Resolving Confusion of Unknowns in Autonomous Vehicles: Types and Perspectives

Kaushik Madala and Hyunsook Do
Department of Computer Science and Engineering, University of North Texas, Denton, U.S.A.

Keywords: Safety Of The Intended Functionality (SOTIF), ISO 21448, UL 4600, Machine Learning Safety, Autonomous Vehicle Safety, SAE Level 5 of Driving Automation.

Abstract: Autonomous vehicles are susceptible to unknowns. In particular, vehicles with SAE level 5 of driving automation, which need to operate in complex operational design domain (ODD) conditions, have a very high chance to face unknowns. While the industrial standards ISO 21448 and UL 4600 hint at analyzing unknowns from the analysts and engineers' perspective, the unknowns from different perspectives such as a autonomous vehicle or a machine learning model within an autonomous vehicle can differ from those perceived by engineers and analysts. In this paper, we discuss the different types of unknowns considering three different perspectives: analysts and engineers, autonomous vehicles, and machine learning (ML) models. We also clarify the often confused concepts of unknown knowns and unknowns unknowns for each perspective. Using a running example, we show how considering unknowns from different perspectives will aid in designing a safe autonomous vehicle.

1 INTRODUCTION

Autonomous vehicles have gained great attention in recent years. The eventual goal of autonomous vehicles is to reach driving automation of SAE level 5 (Committee et al., 2018), where the vehicles shall be able to operate autonomously and safely at any location without human feedback/intervention. Operating anywhere requires engineers of the autonomous vehicles to consider a complex operational design domain (ODD) (BSI/PAS, 2020). This requires engineers to explicitly consider the various environmental factors (e.g., snow, rain), road infrastructural elements (e.g., traffic signs, road arrow markings), road types (e.g., freeway, straight road), and other road users (e.g., skateboarders, pedestrians). Note that every element of ODD can have various attributes.

For example, a pedestrian can have attributes such as race, gender, type of clothing, and color of clothing. Considering all the possible values of these attributes and analyzing their effect on safety of a system can be challenging. Further, it is possible to overlook some of the elements or attributes and to have certain ODD elements that are not widely known as these elements are specific to a few locations. This makes autonomous vehicles susceptible to unknown scenarios and situations (Hejase et al., 2020) - both unknown knowns and unknown unknowns (Pickard et al., 2010). These unknowns can compromise the safety of the vehicles. Hence, autonomous vehicles must be designed taking unknowns into account. In this paper, we simply refer unknown scenarios and situations as *unknowns*.

Finding unknowns can be challenging because unknowns are highly perspective dependent. The industrial autonomy safety standards ISO 21448 (ISO/PAS, 2019) and UL 4600 (ANSI/UL, 2020) stress the need for analyzing unknowns and reducing the hazards that can be caused by the unknowns. Both the standards mostly refer to unknowns from the viewpoint of the engineers and analysts who design and analyze the safety of the system, respectively. However, unknowns can also be considered from the perspective of autonomous vehicles and machine learning models used in those autonomous vehicles.

For example, let us consider a traffic sign detection machine learning model in an autonomous vehicle. If the traffic sign detection model is identifying traffic signs correctly in one frame but not the subsequent frame, then the causes for such behavior can be unknowns to the engineers and analysts. This can be due to one of the following reasons: 1) the engineers and analysts might not be familiar with the factors affecting the traffic sign detection model's inference, and 2) the run-time occurrence of such behaviors by

