Environment-Aware Worker Trajectory Prediction using Surveillance Camera on Modular Construction Sites

Qiuling YANG[1], Qipei MEI[2], Chao FAN[3], Meng MA[4], and Xinming LI[5*]

[1] *Postdoctoral Fellow, Department of Mechanical Engineering, University of Alberta*
[2] *Assistant Professor, Department of Civil & Environmental Engineering, University of Alberta*
[3] *PhD Student, Department of Mechanical Engineering, University of Alberta*
[4] *Researcher, Department of Electrical & Computer Engineering, University of Minnesota*
[5] *Assistant Professor, Department of Mechanical Engineering, University of Alberta*
[*] *Corresponding author's e-mail: xinming.li@ualberta.ca*

ABSTRACT
Modular construction sites are often reported as one of the most hazardous workplaces where the complex environments can lead to near misses and life-threatening collisions. To avoid contact collisions and provide a safe workplace, forecasting workers' trajectories on dynamic construction sites is demanding yet remains challenging. Existing approaches for trajectory prediction are mostly limited to only considering the objects moving information. In this paper, an environment-aware distance worker trajectory prediction model is designed to fully exploit the contextual information on construction sites. Incorporating the interactions among workers and distances between workers and static elements into the prediction model, the proposed approach offers a reliable prediction of worker positions. To further exploit the contextual cues, an environment-aware direction scheme taking directional information of the static elements into account is put forth. Extensive numerical tests on synthetic as well as modular construction datasets showcase the improved prediction performance of the proposed approaches in comparison to several state-of-the-art alternatives.

KEYWORDS
Trajectory prediction; Environment-aware interactions; LSTM; Modular construction site

INTRODUCTION
Building structure components in a factory and transporting them on-site for installation, modular constructions provide a promising way to offer faster implementation, safer manufacturing, and more transportability compared with traditional construction methods (Bertram et al., 2019; Thai et al., 2020). However, its dynamic and information-intensive environment makes ensuring the safety of construction workers becoming a very challenging task (Li et al., 2013). Among various factors which directly and indirectly contribute to fatalities and injuries in construction workspace, contact collisions leading from the proximity of construction resources (i.e., workers, equipment, and materials), is one of the most obvious aspects (Teizer et al., 2008). Due to a lack of proper object coordination and solid planning, the congested modular construction site poses potential hazardous contact collisions and even life-threatening scenarios. For instance, in Fig. 1, the worker in the green bounding box is guiding moving a roof in purple bounding box with the direction in purple arrow, while the worker in the red bounding box is walking forward with the direction in red arrow.

Contact collision may occur when the worker in the red bounding box is focusing on his allocated task and fails to recognize the proximity with the moving roof. To mitigate the potential injuries, accurate forecasting of construction resources' future positions is critical and can serve as a foundation for building the future proactive and informative contact collision prevention systems.

Figure 1. Potential hazard on a modular construction site.

To this end, considerable research efforts have been devoted to getting a reliable prediction of different subjects for various applications (Altche' et al., 2017; Lasota et al., 2017; Alahi et al., 2016; Rudenko et al., 2020; Kothari et al., 2021; Yang et al., 2019; Bartoli et al., 2018; Ma et al., 2018; Cai et al., 2020). To guarantee safely and efficiently driving on public roads, a long short-term memory (LSTM) network is introduced for predicting self-driving vehicles trajectories in (Altche' et al., 2017). Integrating convolution neural networks and LSTM, a hybrid learning model is proposed in (Ding et al., 2018) to detect construction worker unsafe behaviors. In the traffic management domain, an LSTM network with a novel pooling layer to capture the interactions among different pedestrians is pursued in (Alahi et al., 2016). Nonetheless, these approaches suffer limited predictive capability as they do not consider the contextual information of scenes.

