Cycling near misses: A review of the current methods, challenges and the potential of an AI-embedded system

(Manuscript accepted in Transport Reviews Journal)

Mohamed R. Ibrahim[a]*

James Haworth[a]

Nicola Christie[b]

Tao Cheng[a]

Stephen Hailes[c]

a *SpaceTimeLab, Department of Civil, Environmental and Geomatic Engineering, University College London (UCL), London, UK*

b *Centre for Transport Studies (CTS), Department of Civil, Environmental and Geomatic Engineering, University College London (UCL), London, UK*

c *Autonomous Systems Group, Department of Computer Science, University College London (UCL), London, UK*

* Corresponding author. Address: Chadwick Building, Dept Civil, Environmental and Geomatic Engineering,
University College London, Gower Street,
London WC1E 6BT, UK.
mohamed.ibrahim.17@ucl.ac.uk, mibrahim2006@me.com

Acknowledgement

Cycling near misses: A review of the current methods, challenges and the potential of an AI-embedded system

Abstract - Whether for commuting or leisure, cycling is a growing transport mode in many countries. However, cycling is still perceived by many as a dangerous activity. Because the mode share of cycling tends to be low, serious incidents related to cycling are rare. Nevertheless, the fear of getting hit or falling while cycling hinders its expansion as a transport mode and it has been shown that focusing on killed and seriously injured casualties alone only touches the tip of the iceberg. Compared with reported incidents, there are many more incidents in which the person on the bike was destabilised or needed to take action to avoid a crash; so-called near misses. Because of their frequency, data related to near misses can provide much more information about the risk factors associated with cycling. The quality and coverage of this information depends on the method of data collection; from survey data to video data, and processing; from manual to automated. There remains a gap in our understanding of how best to identify and predict near misses and draw statistically significant conclusions, which may lead to better intervention measures and the creation of a safer environment for people on bikes. In this paper, we review the literature on cycling near misses, focusing on the data collection methods adopted, the scope and the risk factors identified. In doing so, we demonstrate that, while many near misses are a result of a combination of different factors that may or may not be transport-related, the current approach of tackling these factors may not be adequate for understanding the interconnections between all risk factors. To address this limitation, we highlight the potential of extracting data using a unified input (images/videos) relying on computer vision methods to automatically extract the wide spectrum of near miss risk factors, in addition to detecting the types of events associated with near misses.

Keywords - Cycling near miss, non-collision incidents, risk factors, machine vision, artificial intelligence

1. Introduction

Cycling has increased in popularity in Europe and elsewhere globally (Marco Dozza et al., 2017). Whether cycling for leisure or commuting, its benefits in terms of public health and the reduction of

environmental pollution have influenced planners and policy-makers to invest in cycling infrastructure (de Hartog et al., 2010; Juhra et al., 2012; Pucher et al., 2010; Steinbach et al., 2011). Globally, various policies, programmes and physical and non-physical interventions have taken place to promote cycling (Pucher et al., 2010; Savan et al., 2017). In the UK, Transport for London (TfL), for instance, has invested in many cycling infrastructure projects such as cycling superhighways, quiet ways, mini-Hollands and cycle hire schemes aimed at promoting a safer environment for people on bikes (TfL, 2018).

Even though the health benefit of cycling exceeds its risk (de Hartog et al., 2010), the risk for people on bikes remains high: *"Compared with car occupants and with regard to time spent traveling, cyclists were 8 times more likely to be injured, 12 to be hospitalized, 16 to be seriously injured, and 3 to be killed"* (Blaizot et al., 2013, p. 43). The experience of near misses can also add to the perception that cycling is dangerous. People on bikes in the UK are likely to face at least one near miss for every six miles of a commute as estimated by (Aldred & Crosweller, 2015). The fear of getting hit or falling whilst cycling hinders the wider adoption of cycling as a transport mode (Aldred, 2016; De Rome et al., 2014; Winters & Branion-Calles, 2017). In the UK, a survey of 244 cyclists and non-cyclists revealed that the perceived lack of safety is one of the main reasons given for not cycling (Gatersleben & Haddad, 2010). Furthermore, Sanders (2015) explained that individuals' concern regarding the risk of cycling and experience of near misses were the main barriers to cycling. In fact, near misses were often associated with several worries related to cycling in traffic, such inattentive drivers, drivers passing too close, being doored, drivers driving too fast, aggressive drivers, or being cut off by a turning driver (Sanders, 2015).

While road casualties are a burden to individuals and society, tracking and recording collisions that involve cyclists remains a challenge. While a number of studies have focused on analysing collisions and fatal incidents (Beck et al., 2016a; Bíl et al., 2010; de Geus et al., 2012; De Rome et al., 2014;

DiGioia et al., 2017; Imprialou & Quddus, 2017; Juhra et al., 2012; Loo & Tsui, 2010; Orsi et al., 2014; Pai, 2011; Shinar et al., 2018; Teschke et al., 2014), they only cover part of the picture. In 2018 in Great Britain, according to police reported casualty data, 17,451 cyclists were either seriously or slightly injured and 99 cyclists were killed (Department for Transport, 2019). However, this data may exclude many cases that involved injured cyclists because cyclist casualties are only reported by the police if this injury includes a crash with a moving vehicle (Winters & Branion-Calles, 2017).

In general, it is challenging to quantitatively analyse the risk of cycling due to the low number of recorded incidents (Aldred, 2018; De Rome et al., 2014), or the impact of the reporting bias in road crash data, in which the less severe the crash, the higher the probability of under-reporting it (Abay, 2015). On the other hand, although many incidents may not result in a hospital visit or being reported to the police, people on bikes still report frequent situations where they need to take direct action to avoid a collision or feel destabilised. Cycling crashes are initiated by 'near miss situations that are not avoided' and therefore result in a crash. By using this analogy, if data on these near misses can be recorded then they can provide a rich source of information with which to study cyclists' crash risk and identify the factors that are most associated with them.

In this paper, we review the literature on cycling near misses and demonstrate that, while many near misses are a result of a combination of different factors that may not be transport-related, the current approaches to tackling these factors may not be adequate to fully understand the genesis of a near miss. Here, we explore the potential of extracting data of different disciplines using a unified input (images/videos) that relies on computer vision methods to automatically extract the wide spectrum of risk factors that may cause potential risk for people on bikes.

This review aims to provide a resource for planners, policy-makers and researchers by 1) defining near misses, their types and their risk factors, 2) reviewing the main methodologies of recording and analysing near misses, and their applicability and limitations, 3) showing a summary of the variation of

near misses and their limitations in understanding the stated issue, 4) introducing a new potential framework understanding near misses through machine vision, in which models can be utilised to extract risk factors and infer near misses, 5) paving the way for developing AI-related automated systems that could be used to tackle crash risk by focusing on near misses, in which we highlight the key enabling technologies and research directions.

The remainder of this review is structured as follows: In section 2, the methodology of the review is described. In section 3, cycling near misses are defined, highlighting the different types. The risk factors and the studies that cover them are reviewed in section 4. Section 5 summarises gaps in our understanding. Section 6 introduces machine vision as an approach for understanding near misses and lastly, section 7 concludes and summarises the research.

2. Review methodology

We adopted a systematic review approach using PRISMA guidelines (PRISMA, 2015). Figure one shows the flow of the information through the different stages of the systematic review.

First, all manuscripts related to cycling near misses were gathered to-date. These manuscripts included peer-reviewed journal articles, governmental and non-governmental reports, and conference proceedings. They covered the different aspects of cycling near misses, methods used, and the risk factors identified. These manuscripts can be accessed from four search engines (Scopus, Google Scholar, and Web of Science) via a combination of 'cycling' or 'road 'with keywords in a Boolean expression such as : Cycling AND near AND miss*, cycling AND perceived AND risk*, perceived AND traffic and risk*, cycling AND near AND collision*, road AND conflict*, cycling AND risk and risk*. The total results for the specified combined terms are 556, 435, and 389 for google scholar, Web of Science, and Scopus respectively. After removing duplicates, the sample size reduced to 531. Second, records were first screened by title and abstracts which reduced the records to 325. The second phase of screening included reviewing each manuscript, which reduced the number to 189 manuscripts that focus

on the various types of conflicts between different road-users, including near misses. At this phase, manuscripts were filtered to exclude studies that involved collisions without addressing near misses were excluded, except where they were required to draw a baseline or lesson learned that could be beneficial for near miss studies (i.e. Beck et al., 2016b; Cho et al., 2009; Orsi et al., 2014). Studies that involved safety policies that address cycling without addressing near misses are excluded. Last, we reduced studies to 19 manuscripts that focus only on cycling near misses which we focused on analysing them in details (See Figure 1).
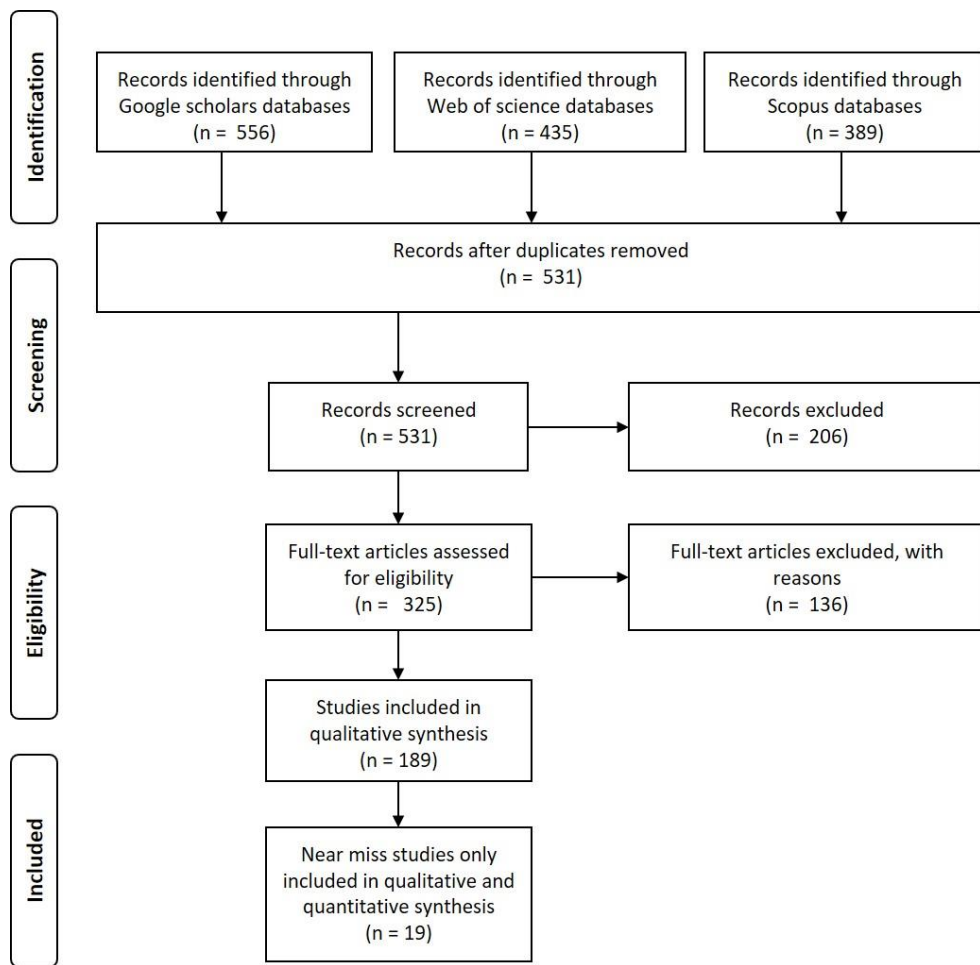
Figure 1: Flow chart for the screened records used for the systematic review

