ARTICLE IN PRESS

Safety Science xxx (xxxx) xxx–xxx

Contents lists available at ScienceDirect

Safety Science

journal homepage: www.elsevier.com/locate/safety

Automated detection of grade-crossing-trespassing near misses based on computer vision analysis of surveillance video data

Zhipeng Zhang[a], Chintan Trivedi[b], Xiang Liu[a,*]

[a] Department of Civil and Environmental Engineering, Rutgers, The State University of New Jersey, United States
[b] Department of Computer Science, Rutgers, The State University of New Jersey, United States

ARTICLE INFO

Keywords:
Grade crossing
Safety
Near miss
Artificial intelligence
Video analytics

ABSTRACT

Grade-crossing trespasses are one of the greatest sources of injuries and fatalities on railways. While there is a wealth of data regarding grade-crossing accidents, near misses (or precursor events) associated with unsafe trespassing on railroad tracks are not reported, and therefore a comprehensive dataset is unavailable. This paper presents a Computer Vision (CV) algorithm to automatically detect trespassing near misses based on surveillance video footage of railway-road grade crossings. The CV algorithm is designed to be robust under changing lighting conditions over the course of the day-night cycle and works well under varying weather conditions. The algorithm is currently implemented based on data from one grade crossing in New Jersey. With minimal configuration changes, the algorithm can be adapted to various other grade crossings. Ultimately, the CV methodology can support data-driven grade-crossing near-miss risk analysis and contribute to proactive safety improvements at grade crossings.

1. Introduction

Highway-railroad grade crossings are intersections where a railroad line crosses a highway. According to US Federal Railroad Administration (FRA) in 2015, there were approximately 210,000 highway-rail grade crossings in the nation, of which 61% were publicly approachable (FRA, 2017a). Each highway-rail grade crossing presents a potential hazard to highway users and train crews. The possible dangers, such as collisions between highway vehicles and trains, can result in property damage, casualties, or even the release of hazardous materials (Chadwick et al., 2014). Collisions between highway vehicles or pedestrians and trains have been the greatest source of injuries and fatalities in the railroad industry since 1996 (FRA, 2017a). From 2006 to 2015, there were over 22,000 highway-rail grade-crossing incidents, resulting in 2717 highway-rail deaths and 9595 highway-rail injuries (FRA, 2017b). Therefore, there is a significant need for safety advancements of highway-rail grade crossing.

Considerable research has been conducted on highway-rail grade-crossings accidents, most of them aiming to reduce accident frequency and/or severity (e.g. number of fatalities and injuries). Observed accident data is useful for safety research; however, a far greater number of risk-prone events (near misses or precursors) occur in which there is no collision or loss of life but which might *possibly* have contributed to such accidents. These events did not result in fatalities, injuries, or property damage to either highway users or trains, but had the propensity to do so if they repeatedly occur. This paper focuses on the near misses due to trespassing at highway-rail grade crossings, which are defined as those incidents where roadway users are found to be in violation of existing laws related to grade crossings. Grade-crossing near misses may precede costly accidents but are not well documented because they did not result in immediate harm. The analysis of near misses can provide additional insight that may contribute to grade-crossing safety improvement. However, very little prior research has focused on near-miss-based grade-crossing data collection and corresponding safety analysis.

This knowledge gap motivates the development of this paper, which aims to develop an adaptable Artificial Intelligence (AI) framework for automatically collecting near-miss data based on surveillance video data. Surveillance cameras are currently deployed at many grade crossings in the United States. The existing video data can be used to track various objects such as pedestrians and wheelchairs that cannot be detected by conventional photoelectric, ultrasonic, and loop coil systems (Fakhfakh et al., 2010). In this study, closed-circuit television (CCTV) cameras are placed at railway grade crossings to monitor real-time activities. Due to limited storage capability, most railroads delete the raw video data every one or two months after relevant data has been collected. To preserve useful near-miss information from these unused big data sources, we collaborated with one rail agency in New

* Corresponding author.
E-mail address: xiang.liu@rutgers.edu (X. Liu).

https://doi.org/10.1016/j.ssci.2017.11.023
Received 13 August 2017; Received in revised form 17 November 2017; Accepted 26 November 2017

Jersey to develop an AI-aided analytic platform for extracting near-miss information from video data. Ultimately, our aim is to build a unique, expanding near-miss database for proactive risk management at grade crossings in the future.

2. Literature review

This study examines the automated detection of near misses due to highway-rail grade-crossing trespassing using AI as a support technology. Three major fields of research are relevant to this project: grade-crossing safety, near misses, and artificial intelligence algorithms for video analytics. The following subsections review some of the most relevant prior studies in each field.

2.1. Grade crossing safety

Considerable efforts have been made to quantify the frequency and severity of highway-rail grade-crossing incidents. The U.S. Department of Transportation (USDOT) Accident Prediction Model is a commonly used model in the U.S. to predict the number of collisions occurring at grade crossings, given specific highway and railroad conditions (Faghri and Demetsky, 1986; FRA, 2007). Saccomanno et al. (2004) performed Poisson regression and Negative Binomial Regression to achieve higher prediction accuracy and found that train speed, exposure, surface width, and number of tracks were significant factors in collision prediction models. Statistical models were developed to analyze the impact of a grade crossing collision on highway traffic (FRA, 2007; Hu et al., 2010) as well as rail traffic (ADL, 2010; Hu et al., 2010; Chadwick et al., 2012). Furthermore, research has also been conducted to examine passenger train car crashworthiness (Simons and Kirkpatrick, 1999; Tyrell et al., 2006).

In addition to the analysis of accident frequency and severity at highway-rail grade crossings, a clear understanding of driver behavior and the identification of the human factors can contribute to the development of better accident-prevention strategies. Caird (2002) summarized a taxonomy of accident contributors in six categories: unsafe actions, individual differences, train visibility, passive signs and markings, active warning systems, and physical constraints. Although each of these issues requires slightly different approaches to reduce undesirable occurrences, engineering, law enforcement, and the education of the public about the risks are three generally effective approaches to the improvement of grade-crossing safety. For more concrete details regarding human factor causes, Chadwick et al. (2014) provided a comprehensive overview of grade-crossing risk research in the United States. The frequency of grade-crossing accidents has declined 80% over the past 20 years, despite an increase of both train and highway traffic. Nevertheless, collisions between highway users and trains are still the greatest source of fatalities and injuries in the U.S. railroad industry (FRA, 2017a). Hence, it is crucial to better understand noncompliant behaviors at grade crossings, through both accident data and near-miss data.

