Context-aware collaborative pushing of heavy objects using skeleton-based intention prediction

Gokhan Solak[1], Gustavo J. G. Lahr[2], Idil Ozdamar[1], Arash Ajoudani[1]

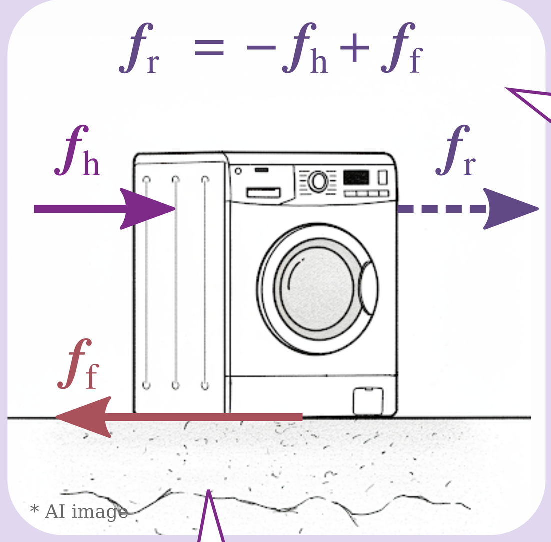
1: Human-Robot Interfaces and Interaction Lab, Italian Institute of Technology (IIT), Genoa
2: Faculdade Israelita de Ciências da Saúde, Hospital Israelita Albert Einstein, São Paulo, Brazil

TORNADO (EU, GA 101189557)

ATLANTA 2025 ICRA

Motivation

Collaboratively pushing heavy objects on frictional surfaces



The sensor measures the compression force between the object and the robot.

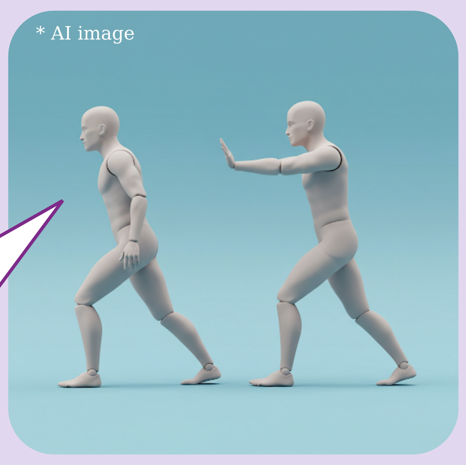
Human force f_h cannot be decoupled!

Thus, admittance control is **not** feasible.

The forces cannot be directly transmitted to the partner since they are subjected to the friction force

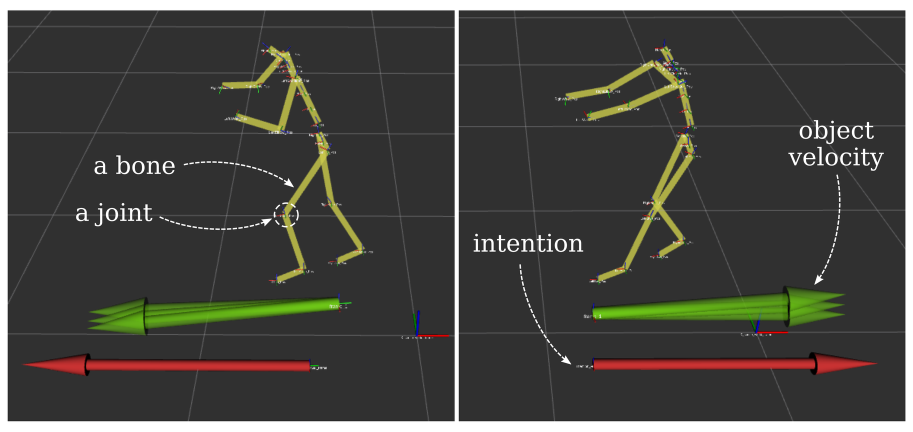
The human has to apply enough force to break static friction