Identification

Records identified through Google scholars databases (n = 556)

Records identified through Web of science databases (n = 435)

Records identified through Scopus databases (n = 389)

Screening

Records after duplicates removed (n = 531)

Records screened (n = 531)

Records excluded (n = 206)

Eligibility

Full-text articles assessed for eligibility (n = 325)

Full-text articles excluded, with reasons (n = 136)

Studies included in qualitative synthesis (n = 189)

Included

Near miss studies only included in qualitative and quantitative synthesis (n = 19)

Figure 1: Flow chart for the screened records used for the systematic review

3. What is a cycling near miss?

3.1 Definition

There are different conceptions and terms that pinpoint the subject of a 'narrowly avoided collision'. It can be defined as: 'perceived crash risk' (Chaurand & Delhomme, 2013; Strauss et al., 2013), 'perceived traffic risk' (Sanders, 2015), 'near collision' (Johnson et al., 2010), or simply 'near miss' (Aldred, 2016; Poulos et al., 2012). While a 'cycling near miss' is a subjective term that may differ based on individuals' experiences and their perceptions of risk, in most cases it is defined as a situation in which a person on a bike was required to act to avoid a crash, such as braking, speeding, swerving or stopping. In some cases, the definition may be extended to include those events that caused the person on the bike to feel unstable or unsafe, such as a close pass or tailgating.

3.2 Types of cycling near miss

Different studies have categorised near misses in different ways. For instance, some studies focus on the conflicts between people on bikes and drivers (Beck et al., 2019), or people on bikes and pedestrians (Paschalidis et al., 2016), in which different types of risky situations are categorised. Another approach is to categorise near misses depending on the type of conflict; either a moving object or stationary (Nelson et al., 2015). The most comprehensive categorisation to date was introduced by Aldred & Goodman (2018), based on an empirical study that included 2586 diaries from a sample of 396 participants over two years (2014-2015). This study summarised the types of cycling near misses were summarised in eight groups. These groups are: 1) a close pass, 2) a near left or right hook, 3) someone pulling in or out, 4) a near-dooring, 5) swerve around an obstruction, 6) pedestrian steps out, 7) someone approaching head-on, or 8) tailgating. Their study found that close pass near misses were the most frequent incident type.

In figure 2, we aim to visualise the different types of cycling near misses addressed in the literature to better understand the road-users or objects involved and their potential impact on people on bikes. In doing so, we aim to highlight the areas that need to be covered for any future method to be comprehensive when addressing cycling near miss.
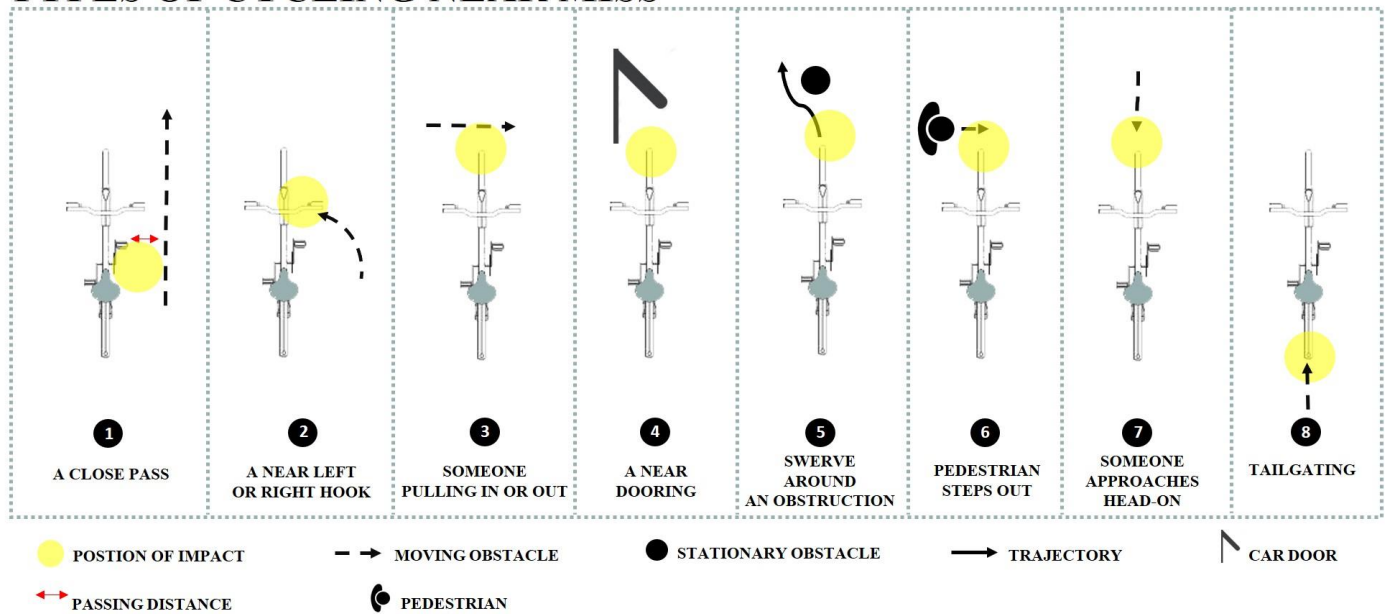
TYPES OF CYCLING NEAR MISS

Figure 2: Types of cycling near miss and their potential impact
Source: Created by the authors based on the literature review

4. Methods and materials used for understanding near misses

Studies have focused on different factors related to cycling near misses and have also used different data collection techniques and methods. These methods can be divided into three types of observational studies: 1) Observational studies relying on self-reported questionnaires, 2) Analysis of video of cyclist behaviour at specific sites, i.e. at an intersection, and 3) analysis of video from cameras used on their bike, the so-called naturalistic study.

4.1 Self-report studies

Self-report studies are the dominant design for most cycling near miss studies. The core element that that identifies a self-report study, in the scope of this research, is how data is gathered, not what types of data are gathered. In an observational study, data are gathered and statistically analysed to show associations and draw conclusions. There are three main ways in which data are collected for an self-reporting study: 1) using a self-reporting mechanism based on a questionnaire survey for a group of participants (Aldred & Crosweller, 2015; Chaurand & Delhomme, 2013; Fuller et al., 2013; Lawson et al., 2013; Paschalidis et al., 2016), or 2) using a self-reporting mechanism based on crowdsourcing platforms where data can be uploaded (Nelson et al., 2015; Poulos et al., 2012). Data gathered based on crowdsourcing have led to significant progress in mapping cycling ridership and safety measures (Jestico et al., 2016; Nelson et al., 2015). However, while observational studies can offer insights about the behaviour of people on bikes over a longer period, the data gathered is limited by potential biases such as over or under-representation of certain cyclist groups or the types of risk factors, in addition to limitation and biases due to manual labelling and processing based on the collectors' interpretations (Dozza & Werneke, 2014).

4.2 Video analysis at specific sites

In a site observational study, video streams for a given context are used that highlight certain safety issues. A focus group of cyclists or non-cyclists participate and observe these video streams to evaluate behaviours. Rather than focussing on near miss events, these studies ask participants to evaluate the level of risk or presence of hazards in the video stream. Vansteenkiste et al. (2016) used site observation to develop a hazard perception test for children, finding that children's reactions to, and interpretations of, hazards are less developed than adults. Lehtonen et al., (2016) asked frequent and infrequent riders to watch video clips and rate risk through a caution estimate, finding that more frequent riders identified a higher number of caution estimate rises. This indicates that awareness of risk increases with rider

experience. Such studies are important because they enable understanding of the differences in exposure to risk between certain groups of riders.

While site observations using cameras may overcome the limitation of interpretation found in self-report studies, the amount of data processing, specifically image processing, limits the scalability of this type of method. Additionally, multiple cameras are required at different positions to capture the entire environment and observe the dynamics of cycling behaviour and the interactions among the different agents (Dozza & Werneke, 2014). However, this approach can potentially enable the analysis of near misses at scale by leveraging the supply of CCTV cameras routinely installed on road networks, particularly in cities. Zangenehpour et al., (2016) used cameras to collect overview footage of intersections and analyse variations in vehicle-bicycle interactions in the presence/absence of cycle tracks. This work was facilitated by automated tracking of road-users using a computer vision approach, which could be applied to CCTV feeds (Zangenehpour et al., 2015).

4.3 Naturalistic study

The naturalistic approach is often perceived as one of the most reliable methods for understanding road-user behaviours and analysing risk factors (Marco Dozza, Bianchi Piccinini, et al., 2016; Marco Dozza & Werneke, 2014; K. Schleinitz et al., 2017). In this method, a group of participants ride instrumented bikes while carrying out their routine activities. Cameras and sensors are fitted to the bike or rider that collect data related to riding behaviour, the surrounding environment and the interactions with other road-users. The types of data gathered vary depending on the purpose of the research and the installed equipment and sensors. The first example of a naturalistic study was in Melbourne, Australia, where riders were given helmet-mounted cameras (Johnson et al., 2010). This camera position was chosen because it gave the rider's perspective and enabled their behaviour to be analysed (e.g. shoulder checks). While this approach enables risk factors to be analysed through manual analysis of the video data, quantitative features such as rider speed, geolocation and acceleration/deceleration cannot be

extracted. More recent studies have used instrumented bikes containing at least a video camera and a GPS device (Gustafsson & Archer, 2013). More sophisticated setups can contain units for inertial measurement, a signal button to record critical incidents and near misses, brake force sensors (Dozza & Werneke, 2014), or even LiDAR sensors for measuring the range of nearby objects such as passing vehicles (Beck et al., 2019). Naturalistic approaches have also been used to study differences in behaviour between pedal bike and e-bike users, which is an emerging area of concern for policymakers (K. Schleinitz et al., 2017).

