

**Manuscript version: Author's Accepted Manuscript**

The version presented in WRAP is the author's accepted manuscript and may differ from the published version or Version of Record.

**Persistent WRAP URL:**

<http://wrap.warwick.ac.uk/173424>

**How to cite:**

Please refer to published version for the most recent bibliographic citation information. If a published version is known of, the repository item page linked to above, will contain details on accessing it.

**Copyright and reuse:**

The Warwick Research Archive Portal (WRAP) makes this work by researchers of the University of Warwick available open access under the following conditions.

Copyright © and all moral rights to the version of the paper presented here belong to the individual author(s) and/or other copyright owners. To the extent reasonable and practicable the material made available in WRAP has been checked for eligibility before being made available.

Copies of full items can be used for personal research or study, educational, or not-for-profit purposes without prior permission or charge. Provided that the authors, title and full bibliographic details are credited, a hyperlink and/or URL is given for the original metadata page and the content is not changed in any way.

**Publisher's statement:**

Please refer to the repository item page, publisher's statement section, for further information.

For more information, please contact the WRAP Team at: [wrap@warwick.ac.uk](mailto:wrap@warwick.ac.uk).

# Review of Graph-Based Hazardous Event Detection Methods for Autonomous Driving Systems

Dannier Xiao<sup>1</sup>, Mehrdad Dianti<sup>1</sup>, William Gonçalves Geiger<sup>1,2</sup>, Roger Woodman<sup>1</sup>

**Abstract**— Automated and autonomous vehicles are often required to operate in complex road environments with potential hazards that may lead to hazardous events causing injury or even death. Therefore, a reliable autonomous hazardous event detection system is a key enabler for highly autonomous vehicles (e.g., Level 4 and 5 autonomous vehicles) to operate without human supervision for significant periods of time. One promising solution to the problem is the use of graph-based methods that are powerful tools for relational reasoning. Using graphs to organise heterogeneous knowledge about the operational environment, link scene entities (e.g., road users, static objects, traffic rules) and describe how they affect each other. Due to a growing interest and opportunity presented by graph-based methods for autonomous hazardous event detection, this paper provides a comprehensive review of the state-of-the-art graph-based methods that we categorise as rule-based, probabilistic, and machine learning-driven. Additionally, we present an in-depth overview of the available datasets to facilitate hazardous event training and evaluation metrics to assess model performance. In doing so, we aim to provide a thorough overview and insight into the key research opportunities and open challenges.

**Index Terms**— Hazardous event, graph neural networks, Bayesian networks, rule-based ontologies, automated vehicles.

## I. INTRODUCTION

An estimated 83% to 94% of traffic accidents are related to human fault [1], [2]. Thus, progressively more driving functions are being automated to improve safety and efficiency. However, due to the complex driving domain, current automated vehicle (AV) systems: have limited functionality (e.g., adaptive cruise control and lane change), are restricted to specific operating conditions defined by an operational design domain (ODD) and require constant human supervision.

Current systems are classified as Level 2 automation by the Society of Automotive Engineers (SAE) [3], and the ultimate goal is to transition from human supervised automation to unsupervised autonomy for all driving functions in any ODD (SAE Level 5). A key enabler to transition past SAE Level 2 systems is the ability to reliably and comprehensively detect scenarios that may cause harm or when the system can no longer safely function, which are classified as hazardous events and defined in section II.A together with other key terms.

To detect hazardous events, AV systems must process a vast amount of perception data (e.g., object detection, classification

and localisation), which can be noisy, uncertain or incomplete, in order to identify potential events that may materialise harm. However, achieving early and robust hazardous event detection remains challenging due to an unlimited variety of edge cases, unseen environments, and the driving domain's overall complexity [4]–[7]. Aggravated by the fact that it is impossible to define all hazardous events manually, a scalable approach is vital to adapt in an ever-changing domain.

Regarding road safety, human behaviour plays a pivotal role but is complex to predict as it is affected by interactions with other actors (i.e., surrounding road users like vehicles and pedestrians) and the environment. With 83-94% of accidents linked to human fault [1], [2], understanding such interactions is vital to enable the transition to SAE Level 3+, where AVs will need to safely control all driving functions. Historically, kinematic and physics-based methods have been proposed to predict colliding trajectories [8], however, such methods can not describe the frequent interactions between road users that cause drivers to suddenly alter trajectories (e.g., lane change).

In response, deep learning models have been proposed that utilise both convolutional neural networks (CNNs) and recurrent neural networks (RNNs) [9]–[11] to detect hazardous events from dashboard video footage. Images are parsed with a CNN to extract spatial features, like the location and appearance of objects in the image and then fed to the RNN to extract temporal features such as how spatial features evolve in time. Despite great advances from CNNs, unguided learning may draw spurious patterns [12]–[14]. In addition, spatial features may omit vital contextual relations that are not observable [15]–[17].

The lack of relational reasoning of the aforementioned methods has given rise to graph-based methods. Such methods allow scene entities to be represented as vertices (i.e., nodes) and related through edges to decompose a complex event into constituent components (e.g., road users, environment and traffic rules) and reason how components affect each other. Not all hazardous events can be directly observed, but through graph-based learning, authors have proposed the contextual inference needed to distinguish simple road debris from a football that may signal a child running onto the road to retrieve it [18], or context to foresee a loss of control from sharp turns by relating the curve radius, speed and max deceleration [19].

\*This research is supported by the Centre for Doctoral Training (CDT) to Advance the Deployment of Future Mobility Technologies at the University of Warwick (<https://warwick.ac.uk/>)

<sup>1</sup>Intelligent Vehicles, WMG, University of Warwick, Coventry, United Kingdom {dannier.xiao, m.dianati, r.woodman, william.goncalves-geiger}@warwick.ac.uk

<sup>2</sup>Assisted and Automated Driving, Jaguar Land Rover Ltd, Gaydon, United Kingdom wgoncalv@jaguarlandrover.com

In the context of deep learning models, graph-based neural networks have been used to learn hazardous lane changes and collision scenes, and demonstrate 16-30% higher accuracy [20], [21], 39-62% earlier prediction [20], [22], and up to 9 times faster inference time [20] compared to non-graph counterparts. However, it is important to note the strength of graphs to encode relational dependencies also brings limitations as researchers must manually define the structure and scene features to sufficiently represent the problem. Thus, creating abstractions that may miss patterns unknown by current human understanding. Nevertheless, given the potential of graph-based methods, the area remains under-explored. To contribute to this gap, this review takes a novel focus on graph-based methods to facilitate relational reasoning.

The topic of hazardous event detection for AVs has been covered by several reviews and survey articles in the literature [8], [23]–[27], although works focus on non-graph methods. In the early works of [25], hazardous event detection methods are surveyed for early collision avoidance systems, which use vehicle detections and distance measurements. This was followed by more advanced methods in a later review [8], which categorised methods by how actor motion was predicted to forecast colliding trajectories or match hazardous behaviour patterns. The key categorisation made by [8], clearly divided the landscape of methods into deterministic physics-based, probabilistic manoeuvre-based and machine learning (ML).

The review in [8] also exemplifies a common limitation: reviews lack coverage on methods that can represent the earlier described actor, environment and regulatory relationships that underpin vehicle behaviour and help explain scene evolution. Usage of graph-based methods to represent such relationships has increased organically and are covered in the more recent reviews of [27], but remain brief and are limited to Bayesian networks, which allow variable relationships to be defined.

Moreover, existing reviews of the literature have restricted domain coverage and primarily review vehicle-based hazards and omit other hazard classes, such as environment (e.g., adverse weather, visibility) or traffic laws (e.g., speeding, illegal manoeuvre). Incomplete coverage limits real-world usage as it does not cover the variety or complexity in the driving domain, exemplified at intersections, which remain the highest site of AV incidents [28], [29].

To the best of the author's knowledge, related survey and review papers also do not cover the datasets to train such methods or the evaluation metrics to assess performance.

Motivated by the above gaps in the literature, this review aims to focus on graph-based methods for hazardous event detection. We present the opportunity of using graphs to decompose a complex operational environment, organise heterogeneous knowledge about the environment and represent relationships between scene entities (e.g., road users, static objects, traffic rules).

To this end, this research contributes a comprehensive literature review of graph-based hazardous event detection methods, evaluation metrics and datasets for training and evaluation. As such, the contributions of this review paper are:

- Comprehensive categorisation scheme of hazard classes
- Unique categorisation and focus on graph-based methodologies for hazardous event detection
- Thorough overview of hazard-focused datasets

- Thorough overview of model performance metrics
- Highlight of key gaps and opportunities within the area

Section I outlines the paper's focus and rationale. Section II defines key terms and concepts of hazardous event detection. Section III, comprises a focused literature review of key graph-based methods, performance metrics and datasets to train and test methods. Section IV then evaluates the advantages and limitations of each method and assesses opportunities for future research. Finally, Section V, concludes with our key findings.

## II. HAZARDOUS EVENT DETECTION

This section presents the key terminology used in the field and is tabulated in TABLE I. Followed by important background concepts of hazardous event: lifecycle, categorisation scheme and detection in the context of the AV pipeline.

### A. Key Terminology

To define the terms used in this research, we present key terms in TABLE I and follow the convention BSI 1890 [30], a vocabulary standard for automated vehicles and ISO 21448 [31], which covers safety standards of road vehicle systems due to external factors. We note that we concentrate on safety assessment at runtime operation, and though this standard applies to system assessment, the definitions are still applicable and consistent with related safety standards [32], [33].

Starting with the formal definition of hazard, it is any "potential source for harm" [31]. In systems engineering, this source is from the system under test, but in our context, we focus on external factors (e.g., other road users or roadside debris) that can cause harm in the form of physical injury or damage to property. Specifically, we focus on naturalistic external hazards without malicious intent. Examples include actor-based hazards (e.g., vehicles, pedestrians), environmental (e.g., adverse weather), and regulatory traffic laws.

However, a hazard alone does not materialise harm. Instead, hazardous events combine a hazard with an operational situation that realises harm [31]. For example, an icy road alone does not cause harm unless combined with a loss of control scenario that materialises harm by causing a collision. In the

TABLE I. KEY TERMS, BSI 1890 [30] AND ISO 21448 [31]

Term	Definition
<b>Hazard</b>	Potential source of harm [31].
<b>Hazardous Event</b>	Combination of a hazard and the operational situation of the vehicle [31].
<b>Hazard Detection</b>	Identification of hazardous events.
<b>Hazardous Event Analysis</b>	Method to identify and categorise hazardous situations and their triggering events to prevent or mitigate harm.
<b>Triggering Event</b>	Specific conditions of a driving scenario that serve as an initiator leading to a hazardous event [31]. E.g., known hazardous environmental conditions or exceeding the limitations of system components [31].
<b>Scenario</b>	Description of the temporal development between several scenes in a sequence of scenes [30] [31].
<b>Scene</b>	Snapshot of the environment including scenery, dynamic elements, all actors and their relations [31].
<b>Ego Vehicle (EV)</b>	Subject vehicle of primary interest in testing, trialling or operational scenarios [30].
<b>Risk</b>	The product of likelihood and severity [31].



Fig. 1. Hazardous event lifecycle example of environmental occlusion that masks the presence of oncoming traffic and cause collision.

previous definition, the term scenario refers to the time-based (i.e., temporal) development of a sequence of scenes. The scene describes the driving environment of actors, elements and scenery. The term actor refers to other roads users surrounding the subject vehicle, called the ego vehicle (EV). Furthermore, as hazardous event detection is a precursor for risk assessment, these terms are not well differentiated in the literature. Hazards are potential sources of harm, whereas risk is the evaluation of a hazardous event in terms of likelihood and severity.

It is vital for hazardous event detection to detect as early as possible while there is time to act; thus, we must identify the causal factors that initiate a hazardous event. Correlation is not causation, and it is the triggering events that need to be identified for early detection.

### B. Hazardous Event Lifecycle

It is important to note that a hazard by itself does not materialise harm. Harm is only realised when a hazard is combined with a driving scenario that materialises the harm. Such scenarios are classed as hazardous events and can be defined in a lifecycle that highlights the importance, but difficulty of early detection as future evolution must be predicted accurately in a complex and dynamic domain. The lifecycle is based on the literature [18], [34], [35] and adopted by the UK Driver and Vehicle Licensing Agency [36]. We define the stages below and exemplify them in Fig. 1.

- **Potential Hazardous Event:** Current scene does not pose immediate harm but may develop into a scenario that may cause harm
- **Developing Hazardous Event:** Current scene is developing into a situation that can cause harm if intervention is not taken soon
- **Materialised Hazardous Event:** Harm realised

As the hazardous event evolves, the detection stage evolves with the event. Beginning with early detection, then imminent detection, which requires immediate intervention to avoid harm and ends with post-detection, after harm is realised.

To illustrate the lifecycle, we present realistic examples that demonstrate the importance but difficulty of early detection as systems require a causal understanding of the triggering event.

Fig. 1 shows three scenes that depict an environment-based hazard of occlusion that impede the detection of oncoming vehicles and thus, could lead to the hazardous event of a collision. In scene 1, the hazardous event is potential at this stage, as the scene does not present immediate harm, but pre-emptive action here is crucial as it may develop to harm. If

undetected, the EV could begin the turn as the incoming vehicle approaches and evolves into a developing event as harm is more imminent in scene 2. Imminent detection at this stage may not allow enough time to act and could result in the collision shown in scene 3. These situations are difficult to formalise as they require an understanding of causality and counterfactuals. Also known as what-if questions, for example, "what if there is an oncoming vehicle behind these trees?".

