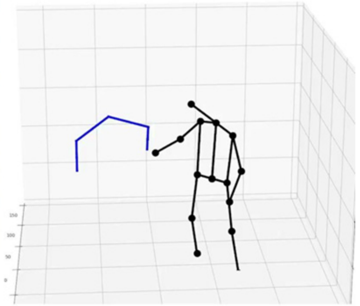
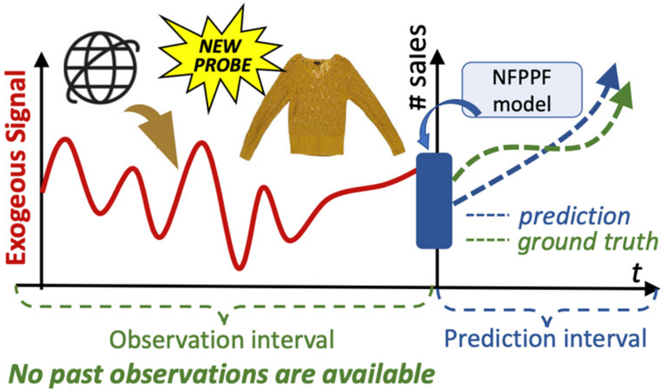




FORECAST THE FORECASTING

MARCO CRISTANI

ICPRAM 2024

13th International Conference on Pattern Recognition Applications and Methods

Rome, Italy

24 - 26 February, 2024

UNIVERSITÀ DEGLI STUDI DI VERONA

TWO PERSPECTIVES

ACADEMY + INDUSTRY

- Professor @UNIVR

UNIVERSITÀ di VERONA
Department of ENGINEERING FOR INNOVATION MEDICINE

INTELLIGO labs

- Focus on forecasting, papers at Intl' Journal of Forecasting, ECCV, CVPR, TPAMI, Pattern Recog.,
- PATCAST workshop at ICPR 2020

- Co-founder of Humatics srl (exit in 2021)
  - Forecasting products sold to:

sirmoney #LiberaMente
DOPPELGÄNGER
NUNALIE
SYS-DAT GROUP

- Co-founder of Qualyco srl ('23)
  - Forecasting for anomaly detection

OUTLINE

- Forecasting, nowadays
- *Forecast the forecasting*

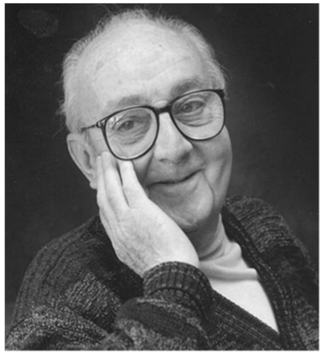
INTRODUCTION TO (STATISTICAL) FORECASTING

SOME HISTORY

(7th century b.C.) looking into the bowels of animals...

$$y'_t = c + \phi_1 y'_{t-1} \dots \phi_p y'_{t-p} + \theta_1 \varepsilon_{t-1} \dots \theta_q \varepsilon_{t-q} + \varepsilon_t$$

ARIMA



George E.P. Box
1919 - 2013

Rob J. Hyndman
1967-

Forecasting: Principles and Practice (2nd ed)
Rob J Hyndman and George Athanasopoulos
Monash University, Australia

https://otexts.com/fpp2/

FORECASTING GOALS, METRICS

- Goal: observe *T* values of a **time series**, predict *N* values

- Error measures:

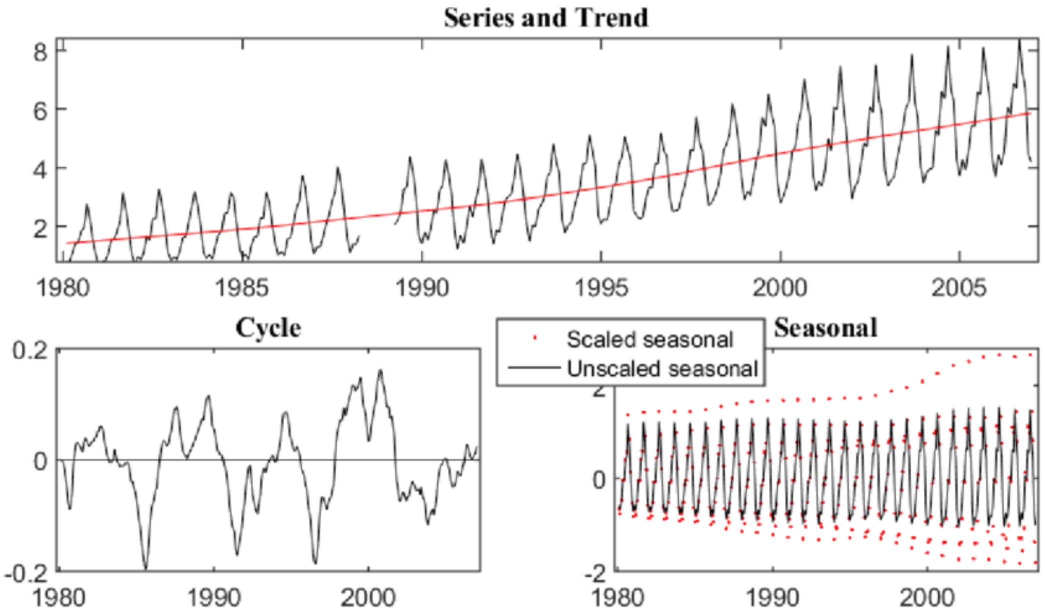
$$E_t = \underset{\substack{\downarrow \\ \text{Ground Truth}}}{Y_t} - \underset{\substack{\downarrow \\ \text{Forecast}}}{F_t}$$

$$MAE = MAD = \frac{\sum_{t=1}^{N} |E_t|}{N}$$

[Box and Jenkins 1970] -- George E.P. Box and Gwilym Jenkins, ‘Time series analysis: forecasting and control’, *Wiley*, 1970

TYPICAL STATISTICAL FORECASTING ELEMENTS

- To properly forecast a signal, patterns have to be identified:
  - **Trend:** a long-term increase or decrease in the data;
  - **Seasonality:** when a time series is affected by seasonal factors such as the time of the year;
  - **Cycle:** when the data exhibit rises and falls at no specific frequency.



[Box and Jenkins 1970] -- George E.P. Box and Gwilym Jenkins, 'Time series analysis: forecasting and control', *Wiley,* 1970

THE WORKHORSE OF STAT. FORECASTING: SARIMAX

Sales

$$F_t = c + \underbrace{\sum_{n=1}^{p} \alpha_n d_{t-n} + \sum_{n=1}^{q} \theta_n \epsilon_{t-n}}_{\text{trend}} + \underbrace{\sum_{n=1}^{r} \beta_n x_{n_t}}_{\text{exogenous}} + \underbrace{\sum_{n=1}^{P} \phi_n d_{t-sn} + \sum_{n=1}^{Q} \eta_n \epsilon_{t-sn} + \epsilon_t}_{\text{seasonality}}$$

++

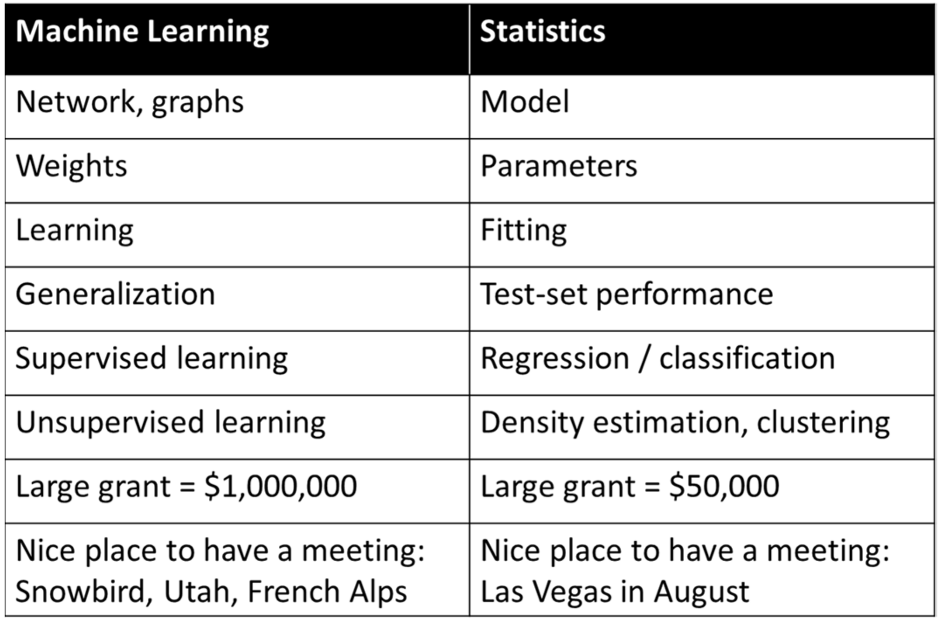
MACHINE LEARNING VS STATISTICAL FORECASTING

A CLASH WITHIN COMMUNITIES

Bontempi, G., Ben Taieb, S., & Le Borgne, Y. A. (2013). Machine learning strategies for time series forecasting. *Business Intelligence: Second European Summer School, eBISS 2012,*

Shereen Elsayed et al. Do We Really Need Deep Learning Models for Time Series Forecasting? (October 2021)

https://towardsdatascience.com/time-series-forecasting-deep-learning-vs-statistics-who-wins-c568389d02df

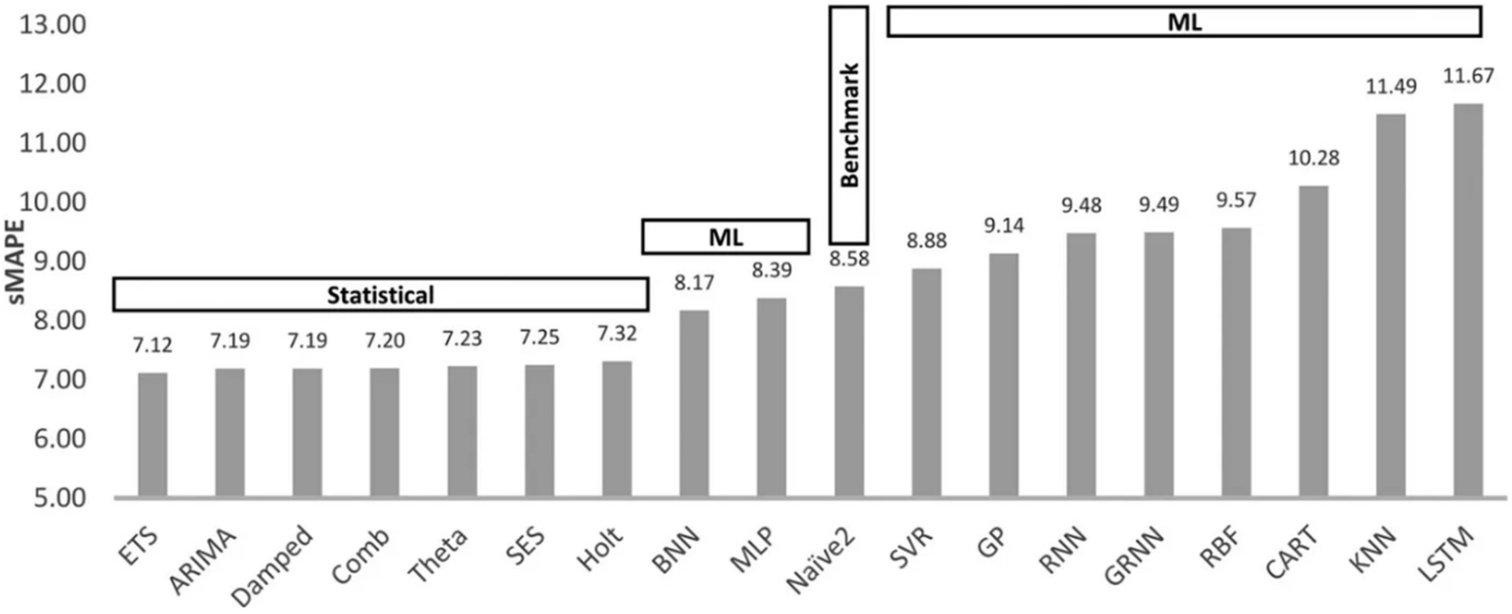


Machine Learning	Statistics
Network, graphs	Model
Weights	Parameters
Learning	Fitting
Generalization	Test-set performance
Supervised learning	Regression / classification
Unsupervised learning	Density estimation, clustering
Large grant = $1,000,000	Large grant = $50,000
Nice place to have a meeting: Snowbird, Utah, French Alps	Nice place to have a meeting: Las Vegas in August

https://windowsontheory.org/2022/06/20/the-uneasy-relationship-between-deep-learning-and-classical-statistics/

MACHINE LEARNING VS STATISTICAL, 2018

FIRST COMPARISON IN 2018, THE M4 COMPETITION



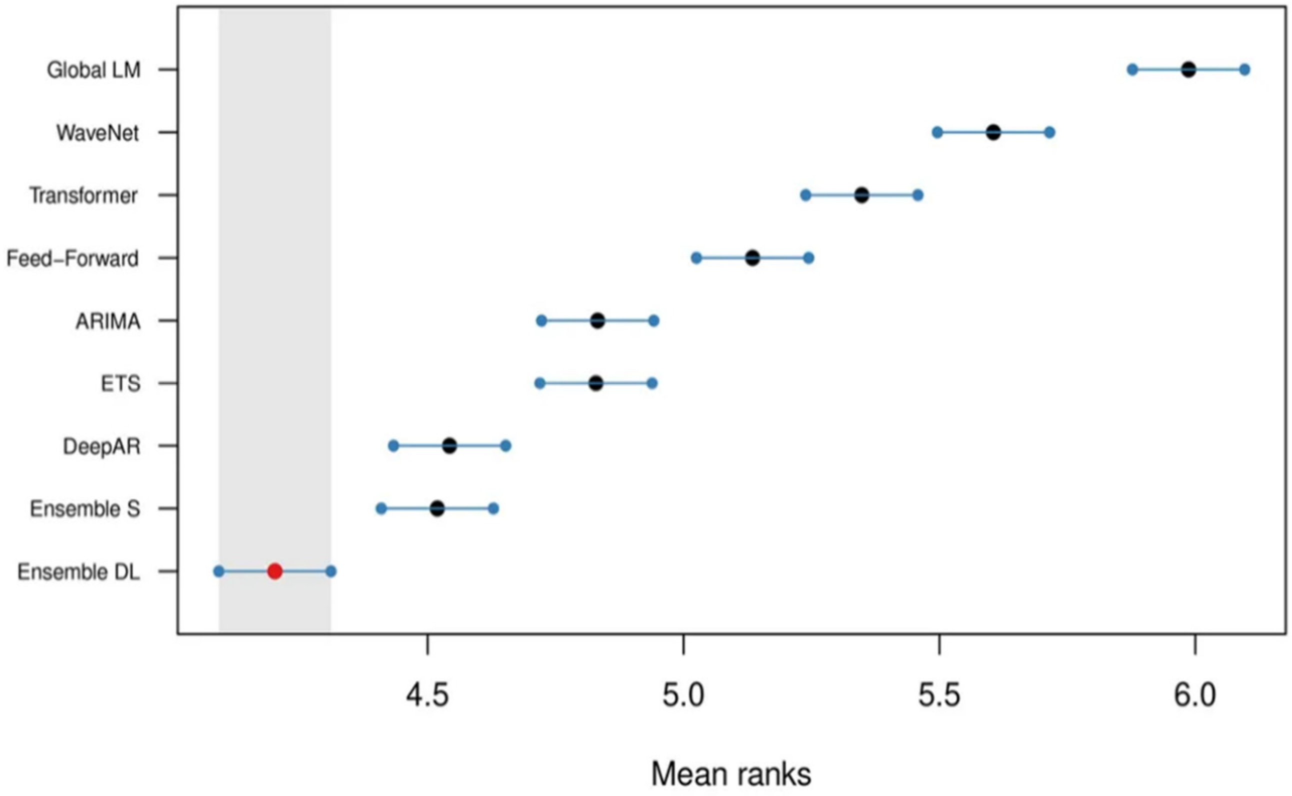
Forecasting accuracy (sMAPE) of the eight statistical and the ten ML forecasting methods examined by Makridakis et al. back in 2018. All ML methods occupied the last places.

