Appendix

Zhonglin Wu

November 13, 2025

1 Appendix

To parse the loss function and region mining, we first denote the set systems in the whole framework, including the k closest neighbor search, labeling, and region extraction phase as class set $\mathbb{U}_{\mathbb{C}}$, k nearest neighborhood set σ_k and region set $\mathbb{R}$, where $\mathbb{U}_C = \{x \in \mathbb{U}_C | l(x) = c\}$ includes all data objects labeled as the same class c, and $\mathbb{U}_R = \{x \in \mathbb{U}_R | \sigma_k(x) \subseteq \mathbb{U}_R \wedge Area(\sigma_k(x)) \leq \tau\}$ includes data objects having dense k nearest neighborhood, and the convex area consisting of k nearest neighbors is constrained by a threshold τ.
According to the definitions, any region set $\mathbb{U}_{\mathbb{R}}$ intersected with more than two $\mathbb{U}_C$ is an overlapping set, the collection of overlapping sets is denoted as $\mathbb{U}_o = \{\mathbb{U}_R | \mathbb{U}_R \cap \mathbb{U}_{C_1} \wedge \mathbb{U}_R \cap \mathbb{U}_{C_2}\}$.
Then the amount of negative and positive pair for each lower approximation and boundary is correlated with the entanglement distribution between $\mathbb{U}_C$ and $\mathbb{U}_R$.
To formulate sample pairs for training, we denote the coverage rates between the region set and the class set as $r = arg\max_{i \in R} |\{\mathbb{U}_{C_j} \cap \mathbb{U}_{R_i} | j \in C\}|$.
As the sampling is conceptually aligned with the sampling principle of usually density-based clustering algorithms, we have

Lemma 1.0.1 *The total amount of sample pairs (x, x^+, x^-) under the sum of every pair of class labels, $\#(x, x^+, x^-) = O(|\mathbb{U}_o|^2) \times \binom{r}{2}$, where $|\mathbb{U}_o| = O(k)$.*

The metric sampling pairs assign region information to the loss function of deep neural network's training by metric structure, in order to sharpen the space supervision other than let the embedding space influenced by varying gradients, we give a theorem to illustrate that the loss function with cross set's separation is heavier structurally than loss only with lower approximation and boundary sampling.

Theorem 1.0.2 *Sampling with cross set's separation has a higher Dasgupta's cost for hierarchical clustering then sampling without separation.*

Proof: Dasgupta's cost for hierarchical clustering is defined as

$$Cost_G(\mathcal{T}) = \sum_{e=(u,v)\in E} w_e |leaves(u \wedge v)|$$

where w_e is the number of edges between two sibling nodes in hierarchical clustering tree $\mathcal{T}$ and $|leaves(u \wedge v)|$ means the number of nodes of the lowest common ancestor subtree of two leaf nodes u and v. The Cost is identically equal for every complete graph, if every pair of leaf nodes is participated in the cost function, and the cost could be reformulated as

$$\begin{aligned} Cost_G(\mathcal{T}) =& (n_k + n_{k-1} + ... + n_0) \times (n_{k-1} + ... + n_0) \times n_k + ...+ \\ & (n_2 + n_1 + n_0) \times (n_1 + n_0) \times n_2 + (n_1 + n_0) \times n_1 \times n_0 \\ =& (\sum_{i=0}^{k} n_i)^3 - \frac{\sum_{i=0}^{k} n_i^3}{3} \end{aligned}$$

If every smallest leaf node n_i is just one node, then every $n_i = 1$, the cost is identically equal. However, in the sampling case, $N^-(x) \geq 1$, so the $n_i \neq 1$, which causes different costs for different leaf node tree's assignments.

