

Anomaly Detection With Multiple Hypotheses Predictions

Duc Tam Nguyen^{1,2}, Zhongyu Lou², Michael Klar², Thomas Brox¹

¹Computer Vision Lab - University of Freiburg, Germany

²Robert Bosch GmbH, Germany



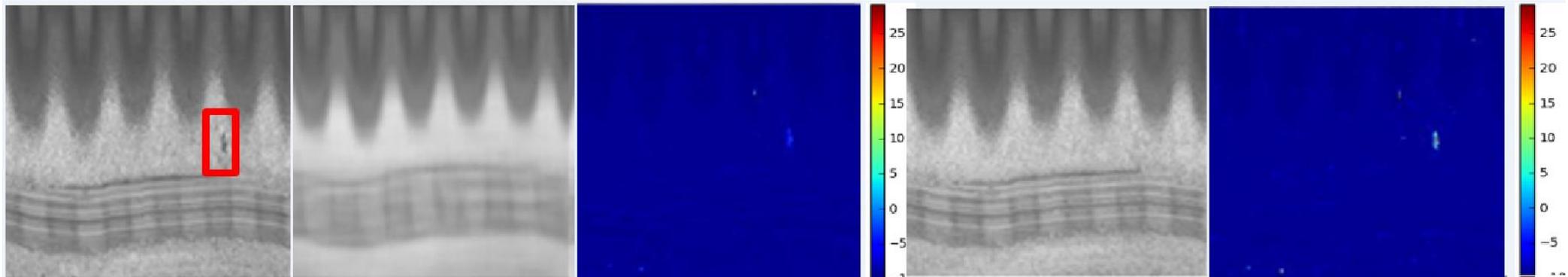
BOSCH

Introduction

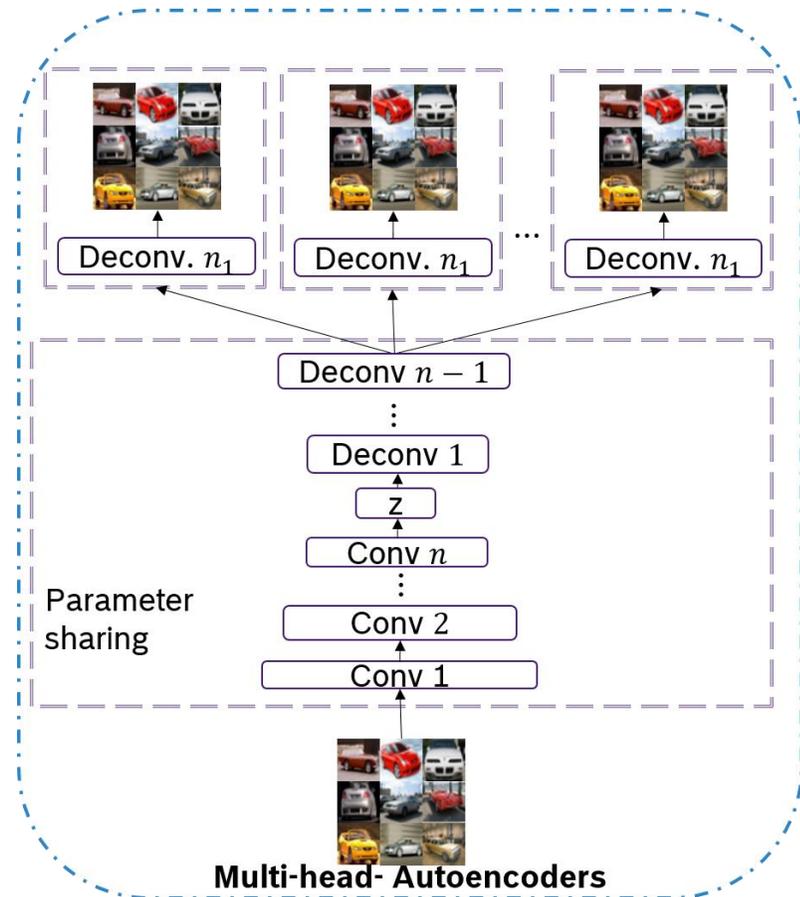
- Anomaly Detection is needed when a subset of classes is extremely rare or some classes are unknown at training time.
- Typically: Learn to approximate “normal” data distribution and measure deviation at test time.
- However:
 - Images inputs are high-dimensional → capturing the complete data density is **difficult** and **data-intensive**
 - In autoencoders, blurry reconstructions have the highest likelihood. But **blurry images** are also **anomalies**!