Makridakis, S., Spiliotis, E., & Assimakopoulos, V. (2018). The M4 Competition: Results, findings, conclusion and way forward. *International Journal of Forecasting*, *34*(4), 802-808.

https://towardsdatascience.com/time-series-forecasting-deep-learning-vs-statistics-who-wins-c568389d02df

MACHINE LEARNING VS STATISTICAL, 2022

COMPARISON ON THE M3 DATA2022

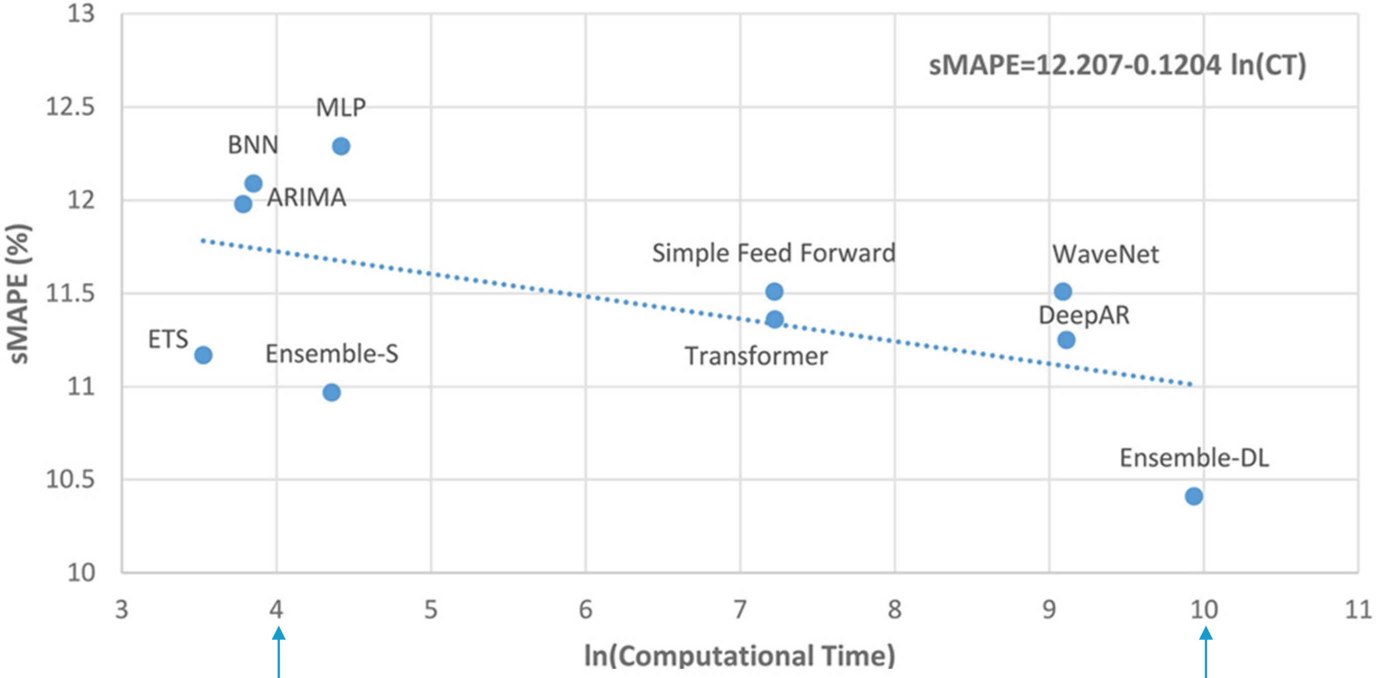


Makridakis et al. Statistical, machine learning and deep learning forecasting methods: Comparisons and ways forward (August 2022)

- The Ensemble-DL consists of 200 models, with 50 models from each category: DeepAR, Transformer, WaveNet, and MLP (github.com/gjmulder/m3-gluonts-ensemble).
- Ensemble-S consists of statistical models.

MACHINE LEARNING VS STATISTICAL, COSTS

COMPUTATIONAL EFFORT



10 minutes

15 days

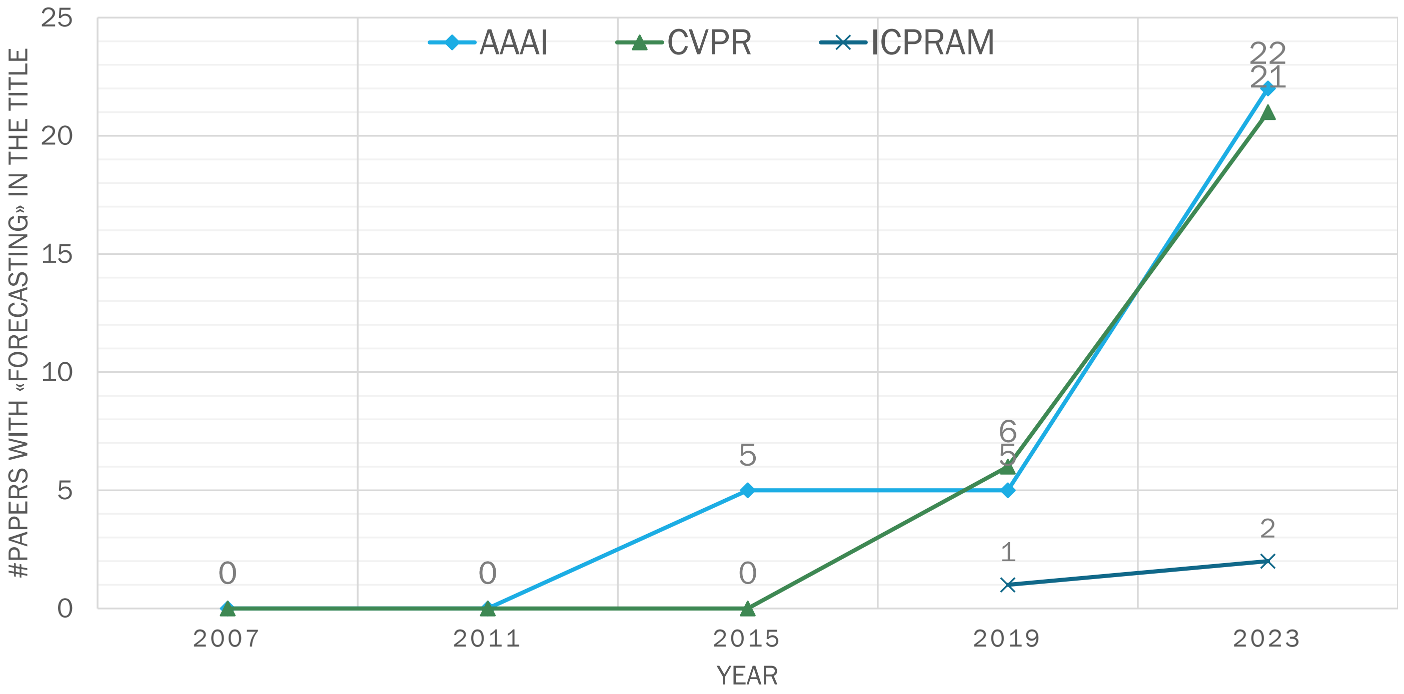
Makridakis et al. Statistical, machine learning and deep learning forecasting methods: Comparisons and ways forward (August 2022)

INTELLIGO labs

	+	−
Machine Learning	High-Dimensional Data Non-linearity Adaptability Feature Discovery Anomaly Detection	Lack of interpretability Large amount of data to train
Statistical	Intepretability Smaller dataset Well-defined relationships Hypothesis testing Stability No overfitting	Low-dimensional data Simpler problems Outliers have an impact

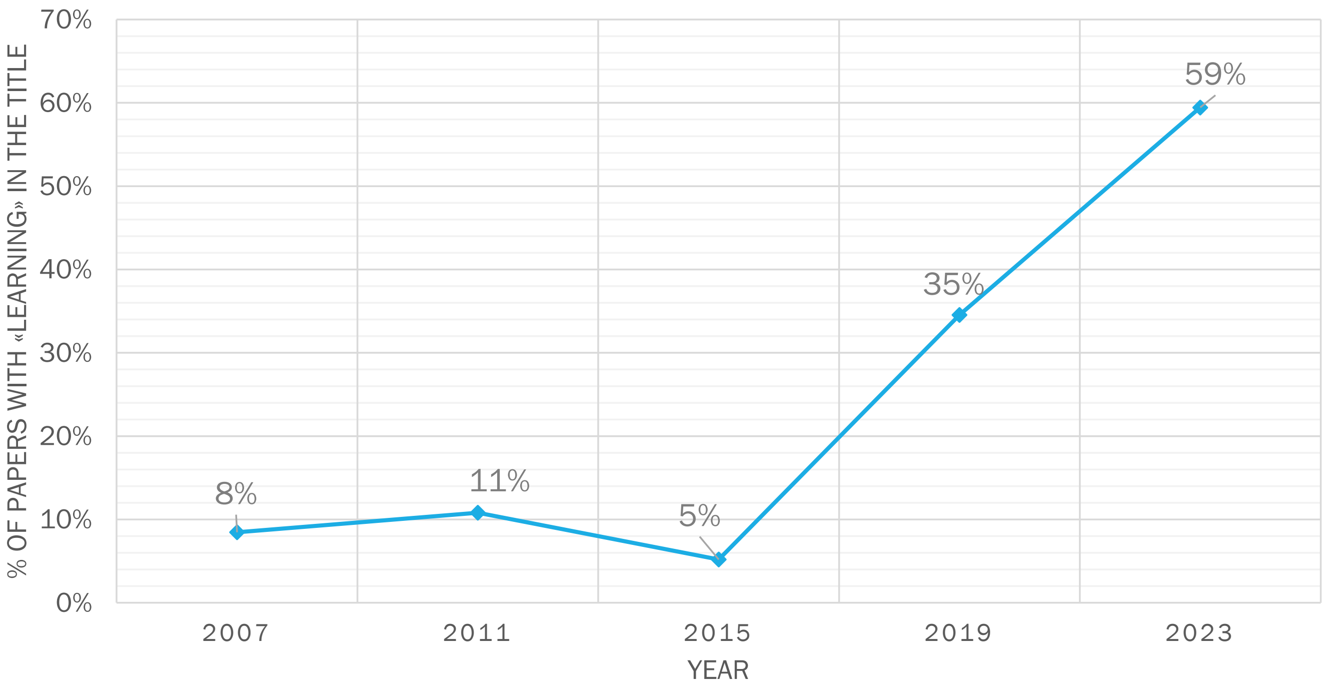
WHERE TO PUBLISH ML-BASED FORECASTING?

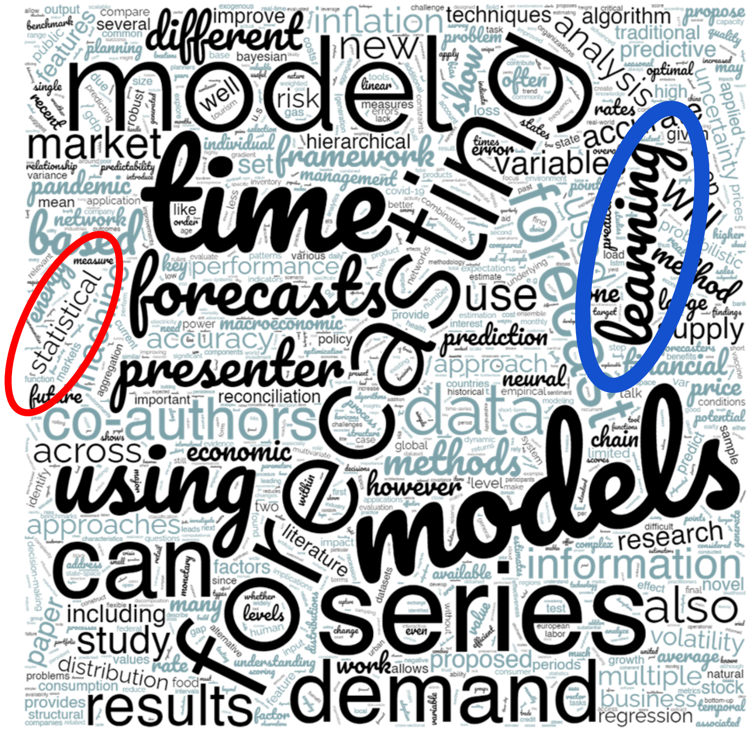
- Machine learning, pattern recognition venues



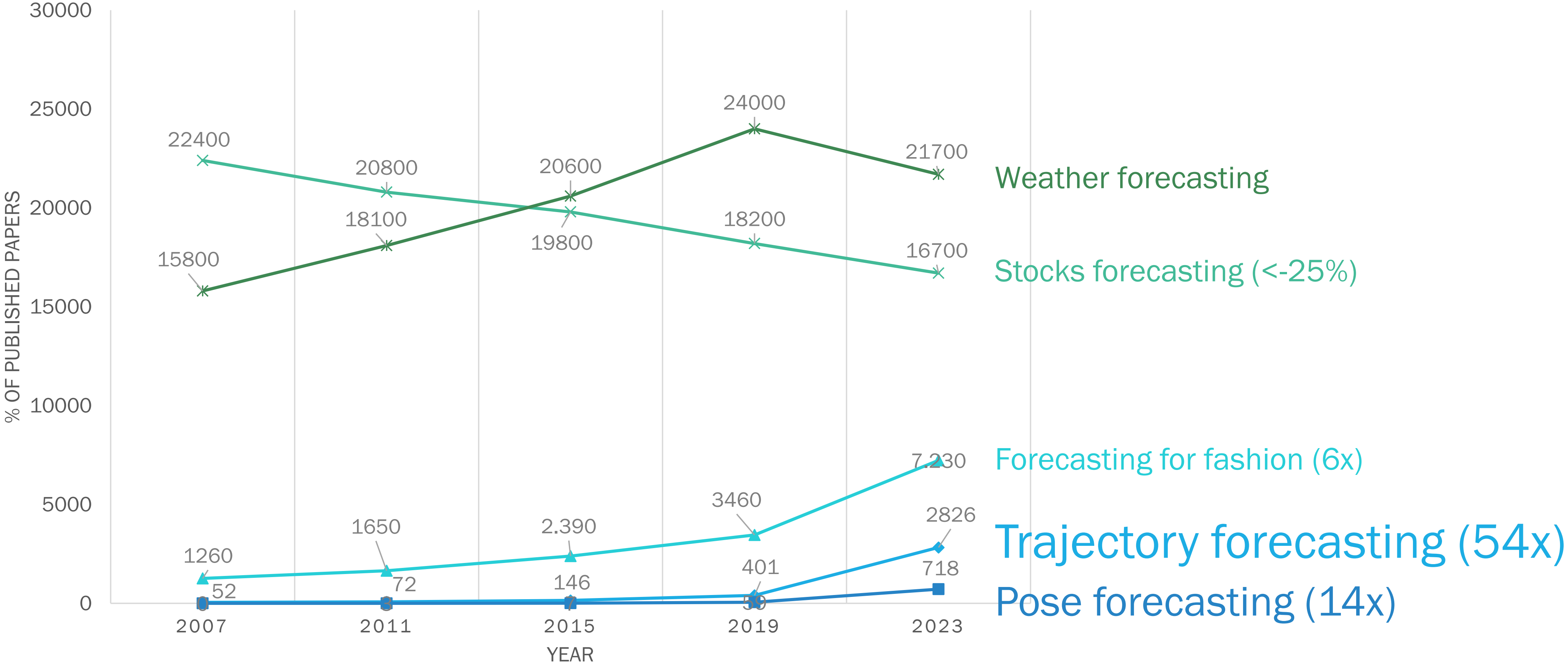
WHERE TO PUBLISH ML-BASED FORECASTING

- International symposium on forecasting





FORECASTING APPLICATIONS



OUTLINE

- Introduction to forecasting
- *Forecast the forecasting*
  - People trajectory forecasting

