Predicting roadway workers' safety behaviour in short-term work zones

Daniel Lu, Semiha Ergan
New York University, United States
dbl299@nyu.edu

Abstract. Towards a more detailed understanding of how roadway workers' instinctive reactions (e.g., moving their entire body, turning their head) to traffic contribute to fatal work zone crashes, there is a need for accurate models of workers' behaviour in these dangerous scenarios. While related studies address similar challenges in vertical building construction sites, this study proposes a deep learning model for worker behaviour in horizontal roadway work zones, building on a previously developed wearable sensor and virtual reality (VR) platform designed to capture safety related behavioural data (e.g., head position and orientation, field of view) on workers exposed to hazardous traffic vehicles as they experience immersive simulations of real-world roadway work zones. Using gated recurrent units (GRUs) trained with behavioural data collected from the platform, the deep learning model's accuracy (i.e., average displacement error) in predicting a worker's future position trajectory will be evaluated.

1. Introduction

Roadway work zones in the United States are now experiencing more than 700 fatal crashes resulting in over 800 deaths annually, with roadway construction workers accounting for around one in seven of these fatalities (American Road & Transportation Builders Association (ARTBA), 2023). While the majority of past research in transportation and construction safety have focused on characteristics of vehicle motorists, only a few studies have focused on the attributes of the roadway work zones, such as work zone layout and construction activities (Ergan et al., 2020). The most recent studies have only begun to investigate workers' instinctive physical reactions to dangerous vehicles, with a significant gap remaining in understanding how roadway workers' individual actions and behaviour contributes to accidents in work zone crashes (Thapa & Mishra, 2021). Traffic safety research and industry efforts to improve work zone safety have led to standards and guidelines on workers' safety behaviour, including use of personal protection equipment (PPE), layout of traffic control devices (e.g., cones, barriers), and attention to hazards during construction activities (Ergan et al., 2020; Thapa & Mishra, 2021). While these guidelines offer workers a clear checklist for preparing roadway work zones, they remain vague on what workers should specifically do during construction work to maintain awareness of surrounding traffic and construction hazards. This awareness of and response to traffic can become a particular concern in short-term work zones where fewer traffic control devices are used.

The lack of actionable worker safety behavioural guidelines is partly due to limited accident data on how specific roadway workers' safety behaviour during construction (e.g., turning their head towards traffic) contributes to rates of work zone fatalities and injuries. To better understand construction workers' safety behaviour, research has proposed approaches for real-world data collection that deploy real-time location systems for workers and equipment, a costly and potentially dangerous endeavour for workers involved (Luo et al., 2016). Data collected from such systems could be used to build accurate and detailed (i.e., high fidelity) statistical models that predict how construction workers move around the work zone (i.e., their walking trajectory) as they work and in response to potential hazards like moving construction equipment or traffic vehicles. Yet in their attempts to address practical implementation challenges, the real-world sensor approaches face constraints in the categories of data they can collect on workers, which can further restrict the classes and accuracy of models for predicting workers' safety behaviour. Recent virtual reality (VR) based studies have started to demonstrate that research can collect a variety of data regarding a roadway worker's behaviour, enabling a wider exploration and development of more data-driven statistical models that relate workers' specific actions to their consequent safety in the work zone (Gao et al., 2022; N. Kim et al., 2021). A key question remains in whether or not models of worker behaviour can improve prediction accuracy with the inclusion of data tracking a worker's context, such as their relative distance to other workers and construction equipment (Cai et al., 2020).

This study evaluates different deep learning models for predicting a time series of a worker's behaviour in short-term roadway work zones, building on a previously developed wearable sensor and virtual reality (VR) platform designed to capture safety related behavioural and context data (e.g., head position, head orientation, field of view) on workers exposed to hazardous traffic vehicles as they experience immersive simulations of real-world roadway work zones. Recurrent neural networks (RNN), using different cells such as long short-term memory (LSTM) and gated recurrent units (GRU), were trained with behavioural data collected from the platform to predict a worker's future position trajectory, accounting for different categories of information such as the worker's previous positions, their head orientation, and other relevant context with respect to traffic. All RNN models were evaluated in terms of their prediction accuracy (i.e., average displacement error) to assess the performance of different network architectures and the usefulness of additional context input data. Overall, this paper provides a case study of how a high-fidelity roadway worker behaviour model can be developed using VR data collection and deep learning networks.

