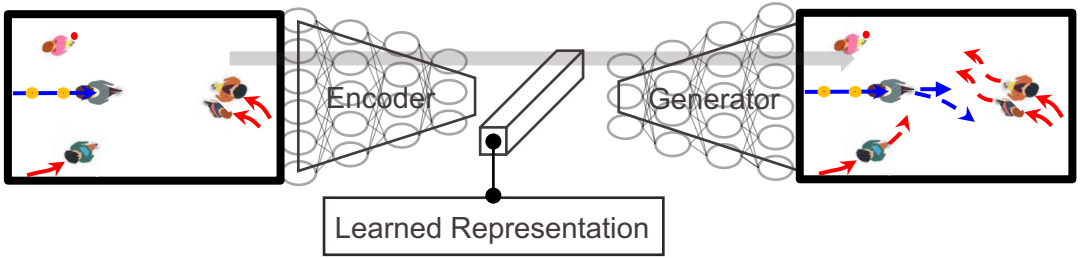


S-GAN[CVPR'18]

[1] S-attack library, https://s-attack.github.io/

Alexandre Alahi, EPFL

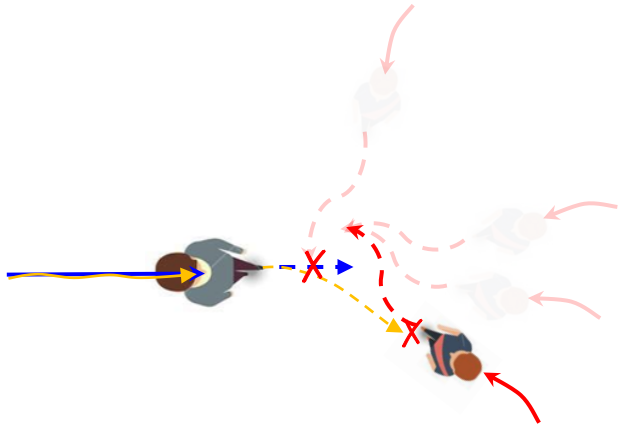
New evaluation protocol



Outcome

✓ **New evaluation** based on realistic adverserial examples [1]

✓ **Robust training**



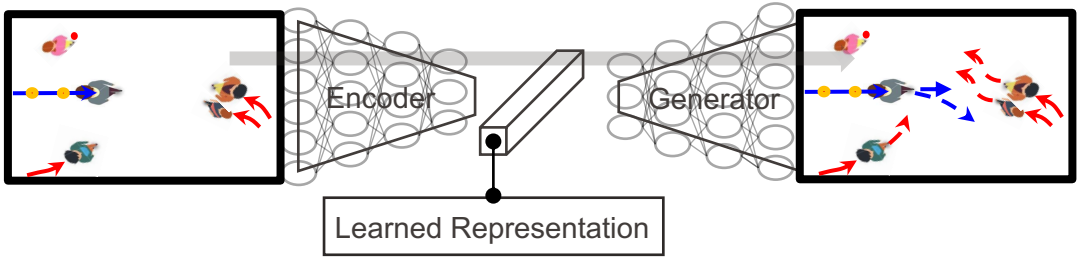
- Observed sequence
- Forecasted sequence by [2]
- Perturbed observation by < 7 cm
- Forecasted sequence leading to collision
- ✗ **Collision**

[1] S-attack library, https://s-attack.github.io/

[2] Ynet, ICCV'21, Top ranked model in Trajnet++ public challenge

Alexandre Alahi, EPFL

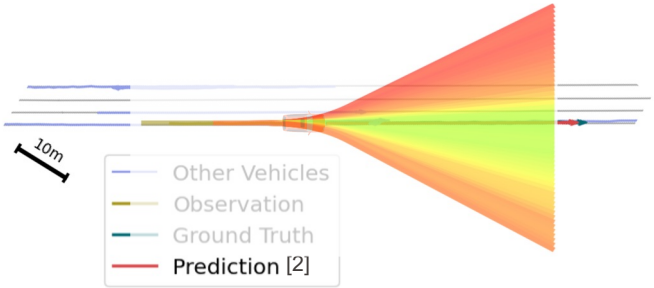
New evaluation protocol

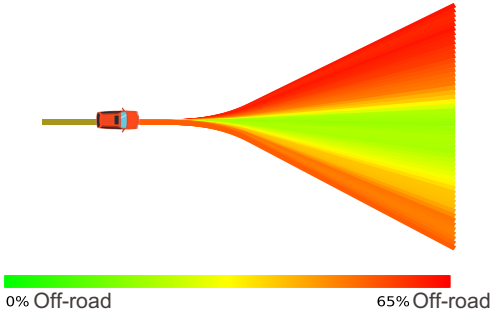


Outcome

✓ **New evaluation** based on realistic adverserial examples [1]

✓ **Robust training**





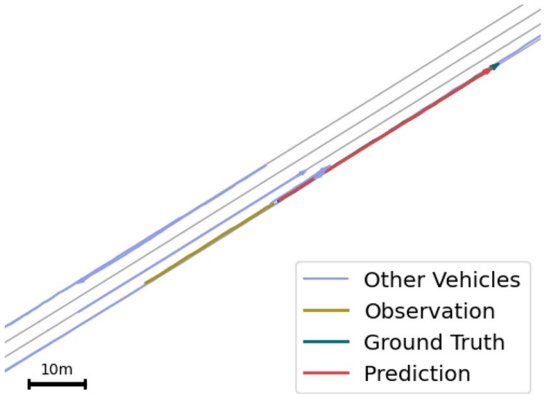
[1] Vehicle trajectory prediction works, but not everywhere, **CVPR'22**

