Digital twin for synthetic data generation - application for automated driving systems

Hassan Hotait[1] and Alexandru Forrai[2]

(1) HAN - University of Applied Sciences, e-mail: H.Hotait@student.han.nl
(2) Siemens Industry Software Netherlands B.V., e-mail: alexandru.forrai@siemens.com

Abstract - *The paper presents how a digital twin (DT) can accelerate the development of automated driving systems. In particular, the verification and validation of sensing and perception system, with focus on safety critical traffic scenarios under parametric uncertainty and different types and variations of vulnerable road users (VRUs). The investigations show that the parametric uncertainty introduced by different VRUs is relatively small, but due to the safety relevance as well as sensitivity of deep neural networks, such a study remains relevant. The uncertainty introduced by the variations in weather and illumination conditions is much larger, which allows a proper robustness assessment of sensing and perception subsystem.*

Keywords: *digital twin, synthetic data generation, automated driving systems, autonomous vehicles, sensing and perception*

Introduction

Digital twin (DT) is a relatively new concept in design, having been around for less than 10 years. Hence, minimal research exists on how DT-driven virtual verification can improve product design and quality. Moreover, much of the technology required to realize DT is in the development, such as big data analytics and Internet of Things (IoT) communication (Tao, et al., 2020).

One of the possible application of DT for autonomous vehicles is to support the development process. In case of automate driving systems (SAE J3016 L3 and L4) the research and development community realized quickly that only real-world testing - using mileage-based coverage - is not feasible, from economical and technical point of view. One of the main reasons is that during real world-driving, safety relevant events, happen very rarely. It became obvious that virtual testing as well as digital twins will play a key role in the certification of automated driving systems (Leitner, Watzenig, and Ibanez-Guzman, 2020).

Therefore, the paper deals with the following research questions:

- How to build a digital twin for automated driving systems, what are the main requirements and main building blocks?
- How synthetic data generation is done using DT and how synthetic data accelerates the verification and validation of sensing and perception systems?

The concept of digital twin

The concept of digital twin (DT) has been introduced by NASA - the DT can be seen as a virtual mirror of the physical counterpart integrating multiple physics and scales, and it employs both dynamic sensor data and historical data from the product life cycle.

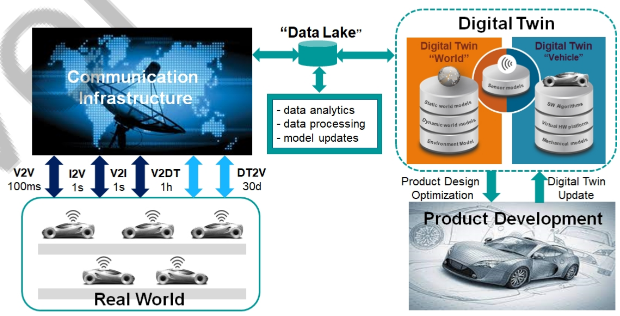

Figure 1: Digital twin for autonomous vehicles.

A digital twin (DT) is characterized by the interaction and convergence of the digital and physical worlds, which could possibly bring many benefits:

- The physical product can be made more intelligent to actively adjust its behaviour in real time according to the simulation by the virtual product (e.g. via software updates over the air).
- The virtual product can be made more realistic to accurately reflect the real states of the physical product.
- The solution enables integrating the real and simulated world at all required levels of integration to support efficient development and production.

A possible DT used for development and deployment of autonomous vehicles is presented in Fig. 1, where the timescales for data communication are also mentioned. For example, the information exchange between the DT and the real vehicle is happening at a times scales of hours (collected data is sent to the DT on hourly basis) and the updates over-the-air (OTA) might happen once a month.

While simulations and digital twins both use digital models to replicate products and processes, there are some key differences between the two. The most notable is that a digital twin creates a virtual environment able to study several simulations, backed up with real-time data and a two-way flow of information between the twin and the sensors that collect this data. This increases the accuracy of predictive analytical models, offering a greater understanding of how products, policies, and procedures, behave in real-world. A simulation replicates what could happen to a product, but a digital twin replicates what is happening to an actual specific product in the real world.

