

Article

Pedestrian Crossing Intention Forecasting at Unsignalized Intersections Using Naturalistic Trajectories

Esteban Moreno ^{1,*}, Patrick Denny ² , Enda Ward ³, Jonathan Horgan ³, Ciaran Eising ², Edward Jones ¹, Martin Glavin ¹, Ashkan Parsi ¹, Darragh Mullins ¹ and Brian Deegan ¹¹ Connaught Automotive Research Group (CAR), University of Galway, H91 TK33 Galway, Ireland² Department of Electronic and Computer Engineering, University of Limerick, V94 T9PX Limerick, Ireland³ Valeo, Tuam, H91 Galway, Ireland

* Correspondence: e.moreno1@universityofgalway.ie

Abstract: Interacting with other roads users is a challenge for an autonomous vehicle, particularly in urban areas. Existing vehicle systems behave in a reactive manner, warning the driver or applying the brakes when the pedestrian is already in front of the vehicle. The ability to anticipate a pedestrian's crossing intention ahead of time will result in safer roads and smoother vehicle maneuvers. The problem of crossing intent forecasting at intersections is formulated in this paper as a classification task. A model that predicts pedestrian crossing behaviour at different locations around an urban intersection is proposed. The model not only provides a classification label (e.g., crossing, not-crossing), but a quantitative confidence level (i.e., probability). The training and evaluation are carried out using naturalistic trajectories provided by a publicly available dataset recorded from a drone. Results show that the model is able to predict crossing intention within a 3-s time window.

Keywords: pedestrian; crossing; intention estimation; infrastructure; forecasting; behaviour



Citation: Moreno, E.; Denny, P.; Ward, E.; Horgan, J.; Eising, C.; Jones, E.; Glavin, M.; Parsi, A.; Mullins, D.; Deegan, B. Pedestrian Crossing Intention Forecasting at Unsignalized Intersections Using Naturalistic Trajectories. *Sensors* **2023**, *23*, 2773. <https://doi.org/10.3390/s23052773>

Academic Editor: Felipe Jiménez

Received: 28 January 2023

Revised: 18 February 2023

Accepted: 28 February 2023

Published: 3 March 2023



Copyright: © 2023 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

According to the World Health Organization, approximately 1.25 million people die every year as a result of a road traffic accident, with pedestrians representing 22% of fatalities [1]. Autonomous vehicles have the potential to reduce traffic fatalities by eliminating accidents that occur due to human error. However, urban environments present significant challenges for autonomous vehicles. Non-signalized intersections present a particular challenge where the ego car must interact with other road users who may behave in complex and unpredictable ways.

Accurate prediction of pedestrian movements in an urban environment is an extremely challenging task. Although pedestrian detection and tracking algorithms have an extensive research history [2–4], detecting and tracking pedestrians is just the first step towards human-like reasoning [5–7] with the ability to answer questions such as: is the pedestrian going to intersect with the ego-vehicle's path? Therefore, prediction of pedestrian behaviour is important for the safe introduction of autonomous vehicles into busy urban areas. Pedestrian behaviour has been studied extensively and related studies have been categorized based on the output (i.e., objective) produced by their algorithms [8]: trajectory prediction, crossing intention, behaviour classification (walking, starting, stopping, standing, walking-crossing, running-crossing) [9,10], movement direction [11] and goal prediction. The focus of this work is on forecasting pedestrian crossing intention.

Pedestrian intent is defined as the future action of crossing (C) or not crossing (NC) the road. Anticipating pedestrian intention might result in smoother and safer maneuvers, such as avoiding an unnecessary stop when a pedestrian at the roadside is not showing a crossing intention. Pedestrian intent prediction is especially challenging in the case of pedestrians because they can change their speed and heading unexpectedly. Current designs of connected and autonomous vehicles rely heavily on sensors fitted to the vehicle

(i.e., on-board sensors) [5–7,10,12–24]. This is extensively reflected in the literature and will be discussed in the next section. Many dangerous situations arise from the fact that the driver's view of a pedestrian is restricted [24]. To help compensate for a the limited field of view from on-board sensors, fixed infrastructure sensors may be used. Furthermore, research initiatives are concerned with short-term and long-term predictions [8]. Short-term approaches predict pedestrians' position between 2.5 and 3 s [25,26], while long-term predictions are more goal-oriented [8].

This study proposes a machine learning approach to predict pedestrian crossing intention up to 3 s before the road crossing takes place. A publicly available dataset recorded from a fixed drone is used to evaluate the model on naturalistic trajectories collected at an intersection. To our knowledge this is the first time pedestrian intention is estimated and evaluated from a drone perspective instead of on-board or ground-level infrastructure sensors.

The main contributions of this work are:

- Utilization of random forest (RF) for detailed pedestrian intention forecasting using naturalistic trajectories that were captured using a drone;
- An algorithm to automate the addition of pedestrian crossing intent labels to the existing dataset (see Section 4.1);
- Extensive and detailed validation and evaluation of pedestrian trajectories extracted from real data to show that the proposed model is applicable at different crossing locations, not just a predefined ROI (e.g., crosswalk). Scenarios where the crossing intention is not clear are also investigated (e.g., pedestrian slowing down before crossing);
- Comparison analysis between an RF and a feed-forward neural network (NN) in the context of pedestrian intention forecasting, verifying that the proposed approach outperforms the NN.

This paper is structured as follows. Section 2 reviews the related literature about pedestrian intent forecasting using data collected from on-board sensors or a fixed node. Section 3 describes the problem formulation and introduces the chosen dataset. Section 4 discusses how the data were pre-processed to add new labels that were used during training. It introduces a case study evaluating different naturalistic pedestrian trajectories (i.e., road users were not aware of measurements taking place) in the given intersection. Also, a comparison analysis between an RF and a feed-forward neural network (NN) is conducted. Section 5 concludes the findings and discusses future work.

2. Related Work

A mature pedestrian crossing intention system should consist of detection, tracking and intention recognition components [27]. A considerable body of literature exists on pedestrian detection and tracking [2–4,28–33] and while there is still scope for improvement, pedestrian detection algorithms are well developed thanks to convolutional neural networks (CNN) and common architectures such as YOLO [28] and Faster R-CNN [29]. Tracking algorithms using deep learning [30,31] are starting to reach the performance levels of traditional and state-of-the-art tracking methods [32,33]. Thus, in this section pedestrian crossing intention prediction is examined exclusively.

There is a clear distinction in the literature between predictions made with data collected from an on-board sensor and data collected from a fixed sensor in the infrastructure [34]. Consequently, the following section is partitioned according to this distinction.

2.1. On-Board Sensors

State-of-the-art automated driving systems employ a range of on-board sensors such as cameras, lidar and radar, although the majority of studies on pedestrian intent are based on video sensors normally placed behind the windshield of the vehicle [5–7,10,12–24].

A vision-based pedestrian intention prediction approach from monocular images is proposed in [7]. Using a multi-stage CNN [35], pedestrian pose is analysed for several frames to determine if they are likely to cross the road. The authors report a classifica-

tion accuracy of 0.88. The evaluation is performed on a publicly available naturalistic dataset focused on driver and pedestrian behaviour. Intention estimation from a monocular dashboard camera is also addressed in [18]. A large-scale dataset designed for pedestrian intention estimation is proposed together with a baseline model for intention and trajectory estimation. The ground truth for crossing intention is collected by conducting a large-scale experiment to determine human reference data for this task. Their intention estimation model, based on a convolutional long short-term memory neural network (LSTM), achieves an accuracy of 0.79. In [6] graph convolution is used to model the spatiotemporal relationship between pedestrians and other objects of interest in the scene. A new dataset, designed specifically for autonomous-driving scenarios in areas with dense pedestrian populations, is also introduced. Results show that their spatiotemporal relationship reasoning model can predict pedestrian intention with an accuracy of over 0.8 on their own dataset one second before the crossing commences.