[2] LaneGCN, ECCV'20, Top ranked model in Argoverse public challenge

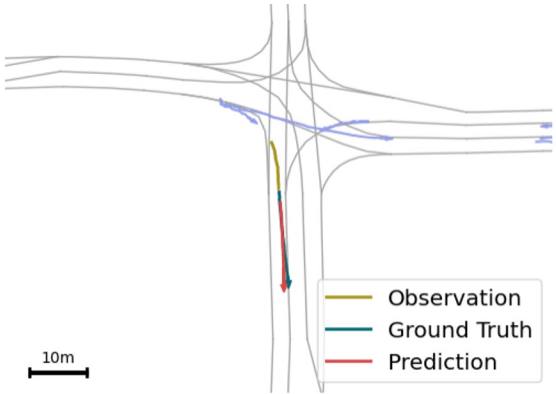
Alexandre Alahi, EPFL

Scene generation

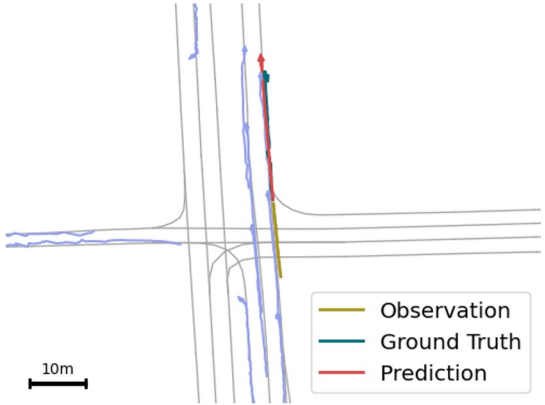
- Atomic scene generation functions



Simple turn



Double turn



Ripple road

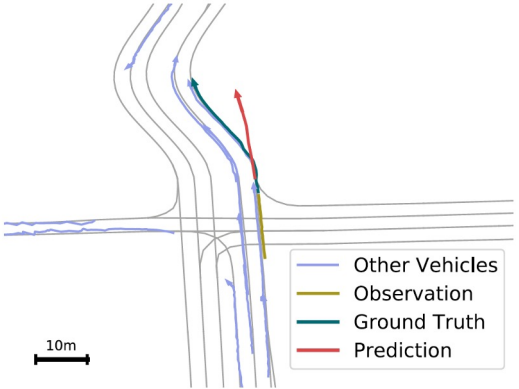
[1] Vehicle trajectory prediction works, but not everywhere, **CVPR'22**

Alexandre Alahi, EPFL

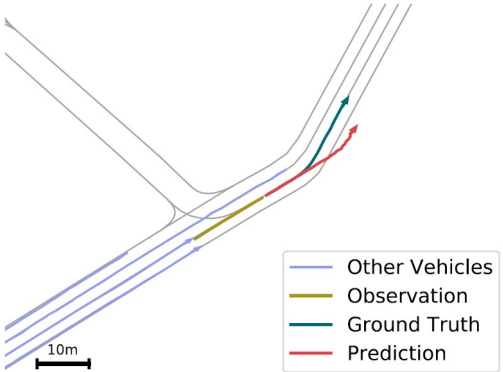
Quantitative results

Baseline	Original off-road	Generated (ours) off-road
DATF (ECCV20)	2%	82%
WIMP (arXiv20)	1%	63%
LaneGCN (ECCV'20)	1%	66%

