



HAL
open science

Enhancing JAAD with Knowledge Graphs for Improved Pedestrian Crossing Predictions

Sandra Victor, Gowrishankar Ganesh, Madalina Croitoru

► **To cite this version:**

Sandra Victor, Gowrishankar Ganesh, Madalina Croitoru. Enhancing JAAD with Knowledge Graphs for Improved Pedestrian Crossing Predictions. CONCEPTS 2024 - 1st International Joint Conference on Conceptual Knowledge Structures, Sep 2024, Cadiz, Spain. pp.319-326, 10.1007/978-3-031-67868-4_21 . hal-04695528

HAL Id: hal-04695528

<https://hal.science/hal-04695528v1>

Submitted on 12 Sep 2024

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Knowledge graph proposed extensions for JAAD for better prediction of road crossing decisions by pedestrians

Sandra Victor, Gowrishankar Ganesh, and Madalina Croitoru

LIRMM, CNRS and University of Montpellier, France
{svictor,gowrishankar,croitoru}@lirmm.fr

Abstract. In this paper we place ourselves in the broad context of autonomous driving and, more precisely, in the context of road crossing decisions by pedestrians. Machine learning models built by the existing work need as training data datasets constructed by semi-automatic annotation of video images. JAAD (Joint Attention in Autonomous Driving) is a popular behavior annotated dataset in this regard. Our contribution is to provide a knowledge graph based proposal of additional features that are important but missing in JAAD dataset in the context of road crossing decisions by pedestrians.

Keywords: Pedestrian crossing · JAAD · Knowledge Graphs.

1 Overview

The contribution of this paper is in the field of Artificial Intelligence (AI) powered autonomous vehicles (AV) and investigates the prediction of road crossing decisions by pedestrians. The problem takes as input data about the environment, the pedestrian and the car and outputs whether the pedestrian will or not cross the street. Addressing this problem, and more specifically identifying the factors that influence decisions and behavior of traffic participants, is crucial for developing algorithms that enable AVs to predict and respond to pedestrian movements accurately. To this end, detailed studies of human actions and reactions in traffic scenarios have been undertaken with a focus on behavioral studies and decision making algorithms development. While significant progress has been made in the field of prediction of road crossing decisions by pedestrians for autonomous driving, this task remains a complex challenge that requires further investigation and refinement. Numerous studies have investigated the pedestrian detection [2, 3, 8] and attempted to forecast crossing intentions using advanced AI techniques such as deep learning [1, 4, 5, 16, 9]. However, the unpredictability and variability of human behavior, coupled with the dynamic nature of traffic environments, present persistent hurdles. As shown by the extensive literature actors such as pedestrian intent, environmental cues, and situational context contribute to the complexity of this task [13, 15, 17, 7, 10].

In this setting datasets like JAAD provide conceptualisations for a critical aspect of autonomous driving: the joint attention that must occur between drivers

and of pedestrians, cyclists or other drivers. The JAAD dataset is an annotated dataset aimed at capturing the complexity of interactions between pedestrians and vehicles in urban environments. It contains high-resolution video sequences accompanied by detailed annotations describing pedestrian movements, their crossing intentions, as well as the actions of drivers and other road actors. This dataset is produced with the intention of demonstrating the behavioral variability of traffic participants and provide a set of features to be considered for the pedestrian crossing problem (such as environmental factors, age, etc.). However, the current literature covers a much larger set of features that have been demonstrated by various studies to directly impact pedestrian behavior. Furthermore, the performance of models tested on JAAD often remains limited [14, 12, 11, 6]. This paper aims to bridge this gap in the current state of the art and provide the following two contributions:

- The first knowledge graph in the literature that formalises the current literature on factors impacting the prediction of road crossing decisions by pedestrians.
- A workflow of how the proposed knowledge graph can be used and integrated with platforms like JAAD for further video annotation and pedestrian crossing classification algorithms accuracy improvement.

2 The knowledge graph of road crossing decisions by pedestrians

In this section we will present the knowledge graph built from the literature survey on the features that impact pedestrian behavior. The section will first present the feature extraction process and then the resulting knowledge graph.