Contextual knowledge using derived features or prior knowledge of the environment [19,20] can improve prediction. A context-based model is introduced in [21] predicting pedestrian crossing behaviours in inner-city locations using a set of twelve carefully chosen features. A supervised learning algorithm is deployed and trained on automatically labelled the pedestrian's crossing intentions. Context is also taken into consideration in [22], modelling pedestrian movement in relation to relevant road edges. In order to determine the relevant road edges, the scene is segmented into four zones comprising: the ego-zone (defined by the ego-vehicle lane), the street-zone (comprising all non-ego-vehicle lanes), the mixed-zone (parking lanes and inlets for buses) and the sidewalk-zone. The reported accuracy using a Support Vector Machine (SVM) is 0.897. Crossing behaviour at signalized intersections is modelled in [23] by incorporating pedestrian physical states and contextual information. The probabilistic model is constructed using a dynamic Bayesian network and is trained with data collected using three monocular cameras mounted on the car. A novel dataset is introduced in [12] to facilitate traffic scene analysis and pedestrian behaviour. The dataset is captured from three cameras positioned inside a car (below the rear-view mirror), providing contextual and behavioural information along with bounding boxes. In order to embed contextual information in the data, each frame is assigned a contextual tag describing the scene (e.g., number of lanes, presence of a zebra crossing, time of day or weather). Results show that classification using only the attention/gait information (i.e., behavioural information) can correctly predict approximately 40% of crossing behaviour observed; however, adding context, such as the width of the street and the presence of the designated crossing, improves the classification by 20%. Two-dimensional human pose features and scene context are investigated in [13] to forecast the crossing intention of pedestrians and other vulnerable road users (i.e., non-motorised road users). A multi-task learning model is proposed to predict actions, crossing intent and trajectories. Finally, pedestrian intent recognition, exploiting various sources of contextual information, is tackled in [14] as a time series modelling and binary classification problem. A novel stacked recurrent neural network (RNN) is proposed, achieving accuracies of 0.844. Moreover, a framework for pedestrian intention estimation is proposed and deployed on a real vehicle in [15]. Results from the field tests showed 0.75 overall accuracy. Additionally, using a Bayesian network the motion and intention of pedestrians near a zebra crossing is estimated in [16] with 91.8% and 84.6% success in motion and intention prediction, respectively, one second into the future.

In more recent work, [17] pedestrian crossing intention is determined through 3D Pose estimation, achieving 0.9128 accuracy. This sets a new standard of performance, although the C/NC labelling is done manually frame by frame. It should also be noted that the test and validation sets have only 7 pedestrians each.

2.2. Infrastructure Sensors

To date, not as much work has been done on crossing intention forecasting from infrastructure sensors (e.g., roadside sensors or cameras installed on aerial vehicles, such

as drones) as there has been with on-board (on vehicle) sensors. Previous studies have focused on the detection and tracking of pedestrians and vehicles [36], pedestrian counting using video surveillance cameras [37], interactions between pedestrians and other agents or path prediction. In [38] interactions between pedestrians and vehicles are studied using video data from two cameras fixed to a lamppost. Their framework, also referred to as the distance-velocity model, is based on vehicle trajectory speed and distance to the crosswalk. In path prediction (i.e., trajectory forecasting), the forecasting of destinations is studied. Although strictly speaking this is not direct pedestrian intention estimation (i.e., binary intention interpretation (C , NC)) it is still classified as pedestrian behaviour prediction. In [26] trajectory prediction is also studied using real-world data collected from a static wide-angle camera and an artificial neural network (ANN). The ANN has two hidden layers and can forecast pedestrian short-time trajectories up to 2.5 s ahead of time for traffic safety applications. Results show that the ANN outperforms other methods such as a constant velocity model Kalman filter (KF) and the extrapolation of fitted polynomials by 21% on average. Using outdoor video recordings and pre-generated cost maps dependent on the environment semantics, an extension to the growing hidden Markov model is proposed in [19] to predict pedestrian positions over a longer time horizon.

Recently, there has been a trend towards formulating the problem of trajectory forecasting as a time series prediction problem using an RNN and its variants [39,40]. In [41] a model where a LSTM is used to predict human movement and forecast future trajectories is proposed. The LSTM method is demonstrated on two publicly available human-trajectory datasets and has been shown to predict motion dynamics in crowded scenes, outperforming methods such as an off-the-shelf KF, or social force model [42]. Based on the same datasets, [43] shows that a simple constant velocity model outperforms some of the most advanced neural network architectures when dealing with linear trajectories.

It is worth mentioning the work by [44,45]. In [44] the authors use quantile regression forests to predict time-to-cross when approaching a crosswalk. Their evaluation is based on vehicle and pedestrian tracks extracted from LIDAR data from a crosswalk in Germany. Because of their aim, prediction of pedestrian crossing behaviour at urban crosswalks in terms of remaining time-to-cross, they only evaluate the trajectories of actual crossing pedestrians. In [45] nine types of pedestrian-vehicle interactions are interpreted by using pose estimation to generate 2 d key points on a pedestrian skeleton and RF. The highest reported accuracy is 0.965 on their own dataset.

Nine studies on crossing intent prediction from a fixed sensor were identified in the literature. Using a database of LIDAR-derived pedestrian trajectories, a large set of features is studied by [46]. The best features are selected to train a binary classifier based on a Support Vector Machine (SVM) achieving an accuracy of 0.9167. In [47], the same authors propose a pedestrian motion prediction method combining pedestrian motion tracking algorithm and a data-driven method, which improves the generalization ability of the model. The pedestrian intention classifier is based on a SVM and mainly serves as a data reduction procedure removing irrelevant pedestrians from the scene, therefore reducing the number of targets to be tracked. For all identified crossing pedestrians, the time-to-cross and crossing point is predicted. Pedestrian intention to cross the road is also investigated by [24] to develop an active pedestrian protection system. Using motion contour histograms of oriented gradients combined with a SVM they estimated the pedestrian crossing intent. Using high-level pre-processed data (e.g., position, velocity, orientation), several ML algorithms are used in [48] to train different classifiers to estimate the intention of a pedestrian to cross at a zebra crossing. Four features are derived from the original dataset and used in the classification task: (i) distance to zebra crossing, (ii) distance between the nearest road border point and the pedestrian position, (iii) distance between the nearest road point and the zebra anchor, measured along the border curve and the angle between a pedestrian's current heading direction, and (iv) pedestrian-to-zebra direction. The k-nearest neighbours is found to be the best performing method for classification, achieving a F2 score (i.e., test accuracy) of 0.9237. Using pre-processed roadside LIDAR

point cloud data, a naïve Bayes based model is presented in [25], predicting the intention to cross between 0.5 and 3 s before the crossing commences. Only 3 features are selected: position, velocity and direction. For evaluation, a test set is collected and divided into 10 C and 10 NC trajectories using a probability threshold of 40%. Results show a high prediction accuracy and higher flexibility than the popular ANN model. LIDAR-based images (i.e., sampled from a point cloud), and different deep learning architectures are used in [49] to estimate pedestrian intention at a crosswalk. A precise map of the environment is utilized, providing information on road geometry and the accurate position of curbs and crosswalks. The best results are achieved using a dense neural network achieving an accuracy greater than or equal to 0.8. The use of LSTM or CNN architectures does not yield the expected improvement although all of them outperform the SVM, which is used as a baseline for performance evaluation. An LSTM is deployed in [50] to predict pedestrians red-light crossing intentions on video data collected at a crosswalk, achieving an accuracy of 0.916. Crossing intention is labelled manually along with other features such as gender, grouping behaviour and walking direction. The model presents a high false-positive rate and according to the authors this problem can be overcome by adding more features related to pedestrian mobility information such as walking speed and acceleration. An extension of the previous work is found in [51] where four machine learning models are used to predict pedestrians' red-light crossing intentions. The best model, an RF, achieves an accuracy of 0.920. An LSTM with attention mechanism (AT-LSTM) is proposed in [27]. Attention mechanism (AT) ensures that greater weight is given to key features. Thanks to the addition of AT and precisely chosen parameter sets using statistical analysis, the model achieves accuracies of 0.9615 and 0.9068, 0 s and 0.6 s prior to crossing, respectively.

In this paper, we propose a model based on an RF that predicts pedestrian crossing intention at different locations around an urban intersection (i.e., crosswalk, pedestrian walking in parallel to the road and a pedestrian walking on other arms of the intersection), not just a predefined ROI. Random forests is a well-established algorithm for statistical learning and has been applied to classification and recently for classification of pedestrian trajectories [48,52,53] and intention [5,51]. We also investigate the addition of a derived categorical feature, which we have called *crossing*, to the original dataset by performing a semantic segmentation of the environment (i.e., partition the image into two segments: sidewalk and road), and evaluate ten crossing and ten non-crossing trajectories. Results show how effective the model is at predicting pedestrian crossing intention when the TTC is above 2 s. Furthermore, we conduct a comparison between an RF and a feed-forward neural network (NN), showing that the former outperforms the latter in terms of accuracy and training times. Finally, we study how the model reacts when the pedestrian demonstrates apparent hesitation (e.g., a significant velocity drop before crossing) and the potential impact of oncoming traffic.

3. Methodology

The problem of crossing intention prediction is formulated as a binary classification task where the objective is to determine if a given pedestrian will cross or not cross the street. The proposed model is able to predict the crossing intention of pedestrians walking at different locations around an unsignalized intersection (e.g., crosswalk, sidewalk or in front of the curb) given a list of selected features and derived features provided by a publicly available dataset and data preprocessing, respectively.