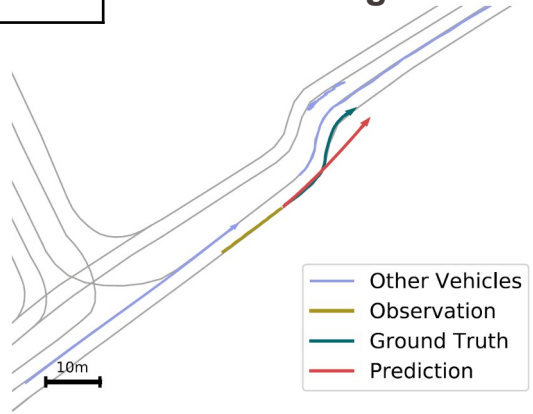
=> 46% w/ aug



(a) DATF



(b) WIMP

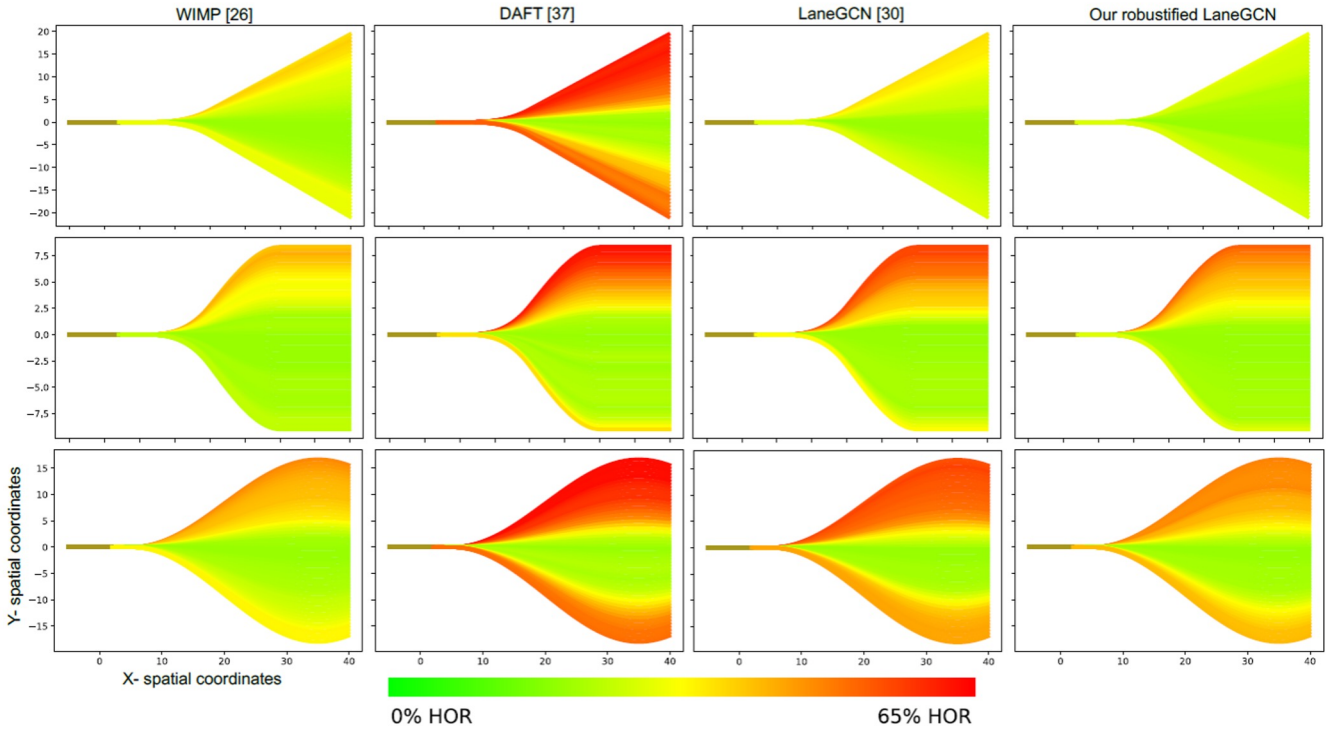


(c) LaneGCN

- [1] Vehicle trajectory prediction works, but not everywhere, **CVPR'22**

Alexandre Alahi, EPFL

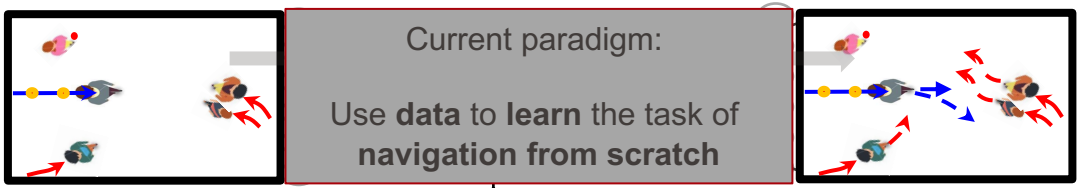
Discussions



- [1] Vehicle trajectory prediction works, but not everywhere, **CVPR'22**

Alexandre Alahi, EPFL

Current paradigm



Our proposed paradigm:

Use **data** to **only learn concepts we struggle to explain** (*e.g.*, our social norms)