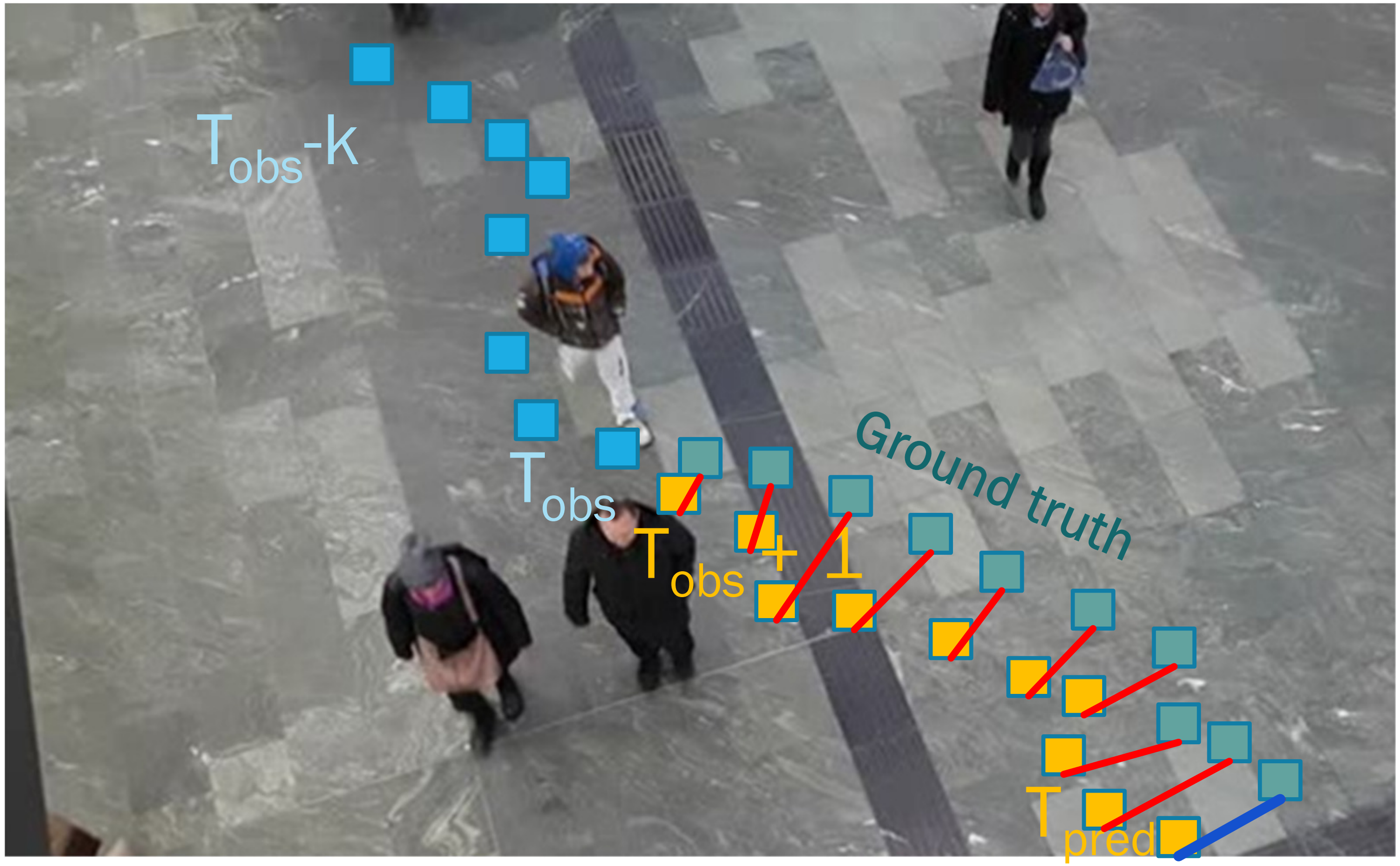
PEOPLE TRAJECTORY FORECASTING

Goal

- Observe *k* steps until T_{obs} and predict T_{pred} steps
- Compute errors over *n* people (ADE, FDE)

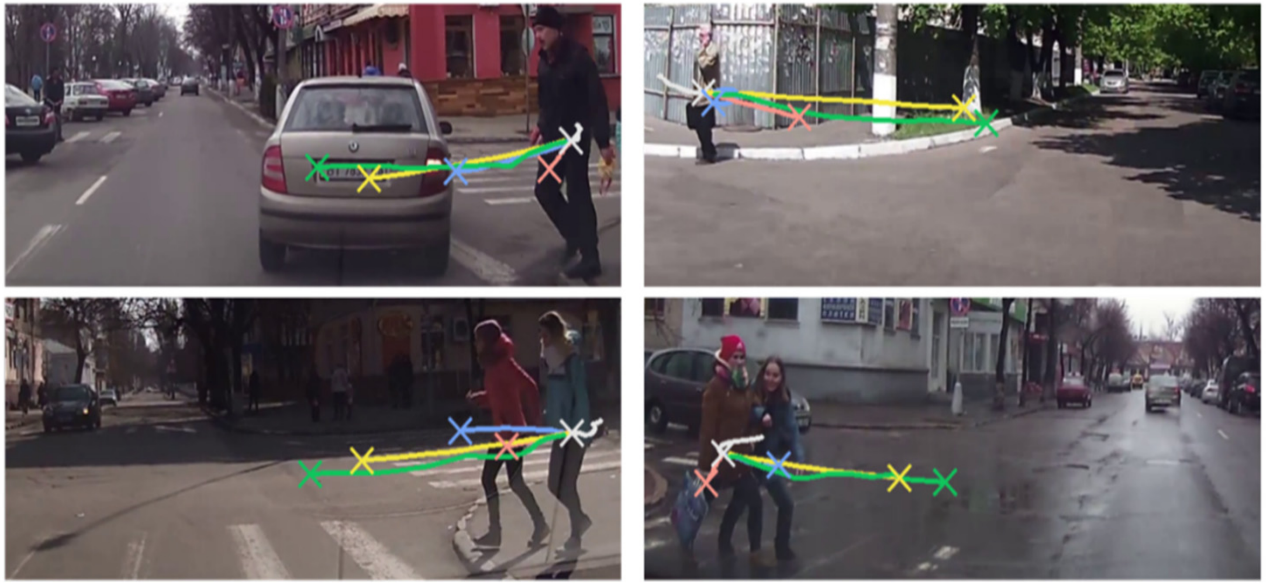
$$ADE = \frac{\sum_{i=1}^{n} \sum_{t=T_{obs}+1}^{T_{pred}} \left[(\hat{x}_i^t - x_i^t)^2 + (\hat{y}_i^t - y_i^t)^2 \right]}{n(T_{pred} - (T_{obs}+1))}$$

$$FDE = \frac{\sum_{i=1}^{n} \sqrt{(\hat{x}_i^{T_{pred}} - x_i^{T_{pred}})^2 + (\hat{y}_i^{T_{pred}} - y_i^{T_{pred}})^2}}{n}$$

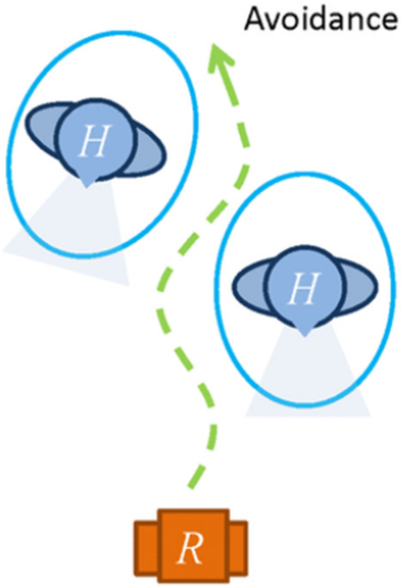


APPLICATIONS OF PEOPLE TRAJ. FORECASTING

https://paperswithcode.com/search?q=author%3AJianru+Xue



Self driving cars H. Ben-Younes, ´ Eloi Zablocki, P. P´erez, M. Cord, Driving behavior explanation with multi-level fusion, Pattern Recognition 123 (2022)



Human-robot interaction

S. Ding, et al., Simultaneous body part and motion identification for human-following robots, Pattern Recognition 50 (2016)

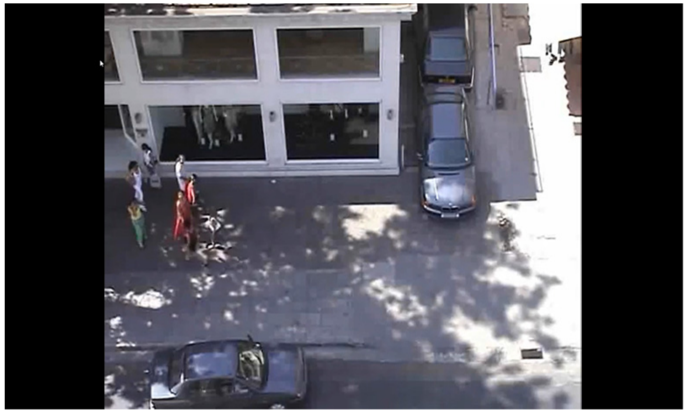
PEOPLE TRAJECTORY FORECASTING 101

MORE COMPLICATED THAN STANDARD FORECASTING? YES

- Key factors to take into account:

	1. The statistics of the single trajectories	S. Zamboni, Z. T. Kefato, S. Girdzijauskas, C. Nor´en, L. Dal Col, Pedestrian trajectory prediction with convolutional neural networks, Pattern Recognition 121 (2022)
SCENE	2. The geometry of the scene	H. Zhao, R. P. Wildes, Where are you heading? dynamic trajectory prediction with expert goal examples, in: ICCV, 2021.
	3. Prior knowledge on target points	Mangalam, K., An, Y., Girase, H., & Malik, J. (2021). From goals, waypoints & paths to long term human trajectory forecasting. In ICCV 2021
PEOPLE	4. People do not collide	A. Alahi, K. Goel, V. Ramanathan, A. Robicquet, L. FeiFei, S. Savarese, Social LSTM: Human trajectory prediction in crowded spaces, in: CVPR, 2016
	5. People could be in a group	Bae, I., Park, J. H., & Jeon, H. G. (2022, October). Learning Pedestrian Group Representations for Multi-modal Trajectory Prediction. In ECCV 2022
	6. Body cues are predictive	Hasan, F. Setti, T. Tsesmelis, V. Belagiannis, S. Amin, A. Del Bue, M. Cristani, F. Galasso, Forecasting people trajectories and head poses by jointly reasoning on tracklets and vislets, IEEE TPAMI (2019)

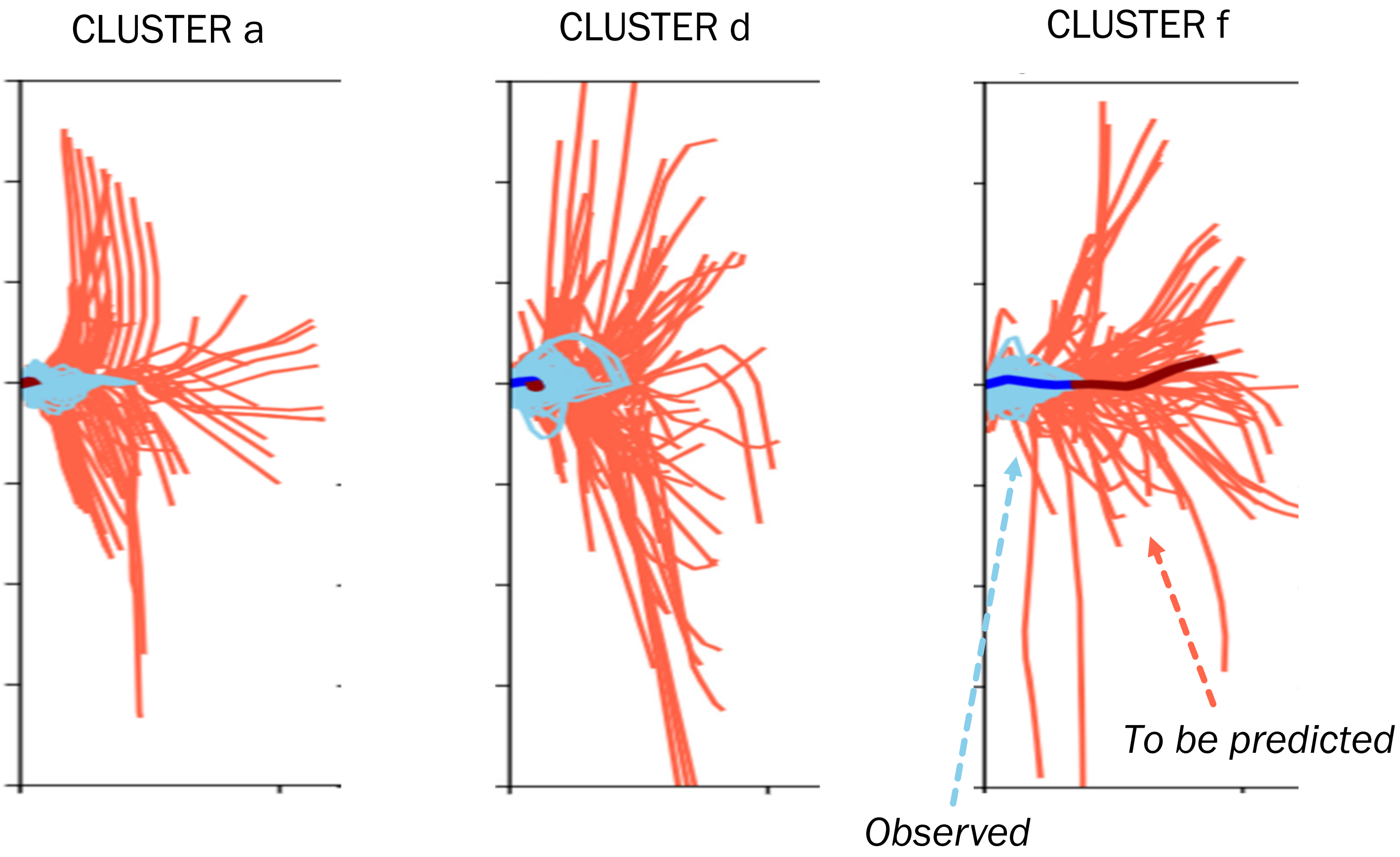
THE STATISTICS OF THE SINGLE TRAJECTORIES



The UCY benchmark ~2000 traj.

Lerner, A., Chrysanthou, Y., & Lischinski, D. (2007, September). Crowds by example. In *Computer graphics forum* (Vol. 26, No. 3)

Franco, L., Placidi, L., Giuliari, F., Hasan, I., Cristani, M., & Galasso, F. (2023). Under the hood of transformer networks for trajectory forecasting. *Pattern Recognition*

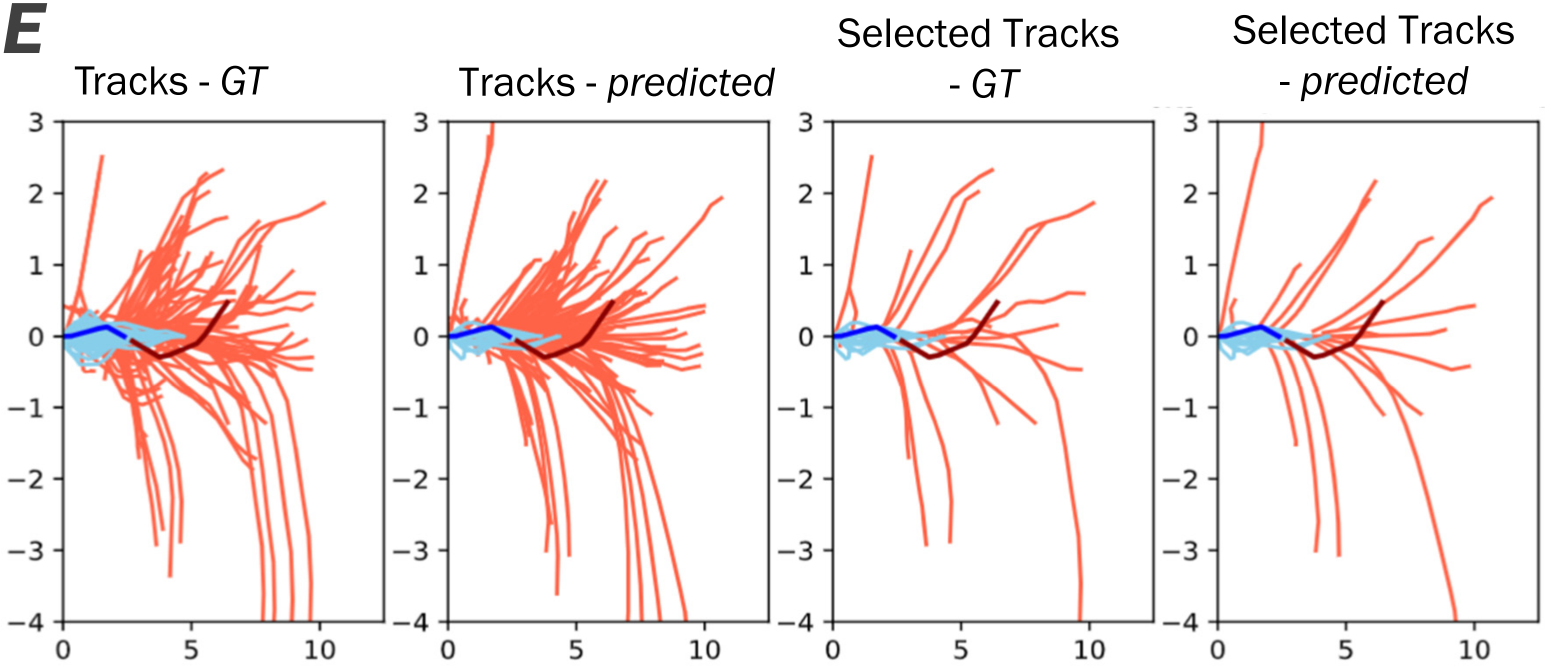


TRAJ. = LANGUAGE

- a trajectory is made by curves, slow down, etc.;

=

- a string is made by symbols



Tracks - *GT*

Tracks - *predicted*

Selected Tracks - *GT*

Selected Tracks - *predicted*

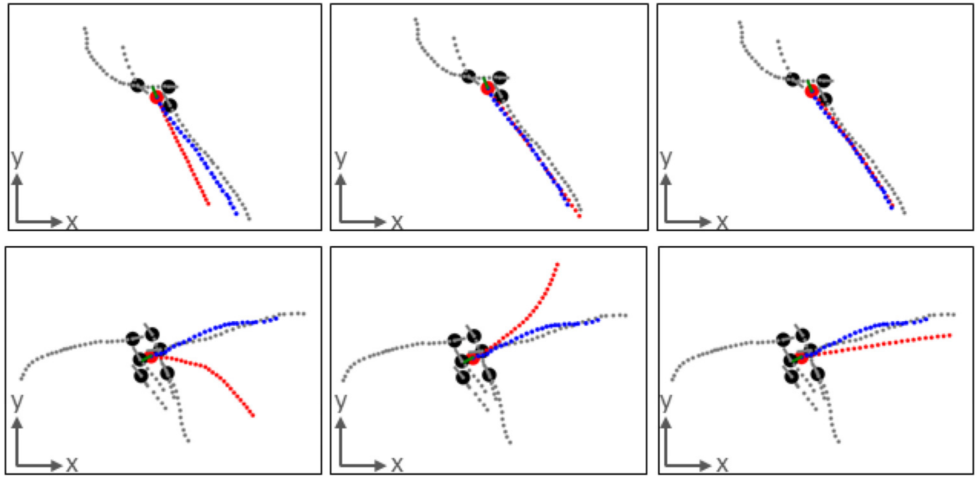
- Trasformer-based architectures are well-suited (ICPR 2020, *Pattern Recognition ‘23*).

code here!

