2018 National Robotics Initiative (NRI) Principal Investigators' Meeting

MICHIGAN ENGINEERING
UNIVERSITY OF MICHIGAN

Scene Understanding and Predictive Monitoring for Safe Human-Robot Collaboration in Unstructured and Dynamic Construction Environments

SangHyun Lee[1], Vineet R. Kamat[1], Jia Deng[2], Daeho Kim[1], Ankit Goyal[2], Ci-Jyun P. Liang[1], and Meiyin Liu[1]

[1]Civil & Environmental Engineering, University of Michigan, Ann Arbor, MI.

[2]Computer Science, Princeton University, Princeton, NJ.

Introduction

1. Autonomous robots have drawn increased attention in the construction industry as a potential means of improving safety and productivity.
2. However, to ensure safe human-robot collaboration on unstructured and dynamic construction sites, advanced detection of impending hazards with timely intervention should be achieved.

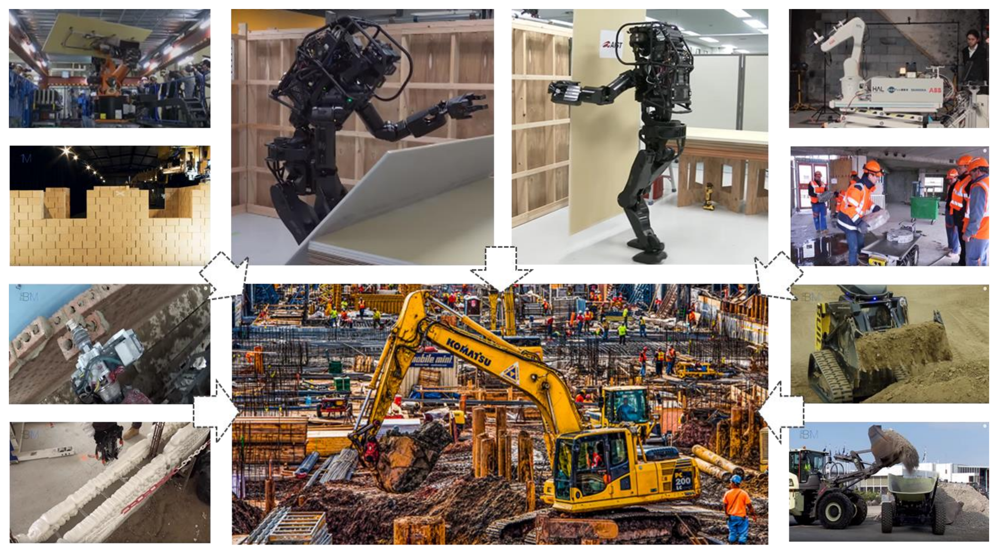

Figure 1. Construction Robots Being Developed in the World

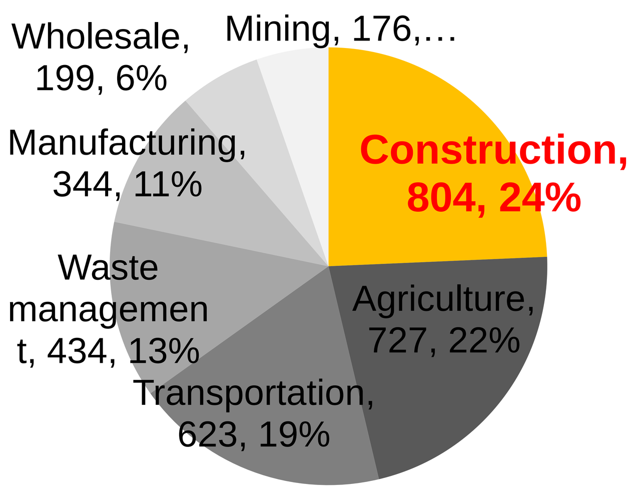

Figure 2. The Number of Struck-by Injuries in U.S. Industries, 2011-2015, CPWR 2017

Problems in current safety monitoring methods

Manual observation: costly and intermittent / Robots' built-in functions: limited sensing range.

Potential alternative: computer vision-based safety monitoring

Economical implementation / continuous monitoring / wide sensing range.

Challenges facing with computer vision methods

1. Lack of holistic scene understanding (e.g., semantic relation shared between entities).
2. Lack of predictive scene understanding (e.g., future locations of entities).
3. Severe occlusions in real construction sites.

Objective: Holistic, and predictive safety monitoring using multi-source cameras, kinematic sensor, computer vision and deep neural network.

