Towards the detection of driver–pedestrian eye contact

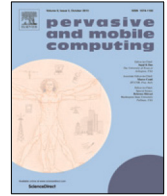Onkhar, V.; Bazilinskyy, P.; Stapel, J. C.J.; Dodou, D.; Gavrila, D.; de Winter, J. C.F.

Important note
To cite this publication, please use the final published version (if applicable).
Please check the document version above.

Towards the detection of driver–pedestrian eye contact

V. Onkhar, P. Bazilinskyy, J.C.J. Stapel, D. Dodou, D. Gavrila, J.C.F. de Winter *

Faculty of Mechanical, Maritime and Materials Engineering, Delft University of Technology, Delft, The Netherlands

A B S T R A C T

Non-verbal communication, such as eye contact between drivers and pedestrians, has been regarded as one way to reduce accident risk. So far, studies have assumed rather than objectively measured the occurrence of eye contact. We address this research gap by developing an eye contact detection method and testing it in an indoor experiment with scripted driver–pedestrian interactions at a pedestrian crossing. Thirty participants acted as a pedestrian either standing on an imaginary curb or crossing an imaginary one-lane road in front of a stationary vehicle with an experimenter in the driver's seat. In half of the trials, pedestrians were instructed to make eye contact with the driver; in the other half, they were prohibited from doing so. Both parties' gaze was recorded using eye trackers. An in-vehicle stereo camera recorded the car's point of view, a head-mounted camera recorded the pedestrian's point of view, and the location of the driver's and pedestrian's eyes was estimated using image recognition. We demonstrate that eye contact can be detected by measuring the angles between the vector joining the estimated location of the driver's and pedestrian's eyes, and the pedestrian's and driver's instantaneous gaze directions, respectively, and identifying whether these angles fall below a threshold of 4°. We achieved 100% correct classification of the trials involving eye contact and those without eye contact, based on measured eye contact duration. The proposed eye contact detection method may be useful for future research into eye contact.

1. Introduction

In 2018, there were over 300,000 pedestrian deaths worldwide [1]. Studies have shown that pedestrian fatalities are growing by the year, especially on urban roads [2], and that most pedestrian casualties occur during the act of street crossing [3]. An area of study with relevance to pedestrian safety is how pedestrians interact with approaching vehicles. Next to formal traffic rules, non-verbal communication plays a role in the safe interaction between pedestrians and drivers [4,5].

1.1. The effect of eye contact on pedestrians

Through interviews and on-site observations [6,7] and recordings of natural driving scenes [7–9], it has been shown that a sizeable percentage of pedestrians use eye contact to negotiate right of way when crossing the road. Additionally, studies have investigated pedestrians' responses to automated vehicles without a driver making eye contact (typically

---

using a Wizard of Oz approach [5,10,11]. In particular, Malmsten Lundgren et al. found that most pedestrians were willing to cross the road when there was eye contact with the driver, whereas only a few were willing when the driver of the automated vehicle was inattentive. There also exists a prevalent belief outside academia that eye contact is of significance to the safety of pedestrians, as evidenced by notices, signs, and advice issued by traffic safety organizations [12–14]. The view that eye contact is important has even led to the development of anthropomorphic external human–machine interfaces (eHMIs) for automated vehicles. Chang et al. [15], for example, tested a novel eHMI with dynamic eyes on the car.

While many studies have shown that pedestrians seek eye contact with drivers [6,7,16], it has been suggested that eye contact is not essential and that pedestrians often cross in front of vehicles by solely relying on vehicle motion cues [6,17,18]. In an online study by AlAdawy et al. [19], participants looking at photos of a car with a driver inside at different distances and under different lighting conditions reported that, in many situations, they could not even see the driver, let alone make eye contact.

1.2. The effect of eye contact on drivers

Next to pedestrians' communication needs at crossings, studies have investigated the effect of pedestrians' communication attempts, including eye contact, on drivers. As early as 1974 [20], Snyder, Grather, and Keller noted in a field experiment on hitchhiking that drivers yielded more often when staged hitchhikers sought eye contact with them. Katz et al. [21] found that drivers slowed down and yielded more often to pedestrians when the pedestrians initiated crossing but were *not* looking in the driver's direction, compared to when they were. More recently, in a field study measuring car speed profiles as a function of eye contact, Ren et al. [22] found that drivers braked earlier for staged pedestrians who attempted to make eye contact than for those who did not. That said, Schmidt and Färber [9] found that participants looking at videos of traffic scenes from a driver's perspective were able to make accurate predictions of pedestrians' crossing intentions even when the pedestrians' heads were occluded, suggesting that eye contact is not essential in traffic.

1.3. Literature gap

From the above, there appears to be a need for further research into the importance of eye contact in traffic. However, as of present, no general conclusions can be obtained due to the variety of measurement methods employed. Measurements such as head orientation, as reported by experimenters standing on the roadside or recorded via cameras inside or outside of the vehicle, can be used to infer eye-contact seeking [7,8,16,23,24]. For example, based on an analysis of video clips from urban driving scenarios, Rasouli et al. [8] suggested that pedestrians looking in the direction of approaching vehicles for longer than 1 s might also seek eye contact. However, head orientation alone does not determine where road users are looking.

Herein, we propose the use of eye tracking to detect eye contact. Eye tracking can establish where road users look without explicitly asking them and without relying on third-party observations. Some research into how drivers look at pedestrians already exists. For example, Walker [25] and Walker and Brosnan [26] reported that drivers gazed at cyclists' faces first and for longer than other body parts. Nathanael et al. [27] used eye tracking to analyze drivers' gaze and concluded that pedestrians' body movement/posture and eye gaze were sufficient to resolve crossing conflicts in the majority of interactions, without the need for eye contact or hand gestures. Diederichs et al. [28] performed eye tracking in a driving simulator study with simulated pedestrians and reported that drivers' pedal responses that indicated an intention to brake were accompanied by eye fixations on pedestrians 0.4–2.4 s earlier. Finally, Borowsky et al. [29] conducted an eye-tracking experiment where participants viewed traffic videos from a driver's perspective and pressed a button when they perceived a hazard. The authors reported that, in general, drivers fixated more often on pedestrians on the road compared to those on the curb.

Research on eye movements in pedestrians exists as well. For example, De Winter et al. [30] conducted an eye-tracking study of pedestrians during interactions with vehicles in a parking lot and found that pedestrians frequently sought eye contact with drivers. Dey et al. [31] used eye tracking of pedestrians at a curb and measured their willingness to cross in front of an oncoming vehicle using a handheld slider. They reported that despite the interior of approaching cars being dark and reflections on the windshield making it difficult to establish eye contact, pedestrians still seek information about the driver's intentions by looking at the windshield when the vehicle was nearby. A recent literature review on eye-tracking studies of pedestrians crossing the road noted that a limitation of the research so far is that the eye-tracking results were not combined with physical measurements such as the distance of the vehicle to the pedestrian [32]. Image recognition combined with eye-tracking could prove a solution to this limitation.

Another important caveat in the above studies involving eye tracking in traffic (e.g., [27,30,31]) is that only one perspective (either the driver's or the pedestrian's) was measured, which provides an incomplete picture because eye contact is a mutual phenomenon (and see [24] for a study on mutual situation awareness of driver and pedestrian; also [33,34], for studies using dual eye tracking in social interaction). In other words, gaze detection of solely one of the two parties is informative about the party's *seeking* of eye contact but cannot tell whether eye contact has been established. These problems in the operationalization of eye contact in the literature have also been reported by Jongerius et al. [35] in their scoping review on eye contact in human–human interaction. This vacuum in the literature could be filled by techniques that detect driver–pedestrian eye contact.

1.4. Study aims

In the current study, we developed a method that detects driver–pedestrian eye contact by means of two eye trackers along with two cameras. Our work's novelties are the use of dual eye tracking in the traffic context, which pinpoints where both the driver and the pedestrian are looking at any given time, and the use of image recognition on video recordings from cameras to estimate the locations of the driver and the pedestrian. We validated the method by means of an indoor experiment with scripted driver–pedestrian interactions at a pedestrian crossing.

Driver–pedestrian eye contact was operationalized as a situation when the driver and the pedestrian are looking at each other at the same time within predefined threshold angles. In the literature, there is an emphasis on the psychological (and hence, subjective) experience of eye contact [35,36]. In this study, we are concerned with the objective detection of eye contact. An underlying assumption in our operationalization is that if two persons are looking at each other's faces, they are looking at each other's eyes in an attempt to make eye contact.

2. Methods

2.1. Participants

Thirty-one persons (23 males, 8 females) took part in the experiment as staged pedestrians. Participants were recruited via social media and personal contacts. Only people with normal visual acuity or corrected with contact lenses were eligible to participate. All participants provided written informed consent. The research was approved by the Human Research Ethics Committee of the Delft University of Technology (reference number 865). One male participant was excluded because of a failure of one eye tracker, resulting in a final sample of 30, with a mean age (*SD*) of 24.8 (2.3) years and with ages ranging from 19 to 31 years.