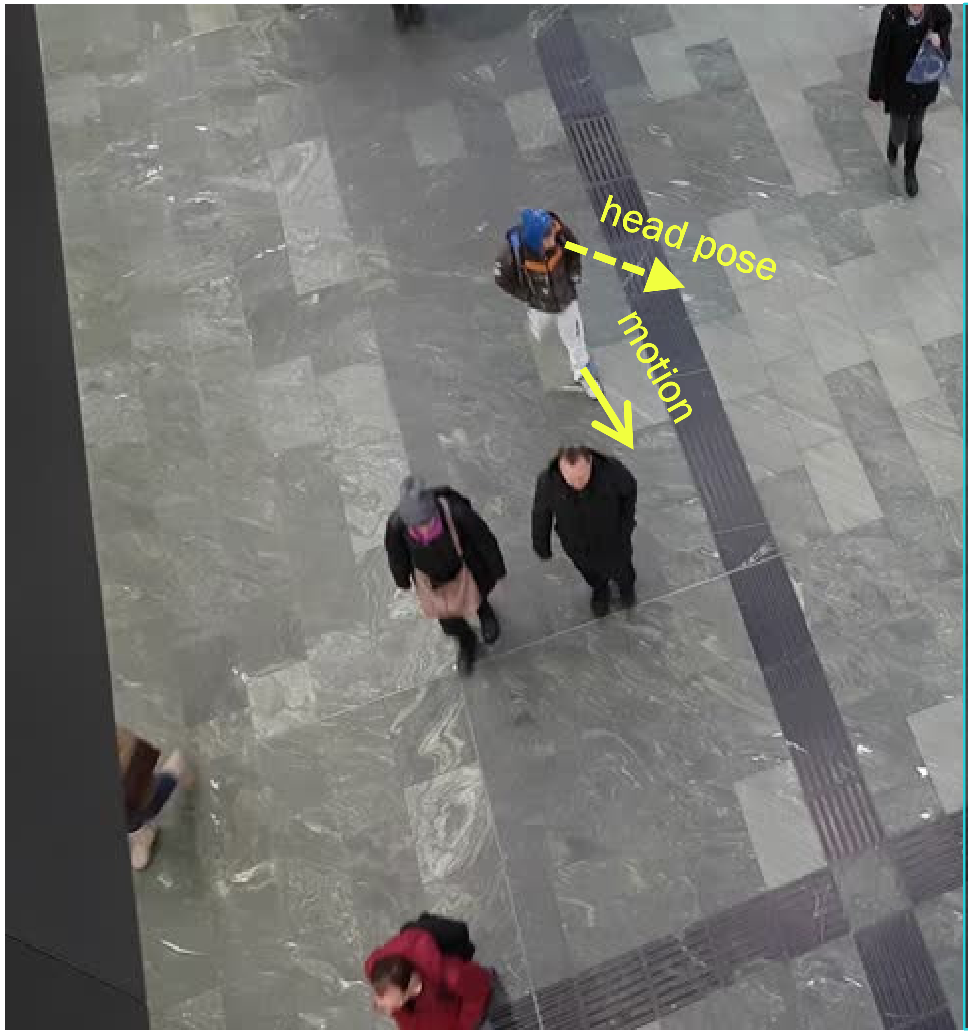
	Linear	LSTM-based					TF-based
	Individual	Individual		Social		Soc.+ map	Individual
	Interpolat.	LSTM [5]	S-GAN-ind [5]	Social LSTM [5]	Soc. Att. [6]	Trajectron++ [7]	Trasformer TF (ours)
avg	0.79/1.59	0.70/1.52	0.74/1.54	*0.72/1.54*	*0.30/2.59*	*0.34/0.84*	**0.54/1.17**

BODY CUES ARE PREDICTIVE

- Main aspect to take into account: **people are not just points!**
- Modeling the body is important since it contains cues which are predictive;
- The head direction is an example;
- Head direction could be captured at low resolutions.



Cormier, M., Clepe, A., Specker, A., & Beyerer, J. (2022). Where are we with human pose estimation in real-world surveillance?. In *Proceedings of the IEEE/CVF Winter Conference on Applications of Computer Vision* (pp. 591-601).

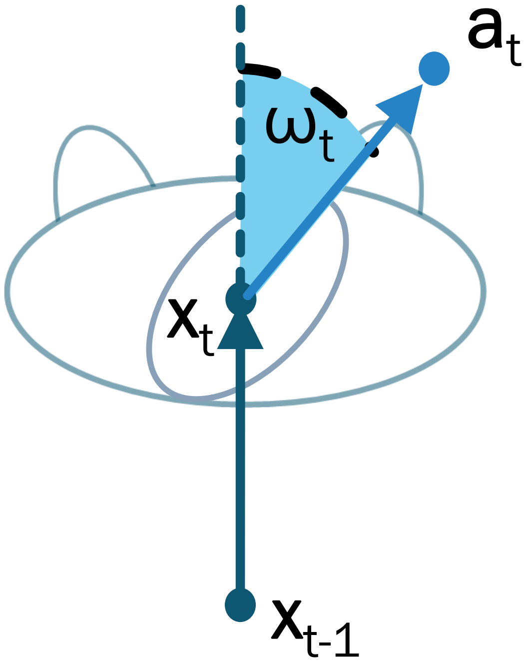


- Investigate the connection between the oriented motion and the head pose (CVPR 2018, PAMI 20)

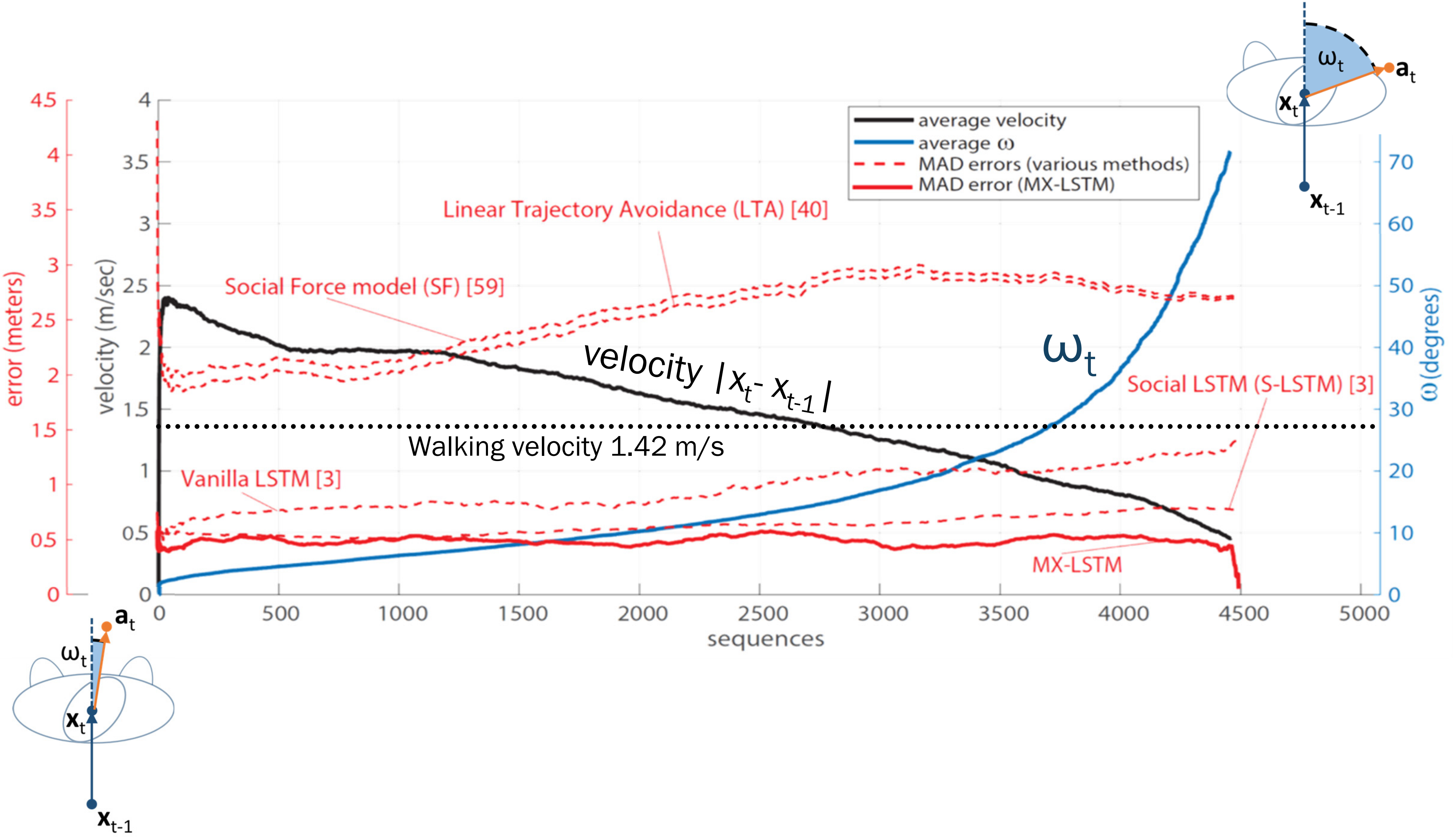
- Hypothesis: the head pose helps in predicting the future trajectory

ALIGNMENT HEAD ORIENTATION - MOTION

A QUANTITATIVE STUDY ON UCY DATASET

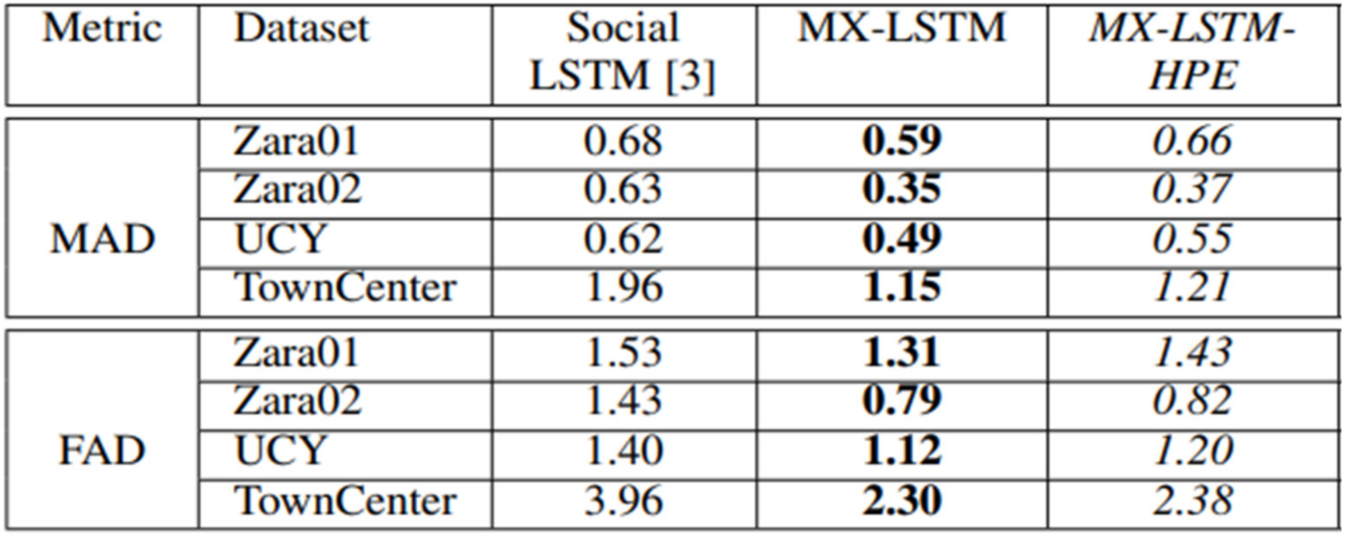


- $x_{t-1}x_t$: motion from *t-1* to *t*
- x_ta_t : head orientation at *t*
- ω_t: alignment



45
4
35
3
25
2
15
1
05
0
error (meters)
4
3.5
3
2.5
2
1.5
1
0.5
0
velocity (m/sec)
0
500
1000
1500
2000
2500
3000
3500
4000
4500
5000
sequences
70
60
50
40
30
20
10
0
ω (degrees)
average velocity
average ω
MAD errors (various methods)
MAD error (MX-LSTM)
Linear Trajectory Avoidance (LTA) [40]
Social Force model (SF) [59]
velocity $|x_t - x_{t-1}|$
ω_t
Social LSTM (S-LSTM) [3]
Walking velocity 1.42 m/s
Vanilla LSTM [3]
MX-LSTM
ω_t
a_t
x_t
x_{t-1}
ω_t
a_t
x_t
x_{t-1}

RESULTS



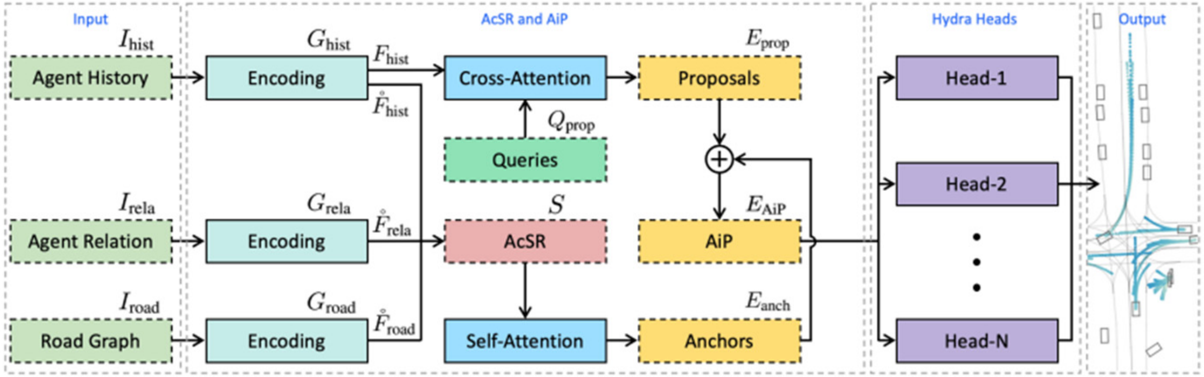
Metric	Dataset	Social LSTM [3]	MX-LSTM	*MX-LSTM-HPE*
MAD	Zara01	0.68	**0.59**	*0.66*
	Zara02	0.63	**0.35**	*0.37*
	UCY	0.62	**0.49**	*0.55*
	TownCenter	1.96	**1.15**	*1.21*
FAD	Zara01	1.53	**1.31**	*1.43*
	Zara02	1.43	**0.79**	*0.82*
	UCY	1.40	**1.12**	*1.20*
	TownCenter	3.96	**2.30**	2.38