✗ **Not** Robust

Because

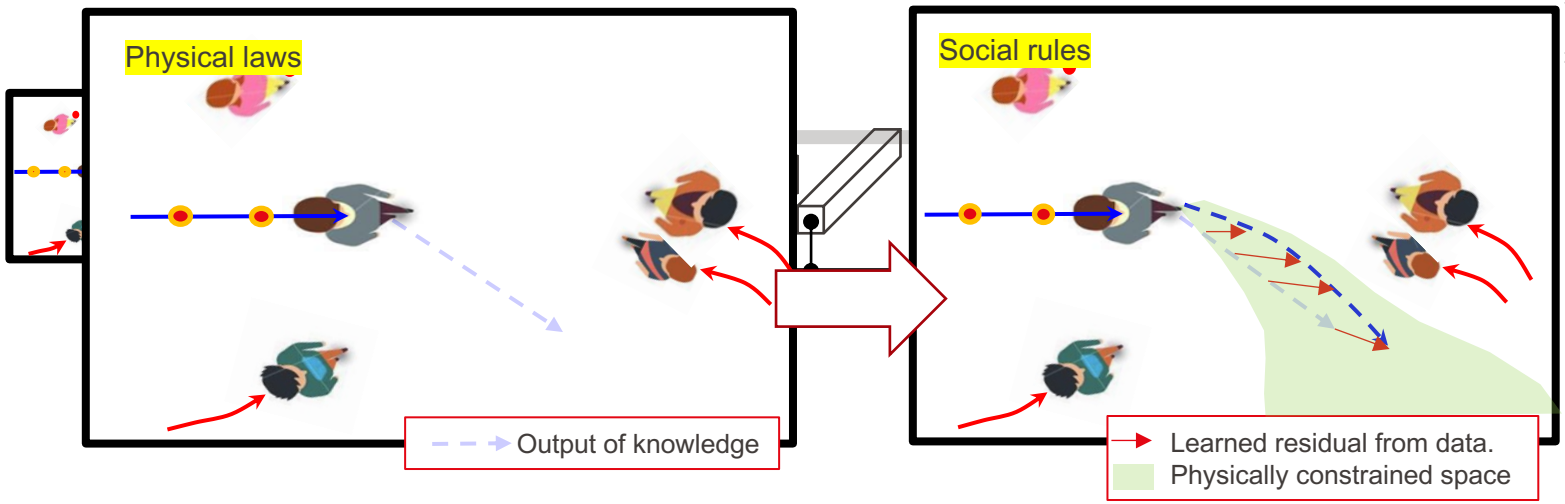
1. Imbalanced/missing data

Solution

- Knowledge-Data

Alexandre Alahi, EPFL

Proposed Knowledge-Data paradigm



Because

1. Imbalanced/missing data

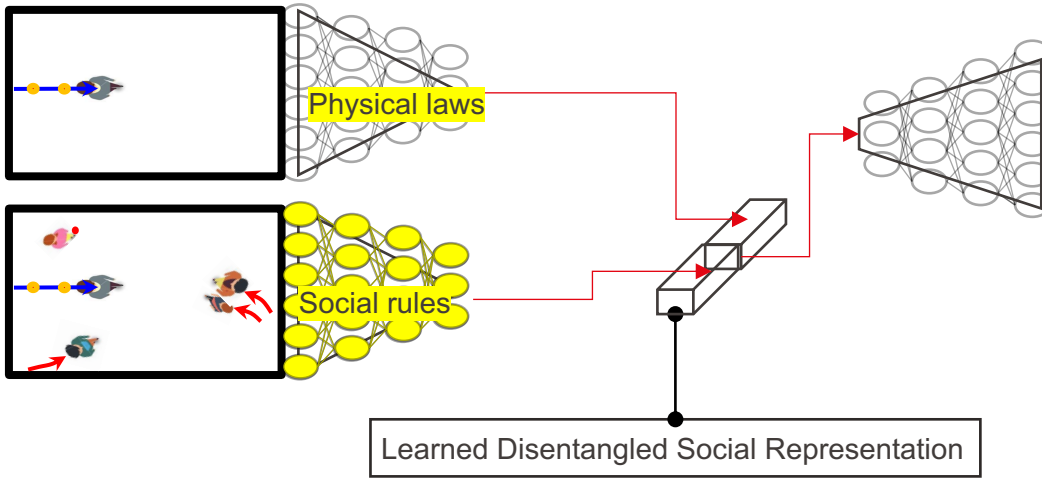
Solution

- Knowledge-Data
  - Knowledge as input

[1] Injecting knowledge in data-driven vehicle trajectory predictors, **TRC'21**

Alexandre Alahi, EPFL

Proposed Knowledge-Data paradigm



Outcome

✓ Generalizable
(low-shot transfer)

Because

1. Imbalanced/missing data

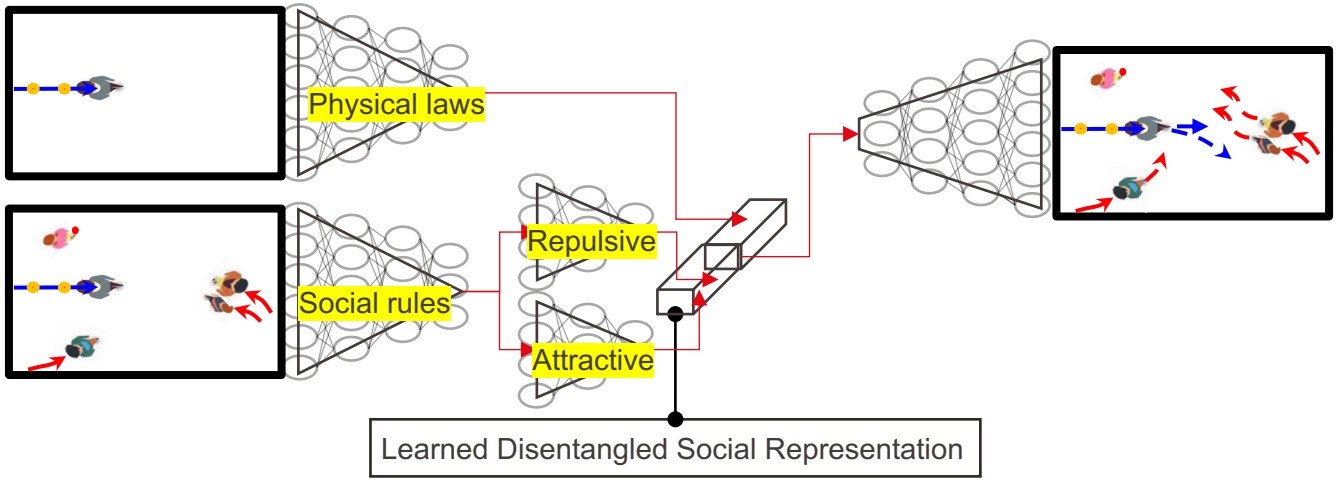
Solution

- Knowledge-Data
  - Knowledge as input
  - Knowledge within

[1] Towards Robust and Adaptive Motion Forecasting: A Causal Representation Perspective, **CVPR'22**

Alexandre Alahi, EPFL

Proposed Knowledge-Data paradigm



Because

1. Imbalanced/missing data

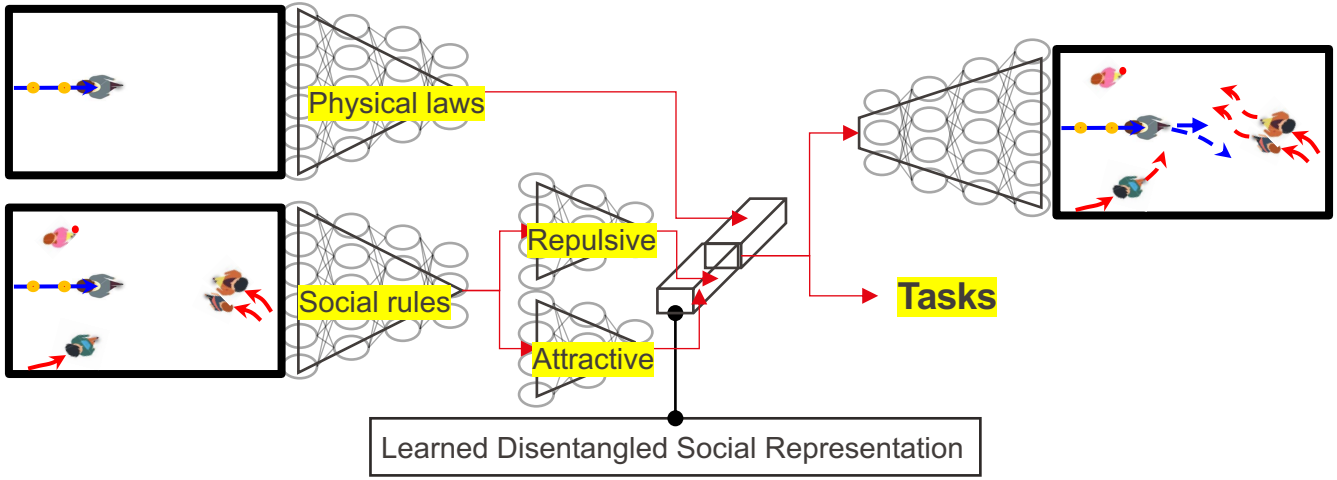
Solution

- Knowledge-Data
  - Knowledge as input
  - Knowledge within
  - Knowledge as supervision