2.2. Near-miss events in transportation safety research

Prior research has been conducted regarding near-miss identification in maritime and aviation sectors. For example, Zhang et al. (2015) proposed a method to identify potential ship-ship collisions using Automatic Identification System (AIS) data. High risk vessel encounters are evaluated by navigational experts to identify a potential near miss. In another study of bird collision hazard to aircraft, Klope et al. (2009) used digital avian tracking radars to automatically monitor and identify near-miss events. It shows that a combined dataset of actual bird-strike incidents and near-misses can provide risk managers with a more responsive metric in assessing the hazards over time than by using only the bird-strike dataset.

To date, almost all prior studies in the field of grade crossing safety have focused on using the reported highway-rail grade-crossing accidents in the available databases, such as the U.S. Federal Railroad Administration Rail Equipment Accident database or Grade Crossing Incident database. However, none of these databases contain near-miss events that did not cause actual damage. Several studies, however, indicated the importance of near-miss data for railroad safety and risk analysis. For example, Wright and Van der Schaaf (2004) stated that accidents are (fortunately) too few in number to support decision making about investing in safety improvements, while the use of near misses is able to dramatically increase the available data to counteract this problem. In addition, most accidents, such as grade-crossing collisions, were preceded by near-miss events and accident prevention should not wait until an accident actually occurs. Nevertheless, a large proportion of published studies have continued to focus on accidents only, while very few included near-miss data. The primary reason behind this deficiency may be the lack of sufficient, available near-miss data. San Kim and Yoon (2013) collected near-miss data from 80 rail accident investigation reports published by an independent accident investigation organization in the UK; however, it is time-consuming to manually judge the type of outcome (accident, incident, or near miss) derived from the title or summary of each report. A similar method was used by Le Coze (2013), in which near misses were identified from published studies or books. Given the limitations of these previous studies, an efficient near-miss collection method with high accuracy is essential to support the study of near misses.

Grade-crossing near-miss data is largely unavailable. Lobb (2006) pointed out that inadequate reporting has hindered the understanding of grade crossing collisions. Understanding and affecting driver behavior and human factors has contributed to about 70% of the decrease in the number of collisions and fatalities at grade crossings over the past 30 years (Mok and Savage, 2005). Behavioral models, such as Signal Detection Theory (SDT), have been used to model motorists' stopping behavior at grade crossings (Richards and Heathington, 1990; Yeh et al., 2009). A FRA report (Raslear, 2015) also discussed the application of SDT in motorist behavior at grade crossings, but also claimed that the models have only been tested against limited data and more field studies about motorist behavior and train arrival times are necessary. Thus, it is crucial to create an algorithm to detect and record grade-crossing near-miss data, so that it can provide sufficient data to test and refine the models within the framework of behavioral risk analysis.

2.3. Artificial Intelligence for video analytics

AI technologies, particularly computer vision, is used to train computer program to "understand" images and videos and identify useful features. AI techniques in computer vision include background subtraction (Shah et al., 2007; Sheikh et al., 2004; Ramesh, 2003; Elgammal et al., 2000; Fakhfakh et al., 2010; Zivkovic, 2004), image segmentation (Sheikh et al., 2004; Elgammal et al., 2000; Salmane, et al., 2015; Sen-Ching and Kamath, 2004), and trajectory prediction using Kalman Filtering (Salmane, et al., 2015; Sen-Ching and Kamath, 2004; Patel and Thakore, 2013). Patel and Thakore (2013) have reviewed the technique of foreground detection by learning the background template of the frame and then applying a background subtraction technique to identify moving objects. After that, objects are detected through segmentation of the isolated pixels, and their movement is tracked using Kalman Filtering. This has proven to be an effective technique for handling this task in computer vision. However, to implement this technique for the detection of highway-rail grade-crossing near-miss events, it is necessary to address several challenging problems in this real-world scenario, such as severe weather conditions and isolation of the train from the rest of the moving objects.

Very little prior work has used these techniques on grade-crossing safety analysis, except the following few studies. Shah et al. (2007) presented a background subtraction approach that detects moving objects on railway tracks, with a high accuracy of object detection. However, the system only performed during the day and turned off automatically when the illumination fell below a certain predefined level. Similarly, Salmane et al. (2015) implemented a security surveillance system in order to detect and evaluate abnormal situations induced by users in grade crossings based on the aforementioned algorithms. This intelligent system allows for automatic recognition and evaluations of noncompliance behaviors in grade-crossing environments. However, that study did not discuss the isolation of the train from other vehicular movement if their algorithm was to be used for detecting grade crossing near misses. Furthermore, the algorithm is expected to run in real time, but there is no information regarding the speed of the algorithm's performance on stored videos. Few previous studies addressed all crucial practical issues related to grade crossing near miss detection.

In this paper, to automate the detection of near misses at grade crossings, a customized AI algorithm is developed and validated using video data provided by one rail agency in New Jersey. The video data contains various traffic and environmental scenarios for the AI algorithm to "see", "understand," and "detect" near misses associated with unsafe trespassing. The collected near-miss data will be used to gain a better understanding of the circumstances and factors for various types of potential collision incidents at grade crossings.

3. Data source and nomenclature

3.1. Data

For the purpose of designing and testing this algorithm, previously captured footage is stored and analyzed. The detection technology can be adapted to real-time video footage with modifications in the future. The video files from the CCTV camera are stored in a proprietary format called Nomad Voice File (NVF). In order to gain access and read these videos for processing, these videos are converted to standard MP4 or AVI formats. This conversion process was performed using the conversion tool provided by the data supplier. Also, to limit the file size of the extremely long videos, the frame rate is reduced to 10 frames per second (fps). The video has been read in the RGB (Red-Green-Blue) color format in the following implementation. Both daytime (Fig. 1a) and nighttime (Fig. 1b) footage is included in the video data.

3.2. Nomenclature

3.2.1. Region of interest

The region of interest (ROI) is defined as the part of the road and rail intersection where roadway traffic is prohibited to enter during the stop signal. In terms of the detection algorithm, this region is defined as an enclosed polygon within every frame of the video, as shown in Fig. 2(a). In addition to the ROI in real frame pictures, Fig. 2(b) displays the ROI in a binary mask that will be used in the following processes.

3.2.2. Grade-crossing trespassing near miss

A grade-crossing-trespassing near miss is defined as an incident in which drivers/pedestrians/cyclists are found to be in violation of existing laws related to grade crossings. If warning signals (particularly the red signal warning roadway users) are active and yet there are vehicles or pedestrians inside the ROI during this time period, although there is no actual injury or damage, a near-miss event has occurred. Such events could lead to potentially hazardous situations or accidents.

