

## 2.3 Implementing requirements using machine learning (ML)

### Practical guidance – automotive

**Author:** Dr Daniele Martini, University of Oxford

#### A diverse dataset

One of the primary contributions of the SAX demonstrator project is the collection of a variegated dataset that includes different combinations of scenes -- urban, rural and off-road -- and hazards -- mixed driving surfaces, adverse weather conditions and presence of other actors. Our broad sensor suite comprises sensors traditionally exploited for AVs and uncommonly-used sensors that show great promise in challenging scenarios. Indeed, we collect cameras, LiDARS, GPS/INS, radar, audio and all the internal states of the vehicle. The different situations contained in the dataset will give us a deeper understanding of the behaviour of the sensors under such conditions, highlighting not only the strengths but especially the limitations. Thus, the aim is first to understand these limitations and overcome them by either fusing the information that the sensors provide in a mutually beneficial way or establishing optimal sensing schedules indicating the most reliable modality to use in each situation.



Figure 1. Vehicle Platform and Sensor Suite and the location in the UK for the collection sites.

Regardless of its dimension and breadth, no dataset can cover all possible operational scenarios an AV can encounter during its working life. Some domain shifts will always be present when deploying an AV in a new environment. Moreover, labelling is a very time-consuming and expensive process, and, as such, it is impractical to annotate every single sensor sample for every task at hand.

#### Generative approaches

We foresee generative approaches as a viable way to tackle the lack of enough data or labels in specific situations and scenarios. They provide a way to synthesise -- i.e. create -- ad hoc new training data that the AV never experienced during the data collection.

Nevertheless, these new hallucinated scenes can benefit the training process since they can represent situations the deployed AV will face during its operation.



Figure 2. A generative approach [2] is applied to synthesise an entirely new scene (right) from the segmentation mask (left) or an actual image (centre).

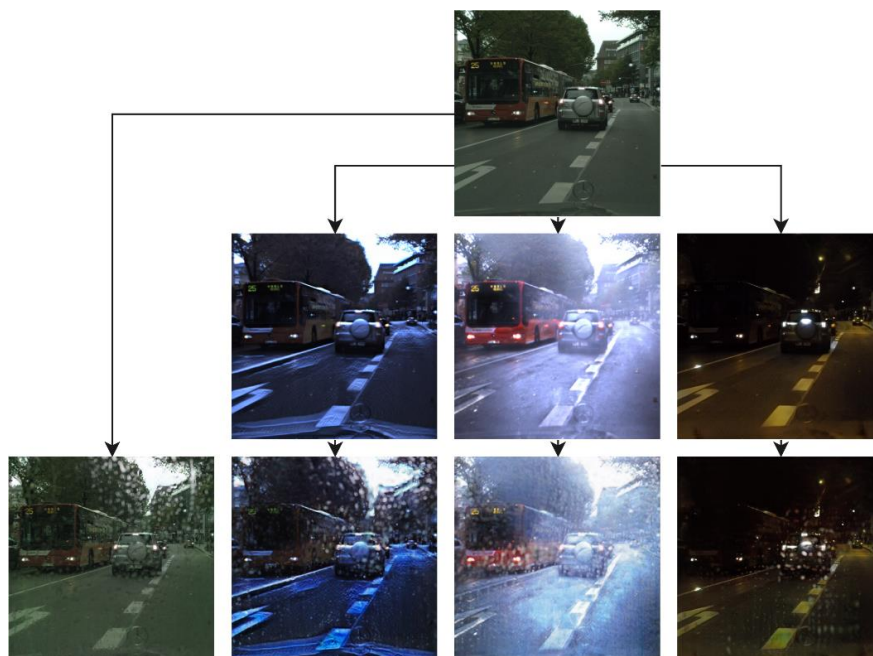


Figure 3. A generative approach [3] is applied to synthesising new weather conditions of the same scene.

### Weakly-supervised learning

Labelling is a costly and time-consuming task, especially for sensing modalities less natural for human operators, e.g. radar data. We can exploit the time and spatial constraints that our scenarios can offer. More specifically, the AVs move in a smooth world, where objects and other agents are consistent in time -- they move smoothly, as does the AV -- and space -- the AV will perceive the same things through different sensing modalities in the same portion of space.

Weakly-supervised learning approaches exploit these constraints to augment the provided labels by either projecting one sensor on top of another or retaining labels through time, significantly increasing the number of available annotations with many -- although imperfect due to the minor, unavoidable inconsistencies -- new samples.

Figures 4 and 5 show two examples of weakly-supervised approaches applied to the radar domain: in figure 4, camera and laser data are projected into the radar scan to create automatic annotations for urban segmentation [4]; figure 5, instead, shows audio labels drawn into a radar scan using the odometry to segment the road surface [5].