Denote three parts of for two objects x_1 and x_2, $N^-(x_1) = B_{N_R}(x_2)$ and $N^-(x_2) = B_{N_R}(x_1)$, $(\mathbb{U}_o(x_1, x_2) \cap \mathbb{U}_{C_1}) \cup (\mathbb{U}_o(x_1, x_2) \cap \mathbb{U}_{C_2}) \neq \emptyset$, $|N^-(x_1) \cap \mathbb{U}_{C_1}| = n_1$, $|N^-(x_2) \cap \mathbb{U}_{C_2}| = n_2$, $|\mathbb{U}_o(x_1, x_2) \cap \mathbb{U}_{C_1}| = n_3$ and $|\mathbb{U}_o(x_1, x_2) \cap \mathbb{U}_{C_2}| = n_4$.

So the loss function consisted by lower approximation's sampling and sampling pairs from set crosses is

$$\begin{aligned} Cost_G(\mathcal{T}_{cross}) &= n_1 \times (n_2 + n_3 + n_4) + n_2 \times (n_3 + n_4) + n_3 \times n_4 \\ &+ \sum_{i=1}^{4} Cost_G(\mathcal{T}_i) \\ Cost_G(\mathcal{T}) &= N_1 \times N_2 + \sum_{i=1}^{2} Cost_G(\mathcal{T}_i) \\ N_1 &= n_1 + n_3, N_2 = n_2 + n_4 \end{aligned}$$

According to basic inequality $\frac{a+b}{2} \geq \sqrt{ab}$, $\sum_{i=1}^{2} Cost_G(\mathcal{T}_i) \geq \sum_{i=1}^{4} Cost_G(\mathcal{T}_i)$, while $Cost_G(\mathcal{T}_{cross}) = Cost_G(\mathcal{T})$, so the set of sampling pairs with cross set is larger than without cross separation.

□

The instantiation of the sampling metric is proceeded by deep learning, more specifically, its gradient, so the triplet loss and hinge loss should guarantee that the network is well driven by the embedding space's region information, to illustrate the dynamic of metric-supervised loss functions, we have the next theorem to quantify the convergence of convex consisted of every k nearest neighbor $\mathbb{U}_R$ by geodesic in loss function's image.

Theorem 1.0.3 *For any $\boldsymbol{p}$-norm distances, where $\boldsymbol{p} \geq 2$, given any margin $\boldsymbol{\lambda} \geq 0$, the area of triangle consisted of representations of positive anchor $\boldsymbol{x_{ap}}$, negative anchor $\boldsymbol{x_{an}}$ and data $\boldsymbol{x}$ is both upper bounded and lower bounded.*

To prove the theorem, we first prove a lemma,

Lemma 1.0.4 *For every positive anchor $\boldsymbol{\sigma}$, there exists a geodesic on continuous conic section if both the distance $\boldsymbol{d_{pn}}$ between $\boldsymbol{\sigma}$ and its paired negative anchor $\boldsymbol{\xi}$ and margin $\boldsymbol{\lambda}$ are fixed.*