2.2. Equipment

A head-mounted Tobii Pro Glasses 2 eye tracker was used to track and record the pedestrian's (i.e., participant's) gaze direction at 50 Hz. The 'Gaze Spot Meter' setting of the Tobii was turned on, as a result of which the exposure of the camera images was automatically adjusted based on where the participant looked. A head-mounted camera built into the Tobii recorded the pedestrian's view as a video at 25 frames per second, a field of view of 90°, and a resolution of 1920 × 1080. Pedestrian gaze calibration was achieved using a card with a printed bull's-eye, and parallax error was corrected automatically by the manufacturer's software.

A dashboard-mounted Smart Eye Pro dx eye tracker installed in a Toyota Prius intelligent vehicle (i.e., with environment-sensing capabilities [37]) was used to track and record the driver's (i.e., experimenter's) gaze direction at 60 Hz. The Smart Eye collected gaze data using a combination of the manufacturer's bundled software and custom C++ programs running inside the Robot Operating System (ROS) on an Ubuntu Linux computer onboard the Toyota Prius. Parallax error was avoided as the Smart Eye was not head-mounted and worked with 3D space rather than 2D projections on an image plane.

Both eye trackers work on the principle of pupil corneal reflection, i.e., by using the angle between the locations of the pupil and reflections of infrared light on a person's cornea to determine their gaze direction [38]. To this end, artificial infrared light sources are employed, and the corneal reflections are captured by infrared cameras. The pedestrian's and the driver's gazes were both taken as the average of the gaze directions from each of their two eyes. An iDS UI-3060CP-C-HQ Rev. 2 stereo camera installed in the car with integrated pedestrian detection recorded the pedestrian's locations using the single-shot detection (SSD [39]) technique at 10 Hz.

The Smart Eye was temporally synchronized with the stereo camera using Network Time Protocol (NTP) clients on a local area network (LAN) with a synchronization buffer, which minimized the difference in capture time between the two sources. The Smart Eye was further spatially calibrated relative to 22 reference points inside the vehicle, the locations of which were known (i.e., the position and orientation of the vehicle's sensors, including that of the stereo camera) obtained by laser scanning the Toyota Prius (methods described by [40]). Driver gaze was calibrated using a different set of 7 known points in the vehicle's interior. Flashes from a NexTorch flashlight (300 lumens) during the experiment were captured by both the Tobii camera and the stereo camera to enable retrospective synchronization of the data from the two eye trackers.

2.3. Experimental procedure

The experiment was conducted indoors in an open lab space, with the area cordoned off to prevent interference from passers-by. The location was well lit by a combination of natural light (indirect and diffused through windows) and fluorescent tubes on the ceiling. The Toyota Prius was parked away from sunlight to reduce windshield glare and angled in such a way as to minimize the chance of unwanted detection of onlookers and passers-by by the stereo camera.

Participants played the role of a pedestrian, and two experimenters conducted the study: the first posed as the driver of the stationary car, and the second instructed the participant about their task and controlled the torch for the

Table 1
Trials in the experiment.

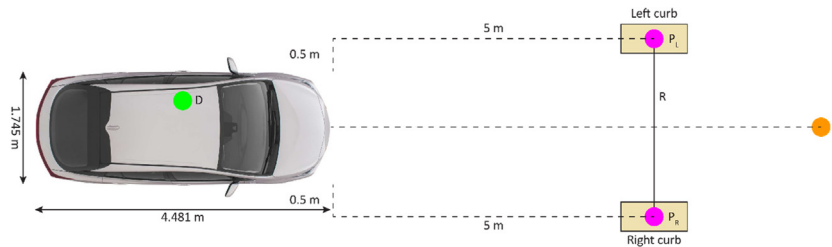| Pedestrian action/position	Eye contact	Abbreviation
Left standing pedestrian	Yes	L-S-EC
Left standing pedestrian	No	L-S-NEC
Right standing pedestrian	Yes	R-S-EC
Right standing pedestrian	No	R-S-NEC
Crossing pedestrian	Yes	C-EC
Crossing pedestrian	No	C-NEC

Fig. 1. Layout of the experiment. The green circle marked 'D' denotes the location of the driver, the orange circle marked 'S' denotes the location of the 'synchronizer' (i.e., the experimenter with the torch), and the magenta circles marked 'P$_L$' and 'P$_R$' denote the locations of the pedestrian on the left and right curbs, respectively. In the experiment, the curbs were rectangles (shown here in light brown), and the road crossing (marked 'R') was a line, all three drawn in chalk on the floor.

synchronization of the two eye trackers. Upon arrival, participants read and signed the consent form, which also contained a brief overview of the experiment, its objectives, and their role in it. They subsequently completed a questionnaire on their height, age, sex, nationality, etc. Next, participants were provided with an oral explanation of how to interact with the car (to supplement what they had read). They were asked to wear the Tobii, which was calibrated using a card with a printed bull's-eye. In the meantime, the driver calibrated his gaze with the Smart Eye using a checkerboard. Participants were instructed to imagine that they were a pedestrian on a curb who had the intention of crossing a one-lane road at an uncontrolled crossing when an approaching car had just slowed down to a stop.

Each participant performed six trials, as summarized in Table 1. The trials involved either a standing pedestrian (on the right or left imaginary curb) or a crossing pedestrian, with instructions to make/not make eye contact with the driver. Three repetitions were conducted per trial. Thus, each participant performed 18 repetitions in total (6 trials x 3 repetitions). The six trials were performed in two blocks: one containing all four standing trials and the other containing the two crossing trials. The order in which the blocks were performed and the order of the trials within the blocks were randomized. The repetitions within each trial were conducted back-to-back.

Before each trial, participants were instructed by the second experimenter about what type of trial would be performed next (i.e., whether/where to stand or cross, and whether to make/not make contact with the driver). In the standing trials, the pedestrian stood on the imaginary curb at a longitudinal distance of 5 m from the front of the car and at a lateral distance of 0.5 m from the side of the car, either to its left or right (see Fig. 1). In the crossing trials, the pedestrian always started on the right curb.

In half of the standing trials, the pedestrians were asked to turn their head to briefly make eye contact with the driver, whereas in the other half, the pedestrians were asked to turn their head to briefly look at the body of the car (at a location of their preference) but refrain from gazing at the driver. In the crossing trials, the pedestrians were instructed to either maintain eye contact with the driver for the whole time as they walked towards the opposite curb across the imaginary road and back (as shown in Fig. 2a) or avoid it by fixating on the body of the car. One repetition in the standing trials involved head turning to look at the car/driver, followed by eye contact/no eye contact, and head-turning to look away from the car/driver. One repetition in the crossing trials involved head turning to look at the car/driver, followed by walking once in front of the vehicle towards the opposite curb and back while maintaining/avoiding eye contact, and head-turning to look away from the car/driver.

The driver briefly sought eye contact with the pedestrian in all standing trials, irrespective of whether the pedestrian was looking back at him. There was no predefined duration for the driver's eye contact seeking in the standing trials as this would be difficult to execute perfectly, so it was left to his discretion and what felt natural. In the crossing trials, the driver followed the eyes of the walking pedestrian with his gaze for the entire duration (as seen in Fig. 2b), irrespective of whether the pedestrian was looking back at him.

Before and after each trial, the participant was instructed to look at the torch held by the experimenter (standing in front of the car a few meters beyond the pedestrian) and wait for a flash, which served as an instantaneous marker for the start and end of that trial, respectively. The driver also knew to look at the torch at the beginning and the end of every trial.

All trials were recorded by both the Tobii and stereo cameras. After completing all the trials, the participants completed a questionnaire about their participation experience on 7-point Likert items (see Table S1 in the supplementary material).
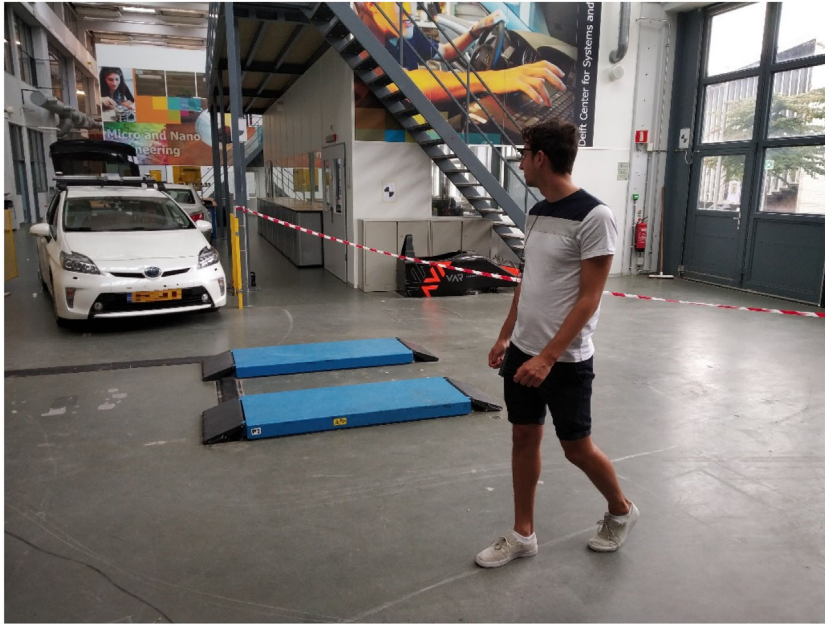
Fig. 2a. A pedestrian crossing towards the right curb, with the driver and pedestrian looking at each other's eyes throughout the interaction. The participant provided permission for the publication of this image.

