Supplementary Material: CTW: Confident Time-Warping for Time-Series Label-Noise Learning

Peitian Ma[1] , **Zhen Liu**[1] , **Junhao Zheng**[1] , **Linghao Wang**[1] and **Qianli Ma**[1,2*]
[1]School of Computer Science and Engineering,
South China University of Technology, Guangzhou, China
[2]Key Laboratory of Big Data and Intelligent Robot
(South China University of Technology), Ministry of Education
ma_scuter@163.com, sezhenliu@mail.scut.edu.cn, junhaozheng47@outlook.com, wlhsama@gmail.com,
qianlima@scut.edu.cn

A Supplementary Experiments Details

Our experiments are conducted on Pytorch 1.11.0 platform with NVIDIA GeForce RTX 3090 GPU and NVIDIA GeForce GTX 1080Ti GPU. The details of 13 benchmark datasets are shown in Table 1. We pad missing values with zero and preprocess each time series with Z-score.

We set $stride = 2$, $padding = 2$, $filters = [128, 128, 256, 256]$ of four convolutional blocks in the encoder. In all the experiments, we train for 300 epochs in total with the Adam optimizer with weight decay of 10^{-4}, momentum parameters set to $\beta_1 = 0.9$ and $\beta_2 = 0.999$, and eps of 10^{-4}, the initial learning rate is set to 10^{-3}. The learning rate is halved every 60 epochs. The batch size is set to: $min(\frac{1}{10} \times dataset\ size, 128)$.

Considering the different quantity of datasets, we set different warm-up epochs for them: if the number of the training set is lower than 1000, set the warm-up epochs to 30; else if the number is lower than 3000, set the warm-up epochs to 15; otherwise, the warm-up epochs is 10. We set $\Delta_{start} = 30$ and $\Delta_{end} = 90$ in SREA [Castellani *et al.*, 2021] and leave other settings unchanged. In Sel-CL [Li *et al.*, 2022], we set k_val as the median of the number of samples in each class. uns_queue_k is set to one fifth of the sample size and $queue\_per\_class$ is set to $uns\_queue\_k\ /\ the\ number\ of\ classes$. For the other hyperparameters, we adopt the default settings, as do the other methods.

B Full Results of Comparison with the State-Of-The-Arts

For a fair comparison, all experiments use the same structure mentioned in the paper. Comparative methods are shown as follows:

- **Vanilla**: A method that does not adopt any technology of label-noise learning.
- **Co-teaching** [Han *et al.*, 2018]: The method trains two networks to select clean samples for each other.
- **Mixup-BMM** [Arazo *et al.*, 2019]: Mixup-BMM fits a beta mixture model to the loss distribution and combines bootstrapping with mixup to implement loss correction.
- **SIGUA** [Han *et al.*, 2020]: SIGUA adopts gradient descent on good data and gradient ascent on bad data.
- **DivideMix** [Li *et al.*, 2020]: DivideMix selects clean samples to form a labeled set and noisy samples to form an unlabeled set, then trains the model on both of them in a semi-supervised manner.
- **SREA** [Castellani *et al.*, 2021]: SREA aims to correct mislabeled samples in a self-supervised fashion.
- **Sel-CL** [Li *et al.*, 2022]: Sel-CL selects confident pairs out of noisy ones for supervised contrastive learning without knowing noise rates.

Table 1: Statistic of 13 benchmark datasets from UCR and UEA repositories

Datasets	#Class	#Count	#Dimension	#Length	Type
ArrowHead	3	211	1	251	Image
CBF	3	930	1	128	Simulated
FaceFour	4	112	1	350	Image
MelbournePedestrian	10	3650	1	24	Traffic
OSULeaf	6	442	1	427	Image
Plane	7	210	1	144	Sensor
Symbols	6	1020	1	398	Image
Trace	4	200	1	275	Sensor
Epilepsy	4	275	3	207	HAR
NATOPS	6	360	24	51	HAR
EthanolConcentration	4	524	3	1751	Other
FaceDetection	2	9414	144	62	-
FingerMovements	2	416	28	50	EEG

For Co-teaching and SIGUA, the noise rate is provided as needed. We describe the relative hyperparameters of baselines in Section A of the supplementary material.

We conduct experiments on benchmark datasets dealing with different symmetric, asymmetric and instance-dependent noise, comparing CTW with baselines. The full results (the average of weighted F1-score) are shown in Table 4 - Table 6. When dealing with 45% symmetric noise, the average rank (Av_Rank) of CTW is worse than Co-teaching because of its poor performance on five UEA datasets. However, CTW outperforms other methods with other noise. Additionally, Table 7 shows Co-teaching obtains the best results on clean datasets.

To be more convincing, we conduct experiments on 128 UCR datasets with 30% symmetric noise and 40% asymmet-

*Qianli Ma is the corresponding author.

ric noise. The full results are shown in Table 8-9. CTW outperforms other methods beyond half of 128 UCR datasets with both two kinds of noise.

DivideMix generally performs better than Co-teaching on image datasets [Li *et al.*, 2020], while fails on time series datasets in our paper. We find that DivideMix discarding noisy samples (*i.e.*, DivideMix w/o Semi-supervised) gets better results, as shown in Tabel 2. For more clarity, we plot precision curves about confident set and co-guessing (Figure 1). As co-guessing generates a soft label, we take the label with the highest probability as the pseudo-label for calculating precisions. As we can see, co-guessing for noisy samples brings amounts of extra mislabeled samples into the training set, which degrades the generalization performance.

