Supplementary Materials: Decision Boundary Computation-based Over-sampling for Imbalance Learning

I. THEORETICAL ANALYSIS OF DBO

In the following, we first give the division of data space for the convenience of analysis. Then, we give the definition of local decision boundary area to explain how to set the α_2. Next, we give the definition of local boundary majority area to explain how to set the α_5. Next, we give the definition of local minority area to explain why DBO conducts two zooming operations on the local decision boundary area. Finally, we give the well performing and poorly performing scenarios of DBO.

The classifier generally divides the data space ($\Re^n$) into different partitions that belong to different classes. For example, the data space of imbalance dataset is divided into 5 partitions, where two partitions belong to the minority class and another three belong to the majority class. We denote the majority area as A_1 and the minority area as A_2, where $A_1 \bigcap A_2 = \emptyset$ and $A_1 \bigcup A_2 = \Re^n$.

A. local decision boundary area of DBO

Definition 1.1 (local decision boundary area): A continuous set S is the local decision boundary area if $S \bigcap A_1 \neq \emptyset$ and $S \bigcap A_2 \neq \emptyset$.

Given one group of samples $G = \{w, \frac{z_1 + w}{2}, \frac{z_2 + w}{2}, ..., \frac{z_k + w}{2}\}$, where w is one boundary minority sample and $z_1, z_2, ..., z_k$ are its k nearest neighbouring majority. Compute the mean and covariance matrix of group set G as $\bar{x}^{DB}$ and U^{DB}, and assign the corresponding length as α_2, obtain the set $S^{DB} = \{x| \ (x - \bar{x}^{DB})^T (U^{DB})^{-1} (x - \bar{x}^{DB}) \leq \alpha_2\}$.

Theorem 1.1: When $w \in S^{DB}$ and $\exists z \in \{z_1, z_2, ..., z_k\}$ makes $z \in S^{DB}$, S^{DB} is the local decision boundary area.

Proof: Since $w, z \in S^{DB}$, thus $S^{DB} \bigcap A_1 \neq \emptyset$ and $S^{DB} \bigcap A_2 \neq \emptyset$ that S^{DB} is the local decision boundary area. ∎

Thus, we set $\alpha_2 = 2 * (w - \bar{x}^{DB})^T (U^{DB})^{-1} (w - \bar{x}^{DB})$. In the one hand, $w \in S^{DB}$. In the other hand, it is of a high probability that $\exists z \in \{z_1, z_2, ..., z_k\}$ makes $z \in S^{DB}$.

Besides, The intersection of $S^{DB} \bigcap A_2$ tends to cover enough minority area.

B. local majority area of DBO

Definition 1.2 (local majority area): A continuous set S is the local majority area if $S \bigcap A_1 \neq \emptyset$, $S \bigcap A_2 = \emptyset$.

Given one group of samples $B_2 = \{z_1, z_2, ..., z_k\}$ where z_k is the k-th neighbouring majority of w. Compute the mean and covariance matrix of group set B_2 as $\bar{z}$ and U^{MAJ}, and assign the corresponding length as α_5, obtain the set $S^{MAJ} = \{x| \ (x - \bar{z})^T (U^{MAJ})^{-1} (x - \bar{z}) \leq \alpha_5\}$.

Theorem 1.2: When $S^{MAJ} \bigcap A_2 = \emptyset$, $\exists z \in \{z_1, z_2, ..., z_k\}$ makes $z \in S^{MAJ}$, S^{MAJ} is the local majority area.

Proof: Since $z \in S^{MAJ}$, thus $S^{MAJ} \bigcap A_1 \neq \emptyset$; and $S^{MAJ} \bigcap A_2 = \emptyset$ that S^{MAJ} is the local majority area. ∎

Thus, we set $\alpha_5 = \alpha_4 + 0.5 \times |\alpha^{min} - \alpha_4|$. Where $\alpha_4 = (z_1 - \bar{z})^T (U^{MAJ})^{-1} (z_1 - \bar{z})$, z_1 is the nearest majority one to w, $\alpha^{min} = (w - \bar{z})^T (U^{MAJ})^{-1} (w - \bar{z})$. On the one hand, it is of a high probability that $\exists z \in \{z_1, z_2, ..., z_k\}$ makes $z \in S^{MAJ}$, thus $S^{MAJ} \bigcap A_1 \neq \emptyset$. On the other hand, imaging that all $z_1, z_2, ..., z_k$ locate in the left side when compared to w in the 2-D data space, resulting $w \notin S^{MAJ}$. Thus, It is of a high probability that $w \notin S^{MAJ}$ and $S^{MAJ} \bigcap A_2 = \emptyset$.

Besides, α_5 makes S^{MAJ} to nearly dividing the local decision boundary area into two halves.

C. local minority area of DBO

Definition 1.3 (local minority area): A continuous set S is the local minority area if $S \bigcap A_1 = \emptyset$ and $S \bigcap A_2 \neq \emptyset$.

Given $S^{DB} = \{x| \ (x - \bar{x}^{DB})^T (U^{DB})^{-1} (x - \bar{x}^{DB}) \leq \alpha_2\}$ and $S^{MAJ} = \{x| \ (x - \bar{z})^T (U^{MAJ})^{-1} (x - \bar{z}) \leq \alpha_5\}$, we obtain the set $S^{MIN} = S^{DB} - S^{DB} \cap S^{Maj}$.

Theorem 1.3: When $(S^{DB} \bigcap A_1) \subseteq (S^{MAJ} \bigcap A_1)$, S^{MIN} is the local minority area.

Proof: Since $(S^{DB} \bigcap A_1) \subseteq (S^{MAJ} \bigcap A_1)$, thus $S \bigcap A_1 = \emptyset$ and $S \bigcap A_2 \neq \emptyset$ that S^{MIN} is the local minority area. ∎

Thus, instead of $\{w, z_1, z_2, ..., z_k\}$, we first use $G = \{w, \frac{z_1+w}{2}, \frac{z_2+w}{2}, ..., \frac{z_k+w}{2}\}$ as the basic group of samples of S^{DB}; we then assign $\alpha_2 = 2 * (w - \bar{x}^{DB})^T (U^{DB})^{-1} (w - \bar{x}^{DB})$ for S^{DB}. Obviously, they are two zooming operations that conducted on S^{DB} where the first operation double down this area and the second operation double up this area. The the first operation make S^{DB} not deeply rooting into the majority area that try to make the set of $(S^{DB} \bigcap A_1)$ belong to the set of $(S^{MAJ} \bigcap A_1)$.

D. well performing and poorly performing scenarios of DBO

For convenience, we divide the majority area as the union of inner majority area (A_{11}) and non-inner (or boundary) majority area (A_{12}), where $A_{11} \bigcap A_{12} = \emptyset$ and $A_{11} \bigcup A_{12} = A_1$. Similarly, we divide the minority area as the union of inner minority area (A_{21}) and non-inner (or boundary) minority area (A_{22}) , where $A_{21} \bigcap A_{22} = \emptyset$ and $A_{21} \bigcup A_{22} = A_2$.

As seen in Theorem 1.3, S^{MIN} is the local minority area when:

$$\begin{aligned}
(S^{DB} \bigcap A_1) \subseteq (S^{MAJ} \bigcap A_1) &\Leftrightarrow (S^{DB} \bigcap (A_{11} \bigcup A_{12}) \subseteq (S^{MAJ} \bigcap (A_{11} \bigcup A_{12})) \\
&\Leftrightarrow (S^{DB} \bigcap A_{11}) \bigcup (S^{DB} \bigcap A_{12}) \subseteq (S^{MAJ} \bigcap A_{11}) \bigcup (S^{MAJ} \bigcap A_{12}) \\
&\Leftrightarrow (S^{DB} \bigcap A_{11}) \subseteq (S^{MAJ} \bigcap A_{11}) \quad \& \quad (S^{DB} \bigcap A_{12}) \subseteq (S^{MAJ} \bigcap A_{12})
\end{aligned} \tag{1}$$

Thus, when $(S^{DB} \bigcap A_{11}) \nsubseteq (S^{MAJ} \bigcap A_{11})$, the new synthetic sample may root in the inner majority area that causes noises; when $(S^{DB} \bigcap A_{12}) \nsubseteq (S^{MAJ} \bigcap A_{12})$, the new synthetic sample may root in the boundary majority area that causes overlapping. Obviously, these two cases are the poorly performing scenarios of DBO. On the contrary, it makes the well performing scenarios of DBO when $(S^{DB} \bigcap A_{11}) \nsubseteq (S^{MAJ} \bigcap A_{11})$ and $(S^{DB} \bigcap A_{12}) \nsubseteq (S^{MAJ} \bigcap A_{12})$. Especially, when $(S^{MAJ} \bigcap A_2) = \emptyset$, DBO would like to generate more synthetic samples in the boundary minority area.

Generally speaking, when the ellipsoid structure of S^{DB} and S^{MAJ} match the borderline distribution ((for example, being convex or elliptical as our ellipsoid structure), the new synthetic sample may not root in the inner or boundary majority area that makes a well performing scenario of DBO.

II. DBO ALGORITHM

Algorithm S1 DBO

Input: Training set: $T = \{T_{maj}, T_{min}\}$; number of samples in majority and minority class: n_{maj} and n_{min}.
Output: Synthetic samples S_{new}

```
Construct the boundary minority set B_1;
for i=1 to n_min do
    Pick up the i-th boundary minority sample w_i from B_1;
    Pick up corresponding boundary majority samples B2_i;
    Estimate the local boundary area S_i^DB;
    Estimate the local boundary majority area S_i^Maj;
    Obtain the local boundary minority area S_i^MIN = S_i^DB − S_i^DB ∩ S_i^Maj;
end for
Compute the number of new synthetic samples: N = n_maj − n_min
j=0;
while j < N do
    Randomly select one local boundary minority area S_i^MIN;
    Randomly generate one temporary data t in S_i^DB;
    Judge whether t ∈ S_i^Maj;
    If is, randomly re-generate again (repetition maximum: 100).
    If not, add the temporary synthetic sample to S_new, update j=j+1;
end while
return S_new
```

III. Basic properties of real-world datasets

TABLE S1: Basic properties of real-world datasets

dataset	dimension	minority, majority class	number	ratio
Cancer wpbc ret	33	-	47:151	3.2
Diabetes absent	8	-	268:500	1.9
Housing MEDV>35	13	-	48:458	9.5
Iris versicolor	4	-	50:100	2.0
Iris virginica	4	-	50:100	2.0
Spectf 0	44	-	95:254	2.7
Thyriod hyperfunction	21	-	191:3581	18.7
Vowel 4	10	-	48:480	10.0
Vowel 5	10	-	48:480	10.0
BreastTissue1	8	1;rest	21:85	4.0
BreastTissue3	8	3;rest	18:88	4.9
Ecoli2	7	im;rest	77:259	3.4
Glass2	9	2;rest	76:138	1.8
Glass3	9	3;rest	17:197	11.6
ImageSegmentation7	19	7;rest	330:1980	6.0
LibrasMovement2	90	2;rest	24:336	14.0
LibrasMovement14	90	14;rest	24:336	14.0
Pageblocks2	10	2;rest	329:5144	15.6
Pageblocks5	10	5;rest	115:5358	46.6
StatlogVehicleSilhouettes2	18	2;rest	217:629	2.9
WallFollowingRobotNavigation4	24	4;rest	328:5128	15.6
Yeast6	8	6;rest	44:1440	32.7
DMEAntiVirus	531	-	72:302	4.2
GLRCWL1	698	hyperplasic;rest	21:55	2.6
GLRCNBI2	698	serrated;rest	15:61	4.1
ParkinsonsDC	754	-	192:564	2.9
Colon 1	1908	-	22:40	1.8
Leukemia 1	3571	-	25:47	1.9
DrivFace1	6399	1;rest	27:579	21.4
ARBT8	8265	BookOfWisdom;rest	19:571	30.1