Fig. 2b. The driver watching a pedestrian who is crossing towards the right curb, with the driver and pedestrian looking at each other's eyes throughout the interaction. Both persons provided permission for the publication of this image.

3. Analysis

3.1. Data export

The gaze data collected by the Tobii, such as the pedestrian's gaze direction and gaze points in pixel coordinates corresponding to its camera video, were exported as Microsoft Excel files via the Tobii Pro Lab software. The data collected

by the Smart Eye and the stereo camera, such as the driver's gaze direction, position of his eyes, and the pedestrian's location, were saved in ROSBAG format and subsequently converted to comma-separated values (CSV) files via ROS.

3.2. Gaze data quality

First, the quality of the raw gaze data of the experimental trials was assessed using a gaze sample percentage, defined as the percentage of the trial duration for which the driver's or pedestrian's gaze direction could be measured. Any instants with missing data on either the driver's end or the pedestrian's end were filled using the previous non-missing entry.

3.3. Driver eye contact seeking

To determine the vector connecting the driver's eyes to the pedestrian's eyes in 3D space (henceforth referred to as the 'ideal driver gaze'), the location of the driver's eyes (i.e., the midpoint between his right and left eye) was measured by the Smart Eye. The location of the pedestrian's eyes (i.e., the midpoint between the right and left eye) was estimated based on the pedestrian's location (x, y) obtained from the stereo camera and the pedestrian's eye height (z) calculated as 0.1 m below the pedestrian's self-reported height. Variation in the pedestrian's eye height due to their gait was assumed to be negligible. The stereo camera also detected the location of the experimenter holding the torch, which counted as false positives in the pedestrian detection. These readings were eliminated using a heuristically determined distance threshold of 7.3 m; that is, detections nearer than the threshold were the pedestrian and therefore retained, and those farther were the experimenter and therefore discarded.

A global coordinate system was defined, and the measurements from the Smart Eye (i.e., driver's instantaneous gaze direction and the position of his eyes) and the stereo camera (i.e., pedestrian's location) were converted to it from the devices' coordinate systems, as seen in Eq. (1). This equation describes the quaternion transformation of a coordinate system.

$$[0, x_{universal}, y_{universal}, z_{universal}] = [s, a, b, c]^* \left([0, x_{device}, y_{device}, z_{device}] - [0, k_x, k_y, k_z]\right)[s, a, b, c], \tag{1}$$

with * denoting the complex conjugate. To accomplish the coordinate transformation, translation of the measurements was performed using a translation matrix $[k_x, k_y, k_z]$, followed by rotation using a quaternion $[s, a, b, c]$, as seen above. Finally, all data were resampled to a common sampling rate of 100 Hz using linear interpolation.

The driver's eye contact seeking for each sampling instant was determined using the angle between the 'ideal driver gaze' vector and the driver's instantaneous gaze direction vector, as shown in Eq. (2). This equation uses Euclidean geometry to find the angle between two vectors in 3D that originate from a common point. Driver eye contact seeking was operationalized as gaze angle error $\theta_{driver} < 4°$. This threshold was based on a visual inspection of the distribution of θ_{driver} for all eye contact (EC) trials combined.

$$\theta_{driver} = \tan^{-1} \left(\frac{\|ideal\ driver\ gaze \times instantaneous\ gaze\ direction\|}{ideal\ driver\ gaze \cdot instantaneous\ gaze\ direction} \right) \tag{2}$$

3.4. Pedestrian eye contact seeking

The location of the driver's eyes in the video recordings from the pedestrian's head-mounted camera was estimated on a frame-by-frame basis using computer vision. First, for increasing computational speed, each frame was resized from 1920 × 1080 to 192 × 108 pixels, and only the blue dimension from the RGB dimensions was retained. Next, the location of the Toyota Prius in each frame was estimated using a normalized two-dimensional cross-correlation technique (i.e., template matching; cf. [41]) with 82 reference images of the car (also consisting of only the blue dimension) with a resolution of 51 × 36 pixels. These reference images were cropped from Tobii recordings when the car was in view under different lighting conditions, angles, and perspectives. Pixel coordinates marking the driver's eyes in the reference images were manually coded. The location of the driver's eyes in each frame was estimated by finding the reference image that offered the maximum correlation (while also being above an empirically determined correlation threshold of 0.73), as shown in Fig. 3.

The pedestrian's gaze point in pixel coordinates in each frame of a recording was available from the Tobii gaze data, and these values were divided by ten to suit the resized frames of 192 × 108 pixels. Pedestrian eye contact seeking was operationalized as $\theta_{pedestrian} < 4°$. $\theta_{pedestrian}$ was approximated by converting the Pythagorean distance in pixels between the location of the driver's eyes and the pedestrian's gaze point into an angle (Eq. (3)). Since the width of the resized frames is 192 pixels and the field of view of the Tobii camera is 90°, an 8.53-pixel distance corresponds to an angle of 4°.[1]

$$\theta_{pedestrian} = \frac{90}{192} \sqrt{\left(driv.head\ location_x - ped.gaze\ point_x\right)^2 + \left(driv.head\ location_y - ped.gaze\ point_y\right)^2} \tag{3}$$

---

[1] The approximation of 4° corresponding to 8.53 pixels relies on assumptions related to the Tobii's camera perspective. We discovered later on that the horizontal field of view of the Tobii camera is about 82°, not 90° (90°, reported by the manufacturer, is likely the diagonal field of view). At the same time, Eq. (3) assumes small angles, which may be untenable. Using manual measurements with the Tobii, we found that at low eye eccentricities of −15° to 15° (i.e., looking ahead), a step of 4° corresponds to 8 to 9 pixels and that at very high eye eccentricities of −35° or 35° (i.e., eyes turned strongly towards the left or right), a step of 4° corresponds to about 12.0 pixels. It may be assumed that participants were mostly looking ahead and that our adopted threshold of 8.53 pixels is indeed close to 4°.
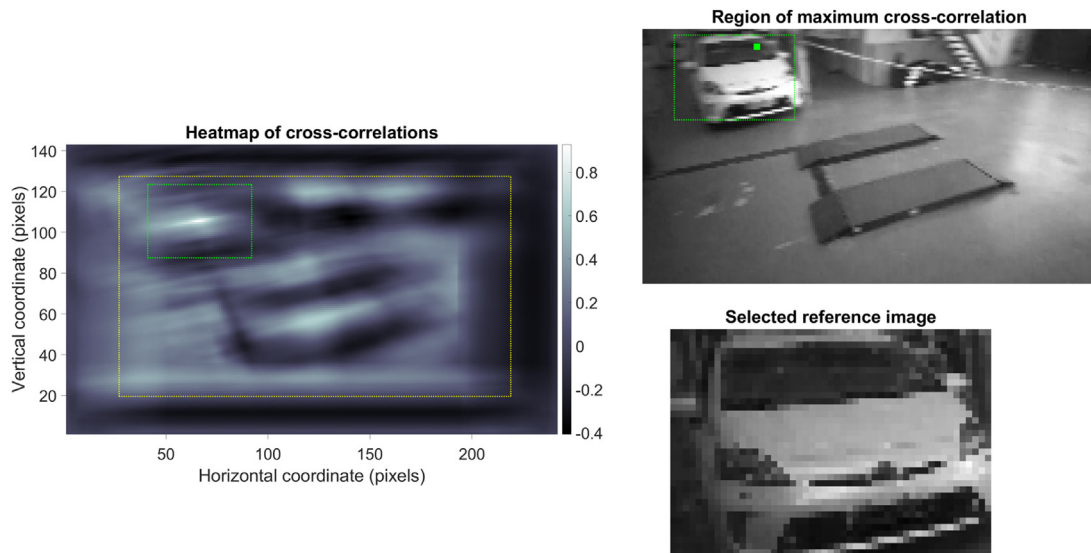
Fig. 3. Heatmap of cross-correlations (left) between a resized Tobii video frame (top right) and the reference image having the highest value of maximum correlation with the frame among all reference images (bottom right). A green box marks the region of maximum correlation. The driver's estimated position is marked with a green marker.