Another common example can be seen in the regulatory category. For example, upon approaching a distant traffic light that has been green for a while. A human driver may slow down as they consider the counterfactual of "what if the light starts to turn now?". However, this pre-emptive reasoning is difficult to formulate as it requires a causal understanding. It is not just the correlation that amber and red require you to yield but teaching the relational significance of green, amber and red.

These examples of pre-emptive reasoning highlight the difficulty of early detection for machines but remain essential to avoid harm. However, early detection remains a current gap and thus, has gained traction in literature due to inferior hazardous event detection performance in complex environments, typical of the real domain [37], [38].

### C. Novel Hazard Categorisation Scheme

A hazard represents a potential source of harm and can originate from various sources during driving. For example, hazards can originate internally from the system due to the occupant (e.g., condition, misuse) or the system (e.g., hardware, software). Alternatively, harm can be external to the system, such as other actors (e.g., vehicles, pedestrians) or the environment (e.g., debris, ice). Harm can also be malicious and security related (e.g., malware, denial-of-service). Furthermore, harm may not be physical, but also economic or legal by violating regulatory traffic laws.

In our survey, no single categorisation scheme covered both internal and external hazards or exhaustively captured their diverse sub-classes. Subsequently, we propose a novel categorisation scheme derived from the literature [26], [39]–[41], ISO standards [31] and large projects such as PEGASUS, which provides a structured categorisation of highway scenes [42], [43]. We note the popular five-layer highway scene segmentation scheme in PEGASUS (L1: Road-level, L2: Traffic infrastructure, L3: Temporary manipulation of previous layers, L4: Objects and L5: Environment) [42], [43]. However, PEGASUS is limited to highways and focuses on categorising the scene components that can be hazards but does not focus on categorising the hazards themselves.

Alternatively, our novel categorisation aims to categorise the hazards irrespective of the scene to provide generalised high-

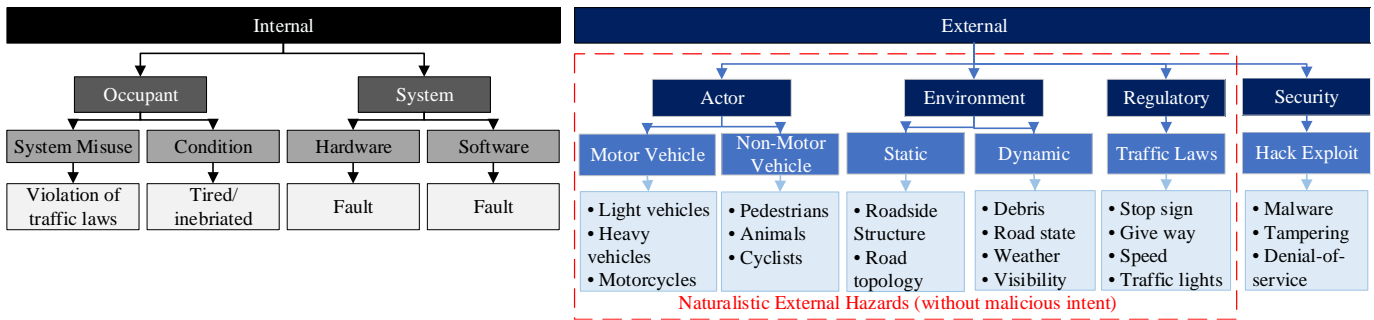


Fig. 2. Categorisation of hazards in the driving domain.

level hazard classes. By focusing on hazards, we aim to allow future systems to attach semantic meaning to class identifiers such as "actor" to better understand their impact and behaviour.

Our research focuses on naturalistic external hazards (e.g., actor, environment and regulatory), highlighted in Fig. 2. Naturalistic hazards that are not from malicious intent, such as security-based hacking factors or internal factors, as we believe these are separate in-depth topics in themselves.

#### Actor

**Motor Vehicle:** This sub-class includes self-powered vehicles such as cars, vans, lorries and motorcycles. Hazardous events are typically collisions or harm from an evasive manoeuvre to avoid a collision. Triggering events can be behaviours such as overtaking and lane change. The behaviour of larger vehicles is typically more predictable due to traffic laws and the structured driving domain.

**Non-Motor Vehicle:** Includes pedestrians, bicycles and animals that can cause collision or harm due to evasive manoeuvres. This class is more challenging to predict as directions can change rapidly with unconstrained movement as many pedestrian accidents occur due to violation of traffic laws and inattentiveness [44].

#### Environment:

**Static:** These hazards are stationary in time and represent more permanent entities. It consists of obstacles, which can be roadside structures such as bollards, islands or buildings. Hazards can also come from road geometry, such as sharp corners that can lead to a loss of control or from complex road topology, such as intersections, which are the largest site of AV incidents [28], [29].

**Dynamic:** These hazards are dynamic in time and consist of adverse weather conditions that can make roads slippery due to rain, ice and snow and can affect visibility for vision-based sensors. Also included are temporary debris and road condition such as uneven or damaged surfaces that must be avoided and can affect the behaviour of other road users.

**Regulatory:** Includes traffic laws and are usually detected using map priors that contain road rules (e.g., right of way, speed limit) or by using live detection of traffic signal infrastructure. Systems must robustly detect the driveable region and be aware of the applicable regulations to avoid violating traffic laws, for example, staying under speed limits and who has the right of way.

#### D. Hazardous Event Detection Process

Due to the prevalence of dangers in the driving domain, a robust hazardous event detection framework is critical to ensure

a safe driving policy for short-term automated driving systems and the long-term transition to autonomous systems. The traditional automated driving pipeline starts with perception, which uses multiple sensors such as RADAR (Radio Detection and Ranging), camera and LIDAR (Light Detection And Ranging) to sense the surrounding environment. These sensors collect low-level data that is then processed with perception algorithms to extract object detections, classification and trajectory predictions of actors and elements in the scene. Scene information and EV intention from the decision-making and planning modules are then passed to the runtime hazardous event detection module to identify all potential sources of harm (i.e., hazardous events). Potential hazardous events are then passed to a risk assessment module which evaluates each in terms of likelihood and severity to produce a risk score for the decision-making module to plan mitigating action.

The hazardous event detection goal is to identify situations where a hazard may lead to a hazardous event either known or unknown to the system. Thus, the challenge is how to process vast amounts of perception data to represent a driving scene with sufficient detail [45], [46] for detection and then how to generalise hazardous events to detect unseen scenarios.

The level of hazard awareness required is dependent on the complexity of the automated driving function, with minimum distance thresholds sufficient for Level 1 emergency braking, but insufficient for Level 2 adaptive cruise control. This survey focuses on the application to help transition to higher levels of autonomy Level 3+, where AVs will need to safely control all driving functions. As human behaviour plays a large factor in road safety, this review aims to present readers with the current landscape of frameworks and inputs to model or learn such interactions between actors, the environment and traffic rules.

Graph-based methods show a promising opportunity to employ relational reasoning to incorporate context and infer interactions, however, the desired input to sufficiently represent a hazardous driving event is yet to be defined. Consequently, this survey aims to collect and present researchers with the current landscape of approaches to inspire applications that refine and contribute to the state-of-the-art.

### III. REVIEW OF GRAPH-BASED METHODS

This section begins with a background to graph theory. We then review the state-of-the-art hazardous event methods and divide into the most prominent categories: knowledge rule-based, probabilistic and machine learning-based and discuss each individually. A summary of each method, its advantages and limitations are provided in Table II. Followed by an

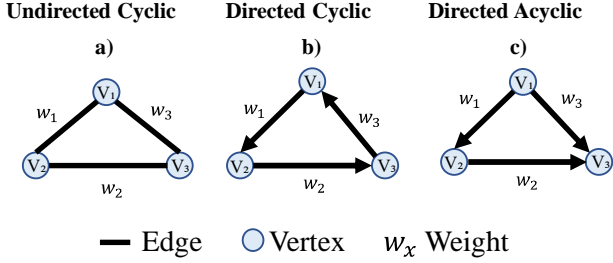


Fig. 3. Graph types: (a) undirected cyclic (b) directed cyclic (c) directed acyclic.

overview of performance metrics and hazard-focused datasets to evaluate methods. The authors also note that as hazardous event detection and risk assessment are so intertwined, many papers adopt the term risk as it encapsulates hazardous event detection, but terms are not well differentiated in literature.

### A. Graph Networks Background

In an inherently interconnected driving domain, this research aims to utilise graph-based methods to represent these relations as a key paradigm shift for knowledge representation and inference [47]–[49]. As is the limitation of non-Euclidean data structures that cannot represent complex relationships using  $n$ -dimensional space without losing essential node or vertex features or the directionality of relations and how they interact. Instead, graph structures can store meaningful representations of the domain, such as driving scenes with variables, such as actors, as vertices (i.e., nodes) and their relationships to other entities as edges.

A graph  $G$  is formally defined as a pair of sets  $G = (V, E)$ , where  $V$  is a set of vertices and  $E$  is a set of edges [50], [51]. Each edge  $e$  connects to at least two vertices described by  $\{x, y\}$  pairs denoting originating and terminating vertices,  $x, y$  respectively. This relationship can be directed or undirected to show the direction of influence, and thus, each  $\{x, y\}$  pair can be ordered or unordered.

Edges can also be weighted ( $w_x$ ) to encode semantic information such as distance between actors or spatial relationships (e.g., in front, behind). This unique data structure allows cause and effect relations to be used directly for hazard event detection or to organise domain data into relational structures for feature learning.

In an undirected graph, the edges are bidirectional, with no directional restrictions. If an undirected graph has connections between all vertices, this also allows cyclical traversal, as shown in Fig. 3a,b. In this configuration, vertex traversal is unrestricted, and thus, edge descriptions  $\{x, y\}$  do not form ordered pairs. These represent symmetric unidirectional relationships such as a two-way road network. Conversely, directed graphs have directional restrictions, so edge descriptions form ordered pairs between originating and terminating vertices  $\{x, y\}$ . A directed graph can be cyclic or acyclic if it is not possible to return to the starting vertex, as visualised in Fig. 3c. The latter has no connection from vertex three back to vertex 1, disconnecting the cycle into an acyclic graph. This ability to map directional relationships allows graph structures to represent cause and effect and isolate triggering events needed to pre-empt hazardous events.

Directionality defines reasoning capability and can be deductive, inductive and abductive. Deductive reasoning is a

TABLE II. SUMMARY OF METHODS

Method	Description	Papers
<b>Rule-Based Knowledge</b>	Knowledge bases are built to describe hazardous events and the driving domain, from which rules are formulated to match known hazardous scenarios to current observations. Rules can range from simple heuristics comparing the current speed to speed limit to combining several factors such as actor location, heading and speed.	[18], [35], [55]–[61]
Ontology		
<b>Probabilistic</b>	Hazardous events modelled by determining the conditional probability of scene elements affecting the hazardous event. Networks can propagate uncertainties from input variables to prediction output.	[19], [67], [69]–[74]
Bayesian Networks		
<b>Machine Learning</b>	Machine learning (ML) approaches learn hazardous event patterns between traffic entities using large datasets. Thus, less interpretable than the other methods but allows better generalisation and scalability.	[20]–[22], [79]–[82]
Graph Neural Networks		

one-directional top-down approach. Starting from a hypothesis, such as the presence of a hazard, and confirming these using observations and known relationships. As it uses known relations, it does not support reasoning outside the scope of learning and to achieve this; bottom-up approaches are used to reason from effect to cause. Inductive reasoning is an example of this which uses detailed observations to empirically identify patterns and relations to infer a hypothesis. In practice, obtaining comprehensive observations is not always viable, and subsequently, abductive reasoning can be used to infer the most plausible conclusion with incomplete data. Abduction uses its limited observations to "explain away" causal hypotheses. For example, predicting whether it is raining; if pedestrians have umbrellas, the likelihood of rain is higher. Alternatively, in countries where umbrellas are used to shade users from the sun, observing both umbrellas and the road being wet is required to reduce the likelihood of sun and reaffirm rain.

Given these properties, graph networks are an ideal platform to model hazardous events by providing a relational data structure that can aggregate heterogeneous data from various sensors while being able to represent traffic objects and their relationships. Using expert knowledge and ML, hazardous events can be decomposed by defining their characteristics and interdependencies for earlier and more robust detection.

This process imitates the high-level reasoning of human drivers that process the constant stream of perception to form a cognitive abstraction of the scenario. Abstraction by understanding the relationships between the objects and events to infer situation semantics. This is comparable to high-level data fusion to reduce both data volume challenges with asynchronous and heterogeneous information fusion to build higher-level hazard descriptions.

### B. Method Categorisation

This survey reviewed three key categories within graph-based methods, as shown in TABLE II. Starting with rule-based knowledge methods that form knowledge bases and use rules to match known hazardous combinations with current scenes. This is followed by probabilistic methods that use a combination of probabilistic events as evidence to update the probability of a hazardous event occurring. Probabilistic methods can incorporate prediction or perception uncertainty which is invaluable in the complex driving domain. Lastly, we review the opportunity presented by the ML approaches that use accident databases to learn features and hazardous event

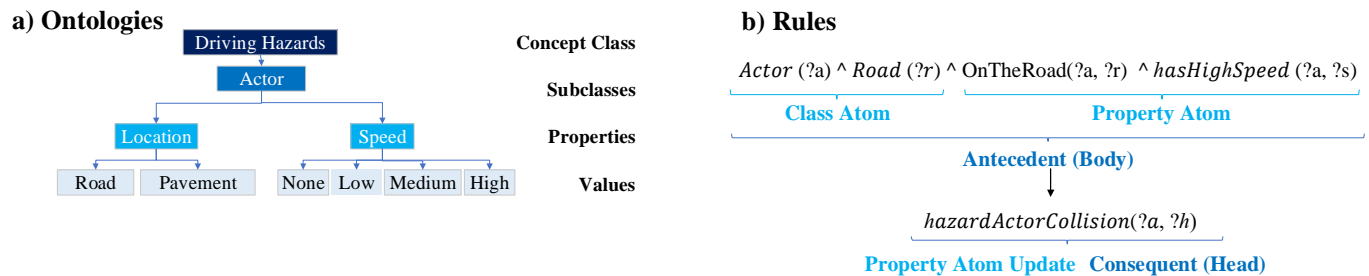


Fig. 4. (a) example of driving hazard class hierarchy (b) SWRL example of actor-based collision rule

patterns between traffic entities. This represents an under-exploited area that could help resolve the intractable problem of explicitly defining all hazardous events through supervised or unsupervised learning.