2.1 Feature Extraction

The feature extraction was primarily done based on the bibliography. We reviewed 52 different papers to gather as much existing information in this domain as possible. Due to space constraints, only the most relevant articles are cited in this paper. These 52 articles included literature reviews, frameworks, pedestrian recognition and prediction models, studies conducted on real-world data, and others experimented in virtual reality. Although there are literature reviews already compiling various information, the rapid advancement of autonomous driving can render them obsolete and lacking in information. In each of these articles, we identified features influencing pedestrian behavior, and using terms already present in the literature, we constructed the following ontology in Fig.1. Subsequently, we turned our attention to the JAAD database, which provides video sequences of real-world driving with annotations for easier utilization of context information in the prediction problem. JAAD offers 346 videos ranging from 5 to 10 seconds each. To clearly identify the represented factors, the video sequences need to be separated into frames

(82,032 in total), and annotations extracted from the XML files provided by JAAD need to be added. JAAD only provides us with separated videos and annotations. You can view the annotation extraction on a sample video at this link <https://gite.lirmm.fr/svictor/jaad2.0-new-features>. Once this task was completed, we were able to meticulously record the annotated features and add them to the ontology.

2.2 The ontology implementation

The ontology presented here was implemented mainly in Protégé, an open source development environment for ontology creation and knowledge management. It is widely used in the field of artificial intelligence, particularly in semantics and knowledge modeling. With Protégé, users can create, edit and visualize ontologies in different languages, such as RDF, OWL and XML.

As shown in Fig.1, features can be categorized into 3 main different classes. Those centered on the pedestrian’s environment (Fig.1a), those related to other pedestrians in the street (Fig.1c), and those concerning the observed pedestrian (Fig.1b). Within these 3 main classes, we can group together features into smaller classes. For example, the environmental features class encompasses certain features of the physical context of the scene, such as road structure, weather, or the presence of traffic lights, which will differ from features regarding the dynamics of the scene, such as the actual speed of vehicles present and their distances from pedestrians. Furthermore, it is noteworthy that some mixed features can represent both environmental and observed pedestrian-related factors, such as the state of the observed pedestrian.

By adding properties to the classes, such as whether they are present in JAAD or if they have been used for a predictive model, we were able to identify missing features that would be relevant to add to JAAD, as shown in Fig.2. They are described below :

- Distance orthogonal to the curb : Orthogonal distance between pedestrian and curb.
- Head orientation : JAAD indicates whether the pedestrian is looking at the car or not, but there is no quantitative information on the pedestrian’s head orientation in degrees.
- Minimal distance to crosswalk : The minimum distance between the pedestrian and the nearest crosswalk.
- Awareness : The pedestrian can be attentive without being aware of the risk.
- Distance car-crosswalk : The distance between the car approaching and the crosswalk near to the pedestrian.
- Relative pedestrian-car speed : The relative speed between the approaching car and the pedestrian.
- Pedestrian speed : The pedestrian’s speed.
- Distance traveled in previous time : The distance traveled previously indicates whether the pedestrian has been waiting for some time or if they were walking and suddenly stopped.