IV. Classification performance on neural network (NN)

TABLE S2: NN: average precision

Cancer wpbc ret	**0.4389±0.2333(1)**	0.3911±0.0626(3)	0.3908±0.0696(4)	0.4005±0.0617(2)	0.3783±0.0677(5)	0.3421±0.0729(8)	0.3651±0.0753(6)	0.3441±0.0848(7)
Diabetes absent	**0.6767±0.0508(1)**	0.5966±0.0401(4)	0.5929±0.0371(7)	0.6047±0.0327(3)	0.6201±0.0381(2)	0.5959±0.0364(5)	0.5931±0.0422(6)	0.5828±0.0284(8)
Housing MEDV>35	**0.6898±0.2075(1)**	0.5383±0.0983(5)	0.5498±0.0788(4)	0.5800±0.0987(2)	0.5696±0.0955(3)	0.3787±0.0458(8)	0.4661±0.0644(6)	0.4236±0.0568(7)
Iris versicolor	0.9047±0.1379(5)	0.9078±0.1353(4)	0.9265±0.0554(2)	0.9233±0.0570(3)	**0.9271±0.0586(1)**	0.8763±0.1045(8)	0.8890±0.0660(6)	0.8826±0.0718(7)
Iris virginica	**0.9259±0.0566(1)**	0.8973±0.0805(5)	0.9162±0.0542(3)	0.9217±0.0567(2)	0.9114±0.0597(4)	0.7737±0.1021(8)	0.8893±0.0762(7)	0.8911±0.0558(6)
Spectf 0	0.5470±0.1766(3)	0.5326±0.0511(7)	0.5616±0.0464(2)	**0.5618±0.0584(1)**	0.5421±0.0586(5)	0.4696±0.0497(8)	0.5436±0.0497(4)	0.5359±0.0499(6)
Thyriod hyperfunction	0.1363±0.2978(7)	0.1507±0.0590(5)	0.1614±0.1029(4)	0.1884±0.1001(2)	**0.3622±0.1371(1)**	0.1294±0.0551(8)	0.1705±0.0668(3)	0.1433±0.0742 (6)
Vowel 4	0.6463±0.2426(3)	0.5455±0.1305(5)	0.6571±0.1377(2)	**0.6592±0.1663(1)**	0.5710±0.1420(4)	0.3108±0.0612(8)	0.5032±0.1063(6)	0.4614±0.0854(7)
Vowel 5	0.3460±0.3122(7)	0.4989±0.1444(4)	**0.6126±0.1452(1)**	0.5859±0.1403(2)	0.5212±0.1433(3)	0.3047±0.0737(8)	0.4078±0.1006(5)	0.3808±0.0958(6)
BreastTissue1	0.6654±0.3182(4)	0.6052±0.1554(7)	0.6782±0.1214(3)	0.6812±0.1457(2)	**0.7004±0.1097(1)**	0.6588±0.1477(5)	0.6260±0.1439(6)	0.6035±0.1213(8)
BreastTissue3	0.0494±0.1061(8)	0.2469±0.0739(4)	0.2465±0.0684(5)	0.2561±0.0537(2)	0.2431±0.0657(6)	0.2274±0.0817(7)	**0.2587±0.0559(1)**	0.2528±0.0619(3)
Ecoli2	**0.7649±0.0782(1)**	0.6445±0.0595(5)	0.6611±0.0547(4)	0.6976±0.0610(2)	0.6821±0.0598(3)	0.6133±0.0628(8)	0.6342±0.0490(6)	0.6219±0.0444(7)
Glass2	**0.5482±0.1881(1)**	0.4901±0.0942(4)	0.4896±0.0722(5)	0.4881±0.1152(6)	0.5183±0.0940(2)	0.4777±0.0866(7)	0.4749±0.0929(8)	0.4998±0.0984(3)
Glass3	0.0757±0.2003(8)	0.1471±0.0877(7)	0.1974±0.1131(2)	**0.2000±0.1219(1)**	0.1832±0.0845(3)	0.1527±0.0671(6)	0.1603±0.0752(5)	0.1622±0.0659(4)
ImageSegmentation7	0.9987±0.0038(2)	0.9941±0.0085(5)	0.9967±0.0074(4)	**0.9988±0.0031(1)**	0.9974±0.0045(3)	0.9935±0.0084(6)	0.9128±0.0819(8)	0.9851±0.0121(7)
LibrasMovement2	0.4279±0.3469(7)	0.5726±0.1574(5)	0.5968±0.1338(2)	**0.6190±0.1324(1)**	0.5837±0.1286(4)	0.3729±0.0760(8)	0.5872±0.1427(3)	0.5674±0.1213(6)
LibrasMovement14	**0.7497±0.2368(1)**	0.6540±0.1962(5)	0.6345±0.1379(7)	0.6979±0.1725(2)	0.6666±0.1528(3)	0.3466±0.0736(8)	0.6645±0.1982(4)	0.6420±0.1448(6)
Pageblocks2	**0.8833±0.0899(1)**	0.5354±0.0478(4)	0.5028±0.0586(5)	0.6444±0.0879 (2)	0.6009±0.0639(3)	0.2708±0.0713(8)	0.4217±0.0585(7)	0.4542±0.0620(6)
Pageblocks5	**0.5930±0.1529(1)**	0.2578±0.0576(3)	0.1972±0.0335(5)	0.2434±0.0359 (4)	0.2923±0.0503(2)	0.0507±0.0162(8)	0.1731±0.0279(7)	0.1907±0.0249(6)
StatlogVehicleSilhouettes2	**0.6067±0.1376(1)**	0.5281±0.0575(5)	0.5306±0.0537(4)	0.5481±0.0406(2)	0.5380±0.0519(3)	0.4815±0.0486(8)	0.5001±0.0476(7)	0.5083±0.0488 (6)
WallFollowingRobotNavigation4	**0.8239±0.0620(1)**	0.7220±0.0484(4)	0.7466±0.0589(3)	0.8002±0.0479(2)	0.6543±0.0603(5)	0.5003±0.0614(8)	0.5238±0.0738(7)	0.6075±0.0614 (6)
Yeast6	**0.5513±0.1668(1)**	0.4020±0.0587(5)	0.4295±0.0667(3)	0.4336±0.0749(2)	0.4246±0.0575(4)	0.2390±0.0247(8)	0.3263±0.0328(7)	0.3270±0.0405(6)
DMEAntiVirus	0.9727±0.0184(3)	0.9725±0.0179(4)	0.9723±0.0179(6)	0.9720±0.0182(7)	**0.9782±0.0189(1)**	0.9494±0.1036(8)	0.9725±0.0179(5)	0.9732±0.0180(2)
GLRCWL1	0.6216±0.2382(6)	0.6535±0.1194(3)	0.6642±0.1297(2)	0.6529±0.1175(4)	0.6493±0.1144(5)	0.4077±0.1353(8)	**0.6933±0.1512(1)**	0.6035±0.1448(7)
GLRCNBI2	0.2170±0.1878(8)	0.2965±0.1257(6)	0.3182±0.1196(4)	0.3270±0.1225(3)	0.3293±0.1268(2)	0.2616±0.1292(7)	**0.3373±0.1339(1)**	0.3181±0.0996(5)
ParkinsonsDC	**0.7249±0.1317(1)**	0.5558±0.0926(6)	0.5919±0.0687(5)	0.5992±0.0704(4)	0.6129±0.0637(2)	0.4459±0.0498(8)	0.6035±0.0649(3)	0.5170±0.0559(7)
Colon 1	0.6552±0.1987(6)	0.6643±0.1532(4)	0.6636±0.1878(5)	0.7024±0.1531(3)	0.7069±0.1525(2)	0.6004±0.1645(8)	**0.7081±0.1740(1)**	0.6420±0.1394(7)
Leukemia 1	0.9248±0.1151(3)	0.9015±0.1148(5)	**0.9475±0.0682(1)**	0.9188±0.1545 (4)	0.6906±0.1045(7)	0.6043±0.2046(8)	0.9449±0.1237(2)	0.7750±0.1364(6)
DrivFace1	**0.9264±0.0857(1)**	0.7700±0.1011(6)	0.8732±0.0910(2)	0.8717±0.0864(3)	0.8659±0.0803(5)	0.6564±0.1268(8)	0.8684±0.0839(4)	0.7104±0.1098(7)
ARBT8	0.0000±0.0000(7.5)	**0.7672±0.2681(1)**	0.6403±0.3931(2)	0.0000±0.0000(7.5)	0.0322±0.0120(6)	0.0698±0.1167(5)	0.1833±0.3678(3)	0.1483±0.0984(4)

A stratified k-fold cross validation (k=2 in experience) is used for 30 times that 60 (2×30) runs are conducted. Thus for each table cell, the mean and standard deviation of corresponding performance on 60 runs are first recorded and then its rank among all methods is followed in one bracket. The best rank for each row is highlighted as bold.

TABLE S3: NN: average recall

Cancer wpbc ret	0.2848±0.2236(8)	0.5739±0.1170(5)	0.5833±0.1396(4)	0.5696±0.1040(6)	0.5645±0.1122(7)	**0.6232±0.1460(1)**	0.5920±0.1276(3)	0.6036±0.1617(2)
Diabetes absent	0.5637±0.0874(8)	0.7454±0.0699(4)	0.7490±0.0646(2)	0.7177±0.0556(6)	0.6973±0.0501(7)	0.7470±0.0516(3)	0.7302±0.0727(5)	**0.7646±0.0490(1)**
Housing MEDV>35	0.5361±0.1932(8)	0.7569±0.0956(6)	0.8021±0.0954(4)	0.7458±0.1062(7)	0.7986±0.0968(5)	**0.9021±0.0641(1)**	0.8472±0.0884(3)	0.8771±0.0753(2)
Iris versicolor	0.9013±0.1383(8)	0.9173±0.1340(7)	0.9320±0.0548(5)	0.9300±0.0529(6)	0.9347±0.0531(4)	0.9460±0.0545(3)	0.9547±0.0468(2)	**0.9667±0.0369(1)**
Iris virginica	0.9180±0.0696(8)	0.9533±0.0557(4)	0.9447±0.0663(6)	0.9387±0.0738(7)	0.9507±0.0680(5)	0.9620±0.0502(2)	0.9573±0.0772(3)	**0.9787±0.0407(1)**
Spectf 0	0.4784±0.2265(8)	0.7968±0.0942(3)	0.7574±0.0913(7)	0.7688±0.0945(5)	0.7649±0.1080(6)	**0.8883±0.0840(1)**	0.7777±0.1077(4)	0.8305±0.0836(2)
Thyriod hyperfunction	0.0030±0.0058(8)	0.5970±0.0712(5)	0.7221±0.0763(3)	0.5714±0.0737(6)	0.7168±0.0714(4)	0.5191±0.2271(7)	**0.7918±0.0778(1)**	0.7567±0.0817 (2)
Vowel 4	0.4840±0.2534(8)	0.8313±0.1194(6)	0.8625±0.0925(5)	0.8104±0.1498(7)	0.8681±0.1093(4)	**0.9542±0.0603(1)**	0.9090±0.0876(3)	0.9375±0.0738(2)
Vowel 5	0.2160±0.2533(8)	0.7472±0.1225(6)	0.7792±0.1024(5)	0.7306±0.1367(7)	0.7833±0.1057(4)	**0.8993±0.1016(1)**	0.8667±0.1029(3)	0.8806±0.0941(2)
BreastTissue1	0.6317±0.3192(8)	0.8650±0.1624(4)	0.8683±0.1127(3)	0.8417±0.1197(7)	0.8550±0.0964(6)	**0.9217±0.0904(1)**	0.8617±0.1462(5)	0.9133±0.0999(2)
BreastTissue3	0.0370±0.0952(8)	0.6056±0.2801(7)	0.6889±0.2626(5)	0.6944±0.2374(4)	0.6704±0.2538(6)	0.7019±0.2996(3)	**0.7278±0.2159(1)**	0.7259±0.2363(2)
Ecoli2	0.7412±0.0948(8)	0.9079±0.0651(4)	0.8877±0.0685(5)	0.8539±0.0785(7)	0.8680±0.0777(6)	0.9162±0.0576(2)	0.9145±0.0651(3)	**0.9303±0.0496(1)**
Glass2	0.4031±0.2244(8)	0.6211±0.1783(7)	0.6702±0.1715(2)	0.6671±0.1815(3)	0.6329±0.1552(6)	0.6390±0.1767(5)	0.6465±0.1875(4)	**0.7202±0.1832(1)**
Glass3	0.0375±0.0839(8)	0.3958±0.2173(7)	0.5729±0.2199(5)	0.4688±0.2028(6)	0.5750±0.2198(4)	0.6604±0.2447(2)	0.6479±0.2272(3)	**0.6708±0.2109(1)**
ImageSegmentation7	0.9911±0.0044(8)	0.9931±0.0052(3)	0.9929±0.0051(4)	0.9914±0.0049 (6.5)	0.9948±0.0050(2)	0.9924±0.0047(5)	0.9914±0.0079(6.5)	**0.9956±0.0050(1)**
LibrasMovement2	0.3500±0.3286(8)	0.7806±0.1910(7)	0.8194±0.1618(5)	0.7917±0.1553(6)	0.8208±0.1615(4)	**0.8958±0.1124(1)**	0.8292±0.1451(3)	0.8653±0.1228(2)
LibrasMovement14	0.6708±0.2546(8)	0.8542±0.1325(4)	0.8500±0.1254(5)	0.8417±0.1299(6)	0.8389±0.1301(7)	**0.9278±0.0901(1)**	0.8611±0.1317(3)	0.9139±0.1063(2)
Pageblocks2	0.6457±0.1150(8)	0.9552±0.0174(3)	0.9578±0.0160(2)	0.9137±0.0239(6)	0.9166±0.0316(5)	0.8308±0.1207(7)	0.9539±0.0268(4)	**0.9638±0.0181(1)**
Pageblocks5	0.3032±0.1488(8)	0.8532±0.0446(5)	0.8886±0.0394(3)	0.8357±0.0556(6)	0.8547±0.0368(4)	0.6871±0.2495(7)	0.8980±0.0325(2)	**0.9094±0.0287(1)**
StatlogVehicleSilhouettes2	0.3974±0.1599(8)	0.7637±0.0645(5)	0.7773±0.0655(3)	0.7579±0.0613(6)	0.7455±0.0773(7)	0.7736±0.0720(4)	0.7935±0.0633(2)	**0.8182±0.0673 (1)**
WallFollowingRobotNavigation4	0.7761±0.1134(8)	0.9002±0.0261(5)	0.8919±0.0300(6)	0.8845±0.0321(7)	0.9254±0.0240(4)	0.9481±0.0253(2)	**0.9533±0.0302(1)**	0.9424±0.0195(3)
Yeast6	0.4371±0.1920(8)	0.9295±0.0851(4)	0.9136±0.0779(6)	0.8985±0.0933(7)	0.9280±0.0628(5)	**0.9932±0.0202(1)**	0.9879±0.0333(2)	0.9841±0.0343(3)
DMEAntiVirus	0.9481±0.0518(6)	0.9704±0.0270(2)	0.9611±0.0351(3)	0.9519±0.0392(5)	0.8986±0.0592(8)	0.9435±0.0599(7)	**0.9708±0.0300(1)**	0.9560±0.0754(4)
GLRCWL1	0.5750±0.2891(8)	0.7417±0.1934(2)	0.7267±0.1930(3)	0.6967±0.1850(6)	0.6817±0.1900(7)	0.7033±0.2277(5)	0.7217±0.2009(4)	**0.8267±0.1494(1)**
GLRCNBI2	0.2286±0.2188(8)	0.4714±0.2472(3)	0.4571±0.2068(4)	0.4524±0.2187(5)	0.4143±0.1998(6)	0.5238±0.2877(2)	0.3881±0.2056(7)	**0.5571±0.1854(1)**
ParkinsonsDC	0.4915±0.1655(8)	0.7245±0.1075(3)	0.7229±0.0633(4)	0.7113±0.0539(6)	0.7122±0.0528(5)	**0.7854±0.0949(1)**	0.7016±0.0649(7)	0.7722±0.0532(2)
Colon 1	0.5955±0.2399(8)	0.6894±0.2126(6)	0.7273±0.2252(3)	0.7242±0.2090(4)	0.7136±0.1929(5)	0.7712±0.2107(1.5)	0.6742±0.2456(7)	0.7712±0.1743(1.5)
Leukemia 1	0.8736±0.1663(6)	**0.9278±0.1132(1)**	0.9208±0.0951(2)	0.9181±0.1484 (4)	0.7472±0.1549(8)	0.8417±0.2463(7)	0.9194±0.0867(3)	0.8931±0.1667(5)
DrivFace1	0.7218±0.1664(8)	0.8462±0.1115(3)	0.8167±0.1322(6)	0.8244±0.1227(4)	0.8167±0.1299(5)	**0.8962±0.0938(1)**	0.7910±0.1252(7)	0.8962±0.0948(2)
ARBT8	0.0000±0.0000(7.5)	0.3204±0.1952(2)	0.1759±0.1691(5)	0.0000±0.0000(7.5)	**0.4926±0.1901(1)**	0.3185±0.3314(3)	0.0315±0.0681(6)	0.2963±0.1513(4)

A stratified k-fold cross validation (k=2 in experience) is used for 30 times that 60 (2×30) runs are conducted. Thus for each table cell, the mean and standard deviation of corresponding performance on 60 runs are first recorded and then its rank among all methods is followed in one bracket. The best rank for each row is highlighted as bold.

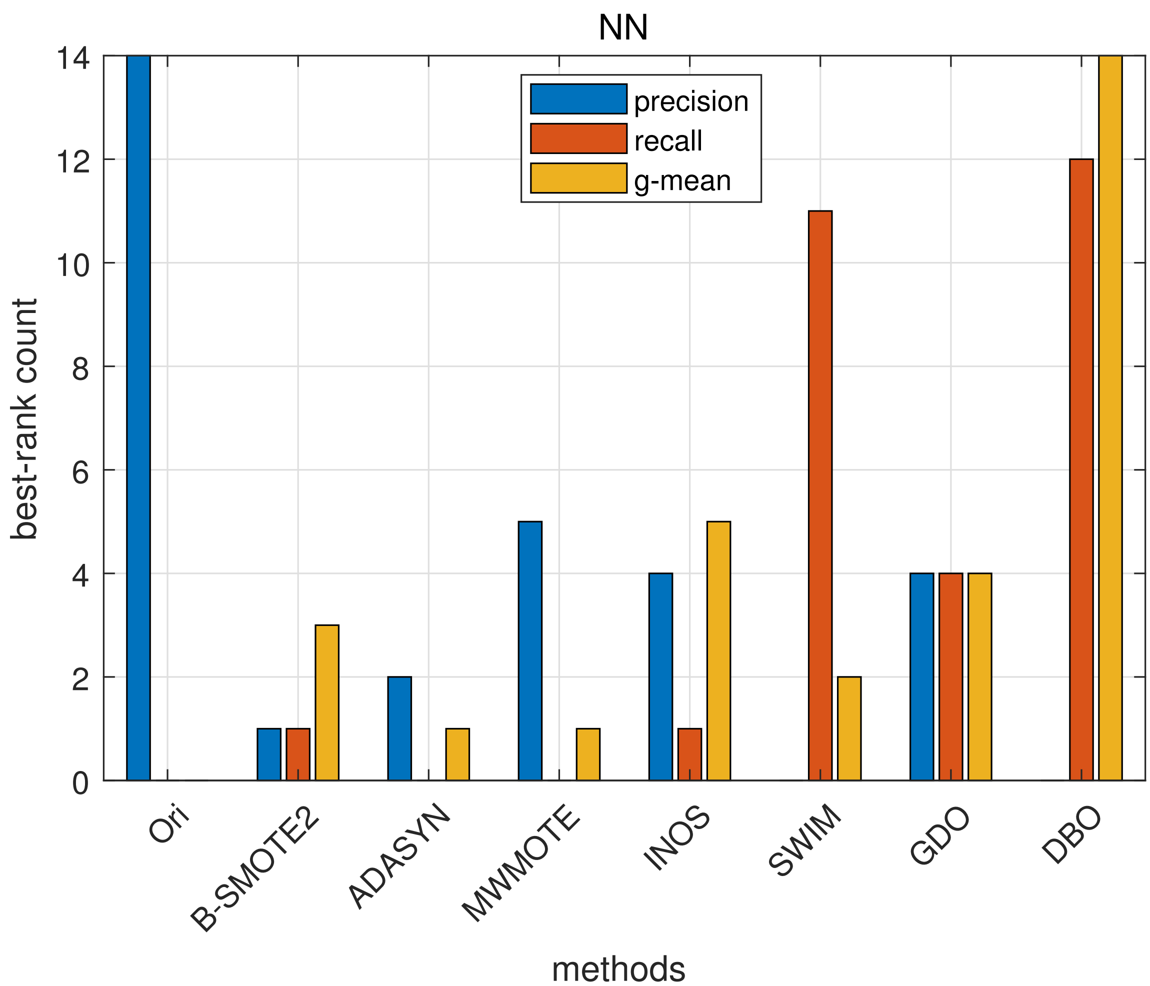

Fig. S1: Counts of best performance on NN. We count the time for each method that achieves the best rank over all datasets.