The main advantage of the naturalistic approach is that detailed information is collected about the behaviours and actions of agents involved in an event, as well as the instantaneous features of the environment. By definition, the naturalistic approach also collects data in which no event occurred, which can be used as counterfactual events in a case-control framework. However, processing data collected in naturalistic studies is labour intensive and automated methods are required if such data are to be collected at scale.

5. Factors related to near misses and their impacts

Different methods have different advantages and limitations when it comes to including and analysing the wide range of risk factors. In general, cycling near misses are transport-related, however, their risk factors may or may not be transport-related (Aldred, 2016; Beck et al., 2016a; De Rome et al., 2014; Vanparijs et al., 2015). This creates challenges for methods currently used on the literature – either in extracting these factors or analysing them - to assess in a single study. Based on the literature, we categorise these factors into aspects related to visibility, physical conditions of the built environment, interaction among different agents (e.g. people or animals), and behavioural and psychological factors related to the cyclist.
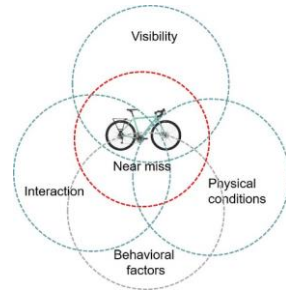
Figure 3: Types of factors related to cycling near misses

Figure 3 shows the interaction of these four categories that could lead to a near miss. Most of the factors belong to behavioural, physical, or visibility related factors. In many cases, near misses occur as a result of a combination of several independent factors represented by the 'interaction' section of the diagram.

Many studies focus on the impacts of one of the risk factors for a near miss while excluding other factors. Poulos et al. (2012) analysed near misses based on the type of cycling infrastructure, including pedestrian footpath, shared path, road with no bicycle lane, bicycle path, and road with a bicycle lane. Johnson et al. (2013) studied collision and near-collision characteristics based on people on bikes and open vehicle doors, highlighting how an open vehicle door could lead to frequent and serious injuries that could sometimes be fatal. Few studies, however, explore the holistic nature of the factors related to near misses.

Table 1 summarises the 19 studies identified in the literature in terms of methods used, near miss type and risk factors covered, broken down by the categories shown in figure 3. The categories are further subdivided according to specific risk factors. The interaction risk factors are combined with the

near miss type in column three. The extent to which each risk factor has been analysed in the literature by each method is discussed in sections 5.1 to 5.4, below.

Table 1: Current near miss literature and covered scope and risk factors

NEAR MISS STUDY	METHOD USED	TYPES OF NEAR MISS OR SCOPE OF THE STUDY	RISK FACTORS COVERED																			
			Visibility									Physical conditions					Behavioural factors						
			Weather conditions					Light conditions and time of the day				Built environment conditions					Cyclist characteristics				Safety equipment		
			Clear	Cloud	Rain	Snow	Fog	Dusk/ Dawn	Day-time	Night-time	lighting conditions	Road types	Road conditions	Existence of bicycle lane	Parked cars	Intersection	Age	Gender	Experience	Trip purpose	Helmet	Hi-vis clothing	Bike light
(Paschalidis et al., 2016)	Self-report studies	Conflicts with cars and pedestrian – 8 types	-	-	-	-	-	-	-	-	-	-	-	x	-	x	x	x	x	x	-	-	-
(Vansteenkiste et al., 2016)	Video analysis at specific sites	Visual awareness, environmental awareness, risk perception, and reaction time	-	-	-	-	-	-	-	-	-	x	-	-	-	x	x	-	-	-	-	-	-
(Gustafsson & Archer, 2013)	Naturalistic study	All types of conflicts divided by 17 types	-	-	-	-	-	x	x	x	-	x	-	x	-	x	x	x	-	x	-	-	-
(Nelson et al., 2015)	Self-report studies	Near miss with stationary or moving object or vehicle	-	-	-	-	-	x	x	x	x	x	x	-	x	-	x	x	-	-	x	-	x
(Branion-Calles et al., 2017)	Self-report studies / comparative study	Near miss with stationary or moving object or vehicle	x	x	x	-	x	x	x	x	x	x	x	x	x	x	x	x	-	x	x	-	x
(Katja	Naturalistic	Conflict with	-	-	-	-	-	x	x	x	-	x	x	x	-	-	x	-	-	-	-	-	-

13

Reference	Study type	Description																						
(Schleinitz et al., 2015)	study	vehicles, cyclists and pedestrians																						
(Aldred, 2016)	Self-report studies/ discussion	Eight types of near misses	-	-	-	-	-	-	-	-	-	x	x	x	x	x	x	x	-	x	-	-	-
(Chaurand & Delhomme, 2013)	Meta-analysis/ Comparative study	Bike-car interaction at road intersections	-	-	-	-	-	-	-	-	-	-	-	-	-	x	x	x	x	x	x	-	-
(Johnson et al., 2013)	Naturalistic study	One type of near misses- door opening	-	-	-	-	-	-	x	x	-	-	-	x	x	-	x	x	-	x	-	-	-
(Lehtonen et al., 2016)	Site observational study	Risk perception of bicyclists in a city environment	-	-	-	-	-	-	-	-	-	-	-	x	-	-	x	x	x	x	-	-	-
(Poulos et al., 2012)	Self-report studies	Near misses- undefined types	-	-	-	-	-	-	-	-	-	x	-	x	-	-	x	x	x	x	-	-	-
(M Dozza et al., 2012)	Naturalistic study	Pilot study showing six unique risky events.	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
(Marco Dozza & Werneke, 2014)	Naturalistic study	Critical events - 44	-	-	-	-	-	-	x	-	-	-	x	x	-	x	x	x		x	-	-	-
(Aldred & Crosweller, 2015)	Self-report studies	Near miss types – 8 types	-	-	-	-	-	x	x	x	-	-	-	-	-	-	x	x	-	x	-	-	-
(Sanders, 2015)	Self-report studies	exploration of perceptions of traffic risk between cyclists and drivers/ 4types of near misses	-	-	-	-	-	-	-	-	-	x	-	x	x	-	x	x	x	x	-	-	-
(Johnson et al., 2010)	Naturalistic study	Collision/ near collision for on-road cyclists	-	-	-	-	-	-	-	-	x	-	-	x	-	x	x	x		x	x	-	-

14

(Lawson et al., 2013)	Self-report studies	Perception of safety for cyclists	-	-	-	-	-	-	-	-	-	x	x	x	x	x	x	x	x	x	x	x	-
(Aldred & Goodman, 2018)	Self-report studies	Near miss types- 8	-	-	-	-	-	-	-	-	-	-	-	-	-	-	x	x	x	-	-	-	-
(Fuller et al., 2013)	Self-report studies	the impact of implementing a public bicycle share program on events of near misses	-	-	-	-	-	-	-	-	-	-	-	-	-	-	x	x	x	-	x	-	-

5.1 Behavioural aspects

Behavioural aspects play an important role in defining the various cycling styles that may influence the types and the frequency of near miss events. They can be subcategorised into three groups: 1) Individual characteristics, 2) trip characteristics, and 3) safety measures.

All of the 19 reviewed studies cover individual characteristics to a greater or lesser extent.

First, most of the cycling near miss studies focus on covering multiple factors related to socio-demographic aspects of people on bikes. It is noticeable from table 1 that self-report studies generally collect more demographic data than site observational and naturalistic studies, which is due to their design.

Safety measures related to individual behaviours are crucial for avoiding near misses. Johnson et al. (2014) studied the behaviours and perception of both drivers and people on bikes towards cycling safety. They found that drivers who also cycle are more likely to have a positive attitude towards cycling. They also highlighted the importance of rethinking driver education when overtaking people on bikes; considering head checks, buffer space, and adequate indications. Walker et al. (2014) found that the appearance of the person and the type of outfit they wore has an insignificant effect on the clearance distance drivers gave when overtaking people on bikes, contrary, police/video-recording jacket is the

only outfit that has a significant correlation with the passing proximities. Nelson et al. (2015) found that the frequency of near misses was higher when cycling without bike lights than when using front and backlights.

5.2 Physical conditions

Fifteen of the 19 studies referred to physical conditions. While there are different factors related to the built environment that influence the choice of cycling routes from a behavioural perspective (Broach et al., 2012), these are also factors that cause potential risk for encountering collisions or a near miss (Cho et al., 2009). As Aldred notes: *"The vast majority of near misses were judged potentially preventable by changes to road user behaviour and/or the cycling environment*'' (Aldred, 2016, p. 78). Therefore, physical conditions play an important role in either experiencing a near miss or avoiding one. Based on the literature review, we have subcategorised these into four groups: 1) Infrastructure, 2) surface conditions, 3) location and 4) surrounding context.

In terms of infrastructure, Parkin and Meyers (2010) found that in the presence of a cycle lane, drivers may drive within their marked lane with less consideration of ensuring a comfortable passing distance for people on bikes in the adjacent cycle lane. This has been confirmed in a recent study by Beck et al. (2019) on close pass events.

Several studies include factors related to surface conditions, such as wet, dry, well-maintained, or deteriorated surfaces (Aldred, 2016; Branion-Calles et al., 2017; Gustafsson & Archer, 2013; Nelson et al., 2015; Katja Schleinitz et al., 2015). Dozza et al. (2012) found that some near misses occurred when the condition was icy. Nelson et al. (2015) found that the near misses mostly took place on dry surfaces (64.6% of the total near misses) with no parking on the road (60.1% of the total near misses). However, these frequencies are based on the count of responses rather than the significance of the result, which may be due to the self-selection of either trip routes or time. The challenge remains in understanding how the individual factors combine with physical conditions to produce risk.