### C. Rule-Based Knowledge Methods

#### 1) Overview

The first set of methods uses rules to detect hazardous events, with implementations that range from simple decision trees which become difficult to scale to more advanced knowledge database implementations that formally describe complex domains and allow rules to be written that define hazardous events as derivatives from the domain. We focus on the latter method as a scalable and widely applicable approach and are referred to as ontology-based methods.

#### 2) Ontology Background

An ontology has many definitions in the literature; for this research, it is the formal description of concepts (i.e., anything that exists) in a domain as a definition of classes and subclasses, each with associated properties and values which are related using a hierarchy [52], [53]. Written in ontology web language (OWL), formulated by the World Wide Web Consortium initially to describe Web data [54]. OWL is a logic-based language that allows knowledge to be used computationally using rules and statements.

Rule-based languages and logic can be similarly used for hazardous event detection to translate hazardous event concepts into first-order logic. To detect specific hazardous events, hazard classes are modelled using all associated properties and relationships. For example, an actor hazard can have properties describing location and speed, as shown in Fig. 4(a). In ontology methods, semantic web rule languages (SWRLs) are commonly used to define rules on top of ontologies, as shown in Fig. 4(b). In this example, an actor collision event rule is defined using parameters referred to as atoms which are linked by conjunction logic denoted  $\wedge$  to form the antecedent (i.e., body) leading to a consequent result (i.e., head), denoted  $\rightarrow$ . This rule can be interpreted as follows: if the domain contains an actor *and* road and the actor is on the road *and* at high speed. The result is a potentially hazardous actor collision. SWRL link ontologies by describing the domain with query rules to identify hazardous events.

The ontologies defining hazardous events are then compared with information describing the current domain from sensor data, stored in a set of assertional axioms to infer hazardous events. Thus, the ontology-based framework in hazardous event detection follows three key steps: (1) Ontology modelling to derive terminological axioms of key rules, classes,

and relations. (2) Description of the current scene with perception input stored as assertional axioms. (3) Inference by comparing axioms with rules to detect hazardous events.

#### 3) Ontology Methods

As a popular method for aggregating perception data to infer semantic information about scenes, ontologies have been used to organise detected traffic entities from perception sensors into an ontology that describes the entities with a hierarchy of relations and interactions. Hazardous events are then inferred using rules defining which combination of entities and interactions describes a hazard.

As proposed in [55], [56], the authors evaluated pedestrian-based hazards in different traffic scenarios with a monocular camera. Pedestrians were detected and modelled as objects with attributes of speed, location and motion direction. SWRL rules were used to define whether pedestrians could cause harm if they were on the road *and* at high speed *and* posed toward the EV. Due to the limited sensor suite of only a monocular camera, this technique was sensitive to illumination with limited accuracy for object localisation. As determining pedestrian distance, speed and location were vital and only pixel and frame data were available, this limits practical usage. However, the SWRL ontology shows a simple method to generalise actor hazards by abstracting motion behaviour and scene semantics.

In addition to actor hazards, environment considerations were modelled in [57] as it focused on off-road hazards. Hazards such as road type (path, path side), adverse weather and occlusion we considered. Particularly for adverse weather, the effects of fog, snow, and rain were modelled as it creates both visibility and road surface traction hazards.

Using SWRL to define hazardous events within the scene, soft body obstacles were given semantic attributes to avoid prompting evasive manoeuvre as they don't impede motion.

The final hazard category is regulatory, and in [58], traffic regulation-based hazards were studied using ontological descriptions of road segments and intersections to contain road sign classifications and generic right-of-way rules. Similarly, SWRL was used to connect ontologies to detect traffic rule violations, using a modular framework to be country agnostic; this method is promising but is yet to be validated with testing.

Some state-of-the-art hazardous event detection works can also be found in recent research applied to construction [59]–[61] and are included due to their highly transferrable methodologies. The majority of the methods [59], [61] use computer vision first to detect scene entities and their attributes which are related in expert-defined ontologies and then queried using rules to identify hazardous events.

In [59], [61], hazardous events are based on spatial features such as the distance of workers from machinery and their



Fig. 5. Example of abductive reasoning, adapted from [21].

positional relations calculated by bounding box overlap (e.g., inside, overlap, outside) to signal hazards such as a PPE violation of a worker not wearing a helmet if the bounding box of the helmet did not overlap with the worker. The key limitation of this approach is only utilising spatial relation, which may not be robust in cases of occlusion or lack of observable data. A hazardous event has more attributes than spatial relations, and thus, [60] added semantic context such as activity and duration to better generalise hazardous events, which were then compared to the current scenario for similarity.

SWRL rules represent a simplistic deductive reasoning approach to detect known hazardous combinations of scene entities which limits reasoning to the scope of learning. To transition to autonomy, there will always be unseen scenarios and inherent uncertainty for which explicit rules may not be sufficient. To find the most plausible conclusion given incomplete observations, abductive reasoning is necessary.

Reasoning from effects to the cause, abductive reasoning is used in the iterative works of [18], [35]. In [18], the authors assessed traffic hazards by compiling a vast knowledge base of 32 causal rules to relate traffic entities, such as "cars must stop at red traffic signals", 2077 ontologies to categorise domain entities and 11 hazardous behaviour rules describing such as "a person can rush out" [18]. The network then assessed semantic scene descriptions for hazards using a natural language abductive reasoner [62] to generate plausible hazardous event hypotheses. An example of abduction can be represented by a ball in the road observation leading to a hazardous event prediction. If there is a ball on the road, an owner could follow the ball. The owner could be a person, and this person could be a child. The owner of the ball could be hidden behind a house and thus, does not see the road. The person may run out to retrieve their ball, which may lead to a hazardous event. This abduction is exemplified in Fig. 5.

The reasoner reasoned with the knowledge base rules and was optimised using ML training to minimise prediction error by favourably weighting knowledge base rules that best identified hazardous events. As the method only uses high-level scene semantics, the inference is only qualitative and lacks rigour. This limitation is addressed by [35], who expanded the method by adding quantitative metrics such as actor position or velocity to derive trajectory and predict intention. The qualitative hazardous event predictions were then re-ranked using a physics simulator to predict actor trajectories based on physical plausibility. This method demonstrates how rule-based methods can be extended with ML approaches and incorporate both qualitative scene semantics and quantitative trajectory level predictions to improve the accuracy of hypotheses.

## D. Probabilistic Methods

### 1) Overview

These methods incorporate domain uncertainty by utilising conditional probabilities that a road element(s) lead to a hazardous event given some assumptions and uncertainties from input variables. With a wide range of probabilistic methods using hidden Markov models and Bayesian networks (BNs), this survey will review BN implementations due to their ability for omnidirectional inference, allowing both deductive and abductive reasoning and the use of continuous variables to represent the dynamic driving domain. Even under uncertain or missing inputs by using conditional probability [63], [64].

### 2) BN Background

Amongst the most popular of the methods, BNs allow omnidirectional (i.e., abductive and deductive) reasoning with uncertain or incomplete data in an interpretable graph structure [63]–[66]. A BN is an acyclic graph using Bayes' theorem shown in (1) to calculate a posterior probability (I). The posterior probability is the probability of hypothesis A, given an observation B, which has already occurred. The posterior probability is calculated by multiplying the probability of hypothesis A (II) with the likelihood (III), divided by the marginal probability (IV). Where the likelihood (III) is the probability of the hypothesis before evidence was updated. The marginal probability (IV) is the probability of the evidence being independent of the hypothesis. Part (III) and (IV) determine the strength of the evidence to affect the prior (II).

$$P(A|B)^I = P(A)^{II} \frac{P(B|A)^{III}}{P(B)^{IV}} \quad (1)$$

**I Posterior Probability:** Probability of hypothesis A, given an observation B, which has already occurred.

**II Prior:** Prior probability of hypothesis A, before evidence of observation B.

**III Likelihood:** Probability of the observation B, given a hypothesis A.

**IV Marginal Probability:** The probability of observation B, is independent of any hypothesis and is not always known, so marginalisation is used, and the unconditional probability of B is usually substituted.

Thus, the network describes a set of random variables  $X$ , with directed relations to other vertices and a conditional probability table (CPT)  $P(X_i | Parents(X_i))$  to quantify parent to current vertex relations. This creates a network of random variables  $X_1 \dots X_n$  and CPTs, which combine to form a joint probability distribution using the chain rule and assuming conditional independence that each variable  $X$  only relates to its parents. The joint distribution (2), with random variables  $X_i$  taking fixed values  $x_i$ .

$$P(x_1 \dots x_n) = \prod_{i=1}^N P(x_i | Parents(X_i)) \quad (2)$$

As a strength of BNs, the CPTs quantify cause and effect between random variables and uncertainty from data quality or



incomplete observations to produce a probability distribution for the prediction. To generate CPTs, authors generally use either expert scoring or ML approaches that relate a statistical occurrence to the probability. However, generation is challenging as either can be biased, with ML approaches proving difficult as the effect of random variables on hazardous events can be challenging to quantify and requires specific datasets that can be incomplete or imbalanced.

To represent relations and causal influence, the vertices are connected by directional arrows. However, the flow of information is unidirectional allowing deduction, induction and abduction [64], [67]. Moreover, the graph does not form cycles to avoid circular reasoning, which would make it challenging to trace causality to a triggering event. Events that evolve in time so have a temporal element that must be considered; BNs can account for this using dynamic Bayesian networks (DBNs) [66], [68]. DBNs model the relationship of discrete random variables between time steps through CPTs that calculate the conditioned dependence of the current step given states at the previous step. By this method, DBN overcomes the limitations of independent time step inference, so are prevalent in interaction focused hazard detection methods [8], [66].

The ability to trace causal relationships, use bidirectional reasoning and incorporate uncertainty makes it ideal for hazardous event detection, which is difficult to formalise as it is interconnected and complex. Thus, decomposing hazardous events by cause and effect makes the formalisation more manageable. In addition, the ability to propagate uncertainty enables the network to detect this as a hazard in itself if the input information or prediction is untrustworthy.

### 3) BN Methods

In an inherently uncertain driving domain, reasoning the influence of a combination of probabilistic events is the strength of BNs. Quantifying uncertainty is crucial for decision making and is utilised clearly in the work of [69]. The authors present a theoretical BN framework for platoon vehicles to detect hazardous events such as a collision between members and speed violations by comparing safe prior with current speed and inter-distance observations. Hazardous events are then to derive system states, from state 0 (safe distance, legal speed) to state 4 (unsafe distance, illegal speed) up to state 5 (insufficient sensor information). The novelty of this method was that each state predefines mitigating actions (continue, decelerate, stop) to intervene. Uncertainty in hazardous event detection is then propagated through probability distributions at each node with an added "DetectionQuality" node to flag unsure detections and incorporate such uncertainties in the probability distributions.

The method in [69] is a good representation of BN reasoning but illustrates a common limitation between methods. Hazardous event probabilities were aggregated to infer system states that reflect both actor collision and speed violation. As events were not separated, this led to a mixed network that does not easily scale. A limitation of many methods in the literature and is due to one graph structure for all reasoning.

An example of hazard category separation can be seen in [70], in which internal EV conditions, external environment and other vehicle-based hazards are modelled as independent subclasses in the BN. Using properties such as visibility, road conditions and vehicle density to reason hazardous environmental events whilst using velocity, distance, vehicle type and acceleration to determine if other vehicles may lead to

a hazardous event. EV conditions are also quantified by the dashboard fault warnings (good 0, moderate 1, bad >1).

Modelling hazardous events independently allows multiple types to be detected, but interactions between types were not captured. This reduces the predictive power of inference as it cannot propagate the effect of events between classes that are highly connected. As explored in [19], the authors focused on interactions between actors and the environment to predict the hazardous events of a vehicle collision and loss of control at corners, Fig 6. Using both measurements and vehicle-to-vehicle (V2V) information, the BN models a loss of control from the EV by considering the upcoming curve radius, driver reaction time, EV speed and max deceleration. The last three factors of this class also affect the probability of rear-end collision and thus are interlinked between hazard classes. Using these interrelations, rear-end collision also utilised front vehicle speed, distance and braking intention. The interaction of the environment is especially explored, with weather conditions (e.g., temperature, precipitation) affecting the modelled road state and maximum deceleration, which is used for prediction.

Another vital topic in the literature is the selection of random variables to model hazardous events and their associated CPTs to quantify the strength of relations. The random variables at the vertices are generally defined using expert knowledge and the CPTs through either experimental data or expert scoring. This means that models and their CPTs can be difficult to validate, especially as it is challenging to quantify the effect of random variables on event materialisation. To explore this challenge, we then evaluate methods that propose a data-driven model and CPT generation.