3.5. Driver–pedestrian eye contact establishment

To recap, driver eye contact seeking was operationalized as looking at the pedestrian with a gaze angle error smaller than 4°. Similarly, pedestrian eye contact seeking was operationalized as looking at the driver with a gaze angle error smaller than 4°. Eye contact was established when driver eye contact seeking and pedestrian eye contact seeking occurred concurrently. Finally, a classification between trials involving eye contact and those involving no eye contact was made by examining, per participant, whether the eye contact duration of the former type of trial exceeded the corresponding eye contact duration of the latter type of trial.

4. Results

4.1. Gaze data quality

Gaze data quality during the experiment was found to be high. The mean gaze sample percentage (SD) of the Smart Eye (i.e., driver) was 99.6% (0.7%) and 99.7% (0.6%) for the standing and crossing trials, respectively. The corresponding values of the Tobii (i.e., pedestrian) were 91.7% (5.3%) and 94.4% (4.4%) for the standing and crossing trials, respectively.

4.2. Driver eye contact seeking

Fig. 4 depicts the frequency distributions of driver gaze angle error (θ_{driver}) for all standing and crossing trials. A distinction is made between trials involving eye contact (EC), shown in blue, and trials involving no eye contact (NEC), shown as black dotted lines. There is no marked difference in the curves in terms of the number of data points, which can be explained by the fact that the driver was not asked to adjust his gaze behavior based on the eye contact seeking behavior of the pedestrian. There was more eye contact in the crossing trials compared to the standing trials (see the ratio of the two peaks of the bimodal distribution), because the driver continuously tracked the pedestrian in the former case, whereas in the latter case, he only briefly sought eye contact with the pedestrian and looked away, and repeated this process thrice. From Fig. 4, it can be seen that our threshold of 4° is a suitable choice, as it captures the large peak in the number of data points, which correspond to eye contact. More specifically, on combining the four distributions depicted in Fig. 4, there were 2480 s of data for $0° < \theta_{driver} \leq 4°$ and only 60 s of data for $4° < \theta_{driver} \leq 8°$. When the driver was not looking at the pedestrian, he was either looking at the torch (at the start and end of every trial) or at the experimenter (after each repetition in the standing trials); this explains the second peak in the distributions, between 8° and 14°.
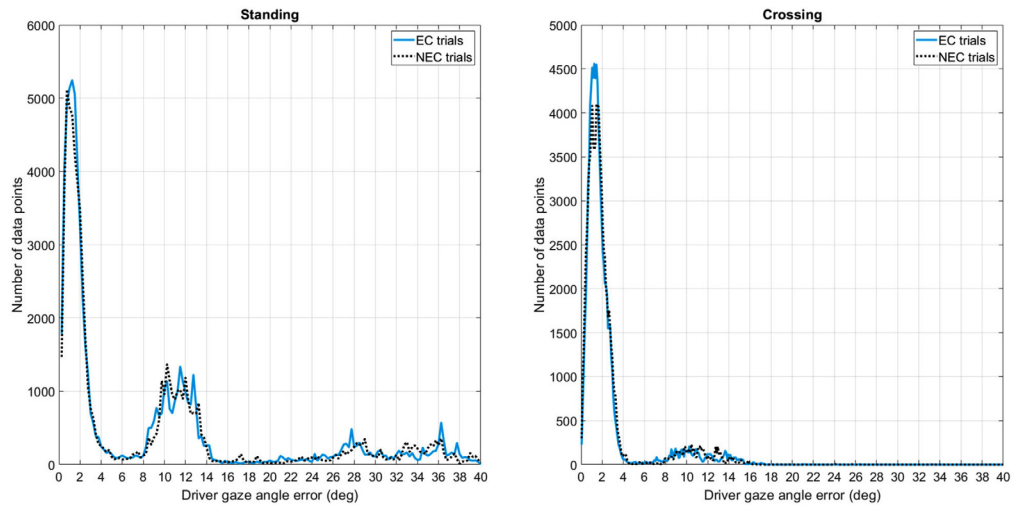
Fig. 4. Distribution of the driver gaze angle error θ_{driver} (i.e., the angle between the instantaneous gaze direction of the driver and the 'ideal driver gaze') for eye contact (EC) and no eye contact (NEC) trials. The sampling rate was 100 Hz, and the resolution of the distribution was 0.25°. Left: Standing trials, Right: Crossing trials.

4.3. Pedestrian eye contact seeking

Fig. 5 shows similar information as Fig. 4, but from the pedestrian's perspective. In contrast to the driver gaze angle error, the distributions of the pedestrian gaze angle error ($\theta_{pedestrian}$) are not bimodal but unimodal, which can be explained by individual differences, task instructions, and head and body rotation. That is, during the standing EC trials, participants were asked to seek eye contact thrice and look away after each time, without instructions about where to look when looking away. Thus, due to varying levels of head and body rotation when looking away, the car (and therefore also the driver) was located at widely different parts of the visual field of the Tobii camera, or even completely outside of it. Accordingly, there are no distinct second peaks in the distribution of $\theta_{pedestrian}$, as was the case for driver gaze angle error shown in Fig. 4. Fig. 5 further shows a clear distinction in the pedestrian gaze angle error between the EC and NEC trials, with the mode of the distribution being about 1.5° (i.e., very close to the driver's eyes) for EC trials and about 12° (i.e., corresponding to some location on the car, as per the participants' task instructions) for NEC trials. The latter observation is corroborated by manual geometric calculations, which show that the pedestrian gaze angle error in an NEC trial when they are looking at the car's number plate (a plausible scenario) is 10–12°. This range of values is well above the 4° threshold, thereby reducing the likelihood of there being many false positives of eye contact in the NEC trials, assuming that the participants carried out the instructions faithfully. Our chosen $\theta_{pedestrian}$ threshold of 4° also represents a good trade-off for making a classification between EC and NEC trials, as gaze angle errors below this value capture a large portion of the samples in the EC trials (pedestrians were expected to exhibit small $\theta_{pedestrian}$ values for a portion of the EC trials) while capturing relatively few samples in the NEC trials (pedestrians were expected not to exhibit small $\theta_{pedestrian}$ values, since they were instructed not to make eye contact). Figure S1 in the supplementary material provides further justification for the 4° threshold through a sensitivity analysis.

