

Introspection of 2D Object Detection using Processed Neural Activation Patterns in Automated Driving Systems

Hakan Yekta Yatbaz, Mehrdad Dianati, Konstantinos Koufos, Roger Woodman

Overview

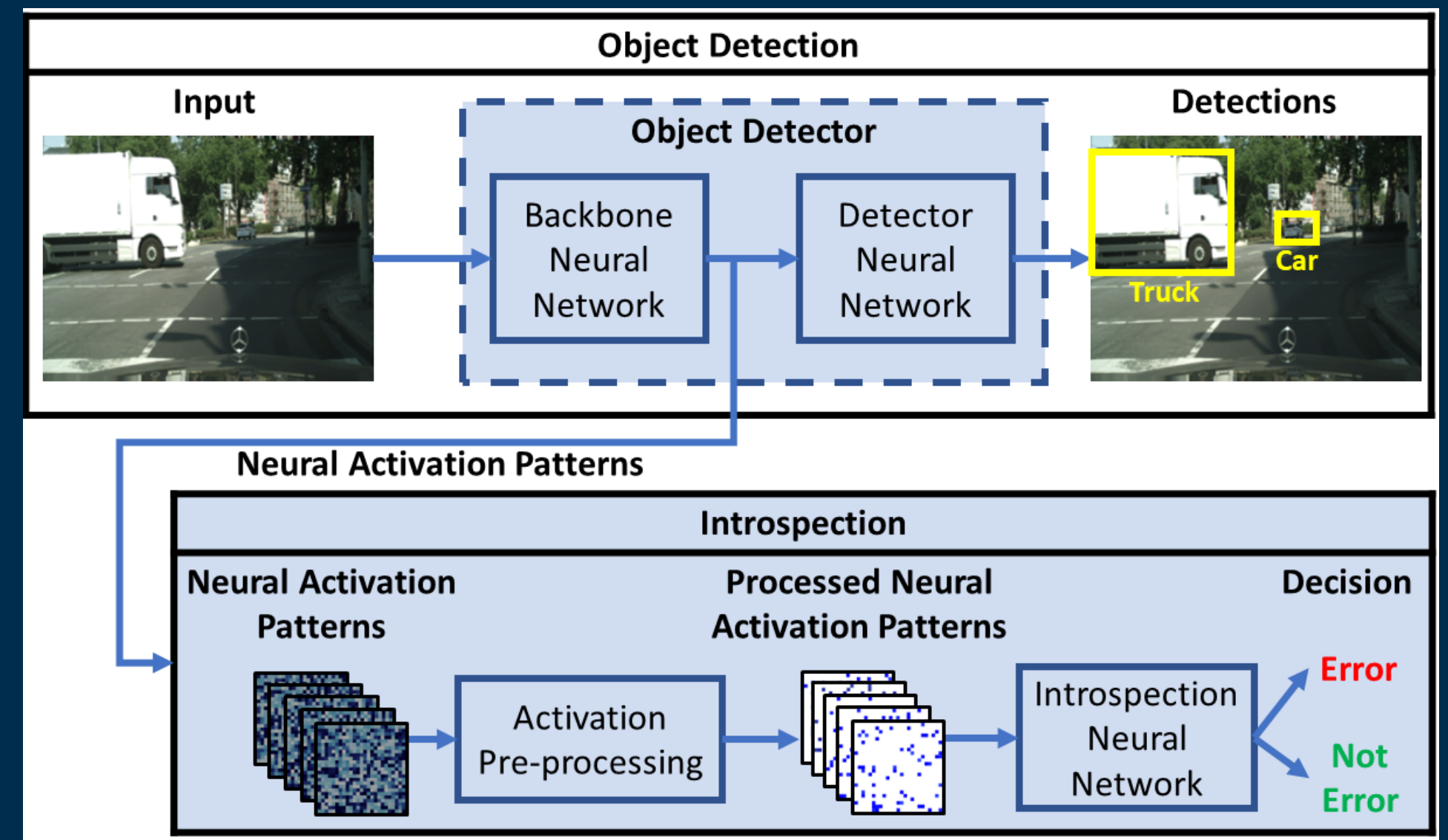
Introspection: Vital for DNN error detection.

Recent Findings:

- Neural patterns and metrics enable introspection.
- Simplified patterns improve out-of-distribution images for classification

Contribution:

- **Activation Simplification:** Adapts neural pattern pre-processing technique for 2D object detection introspection.
- **Four-Stage Framework:** Structured approach for introspection in driving datasets.
- **Evaluation:** Benchmarks with state-of-the-art methods.