In [71], the authors searched for hazardous events at automatic railway level crossings and determined the most influential factors that cause vehicle accidents. The method uses real-world data on railway crossing accidents which are processed for causal discovery of hazardous events. First, causes were learnt through automatic structure learning methods such as Bayesian Search and augmented Naïve Bayes (ANB), which generates a BN by defining the hazard to be studied as the parent variable from which all remaining dataset variables are considered features and learns corresponding connections between features and CPTs from data. Six search models were tested using a dataset of 4,200 accidents, and each model was then manually optimised to remove trivial factors. The level of causality of hazardous event variables were then classified into three levels: primary, secondary and tertiary, along with the CPT tables to quantify the influence of variables.

Unfortunately, the limitation of empirical structure learning is that the algorithms tend to find observable correlations rather

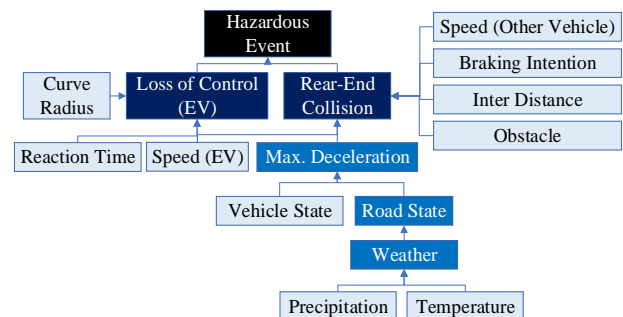


Fig 6. Hazardous event BN model, adapted from [19].

than causations which sometimes need to be inferred. As such, search algorithms are used in preliminary search stages with constraints to guide learning and require a manual post-review. Although algorithms do show progress, as [71] demonstrated promising results using ANB methods, achieving only a 5-13% performance difference (accuracy and AUC) to the expertly-reviewed network that culled erroneous or trivial connections. Alternatively, [72] investigated search methods to find highway hazards by first using expert knowledge to build the initial network and then a K2 greedy vertex search to expand the network. CPTs are then generated via an expectation-maximisation algorithm, and highway hazards are inferred using a junction tree method. The final method modelled five different hazard types (driver, vehicle, road, environment and management) with interactions between categories.

The methods above are limited to single time steps, but the temporal element is required and modelled using DBNs to model evolving interactions. To show this, [73] modelled driver intention using the temporal capability of DBNs. To detect vehicle collision hazards at intersections, the network inferred the expected action according to traffic regulation and compared it with temporal driver intention. The network utilised both observable vertices of vehicle position, heading, and speed and hidden vertices to represent behavioural variables that are not directly observable but inferred. As one of the few methods tested in the field at a controlled intersection, data on vehicle metrics was gained through V2X, limiting practical applicability if this is unavailable or not reliably detectable. Nevertheless, it highlighted the key that hazardous event detection requires knowledge of developing interactions for robust prediction [73].

As the importance of interactions becomes more apparent, so does DBN implementation with further development, as seen in [74], in which the authors explicitly model inter-vehicle dependencies by grouping scene information in organised layers. Starting with a first layer containing observable sensor measurements followed by a kinematic layer describing ego position, speed and heading. This then feeds into a driving context layer that reasons the hazardous events and associated risks. Using the DBN to process the temporal data from previous layers, hazardous events on a vehicle level are then detected from geometric and dynamic relationships between vehicle motion. Scene-level hazards are also identified using ML classifiers to identify potentially hazardous traffic conditions by considering the average speed of scene actors and vehicle density. This method allows the network to not identify potential vehicle collisions from their temporal interactions but pre-empt hazardous driving conditions.

## E. Machine Learning Methods

### 1) Overview

This category represents an opportunity for further research to use feature learning over graph networks. A feature is representative of a property or characteristic that is learnt and provides a level of abstraction [75]. As it is intractable to specify all hazardous events, this could synergise with current methods by finding new causal links or generalising using data.

Traditional ML methods cannot be used directly on graph networks due to the irregular data structure, no fixed vertex order and dependency between vertices which need to be considered during learning. For reasoning using graph

networks, a key approach is using Graph Neural Networks (GNNs) [76]–[78]. As an area that we feel is under-exploited, GNNs incorporate ML techniques that can harness the relational and semantic data uniquely captured by graphs.

### 2) GNN Background

GNNs are a group of neural networks (NNs) that perform a learning task with graph data. Each NN represents a learning algorithm that uses a network of functions (neurons) to learn a representation of a given input, as inspired by neural activity within the brain. GNNs present a compelling opportunity to learn the complex evolution of hazardous events due to their ability to apply ML to extract features from graph data that depicts a relation hierarchy linking the driving scene [76]–[78].

GNNs take graph structures as input to learn the spatial features by using information aggregated from neighbouring vertices which are then combined using NNs repeating over all neighbouring vertices to preserve connectivity. The NNs perform permutation invariant aggregation of neighbour features, combined using a separate NN into an embedding that is a generalised representation of the graph [78].

GNNs come in many variations, such as convolutional for vertex and graph classification, autoencoder-based for link prediction and temporal for time-series forecasting [77]. The methods within this survey were primarily convolutional, spatio-temporal or a combination of methods.

- **Convolutional Graph Neural Networks (CGNNs):** Methods transform the convolutional operation from grid to graph data. Convolutions are used to aggregate a vertices' features with their neighbour's features to form an aggregated representation. Due to their ability to aggregate local vertex representations, these methods form a basis for building other models.
- **Spatio-Temporal Graph Neural Networks (STGNNs):** Methods consider both the spatial and temporal features. For spatial learning, the model processes each vertex to create vertex representations that contain information about each vertex, and connected edges interact and relate with their neighbours to create generalised representations. Using the generated vertex representations, the temporal dimension then learns their evolution over time to create a combined spatio-temporal embedding for prediction.

### 3) GNN Methods

Not all hazardous events can be explicitly defined, and thus, the challenge is learning features to characterise new or existing hazardous events to continuously improve detection. To be robust in the driving scene, we must extract features from the complex interactions between the environment and vehicles.

With an ability to learn patterns from relational data, GNNs are prevalent in related topics of scene understanding and vehicle interaction by relating scene semantics in a graph structure [79], [80]. By focusing on hazardous interactions, GNNs can easily extend to hazardous event detection by representing actors and traffic objects in a relational graph and learning spatio-temporal patterns to predict hazardous events. A few authors have demonstrated this by using Bayesian NNs [22] or long short-term memory (LSTM) networks in [21], and have shown promising results but remain under-exploited.

Detecting dynamic hazards, such as other actors, requires spatio-temporal data to track their trajectory relative to EV intention. As the optimal solution is unclear, the extraction and encoding methods to represent spatio-temporal evolution from video inputs have evolved over time. In the related works of [79], [80], the authors had the aim of classifying driver behaviour (e.g., lane change, overtake) and represented the scene as a graph of actors as vertices and used edges to encode semantic spatial (e.g., right, left) and semantic temporal data between actors (e.g., moving forward). At each time step, monocular camera images were used to detect, track and project actors to bird's eye view (BEV). BEV allowed spatio-temporal calculations represented in an interaction graph and encoded using a multi-relational graph convolution network (MRGCN) for an LSTM to predict temporal vehicle behaviours. The novelty in this multi-layer approach was using one layer for each actor relation to which graph convolutions were applied to generalise the combination of temporal relations that could predict vehicle behaviours. However, complexity was limited as spatial relations were semantic, with only vehicles and lanes.

As spatio-temporal relations are key for describing hazard evolution, methods have naturally expanded to capture more complex evolutions. In [81], authors classify more complex vehicle interactions to predict actor goal's (e.g., give way) and cause (e.g., crossing vehicle). Separate spatial relation scene graphs are encoded to reason dynamic actor interactions (e.g., car, pedestrian, bicycle, bus, train) and the effect of static traffic regulation objects on interactions (e.g., road markings, traffic signs, road infrastructure). This is done to explicitly define which actor interactions or objects affect actor behaviour. Secondly, this differentiation is done as traffic infrastructure can be diverse and thus difficult to detect or contain within rectangular bounding boxes. Using video as input, instance and semantic segmentation are used to detect actors and objects from which spatial relations between them are calculated using Euclidean distance. Graphs are then formed to model actors in one graph and regulation objects in a separate graph. CGNNs were then applied to encode graphs into vector representations, followed by temporal fusion with element-wise max pooling and finally fed to the classifier.

Previous authors calculated spatial information from monocular image input, but this contains limited depth information. In [82], again, behaviour is forecasted, but using LiDAR and HD maps to help calculate more accurate spatial and motion predictions from LiDAR point clouds. What is novel about this study is the probabilistic GNN, inspired by Gaussian Markov random fields to forecast motion by modelling spatio-temporal interactions. The method takes actors as vertices to build a fully connected directed graph trained to model the cause and effect of each pair of actors. Probabilistic prediction also allowed uncertainty to be propagated to prediction for more informed interpretation.

With advanced spatio-temporal and GNN processing in related fields, methods can extend to hazardous event detection by capturing complex spatio-temporal evolutions. As demonstrated in [22], a probabilistic GNN was created by pairing a CGNN with a Bayesian NN, with a focus on preemptively detecting actor-based collisions in traffic accident videos. Object detections were fed as input into a CGNN to encode spatial embeddings for a recurrent neural network (RNN) to capture temporal patterns. Actor collisions are then predicted with the Bayesian NN to propagate uncertainty and

evaluate accident likelihood; however, the limitation was interpretability due to the RNN hidden states.

Alternatively, in the iterative works of [20], [21], detailed semantic descriptions and attention mechanisms were used to improve the interpretability of cause and effect. The authors identified hazardous events within videos of lane-change manoeuvres using both real and synthetic data for transfer learning. Dynamic actors (e.g., vehicles and pedestrians) and static traffic objects (e.g., lane markings and traffic signs) were detected in image frames and extrapolated into BEV to calculate spatial relations. Actors and objects were then represented as graph vertices with semantic spatial relations stored as edges. Semantic relations included distance (e.g., near ~5m, super\_near ~2m), relation (e.g., rear\_right, right\_front) and an "isIn" relation to describe which lane actors are within (e.g., left lane, middle lane, or right). A MRGCN was then used to convert the graphs structures to a vectorised representation and processed with a LSTM network to output a spatio-temporal sequence and then classified with a multi-layer perceptron. Hazardous events from spatial and temporal features are then learned through training data. In addition, interpretability was increased by adding a temporal attention layer to the LSTM to rank the effect of hidden states on classification outcomes. Thus, this method uses detailed semantic spatial descriptions and attention mechanisms to improve the interpretability of hazardous event detection.

#### F. Hazardous Event Quantification Metrics

With the variety of different hazardous events studied in this survey, each utilised different metrics to identify hazardous events. Due to this variety, we examined the metrics utilised for each event category and summarised these in TABLE III.

**Vehicle Actor:** These hazardous events involve collisions, which are usually quantified based on intersecting trajectory or predicted behaviour. As such, metrics centre around spatial and temporal features. Over half of the methods within this survey used spatial features to either semantically describe the relative location (e.g., front, behind, left, right) in the case of [18] or

TABLE III. HAZARDOUS EVENT QUANTIFICATION METRICS

Actor		Environment		Regulatory
Vehicle	Non-Vehicle	Static	Dynamic	Traffic Laws
<b>Spatial</b>	<b>Spatial</b>	<b>Road Type</b>	<b>Weather</b>	<b>Rules</b>
[18], [19], [21], [22], [35], [57], [58], [69], [70], [73], [74], [79]–[82]	[18], [22], [35], [55], [56]	[19], [57], [58], [72]	[18], [19], [57], [70], [72]	[58], [69], [72]
	<b>Temporal</b>	<b>Geometry</b>	<b>Visibility</b>	
	[22], [35], [55], [56]	[19]	[57], [70]	
<b>Temporal</b>	<b>Behaviour</b>			
[19], [21], [22], [35], [69], [70], [73], [79]–[82]	[18], [35]			
<b>Behaviour</b>				
[18], [19], [35], [73], [79]–[82]				

quantify actor location and pose in [35] and inter-vehicle distance within [19], [69]. Similarly, temporal features were used to represent speed, velocity and trajectory. Spatio-temporal features were then usually processed to derive physics-based colliding trajectory predictions such as TTC and equivalents or used to predict future behaviour such as braking intention [19] or intended manoeuvres [18], [35].

**Non-Vehicle Actor:** Similar to vehicle actors, a collision event is quantified based on spatio-temporal features to search for colliding trajectories. The key difference is the unstructured nature of pedestrian trajectory, which is slower but can change direction rapidly and is less predictable than vehicles restricted to road rules and lanes.

**Static Environment:** To detect static hazards within the environment, [19] utilised road geometry to predict a loss of control event at sharp corners by considering curve radius, current speed and maximum deceleration. Road type can also be used to predict hazardous regions within specific domains such as intersections, which have more complex interactions, traffic density and regulation or [57] who studied hazards in off-road driving. Off-paths were also classified as more hazardous with more obstacles and uneven surfaces.

**Dynamic Environment:** Methods reviewed in this survey primarily focused on adverse weather conditions as they can affect visibility and road traction. As considered in [18], who detected rain, snow and ice with more complex implementations using temperature and precipitation to predict road traction to forecast max deceleration and avoid rear-end collision in [19]. Along with creating adverse road conditions, weather conditions can also change visibility which affects vision-based sensors, and as such, methods also quantified this hazardous event by semantically classifying the range of sight (short, near, far) [70] and phenomena such as fog, snow and rain in [57] to trigger more cautious driving policy.