[1] Interpretable Social Anchors, **CVPR'21**

Alexandre Alahi, EPFL

Proposed Knowledge-Data paradigm



Because

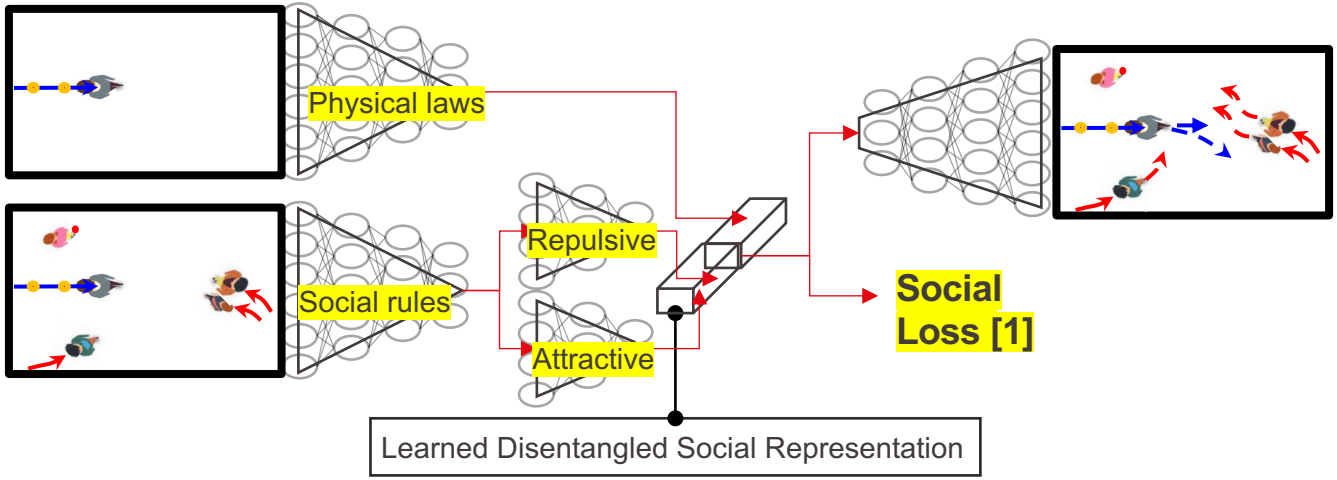
1. Imbalanced/missing data

Solution

- Knowledge-Data
  - Knowledge as input
  - Knowledge within
  - Knowledge as supervision

Alexandre Alahi, EPFL

Proposed Knowledge-Data paradigm



Because

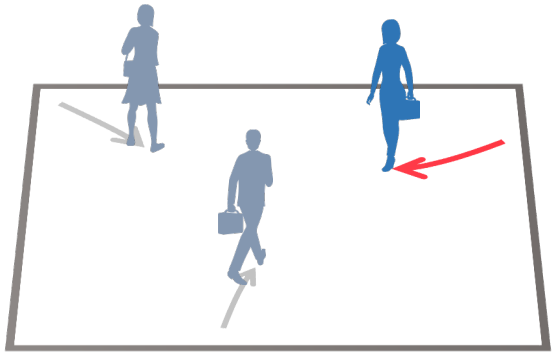
1. Imbalanced/missing data
2. Positive examples only

Solution

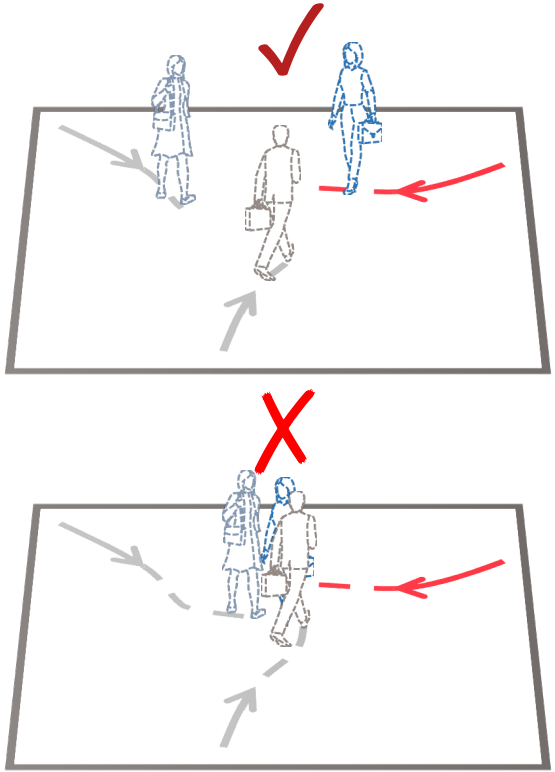
- Knowledge-Data
- w/ Opposite principle