Proof: Let the dimension of representation space is d, denote the vector from σ to data x as $\overrightarrow{PX}$, the vector from σ to ξ as $\overrightarrow{PN}$, and the included angle between these two vectors is α, then the optimized margin satisfies $(\|\overrightarrow{PN}\| - \|\overrightarrow{PX}\|cos(\alpha))^p + (\|\overrightarrow{PX}\|sin(\alpha))^p - \|\overrightarrow{PX}\|^p \geq \lambda^p$.
Due to the margin loss penalizes the maximal distance range generated by two anchors, let's first calculate the distance range contours on the vector $\overrightarrow{PN}$, then expand the contour to full d-dimensional space by the included angle.
Let the point on $\overrightarrow{PN}$ be denoted as $\hat{x}$, then $\|\hat{x} - \xi\|^p - \|\hat{x} - \sigma\|^p \in [0, \frac{\|\overrightarrow{PN}\|^p}{2^p}]$, if $\hat{x}$ has positive label.
Due to the continuity of measurement, if $\lambda < \frac{\|\overrightarrow{PN}\|}{2}$, there exists a point x_λ satisfies $\|x_\lambda - \xi\|^p - \|x_\lambda - \sigma\|^p = \lambda^p$, which means there exists an angle interval $\alpha \in [0, \alpha_\lambda]$ for every $x \in (x_\lambda, \xi]$, satisfies $(\|\overrightarrow{PN}\| - \|\overrightarrow{PX}\|cos(\alpha))^p + (\|\overrightarrow{PX}\|sin(\alpha))^p - \|\overrightarrow{PX}\|^p \geq \lambda^p$. We denote the solutions of the equation $(\|\overrightarrow{PN}\| - \|\overrightarrow{PX}\|cos(\alpha))^p + (\|\overrightarrow{PX}\|sin(\alpha))^p - \|\overrightarrow{PX}\|^p = \lambda^p$ as $\{A_x, X_\alpha\}$. Then given a fixed magnitude of x, there's a $(d-1)$-dimensional curved surface, as a boundary for a $(d-1)$-dimensional conic section when $-x_\alpha \leq \alpha \leq x_\alpha$. And the monotonicity of $\alpha \in A_x$ and $x \in X_\alpha$ is inverse.
With the inverse monotonicity and the boundedness of α and x, the geodesic of λ across the d-dimensional hypersphere is not on the surface, because the curvature of the surface is a variable about p, so there's a geodesic on a continuous conic section.
□
Let denote the triangle between data x and positive anchor x_{ap}, negative anchor x_{an} as $\triangle_{xpn}$, then the three edges of Δ_{xpn} are the vector $\overrightarrow{PN}$, $\overrightarrow{PX}$ and $\overrightarrow{XN}$. Due to the lemma is already proved the monotonicity of α and x, the monotonicity of the magnitude of $\overrightarrow{PX}$ and the cosine value of $\langle \overrightarrow{PX}, \overrightarrow{PN} \rangle$ is inverse, when the magnitude of $\overrightarrow{PN}$ is fixed. Given a range for $\|\overrightarrow{PN}\|$, the area of Δ_{xpn} is both lower and upper bounded.

deduction 1.0.5 *In region set which has more than two intersected class set, the number of sample pair, the triangle that has area more than its upper bound, is $O(k \times \binom{r}{2})$.*

Because of the boundness of every sample pair in loss function, every pair of objects with incorrect distance bias, is probably out of the upper bound and trained to shrink its area in between the k nearest neighbor convex, due to the area of region would not be varied too violently, most of the $\mathbb{U}_R$ will be consistent on its area, so the minimum of the number of under margin sample pair in a k-vertices convex with area having low standard deviation should be $O(k)$, considering the number of class labels, in each epoch, the average sampling pair would be $O(k \times \binom{r}{2})$. This also makes the gradient more stable and gives the model more time to parse every region.

2 Detailed Experimental Results

2.1 Balanced accuracy result

Table 1: The balanced accuracy performance of all semi-supervised algorithms across datasets with different label ratios.