Observation period is 3.2s and the forecasting horizon is 4.8s.
ADE errors, in meters. Less is better ↓

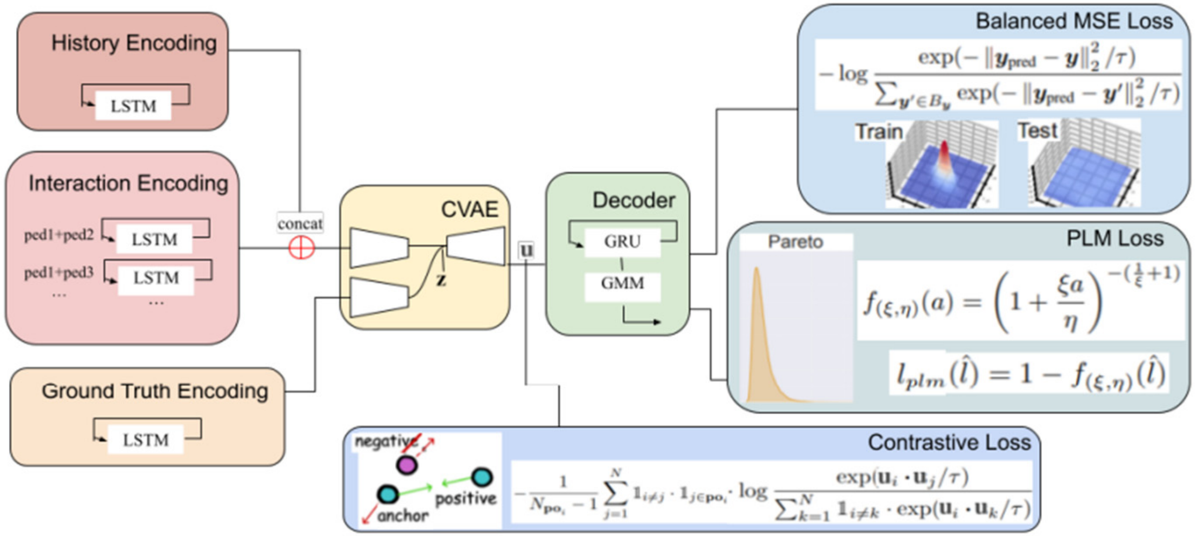
Hasan, I., Setti, F., Tsesmelis, T., Del Bue, A., Galasso, F., & Cristani, M. (2018). Mx-lstm: mixing tracklets and vislets to jointly forecast trajectories and head poses. In *CVPR 2018*

SOTA

- SotA is made by very structured approaches.
- Architectures which are very effective backbones: GCNN (LaneGCN ‘20, BANet ’22, PAGA ‘22), Transformer (vanilla ’20, SceneTransformer ‘21, Wayformer ‘22)
- Models which seem promising: TrackGPT (Feb.’24), Diffusion models (Jan ‘24), long tailed learning (Feb.’24)



Wang, X., Su, T., Da, F., & Yang, X. (2023). ProphNet: Efficient agent-centric motion forecasting with anchor-informed proposals. In *CVPR*

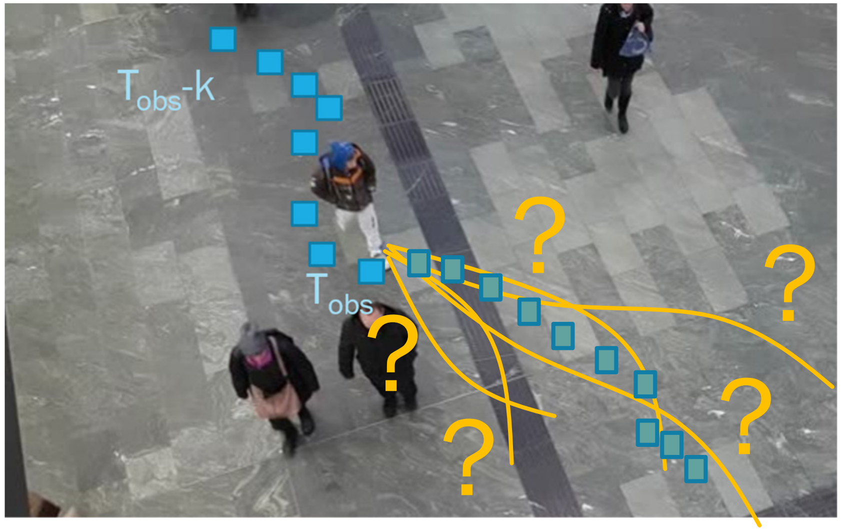


Thuremella, D., & Kunze, L. Evaluating Long-Tailed Learning Techniques on Pedestrian Trajectory Prediction.

QUANTITATIVE ANALYSIS – MY POINT OF VIEW

MULTI PREDICTION METRICS

Multi prediction

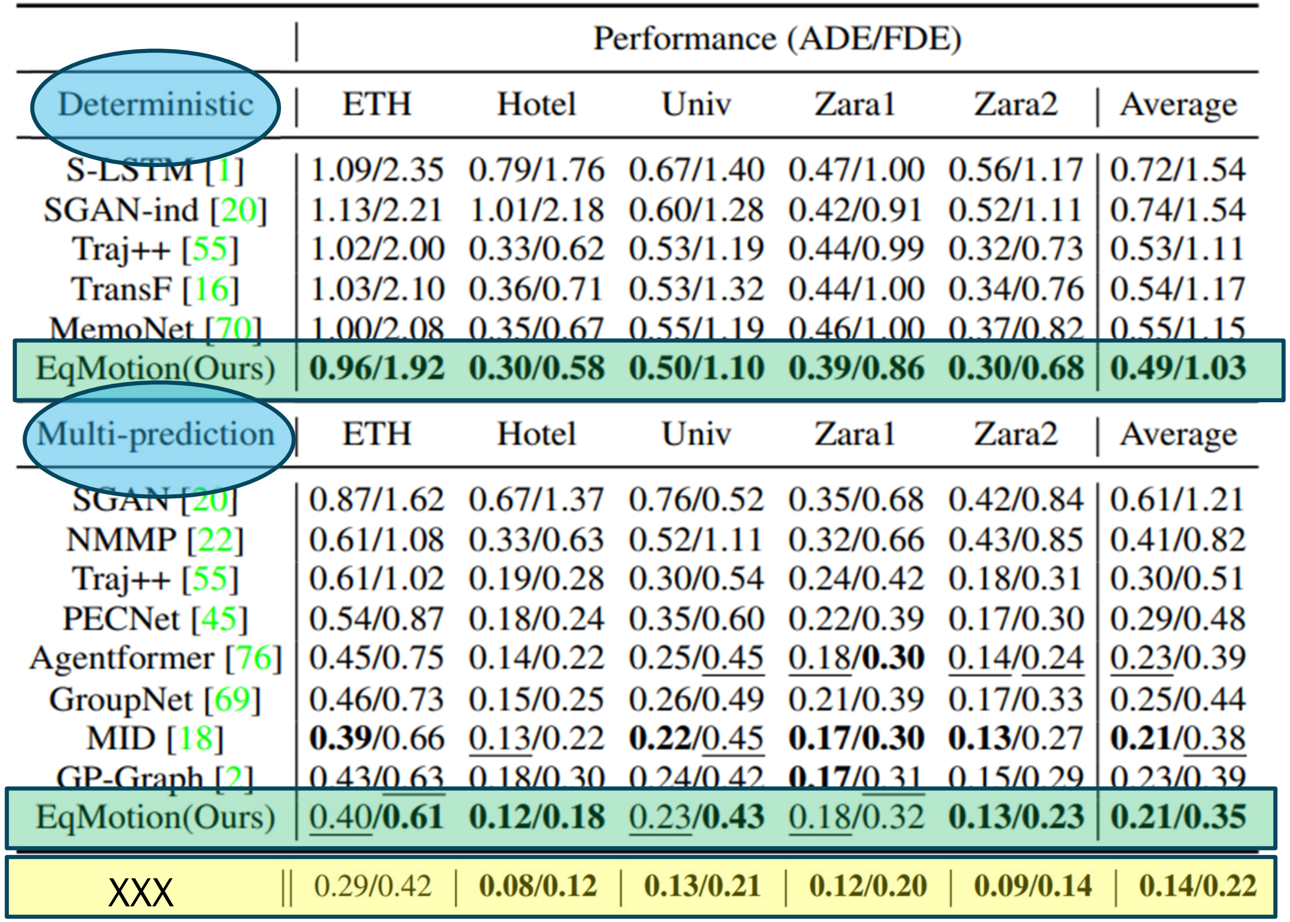


$$ADE = \frac{\sum_{i=1}^{n} \sum_{t=T_{obs}+1}^{T_{pred}} \left[(\hat{x}_i^t - x_i^t)^2 + (\hat{y}_i^t - y_i^t)^2 \right]}{n(T_{pred} - (T_{obs}+1))}$$

Best-of-K ADE metrics

QUANTITATIVE ANALYSIS - A CRITICISM

- Best-of-k is giving better results
  - Observation period is 3.2s and the forecasting horizon is 4.8s
  - ADE/FDE errors, in meters; Less is better ↓
- Using only best-of-k is not fair!!!



	Performance (ADE/FDE)					
Deterministic	ETH	Hotel	Univ	Zara1	Zara2	Average
S-LSTM [1]	1.09/2.35	0.79/1.76	0.67/1.40	0.47/1.00	0.56/1.17	0.72/1.54
SGAN-ind [20]	1.13/2.21	1.01/2.18	0.60/1.28	0.42/0.91	0.52/1.11	0.74/1.54
Traj++ [55]	1.02/2.00	0.33/0.62	0.53/1.19	0.44/0.99	0.32/0.73	0.53/1.11
TransF [16]	1.03/2.10	0.36/0.71	0.53/1.32	0.44/1.00	0.34/0.76	0.54/1.17
MemoNet [70]	1.00/2.08	0.35/0.67	0.55/1.19	0.46/1.00	0.37/0.82	0.55/1.15
EqMotion(Ours)	**0.96/1.92**	**0.30/0.58**	**0.50/1.10**	**0.39/0.86**	**0.30/0.68**	**0.49/1.03**
Multi-prediction	ETH	Hotel	Univ	Zara1	Zara2	Average
SGAN [20]	0.87/1.62	0.67/1.37	0.76/0.52	0.35/0.68	0.42/0.84	0.61/1.21
NMMP [22]	0.61/1.08	0.33/0.63	0.52/1.11	0.32/0.66	0.43/0.85	0.41/0.82
Traj++ [55]	0.61/1.02	0.19/0.28	0.30/0.54	0.24/0.42	0.18/0.31	0.30/0.51
PECNet [45]	0.54/0.87	0.18/0.24	0.35/0.60	0.22/0.39	0.17/0.30	0.29/0.48
Agentformer [76]	0.45/0.75	0.14/0.22	0.25/0.45	0.18/**0.30**	0.14/0.24	0.23/0.39
GroupNet [69]	0.46/0.73	0.15/0.25	0.26/0.49	0.21/0.39	0.17/0.33	0.25/0.44
MID [18]	**0.39**/0.66	0.13/0.22	**0.22**/0.45	**0.17/0.30**	**0.13**/0.27	**0.21**/0.38
GP-Graph [2]	0.43/0.63	0.18/0.30	0.24/0.42	**0.17**/0.31	0.15/0.29	0.23/0.39
EqMotion(Ours)	0.40/**0.61**	**0.12/0.18**	0.23/**0.43**	0.18/0.32	**0.13/0.23**	**0.21/0.35**
XXX	0.29/0.42	**0.08/0.12**	**0.13/0.21**	**0.12/0.20**	**0.09/0.14**	**0.14/0.22**

Xu, C., Tan, R. T., Tan, Y., Chen, S., Wang, Y. G., Wang, X., & Wang, Y. (2023). EqMotion: Equivariant Multi-agent Motion Prediction with Invariant Interaction Reasoning. In *CVPR*

THE MARKET SIDE – MY POINT OF VIEW

- *Automotive:* very hot but *very specific*
- There is a stream of long-term trajectory forecasting (even between different buildings)

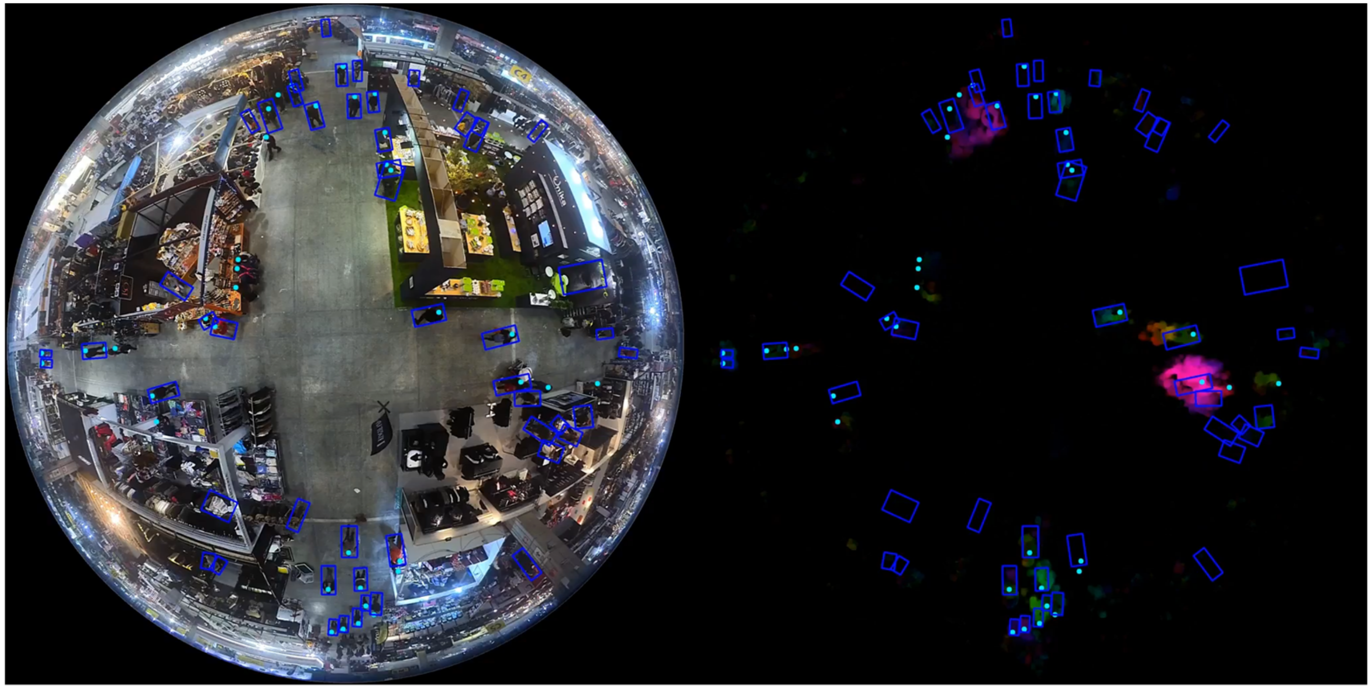
veronafiere
Trade shows & events since 1898

vinitaly

Fieracavalli
a Verona dal 1898

MARMO+MAC

MBE
MOTOR BIKE EXPO



TAKE-HOME MESSAGES – TRAJECTORY FORECASTING

- Trajectory forecasting is not as the standard time series forecasting, since many additional factors have to be taken into account.
- Transformers as inference engine for trajectory forecasting seems promising, so LLM could give the expected boost, but…
- People are not points, and we need to model the body for getting better results.
- We need to take care about the evaluation metrics. Best-of-K.. is it really realistic?