Different studies have focused on the location of cycling, highlighting a higher exposure to near misses at intersections (Branion-Calles et al., 2017). Strauss et al. (2013) analysed 650 intersections in Montreal in which cyclists were exposed to injury incidents. They found that more cyclists tended to suffer injuries at junctions but with a lower injury rate due to the non-linear correlation between injury occurrence and bicycle volume. They also highlighted that the frequency of cycling crashes is associated with the changes in the flows of both vehicles and bicycles based on injury data of people on bikes between 2003 and 2008 in Montreal, Canada. Additionally, crashes were more likely to happen at intersections that include a bus stop. The important role of the built environment was underscored, in which a change in conditions (i.e. presence of cycle lane, land use mix, presence of a school, etc.) is more likely to cause a direct impact on cyclist activity and safety.

There is still an absence of near miss studies that directly investigate the impact of the surrounding context that may densify the flow of traffic for cyclists, pedestrians, and vehicles in a certain location in cities, which may cause potential risk exposure for the people on bikes. This has been outlined by Vanparijs et al. (2015) in their review of studies related to exposure measurement. They reviewed the different methods to measure cycling exposure to incidents, including time, distance, and trips as exposure units. They showed that the lack of exposure data hinders the ability to draw significant conclusions, noting that analyses that do use exposure measures often neglect minor incidents, meaning near miss events are more likely to be missing as well. This makes it difficult to understand safety levels between different types of infrastructure, and the age categories of people on bikes.

5.3 Visibility-related conditions

Visibility related conditions play a role in crashes and near misses (Lacherez et al., 2013), particularly, factors related to 1) the time of the day, 2) weather conditions, and 3) the level of illumination, including the existence of glare. Of the reviewed studies, only two cover weather conditions, while nine cover lighting conditions or time of day. (Branion-Calles et al., 2017) use

secondary data sources to infer weather and lighting conditions, using time of sunrise/sunset and meteorological data at a single location in the two cities studied. This may fail to capture local variations in weather, street lighting and glare caused by the direction of travel. (M Dozza et al., 2012) highlights some factors related to six unique events but does not analyse risk factors.

There are a limited number of studies that included factors related to weather conditions to address their impact on the occurrence of near miss events. Branion-Calles et al. (2017) addressed weather conditions such as clear, cloudy, rain, and fog in analysing near misses. Based on descriptive analysis, they found that a higher frequency of near misses and crashes are reported when the weather is rainy, snowy or foggy compared with cloudy or clear conditions.

While, the time of the day could be a significant risk factor in cycling crashes (Johnson et al., 2013), many studies neglect the issue of time completely (Aldred, 2016; Aldred & Goodman, 2018; Chaurand & Delhomme, 2013; Fuller et al., 2013; Lawson et al., 2013; Lehtonen et al., 2016; Paschalidis et al., 2016; Poulos et al., 2012; Sanders, 2015; Vansteenkiste et al., 2016). Other studies use a binary classification of day and night, without considering more nuanced effects on the lighting conditions such as those caused by direct sunlight at dawn and dusk. For instance, Branion-Calles et al. (2017) studied near misses according to the time of either peak hours or off-peak hours. Gustafsson and Archer (2013) categorised time to 'morning (06:30-9:30)' and 'afternoon (15:30-18:30)', highlighting that more incidents and safety issues occur in the morning.

The impact of the lighting conditions -including glare- on near miss events remains under-investigated. Branion-Calles et al. (2017) mentioned lighting conditions as risk factors represented in two classes of day or night-time only. Dozza et al. (2012) mentioned the effect of glare on the quality of the video streams, without studying its impact on the frequencies of near misses. Even though Nelson et al. (2015) had looked at glare based on self-reporting data, this data only provides a subjective account

of the incident and glare may have been present but not mentioned. Therefore the data may not represent a reliable source of information to study the impact of glare.

5.4 Interaction between road-users

While near misses can involve a single individual riding a bike, they are often the result of interactions between the rider and other road-users, such as people in cars, people driving or other people on bikes. During a journey, a bike rider will have many safe interactions with other road-users, which makes it difficult to define those situations that lead to elevated risk. Therefore, quantitative studies in this area have tended to focus on a single type of interaction. A notable example is (Beck et al., 2019), who focussed on passing distance. Their study used a distance sensor attached to a bicycle to measure the range of passing vehicles. Using a passing distance of 1 metre at <= 60km/h (1.5m at >= 60km/h) informed by Australian legislation, they identified that 1 in 17 passing events was a close pass. As mentioned in section 5.2, they found that the presence of a bike lane was associated with closer passing distance. This type of study is important in identifying a particular type of risky behaviour, but the use of a 1 metre passing distance is somewhat arbitrary. For example, in the UK, Operation Close Pass uses 1.5 metres as a safe passing distance; if this figure was used in (Beck et al., 2019) it would change the interpretation of results. In general, there is no good evidence on what is safe under what circumstances, and as the authors, note: *"It is important to understand how cyclists' subjective experiences align with quantified passing distances"* (Beck et al., 2019, p. 259).

Some studies focused on a specific type of interaction that may result in a near miss if the safety measures are not considered properly. For instance, Dozza et al. (2016) provided an in-depth analysis of how drivers overtake bicycles during passing events. They found that manoeuvres, especially on rural roads, are often more critical since they happen at higher speeds (approx. 70km/h) with less time to avoid collisions (less than 2s) if critical or unforeseen events took place.

5.5 Combined factors

The various permutations of factors described above indicate the complexity of understanding how risk factors interact to cause near misses. Figure 4 shows a dendrogram of the different factors that may be involved. Factors highlighted in black are those identified from the reviewed studies, while those in blue are additional factors that could be considered. The number of potential factors illustrates that, theoretically, even if one of the individual factors is not statistically significant from a linear perspective, it may influence the occurrence of a near miss from a nonlinear perspective. Consequently, due to the potential of the existence of nonlinearity in near-miss research, the types of methods used to conduct near miss research may vary depending on which factors are addressed and how the data are collected.
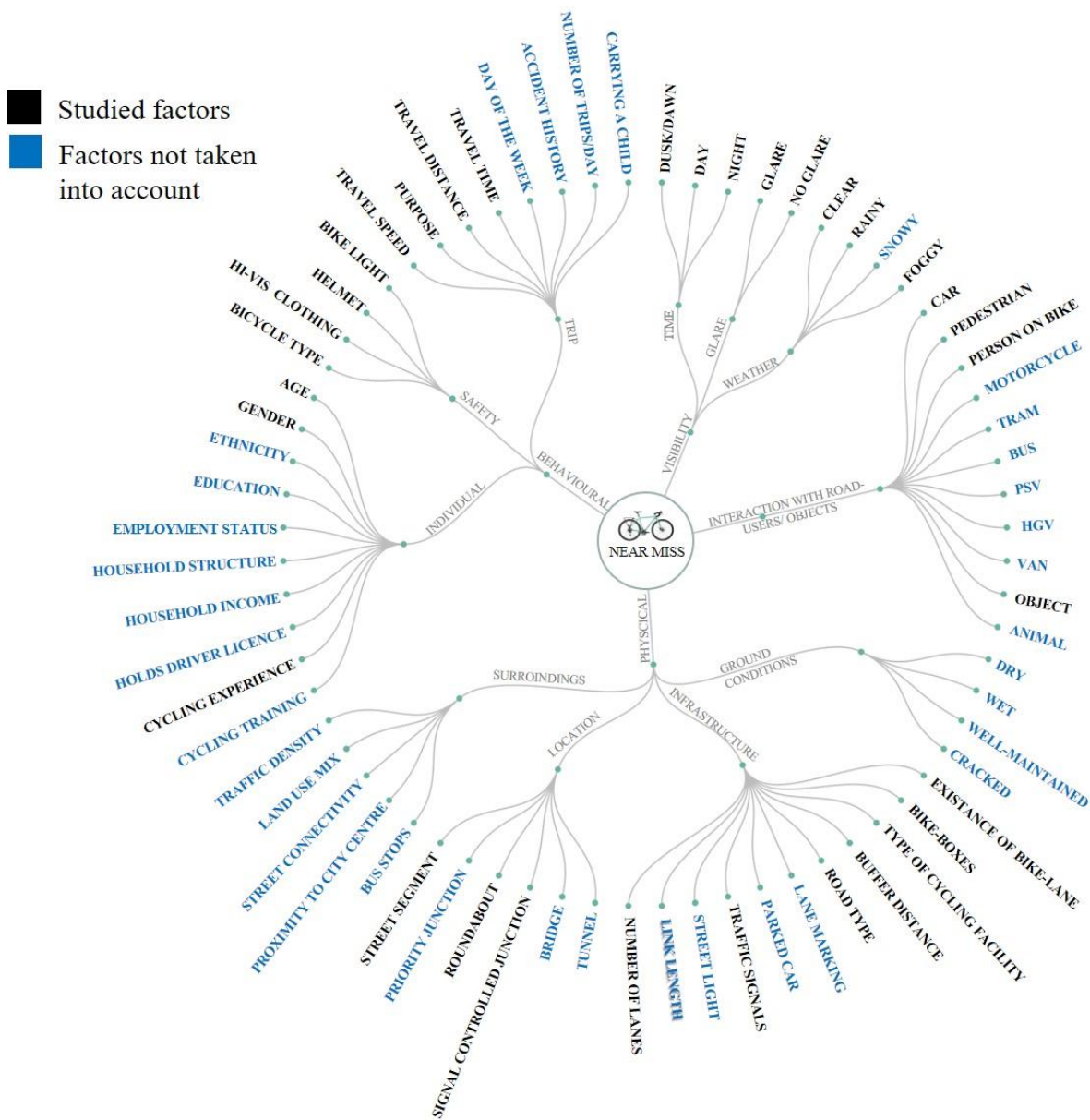
Figure 4: Risk factors related to cycling near misses
Source: Created by the authors based on the literature review

5.6 Summary

Table 2 summarises the risk factors covered in the 19 studies, broken down by method, where the number in each cell is the count of studies of that type that refer to that risk factor. In general, the built environment and cyclist characteristics are covered by more studies because they are static or slowly changing variables. It should be noted that, while the same factors are covered by many studies,

the form and quality of the data can be very different. For example, a self-report study may include a question on presence/absence of cycle lane, while a naturalistic study using video will allow interpretation of the type of cycle lane, the road condition and the surrounding context, albeit usually with expert interpretation.