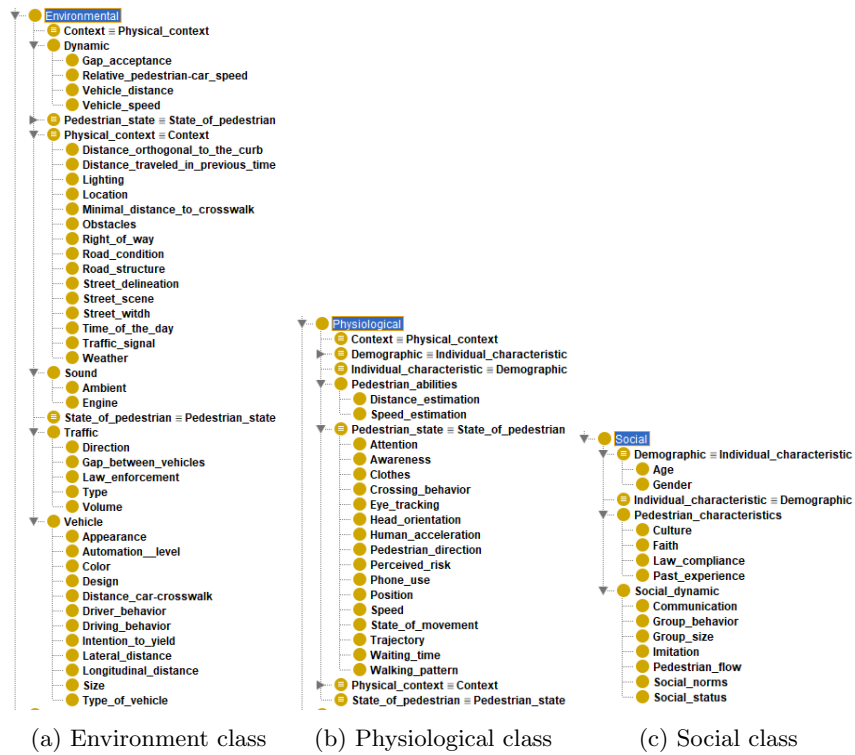


Fig. 1: Road crossing decision by pedestrians knowledge graph

- Longitudinal distance : The distance between the car and the pedestrian.
- Obstacles : The existence of obstacles that could influence the pedestrian’s trajectory.
- Human acceleration : The pedestrian’s acceleration.
- Pedestrian position : The pedestrian’s position.
- Street width : The width of the street.



(a) Annotations in JAAD. (b) Annotations used for prediction.

Fig. 2: Differences between features already annotated in JAAD and those used in the prediction of road crossing decision by pedestrians.

3 JAAD annotation workflow

In this section we provide the workflow based on the proposed knowledge graph (Fig.3), used for the aim of improving the annotation process of JAAD.

A workflow consists of a sequence of tasks that process information, designed to accomplish a specific objective (in our case knowledge graph powered image annotation for JAAD videos). It involves automated processes, systems, or software applications that enable the tasks to be executed in a prescribed order, allowing for the seamless transfer of data and control between different parts of a process. This is crucial in environments where tasks are dependent on the output of different systems or software tools.

RDF knowledge graphs represent information in a structured form as a set of triples: subject, predicate, and object. This facilitates the description of concepts and relationships within a particular domain, making it possible to model complex information in an interoperable format. RDF knowledge graphs can then serve as a unifying model that facilitates data integration and interoperability.

By mapping different data sources to a common semantic model, we can seamlessly operate across diverse data sets. By means of an example we show below how the proposed workflow can dynamically adjust its behavior based on the relationships and attributes defined in the knowledge graph. Let us first describe the workflow in Fig.3.

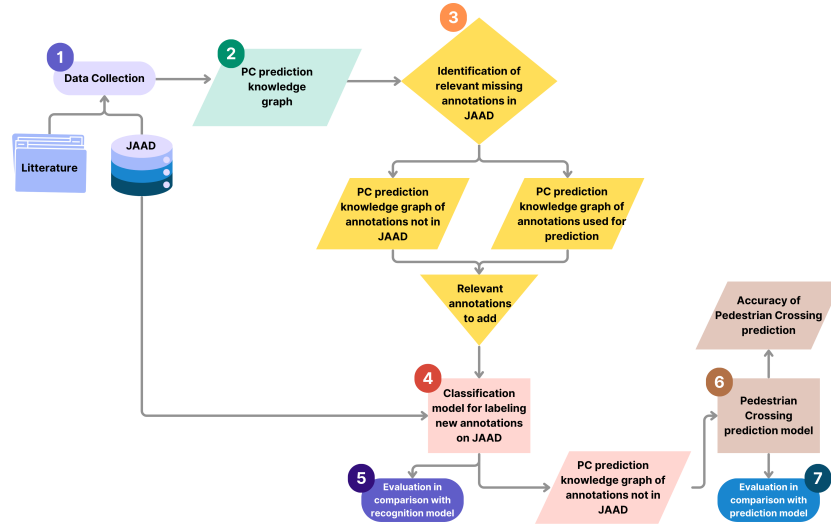


Fig. 3: The knowledge graph based workflow of road crossing decision by pedestrian (PC).