OUTLINE

- Introduction to forecasting
- *Forecast the forecasting*
  - People trajectory forecasting
  - **Human pose forecasting**

HUMAN 3D POSE FORECASTING

GOAL *Predict the future locations of 3D joints of a skeletal structure.*

T observed frames → *N_F predicted frames*

Ground Truth Predicted

$$\mathcal{X}_{in} = [X_1, X_2 \ldots, X_T]$$

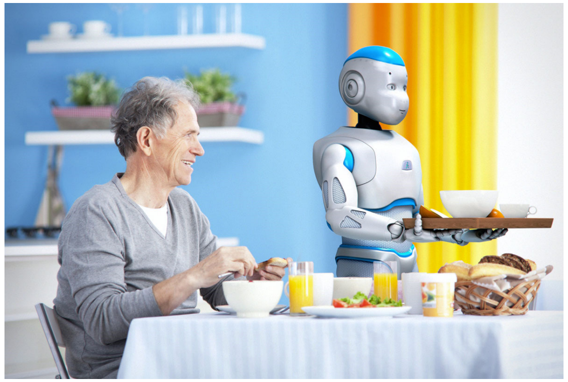
$$X_i \in \mathbb{R}^{3\times N_j}$$

$$MPJPE = \frac{1}{N_F}\frac{1}{N_J}\sum_{f,j}\|p_{f,j} - \hat{p}_{f,j}\|_2$$

APPLICATIONS OF 3D POSE FORECASTING

- Human robot interaction and cooperation, domotics/assistive

  H. S. Koppula and A. Saxena. Anticipating human activities for reactive robotic response. In The IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS), pages 2071–2071, 2013.



- Entertainment

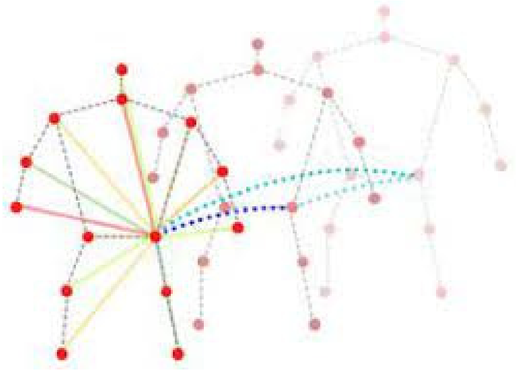
  Zhu, Y., Doermann, D., Zhang, Y., Liu, Q., & Girgensohn, A. (2021, January). What and how? jointly forecasting human action and pose. In *2020 25th International Conference on Pattern Recognition (ICPR)* (pp. 771-778). IEEE.

METHODS FOR 3D POSE FORECASTING

2 MAINSTREAMS

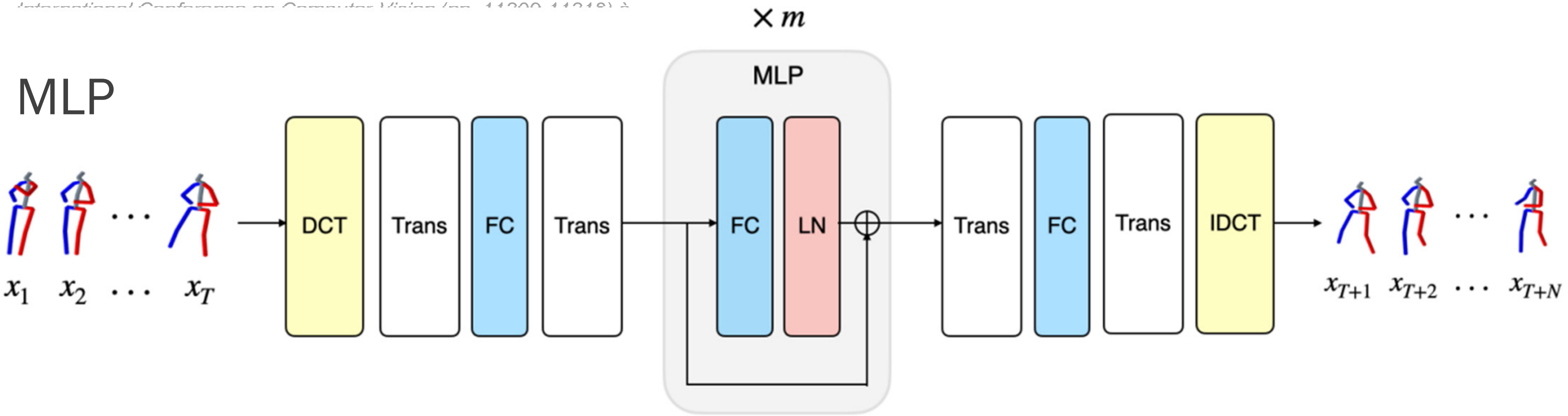
- GCN

$$\mathcal{H}^{(l+1)} = \sigma(A^{s-(l)} A^{t-(l)} \mathcal{H}^{(l)} W^{(l)}) \quad \mathcal{H}^{(1)} = \mathcal{X}_{in}$$



Sofianos, T., Sampieri, A., Franco, L., & Galasso, F. (2021). Space-time-separable graph convolutional network for pose forecasting. In *Proceedings of the IEEE/CVF International Conference on Computer Vision* (pp. 11209-11218).

- MLP



Guo, W., Du, Y., Shen, X., Lepetit, V., Alameda-Pineda, X., & Moreno-Noguer, F. (2023). Back to mlp: A simple baseline for human motion prediction. In *Proceedings of the IEEE/CVF Winter Conference on Applications of Computer Vision* (pp. 4809-4819).

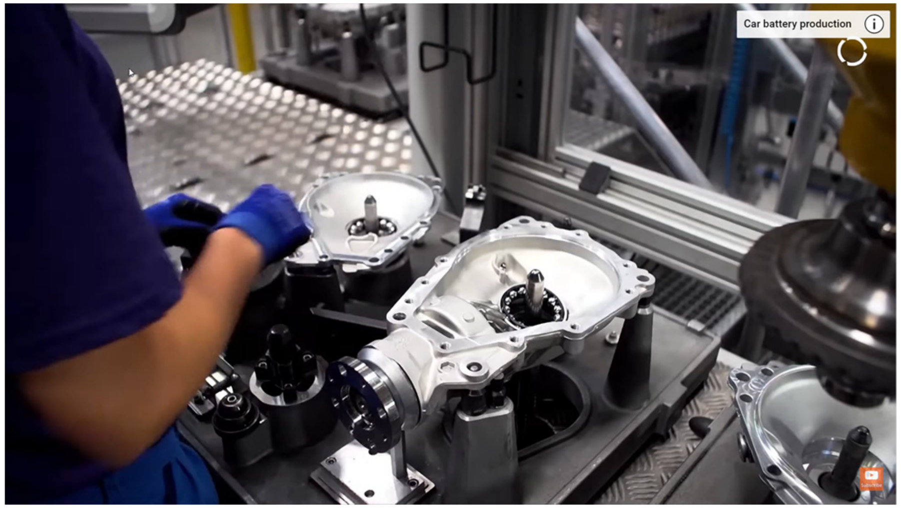
POSE FORECASTING – OUR EFFORT

SOMETHING IS MISSING

- I'm interested in *human robot interaction* (for Industry 5.0);
- Issues: interaction at low speed, collisions are harmful;

Universal Robots UR5e

Boston Dynamics' Spot

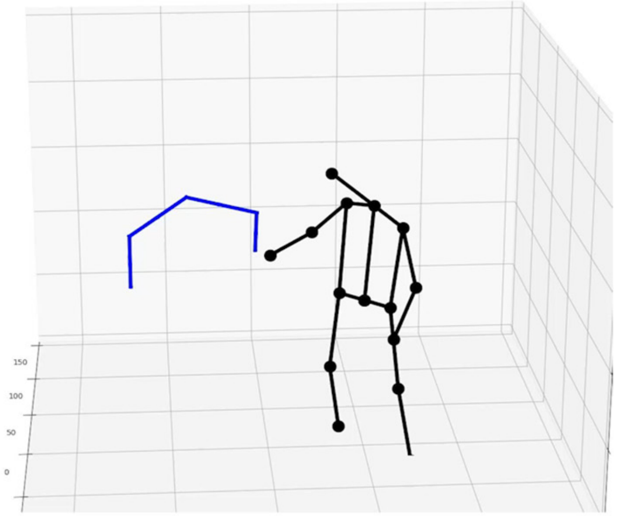


- Help avoiding collisions with 3D pose forecasting;
- What is missing? <u>To put the robot into play</u>

A MAN + A ROBOTIC ARM: CHICO DATASET

SEEING HUMAN + ROBOT FROM AN EXTERNAL POV

- Acquisition details:
  - POV: 3 wall-mounted HD-RGB cameras
  - A UR5 cobot
- Statistics:
  - Markerless
  - 7 Industrial actions
  - 226 annotated collisions
  - 20 actors
- Tasks:
  - Pose Forecasting
  - Collision Detection



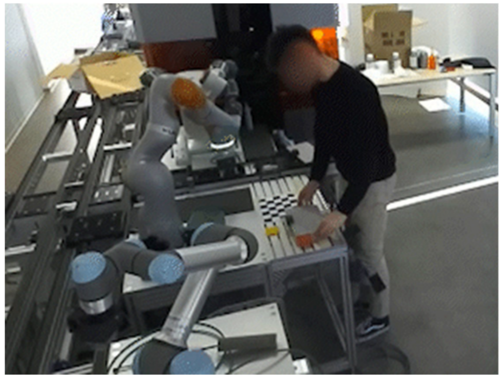
dataset here!

CHICO DATASET

SAMPLE VIDEOS

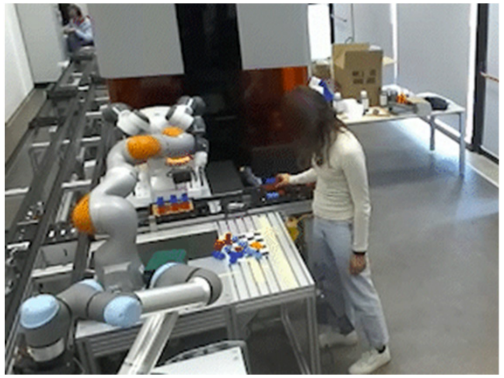
CHICO DATASET

SAMPLE VIDEOS

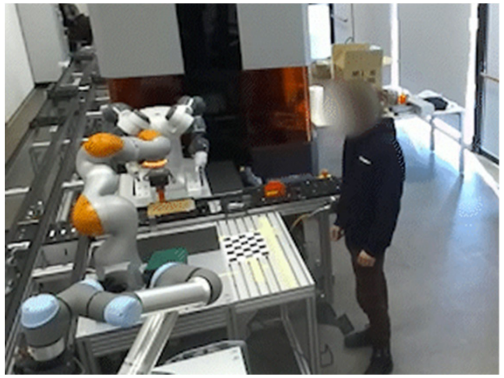


Span Heavy

Heavy P&P

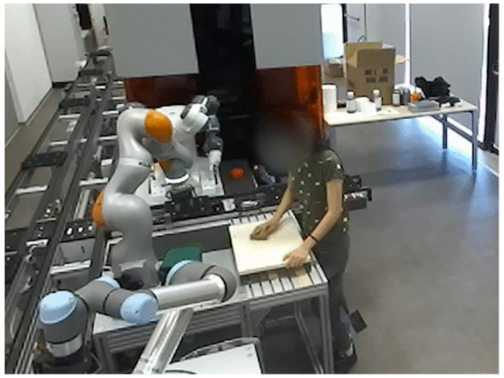


Random P&P

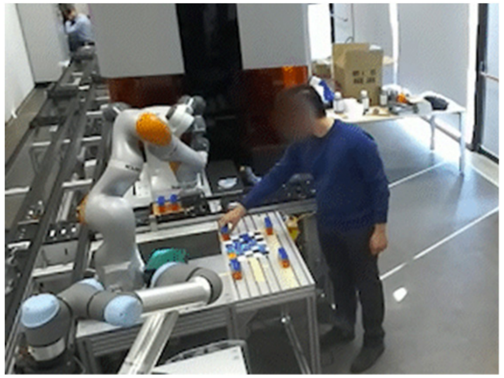


Light P&P

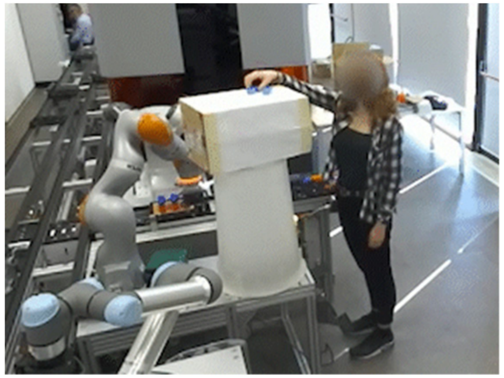
Hammer



Polishng



Precise P&P

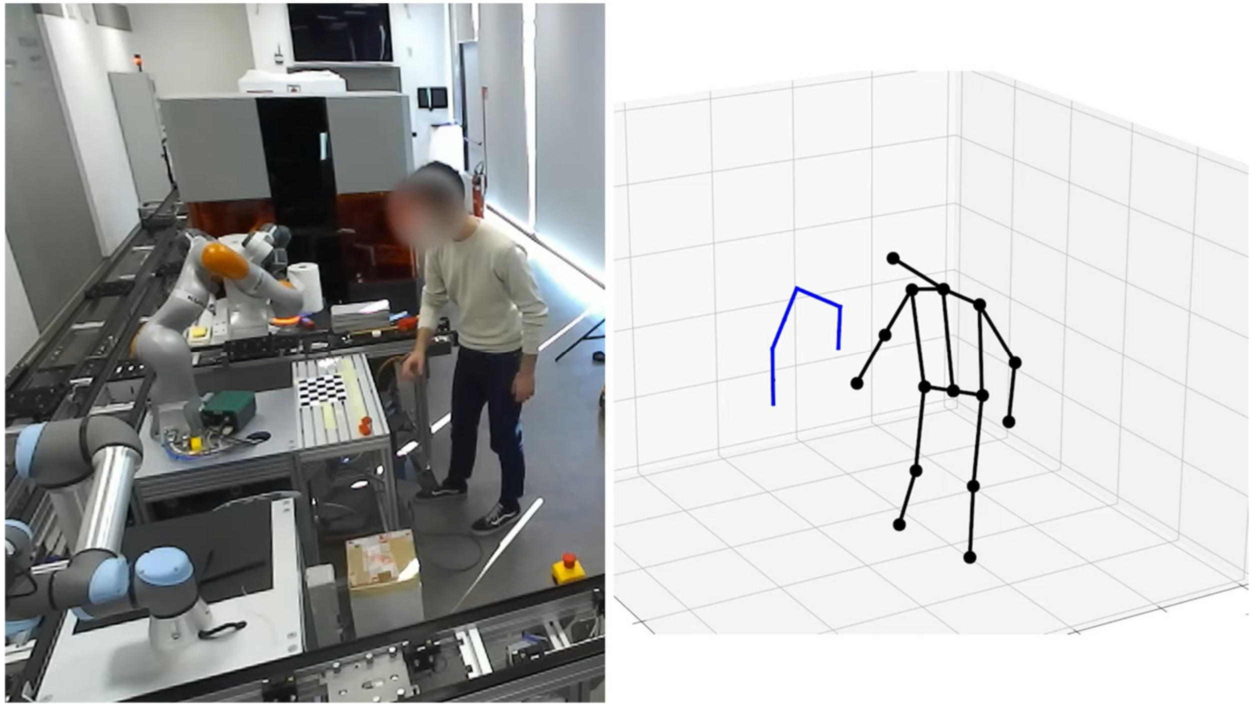


CHICO DATASET

COLLISIONS

CHICO DATASET

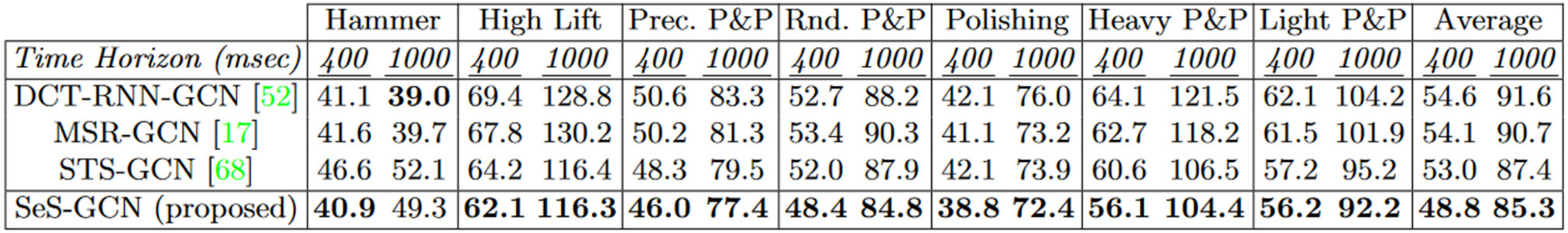
SAMPLES OF 3D ANNOTATIONS



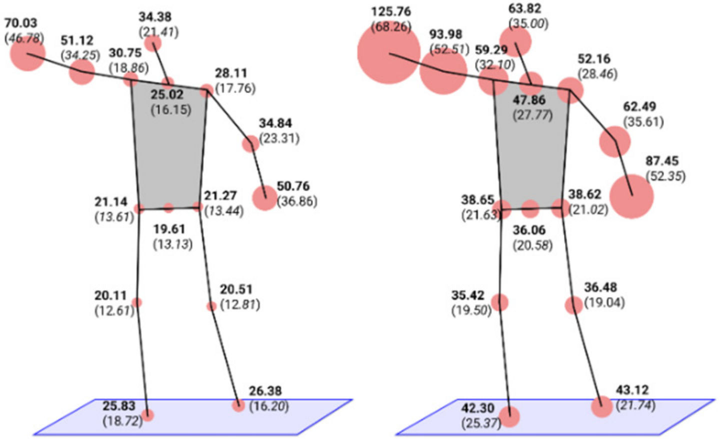
Span Light

RESULTS

$$L_{\mathrm{MPJPE}} = \frac{1}{V}\sum_{v=1}^{V} \|\hat{\boldsymbol{x}}_{vt} - \boldsymbol{x}_{vt}\|_2,$$ in millimeters