Methodology

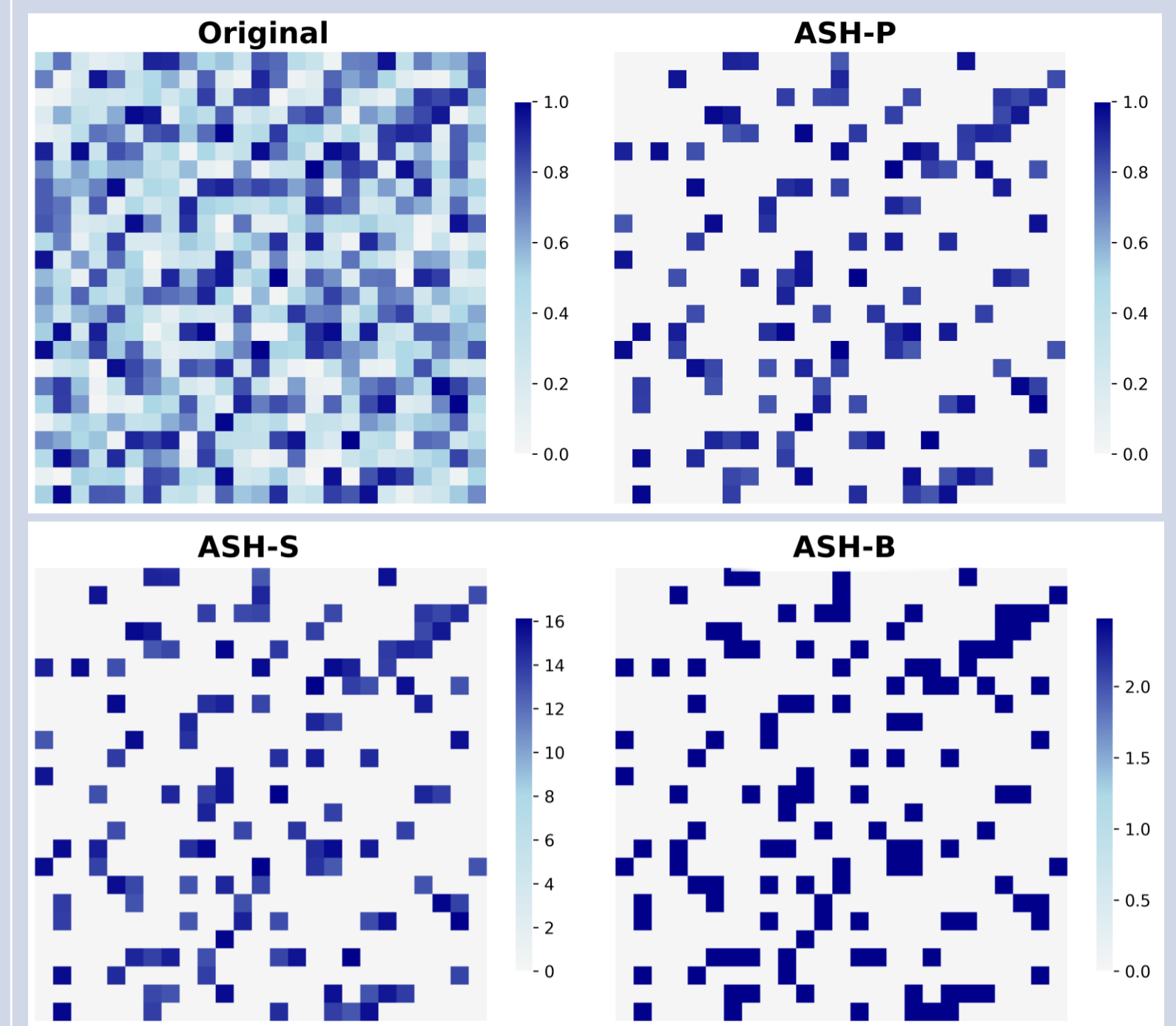
Pre-processing Method: Adopts "Activation Shaping" [1] to simplify raw neural patterns for Out-of-Distribution (OOD) detection.

Two-Stage Approach:

- Zeroes activation elements below a percentile threshold.
- Processes remaining activations with specific rules.

Modes:

- **ASH-P:** Keeps remaining activations as-is.
- **ASH-B:** Sets all values to a positive constant, calculated based on the sum of all activations.
- **ASH-S:** Scales up activations by a ratio calculated from the sum of activations before and after pruning.



Results

Mode Evaluation: Pruning alone yields the best results for introspection, contrasting with previous findings that favored scaling for out-of-distribution detection.

ASH-B: Excluded from evaluation due to consistent misclassification.

Comparison with SOTA: For the BDD dataset, pruning reduced the FNR by 14% and increased overall performance by 3%.



EVENTS project has received funding under grant agreement No 101069614. It is funded by the European Union. Views and opinions expressed are however those of the author(s) only and do not necessarily reflect those of the European Union or European Commission. Neither the European Union nor the granting authority can be held responsible for them.

Dataset	Type	Percentile	AUROC	FNR
BDD	S	90	0.7994	0.3302
		85	0.8057	0.2996
		80	0.7612	0.0180
		75	0.8021	0.0952
		70	0.7971	0.1114
	P	90	0.8009	0.2611
		85	0.8068	0.3521
		80	0.7972	0.2374
		75	0.8103	0.1069
		70	0.7999	0.2306
Dataset	Method	Percentile	AUROC	FNR
BDD	with ASH	75	0.8103	0.1069
	w/o ASH	-	0.7793	0.2439
	[2]	-	0.795	0.476

[1] Andrija Djuric, Nebojsa Bozanic, Arjun Ashok, and Rosanne Liu. Extremely simple activation shaping for out-of-distribution detection. 2022.

[2] Quazi Marufur Rahman, Niko S'underhauf, and Feras Dayoub. Per-frame map prediction for continuous performance monitoring of object detection during deployment. In IEEE Winter Conference on Applications of Computer Vision Workshops, pages 152–160, 2021.