3.2.3. Algorithm reading frame rate (r_1)

The reading rate of the algorithm (r_1) is defined as the time interval between the times the algorithm checks the video for active stop signals. For example, if the reading rate is 10 s, the video will check for active signals every 10 s rather than checking every frame. This is to avoid unnecessary processing of video frames to accomplish the identification task.

3.2.4. Skipping time frame (r_2)

The skipping time frame (r_2) is used to speed up processing based on train schedule. It is defined as the time interval that can be skipped during video processing immediately after a stop signal ends. The length of this time interval (r_2) depends on the train schedule, or more concretely, it should be smaller than the minimum possible time gap between two trains crossing a certain grade crossing based on the local train schedule. For example, for the chosen grade crossing, it has been observed that no stop signal is followed by another for at least 10 min. Hence, the value of the skipping time frame (r_2) in this study is set at 5 min. It is a rational parameter in most cases, especially for tracks that are operated by freight railroads or main passenger train lines. Some passenger train lines may need a revised parameter for the skipping time frame (r_2).

4. Algorithmic framework

A generalized AI algorithm was developed here to detect near misses from video footage of a grade crossing. This AI algorithm reads the video file, looks for a red signal, processes the image (details will be presented later), and evaluates whether a near-miss has occurred. The basic procedural steps of the algorithm are shown in Fig. 3. Furthermore, the developed computer vision algorithm should be trained to test and verify its robustness. A training program for a computer vision application for railroads would require the development of an initial algorithm with established environmental parameters. This initial algorithm would analyze a training set of data in comparison to the known to ensure that the algorithm can properly "see" and "understand" the images of trains and pedestrians independently from the background. The computer vision algorithm should also be refined and tested with various weather conditions and diverse daylight conditions, such as dawn, day, dusk, and dark. After undergoing this training, testing and continual parameter refinement, a computer vision application can capture the images and moving paths of trains and highway users (e.g. cars, pedestrians, bicyclists) under diverse conditions. Then

(a) Daytime (b) Nighttime

Fig. 1. Example of footage at (a) daytime (b) nighttime.

(a) Real frame picture (b) Binary mask

Fig. 2. Region of Interest (ROI) as indicated by (a) the blue polygonal area on the tracks, and (b) binary polygon mask.

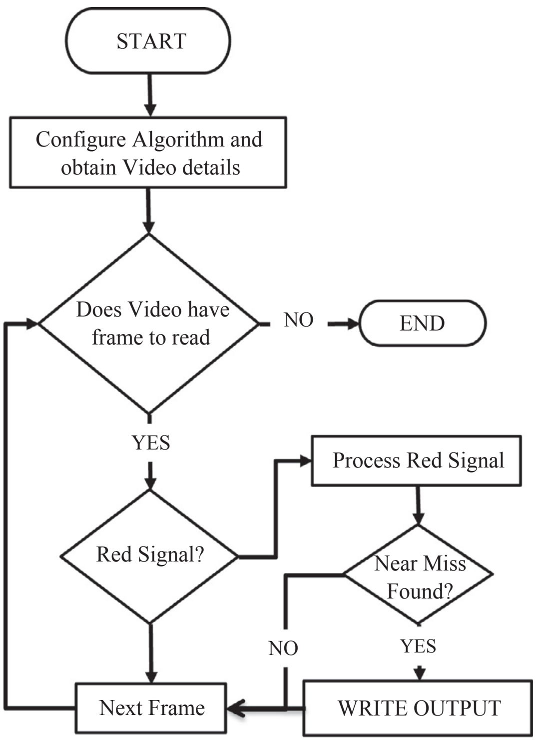


Fig. 3. General algorithm flowchart.

the developed algorithm can be used to automatically process video data, and compile the near miss information into a database for future safety study.

Further implementation details for every step of the procedure are as follows.

Step 1: Configure Algorithm and Video Settings

The first step is to extract the meta-data of the video to be processed, as well as necessary information, such as the duration, frame-rate, and the resolution of the video. Parameters of the algorithm are initialized in order to define the ROI and the location of the stop signal in the frame. In addition, the two hyper-parameters of the algorithm, algorithm reading frame rate (r_1) and skip time frame (r_2), are also initialized at this step.

Step 2: Read Video Frames

Next, the algorithm starts to read the video file from the first to the last frame of the video. During this reading, the prime objective is to detect whether the active signalized crossing light has been triggered. To optimize processing speed, these frames are not read in a continuous manner; instead, a frame-skip segment is conducted, which advances the reading in a fixed reading frame rate (r_1), 10-s intervals, and stops when a red light is detected. This is practical in this application, since a stop signal normally lasts for more than 10 s, so there is no risk of skipping any active signal. Most importantly, this efficient frame-skip algorithm supports adaptability to high frame rate surveillance video with acceptable analysis time.

Step 3: Check Stop Signal

Once a frame has been extracted, the algorithm checks whether the (red) stop signal has been triggered in that frame. This is achieved simply by checking the red pixel values in the small area of the frame where the signal is located. If a stop signal is detected, the algorithm performs a frame-by-frame check backward to find the initial activation of the stop signal, and then activates a subroutine procedure to detect near-miss events during that red signal.

Step 4: Background Learning Model

Before the process of near miss detection can begin, the algorithm must learn to subtract the background template established at the initial activation of the stop signal. Instead of using a single background template throughout the whole video, the algorithm learns a new template for every stop signal encountered throughout the video. This overcomes the challenge of gradually changing light levels over the course of the day. In addition, this ensures that the algorithm also takes into consideration temporary changes, such as short-term rainstorms, cars parked in the background, or other elements that may not remain static throughout the day.

Step 5: Detection of Moving Objects

Based on the background templates that are learned in the previous step, the algorithm initializes the detection of moving objects with the background subtraction technique. For every frame of video during the stop signal, the number of moving pixels is tracked and recorded. This detection activity continues until the stop signal turns off.

Step 6: Identification of Near-Miss Events

Finally, based on moving objects detected in the form of a set of moving pixels, the algorithm analyzes the recorded values of moving pixels to find near-miss events. The main challenge in this step is to separate the moving pixels of a near miss from those of a train or other static noise which is not included in the background template. The number of pixels that a train occupies in the foreground during a crossing is much larger than that of a pedestrian or vehicle, and accordingly, it is necessary to establish a threshold based on a set of training data. Then, if a segment is identified as and confirmed to be a near miss, all the frames within the duration of the red signal are extracted to an output video file for further review and study. Once the processing of one stop signal period concludes, the algorithm has a skip time frame (r_2), such as 5 min, and then continues to process the rest of the video from Step 1. This 5-min skip further reduces processing time and does not compromise accuracy, since no stop signal would occur again within such a short an interval in this case study. This parameter can be easily modified accordingly for different applications, such as a busy commuter-train crossing or short-line passenger trains.