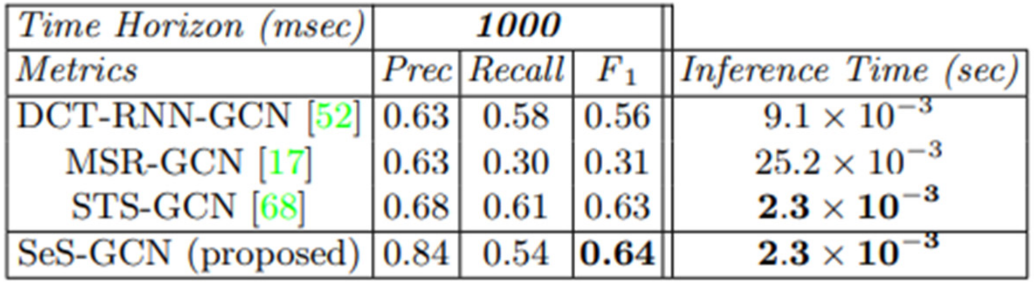


	Hammer		High Lift		Prec. P&P		Rnd. P&P		Polishing		Heavy P&P		Light P&P		Average	
Time Horizon (msec)	*400*	*1000*	*400*	*1000*	*400*	*1000*	*400*	*1000*	*400*	*1000*	*400*	*1000*	*400*	*1000*	*400*	*1000*
DCT-RNN-GCN [52]	41.1	**39.0**	69.4	128.8	50.6	83.3	52.7	88.2	42.1	76.0	64.1	121.5	62.1	104.2	54.6	91.6
MSR-GCN [17]	41.6	39.7	67.8	130.2	50.2	81.3	53.4	90.3	41.1	73.2	62.7	118.2	61.5	101.9	54.1	90.7
STS-GCN [68]	46.6	52.1	64.2	116.4	48.3	79.5	52.0	87.9	42.1	73.9	60.6	106.5	57.2	95.2	53.0	87.4
SeS-GCN (proposed)	**40.9**	49.3	**62.1**	**116.3**	**46.0**	**77.4**	**48.4**	**84.8**	**38.8**	**72.4**	**56.1**	**104.4**	**56.2**	**92.2**	**48.8**	**85.3**



Short term

Long term

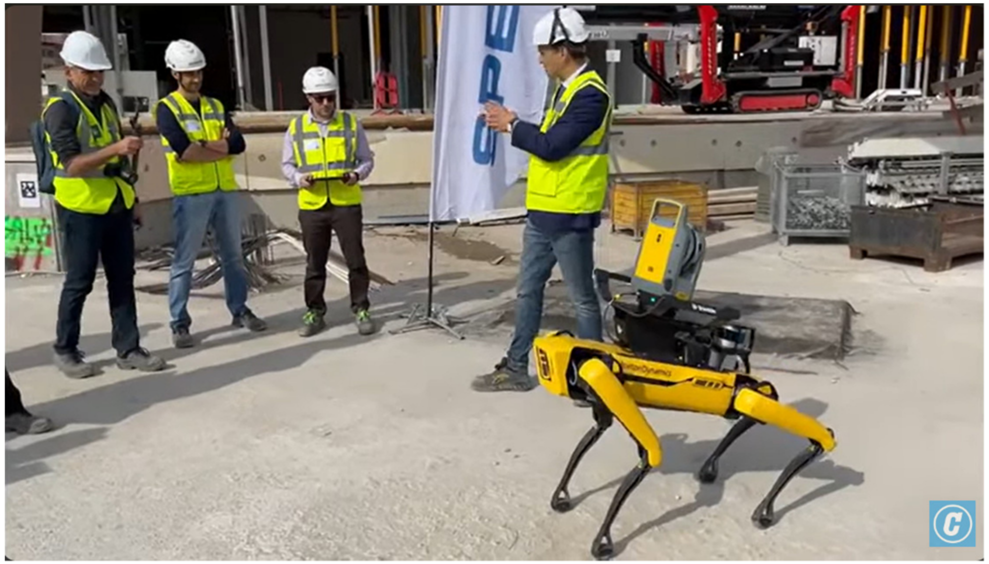


Time Horizon (msec)	*1000*			
Metrics	*Prec*	*Recall*	*F_1*	*Inference Time (sec)*
DCT-RNN-GCN [52]	0.63	0.58	0.56	9.1×10^{-3}
MSR-GCN [17]	0.63	0.30	0.31	25.2×10^{-3}
STS-GCN [68]	0.68	0.61	0.63	$\mathbf{2.3 \times 10^{-3}}$
SeS-GCN (proposed)	0.84	0.54	**0.64**	$\mathbf{2.3 \times 10^{-3}}$

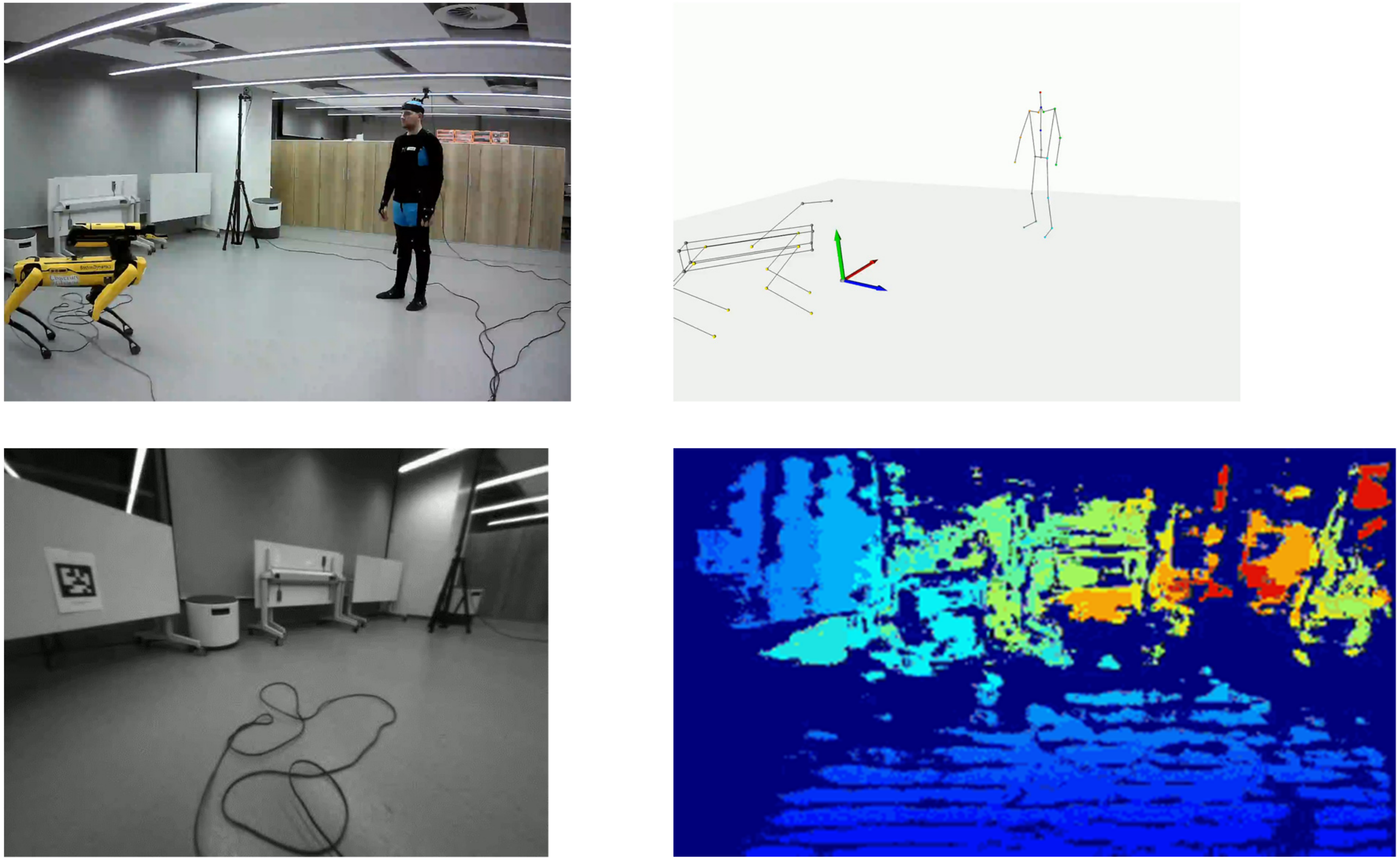
HARPER: HUMANS FROM AN ARTICULATED ROBOT PERSPECTIVE

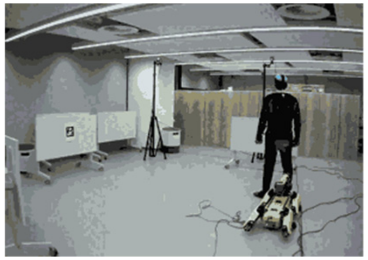
PUT THE 3D POSE FORECASTING MODEL ON A ROBOT

- *A double point of view: External (like CHICO) + Internal (on the robot!);*
  - Collaboration with University of Glasgow (*being submitted at IROS 24*);
  - Exploit the *Spot* (by Boston Dynamics) quadruped robot.



- *Multiple point of views: System + Robot POV*
- *MOCAP data for the human AND the robot*
- *Synchronized data*

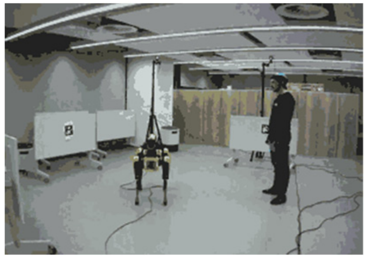




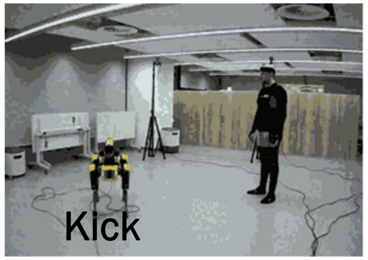
Collision in movement

Collision Avoidance

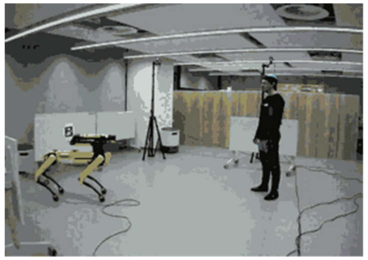
Touch



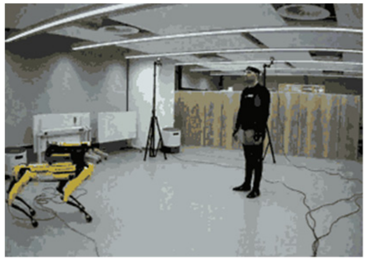
Push



Kick



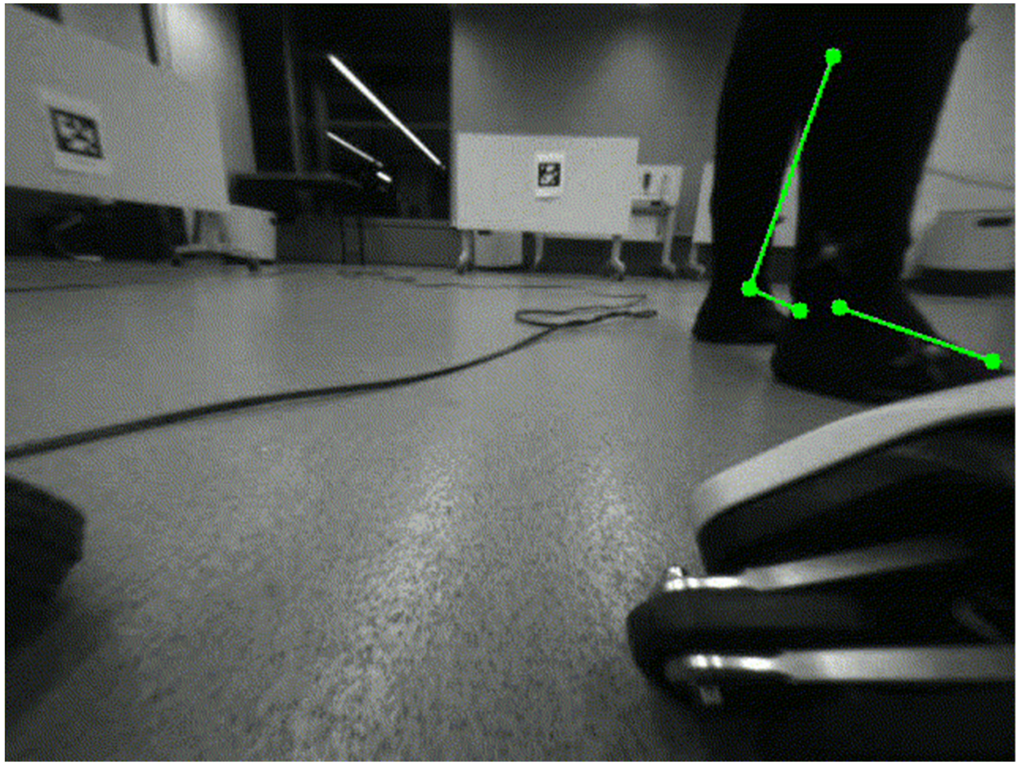
Proximity



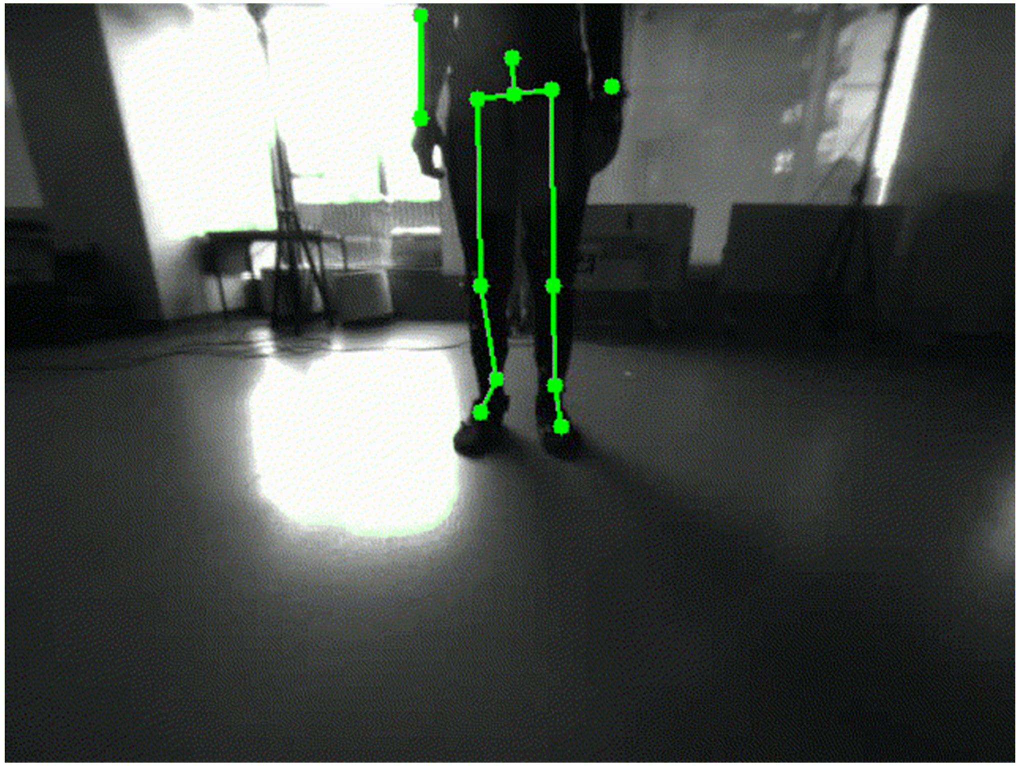
Sudden stop

Accidental Collision

POINT OF VIEW OF THE ROBOT



Human-robot Collision



Kick

RESULTS

- Finalizing them…

TAKE-HOME MESSAGES – HUMAN POSE FORECASTING

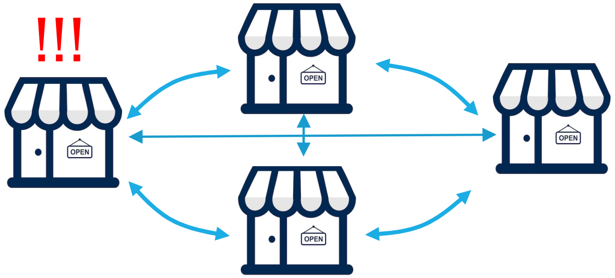
- Pose forecasting is crucial for specific applications, and human robot interaction seems to be very exciting
- Graph convolutional network appear to be the most suited tools to deal with. MLP seems to be a valid alternative, with way less parameters
- Emphasis now should be on how to put this framework in a real-time scenario, to show the real capabilities / limitations.