TABLE S4: NN: mean ranks of recall, precision and g-mean

Measurement	Actual value(Friedman test)	Ori	B-SMOTE2	ADASYN	MWMOTE	INOS	SWIM	GDO	DBO
precision	87.89(reject)	3.38	4.67	3.60	**2.75**	3.33	7.47	4.83	5.97
recall	119.01(reject)	7.85†	4.43†	4.17†	5.87†	5.23†	2.95	3.62†	**1.88**
g-mean	92.18(reject)	7.72†	4.47†	3.53	4.68†	3.83	5.60†	3.93	**2.23**
the Friedman Test: F=14.07, (n=8-1,alpha=0.05)									
the Bonferroni-Dunn test: critical values=1.70, (k=8,alpha=0.05)									

The best mean rank is highlighted as bold.
Using the Friedman Test with F=14.07 under (n=8-1,alpha=0.05) as one of the statistic test. If the actual value is larger than 14.07, reject the original hypothesis that significant difference exists among those methods.
Using the Bonferroni-Dunn test with critical values=1.70 under (k=8, alpha=0.05) as one of the statistic test. If the gap of mean rank between two methods is larger than 1.70, reject the original hypothesis that significant difference exists between two methods and marked with †. There, we only consider the measurement in which the proposed method is of the best mean rank , and compare it with any one of other methods.

TABLE S5: NN: Wilcoxon paired signed-rank test for pairwise comparisons

recall		g-mean	
Ours vs.	p-V	Ours vs.	p-V
SWIM	0.1358	ADASYN	**0.0047**
		INOS	**0.0034**
		GDO	**0.0057**

p-value under 0.05 is highlighted as bold which means the proposed one outperforms the compared one when using the Wilcoxon rank test (alpha=0.05)).

V. Classification performance on support vector machine (SVM)

TABLE S6: SVM: average precision

Cancer wpbc ret	**0.5727±0.1122(1)**	0.4550±0.0829(5)	0.4707±0.0824(4)	0.4772±0.0763(2)	0.4734±0.0754(3)	0.4431±0.0639(7)	0.4500±0.0934(6)	0.4195±0.0679(8)
Diabetes absent	**0.7865±0.0407(1)**	0.6859±0.0476(7)	0.6934±0.0431(5)	0.6917±0.0404(6)	0.7096±0.0484(2)	0.7014±0.0437(3)	0.6939±0.0442(4)	0.6837±0.0413(8)
Housing MEDV>35	**0.8694±0.1137(1)**	0.5848±0.0995(4)	0.5767±0.0699(5)	0.6439±0.1015(2)	0.6238±0.0810(3)	0.4805±0.0523(7)	0.5289±0.0659(6)	0.4561±0.0549(8)
Iris versicolor	**0.6797±0.1675(1)**	0.5796±0.0797(6)	0.5937±0.0705(3)	0.5898±0.0691(5)	0.6056±0.0750(2)	0.5913±0.0772(4)	0.5703±0.0812(7)	0.5471±0.0787(8)
Iris virginica	**0.9866±0.0247(1)**	0.9500±0.0420(5)	0.9513±0.0418(4)	0.9619±0.0363(3)	0.9694±0.0340(2)	0.9058±0.0673(8)	0.9396±0.0447(6)	0.9224±0.0400(7)
Spectf 0	**0.6757±0.0683(1)**	0.6210±0.0601(5)	0.6363±0.0575(3)	0.6413±0.0667(2)	0.6283±0.0565(4)	0.5322±0.0486(8)	0.6151±0.0594(6)	0.5856±0.0511(7)
Thyriod hyperfunction	0.0000±0.0000(8)	0.4238±0.0981(6)	**0.5774±0.0269(1)**	0.5172±0.0473(2)	0.4802±0.0440(4)	0.2229±0.0653(7)	0.4325±0.0718(5)	0.5106±0.0662(3)
Vowel 4	**0.4499±0.4368(1)**	0.3759±0.0592(3)	0.3750±0.0388(4)	0.3772±0.0521(2)	0.3636±0.0513(5)	0.2975±0.0346(8)	0.3455±0.0390(6)	0.3165±0.0385(7)
Vowel 5	0.0000±0.0000(8)	**0.2250±0.0385(1)**	0.2088±0.0229(7)	0.2089±0.0322(6)	0.2169±0.0336(3)	0.2114±0.0282(5)	0.2172±0.0335(2)	0.2142±0.0320(4)
BreastTissue1	**0.8662±0.1228(1)**	0.7077±0.1339(7)	0.7457±0.1281(5)	0.7927±0.1179(3)	0.8094±0.1104(2)	0.7495±0.1182(4)	0.7263±0.1389(6)	0.6914±0.1308(8)
BreastTissue3	0.0000±0.0000(8)	0.2469±0.0662(6)	0.2649±0.0539(3)	0.2583±0.0598(5)	0.2416±0.0719(7)	**0.2769±0.0514(1)**	0.2634±0.0513(4)	0.2733±0.0494(2)
Ecoli2	**0.8726±0.0510(1)**	0.6707±0.0413(5)	0.6793±0.0461(4)	0.7080±0.0644(3)	0.7170±0.0636(2)	0.6603±0.0390(8)	0.6641±0.0403(7)	0.6650±0.0412(6)
Glass2	0.3351±0.2657(8)	0.4453±0.0642(7)	0.4652±0.0633(4)	**0.5068±0.0621(1)**	0.4721±0.0752(3)	0.4583±0.0839(6)	0.4606±0.0769(5)	0.4903±0.0559(2)
Glass3	0.0000±0.0000(8)	0.2475±0.1324(5)	**0.3250±0.1010(1)**	0.3135±0.1331(2)	0.3076±0.1118(3)	0.2454±0.0598(6)	0.2436±0.0997(7)	0.2675±0.0811(4)
ImageSegmentation7	1.0000±0.0000(1.5)	0.9972±0.0069(6)	0.9974±0.0073(5)	1.0000±0.0000 (1.5)	0.9991±0.0031(3)	0.9988±0.0031(4)	0.9313±0.0482(8)	0.9859±0.0117(7)
LibrasMovement2	0.3386±0.3266(3)	**0.3502±0.1314(1)**	0.2950±0.1031(6)	0.3387±0.1291(2)	0.3044±0.1214(5)	0.2659±0.0761(7)	0.3091±0.1007(4)	0.2421±0.0676(8)
LibrasMovement14	**0.7764±0.1361(1)**	0.5376±0.1290(6)	0.5640±0.1793(5)	0.6626±0.1261(2)	0.6365±0.1621(3)	0.3614±0.0572(8)	0.5727±0.1769(4)	0.3843±0.0785(7)
Pageblocks2	**0.9377±0.0176(1)**	0.6641±0.0439(4)	0.6398±0.0346(5)	0.7568±0.0236 (2)	0.7319±0.0251(3)	0.5651±0.0732(8)	0.6252±0.0576(6)	0.5981±0.0367(7)
Pageblocks5	**0.8132±0.0991(1)**	0.3718±0.0614(4)	0.3104±0.0366(6)	0.4115±0.0425 (3)	0.4302±0.0452(2)	0.2117±0.0758(8)	0.3038±0.0315(7)	0.3107±0.0316(5)
StatlogVehicleSilhouettes2	**0.7840±0.0766(1)**	0.5746±0.0295(6)	0.5765±0.0277(4)	0.5810±0.0264(3)	0.5847±0.0318(2)	0.5722±0.0260(7)	0.5762±0.0318(5)	0.5569±0.0272 (8)
WallFollowingRobotNavigation4	**0.7481±0.0541(1)**	0.3638±0.0144(4)	0.3617±0.0233(5)	0.4130±0.0146(2)	0.3891±0.0168(3)	0.3135±0.0119(8)	0.3522±0.0257(6)	0.3520±0.0156 (7)
Yeast6	**0.7004±0.1704(1)**	0.4197±0.0370(5)	0.4482±0.0376(4)	0.4491±0.0411(3)	0.4543±0.0408(2)	0.2893±0.0226(8)	0.3962±0.0345(6)	0.3760±0.0338(7)
DMEAntiVirus	0.9711±0.0185(6)	**0.9723±0.0181(1)**	0.9711±0.0185(6)	0.9711±0.0185(6)	0.9711±0.0185(8)	0.9712±0.0185(4)	0.9712±0.0184(3)	0.9718±0.0182(2)
GLRCWL1	0.7039±0.1409(4)	0.6883±0.1125(6)	0.7039±0.1409(4)	0.7039±0.1409(4)	**0.7096±0.1352(1)**	0.6488±0.1537(7)	0.7090±0.1407(2)	0.6348±0.0964(8)
GLRCNBI2	0.2978±0.2378(6)	0.3154±0.1658(2)	0.2978±0.2378(6)	0.2978±0.2378(6)	0.2986±0.2379(4)	0.3099±0.1240(3)	0.2867±0.2075(8)	**0.3214±0.1016(1)**
ParkinsonsDC	0.7242±0.0536(2)	0.6369±0.0495(6)	0.7242±0.0536(2)	0.7242±0.0536(2)	0.7122±0.0508(5)	0.4357±0.0399(8)	0.7218±0.0542(4)	0.5754±0.0346(7)
Colon 1	0.8050±0.2154(4)	**0.8160±0.1222(1)**	0.8050±0.2154(4)	0.8050±0.2154(4)	0.8055±0.2152(2)	0.8030±0.1489(7)	0.8046±0.2153(6)	0.7811±0.0834(8)
Leukemia 1	1.0000±0.0000(3)	0.9918±0.0284(6)	1.0000±0.0000(3)	1.0000±0.0000(3)	1.0000±0.0000(3)	0.8249±0.1675(8)	1.0000±0.0000(3)	0.9622±0.0534(7)
DrivFace1	0.9845±0.0407(2.5)	0.8627±0.0874(6)	0.9845±0.0407(2.5)	0.9845±0.0407(2.5)	0.9758±0.0531(5)	0.8389±0.0770(8)	0.9845±0.0407(2.5)	0.8535±0.0744(7)
ARBT8	0.0000±0.0000(5)	0.0000±0.0000(5)	0.0000±0.0000(5)	0.0000±0.0000(5)	0.0000±0.0000(5)	**0.1687±0.2890(1)**	0.0000±0.0000(5)	0.0000±0.0000(5)

A stratified k-fold cross validation (k=2 in experience) is used for 30 times that 60 (2×30) runs are conducted. Thus for each table cell, the mean and standard deviation of corresponding performance on 60 runs are first recorded and then its rank among all methods is followed in one bracket. The best rank for each row is highlighted as bold.

TABLE S7: SVM: average recall

Cancer wpbc ret	0.3080±0.1069(8)	0.4913±0.0955(4)	0.5007±0.1005(3)	0.4732±0.0927(6)	0.4877±0.0859(5)	0.5493±0.0993(2)	0.4435±0.1100(7)	**0.5725±0.0948(1)**
Diabetes absent	0.4230±0.0517(8)	0.5935±0.0432(2)	0.5833±0.0441(4)	0.5843±0.0435(3)	0.5517±0.0488(7)	0.5697±0.0450(6)	0.5739±0.0463(5)	**0.6035±0.0488(1)**
Housing MEDV>35	0.4799±0.0984(8)	0.7049±0.1013(6)	0.7535±0.1075(4)	0.6910±0.1141(7)	0.7465±0.1007(5)	0.8382±0.0656(2)	0.7910±0.0932(3)	**0.8451±0.0783(1)**
Iris versicolor	0.2507±0.1173(8)	**0.6847±0.1543(1)**	0.6153±0.1741(5)	0.6413±0.1518(3)	0.6113±0.1569(6)	0.6033±0.1691(7)	0.6327±0.1864(4)	0.6647±0.1976(2)
Iris virginica	0.8267±0.1017(8)	0.9440±0.0684(2)	0.9413±0.0723(3)	0.9240±0.0789(5)	0.9073±0.0828(7)	0.9173±0.0560(6)	0.9407±0.0623(4)	**0.9700±0.0408(1)**
Spectf 0	0.5376±0.0775(8)	0.6777±0.0842(3)	0.6738±0.0835(4)	0.6613±0.0857(5)	0.6543±0.0867(6)	0.7280±0.0763(2)	0.6500±0.0818(7)	**0.7348±0.0889(1)**
Thyriod hyperfunction	0.0000±0.0000(8)	0.6491±0.0854(3)	0.7298±0.0570(2)	0.5896±0.0678(5)	0.4582±0.0804(6)	0.3789±0.0740(7)	0.6137±0.0532(4)	**0.8177±0.0670 (1)**
Vowel 4	0.0625±0.0671(8)	0.7424±0.1137(5)	0.7868±0.1091(2)	0.7243±0.1228(7)	0.7264±0.1202(6)	0.7792±0.0977(4)	0.7868±0.1047(3)	**0.8111±0.1118(1)**
Vowel 5	0.0000±0.0000(8)	0.5674±0.1329(5)	**0.6139±0.1203(1)**	0.5729±0.1339(4)	0.5368±0.1181(7)	0.5757±0.1200(3)	0.5493±0.1224(6)	0.5958±0.1277(2)
BreastTissue1	0.6283±0.1303(8)	0.8417±0.0996(2)	0.8250±0.1002(5)	0.8150±0.0936(6)	0.7917±0.1046(7)	0.8400±0.0887(3)	0.8383±0.0922(4)	**0.8467±0.0812(1)**
BreastTissue3	0.0000±0.0000(8)	0.5630±0.2230(7)	0.6778±0.1970(3)	0.6167±0.2110(5)	0.5907±0.2355(6)	**0.7611±0.1805(1)**	0.6519±0.1869(4)	0.7111±0.1574(2)
Ecoli2	0.6491±0.0803(8)	0.8855±0.0876(3)	0.8421±0.0893(5)	0.7912±0.0992(6)	0.7890±0.0847(7)	**0.8952±0.0721(1)**	0.8697±0.0770(4)	0.8868±0.0770(2)
Glass2	0.0833±0.0772(8)	0.4272±0.1719(6)	0.4947±0.1547(3)	0.5513±0.1167(2)	0.4110±0.1544(7)	0.4614±0.1840(4)	0.4417±0.1815(5)	**0.5974±0.1378(1)**
Glass3	0.0000±0.0000(8)	0.4479±0.2125(7)	0.5917±0.1576(3)	0.4604±0.2052(6)	0.5292±0.1728(4)	**0.7125±0.1860(1)**	0.4771±0.2103(5)	0.6750±0.1969(2)
ImageSegmentation7	0.9895±0.0043(6.5)	0.9917±0.0055(5)	0.9924±0.0051(4)	0.9895±0.0043 (6.5)	0.9931±0.0047(2)	0.9926±0.0050(3)	0.9698±0.0266(8)	**0.9953±0.0057(1)**
LibrasMovement2	0.1153±0.1275(8)	0.4861±0.1536(3)	0.4431±0.1410(5)	0.4403±0.1736(6)	0.4236±0.1474(7)	0.5167±0.1318(2)	0.4431±0.1553(4)	**0.5597±0.1356(1)**
LibrasMovement14	0.6403±0.1732(8)	0.7903±0.1385(5)	0.8056±0.1326(3)	0.7806±0.1423(6)	0.7556±0.1282(7)	**0.9083±0.0929(1)**	0.7986±0.1340(4)	0.9028±0.1150(2)
Pageblocks2	0.6657±0.0364(8)	0.9484±0.0224(3)	0.9561±0.0184(2)	0.9369±0.0205(4)	0.9307±0.0212(6)	0.8824±0.0244(7)	0.9338±0.0207(5)	**0.9689±0.0163(1)**
Pageblocks5	0.1810±0.0762(8)	0.8234±0.0444(5)	0.8471±0.0470(3)	0.8237±0.0469(4)	0.8088±0.0468(6)	0.8073±0.1083(7)	0.8632±0.0391(2)	**0.8836±0.0301(1)**
StatlogVehicleSilhouettes2	0.2852±0.0673(8)	0.7139±0.0538(4)	0.7227±0.0538(2)	0.6949±0.0605(5)	0.6752±0.0495(6)	0.7173±0.0493(3)	0.6685±0.0514(7)	**0.7660±0.0466 (1)**
WallFollowingRobotNavigation4	0.2261±0.0482(8)	0.9029±0.0283(3)	0.8229±0.0530(6)	0.8567±0.0293(5)	0.8735±0.0262(4)	**0.9510±0.0153(1)**	0.7834±0.0338(7)	0.9221±0.0232(2)
Yeast6	0.2455±0.1072(8)	0.8947±0.1096(4)	0.8621±0.1021(6)	0.8455±0.1151(7)	0.8667±0.1048(5)	**1.0000±0.0000(1)**	0.9788±0.0447(3)	0.9864±0.0413(2)
DMEAntiVirus	0.9130±0.0564(6)	**0.9620±0.0343(1)**	0.9130±0.0564(6)	0.9130±0.0564(6)	0.9125±0.0553(8)	0.9176±0.0565(3)	0.9167±0.0551(4)	0.9468±0.0373(2)
GLRCWL1	0.4467±0.1702(6)	0.5967±0.1484(2)	0.4467±0.1702(6)	0.4467±0.1702(6)	0.4400±0.1659(8)	0.5283±0.1823(3)	0.4483±0.1682(4)	**0.7467±0.1308(1)**
GLRCNBI2	0.1667±0.1537(7)	0.3381±0.1950(3)	0.1667±0.1537(7)	0.1667±0.1537(7)	0.1738±0.1623(4)	0.3405±0.1880(2)	0.1690±0.1522(5)	**0.4286±0.1803(1)**
ParkinsonsDC	0.5014±0.0651(6)	0.5993±0.0555(3)	0.5014±0.0651(6)	0.5014±0.0651(6)	0.4990±0.0627(8)	**0.7234±0.0587(1)**	0.5142±0.0631(4)	0.6472±0.0592(2)
Colon 1	0.4152±0.2204(6)	0.5848±0.1790(2)	0.4152±0.2204(6)	0.4152±0.2204(6)	0.4167±0.2222(4)	0.5061±0.2133(3)	0.4121±0.2192(8)	**0.7409±0.1664(1)**
Leukemia 1	0.5903±0.1322(6)	0.7847±0.1366(3)	0.5903±0.1322(6)	0.5903±0.1322(6)	0.5861±0.1337(8)	**0.9403±0.0934(1)**	0.5917±0.1335(4)	0.9403±0.0856(2)
DrivFace1	0.5615±0.1269(6)	0.7526±0.1329(3)	0.5615±0.1269(6)	0.5615±0.1269(6)	0.5269±0.1350(8)	**0.8090±0.1006(1)**	0.5667±0.1266(4)	0.7936±0.1102(2)
ARBT8	0.0000±0.0000(5)	0.0000±0.0000(5)	0.0000±0.0000(5)	0.0000±0.0000(5)	0.0000±0.0000(5)	**0.4296±0.3953(1)**	0.0000±0.0000(5)	0.0000±0.0000(5)