5. Application and implementation

To illustrate and validate the feasibility of AI-aided video analytics, a customized algorithm was developed and implemented for the detection of trespasses resulting in near misses using data at one grade crossing in New Jersey. Near misses detected by this AI algorithm can provide researchers and policy makers with additional information for safety improvement. The following section details the parameters and process of using AI to automatically detect near misses from grade-crossing surveillance video data.

5.1. Stop signal detection

Indication of the on/off state of the stop signal is derived by focusing on the stop signal post. The stop signal consists of a left lamp and a right lamp emitting red light. Two small square windows in RGB color scale with a size of 7 × 7 pixels are extracted from the left signal lamp (L_{RGB}) and the right signal lamp (R_{RGB}) respectively, in one frame. L_{RGB} and R_{RGB} are converted to gray-scale using a weighted sum transformation $\Phi(x_{RGB}) = x_{GRAY}$ of the individual red, blue, and green color channels (Eq. (1)). As a standard color-to-grayscale conversion method, the weights for each color have been widely used in previous research (Bala and Braun, 2004; Grundland and Dodgson, 2007).

$$\Phi(x_{RGB}) = x_{GRAY} = (0.2989) \times x_R + (0.5870) \times x_G + (0.1140) \times x_B \tag{1}$$

where x_R, x_G, x_B, x_{GRAY} are color values for red, green, blue, and gray respectively.

Based on this transformation for each element of the two RGB windows L and R, the gray scale representations are defined as $L_{GRAY} = (l_{ij}) \in [0,1]^{7\times7}$ and $R_{GRAY} = (l_{ij}) \in [0,1]^{7\times7}$. Here, the on/off state for each signal lamp is identified based on the median values, which are represented as Θ_L in L_{GRAY} and Θ_R in R_{GRAY}.

From observation of the stop signal in the actual surveillance video data, it was noticed that the two lamps light up alternately with a rapid flicker between the two lamps. As a result, at any given moment, one of the two lamps is brighter than the other in intensity. Thus, the absolute intensity difference δ between the two lamps gives a positive indication of the state of the entire signal, $\delta = |\Theta_L - \Theta_R|$. This technique of evaluating intensity difference makes signal state detection robust under nighttime conditions, in which the headlights of vehicles bounce off both signal lamps, giving the false appearance that they are active. The state of the entire stop signal is chosen to be off ($\Theta_s = 0$) or on ($\Theta_s = 1$) depending on a particular threshold α:

$$\Theta_S = \begin{cases} 0, & \text{if}\, \delta < \alpha \\ 1, & \text{otherwise} \end{cases} \tag{2}$$

As shown in Fig. 4, the intensity difference is easier to detect in the daytime as compared to nighttime, when the lamps look extremely bright. Thus, accounting for both daytime and nighttime conditions, the intensity difference threshold of α has been fixed at 0.3 after trial and error. The practicability of this threshold is validated in the test of the detection algorithm.

5.2. Background learning

The first step in detecting moving objects throughout the video is to learn the pattern of static or non-moving objects in the video frames. This is done by learning the background template for a short time duration. The learning of background template B depends on N consecutive video frames, starting from the initial activation of the stop signal. Each video frame is divided into $l \times d$ RGB images and the n^{th}video frame in the set of video frames is denoted as $F(n) = (f_{ijk}(n)) \in [0,1]^{l\times d\times 3}$. The background template is learned by taking the mean value over N consecutive video frames for every pixel position in the background frame:

$$b_{ijk} = \frac{1}{N}\sum_{n=1}^{N} f_{ijk}(n) \quad \forall\ i,j,k \tag{3}$$

where $i \in \{1,2,...,l\}, j \in \{1,2,...,d\}$ and $k \in \{1,2,3\}$.

In this study, N (number of frames) is set at 200, which is equivalent to watching the first 20 s of the video to determine the positions of non-moving objects using the frame rate of 10 frames per second as described above. Fig. 5 shows a sequence of learned frames from video footage and the corresponding background template generated from them.

5.3. Daytime object tracking

In order to track moving objects during the daytime, Background Subtraction is applied to create a moving-object binary "mask" (M) for every frame. Background Subtraction here performs subtraction with binary thresholding for every pixel of the frame (F) and the background template (B) as shown here:

$$m_{ij} = \begin{cases} 1, & \text{if } \max(0, \Phi(f_{ijk} - b_{ijk})) > \lambda \\ 0, & \text{otherwise} \end{cases} \tag{4}$$

where $i \in \{1,2,...,l\}, j \in \{1,2,...,d\}$ and $k \in \{1,2,3\}$. The threshold in object tracking $\lambda = 0.3$ has been chosen.

After obtaining the noise-free mask (Fig. 6), the algorithm performs a check to see whether these moving objects lie inside the ROI. To achieve this, a pixel-wise "and" operation is conducted between the foreground mask M and the ROI mask P to obtain a binary image I indicating objects moving on the rail tracks:

$$I = (m_{ij} \cap p_{ij}) \in \{0,1\}^{l\times d} \tag{5}$$

5.4. Nighttime object tracking

The night time footage presents a big challenge in the detection of moving objects due to extreme illumination in various regions, largely arising from the light emitted from vehicle head lights. Another challenge is the static noise produced at nighttime due to the low quality of the CCTV camera. Both of these problems are apparent in Fig. 7(a).