Madala, K. and Do, H.
Resolving Confusion of Unknowns in Autonomous Vehicles: Types and Perspectives.
DOI: 10.5220/0010482106460653
In *Proceedings of the 7th International Conference on Vehicle Technology and Intelligent Transport Systems (VEHITS 2021)*, pages 646-653
ISBN: 978-989-758-513-5

the traffic sign detection model might not have been exposed to engineers and analysts. On the other hand, if the machine learning model is not trained to detect pedestrians, then pedestrians are unknowns to the traffic sign detection model. However, this does not imply pedestrians are unknowns to the autonomous vehicle as it can have another machine learning model which can detect pedestrians. In this example, we can observe that identifying unknowns of a machine learning model helps us to understand what additional machine learning models we will need to ensure that we can identify objects that are part of ODD and can compromise safety. Exposing unknowns to autonomous vehicles will aid in making better design decisions. Similarly, exposing unknowns to engineers and analysts helps to make better safety solutions, gather better data, and revise ODD. These observations lead to our central research question: **"Should we consider unknowns from different perspectives to ensure safety of an autonomous vehicle?"**. We address this central research question by answering the following research questions:

RQ1. What are unknowns from the perspective of machine learning models, autonomous vehicles, and engineers and analysts, respectively?

RQ2. What are similarities and dissimilarities among the unknowns from the three perspectives?

In this paper, we address these research questions by comparing the unknowns from each perspective with others using a running example. We discuss unknowns and their sub-categories (i.e., unknown knowns and unknown unknowns) from the perspectives of engineers and analysts, autonomous vehicles, and machine learning models in an autonomous vehicles. We discuss similarities and dissimilarities among the unknowns from the three perspectives and provide the details on how sub-categories of unknowns from one perspective can relate and differ to sub-categories of unknowns from other perspectives.

The rest of the paper is organized as follows. Section 2 discusses about the classification of knowns and unknowns. Section 3 details about the importance of unknowns with respect to ISO 21448 and UL 4600. Section 4 discusses the types of unknowns with respect to the three perspectives mentioned earlier with an example. Section 5 provides insights and observations, and we finally conclude in Section 6.

2 KNOWNS AND UNKNOWNS

We can classify knowns and unknowns based on the knowledge possessed by an intelligent agent/person (or a group of intelligent agents/persons) such as an engineer, analyst, autonomous system or organization. We can consider a set U which represents an entire universal knowledge. Considering each intelligent agent, we can divide U into two subsets: a set K denoting the knowledge possessed by the agent and a set N denoting the rest of universal knowledge not possessed by the agent. We can represent this mathematically in Equation 1, where $K \subset U$ and $N \subset U$.

$$U = K \cup N \tag{1}$$

We can further divide a set K into subsets. There are many classifications of K proposed by the existing literature (Smith, 2001; McCormick, 1997; De Jong and Ferguson-Hessler, 1996) (e.g., explicit and tacit knowledge (Smith, 2001); factual, procedural, conceptual and meta-cognitive knowledge (McCormick, 1997)). In this paper, to focus on knowns and unknowns, we classify K into following subsets: 1) a set D representing the direct knowledge, i.e., knowledge which an intelligent agent can comprehend and/or analyze easily after seeing an object, action or event, and 2) a set I representing the indirect knowledge, i.e., knowledge inferred using the knowledge from D. We can represent the relation between K, D, and I using Equation 2.

$$K = D \cup I \tag{2}$$

To understand the relation between sets D and I, let us consider '$\mathcal{P}$ (D)' which is the power set of D and a set V = {valid, invalid}. We can define a set I as shown in Equation 3, where f(x) is a function which provides the inference that can be generated based on input x, g(f(x)) is a function which tells if the inference output given from f(x) is a valid or invalid inference. Each element in I must have a mapping to only one of the elements in V, i.e., the intelligent agent shall be able to derive valid or invalid inferences from D. We mentioned element 'x' belongs to $\mathcal{P}(D)$ because I cannot exist without D, i.e., without the knowledge from a set D, we cannot infer the corresponding knowledge in I.

$$I = \{f(x) \mid x \in \mathcal{P}(D) \text{ and } g(f(x)) \in V\} \tag{3}$$

Based on the knowledge possessed and information recognized by an intelligent agent, we can classify the knowns and unknowns similar to ones classified by the existing literature (Pickard et al., 2010; Collins and Cruickshank, 2014; Jensen et al., 2017) as shown in Figure 1. We also illustrate the relation between these classifications and sets U, K, N, D and I in Figure 2. The figure shows four widely recognized classifications as follows.