1. We manually extract data (from existing literature and annotations already present in JAAD).
2. We construct an RDF knowledge graph formalising all extracted features from the literature.
3. We identify relevant features to extend the Joint Attention and Action Dataset (JAAD) and explore features never before utilized in predicting Pedestrian Crossing.
4. For the new annotations we run a classification model to identify them within the JAAD database.
5. We compare the classification model with existing models used to recognize other features in JAAD.
6. We input these newly annotated video sequences into our prediction model to assess pedestrian crossing accuracy.
7. Evaluation of this model will involve comparing it with prediction models previously employed by JAAD [12], including one focusing solely on pedestrian actions and another incorporating pedestrian actions with contextual information.

4 Discussion

The main motivation of this paper stems from the pressing need for continued research and development efforts to enhance the accuracy and reliability of pre-

diction of road crossing decision by pedestrians algorithms. Addressing these challenges will be pivotal in ensuring the safety and efficacy of autonomous driving systems in real-world scenarios. To address the challenge of understanding pedestrian decision-making in autonomous driving scenarios, a knowledge rich approach can determine coverage and identify gaps in the literature. For example existing work aims to identify which features are already present in datasets like JAAD or Nuscenes and which ones are missing. Such structured approaches have been shown to guide future data collection efforts in order to ensure that all relevant features are adequately represented [18]. Immediate future work is looking at implementing the annotation workflow. We are currently investigating the use of AWS Step Functions, a cloud service from Amazon Web Services allowing to coordinate the components of distributed applications and microservices using visual workflows. The process is less straightforward than what could be believed at a first glance because of the compatibility of the JAAD data with the AWS environment. Another venue investigated is the simple use of Bash/Shell Scripting. For simpler workflows, bash or shell scripting can automate sequences of commands and basic data manipulation tasks but this is more difficult to integrate with Java environments, making it also difficult to leverage for the use of the heterogeneous and industrial projects (such as AI4CCAM).

5 Acknowledgements

This work is funded by the European Commission through the AI4CCAM project under grant agreement No 101076911.

References

1. Abughalieh, K.M., Alawneh, S.G.: Predicting pedestrian intention to cross the road. *IEEE Access* **8**, 72558–72569 (2020). <https://doi.org/10.1109/ACCESS.2020.2987777>, conference Name: IEEE Access
2. Brouwer, N., Kloeden, H., Stiller, C.: Comparison and evaluation of pedestrian motion models for vehicle safety systems. In: 2016 IEEE 19th International Conference on Intelligent Transportation Systems (ITSC). p. 2207–2212 (Nov 2016). <https://doi.org/10.1109/ITSC.2016.7795912>
3. Fang, Z., López, A.M.: Is the pedestrian going to cross? answering by 2d pose estimation. In: 2018 IEEE Intelligent Vehicles Symposium (IV). p. 1271–1276 (Jun 2018). <https://doi.org/10.1109/IVS.2018.8500413>
4. Fuest, T., Schmidt, E., Bengler, K.: Comparison of methods to evaluate the influence of an automated vehicle’s driving behavior on pedestrians: Wizard of oz, virtual reality, and video. *Information* **11**(6), 291 (Jun 2020). <https://doi.org/10.3390/info11060291>, <https://www.mdpi.com/2078-2489/11/6/291>, number: 6 Publisher: Multidisciplinary Digital Publishing Institute
5. Gulzar, M., Muhammad, Y., Muhammad, N.: A survey on motion prediction of pedestrians and vehicles for autonomous driving. *IEEE Access* **9**, 137957–137969 (2021). <https://doi.org/10.1109/ACCESS.2021.3118224>