Introduction

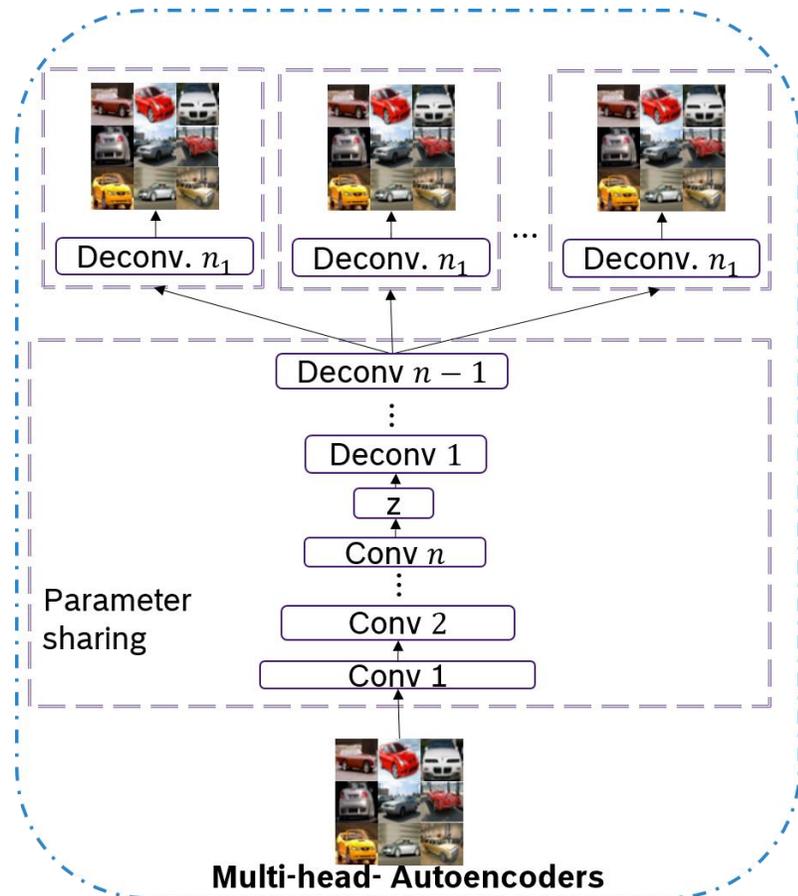
- Anomaly Detection is needed when a subset of classes is extremely rare or some classes are unknown at training time.
- Typically: Learn to approximate “normal” data distribution and measure deviation at test time.
- However:
 - Images inputs are high-dimensional → capturing the complete data density is **difficult** and **data-intensive**
 - In autoencoders, blurry reconstructions have the highest likelihood. But **blurry images** are also **anomalies!**