These differences can be further explained as follows:

- Static vs. active: A simulation model is static as it won't change or develop unless a designer introduces more elements. However, while a digital twin will begin much the same as a simulation model, the introduction of real-time data means that the twin can change and develop to provide a more active simulation.
- Possible vs. actual: A simulation replicates what could happen to a product, but a digital twin replicates what is happening to an actual specific product in the real world. Any changes to a simulation are limited to the imagination of a designer who needs to input any changes. However, because a digital twin offers real feedback, the designer can see if it is working as intended and then determine any improvements based on actual use.
- Scope of use: The final key difference is the scope of use that is offered by simulations vs digital twins. Simulations allows designers to test different scenarios against set parameters, making it useful for product design purposes.

A DT is not only highly relevant for development of autonomous vehicles, but they are becoming increasingly more important in the high-tech industry. Without DT the high-tech industry would not be able to accurately design and monitor complex systems, e.g. detecting system failures or degradation as early as possible.

DT for autonomous vehicles

Developing a DT for verification and validation of autonomous vehicles has several benefits and possible applications. A typical, data-driven workflow (continuous integration and continuous deployment) is presented in Fig. 2 (ASAM, 2023b), in which the DT could play a key role, since DT-driven virtual verification is capable of high-fidelity simulations, allowing designers to further improve and refine the design.

Due to the ability to model real-life scenarios, designers can obtain feedback on how a concept will behave in the real-world (e.g. in case of safety critical traffic scenarios, large variation of vulnerable road users).

Furthermore, the generated synthetic data set allows verification and validations at early stage of developments.

According to (Andrews, 2022) synthetic data is information that's artificially generated rather than produced by real-world events. Typically created using algorithms, synthetic data can be deployed to validate mathematical models and to train machine learning models.

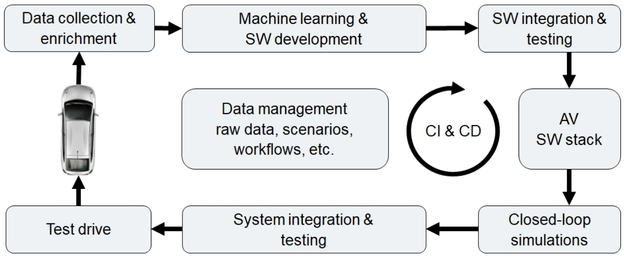

Figure 2: Data driven development of autonomous vehicles.

Data generated by a computer simulation can be seen as synthetic data. This encompasses most applications of physical modeling, such as music synthesizers or flight simulators. The output of such systems approximates the real thing, but is fully algorithmically generated (Nowruzi, et al., 2019).

Synthetic data is generated to meet specific needs or certain conditions that may not be found in the original, real data. Synthetic data are often generated to represent the authentic data and allows a baseline to be set. Another benefit of synthetic data is to protect the privacy and confidentiality of authentic data (Barse, Kvarnström, and Jonsson, 2003).

In this paper the presented DT is built using Simcenter Prescan, considering physics-based sensor models, Simcenter Amesim, considering a 15DOF vehicle dynamics model, SUMO traffic simulator.

Simcenter Prescan allows the definition and parametrization of the scenario in a flexible way. For vehicles, different models and colours can be selected, while for humans: gender, race, age can be chosen. Furthermore, environmental and illumination conditions specific for the operational design domain can be easily specified - see Fig 3. There are three essential components to provide a good physics-based simulation: sensor specific information, simulation engine and digital twin of the world/environment.

In this paper, the synthetic data generated by DT is stored in KITTI data format - see Fig. 4 (Geiger, et

Figure 3: Digital twin in Simcenter Prescan.

al., 2013) and is going to be extended to OpenLabel (ASAM, 2023a) data format. Hereby, as an application of the DT the robustness assessment of the sensing and perception stack, which relies on Yolov3 (Redmon and Farhadi, 2018) is presented.