Table 2: DivideMix vs DivideMix w/o Semi-supervised. The best results are in **bold**.

Methods	Sym 30%	Sym 60%	Asym 20%	Asym 40%
Dividemix	0.42(0.067)	0.345(0.078)	0.414(0.059)	0.422(0.08)
Dividemix w/o Semi-supervised	**0.722(0.056)**	**0.512(0.073)**	**0.73(0.054)**	**0.615(0.068)**

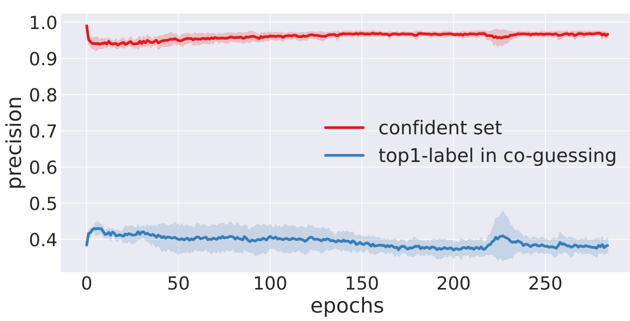


Figure 1: Precision about confident set and co-guessing on MelbournePedestrian with 30% symmetric noise. The red line means the precision.

C Augmentation Works for Expanding the Distribution of Confident Instances

We show full results (the average of weighted F1-score) of AAug (Augment on all data) and CAug (Augment on confident data) here. Both models are equipped with multiple augmentation methods: Time-Warping (TimeWarp), Add Gaussian noise (GaussNoise), Convolve, Drift, Oversample, Crop[1], and Manifold Mixup (MF_Mixup) [Verma *et al.*, 2019]. The settings of the above methods are shown in Table 3.

Table 3: The settings of augmentation methods. (*) is implemented by *tsaug* mentioned in section C.

Augmentation	Settings
TimeWarp*	n_speed_change = 5, max_speed_ratio = 3.
GaussNoise*	scale=0.015
Convolve*	window='flattop', size=10
Drift*	max_drift=0.2, n_drift_points=5
Crop*	size = origin_len×2/3, resize = origin_len
Oversample	Keep unchanged
Manifold Mixup	beta(0.2, 0.2). For simplified, we implement mixup on embeddings space.

Table 12-13 show that Time-Warping obtains the best result in both models. In addition, AAug performs worse than CAug on the overall benchmark datasets as fewer results are underlined in Table 13. Figure 4 shows that CAug helps the model learn more compact representations.

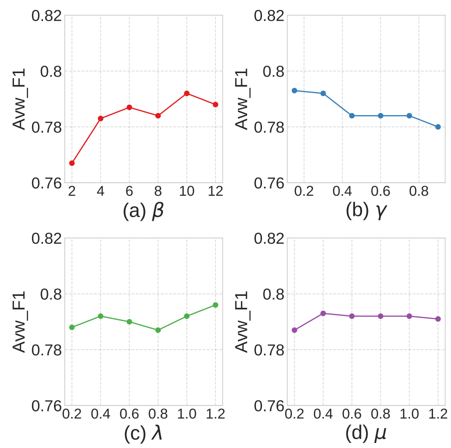


Figure 2: Hyperparameter analysis on benchmark datasets with 30% symmetric noise. β: the number of historical losses. γ: the proportion of the losses at the current epoch. λ: the coefficient of $\mathcal{L}_{aug}$. μ: the coefficient of $\mathcal{L}_{rec}$.

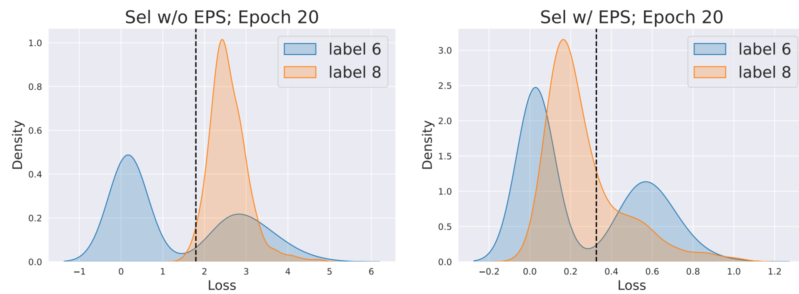
D Sample Selection Analysis

Table 15-16 show the full results (the average of weighted F1-score) of the sample selection analysis in the experimental section of the main text. CTW does not perform as well as the method which uses GMM to fit loss class by class [Huang *et al.*, 2022] at 60% symmetric noise but outperforms other methods at 40% asymmetric noise. The method that selects the same proportion R of instances in each class [Karim *et al.*, 2022] performs worse with both two kinds of noise rates. It should be noted that we only compare the sample selection practices of the above two works. In addition, CTW* (selects confident instances from the whole dataset instead of each batch) has the best result in Av_Rank with both noises.

