

Motivation

Model-agnostic explanation methods for deep learning models are flexible regarding usability and availability. However, they can only manipulate input to see changes in output, which is why they suffer from weak performance when complex models like object detectors are investigated. In this work, we want to assess, which body parts are more or less relevant for the detection of a pedestrian regardless of the ML model architecture. We base our relevance assessment method on the *KernelSHAP* method and make adaptations to improve performance and usability for pedestrian detection.

Measures

We define the prediction quality of the object detection model as

$$q_p = \frac{2|P \cap T|}{|P| + |T|} \cdot c_p$$

Sørensen-Dice Coefficient
between predicted and true box
Classification score
of object classifier

Methods

Superpixel model

- Use segmented body parts as image regions (“superpixels”) and mask the pedestrian (inpainting or noise masking)
- Blend original image and mask with blending factor $\pi_i \in [0,1]$ individually for each body part i

Relevance assessment

- Use π_i as features and measure their attribution with *KernelSHAP*
- New sampling approach that samples $\vec{\pi}$ from Beta distribution which allows non-binary π_i 's (gradual blending)
- Results in relevance maps for detected pedestrians (i.e., contribution of each body part to the prediction quality)

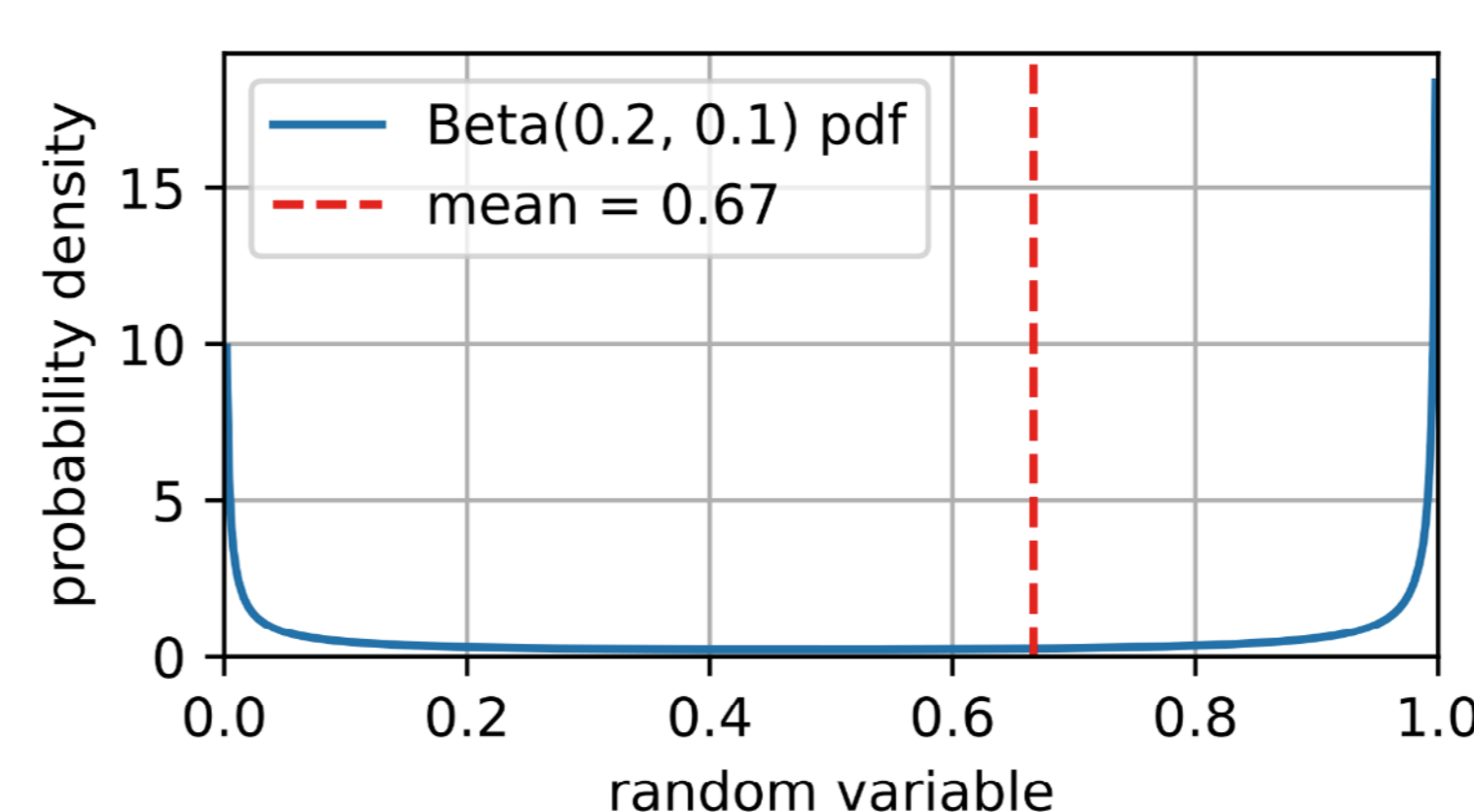


Figure 1: Beta distribution used by our sampling method. (© Fraunhofer IAIS)