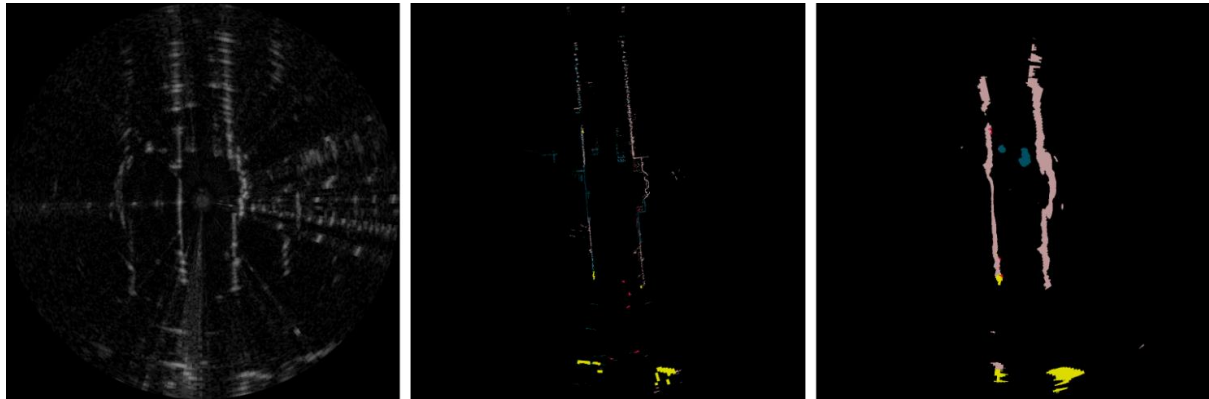


Figure 4. A radar scan (left) is labelled using camera and lidar data (centre) for training a semantic segmentation pipeline (right) [4].

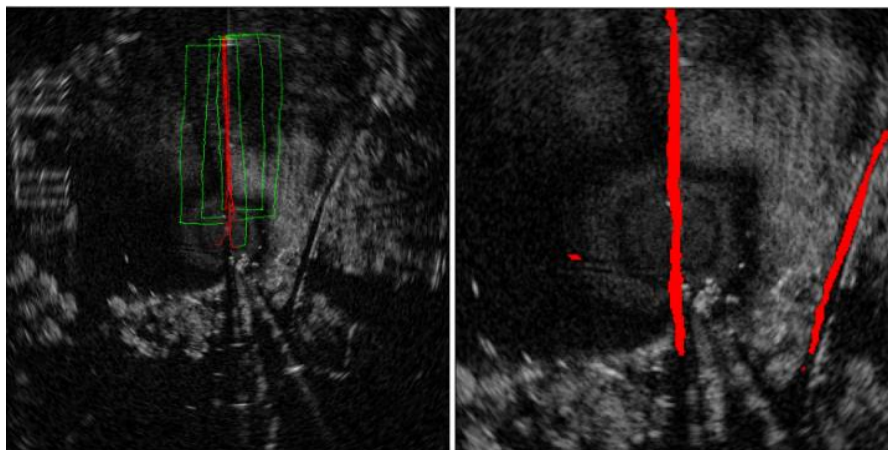


Figure 5. Audio data is used to label a radar scan (left) for training a semantic segmentation pipeline (right) to distinguish between road surfaces [5].

## Verification of the learned model

Substantially, we provide ground-truth labels for different tasks to train and validate state-of-the-art algorithms, to highlight their strengths and weaknesses in diverse operating scenarios. We include object detection/segmentation, drivable surface segmentation, odometry and decisions taken from a human driver. Such accurate ground-truth signals on different sensors, scenarios, and tasks will help assess sensors' performances and approaches in specific operations.

## Summary

In summary, we present an overview of AV training requirements and approaches to tackle the lack of specific sensing combinations or labels. Ideally, we would like our vehicles to be deployable and performant in any situation. The sensing capability of the AV plays a critical

role, and to evaluate the suitability of specific sensors in specific scenarios, we collected a dataset with broad combinations of environments and weather conditions. Nevertheless, no dataset can cover any possible situation an AV deployed in the wild can face and provide all the labels needed for any training pipeline. For this reason, we propose to expand the training procedure through generative approaches -- to synthesise new training scenes, never experienced by the AV -- and weakly or unlabelled data.

## References

- [1] Gadd, Matthew, et al. "Sense-Assess-eXplain (SAX): Building Trust in Autonomous Vehicles in Challenging Real-World Driving Scenarios." *2020 IEEE Intelligent Vehicles Symposium (IV)*. IEEE, 2020.
- [2] Sushko, Vadim, et al. "You Only Need Adversarial Supervision for Semantic Image Synthesis." *arXiv preprint arXiv:2012.04781* (2020).
- [3] Musat, Valentina, et al. "Multi-weather city: Adverse weather stacking for autonomous driving." *Proceedings of the IEEE/CVF International Conference on Computer Vision*. 2021.
- [4] Kaul, Prannay, et al. "Rss-net: Weakly-supervised multi-class semantic segmentation with FMCW radar." *2020 IEEE Intelligent Vehicles Symposium (IV)*. IEEE, 2020.
- [5] Williams, David, et al. "Keep off the grass: Permissible driving routes from radar with weak audio supervision." *2020 IEEE 23rd International Conference on Intelligent Transportation Systems (ITSC)*. IEEE, 2020.