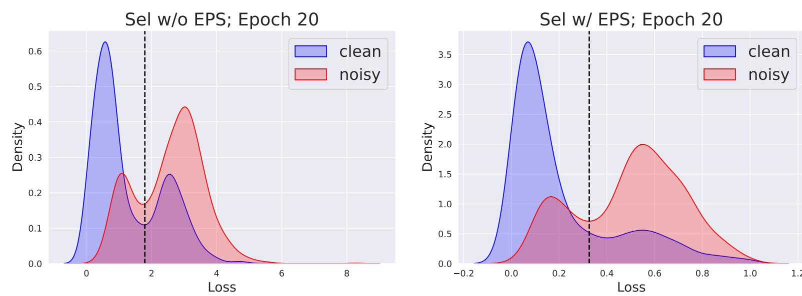
E Ablation Analysis

Table 17 shows the full results (the average of weighted F1-score) of ablation analysis on benchmark datasets with 30% symmetric noise. Additionally, we reconstruct all augmented samples with different augmentation methods on benchmark datasets (Table 14). Comparing the results with Table 12, it can be found that the results of TimeWarp, GaussNoise, Convolve, and Crop without reconstruction are better in Avw_F1.

[1]All from tsaug: https://github.com/odnura/tsaug



(a) The loss distribution of samples from two classes in MelbournePedestrian



(b) The loss distribution of samples from all classes in MelbournePedestrian

Figure 3: These experiments are based on one encoder with 40% asymmetric noise. *EPS*: Eliminate the preference of selection. *Sel w/o EPS*: Select confident samples without normalizing the training loss of each class separately. *Sel w/ EPS*: Select confident samples without normalizing the training loss of each class separately. The dashed line indicates the threshold for selecting samples, and we can find that Sel w/o EPS almost does not learn samples of label 8, while the other method is able to learn most of the samples of label 8. At epoch 20, the confident set selected by Sel w/o EPS includes 79.3% clean data, while the other method includes 82.2% clean data.

F Hyperparameter Analysis

The dataset obtained in the real scenario contains noisy labels, so there is no clean validation set for selecting the hyperparameters. However, we perform a hyperparameter analysis to show how the model is affected by the variation of the hyperparameter.

We keep default settings: $\beta = 10$, $\gamma = 0.3$, $\lambda = 1$, $\mu = 1$, and change one hyperparameter each time for analysis. The experiments are conducted on benchmark datasets with 30% symmetric noise. Figure 2 shows that CTW with bigger β and smaller γ performs better because more historical loss information can help the model make more accurate judgments when selecting clean samples. λ and μ are insensitive in Avw_F1. Complete data are shown in Table 18-21.

G Augmentation Analysis without Noisy Labels

In this section, we compare the classification ability of Time-Warping with other augmentation methods (Gaussian Noise, Convolve, Drift, Oversample, Crop and Manifold Mixup) on 128 UCR datasets without noisy labels. We adopt the backbone mentioned in our paper and don't use other tricks (e.g., sample selection and reconstruction). Figure 5 shows Time-Warping performs better than other augmentation methods on clean datasets. It supports us in adopting Time-Warping on the confident set.

H Sample Selection Preference

We conduct experiments based on one encoder mentioned in the paper with 40% asymmetric noise for discussing why we need to eliminate the preference of sample selection. We adopt the same threshold as in the paper. Figure 3(a) shows that Sel w/o EPS learns almost no samples of label 8 at epoch 20 since most of the sample loss of label 8 is located on the right side of the dashed line. However, Sel w/ EPS is able to learn most of the samples of label 8. Then the confident set selected by Sel w/o EPS includes 79.3% clean data, while the other method includes 82.2% clean data.

Table 4: The average of weighted F1-score (standard deviation) of methods on 13 benchmark datasets dealing with symmetric noise. The best results are in **bold**.