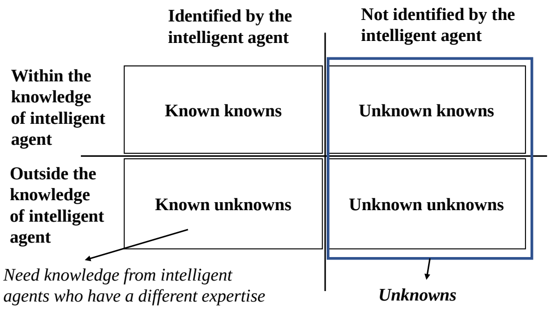

Figure 1: Classification of knowns and unknowns for an intelligent agent.

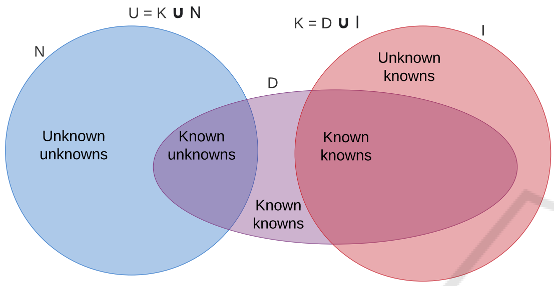

Figure 2: Venn diagram showing the different types of knowns and unknowns and their relation with sets U, K, N, D and I.

1. **Known Knowns:** These refer to the concepts and information that are present within the scope of the knowledge possessed by an intelligent agent. In the knowledge classification, we discussed prior, known knowns are a subset of the set K. An example of known known is the object correctly recognized by an individual agent.

2. **Known Unknowns:** These refer to the concepts and information that are identified by an intelligent agent but not within the scope of their knowledge. For an intelligent agent to identify a concept or information that is out of scope their knowledge, the agent should have a basic knowledge about the existence of the concepts, but does not need to have expertise enough to make observations or inferences from it. Known unknowns are sets which have few elements in the set D and no elements in the set I. Since only few elements in D help us in recognizing the unknown concepts, the knowledge that can be inferred from these elements is mostly part of the set N as we will need more knowledge to make a meaningful inference. An example of known unknown is a runtime monitor identifying that the machine learning model is uncertain of its input (Weiss and Tonella, 2021).

3. **Unknown Knowns:** These refer to concepts and information which are within the scope of knowledge possessed by an intelligent agent but are not identified. Unknown knowns mostly belong to the set I as they mostly represent the overlooked information by an intelligent agent either due to the insufficient analysis, a lack of awareness about presuppositions we have, a lack of proper inference, or due to the effect from mental factors such as stress, anxiety, and others. An example of unknown known is the misprediction of an object by a machine learning model which was recognized in the previous frame. In this case, the model knows the object and has knowledge to identify it, but was unable to identify it. An unknown known when exposed and analyzed will change into a known known.

4. **Unknown Unknowns:** These are concepts and information which are not in the scope of knowledge of an intelligent agent and not identified by the agent. For the misprediction example we used for unknown known, the cause that resulted in a misprediction of an object, which was correctly recognized in the previous frame, can be an unknown unknown to engineers until the frame is exposed and analyzed. All unknown unknowns belong to the set N. When exposed, unknown unknowns change mostly into known unknowns. However, with sufficient knowledge and expertise (potentially with the help from other experts), unknown unknowns when exposed can change to known knowns.

Related Work. To date many researchers (Pickard et al., 2010; Collins and Cruickshank, 2014; Jensen et al., 2017) have discussed the need for finding unknowns and proposed approaches to identify or expose unknowns. The application domains for such approaches mostly include biomedical applications (Collins and Cruickshank, 2014; Hoskisson and Seipke, 2020), software security (Al-Zewairi et al., 2020; Rashid et al., 2016), system design (Jensen et al., 2017), complex systems (Pickard et al., 2010), and autonomous vehicles (Wong et al., 2020; Hejase et al., 2020; Zhu et al., 2020). For example, Zhu et al. (Zhu et al., 2020) proposed a weakly-hard paradigm framework to model as well as mitigate time-based uncertainties for autonomous software. Even though all these current approaches discuss identification of unknowns or their mitigations, they do not consider different perspectives, which are the main concerns when we identify unknowns and understand reasons why such consideration is important.