To address this limitation, two environment-aware trajectory prediction algorithms for construction workers are proposed in this paper. Specifically, explicitly considering workers' movement as well as the contextual information, an environment-aware distance (EA-Distance) forecasting scheme is first developed. Every worker path is modelled through an LSTM network with a novel pooling that captures the interactions among workers and the distance with surrounding static objects. Furthermore, leveraging the directional information of the surrounding static objects relative to the worker, an environment-aware direction (EA-Direction) model is proposed. Different from the previous works (Alahi et al., 2016; Kothari et al., 2021), our proposed methods offer a systematic and flexible framework to incorporate more general information into the prediction model on construction sites.

METHODS

When conducting construction activities, workers are capable of navigating their behavior to avoid potential collisions by sensing the surrounding contextual information. In this study, taking these rich contextual features into account, two environment-aware trajectory prediction models are built to forecast future movements of workers in the construction workspace using LSTM techniques.

LSTM network has demonstrated remarkable performance in many sequential problems, such as machine language translation, acoustic modelling, and activity recognition (Yu et al., 2019). Compared to a vanilla recurrent neural network structure, LSTM network contains learnable gating functions, that are, input gate contributing new information to be stored in the memory, forget gate deciding the forget degree of the internal state, and output gate computing output from previous states. This gate design allows the LSTM network to manage the internal memory state when dealing with long time series (Hochreiter et al, 1997).

Given a construction video sequence with N workers, the position of worker $i \in N$ at time instant t is represented in the form of its spatial coordinates (x_i^t, y_i^t). Having the observations at time instants $t = 1, \ldots, T_{obs}$, i.e., $\{(x_i^t, y_i^t)\}_{i=1:N}^{t=1:T_{obs}}$, the goal of worker trajectory prediction is to forecast each worker position at the future time instants $t = T_{obs}+1, \ldots, T_{pred}$, that is $\{(x_i^t, y_i^t)\}_{i=1:N}^{t=T_{obs}+1:T_{pred}}$. The LSTM cell of worker i is updated as following

$$\begin{aligned}
\boldsymbol{i}^t &= \sigma(\boldsymbol{W}_{ii}\boldsymbol{s}_i^t + \boldsymbol{b}_{ii} + \boldsymbol{W}_{hi}\boldsymbol{h}^{t-1} + \boldsymbol{b}_{hi}) \\
\boldsymbol{f}^t &= \sigma(\boldsymbol{W}_{if}\boldsymbol{s}_i^t + \boldsymbol{b}_{if} + \boldsymbol{W}_{hf}\boldsymbol{h}^{t-1} + \boldsymbol{b}_{hf}) \\
\boldsymbol{g}^t &= tanh(\boldsymbol{W}_{ig}\boldsymbol{s}_i^t + \boldsymbol{b}_{ig} + \boldsymbol{W}_{hg}\boldsymbol{h}^{t-1} + \boldsymbol{b}_{hg}) \\
\boldsymbol{o}^t &= \sigma(\boldsymbol{W}_{io}\boldsymbol{s}_i^t + \boldsymbol{b}_{io} + \boldsymbol{W}_{ho}\boldsymbol{h}^{t-1} + \boldsymbol{b}_{ho}) \\
\boldsymbol{c}^t &= \boldsymbol{f}^t \odot \boldsymbol{c}^{t-1} + \boldsymbol{i}^t \odot \boldsymbol{g}^t \\
\boldsymbol{h}^t &= \boldsymbol{o}^t \odot tanh(\boldsymbol{c}^t)
\end{aligned} \tag{1}$$