Noise	Datasets	Vanilla	SIGUA	Co-teaching	Mixup-BMM	Dividemix	Sel-CL	SREA	CTW
15%	ArrowHead	**0.886(0.041)**	0.871(0.074)	0.876(0.051)	0.817(0.047)	0.207(0.125)	0.538(0.092)	0.883(0.043)	0.877(0.043)
	CBF	0.942(0.013)	0.945(0.029)	0.952(0.017)	0.751(0.056)	0.553(0.002)	0.919(0.027)	0.982(0.006)	**0.997(0.002)**
	FaceFour	0.917(0.045)	0.875(0.074)	0.963(0.036)	0.937(0.045)	0.421(0.199)	0.647(0.109)	0.956(0.027)	**0.964(0.001)**
	MelbournePedestrian	0.852(0.012)	0.888(0.019)	0.898(0.008)	0.895(0.01)	0.558(0.072)	0.693(0.009)	0.904(0.007)	**0.92(0.008)**
	OSULeaf	0.753(0.058)	0.675(0.076)	0.826(0.044)	0.703(0.076)	0.38(0.034)	0.598(0.125)	0.864(0.05)	**0.909(0.045)**
	Plane	0.869(0.045)	0.955(0.024)	0.951(0.047)	0.985(0.02)	0.924(0.152)	**0.995(0.01)**	0.91(0.032)	0.99(0.01)
	Symbols	0.872(0.031)	0.952(0.015)	0.967(0.014)	0.975(0.009)	0.763(0.019)	0.94(0.016)	**0.99(0.008)**	0.989(0.003)
	Trace	0.949(0.028)	0.954(0.07)	0.995(0.01)	0.918(0.129)	0.867(0.163)	0.948(0.104)	0.98(0.019)	**1(0)**
	Epilepsy	0.87(0.049)	0.885(0.046)	0.923(0.032)	0.881(0.045)	0.808(0.075)	0.865(0.064)	0.916(0.049)	**0.924(0.026)**
	NATOPS	0.752(0.069)	0.794(0.048)	0.827(0.048)	0.79(0.039)	0.77(0.045)	0.826(0.023)	0.828(0.034)	**0.831(0.023)**
	EthanolConcentration	0.174(0.042)	0.192(0.033)	0.191(0.04)	0.125(0.026)	0.109(0.02)	**0.231(0.034)**	0.1(0.003)	0.181(0.042)
	FaceDetection	0.619(0.021)	0.612(0.031)	**0.623(0.009)**	0.606(0.018)	0.373(0.079)	0.505(0.011)	0.568(0.021)	0.611(0.023)
	FingerMovements	0.531(0.051)	0.505(0.055)	0.521(0.052)	0.517(0.051)	0.456(0.104)	0.504(0.059)	0.544(0.02)	**0.555(0.04)**
	Avw_F1	0.768(0.039)	0.777(0.046)	0.809(0.031)	0.762(0.044)	0.553(0.084)	0.708(0.053)	0.802(0.024)	**0.827(0.02)**
	Av_Rank	5.00	4.54	3.00	5.08	7.69	5.62	3.31	**1.77**
30%	ArrowHead	0.776(0.053)	0.837(0.035)	0.842(0.073)	0.822(0.057)	0.171(0.052)	0.564(0.059)	0.829(0.061)	**0.851(0.045)**
	CBF	0.789(0.042)	0.868(0.047)	0.894(0.02)	0.687(0.058)	0.552(0.001)	0.924(0.022)	0.866(0.007)	**0.958(0.02)**
	FaceFour	0.744(0.065)	0.743(0.147)	0.901(0.045)	0.807(0.053)	0.331(0.116)	0.71(0.115)	0.909(0.06)	**0.965(0.032)**
	MelbournePedestrian	0.745(0.012)	0.85(0.012)	0.877(0.014)	**0.892(0.008)**	0.595(0.015)	0.683(0.014)	0.884(0.01)	0.88(0.006)
	OSULeaf	0.64(0.037)	0.624(0.022)	0.738(0.044)	0.686(0.05)	0.319(0.083)	0.545(0.066)	0.771(0.058)	**0.851(0.054)**
	Plane	0.794(0.091)	0.904(0.032)	0.907(0.029)	0.985(0.02)	0.852(0.213)	0.981(0.018)	0.836(0.047)	**0.995(0.01)**
	Symbols	0.782(0.025)	0.923(0.046)	0.951(0.023)	0.952(0.03)	0.775(0.009)	0.915(0.022)	**0.989(0.006)**	0.978(0.007)
	Trace	0.688(0.093)	0.758(0.085)	0.867(0.087)	0.918(0.063)	0.867(0.163)	0.845(0.109)	0.878(0.059)	**0.99(0.02)**
	Epilepsy	0.755(0.051)	0.768(0.106)	0.776(0.095)	0.754(0.101)	0.622(0.106)	0.83(0.067)	**0.865(0.068)**	0.751(0.034)
	NATOPS	0.674(0.075)	0.673(0.083)	0.714(0.024)	0.68(0.058)	0.758(0.088)	**0.811(0.041)**	0.743(0.049)	0.747(0.043)
	EthanolConcentration	0.196(0.024)	0.192(0.037)	0.185(0.026)	0.112(0.028)	0.147(0.038)	**0.237(0.024)**	0.115(0.032)	0.202(0.024)
	FaceDetection	0.554(0.014)	0.549(0.012)	**0.564(0.018)**	0.546(0.007)	0.456(0.101)	0.508(0.012)	0.523(0.022)	0.536(0.016)
	FingerMovements	0.535(0.032)	0.529(0.037)	0.519(0.039)	0.495(0.032)	0.394(0.081)	**0.544(0.053)**	0.501(0.039)	0.519(0.039)
	Avw_F1	0.667(0.047)	0.709(0.054)	0.749(0.041)	0.718(0.043)	0.526(0.082)	0.7(0.048)	0.747(0.04)	**0.786(0.027)**