Overview of Research Methods

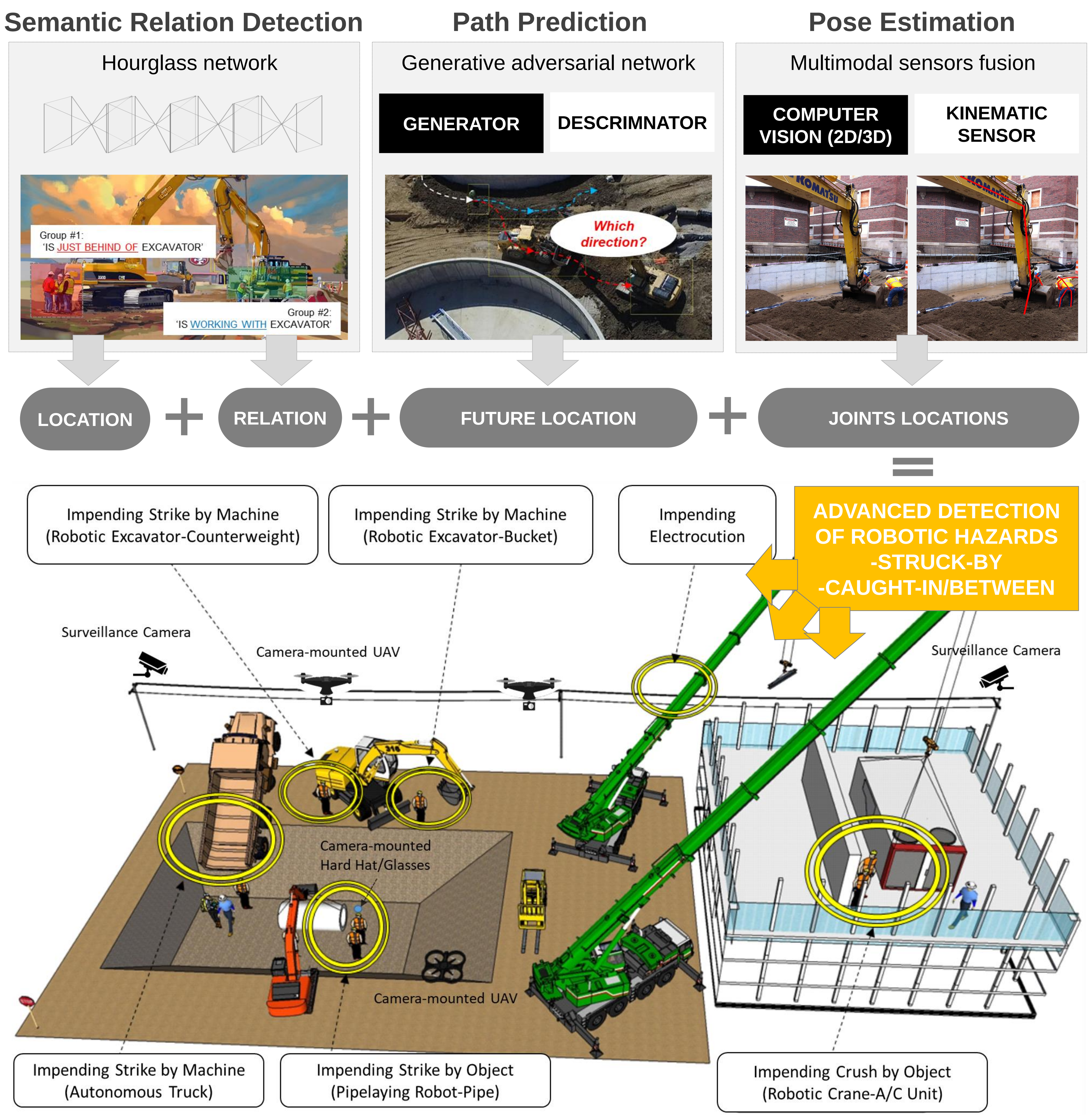

Figure 3. Overview of Research Methods

Semantic Relation Detection – Current Progress

One-stage recall@100: 88.89% & Two-stage recall@100:91.54%.