In the next sections, we will discuss why we need to perform unknown analysis for autonomous vehi-

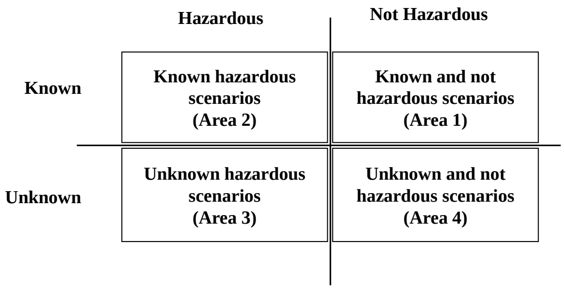

Figure 3: Scenario categories described in ISO 21448 standard.

cles and different perspectives we need to consider for performing such an analysis.

3 INDUSTRIAL STANDARDS AND UNKNOWNS

Industrial standards for safety of autonomy such as ISO 21448 (ISO/PAS, 2019) and UL 4600 (ANSI/UL, 2020) stress the need for analyzing unknowns to reduce the risks for autonomous vehicles. As autonomous vehicles with SAE level 5 of driving automation operate in a complex and changing ODD, it is possible to overlook some characteristics of ODD, which might be uncovered over time or when an unsafe situation is exposed.

ISO 21448 (ISO/PAS, 2019) is an industrial standard that details steps to analyze the safety of the intended functionality (SOTIF) of autonomous vehicles and thereby achieve compliance with respect to SOTIF. The standard focuses on identifying the gaps in nominal requirements and in reducing the safety issues for a system when exposed to unknown conditions. The standard illustrates a four quadrant structure for scenario categories as shown in Figure 3. We can observe that the categories are based on whether the scenarios are known or unknown and if they are hazardous or not. Each quadrant in Figure 3 represents a category and the corresponding area number that is assigned. For example, Area 1 implies scenarios which are known and non-hazardous. The goal of SOTIF is to increase Area 1 and reduce Areas 2 and 3. This illustrates the importance the ISO 21448 has given in need for analyzingunknown scenarios and proposing mechanisms to ensure autonomous vehicles operate safely in such scenarios.

Similarly, UL 4600 (ANSI/UL, 2020) which details a process of creating a safety case for autonomous vehicles also stresses the need for considering unknowns as a part of assuring safety of an autonomous vehicle. The standard mentions unknowns when discussing the safety case and arguments, autonomy function and support, dependability, lifecycle concerns, metrics and safety performance indicators, and assessment. The standard suggests to use feedback loops to keep track of unknowns. By doing so, we can accumulate knowledge of unknowns overtime and thus change design as needed. Hence, we can conclude analyzing unknowns for autonomous vehicles plays a vital role in assuring its safety for its operation in a complex ODD.

Note that it is not possible to identify all unknowns, and it might be very difficult to replicate some of the unknowns that are exposed in the real world. Hence, the standards focus more on proving that the occurrence of unknowns is rare and that we have mechanisms to accumulate the knowledge of unknowns over time as they occur. In addition to these standards, there is a functional safety standard ISO 26262 (ISO, 2018) for automotive, which deals only with malfunctions of electrical and electronic systems used in the vehicles but not unknowns.