Dynamic variables such as lighting and weather conditions are more difficult to capture using each of the methods, but for different reasons. In self-reporting studies, they require recall and may be subjective, which means that are usually incorporated in simple terms such as day/night time or rain/no rain. Naturalistic studies and site video analysis can capture more nuanced factors but they require labelling of video data, which is usually manual and time-consuming.

Table 2: summary of near miss studies

Method (Count)	Weather conditions					Light conditions and time of the day				Built environment conditions					Cyclist characteristics				Safety equipment		
	Clear	Cloud	Rain	Snow	Fog	Dusk/ Dawn	Day-time	Night-time	lighting conditions	Road types	Road conditions	Existence of bicycle lane	Parked cars	Intersection	Age	Gender	Experience	Trip purpose	Helmet	Hi-vis clothing	Bike light
Self-reporting (10)	1	1	1	0	1	3	3	3	2	6	4	6	5	4	10	10	6	7	4	1	2
Site VideoAnalysis (2)	0	0	0	0	0	0	0	0	0	1	0	1	0	1	2	1	1	1	NA	NA	NA
Naturalistic (5)	1	0	1	0	0	2	5	4	2	2	3	5	1	3	6	4	0	4	1	0	0
Total	**2**	**1**	**2**	**0**	**1**	**5**	**8**	**7**	**4**	**9**	**7**	**12**	**6**	**8**	**18**	**15**	**7**	**12**	**5**	**1**	**2**

A combination of naturalistic studies with self-reporting has the potential to capture the broadest range of information, while site video analysis is severely limited in its ability to capture demographic factors. It is also important to note that, while self-reporting captures the characteristics of the rider, it cannot easily capture the characteristics of the other road-users involved. In the next section, a detailed discussion of the limitations of current methods and gaps in the literature is presented.

6. Gaps in the literature

There are methodological challenges for collecting and analysing road safety data and its risk factors (Schlögl & Stütz, 2017). There are several limitations in the existing methods used to understand near misses. These limitations are: 1) How to eliminate factors due to manual labelling of data, 2) The current lack of functionality of sensors, 3) small data samples, 4) limited scope of studies, 5) The absence of a unified method for understanding near miss, which leads to 6) lack of understanding of the impact of the risk factors.

6.1 Elimination of factors due to manual labelling

One of the crucial drawbacks in current studies is that the analysis of data gathered has been dependent on manual labelling, which is highly time-consuming and may also introduce some bias in how data sets are labelled, limiting the transferability of findings to a different context.

6.2 Limitations of sensor data

The current sensors used have not been developed to measure the range of possible factors that might influence near misses. In particular, current approaches mean that it is difficult to compare sensor data from one location to another, especially where the environmental context can be very different (e.g. urban vs rural, hot vs cold climate, etc.).

6.3 Limitation of data sample size

Current studies have involved small data sets making it difficult to draw statistically significant conclusions, which again limits the ability to draw conclusions relevant across a range of contexts. Even though the naturalistic approach shows progress in collecting rich data on the context and factors related to near misses, the current approach of labelling data manually reduces its potential for large scale implementation without automated data processing.

6.4 Limitation of the study scope

Most studies focus either on certain types of near misses, i.e. passing events, with a wide range of risk factors (Beck et al., 2019), or on a range of near-miss types with a limited number of analysed risk factors (Aldred, 2016; Aldred & Crosweller, 2015).

6.5 The absence of a unified framework for understanding near misses

Previous studies of cycling near misses have lacked a method for understanding near misses regardless of the context, the types of near misses, or the related risk factors associated with these incidents. Overcoming this limitation may lead to a deeper understanding of near misses and for drawing transferable guidelines.

6.6 Limitation in understanding the impact of the risk factors

Understanding causality and effect of the given risk factors on near misses from a statistical model would require the variables to be random and controlled via unbiased variables that show a direct effect on both risk factors and near misses. Given the limitation in extracting a wide range of risk factors with a large sample size as aforementioned, we cannot currently quantify and assess the impact of risk factors in an unbiased and systematic way.

7. Computer vision for recognising near misses and their risk factors

Cycling near misses can be viewed in the wider context of the place and time in which they occur. Our knowledge about near misses and their related risk factors could be built through video streams that may provide a very effective approach to identifying patterns related to their occurrence. In general, the success of machine intelligence and computer vision in pattern recognition in the last decade, (LeCun et al., 2015) has added a new dimension towards understanding cities generally. Machine vision has the potential to understand cycling near misses by extracting safety-related features from still or multi-frame images in complex daily life scenes. However, we argue that near miss studies should not focus solely

on extracting known risk factors. All the different layers of cities (the built environment, human interaction, transportation and traffic, the natural environment and infrastructure) can have potential impacts on the experience of a rider and should not be discounted (Ibrahim et al., 2020). Using the different algorithms of computer vision, coupled with deep learning, we can extract and analyse these features to develop automated systems that can be applied to multiple tasks, functioning at different scales. This will help draw significant conclusions and assist the process of policy-making.

To develop an autonomous and multi-tasking system to detect near miss scenes, their types, and the associated risk factors, seven steps need to be considered: 1) Sensing and classifying the physical environment, 2) detecting objects and obstacles, 3) inferring distance and detecting safety measures, 4) recognising motion, 5) recognising actions and inferring behaviours, 6) Inferring individual characteristics. Such a system would make use of a range of computer vision techniques, such as image classification, segmentation (classifying an image at a pixel level), object detection, action recognition, scene awareness and understanding the underlying gist of a scene. By embedding these technologies within sensors, this approach would move beyond naturalistic studies towards automatic quantification and analysis of risk, and 7) Integrating all algorithms, which could allow inferring causation and effect of the different risk factors.

Figure 5 shows how deep learning and computer vision algorithms can be integrated to identify: 1) The risk factors, 2) near miss scenes, 3) types of near misses, and 4) the impact of the different risk factors on the different types of near misses. We discuss each of the components in turn in the following subsections.
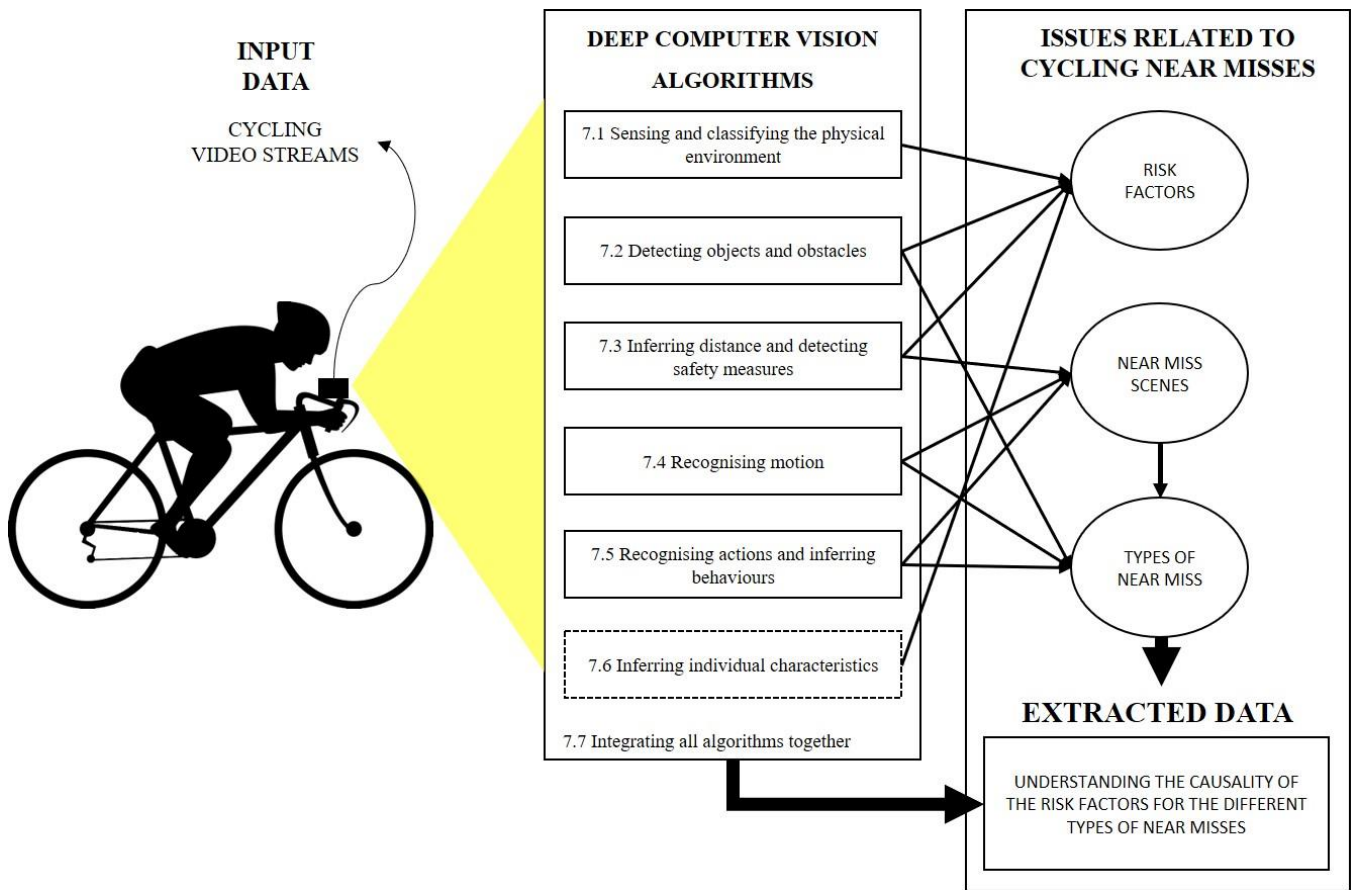
Figure 5: Conceptual framework for an AI-embedded system to understand cycling near miss

7.1 Sensing and classifying the physical environment

Convolutional Neural Network (CNN) models make it possible to tackle risk factors related to cycling near misses, particularly those related to the physical environment and the visual conditions. For instance, models such as AlexNet (Krizhevsky et al., 2012), VGGNet (Simonyan & Zisserman, 2014), GoogLeNet (Szegedy et al., 2015), ResNet (K. He et al., 2015) and DenseNet (Huang et al., 2017) have been successful in recognising thousands of objects in large databases, such as ImageNet, with high accuracy and precision.