3.1. Random Forest

A random forest [54] is a supervised machine learning algorithm used for classification and regression. As an ensemble learning method, many decision trees are constructed during training using bagging [55] and feature randomness. Each decision tree, consisting of split and leaf nodes, is fitted with random sub-samples of the dataset. Once the RF is trained, the dependent variable is computed by averaging the prediction of all trees in the ensemble (i.e., forest), in the case of regression, or by majority vote of the predicted class

label, for classification. This paper focuses on the latter, the classification task. Random forests can also deal with missing and categorical data [56], are less computationally expensive, are more interpretable and are less prone to overfit than a (NN) [57,58]. Neural networks trained models are difficult to interpret, consisting of hundreds to millions of parameters. This black box problem makes it complicated to use NN in applications where trustworthiness and reliability of the predictions are of great importance. Humans have to understand what models are doing, especially when they are responsible for the consequences of their application [59].

Given a set of input features v , a trained RF classifier will output the class c they belong to and the probability of it being in that class. The predicted class probabilities of an input feature vector are computed as the mean predicted class probabilities of the trees in the forest, denoted by T . The class probability of a single tree is the fraction of samples of the same class in a leaf [60].

$$p(c|v) = \frac{1}{T} \sum_t^T p_t(c|v) \quad (1)$$

For the case of an RF, important hyperparameters include: number of decision trees in the forest and number of features considered by each tree when splitting a node. For the experiments, 500 tree estimators with \sqrt{v} features in each split are used. To measure the quality of the split, entropy is used.

$$Entropy = \sum_{i=1}^{\#Classes} -p_i * \log_2(p_i) \quad (2)$$

3.2. Dataset

This study relies on a publicly available dataset [61] to train and test the proposed model. Pedestrian trajectories are extracted from drone video recordings taken at four German road intersections. The videos were recorded in 4K (4096 × 2160 pixel) resolution at 25 frames per second. Every recording has a duration of approximately 20 min covering areas of 80 × 40 m to 140 × 70 m. All four intersections are unsignalized and differ in shape (e.g., T-junction, four-armed), right-of-way rules, and the number and types of lanes and traffic composition. In this study, the intersection at *Frankenburg* is chosen (see Figure 1) because it contains the largest number of pedestrians, due to its location near the city center. There are 12 recordings (#18–#29) of this intersection with a total duration of ~242.5 min. An image and three CSV files are provided per recording. The image shows the intersection from a drone's viewpoint. Two of the CSV files contain meta data about the road users and the recording itself (e.g., recording number, track id, class, frame rate, duration, location, number of pedestrians) while the third contains the trajectory information. The features provided per frame are: *id*, *frame*, *xPosition*, *yPosition*, *xVelocity*, *yVelocity*, *Heading*, *xAcceleration* and *yAcceleration*. The dataset contains trajectories of vehicles including cars, buses and trucks, as well as VRU such as pedestrians and cyclists; however, in this work pedestrian trajectories are the primary interest.

Pedestrian crossing intention is inferred based on their features (*xPosition*, *yPosition*, *xVelocity*, *yVelocity*, *Heading*) and a derived feature *distanceToRoad*. As discussed in Section 2, derived features and semantic segmentation of the environment can improve prediction. Here, the derived feature *distanceToRoad* is extracted along with *x-yPosition*, *x-yVelocity* and *Heading*, to train an RF. Given a trajectory sample, the RF will output a class label (*C*, *NC*) and its probability. The next section discusses how the data are processed to extract the labels (*C* and *NC*) and presents results in two cases: when the model is applied to the overall dataset and when the model is applied to a particular subset of selected trajectories, belonging to one of the recordings of the chosen intersection.

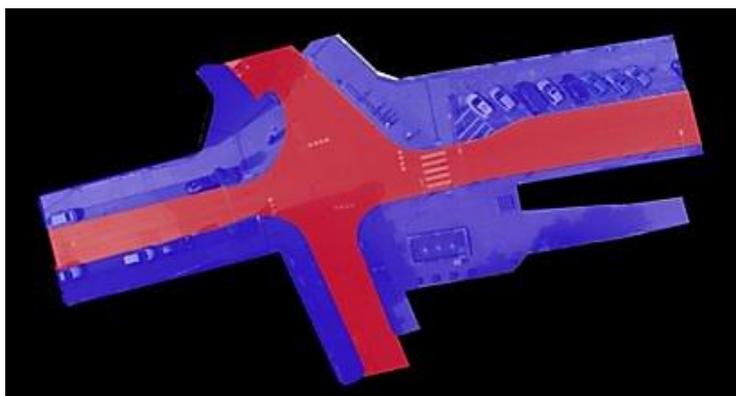


Figure 1. Segmented intersection. Blue represents the sidewalk and red the road.

4. Experiments and Results

4.1. Data Preprocessing

Since the original dataset does not provide labels for pedestrian crossing intention, an automatic labelling strategy (i.e., an algorithm that provides the ground truth at each timestep) was adopted. A similar approach was carried out by [21], although their data were collected from an on-board camera and they presented their results from the ego lane perspective.

In this paper, in order to add the derived categorical feature *crossing* to the original dataset, a basic semantic segmentation of the environment is needed to partition the image into two segments: sidewalk and road. This segmentation was done manually since the dataset did not provide a high-definition map, although it could have been done automatically using data from OpenStreetMaps [62] and a CNN [63] or using roadside LIDAR data [64]. Figure 1 shows the segmentation applied to the chosen intersection. As a result of this segmentation and a derived feature, which we have called *distanceToRoad*, it can be determined if a pedestrian commences crossing the road. The *distanceToRoad* feature can be calculated by casting a vector relative to the pedestrian's *heading* from its current position to the road and measuring the vector length (Figure 2). When a crossing action is detected, each previous timestep (up to 3 s) is annotated as a C behaviour. The rest are considered NC except for the trajectories where the pedestrian is on the road (i.e., being on the road is always considered C) which are ignored and not used during training or testing.

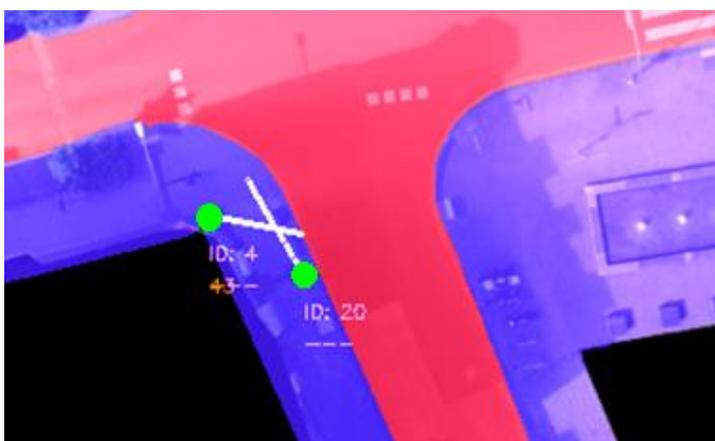


Figure 2. Distance to the road from current pedestrian position. The pedestrian current position is represented by the point shown in green and the heading direction by the white line.

In the dataset containing the derived categorical feature *crossing*, there is a significant imbalance among classes (C, NC). This imbalance can be observed in Figure 3 for recording

#18, where ~7700 C samples and ~98,990 NC samples. The imbalance was addressed by randomly down sampling the majority class [65], NC in this case.

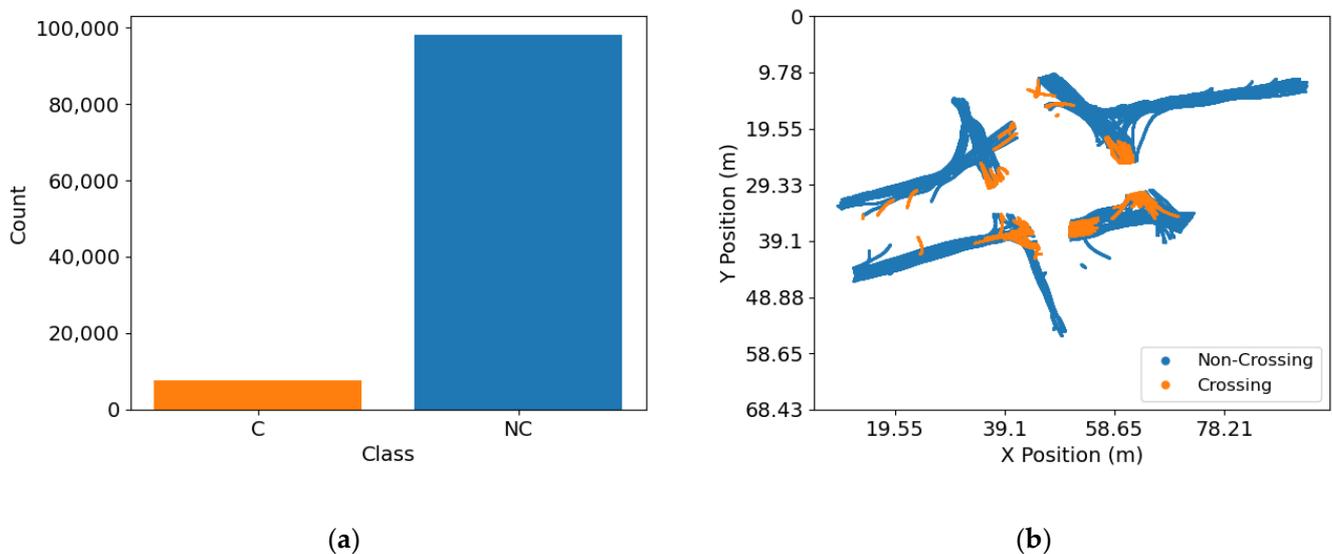


Figure 3. Dataset containing the derived categorical feature *crossing* for recording #18. (a) Imbalance among classes; (b) color-coded trajectories of C and NC trajectory samples.