[1] Social NCE, **ICCV'21**

Negative data augmentation



History Observation



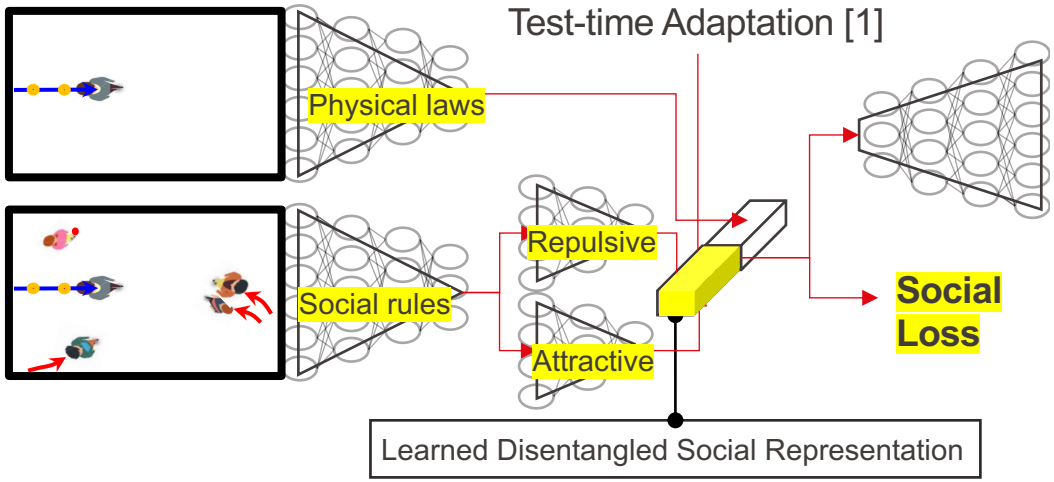
Future Prediction

[1] Social NCE, **ICCV'21**

Alexandre Alahi, EPFL

Alexandre Alahi, EPFL

Proposed Knowledge-Data paradigm



Outcome

- ✓ Robust
- ✓ Generalizable
- ✓ Interpretable

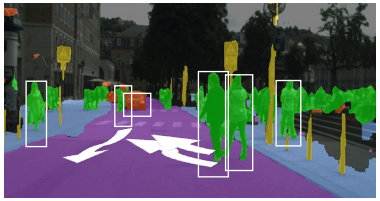
Because

1. Imbalanced/missing data
2. Positive examples only
3. Distributional shifts

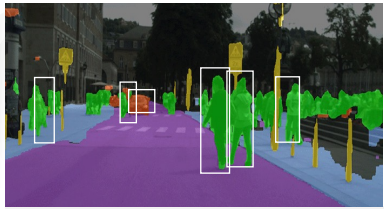
Solution

- Knowledge-Data
- w/ Opposite principle
- w/ Low-rank principle

[1] Test Time Training++, **NeurIPS'21**

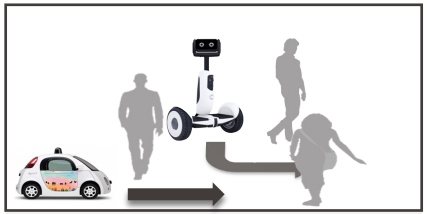


Social Forecasting



Perceiving

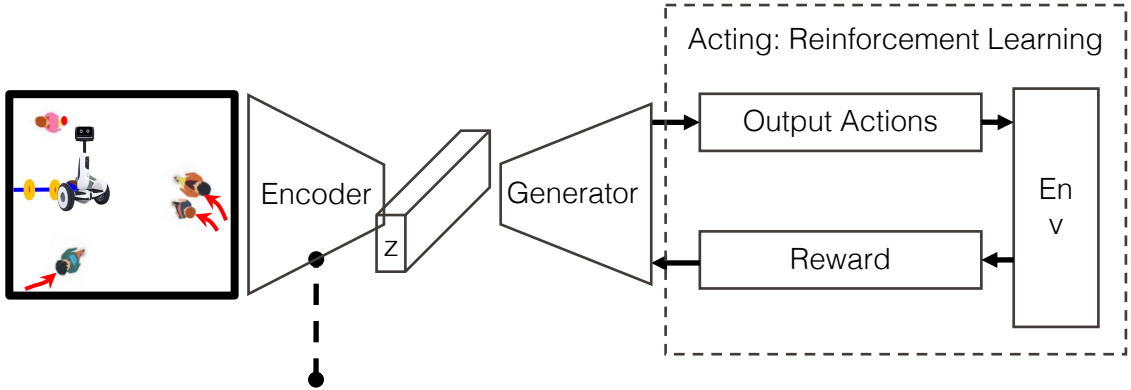
Socially-aware AI



Planning

Alexandre Alahi, EPFL

Crowd-Robot Interaction [1]



Human-Robot Interaction [1,2]
&
Our Human-Human Interaction [3]