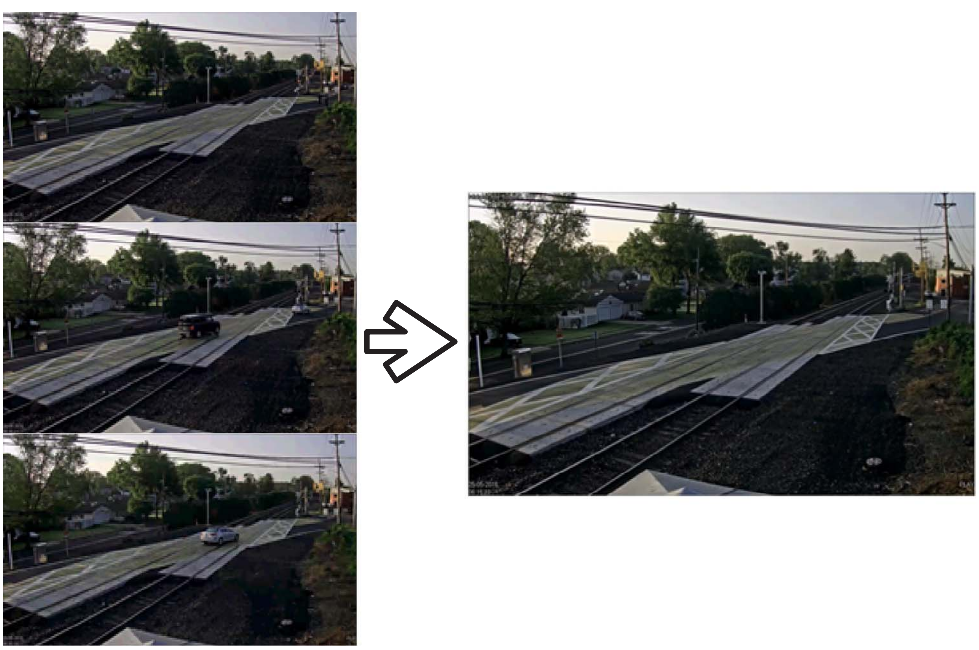
To solve these two difficulties, the algorithm makes use of the fact that light makes any object brighter when it falls upon it, thereby leading to an increase in that object's pixel intensity in the frame. Hence, to eliminate illumination, the mask is developed using a different process than the usual Background Subtraction approach. Here, all the pixels whose intensity is higher in comparison to the background model are eliminated from the mask, since they may correspond to illumination (Eq. (6)). While this approach loses some of the moving objects' resolution, sufficient information remains in the mask to detect vehicular movement in spite of the head lights. A comparison of this mask and the background subtraction mask are shown in Fig. 7(b) and (c).

$$m_{ij} = \begin{cases} 1, & \text{if } \max(0, \Phi(b_{ijk} - f_{ijk})) > \lambda \\ 0, & \text{otherwise} \end{cases} \tag{6}$$

5.5. Tracking under rainy conditions

Previous researchers have pointed out that one of limitations of their studies is the inability to operate the algorithm in rainy or other

Fig. 4. Intensity difference of stop signal during day and night.

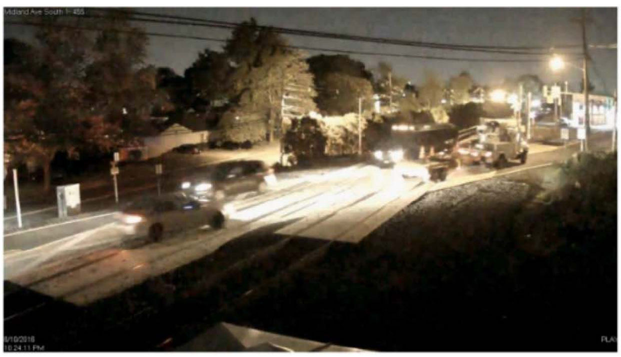


(a) Input frames (b) Output background

Fig. 5. Background learning (a) input frames to determine (b) output background.

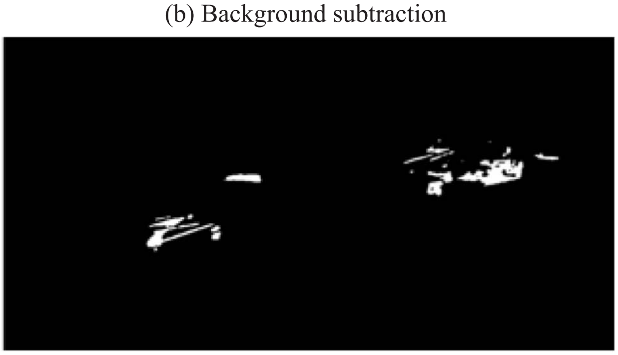
adverse weather conditions (Shah et al., 2007). When tracking objects move in rainy weather, the algorithm may incorrectly identify raindrops as moving objects that count as "trespassing". To obtain high detection accuracy, these raindrops must be filtered out of the moving-object mask. Accordingly, adapting principles similar to those used to eliminate light in nighttime tracking, minor modifications are made to the algorithm to eliminate raindrops from the mask. For light rain, the noise removal process of the algorithm will filter rain from the binary mask based on the usual threshold value of λ, 0.3. However, when considering heavy rain and strong wind, the threshold value λ must be adjusted to eliminate the extreme noise in the mask. In the currently available video database, there is no footage recording such weather conditions and $\lambda = 0.3$ is sufficient under all conditions. Fig. 8 shows the foreground mask for rainy conditions, and the moving pixels corresponding to rain drops have been eliminated using the noise removal process.

Fig. 6. Moving-Object Mask - Daytime.



(a) Moving objects at night time

(b) Background subtraction



(c) Illumination elimination

Fig. 7. Moving objects in (a) raw mask and masks obtained by (b) background subtraction and (c) illumination elimination.

5.6. Separation of the train from other moving objects

Another challenge of this application is to make a distinction between moving trains and uncompliant highway users in the mask. The percentage of pixels corresponding to the moving objects in the mask is used as the metric to achieve the distinction. Fig. 9 shows a certain binary mask obtained from a frame in which the train is passing. As evidenced from the mask, the train covers most of the frame. In other words, the percentage of pixels corresponding to this single moving object is significantly higher than that of highway users in the ROI, such as motorists, pedestrians, and bicyclists. Hence, it is assumed that during an active stop signal, the time frame containing considerable activity in the mask corresponds to a passing train.

5.7. Algorithm validation

The performance of the proposed grade-crossing near-miss detection algorithm was tested using a footage dated May 25, 2016. This is a testing database and was not used for model development. Two near-miss events of interest were successfully extracted and recorded by this algorithm. It took around 5 min of processing time to go through the 210-min segment extending from 5:30 AM to 9:00 AM. This means that the processing of the surveillance video takes roughly 2–3% of the total video duration to complete. However, the exact computational efficiency depends upon the number of active stop signals throughout the video. Less processing time is needed for fewer stop signal activations.

In terms of accuracy, there are four possible results: (1) an illegal trespass occurs and a detection is recorded (correct); (2) no illegal trespass occurs, but a detection is recorded (false positive); (3) an illegal trespass occurs, but there is no detection (false negative); and (4) there is no illegal trespass and no resulting detection (correct). To check the algorithm results with actual conditions, surveillance footage was reviewed manually and then compared with the output of AI detection. In this short-term case study, the AI algorithm detects all near misses accurately without any false positives or false negative.

5.8. Description of detected near misses

Two AI-detected near-miss events occurred within the active period of a single stop signal in the early morning. In the first near miss, before the arrival of the train, two pedestrians crossed inside the zone of danger while long arm gates were open and stop signals were activated (Fig. 10a). Five seconds after the two pedestrians crossed the track, the train arrived. In the second near miss, a bicyclist had stopped in front of the deployed arm gate and stop signal while the train passed this grade crossing. However, immediately after seeing the train had passed, this bicyclist crossed the tracks without waiting for the signal to be deactivated and without even checking whether there was a second train on this multiple-track territory (Fig. 10b).