4.2. Training

Using crossing labels and pedestrian features *xPosition*, *yPosition*, *xVelocity*, *yVelocity*, *Heading* and the derived feature *distanceToRoad*, an RF was trained with the hyperparameters discussed in Section 3. For training and evaluation Scikit-learn was used [66] which is a Python machine learning library. The trajectory data were shuffled and split into training (70%) and test sets (30%) for model fitting and evaluation purposes. We also ensured the test set did not contain any duplicates of trajectory samples found in the training set. Eleven recordings from the intersection located at Frankenburg (#19–#29) were used. The pedestrian trajectory data from the recording that was not used for training (#18) was used to perform further evaluation.

4.3. Evaluation

Since a binary classifier was used, performance can be expressed by using a confusion matrix. This gives the number of correctly and incorrectly classified instances: true positives (TP), true negatives (TN), false positives (FP) and false negatives (FN). The normalized confusion matrix for the test set is illustrated in Figure 4.

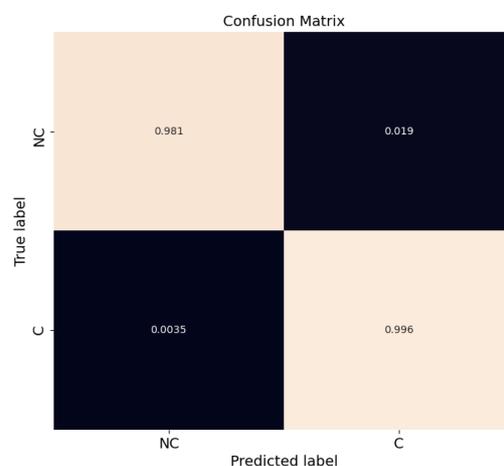


Figure 4. Confusion matrix for RF classification for test set (recordings #19 to #29).

Classification accuracy is also computed according to the widely accepted definition:

$$Accuracy = \frac{(TP + TN)}{(TP + FP + FN + TN)} \quad (3)$$

K -fold cross-validation with $k = 5$ is used, which results in an average accuracy of 0.9883 with a standard deviation of 0.0006.

The confusion matrix in Figure 4 shows that the model is able to classify C and NC trajectory samples successfully, with TP s and TN s above 97%. In terms of FN s, which is an important safety concern, the number is below 1%. The FP s are also low at ~2%.

To further evaluate the model performance, ten crossing and ten non-crossing trajectories, previously unseen by the model during training, are chosen in an experiment similar to that performed in [25]. The trajectories are chosen from recording #18. Figures 5 and 6 show a subset of the selected C and NC trajectories along with the crossing intent prediction and its probability. The possible predicted intention values are 0 or 1, C and NC , respectively.

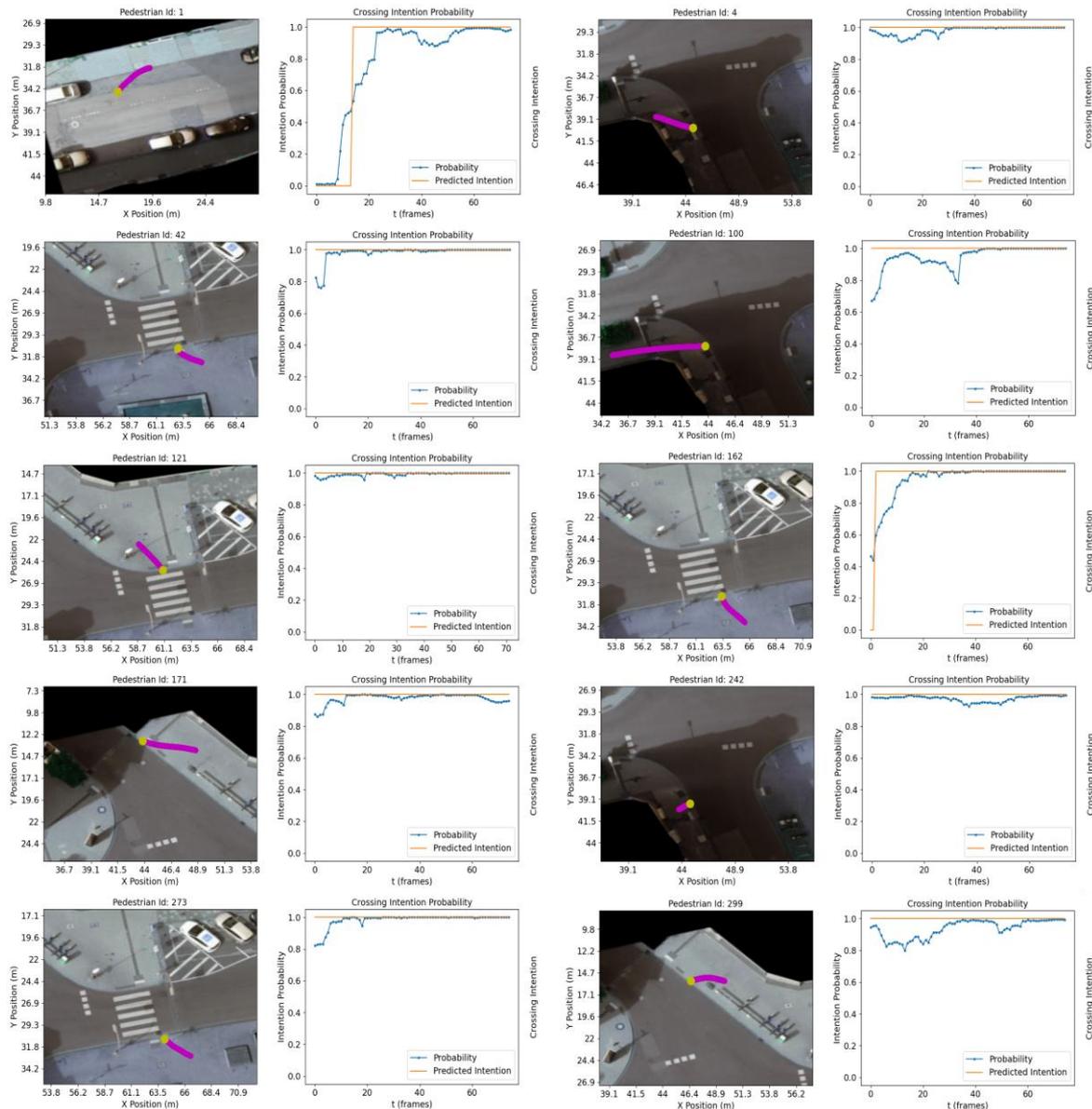


Figure 5. Classification results for 10 crossing trajectories. Pedestrian trajectories are shown in magenta. The trajectory point shown in yellow indicates the last trajectory sample in the sequence.