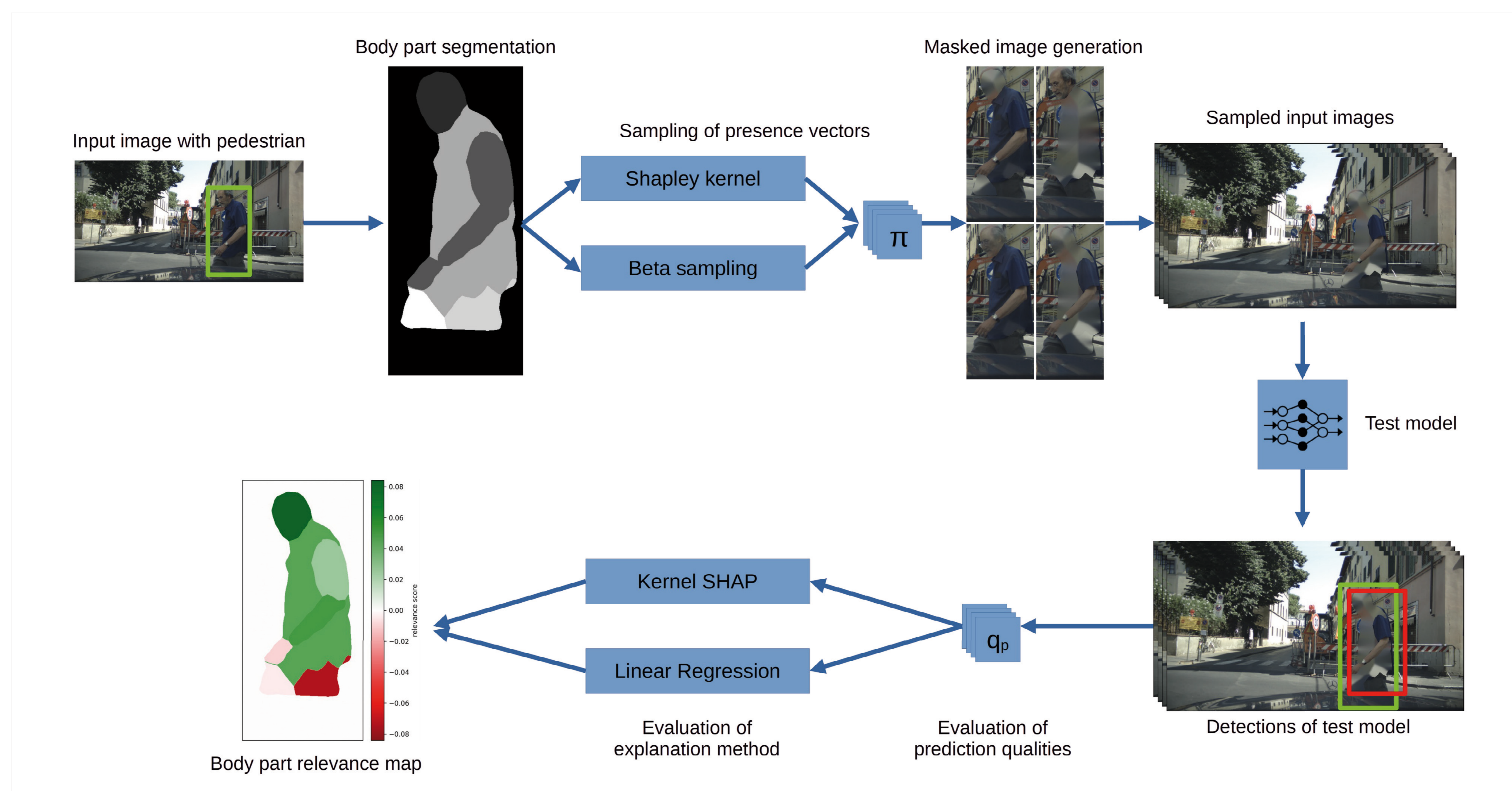


Figure 2: Sketch of our body part relevance assessment approach. (© Fraunhofer IAIS)