dataset	Label%	LNP	OReSSL	CDSC-AL	DEVDAN	ParsNet	CPSSDS	HDDS	Ours
	10	88.69 $\pm$ 0.69	93.07 $\pm$ 0.29	94.59 $\pm$ 0.54	71.41 $\pm$ 1.75	60.76 $\pm$ 1.14	75.52 $\pm$ 0.13	81.43 $\pm$ 1.35	**95.72 $\pm$ 0.43**
Syn-1	15	89.39 $\pm$ 1.90	93.07 $\pm$ 0.16	94.95 $\pm$ 1.49	71.42 $\pm$ 0.68	65.57 $\pm$ 0.24	75.03 $\pm$ 1.07	82.98 $\pm$ 0.81	**97.63 $\pm$ 0.71**
	20	90.43 $\pm$ 1.10	93.33 $\pm$ 1.28	96.34 $\pm$ 0.91	76.71 $\pm$ 1.50	73.06 $\pm$ 1.48	75.48 $\pm$ 0.21	84.79 $\pm$ 1.05	**98.15 $\pm$ 0.44**
	10	83.38 $\pm$ 1.54	84.81 $\pm$ 0.22	84.59 $\pm$ 1.46	40.69 $\pm$ 1.42	42.51 $\pm$ 1.62	47.19 $\pm$ 0.39	73.40 $\pm$ 0.39	**90.42 $\pm$ 1.93**
Syn-2	15	84.16 $\pm$ 0.58	85.42 $\pm$ 0.03	85.49 $\pm$ 0.56	44.85 $\pm$ 1.89	44.31 $\pm$ 0.95	47.30 $\pm$ 0.52	74.55 $\pm$ 1.15	**92.85 $\pm$ 1.21**
	20	85.62 $\pm$ 0.02	86.34 $\pm$ 0.46	86.16 $\pm$ 1.29	54.86 $\pm$ 0.44	49.35 $\pm$ 0.67	47.18 $\pm$ 0.35	75.51 $\pm$ 0.83	**93.51 $\pm$ 1.50**
	10	64.79 $\pm$ 1.04	88.41 $\pm$ 1.30	89.16 $\pm$ 0.10	10.40 $\pm$ 0.64	18.83 $\pm$ 0.16	48.07 $\pm$ 0.42	90.64 $\pm$ 0.90	**90.65 $\pm$ 0.32**
GSD	15	67.56 $\pm$ 0.49	92.02 $\pm$ 1.37	92.91 $\pm$ 1.75	11.03 $\pm$ 0.81	22.43 $\pm$ 1.14	51.51 $\pm$ 0.04	91.02 $\pm$ 1.80	**92.94 $\pm$ 0.17**
	20	67.91 $\pm$ 1.84	92.10 $\pm$ 0.12	92.93 $\pm$ 0.99	12.34 $\pm$ 0.79	32.15 $\pm$ 0.50	53.42 $\pm$ 0.86	91.53 $\pm$ 1.40	**94.02 $\pm$ 0.33**
	10	41.72 $\pm$ 1.78	47.09 $\pm$ 1.53	47.44 $\pm$ 0.78	33.66 $\pm$ 1.19	35.04 $\pm$ 1.74	47.96 $\pm$ 1.70	**55.42 $\pm$ 1.96**	48.39 $\pm$ 0.47
Shuttle	15	42.59 $\pm$ 1.71	50.04 $\pm$ 1.74	50.15 $\pm$ 0.80	32.76 $\pm$ 0.67	35.30 $\pm$ 1.56	50.19 $\pm$ 1.99	**59.40 $\pm$ 1.90**	48.39 $\pm$ 0.52
	20	44.36 $\pm$ 1.31	51.97 $\pm$ 0.65	52.23 $\pm$ 1.37	35.55 $\pm$ 1.12	35.97 $\pm$ 0.29	52.43 $\pm$ 0.43	**62.28 $\pm$ 1.96**	50.49 $\pm$ 0.58