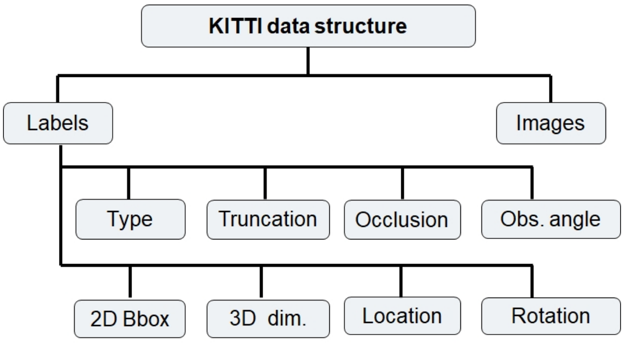

Figure 4: KITTI data structure.

Robustness assessment

Let us assume that in case of camera-based sensing and perception system the misdetection are associated with a failure. The number of misdetections depend of the operating conditions of the sensors, therefore the input data set $\forall x \in X_n$ (e.g. in case of a camera sensor the image data set) shall be representative capturing a wide variety of operating conditions (e.g. illumination conditions, weather conditions, etc.).

Therefore, it is relevant to investigate the robustness of the system, e.g. how the number of failures are changing, when the sensing system is exposed to an augmented data set $\forall x \in X$ - the nominal data set X_n is augmented by aleatoric and parametric uncertainty.

Robustness assessment of camera-based sensing and perception stack requires the definition of a reference data set, an augmented data set, a distance metric to assess the similarity between the reference data set and augmented data set as well as a performance metric such as the validation accuracy of the network.

A block diagram of the robustness assessment of the sensing and perception subsystem is presented in Fig. 5.

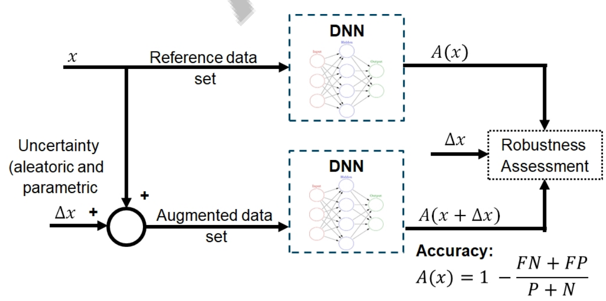

Figure 5: Robustness assessment.

The structural similarity index measure (SSIM) is a method for predicting the perceived quality of digital images and can be used for measuring the similarity between two images x, y - see Wang, et al., 2004.

$$SSIM(x,y) = \frac{(2\mu_x\mu_y + c_1)(2\sigma_{xy} + c_2)}{(\mu_x^2 + \mu_y^2 + c_1)(\sigma_x^2 + \sigma_y^2 + c_2)} \quad (1)$$

where μ_x and μ_y are the the pixel sample mean of x and y, σ_x^2 and σ_y^2 are the variance of x and y, σ_{xy} is the cross-correlation of x and y and c_1 and c_2 are two variables to stabilize the division with weak denominator.

The resultant SSIM index is a decimal value between -1 and 1, where 1 indicates perfect similarity, 0 indicates no similarity, and -1 indicates perfect anti-correlation.

Furthermore, a structural dissimilarity (DSSIM) may be derived from SSIM, though it does not constitute a distance function as the triangle inequality is not necessarily satisfied.

$$DSSIM(x,y) = (1 - SSIM(x,y))/2 \quad (2)$$

In addition to $DSSIM$, there are three widely used distance metrics for measuring uncertainty, all of which are p-norms, defined for matrices (since the focus is on camera images). Ideally, two inputs with smaller distance should be more similar with respect to human perception ability. Hereby, only the ∞-norm of a matrix is defined as the maximum absolute sum of the matrix rows.

$$||A||_\infty = \max_{1 \le i \le n} \sum_{j=1}^{n} |a_{ij}| \quad (3)$$

As a remark, the norm above is normalized in such a way that is independent of image size as well as number of layers.