	Av_Rank	5.38	4.85	3.46	4.54	7.00	4.38	3.85	**2.54**
45%	ArrowHead	0.509(0.089)	0.557(0.073)	0.711(0.054)	0.674(0.108)	0.145(0.002)	0.543(0.103)	0.629(0.068)	**0.782(0.067)**
	CBF	0.589(0.043)	0.639(0.044)	0.726(0.064)	0.576(0.07)	0.55(0.005)	**0.931(0.026)**	0.63(0.047)	0.798(0.035)
	FaceFour	0.622(0.097)	0.425(0.139)	0.645(0.138)	0.533(0.111)	0.31(0.163)	0.509(0.089)	0.731(0.157)	**0.822(0.125)**
	MelbournePedestrian	0.626(0.016)	0.806(0.05)	0.863(0.011)	**0.863(0.012)**	0.652(0.068)	0.665(0.016)	0.849(0.012)	0.823(0.012)
	OSULeaf	0.497(0.072)	0.475(0.06)	0.66(0.035)	0.565(0.077)	0.429(0.143)	0.387(0.042)	0.6(0.041)	**0.664(0.041)**
	Plane	0.633(0.075)	0.797(0.125)	0.85(0.088)	0.856(0.099)	0.738(0.143)	**0.98(0.03)**	0.715(0.042)	0.911(0.076)
	Symbols	0.661(0.054)	0.886(0.052)	0.898(0.04)	0.948(0.011)	0.759(0.019)	0.891(0.053)	**0.985(0.005)**	0.911(0.034)
	Trace	0.588(0.079)	0.694(0.127)	0.848(0.029)	0.679(0.086)	**0.867(0.163)**	0.774(0.07)	0.739(0.04)	0.86(0.092)
	Epilepsy	0.633(0.072)	0.58(0.086)	0.726(0.039)	0.66(0.021)	0.369(0.1)	**0.745(0.113)**	0.693(0.064)	0.635(0.058)
	NATOPS	0.507(0.043)	0.565(0.019)	0.587(0.072)	0.509(0.03)	**0.75(0.049)**	0.733(0.039)	0.58(0.046)	0.561(0.041)
	EthanolConcentration	0.184(0.031)	0.218(0.06)	0.211(0.048)	0.162(0.064)	0.129(0.019)	**0.241(0.02)**	0.125(0.049)	0.2(0.049)
	FaceDetection	0.511(0.012)	0.512(0.01)	**0.513(0.009)**	0.499(0.01)	0.401(0.083)	0.504(0.013)	0.51(0.014)	0.509(0.011)
	FingerMovements	0.498(0.032)	0.508(0.047)	0.506(0.032)	0.486(0.044)	0.355(0.041)	0.488(0.071)	**0.513(0.042)**	0.492(0.058)
	Avw_F1	0.543(0.055)	0.589(0.069)	0.673(0.051)	0.616(0.057)	0.496(0.077)	0.645(0.053)	0.638(0.048)	**0.69(0.054)**
	Av_Rank	6.15	4.77	**2.69**	4.85	6.46	4.08	3.85	3.15
60%	ArrowHead	0.368(0.182)	0.377(0.132)	0.3(0.086)	0.329(0.075)	0.158(0.026)	0.277(0.043)	0.392(0.122)	**0.441(0.107)**
	CBF	0.388(0.01)	0.457(0.059)	0.45(0.042)	0.395(0.021)	0.276(0.153)	**0.88(0.078)**	0.429(0.022)	0.537(0.056)
	FaceFour	0.442(0.096)	0.309(0.121)	0.351(0.104)	0.426(0.095)	0.203(0.12)	0.382(0.108)	0.49(0.111)	**0.532(0.136)**
	MelbournePedestrian	0.446(0.01)	0.791(0.012)	**0.821(0.011)**	0.819(0.008)	0.668(0.02)	0.628(0.016)	0.82(0.005)	0.713(0.01)
	OSULeaf	0.398(0.032)	0.369(0.092)	0.498(0.052)	0.511(0.059)	0.134(0.091)	0.407(0.046)	0.42(0.083)	**0.544(0.073)**
	Plane	0.445(0.076)	0.601(0.128)	0.714(0.111)	0.677(0.087)	0.542(0.148)	**0.956(0.029)**	0.53(0.06)	0.722(0.101)
	Symbols	0.451(0.034)	0.777(0.149)	0.926(0.017)	0.9(0.08)	0.724(0.025)	0.871(0.046)	**0.978(0.01)**	0.738(0.034)
	Trace	0.332(0.039)	0.282(0.033)	0.472(0.049)	0.483(0.066)	0.593(0.097)	**0.682(0.162)**	0.412(0.088)	0.559(0.097)
	Epilepsy	0.371(0.032)	0.344(0.08)	0.534(0.08)	0.427(0.045)	0.292(0.108)	**0.593(0.134)**	0.484(0.082)	0.428(0.081)
	NATOPS	0.355(0.043)	0.346(0.055)	0.467(0.082)	0.388(0.069)	**0.637(0.061)**	0.636(0.033)	0.42(0.08)	0.399(0.061)
	EthanolConcentration	0.205(0.034)	0.153(0.029)	0.195(0.062)	0.156(0.05)	0.138(0.031)	**0.259(0.047)**	0.1(0.003)	0.219(0.037)
	FaceDetection	0.473(0.013)	0.467(0.012)	0.418(0.059)	0.461(0.014)	0.389(0.067)	**0.499(0.011)**	0.498(0.007)	0.48(0.006)
	FingerMovements	0.501(0.028)	**0.503(0.03)**	0.475(0.023)	0.45(0.075)	0.347(0.031)	0.468(0.064)	0.461(0.076)	0.478(0.056)
	Avw_F1	0.398(0.048)	0.444(0.072)	0.509(0.06)	0.494(0.057)	0.392(0.075)	**0.58(0.063)**	0.495(0.058)	0.522(0.066)
	Av_Rank	5.62	5.31	3.85	4.62	6.54	3.15	4.00	**2.92**