	10	29.23 $\pm$ 1.71	84.37 $\pm$ 0.54	47.33 $\pm$ 0.39	29.97 $\pm$ 1.46	68.61 $\pm$ 1.16	51.36 $\pm$ 0.56	86.67 $\pm$ 0.18	**89.52 $\pm$ 0.46**
HAR	15	39.73 $\pm$ 0.52	88.59 $\pm$ 0.02	65.24 $\pm$ 0.19	31.97 $\pm$ 0.69	71.81 $\pm$ 0.49	51.98 $\pm$ 1.74	87.04 $\pm$ 0.32	**89.58 $\pm$ 0.23**
	20	44.74 $\pm$ 1.24	88.97 $\pm$ 1.58	76.88 $\pm$ 1.76	33.75 $\pm$ 0.98	72.46 $\pm$ 0.08	52.15 $\pm$ 0.36	87.41 $\pm$ 0.99	**90.36 $\pm$ 0.27**
	10	67.92 $\pm$ 0.44	48.52 $\pm$ 0.27	65.81 $\pm$ 1.12	72.66 $\pm$ 1.39	70.89 $\pm$ 1.05	73.92 $\pm$ 0.07	49.46 $\pm$ 1.03	**74.25 $\pm$ 0.47**
Wave_Su	15	68.75 $\pm$ 1.22	63.31 $\pm$ 1.78	66.33 $\pm$ 1.66	74.69 $\pm$ 0.80	71.01 $\pm$ 1.04	73.94 $\pm$ 1.69	50.31 $\pm$ 0.39	**75.82 $\pm$ 0.09**
	20	69.29 $\pm$ 0.16	71.92 $\pm$ 1.18	67.38 $\pm$ 1.37	76.21 $\pm$ 0.95	74.48 $\pm$ 0.18	74.08 $\pm$ 0.84	48.16 $\pm$ 1.20	**77.68 $\pm$ 0.04**
	10	65.35 $\pm$ 0.54	64.79 $\pm$ 0.11	57.25 $\pm$ 1.21	72.42 $\pm$ 0.58	70.80 $\pm$ 1.36	75.31 $\pm$ 0.50	46.08 $\pm$ 0.76	**77.53 $\pm$ 1.58**
Wave_Gr	15	65.98 $\pm$ 1.09	66.25 $\pm$ 1.62	61.30 $\pm$ 1.18	73.11 $\pm$ 0.11	72.92 $\pm$ 1.36	75.34 $\pm$ 0.79	46.84 $\pm$ 1.31	**78.08 $\pm$ 1.77**
	20	67.52 $\pm$ 1.24	72.27 $\pm$ 0.98	62.58 $\pm$ 0.44	75.11 $\pm$ 0.97	74.76 $\pm$ 1.56	75.46 $\pm$ 1.42	48.76 $\pm$ 0.62	**78.32 $\pm$ 1.36**
	10	18.56 $\pm$ 0.48	18.25 $\pm$ 1.54	37.33 $\pm$ 0.79	24.21 $\pm$ 0.42	22.45 $\pm$ 1.81	35.63 $\pm$ 1.15	37.41 $\pm$ 1.80	**44.92 $\pm$ 0.55**
HB	15	19.83 $\pm$ 0.89	19.71 $\pm$ 0.33	45.73 $\pm$ 0.41	25.18 $\pm$ 1.11	23.86 $\pm$ 0.08	36.97 $\pm$ 0.67	39.14 $\pm$ 1.33	**45.98 $\pm$ 0.45**
	20	21.47 $\pm$ 1.33	22.75 $\pm$ 1.00	46.28 $\pm$ 1.94	24.77 $\pm$ 1.28	24.90 $\pm$ 1.22	37.81 $\pm$ 0.91	38.52 $\pm$ 1.60	**47.98 $\pm$ 0.07**