These two near-miss events epitomize two types of highway user, pedestrian and bicyclist, and two typical non-compliance behaviors. More specifically, the two pedestrians timed the arrival of the train using their visual judgment and were confident in their ability to cross the track before the train arrived under the current railroad conditions. However, there may be considerable uncertainty when visually estimating train speed and time of arrival, especially the evaluation of train speed from their perpendicular perspective. Similar human error of misjudging speed of oncoming vehicle, was found by Stanton and Salmon (2009). In regard to the bicyclist trespassing, the second near miss illustrates the common assumption that no other train will pass directly after another, despite the presence of multiple tracks and the continuation of active signals at the grade crossing. This assumption is not always true. Yeh and Multer (2008) highlighted the safety of drivers and pedestrians at grade crossings with multiple tracks, and concluded that significantly more crashes appeared to occur with multiple tracks than that at crossings with single tracks. Both of these near misses represent risky non-compliance behaviors and can potentially cause catastrophic consequences, as evidenced in past accident data (Ogden, 2007).

6. Counting traffic exposure and calculating near-miss rate

As shown in the two aforementioned near-miss examples, the

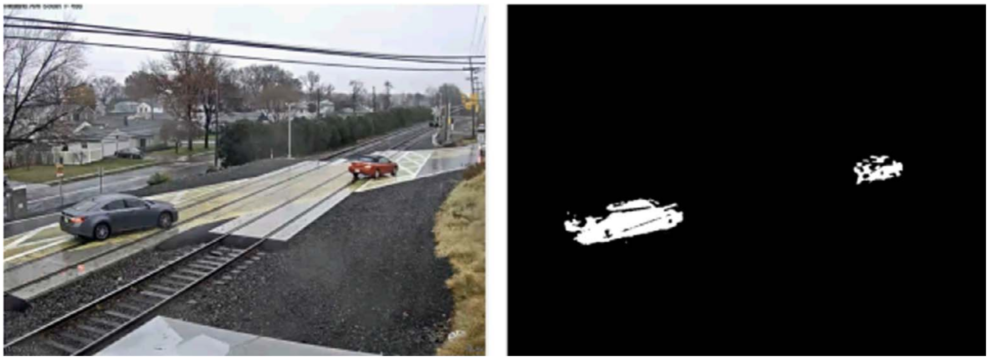


Fig. 8. Moving-Objects Mask in Rain.

algorithm detects and tracks noncompliance behaviors based on the surveillance video. Apart from the detection of near-miss events, it is also essential to count traffic objects, such as vehicles and pedestrians, in the output video. The tally results can be used to determine traffic volume and calculate the rate of near-miss events, which is equal to the number of near misses normalized by traffic volume. In train accident studies, accident rate is a major empirical model and can be implemented for high-level rail operational safety analysis (Liu, 2016). Similarly, near-miss event rates can provide grade-crossing safety conditions and comparisons between crossings involving various traffic volumes. In addition, it can act as an essential standard to evaluate potentially high-risk time periods or locations, where special instruction and prevention efforts can be implemented. Object Path-Tracking is one commonly used vision-based counting approach (Barandiaran et al., 2008; Cong et al., 2009; Kocamaz et al., 2016). It is adapted here to estimate the number of traffic objects in highway-rail grade crossing areas over a specific time period. To distinguish the types of traffic objects (pedestrian or vehicle), bounding boxes can be assigned to moving objects using a series of tracking points and employing a combination of Background Subtraction, Kalman Filtering (Coifman et al., 1998; Chen, 2015), and the Hungarian cost optimization algorithms (Kuhn, 1955; Torresani et al., 2008). The parameters of these bounding boxes (height, width, and aspect ratio) and a confidence score based on series of points are two metrics capable of distinguishing vehicles from pedestrians. For example, a relatively larger bounding box or a score above a certain threshold indicates a motor vehicle (Fig. 11). Nevertheless, there are several common technical challenges, involving occlusion caused by multiple objects, maintaining the identity of a moving object across multiple frames given low camera quality, etc., which call for future research using more video data.

For illustration, we test a similar AI algorithm to count traffic exposure (e.g. pedestrian, vehicle) from 5:30 AM to 9:00 AM on the same day described above. Taking pedestrians as an example, the traffic count algorithm shows that 25 pedestrians and 2335 vehicles crossed this grade crossing, both of which are almost identical with the authors' manual counting. There were only 8 bicycle passes during this period. Given this low traffic exposure, we did not calculate the near miss rate for bicycle-related trespassing, which can be studied in the future using more data. As described above, one near miss occurred due to pedestrians and no near miss resulting from vehicular traffic. In this paper, a group of people crossing the track together is treated as one trespassing activity. Based on these information, there are an average of 4 near-miss events per 100 pedestrians at this grade crossing. On average, for an individual, the estimated probability of unsafe trespassing is 0.04 assuming that the unsafe trespass for each individual entering the track region follows an independent, identical Bernoulli process. If we assume that the number of near misses approximately follows a Poisson distribution, the probability of any possible number of near misses can be estimated as $p(n) = \frac{4^n e^{-4}}{n!}$, where n is number of near misses per 100 pedestrians. For example, if 100 pedestrians cross the track within the studied 210-min interval, the probability of having at least 10 near misses is around 0.0081 ($1-\sum_{n=0}^{9}\frac{4^n e^{-4}}{n!}$). For all types of motor vehicles, near misses appear to be rare events, and so far there was no recorded occurrence within the traffic exposure covered by this study. For situations where there are zero near-miss events within the study interval, Quigley and Revie (2011) estimated that the rate of this event occurring can be simply approximated with $\frac{1}{2.5n}$, where n is the number of trials. Under this estimator, the probability of a vehicle near miss here is 0.00017 ($\frac{1}{2.5 \times 2335}$) using the data in the studied period. Long-term, additional video data will advance the estimation of the near-miss rate. The locations or time periods with higher near-miss rates may need special preventative efforts or enforcement.