where σ represents the sigmoid function; $\odot$ is Hadamard product; the input data $\mathbf{s}_i^t$ is obtained by feeding the position (x_i^t, y_i^t) of worker i at t into a ReLU network with a weight matrix $\mathbf{W}_{xy}$, i.e., $\mathbf{s}_i^t = \phi(x_i^t, y_i^t; \mathbf{W}_{xy})$; $\mathbf{i}^t$, $\mathbf{f}^t$, $\mathbf{g}^t$ and $\mathbf{o}^t$ represent input gate, forget gate, cell gate, and output gate, respectively; $\mathbf{c}^t$ is the cell state; and $\mathbf{h}^t$ is the hidden state at time instant t. Parameters $\mathbf{W}_{ii}$, $\mathbf{b}_{ii}$, $\mathbf{W}_{hi}$, $\mathbf{b}_{hi}$, $\mathbf{W}_{if}$, $\mathbf{b}_{if}$, $\mathbf{W}_{hf}$, $\mathbf{b}_{hf}$, $\mathbf{W}_{ig}$, $\mathbf{b}_{ig}$, $\mathbf{W}_{hg}$, $\mathbf{b}_{hg}$, $\mathbf{W}_{io}$, $\mathbf{b}_{io}$, $\mathbf{W}_{ho}$, and $\mathbf{b}_{ho}$ are learnable in each LSTM cell. For simplicity of future illustration, let us concatenate these parameters into a vector $\boldsymbol{\theta}$. With various gates, LSTM can transfer relative information down a sequence chain after learning what information is relevant to remember and what is irrelevant to forget.

To predict the position of worker i at next time instant $t+1$, i.e., (x_i^{t+1}, y_i^{t+1}), a bivariate Gaussian distribution $\mathrm{N}(\mu_i^{t+1}, \sigma_i^{t+1}, \rho_i^{t+1})$ is considered, where μ_i^{t+1}, σ_i^{t+1}, and ρ_i^{t+1} is the mean, standard deviation, and correlation coefficient, respectively. These Gaussian distribution parameters are estimated by applying a linear transformation $\mathbf{W}_o$ to the output hidden state $\mathbf{h}_i^t$, i.e., $[\mu_i^{t+1}, \sigma_i^{t+1}, \rho_i^{t+1}] = \mathbf{W}_o \mathbf{h}_i^t$. Therefore, a meaningful approach to learning the parameters of worker i entails minimizing the negative log-Likelihood loss as follows (Alahi et al., 2016)

$$L_i(\mathbf{W}_{xy}, \boldsymbol{\theta}, \mathbf{W}_o) = -\sum_{t=T_{obs}+1}^{T_{pred}} \log\left(\mathrm{P}(x_i^t, y_i^t \mid \mu_i^t, \sigma_i^t, \rho_i^t)\right). \tag{2}$$

As a powerful tool for modeling time-dependent data, this vanilla LSTM network provides a good prediction for worker trajectories. However, the environment of modular construction site is always complicated, and multiple aspects can influence the path of workers. The vanilla LSTM algorithm fails to capture the rich interactions among workers and their surroundings. To address this issue, we propose environment-aware trajectory prediction algorithms to incorporate context information in the model.

Worker-to-worker interactions

As workers keep interacting with each other to collaborate on a certain task, worker-to-worker interactions are important for predicting the worker's path. In reality, each worker of interest has a different number of neighbors, handling the variable number of neighbors is a challenge for aggregating the interaction information. To this end, a local grid is constructed around the worker with each cell of the grid standing for the information of neighborsin this grid. Two main forms of the grid information are LSTM with occupancy maps (O-LSTM) where every grid cell denotes the position of neighbors, and social LSTM (S-LSTM) where every grid cell indicates the LSTM hidden state of the neighbors.

Having all neighbors positions $\{(x_j^t, y_j^t)\}_{j\in \mathrm{N}_i}$ of worker i, with N_i representing the collection of neighbors of worker i, the occupancy map pooling can be mathematically expressed as follows

$$O_i^t(m,n) = \sum_{j\in \mathrm{N}_i} \mathbf{1}_{mn}[x_j^t - x_i^t, y_j^t - y_i^t] \tag{3}$$