4 UNKNOWN TYPES AND DIFFERENT PERSPECTIVES

Although both ISO 21448 and UL 4600 stress the need for analyzing unknowns, the focus of these standards is often interpreted as analyzing unknowns from the engineers and analysts' perspective. This interpretation comes from the assumption that unknowns perceived by engineers and analysts are not different from the ones by autonomous vehicles or machine learning models' perspective. However, this is not always true. Machine learning algorithms, which are intended for effective generalization based on data on which they are trained, could produce correct outputs for instances that might be unknowns to engineers and analysts. Hence, it is possible for unknowns to differ from different perspectives such as machine learning models used in autonomous vehicles, autonomous vehicles (i.e., vehicle-level unknowns), and engineers and analysts. This leads to our central research question stated in Section 1: **Should we consider unknowns from different perspectives to ensure safety of an autonomous vehicle?**

To answer our central research question, we specifically investigate the following two research questions:

RQ1. What are unknowns from the perspective of machine learning models, autonomous vehicles, and engineers and analysts, respectively?

RQ2. What are similarities and dissimilarities among the unknowns from these three perspectives?

After finding answers from these two research questions, we will conclude by answering our central question.

Running Example: To better illustrate unknowns from different perspectives, let us consider a running example of a level 5 autonomous vehicle which has cameras and a LIDAR and travels in city roads that have high density of pedestrians. Let us also assume that the pedestrians are multicultural and diverse in nature.

4.1 RQ1: Unknowns from the Three Perspectives

4.1.1 Unknowns with Respect to Machine Learning Models

Machine learning models play a vital role for perception and motion planning in autonomous vehicles. Tasks for which machine learning models are used include road object detection (Ashraf et al., 2016), pedestrian detection (Yang et al., 2020), traffic sign detection (Ayachi et al., 2020), unknown object detection (Wong et al., 2020) and trajectory estimation (Rozumnyi et al., 2020). The results of a machine learning model are highly dependent on its data, algorithm, and the training process followed. An unknown with respect to a machine learning algorithm is something which it cannot identify or detect. A machine learning model is susceptible to both unknown knowns and unknown unknowns.

Unknown Knowns for Machine Learning Models: Unknowns knowns for a machine learning model implies that the predictions regarding an object or an action are correct for the previous inputs or subsequent inputs, but are not correct for the current input. To better understand unknown knowns, let us consider the running example in which the autonomous vehicle detects pedestrians on city roads. If we assume the vehicle relies on cameras for pedestrian detection, then the input to the machine learning model that detects pedestrians will be a sequence of frames. If a pedestrian is identified in one frame but not identified in the subsequent frame, then we can refer to the latter case as an unknown known with respect to the machine learning model. This is because the machine learning model has knowledge to identify the pedestrian, but it might not always be able to correctly infer it.

Unknown Unknowns for Machine Learning Models: A machine learning model cannot detect or perform something that is never trained for. For example, if the training data of pedestrians for our running example is missing data on pedestrians belonging to a certain race and/or gender or pedestrians wearing a certain type of clothing, then the model might not be able to identify such pedestrians because it never knows such pedestrians would exist in the first place. Such instances for which machine models are never trained for are considered as unknown unknowns for machine learning models.

4.1.2 Unknowns with Respect to Autonomous Vehicles

When we consider unknowns with respect to autonomous vehicles, we refer to situations or events that an autonomous vehicle might not be familiar with or might have misinterpreted the situations or events as something else. Unknowns with respect to an autonomous vehicle are highly dependent on the architecture and the nature of sensor fusion. Similar to machine learning models, an autonomous vehicle also has a possibility of facing unknown knowns and unknown unknowns.

Unknown Knowns for Autonomous Vehicles: An autonomous vehicle takes a decision based on information it gathers from different sensors. If the autonomous vehicle's algorithm prioritizes one sensor over the other and the least prioritized sensor (not the highly prioritized one) provides correct information for a potential collision, then we can refer to such a situation as unknown knowns to autonomous vehicles. In our running example, if we consider the autonomous vehicle to have cameras and a LIDAR, and the cameras are prioritized over a LIDAR, then a pedestrian not detected by the machine models in the cameras but correctly detected by a LIDAR might be ignored by the vehicle. In this situation, the vehicle has the knowledge of the presence of a pedestrian, but the priority of sensors in the algorithm made it to ignore the pedestrian detected by the LIDAR. Since the motion planner takes action based on the output from camera models, the occurrence of the pedestrian will be an unknown known to the autonomous vehicle.