7. Potential use of near-miss data for grade-crossing risk analysis

The grade-crossing near misses detected and recorded in this study can add to research into the improvement of grade-crossing safety, since prior research has been based on reported accidents only. This algorithm can efficiently record near-miss outputs that can ultimately provide detailed characteristics of near-miss events for behavioral risk analysis (Raslear, 2015). Even though forms of noncompliance have been hypothesized, a lack of actual behaviors captured at the time of violation has hindered further studies. Yeh et al. (2009) have confirmed that grade crossing warning devices are effective in encouraging drivers to stop and behave more cautiously using the FRA Highway-Rail Grade Crossing Accident/Incident database. Nevertheless, a database involving accidents may provide limited information to clarify behavioral characteristics. In the study of Signal Detection Theory, Raslear (2015) similarly raised the concern that models have only been tested against

Fig. 9. Mask corresponding to a frame containing the train.

(a) Two pedestrian

(b) One bicyclist

Fig. 10. Two trespassing near-miss events detected by the AI algorithm: (a) two pedestrians and (b) one bicyclist.

Fig. 11. Traffic counting with bounding boxes and tracking.

limited data and there is a need for more information about motorist behavior. Thus, sufficient near-miss data detected through the proposed AI-based algorithm in this paper can contribute to a specific behavior risk analysis. Highway users' decision-making strategies modeled by behavioral risk analysis can then help clarify the factors that influence such noncompliant decisions and establish a comprehensive framework for evaluating the impact of proposed countermeasures.

The detected near misses at grade crossings can be used to support more detailed cognitive and sociological analyses. For example, some studies have found that prior behaviors can affect future decisions (Hastie and Dawes, 2010; Tinsley et al., 2012). Also, people's activities may depend, in part, on what others decide to do (Fitzpatrick and Mileti, 1991). For instance, one near miss identified in this study shows that two pedestrians chatted and unsafely cross the grade crossing together. There might be clustering of near misses attributable to sociological interaction.

The video data analyzed by computer vision technology can provide valuable information for prospective, in-depth, cognitive and sociological analyses. Researchers from multiple disciplines, including psychology, cognitive science, statistics, risk analysis, and transportation engineering, may collaborate to better understand human behaviors and root causes of near misses. Furthermore, it would be interesting to better understand the relationship between near misses and accident occurrences. If this relationship exists, near miss might be used as surrogate safety measure for evaluating grade crossing safety risk and thus prioritizing risk mitigation resources (Tinsley et al., 2012).

All these prospective analyses can shed new light on grade crossing safety improvement through Engineering, Education or Enforcement. For example, if certain types of grade crossing configuration characteristics are correlated with more near misses, engineering improvement strategies might be identified. Also, if near misses cluster in certain locations or time periods, enforcement may be optimally arranged to prevent them. Also, the near miss video data can be used to educate the public about the risk of unsafely trespassing grade crossings.

8. Conclusion

There is an increasing amount of surveillance video data available from railroads that provide opportunities for near-miss-driven safety analysis. This paper presents an AI-aided computer vision technique that is able to automatically identify and collect data regarding near-miss events due to unsafe trespassing of highway-rail grade crossings. This grade-crossing detection algorithm is capable of processing video data in a reasonably short period of time with adequate accuracy in the real-world scenario. Moreover, the near-miss detection can efficiently work under various weather and visibility conditions. The automated collection of near-miss data supported by the AI technology in this paper can be used for the development of a grade-crossing near-miss database and provide opportunities to study and improve grade-crossing safety. The near-miss database can be used for behavioral risk analysis and the development of risk-informed studies, to prevent the occurrence of risk-prone behaviors and resultant accidents. Ultimately, an interdisciplinary safety study between psychology, statistics, and railway engineering can contribute to insightful risk reduction strategies and significantly improved the safety at highway-rail grade crossings.

9. Future work

In the future, the near-miss detection algorithm can be trained and adapted based on the increasing duration of the surveillance video data covering more environmental conditions and various noncompliant crossings. The next step is to establish a significant near-miss database that includes all grade-crossing near misses captured by the detection algorithm based on multiple grade crossings in several months. One area of future research is to better understand human behaviors and root causes of near misses at grade crossings using a nexus of multidisciplinary approaches, such as cognitive and sociological theory, transportation engineering and statistical risk analysis. In addition, the AI algorithm can be adapted to other relevant areas in railroad safety research, such as in-cab video analysis for distraction detection or security surveillance in railway stations. Forthcoming applications of AI algorithms can be developed in support of identifying precursors and prioritizing risk mitigation strategies related to engineering, education and enforcement.

Acknowledgement

The authors are partially funded by the USDOT Federal Railroad Administration at the time of writing this paper. However, all the views and analyses are solely our own. The authors thank Mr. Kshitij Shah for his assistance.

References

Arthur D. Little (ADL), 2010. Development of the All Level Crossing Risk Model: A History, 1993-2013. United Kingdom.

Bala, R., Braun, K.M., 2004. Color-to-grayscale conversion to maintain discriminability.