where $\mathbf{1}_{mn}[a,b]$ is an indicator function taking value 1 if (a,b) is in the (m,n) cell of the gird, and 0 otherwise. The occupancy map $\mathbf{O}_i^t$ is embedded into a vector $\mathbf{e}_{oi}^t = \phi(\mathbf{O}_i^t, \mathbf{W}_{eo})$, which constitutes a new input $\tilde{\mathbf{s}}_i^t$ by concatenating with the original input vector $\mathbf{s}^t$, i.e., $\tilde{\mathbf{s}}_i^t = [\mathbf{s}_i^t \;\; \mathbf{e}_{oi}^t]$. Consequently, worker i trajectory can be predicted by O-LSTM model by feeding this new input $\tilde{\mathbf{s}}_i^t$ is into the LSTM model (1). Leveraging position information of neighbors, the O-LSTM model can make predictions and avoid immediate collisions. Along this idea, S-LSTM model takes one step further to get a smooth prediction by containing the history information of neighbors in the pooling layer. Since the hidden states of an LSTM network can capture history time varying motion-properties of a worker, the pooling layer takes into account all neighbor hidden states as

$$H_i^t(m,n,:) = \sum_{j\in \mathrm{N}_i} \mathbf{1}_{mn}\left[x_j^t - x_i^t, y_j^t - y_i^t\right]\mathbf{h}_j^{t-1}. \tag{4}$$

Similarly, this social pooling $\mathbf{H}_i^t$ is fed into a ReLU network to get the embedding vector $\mathbf{e}_{hi}^t = \phi(\mathbf{H}_i^t, \mathbf{W}_{eh})$. Substituting $\mathbf{e}_{hi}^t$ for $\mathbf{e}_{oi}^t$ in the input vector $\tilde{\mathbf{s}}_i^t$, the predictions of the worker i trajectory can be obtained by updating LSTM model using (1).

Environment-to-worker interactions

Worker behavior is collision avoidance-based, which is not only influenced by the factor of interactions with neighbors, but also the static objects located in the environment in which a worker is moving. It is expected that the worker will inherently circumvent these static objects to conduct the allocated tasks. Thus, in addition to the worker-to-worker interaction, another common interaction in the construction workspace is the environment-to-worker interaction. This paper takes the static obstacle as an example of the objects in the environment for easy illustration.

As the obstacle is static and sparse in the environment, the spatial coordinates of these obstacles can be obtained as prior information and no further prediction is needed. A straightforward way to take obstacles into account in the problem is by extending the grid cell information in (3) or (4). In the construction workspace, however, those static objects usually are located sparsely. Modeling environment-to-worker interactions of all obstacles can lead to the model learning spurious correlations. Thus, we propose to consider only the top closest obstacles relative to the worker of interest. Specifically, considering top-K static obstacles around worker i, and letting (x_{dk}, y_{dk}) denote the position of k obstacle, $k \in K$, the Euclidean distance between worker i and obstacle

k can be expressed as

$$d_i^t(k) = \sqrt{(x_i^t - x_{dk})^2 + (y_i^t - y_{dk})^2}. \tag{5}$$

The embedding vector of the distance model $\mathbf{e}_{di}^t$ can be obtained by $\mathbf{e}_{di}^t = \phi(\mathbf{d}_i^t, \mathbf{W}_{ed})$. The proposed EA-Distance model entails concatenating worker-to-worker interaction vector $\mathbf{e}_{hi}^t$ as well as environment-to-worker interactions vector $\mathbf{e}_{di}^t$ to constitute a new input $\tilde{\mathbf{s}}_i^t$, that is

$$\tilde{\mathbf{s}}_i^t = [\mathbf{s}_i^t \;\; \mathbf{e}_{hi}^t \;\; \mathbf{e}_{di}^t]. \tag{6}$$

This new input $\tilde{\mathbf{s}}_i^t$ is fed into LSTM model (1) to get predictions of worker i. Notice in the back-propagation, the loss function of worker i in (2) introduces new parameters $\mathbf{W}_{eh}$ and $\mathbf{W}_{ed}$ as

$$L_i(\mathbf{W}_{xy}, \mathbf{W}_{eh}, \mathbf{W}_{ed}, \boldsymbol{\theta}, \mathbf{W}_o) = -\sum_{t=T_{obs}+1}^{T_{pred}} \log\left(\mathrm{P}(x_i^t, y_i^t \mid \mu_i^t, \sigma_i^t, \rho_i^t)\right). \tag{7}$$