A stratified k-fold cross validation (k=2 in experience) is used for 30 times that 60 (2×30) runs are conducted. Thus for each table cell, the mean and standard deviation of corresponding performance on 60 runs are first recorded and then its rank among all methods is followed in one bracket. The best rank for each row is highlighted as bold.

TABLE S8: SVM: average g-mean

Cancer wpbc ret	0.5243±0.0951(8)	0.6279±0.0589(5)	0.6370±0.0615(3)	0.6252±0.0586(6)	0.6322±0.0528(4)	**0.6524±0.0570(1)**	0.6000±0.0779(7)	0.6514±0.0517(2)
Diabetes absent	0.6282±0.0352(8)	0.7099±0.0210(2)	0.7074±0.0237(4)	0.7076±0.0238(3)	0.6946±0.0289(7)	0.7024±0.0255(6)	0.7025±0.0253(5)	**0.7145±0.0248(1)**
Housing MEDV>35	0.6861±0.0721(8)	0.8130±0.0560(6)	0.8391±0.0582(5)	0.8102±0.0650(7)	0.8404±0.0557(4)	**0.8691±0.0328(1)**	0.8529±0.0480(3)	0.8672±0.0394(2)
Iris versicolor	0.4638±0.1156(8)	**0.6928±0.0542(1)**	0.6761±0.0775(4)	0.6880±0.0571(2)	0.6824±0.0683(3)	0.6680±0.0638(5)	0.6656±0.0745(6)	0.6568±0.0549(7)
Iris virginica	0.9044±0.0556(8)	0.9578±0.0336(2)	0.9567±0.0353(3)	0.9506±0.0383(5)	0.9440±0.0425(6)	0.9325±0.0324(7)	0.9537±0.0320(4)	**0.9636±0.0203(1)**
Spectf 0	0.6940±0.0474(8)	0.7547±0.0471(3)	0.7572±0.0444(2)	0.7526±0.0500(4)	0.7454±0.0463(5)	0.7418±0.0382(6)	0.7401±0.0458(7)	**0.7665±0.0417(1)**
Thyriod hyperfunction	0.0000±0.0000(8)	0.7837±0.0576(3)	0.8415±0.0327(2)	0.7553±0.0439(5)	0.6655±0.0586(6)	0.5896±0.0602(7)	0.7650±0.0350(4)	**0.8841±0.0394 (1)**
Vowel 4	0.1825±0.1714(8)	0.8016±0.0555(4)	**0.8237±0.0542(1)**	0.7930±0.0602(6)	0.7904±0.0576(7)	0.7938±0.0450(5)	0.8147±0.0500(2)	0.8140±0.0521(3)
Vowel 5	0.0000±0.0000(8)	0.6668±0.0657(3)	**0.6814±0.0551(1)**	0.6613±0.0667(5)	0.6515±0.0611(7)	0.6664±0.0614(4)	0.6557±0.0589(6)	0.6749±0.0611(2)
BreastTissue1	0.7774±0.0837(8)	0.8704±0.0486(6)	0.8707±0.0524(4)	0.8748±0.0508(2)	0.8652±0.0587(7)	**0.8800±0.0474(1)**	0.8731±0.0485(3)	0.8707±0.0448(5)
BreastTissue3	0.0000±0.0000(8)	0.5854±0.1115(7)	0.6380±0.0903(3)	0.6130±0.1014(5)	0.5894±0.1354(6)	**0.6628±0.0822(1)**	0.6297±0.0848(4)	0.6545±0.0753(2)
Ecoli2	0.7922±0.0476(8)	0.8755±0.0346(2)	0.8589±0.0372(5)	0.8408±0.0426(7)	0.8426±0.0366(6)	**0.8775±0.0306(1)**	0.8677±0.0319(4)	0.8749±0.0308(3)
Glass2	0.2238±0.1603(8)	0.5248±0.0701(7)	0.5671±0.0670(3)	**0.6130±0.0517(1)**	0.5348±0.0737(6)	0.5392±0.0697(5)	0.5403±0.0873(4)	0.6121±0.0583(2)
Glass3	0.0000±0.0000(8)	0.5925±0.1994(7)	0.7193±0.1007(3)	0.6182±0.1880(6)	0.6734±0.1387(4)	**0.7543±0.0987(1)**	0.6194±0.1785(5)	0.7426±0.1113(2)
ImageSegmentation7	0.9947±0.0022(6.5)	0.9956±0.0026(5)	0.9960±0.0025(4)	0.9947±0.0022 (6.5)	**0.9965±0.0023(1)**	0.9962±0.0025(3)	0.9787±0.0163(8)	0.9964±0.0024(2)
LibrasMovement2	0.2572±0.2182(8)	0.6626±0.1107(3)	0.6277±0.1102(5)	0.6198±0.1616(6)	0.6110±0.1341(7)	0.6719±0.0853(2)	0.6289±0.1166(4)	**0.6900±0.0879(1)**
LibrasMovement14	0.7854±0.1178(8)	0.8601±0.0723(6)	0.8690±0.0724(3)	0.8652±0.0811(5)	0.8499±0.0741(7)	0.8930±0.0450(2)	0.8667±0.0753(4)	**0.8937±0.0530(1)**
Pageblocks2	0.8144±0.0222(8)	0.9586±0.0115(3)	0.9607±0.0089(2)	0.9585±0.0104(4)	0.9541±0.0103(5)	0.9180±0.0158(7)	0.9484±0.0100(6)	**0.9634±0.0082(1)**
Pageblocks5	0.4170±0.0834(8)	0.8929±0.0227(5)	0.9010±0.0244(3)	0.8955±0.0242(4)	0.8884±0.0248(6)	0.8595±0.0463(7)	0.9087±0.0195(2)	**0.9197±0.0145(1)**
StatlogVehicleSilhouettes2	0.5221±0.0604(8)	0.7629±0.0251(4)	0.7675±0.0261(2)	0.7571±0.0292(5)	0.7495±0.0234(6)	0.7634±0.0197(3)	0.7439±0.0248(7)	**0.7769±0.0208 (1)**
WallFollowingRobotNavigation4	0.4712±0.0542(8)	0.9007±0.0132(3)	0.8630±0.0249(6)	0.8886±0.0143(5)	0.8924±0.0123(4)	**0.9078±0.0082(1)**	0.8428±0.0177(7)	0.9063±0.0106(2)
Yeast6	0.4818±0.1113(8)	0.9256±0.0571(4)	0.9115±0.0550(6)	0.9024±0.0638(7)	0.9141±0.0566(5)	0.9614±0.0043(3)	0.9660±0.0219(2)	**0.9674±0.0204(1)**
DMEAntiVirus	0.9518±0.0292(6)	**0.9774±0.0175(1)**	0.9518±0.0292(6)	0.9518±0.0292(6)	0.9516±0.0286(8)	0.9543±0.0294(3)	0.9538±0.0284(4)	0.9696±0.0194(2)
GLRCWL1	0.6301±0.1218(6)	0.7231±0.0859(2)	0.6301±0.1218(6)	0.6301±0.1218(6)	0.6268±0.1192(8)	0.6666±0.1163(3)	0.6322±0.1215(4)	**0.7845±0.0695(1)**
GLRCNBI2	0.3168±0.2283(7)	0.4904±0.1967(3)	0.3168±0.2283(7)	0.3168±0.2283(7)	0.3221±0.2339(4)	0.4984±0.1499(2)	0.3216±0.2239(5)	**0.5605±0.1158(1)**
ParkinsonsDC	0.6829±0.0441(6)	0.7262±0.0337(2)	0.6829±0.0441(6)	0.6829±0.0441(6)	0.6799±0.0411(8)	0.6982±0.0317(3)	0.6908±0.0423(4)	**0.7348±0.0309(1)**
Colon 1	0.5907±0.1869(6)	0.7199±0.1155(2)	0.5907±0.1869(6)	0.5907±0.1869(6)	0.5916±0.1877(4)	0.6627±0.1481(3)	0.5886±0.1857(8)	**0.8004±0.0898(1)**
Leukemia 1	0.7633±0.0882(6)	0.8799±0.0815(3)	0.7633±0.0882(6)	0.7633±0.0882(6)	0.7604±0.0897(8)	0.8924±0.0913(2)	0.7641±0.0889(4)	**0.9586±0.0496(1)**
DrivFace1	0.7441±0.0877(6)	0.8611±0.0827(3)	0.7441±0.0877(6)	0.7441±0.0877(6)	0.7193±0.0966(8)	**0.8943±0.0557(1)**	0.7476±0.0876(4)	0.8858±0.0616(2)
ARBT8	0.0000±0.0000(5)	0.0000±0.0000(5)	0.0000±0.0000(5)	0.0000±0.0000(5)	0.0000±0.0000(5)	**0.4300±0.3475(1)**	0.0000±0.0000(5)	0.0000±0.0000(5)

A stratified k-fold cross validation (k=2 in experience) is used for 30 times that 60 (2×30) runs are conducted. Thus for each table cell, the mean and standard deviation of corresponding performance on 60 runs are first recorded and then its rank among all methods is followed in one bracket. The best rank for each row is highlighted as bold.

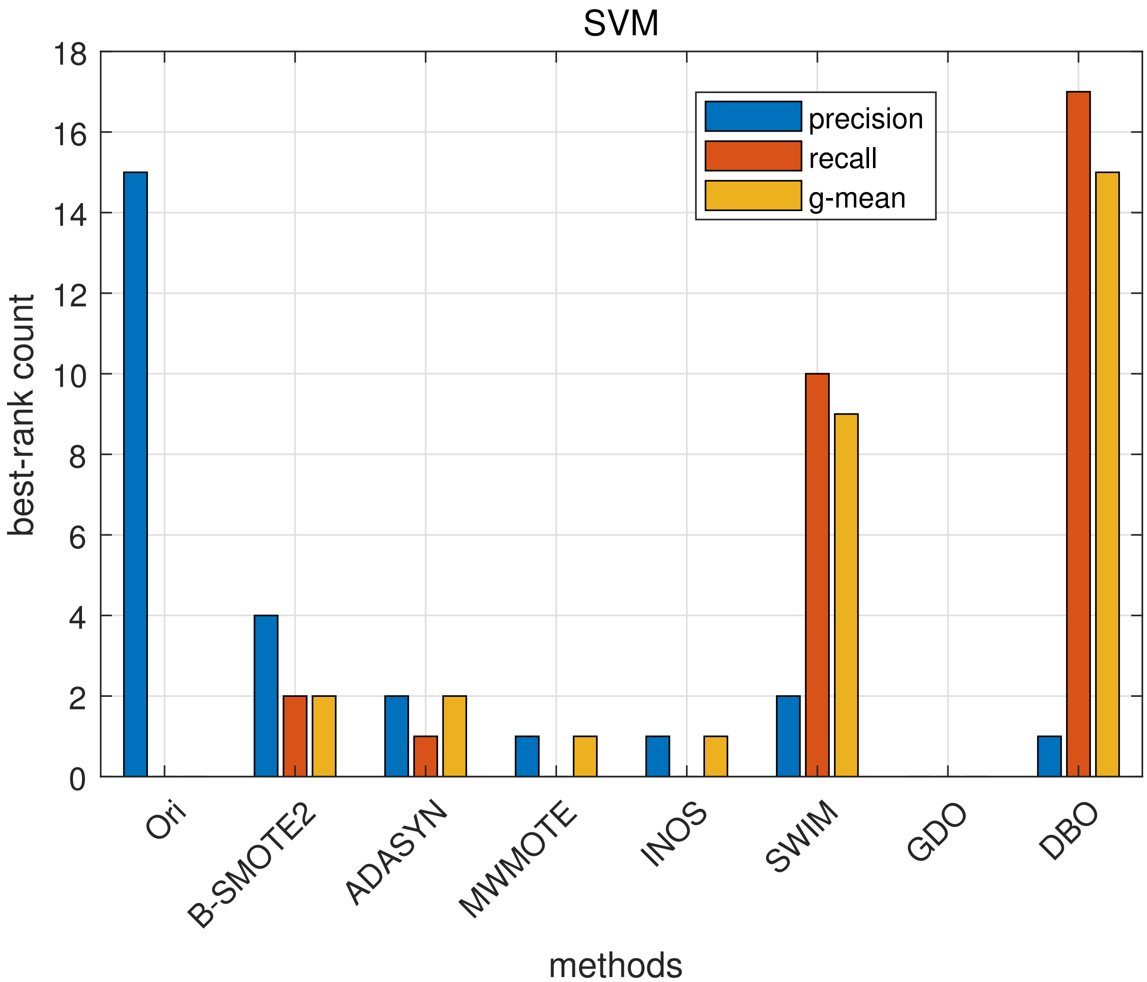

Fig. S2: Counts of best performance on SVM. We count the time for each method that achieves the best rank over all datasets.