Capture multiple data modes with Multi-hypotheses-networks



Capture multiple data modes with Multi-hypotheses-networks

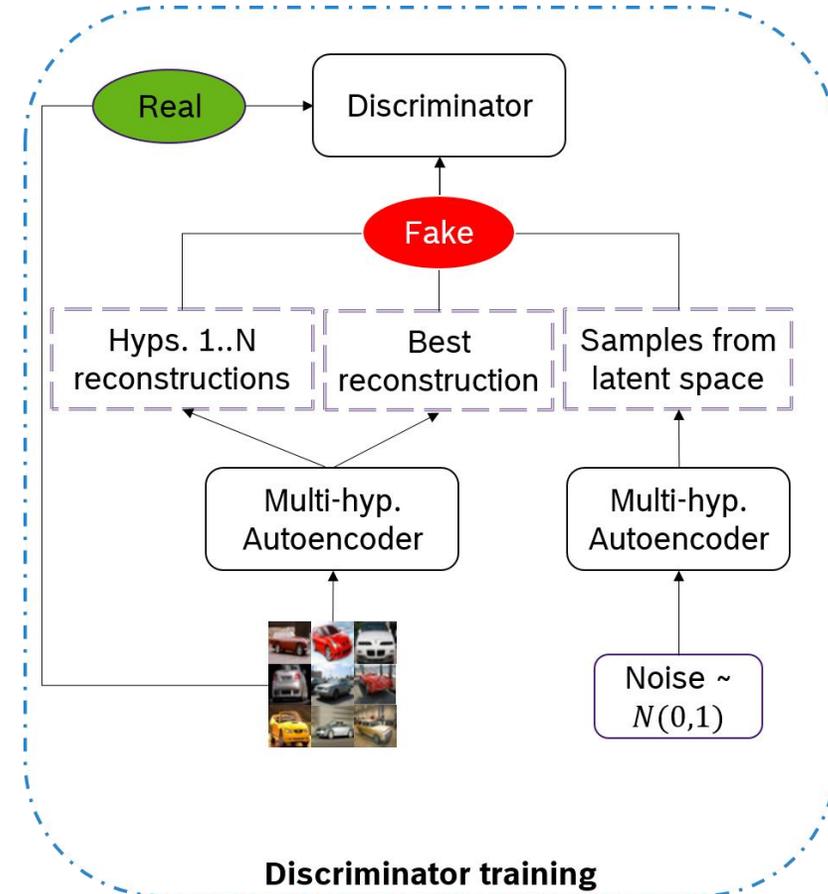
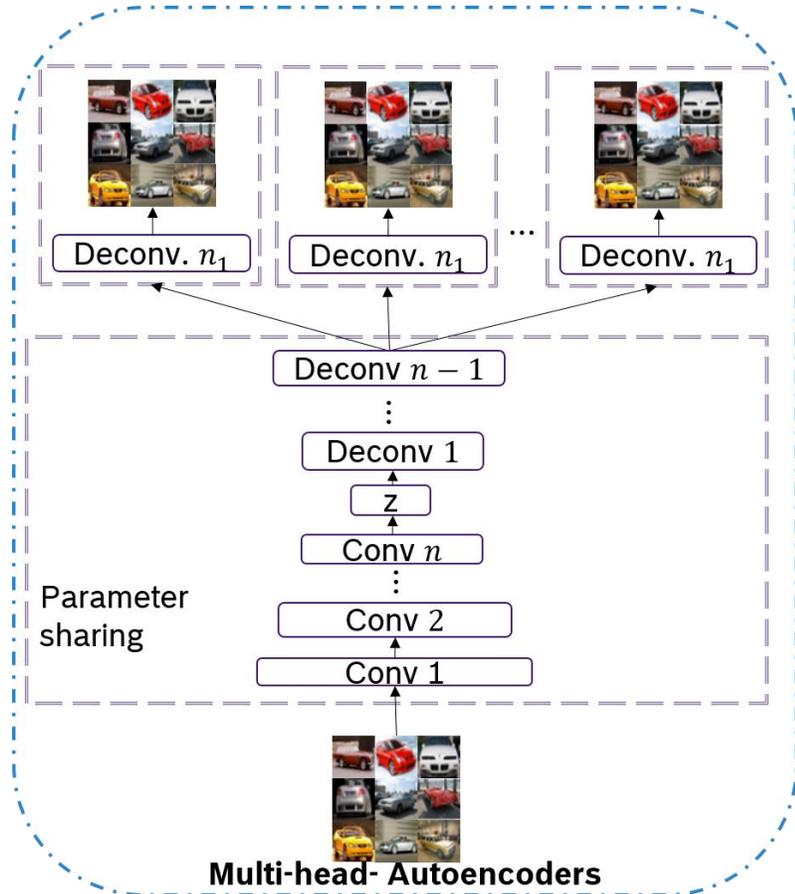


← Hypotheses could support non-existing data mode!