OUTLINE

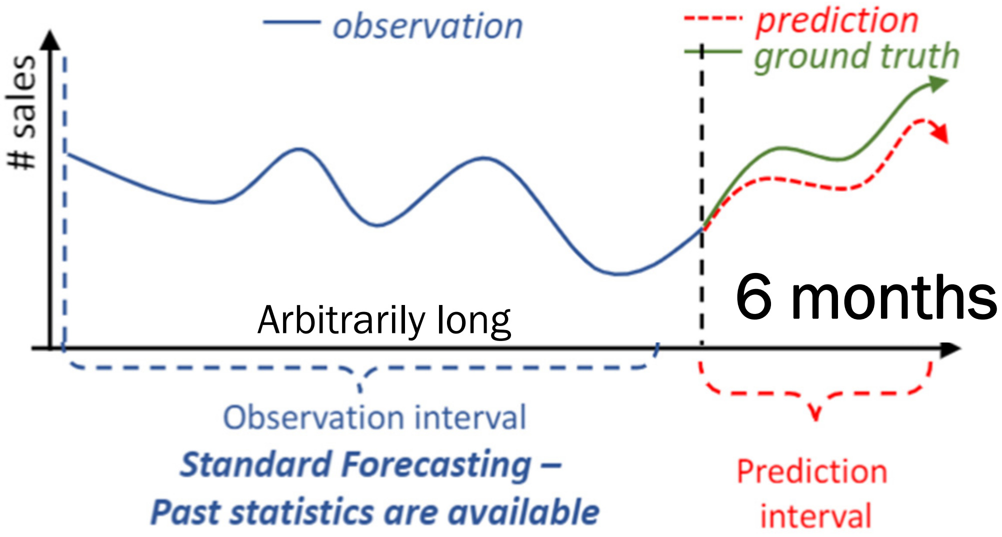
- Introduction to forecasting
- *Forecast the forecasting*
  - People trajectory forecasting
  - Human pose forecasting
  - **Forecasting for fashion**

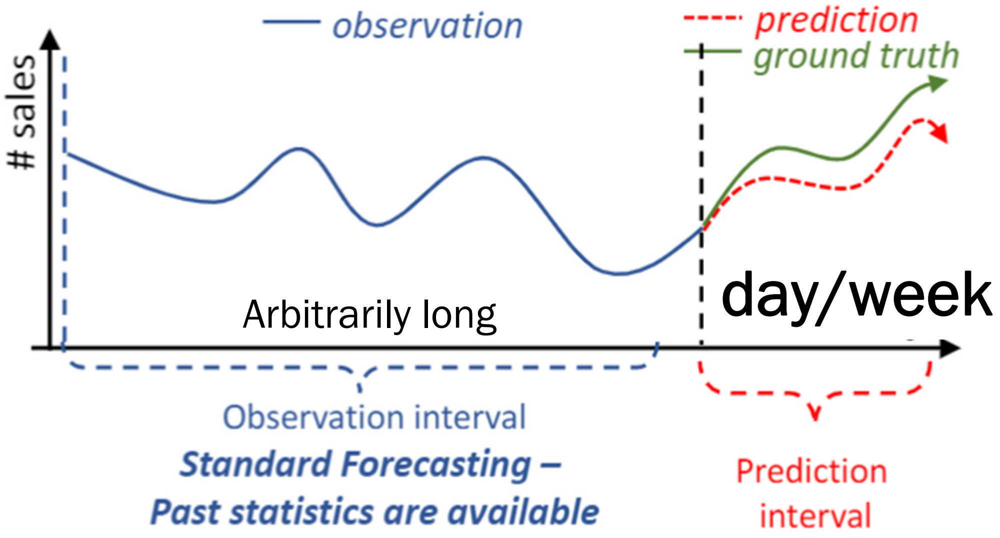
FORECASTING IN FASHION: GOALS

- Fashion is a *complex concept* and so are the business processes related to it.
- Common issues that are a direct consequence of the management of these processes:
  - **Overproduction:** warehouses are full after the season;
  - **Missing items:** shop are empty during the season;
  - (many others)!
- Forecasting is an optimal analytic tool to tackle the two problems mentioned above.



FORECASTING FOR FASHION: 3 CHALLENGES





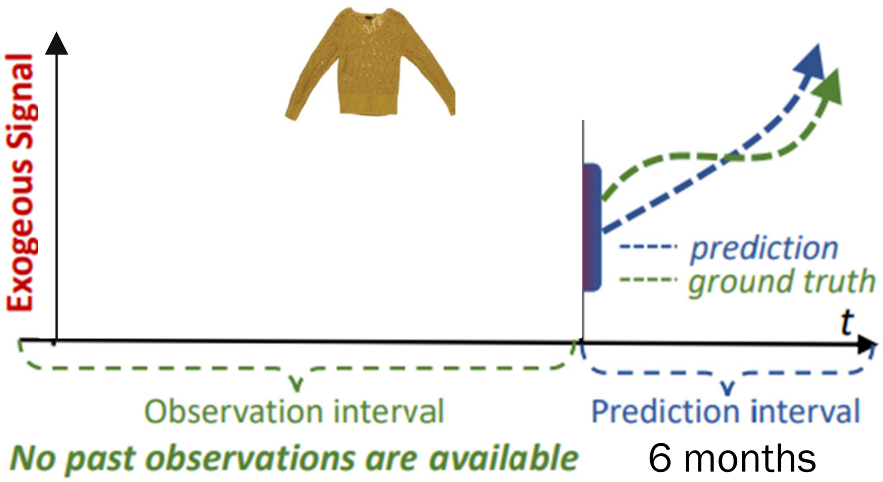
Long-term forecasting

- *first-order with suppliers*

Short-term forecasting

- *replenishment*

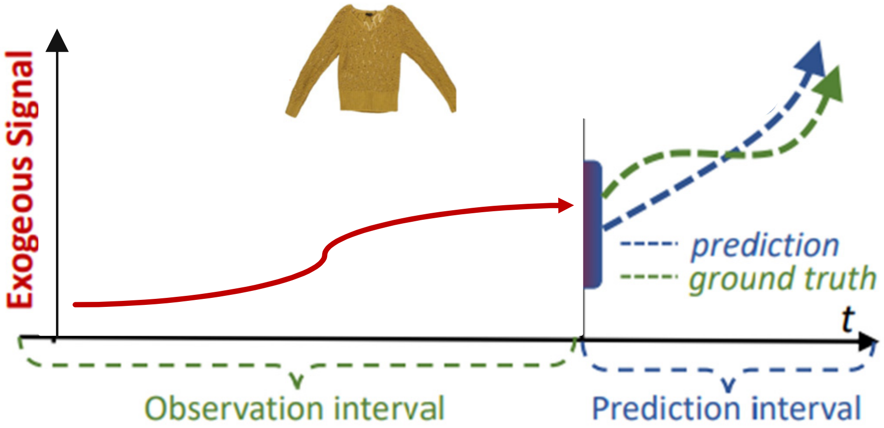
FORECASTING FOR FASHION: 3 CHALLENGES



- New product performance forecasting (NPPF):
  - This is crucial buying the correct number of items, to have an effective promotion etc.

NEW FASHION PRODUCT PERFORMANCE FORECASTING (NPPF)

- Adoption of exogenous signals;
- NFPPF models work under the rationale that new products will perform comparably to aesthetically similar, older products.



No past observations are available

6 months

Ekambaram, Vijay, et al. "Attention based multi-modal new product sales time-series forecasting." Proceedin

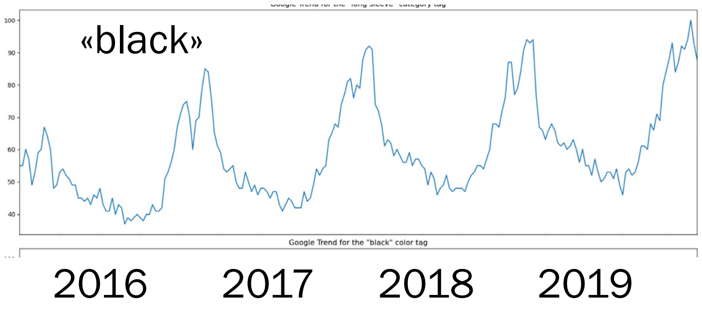
VISUELLE 2.0: MULTIMODAL + SALES

- Visuelle 2.0contains real data for 5355 clothing products, sold from 2016-2020, of a retail, Italian fast-fashion company: Nuna Lie.
- Each product-shop pair in our dataset contains:
  1. An HD image;
  2. Sales, inventory and discount time series
  3. Textual tags related to category, color and fabric;
  4. Exogenous time series related to weather and online popularity (represented by Google Trends).
- We also provide purchase data for 667086 fidelity customers.

dataset

NUNA LIE

Category: Long sleeve
Color: Black
Fabric: Acrylic



POP SIGNAL: A RAG *ANTE-LITTERAM*

ECCV2022

Joppi, C., Skenderi, G., & Cristani, M. (2022, October). POP: Mining POtential Performance of new fashion products via webly cross-modal query expansion. In *ECCV*

START

1° Extract visual tags:

«Yellow pullover»

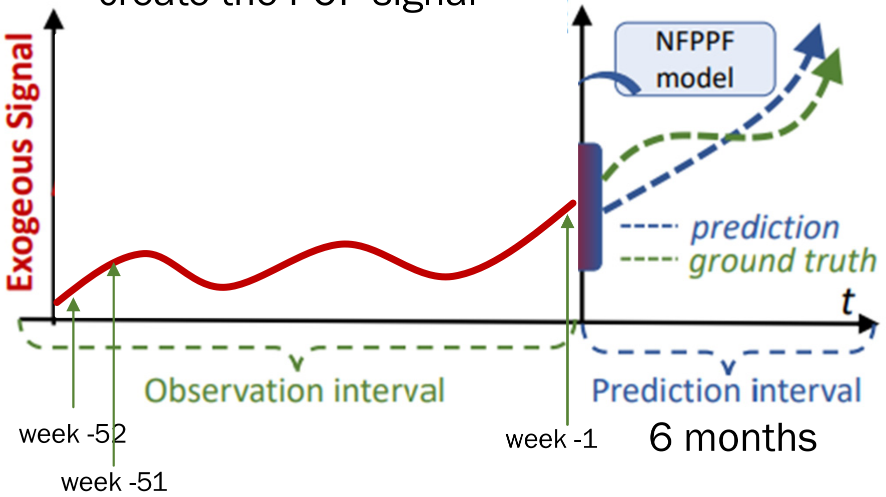
2° Search images over internet in the past

«*Nice* Yellow pullover»

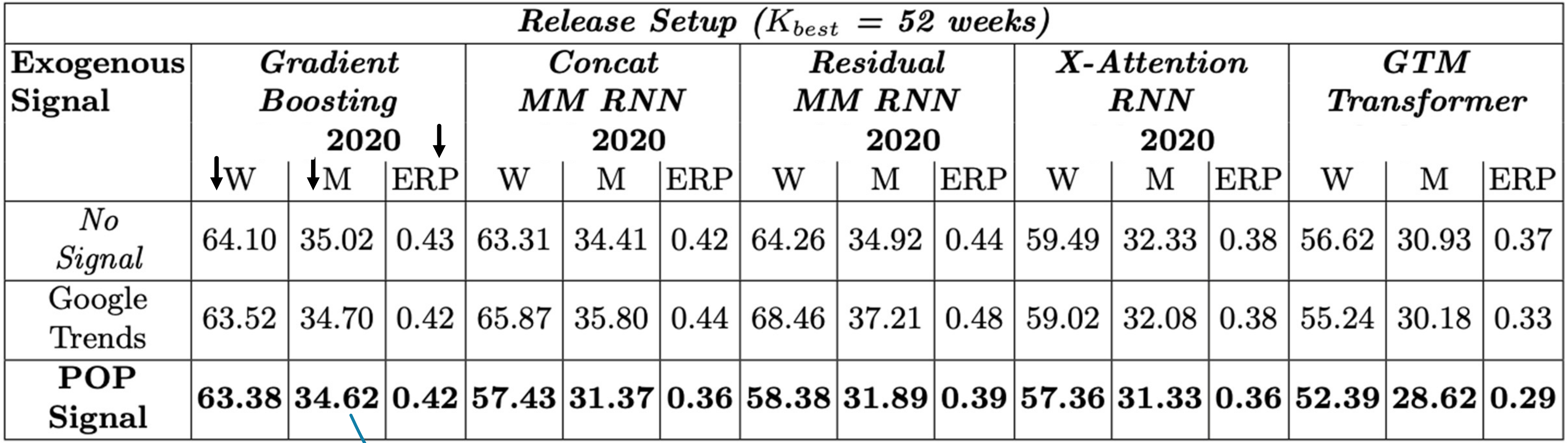
SIMILARITY

3°

4° create the POP signal



RESULTS: NEW FASHION PRODUCT SALES CURVE PREDICTION



Release Setup (K_{best} = 52 weeks)															
Exogenous Signal	***Gradient Boosting*** **2020**			***Concat MM RNN*** **2020**			***Residual MM RNN*** **2020**			***X-Attention RNN*** **2020**			***GTM Transformer***		
	W	M	ERP	W	M	ERP	W	M	ERP	W	M	ERP	W	M	ERP
No Signal	64.10	35.02	0.43	63.31	34.41	0.42	64.26	34.92	0.44	59.49	32.33	0.38	56.62	30.93	0.37
Google Trends	63.52	34.70	0.42	65.87	35.80	0.44	68.46	37.21	0.48	59.02	32.08	0.38	55.24	30.18	0.33
POP Signal	**63.38**	**34.62**	**0.42**	**57.43**	**31.37**	**0.36**	**58.38**	**31.89**	**0.39**	**57.36**	**31.33**	**0.36**	**52.39**	**28.62**	**0.29**

$\underline{W}APE$

$$\underline{M}APE = 100 * \frac{\sum_{t=1}^{N} \left|\frac{E_t}{Y_t}\right|}{N}$$

average abs. #
per-item per-shop
that I mispredicted

Ekambaram, V., Manglik, K., Mukherjee, S., Sajja, S.S.K., Dwivedi, S., Raykar, V.: Attention based Multi-Modal New Product Sales Time-series Forecasting. In SIGKDD 2020

Skenderi, G., Joppi, C., Denitto, M., & Cristani, M. (2024). Well googled is half done: Multimodal forecasting of new fashion product sales with image-based google trends. *International Journal of Forecasting*.

TAKE-HOME MESSAGES – NEW PRODUCT PERFORMANCE FORECASTING

- NPPF problem is definitely the hot topic in forecasting for fashion
- Inevitably, is multimodal: an image is worth 1000 words
- The intuition is to create a past signal predictive for the item image
- Extend to LLM and proper RAG approaches

TAKE-HOME MESSAGES – FORECASTING WITH MACHINE LEARNING

- Forecasting spans over different fields and applications
- Machine learning models are promising, but take care of guarantee/interpretability
- More than models, is about (exogenous) *data:* forecasting is *data centric*
- Large language models need to be incorporated

INTELLIGO labs

THANKS FOR YOUR ATTENTION!

Francesco Setti

Assistant Professor

Christian Joppi

Computer Vision & Deep Learning Specialist

Andrea Avogaro

Federico Cunico

HUMATICS
SYS-DAT GROUP

NUNALIE

Andrea Toaiari

Geri Skenderi