Table 5: The average of weighted F1-score (standard deviation) of methods on 13 benchmark datasets dealing with asymmetric noise. The best results are in **bold**

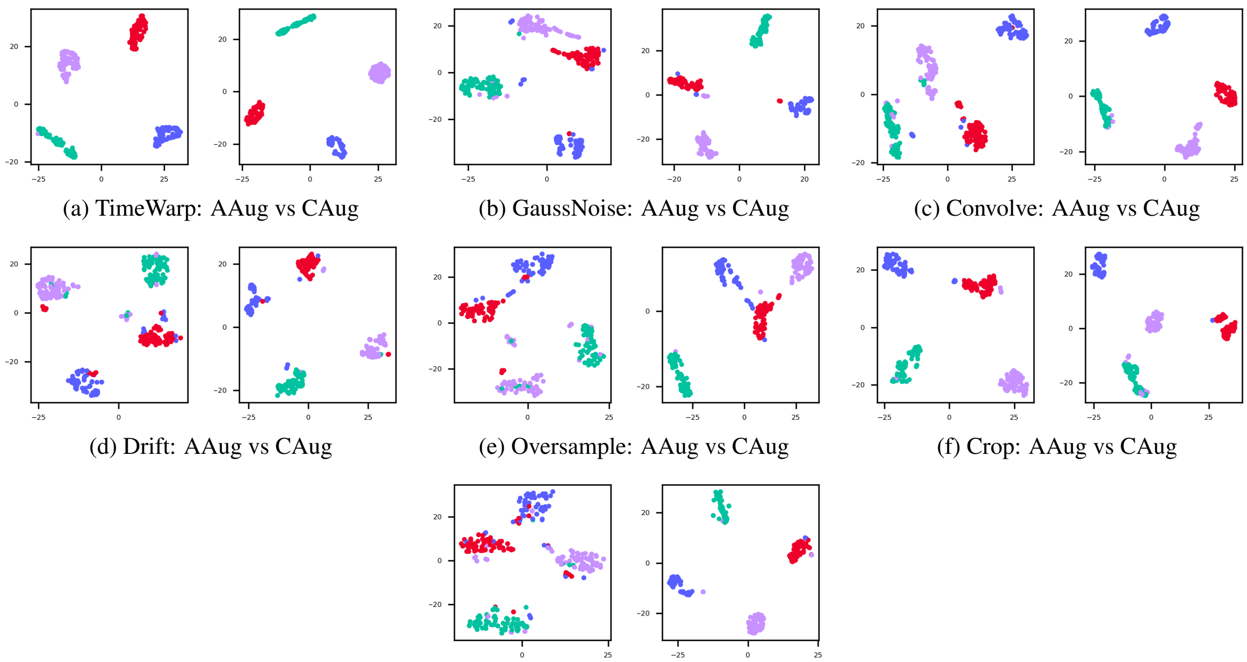
Noise	Datasets	Vanilla	SIGUA	Co-teaching	Mixup-BMM	Dividemix	Sel-CL	SREA	CTW
10%	ArrowHead	0.836(0.056)	0.843(0.051)	0.852(0.047)	0.78(0.081)	0.2(0.11)	0.53(0.078)	0.84(0.058)	**0.89(0.044)**
	CBF	0.957(0.022)	0.969(0.019)	0.989(0.006)	0.747(0.045)	0.551(0.006)	0.924(0.027)	0.986(0.011)	**0.999(0)**
	FaceFour	**0.972(0.037)**	0.914(0.074)	0.926(0.063)	0.915(0.056)	0.457(0.129)	0.743(0.067)	0.964(0.018)	0.965(0.018)
	MelbournePedestrian	0.875(0.011)	0.896(0.016)	0.91(0.009)	0.899(0.013)	0.509(0.087)	0.695(0.011)	0.918(0.007)	**0.923(0.009)**
	OSULeaf	0.818(0.019)	0.796(0.051)	0.849(0.046)	0.723(0.061)	0.388(0.083)	0.611(0.056)	0.894(0.036)	**0.944(0.015)**
	Plane	0.903(0.031)	0.962(0.012)	0.972(0.018)	**0.99(0.012)**	0.735(0.186)	**0.99(0.012)**	0.936(0.025)	**0.99(0.01)**
	Symbols	0.91(0.023)	0.954(0.017)	0.961(0.017)	0.934(0.084)	0.556(0.184)	0.924(0.018)	**0.99(0.003)**	**0.99(0.007)**
	Trace	0.98(0.01)	0.98(0.019)	0.995(0.01)	0.933(0.133)	**1(0)**	**1(0)**	0.98(0.01)	**1(0)**
	Epilepsy	0.912(0.048)	0.923(0.044)	0.927(0.047)	0.91(0.063)	0.688(0.174)	0.869(0.064)	0.913(0.038)	**0.956(0.037)**
	NATOPS	0.794(0.053)	0.79(0.064)	0.825(0.015)	0.812(0.035)	0.806(0.023)	**0.862(0.014)**	0.858(0.018)	0.848(0.013)
	EthanolConcentration	0.195(0.022)	0.172(0.035)	0.188(0.029)	0.122(0.025)	0.141(0.023)	**0.233(0.058)**	0.099(0.001)	0.194(0.015)
	FaceDetection	0.634(0.02)	0.639(0.02)	**0.648(0.019)**	0.607(0.015)	0.506(0.089)	0.513(0.009)	0.583(0.025)	0.623(0.016)
	FingerMovements	0.52(0.059)	0.498(0.031)	0.537(0.022)	0.525(0.044)	0.4(0.061)	**0.566(0.015)**	0.476(0.032)	0.545(0.033)
	Avw_F1	0.793(0.031)	0.795(0.035)	0.814(0.027)	0.761(0.051)	0.534(0.089)	0.728(0.033)	0.803(0.022)	**0.836(0.017)**
	Av_Rank	4.85	4.77	3.00	5.38	7.23	4.69	4.19	**1.88**
20%	ArrowHead	0.784(0.069)	0.809(0.093)	0.89(0.05)	0.807(0.06)	0.145(0.002)	0.53(0.098)	0.826(0.061)	**0.9(0.044)**
	CBF	0.897(0.025)	0.905(0.037)	0.952(0.017)	0.729(0.039)	0.58(0.055)	0.94(0.027)	0.936(0.014)	**0.993(0.007)**
	FaceFour	0.847(0.069)	0.832(0.05)	0.819(0.104)	0.814(0.024)	0.272(0.122)	0.623(0.089)	0.879(0.053)	**0.948(0.041)**
	MelbournePedestrian	0.801(0.013)	0.869(0.01)	0.892(0.01)	0.898(0.003)	0.616(0.118)	0.696(0.013)	0.896(0.01)	**0.898(0.013)**
	OSULeaf	0.736(0.085)	0.661(0.044)	0.786(0.034)	0.712(0.071)	0.285(0.063)	0.581(0.094)	0.826(0.042)	**0.926(0.022)**
	Plane	0.825(0.073)	0.946(0.048)	0.976(0.022)	0.981(0.028)	0.733(0.152)	0.981(0.018)	0.858(0.049)	**0.995(0.01)**
	Symbols	0.867(0.024)	0.928(0.046)	0.955(0.01)	0.973(0.007)	0.788(0.026)	0.927(0.02)	0.984(0.008)	**0.99(0.004)**
	Trace	0.847(0.041)	0.897(0.072)	0.918(0.053)	0.98(0.04)	**1(0)**	0.99(0.02)	0.917(0.069)	**1(0)**
	Epilepsy	0.827(0.113)	0.814(0.05)	0.865(0.055)	0.78(0.09)	0.635(0.076)	0.81(0.077)	0.859(0.063)	**0.931(0.027)**
	NATOPS	0.748(0.019)	0.756(0.047)	0.758(0.032)	0.772(0.043)	0.806(0.034)	**0.826(0.055)**	0.794(0.045)	0.796(0.029)
	EthanolConcentration	0.189(0.057)	0.217(0.056)	0.192(0.018)	0.121(0.039)	0.144(0.035)	**0.225(0.017)**	0.117(0.024)	0.191(0.02)
	FaceDetection	0.599(0.01)	0.604(0.015)	**0.609(0.015)**	0.575(0.011)	0.478(0.077)	0.508(0.009)	0.554(0.019)	0.578(0.01)
	FingerMovements	0.524(0.044)	0.521(0.042)	0.519(0.03)	0.513(0.038)	0.381(0.093)	**0.533(0.045)**	0.489(0.044)	0.506(0.054)