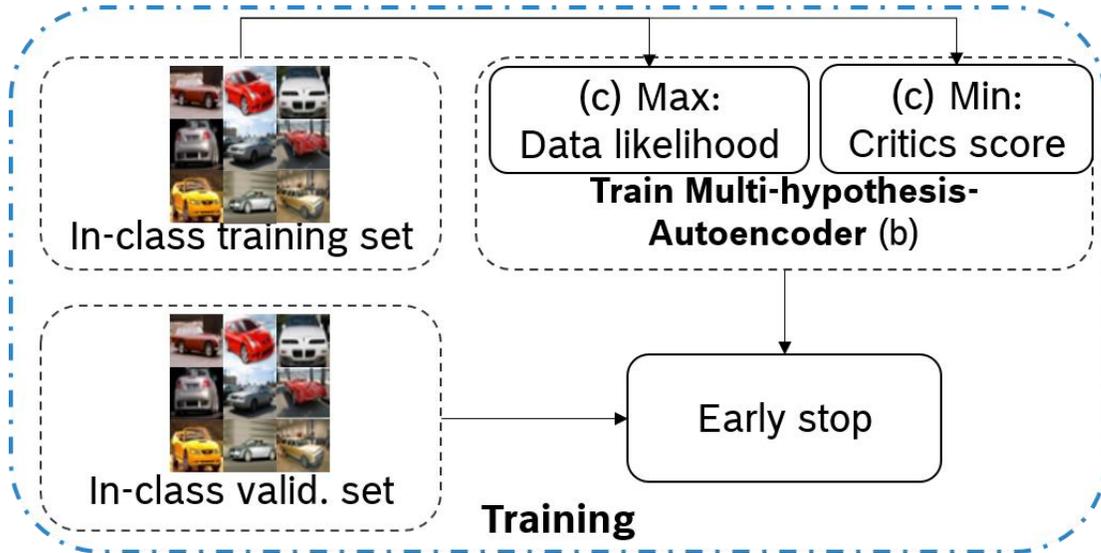
← Diversity among hypotheses?

→ **Unsuitable for anomaly detection** in this form !