On the other hand, the directional relationship between obstacles and workers also has different influences on worker movement. For instance, the obstacle has much less influence on worker paths when workers are deviating it than when workers are approaching it. Therefore, we further propose an environment-aware direction model, which entails a directional vector $\mathbf{q}_i^t(k)$ to capture this directional information of obstacle k relative to worker i as

$$\mathbf{q}_i^t(k) = \left(\frac{x_i^t - x_{dk}}{\| x_i^t - x_{dk} \|}, \frac{y_i^t - y_{dk}}{\| y_i^t - y_{dk} \|}\right). \tag{8}$$

The direction embedding vector is $\mathbf{e}_{qi}^t = \phi(\mathbf{q}_i^t, \mathbf{W}_{eq})$. Replacing the distance embedding vector $\mathbf{e}_{di}^t$ $\mathbf{e}_{qi}^t$ in (6), the proposed EA-Direction model can be achieved by (1).

The proposed environment-aware worker trajectory prediction schemes EA-Distance and EA-Direction are implemented in two phases, namely offline training as well as online inference phases. Specifically, in the training phase, the model is fed with the set of observations $\{\mathbf{x}^t, \mathbf{y}^t\}^{t=1:T_{obs}}$ of all the workers at time instants $t = 1, \ldots, T_{obs}$ and trained to output the forecasting. In the inference process, the observations of all workers $\{\mathbf{x}^t, \mathbf{y}^t\}^{t=1:T_{obs}}$ are fed into the trained model, and the near future forecasting $\{\mathbf{x}^t, \mathbf{y}^t\}^{t=T_{obs}+1:T_{pred}}$ can be obtained by considering worker-to-worker interactions and environment-to-worker interactions.

RESULTS AND DISCUSSION

In this section, extensive experiments on both synthetic dataset and real constructionsite are implemented to showcase the effectiveness of our proposed methods.

Implementation details and evaluation metrics

Consisting of relative work (Kothari et al., 2021), the length of observation is 3.6 seconds, while the forecasting length is 4.8 seconds. With frame rate setting as 0.4, we use the past 9 frames to predict 12 future frames. Two most widely used metrics of trajectory prediction are entailed for evaluation, that are, average displacement error (ADE) as well as final displacement error (FDE)

(Pellegrini et al., 2009; Alahi et al., 2016). ADE is the mean square error overall predicted positions of the path and the ground-truth path as follows

$$ADE := \frac{\sum_{i=1}^{N} \sum_{t=T_{obs}+1}^{T_{pred}} \| (\hat{x}_i^t, \hat{y}_i^t) - (x_i^t, y_i^t) \|}{N \times (T_{pred} - T_{obs} - 1)} \tag{9}$$

where $(\hat{x}_i^t, \hat{y}_i^t)$ and (x_i^t, y_i^t) are the predicated coordinate and ground-truth coordinate, respectively. FDE is the distance between the predicated destination and the ground-truth one as

$$FDE := \frac{\sum_{i=1}^{N} \| (\hat{x}_i^{T_{pred}}, \hat{y}_i^{T_{pred}}) - (x_i^{T_{pred}}, y_i^{T_{pred}}) \|}{N}. \tag{10}$$

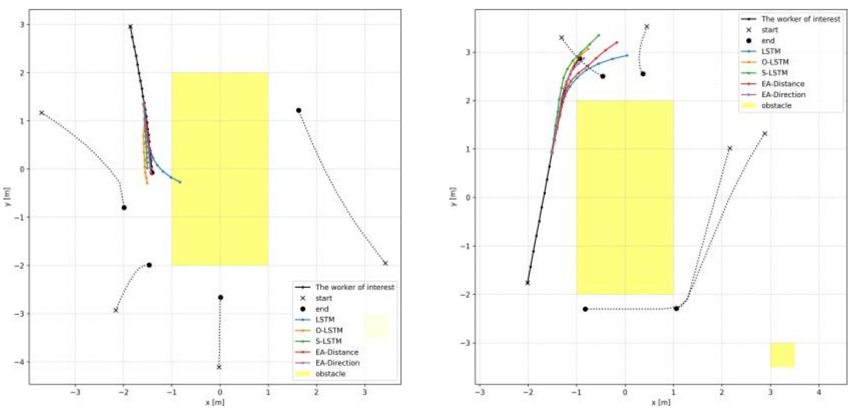

Figure 2: Two example scenes of synthetic data with different prediction models.

