1.13 Object Detection Plausibility with Concept-Bottleneck Models [1]

Mert Keser, Gesina Schwalbe, Azarm Nowzad | Continental AG
Alois Knoll | Technical University of Munich

Detection

Motivation

- Despite the unchallenged performance, deep neural network (DNN) based object detectors (OD) for computer vision have inherent, hard-to-verify limitations like brittleness, opacity, and unknown behavior on corner cases.
- Operation-time safety measures like monitors will be inevitable—even mandatory—for use in safety-critical applications like automated driving (AD).

Solution

- We proposed an approach for plausibilization of OD detections using a small model-agnostic, robust, interpretable, and domain-invariant image classification model.
- The safety requirements of interpretability and robustness are achieved by using a small concept bottleneck model (CBM), a DNN intercepted by interpretable intermediate outputs.
- Concept extraction with Color-Invariant Convolutional layer (CIConv) [2] approach increased the robustness of Concept Bottleneck Models (CBMs) [3] against domain shifts, thus enabling their application to diverse Automated Driving (AD) settings.

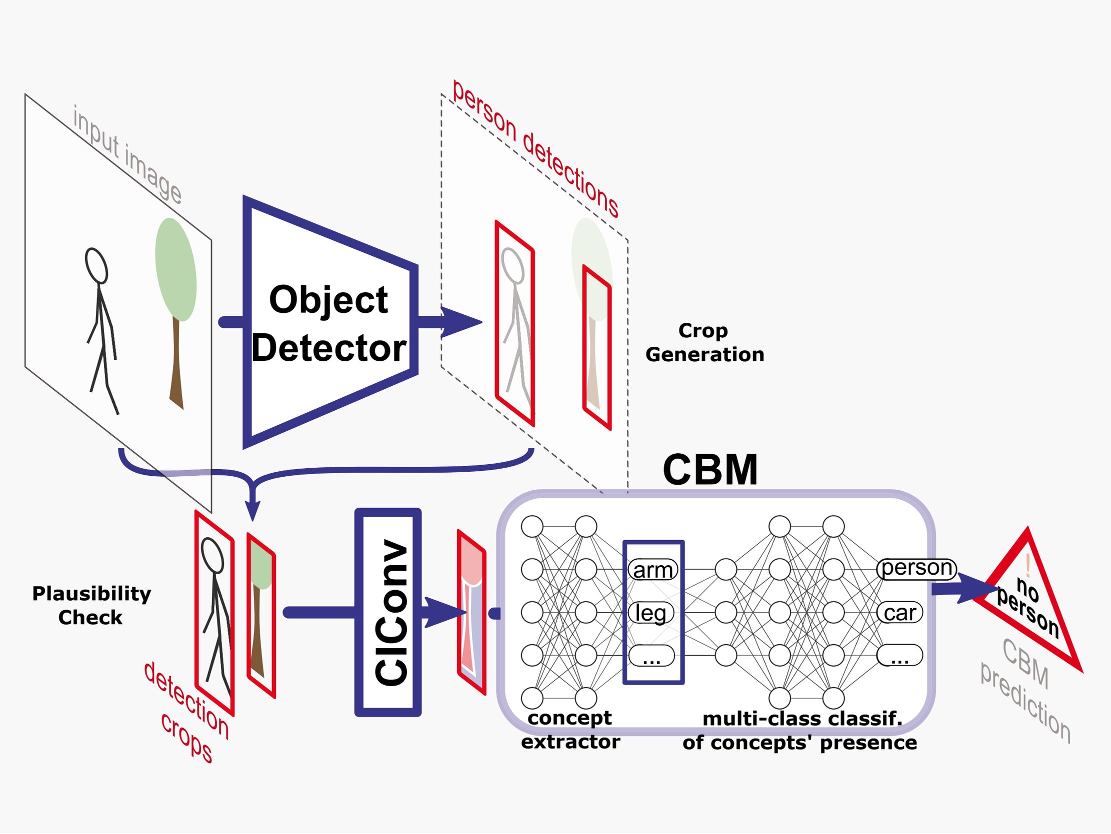

Figure 1: Proposed interpretable, model-agnostic monitoring approach for identification of false positive person detections. It uses an interpretable concept bottleneck model (CBM) as independent classifier on the color-invariant representations (CI-repr) of the object detections (here: for class person) (© Continental AG)

Experiments

- Learning Robust Concept Representations, Comparison of Object Class Prediction Accuracy on Broden Test Data [4]