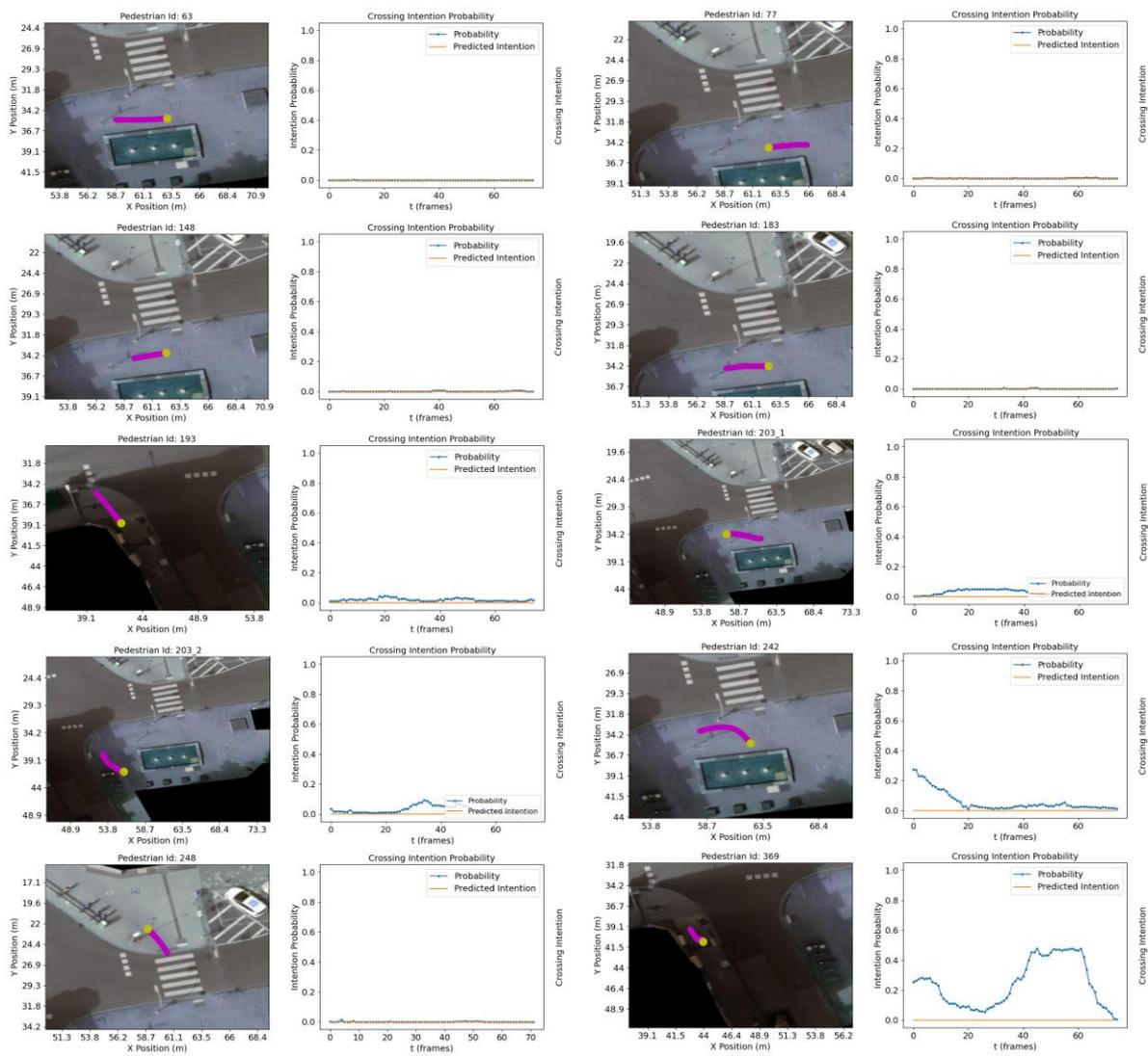


Figure 6. Classification results for 10 not crossing trajectories. Pedestrian trajectories are shown in magenta. The trajectory point shown in yellow indicates the last trajectory sample in the sequence.

Our results show a higher confidence level than [25] when the TTC is longer than 2 s. As expected, the crossing probability increases as the pedestrian approaches the crossing facilities (e.g., zebra crossing or road edge). Prediction accuracy for the ten crossing trajectories is 0.97, with a decision boundary of 0.5. As in [25], intent prediction is computed every 0.5 s. In the evaluation for the ten non-crossing trajectories, it is not possible to replicate the experiment meticulously due to differences in the road structure (e.g., size of the sidewalk when approaching decision making point). Instead, NC trajectories capturing a range of scenarios are presented (e.g., pedestrians turning when approaching any of the crossing areas, pedestrians walking in parallel to the road, pedestrians walking away from the pedestrian crossing or changing direction abruptly).

Figure 5 shows how the model is capable of predicting pedestrian intention for different crossing scenarios in the intersection:

- C Pedestrian at the zebra crossing, ID: 42, 121, 162, 273;
- C Pedestrian at the bottom arm of the intersection, ID: 4, 100, 242;
- C Pedestrians on less utilized (i.e., walked) areas (left and top arms of the intersection), ID: 1, 171, 299.

Conversely, Figure 6 shows the classification results for non-crossing trajectories:

- Pedestrians walking on the sidewalk with NC intention, ID: 193, 369;

- NC while walking in front of the zebra-crossing, ID: 63, 77, 148, 183, 203_1, 242;
- Pedestrian walking away from the crosswalk, ID: 248;
- Pedestrian turning and NC in front of the road, ID: 203_2.

Finally, for the sake of completeness, a confusion matrix of all trajectory samples for intersection recording number 18 is provided in Figure 7. Results show that the model is able to classify C and NC trajectory samples successfully, with TPs and TNs counts of 96% and 95%, respectively. Although there has been an increase in the number of FNs and FPs with respect to Figure 4, their values are below 5%.

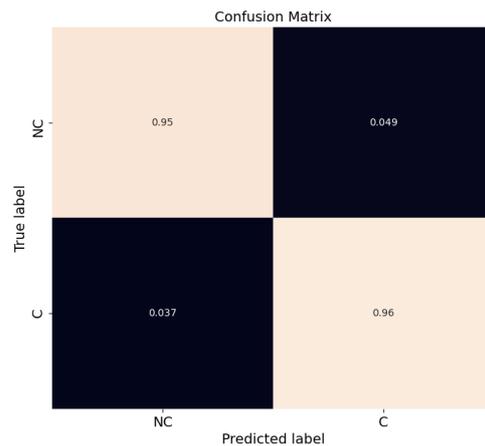


Figure 7. Confusion matrix for RF classification for all samples at recording #18.

To further evaluate the proposed model, we compare the classification accuracy of the RF against a NN. We train a multi-layer perceptron (MLP) [67] which stands for a NN with one or more hidden layers [68], by using the same training and testing sets discussed in section IV (B). In order to tune the hyperparameters of the NN, we use grid search (GS) [69] which let us define a search space grid of hyperparameter values and evaluate every position in the grid to find the best combination that achieves the higher classification accuracy. In our case the (GS) is parameterized for training a MLP classifier using 5-fold cross-validation with the following hyperparameter values:

- *Number of hidden layers* (1, 2, 3)
- *Number of units per layer* (10, 100, 200, 500, 1000)
- *Activation function* (rectified linear unit (ReLU), hyperbolic tangent)
- *Alpha* (0.0001, 0.05) [70]
- *Learning rate* (constant, adaptive) [70]
- *Solver* (stochastic gradient descent, Adam [71])

The best combination of hyperparameters is shown in Table 1 and achieves a mean accuracy of 0.97, which is moderately lower than the accuracy obtained by the RF. Both models are implemented in Python running on an Intel Desktop (Intel Core i7-8700 CPU @3.20 Ghz, 16 GB RAM). The training time of the RF is 6.9 s, which is significantly lower than the 861.2 s required to train the NN. Finally, the RF did not require a grid search, which took 28,150 s (~7.8 h) for a NN on the desktop system.

Table 1. Grid search results.

Hyperparameter	Value
Number of hidden layers	3
Number of units per layer	1000
Activation function	ReLU
Solver	Adam
Alpha	0.05
Learning rate	Constant

4.4. When Pedestrian Intention Is Not Clear

The previous section investigated the crossing intention in cases where the pedestrian maintained a steady walking velocity. This section studies how the model reacts when the pedestrian demonstrates apparent hesitation (e.g., a significant velocity drop before crossing). Figure 8 shows the crossing intention of two pedestrians that show hesitation before crossing the street (i.e., C intention). In order to investigate the potential impact of oncoming traffic, an extra plot displaying both the pedestrian (blue) and approaching vehicle (black) velocities is added.