4.4. Occurrence and reconstruction of eye contact

Fig. 6 shows four time-synchronized plots from a particular instant of a crossing trial with instructed eye contact (C-EC). In the bottom half, animations (comprising of a top view on the left and a side view on the right) of the trial are shown, created by plotting the position of the driver's eyes (green marker), his gaze direction (green dashed line), and the position of the pedestrian's eyes (magenta marker), along with an image of the Toyota Prius. Along the axes, X, Y, and Z represent the longitudinal, lateral, and vertical position in the world. The pedestrian's gaze direction is intentionally not plotted, as there was no sufficiently accurate way to translate it from a 2D pixel coordinate in the Tobii recordings to a 3D gaze direction for the animations. However, this omission does not have a bearing on detecting eye contact. The left top shows a Tobii camera screenshot with the pedestrian's view during a crossing trial, resized to a resolution of 192 × 108 pixels, and with the pedestrian's gaze point and the driver's estimated location overlaid as magenta and green markers, respectively. The top right part shows the driver's and pedestrian's gaze angle error plotted over time together with the 4° threshold of eye contact. Eye contact occurs when the values for the driver (green) and pedestrian (magenta) are both under the horizontal threshold line at the same time. It can be seen that for most of the trial, both the driver and the pedestrian sought eye contact.
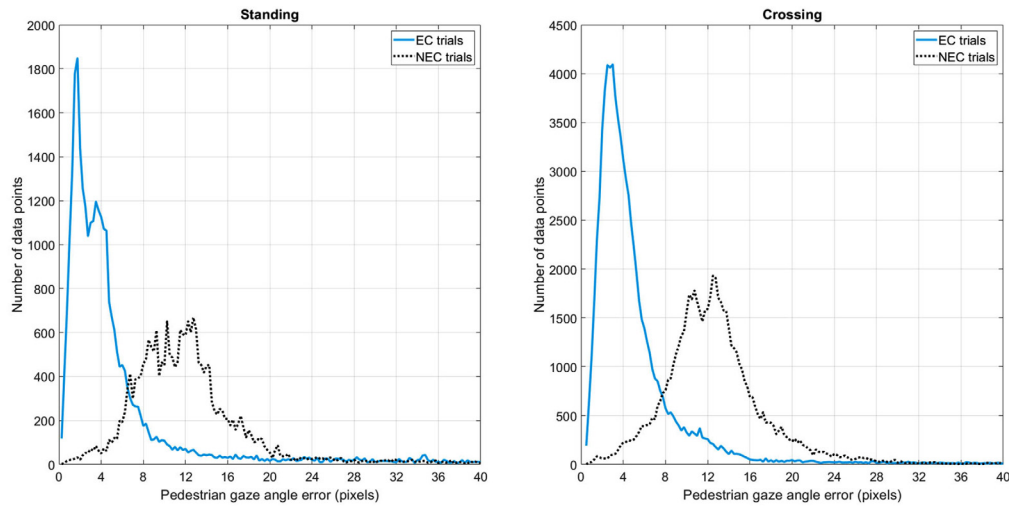
Fig. 5. Distribution of the pedestrian gaze angle error $\theta_{pedestrian}$ (i.e., the angle between the gaze point of the pedestrian and the estimated position of the driver's eyes) for eye contact (EC) and no eye contact (NEC) trials. The sampling rate was 100 Hz, and the resolution of the distribution was 0.25°. Left: Standing trials, Right: Crossing trials.
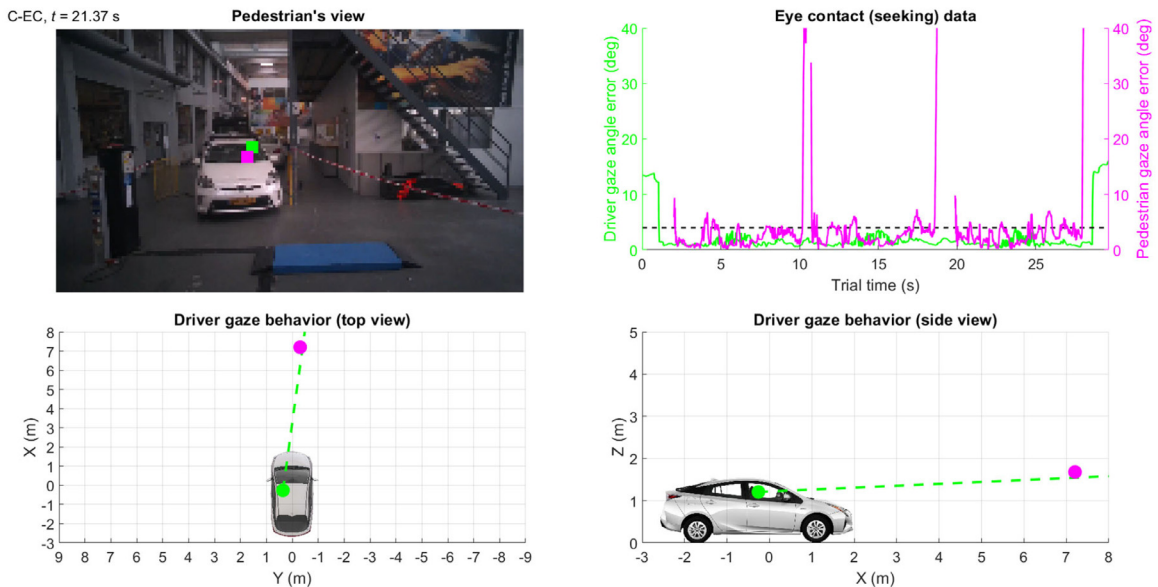

Fig. 6. Eye contact overview of one trial (Crossing, Eye Contact; C-EC). Top left figure: Pedestrian's view. Top right figure: Driver and pedestrian gaze angle error as a function of time. Bottom left figure: Top view. Bottom right figure: Side view. Green marker: Estimated location of the driver's eyes. Magenta marker: Estimated location of the pedestrian's eyes (bottom figures) or pedestrian's gaze point (top left figure).

The large gaze angle errors in the first and last two seconds of the trial are caused by the driver and pedestrian looking away from and towards the torch flash. The other sharp spikes at around 10 s and 19 s in the pedestrian's eye contact seeking graph are due to looking away from the driver upon completing a repetition (i.e., walking towards the opposite curb and back once). The gaps accompanying the spikes are due to the gaze angle error being undefined because the car (and therefore also the driver) being out of view in the Tobii camera.
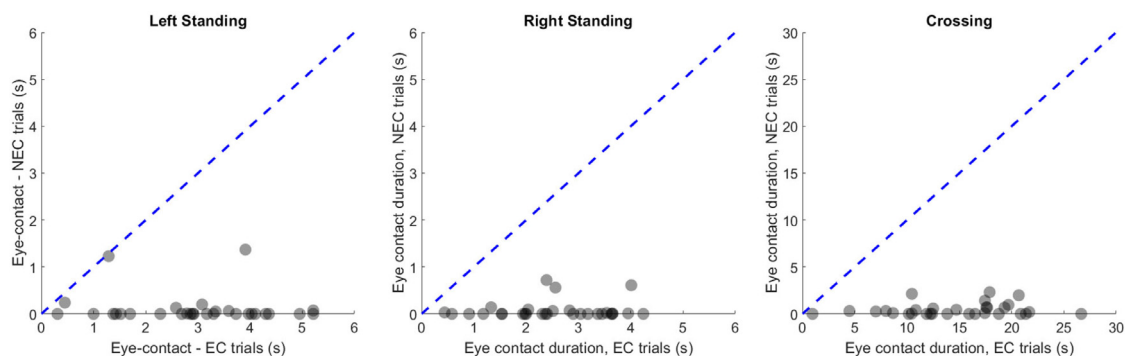
4.5. Eye contact durations

Table 2 shows the means and standard deviations across the 30 participants for eye contact measures for the driver and the pedestrian. The driver sought eye contact about 55% of the time in the standing trials and in about 90% of the time in the crossing trials. Pedestrians sought eye contact for 20%–25% of the duration of the standing trials with instructed

Table 2

Means and standard deviations of eye contact measures for the standing trials and crossing trials. One trial consists of three repetitions.

	L-S-EC Mean	L-S-NEC Mean	R-S-EC Mean	R-S-NEC Mean	L-S-EC SD	L-S-NEC SD	R-S-EC SD	R-S-NEC SD
Trial duration (s)	12.35	11.97	12.38	11.81	3.25	2.90	2.72	2.71
Driver eye contact seeking (s)	6.82	6.59	6.80	6.47	2.18	1.92	1.63	1.79
Driver eye contact seeking (% of time)	55.0%	54.8%	55.1%	54.4%	5.6%	6.6%	5.9%	5.6%
Pedestrian eye contact seeking (s)	3.18	0.12	2.75	0.09	1.50	0.33	1.16	0.21
Pedestrian eye contact seeking (% of time)	26.0%	1.0%	22.6%	0.8%	11.1%	2.9%	9.2%	1.8%
Eye contact (s)	2.94	0.11	2.52	0.08	1.35	0.33	1.04	0.19
Eye contact (% of time)	24.3%	0.9%	20.8%	0.7%	10.7%	2.9%	8.7%	1.7%

	C-EC Mean	C-NEC Mean	C-EC SD	C-NEC SD
Trial duration (s)	30.83	30.94	4.70	4.90
Driver eye contact seeking (s)	27.94	28.05	4.11	4.41
Driver eye contact seeking (% of time)	90.7%	90.7%	2.4%	3.1%
Pedestrian eye contact seeking (s)	14.94	0.51	5.91	0.77
Pedestrian eye contact seeking (% of time)	49.0%	1.7%	19.5%	2.6%
Eye contact (s)	14.61	0.45	5.72	0.67
Eye contact (% of time)	47.9%	1.5%	18.9%	2.2%

Fig. 7. Comparison of pedestrian eye contact durations in the trials with and without instructed eye contact. Each marker represents one participant.

eye contact and roughly 45% of the duration of the corresponding crossing trials. These values are consistent with the experimental protocol, where both parties sought intermittent eye contact with each other in the standing trials and visually tracked each other in the crossing trials. The pedestrian's percentages are lower than the driver's, as expected, because of the greater extent of head and body rotations by the pedestrian, thereby taking longer to (re-)establish eye contact.

Mean pedestrian eye contact seeking duration was 3.18 s in the L-S-EC trials, 2.75 s in the R-S-EC trials, and 14.94 s in the C-EC trials (see Table 2). Since each trial consisted of three repetitions, pedestrians sought eye contact for an average of 1.06 s, 0.92 s, and 4.98 s in a single repetition, respectively. Pedestrians crossed the road twice in a single repetition, so the average pedestrian eye contact seeking duration in only one direction was half of 4.98 s, or 2.49 s.

The mean durations of eye contact were 2.94 s, 2.52 s, and 14.61 s for the L-S-EC, R-S-EC, and C-EC trials, respectively (see Table 2). This meant that in one repetition, eye contact lasted for 0.98 s, 0.84 s, and 4.87 s, in that order. Dividing the third value by two gives the average eye contact duration per repetition while crossing the road one-way, that is, 2.43 s.

4.6. Classification of trials

Fig. 7 illustrates the classification performance of trials with and without instructed eye contact based on eye contact duration. It can be seen that the type of trial is distinguished with 100% accuracy within-subject. In other words, all markers in Fig. 7 lie below the diagonal line.