2.2 Fmac result

Table 2: The Fmac performance of all semi-supervised algorithms across datasets with different label ratios.

dataset	Label%	LNP	OReSSL	CDSC-AL	DEVDAN	ParsNet	CPSSDS	HDDS	Ours
	10	79.39 $\pm$ 0.56	93.18 $\pm$ 0.15	94.90 $\pm$ 0.72	72.13 $\pm$ 1.37	57.16 $\pm$ 0.92	74.47 $\pm$ 1.37	67.23 $\pm$ 0.26	**95.71 $\pm$ 0.16**
Syn-1	15	81.26 $\pm$ 1.82	93.18 $\pm$ 0.89	95.37 $\pm$ 1.88	72.13 $\pm$ 0.15	63.75 $\pm$ 1.80	74.15 $\pm$ 0.38	68.74 $\pm$ 1.08	**97.23 $\pm$ 0.24**
	20	83.43 $\pm$ 1.77	94.28 $\pm$ 0.42	96.84 $\pm$ 0.68	76.75 $\pm$ 0.67	73.06 $\pm$ 0.79	75.42 $\pm$ 1.70	69.87 $\pm$ 0.55	**98.11 $\pm$ 0.47**
	10	63.75 $\pm$ 1.77	78.99 $\pm$ 1.21	81.49 $\pm$ 0.23	35.67 $\pm$ 1.65	34.76 $\pm$ 1.18	38.39 $\pm$ 1.95	51.07 $\pm$ 1.10	**91.01 $\pm$ 1.95**
Syn-2	15	65.98 $\pm$ 0.71	79.37 $\pm$ 1.55	82.87 $\pm$ 0.39	42.68 $\pm$ 0.67	38.37 $\pm$ 1.81	38.20 $\pm$ 1.48	51.61 $\pm$ 0.95	**93.46 $\pm$ 1.65**
	20	68.62 $\pm$ 0.16	79.88 $\pm$ 0.99	83.67 $\pm$ 0.23	52.92 $\pm$ 0.97	42.96 $\pm$ 1.61	38.46 $\pm$ 0.66	52.19 $\pm$ 1.28	**93.81 $\pm$ 1.93**
	10	66.11 $\pm$ 1.37	86.74 $\pm$ 1.10	89.95 $\pm$ 0.32	8.71 $\pm$ 1.36	16.67 $\pm$ 0.87	48.29 $\pm$ 1.08	90.50 $\pm$ 1.00	**90.52 $\pm$ 0.42**
GSD	15	68.58 $\pm$ 0.19	92.38 $\pm$ 1.27	92.99 $\pm$ 0.60	9.87 $\pm$ 0.93	19.91 $\pm$ 0.79	50.62 $\pm$ 1.78	91.06 $\pm$ 0.63	**93.13 $\pm$ 0.27**
	20	68.92 $\pm$ 0.45	92.48 $\pm$ 1.88	93.01 $\pm$ 1.41	11.71 $\pm$ 0.44	32.61 $\pm$ 1.89	53.28 $\pm$ 0.03	91.64 $\pm$ 1.74	**93.80 $\pm$ 0.32**
	10	41.19 $\pm$ 1.07	48.62 $\pm$ 0.82	47.89 $\pm$ 0.74	33.13 $\pm$ 1.56	36.36 $\pm$ 0.54	40.77 $\pm$ 0.38	**60.01 $\pm$ 1.61**	51.29 $\pm$ 0.45
Shuttle	15	42.56 $\pm$ 1.68	51.93 $\pm$ 1.80	52.05 $\pm$ 0.07	32.58 $\pm$ 1.02	36.67 $\pm$ 1.76	41.63 $\pm$ 0.41	**62.35 $\pm$ 1.32**	52.18 $\pm$ 1.09