Using these methods, different models have been developed, often in other research domains, that can be used to extract risk factors related to cycling near misses. Most recently, models that are

capable of detecting weather conditions and time of the day (Ibrahim et al., 2019b) and understanding the overall deterioration of the environment have been developed (Ibrahim et al., 2019a). Given sufficient training data, bespoke models can be built to detect the different dynamics of the physical environment.

7.2 Detecting objects and obstacles

Detecting road-users or objects, such as a car door, is indispensable for understanding the risk factors related to cycling near misses and for classifying the different types of near misses as previously discussed (Section 3.2). Localisation is the process of identifying multiple objects in a single frame and is also applicable to video streams. These models use a single deep model in an end-to-end fashion to localise different objects with a given confidence. Similar to classification, different models of different convolutional structures have been developed to segment and localise objects in a single frame of an image such as You Only Look Once (YOLO) (Redmon & Farhadi, 2017, 2018), and MultiBox Detectors (SSD) (Liu et al., 2016).

Relying on this type of algorithm, extracting pedestrian and transport modes from complex urban settings and mapping them has been achieved (Ibrahim et al., 2019a), which can be used to recognise the different agents. This is the first step in understanding the interaction of people on bikes with other road-users.

7.3 Inferring distance and detecting safety measures

Understanding what is a safe or unsafe distance when passing or overtaking a bicycle is crucial for detecting near miss scenes and identifying their types. CNN based computer vision techniques can be trained to infer the distance of objects from the camera just by looking at a still image or a video stream. For instance, Cao, Wu, & Shen (2017) trained deep CNN models to estimate the depth in a single image

by labelling the different depths on the image. Also, He, Wang, & Hu (2018) trained a deep CNN model to estimate the depth of a monocular image. In the context of near misses, CNN algorithms can be trained with images labelled using Light Detection and Ranging (LiDAR), or ultrasonic range sensors. By utilising the depth estimation algorithms, a standard buffer distance can be measured and detected from cycling video streams, which would contribute to understanding the different types of near misses.

Furthermore, it is foreseeable that once trained, these algorithms will no longer require the range sensors and can be applied to video streams collected in isolation. This opens up near miss study to the vast amounts of data that are routinely collected by riders who use action cameras.

7.4 Recognising motion

Understanding the overall motions of the different objects in the scene is another key role in understanding scenes of near misses and identifying their types. Various computer vision models rely on estimating the change in motion for a sequential frame of images, or so-called optical flow (Alvarez et al., 2007; Andrade et al., 2006; Ayvaci et al., 2012; Baker et al., 2011; Butler et al., 2012; Enkelmann, n.d.; Mallot et al., 1991; Sun et al., 2010).

Estimating optical flow from cycling video streams would enable the differentiation of the moving objects from the stationary part of the scene, in addition to understanding obstacles and occlusion. Thus, it would allow a better understanding of such instant actions as near misses.

7.5 Recognising actions and inferring behaviours

Moving from tracking the motion of objects in a complex scene towards classifying multiple actions while tracking, deep models also have shown continuous progress (Bilen et al., 2016; Wang et al., 2015; Zhang et al., 2016). In fact, deep computer vision models have been successful in understanding human behaviours based on the poses of human skeletons and how they interact with

other objects in a complex environment (El-Nouby & Taylor, 2018; Saha et al., 2016; Soomro & Shah, 2017; Weinzaepfel et al., 2016). By utilising action-recognition models to detect the overall motion, the unsafe riding scenes and near misses can be automatically detected and categorised.

7.6 Inferring individual characteristics

Computer vision coupled with deep learning shows good potential for extracting information related to individual characteristics (as mentioned in Section 5.1). However, this topic remains the most significant bottleneck in achieving the system outlined in figure 5. While it is feasible to collect personal characteristics of people using an AI-embedded device, inferring the characteristics of those they come into contact with is more of a challenge and is a common issue for all near-miss study methodologies. Various deep models have shown good potential in extracting information in this regard, such as recognising gender (Levi & Hassncer, 2015; Narang & Bourlai, 2016), age (Levi & Hassncer, 2015), facial emotions (Minaee & Abdolrashidi, 2019), and ethnicity (Narang & Bourlai, 2016). Building on these models could be beneficial for understanding the various characteristics of people on bikes and other road-users that would help to understand the impacts of near misses, along with their different types and the other risk factors.

7.7 Integrating all algorithms together

By putting all algorithms together, a rich data set can be generated that includes information related to risk factors and the frequency and type of near misses. This would respond to the current knowledge gap in the applied methods used for studying cycling near misses. Moreover, better regression models can be used to understand the causality and the impact of the various risk factors toward the different types of near misses, in which drawing conclusions could be more effective for both people on bikes and policy-makers.

8.   Conclusion

Although serious incidents related to cycling leading to major injuries are rare, it has been shown that focusing on them only touches the tip of the iceberg (Jones et al., 1999), and near miss data can provide much more information about potential problems and how to avoid a risky situation that may lead to serious incidents.

Different studies focus on analysing the perceived risk related to the various types of near misses and particular progress has been achieved related to understanding the most frequent type of near miss - close pass - where people on bikes are passed in unsafe or uncomfortable manoeuvres by other road-users. There are various challenges related to the current methods used to collect, analyse, and produce evidence that could assist policy-making towards minimising risky situations. In this article, we reviewed the different studies, highlighting the methods and scope to underline the knowledge gap in which further work is needed.

While the approach of the naturalistic study seems to be promising in understanding instant and risky situations such as near misses, the current methods related to collecting and analysing video streams remain a challenge for drawing significant conclusions. In this article, we propose a framework for how machine intelligence, relying specifically on the domain of computer vision, may overcome the current limitations and the move towards a unified method of understanding near misses regardless of the limitation of labelled data or the configurations of the context.

Our conceptual framework explains how different deep models can be trained and utilised to reach an AI-embedded system that automates the detection of near miss scenes, analyses their types, their risk factors, and draws significant conclusions based on the causality of risk factors, behaviours of people on bikes, and their mutual interaction with other road-users.

References

Abay, K. A. (2015). Investigating the nature and impact of reporting bias in road crash data.

Transportation Research Part A: Policy and Practice, *71*, 31–45.

https://doi.org/10.1016/j.tra.2014.11.002

Aldred, R. (2016). Cycling near misses: Their frequency, impact, and prevention. *Transportation*

Research Part A: Policy and Practice, *90*, 69–83. https://doi.org/10.1016/j.tra.2016.04.016

Aldred, R. (2018). Inequalities in self-report road injury risk in Britain: A new analysis of National

Travel Survey data, focusing on pedestrian injuries. *Journal of Transport & Health*, *9*, 96–104.

https://doi.org/10.1016/j.jth.2018.03.006

Aldred, R., & Crosweller, S. (2015). Investigating the rates and impacts of near misses and related

incidents among UK cyclists. *Journal of Transport & Health*, *2*(3), 379–393.

https://doi.org/10.1016/j.jth.2015.05.006

Aldred, R., & Goodman, A. (2018). Predictors of the frequency and subjective experience of cycling

near misses: Findings from the first two years of the UK Near Miss Project. *Accident Analysis &*

Prevention, *110*, 161–170. https://doi.org/10.1016/j.aap.2017.09.015

Alvarez, L., Deriche, R., Papadopoulo, T., & Sánchez, J. (2007). Symmetrical Dense Optical Flow

Estimation with Occlusions Detection. *International Journal of Computer Vision*, *75*(3), 371–

385. https://doi.org/10.1007/s11263-007-0041-4

Andrade, E. L., Blunsden, S., & Fisher, R. B. (2006). Modelling Crowd Scenes for Event Detection.

18th International Conference on Pattern Recognition (ICPR'06), 175–178.

https://doi.org/10.1109/ICPR.2006.806

Ayvaci, A., Raptis, M., & Soatto, S. (2012). Sparse Occlusion Detection with Optical Flow.

International Journal of Computer Vision, *97*(3), 322–338. https://doi.org/10.1007/s11263-011-

0490-7

Baker, S., Scharstein, D., Lewis, J. P., Roth, S., Black, M. J., & Szeliski, R. (2011). A Database and

    Evaluation Methodology for Optical Flow. *International Journal of Computer Vision*, *92*(1), 1–

    31. https://doi.org/10.1007/s11263-010-0390-2

Beck, B., Chong, D., Olivier, J., Perkins, M., Tsay, A., Rushford, A., Li, L., Cameron, P., Fry, R., &

    Johnson, M. (2019). How much space do drivers provide when passing cyclists? Understanding

    the impact of motor vehicle and infrastructure characteristics on passing distance. *Accident*

    Analysis & Prevention, *128*, 253–260. https://doi.org/10.1016/j.aap.2019.03.007

Beck, B., Stevenson, M., Newstead, S., Cameron, P., Judson, R., Edwards, E. R., Bucknill, A., Johnson,

    M., & Gabbe, B. (2016a). Bicycling crash characteristics: An in-depth crash investigation study.

    Accident Analysis & Prevention, *96*, 219–227. https://doi.org/10.1016/j.aap.2016.08.012

Beck, B., Stevenson, M., Newstead, S., Cameron, P., Judson, R., Edwards, E. R., Bucknill, A., Johnson,

    M., & Gabbe, B. (2016b). Bicycling crash characteristics: An in-depth crash investigation study.

    Accident Analysis & Prevention, *96*, 219–227. https://doi.org/10.1016/j.aap.2016.08.012

Bíl, M., Bílová, M., & Müller, I. (2010). Critical factors in fatal collisions of adult cyclists with

    automobiles. *Accident Analysis and Prevention*, 5.