5. Discussion

5.1. Main findings

This study aimed to develop an eye contact detection method to address the research gap in the objective measurement of eye contact in the traffic context. Our method's main innovation was the use of two eye trackers to detect driver–pedestrian eye contact. The use of computer vision techniques to estimate the driver's and pedestrian's locations

eliminates the need for manually coding the areas of interest. Compared to existing techniques such as self-reports and button press responses for recording eye contact, our dual eye-tracking method is accurate, since it does not rely on subjective perceptions of eye contact occurrence and is not influenced by reaction times. Accordingly, our setup may be useful for experimental research in staged scenarios and may form the first step towards real-time eye contact detection in crossing conflicts.

We provided a new operationalization for eye contact, namely that it occurred when the gaze angle errors of the driver and the pedestrian were both below 4°. The 4° threshold was determined heuristically based on the angle error distributions. However, the selected threshold also appears to have psychological significance: A study using animated faces by Gamer and Hecht [42] showed that eye gaze around faces (at a distance of 5 m) was in the form of a cone of angular width up to approximately 8°, which translates to a gaze angle error of up to 4°.

Our method was validated using staged interactions with and without eye contact and yielded perfect within-subject classification. Furthermore, we generated animations of the driver–pedestrian interactions, demonstrating that traffic encounters could be reconstructed using only the information obtained from cameras and eye trackers. Such a visualization could prove useful to enhance the situational awareness of occupants of the vehicle (see [43] for a top–down display that enhances situational awareness in automated vehicles).

Participants were instructed to seek eye contact with the driver briefly in the standing trials (but were not told to look for a specific amount of time) and to walk in front of the car in the crossing trials (but were not told how fast to walk). It is worth noting that pedestrians have mostly been observed seeking eye contact when a vehicle is close to them and moving at low speeds or stopped [31], which resembles our experimental setting. However, strong conclusions about how long each party seeks eye contact or how long eye contact lasts in a real driver–pedestrian interaction cannot be made using our measurements. The observed eye contact durations (0.9 s and 2.4 s for a standing and crossing pedestrian, respectively) may be higher than what one might expect in real traffic. At pedestrian crossings in real traffic, road users look at various elements of the scene, including signs and road markings [44], not just the other party's eyes. Furthermore, pedestrians may stop glancing at the car when it has become clear that the pedestrian can cross before the car (see [45], for a similar phenomenon in pedestrian–pedestrian interaction).

As a corollary of our operationalization of eye contact (i.e., a logical AND of the driver's and pedestrian's gaze behavior), eye contact duration was always less than or equal to the lower of the two eye contact seeking durations. Of note, the mean eye contact durations for the trials were closer in magnitude to the pedestrian's mean durations of eye contact seeking than those of the driver's. This was probably due to the driver constantly tracking the pedestrian in the crossing trials, whereas the pedestrian had to turn around and look away.

5.2. Limitations

Our method has a few limitations. First, our operationalization of eye contact does not include the subjective awareness that eye contact is occurring. This could have been measured using a think-aloud method or with event recorders (as noted by [35]) in the hands of both parties. It would be interesting to examine the association between objective and subjective driver–pedestrian eye contact, something that has become possible through our eye contact detection method. Pedestrians in the present study reported being highly involved in the task (see Table S1 in the supplementary material), so close congruence between subjective and objective eye contact is expected. In more demanding scenarios, it may be the case that the driver and the pedestrian are objectively looking at each other but not subjectively aware that they are making eye contact, i.e., the 'looking but not seeing' phenomenon [46].

A second limitation is that aspects of synchronization, image recognition, and data processing were still performed manually. It is noted that our algorithm, though we ran it offline, processes data on a frame-by-frame basis (i.e., without forward-looking filters), and therefore can be made to run in real-time. The Smart Eye and the stereo camera already reported driver gaze and pedestrian location in real-time. If the Tobii and its camera could also be configured to do the same via the API provided by the manufacturer, real-time eye contact detection is a viable target. We are currently developing a real-time pedestrian feedback system based on this proposed setup, where auditory feedback is provided to the pedestrian depending on whether the pedestrian has or has not looked at the car.

A third limitation concerns the artificial setup of our experiment. Our study involved a staged indoor experiment with a stationary vehicle, not to mention that most pedestrians are not equipped with wearable eye trackers [47]. These issues may be solved in the future with greater affordability of eye trackers and the advent of smart glasses or eye-tracking contact lenses [48]. Findings of a pilot study revealed that the Tobii performed poorly at tracking the user's gaze outdoors, presumably due to infrared radiation in sunlight [49]. The pilot test also found that the visibility of the driver was compromised because of windshield glare outside. Windshield glare appears to be a factor that prevents eye contact in traffic [19], suggesting a need for synthetic eye contact detection, such as ours.

A final limitation is that our method used basic computer vision techniques for detecting the pedestrian and vehicle. Although road user detection worked reliably in our case, more sophisticated methods, which are now becoming available for less than $100 (e.g., Nvidia Jetson [50]), would be required to make our method work with a wider variety of road users. As a proof of concept, we applied recent object recognition software intended for real-time usage (YOLOv5 [51]) on one of our experimental videos captured with the Tobii's camera. Results showed that the algorithm detected the target car, a car in the background, and persons outside the car, but not the experimenter in the car (see Fig. 8). Although the detection
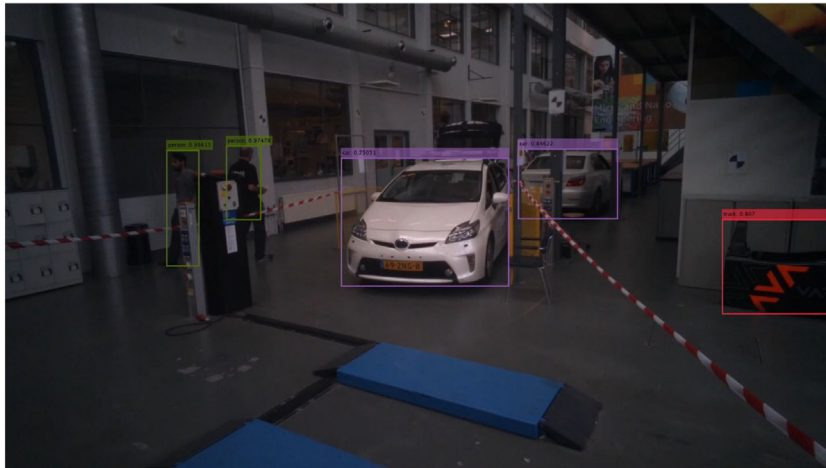
Fig. 8. Screenshot of object recognition applied to one of our Tobii videos.

was not as robust as our template matching approach (the target car was often labeled a truck, bus, or train), there clearly appears to be potential for real-time usage in situations with multiple different vehicles. It is worth remembering that our image technique also does not detect the driver in the car, and hence the choice of image recognition algorithm does not largely affect the detection of eye contact.

5.3. Outlook & conclusion

There is ample scope for further research and applications. The topic of driver–pedestrian eye contact is not only of interest to manual and semi-automated driving (SAE Levels 0–2), it is also relevant to automated vehicles in which the driver is intermittently inattentive (SAE Levels 3 and 4). The vacuum created by missing eye contact in road interactions opens up possibilities to artificially substitute it. The anthropomorphic eye contact eHMI variants proposed by Ochiai and Toyoshima [52], Chang et al. [15], and Rover [53] are one way to achieve this. Eye contact could also be used as an objective input in vehicles or wearables for providing warnings (e.g., 'mind the pedestrian', 'watch out, the driver is distracted') or in automated vehicle control (e.g., braking earlier if there is no eye contact).

As pointed out above, an issue is that drivers and pedestrians are currently not equipped with eye trackers. As an alternative to wearable eye-tracking, vehicle-based eye-contact estimation may be possible through pedestrian head orientation estimation combined with contextual information (e.g., [54–56]; for a survey of methods, see [57]). However, for the time being, our method may be most useful for research purposes in staged scenarios. For example, our method could be applied to outdoor experiments (in cloudy weather) to study eye contact in situations where right-of-way is not clear (e.g., [58]).

To conclude, the present study validated a novel eye contact detection method. Our method may stimulate further research that aims to obtain a deeper understanding of eye contact and its role in traffic.

Supplementary material

The two questionnaires, a demo video corresponding to Fig. 6, the data, and MATLAB codes used for the analyses are available at https://doi.org/10.4121/15134037.

CRediT authorship contribution statement

V. Onkhar: Conceived the detection method and experiment design, Executed the experiment, Formal analysis, Writing – original draft. **P. Bazilinskyy:** Conceived the detection method and experiment design, Writing – review & editing. **J.C.J. Stapel:** Executed the experiment, Formal analysis, Writing – review & editing. **D. Dodou:** Formal analysis, Writing – original draft. **D. Gavrila:** Provided access to equipment, Writing – review & editing. **J.C.F. de Winter:** Conceived the detection method and experiment design, Formal analysis, Writing – original draft.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgments

References

[1] World Health Organization, Global status report on road safety, 2018, Retrieved from: https://www.who.int/violence_injury_prevention/road_safety_status/2018/en/.

