Supplementary Materials for FlexUOD: The Answer to Real-world Unsupervised Image Outlier Detection

Zhonghang Liu
SMU, Singapore
zhliu.2020@phdcs.smu.edu.sg

Kun Zhou
SmartMore Corporation, China
zhoukun303808@gmail.com

Changshuo Wang
NTU, Singapore
wangchangshuo1@gmail.com

Wen-Yan Lin
SMU, Singapore
linwenyan.daniel@gmail.com

Jiangbo Lu
SmartMore Corporation, China
jiangbo.lu@gmail.com

A. Bray-Curtis Distance

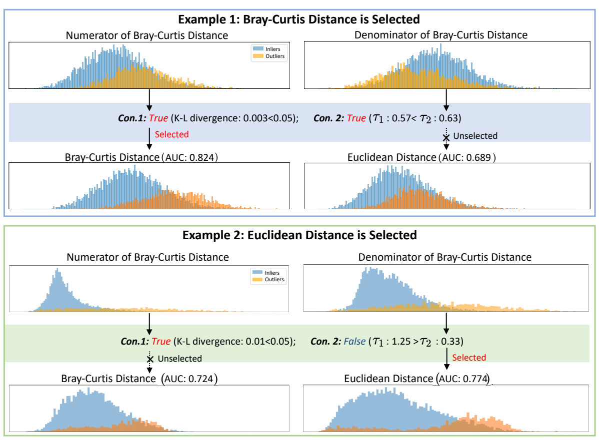

Figure 1. In the upper plot, its numerator and denominator exhibit near-symmetry, prompting the selection of the Bray-Curtis distance. In the below plot, the two conditions cannot be both satisfied so the default Euclidean distance will be selected. The AUC results also validate the accuracy of our distance selection.

In this study, we introduce Bray-Curtis distance, which is widely utilized in the fields of bioinformatics [1], to advance Euclidean distance. To our best knowledge, this is the first usage of Bray-Curtis distance for UOD. So we provide its empirical analysis for better implementation.

$$\text{Bray-Curtis}(x_i) = \frac{||x_i - m_{\mathbf{X}}||_1}{||x_i + m_{\mathbf{X}}||_1} = \frac{||x_i - m_{\mathbf{X}}||_1}{||x_i - (-m_{\mathbf{X}})||_1}. \tag{1}$$

Based on Eq. 1, the Bray-Curtis distance can be decomposed into its numerator $\text{nume}(x_i) = ||x_i - m_{\mathbf{X}}||_1$ and the denominator $\text{deno}(x_i) = ||x_i - (-m_{\mathbf{X}})||_1$, which show potential symmetry based on their similar structures, shown in Fig. 1 (upper plot). We first assume an ideal case (feature representation is separable), all outliers are located at greater ℓ_1-norm values from the mean $m_{\mathbf{X}}$ compared to any inlier, i.e.,

$$\begin{aligned}&\max(\{||x_i - m_{\mathbf{X}}||_1 | x_i \in \mathbf{X}_{in}\}) < \\ &\min(\{||x_i - m_{\mathbf{X}}||_1 | x_i \in \mathbf{X}_{out}\})\end{aligned} \tag{2}$$

Leveraging the symmetry between the numerator and denominator, we have:

$$\begin{aligned}&\max(\{||x_i - (-m_{\mathbf{X}})||_1 | x_i \in \mathbf{X}_{in}\}) > \\ &\min(\{||x_i - (-m_{\mathbf{X}})||_1 | x_i \in \mathbf{X}_{out}\})\end{aligned} \tag{3}$$

Thus, we have:

$$\begin{aligned}&\frac{\max(\{||x_i - m_{\mathbf{X}}||_1 | x_i \in \mathbf{X}_{in}\})}{\max(\{||x_i - (-m_{\mathbf{X}})||_1 | x_i \in \mathbf{X}_{in}\})} \ll \\ &\frac{\min(\{||x_i - m_{\mathbf{X}}||_1 | x_i \in \mathbf{X}_{out}\})}{\min(\{||x_i - (-m_{\mathbf{X}})||_1 | x_i \in \mathbf{X}_{out}\})},\end{aligned} \tag{4}$$

i.e., the separability between inliers and outliers will be enhanced, if the symmetry property is satisfied. To empirically define the symmetry, $\{\text{nume}(x_i)\}_{i=1}^n$ and $\{\text{deno}(x_i)\}_{i=1}^n$ should satisfy two following conditions:
Condition 1: Similar distributions.
To measure the similarity, we utilize K-L divergence [2], constraining its value as less than $\mathbf{0.05}$.