Bilen, H., Fernando, B., Gavves, E., Vedaldi, A., & Gould, S. (2016). Dynamic Image Networks for

    Action Recognition. *2016 IEEE Conference on Computer Vision and Pattern Recognition*

    (CVPR), 3034–3042. https://doi.org/10.1109/CVPR.2016.331

Blaizot, S., Papon, F., Haddak, M. M., & Amoros, E. (2013). Injury incidence rates of cyclists compared

    to pedestrians, car occupants and powered two-wheeler riders, using a medical registry and

    mobility data, Rhône County, France. *Accident Analysis & Prevention*, *58*, 35–45.

    https://doi.org/10.1016/j.aap.2013.04.018

Branion-Calles, M., Nelson, T., & Winters, M. (2017). Comparing Crowdsourced Near-Miss and

    Collision Cycling Data and Official Bike Safety Reporting. *Transportation Research Record:*

    Journal of the Transportation Research Board, *2662*(1), 1–11. https://doi.org/10.3141/2662-01

Broach, J., Dill, J., & Gliebe, J. (2012). Where do cyclists ride? A route choice model developed with

    revealed preference GPS data. *Transportation Research Part A: Policy and Practice*, *46*(10),

    1730–1740. https://doi.org/10.1016/j.tra.2012.07.005

Butler, D. J., Wulff, J., Stanley, G. B., & Black, M. J. (2012). A Naturalistic Open Source Movie for

    Optical Flow Evaluation. In A. Fitzgibbon, S. Lazebnik, P. Perona, Y. Sato, & C. Schmid (Eds.),

    Computer Vision – ECCV 2012 (Vol. 7577, pp. 611–625). Springer Berlin Heidelberg.

    https://doi.org/10.1007/978-3-642-33783-3_44

Cao, Y., Wu, Z., & Shen, C. (2017). Estimating Depth from Monocular Images as Classification Using

    Deep Fully Convolutional Residual Networks. *ArXiv:1605.02305 [Cs]*.

    http://arxiv.org/abs/1605.02305

Chaurand, N., & Delhomme, P. (2013). Cyclists and drivers in road interactions: A comparison of

    perceived crash risk. *Accident Analysis and Prevention*, 9.

Cho, G., Rodríguez, D. A., & Khattak, A. J. (2009). The role of the built environment in explaining

    relationships between perceived and actual pedestrian and bicyclist safety. *Accident Analysis &*

    Prevention, *41*, 692–702. https://doi.org/10.1016/j.aap.2009.03.008

de Geus, B., Vandenbulcke, G., Int Panis, L., Thomas, I., Degraeuwe, B., Cumps, E., Aertsens, J., Torfs,

    R., & Meeusen, R. (2012). A prospective cohort study on minor accidents involving commuter

    cyclists in Belgium. *Accident Analysis & Prevention*, *45*, 683–693.

    https://doi.org/10.1016/j.aap.2011.09.045

de Hartog, J. J., Boogaard, H., Nijland, H., & Hoek, G. (2010). Do the Health Benefits of Cycling

    Outweigh the Risks? *Environmental Health Perspectives*, *118*(8), 1109–1116.

    https://doi.org/10.1289/ehp.0901747

De Rome, L., Boufous, S., Georgeson, T., Senserrick, T., Richardson, D., & Ivers, R. (2014). Bicycle

    Crashes in Different Riding Environments in the Australian Capital Territory. *Traffic Injury*

    Prevention, *15*(1), 81–88. https://doi.org/10.1080/15389588.2013.781591

Department for Transport. (2019). *Reported road casualties in Great Britain: 2018 annual report* (p.

    47). UK Department for Transport.

DiGioia, J., Watkins, K. E., Xu, Y., Rodgers, M., & Guensler, R. (2017). Safety impacts of bicycle

    infrastructure: A critical review. *Journal of Safety Research*, *61*, 105–119.

    https://doi.org/10.1016/j.jsr.2017.02.015

Dozza, M, Werneke, J., & Fernandez, A. (2012). *Piloting the Naturalistic Methodology on Bicycles*. 11.

Dozza, Marco, Bianchi Piccinini, G. F., & Werneke, J. (2016). Using naturalistic data to assess e-cyclist

    behavior. *Transportation Research Part F: Traffic Psychology and Behaviour*, *41*, 217–226.

    https://doi.org/10.1016/j.trf.2015.04.003

Dozza, Marco, Schindler, R., Bianchi-Piccinini, G., & Karlsson, J. (2016). How do drivers overtake

    cyclists? *Accident Analysis & Prevention*, *88*, 29–36. https://doi.org/10.1016/j.aap.2015.12.008

Dozza, Marco, Schwab, A., & Wegman, F. (2017). Safety Science Special Issue on Cycling Safety.

    Safety Science, *92*, 262–263. https://doi.org/10.1016/j.ssci.2016.06.009

Dozza, Marco, & Werneke, J. (2014). Introducing naturalistic cycling data: What factors influence

    bicyclists' safety in the real world? *Transportation Research Part F: Traffic Psychology and*

    Behaviour, *24*, 83–91. https://doi.org/10.1016/j.trf.2014.04.001

El-Nouby, A., & Taylor, G. W. (2018). Real-Time End-to-End Action Detection with Two-Stream

    Networks. *ArXiv:1802.08362 [Cs]*. http://arxiv.org/abs/1802.08362

Enkelmann, W. (n.d.). *Obstacle detection by evaluation of optical flow fields from image sequences*. 5.

Fuller, D., Gauvin, L., Morency, P., Kestens, Y., & Drouin, L. (2013). The impact of implementing a public bicycle share program on the likelihood of collisions and near misses in Montreal, Canada. *Preventive Medicine*, *57*(6), 920–924. https://doi.org/10.1016/j.ypmed.2013.05.028

Gatersleben, B., & Haddad, H. (2010). Who is the typical bicyclist? *Transportation Research Part F: Traffic Psychology and Behaviour*, *13*(1), 41–48. https://doi.org/10.1016/j.trf.2009.10.003

Gustafsson, L., & Archer, J. (2013). A naturalistic study of commuter cyclists in the greater Stockholm area. *Accident Analysis & Prevention*, *58*, 286–298. https://doi.org/10.1016/j.aap.2012.06.004

He, K., Zhang, X., Ren, S., & Sun, J. (2015). Deep Residual Learning for Image Recognition. *ArXiv:1512.03385v1*. https://arxiv.org/pdf/1512.03385.pdf

He, L., Wang, G., & Hu, Z. (2018). Learning Depth from Single Images with Deep Neural Network Embedding Focal Length. *IEEE Transactions on Image Processing*, *27*(9), 4676–4689. https://doi.org/10.1109/TIP.2018.2832296

Huang, G., Liu, Z., Weinberger, K. Q., & van der Maaten, L. (2017). Densely connected convolutional networks. *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, *1*, 3.

Ibrahim, M. R., Haworth, J., & Cheng, T. (2019a). URBAN-i: From urban scenes to mapping slums, transport modes, and pedestrians in cities using deep learning and computer vision. *Environment and Planning B: Urban Analytics and City Science*, 239980831984651. https://doi.org/10.1177/2399808319846517

Ibrahim, M. R., Haworth, J., & Cheng, T. (2019b). WeatherNet: Recognising Weather and Visual Conditions from Street-Level Images Using Deep Residual Learning. *ISPRS International Journal of Geo-Information*, *8*(12), 549. https://doi.org/10.3390/ijgi8120549

Ibrahim, M. R., Haworth, J., & Cheng, T. (2020). Understanding cities with machine eyes: A review of deep computer vision in urban analytics. *Cities*, *96*, 102481. https://doi.org/10.1016/j.cities.2019.102481

Imprialou, M., & Quddus, M. (2017). Crash data quality for road safety research: Current state and future directions. *Accident Analysis & Prevention*. https://doi.org/10.1016/j.aap.2017.02.022

Jestico, B., Nelson, T., & Winters, M. (2016). Mapping ridership using crowdsourced cycling data. *Journal of Transport Geography*, *52*, 90–97. https://doi.org/10.1016/j.jtrangeo.2016.03.006

Johnson, M., Charlton, J., Oxley, J., & Newstead, S. (2010). Naturalistic Cycling Study: Identifying Risk Factors for On-Road Commuter Cyclists. *Annals of Advances in Automotive Medicine / Annual Scientific Conference*, *54*, 275–283.

Johnson, M., Newstead, S., Oxley, J., & Charlton, J. (2013). Cyclists and open vehicle doors: Crash characteristics and risk factors. *Safety Science*, *59*, 135–140. https://doi.org/10.1016/j.ssci.2013.04.010

Johnson, M., Oxley, J., Newstead, S., & Charlton, J. (2014). Safety in numbers? Investigating Australian driver behaviour, knowledge and attitudes towards cyclists. *Accident Analysis & Prevention*, *70*, 148–154. https://doi.org/10.1016/j.aap.2014.02.010

Jones, S., Kirchsteiger, C., & Bjerke, W. (1999). The importance of near miss reporting to further improve safety performance. *Journal of Loss Prevention in the Process Industries*, 9.

Juhra, C., Wieskötter, B., Chu, K., Trost, L., Weiss, U., Messerschmidt, M., Malczyk, A., Heckwolf, M., & Raschke, M. (2012). Bicycle accidents – Do we only see the tip of the iceberg? *Injury*, *43*(12), 2026–2034. https://doi.org/10.1016/j.injury.2011.10.016

Krizhevsky, A., Sutskever, I., & Hinton, G. E. (2012). Imagenet classification with deep convolutional neural networks. *Proceeding NIPS'12 Proceedings of the 25th International Conference on Neural Information Processing Systems*, *1*, 1097–1105.

Lacherez, P., Wood, J. M., Marszalek, R. P., & King, M. J. (2013). Visibility-related characteristics of crashes involving bicyclists and motor vehicles – Responses from an online questionnaire study. *Transportation Research Part F: Traffic Psychology and Behaviour*, *20*, 52–58. https://doi.org/10.1016/j.trf.2013.04.003

Lawson, A. R., Pakrashi, V., Ghosh, B., & Szeto, W. Y. (2013). Perception of safety of cyclists in Dublin City. *Accident Analysis & Prevention*, *50*, 499–511. https://doi.org/10.1016/j.aap.2012.05.029

LeCun, Y., Bengio, Y., & Hinton, G. (2015). Deep learning. *Nature*, *521*(7553), 436–444. https://doi.org/10.1038/nature14539

Lehtonen, E., Havia, V., Kovanen, A., Leminen, M., & Saure, E. (2016). Evaluating bicyclists' risk perception using video clips: Comparison of frequent and infrequent city cyclists. *Transportation Research Part F: Traffic Psychology and Behaviour*, *41*, 195–203. https://doi.org/10.1016/j.trf.2015.04.006

Levi, G., & Hassncer, T. (2015). Age and gender classification using convolutional neural networks. *2015 IEEE Conference on Computer Vision and Pattern Recognition Workshops (CVPRW)*, 34–42. https://doi.org/10.1109/CVPRW.2015.7301352

Liu, W., Anguelov, D., Erhan, D., Szegedy, C., Reed, S., Fu, C.-Y., & Berg, A. C. (2016). Ssd: Single shot multibox detector. *European Conference on Computer Vision*, 21–37.