Unknown Unknowns for Autonomous Vehicles: An autonomous vehicle can face situations or ODD conditions which it never faced before. Such situations and ODD conditions, which autonomous vehicles are completely unfamiliar with and either cannot process the corresponding information or ignore them, are considered as unknown unknowns to autonomous vehicles. For example, let us assume the autonomous vehicle is never trained to operate it in snowy weather due to the low probability of snow in that region. If such a vehicle faces an unexpected snowfall in the region of its operation, then it is an unknown unknown condition with respect to the autonomous vehicle.

4.1.3 Unknowns with Respect to Engineers and Analysts

Since level 5 autonomous vehicles need to be able to travel everywhere without human intervention, the engineers and analysts of the autonomous vehicle will need to consider complex ODD conditions. However, it is not possible for the engineers and analysts to know all possible conditions that occur in the world and the ones which might affect the behavior of the autonomous vehicle. Hence, there might be some aspects that are overlooked or unknown to engineers and analysts. While considering engineers and analysts from diverse backgrounds and experiences across various regions can reduce the total number of unknowns, it is possible that we can have new potential unknowns. Also, as environments change, new unknowns can occur over time, which the engineers and analysts might not be familiar with unless such conditions or situations are exposed to them. This can result in a lack of consideration of such scenarios or situations while verifying the system as well as a lack of presence of such environments in simulation tools. Similar to machine learning models and autonomous vehicles, there can be unknown knowns and unknown unknowns with respect to engineers and analysts of an autonomous vehicle.

Unknown Knowns for Engineers and Analysts: It is possible for engineers and analysts to overlook some aspects when designing an intelligent autonomous vehicle. The reason behind overlooking the aspects can be intentional (e.g., to reduce complexity/scope) or unintentional. We refer to such aspects as unknown knowns. For our running example, if the engineers and analysts did not consider pedestrians' race when training the pedestrian detection machine learning model or analyzing the model, and the model did not predict people of a certain race properly, then it is an unknown known to the engineers and analysts. The engineers despite having knowledge of different races of people could have assumed that the machine learning model will be effective in identifying pedestrians of all races, thereby making the missing predictions of pedestrians belonging to a specific race as unknown knowns to engineers and analysts.

Unknown Unknowns for Engineers and Analysts: The unknown unknowns for engineers and analysts can occur mostly either due to a lack of knowledge about some ODD elements and conditions or due to a lack of appropriate understanding/analysis of machine learning models used in autonomous vehicles. For example, in our running example, if a pedestrian wore a reflective costume which resulted in a collision, such an occurrence can be an unknown unknown to the engineers and analysts. This is because the engineers and analysts have never seen such an occurrence before and hence did not take into account pedestrians with reflective costumes when designing the autonomous vehicle. Another example of unknown unknowns for engineers and analysts is the root causes for the non-deterministic behavior of a machine learning model, which identified an object in one frame but did not identify the same object in the subsequent frame for multiple inputs. This implies the machine learning model has knowledge to recognize the object, but the engineers and analysts do not know why it produces the correct output for one frame and the incorrect one for the other despite having knowledge about recognizing objects.

As mentioned earlier, once exposed unknown unknowns translate to known unknowns or known knowns depending on knowledge gained by the engineers and analysts.

4.2 RQ2: Comparing the Unknowns from Different Perspectives

So far, we have discussed unknowns with respect to machine learning models, autonomous vehicles, and engineers and analysts. However, are all these unknowns same or different? If they are different, how do they differ from each other? We shall now compare these unknowns from different perspectives.