Px2Graph network (Newell and Deng 2017, NIPS 2017)

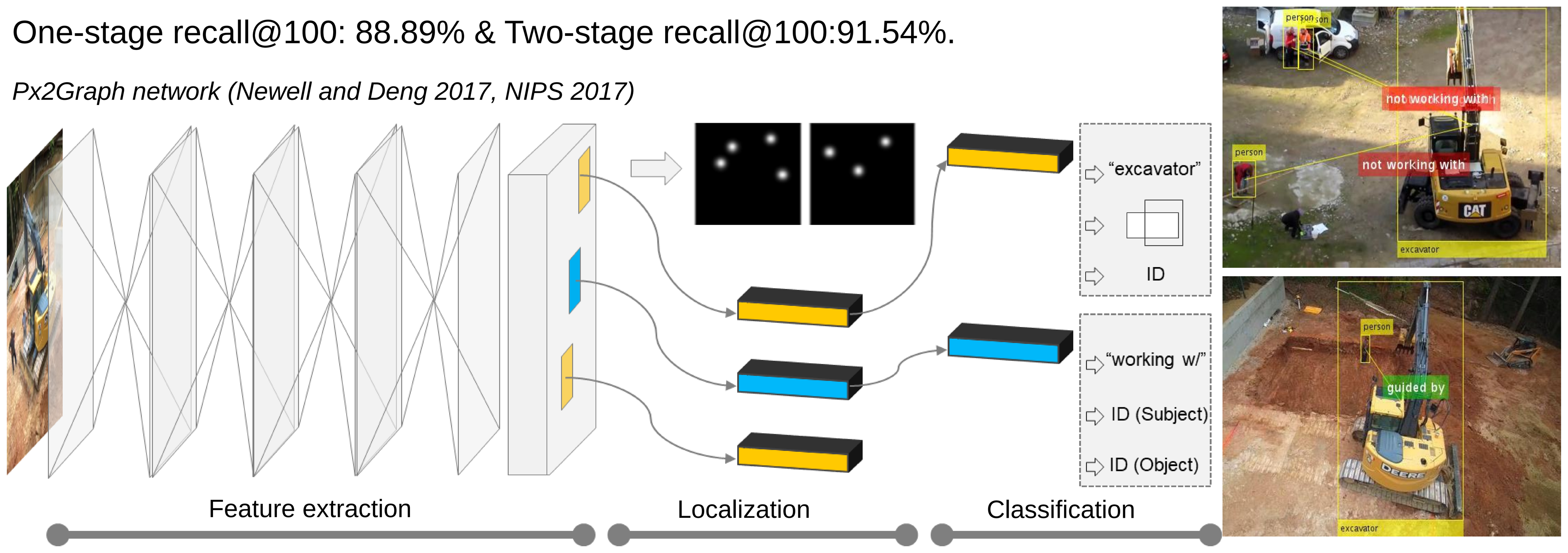

Figure 4. Semantic Relation Detection: Network Architecture & Test Examples

Path Prediction – Current Progress

Average distance error: 0.49 meters & Final destination error: 0.86 meters.
Prediction length: 80 frames (2.67 seconds).

Social GAN (Gupta et al. 2018, CVPR 2018)