2. Related Work

2.1 VR Studies on Roadway Worker Safety Behaviour

While VR is being widely experimented with as a construction safety training tool, only a few recent studies have explored using VR for analyzing workers' unsafe behaviour in construction sites (Ergan et al., 2020; Gao et al., 2022; N. Kim et al., 2021). Utilizing a VR headset as a wearable sensor that tracks a person's head location, orientation, and visual attention (i.e., eye tracking), it is possible for researchers to analyze a workers' second-by-second behaviour and instinctive reactions to safety hazards. While there are numerous studies using VR to model vertical building construction sites, roadway work zones have only become a focus of VR-based studies in recent years. In a 2021 study, Kim, et. al. specifically used VR to simulate roadway work zones and tracked how far away vehicles were to an individual worker at the moment that worker raised their head to check if the vehicle was too close (i.e., checking distance). Using multi-level models estimating a worker's checking distance over time, this study examined the rate at which roadway workers became less vigilant and more habituated to risks of vehicle crashes. Overall, past VR studies build statistical and machine learning models to analyze trends within workers' safety behaviour but do not utilize the models to predict future worker behaviour. In contrast, this paper utilizes a VR environment to collect a time series of data on roadway worker's safety behaviour and context to develop a prediction model of how roadway workers will move (i.e., trajectory) while they work and in response to surrounding traffic.

2.2 Human Trajectory Prediction with Deep Learning Methods

Human trajectory prediction involves estimating how a person moves in a specified time frame, given that person's prior movement and other potentially relevant past information. Previous research in both vertical construction worker and on-road pedestrian trajectory prediction have evaluated a variety of different deep learning model architectures (Deo & Trivedi, 2017; D. Kim et al., 2019; Tang et al., 2019). Due to the sequence-to-sequence nature of trajectory problems, recurrent neural networks (RNNs) have been considered, given their well-established performance in natural language and machine translation applications (Cho et al., 2014; Sak et al., 2014). Long short-term memory (LSTM) and gated recurrent unit (GRU) cells are two commonly used units in recurrent networks and can be simply used to generate predictions recursively (Figure 1, top). The potential drawback of this recursive approach is that any error in the first predicted position can lead to greater errors in subsequent position predictions (i.e., error accumulation). A well-known improvement to this approach is the encoder-decoder RNN architecture, which tries to reduce those errors by using one recurrent network (i.e., encoder) to interpret the entire input sequence before using a separate recurrent network (i.e., decoder) to predict an output sequence (Cho et al., 2014). The encoder's final hidden state, which can better capture the internal structure of input data, is passed to the decoder network to make more accurate predictions (Figure 1, bottom).

In the construction worker domain, a 2020 study by Cai, et. al. evaluated LSTM units in both recursive and encoder-decoder architectures to predict construction workers' trajectory accounting for their prior positions and other context information, such as the worker's social group and general head direction (e.g., north, south, east, west). While encoder-decoder LSTM networks achieved higher accuracy than recursive prediction architectures, the accuracy of models using extra context input data were, on average, comparable to those which only accounted for a workers' past positions. A recent virtual reality domain paper focused more on predicting people's movement while using VR systems found that GRU-based encoder-decoder network architectures can outperform LSTM-based ones in human trajectory prediction (Lemic et al., 2022). Motivated by these findings, this paper explores a novel application of both LSTM and GRU-based models to predict the trajectory of roadway workers, given their prior positions and safety context with respect to traffic.