$$\begin{aligned}&\mathrm{D_{KL}}(\{\text{nume}(x_i)\}_{i=1}^n || \{\text{deno}(x_i)\}_{i=1}^n) \\ &= \sum_{i=1}^{n} \text{nume}(x_i) \log\left(\frac{\text{nume}(x_i)}{\text{deno}(x_i)}\right).\end{aligned} \tag{5}$$

Condition 2: Independent with each other.
In statistics, the "3-sigma" rule [3] suggests data points with more than three standard deviations σ from the mean μ can be considered as out-of-distribution instances. So we test whether:

$$\tau_1 < \tau_2, \tag{6}$$

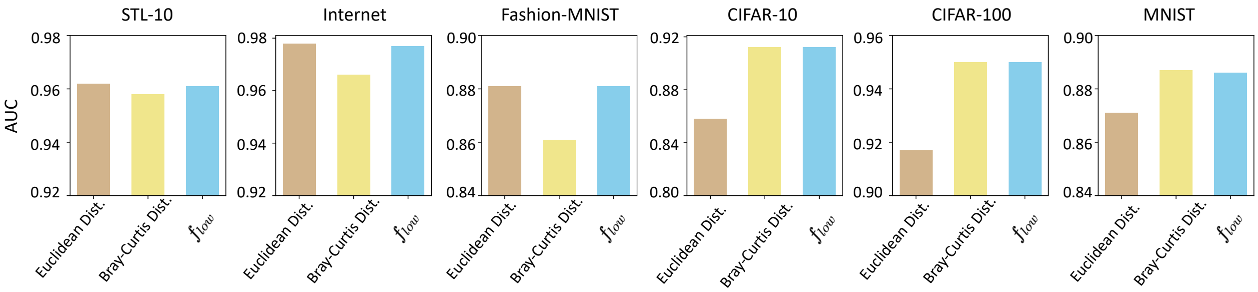

Figure 2. Average AUC results for different distance metrics.

where $\tau_1 = \mu(\{\text{nume}(x_i)\}_{i=1}^n) + 3 \cdot \sigma(\{\text{nume}(x_i)\}_{i=1}^n)$ and $\tau_2 = \mu(\{\text{deno}(x_i)\}_{i=1}^n) - 3 \cdot \sigma(\{\text{deno}(x_i)\}_{i=1}^n)$. If condition 1 and condition 2 are both satisfied, $\omega = 1$, Bray-Curtis distance (Eq. 6 in the main paper) will be selected; otherwise, $\omega = 0$, i.e., Euclidean distance (Eq. 5 in the main paper) will be selected. Fig. 2 demonstrates that our f_{low} always selects the optimal distance metrics across different scenarios.

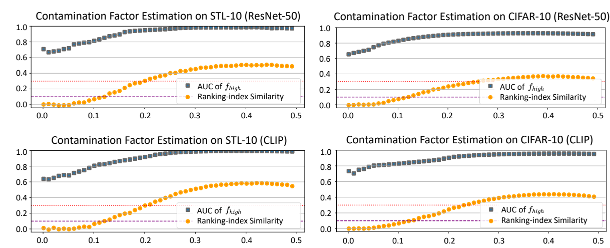

Figure 3. Contamination factor estimation on different datasets.

B. Comparison with LVAD at low γ

When γ is low (0.005, 0.01), FlexUOD is more stable on various datasets as well as higher efficiency (both running time and memory cost) compared with LVAD, as shown in Tab. 1.

Table 1. Multiple comparison with LVAD at **low-**γ scenarios.

Method	Different Dataset (Avg (ResNet-50, CLIP))				Efficiency (CPU)	
	STL	Internet	CIFAR-10	Average	Timing (#5000)	Memory Cost
LVAD	0.978	0.974	0.940	0.964	8.057 (sec.)	0.254 (MB)
FlexUOD	0.974	0.982	0.960	**0.972**	**0.118** (sec.)	**0.204** (MB)

C. Contamination Factor Estimation

Fig. 3 presents more examples for contamination factor estimation. First, the introduced f_{high} always performs well on high-γ scenarios while its performance significantly degrades in low-γ settings. Besides, the linearly-correlated property of ranking-index similar maintains on different target datasets.

References

[1] K Robert Clarke, Paul J Somerfield, and M Gee Chapman. On resemblance measures for ecological studies, including taxonomic dissimilarities and a zero-adjusted bray–curtis coefficient for denuded assemblages. *J. Exp. Mar. Biol. Ecol.*, pages 55–80, 2006. 1

[2] Thomas M Cover. *Elements of information theory*. 1999. 1

[3] Friedrich Pukelsheim. The three sigma rule. *The American Statistician*, pages 88–91, 1994. 1