TABLE S9: SVM: mean ranks of recall, precision and g-mean

Measurement	Actual value(Friedman test)	Ori	B-SMOTE2	ADASYN	MWMOTE	INOS	SWIM	GDO	DBO
precision	58.59(reject)	**3.07**	4.70	4.18	3.17	3.37	6.20	5.22	6.10
recall	122.89(reject)	7.42†	3.67†	4.20†	5.38†	6.07†	2.97	4.77†	**1.53**
g-mean	100.76(reject)	7.42†	3.73†	4.07†	5.15†	5.73†	3.23	4.73†	**1.93**
the Friedman Test: F=14.07, (n=8-1,alpha=0.05)									
the Bonferroni-Dunn test: critical values=1.70, (k=8,alpha=0.05)									

The best mean rank is highlighted as bold.
Using the Friedman Test with F=14.07 under (n=8-1,alpha=0.05) as one of the statistic test. If the actual value is larger than 14.07, reject the original hypothesis that significant difference exists among those methods.
Using the Bonferroni-Dunn test with critical values=1.70 under (k=8, alpha=0.05) as one of the statistic test. If the gap of mean rank between two methods is larger than 1.70, reject the original hypothesis that significant difference exists between two methods and marked with †. There, we only consider the measurement in which the proposed method is of the best mean rank , and compare it with any one of other methods.

TABLE S10: SVM: Wilcoxon paired signed-rank test for pairwise comparisons

recall		g-mean	
Ours vs.	p-V	Ours vs.	p-V
SWIM	**0.0285**	SWIM	**0.0073**

p-value under 0.05 is highlighted as bold which means the proposed one outperforms the compared one when using the Wilcoxon rank test (alpha=0.05)).

VI. Classification performance on AdaBoostM1

TABLE S11: AdaBoostM1: average precision

Cancer wpbc ret	**0.5003±0.2566(1)**	0.4129±0.1152(8)	0.4365±0.1008(4)	0.4344±0.0906(5)	0.4276±0.0860(6)	0.4260±0.1218(7)	0.4540±0.0932(3)	0.4552±0.1244(2)
Diabetes absent	**0.7398±0.0694(1)**	0.6907±0.0577(6)	0.6847±0.0563(8)	0.6904±0.0535(7)	0.7144±0.0575(2)	0.7122±0.0484(3)	0.7113±0.0647(4)	0.6954±0.0708(5)
Housing MEDV>35	**0.8817±0.0960(1)**	0.6573±0.0996(5)	0.6713±0.0810(4)	0.6990±0.0872(3)	0.7026±0.0941(2)	0.4945±0.2325(8)	0.6370±0.1163(6)	0.5755±0.1272(7)
Iris versicolor	**0.9414±0.0545(1)**	0.9210±0.0489(4)	0.9225±0.0498(3)	0.9327±0.0549(2)	0.9196±0.0561(5)	0.9170±0.0603(6)	0.9100±0.0543(7)	0.9086±0.0637(8)
Iris virginica	0.9051±0.1755(5)	0.9126±0.1265(4)	0.9013±0.1733(7)	0.8982±0.1750(8)	0.9164±0.1268(2)	0.9024±0.0666(6)	**0.9203±0.0481(1)**	0.9131±0.0543(3)
Spectf 0	**0.6958±0.1122(1)**	0.6401±0.0728(6)	0.6350±0.0598(7)	0.6436±0.0640(4)	0.6403±0.0681(5)	0.5119±0.1750(8)	0.6509±0.0614(3)	0.6550±0.0705(2)
Thyriod hyperfunction	0.0761±0.2547(8)	0.3882±0.0213(6)	0.6711±0.0902(2)	0.6224±0.0960(3)	**0.8218±0.0745(1)**	0.4062±0.3395(5)	0.2599±0.2834(7)	0.5621±0.2696 (4)
Vowel 4	**0.8692±0.2086(1)**	0.5376±0.1046(4)	0.5555±0.0921(2)	0.5395±0.0937(3)	0.4911±0.0852(5)	0.3530±0.0968(8)	0.4799±0.0780(6)	0.4366±0.0653(7)
Vowel 5	0.0389±0.1483(8)	0.2821±0.1091(4)	**0.3175±0.1050(1)**	0.3149±0.1099(2)	0.2288±0.1299(6)	0.1559±0.1498(7)	0.2978±0.0941(3)	0.2801±0.1141(5)
BreastTissue1	**0.8108±0.1867(1)**	0.7551±0.1467(7)	0.7715±0.1834(5)	0.7823±0.1858(4)	0.7927±0.1548(2)	0.7844±0.1099(3)	0.7564±0.1193(6)	0.7546±0.1094(8)
BreastTissue3	0.1372±0.2647(8)	0.2938±0.1415(2)	0.2787±0.1173(4)	0.2871±0.1529(3)	**0.3008±0.1434(1)**	0.1927±0.2133(7)	0.2372±0.1773(6)	0.2754±0.1060(5)
Ecoli2	0.8581±0.0772(3)	0.7507±0.1106(7)	0.7329±0.0788(8)	0.7636±0.0770(6)	0.8004±0.0872(5)	0.8348±0.0948(4)	0.8739±0.0794(2)	**0.8756±0.0746(1)**
Glass2	0.6712±0.1477(2)	0.6459±0.1060(6)	**0.6754±0.0955(1)**	0.6568±0.0892(5)	0.6705±0.1336(3)	0.6060±0.1658(7)	0.6058±0.2033(8)	0.6625±0.1067(4)
Glass3	0.0167±0.1291(8)	0.2479±0.2548(3)	0.2530±0.1883(2)	**0.2872±0.3126(1)**	0.2270±0.2256(4)	0.1063±0.2149(7)	0.1534±0.2921(6)	0.2024±0.1172(5)
ImageSegmentation7	0.9321±0.2513(3.5)	0.9314±0.2511(5)	0.9305±0.2509(7)	0.9321±0.2513 (3.5)	0.9311±0.2510(6)	0.9661±0.1809(2)	0.9217±0.2171(8)	**0.9986±0.0037(1)**
LibrasMovement2	0.1403±0.2307(8)	**0.3648±0.1355(1)**	0.3448±0.1200(2)	0.3382±0.1450(3)	0.3038±0.1624(6)	0.2177±0.1681(7)	0.3345±0.1434(4)	0.3108±0.1173(5)
LibrasMovement14	**0.8431±0.1443(1)**	0.7122±0.1628(3)	0.5785±0.1527(6)	0.7283±0.1614(2)	0.6075±0.1125(5)	0.4103±0.0862(8)	0.6417±0.1470(4)	0.5655±0.1302(7)
Pageblocks2	**0.8943±0.0461(1)**	0.7406±0.0401(5)	0.7041±0.0645(7)	0.7512±0.0362 (4)	0.7364±0.0391(6)	0.7692±0.1466(3)	0.8112±0.0662(2)	0.6648±0.0407(8)
Pageblocks5	**0.8025±0.1327(1)**	0.3960±0.0614(3)	0.3490±0.0576(6)	0.3573±0.0429 (5)	0.4169±0.0581(2)	0.1973±0.0972(8)	0.3765±0.1112(4)	0.3140±0.0368(7)
StatlogVehicleSilhouettes2	0.5189±0.3337(8)	0.5560±0.0770(6)	0.5597±0.0823(5)	0.5605±0.0806(4)	0.5392±0.0791(7)	0.5614±0.1395(3)	0.5644±0.1020(2)	**0.5674±0.1177 (1)**
WallFollowingRobotNavigation4	**0.9622±0.0411(1)**	0.6690±0.0644(5)	0.7927±0.0492(3)	0.7970±0.0920(2)	0.6907±0.0684(4)	0.6120±0.0539(7)	0.6032±0.0672(8)	0.6457±0.0712 (6)
Yeast6	**0.6380±0.2688(1)**	0.4867±0.0887(6)	0.5109±0.0841(4)	0.5087±0.0847(5)	0.5620±0.1390(2)	0.0813±0.1848(8)	0.3739±0.3107(7)	0.5320±0.0725(3)
DMEAntiVirus	0.9401±0.1769(6)	0.9400±0.1770(8)	0.9402±0.1769(4)	0.9402±0.1769(3)	0.9402±0.1769(5)	0.9401±0.1769(7)	**0.9723±0.0182(1)**	0.9403±0.1769(2)
GLRCWL1	0.7311±0.1612(2)	0.6968±0.1457(5)	**0.7317±0.1192(1)**	0.7261±0.1309(3)	0.6851±0.1306(6)	0.6512±0.1529(8)	0.7115±0.1328(4)	0.6695±0.1433(7)
GLRCNBI2	0.3162±0.2578(5)	0.3195±0.1931(4)	0.2838±0.2221(8)	0.3079±0.1927(6)	0.3382±0.2090(3)	0.3385±0.2348(2)	**0.3488±0.2065(1)**	0.2915±0.2030(7)
ParkinsonsDC	0.7468±0.0781(2)	0.6104±0.0672(7)	0.6231±0.0695(6)	0.6231±0.0612(5)	0.7066±0.0698(3)	0.7024±0.0873(4)	**0.7585±0.0856(1)**	0.5985±0.0711(8)
Colon 1	0.6908±0.1594(7)	0.6920±0.1465(6)	0.6957±0.1519(5)	0.6797±0.1674(8)	0.7193±0.1538(2)	0.7116±0.0992(4)	**0.7233±0.1206(1)**	0.7132±0.1044(3)
Leukemia 1	0.4885±0.4935(7)	0.7200±0.4069(4)	0.4907±0.4954(6)	0.4812±0.4870(8)	0.5554±0.4752(5)	**0.8840±0.1009(1)**	0.7911±0.3430(3)	0.8236±0.2909(2)
DrivFace1	**0.8604±0.1557(1)**	0.7412±0.1070(5)	0.7891±0.1496(2)	0.7807±0.1416(3)	0.7532±0.1051(4)	0.5482±0.0888(8)	0.7119±0.1099(7)	0.7320±0.1059(6)
ARBT8	0.8053±0.2249(1.5)	0.6224±0.2405(7)	0.7419±0.2453(4)	0.8053±0.2249(1.5)	0.5908±0.3137(8)	0.6372±0.3128(6)	0.6956±0.3242(5)	0.7574±0.2725(3)

A stratified k-fold cross validation (k=2 in experience) is used for 30 times that 60 (2×30) runs are conducted. Thus for each table cell, the mean and standard deviation of corresponding performance on 60 runs are first recorded and then its rank among all methods is followed in one bracket. The best rank for each row is highlighted as bold.

TABLE S12: AdaBoostM1: average recall

Cancer wpbc ret	0.1428±0.0934(8)	0.3428±0.1135(5)	**0.3848±0.1157(1)**	0.3819±0.0877(2)	0.3783±0.1087(3)	0.3275±0.1131(6)	0.3022±0.1145(7)	0.3471±0.1304(4)
Diabetes absent	0.3607±0.0692(8)	0.5090±0.0806(3)	**0.5143±0.0684(1)**	0.5116±0.0746(2)	0.4622±0.0747(4)	0.3956±0.0741(6)	0.3924±0.0787(7)	0.4302±0.0972(5)
Housing MEDV>35	0.6347±0.0978(5)	0.7736±0.0959(3)	**0.8063±0.1032(1)**	0.7785±0.0870(2)	0.7507±0.1005(4)	0.1861±0.1769(8)	0.4618±0.2188(6)	0.3583±0.1914(7)
Iris versicolor	0.8953±0.0776(8)	0.9333±0.0597(2)	0.9293±0.0626(5)	0.9120±0.0654(7)	0.9307±0.0589(3)	0.9280±0.0579(6)	0.9300±0.0816(4)	**0.9540±0.0447(1)**
Iris virginica	0.8647±0.1771(8)	0.9040±0.1322(4)	0.8867±0.1778(6)	0.8847±0.1754(7)	0.9033±0.1357(5)	**0.9333±0.0578(1)**	0.9313±0.0599(2)	0.9267±0.0626(3)
Spectf 0	0.3528±0.1270(8)	0.5890±0.1008(4)	**0.6046±0.0976(1)**	0.5869±0.1039(5)	0.6000±0.1079(3)	0.5089±0.2355(7)	0.5433±0.1130(6)	0.6039±0.1197(2)
Thyriod hyperfunction	0.0833±0.2787(6)	**1.0000±0.0000(1)**	0.9967±0.0120(3.5)	0.9967±0.0165(3.5)	0.9982±0.0052(2)	0.0416±0.0548(7)	0.0165±0.0223(8)	0.5247±0.4932(5)
Vowel 4	0.1958±0.0902(8)	0.7215±0.1028(5)	0.7299±0.1262(4)	0.6785±0.1258(6)	0.7556±0.0963(2)	0.6646±0.2261(7)	0.7535±0.1041(3)	**0.8021±0.0926(1)**
Vowel 5	0.0042±0.0166(8)	0.4000±0.2132(6)	**0.4583±0.2000(1)**	0.4278±0.1896(4)	0.4021±0.2598(5)	0.3542±0.3602(7)	0.4403±0.2029(3)	0.4424±0.2192(2)
BreastTissue1	0.7483±0.1961(8)	0.8533±0.1702(4)	0.8250±0.1910(6)	0.8217±0.1887(7)	0.8450±0.1512(5)	0.9067±0.0841(2)	0.8833±0.1122(3)	**0.9117±0.1121(1)**
BreastTissue3	0.0500±0.0902(8)	0.2870±0.1507(4)	0.3222±0.1916(3)	0.2852±0.1522(5)	0.3352±0.1967(2)	0.1926±0.1971(7)	0.2352±0.1995(6)	**0.3889±0.2081(1)**
Ecoli2	0.6320±0.1200(8)	0.7851±0.1282(3)	**0.8325±0.0905(1)**	0.8035±0.0916(2)	0.7662±0.0999(4)	0.6386±0.1096(6)	0.6351±0.1306(7)	0.6425±0.0795(5)
Glass2	0.3000±0.1378(6)	0.3662±0.1433(4)	0.4154±0.1291(3)	**0.4228±0.1111(1)**	0.3202±0.1396(5)	0.2272±0.1367(7)	0.1974±0.1264(8)	0.4219±0.1447(2)
Glass3	0.0021±0.0161(8)	0.2125±0.1664(3)	0.2604±0.1816(2)	0.1917±0.1602(4)	0.1833±0.1482(5)	0.0396±0.0586(7)	0.0417±0.0594(6)	**0.3688±0.2385(1)**
ImageSegmentation7	0.9256±0.2495(6.5)	0.9260±0.2496(4)	0.9256±0.2495(6.5)	0.9259±0.2496(5)	0.9266±0.2498(3)	0.9585±0.1795(2)	0.8849±0.2494(8)	**0.9920±0.0052 (1)**
LibrasMovement2	0.0458±0.0818(8)	0.4347±0.1770(4)	0.4708±0.2207(3)	0.3667±0.1887(7)	0.4069±0.2099(6)	0.4194±0.3211(5)	**0.5153±0.2123(1)**	0.4861±0.2287(2)
LibrasMovement14	0.5125±0.1470(8)	0.7333±0.1446(6)	0.7667±0.1462(3)	0.6875±0.1562(7)	0.7625±0.1494(4)	**0.8556±0.1105(1)**	0.7528±0.1447(5)	0.8194±0.1200(2)
Pageblocks2	0.6486±0.1081(7)	0.9116±0.0279(3)	0.9256±0.0234(2)	0.8937±0.0266(5)	0.8945±0.0375(4)	0.6461±0.1037(8)	0.7826±0.0761(6)	**0.9331±0.0213(1)**
Pageblocks5	0.3102±0.1140(8)	0.7772±0.0511(3)	0.7930±0.0598(2)	0.7716±0.0531(4)	0.7386±0.0623(5)	0.3251±0.1257(7)	0.6216±0.1261(6)	**0.8018±0.0464(1)**
StatlogVehicleSilhouettes2	0.0989±0.0864(8)	0.3026±0.1060(2)	0.2750±0.0943(4)	**0.3103±0.0882(1)**	0.2670±0.0837(5)	0.1949±0.1159(7)	0.2043±0.0955(6)	0.2884±0.1182(3)
WallFollowingRobotNavigation4	0.8489±0.0786(8)	**0.9812±0.0266(1)**	0.9722±0.0181(3)	0.9282±0.0305(7)	0.9548±0.0205(5)	0.9512±0.0180(6)	0.9694±0.0196(4)	0.9775±0.0175 (2)
Yeast6	0.3750±0.2509(7)	0.8652±0.1028(4)	**0.8939±0.0801(1)**	0.8697±0.0911(3)	0.8462±0.1175(5)	0.1523±0.3444(8)	0.4636±0.4237(6)	0.8818±0.1039(2)
DMEAntiVirus	0.9106±0.1791(7.5)	0.9250±0.1793(2)	0.9148±0.1780(6)	0.9153±0.1804(5)	0.9167±0.1794(4)	0.9106±0.1788(7.5)	**0.9421±0.0483(1)**	0.9190±0.1776(3)
GLRCWL1	0.5300±0.1769(8)	**0.6433±0.1731(1)**	0.6033±0.1697(5)	0.5933±0.1736(6)	0.6333±0.1980(2)	0.5850±0.1885(7)	0.6033±0.1737(4)	0.6150±0.1716(3)
GLRCNBI2	0.1738±0.1443(8)	**0.2905±0.1764(1)**	0.2452±0.1970(7)	0.2881±0.1994(2)	0.2762±0.1597(3)	0.2548±0.1824(4)	0.2548±0.1727(5)	0.2476±0.1702(6)
ParkinsonsDC	0.3734±0.0779(6)	0.5267±0.0697(2)	**0.5372±0.0729(1)**	0.5226±0.0761(3)	0.4405±0.0763(5)	0.3259±0.1128(7)	0.2687±0.1082(8)	0.4432±0.1036(4)
Colon 1	0.6030±0.1945(8)	0.6515±0.1967(3)	0.6242±0.1694(7)	0.6303±0.2015(6)	0.6318±0.1983(5)	0.6955±0.1689(2)	0.6424±0.1929(4)	**0.7000±0.1452(1)**
Leukemia 1	0.4722±0.4792(7)	0.7306±0.4109(4)	0.4681±0.4749(8)	0.4750±0.4806(6)	0.5486±0.4707(5)	**0.9042±0.1039(1)**	0.8028±0.3448(3)	0.8597±0.2953(2)
DrivFace1	0.6077±0.1698(8)	0.7462±0.1261(5)	0.7346±0.1684(6)	0.7244±0.1635(7)	0.7769±0.1449(3)	**0.8667±0.1087(1)**	0.7872±0.1246(2)	0.7577±0.1289(4)
ARBT8	0.2907±0.1343(3.5)	**0.3463±0.1419(1)**	0.2741±0.1317(5)	0.2907±0.1343(3.5)	0.2259±0.1402(6)	0.2259±0.1417(7)	0.2111±0.1413(8)	0.3315±0.1643(2)

