Supplementary Material: Locally Varying Distance Transform for Unsupervised Visual Anomaly Detection

1 Comparison with Semi-supervised Methods

The main paper focused on unsupervised anomaly detection, where no training data is available. Many anomaly detection papers focus on a different scenario, where no example of anomalies are provided a-priori; but a a clean set of normal data is available for training. This is termed semi-supervised anomaly detection or one-class learning. To put LVAD's results into perspective, we compare it against the current state-of-the-art in sem-supervised anomaly detection.

Following the experimental protocol established in Sec. 4.1 of the main paper, the definition of normality is rotated through each class of the dataset. When a class is designated as normal, an anomaly detector is trained on a subset of its instances. The remaining normal instances are mixed with instances from other (anomalous) classes to form the test set. The anomaly detector is tasked with identifying the anomalous instances. Performance is measured in terms of the Area Under the Receiver Operator Characteristic Curve.

Results are reported in Table 1, where LVAD and OCSVM [2] from the main paper are compared with state- deep learned techniques like Deep Structured Energy-based Model (DSEBM) [8], Anomaly Detection with a Generative Adversarial Network (AD-GAN) [3], Deep Anomaly Detection Using Geometric Transformations (Geo. Transform) [4] and Deep One-Class Classification (Deep SVDD) [7]. Similar to the main paper, for MNIST and Fashion-MNIST, LVAD used ratzerized pixels as its image representation; for the remaining datasets, images were represented using ResNet-50 [5] features. Semi-supervised anomaly detection were allowed to learn an appropriate representation from the training data.

Table 1 shows that although LVAD is not designed semi-supervised anomaly detection, its formulation generalizes well to this scenario. In Table 1, LVAD is consistently the best or close to the best algorithm. This is even true on MNIST and Fashion-MNIST, where its performance using raw pixels is respectably close to the current state-of-the-art.

2 Ablation Study: Effects of Instance Normalization

In the main paper, instance normalization [1] is applied to the data whenever possible. Table 2 illustrates the performance if normalization is not applied; for the reader's convenience, the normalized scores are copied from the main paper.

Algorithm	MNIST	Fashion-MNIST	STL-10	Internet-STL-10	MIT-Places-5	CatVsDog	CIFAR-10
LVAD (Ours)	0.948	0.905	**0.981**	**0.976**	**0.923**	**0.985**	**0.905**
DSEBM [8]	0.850	0.907	0.637	0.533	0.493	0.522	0.625
AD-GAN [3]	0.799	0.896	0.637	0.474	0.611	0.540	0.605
Geo. Transform [4]	**0.980**	**0.924**	0.887	0.829	0.638	0.809	0.885
Deep SVDD [7]	0.948	0.871	0.590	0.519	0.511	0.481	0.498

Table 1. Average AUROC of semi-supervised visual anomaly detectors. Comparing LVAD with state-of-the-art, deep-learned anomaly detectors.

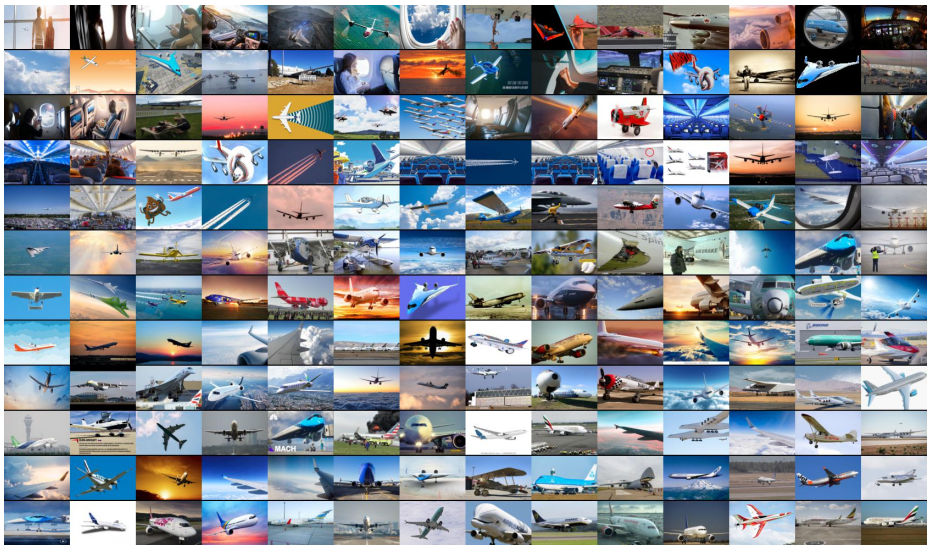
Instance normalization generally improves the performance of anomaly detectors. In the context of LVAD and OC-SVM, improvements were present in almost all datasets. For RSRAE, the improvements are not consistent, with performance gains mostly concentrated on the difficult datasets like on MNIST.