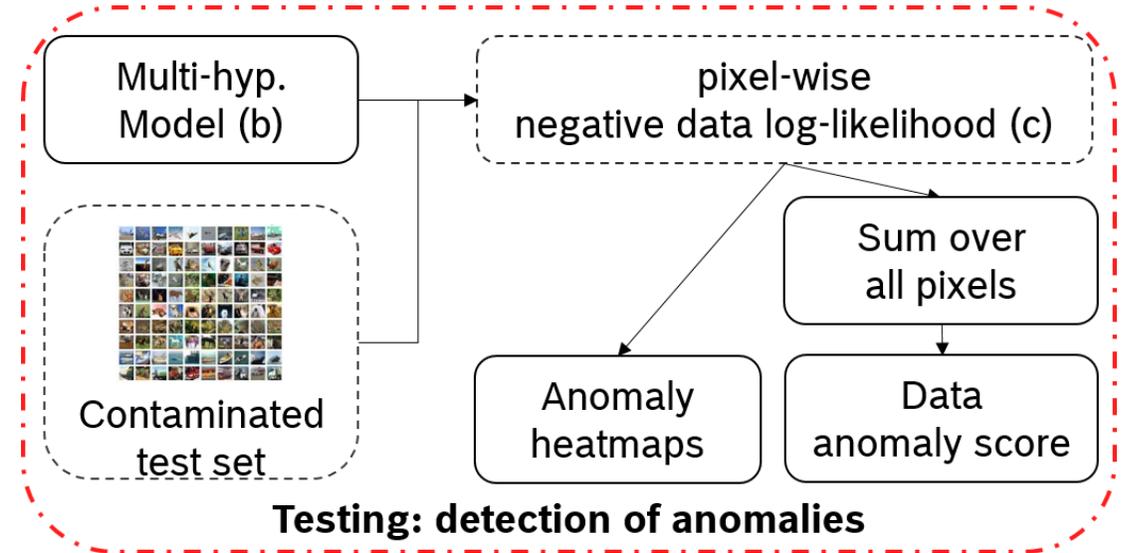
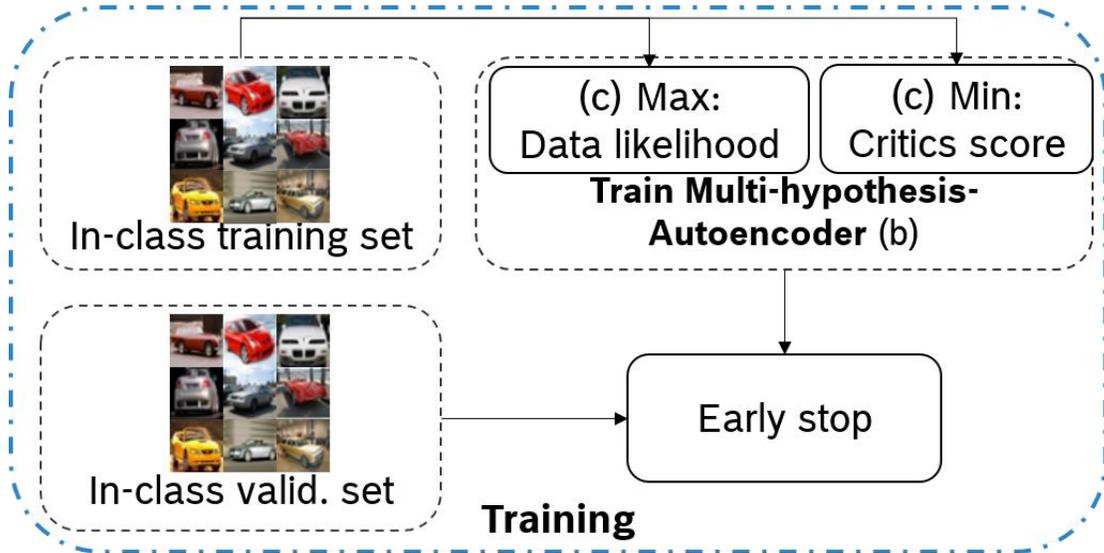
Adversarial regularizer



Anomaly detection with multiple hypotheses



Anomaly detection with multiple hypotheses



Experimental Results

Table 2. Anomaly detection on CIFAR-10, performance measured in AUROC. Each class is considered as the normal class once with all other classes being considered as anomalies, resulting in 10 one-vs-nine classification tasks. Performance is averaged for all ten tasks and over three runs each (see Appendix for detailed performance). Our approach significantly outperforms previous non-Deep Learning and Deep Learning methods.

TYPE	MODELS			
NON-DL.	KDE-PCA	OC-SVM-PCA	IF	GMM
	59.0	61.0	55.8	58.5
DL	ANoGAN	OC-D-SVDD	ADGAN	CONAD
	61.2	63.2	62.0	67.1

Experimental Results

Table 2. Anomaly detection on CIFAR-10, performance measured in AUROC. Each class is considered as the normal class once with all other classes being considered as anomalies, resulting in 10 one-vs-nine classification tasks. Performance is averaged for all ten tasks and over three runs each (see Appendix for detailed performance). Our approach significantly outperforms previous non-Deep Learning and Deep Learning methods.