Intention prediction through human pose data overcomes these issues



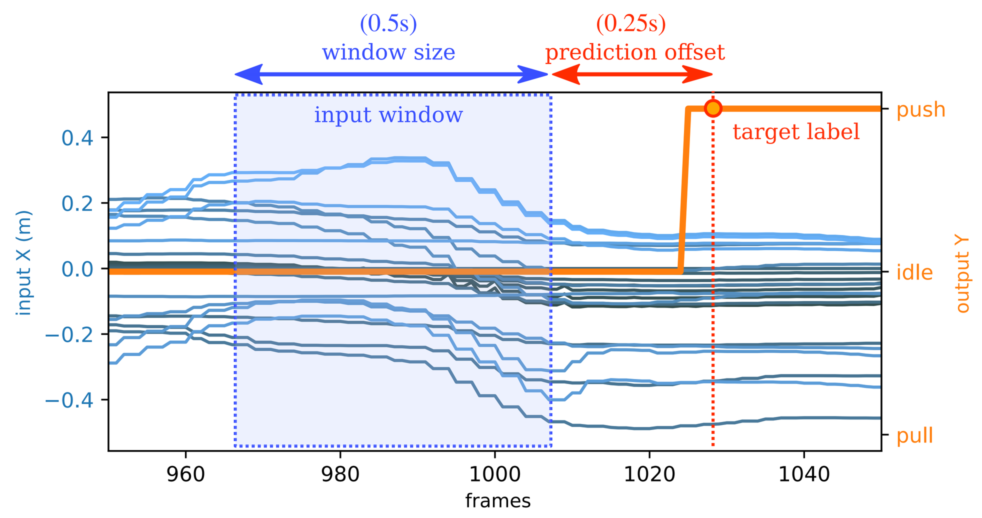
Methodology

Skeleton-based intention prediction using directed graph neural networks (DGNN) for assistive force control