3 Qualitative Results

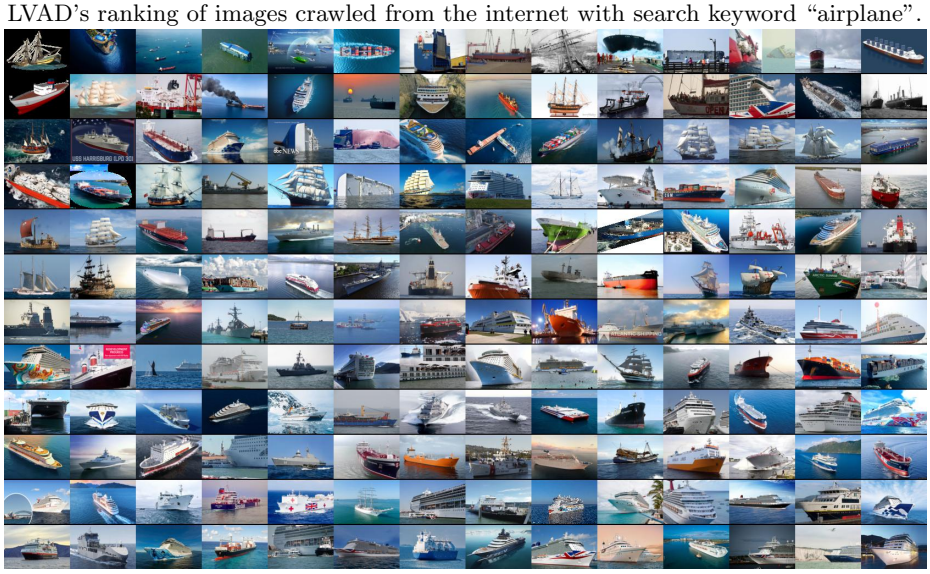
Finally, Fig. 1 and Fig. 2 provide qualitative evaluation of LVAD. Images are ranked by their normality scores, which increases from left to right, top to bottom. LVAD's scoring seems intuitive, with anomalous images arranged in the top rows and the normal ones congregating at the bottom.

Dataset	**Algorithm**	Anomaly Percentage (%)					**Ave.**	**Diff.**
		0.1	1	10	20	30		
STL-10 (ResNet-50)	LVAD (Ours) (Normalized)	0.998	0.993	0.979	0.983	0.977	**0.986**	**0.021**
	LVAD (Ours) (No norm.)	0.919	0.935	0.921	0.912	0.901	0.917	0.033
	RSRAE [6] (Normalized)	0.995	0.992	0.972	0.971	0.944	0.975	0.051
	RSRAE [6] (No norm.)	0.991	0.995	0.971	0.969	0.941	0.973	0.054
	OC-SVM [2] (Normalized)	0.996	0.995	0.967	0.877	0.777	0.922	0.219
	OC-SVM [2] (No norm.)	0.901	0.927	0.861	0.790	0.724	0.841	0.203
Internet STL-10 (ResNet-50)	LVAD (Ours) (Normalized)	0.997	0.997	0.996	0.985	0.981	**0.991**	**0.016**
	LVAD (Ours) (No norm.)	0.906	0.950	0.943	0.922	0.899	0.924	0.051
	RSRAE [6] (Normalized)	0.998	0.997	0.979	0.993	0.973	0.988	0.025
	RSRAE [6] (No norm.)	0.998	0.997	0.992	0.985	0.977	0.990	0.021
	OC-SVM [2] (Normalized)	0.999	0.997	0.985	0.908	0.817	0.941	0.182
	OC-SVM [2] (No norm.)	0.919	0.949	0.891	0.818	0.761	0.868	0.188
MIT-Places-5 (RestNet-50)	LVAD (Ours) (Normalized)	0.955	0.941	0.922	0.891	0.867	**0.915**	**0.088**
	LVAD (Ours) (No norm.)	0.728	0.679	0.630	0.570	0.536	0.628	0.191
	RSRAE [6] (Normalized)	0.965	0.928	0.893	0.686	0.605	0.815	0.360
	RSRAE [6] (No norm.)	0.967	0.916	0.903	0.814	0.663	0.853	0.304
	OC-SVM [2] (Normalized)	0.966	0.908	0.834	0.727	0.683	0.824	0.283
	OC-SVM [2] (No norm.)	0.620	0.727	0.617	0.571	0.563	0.620	0.164
CIFAR-10 (RestNet-50)	LVAD (Ours) (Normalized)	0.930	0.940	0.903	0.854	0.816	**0.889**	**0.124**
	LVAD (Ours) (No norm.)	0.781	0.853	0.810	0.756	0.714	0.783	0.139
	RSRAE [6] (Normalized)	0.901	0.911	0.800	0.814	0.739	0.833	0.172
	RSRAE [6] (No norm.)	0.919	0.912	0.796	0.787	0.755	0.834	0.164
	OC-SVM [2] (Normalized)	0.913	0.922	0.869	0.801	0.742	0.849	0.180
	OC-SVM [2] (No norm.)	0.749	0.829	0.787	0.728	0.678	0.754	0.152
CatVsDog (RestNet-50)	LVAD (Ours) (Normalized)	0.981	0.978	0.927	0.851	0.780	0.903	0.201
	LVAD (Ours) (No norm.)	0.827	0.801	0.739	0.685	0.651	0.741	0.175
	RSRAE [6] (Normalized)	0.982	0.981	0.961	0.917	0.835	**0.935**	**0.147**
	RSRAE [6] (No norm.)	0.981	0.959	0.923	0.889	0.797	0.910	0.184
	OC-SVM [2] (Normalized)	0.989	0.982	0.892	0.799	0.737	0.880	0.252
	OC-SVM [2] (No norm.)	0.794	0.810	0.743	0.682	0.642	0.734	0.168
MNIST (Rasterized Pixels)	LVAD (Ours) (Normalized)	0.974	0.948	0.938	0.923	0.904	**0.937**	**0.070**
	LVAD (Ours) (No norm.)	0.956	0.905	0.892	0.873	0.846	0.894	0.110
	RSRAE [6] (Normalized)	0.966	0.948	0.851	0.794	0.763	0.864	0.203
	RSRAE [6] (No norm.)	0.882	0.851	0.788	0.795	0.774	0.818	0.108
	OC-SVM [2] (Normalized)	0.937	0.901	0.885	0.856	0.824	0.881	0.113
	OC-SVM [2] (No norm.)	0.099	0.260	0.334	0.436	0.509	0.328	0.410
Fashion-MNIST (Rasterized Pixels)	LVAD (Ours) (Normalized)	0.896	0.909	0.899	0.884	0.868	**0.891**	**0.041**
	LVAD (Ours) (No norm.)	0.946	0.904	0.895	0.874	0.830	0.890	0.116
	RSRAE [6] (Normalized)	0.900	0.854	0.748	0.711	0.689	0.780	0.211
	RSRAE [6] (No norm.)	0.897	0.882	0.872	0.814	0.768	0.847	0.129
	OC-SVM [2] (Normalized)	0.875	0.898	0.889	0.867	0.843	0.874	0.055
	OC-SVM [2] (No norm.)	0.142	0.286	0.637	0.589	0.490	0.429	0.496