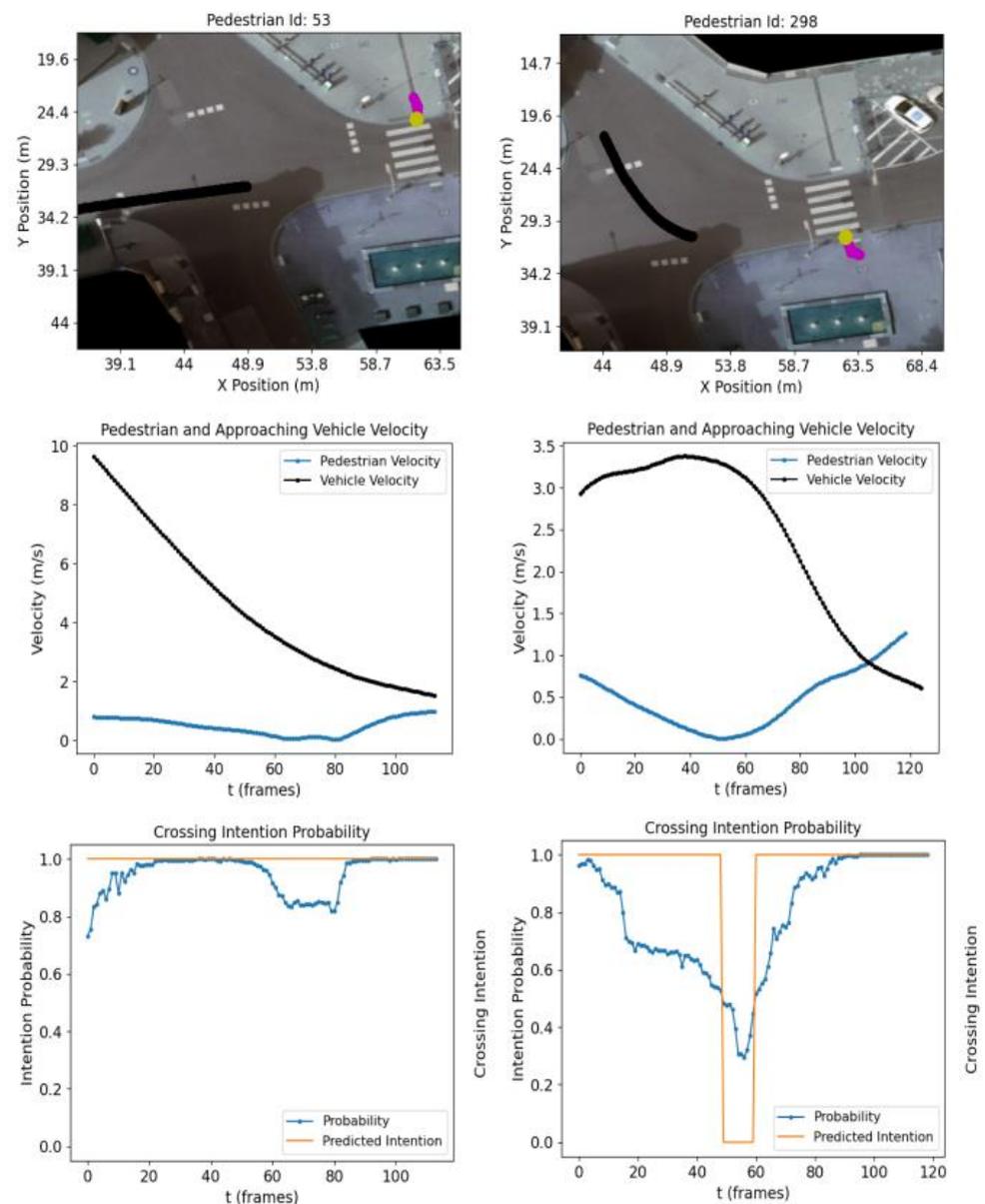


Figure 8. Intent prediction and velocity (plus approaching vehicle velocity) for pedestrians 53 (left) and 298 (right) from recording #18.

Analysing the results in Figure 8 shows that the model reacts with a decrease in crossing intention probability for pedestrians ID 53 and 298 when the pedestrian velocity drops below 0.1 m/s. In the case of pedestrian ID 298, the probability goes below the decision boundary predicting NC for ~12 frames. Other features such as heading, position and distance to the road are used to infer pedestrian crossing intention. Pedestrian ID 298 changes direction and speed suddenly and therefore distance to the road, which contributes to the probability decrease seen in the initial frames (frames 1–20). A similar sudden change

of direction with the consequent probability decrease is also observed for pedestrian ID 369 in Figure 6.

Finally, in an attempt to infer the cause of the significant pedestrian change of velocity before crossing, we further inspect the scenarios for pedestrians ID 53 and 298 by extracting the trajectory and velocity of the closest vehicle that is about to encounter the pedestrian. Pedestrian ID 53 decreases their velocity as the vehicle approaches the zebra crossing at above 9 m/s. The pedestrian velocity drop is lower at the beginning, possibly due to the distance of the vehicle to the zebra crossing and, although the driver is decreasing the vehicle speed, potentially encouraging the pedestrian to cross, we observe a pedestrian velocity drop (below 0.1 m/s) up to frame ~80 when the vehicle speed is below ~3 m/s. In contrast, the vehicle approaching pedestrian ID 298 does not decrease its velocity until frame ~40, forcing the pedestrian to decrease their velocity more aggressively. This supports the importance of studying pedestrian interaction with other agents in the scene. The model could also be improved by adding other contextual information (e.g., extracting further pedestrians features such as body pose or gaze direction).

5. Conclusions and Future Work

This paper has presented a random forest classifier to predict pedestrian crossing intention at a busy inner-city intersection. Using pedestrian features such as position, velocity and heading, provided by a publicly available dataset, and derived features, the model was able to predict crossing and non-crossing intention within a 3-s time window and an accuracy of 0.98. Due to the nature of the dataset used, providing only information related to the trajectory of the tracked pedestrian, detailed features such as body pose or gaze direction that might potentially improve the model performance could not be extracted. In future, this could be improved by building a dataset recorded from a fixed sensor and extracting these features. The authors also conducted a comparison analysis between an RF and a NN showing that the RF outperformed the latter in terms of accuracy and training times. Scenarios where the crossing intention was not clear were also investigated (e.g., pedestrian slowing down before crossing), revealing the potential importance of studying pedestrian interaction with other agents in the scene in future work.

Author Contributions: Conceptualization, E.M., M.G., E.J., B.D., E.W. and J.H.; methodology, E.M.; software, E.M.; validation, E.M.; formal analysis, E.M.; investigation, E.M.; resources, E.M.; data curation, E.M.; writing—original draft preparation, E.M.; writing—review and editing, E.M., M.G., E.J., B.D. and D.M.; visualization, E.M.; supervision, M.G., E.J., B.D., D.M., A.P., P.D., C.E., E.W. and J.H.; project administration, M.G. and E.J.; funding acquisition, M.G. and E.J. All authors have read and agreed to the published version of the manuscript.

Funding: This work was supported with the financial support of the Science Foundation Ireland grant 13/RC/2094_P2 and co-funded under the European Regional Development Fund through the Southern & Eastern Regional Operational Programme, Lero—the Science Foundation Ireland Research Centre for Software (www.lero.ie (accessed on 9 September 2019)).

Conflicts of Interest: The authors declare no conflict of interest.

References

1. World Health Organization. Global status report on road safety. *Inj. Prev.* **2015**. Available online: http://www.who.int/violence_injury_prevention/road_safety_status/2015/en/ (accessed on 20 May 2020).
2. Cao, J.; Pang, Y.; Xie, J.; Khan, F.S.; Shao, L. From Handcrafted to Deep Features for Pedestrian Detection: A Survey. *IEEE Trans. Pattern Anal. Mach. Intell.* **2021**, *44*, 4913–4934. [[CrossRef](#)] [[PubMed](#)]
3. Yilmaz, A.; Javed, O.; Shah, M. Object tracking: A survey. *ACM Comput. Surv.* **2006**, *38*, 13. [[CrossRef](#)]
4. Gawande, U.; Hajari, K.; Golhar, Y. Pedestrian Detection and Tracking in Video Surveillance System: Issues, Comprehensive Review, and Challenges. In *Recent Trends in Computational Intelligence*; Elsevier: Amsterdam, The Netherlands, 2020.
5. Fang, Z.; Vázquez, D.; López, A.M. On-Board Detection of Pedestrian Intentions. *Sensors* **2017**, *17*, 2193. [[CrossRef](#)]
6. Liu, B.; Adeli, E.; Cao, Z.; Lee, K.-H.; Sheno, A.; Gaidon, A.; Niebles, J.C. Spatiotemporal Relationship Reasoning for Pedestrian Intent Prediction. *IEEE Robot. Autom. Lett.* **2020**, *5*, 3485–3492. [[CrossRef](#)]