**Regulatory:** To avoid violating traffic laws, methods used current speed information to check against known speed limits [69], and more complex methods linked traffic regulations into different road types to interrogate whether the current traffic participants were abiding by rules [58].

### G. Performance Metrics

As hazardous event detection aims to predict hazardous events from scene entities, this is both a classification problem to detect if there is potential harm and a time-based problem to detect early. As such, classification metrics such as accuracy, precision and recall are used, and metrics like time-before-collision for time-based evaluation. It is noted that a common limitation is the subjectivity of "hazardous events" if harm is not realised. Therefore, qualitative analysis and comparisons are also commonplace to contextualise performance.

Regarding classification tasks, it is important to note key terms of positive and negative detections. In the context of hazardous event classification, a positive detection would be a hazardous event detected, whereas negative detections refer to a non-hazardous classification. Whether these predictions are true or false are essential to distinguish and can be represented in a confusion matrix, which has a total of four outcomes that are used to calculate performance metrics:

- **True Positive (TP):** Predicted hazardous and the scene is hazardous.
- **False Positive (FP):** Predicted hazardous but is safe.

TABLE IV. PERFORMANCE METRICS

Category	Description	Benefit	Limits	Paper
<b>Qualitative</b>	Based on expert verdict against intended functionality	Tailored to specific use case	Can be subjective, bias and hard to compare between studies	[19], [55]–[61], [69], [70], [73]
<b>Time-Before-Collision (TBC)</b>	Time between prediction and collision	Evaluates ability for early detection	Does not consider missed detections	[20], [22]
<b>Accuracy</b> $\frac{TP + TN}{(TP + TN + FP + FN)}$	Ratio of correct predictions over all predictions	Intuitive and shows general model ability	Misleading when data is not balanced	[20], [21], [35], [55], [56], [74], [79], [80]
<b>Precision</b> $\frac{TP}{TP + FP}$	Ratio of true positives over all positive detections	Can evaluate quality of positive detection	Does not consider missed detections	[18], [59], [79], [80]
<b>Mean Average Precision (mAP)</b>	Mean across average precision scores for each hazard type	Gives an overview across class types	Does not consider missed detections	[22], [81], [82]
<b>Recall</b> $\frac{TP}{(TP + FN)}$	Ratio of true positives over all positives (TP & FN)	Evaluates ability to detect a class	Does not consider confidence	[18], [59], [74], [79], [80]
<b>Area Under the Receiver Operating Curve (AUROC/AUC)</b>	Area under ROC curve that plots true positives over false positives	Evaluates ability to separate classes and considers	Skewed sample distribution can lead to misleading results	[20], [21], [71]
<b>False Negative Rate (FNR)</b> $\frac{FN}{FN + TP}$	Ratio of false negatives over the sum of false negatives and true positives	Evaluates how many hazardous events missed as could of led to harm	A system can predicts all samples as positive to achieves a perfect score	[74]

- **True Negative (TN):** Predicted safe and is safe.
- **False Negative (FN):** Predicted safe but is hazardous.

**Classification Metrics:** To evaluate classification, the key performance indicators (KPIs) are accuracy, precision and recall [83], [84]. Accuracy is the ratio of correct predictions over all predictions  $[(TP + TN)/(TP + TN + FP + FN)]$ , but this does not consider individual categories. Thus, a detector may not detect environmental or regulatory hazards but still score highly if those types rarely occur.

As accuracy does not consider performance over different classes, it is also important to review Precision  $[TP / (TP + FP)]$ . For example, in predicting actor-based hazardous events, the precision is the ratio of correct actor predictions over all predictions for that class; however, this does not consider missed detections. Thus, Recall  $[TP / (TP + FN)]$  is required to assess how many detections were missed in a scene, which is

the ratio of correct predictions over all the positive class detections in the scene. There is usually a trade-off between precision and recall, so these metrics can be plotted on a precision-recall curve with precision on the y- and recall on the x-axis. This can then be used to calculate Average Precision (AP) as the area under the curve and the mean Average Precision (mAP), which is the mean of all APs for each class.

As hazardous event detection is safety-critical, it is preferable to have a system that is overly cautious (i.e., False Positive) than a system that misses hazards (i.e., False Negative). As such, False Negative Rate (FNR) is an important measure to evaluate the ratio of missed detections that could have led to harm by calculating the ratio of false negatives over the sum of false negatives and true positives  $[FN / (FN + TP)]$ . As this is a binary classification, it can also be plotted with a Receiver Operating Characteristic (ROC) curve with the True Positive (i.e., probability of detection) rate in the y- and the False Positive rate (i.e., probability of false alarm) in the x-axis at various thresholds. Using this, the Area Under the Curve (AUC) can be calculated. The higher the area, the better the model can correctly separate hazards from non-hazards.

**Temporal Forecasting Metrics:** To evaluate how far in advance prediction occurs, Time-Before-Collision (TBC) is a metric used to calculate how many seconds before an accident the system can identify the hazardous event and is commonly used to evaluate the pre-emptive ability of detection.

**Utilisation:** Overall, the most popular metrics were accuracy, precision and recall as the hazardous event detection task is commonly framed as a binary classification problem (i.e., safe 0 or hazardous 1). When reporting performance, authors utilised multiple datasets to evaluate the model's capacity to learn, yet only a minority reported on dataset shift to assess generalisability to unseen data [20], [21], [79], [80].

In addition, self-collected simulation-based datasets were utilised as real crash scenes are rare. While simulation is an important tool, researchers should be aware of mismatches in domain complexity and extractable scene data as one study dropped from 91.1% accuracy and 96.2% AUC to 65.3% and 71.1%, respectively when trained on real-world footage [20].

Furthermore, even though the key to hazardous event detection is to detect before harm, few methods [20], [22] evaluated the time-before-collision metric. Though, methods show promising results of 4.9-30.9s TBC over short and medium duration clips of an average of 5-37.3s respectively. A TBC that allows ample reaction time for a human takeover, cited to require 2.8-23.8s in best case non-critical takeovers [85]. However, evaluation of TBC requires careful consideration to avoid scenarios of largely false positive detections that would erroneously lead to high TBC. A consideration mitigated by [22], which reported TBC with an average precision of at least 80%. Similarly, as the hazardous event detection task is critical for safety, false negative detections are more severe than false positives and represent events that could have been allowed to materialise harm. Yet, only one reviewed study evaluated the false negative rate [74].

## H. Datasets

Training and testing hazardous event detection require carefully annotated datasets of hazardous events such as vehicle

TABLE V. HAZARD FOCUSED DATASETS

Actor		Environment		Regulatory
Motor Vehicles	Non-Motor Vehicles	Static	Dynamic	Traffic Laws
<u>Collision</u>	<u>Collision/</u>	<u>Obstacle</u>	<u>Object</u>	<u>Rule</u>
<u>/Near</u>	<u>Near Miss</u>	BDD	A3D [89],	<u>Breach</u>
<u>Miss</u>	A3D <sup>123</sup>	[96],	BDD [96],	VIENA
A3D	[89],	HDD	TAB [91],	[93]
[89],	AAAD <sup>12</sup>	[95]	Lost and	
AAD	[88],		Found [98]	<u>Traffic</u>
[88],	Collision <sup>12</sup>	<u>Road</u>		<u>Lights</u>
CCD	[90],	<u>Type</u>	<u>Adverse</u>	Bosch
[22],	NIBD <sup>12</sup>	A3D	<u>Weather</u>	Traffic
Collision	[87],	[89],	A3D [89],	Lights
[90],	VIENA <sup>1</sup>	BDD	ACDC[99],	[105]
GTACras	[93]	[96],	CADCD[100]	
h [92],		Collision	CCD [22],	<u>Traffic</u>
NIBD		[90],	DAWN[101]	<u>Signs</u>
[87],		NIBD		LISA
TAB		[87],	<u>Illumination</u>	Traffic
[91],		Málaga	Collision	Signs
VIENA		[97]	[90],	[103]
[93]			NIBD	German
			[87],	Traffic
			NightOwls	Signs
			[102]	[104]

<sup>1</sup> Pedestrian | <sup>2</sup> Bicycle | <sup>3</sup> Animal

near-miss or collision events that are scarce. With the majority of large traffic scene datasets lacking these scenes [4], [86], [87], this slowed progress as researchers had to manually collect it, filter existing datasets or revert to simulation. With increased focus on robust safety testing and limited hazard datasets for training, there has been a promising increase in recent years. Arranged by hazard categories, the relevant datasets are compiled in TABLE V.

**Vehicle Actor:** Datasets contain near-miss scenes in which a hazardous event was narrowly avoided and scenes in which the harm was realised. Datasets use either real dashcam footage or are simulation-based to reduce collection time and cost. Of the real datasets these include: Anticipating Accidents Dataset [88], AnAn Accident Detection (A3D) [89], Car Crash Dataset (CCD) [22], Collision [90], Near-miss Incident Database (NIBD) [87] and Traffic Accident Benchmark (TAB) [91].

The largest real dataset is NIBD [87], with 4595 near-miss videos, each 10-15s. NIBD is captured over ten years with one hundred taxis using a monocular dashcam and annotated with either a high or low level of danger to represent the proximity of collision. This allows systems to be trained to understand immediate hazardous events that require immediate action (TTC < 0.5s, high danger) or potential hazardous events (TTC > 2s, low danger). With high scene diversity of intersections, highways, residential and parking areas and a roughly even balance between day and night illumination. This diversity makes this dataset multipurpose for other event categories such as road types and illumination contrast.

Of the synthetic datasets, GTACrash [92] and VIENA [93] consist of near-miss and collision scenarios for training. It is also common to use driving simulation platforms such as Carla [94] for bespoke training/testing. The advantage is creating hazardous events that may not readily occur, but the limitation is simulated sensory data and predetermined actor behaviour.

**Non-Vehicle Actor:** This category contains almost all the same datasets as the vehicle category as they extended to include pedestrian, bicycle and even animal actors in the case of A3D [89]. In [89], 1500 hazardous event videos are captured in differing weather conditions (e.g., rain, snow) and locations

in East Asia. Authors used expert judgement to annotate videos when developing hazardous events transitioned to materialised.

**Static Environment:** Containing datasets of a naturalistic nature as these hazardous events are derived from temporally static scene entities. It contains entities such as obstacles from parked vehicles, as seen in the Honda Research Institute Driving Dataset (HDD) [95], or objects that cause occlusion, as seen in Berkley DeepDrive (BDD) [96]. Another static attribute is road type, as each type can have varying traffic conditions, such as traffic density and location-specific hazards such as children in residential areas that suddenly appear behind parked cars. These hazards can be trained using datasets with large scene diversity, as seen in A3D [89], BDD [96], Collision [90], NIBD [87] and Málaga [97].

**Dynamic Environment:** These hazardous events are temporally dynamic in time, featuring dynamic objects (e.g., road debris) that need to be safely avoided but can be hard to identify due to their irregular shape or can appear suddenly. Seen within A3D [89], BDD [96], TAB [91] and in particular, Lost and Found [98], which focuses on small road objects.

In addition, this category also contains adverse weather scenes, which can be seen lightly within A3D [89] and CCD [22]. Recently, more focused datasets include: Adverse Conditions Dataset with Correspondences (ACDC) [99], Canadian Adverse Driving Conditions Dataset (CADCD) [100] and Detection in Adverse Weather Nature (DAWN) [101]. These allow more robust training, which is of particular importance due to heavy reliance on vehicle sensors which can be impaired with adverse weather (e.g., fog, rain, snow). Similarly, changes in illumination can adversely affect vision-based sensors and as such, varying day and night scenes can be tested within Collision [90], NIDB [87] and NightOwls [102], which focuses on pedestrian detections in dark, low visibility night scenes.

**Traffic Laws:** As traffic sign and light recognition is vital but highly varied between regions, this category contains datasets of scenes containing diverse traffic signs. Found in LISA Traffic Signs [103] and German Traffic Signs [104], which focus on danger, mandatory and prohibitory traffic signs. Traffic lights can also vary in shape and size, with 13 types collected in Bosch Small Traffic Lights [105] and sizes as small as 2 pixels to train far detection. Such datasets assist systems in avoiding the hazardous event of EV traffic violation and in detecting other actors in violation. For this, examples such as VIENA [93] collect scenes with other actors committing red light violations and driving on the wrong side of the road.

#### IV. DISCUSSION AND RESEARCH OPPORTUNITIES

This section critically evaluates the reviewed graph-based methods. Followed by the key research challenges and opportunities categorised into: domain complexity, data input, methodology and testing. Finally, we conclude with the limitations of the study.

##### A. Evaluation of Methods

Graph methods are shown as an explainable and scalable methodology for knowledge representation, allowing authors to model causal factors using graph connectivity. This relational representation is key to the early detection of hazardous events as they develop but remains challenging due to complex interactivity between road users and the environment (e.g., weather, road condition), which are not sufficiently understood.