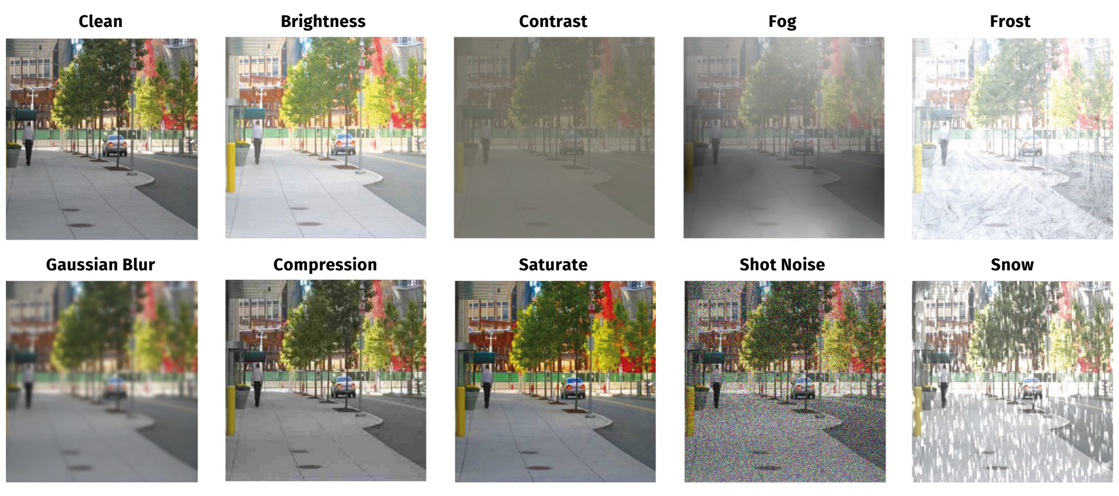

Figure 2: Different perturbations on the Broden Test Data [4] (© Continental AG)

	Person, Car			Person, Car	
Corruptions	**CBM**	**CIConv-CBM**	**Corruptions**	**CBM**	**CIConv-CBM**
Clean	**89.3%, 91.4%**	87.4%, 90.3%	**Compr.**	35.06%, 69.74%	**83.84%, 87.55%**
Brightness	33.88%, 64.92%	**85.69%, 88.77%**	**Saturate**	34.91%, 68.89%	**85.65%, 89.27%**
Contrast	33.03%, 52.17%	**85.60%, 87.46%**	**Shot Noise**	34.27%, 42.13%	**65.53%, 74.55%**
Fog	34.62%, 69.38%	**84.74%, 87.30%**	**Snow**	35.27%, 68.33%	**65.37%, 77.90%**

Table 1: Object class prediction accuracy comparison between vanilla CBM and CBM with CIConv layer on Broden test data with applied corruptions (severity=3). Bold numbers highlight the best prediction performance for each class and corruption type.

Figure 3: Different type of false positive detections (© Continental AG)

- Comparison of fine-tuning (FT) and zero-shot false positive monitoring for SqueezeDet (SDet) [5] on KITTI easy [6].

Data, Model	Task	IoU	Precision	Recall
KITTI, SDet	**Car**	**0.7**	**0.96**	0.07
KITTI, SDet	(FT) Car	**0.7.**	0.81	**0.56**
KITTI, SDet	Ped	**0.5**	**0.83**	0.01
KITTI, SDet	(FT) Ped	**0.5**	0.72	**0.95**

Table 2: Comparison of fine-tuning (FT) and zero-shot false positive monitoring for SqueezeDet (SDet) on KITTI easy. Bold numbers highlight best-performing method for each metric and task. (© Continental AG)

Conclusion

- The plausibilisation with monitoring approach was identified as a promising method for error detection in object detection during operation.

References:

[1] Keser et al. Interpretable Model-Agnostic Plausibility Verification for 2D Object Detectors Using Domain-Invariant Concept Bottleneck Models, CVPRW, 2023
[2] Lengyel, Attila, et al. "Zero-shot day-night domain adaptation with a physics prior." Proceedings of the IEEE/CVF International Conference on Computer Vision. 2021
[3] Koh, Pang Wei, et al. "Concept bottleneck models." International Conference on Machine Learning. PMLR, 2020
[4] Bau, David, et al. "Network dissection: Quantifying interpretability of deep visual representations." Proceedings of the IEEE conference on computer vision and pattern recognition. 2017.
[5] Wu, Bichen, et al. "Squeezedet: Unified, small, low power fully convolutional neural networks for real-time object detection for autonomous driving." Proceedings of the IEEE conference on computer vision and pattern recognition workshops. 2017
[6] Geiger, Andreas, et al. "Vision meets robotics: The kitti dataset." The International Journal of Robotics Research 32.11 (2013): 1231-1237.

Partners

External partners

For more information contact:
mert.keser@continental-corporation.com

KI Wissen is a project of the KI Familie. It was initiated and developed by the VDA Leitinitiative autonomous and connected driving and is funded by the Federal Ministry for Economic Affairs and Climate Action.

Funded by the European Union
NextGenerationEU
Supported by:
Federal Ministry for Economic Affairs and Climate Action
on the basis of a decision by the German Bundestag