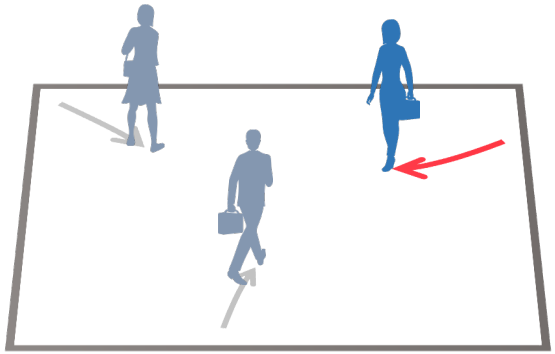
Previous works

[1] HRI, Chen, C., *et al.,*	IROS'17
[2] HRI, Everett, M., *et al.*,	IROS'18

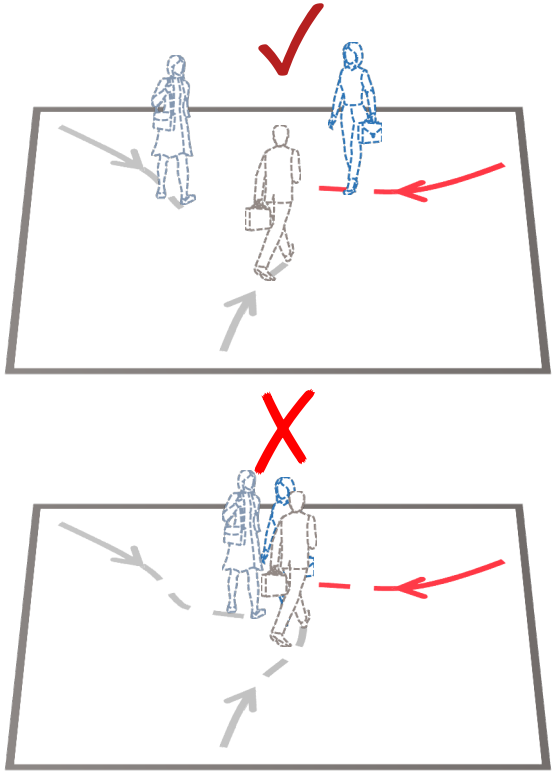
Our work

[3] Crowd-Robot Interaction,	ICRA'19

Negative data augmentation



History Observation



Future Prediction

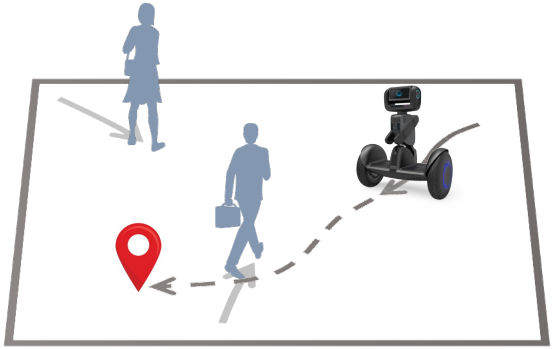
[1] Social NCE, **ICCV'21**

Alexandre Alahi, EPFL

Alexandre Alahi, EPFL

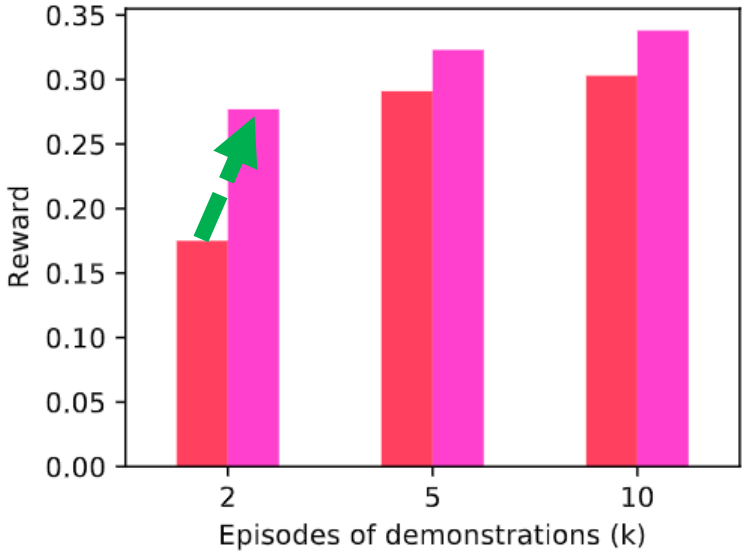
Experiments

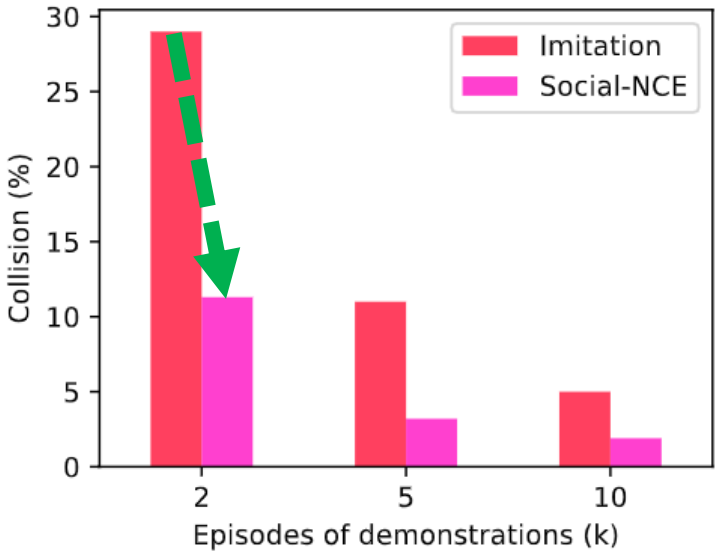
1 – Imitation Learning



[1] Social NCE, **ICCV'21**

Imitation Learning





[1] Social NCE, **ICCV'21**

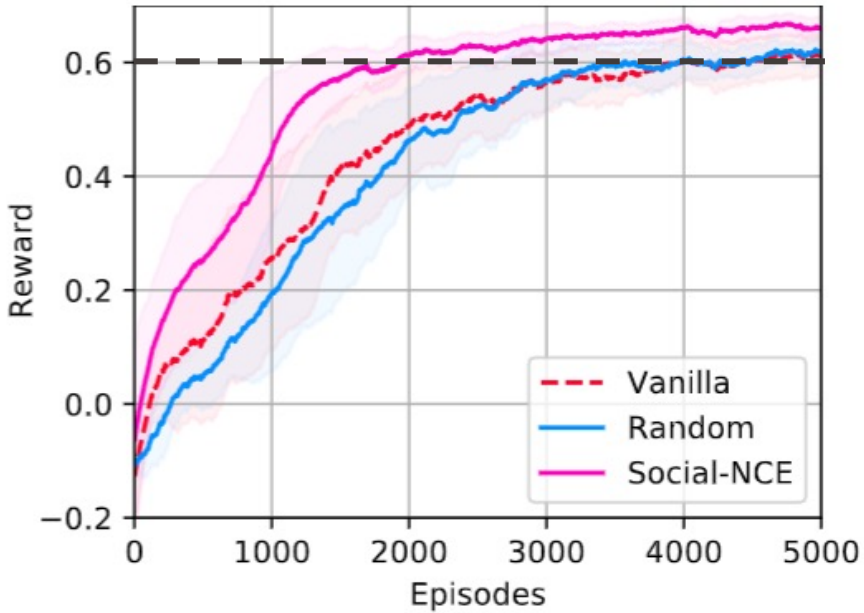
Alexandre Alahi, EPFL

Alexandre Alahi, EPFL

Experiments

2 – Reinforcement Learning

Reinforcement Learning

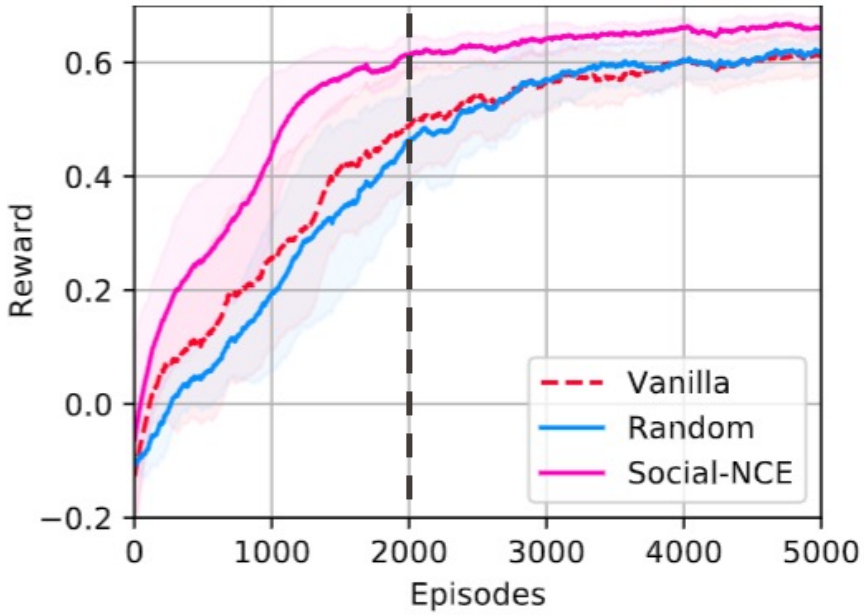


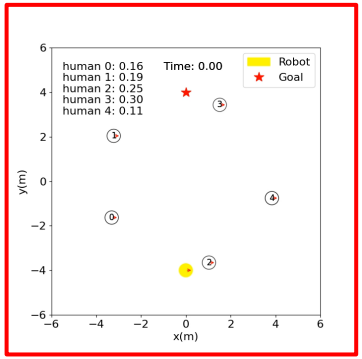
⟵ Collision-free

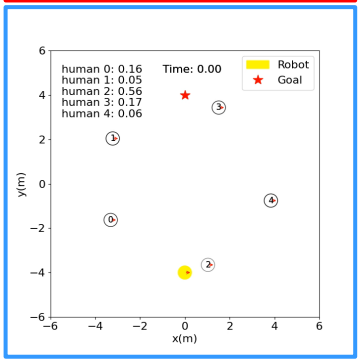
Alexandre Alahi, EPFL

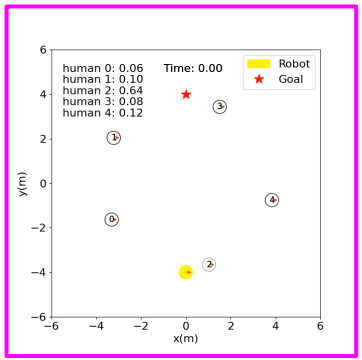
[1] Social NCE, **ICCV'21**

Reinforcement Learning



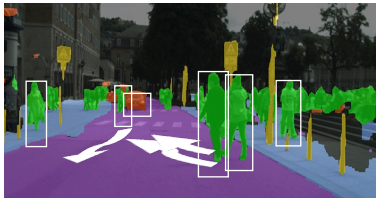




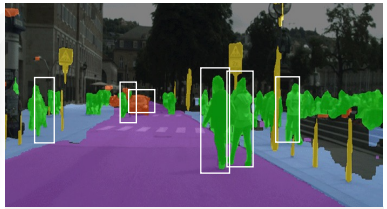


- [1] Social NCE, **ICCV'21**

Alexandre Alahi, EPFL

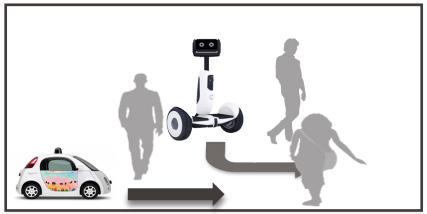


Social Forecasting



Perceiving