	Avw_F1	0.73(0.049)	0.751(0.047)	0.779(0.034)	0.743(0.038)	0.528(0.066)	0.705(0.045)	0.764(0.038)	**0.819(0.022)**
	Av_Rank	5.31	4.62	3.46	4.85	6.88	4.54	4.31	**2.04**
30%	ArrowHead	0.737(0.096)	0.766(0.045)	0.829(0.048)	0.756(0.089)	0.171(0.032)	0.512(0.153)	0.772(0.112)	**0.865(0.064)**
	CBF	0.814(0.045)	0.851(0.082)	0.9(0.025)	0.689(0.051)	0.551(0.003)	0.926(0.031)	0.837(0.02)	**0.926(0.013)**
	FaceFour	0.743(0.07)	0.693(0.088)	0.828(0.078)	0.727(0.061)	0.355(0.037)	0.635(0.083)	0.802(0.087)	**0.899(0.084)**
	MelbournePedestrian	0.71(0.006)	0.832(0.022)	0.88(0.009)	0.871(0.009)	0.587(0.146)	0.685(0.019)	0.864(0.01)	**0.883(0.008)**
	OSULeaf	0.644(0.072)	0.618(0.092)	0.758(0.033)	0.724(0.059)	0.271(0.124)	0.519(0.067)	0.688(0.07)	**0.82(0.056)**
	Plane	0.704(0.046)	0.868(0.036)	0.876(0.031)	0.976(0.031)	0.618(0.158)	**0.981(0.018)**	0.809(0.04)	0.918(0.037)
	Symbols	0.732(0.051)	0.914(0.036)	0.945(0.008)	0.946(0.031)	0.836(0.082)	0.93(0.022)	**0.992(0.004)**	0.978(0.004)
	Trace	0.702(0.06)	0.796(0.02)	0.833(0.03)	0.858(0.081)	0.911(0.134)	0.975(0.032)	0.838(0.044)	**0.98(0.025)**
	Epilepsy	0.761(0.104)	0.79(0.081)	0.765(0.072)	0.759(0.075)	**0.807(0.124)**	0.799(0.105)	0.787(0.123)	0.794(0.078)
	NATOPS	0.69(0.059)	0.716(0.05)	0.734(0.053)	0.699(0.069)	0.774(0.062)	**0.791(0.051)**	0.735(0.027)	0.721(0.038)
	EthanolConcentration	0.191(0.028)	0.141(0.046)	**0.203(0.041)**	0.153(0.048)	0.131(0.031)	0.201(0.027)	0.107(0.017)	0.199(0.028)
	FaceDetection	0.552(0.013)	0.535(0.015)	**0.556(0.008)**	0.534(0.01)	0.521(0.016)	0.51(0.011)	0.524(0.017)	0.553(0.012)
	FingerMovements	0.478(0.021)	0.465(0.039)	0.489(0.022)	0.487(0.017)	0.397(0.087)	0.471(0.074)	**0.504(0.05)**	0.481(0.043)
	Avw_F1	0.651(0.052)	0.691(0.05)	0.738(0.035)	0.706(0.048)	0.533(0.08)	0.687(0.053)	0.712(0.048)	**0.771(0.038)**
	Av_Rank	5.92	5.38	3.00	4.62	6.38	4.38	4.15	**2.15**
40%	ArrowHead	0.581(0.112)	0.601(0.067)	**0.746(0.063)**	0.662(0.12)	0.183(0.033)	0.388(0.214)	0.661(0.066)	0.741(0.063)
	CBF	0.637(0.064)	0.702(0.122)	0.717(0.028)	0.612(0.031)	0.446(0.124)	**0.937(0.017)**	0.667(0.029)	0.767(0.018)
	FaceFour	0.564(0.091)	0.488(0.129)	0.569(0.098)	0.591(0.062)	0.414(0.107)	0.559(0.081)	0.583(0.081)	**0.754(0.114)**
	MelbournePedestrian	0.595(0.014)	0.757(0.055)	**0.842(0.011)**	0.826(0.021)	0.683(0.021)	0.682(0.017)	0.81(0.012)	0.796(0.011)
	OSULeaf	0.502(0.055)	0.549(0.126)	0.598(0.04)	0.605(0.061)	0.182(0.122)	0.449(0.038)	0.599(0.07)	**0.722(0.031)**
	Plane	0.583(0.043)	0.723(0.119)	0.764(0.033)	0.815(0.074)	0.442(0.149)	0.77(0.093)	0.631(0.051)	**0.906(0.022)**
	Symbols	0.612(0.042)	0.801(0.072)	0.928(0.031)	0.814(0.194)	0.873(0.114)	0.905(0.047)	**0.985(0.008)**	0.911(0.039)
	Trace	0.588(0.046)	0.694(0.075)	0.676(0.047)	0.657(0.03)	0.827(0.15)	0.764(0.132)	0.631(0.055)	**0.95(0.036)**
	Epilepsy	0.585(0.074)	0.563(0.071)	0.594(0.071)	0.598(0.079)	0.603(0.148)	**0.666(0.099)**	0.615(0.079)	0.612(0.051)
	NATOPS	0.633(0.066)	0.553(0.058)	0.61(0.036)	0.593(0.06)	**0.752(0.064)**	0.74(0.039)	0.619(0.029)	0.655(0.037)
	EthanolConcentration	0.207(0.025)	0.166(0.018)	0.202(0.027)	0.154(0.045)	0.14(0.034)	**0.226(0.036)**	0.113(0.029)	0.183(0.016)
	FaceDetection	**0.515(0.012)**	0.505(0.006)	0.513(0.011)	0.502(0.004)	0.432(0.082)	0.51(0.019)	0.513(0.009)	0.505(0.019)
	FingerMovements	0.503(0.045)	0.496(0.072)	0.501(0.088)	**0.516(0.017)**	0.385(0.065)	0.505(0.074)	0.507(0.038)	0.495(0.069)
	Avw_F1	0.547(0.053)	0.584(0.076)	0.635(0.045)	0.611(0.061)	0.489(0.093)	0.623(0.07)	0.61(0.043)	**0.692(0.041)**
	Av_Rank	5.54	5.77	3.54	4.31	6.23	3.77	4.00	**2.85**