An unknown for a machine learning model in an autonomous vehicle need not to be unknown to other machine learning models or algorithms that rely on a sensor different from the sensor being used by the current machine learning model. This implies unknowns of a machine learning model and unknowns of an autonomous vehicle are not necessarily the same. The unknowns of one machine learning model can be knowns to other machine learning models or can be exposed using sensors that are different from the input sensor to the machine learning model. The causes of unknown knowns for a machine learning model can be unknown unknowns to engineers and analysts. For example, if a machine learning model is able to detect an object in the previous frame but not the current frame, then we consider it to be an unknown known for the machine learning model. However, the engineers and analysts might not be familiar with such behavior of the machine learning model until it is exposed to them. Hence, it becomes unknown unknown for engineers and analysts.

Any ODD elements which are not considered by engineers and analysts because they do not know about their occurrence in the location considered are unknown unknowns, and remain unknown unknowns with respect to both an autonomous vehicle and machine learning model. This is because if the engineers and analysts do not know about the existence of an ODD element or condition in the first place, the corresponding design criterion might never have been taken into account. With respect to unknowns for an autonomous vehicle, however, they are not necessarily the same as unknowns for engineers and analysts. This is because the machine learning models used in autonomous vehicles are meant to support generalization to some extent. The process of generalization involves a machine learning model providing the expected output for an input which it has not been trained on or validated on. Hence, it is possible for machine learning models to work for some cases, even if they are not part of training data. Hence, unknowns for the autonomous vehicles need not to be the same as unknowns for engineers and analysts.

From the analysis of RQ1 and RQ2, we can conclude that,in the case of level 5 autonomous vehicles, *unknowns for engineers and analysts, unknowns for autonomous vehicles, and unknowns for a machine learning model used in the autonomous vehicle are not necessarily the same.*

5 DISCUSSION

Our comparison among the unknowns considering three different perspectives (autonomous vehicles, machine learning models within autonomous vehicles, and engineers and analysts) has shown that the unknowns perceived by these different perspectives can differ. This difference among these perspectives tells us the need to analyze unknowns considering all three, rather than following the typical practices, which tend to focus on unknowns identified by the engineers and analysts. If we use algorithms that are robust and not stochastic in the autonomous vehicle software, the unknowns identified from the autonomous vehicle's perspective will be the same as ones from the engineers and analysts' perspective. However, all autonomous vehicle heavily rely on machine learning models which are stochastic and not very robust in nature. Hence, for the central research question which aims to clarify if we need to consider unknowns from different perspectives to ensure safety of the autonomous vehicle, we can conclude that we will need to consider unknowns from all three different perspectives.

We believe that by exposing unknowns from the three perspectives, we can design the system better to handle unknown circumstances, as well as make the probability of facing unknown situations extremely rare. Further, by analyzing unknowns for machine learning models, we can understand if we need to add examples with certain elements into data and retrain a machine learning model or if we need to use other machine learning models and other sensor modalities to ensure safe operation of the system. Analyzing unknowns from the three perspectives also help the simulation engineers in enhancing existing simulation environments as well as building environments for evaluating new scenarios and situations.

6 CONCLUSION

In this research, we discussed how unknowns in an autonomous vehicle can be considered from different perspectives and their differences. In this paper, we focused on level 5 autonomous vehicles without considering vehicle-to-everything (V2X) connectivity (Hobert et al., 2015). We also used only a small running example as a proof of concept for the proposed idea. As a part of future work, we plan to analyze unknowns comprehensively by considering V2X communication. We also plan to conduct an extensive analysis to better understand how unknowns differ for the three different perspectives and how identifying them will play a major role in assuring safety of the system, even when operating in a complex ODD.

ACKNOWLEDGEMENTS

We would like to thank kVA by UL for their feedback on the topic.

REFERENCES

Al-Zewairi, M., Almajali, S., and Ayyash, M. (2020). Unknown security attack detection using shallow and deep ann classifiers. *Electronics*, 9(12):2006.

ANSI/UL (2020). 4600. *Standard for Evaluation of Autonomous Products*.

Ashraf, K., Wu, B., Iandola, F. N., Moskewicz, M. W., and Keutzer, K. (2016). Shallow networks for high-accuracy road object-detection. *arXiv preprint arXiv:1606.01561*.