TYPE	MODELS			
NON-DL.	KDE-PCA	OC-SVM-PCA	IF	GMM
	59.0	61.0	55.8	58.5
DL	ANOGAN	OC-D-SVDD	ADGAN	CONAD
	61.2	63.2	62.0	67.1

Table 6. Anomaly detection performance and their standard variance on the Metal Anomaly dataset. To reduce noisy residuals due to the high-dimensional input domain, only 10% of maximally abnormal pixels with the highest residuals are summed to form the total anomaly score. AUROC is computed on an unseen test set, a combination of normal and anomaly data. For more detailed results see Appendix. The anomaly detection performance of plain MHP rapidly breaks down with an increasing number of hypotheses.

MODEL	HYPOTHESES			
	1	2	4	8
MHP		98.0 (0.5)	97.0 (1.0)	95.0 (0.2)
MHP+WTA	94.2 (1.4)	98.0 (0.9)	98.0 (0.1)	94.6 (3.3)
MDN		90.0 (1.1)	91.0 (1.9)	91.6 (3.5)
MDN+GAN	93.6 (0.7)	94.2 (1.6)	91.3 (1.9)	94.3 (1.1)
CONAD		98.5 (0.1)	97.7 (0.5)	96.5 (0.2)

Visit our poster

- Our PosterID : #135 @ Pacific Ballroom



BOSCH

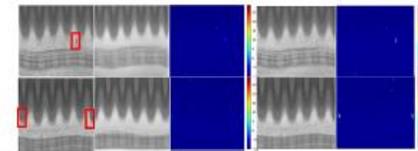


Anomaly Detection With Multiple Hypotheses Predictions

Duc Tam Nguyen^{1,2}, Zhongyu Lou², Michael Klar², Thomas Brox¹
¹Computer Vision Lab - University of Freiburg, Germany
²Robert Bosch GmbH, Germany

Introduction

- Anomaly Detection is needed when a subset of classes is extremely rare or some classes are unknown at training time.
- Typically: Learn to approximate "normal" data distribution and measure deviation at test time.
- However:
 - Images inputs are high-dimensional → capturing the complete data density is **difficult** and **data-intensive**
 - In autoencoders, blurry reconstructions have the highest likelihood. But **blurry images** are also **anomalies!**



Experiments

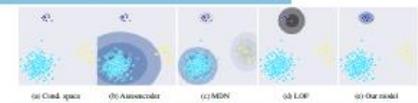


Figure 3. (a) Anomaly samples on Mind Anomaly data-set. (b) Anomalies are highlighted. (c) shows maximum likelihood reconstructions under a Variational Autoencoder and the corresponding anomaly heatmaps based on negative log-likelihood. (d) shows the reconstructions and anomaly maps for ConAD. Heat maps show the maximum likelihood expectations under the autoencoder model. Blurry and distorted samples are seen as an anomaly. Contrary, under our model, the maximum likelihood expectations of the input to match closely to the input and normal results. Due to the fine-grained learning, the anomaly heatmaps could robustly identify the location and strength of possible anomalies.

Table 2. Anomaly detection on CIFAR-10, performance measured in AUROC. Each class is considered as the normal class also with all other classes being considered as anomalies, resulting in 10 one-vs-one classification tasks. Performance is averaged for all one-tasks and over three runs each. Appendix for detailed performance. Our approach significantly outperforms previous state-of-the-art methods.

Type	Method	AUROC
Non-DL	KDE-PCA	0.19
	OC-SVM	0.18
	PCN	0.18
	GHM	0.18
DL	AutoGAN	0.12
	ADGAN	0.12

Table 3. Ablation study of our approach ConAD on CIFAR-10, measured in anomaly detection performance (AUROC-scores on 10 one-vs-one classification classes).