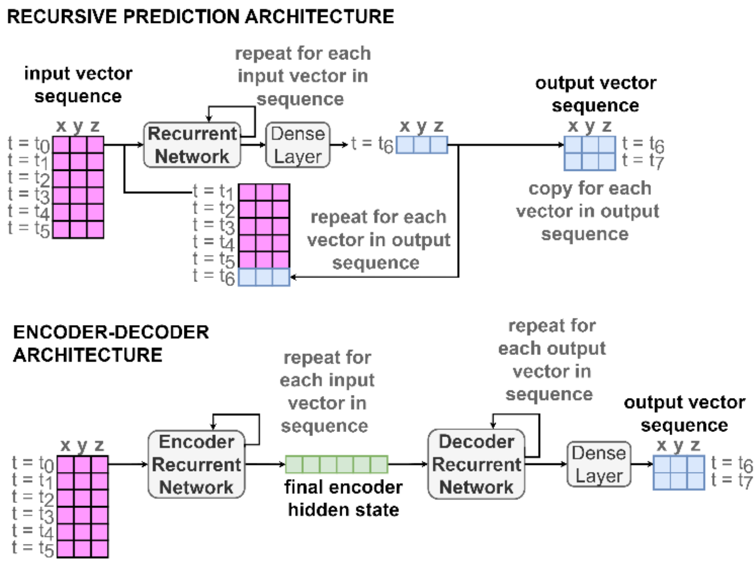

Figure 1: Different recurrent network approaches for predicting the same length of output sequence (i.e., worker's future trajectory).

3. Methodology

The approach for this study was to collect time series data on roadway worker behaviour on an integrated VR-traffic simulation platform and then use that collected train RNNs for perform worker trajectory prediction (i.e., offline prediction). Details on data collection, pre-processing that data for deep learning models, RNN architectures considered, and training model hyperparameters are discussed in the following sections.

3.1 VR Data Collection Platform

Prior work by the research team developed an integrated VR and traffic simulation platform, where vehicle movements in a Unity game engine virtual environment are synchronized with calibrated microsimulations of traffic flow in Simulation of Urban Mobility (SUMO) software. A single roadway worker wearing a VR headset (HTC Vive Pro) can perform simulated versions of roadway construction work in a virtual model of a work zone while traffic vehicles realistically move around it. The realism of these simulations is enhanced when virtual model dimensions match 3D point cloud scans of real-world roadways and work zones. As shown in Figure 2, this platform can enable researchers to collect detailed data on workers' behaviour in realistic urban roadway environments and traffic conditions while safely remaining in a controlled lab setting. Details of the VR-traffic simulation platform's full capabilities can be found in previous publications (Ergan et al., 2022). All VR user studies were performed with an HTC Vive Pro VR system running simulations in Unity 2019 and SteamVR on an Alienware m17 laptop with Intel Core i7-9750H CPU and NVIDIA GeForce RTX 2080 MaxQ GPU. Subsequent sections will outline the specific implementation of the roadway work zone virtual environment and the types of worker behaviour data collected from VR user studies.

3.2 User Studies on Data Collection Platform

Roadway worker behaviour data was collected for this study on people performing a PolyVinyliDene Fluoride (PVDF) weigh-in-motion (WIM) sensor installation on a highway in a virtual 3D replica of an urban highway near New York University's Tandon campus in Brooklyn, New York. The scenario was especially of interest to the research team since it involved shorter term duration roadwork on a highway, where relatively fewer traffic control devices serve as conventional safety protection for workers (Ergan et al., 2020). Screenshots of the virtual roadway work zone and roadwork activities are shown in Figure 3, including pushing a road pavement saw-cutting machine, inserting the actual WIM sensor cable, and distributing grout to embed the sensor into the pavement. New York University's IRB approved human

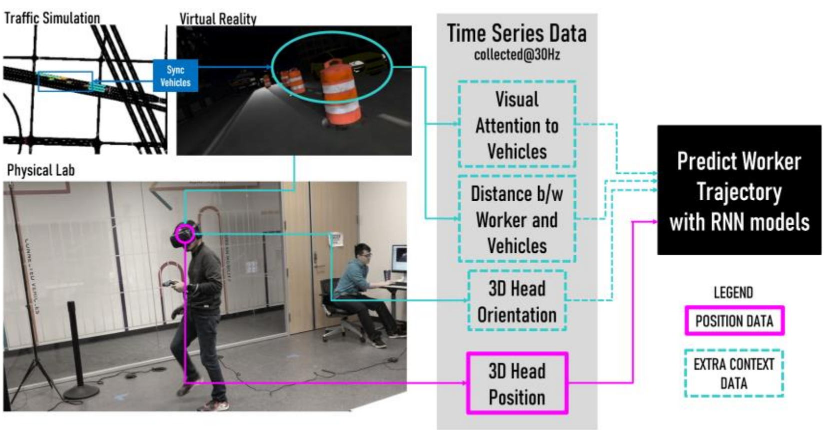

Figure 2: VR-traffic simulation platform for collecting data on roadway workers' safety behaviour for predicting their walking trajectory