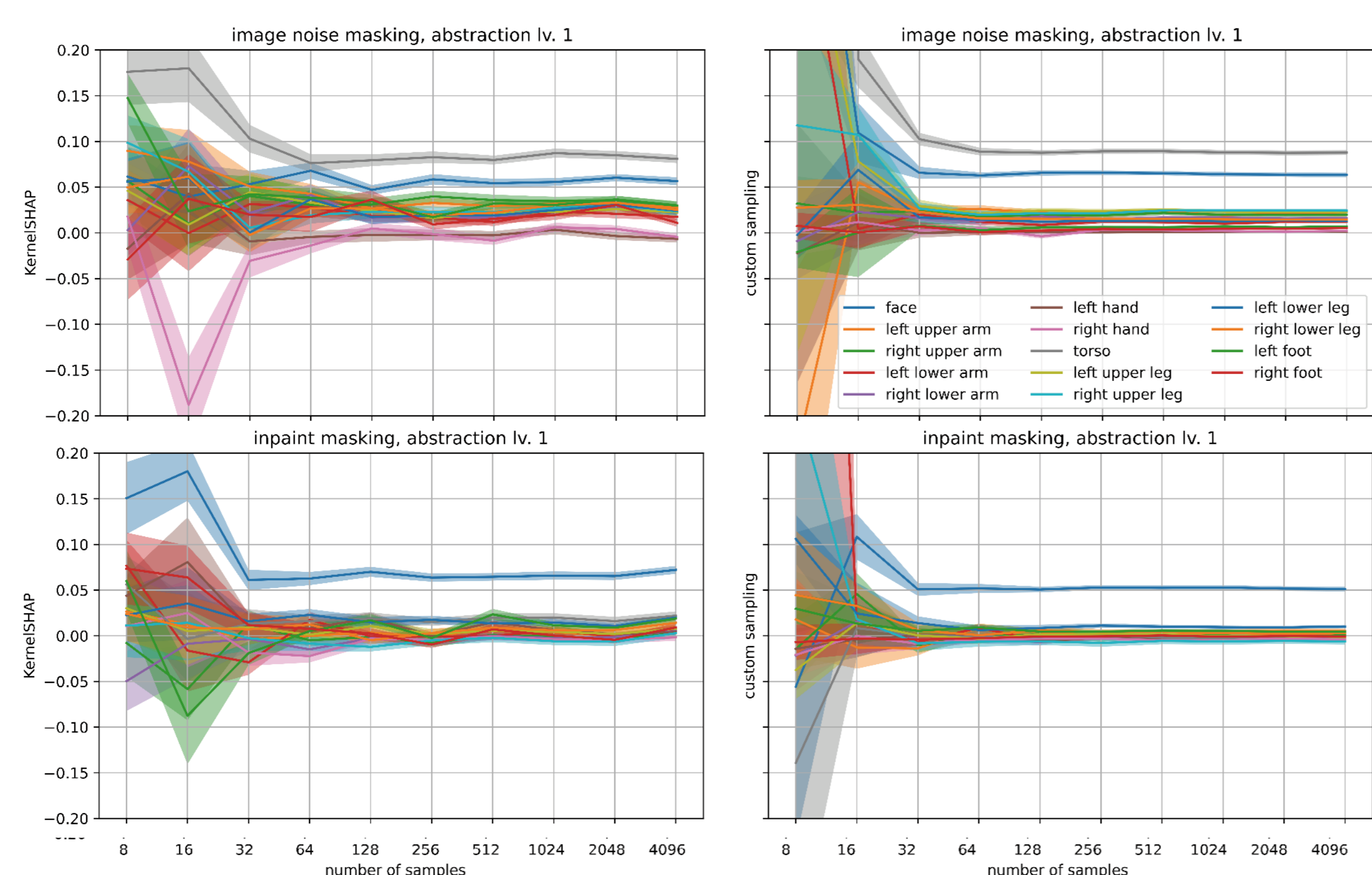


Figure 3: Comparison of *KernelSHAP* (left) and our *Beta sampling* (right) for different masking methods and abstraction levels. (© Fraunhofer IAIS)

Results

We track the relevance scores of a batch of pedestrians vs. number of samples and observe the convergence. Our *Beta sampling* method approximates the Shapley values given by *KernelSHAP* well and converges to stable results for fewer samples.

Read the full paper here!



Günder, M., Banerjee, S., Sifa, R., & Bauckhage, C. (2023). Model-agnostic Body Part Relevance Assessment for Pedestrian Detection. arXiv preprint arXiv:2311.15679.

Partners



External partners



For more information contact:

maurice.guender@iais.fraunhofer.de