TABLE VI: METHOD ADVANTAGES AND LIMITATIONS

Method	Advantages	Limitations
Knowledge Rule-Based (Ontology)	<ul style="list-style-type: none"> <li>• Simple</li> <li>• Few assumptions</li> <li>• Interpretable</li> <li>• Simple to model</li> <li>• Low complexity</li> <li>• Fast inference</li> </ul>	<ul style="list-style-type: none"> <li>• Models simple behaviour</li> <li>• Short predictive horizon</li> <li>• Cannot propagate uncertainties</li> <li>• Manual network definition</li> <li>• Cannot consider temporal evolutions</li> </ul>
Probabilistic (Bayesian)	<ul style="list-style-type: none"> <li>• Probabilistic prediction</li> <li>• Uncertainty propagation</li> <li>• Interpretable</li> <li>• Deductive and abductive reasoning</li> <li>• Counterfactual reasoning</li> </ul>	<ul style="list-style-type: none"> <li>• Difficult to scale</li> <li>• CPTs can be biased if based on expert opinion or difficult from data, as quantifying random variable effect on hazardous events can be challenging to isolate</li> <li>• Must explicitly define network</li> <li>• Cannot model cyclic relations</li> </ul>
Machine Learning (GNN)	<ul style="list-style-type: none"> <li>• Model complex behaviour</li> <li>• Ability to generalise</li> <li>• Ability for offline and potential online learning</li> <li>• Can integrate with other methods</li> </ul>	<ul style="list-style-type: none"> <li>• Lack of interpretability</li> <li>• Reduced training data</li> <li>• Dataset annotation can be biased or subjective</li> <li>• Models overfit to small training datasets</li> <li>• Poor prediction outside the scope of training can lead to dangerous behaviour</li> </ul>

In the past, authors have tried to incorporate causal dependencies by combining expert knowledge with empirical ML [18], [35], [71], [72], but developing a robust approach is yet to be realised. Methods focus on semantic scene description for better hazardous event generalisation and data reduction by up to an order of magnitude [106]. Each method reviewed had unique advantages and limitations, summarised in TABLE VI.

##### 1) Rule-Based Ontology Methods

Rule-based ontology methods are popular due to their high interpretability and simplicity. They are typically used by formulating known hazardous scenario patterns into rules for explainable real-time detection. Methods are simple to formulate and the graph-based logic makes decisions easy to interpret. However, the reasoning is limited to human knowledge and is unable to deal with uncertainty. Furthermore, user definition can make such methods difficult to scale and rules sensitive to bias, and modelled events tend to be simplified to allow formulation but leads to shorter predictive horizons.

In addition, methods primarily used deductive reasoning over observable information, which can be occluded or otherwise incomplete. To reason with incomplete data, some approaches demonstrated the use of abductive reasoning from effect to cause [18][35]. Although, demonstrated varying performance due to uncertainty over incomplete observations.

The inclusion of uncertainty or time history is another limitation of rule-based ontologies. Methods had no mechanism to incorporate the uncertainty of sensor information or store how behaviours evolved in time, which is particularly important to model interactions. Consequently, in an inherently uncertain and temporally dynamic driving domain, these key limitations reduce its application for complex hazards but represent a method that is both scalable, interpretable and easy to implement for hazards that are clear to formulate.

##### 2) Probabilistic Bayesian Network Methods

Similar to the rule-based methods, causal relationships can be defined using BNs through the manual definition but are again limited to current understanding. Unlike rule-based

methods, probabilistic BNs are difficult to scale as each hazardous event needs to be modelled individually in a graph structure with potential interdependencies.

The relational dependencies also make calculating the CPTs challenging as expert scoring can be biased, and data-driven approaches require well-annotated specialist datasets, which can be challenging to source. This complexity can also spiral if using continuous variables over discrete variables.

On the positives, BNs offer probabilistic predictions with the mathematical rigour of Baye's theorem and can be extended to consider temporal evolutions using DBNs. BNs also have the unique ability for both deductive and abductive reasoning due to the bidirectional flow of information. Being able to reason in this way conjoins the use of domain expertise (i.e., deduction) with logical reasoning (i.e., abduction) in unknown scenarios, which gives BNs a distinct advantage over other methods. The relations captured using graph connectivity and the CPTs that govern such relations also allow counterfactual reasoning. This gives BNs a unique ability to understand causal intervention, which is critical for hazard mitigation in later processes.

Another advantage is the ability to propagate uncertainty, as AVs rely heavily on a multitude of sensor data with inherent noise and computer vision. BNs can incorporate uncertainty by revising the probability distributions of all unknown variables, given new evidence in known areas of the network. This allows BNs to provide a probability distribution over prediction possibilities instead of a singular output.

### 3) Machine Learning Graph Neural Network Methods

Looking at GNN implementations, these methods extracted actors and objects from a list of detections into a relational graph network, used to learn features to detect hazardous events based on the spatial and temporal sequences. Automatic feature learning gives GNNs superior generalising ability, which is crucial in a domain where we cannot manually define all hazardous events. Although, this non-deterministic learning approach presents a verification hurdle as it is difficult to make guarantees about system behaviour in unseen scenarios. These approaches are also inherently less interpretable than previous methods and dependent on data for training but may synergise well in a hybrid methodology.

A further limitation of the ML approach is that learnt hazardous event patterns are based on detected patterns in the input data. Methods can mistake weak correlations for causal triggers and can be challenging to distinguish, as these weak correlations may overfit to training data but may be brittle in the real domain. To tackle this, some authors have implemented attention-based explainability to allow human validation.

GNNs also show great room for expansion, as the inference ability is directly related to the knowledge graph used to represent scene information; representations can be tailored to reason hidden relational dependencies to characterise known events and help generalise for unknown hazardous events.

For these reasons, GNNs show a promising research opportunity to learn features from relational data that can help us identify complex hazardous events not achievable otherwise. Given the learning ability of GNNs, this also presents an opportunity to synergise with other methods such as rule-based or probabilistic BNs. This synergistic approach is especially relevant to capturing edge cases that are easier to scale with ML methods to produce rough generalisations that are refined by

expert scoring and reasoned with the other methods into more robust frameworks for real-world usage.

## B. Challenges and Opportunities

Hazardous event detection faces many hurdles and opportunities, which we categorise under domain, data, methodology and testing to represent the domain complexity, data challenges, methodology limits and testing hurdles.

### 1) Domain Complexity

**Dynamic:** The driving domain is a dynamic network of traffic elements that interact in time and space. The challenge is decomposing this complex network into its component parts and representing how each interacts. How to represent these components and interactions remains an open question, but we believe graph structures to be a key enabler to this question.

**Interrelated Interactions:** Hazardous event evolution can be affected by many factors in the driving scene (e.g., road users, environment and traffic regulations). In addition, events can affect each other and co-occur due to spatial and temporal interactions that can be hard to formalise but must be understood for robust detection [4], [107], [108]. Authors have tried to understand causal dependencies by combining expert knowledge with empirical ML methods [18], [35], [71], [72], but developing a robust approach is yet to be realised.

### 2) Data Input

**Heterogeneous:** AVs receive vast amounts of data from sensors, communication systems and data streams from the cloud, road infrastructure and other vehicles. Thus, the challenge is aggregating, filtering, and representing this data for detection. We believe graph structures can help aggregate irregular inputs and map those inputs and their relationships. To demonstrate this, we present influential papers in the literature to show the utility of this emerging form of data representation.

**Uncertainty:** The data AVs receive may contain vast amounts of noisy, incomplete or uncertain data from many different sources, ranging from sensor noise to uncertainty from the computer vision systems due to complexity and lack of interpretability [109]–[111]. Thus, an opportunity is utilising the probabilistic methods covered in this survey to quantify and cascade uncertainty and modelling limitations through CPTs.

**Sensor Modality:** A large number of the methods only utilised observable features from an RGB camera that lacks depth encoding, which adversely affects localisation ability. The lack of accurate depth estimation makes it difficult to accurately extract the spatial features necessary to describe vehicle movements. Alternatively, simulation is used to extract distance from ground truth but has its limitations to practical transferability to real behaviour. Thus, accurately extracting spatial features from raw image data remains a key opportunity.

### 3) Methodology

**Scene Encoding:** The knowledge graph representing the driving scene must be comprehensive enough to model the system dynamics without becoming too complex for real-time inference. Hence, there will be an opportunity to test different approaches and investigate which scene features are most meaningful for prediction and which relations to represent.

**Scalable Framework:** As all hazardous events cannot be explicitly defined, a vital research question is how we can develop a scalable approach to adapt to an ever-changing domain. Unknown hazardous events pose the greatest danger as they can go undetected and unmitigated.

**Generalisation:** The endless types of hazardous events make it vital for systems to independently reason unseen dangers. The opportunity is to create a generalised hazardous event reasoning system using a combination of expert knowledge and data-driven modelling for generalised inference. For example, we do not know all hazardous events, but we have vast databases of non-hazardous driving scenes that may be investigated to allow anomaly detection and signal novel hazardous events.

**Multi-Type Detection:** Methods are generally scenario-specific, which does not sufficiently cover the driving domain. Typically, only actor and simple regulatory hazards are considered as they are clear to formulate using colliding trajectory or speed checks. In reality, hazardous events do not appear in isolation and affect one another, so it is essential to understand interrelated causality. For example, adverse weather conditions (e.g., fog, haze, rain) can affect visibility and vehicle handling and cause other hazardous events as a result.

**Early Detection:** The difficulty in hazardous event detection is early identification while there is time to react. Rule-based methods can only formalise limited complexity events, unlike ML approaches that show promising results generalising complex events but require sufficient training data and manual validation. Early detection requires causal relations to isolate the triggering events, but how to achieve this remains unsolved.

**Causal Understanding:** Correlation is not causation; sometimes, the causal drivers that affect behaviour cannot be observed and must be inferred. By focusing on causality, we can identify the variables with the most predictive power and intervene sooner. In addition, to pre-empt non-observable hazardous events such as occluded vehicles and to realise AV actions can mitigate an event developing further. Graphs provide a framework to build a relational network for inference, however, realising a robust framework remains an opportunity.

**Hybrid Method:** Given the advantages and limitations of each method, future work could investigate a hybrid approach to synergise the learning ability of GNNs with the mathematical rigour and uncertainty propagation of BNs to determine a probabilistic prediction. Subsequently, method synergy and cooperation remain an open but vital research question to combine the advantages whilst mitigating their limitations.

**Online learning:** In a dynamic and continually evolving domain, explicitly defining all hazardous events is intractable. Thus, one opportunity is to create a ML-based approach capable of online learning from live operations to learn continuously.

#### 4) Testing

**Dataset Availability and Completeness:** Vast amounts of naturalistic driving data are available, but scenes lack hazardous or collision scenarios due to scarcity of occurrences. This makes training difficult as researchers must synthesise specific edge cases and find an optimal number of tests to run to guarantee reasonable domain coverage. More hazardous event-focused datasets have emerged but lack vital depth information or diverse sensor suite information. Thus, an opportunity is to generate hazard-focused datasets with rich sensor suite

coverage that also sufficiently covers the range of possible scenarios and edge cases to certify safe behaviour.

**Dataset Annotation:** Annotating datasets accurately can be challenging as quantifying potential harm can be subjective if harm is not realised. In the literature, authors commonly used three experts [87], [89], [95], but standardisation, annotating hazardous events can be inconsistent, and this makes early detection training difficult. One approach was to associate hazardous events with a quantifiable metric such as TTC, as seen in NIBD [87]. High danger is represented by a  $TTC < 0.5s$  and low danger with danger with  $TTC > 2s$ ; however, implementations widely vary and are not valid for adverse weather or regulatory hazards. This raises the question of how to define a unified standard for annotating hazardous events.

**KPI Standardisation:** As detection is typically framed as a classification problem (e.g., hazardous event/ not), many authors use accuracy along with precision and recall to mitigate the effects of data imbalance in testing. Detection also has a temporal element as we want to detect when harm will occur, as seen with the TBC metric adopted in [22], but is not evaluated in other works. With each author using different metrics, an opportunity would be to propose a unified standard to allow meaningful comparison and establish a minimum threshold to reasonably guarantee safety.

**Safety Assurance:** Some methods lack a mathematical basis for safety certification as rule-based methods are commonly based on expert knowledge, and ML methods lack interpretability. Probabilistic Bayesian methods use conditional probability tables but how to generate trustworthy CPTs remains unclear. Therefore, an opportunity to synergise data-driven methods to produce CPTs through empirical data arises and could be explored in the future.

**Real-Time Operation:** The larger the input, the more memory and computational complexity are required for processing. To reduce inference costs in terms of time and memory, graph structures allow input data to be compressed as users define an abstraction of the scene. Differing from CNN-based methods that input raw 2D images, with one to three colour channels. Oppositely, the compact representations used by GNNs have exhibited up to 9.3 times faster inference, whilst up to 29% higher accuracy over non-graph counterparts [20].

Further improvements can also be explored, such as how to prioritise or filter nearby scene entities to reduce computational complexity. However, what to prioritise or filter requires investigation, as improper filtering may cause the model to miss patterns that are unknown by human understanding.

Yet, given the importance of the real-time operation, many real-time theoretical frameworks suggested did not test with real perception sensors or as part of live systems [67], [69]. This is sufficient for proof of concept but lacks rigour for real implementation. Consequently, the next step is to assess system real-time performance to validate if systems can process sensor information and provide detections during live operations.

**Realistic Simulations:** Due to the hurdles of live system prototyping and the dangers of hazardous event testing, simulation is popular among authors for speed of development. Although authors commonly utilise unrealistic access to ground truth or perfect sensor measurements, not representative of the real domain. An opportunity is to develop realistic simulation



environments by incorporating incomplete, erroneous and uncertain sensor data, which may reveal insightful system limitations or unexpected system behaviour.

### C. Review Limitations and Mitigation Scheme

The emerging nature of graph-based methods for hazardous event detection narrowed the availability of research for review. Therefore, a searching strategy was defined to best capture key work and minimise bias. To capture key works, search terms were defined using ISO definitions [30], [31] with synonym variations and explored in a diverse set of bibliographic databases. In addition to backward and forward citation searching and citation graph clustering to construct clusters of related work through co-citation and bibliographic coupling [112]. Likewise, to avoid recency bias, the authors took care to include key works in the area ranging from 2012 to 2022.