Figure 3: Virtual model overview (left) and example user interactive tasks in VR (right) for short-term highway work zone for installing PVDF sensor.

subject experiments on the integrated VR-traffic platform (IRB-FY2020-3946). Twenty participants (n = 20) were recruited for VR user studies involving this scenario. Participant demographics are discussed in the Results section. Once participants were given a tutorial of how to perform the sensor installation tasks in the VR environment, time series data on the roadway worker's safety behaviour (e.g., maintaining distance from vehicles, visual attention to vehicles) was collected in one trial for each participant while night-time highway traffic passed by the work zone. The categories of data collected on each participant's safety behaviour are illustrated in Figure 2 and explained in detail in the subsequent section.

3.3 Categories of Data Collected in VR User Studies

All data was collected directly from the Unity game engine with SteamVR plugin, which essentially remaps a VR headset's physical position and orientation to a virtual camera's position and orientation in the virtual environment. Since VR simulations of the work zone did not use any re-directed walking or teleportation within the virtual environment, all of the participant's physical movements remained within the 2.29 m (7.5 ft) by 2.08 m (6.83 ft) lab space. Since a unit length in VR model units is equivalent to 1 meter (3.28 ft), all 3D locations of the virtual camera in Unity were treated as capturing the worker's real 3D **head position**. Additional data was collected to serve as extra contextual information for worker behaviour models to make trajectory predictions. This included the participant's **head orientation**, which was collected during VR user studies directly from the Unity virtual camera's quaternion rotations, a four-dimensional complex number orientation convention. Context data also included the virtual **distance between the worker and nearest vehicle**, given that the worker's head position and the position of all traffic vehicles are known within the Unity VR model. Using raycasting over virtual camera's entire field of view, the **number of vehicles in worker's field of view** could also be counted by tracking the number of unique traffic vehicles hit by the virtual camera rays. This data combined with the minimum vehicle distance and head orientation results in six extra dimensions for contextual information regarding a roadway worker's safety. All numerical data was stored into a time series at every VR display frame update, at about 30 Hz.

3.4 Data Pre-processing

One user trial's data had to be dropped due to a momentary loss of headset tracking by the VR system. Data collected from the remaining nineteen (n=19) user study trials was then pre-processed for recurrent network model training and evaluation. All time-series data collected from VR user studies (worker's head position, orientation, minimum vehicle distance, number of vehicles seen) were typically recorded at 30 Hz but actual sampling rate varied depending on the VR headset's display framerate and varying number of computations for counting vehicles in the user's field of view. All recorded raw experiment data was then interpolated at a uniform frequency of 10 Hz (i.e., data recorded every 0.1 second), in order for the machine learning training data to resemble a consistent data stream while requiring a shorter training time. After interpolation, the data values in each dimension were normalized between 0 and 1 based on the minimum and maximum value in each column (i.e., column-wise re-scale). The overall time series dataset was then subdivided into pairs of input and target (i.e., ground truth) vector sequences for recurrent network model training and evaluation. Based on previous trajectory prediction studies showing significantly increasing error with longer future prediction time frames, it was decided that the RNN models should try to predict the next one (1) second of a worker's position/trajectory based on the previous three (3) seconds of that worker's observed trajectory (Deo & Trivedi, 2017; D. Kim et al., 2019). In other words, at a 10Hz data sampling rate, a sequence of 30 vectors would serve as RNN model inputs, and model predictions would be trained and evaluated on a sequence of 10 target vectors. Pre-processing the dataset resulted in 46,690 pairs of 3 second observed input and 1 second target data were produced for RNN model training, validation, and testing. Data was then split into respective 80-10-10 percent training, validation, and test subsets based on the VR experiment trial that the input-target pair belonged to. This means that among the 19 trials of recorded VR data, 17 trials' data was used for training, 1 trial was used for validation, and 1 trial's data was used for testing. After this splitting approach, actual numbers of input-target pairs for training, validation, and testing are n = 42938, 2061, and 1691, respectively.