Accuracy is one metric for evaluating classification models. Accuracy is defined as a ratio between the number of correct predictions and the total number of predictions. For binary classification, accuracy can also be calculated in terms of positives and negatives as follows:

$$A(x) = \frac{TP + TN}{TP + TN + FP + FN} \quad (4)$$

where TP, TN, FP, FN denote True Positives, True Negatives, False Positives and False Negatives.

According to the ethics guidelines for trustworthy AI (EU-study/report, 2019) the development, deployment and use of AI systems should meet the seven key requirements for trustworthy AI: (1) human agency and oversight, (2) technical robustness and safety, (3) privacy and data governance, (4) transparency, (5) diversity, non-discrimination and fairness, (6) environmental and societal well-being and (7) accountability.

Therefore, one of the investigations presented hereby, are related to technical robustness and safety considering diversity, non-discrimination and fairness. In this sense, the robustness of the sensing and perception stack is studied, when different vulnerable road users, which belong to different age, gender and ethnic categories are considered.

In Simcenter Prescan (Siemens-DISW, 2023), a large number of different vulnerable road users (VRUs) are already predefined. In the current study, the following VRUs, have been considered: VRUs = $\{$*Male Regular (as reference), Female Regular, Male African, Male CyclingCyclist, Child Regular, Female wBuggy, Male Old White WithStick, Female wShoppingCart*$\}$. The considered scenario is an urban driving scenario, when the VRU is crossing in front of the vehicle..

Furthermore, the three different diffculty classes - $Easy, Moderate, Hard$ - are defined based on object 2D bounding box size, object truncation and object occlusion. Smaller objects, which are truncated or partly occluded are more difficult to detect than larger objects, which are fully visible.

The performed investigations show that the normalized uncertainty - considering different VRUs - belongs to a narrow interval $\Delta x \in [0, 0.05]$ and the validation accuracy is $A(x) \in [0.6, 0.8]$. If we classify the detected objects in three different classes based on the difficulty of the detections such as: $Easy, Moderate, Hard$, we could observe - as shown in Fig. 6 - that child detection is more difficult in comparison with adult detection (mainly due to the size).

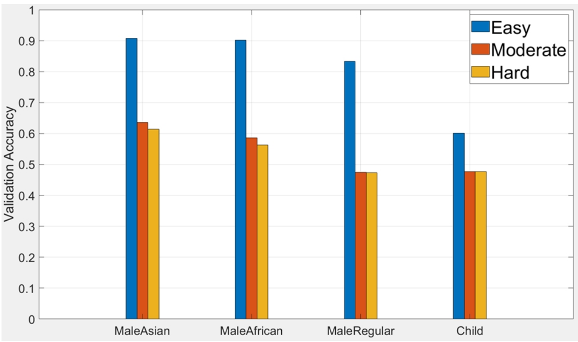

Figure 6: Validation accuracy for different VRUs.

Finally, the robustness of the network against aleatoric and parametric uncertainties has been investigated, where the uncertainty is quantified using the structural dissimilarity matrix. Hereby, only the results related to VRUs, in case of parametric uncertainty, introduced by the weather condition variations (fog) are shown - see Fig. 7 . It observed that as uncertainty is increasing the accuracy is decreasing, with a sharper drop, when the uncertainty exceeds a certain threshold $\Delta x \geq 0.275$.