Socially-aware AI



Planning

#Open Science

Perception:
[1] S. Kreiss et al., OpenPifPaf **library** for pose estimation, **CVPR'19, ICCV'21 (licensed)**
[2] L. Bertoni et al., 3D perception **library**, **ICCV'19, ICRA'21**
[3] L. Bertoni et al., Perceiving Social Distancing, **ITS'20**
[4] G. Adaimi et al., Deep Visual Re-identification with Confidence, **TRC'21**
[5] T. Mordan et al., Detecting 32 human attributes, **ITS'21**
Prediction:
[6] Kothari et al., Trajnet++ **library** for spatio-temporal forecasting tasks (>15 implemented models)
[7] Kothari et al., Social Anchor, **ICCV'21**
[8] Liu et al., Social NCE, **ICCV'21**
Planning:
[9] C. Chen et al., Crowd-Robot Interaction, **ICRA'19**
Generative models:
[10] Y. Liu* et al., Collaborative Sampling in GAN, **AAAI'20**
[11] A. Carlier et al., Deep SVG, **NeurIPS'20**
DCM + NN
[12] B. Sifringer et al., L-MNL, **TRB'20**
Test-time training:
[13] Y. Liu* et al., TTT++, **NeurIPS'21**
Tools
[14] Video Ultimate labeling

GitHub

Code on-line: **vita.epfl.ch/code**

Alexandre Alahi, EPFL

EPFL VITA

Visual Intelligence for Transportation

SAMSUNG Schindler RICHEMONT CarPostal La classe jaune. HITACHI HONDA The Power of Dreams Valeo

Innosuisse FNSNF HASLERSTIFTUNG Horizon 2020