Ayachi, R., Afif, M., Said, Y., and Atri, M. (2020). Traffic signs detection for real-world application of an advanced driving assisting system using deep learning. *Neural Processing Letters*, 51(1):837–851.

BSI/PAS (2020). 1883:2020. *Operational Design Domain (ODD) taxonomy for an automated driving system (ADS) – Specification*.

Collins, R. A. and Cruickshank, R. H. (2014). Known knowns, known unknowns, unknown unknowns and unknown knowns in dna barcoding: a comment on dowton et al. *Systematic biology*, 63(6):1005–1009.

Committee, S. O.-R. A. V. S. et al. (2018). Taxonomy and definitions for terms related to driving automation systems for on-road motor vehicles. *SAE Standard J*, 3016.

De Jong, T. and Ferguson-Hessler, M. G. (1996). Types and qualities of knowledge. *Educational psychologist*, 31(2):105–113.

Hejase, M., Barbier, M., Ozguner, U., Ibanez-Guzman, J., and Acarman, T. (2020). A validation methodology for the minimization of unknown unknowns in autonomous vehicle systems. In *2020 IEEE Intelligent Vehicles Symposium (IV)*, pages 114–119. IEEE.

Hobert, L., Festag, A., Llatser, I., Altomare, L., Visintainer, F., and Kovacs, A. (2015). Enhancements of v2x communication in support of cooperative autonomous driving. *IEEE communications magazine*, 53(12):64–70.

Hoskisson, P. A. and Seipke, R. F. (2020). Cryptic or silent? the known unknowns, unknown knowns, and unknown unknowns of secondary metabolism. *mBio*, 11(5).

ISO (2018). 26262. *Road Vehicles – Functional Safety*.

ISO/PAS (2019). 21448. *Road vehicles - Safety of the intended functionality*.

Jensen, M. B., Elverum, C. W., and Steinert, M. (2017). Eliciting unknown unknowns with prototypes: Introducing prototrials and prototrial-driven cultures. *Design Studies*, 49:1–31.

McCormick, R. (1997). Conceptual and procedural knowledge. *International journal of technology and design education*, 7(1-2):141–159.

Pickard, A. C., Nolan, A. J., and Beasley, R. (2010). 8.3. 1 certainty, risk and gambling in the development of complex systems. In *INCOSE International Symposium*, volume 20, pages 1051–1061. Wiley Online Library.

Rashid, A., Naqvi, S. A. A., Ramdhany, R., Edwards, M., Chitchyan, R., and Babar, M. A. (2016). Discovering "unknown known" security requirements. In *Proceedings of the 38th International Conference on Software Engineering*, ICSE '16, page 866–876, New York, NY, USA. Association for Computing Machinery.

Rozumnyi, D., Matas, J., Sroubek, F., Pollefeys, M., and Oswald, M. R. (2020). Fmodetect: Robust detection and trajectory estimation of fast moving objects. *arXiv preprint arXiv:2012.08216*.

Smith, E. A. (2001). The role of tacit and explicit knowledge in the workplace. *Journal of knowledge Management*.

Weiss, M. and Tonella, P. (2021). Fail-safe execution of deep learning based systems through uncertainty monitoring. In *2021 IEEE 14th International Conference on Software Testing, Validation and Verification (ICST). IEEE*.

Wong, K., Wang, S., Ren, M., Liang, M., and Urtasun, R. (2020). Identifying unknown instances for autonomous driving. In *Conference on Robot Learning*, pages 384–393. PMLR.

Yang, P., Zhang, G., Wang, L., Xu, L., Deng, Q., and Yang, M.-H. (2020). A part-aware multi-scale fully convolutional network for pedestrian detection. *IEEE Transactions on Intelligent Transportation Systems*.

Zhu, Q., Li, W., Kim, H., Xiang, Y., Wardega, K., Wang, Z., Wang, Y., Liang, H., Huang, C., Fan, J., et al. (2020). Know the unknowns: addressing disturbances and uncertainties in autonomous systems. In *Proceedings of the 39th International Conference on Computer-Aided Design*, pages 1–9.