A stratified k-fold cross validation (k=2 in experience) is used for 30 times that 60 (2×30) runs are conducted. Thus for each table cell, the mean and standard deviation of corresponding performance on 60 runs are first recorded and then its rank among all methods is followed in one bracket. The best rank for each row is highlighted as bold.

TABLE S13: AdaBoostM1: average g-mean

Cancer wpbc ret	0.3389±0.1409(8)	0.5287±0.0914(5)	**0.5616±0.0880(1)**	0.5605±0.0619(2)	0.5554±0.0717(3)	0.5179±0.0900(6)	0.5036±0.0951(7)	0.5343±0.1070(4)
Diabetes absent	0.5749±0.0512(8)	0.6640±0.0448(3)	**0.6662±0.0364(1)**	0.6658±0.0415(2)	0.6411±0.0458(4)	0.5973±0.0504(6)	0.5938±0.0523(7)	0.6135±0.0575(5)
Housing MEDV>35	0.7902±0.0620(5)	0.8574±0.0517(3)	**0.8760±0.0541(1)**	0.8643±0.0471(2)	0.8486±0.0556(4)	0.3808±0.1919(8)	0.6485±0.1603(6)	0.5666±0.1602(7)
Iris versicolor	0.9307±0.0408(8)	0.9446±0.0262(2)	0.9427±0.0258(3)	0.9368±0.0294(7)	0.9425±0.0254(4)	0.9400±0.0225(5)	0.9393±0.0429(6)	**0.9504±0.0207(1)**
Iris virginica	0.8980±0.1713(8)	0.9243±0.1240(5)	0.9087±0.1728(6)	0.9066±0.1718(7)	0.9248±0.1251(4)	0.9385±0.0296(3)	**0.9434±0.0274(1)**	0.9387±0.0268(2)
Spectf 0	0.5622±0.1076(8)	0.7132±0.0556(5)	0.7220±0.0565(2)	0.7141±0.0610(4)	0.7192±0.0583(3)	0.6073±0.2151(7)	0.6905±0.0679(6)	**0.7237±0.0690(1)**
Thyriod hyperfunction	0.0831±0.2780(8)	0.9571±0.0038(4)	0.9844±0.0081(2)	0.9811±0.0106(3)	**0.9931±0.0045(1)**	0.1511±0.1372(6)	0.0907±0.0914(7)	0.5719±0.4360 (5)
Vowel 4	0.4265±0.1140(8)	0.8179±0.0553(5)	0.8233±0.0683(4)	0.7932±0.0709(6)	0.8299±0.0503(2)	0.7391±0.1530(7)	0.8267±0.0495(3)	**0.8445±0.0484(1)**
Vowel 5	0.0164±0.0628(8)	0.5584±0.2117(5)	**0.6127±0.1785(1)**	0.5933±0.1779(3)	0.5140±0.2866(6)	0.3929±0.3746(7)	0.5963±0.1832(2)	0.5803±0.2343(4)
BreastTissue1	0.8279±0.1715(8)	0.8801±0.1347(5)	0.8633±0.1701(7)	0.8639±0.1704(6)	0.8835±0.1265(4)	**0.9190±0.0422(1)**	0.9013±0.0584(3)	0.9142±0.0571(2)
BreastTissue3	0.1153±0.1841(8)	0.4595±0.1676(5)	0.4822±0.1551(3)	0.4596±0.1526(4)	0.4912±0.1843(2)	0.3021±0.2508(7)	0.3687±0.2507(6)	**0.5199±0.1725(1)**
Ecoli2	0.7768±0.0707(7)	0.8410±0.0560(4)	**0.8652±0.0336(1)**	0.8582±0.0395(2)	0.8449±0.0438(3)	0.7774±0.0559(6)	0.7764±0.0966(8)	0.7875±0.0454(5)
Glass2	0.5067±0.1379(6)	0.5568±0.1050(4)	0.5955±0.0896(2)	**0.5999±0.0737(1)**	0.5209±0.1175(5)	0.4320±0.1443(7)	0.3984±0.1495(8)	0.5950±0.1136(3)
Glass3	0.0059±0.0456(8)	0.3817±0.2207(3)	0.4329±0.2161(2)	0.3548±0.2223(4)	0.3470±0.2221(5)	0.1100±0.1629(7)	0.1162±0.1657(6)	**0.5120±0.2086(1)**
ImageSegmentation7	0.9293±0.2505(6)	0.9295±0.2505(5)	0.9292±0.2505(7)	0.9295±0.2505(4)	0.9298±0.2506(3)	0.9625±0.1803(2)	0.9081±0.2402(8)	**0.9959±0.0025(1)**
LibrasMovement2	0.1155±0.1793(8)	0.6226±0.1299(4)	0.6408±0.1544(2)	0.5638±0.1670(6)	0.5817±0.1771(5)	0.5057±0.3413(7)	**0.6647±0.1578(1)**	0.6393±0.1796(3)
LibrasMovement14	0.7052±0.1074(8)	0.8401±0.0848(6)	0.8498±0.0815(4)	0.8143±0.0978(7)	0.8515±0.0841(3)	0.8766±0.0510(2)	0.8472±0.0847(5)	**0.8789±0.0625(1)**
Pageblocks2	0.8002±0.0686(7)	0.9447±0.0137(3)	0.9495±0.0112(2)	0.9362±0.0132(4)	0.9357±0.0181(5)	0.7955±0.0657(8)	0.8779±0.0409(6)	**0.9511±0.0105(1)**
Pageblocks5	0.5430±0.1225(8)	0.8694±0.0280(3)	0.8751±0.0325(2)	0.8645±0.0288(4)	0.8487±0.0350(5)	0.5459±0.1248(7)	0.7729±0.0840(6)	**0.8778±0.0249(1)**
StatlogVehicleSilhouettes2	0.2544±0.1781(8)	0.5144±0.0845(2)	0.4941±0.0874(4)	**0.5250±0.0674(1)**	0.4872±0.0700(5)	0.4011±0.1356(7)	0.4235±0.1031(6)	0.4984±0.1177(3)
WallFollowingRobotNavigation4	0.9192±0.0442(8)	0.9745±0.0127(2)	**0.9777±0.0095(1)**	0.9554±0.0147(7)	0.9632±0.0109(5)	0.9559±0.0093(6)	0.9636±0.0088(4)	0.9709±0.0089 (3)
Yeast6	0.5525±0.2613(6)	0.9143±0.0549(4)	**0.9312±0.0413(1)**	0.9184±0.0490(3)	0.9053±0.0636(5)	0.1568±0.3537(8)	0.5185±0.4348(7)	0.9255±0.0560(2)
DMEAntiVirus	0.9347±0.1772(8)	0.9422±0.1780(2)	0.9370±0.1771(6)	0.9371±0.1778(5)	0.9379±0.1776(4)	0.9348±0.1772(7)	**0.9671±0.0245(1)**	0.9392±0.1772(3)
GLRCWL1	0.6868±0.1246(8)	**0.7445±0.1047(1)**	0.7299±0.0978(3)	0.7214±0.0959(6)	0.7375±0.1188(2)	0.7036±0.1192(7)	0.7284±0.1075(4)	0.7227±0.1089(5)
GLRCNBI2	0.3324±0.2130(8)	**0.4549±0.1916(1)**	0.4009±0.2125(7)	0.4505±0.2035(3)	0.4536±0.1795(2)	0.4134±0.2138(5)	0.4290±0.1926(4)	0.4059±0.2191(6)
ParkinsonsDC	0.5935±0.0607(6)	0.6797±0.0409(2)	**0.6874±0.0407(1)**	0.6788±0.0416(3)	0.6386±0.0510(4)	0.5459±0.0917(7)	0.4982±0.1008(8)	0.6239±0.0655(5)
Colon 1	0.6957±0.1397(8)	0.7241±0.1418(4)	0.7155±0.1304(5)	0.7069±0.1480(7)	0.7149±0.1393(6)	0.7503±0.0849(2)	0.7253±0.1183(3)	**0.7593±0.0816(1)**
Leukemia 1	0.4822±0.4870(6)	0.7336±0.4102(4)	0.4809±0.4857(8)	0.4817±0.4862(7)	0.5573±0.4757(5)	**0.9136±0.0572(1)**	0.8082±0.3444(3)	0.8548±0.2901(2)
DrivFace1	0.7650±0.1410(8)	0.8550±0.0732(5)	0.8423±0.1366(6)	0.8362±0.1356(7)	0.8717±0.0834(3)	**0.9129±0.0581(1)**	0.8773±0.0701(2)	0.8611±0.0754(4)
ARBT8	0.5201±0.1404(3.5)	**0.5723±0.1254(1)**	0.5016±0.1476(5)	0.5201±0.1404(3.5)	0.4300±0.2013(7)	0.4332±0.1950(6)	0.4203±0.1856(8)	0.5424±0.1916(2)

A stratified k-fold cross validation (k=2 in experience) is used for 30 times that 60 (2×30) runs are conducted. Thus for each table cell, the mean and standard deviation of corresponding performance on 60 runs are first recorded and then its rank among all methods is followed in one bracket. The best rank for each row is highlighted as bold.

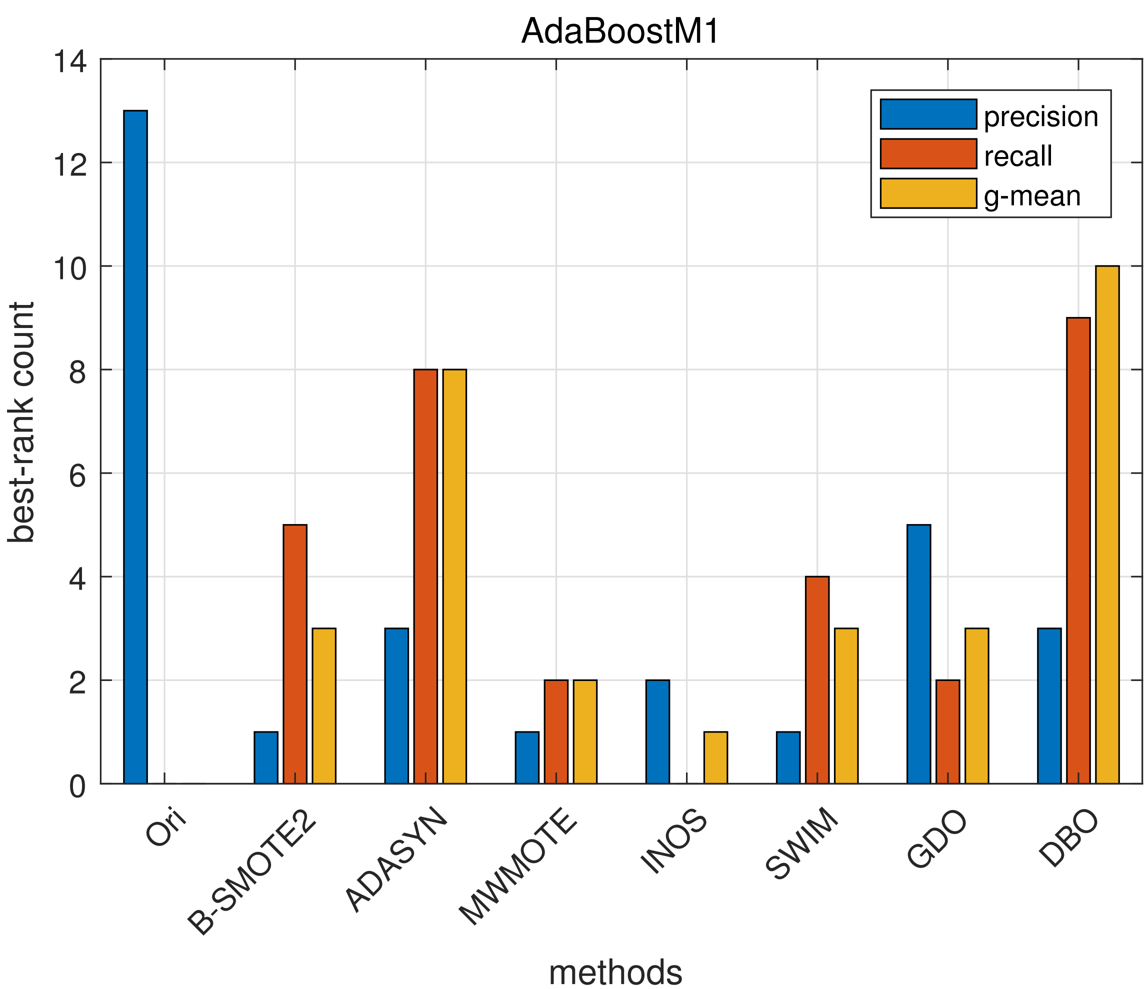

Fig. S3: Counts of best performance on AdaBoost. We count the time for each method that achieves the best rank over all datasets.