3.5 Recurrent Network Model Implementations

This study evaluates three different RNN architectures using either LSTM or GRU cells: 1) a recursive prediction network that considers only a worker's previous positions, 2) an encoder-decoder network that considers only a worker's previous positions, and 3) an encoder-decoder network that considers worker's previous positions and extra contextual information. All model input vectors have three (3) dimensions for previous positions. The encoder-decoder model with extra context also uses head orientation (4 dimensions), the minimum distance between worker and nearest vehicle (1), and the number of vehicles in the worker's field of view (1) as extra context information included in their input sequence, giving their input vectors a total of nine (9) dimensions. All models considered in this study will only predict a future trajectory of the worker's head position, meaning all target and prediction output vectors' have three dimensions (3). Hyperparameter tuning the RNN models settled on applying the same hyperparameters to all network LSTM/GRU cell modules: a hidden size of 20, 4 stacked LSTM/GRU layers, and a dropout rate of 0.5.

3.6 Recurrent Network Model Training and Evaluation

All six RNN models (i.e., LSTM and GRU cell implementations of the three network architectures) are trained on the mean square error (MSE) loss function, which measures the average squared difference between, $\boldsymbol{\hat{y}_i}'$, each of the models' predicted *normalized* worker positions (i.e., every position in every predicted trajectory), and $\boldsymbol{y_i}'$, the corresponding target *normalized* worker positions (i.e., every position in every ground truth trajectory) from the pre-processed VR user trial data (Equation 1).

$$MSE = \frac{\sum_i^N \left(\boldsymbol{\hat{y}_i}' - \boldsymbol{y_i}'\right)^2}{N} \quad (1)$$

Towards evaluating these prediction models with physically relatable metrics for roadway worker safety, the average displacement error (ADE) is also measured for each predicted trajectory. That is, the actual Euclidean distance (i.e., not normalized) between predicted positions ($\widehat{\boldsymbol{y}}_j$) and ground truth positions ($\boldsymbol{y}_j$) averaged over each 1 second timeframe is calculated as a measure of a model prediction's accuracy (Equation 2).

$$ADE = \frac{\sum_{j}^{n_{future}} \sqrt{\left(\widehat{\boldsymbol{y}}_j - \boldsymbol{y}_j\right)^2}}{n_{future}} \quad (2)$$

In the above equation, $n_{future} = 10$ since the output data encompasses 1 second sampled at a rate of 10Hz. This metric can be expressed in units of physical distance (e.g., meters) to get a real-world sense of how accurate the models are in predicting worker's position in the future, since 1 virtual unit of length in Unity corresponds to 1 meter in the physical world. All RNN models were trained over 1000 epochs with the same training loop parameters: a batch size of 32, learning rate of 0.00005, and Adam optimizer with weight decay regularization. Though all models were trained and validated in the same number of epochs, only the learned model parameters that achieved the lowest validation loss (i.e., minimum validation MSE over all epochs) were used for each model's evaluation on the test dataset. Models were implemented in PyTorch version 1.13.1 and trained on New York University's High Performance Computing Greene cluster using two Intel Xeon Platinum 8268 CPU cores, 8 GB memory, and an NVIDIA RTX 8000 GPU. Training time for each model ranged between 3.72 and 6.7 hours.

4. Results

The final nineteen (n=19) participants' VR user study trial data consisted of affiliated NYU students, ages between 22 and 28. Only a select fraction (n = 4, 21%) of participants had previous construction and roadwork experience, ranging between 2 weeks to 2 years. Participants took between 2.5 to 10 minutes to complete the entire procedure within each virtual simulation data collection trial. While this data collection sample is not representative of the roadway construct workforce, an initial analysis comparing the construction-experienced participants' VR positions during each trial to that of participants without construction experience revealed relatively similar movement patterns, especially since all participants had to perform the same roadwork tasks. All participants felt immersed in the realistic VR traffic environment such that their behaviour working safely around traffic is still worth using for training and evaluating human trajectory prediction models.