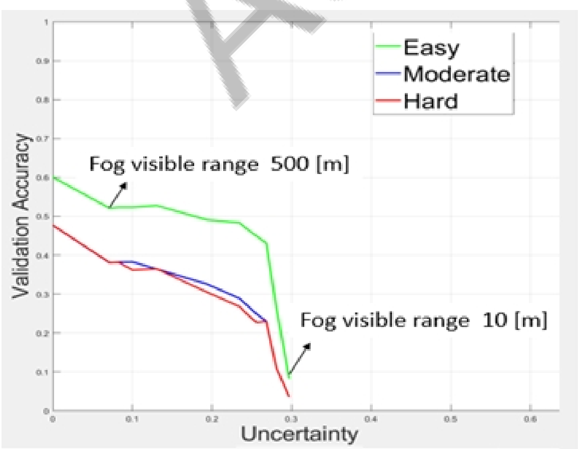

Figure 7: Network robustness against fog.

Conclusions

The DT and the generated synthetic data allow, efficient robustness assessment of the sensing and perception subsystem, considering aleatoric (e.g. simulated hardware faults) and parametric uncertainty (e.g. parameter variations of vulnerable road users, variations of environmental and illumination conditions).

A digital twin matures through a product lifecycle as more data is collected and analysed, offering different information that is not available with a static simulation. The scope of a digital twin reaches much further than simulations and includes all stages of a product's lifecycle. This increased scope means that digital twin can find uses outside of design and can help improve processes and make wider business decisions.

The investigations show that the parametric uncertainty introduced by different VRUs is relatively small, but due to the safety relevance as well as sensitivity of deep neural networks, such a study remains relevant. The uncertainty introduced by the variations in weather and illumination conditions is much larger, which allows a proper robustness assessment of sensing and perception subsystem.

References

Andrews, G., 2022. *What is synthetic data?* Accessed on 08 09, 2022. Available at: ¡https://blogs.nvidia.com/blog/2021/06/08/what-is-synthetic-data¿.

ASAM, 2023a. *ASAM OpenLABEL*. Accessed on 14 04, 2023. Available at: ¡https://www.asam.net/standards/detail/openlabel/¿.

— 2023b. *ASAM OpenTEST*. Accessed on 14 04, 2023. Available at: ¡https://www.asam.net/members/product-directory/detail/open-test-framework/¿.

Barse, E., Kvarnström, H., and Jonsson, E., 2003. Synthesizing test data for fraud detection systems. In: *IEEE Proceedings of the 19th Annual Computer Security Applications Conference*. doi:10.1109/CSAC.2003.1254343.

EU-study/report, 2019. Ethics guidelines for trustworthy AI. In: *https://digital-strategy.ec.europa.eu/en/library/ethics-guidelines-trustworthy-ai*.

Geiger, A., Lenz, P., Stiller, C., and Urtasun, R., 2013. Vision meets robotics: The KITTI dataset. In: *The International Journal of Robotics Research*. Vol. 11). 32, pp. 1231–1237.

Leitner, A., Watzenig, D., and Ibanez-Guzman, J. E., 2020. *Validation and Verification of Automated Systems, Results of the ENABLE-S3 Project*. Springer.

Nowruzi, F., Kapoor, P., Kolhatkar, D., Hassanat, F., Laganiere, R., and Rebut, J., 2019. How much real data do we actually need: Analyzing object detection performance using synthetic and real data. In: *arXiv preprint*. arXiv:1907.07061.

Redmon, J. and Farhadi, A., 2018. *YOLOv3: An Incremental Improvement*. arXiv: 1804.02767 [cs.CV].

Siemens-DISW, 2023. *Simcenter Prescan*. Accessed on 14 04, 2023. Available at: ¡https://plm.sw.siemens.com/en-US/simcenter/autonomous-vehicle-solutions/prescan/¿.

Tao, F., Liu, A., Hu, T., and Nee, A., 2020. *Digital Twin Driven Smart Design*. Academic Press, Elsevier.

Wang, Z., Bovik, A., Sheikh, H., and Simoncelli, E., 2004. Image quality assessment: from error visibility to structural similarity. In: *IEEE Transactions on Image Processing*. Vol. 13). 4, pp. 600–612.

Acknowledgements:

This work has received funding from the European Union's Horizon Europe research and innovation program under grant agreement No 101076754 - Althena project.