In Proc. of SPIE Vol (Vol. 5293, p. 197).
Barandiaran, J., Murguia, B., Boto, F., 2008. Real-time people counting using multiple lines. In International Workshop on Image Analysis for Multimedia Interactive Services, pp.159–162.
Caird, J.K., 2002. A Human Factors Analysis of Highway-Railway Grade Crossing Accidents in Canada. Transportation Development Centre, Montreal, Canada.
Chadwick, S.G., Saat, M.R., Barkan, C.P., 2012. Analysis of factors affecting train derailments at highway-rail grade crossings. In Proceedings of Transportation Research Board 91st Annual Meeting.
Chadwick, S.G., Zhou, N., Saat, M.R., 2014. Highway-rail grade crossing safety challenges for shared operations of high-speed passenger and heavy freight rail in the US. Safety Sci. 68, 128–137.
Chen, C.H. (Ed.), 2015. Handbook of pattern recognition and computer vision. World Scientific.
Coifman, B., Beymer, D., McLauchlan, P., Malik, J., 1998. A real-time computer vision system for vehicle tracking and traffic surveillance. Trans. Res. Part C: Emerg. Technol. 6 (4), 271–288.
Cong, Y., Gong, H., Zhu, S.C., Tang, Y., 2009. Flow mosaicking: Real-time pedestrian counting without scene-specific learning. In Computer Vision and Pattern Recognition, 2009. CVPR 2009. IEEE Conference on (pp. 1093–1100). IEEE.
Elgammal, A., Harwood, D., Davis, L., 2000. Non-parametric model for background subtraction. Computer Vision—ECCV 2000, 751–767.
Faghri, A., Demetsky, M.J., 1986. Evaluation of methods for predicting rail highway crossing hazards (No. VHTRC86-R32).
Fakhfakh, N., Khoudour, L., El-Koursi, E.M., Bruyelle, J. L., Dufaux, A., Jacot, J., 2010. Background subtraction and 3D localization of moving and stationary obstacles at level crossings. In Image Processing Theory Tools and Applications (IPTA), 2010 2nd International Conference on (pp. 72–78). IEEE.
Federal Railroad Administration (FRA), 2007. Railroad-highway Grade Crossing Handbook-revised, second ed.
Federal Railroad Administration (FRA), 2017a. Highway-Rail Grade Crossings Overview. < https://www.fra.dot.gov/Page/P0156 > (accessed February 2017).
Federal Railroad Administration (FRA), 2017b. Accident Data as Reported by Railroads. < http://safetydata.fra.dot.gov/OfficeofSafety/publicsite/on_the_fly_download.aspx > (accessed March 2017).
Fitzpatrick, C., Mileti, D.S., 1991. Motivating public evacuation. Int. J. Mass Emergencies Disaster 9 (2), 137–152.
Grundland, M., Dodgson, N.A., 2007. Decolorize: fast, contrast enhancing, color to grayscale conversion. Pattern Recogn. 40 (11), 2891–2896.
Hastie, R., Dawes, R.M., 2010. Rational choice in an uncertain world: The psychology of judgment and decision making. Sage.
Hu, S.R., Li, C.S., Lee, C.K., 2010. Investigation of key factors for accident severity at railroad grade crossings by using a logit model. Safety Sci. 48 (2), 186–194.
Klope, M.W., Beason, R.C., Nohara, T.J., Begier, M.J., 2009. Role of near-miss bird strikes in assessing hazards. Human-Wildlife Conflicts 3 (2), 208–215.
Kocamaz, M.K., Gong, J., Pires, B.R., 2016. Vision-based counting of pedestrians and cyclists. In Applications of Computer Vision (WACV), 2016 IEEE Winter Conference on (pp. 1–8). IEEE.
Kuhn, H.W., 1955. The Hungarian method for the assignment problem. Naval Research Logistics (NRL), 2(1–2), 83–97.
Le Coze, J.C., 2013. What have we learned about learning from accidents? Post-disasters reflections. Safety Sci. 51 (1), 441–453.
Liu, X., 2016. Analysis of collision risk for freight trains in the United States. Trans. Res. Record: J. Trans. Res. Board 2546, 121–128.
Lobb, B., 2006. Trespassing on the tracks: a review of railway pedestrian safety research. J. Safety Res. 37 (4), 359–365.
Mok, S.C., Savage, I., 2005. Why has safety improved at rail-highway grade crossings? Risk Anal. 25 (4), 867–881.
Ogden, B.D., 2007. Railroad-highway grade crossing handbook.
Patel, H.A., Thakore, D.G., 2013. Moving object tracking using kalman filter. Int. J. Comput. Sci. Mobile Comput. 2 (4), 326–332.
Quigley, J., Revie, M., 2011. Estimating the probability of rare events: addressing zero failure data. Risk Anal. 31 (7), 1120–1132.
Raslear, T.G., 2015. Using Signal Detection Theory to Understand Grade Crossing Warning Time and Motorist Stopping Behavior (No. DOT/FRA/ORD-15/02).
Ramesh, V., 2003. Background modeling and subtraction of dynamic scenes. In Computer Vision, 2003. Proceedings. Ninth IEEE International Conference on (pp. 1305–1312). IEEE.
Richards, S.H., Heathington, K.W., 1990. Assessment of warning time needs at railroad-highway grade crossings with active traffic control. Trans. Res. Record 1254, 72–84.
Saccomanno, F., Fu, L., Miranda-Moreno, L., 2004. Risk-based model for identifying highway-rail grade crossing blackspots. Trans. Res. Record: J. Trans. Res. Board 1862, 127–135.
Salmane, H., Khoudour, L., Ruichek, Y., 2015. A video-analysis-based railway–road safety system for detecting hazard situations at level crossings. IEEE Trans. Intell. Trans. Syst. 16 (2), 596–609.
San Kim, D., Yoon, W.C., 2013. An accident causation model for the railway industry: application of the model to 80 rail accident investigation reports from the UK. Safety Sci. 60, 57–68.
Sen-Ching, S.C., Kamath, C., 2004. Robust techniques for background subtraction in urban traffic video. In Electronic Imaging 2004, International Society for Optics and Photonics, 881–892.
Shah, M., Javed, O., Shafique, K., 2007. Automated visual surveillance in realistic scenarios. IEEE MultiMedia 14 (1).
Sheikh, Y., Zhai, Y., Shafique, K., Shah, M., 2004. Visual monitoring of railroad grade crossing. In Proc. of SPIE Vol (Vol. 5403, p. 655).
Simons, J.W., Kirkpatrick, S.W., 1999. High-speed passenger train crashworthiness and occupant survivability. Int. J. Crashworthiness 4 (2), 121–132.
Stanton, N.A., Salmon, P.M., 2009. Human error taxonomies applied to driving: a generic driver error taxonomy and its implications for intelligent transport systems. Safety Sci. 47 (2), 227–237.
Tinsley, C.H., Dillon, R.L., Cronin, M.A., 2012. How near-miss events amplify or attenuate risky decision making. Manage. Sci. 58 (9), 1596–1613.
Torresani, L., Kolmogorov, V., Rother, C., 2008. Feature correspondence via graph matching: Models and global optimization. Computer Vision–ECCV 2008, 596–609.
Tyrell, D., Martinez, E., Severson, K., Jacobsen, K., Parent, D., Priante, M., Perlman, A.B., 2006. Overview of a crash energy management specification for passenger rail equipment. In Rail Conference, 2006. Proceedings of the 2006 IEEE/ASME Joint (pp. 131–140). IEEE.
Wright, L., Van der Schaaf, T., 2004. Accident versus near miss causation: a critical review of the literature, an empirical test in the UK railway domain, and their implications for other sectors. J. Hazard. Mater. 111 (1), 105–110.
Yeh, M., Multer, J., 2008. Driver behavior at highway-railroad grade crossings: a literature review from 1990–2006 (No. DOT/FRA/ORD-08/03).
Yeh, M., Multer, J., Raslear, T., 2009. An application of signal detection theory for understanding driver behavior at highway-rail grade crossings. In Proceedings of the Human Factors and Ergonomics Society Annual Meeting (Vol. 53, No. 23, pp. 1776–1780). Sage CA, Los Angeles, CA, Sage Publications.
Zhang, W., Goerlandt, F., Montewka, J., Kujala, P., 2015. A method for detecting possible near miss ship collisions from AIS data. Ocean Eng. 107, 60–69.
Zivkovic, Z., 2004. Improved adaptive Gaussian mixture model for background subtraction. In: Proceedings of the 17th International Conference on Pattern Recognition, 2004. ICPR 2004. Vol. 2, IEEE, pp. 28–31.