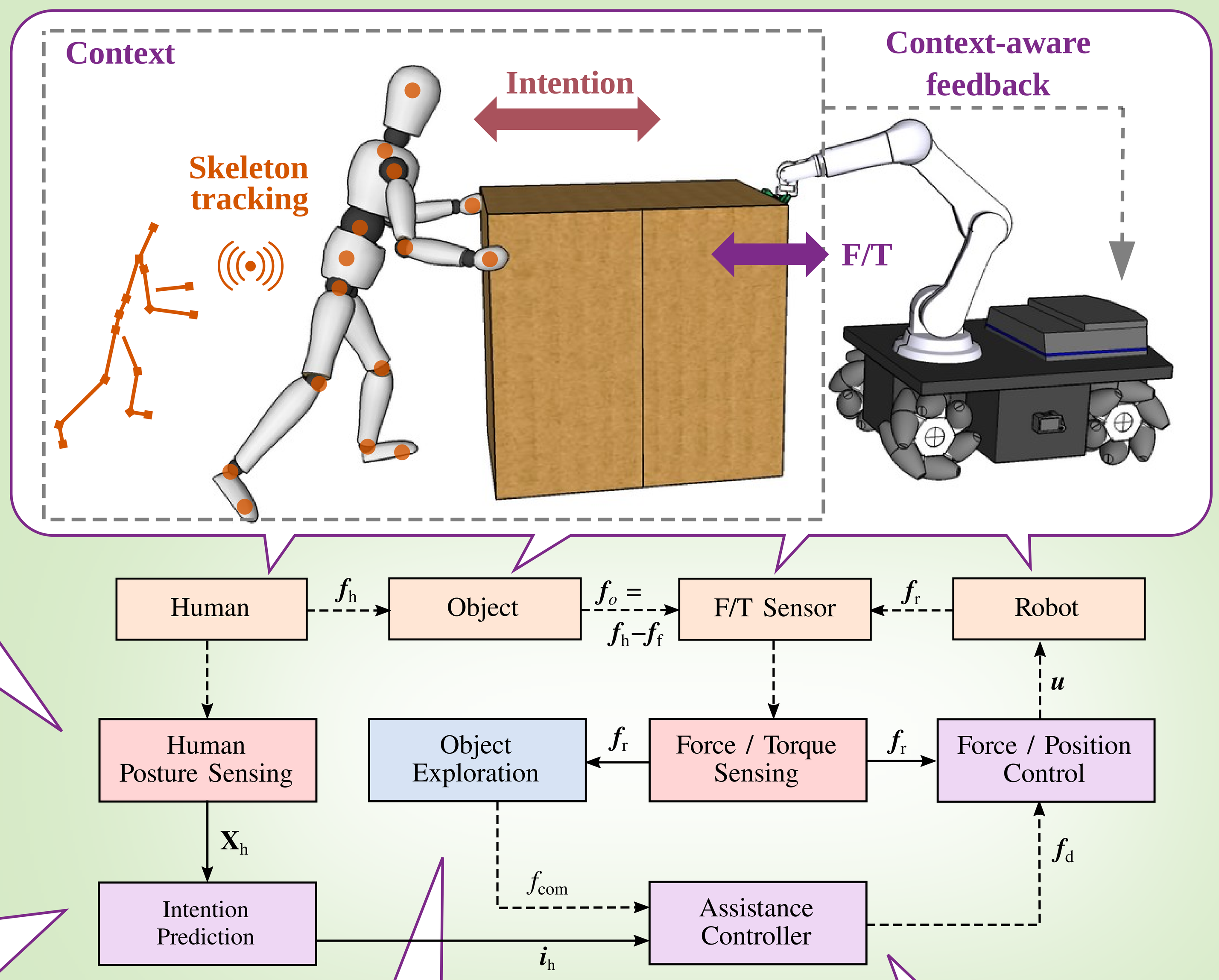


- The **skeleton data** is represented as edges and nodes (**bones and joints**) of a graph.
- We collected skeleton tracking data using Xsens, **IMU-based body suit**.
- Intention labels are created from the object velocity data.

- Given: time window graph data $\mathbf{X}_h$
- Predict: Intention class $\boldsymbol{i}_h \in \{-1, 0, +1\}$
- Using **DGNN** algorithm [1]



- DGNN uses graph convolution and temporal convolution on the connected joints and bones
- We transformed all joint data to be on the pelvis frame to remove dependency on the starting position



- The exploration is done once per object
- The object is pushed without the human
- The absolute force that breaks the static friction is recorded: f_{com}

- The robot acts as a **follower**
- Applying force in the same direction as the predicted human intention:

$$f_d = \boldsymbol{i}_h \, f_{com}$$

- We linearly change the desired force to avoid jumps
- The intention output is filtered (Moving average)
- Force control:

$$\boldsymbol{u} = K_f\,(f_d - f_r)$$

Experiments

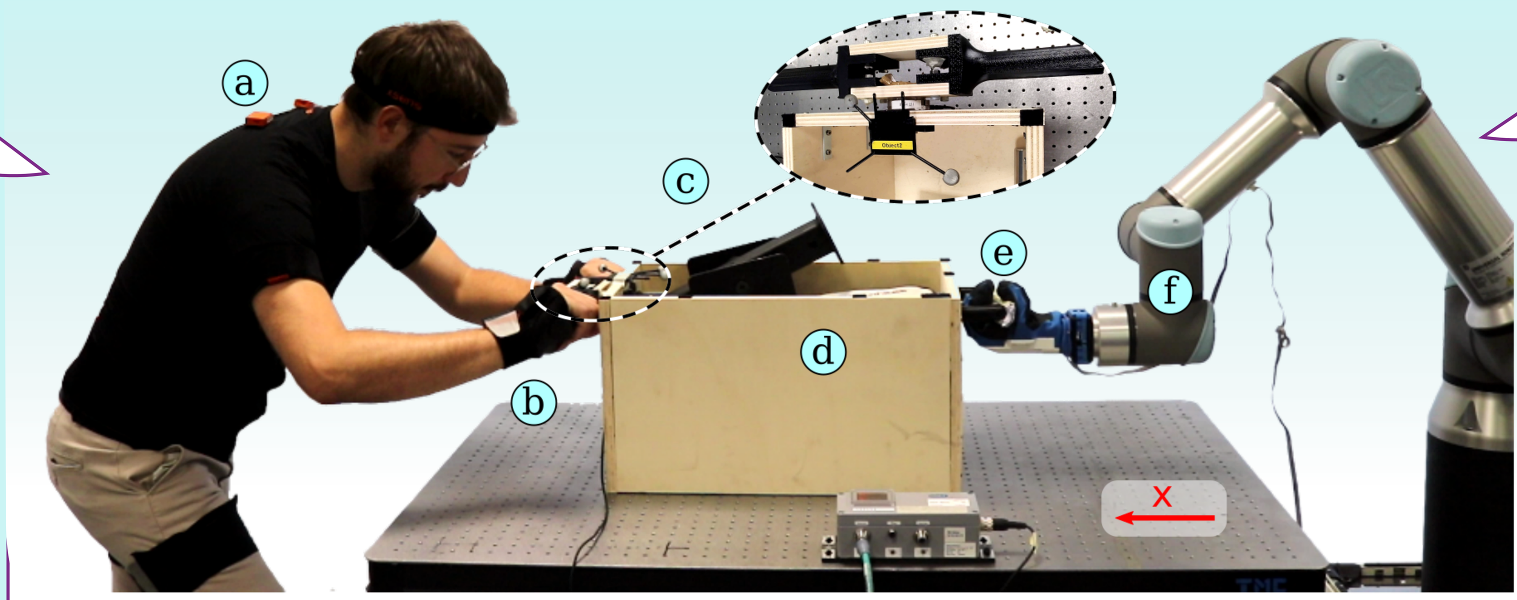
TABLE I: Description of the experiments

Exp.	Participant	m_{box} [kg]	Task time [s]	f_{comp} [N]
1	1	27.7	6	65
2	1	27.7	10	65
3	1	36.0	6	80
4	1	36.0	10	80
5	2	36.0	6	80
6	2	36.0	10	80

Experiment cases

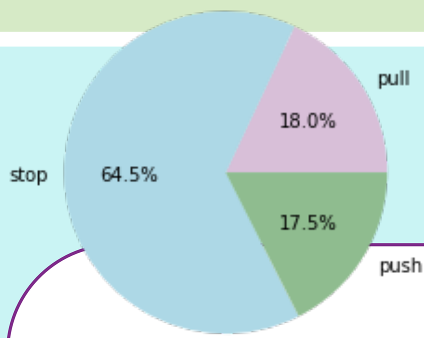
- Changing **box weights**
- Changing **participants**
- Changing task durations (**speed**)
- Constant distance (30 cm)
- Single DGNN model for all
- f_{com} is reidentified for different boxes

- **5 repetitions** of pull and push
- Two approaches:
  - ***assisted*** (proposed)
  - ***dry*** (no assistance)



Experiment and data collection setup:

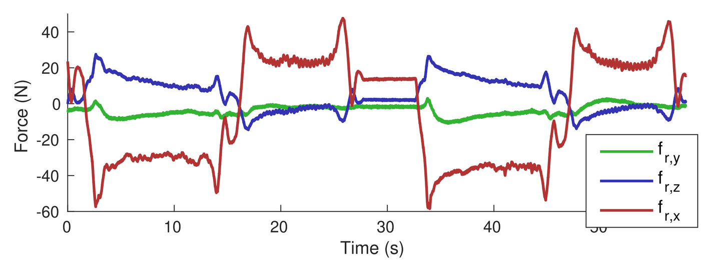
a) Xsens IMU markers
b) Human-side F/T sensor (evaluation-only)
c) Object tracking marker (data collection-only)
d) Heavy rigid box (27.7-36.0 kg)
e) Compliant 1-DoF hand
f) 6-DoF robot arm with an F/T sensor at the wrist



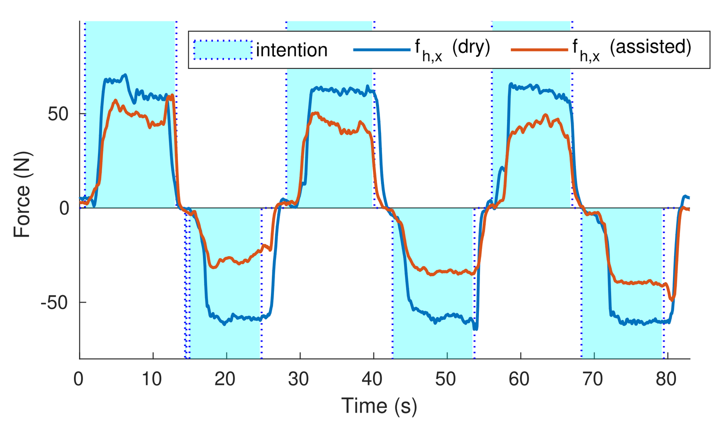
Dataset collection

- Data to train our prediction model
- Recorded:
  - Box pose $\mathbf{X}_b$
  - Human skeleton data $\mathbf{X}_h$
  - Human-side Force/Torque f_h
  - Robot-side Force/Torque f_r
  - Intention labels $\boldsymbol{i}_h$ (velocity threshold)
- Different speed, box weight, subject
- Idle actions (negatives)
- **250~ pull/push actions** in 28 sessions
- 22 for training 6 for validation
- 98k~ frames at 100 *Hz*