Configuration	AUROC
CONAD (w/o HYPOTHESES)	0.11
- LEARN HYPOTHESES (2)	0.15
- TEST DISCRIMINATOR	0.19
- WARMER (ALL LOGS (WTA))	0.18
- WTA & LOOF (WTF COUPLING)	0.11
- MULTIPLE HYPOTHESES	0.17
- MULTIPLE HYPOTHESES & DISCRIMINATOR	0.10

Table 4. Anomaly detection performance on CIFAR-10 (averaged on multiple hypotheses prediction models and hypothesis number). Performance averaged over tasks and its multiple runs each.

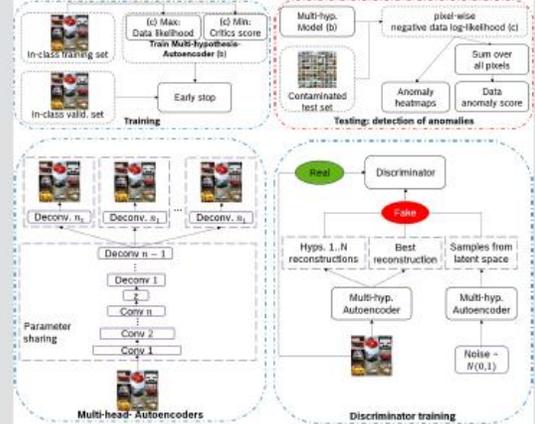
Model	HYPOTHESES				
	1	2	4	8	16
MHP	0.19	0.19	0.18	0.18	0.17
MHP+WTA	0.12	0.22	0.19	0.24	0.24
MDN	0.09	0.10	0.10	0.10	0.09
MDN+GAN	0.17	0.21	0.21	0.21	0.14
CONAD	0.13	0.20	0.21	0.21	0.24

Table 5. Anomaly detection performance on three standard tasks on the Mind Anomaly dataset. To make sure anomaly detection is the high-dimensional space domain, only 10% of anomaly observations with the highest probability are examined from the total anomaly score. ConAD is compared on an unseen test set, a combination of normal and anomaly data. For more detailed results see Appendix. The anomaly detection performance of plain-MHP equally benefits from an increasing number of hypotheses.

Model	HYPOTHESES				
	1	2	4	8	16
MHP	0.12	0.19	0.19	0.19	0.19
MHP+WTA	0.12	0.20	0.21	0.24	0.24
MDN	0.09	0.10	0.10	0.10	0.10
MDN+GAN	0.17	0.21	0.21	0.21	0.14
CONAD	0.13	0.20	0.21	0.21	0.24

Our approach

- We propose the use of multiple-hypotheses networks (MHP) (Rupprecht et al., 2016; Chen & Koltun, 2017; Ilg et al., 2018; Bhattacharyya et al., 2018) for anomaly detection
- It provides a more fine-grained description of the data distribution than with a single-headed network.
- We identify and address fake-data-support of MHP-techniques, which make them unsuitable for anomaly detection.
- Solution: ConAD in combination with a discriminator as a solution to avoid support of non-existent data regions and amplify the coverage of real data modes.



Conclusion

We propose an anomaly-detection approach that combines modeling the foreground class via multiple local densities with adversarial training. It results in significantly better anomaly detection performance.

References

- Bhattacharyya, A., Schiele, B., and Fritz, M. Accurate and diverse sampling of sequences based on a best of many sample objective. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, pp. 8485–8493, 2018.
- Rupprecht, C., Laina, I., DiPietro, R., Baust, M., Tombari, F., Navab, N., and Hager, G. D. Learning in an Uncertain World: Representing Ambiguity Through Multiple Hypotheses. In International Conference on Computer Vision (ICCV), 2017
- Ilg, E., C. Leik, O., Galesso, S., Klein, A., Makansi, O., Hutter, F., and Brox, T. Uncertainty Estimates with Multi-Hypotheses Networks for Optical Flow. In European Conference on Computer Vision (ECCV), 2018.