Synthetic data experiments

The first experiment builds on a synthetic dataset generated from Optimal Reciprocal Collision Avoidance (ORCA), which provides sufficient conditions for collision avoidance motion (Van et al., 2008). To assess the performance of the proposed methods, we have simulated three trajectory prediction policies as baselines, namely vanilla LSTM model, O-LSTM model (Alahi et al., 2016), and S-LSTM model (Alahi et al., 2016). Table 1 illustrates experimental results for the five different models. Evidently, by incorporating worker-to-worker interactions and environment-to-worker interactions into the prediction model, the proposed EA-Distance and EA-Direction schemes can attain smaller errors compared with the alternatives. Fig.2 depicts the qualitative results of the proposed schemes in comparison to alternatives. Curves show the prediction trajectories of our proposed schemes are close to the ground-truth.

Table 1. Prediction performance of different models on a synthetic dataset.

Model name	**ADE**	**FDE**
LSTM	0.35	0.87
O-LSTM (Alahi et al., 2016)	0.31	0.74
S-LSTM (Alahi et al., 2016)	0.27	0.64
EA-Distance (proposed model)	0.26	0.60
EA-Direction (proposed model)	0.23	0.51

Construction data experiments

The second experiment entails real modular construction data from a leading offsite construction facility located in Edmonton, Canada. The dataset consists of 7 moving workers as well as 2 static obstacles in a total of 1226 scenes. The surveillance video is obtained by an oblique view, therefore a homography matrix transferring image to real-world coordinatesis estimated from four manually selected points on the ground with estimated measurements. Upon this homography matrix, positions of workers and static obstacles are extracted. ORCA is implemented to generate training data with the same static obstacles' information, while in the testing phase, the construction dataset is adopted to evaluate the performance of different models. Table 2 provides the performance of the five different prediction methods on the construction dataset, while Fig. 3 illustrates two example scenes. Result shows the proposed environment-aware based prediction models can achieve competitive performance relative to the state-of-art alternatives.

Table 2. Prediction performance of different models on a modular construction dataset.

Model name	**ADE**	**FDE**
LSTM	1.68	2.87
O-LSTM (Alahi et al., 2016)	1.52	2.45
S-LSTM (Alahi et al., 2016)	1.40	1.92
EA-Distance (proposed model)	1.43	1.99
EA-Direction (proposed model)	1.33	1.82

Figure 3: Two example scenes of construction data with different prediction models.

CONCLUSION

This research provides a systematic and flexible framework to incorporate rich contextual information into trajectory prediction model to improve the current practice of worker trajectory forecasting in construction sites. Specifically, considering environment-to-worker interactions by investigating collision avoidance of static obstacles, two different environment-aware based prediction models, namely EA-Distance and EA-Direction, are devised to predict immediate future positions of workers. Notice though the present work focuses on obstacles, the proposed framework can account for other types of static objects in the environment as well. Numerical tests showcase the competitive performance of our proposed prediction models relative to several existing ones. Future directions include investigating the influence moving equipment makes on worker trajectories as well as developing hazard detection systems.

ACKNOWLEDGMENTS

This study was funded by the Natural Sciences and Engineering Research Council of Canada (NSERC) through Alliance Grant with Alberta Innovates (File No. ALLRP 561120 - 20). The authors would also like to acknowledge the support from their industry partner in Canada.

REFERENCES

Alahi, A., Goel, K., Ramanathan, V., Robicquet, A., Fei-Fei, L., & Savarese, S. (2016). Social LSTM: human trajectory prediction in crowded spaces. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition* (pp. 961–971).