TABLE S14: AdaBoostM1: mean ranks of recall, precision and g-mean

Measurement	Actual value(Friedman test)	Ori	B-SMOTE2	ADASYN	MWMOTE	INOS	SWIM	GDO	DBO
precision	16.38(reject)	**3.50**	5.07	4.47	4.07	4.10	5.73	4.33	4.73
recall	78.76(reject)	7.38†	3.23	3.60	4.50†	4.07	5.48†	5.10†	**2.63**
g-mean	74.17(reject)	7.35†	3.57	3.33	4.35	3.97	5.53†	5.07†	**2.83**
the Friedman Test: F=14.07, (n=8-1,alpha=0.05)									
the Bonferroni-Dunn test: critical values=1.70, (k=8,alpha=0.05)									

The best mean rank is highlighted as bold.
Using the Friedman Test with F=14.07 under (n=8-1,alpha=0.05) as one of the statistic test. If the actual value is larger than 14.07, reject the original hypothesis that significant difference exists among those methods.
Using the Bonferroni-Dunn test with critical values=1.70 under (k=8, alpha=0.05) as one of the statistic test. If the gap of mean rank between two methods is larger than 1.70, reject the original hypothesis that significant difference exists between two methods and marked with †. There, we only consider the measurement in which the proposed method is of the best mean rank , and compare it with any one of other methods.

TABLE S15: AdaBoostM1: Wilcoxon paired signed-rank test for pairwise comparisons

recall		g-mean	
Ours vs.	p-V	Ours vs.	p-V
B-SMOTE2	0.2210	B-SMOTE2	0.4048
ADASYN	0.2059	ADASYN	0.4653
INOS	**0.0300**	MWMOTE	0.2452
		INOS	0.0897

p-value under 0.05 is highlighted as bold which means the proposed one outperforms the compared one when using the Wilcoxon rank test (alpha=0.05)).

VII. Ablation study between linear interpolation and surrounding area

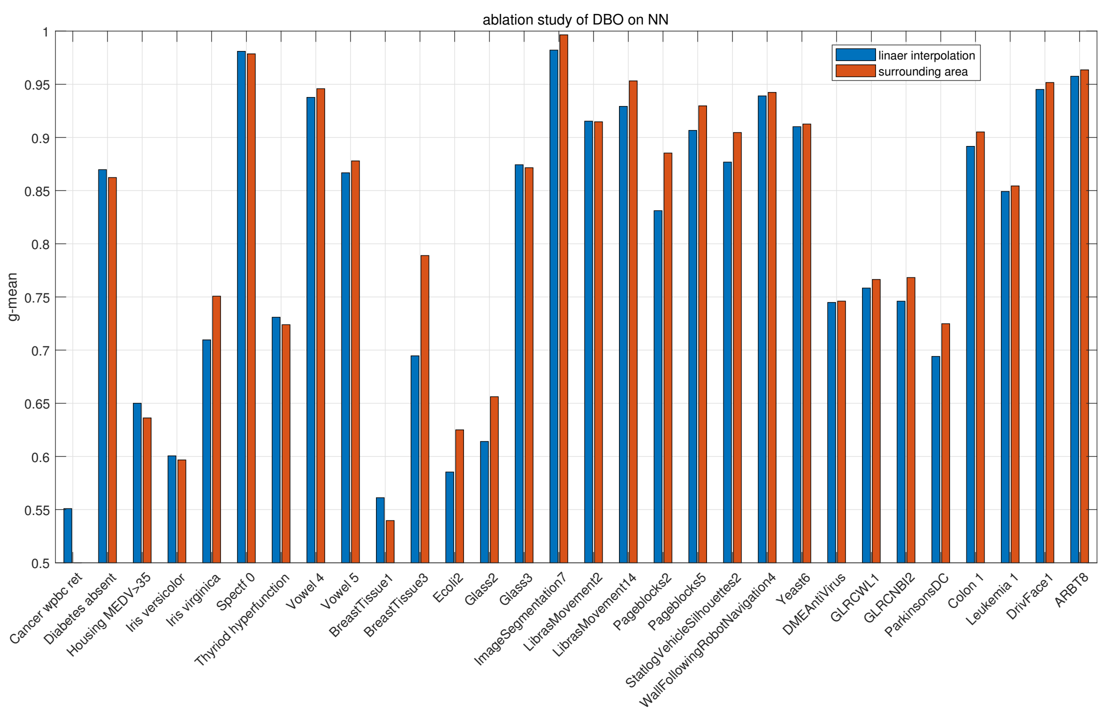

Fig. S4: Ablation study between linear interpolation and surrounding area.

VIII. Robust to outliers

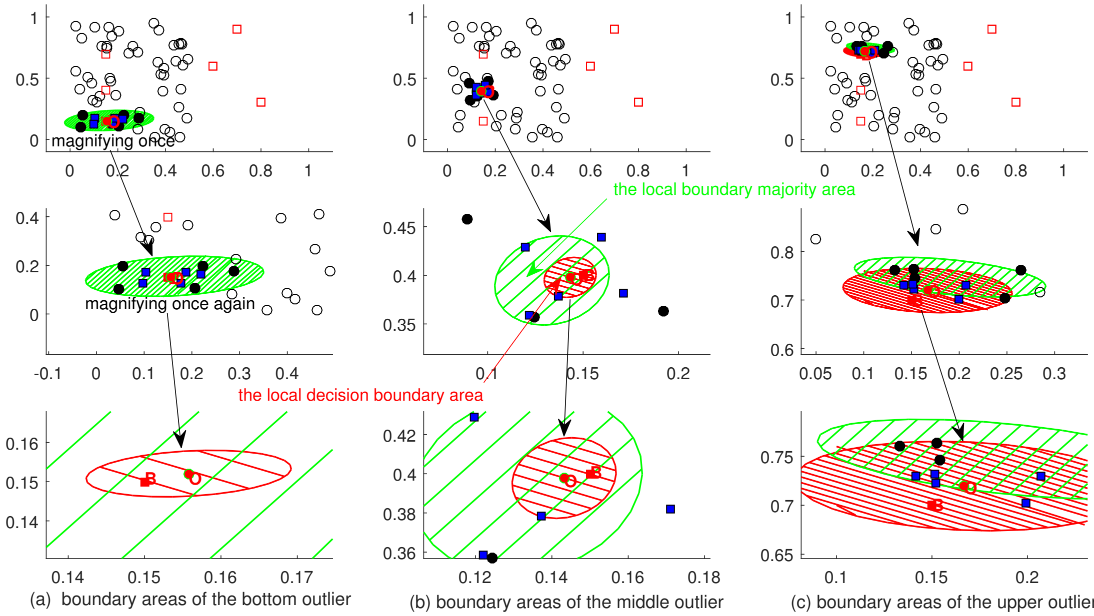

Fig. S5: Robustness of DBO to outliers. The black circle denotes majority samples and the red square denotes minority samples in the original data, the solid red square denotes the target boundary minority sample B, the solid black circle denote k nearest majority samples for B, the solid blue square denotes mean points between B and its k nearest boundary majority samples, the solid red circle denotes the center of those mean points and B, the red slash area denotes the local decision boundary area and the blue slash area denotes the local boundary majority area. For the first and second outlier, no local boundary minority area exists for the reason that the local boundary majority area covers the whole local decision boundary area. For the third outlier, middle size of local boundary minority area exists for the reason that its local boundary majority area covers nearly half of corresponding local decision boundary area. Obviously, k nearest majority samples of the third outlier are of the same side when compared to it, while k nearest majority samples of the first and second outlier are surrounding them. In other word, the remained local boundary minority area of the third outlier distributes in the non-majority existed region (a certain region that no majority samples existed).

IX. Parameter setting of DBO

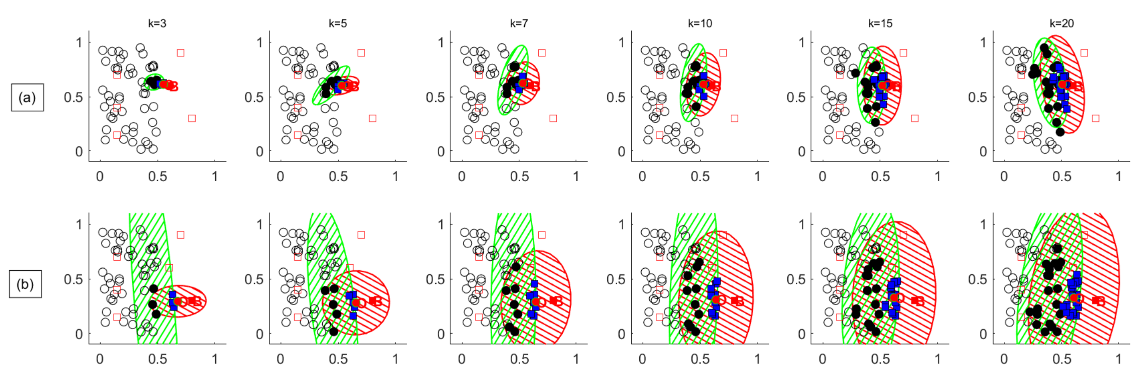

Fig. S6: Different size of boundary area with varying k. The parameter k, which means k nearest majority samples of the target boundary minority sample, is involved in DBO. Different k values are tested for two boundary minority samples respectively in (a) and (b). Obviously, larger the value of k is, larger sizes of both the local decision boundary area and local boundary minority area are.

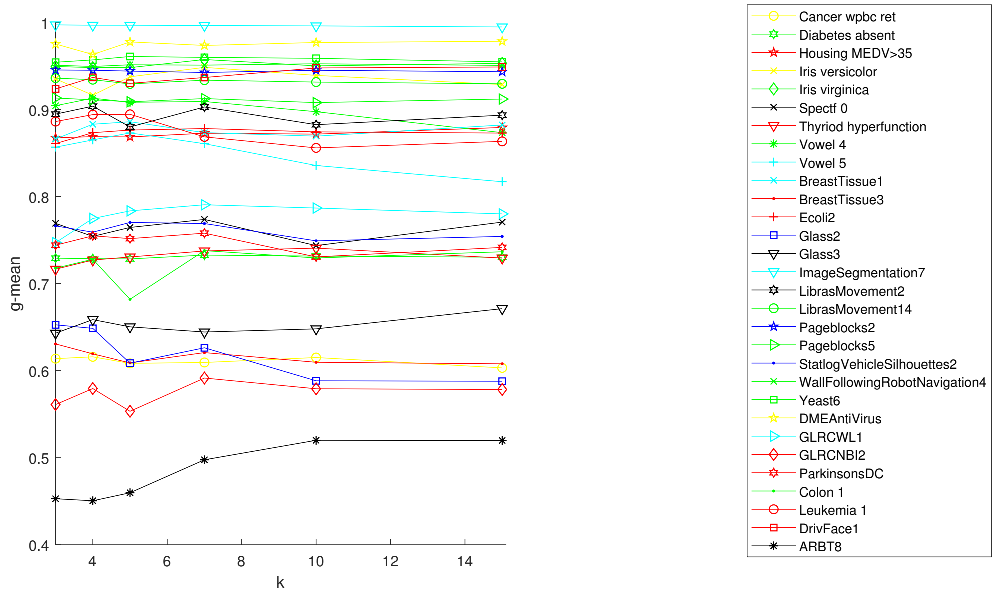

Fig. S7: g-mean on NN: parameter $k(=[3, 4, 5, 7, 10, 15])$. In experience, we set $k = 7$.

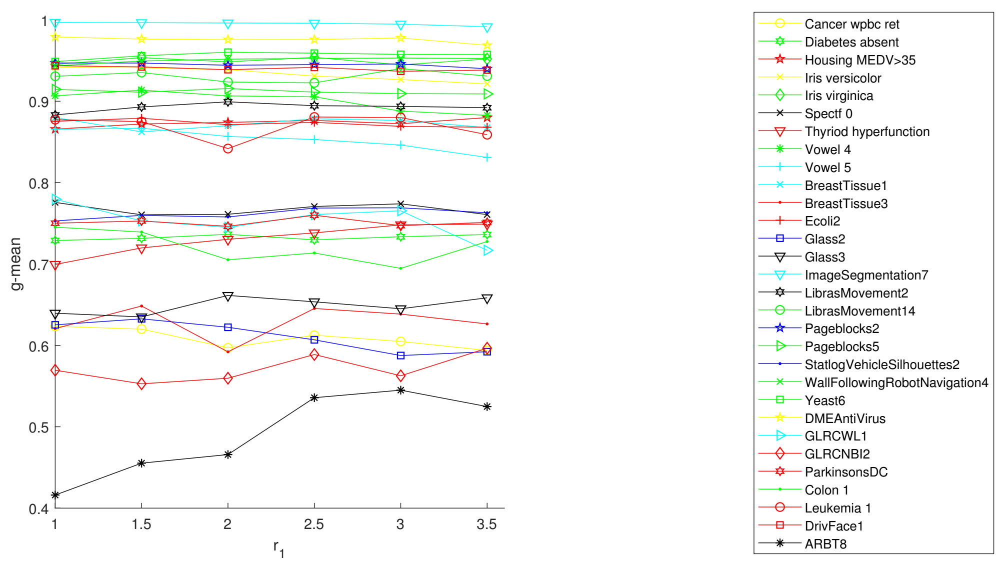

Fig. S8: g-mean on NN: parameter $r_1(=[1, 1.5, 2, 2.5, 3, 3.5])$. We use r_1 to control the length (α_2) of local decision boundary area where $\alpha_2 = r_1 * \alpha_1$. In experience, we set $r_1 = 2$.

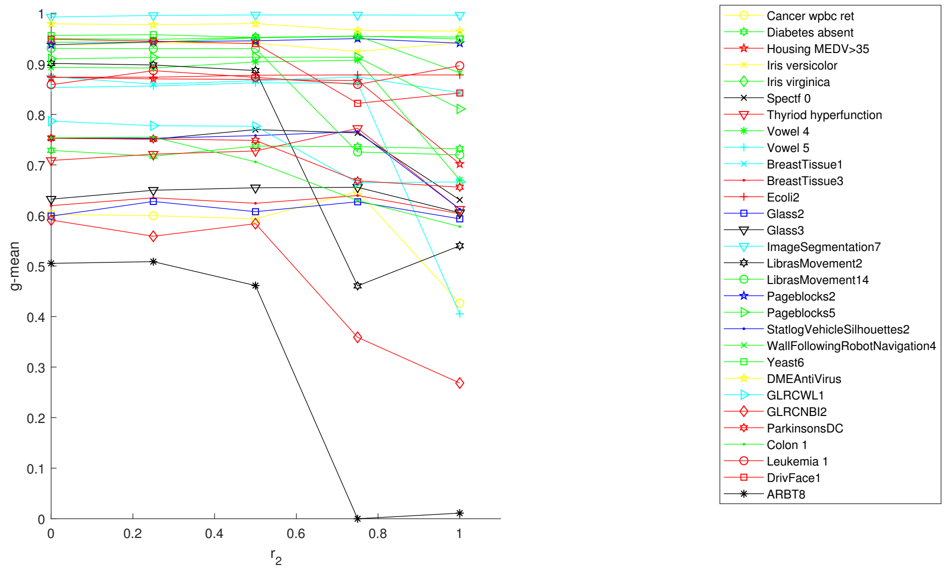

Fig. S9: g-mean on NN: parameter $r_2(=[0, 0.25, 0.5, 0.75, 1],)$. We use r_2 to control the length (α_5) of local boundary majority area where $\alpha_5 = \alpha_4 + r_2 * |\alpha_3 - \alpha_4|$. In experience, we set $r_2 = 0.5$.