	20	46.43 $\pm$ 0.19	53.61 $\pm$ 1.83	54.55 $\pm$ 0.60	36.26 $\pm$ 1.06	37.14 $\pm$ 1.38	47.09 $\pm$ 1.51	**64.62 $\pm$ 1.68**	54.70 $\pm$ 1.09
	10	25.49 $\pm$ 0.68	84.34 $\pm$ 0.26	40.12 $\pm$ 0.93	28.21 $\pm$ 0.79	68.60 $\pm$ 0.96	43.47 $\pm$ 0.22	86.65 $\pm$ 1.82	**89.56 $\pm$ 1.81**
HAR	15	37.84 $\pm$ 0.60	88.62 $\pm$ 1.59	61.00 $\pm$ 1.83	29.94 $\pm$ 1.78	71.64 $\pm$ 0.61	44.27 $\pm$ 1.89	87.08 $\pm$ 1.93	**89.64 $\pm$ 0.32**
	20	44.30 $\pm$ 0.05	89.06 $\pm$ 0.93	74.80 $\pm$ 0.64	31.12 $\pm$ 1.04	72.48 $\pm$ 0.10	44.45 $\pm$ 0.61	87.47 $\pm$ 0.36	**90.41 $\pm$ 0.19**
	10	70.86 $\pm$ 1.35	49.89 $\pm$ 0.14	57.81 $\pm$ 0.33	73.33 $\pm$ 0.29	69.01 $\pm$ 0.91	73.13 $\pm$ 1.33	30.18 $\pm$ 0.32	**73.75 $\pm$ 0.34**
Wave_Su	15	70.99 $\pm$ 0.37	59.08 $\pm$ 0.05	61.07 $\pm$ 0.99	73.80 $\pm$ 1.57	69.52 $\pm$ 1.21	74.02 $\pm$ 1.00	30.01 $\pm$ 0.09	**74.14 $\pm$ 0.31**
	20	71.64 $\pm$ 1.62	76.48 $\pm$ 0.67	63.01 $\pm$ 0.87	74.88 $\pm$ 0.76	73.58 $\pm$ 0.42	74.37 $\pm$ 1.51	28.71 $\pm$ 0.52	**76.50 $\pm$ 0.42**
	10	67.35 $\pm$ 1.40	62.03 $\pm$ 0.66	46.53 $\pm$ 1.72	74.02 $\pm$ 0.15	71.58 $\pm$ 1.85	**74.26 $\pm$ 0.42**	27.16 $\pm$ 0.25	65.75 $\pm$ 0.45
Wave_Gr	15	67.68 $\pm$ 1.64	71.73 $\pm$ 1.18	51.05 $\pm$ 1.43	74.29 $\pm$ 0.15	71.37 $\pm$ 1.08	74.82 $\pm$ 1.94	28.02 $\pm$ 0.83	**75.16 $\pm$ 0.13**
	20	69.15 $\pm$ 0.46	77.02 $\pm$ 1.25	54.10 $\pm$ 0.99	74.58 $\pm$ 0.83	72.95 $\pm$ 0.64	75.51 $\pm$ 0.46	28.87 $\pm$ 0.34	**77.19 $\pm$ 0.39**
	10	18.22 $\pm$ 1.33	18.07 $\pm$ 0.44	37.74 $\pm$ 1.36	22.84 $\pm$ 0.56	26.31 $\pm$ 1.80	37.43 $\pm$ 1.27	27.31 $\pm$ 1.06	**46.43 $\pm$ 1.62**
HB	15	19.59 $\pm$ 0.40	19.13 $\pm$ 0.08	44.51 $\pm$ 0.06	23.53 $\pm$ 0.83	28.32 $\pm$ 1.49	38.26 $\pm$ 1.43	30.49 $\pm$ 1.11	**46.89 $\pm$ 1.24**
	20	21.21 $\pm$ 0.69	23.90 $\pm$ 0.82	45.00 $\pm$ 1.10	25.45 $\pm$ 0.95	30.29 $\pm$ 1.34	39.74 $\pm$ 1.85	27.04 $\pm$ 1.07	**48.75 $\pm$ 1.50**