To evaluate the accuracy of these different RNN models' predictions, the ADE of every model's 1 second trajectory predictions were measured against the corresponding 1 second ground truth trajectory in the test dataset. Figure 4 shows box plot distributions of each model's one second trajectory prediction ADE. The encoder-decoder networks that used additional worker context data in their input data achieved relatively similar spreads of ADE to the equivalent networks that only used a worker's past position data. Table 1 shows the minimum, maximum and average ADE of each model's one second trajectory predictions. Both LSTM and GRU encoder-decoder networks respectively achieved 55% and 23% lower ADE compared to their recursive counterparts. While the LSTM encoder-decoder network achieved the lowest ADE over all predictions by all models at 0.0394 m (0.129 ft), the GRU encoder-decoder network achieved the lowest average ADE of 0.1885 m (0.618 ft). In particular, the GRU encoder-decoder network with context input data reduced its average ADE by 9.4% compared to the same architecture with only position data inputs, suggesting that contextual information can improve a GRU model's ability to predict workers' future safety behaviour.

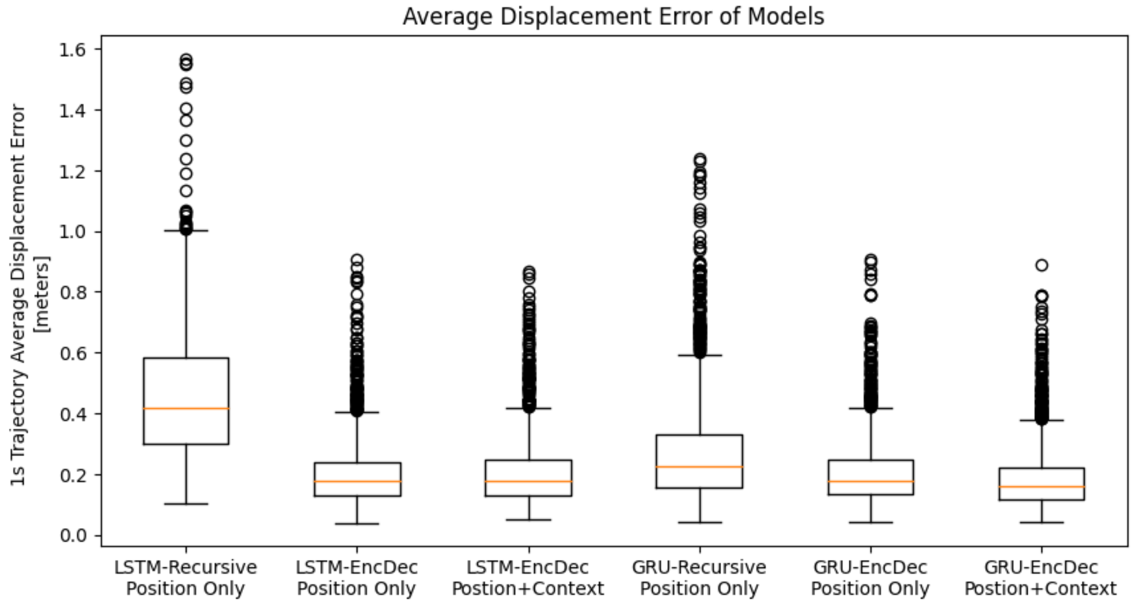

Figure 4: Distribution of average displacement errors of each one second future trajectory prediction by the different models on the test dataset.

Trajectory predictions are visualized side by side for the same input-target trajectory pairs (i.e., same timeframe in the test data experiment). Figures 5 (a) and (b) show "birds eye views" of the worker's head position in the virtual work zone, where each plot has a blue line indicating the three seconds of observed worker trajectory (i.e., model input), a green line indicating how the worker actual moved in the one second after (i.e., ground truth), and different red lines showing different GRU network models' trajectory predictions. Figure 5 (a) shows one of the best predictions (i.e., among the lowest ADE) made by the GRU encoder-decoder network that used context data in its input. While encoder-decoder networks appear to anticipate a change in a worker's walking direction (i.e., worker turns around to walk in a different way), recursive prediction networks predictions are generally straight. The right-most plot of Figure 5 (b) shows the trajectory prediction with the lowest ADE by the GRU encoder-decoder network that used context data in its input. While the GRU encoder-decoder-context data network is able to achieve this low ADE because a few select predicted positions are very close to the ground truth positions, visual inspection reveals that it fails to capture the relatively straight path (i.e., no curvature) of the worker's actual trajectory that other GRU networks are able to match. Further refinements to the GRU models' training based on this curvature could improve their accuracy.