X. Time consuming of different methods

TABLE S16: Time consuming of different methods (seconds)

Dataset	B-SMOTE2	ADASYN	MWMOTE	INOS	SWIM	GDO	DBO
Cancer wpbc ret	0.00	0.00	0.00	0.01	0.00	0.00	0.03
Diabetes absent	0.00	0.00	0.01	0.02	0.00	0.01	0.08
Housing MEDV>35	0.00	0.00	0.00	0.02	0.00	0.01	0.04
Iris versicolor	0.00	0.00	0.00	0.00	0.00	0.00	0.01
Iris virginica	0.00	0.00	0.00	0.00	0.00	0.00	0.01
Spectf 0	0.00	0.00	0.00	0.01	0.00	0.00	0.05
Thyriod hyperfunction	0.00	0.00	0.01	0.71	0.00	0.07	0.35
Vowel 4	0.00	0.00	0.00	0.02	0.00	0.01	0.04
Vowel 5	0.00	0.00	0.00	0.02	0.00	0.01	0.04
BreastTissue1	0.00	0.00	0.00	0.00	0.00	0.00	0.01
BreastTissue3	0.00	0.00	0.00	0.01	0.00	0.00	0.01
Ecoli2	0.00	0.00	0.00	0.01	0.00	0.00	0.03
Glass2	0.00	0.00	0.00	0.00	0.00	0.00	0.02
Glass3	0.00	0.00	0.00	0.02	0.00	0.00	0.02
ImageSegmentation7	0.13	0.01	0.01	0.28	0.00	0.03	0.15
LibrasMovement2	0.00	0.00	0.00	0.15	0.00	0.01	0.12
LibrasMovement14	0.00	0.00	0.00	0.14	0.00	0.01	0.08
Pageblocks2	0.02	0.01	0.03	0.90	0.00	0.08	0.41
Pageblocks5	0.04	0.03	0.06	8.70	0.00	0.08	0.45
StatlogVehicleSilhouettes2	0.01	0.02	0.04	1.23	0.00	0.01	0.10
WallFollowingRobotNavigation4	1.53	0.12	0.21	26.19	0.00	0.11	1.35
Yeast6	0.01	0.01	0.02	2.80	0.00	0.02	0.09
DMEAntiVirus	0.04	0.07	0.11	1.30	0.03	0.27	1.08
GLRCWL1	0.02	0.01	0.06	0.32	0.02	0.07	0.99
GLRCNBI2	0.03	0.03	0.06	0.44	0.03	0.09	1.16
ParkinsonsDC	0.03	0.16	0.27	3.91	0.04	0.88	4.78
Colon 1	0.02	0.03	0.07	0.82	0.21	0.27	2.12
Leukemia 1	0.02	0.06	0.11	5.08	2.42	1.46	18.51
DrivFace1	0.27	0.09	0.18	138.07	17.68	141.93	125.75
ARBT8	0.37	0.07	0.03	356.30	17.47	301.03	401.52

XI. Classification performance on image datasets

TABLE S17: Basic properties of image datasets

dataset	minority, majority class	training number of minority : majority	testing number of minority : majority	ratio
$cifar10_4$	4; rest	5000:45000	1000:9000	9.0
$cifar10_2$	2; rest	5000:45000	1000:9000	9.0
$cifar100_{3R88}$	3; 88	250:500	50:100	2.0
$cifar100_{10R17}$	10; 17	250:500	50:100	2.0
$flower102_{1R3}$	1; 3	5:10	10:20	2.0
$flower102_{8R71}$	8; 71	5:10	29:58	2.0

In experience, we first sent the trained images to train a Convolutional Neural Network (CNN); then use the trained model to extract corresponding features of both trained and test images, called them as features of train data and features of test data; finally, we use the over-sampling methods on features of train data, and test their classification performance on features of test data.

TABLE S18: neural network (NN)

Dataset	Ori	B-SMOTE2	ADASYN	MWMOTE	INOS	SWIM	GDO	DBO
				Precision				
$cifar10_4$	0.0000(8)	0.2914(3)	0.2892(4)	0.2677(6)	0.3199(2)	**0.6735(1)**	0.2889(5)	0.2423(7)
$cifar10_2$	**0.8820(1)**	0.5721(4)	0.4682(6)	0.7045(2)	0.6187(3)	0.3202(8)	0.4787(5)	0.4558(7)
$cifar100_{3R88}$	0.8214(2)	0.7931(6)	0.7869(7)	0.8070(4.5)	0.8182(3)	**0.8679(1)**	0.8070(4.5)	0.7833(8)
$cifar100_{10R17}$	0.7234(2)	0.7193(3)	0.6833(6)	0.6167(8)	0.7000(4)	**0.8889(1)**	0.6885(5)	0.6618(7)
$flower102_{1R3}$	1.0000(3)	1.0000(3)	0.4000(8)	1.0000(3)	0.7500(6.5)	1.0000(3)	1.0000(3)	0.7500(6.5)
$flower102_{8R71}$	0.8214(4)	0.8214(4)	0.8214(4)	0.8462(2)	0.8095(6)	**0.9048(1)**	0.7826(7)	0.7059(8)
				Recall				
$cifar10_4$	0.0000(8)	0.6260(4)	0.6290(3)	0.6700(2)	0.5550(6)	0.0330(7)	0.6050(5)	**0.7380(1)**
$cifar10_2$	0.7920(8)	0.9320(5)	0.9570(3)	0.8870(7)	0.9200(6)	**0.9720(1)**	0.9560(4)	0.9590(2)
$cifar100_{3R88}$	0.9200(5)	0.9200(5)	**0.9600(1)**	0.9200(5)	0.9000(8)	0.9200(5)	0.9200(5)	0.9400(2)
$cifar100_{10R17}$	0.6800(7)	0.8200(4.5)	0.8200(4.5)	0.7400(6)	0.8400(2.5)	0.6400(8)	0.8400(2.5)	**0.9000(1)**
$flower102_{1R3}$	0.6000(2)	0.4000(5)	0.2000(8)	0.5000(4)	0.3000(6.5)	0.6000(2)	0.3000(6.5)	0.6000(2)
$flower102_{8R71}$	0.7931(3)	0.7931(3)	0.7931(3)	0.7586(5)	0.5862(8)	0.6552(6)	0.6207(7)	**0.8276(1)**
				G-mean				
$cifar10_4$	0.0000(8)	0.7212(4)	0.7218(3)	0.7304(2)	0.6944(6)	0.1815(7)	0.7106(5)	**0.7408(1)**
$cifar10_2$	0.8847(7)	0.9273(2)	0.9173(5)	0.9221(3)	**0.9285(1)**	0.8655(8)	0.9195(4)	0.9149(6)
$cifar100_{3R88}$	0.9099(3)	0.8998(8)	0.9139(2)	0.9049(4.5)	0.9000(7)	**0.9250(1)**	0.9049(4.5)	0.9043(6)
$cifar100_{10R17}$	0.7692(7)	0.8299(2.5)	0.8150(5)	0.7549(8)	0.8299(2.5)	0.7838(6)	0.8249(4)	**0.8325(1)**
$flower102_{1R3}$	0.7746(1.5)	0.6325(5)	0.4123(8)	0.7071(4)	0.5339(7)	0.7746(1.5)	0.5477(6)	0.7348(3)
$flower102_{8R71}$	0.8513(2)	0.8513(2)	0.8513(2)	0.8404(4)	0.7388(8)	0.7953(6)	0.7531(7)	0.8276(5)

The best rank for each row is highlighted as bold.

TABLE S19: support vector machine (SVM)

Dataset	Ori	B-SMOTE2	ADASYN	MWMOTE	INOS	SWIM	GDO	DBO
Precision								
$cifar10_4$	0.0000(7.5)	0.2503(2)	0.1642(4)	**0.2952(1)**	0.2279(3)	0.0000(7.5)	0.1560(5)	0.1463(6)
$cifar10_2$	**0.9393(1)**	0.5656(4)	0.4115(5)	0.7853(2)	0.7098(3)	0.1521(8)	0.3230(6)	0.3013(7)
$cifar100_{3R88}$	0.8846(1.5)	0.8519(3.5)	0.8070(6.5)	0.8519(3.5)	0.8214(5)	0.8846(1.5)	0.8070(6.5)	0.8000(8)
$cifar100_{10R17}$	0.8824(1.5)	0.7778(5)	0.7451(7)	0.8049(4)	0.8140(3)	0.8824(1.5)	0.7547(6)	0.7358(8)
$flower102_{1R3}$	0.0000(5)	0.0000(5)	0.0000(5)	0.0000(5)	0.0000(5)	0.0000(5)	0.0000(5)	**1.0000(1)**
$flower102_{8R71}$	0.8571(4)	0.8571(4)	0.8571(4)	0.8571(4)	0.8571(4)	0.8571(4)	0.8571(4)	0.8000(8)
Recall								
Dataset	Ori	B-SMOTE2	ADASYN	MWMOTE	INOS	SWIM	GDO	DBO
$cifar10_4$	0.0000(7.5)	0.7010(2)	**0.7200(1)**	0.6810(5)	0.6650(6)	0.0000(7.5)	0.6930(3)	0.6910(4)
$cifar10_2$	0.7270(8)	0.8840(2)	**0.8970(1)**	0.8670(6)	0.8780(3)	0.7300(7)	0.8730(4)	0.8680(5)
$cifar100_{3R88}$	0.9200(5)	0.9200(5)	0.9200(5)	0.9200(5)	0.9200(5)	0.9200(5)	0.9200(5)	**0.9600(1)**
$cifar100_{10R17}$	0.6000(7.5)	0.7000(4.5)	0.7600(3)	0.6600(6)	0.7000(4.5)	0.6000(7.5)	**0.8000(1)**	0.7800(2)
$flower102_{1R3}$	0.0000(5)	0.0000(5)	0.0000(5)	0.0000(5)	0.0000(5)	0.0000(5)	0.0000(5)	**0.6000(1)**
$flower102_{8R71}$	0.2069(5)	0.2069(5)	0.2069(5)	0.2069(5)	0.2069(5)	0.2069(5)	0.2069(5)	**0.8276(1)**
G-mean								
$cifar10_4$	0.0000(7.5)	0.7331(2)	0.6532(4)	**0.7470(1)**	0.7061(3)	0.0000(7.5)	0.6358(5)	0.6177(6)
$cifar10_2$	0.8504(5)	0.9041(3)	0.8770(4)	**0.9188(1)**	0.9181(2)	0.6324(8)	0.8340(6)	0.8209(7)
$cifar100_{3R88}$	0.9299(1.5)	0.9200(3.5)	0.9049(7.5)	0.9200(3.5)	0.9099(6)	0.9299(1.5)	0.9049(7.5)	0.9191(5)
$cifar100_{10R17}$	0.7589(7.5)	0.7937(5)	0.8131(3)	0.7792(6)	0.8025(4)	0.7589(7.5)	**0.8343(1)**	0.8190(2)
$flower102_{1R3}$	0.0000(5)	0.0000(5)	0.0000(5)	0.0000(5)	0.0000(5)	0.0000(5)	0.0000(5)	**0.7746(1)**
$flower102_{8R71}$	0.4509(5)	0.4509(5)	0.4509(5)	0.4509(5)	0.4509(5)	0.4509(5)	0.4509(5)	**0.8614(1)**

The best rank for each row is highlighted as bold.

TABLE S20: AdaBoostM1

Dataset	Ori	B-SMOTE2	ADASYN	MWMOTE	INOS	SWIM	GDO	DBO
Precision								
$cifar10_4$	0.0000(6.5)	**0.3338(1)**	0.2342(4)	0.2691(3)	0.2812(2)	0.0000(6.5)	0.0000(6.5)	0.0000(6.5)
$cifar10_2$	0.8702(2)	0.6496(6)	0.6446(7)	0.7072(4)	0.6554(5)	0.8389(3)	**0.9661(1)**	0.4509(8)
$cifar100_{3R88}$	0.8800(1.5)	0.8364(4.5)	0.7759(8)	0.8302(6)	0.7797(7)	0.8800(1.5)	0.8364(4.5)	0.8679(3)
$cifar100_{10R17}$	0.8049(3.5)	0.7660(6)	0.7347(8)	0.8372(2)	0.7800(5)	0.8049(3.5)	0.7609(7)	**0.9118(1)**
$flower102_{1R3}$	0.0000(4.5)	0.0000(4.5)	0.0000(4.5)	0.0000(4.5)	0.0000(4.5)	0.0000(4.5)	0.0000(4.5)	0.0000(4.5)
$flower102_{8R71}$	0.0000(4.5)	0.0000(4.5)	0.0000(4.5)	0.0000(4.5)	0.0000(4.5)	0.0000(4.5)	0.0000(4.5)	0.0000(4.5)
Recall								
$cifar10_4$	0.0000(6.5)	0.5250(4)	**0.7620(1)**	0.6760(2)	0.6260(3)	0.0000(6.5)	0.0000(6.5)	0.0000(6.5)
$cifar10_2$	0.7640(7)	0.8900(3.5)	0.8940(2)	0.8840(5)	0.8900(3.5)	0.7760(6)	0.5130(8)	**0.9370(1)**
$cifar100_{3R88}$	0.8800(7)	0.9200(2.5)	0.9000(5)	0.8800(7)	0.9200(2.5)	0.8800(7)	0.9200(2.5)	0.9200(2.5)
$cifar100_{10R17}$	0.6600(6.5)	0.7200(3)	0.7200(3)	0.7200(3)	**0.7800(1)**	0.6600(6.5)	0.7000(5)	0.6200(8)
$flower102_{1R3}$	0.0000(4.5)	0.0000(4.5)	0.0000(4.5)	0.0000(4.5)	0.0000(4.5)	0.0000(4.5)	0.0000(4.5)	0.0000(4.5)
$flower102_{8R71}$	0.0000(4.5)	0.0000(4.5)	0.0000(4.5)	0.0000(4.5)	0.0000(4.5)	0.0000(4.5)	0.0000(4.5)	0.0000(4.5)
G-mean								
$cifar10_4$	0.0000(6.5)	0.6811(4)	**0.7423(1)**	0.7336(2)	0.7174(3)	0.0000(6.5)	0.0000(6.5)	0.0000(6.5)
$cifar10_2$	0.8685(7)	0.9179(4)	0.9193(2)	**0.9209(1)**	0.9185(3)	0.8736(6)	0.7155(8)	0.9045(5)
$cifar100_{3R88}$	0.9095(4.5)	0.9150(2.5)	0.8849(8)	0.8949(6)	0.8947(7)	0.9095(4.5)	0.9150(2.5)	**0.9250(1)**
$cifar100_{10R17}$	0.7792(6.5)	0.8005(3)	0.7915(4)	0.8183(2)	**0.8332(1)**	0.7792(6.5)	0.7893(5)	0.7755(8)
$flower102_{1R3}$	0.0000(4.5)	0.0000(4.5)	0.0000(4.5)	0.0000(4.5)	0.0000(4.5)	0.0000(4.5)	0.0000(4.5)	0.0000(4.5)
$flower102_{8R71}$	0.0000(4.5)	0.0000(4.5)	0.0000(4.5)	0.0000(4.5)	0.0000(4.5)	0.0000(4.5)	0.0000(4.5)	0.0000(4.5)

The best rank for each row is highlighted as bold.

TABLE S21: Time consuming on image datasets (seconds)

Dataset	Ori	B-SMOTE2	ADASYN	MWMOTE	INOS	SWIM	GDO	DBO
$cifar10_4$	0.00	17.26	18.18	44.63	9691.46	0.68	224.08	285.61
$cifar10_2$	0.00	1667.69	18.01	38.40	2843.06	0.71	223.94	216.36
$cifar100_{3R88}$	0.00	0.19	0.04	0.08	0.56	0.03	0.49	1.00
$cifar100_{10R17}$	0.00	0.19	0.02	0.04	0.56	0.02	0.48	1.42
$flower102_{1R3}$	0.00	0.00	0.01	0.00	0.03	0.01	0.01	0.09
$flower102_{8R71}$	0.00	0.00	0.00	0.00	0.02	0.01	0.01	0.06