Table 2. Average AUROC of unsupervised visual anomaly detectors, with and without instance normalization. Ave. is the average score on a dataset; a high Ave. indicates accuracy. Diff. is the difference between the highest and lowest scores on a dataset; a small Diff. indicates stability.



LVAD's ranking of images crawled from the internet with search keyword "airplane".



LVAD's ranking of images crawled from the internet with search keyword "ship".

Fig. 1. Normality scores increase from left to right, top to bottom.

LVAD's ranking of images crawled from the internet with search keyword "bird".

LVAD's ranking of images crawled from the internet with search keyword "truck".

Fig. 2. Normality scores increase from left to right, top to bottom.

References

1. Ba, J.L., Kiros, J.R., Hinton, G.E.: Layer normalization. arXiv preprint arXiv:1607.06450 (2016)
2. Chen, Y., Zhou, X.S., Huang, T.S.: One-class svm for learning in image retrieval. In: ICIP. pp. 34–37. Citeseer (2001)
3. Deecke, L., Vandermeulen, R., Ruff, L., Mandt, S., Kloft, M.: Anomaly detection with generative adversarial networks. https://openreview.net/forum?id=S1EfylZ0Z (2018)
4. Golan, I., El-Yaniv, R.: Deep anomaly detection using geometric transformations. arXiv preprint arXiv:1805.10917 (2018)
5. He, K., Zhang, X., Ren, S., Sun, J.: Deep residual learning for image recognition. In: Proceedings of the IEEE conference on computer vision and pattern recognition. pp. 770–778 (2016)
6. Lai, C.H., Zou, D., Lerman, G.: Robust subspace recovery layer for unsupervised anomaly detection. arXiv preprint arXiv:1904.00152 (2019)
7. Ruff, L., Vandermeulen, R.A., Görnitz, N., Deecke, L., Siddiqui, S.A., Binder, A., Müller, E., Kloft, M.: Deep one-class classification. In: Proceedings of the 35th International Conference on Machine Learning. vol. 80, pp. 4393–4402 (2018)
8. Zhai, S., Cheng, Y., Lu, W., Zhang, Z.: Deep structured energy based models for anomaly detection. arXiv preprint arXiv:1605.07717 (2016)