[2] National Highway Traffic Safety Administration (NHTSA), 2017 fatal motor vehicle crashes: overview, 2018, Retrieved from: https://crashstats.nhtsa.dot.gov/Api/Public/ViewPublication/812603.

[3] M.P. DaSilva, J.D. Smith, W.G. Najm, Analysis of Pedestrian Crashes, Report No. DOT-VNTSC-NHTSA-02-02, John A. Volpe National Transportation Systems Center, Cambridge, MA, 2004, Retrieved from: http://www.nhtsa.gov/DOT/NHTSA/NRD/Multimedia/PDFs/Crash%20Avoidance/2003/DOTHS809585.pdf.

[4] B. Färber, Communication and communication problems between autonomous vehicles and human drivers, in: M. Maurer, J. Gerdes, B. Lenz, H. Winner (Eds.), Autonomous Driving, Springer, Berlin, Heidelberg, 2016, pp. 125–144, http://dx.doi.org/10.1007/978-3-662-48847-8_7.

[5] A. Habibovic, V. Malmsten Lundgren, J. Andersson, M. Klingegård, T. Lagström, A. Sirkka, et al., Communicating intent of automated vehicles to pedestrians, Front. Psychol. 9 (2018) 1336, http://dx.doi.org/10.3389/fpsyg.2018.01336.

[6] Y.M. Lee, R. Madigan, O. Giles, L. Garach-Morcillo, G. Markkula, C. Fox, et al., Road users rarely use explicit communication when interacting in today's traffic: implications for automated vehicles, Cogn. Technol. Work 23 (2021) 367–380, http://dx.doi.org/10.1007/s10111-020-00635-y.

[7] M. Sucha, D. Dostal, R. Risser, Pedestrian-driver communication and decision strategies at marked crossings, Accid. Anal. Prev. 102 (2017) 41–50, http://dx.doi.org/10.1016/j.aap.2017.02.018.

[8] A. Rasouli, I. Kotseruba, J.K. Tsotsos, Agreeing to cross: How drivers and pedestrians communicate, in: Proceedings of the 2017 IEEE Intelligent Vehicles Symposium, Los Angeles, CA, 2017, pp. 264–269, http://dx.doi.org/10.1109/IVS.2017.7995730.

[9] S. Schmidt, B. Färber, Pedestrians at the kerb–recognising the action intentions of humans, Transp. Res. F Traffic Psychol. Behav. 12 (2009) 300–310, http://dx.doi.org/10.1016/j.trf.2009.02.003.

[10] V. Malmsten Lundgren, A. Habibovic, J. Andersson, T. Lagström, M. Nilsson, A. Sirkka, et al., Will there be new communication needs when introducing automated vehicles to the urban context? in: N. Stanton, S. Landry, G. Di Bucchianico, A. Vallicelli (Eds.), Advances in Human Aspects of Transportation, Springer, Cham, 2017, pp. 485–497, http://dx.doi.org/10.1007/978-3-319-41682-3_41.

[11] D. Rothenbücher, J. Li, D. Sirkin, B. Mok, W. Ju, Ghost driver: a platform for investigating interactions between pedestrians and driverless vehicles, in: Adjunct Proceedings of the 7th International Conference on Automotive User Interfaces and Interactive Vehicular Applications, Nottingham, UK, 2015, pp. 44–49, http://dx.doi.org/10.1145/2809730.2809755.

[12] london.ca, Make eye contact. Vision zero. London road safety strategy, 2019, Retrieved from http://www.london.ca/residents/Roads-Transportation/Road-Safety/Pages/make-eye-contact.aspx.

[13] National Highway Traffic Safety Administration (NHTSA), Pedestrian safety, 2018, Retrieved from https://www.nhtsa.gov/road-safety/pedestrian-safety.

[14] Veiligverkeer, Verkeersregels [Traffic rules], 2020, Retrieved from https://verkeersregels.vvn.nl/situatie/fietspad-voorrang.

[15] C.M. Chang, K. Toda, D. Sakamoto, T. Igarashi, Eyes on a car: an interface design for communication between an autonomous car and a pedestrian, in: Proceedings of the 9th International Conference on Automotive User Interfaces and Interactive Vehicular Applications, Oldenburg, Germany, 2017, pp. 65–73, http://dx.doi.org/10.1145/3122986.3122989.

[16] F. Schneemann, I. Gohl, Analyzing driver-pedestrian interaction at crosswalks: A contribution to autonomous driving in urban environments, in: Proceedings of the 2016 IEEE Intelligent Vehicles Symposium (IV), Gothenburg, Sweden, 2016, pp. 38–43, http://dx.doi.org/10.1109/IVS.2016.7535361.

[17] D. Moore, R. Currano, G.E. Strack, D. Sirkin, The case for implicit external human-machine interfaces for autonomous vehicles, in: Proceedings of the 11th International Conference on Automotive User Interfaces and Interactive Vehicular Applications, Utrecht, the Netherlands, 2019, pp. 295–307, http://dx.doi.org/10.1145/3342197.3345320.

[18] D. Dey, J. Terken, Pedestrian interaction with vehicles: roles of explicit and implicit communication, in: Proceedings of the 9th International Conference on Automotive User Interfaces and Interactive Vehicular Applications, Oldenburg, Germany, 2017, pp. 109–113, http://dx.doi.org/10.1145/3122986.3123009.

[19] D. AlAdawy, M. Glazer, J. Terwilliger, H. Schmidt, J. Domeyer, B. Mehler, et al., Eye contact between pedestrians and drivers, in: Proceedings of the Tenth International Driving Symposium on Human Factors in Driver Assessment, Santa Fe, New Mexico, 2019, pp. 301–307. http://dx.doi.org/10.17077/drivingassessment.1710.

[20] M. Snyder, J. Grather, K. Keller, Staring and compliance: A field experiment on hitchhiking, J. Appl. Soc. Psychol. 4 (1974) 165–170, http://dx.doi.org/10.1111/j.1559-1816.1974.tb00666.x.

[21] A. Katz, D. Zaidel, A. Elgrishi, An experimental study of driver and pedestrian interaction during the crossing conflict, Hum. Factors 17 (1975) 514–527, http://dx.doi.org/10.1177/001872087501700510.

[22] Z. Ren, X. Jiang, W. Wang, Analysis of the influence of pedestrians' eye contact on drivers' comfort boundary during the crossing conflict, Procedia Eng. 137 (2016) 399–406, http://dx.doi.org/10.1016/j.proeng.2016.01.274.

[23] I. Kotseruba, A. Rasouli, J.K. Tsotsos, Joint attention in autonomous driving (JAAD), 2016, arXiv. Retrieved from https://arxiv.org/abs/1609.04741.

[24] M. Roth, F. Flohr, D.M. Gavrila, Driver and pedestrian awareness-based collision risk analysis, in: Proceedings of the 2016 IEEE Intelligent Vehicles Symposium (IV), Gothenburg, Sweden, 2016, pp. 454–459, http://dx.doi.org/10.1109/IVS.2016.7535425.

[25] I. Walker, Signals are informative but slow down responses when drivers meet bicyclists at road junctions, Accid. Anal. Prev. 37 (2005) 1074–1085, http://dx.doi.org/10.1016/j.aap.2005.06.005.

[26] I. Walker, M. Brosnan, Drivers' gaze fixations during judgements about a bicyclist's intentions, Transp. Res. F Traffic Psychol. Behav. 10 (2007) 90–98, http://dx.doi.org/10.1016/j.trf.2006.06.001.

[27] D. Nathanael, E. Portouli, V. Papakostopoulos, K. Gkikas, A. Amditis, Naturalistic observation of interactions between car drivers and pedestrians in high density urban settings, in: S. Bagnara, R. Tartaglia, S. Albolino, T. Alexander, Y. Fujita (Eds.), Proceedings of the 20th Congress of the International Ergonomics Association (IEA 2018), IEA 2018, Springer, Cham, 2018, pp. 389–397, http://dx.doi.org/10.1007/978-3-319-96074-6_42.

[28] F. Diederichs, T. Schüttke, D. Spath, Driver intention algorithm for pedestrian protection and automated emergency braking systems, in: 2015 IEEE 18th International Conference on Intelligent Transportation Systems, 2015, pp. 1049–1054, http://dx.doi.org/10.1109/ITSC.2015.174.

[29] A. Borowsky, T. Oron-Gilad, A. Meir, Y. Parmet, Drivers' perception of vulnerable road users: A hazard perception approach, Accid. Anal. Prev. 44 (2012) 160–166, http://dx.doi.org/10.1016/j.aap.2010.11.029.