Table 1: Average displacement error of predicted 1s worker trajectory in test dataset

Network Cell	**Network Type**	**Input**	**Min ADE [m]**	**Max ADE [m]**	**Average ADE [m]**
LSTM	Recursive	position only	0.1040	1.5685	0.4595
LSTM	Encoder Decoder	position only	0.0394	0.9071	0.2035
LSTM	Encoder Decoder	position + context	0.0504	0.8675	0.2070
GRU	Recursive	position only	0.0401	1.2384	0.2709
GRU	Encoder Decoder	position only	0.0431	0.9084	0.2079
GRU	Encoder Decoder	position + context	0.0437	0.8882	0.1885

5. Discussion

This study's findings contribute toward models that can anticipate a roadway worker's safety behaviour by utilizing wearable sensor data. Results indicate that extra context data can improve the accuracy of a GRU encoder-decoder network's ability to predict a workers' future movement. A qualitative review of screen capture videos of what participants saw in VR and how they moved their heads during the 4 second timeframes captured in Figure 5 (a) generally revealed that accounting for a worker's head orientation was likely the most beneficial towards more accurate encoder-decoder network predictions. ADE values in Table 1 also indicate that encoder-decoder recurrent architectures are more accurate than their recursive prediction counter parts, and GRU-based networks occasionally perform better than LSTM-based networks. Figure 5 (b) illustrates how the ADE may not be the best metric of a trajectory prediction model's accuracy, since a model can predict a few points close to the ground truth trajectory while not capturing that trajectory's general curvature. Alternatives to the ADE metric that can account for this curvature warrant further investigation. However, in demonstrating accurate predictions of a worker's future positional trajectory, this study's findings are a step towards a more comprehensive high-fidelity model of roadway worker safety behaviour in relation to traffic.

6. Conclusion

This paper compares the accuracy of different RNN models for predicting a roadway worker's future walking trajectory, given data collected on human behaviour during realistic VR simulations of a highway work zone. Most importantly, the inclusion of extra contextual information regarding a roadway worker's safety (e.g., their head orientation, vehicles in their field of view, relative distance to vehicles) was found to potentially contribute to better accuracy in a GRU encoder-decoder models' trajectory predictions, suggesting that sensors in real-world work zones capturing similar context data can contribute to higher fidelity models of roadway workers' safety behaviour. Results also indicate that the GRU encoder-decoder network architecture can outperform LSTM-equivalents and simpler RNN architectures. Overall, the approaches explored in this study for using VR-traffic simulation platforms to collect data for training workers' trajectory models should lead to a better understanding of roadway workers' safety behaviour and ways to improve it.

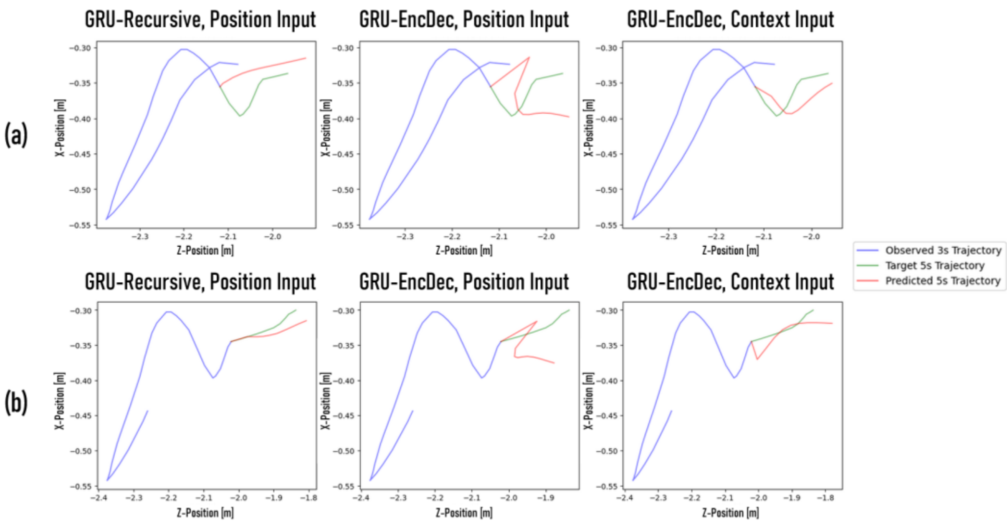

Figure 5: Birds-eye view (Unity X-Z coordinates) of different networks' trajectory predictions for the same 3 second observed worker trajectories in (a) and (b).