Altché, F., & de La Fortelle, A. (2017). An LSTM network for highway trajectory prediction. *IEEE International Conference on Intelligent Transportation Systems* (pp. 353–359).

Bartoli, F., Lisanti, G., Ballan, L., & Del Bimbo, A. (2018). Context-aware trajectory prediction. In *2018 24th International Conference on Pattern Recognition* (pp. 1941–1946).

Bertram, N., Fuchs, S., Mischke, J., Palter, R., Strube, G., & Woetzel, J. (2019). Modular construction: From projects to products. *Capital Projects*, 1–34.

Cai, J., Zhang, Y., Yang, L., Cai, H., & Li, S. (2020). A context-augmented deep learningapproach for worker trajectory prediction on unstructured and dynamic construction sites. *Advanced Engineering Informatics*, *46*, 101173. doi: 10.1016/j.aei.2020.101173

Ding, L., Fang, W., Luo, H., Love, P. E. D., Zhong, B., & Ouyang, X. (2018).A deep hybrid learning model to detect unsafe behavior: Integrating convolution neural networks and long short-term memory. *Automation in Construction*, *86*, 118–124.

Hochreiter, S., & Schmidhuber, J. (1997). Long Short-Term Memory. *Neural Computation*, *9*(8), 1735–1780. doi: 10.1162/neco.1997.9.8.1735

Kothari, P., Kreiss, S., & Alahi, A. (2021). Human trajectory forecasting in crowds: A deep learning perspective. *IEEE Transactions on Intelligent Transportation Systems*, 1–15.

Lasota, P. A., Fong, T., & Shah, J. A. (2017). A survey of methods for safe human-robot interaction. *Foundations and Trends in Robotics*, *5*(3), 261–349. doi: 10.1561/2300000052

Li, H. X., Al-Hussein, M., Lei, Z., & Ajweh, Z. (2013). Risk identification and assessment of modular construction utilizing fuzzy analytic hierarchy process (AHP) and simulation. *Canadian Journal of Civil Engineering*, *40*(12), 1184–1195.

Ma, M., Nikolakopoulos, A.N. and Giannakis, G.B., 2018. Hybrid ADMM: a unifying and fast approach to decentralized optimization. *EURASIP Journal on Advances in Signal Processing*, 2018(1), pp.1-17.

Pellegrini, S., Ess, A., Schindler, K., & Van Gool, L. (2009). You'll never walk alone: Modeling social behavior for multi-target tracking. *IEEE international conference on computer vision* (pp. 261–268).

Rudenko, A., Palmieri, L., Herman, M., Kitani, K. M., Gavrila, D. M., & Arras, K. O. (2020, July). Human motion trajectory prediction: a survey. *The International Journal of Robotics Research*, *39*(8), 895–935. doi: 10.1177/0278364920917446

Teizer, J., Venugopal, M., & Walia, A. (2008). Ultrawideband for automated real-time three-dimensional location sensing for workforce, equipment, and material positioning and tracking. *Transportation Research Record*, *2081*(1), 56–64.

Thai, H.-T., Ngo, T., & Uy, B. (2020). A review on modular construction for high-risebuildings. *Structures*, *28*, 1265–1290.

Van Den Berg, J., Lin, M., & Manocha, D. (2008). Reciprocal Velocity Obstacles for real-time multi-agent navigation. *IEEE International Conference on Robotics and Automation* (pp. 1928–1935). doi: 10.1109/ROBOT.2008.4543489

Yang, Q., Wang, G., Sadeghi, A., Giannakis, G. B., and Sun, J., "Two-timescale voltage control in distribution grids using deep reinforcement learning," *IEEE Transactions on Smart Grid*, vol. 11, no. 3, pp. 2313–2323, 2019.

Yu, Y., Si, X., Hu, C., & Zhang, J. (2019). A review of recurrent neural networks:LSTM cells and network architectures. *Neural Computation*, *31*(7), 1235–1270.