6. Kotseruba, I., Rasouli, A., Tsotsos, J.K.: Joint attention in autonomous driving (jaad) (2020)
7. Lévêque, L., Ranchet, M., Deniel, J., Bornard, J.C., Bellet, T.: Where do pedestrians look when crossing? a state of the art of the eye-tracking studies. *IEEE Access* **8**, 164833–164843 (2020). <https://doi.org/10.1109/ACCESS.2020.3021208>, conference Name: IEEE Access
8. Mordan, T., Cord, M., Perez, P., Alahi, A.: Detecting 32 pedestrian attributes for autonomous vehicles. *IEEE Transactions on Intelligent Transportation Systems* **23**(8), 11823–11835 (Aug 2022). <https://doi.org/10.1109/tits.2021.3107587>, <http://dx.doi.org/10.1109/TITS.2021.3107587>
9. Mozaffari, S., Al-Jarrah, O.Y., Dianati, M., Jennings, P., Mouzakitis, A.: Deep learning-based vehicle behavior prediction for autonomous driving applications: A review. *IEEE Transactions on Intelligent Transportation Systems* **23**(1), 33–47 (Jan 2022). <https://doi.org/10.1109/tits.2020.3012034>, <http://dx.doi.org/10.1109/TITS.2020.3012034>
10. Nuñez Velasco, J.P., Farah, H., van Arem, B., Hagenzieker, M.P.: Studying pedestrians’ crossing behavior when interacting with automated vehicles using virtual reality. *Transportation Research Part F: Traffic Psychology and Behaviour* **66**, 1–14 (Oct 2019). <https://doi.org/10.1016/j.trf.2019.08.015>, <https://www.sciencedirect.com/science/article/pii/S136984781930066X>
11. Rasouli, A., Kotseruba, I., Tsotsos, J.K.: Agreeing to cross: How drivers and pedestrians communicate. In: *IEEE Intelligent Vehicles Symposium (IV)*. pp. 264–269 (2017)
12. Rasouli, A., Kotseruba, I., Tsotsos, J.K.: Are they going to cross? a benchmark dataset and baseline for pedestrian crosswalk behavior. In: *Proceedings of the IEEE International Conference on Computer Vision (ICCV) Workshops* (Oct 2017)
13. Rasouli, A., Tsotsos, J.K.: Autonomous vehicles that interact with pedestrians: A survey of theory and practice (arXiv:1805.11773) (May 2018). <https://doi.org/10.48550/arXiv.1805.11773>, <http://arxiv.org/abs/1805.11773>, arXiv:1805.11773 [cs]
14. Razali, H., Mordan, T., Alahi, A.: Pedestrian intention prediction: A convolutional bottom-up multi-task approach. *Transportation Research Part C: Emerging Technologies* **130**, 103259 (Sep 2021). <https://doi.org/10.1016/j.trc.2021.103259>, <https://www.sciencedirect.com/science/article/pii/S0968090X21002710>
15. Ridet, D., Rehder, E., Lauer, M., Stiller, C., Wolf, D.: A literature review on the prediction of pedestrian behavior in urban scenarios. In: *2018 21st International Conference on Intelligent Transportation Systems (ITSC)*. p. 3105–3112 (Nov 2018). <https://doi.org/10.1109/ITSC.2018.8569415>
16. Schulz, A.T., Stiefelhagen, R.: Pedestrian intention recognition using latent-dynamic conditional random fields. In: *2015 IEEE Intelligent Vehicles Symposium, IV 2015, Seoul, South Korea, June 28 - July 1, 2015*. pp. 622–627. IEEE (2015). <https://doi.org/10.1109/IVS.2015.7225754>, <https://doi.org/10.1109/IVS.2015.7225754>
17. Tran, T.T.M., Parker, C., Tomitsch, M.: A review of virtual reality studies on autonomous vehicle–pedestrian interaction. *IEEE Transactions on Human-Machine Systems* **51**(6), 641–652 (Dec 2021). <https://doi.org/10.1109/THMS.2021.3107517>, conference Name: IEEE Transactions on Human-Machine Systems
18. Wickramarachchi, R., Henson, C., Sheth, A.: Knowledge-infused learning for entity prediction in driving scenes. *Frontiers in Big Data* **4** (2021), <https://www.frontiersin.org/articles/10.3389/fdata.2021.759110>