7. Fang, Z.; Lopez, A.M. Is the Pedestrian going to Cross? Answering by 2D Pose Estimation. In Proceedings of the 2018 IEEE Intelligent Vehicles Symposium (IV), Changshu, China, 26–30 June 2018; pp. 1271–1276. [CrossRef]
8. Ridet, D.; Rehder, E.; Lauer, M.; Stiller, C.; Wolf, D. A Literature Review on the Prediction of Pedestrian Behavior in Urban Scenarios. In Proceedings of the 2018 21st International Conference on Intelligent Transportation Systems (ITSC) 2018, Maui, HI, USA, 4–7 November 2018; pp. 3105–3112. [CrossRef]
9. Kwak, J.-Y.; Ko, B.C.; Nam, J.-Y. Pedestrian intention prediction based on dynamic fuzzy automata for vehicle driving at nighttime. *Infrared Phys. Technol.* **2017**, *81*, 41–51. [CrossRef]
10. Quintero, R.; Parra, I.; Llorca, D.F.; Sotelo, M.A. Pedestrian Intention and Pose Prediction through Dynamical Models and Behaviour Classification. In Proceedings of the 2015 IEEE 18th International Conference on Intelligent Transportation Systems 2015, Gran Canaria, Spain, 15–18 September 2015; pp. 83–88. [CrossRef]
11. Dominguez-Sanchez, A.; Cazorla, M.; Orts-Escolano, S. Pedestrian Movement Direction Recognition Using Convolutional Neural Networks. *IEEE Trans. Intell. Transp. Syst.* **2017**, *18*, 3540–3548. [CrossRef]
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 Workshops, Venice, Italy, 22–29 October 2017; pp. 206–213. [CrossRef]
13. Ranga, A.; Giruzzi, F.; Bhanushali, J.; Wirbel, E.; Pérez, P.; Vu, T.-H.; Perotton, X. VRUNet: Multi-Task Learning Model for Intent Prediction of Vulnerable Road Users. *Electron. Imaging* **2020**, *32*, art00012. [CrossRef]
14. Rasouli, A.; Tsotsos, J.K. Pedestrian Action Anticipation using Contextual Feature Fusion in Stacked RNNs. *arXiv* **2020**, arXiv:2005.06582.
15. Alvarez, W.M.; Moreno, F.M.; Sipele, O.; Smirnov, N.; Olaverri-Monreal, C. Autonomous Driving: Framework for Pedestrian Intention Estimation in a Real World Scenario. In Proceedings of the 2020 IEEE Intelligent Vehicles Symposium (IV), Las Vegas, NV, USA, 19 October–13 November 2020; pp. 39–44. [CrossRef]
16. Škovierová, J.; Vobecký, A.; Uller, M.; Škoviera, R.; Hlaváč, V. Motion Prediction Influence on the Pedestrian Intention Estimation Near a Zebra Crossing. In Proceedings of the VEHITS, Crete, Greece, 16–18 March 2018; pp. 341–348. [CrossRef]
17. Kim, U.-H.; Ka, D.; Yeo, H.; Kim, J.-H. A Real-time Vision Framework for Pedestrian Behavior Recognition and Intention Prediction at Intersections Using 3D Pose Estimation. 2020, pp. 1–12. Available online: <http://arxiv.org/abs/2009.10868> (accessed on 1 February 2021).
18. Rasouli, A.; Kotseruba, I.; Kunic, T.; Tsotsos, J. PIE: A Large-Scale Dataset and Models for Pedestrian Intention Estimation and Trajectory Prediction. In Proceedings of the IEEE/CVF International Conference on Computer Vision 2019, Seoul, Republic of Korea, 27 October–2 November 2019. [CrossRef]
19. Vasishtha, P.; Vaufreydaz, D.; Spalanzani, A. Building Prior Knowledge: A Markov Based Pedestrian Prediction Model Using Urban Environmental Data. In Proceedings of the 2018 15th International Conference on Control, Automation, Robotics and Vision (ICARCV), Singapore, 18–21 November 2018; pp. 247–253. [CrossRef]
20. Lisotto, M.; Coscia, P.; Ballan, L. Social and Scene-Aware Trajectory Prediction in Crowded Spaces. In Proceedings of the IEEE/CVF International Conference on Computer Vision Workshops 2019, Seoul, Republic of Korea, 27–28 October 2019; pp. 2567–2574. [CrossRef]
21. Bonnin, S.; Weisswange, T.H.; Kummert, F.; Schmuelderich, J. Pedestrian crossing prediction using multiple context-based models. In Proceedings of the 17th International IEEE Conference on Intelligent Transportation Systems (ITSC) 2014, Qingdao, China, 8–11 October 2014; pp. 378–385. [CrossRef]
22. Schneemann, F.; Heinemann, P. Context-based detection of pedestrian crossing intention for autonomous driving in urban environments. In Proceedings of the 2016 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS) 2016, Daejeon, Republic of Korea, 9–14 October 2016; pp. 2243–2248. [CrossRef]
23. Hashimoto, Y.; Gu, Y.; Hsu, L.-T.; Iryo-Asano, M.; Kamijo, S. A probabilistic model of pedestrian crossing behavior at signalized intersections for connected vehicles. *Transp. Res. Part C Emerg. Technol.* **2016**, *71*, 164–181. [CrossRef]
24. Kohler, S.; Schreiner, B.; Ronalter, S.; Doll, K.; Brunsmann, U.; Zindler, K. Autonomous evasive maneuvers triggered by infrastructure-based detection of pedestrian intentions. In Proceedings of the 2013 IEEE Intelligent Vehicles Symposium (IV) 2013, Gold Coast, Australia, 23–26 June 2013; pp. 519–526. [CrossRef]
25. Zhao, J.; Li, Y.; Xu, H.; Liu, H. Probabilistic Prediction of Pedestrian Crossing Intention Using Roadside LiDAR Data. *IEEE Access* **2019**, *7*, 93781–93790. [CrossRef]
26. Goldhammer, M.; Doll, K.; Brunsmann, U.; Gensler, A.; Sick, B. Pedestrian’s Trajectory Forecast in Public Traffic with Artificial Neural Networks. In Proceedings of the International Conference on Pattern Recognition, Stockholm, Sweden, 24–28 August 2014; pp. 4110–4115. [CrossRef]
27. Zhang, H.; Liu, Y.; Wang, C.; Fu, R.; Sun, Q.; Li, Z. Research on a Pedestrian Crossing Intention Recognition Model Based on Natural Observation Data. *Sensors* **2020**, *20*, 1776. [CrossRef] [PubMed]
28. Redmon, J.; Divvala, S.; Girshick, R.; Farhadi, A. You Only Look Once: Unified, Real-Time Object Detection. In Proceedings of the 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Las Vegas, NV, USA, 27–30 June 2016; pp. 779–788. [CrossRef]
29. Ren, S.; He, K.; Girshick, R.; Sun, J. Faster R-CNN: Towards real-time object detection with region proposal networks. *Adv. Neural Inf. Process. Syst.* **2017**, *30*, 1137–1149. [CrossRef]