References

American Road & Transportation Builders Association (ARTBA), (2023). *Work Zone Data*. https://workzonesafety.org/work-zone-data/, accessed April 2023.

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

Cho, K., van Merrienboer, B., Gulcehre, C., Bahdanau, D., Bougares, F., Schwenk, H., & Bengio, Y. (2014). *Learning Phrase Representations using RNN Encoder-Decoder for Statistical Machine Translation*.

Deo, N., & Trivedi, M. M. (2017). Learning and predicting on-road pedestrian behavior around vehicles. *2017 IEEE 20th International Conference on Intelligent Transportation Systems (ITSC)*, 1–6. https://doi.org/10.1109/ITSC.2017.8317865

Ergan, S., Ozbay, K., Zou, Z., Bernardes, S. D., & Shen, Y. (2020). *Increasing Work Zone Safety: Worker Behavioral Analysis With Integration of Wearable Sensors and Virtual Reality*. https://rosap.ntl.bts.gov/view/dot/58703

Ergan, S., Zou, Z., Bernardes, S. D., Zuo, F., & Ozbay, K. (2022). Developing an integrated platform to enable hardware-in-the-loop for synchronous VR, traffic simulation and sensor interactions. *Advanced Engineering Informatics*, *51*, 101476. https://doi.org/10.1016/j.aei.2021.101476

Gao, Y., González, V. A., Yiu, T. W., Cabrera-Guerrero, G., Li, N., Baghouz, A., & Rahouti, A. (2022). Immersive virtual reality as an empirical research tool: exploring the capability of a machine learning model for predicting construction workers' safety behaviour. *Virtual Reality*, *26*(1), 361–383. https://doi.org/10.1007/s10055-021-00572-9

Kim, D., Liu, M., Lee, S., & Kamat, V. R. (2019, May 24). *Trajectory Prediction of Mobile Construction Resources Toward Pro-active Struck-by Hazard Detection*. https://doi.org/10.22260/ISARC2019/0131

Kim, N., Anderson, B. A., & Ahn, C. R. (2021). Reducing Risk Habituation to Struck-By Hazards in a Road Construction Environment Using Virtual Reality Behavioral Intervention. *Journal of Construction Engineering and Management*, *147*(11). https://doi.org/10.1061/(ASCE)CO.1943-7862.0002187

Lemic, F., Struye, J., & Famaey, J. (2022). Short-Term Trajectory Prediction for Full-Immersive Multiuser Virtual Reality with Redirected Walking. *GLOBECOM 2022 - 2022 IEEE Global Communications Conference*, 6139–6145. https://doi.org/10.1109/GLOBECOM48099.2022.10001349

Luo, X., Li, H., Huang, T., & Rose, T. (2016). A field experiment of workers' responses to proximity warnings of static safety hazards on construction sites. *Safety Science*, *84*, 216–224. https://doi.org/10.1016/j.ssci.2015.12.026

Sak, H., Senior, A., & Beaufays, F. (2014). *Long Short-Term Memory Based Recurrent Neural Network Architectures for Large Vocabulary Speech Recognition*.

Tang, S., Golparvar-fard, M., Naphade, M., & Gopalakrishna, M. M. (2019). Video-Based Activity Forecasting for Construction Safety Monitoring Use Cases. *Computing in Civil Engineering 2019*, 204–210. https://doi.org/10.1061/9780784482445.026

Thapa, D., & Mishra, S. (2021). Using worker's naturalistic response to determine and analyze work zone crashes in the presence of work zone intrusion alert systems. *Accident Analysis & Prevention*, *156*, 106125. https://doi.org/10.1016/j.aap.2021.106125