It is important to note that though each study relates to hazardous event detection, a few works applied to non-vehicle applications. However, these methods exemplified state-of-the-art developments with directly transferrable methodologies. Specifically, the recent ontology-based hazardous event detection within construction safety monitoring [59], [61]. Demonstrating how perception inputs can be related in graph ontologies and queried using rules to identify hazardous events.

## V. CONCLUSION

In an interconnected driving domain, the evolution of hazardous events can involve many variables. The challenge of detection is to process vast amounts of heterogeneous, uncertain and incomplete scene information to detect such events early and reliably. With the majority of accidents due to human error, understanding human behaviour is vital to enable the transition to SAE Level 3+, where AVs will need to safely control all driving functions. However, behaviour is complex and affected by other actors, the environment, and traffic rules. To tackle this challenge, this review takes a novel focus on graph methods to decompose a complex driving scene, define causal relations and aggregate heterogeneous data. To which, we present an overview of approaches from handcrafted graph models to recent GNN-based methods to learn multi-variate interactions.

From the vast range of methods reviewed, the authors found that the majority of works identified specific hazardous events with success; however, how to combine and connect events for comprehensive detection remains unclear. In a domain where it is intractable to manually define all events, GNNs presented the most scalable approach. However, this non-deterministic method presents a validation hurdle as it is less interpretable and highly dependent on data for training. Therefore, future methods may investigate a hybrid approach to synergise the learning ability of GNNs with the mathematical rigour and uncertainty propagation of the probabilistic Bayesian methods.

Ultimately, we need to enable comprehensive and reliable hazardous event detection. As how safely vehicles handle unknown unsafe scenarios will define the transition from human-supervised automated systems to unsupervised autonomy, and we present graph-based networks as a promising direction to achieve this.

## REFERENCES

- [1] UK Government, "Road accidents and safety statistics - GOV.UK," 2020. <https://www.gov.uk/government/collections/road-accidents-and-safety-statistics#road-casualty-annual-statistics>
- [2] US Department of Transportation, "Critical Reasons for Crashes Investigated in the National Motor Vehicle Crash Causation Survey (2018)," Washington, DC, 2018.
- [3] SAE International, "Taxonomy and Definitions for Terms Related to Driving Automation Systems for On-Road Motor Vehicles," 2018.
- [4] J. Guo, U. Kurup, and M. Shah, "Is it Safe to Drive? An Overview of Factors, Metrics, and Datasets for Driveability Assessment in Autonomous Driving," *IEEE Trans. Intell. Transp. Syst.*, 2020
- [5] S. Riedmaier, T. Ponn, D. Ludwig, B. Schick, and F. Diermeyer, "Survey on Scenario-Based Safety Assessment of Automated Vehicles," *IEEE Access*, 2020
- [6] E. Yurtsever, J. Lambert, A. Carballo, and K. Takeda, "A Survey of Autonomous Driving: Common Practices and Emerging Technologies," *IEEE Access*, 2020
- [7] D. Xiao, W. G. Geiger, H. Y. Yatbaz, M. Dianati, and R. Woodman, "Detecting Hazardous Events: A Framework for Automated Vehicle Safety Systems," in *IEEE Conf. on Intell. Transp. Sys. (ITSC)*, Oct. 2022.
- [8] S. Lefèvre, D. Vasquez, and C. Laugier, "A survey on motion prediction and risk assessment for intelligent vehicles," *ROBOMECH J.*, Dec. 2014
- [9] M. Strickland, G. Fainekos, and H. Ben Amor, "Deep Predictive Models for Collision Risk Assessment in Autonomous Driving," in *International Conference on Robotics and Automation*, May 2018.
- [10] E. Yurtsever *et al.*, "Risky Action Recognition in Lane Change Video Clips using Deep Spatiotemporal Networks with Segmentation Mask Transfer," in *2019 IEEE Trans. Intell. Transp. Syst. Conf.*, Oct. 2019.
- [11] C. Li, S. H. Chan, and Y. T. Chen, "Who make drivers stop? Towards driver-centric risk assessment: Risk object identification via causal inference," in *IEEE Conf. Int. Robots and Systems*, 2020.
- [12] M. T. Ribeiro, S. Singh, and C. Guestrin, "Why should I trust you? Explaining the predictions of any classifier," in *Int. Conf. Proc. ACM on Knowledge Discovery and Data Mining*, Aug. 2016.
- [13] G. Plumb, M. T. Ribeiro, and A. Talwalkar, "Finding and Fixing Spurious Patterns with Explanations," *Transactions on Machine Learning Research*, Jun. 2022.
- [14] S. Sagawa, A. Raghunathan, P. Wei Koh, and P. Liang, "An Investigation of Why Overparameterization Exacerbates Spurious Correlations," in *Proceedings of Machine Learning Research*, 2020.
- [15] G. Puebla and J. S. Bowers, "Can deep convolutional neural networks support relational reasoning in the same-different task?," *Journal of Vision*, Sep. 2022
- [16] S. Stabinger, D. Peer, J. Piater, and A. Rodríguez-Sánchez, "Evaluating the progress of deep learning for visual relational concepts," *Journal of Vision*, Oct. 2021
- [17] J. Kim, M. Ricci, and T. Serre, "Not-So-CLEVR: Learning same-different relations strains feedforward neural networks," *Journal Interface Focus*, Aug. 2018
- [18] N. Inoue, Y. Kuriya, S. Kobayashi, and K. Inui, "Recognizing Potential Traffic Risks through Logic-based Deep Scene Understanding," *22nd ITS World Congr.*, 2015.
- [19] J.-N. Russo, T. Sproesser, F. Drouhin, and M. Basset, "Risk Level Assessment for Rear-End Collision with Bayesian Network," *Int. Fed. Autom. Control*, Jul. 2017
- [20] A. V. Malawade, S. Y. Yu, B. Hsu, D. Muthirayan, P. P. Khargonekar, and M. A. A. Faruque, "Spatio-Temporal Scene-Graph Embedding for Autonomous Vehicle Collision Prediction," *IEEE Internet Things J.*, 2022
- [21] S. Yu, A. V. Malawade, D. Muthirayan, P. P. Khargonekar, and M. A. A. Faruque, "Scene-Graph Augmented Data-Driven Risk Assessment of Autonomous Vehicle Decisions," *IEEE Trans. Intell. Transp. Syst.*, 2021
- [22] W. Bao, Q. Yu, and Y. Kong, "Uncertainty-based Traffic Accident Anticipation with Spatio-Temporal Relational Learning," in *Proc. of 28th ACM Int. Conf. on Multimedia*, Oct. 2020.
- [23] P. Feth, "Dynamic Behavior Risk Assessment for Autonomous Systems," Technische Universität Kaiserslautern, 2020.
- [24] Y. Wang, W. Xu, Y. Zhang, Y. Qin, W. Zhang, and X. Wu, "Machine

- Learning Methods for Driving Risk Prediction,” in *Proc. 3rd ACM SIGSPATIAL Workshop on Emerg. Manag.*, 2017.
- [25] A. Mukhtar, L. Xia, and T. B. Tang, “Vehicle Detection Techniques for Collision Avoidance Systems: A Review,” *IEEE Trans. Intell. Transp. Syst.*, 2015
- [26] S. Dogramadzi, M. E. Giannaccini, C. Harper, M. Sobhani, R. Woodman, and J. Choung, “Environmental Hazard Analysis - a Variant of Preliminary Hazard Analysis for Autonomous Mobile Robots,” *J. Intell. Robot. Syst. Theory Appl.*, 2013
- [27] J. Dahl, G. R. De Campos, C. Olsson, and J. Fredriksson, “Collision Avoidance: A Literature Review on Threat-Assessment Techniques,” *IEEE Trans. Intell. Veh.*, 2019
- [28] C. Xu, Z. Ding, C. Wang, and Z. Li, “Statistical analysis of the patterns and characteristics of connected and autonomous vehicle involved crashes,” *J. Safety Res.*, 2019
- [29] A. M. Boggs, R. Arvin, and A. J. Khattak, “Exploring the who, what, when, where, and why of automated vehicle disengagements,” *Accident Analysis. Prevention*, 2020
- [30] British Standards Institution, “BSI 1890 Connected and Automated Vehicles – Vocabulary.” London, 2020.
- [31] International Organization for Standardization, “ISO 21448:2019 Road vehicles — Safety of the intended functionality.” Geneva, Switzerland, 2019. [Online]. Available: <https://www.iso.org/standard/70939.html>
- [32] International Organization for Standardization, “ISO 26262-1:2018, Road vehicles — Functional safety.” Geneva, Switzerland, 2018. [Online]. Available: <https://www.iso.org/obp/ui/#iso:std:iso:26262-1:ed-2:v1:en>
- [33] International Organization for Standardization, “PD ISO GUIDE 51: 2014 Safety aspects — Guidelines for their inclusion in standards.” Geneva, Switzerland, 2014.
- [34] E. Schwalb, “Analysis of Hazards for Autonomous Driving,” *J. Auton. Veh. Syst.*, 2021
- [35] R. Takahashi, N. Inoue, Y. Kuriya, S. Kobayashi, and K. Inui, “Explaining Potential Risks in Traffic Scenes by Combining Logical Inference and Physical Simulation,” *J. Mach. Learn. Comput.*, 2016
- [36] UK Government, “Theory test: cars: Hazard perception test - GOV.UK,” *Hazard perception test*, 2021. <https://www.gov.uk/theory-test/hazard-perception-test> (accessed Jun. 07, 2021).
- [37] R. Koganti *et al.*, “Machine Learning Algorithm for Automotive Collision Avoidance,” in *SAE WCX Digital Summit*, Apr. 2021.
- [38] S. Mozaffari *et al.*, “Deep Learning-Based Vehicle Behavior Prediction for Autonomous Driving Applications: A Review,” *IEEE Trans. Intell. Transp. Syst.*, 2020
- [39] G. Bagschik, A. Reschka, T. Stolte, and M. Maurer, “Identification of potential hazardous events for an Unmanned Protective Vehicle,” *IEEE Intell. Veh. Symp. Proc.*, 2018
- [40] R. Krishnan and S. V. Bhada, “Systems thinking approach to hazard analysis for complex robots,” *IEEE Int. Symp. Syst. Eng.*, 2019
- [41] G. Bagschik, T. Menzel, and M. Maurer, “Ontology based Scene Creation for the Development of Automated Vehicles,” *IEEE Intell. Veh. Symp. Proc.*, 2018
- [42] German Federal Ministry for Economic Affairs and Energy, “PEGASUS Project Overview,” 2019. [Online]. Available: <https://www.pegasusprojekt.de/en/home>
- [43] M. Scholtes *et al.*, “6-Layer Model for a Structured Description and Categorization of Urban Traffic and Environment,” *IEEE Access*, 2021
- [44] A. Rasouli and J. K. Tsotsos, “Autonomous vehicles that interact with pedestrians: A survey of theory and practice,” *IEEE Trans. Intell. Transp. Syst.*, 2020
- [45] J. Van Brummelen, M. O’Brien, D. Gruyer, and H. Najjaran, “Autonomous vehicle perception: The technology of today and tomorrow,” *Transp. Res. Part C Emerg. Technol.*, 2018
- [46] B. Ranft and C. Stiller, “The Role of Machine Vision for Intelligent Vehicles,” *IEEE Trans. Intell. Veh.*, Mar. 2016
- [47] C. Michael, C. Madalina, M. Pierre, and R. Sebastian, *Graph Structures for Knowledge Representation and Reasoning*. Cham: Springer International Publishing, 2021.
- [48] R. Dechter, “Reasoning with Probabilistic and Deterministic Graphical Models: Exact Algorithms, Second Edition,” *Synth. Lect. Artif. Intell. Mach. Learn.*, Feb. 2019
- [49] C. Michel and M. Marie-Laure, *Graph-based Knowledge Representation*. London: Springer London, 2008.
- [50] M. Z. Al-Taie and S. Kadry, *Graph Theory*. Springer, 2017.
- [51] R. J. Trudeau, *Introduction to Graph Theory*. Dover Pub., 2003.
- [52] R. Arp, B. Smith, and A. D. Spear, *Building Ontologies with Basic Formal Ontology*. The MIT Press, 2015.
- [53] Peter Bernus, J. Błażewicz, G. Schmidt, and M. Shaw, *Handbook on Ontologies*. Springer Berlin Heidelberg, 2009.
- [54] I. Harrocks, P. Patel-Schneider, H. Boley, S. Tabet, B. Groszof, and M. Dean, “SWRL: A Semantic Web Rule Language Combining OWL and RuleML,” 2004. Accessed: Jun. 19, 2021. [Online]. Available: <https://www.w3.org/Submission/SWRL/>
- [55] M. A. Mohammad, I. Kaloskampis, Y. Hicks, and R. Setchi, “Ontology-based Framework for Risk Assessment in Road Scenes Using Videos,” *Procedia Comput. Sci.*, 2015
- [56] M. A. Mohammad, “Video-based situation assessment for road safety,” Cardiff University, 2016.
- [57] P. Wolf, T. Ropertz, P. Feldmann, and K. Berns, “Combining ontologies and behavior-based control for aware navigation in challenging off-road environments,” *Proc. 16th Int. Conf. Informatics Control. Autom. Robot.*, 2019
- [58] M. Buechel, G. Hinz, F. Ruehl, H. Schroth, C. Gyoeri, and A. Knoll, “Ontology-based traffic scene modeling, traffic regulations dependent situational awareness and decision-making for automated vehicles,” *IEEE Intell. Veh. Symp.*, 2017
- [59] W. Fang, L. Ma, P. E. D. Love, H. Luo, L. Ding, and A. Zhou, “Knowledge graph for identifying hazards on construction sites: Integrating computer vision with ontology,” *Autom. Constr.*, 2020
- [60] B. Zhong, H. Li, H. Luo, J. Zhou, W. Fang, and X. Xing, “Ontology-Based Semantic Modeling of Knowledge in Construction: Classification and Identification of Hazards Implied in Images,” *J. Constr. Eng. Manag.*, 2020
- [61] M. Zhang, M. Zhu, and X. Zhao, “Recognition of High-Risk Scenarios in Building Construction Based on Image Semantics,” *J. Comput. Civ. Eng.*, 2020
- [62] K. Yamamoto, N. Inoue, K. Inui, Y. Arase, and J. Tsujii, “Boosting the Efficiency of First-Order Abductive Reasoning Using Pre-estimated Relatedness between Predicates,” *Int. J. Mach. Learn. Comput.*, 2015
- [63] R. Torres, M. Molina, and P. Campoy, “Survey of bayesian network applications to intelligent autonomous vehicles (IAVs),” *arXiv*. 2019.
- [64] N. Fenton and M. Neil, *Risk assessment and decision analysis with bayesian networks*. 2012.
- [65] Y. Zheng, Y. Xie, and X. Long, “A comprehensive review of Bayesian statistics in natural hazards engineering,” *Nat. Hazards*, Aug. 2021
- [66] P. Weber and C. Simon, *Benefits of Bayesian Network Models*. Hoboken, NJ, USA: John Wiley & Sons, Inc., 2016.
- [67] J. Reich and M. Trapp, “SINADRA: Towards a Framework for Assurable Situation-Aware Dynamic Risk Assessment of Autonomous Vehicles,” *16th Dependable Comput. Conf.*, 2020
- [68] K. P. Murphy, “Dynamic Bayesian Networks: Representation, Inference and Learning,” University of California, Berkeley, 2002. Accessed: Jun. 24, 2021. [Online]. Available: <https://www.cs.ubc.ca/~murphyk/Thesis/thesis.pdf>
- [69] S. Kabir *et al.*, “A Runtime Safety Analysis Concept for Open Adaptive Systems,” in *Model-Based Safety and Assessment*, Springer, 2019.
- [70] M. Cen, Y. Guo, and K. Lu, “Bayesian network based threat assessment method for vehicle,” *J. Comput.*, 2012
- [71] C. Liang, M. Ghazel, O. Cazier, and L. Bouillaut, “Advanced model-based risk reasoning on automatic railway level crossings,” *Saf. Sci.*, 2020
- [72] L. Wenhui, C. Fengtian, W. Chuna, and M. Xing kai, “Bayesian Network-Based Knowledge Graph Inference for Highway Transportation Safety Risks,” *Adv. Civ. Eng.*, 2021
- [73] S. Lefèvre, C. Laugier, and J. Ibañez-Guzmán, “Risk assessment at road intersections: Comparing intention and expectation,” *IEEE Intell. Veh. Symp. Proc.*, 2012
- [74] C. Katrakazas, M. Quddus, and W. H. Chen, “A new integrated collision risk assessment methodology for autonomous vehicles,” *Accid. Anal. Prev.*, 2019
- [75] V. Honavar and L. Uhr, “Symbolic Artificial Intelligence, Connectionist Networks & Beyond,” 1994. [Online]. Available: [http://lib.dr.iastate.edu/cs\\_techreports/76](http://lib.dr.iastate.edu/cs_techreports/76)