[30] J.C.F. De Winter, P. Bazilinskyy, D. Wesdorp, V. De Vlam, B. Hopmans, J. Visscher, D. Dodou, How do pedestrians distribute their visual attention when walking through a parking garage? An eye-tracking study, Ergonomics 64 (2021) 793–805, http://dx.doi.org/10.1080/00140139.2020.1862310.

[31] D. Dey, F. Walker, M. Martens, J. Terken, Gaze patterns in pedestrian interaction with vehicles: towards effective design of external human-machine interfaces for automated vehicles, in: Proceedings of the 11th International Conference on Automotive User Interfaces and Interactive Vehicular Applications, Utrecht, the Netherlands, 2019, pp. 369–378, http://dx.doi.org/10.1145/3342197.3344523.

[32] L. Lévêque, M. Ranchet, J. Deniel, J.C. Bornard, T. Bellet, Where do pedestrians look when crossing? A state of the art of the eye-tracking studies, IEEE Access 8 (2020) 163833–164833, http://dx.doi.org/10.1109/ACCESS.2020.3021208.

[33] F. Broz, H. Lehmann, C.L. Nehaniv, K. Dautenhahn, Mutual gaze, personality, and familiarity: Dual eye-tracking during conversation, in: IEEE RO-MAN: The 21st IEEE International Symposium on Robot and Human Interactive Communication, 2012, pp. 858–864, http://dx.doi.org/10.1109/ROMAN.2012.6343859.

[34] S.L. Rogers, C.P. Speelman, O. Guidetti, M. Longmuir, Using dual eye tracking to uncover personal gaze patterns during social interaction, Sci. Rep. 8 (2018) 4271, http://dx.doi.org/10.1038/s41598-018-22726-7.

[35] C. Jongerius, R.S. Hessels, J.A. Romijn, E.M.A. Smets, M.A. Hillen, The measurement of eye contact in human interactions: A scoping review, J. Nonverbal Behav. 44 (2020) 363–389, http://dx.doi.org/10.1007/s10919-020-00333-3.

[36] J. Heron, The phenomenology of social encounter: The gaze, Philos. Phenomenol. Res. 31 (1970) 243–264, https://www.jstor.org/stable/2105742.

[37] L. Ferranti, B. Brito, E. Pool, Y. Zheng, R.M. Ensing, R. Happee, et al., SafeVRU: A research platform for the interaction of self-driving vehicles with vulnerable road users, in: Proceedings of the 2019 IEEE Intelligent Vehicles Symposium (IV), Paris, France, 2019, pp. 1660–1666, http://dx.doi.org/10.1109/IVS.2019.8813899.

[38] Tobii AB, How do tobii eye trackers work? 2021, Retrieved from: https://www.tobiipro.com/learn-and-support/learn/eye-tracking-essentials/how-do-tobii-eye-trackers-work/.

[39] M. Braun, S. Krebs, F. Flohr, D.M. Gavrila, EuroCity persons: A novel benchmark for person detection in traffic scenes, IEEE Trans. Pattern Anal. Mach. Intell. 41 (2019) 1844–1861, http://dx.doi.org/10.1109/TPAMI.2019.2897684.

[40] J. Domhof, J.F.P. Kooij, D.M. Gavrila, An extrinsic calibration tool for radar, camera and lidar, in: International Conference on Robotics and Automation, ICRA, 2019, pp. 8107–8113, http://dx.doi.org/10.1109/ICRA.2019.8794186.

[41] K. Briechle, U.D. Hanebeck, Template matching using fast normalized cross correlation, in: Optical Pattern Recognition XII, Vol. 4387, International Society for Optics and Photonics, 2001, pp. 95–102, http://dx.doi.org/10.1117/12.421129.

[42] M. Gamer, H. Hecht, Are you looking at me? Measuring the cone of gaze, J. Exp. Psychol. [Hum. Percept.] 33 (2007) 705–715, http://dx.doi.org/10.1037/0096-1523.33.3.705.

[43] C.C. Chang, R.A. Grier, J. Maynard, J. Shutko, M. Blommer, R. Swaminathan, R. Curry, Using a situational awareness display to improve rider trust and comfort with an AV taxi, in: Proceedings of the Human Factors and Ergonomics Society Annual Meeting, Vol. 63, 2019, pp. 2083–2087, http://dx.doi.org/10.1177/1071181319631428.

[44] A. Bichicchi, F. Mazzotta, C. Lantieri, V. Vignali, A. Simone, G. Dondi, M. Costa, The influence of pedestrian crossings features on driving behavior and road safety, in: Transport Infrastructure and Systems, CRC Press, 2017, pp. 741–746, http://dx.doi.org/10.1201/9781315281896-95.

[45] J.L. Croft, D. Panchuk, Watch where you're going? Interferer velocity and visual behavior predicts avoidance strategy during pedestrian encounters, J. Motor Behav. 50 (2017) 353–363, http://dx.doi.org/10.1080/00222895.2017.1363695.

[46] C.B. White, J.K. Caird, The blind date: The effects of change blindness, passenger conversation and gender on looked-but-failed-to-see (LBFTS) errors, Accid. Anal. Prev. 42 (2010) 1822–1830, http://dx.doi.org/10.1016/j.aap.2010.05.003.

[47] W. Tabone, J.C.F. De Winter, C. Ackermann, J. Bärgman, M. Baumann, S. Deb, C. Emmenegger, A. Habibovic, M. Hagenzieker, P.A. Hancock, R. Happee, J. Krems, J.D. Lee, M. Martens, N. Merat, D.A. Norman, T.B. Sheridan, N.A. Stanton, Vulnerable road users and the coming wave of automated vehicles: Expert perspectives, Transp. Res. Interdiscip. Perspect. 9 (2021) 100293, http://dx.doi.org/10.1016/j.trip.2020.100293.

[48] A. Khaldi, E. Daniel, L. Massin, C. Kärnfelt, F. Ferranti, C. Lahuec, et al., A laser emitting contact lens for eye tracking, Sci. Rep. 10 (2020) http://dx.doi.org/10.1038/s41598-020-71233-1.

[49] Tobii AB, Tobii pro glasses 2 user's manual, version 1.1.3, 2020, Retrieved from: https://www.tobiipro.com/siteassets/tobii-pro/user-manuals/tobii-pro-glasses-2-user-manual.pdf.

[50] A.A. Süzen, B. Duman, B. Şen, Benchmark analysis of Jetson TX2, Jetson Nano and Raspberry PI using Deep-CNN, in: International Congress on Human-Computer Interaction, Optimization and Robotic Applications, HORA, 2020, http://dx.doi.org/10.1109/HORA49412.2020.9152915.

[51] Cui, yoloV5-matlab [Computer software], 2021, https://www.mathworks.com/matlabcentral/fileexchange/89012-yolov5-matlab.

[52] Y. Ochiai, K. Toyoshima, Homunculus: the vehicle as augmented clothes, in: Proceedings of the 2nd Augmented Human International Conference, Tokyo, Japan, 2011, http://dx.doi.org/10.1145/1959826.1959829.

[53] Jaguar Land Rover, The virtual eyes have it, 2018, Retrieved from: https://www.jaguarlandrover.com/2018/virtual-eyes-have-it#::text=The%20'eyes'%20have%20been%20devised,intends%20to%20take%20avoiding%20action.

[54] R. Quintero, J. Almeida, D.F. Llorca, M.A. Sotelo, Pedestrian path prediction using body language traits, in: Proceedings of the 2014 IEEE Intelligent Vehicles Symposium Proceedings, Dearborn, MI, 2014, pp. 317–323, http://dx.doi.org/10.1109/IVS.2014.6856498.

[55] M. Raza, Z. Chen, S.U. Rehman, P. Wang, P. Bao, Appearance based pedestrians' head pose and body orientation estimation using deep learning, Neurocomputing 272 (2018) 647–659, http://dx.doi.org/10.1016/j.neucom.2017.07.029.

[56] D.A. Ridel, N. Deo, D. Wolf, M. Trivedi, Understanding pedestrian-vehicle interactions with vehicle mounted vision: An LSTM model and empirical analysis, in: Proceedings of the 2019 IEEE Intelligent Vehicles Symposium (IV), Paris, France, 2019, pp. 913–918, http://dx.doi.org/10.1109/IVS.2019.8813798.

[57] A. Rudenko, L. Palmieri, M. Herman, K.M. Kitani, D.M. Gavrila, K.O. Arras, Human motion trajectory prediction: A survey, Int. J. Robot. Res. 39 (2020) 895–935, http://dx.doi.org/10.1177/0278364920917446.

[58] T. Fu, W. Hu, L. Miranda-Moreno, N. Saunier, Investigating secondary pedestrian-vehicle interactions at non-signalized intersections using vision-based trajectory data, Transp. Res. C 105 (2019) 222–240, http://dx.doi.org/10.1016/j.trc.2019.06.001.