Loo, B. P. Y., & Tsui, K. L. (2010). Bicycle crash casualties in a highly motorized city. *Accident Analysis & Prevention*, *42*(6), 1902–1907. https://doi.org/10.1016/j.aap.2010.05.011

Mallot, H. A., Bülthoff, H. H., Little, J. J., & Bohrer, S. (1991). Inverse perspective mapping simplifies optical flow computation and obstacle detection. *Biological Cybernetics*, *64*(3), 177–185. https://doi.org/10.1007/BF00201978

Minaee, S., & Abdolrashidi, A. (2019). Deep-Emotion: Facial Expression Recognition Using Attentional Convolutional Network. *ArXiv:1902.01019 [Cs]*. http://arxiv.org/abs/1902.01019

Narang, N., & Bourlai, T. (2016). Gender and ethnicity classification using deep learning in heterogeneous face recognition. *2016 International Conference on Biometrics (ICB)*, 1–8. https://doi.org/10.1109/ICB.2016.7550082

Nelson, T. A., Denouden, T., Jestico, B., Laberee, K., & Winters, M. (2015). BikeMaps.org: A Global Tool for Collision and Near Miss Mapping. *Frontiers in Public Health*, *3*. https://doi.org/10.3389/fpubh.2015.00053

Orsi, C., Ferraro, O. E., Montomoli, C., Otte, D., & Morandi, A. (2014). Alcohol consumption, helmet use and head trauma in cycling collisions in Germany. *Accident Analysis & Prevention*, *65*, 97–104. https://doi.org/10.1016/j.aap.2013.12.019

Pai, C.-W. (2011). Overtaking, rear-end, and door crashes involving bicycles: An empirical investigation. *Accident Analysis & Prevention*, *43*(3), 1228–1235. https://doi.org/10.1016/j.aap.2011.01.004

Parkin, J., & Meyers, C. (2010). The effect of cycle lanes on the proximity between motor traffic and cycle traffic. *Accident Analysis & Prevention*, *42*(1), 159–165. https://doi.org/10.1016/j.aap.2009.07.018

Paschalidis, E., Basbas, S., Politis, I., & Prodromou, M. (2016). *'"Put the blame on. . .others!"': The battle of cyclists against pedestrians and car drivers at the urban environment. A cyclists' perception study*. 18.

Poulos, R. G., Hatfield, J., Rissel, C., Grzebieta, R., & McIntosh, A. S. (2012). Exposure-based cycling crash, near miss and injury rates: The Safer Cycling Prospective Cohort Study protocol: Figure 1. *Injury Prevention*, *18*(1), e1–e1. https://doi.org/10.1136/injuryprev-2011-040160

PRISMA. (2015). *TRANSPARENT REPORTING of SYSTEMATIC REVIEWS and META-ANALYSES*.

    PRISMA. http://prisma-statement.org/

Pucher, J., Dill, J., & Handy, S. (2010). Infrastructure, programs, and policies to increase bicycling: An

    international review. *Preventive Medicine*, *50*, S106–S125.

    https://doi.org/10.1016/j.ypmed.2009.07.028

Redmon, J., & Farhadi, A. (2017). YOLO9000: Better, Faster, Stronger. *2017 IEEE Conference on*

    Computer Vision and Pattern Recognition (CVPR), 6517–6525.

    https://doi.org/10.1109/CVPR.2017.690

Saha, S., Singh, G., Sapienza, M., Torr, P. H. S., & Cuzzolin, F. (2016). Deep Learning for Detecting

    Multiple Space-Time Action Tubes in Videos. *ArXiv:1608.01529 [Cs]*.

    http://arxiv.org/abs/1608.01529

Sanders, R. L. (2015). Perceived traffic risk for cyclists: The impact of near miss and collision

    experiences. *Accident Analysis & Prevention*, *75*, 26–34.

    https://doi.org/10.1016/j.aap.2014.11.004

Savan, B., Cohlmeyer, E., & Ledsham, T. (2017). Integrated strategies to accelerate the adoption of

    cycling for transportation. *Transportation Research Part F: Traffic Psychology and Behaviour*,

    46, 236–249. https://doi.org/10.1016/j.trf.2017.03.002

Schleinitz, K., Petzoldt, T., Franke-Bartholdt, L., Krems, J., & Gehlert, T. (2017). The German

    Naturalistic Cycling Study – Comparing cycling speed of riders of different e-bikes and

    conventional bicycles. *Safety Science*, *92*, 290–297. https://doi.org/10.1016/j.ssci.2015.07.027

Schleinitz, Katja, Petzoldt, T., Franke-Bartholdt, L., Krems, J. F., & Gehlert, T. (2015). Conflict partners

    and infrastructure use in safety critical events in cycling – Results from a naturalistic cycling

    study. *Transportation Research Part F: Traffic Psychology and Behaviour*, *31*, 99–111.

    https://doi.org/10.1016/j.trf.2015.04.002

Schlögl, M., & Stütz, R. (2017). Methodological considerations with data uncertainty in road safety analysis. *Accident Analysis & Prevention*. https://doi.org/10.1016/j.aap.2017.02.001

Shinar, D., Valero-Mora, P., van Strijp-Houtenbos, M., Haworth, N., Schramm, A., De Bruyne, G., Cavallo, V., Chliaoutakis, J., Dias, J., Ferraro, O. E., Fyhri, A., Sajatovic, A. H., Kuklane, K., Ledesma, R., Mascarell, O., Morandi, A., Muser, M., Otte, D., Papadakaki, M., … Tzamalouka, G. (2018). Under-reporting bicycle accidents to police in the COST TU1101 international survey: Cross-country comparisons and associated factors. *Accident Analysis & Prevention*, *110*, 177–186. https://doi.org/10.1016/j.aap.2017.09.018

Simonyan, K., & Zisserman, A. (2014). Very deep convolutional networks for large-scale image recognition. *ArXiv Preprint ArXiv:1409.1556*.

Soomro, K., & Shah, M. (2017). Unsupervised Action Discovery and Localization in Videos. *2017 IEEE International Conference on Computer Vision (ICCV)*, 696–705. https://doi.org/10.1109/ICCV.2017.82

Steinbach, R., Green, J., Datta, J., & Edwards, P. (2011). Cycling and the city: A case study of how gendered, ethnic and class identities can shape healthy transport choices. *Social Science & Medicine*, *72*(7), 1123–1130. https://doi.org/10.1016/j.socscimed.2011.01.033

Strauss, J., Miranda-Moreno, L. F., & Morency, P. (2013). Cyclist activity and injury risk analysis at signalized intersections: A Bayesian modelling approach. *Accident Analysis & Prevention*, *59*, 9–17. https://doi.org/10.1016/j.aap.2013.04.037

Sun, D., Roth, S., & Black, M. J. (2010). Secrets of optical flow estimation and their principles. *2010 IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, 2432–2439. https://doi.org/10.1109/CVPR.2010.5539939

Szegedy, C., Liu, W., Jia, Y., Sermanet, P., & Reed, S. (2015). *Going Deeper with Convolutions*. https://www.cs.unc.edu/~wliu/papers/GoogLeNet.pdf

Teschke, K., Frendo, T., Shen, H., Harris, M. A., Reynolds, C. C., Cripton, P. A., Brubacher, J.,

   Cusimano, M. D., Friedman, S. M., Hunte, G., Monro, M., Vernich, L., Babul, S., Chipman, M.,

   & Winters, M. (2014). Bicycling crash circumstances vary by route type: a cross-sectional

   analysis. *BMC Public Health*, *14*(1). https://doi.org/10.1186/1471-2458-14-1205

TfL. (2018). *Cycling action plan: Making London the world's best big city for cycling* (p. 59).

Vanparijs, J., Int Panis, L., Meeusen, R., & de Geus, B. (2015). Exposure measurement in bicycle safety

   analysis: A review of the literature. *Accident Analysis & Prevention*, *84*, 9–19.

   https://doi.org/10.1016/j.aap.2015.08.007

Vansteenkiste, P., Zeuwts, L., Cardon, G., & Lenoir, M. (2016). A hazard-perception test for cycling

   children: An exploratory study. *Transportation Research Part F: Traffic Psychology and

   Behaviour*, *41*, 182–194. https://doi.org/10.1016/j.trf.2016.05.001

Walker, I., Garrard, I., & Jowitt, F. (2014). The influence of a bicycle commuter's appearance on

   drivers' overtaking proximities: An on-road test of bicyclist stereotypes, high-visibility clothing

   and safety aids in the United Kingdom. *Accident Analysis & Prevention*, *64*, 69–77.

   https://doi.org/10.1016/j.aap.2013.11.007

Wang, L., Qiao, Y., & Tang, X. (2015). Action recognition with trajectory-pooled deep-convolutional

   descriptors. *2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 4305–

   4314. https://doi.org/10.1109/CVPR.2015.7299059

Weinzaepfel, P., Martin, X., & Schmid, C. (2016). Human Action Localization with Sparse Spatial

   Supervision. *ArXiv:1605.05197 [Cs]*. http://arxiv.org/abs/1605.05197

Winters, M., & Branion-Calles, M. (2017). Cycling safety: Quantifying the under reporting of cycling

   incidents in Vancouver, British Columbia. *Journal of Transport & Health*, *7*, 48–53.

   https://doi.org/10.1016/j.jth.2017.02.010

Zangenehpour, S., Miranda-Moreno, L. F., & Saunier, N. (2015). Automated classification based on

    video data at intersections with heavy pedestrian and bicycle traffic: Methodology and

    application. *Transportation Research Part C: Emerging Technologies*, *56*, 161–176.

    https://doi.org/10.1016/j.trc.2015.04.003

Zangenehpour, S., Strauss, J., Miranda-Moreno, L. F., & Saunier, N. (2016). Are signalized intersections

    with cycle tracks safer? A case–control study based on automated surrogate safety analysis using

    video data. *Accident Analysis & Prevention*, *86*, 161–172.

    https://doi.org/10.1016/j.aap.2015.10.025

Zhang, B., Wang, L., Wang, Z., Qiao, Y., & Wang, H. (2016). Real-time Action Recognition with

    Enhanced Motion Vector CNNs. *ArXiv:1604.07669 [Cs]*. http://arxiv.org/abs/1604.07669