Robot-side forces

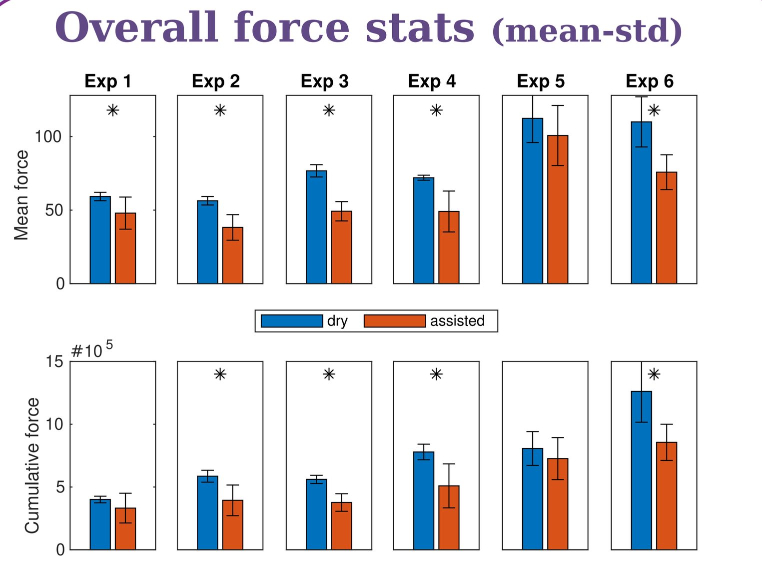


Intention & human force

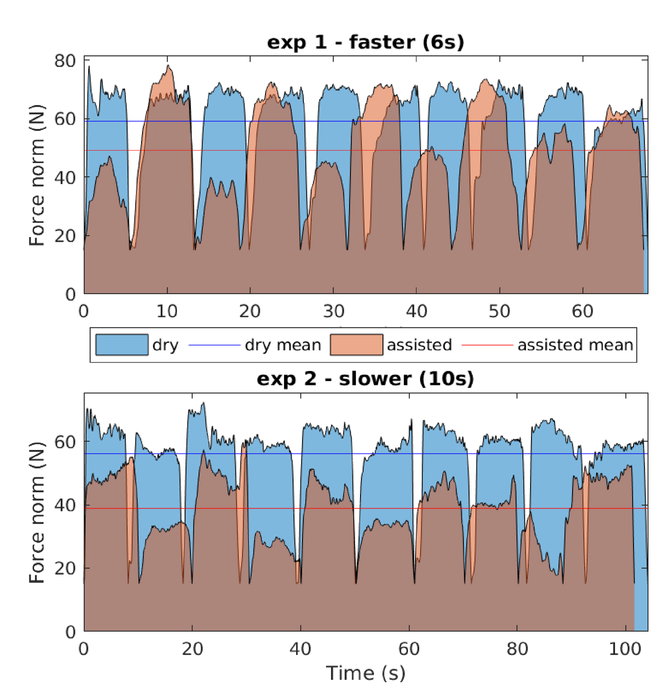


94,1% balanced accuracy on dataset

Overall force stats (mean-std)



Human effort comparison



Scan for the video!

Conclusion

Our context-aware assistive robotic system effectively reduces the human effort in the real-world experiments which presents a valuable benchmark solution.

We designed and analyzed the novel problem of collaboratively transporting heavy objects which is interesting both for its technical challenges and its practical use in industrial and domestic settings

Future works require analyzing generalization to new subjects and the effect of familiarity with the system in multi-dimensional object transportation.

References

[1]: Shi, Lei, et al. "Skeleton-based action recognition with directed graph neural networks." Proceedings of the IEEE/CVF conference on computer vision and pattern recognition. 2019.