2.3 G-mean result

Table 3: The G-mean performance of all semi-supervised algorithms across datasets with different label ratios.

dataset	Label%	LNP	OReSSL	CDSC-AL	DEVDAN	ParsNet	CPSSDS	HDDS	Ours
Syn-1	10	83.21 ± 1.69	95.53 ± 0.97	82.92 ± 1.27	82.98 ± 0.43	75.90 ± 1.30	85.20 ± 1.09	81.27 ± 0.01	**97.57 ± 0.95**
	15	83.46 ± 0.32	97.48 ± 1.99	83.30 ± 0.53	82.99 ± 0.96	79.12 ± 0.27	85.47 ± 0.80	82.14 ± 1.23	**98.66 ± 0.25**
	20	83.89 ± 1.44	98.01 ± 0.44	83.69 ± 1.20	86.30 ± 0.30	83.94 ± 0.10	85.51 ± 0.64	83.11 ± 0.94	**98.47 ± 0.71**
Syn-2	10	62.36 ± 1.12	90.24 ± 1.60	80.86 ± 0.44	61.69 ± 0.75	59.78 ± 0.32	66.26 ± 0.60	69.49 ± 1.98	**94.54 ± 1.69**
	15	63.51 ± 0.29	92.46 ± 1.08	83.19 ± 0.31	65.10 ± 0.12	61.18 ± 0.05	66.29 ± 1.86	70.07 ± 1.96	**95.95 ± 1.85**
	20	68.74 ± 1.75	93.81 ± 1.28	84.04 ± 0.95	72.33 ± 0.39	64.89 ± 1.14	66.35 ± 1.89	70.55 ± 0.89	**96.32 ± 1.78**
GSD	10	71.48 ± 1.28	93.94 ± 0.17	77.80 ± 0.64	29.26 ± 1.23	39.98 ± 1.22	65.73 ± 1.70	94.21 ± 0.29	**94.32 ± 0.24**
	15	73.59 ± 0.13	94.39 ± 1.49	79.36 ± 1.79	30.09 ± 1.38	49.59 ± 1.30	68.25 ± 0.09	94.46 ± 0.13	**95.66 ± 0.24**
	20	75.42 ± 0.53	94.67 ± 1.26	80.65 ± 0.32	31.90 ± 0.02	52.21 ± 0.48	69.61 ± 0.70	94.73 ± 1.14	**96.37 ± 0.27**
Shuttle	10	67.72 ± 0.04	66.33 ± 0.08	19.17 ± 1.96	55.28 ± 0.45	57.48 ± 0.70	69.50 ± 1.22	**74.28 ± 1.44**	69.39 ± 0.83
	15	68.66 ± 1.67	66.61 ± 1.46	19.45 ± 0.65	56.14 ± 1.27	57.71 ± 0.99	70.45 ± 1.44	**76.98 ± 1.51**	70.50 ± 1.72
	20	69.71 ± 1.47	66.70 ± 1.55	19.86 ± 1.17	58.06 ± 1.78	58.41 ± 0.79	70.52 ± 0.63	**78.84 ± 1.63**	71.78 ± 1.56
HAR	10	49.96 ± 0.37	90.42 ± 0.33	80.16 ± 1.19	50.88 ± 0.26	80.25 ± 1.23	67.87 ± 1.43	91.80 ± 1.89	**93.24 ± 0.31**
	15	52.74 ± 1.48	91.01 ± 0.69	81.71 ± 0.39	54.06 ± 1.25	82.39 ± 1.15	68.33 ± 1.68	92.03 ± 1.73	**93.68 ± 0.46**
	20	53.97 ± 1.34	92.26 ± 0.23	83.21 ± 0.38	54.64 ± 0.15	82.83 ± 1.51	68.45 ± 1.32	92.29 ± 1.39	**94.45 ± 1.46**
Wave_Su	10	75.37 ± 1.83	65.90 ± 0.69	72.42 ± 1.63	79.96 ± 1.84	77.79 ± 1.69	75.14 ± 0.19	59.11 ± 0.26	**81.01 ± 0.06**
	15	77.58 ± 1.93	72.67 ± 0.84	73.64 ± 1.57	81.55 ± 1.61	78.41 ± 0.64	75.32 ± 1.45	59.84 ± 0.40	**82.44 ± 0.16**
	20	79.34 ± 0.73	80.15 ± 1.53	74.69 ± 1.16	82.88 ± 0.46	81.29 ± 1.61	75.91 ± 0.14	58.14 ± 1.73	**83.44 ± 0.39**
Wave_Gr	10	72.41 ± 1.24	74.12 ± 0.83	63.28 ± 0.76	79.29 ± 1.93	79.25 ± 0.76	78.37 ± 0.51	56.49 ± 0.59	**82.59 ± 1.21**
	15	74.68 ± 1.75	74.54 ± 1.23	64.78 ± 0.88	82.46 ± 0.03	79.82 ± 1.04	79.39 ± 0.49	57.16 ± 1.19	**84.20 ± 0.85**
	20	77.37 ± 0.20	80.84 ± 1.51	66.34 ± 0.23	82.56 ± 0.89	80.47 ± 0.48	80.47 ± 1.23	58.75 ± 1.09	**85.53 ± 1.79**
HB	10	38.27 ± 1.69	40.12 ± 0.06	52.87 ± 0.50	46.02 ± 0.93	45.48 ± 1.09	53.37 ± 1.18	54.13 ± 1.08	**65.48 ± 0.59**
	15	39.46 ± 0.17	41.88 ± 1.21	53.94 ± 1.02	47.19 ± 0.24	46.31 ± 0.75	56.75 ± 1.94	58.16 ± 0.75	**66.27 ± 0.61**
	20	42.33 ± 1.04	45.15 ± 1.11	55.68 ± 0.27	49.37 ± 1.79	47.93 ± 0.31	58.62 ± 1.48	54.71 ± 0.99	**67.40 ± 0.90**