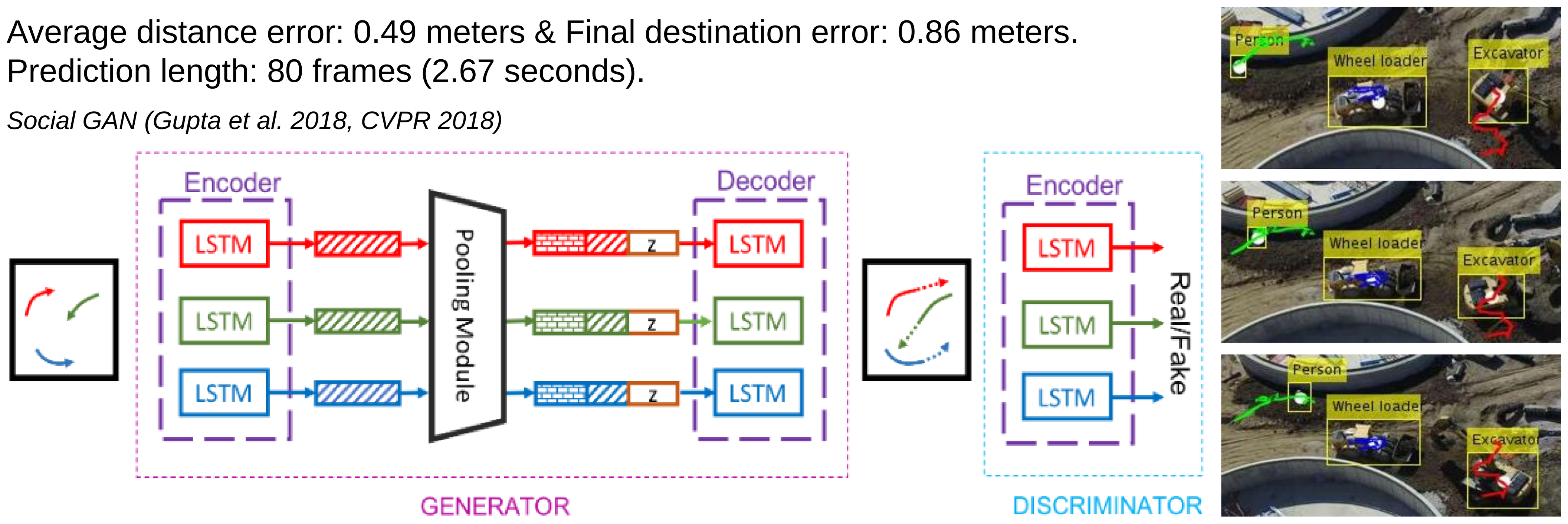

Figure 5. Path Prediction: Network Architecture & Test Examples

Pose Estimation – Current Progress

Average distance error: 4.06 cm in laboratory data and 71.84 pixels on real-site data

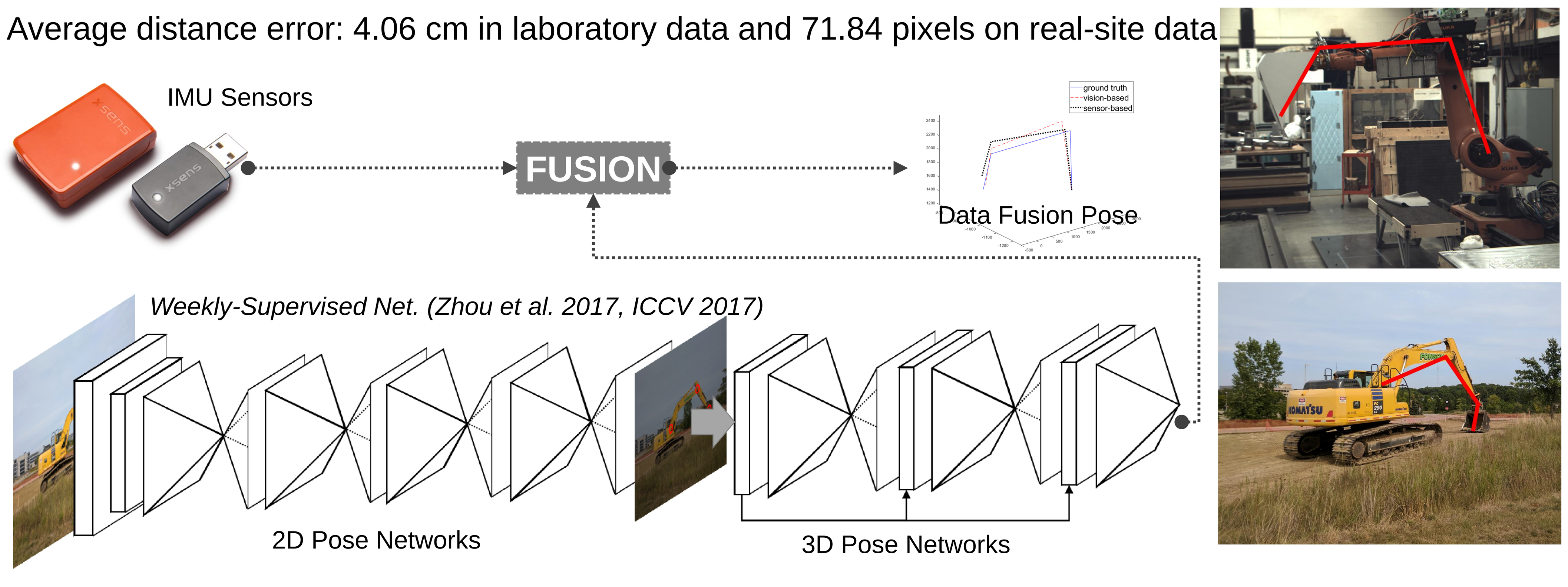

Figure 6. Pose Estimation: Framework & Test Examples

Conclusions

1. The models for semantic relation detection and path prediction can likely be enhanced through successful architecture modification and fine-tuning for construction settings.
2. The method for pose estimation can likely be improved by developing a fusion algorithm that can optimally integrate sensor-based and 3D vision-based pose estimation.
3. The outcomes from this research promise support for holistic and predictive safety monitoring in construction. It will ultimately help to promote safer human-robot collaboration in unstructured and dynamic construction sites.

Acknowledgments

This research is supported financially by a National Science Foundation Award (IIS-1734266). The researcher team would also like to acknowledge Mr. Weston Tanner and Mr. John McGlennon from the Walsh Construction Company for their assistance in collecting onsite construction operations data.