30. Wojke, N.; Bewley, A.; Paulus, D. Simple online and realtime tracking with a deep association metric. In Proceedings of the 2017 IEEE International Conference on Image Processing (ICIP), Beijing, China, 17–20 September 2017; pp. 3645–3649.
31. Held, D.; Thrun, S.; Savarese, S. Learning to track at 100 fps with deep regression networks. In Proceedings of the European Conference on Computer Vision, Amsterdam, The Netherlands, 8–16 October 2016; pp. 749–765.
32. Ghasemi, A.; Kumar, C.R. A Survey of Multi Object Tracking and Detecting Algorithm in Real Scene use in video surveillance systems. *Int. J. Comput. Trends Technol.* **2015**, *29*, 31–39. [[CrossRef](#)]
33. Ciaparrone, G.; Sánchez, F.L.; Tabik, S.; Troiano, L.; Tagliaferri, R.; Herrera, F. Deep learning in video multi-object tracking: A survey. *Neurocomputing* **2020**, *381*, 61–88. [[CrossRef](#)]
34. Jurgen, R.K. (Ed.) *V2V/V2I Communications for Improved Road Safety and Efficiency*; SAE International: Warrendale, PA, USA, 2012.
35. Cao, Z.; Simon, T.; Wei, S.E.; Sheikh, Y. Realtime Multi-Person 2D Pose Estimation Using Part Affinity Fields. In Proceedings of the 30th IEEE Conference on Computer Vision and Pattern Recognition, Honolulu, HI, USA, 21–26 July 2017; pp. 1302–1310. [[CrossRef](#)]
36. Zhao, J.; Xu, H.; Liu, H.; Wu, J.; Zheng, Y.; Wu, D. Detection and tracking of pedestrians and vehicles using roadside LiDAR sensors. *Transp. Res. Part C Emerg. Technol.* **2019**, *100*, 68–87. [[CrossRef](#)]
37. Chen, Y.-Y.; Chen, N.; Zhou, Y.-Y.; Wu, K.-H.; Zhang, W.-W. Pedestrian Detection and Tracking for Counting Applications in Metro Station. *Discret. Dyn. Nat. Soc.* **2014**, *2014*, 712041. [[CrossRef](#)]
38. Fu, T.; Miranda-Moreno, L.; Saunier, N. A novel framework to evaluate pedestrian safety at non-signalized locations. *Accid. Anal. Prev.* **2018**, *111*, 23–33. [[CrossRef](#)]
39. Gupta, A.; Johnson, J.; Fei-Fei, L.; Savarese, S.; Alahi, A. Social GAN: Socially Acceptable Trajectories with Generative Adversarial Networks. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition 2018, Salt Lake City, UT, USA, 18–22 June 2018; pp. 2255–2264. [[CrossRef](#)]
40. Zhang, P.; Ouyang, W.; Zhang, P.; Xue, J.; Zheng, N. SR-LSTM: State Refinement for LSTM Towards Pedestrian Trajectory Prediction. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, Long Beach, CA, USA, 15–20 June 2019; pp. 12077–12086. [[CrossRef](#)]
41. Alahi, A.; Goel, K.; Ramanathan, V.; Robicquet, A.; Fei-Fei, L.; Savarese, S. Social LSTM: Human Trajectory Prediction in Crowded Spaces. In Proceedings of the 29th IEEE Conference on Computer Vision and Pattern Recognition, Las Vegas, NV, USA, 26 June–1 July 2016; pp. 961–971. [[CrossRef](#)]
42. Helbing, D.; Molnár, P. Social force model for pedestrian dynamics. *Phys. Rev. E* **1995**, *51*, 4282–4286. [[CrossRef](#)]
43. Scholler, C.; Aravantinos, V.; Lay, F.; Knoll, A. What the Constant Velocity Model Can Teach Us About Pedestrian Motion Prediction. *IEEE Robot. Autom. Lett.* **2020**, *5*, 1696–1703. [[CrossRef](#)]
44. Volz, B.; Mielenz, H.; Siegwart, R.; Nieto, J. Predicting pedestrian crossing using Quantile Regression forests. In Proceedings of the 2016 IEEE Intelligent Vehicles Symposium (IV) 2016, Gotenburg, Sweden, 19–22 June 2016; pp. 426–432. [[CrossRef](#)]
45. Yang, J.; Gui, A.; Wang, J.; Ma, J. Pedestrian Behavior Interpretation from Pose Estimation. In Proceedings of the 2021 IEEE International Intelligent Transportation Systems Conference (ITSC), Indianapolis, IN, USA, 19–22 September 2021; pp. 3110–3115. [[CrossRef](#)]
46. Volz, B.; Mielenz, H.; Agamennoni, G.; Siegwart, R. Feature Relevance Estimation for Learning Pedestrian Behavior at Crosswalks. In Proceedings of the 2015 IEEE 18th International Conference on Intelligent Transportation Systems 2015, Gran Canaria, Spain, 15–18 September 2015; pp. 854–860. [[CrossRef](#)]
47. Volz, B.; Mielenz, H.; Gilitschenski, I.; Siegwart, R.; Nieto, J. Inferring Pedestrian Motions at Urban Crosswalks. *IEEE Trans. Intell. Transp. Syst.* **2018**, *20*, 544–555. [[CrossRef](#)]
48. Sikudova, E.; Malinowska, K.; Skoviera, R.; Skovierova, J.; Uller, M.; Hlavac, V. Estimating pedestrian intentions from trajectory data. In Proceedings of the 2019 IEEE 15th International Conference on Intelligent Computer Communication and Processing (ICCP) 2019, Cluj-Napoca, Romania, 5–7 September 2019; pp. 19–25. [[CrossRef](#)]
49. Völz, B.; Behrendt, K.; Mielenz, H.; Gilitschenski, I.; Siegwart, R.; Nieto, J. A data-driven approach for pedestrian intention estimation. In Proceedings of the 19th International Conference on Intelligent Transportation System (ITSC), Rio de Janeiro, Brazil, 1–4 November 2016; pp. 2607–2612.
50. Zhang, S.; Abdel-Aty, M.; Yuan, J.; Li, P. Prediction of Pedestrian Crossing Intentions at Intersections Based on Long Short-Term Memory Recurrent Neural Network. *Transp. Res. Rec. J. Transp. Res. Board* **2020**, *2674*, 57–65. [[CrossRef](#)]
51. Zhang, S.; Abdel-Aty, M.; Wu, Y.; Zheng, O. Pedestrian Crossing Intention Prediction at Red-Light Using Pose Estimation. *IEEE Trans. Intell. Transp. Syst.* **2021**, *23*, 2331–2339. [[CrossRef](#)]
52. Papathanasopoulou, V.; Spyropoulou, I.; Perakis, H.; Gikas, V.; Andrikopoulou, E. Classification of pedestrian behavior using real trajectory data. In Proceedings of the 2021 7th International Conference on Models and Technologies for Intelligent Transportation Systems (MT-ITS) 2021, Heraklion, Greece, 16–17 June 2021; pp. 1–6. [[CrossRef](#)]
53. Fuhr, G.; Jung, C.R. Collective Behavior Recognition Using Compact Descriptors. no. 2014. 2018, pp. 1–8. Available online: <http://arxiv.org/abs/1809.10499> (accessed on 12 September 2022).
54. Breiman, L. Random forests. *Mach. Learn.* **2001**, *45*, 5–32. [[CrossRef](#)]
55. Hastie, J.; Tibshirani, T.; Friedman, R. *The Elements of Statistical Learning*, 2nd ed.; Springer: Berlin/Heidelberg, Germany, 2017.
56. Tang, F.; Ishwaran, H. Random forest missing data algorithms. *Stat. Anal. Data Min. ASA Data Sci. J.* **2017**, *10*, 363–377. [[CrossRef](#)]

57. Wang, S.; Aggarwal, C.; Liu, H. Random-Forest-Inspired Neural Networks. *ACM Trans. Intell. Syst. Technol.* **2018**, *9*, 1–25. [[CrossRef](#)]
58. Fern, M.; Cernadas, E. Do we Need Hundreds of Classifiers to Solve Real World Classification Problems? *J. Mach. Learn. Res.* **2014**, *15*, 3133–3181.
59. Roßbach, P. Neural Networks vs. Random Forests—Does it always have to be Deep Learning? *Frankfurt Sch. Financ. Manag.* **2018**. Available online: <https://blog.frankfurt-school.de/wp-content/uploads/2018/10/Neural-Networks-vs-Random-Forests.pdf> (accessed on 15 September 2022).
60. Pedregosa, F.; Varoquaux, G.; Gramfort, A.; Michel, V.; Thirion, B.; Grisel, O.; Blondel, M.; Prettenhofer, P.; Weiss, R.; Dubourg, V.; et al. RandomForestClassifier in Scikit-Learn. 2011. Available online: <https://scikit-learn.org/stable/modules/generated/sklearn.ensemble.RandomForestClassifier.html> (accessed on 27 February 2023).
61. Bock, J.; Krajewski, R.; Moers, T.; Runde, S.; Vater, L.; Eckstein, L. The inD Dataset: A Drone Dataset of Naturalistic Road User Trajectories at German Intersections. In Proceedings of the 2020 IEEE Intelligent Vehicles Symposium (IV) 2020, Las Vegas, Nevada, USA, 19 October–13 November 2020; pp. 1929–1934. [[CrossRef](#)]
62. Curran, K.; Crumlish, J.; Fisher, G. OpenStreetMap. *Int. J. Interact. Commun. Syst. Technol.* **2012**, *2*, 69–78. [[CrossRef](#)]
63. Kaiser, P.; Wegner, J.D.; Lucchi, A.; Jaggi, M.; Hofmann, T.; Schindler, K. Learning Aerial Image Segmentation From Online Maps. *IEEE Trans. Geosci. Remote. Sens.* **2017**, *55*, 6054–6068. [[CrossRef](#)]
64. Wu, J.; Xu, H.; Zheng, J. Automatic background filtering and lane identification with roadside LiDAR data. In Proceedings of the 2017 IEEE 20th International Conference on Intelligent Transportation Systems (ITSC), Yokohama, Japan, 16–19 October 2017; pp. 1–6. [[CrossRef](#)]
65. More, A.S.; Rana, D.P. Review of random forest classification techniques to resolve data imbalance. In Proceedings of the 2017 1st International Conference on Intelligent Systems and Information Management (ICISIM) 2017, Aurangabad, India, 5–6 October 2017; pp. 72–78. [[CrossRef](#)]
66. Pedregosa, F.; Varoquaux, G.; Gramfort, A.; Michel, V.; Thirion, B.; Grisel, O.; Blondel, M.; Prettenhofer, P.; Weiss, R.; Dubourg, V.; et al. Scikit-learn: Machine learning in Python. *J. Mach. Learn. Res.* **2011**, *12*, 2825–2830.
67. Theodoridis, S.; Koutroumbas, K. *Pattern Recognition*, 4th ed.; Academic Press: San Diego, CA, USA, 2009.
68. Haykin, S. *Neural Networks and Learning Machines*; Pearson: London, UK, 2008.
69. Liashchynskiy, P.; Liashchynskiy, P. Grid Search, Random Search, Genetic Algorithm: A Big Comparison for NAS. No. 2017. 2019, pp. 1–11. Available online: <http://arxiv.org/abs/1912.06059> (accessed on 7 October 2022).
70. Pedregosa, F.; Varoquaux, G.; Gramfort, A.; Michel, V.; Thirion, B.; Grisel, O.; Blondel, M.; Prettenhofer, P.; Weiss, R.; Dubourg, V.; et al. GridSearch in Scikit-Learn. 2011. Available online: https://scikit-learn.org/stable/modules/generated/sklearn.model_selection.GridSearchCV.html (accessed on 7 October 2022).
71. Kingma, D.P.; Ba, J.L. Adam: A method for stochastic optimization. In Proceedings of the 3rd International Conference on Learning Representations, ICLR 2015—Conference Track Proceedings, San Diego, CA, USA, 7–9 May 2015.

Disclaimer/Publisher’s Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.