- [76] Z. Liu and J. Zhou, *Introduction to Graph Neural Networks*. 2020.
- [77] Z. Wu et al., "A Comprehensive Survey on Graph Neural Networks," *IEEE Trans. Neural Networks Learn. Syst.*, 2021
- [78] K. Xu, W. Hu, J. Leskovec, and S. Jegelka, "How Powerful are Graph Neural Networks?," in *7th International Conference on Learning Representations*, 2018.
- [79] S. Mylavarapu, M. Sandhu, P. Vijayan, K. M. Krishna, B. Ravindran, and A. Nambodiri, "Understanding dynamic scenes using graph convolution networks," *IEEE Int. Conf. Intell. Robot. Syst.*, 2020
- [80] S. Mylavarapu, M. Sandhu, P. Vijayan, K. M. Krishna, B. Ravindran, and A. Nambodiri, "Towards Accurate Vehicle Behaviour Classification With Multi-Relational Graph Convolutional Networks," in *IEEE Intelligent Vehicles Symposium (IV)*, Oct. 2020.
- [81] C. Li, Y. Meng, S. H. Chan, and Y. T. Chen, "Learning 3D-aware Egocentric Spatial-Temporal Interaction via Graph Convolutional Networks," in *Proceedings - IEEE International Conference on Robotics and Automation*, 2020.
- [82] S. Casas, C. Gulino, R. Liao, and R. Urtasun, "SpAGNN: Spatially-Aware Graph Neural Networks for Relational Behavior Forecasting from Sensor Data," in *Proceedings - IEEE International Conference on Robotics and Automation*, 2020.
- [83] J. Zhang, H. Yuan, and H. Dong, *Intro. to deep learning*. Springer, 2020.
- [84] I. Goodfellow, Y. Bengio, and A. Courville, *Deep Learning*. MIT Press, 2016.
- [85] A. Eriksson and N. A. Stanton, "Takeover Time in Highly Automated Vehicles: Noncritical Transitions to and From Manual Control," *J. Hum. Factors Ergon.*, Jun. 2017
- [86] A. M. Ibrahim, R. M. Hassan, A. E. Tawfiles, T. Ismail, and M. S. Darweesh, "Real-Time Collision Warning System Based on Computer Vision Using Mono Camera," *2nd Nov. Intell. Lead. Emerg. Sci. Conf. NILES 2020*
- [87] H. Kataoka, T. Suzuki, S. Oikawa, Y. Matsui, and Y. Satoh, "Drive Video Analysis for the Detection of Traffic Near-Miss Incidents," *Proc. - IEEE Int. Conf. Robot. Autom.*, 2018
- [88] F.-H. Chan, Y.-T. Chen, Y. Xiang, and M. Sun, "Anticipating Accidents in Dashcam Videos," in *Conf. on Computer Vision*, 2017.
- [89] Y. Yao, M. Xu, Y. Wang, D. J. Crandall, and E. M. Atkins, "Unsupervised Traffic Accident Detection in First-Person Videos," *IEEE Int. Conf. Intell. Robot. Syst.*, 2019
- [90] R. Herzig et al., "Spatio-Temporal Action Graph Networks," in *IEEE/CVF Int. Conf. on Comp. Vision*, Oct. 2019.
- [91] T. You and B. Han, *Traffic Accident Benchmark for Causality Recognition*. Springer International Publishing, 2020.
- [92] H. Kim, K. Lee, G. Hwang, and C. Suh, "Crash to Not Crash: Learn to Identify Dangerous Vehicles Using a Simulator," *Proc. AAAI Conf. Artif. Intell.*, Jul. 2019
- [93] M. S. Aliakbarian, F. S. Saleh, M. Salzmann, B. Fernando, L. Petersson, and L. Andersson, "VIENA: A Driving Anticipation Dataset," in *Computer Vision - ACCV 2018*, Springer, 2019.
- [94] A. Dosovitskiy, G. Ros, F. Codevilla, A. Lopez, and V. Koltun, "CARLA: An Open Urban Driving Simulator," in *1st Conference on Robot Learning (CoRL)*, 2017.
- [95] V. Ramanishka, Y. T. Chen, T. Misu, and K. Saenko, "Toward Driving Scene Understanding: A Dataset for Learning Driver Behavior and Causal Reasoning," *Proc. IEEE Comput. Soc. Conf. Comput. Vis. Pattern Recognit.*, 2018
- [96] F. Yu et al., "BDD100K: A Diverse Driving Dataset for Heterogeneous Multitask Learning," *Proc. IEEE Comput. Soc. Conf. Comput. Vis. Pattern Recognit.*, 2020
- [97] J. L. Blanco-Claraco, F. Á. Moreno-Dueñas, and J. González-Jiménez, "The Málaga urban dataset: High-rate stereo and LiDAR in a realistic urban scenario," *Int. J. Rob. Res.*, 2014
- [98] P. Pinggera, S. Ramos, S. Gehrig, U. Franke, C. Rother, and R. Mester, "Lost and found: Detecting small road hazards for self-driving vehicles," *IEEE Int. Conf. Intell. Robot. Syst.*, 2016
- [99] C. Sakaridis, D. Dai, and L. Van Gool, "ACDC: The Adverse Conditions Dataset with Correspondences for Semantic Driving Scene Understanding," in *2021 IEEE/CVF International Conference on Computer Vision*, Oct. 2021.
- [100] M. Pitropov et al., "Canadian Adverse Driving Conditions dataset," *Int. Journal of Robotics Research*, 2021
- [101] M. A. Kenk and M. Hassaballah, "DAWN: Vehicle Detection in Adverse Weather Nature Dataset," *arXiv*. arXiv, 2020.
- [102] L. Neumann et al., "NightOwls: A Pedestrians at Night Dataset," *Lect. Notes Comput. Sci. (including Subser. Lect. Notes Artif. Intell. Lect. Notes Bioinformatics)*, 2019
- [103] A. Møgelmoose, M. Trivedi, and T. Moeslund, "Vision-based traffic sign detection and analysis for intelligent driver assistance systems: Perspectives and survey," *IEEE Trans. Intell. Transp. Syst.*, 2012
- [104] S. Houben, J. Stallkamp, J. Salmen, M. Schlipsing, and C. Igel, "Detection of traffic signs in real-world images: The German traffic sign detection benchmark," *IEEE Conf. Neural Networks*, 2013
- [105] K. Behrendt, L. Novak, and R. Botros, "A deep learning approach to traffic lights: Detection, tracking, and classification," *Proc. - IEEE Int. Conf. Robot. Autom.*, 2017
- [106] S. Chen, Z. Jian, Y. Huang, Y. Chen, Z. Zhou, and N. Zheng, "Autonomous driving: cognitive construction and situation understanding," *Sci. China Inf. Sci.*, 2019
- [107] I. Elleuch, A. Makni, and R. Bouaziz, "Towards a cooperative intelligent system for unpredictable and predictable Road hazard detection," *J. Comput. Inf. Syst.*, 2019.
- [108] P. Ji, R. Li, Y. Xue, Q. Dong, L. Xiao, and R. Xue, "Perspective, Survey and Trends: Public Driving Datasets and Toolsets for Autonomous Driving Virtual Test," in *IEEE International Intelligent Transportation Systems Conference*, Sep. 2021.
- [109] M. Abdar et al., "A review of uncertainty quantification in deep learning: Techniques, applications and challenges," *Inf. Fusion*, 2021
- [110] N. Ståhl, G. Falkman, A. Karlsson, and G. Mathiason, "Evaluation of Uncertainty Quantification in Deep Learning," in *Communications in Computer and Information Science*, 2020.
- [111] P. Koopman, B. Osyk, and J. Weast, *Autonomous vehicles meet the physical world: RSS, variability, uncertainty, and proving safety*. Springer International Publishing, 2019.
- [112] A. T. Eitan, E. Smolyansky, I. K. Harpaz, and S. Perets, "Find and explore academic papers," *Connected Papers*, 2020. <https://www.connectedpapers.com> (accessed Jun. 03, 2020).

**Dannier Xiao** received his B.Eng. in Mechanical Engineering from the University of Southampton in 2018 and the M.Sc. in Advanced Mechanical Engineering from Imperial College London in 2020. He is currently pursuing the Ph.D. degree at the University of Warwick, within the Intelligent Vehicles group. He has worked in the industry for more than two years in engineering, manufacturing and finance at Siemens, Rolls-Royce Motor Cars and Goldacre, respectively. Among his research interests, are hazardous event detection, hazard modelling, graph neural networks and autonomous vehicles. In addition to research, he is a lab teacher for modules on machine intelligence and data science.

**Mehrdad Dianati** (Senior Member, IEEE) is currently a Professor of autonomous and connected vehicles with the Warwick Manufacturing Group (WMG), University of Warwick, and a Visiting Professor with the 5G Innovation Centre (5GIC), University of Surrey, where he was previously a Professor. He has been involved in a number of national and international projects as the project leader and the work-package leader in recent years. Prior to his academic endeavour, he has worked in the industry for more than nine years as a Senior Software/Hardware Developer and the Director of Research and Development. He frequently provides voluntary services to the research community in various editorial roles; for example, he has served as an Associate Editor for the *IEEE Transactions On Vehicular Technology*.

**Roger Woodman** is an Assistant Professor and Human Factors research lead at WMG. He received his PhD from Bristol Robotics Laboratory and has more than 20 years of experience working in industry and academia. Among his research interests, are self-driving vehicles, shared mobility, and human-machine interaction. He has several scientific papers published in the field of connected and autonomous vehicles. He lectures on the topic of Human Factors of Future Mobility and is the Deputy Director of the Centre for Doctoral Training at WMG, training doctoral researchers in the areas of intelligent and electrified mobility systems.

**William Gonçalves Geiger** received his B.E. degree in Mechanical Engineering from University of Mogi das Cruzes, Brazil, in 2015. He received the M.E. in Automotive Engineering from the University of Applied Sciences Ingolstadt, Germany, in 2018. He is currently pursuing the Ph.D. degree with the University of Warwick, U.K. He is also a System Engineer in the Assisted and Automated Driving department, Jaguar Land Rover. His research interests include runtime risk assessment, risk modelling, decision making, connected and autonomous vehicles.