Table 6: The average of weighted F1-score (standard deviation) of methods on 13 benchmark datasets dealing with instance-dependent noise. The best results are in **bold**.

Noise	Datasets	Vanilla	SIGUA	Co-teaching	Mixup-BMM	Dividemix	Sel-CL	SREA	CTW
30%	ArrowHead	0.7(0.033)	0.642(0.078)	0.808(0.062)	0.744(0.034)	0.171(0.032)	0.507(0.07)	0.747(0.073)	**0.85(0.092)**
	CBF	0.78(0.023)	0.832(0.063)	0.875(0.023)	0.717(0.061)	0.553(0.003)	0.922(0.02)	0.847(0.015)	**0.937(0.026)**
	FaceFour	0.67(0.091)	0.596(0.15)	0.75(0.131)	0.677(0.126)	0.353(0.217)	0.601(0.12)	0.804(0.079)	**0.873(0.078)**
	MelbournePedestrian	0.743(0.009)	0.856(0.014)	**0.885(0.004)**	0.882(0.012)	0.58(0.137)	0.684(0.015)	0.868(0.006)	0.875(0.013)
	OSULeaf	0.668(0.066)	0.63(0.055)	0.718(0.087)	0.636(0.119)	0.29(0.09)	0.518(0.08)	0.757(0.055)	**0.826(0.016)**
	Plane	0.65(0.035)	0.74(0.084)	0.841(0.042)	0.845(0.094)	0.821(0.188)	**0.995(0.01)**	0.741(0.068)	0.931(0.041)
	Symbols	0.724(0.04)	0.898(0.038)	0.922(0.018)	0.973(0.01)	0.774(0.062)	0.908(0.04)	**0.983(0.008)**	0.966(0.017)
	Trace	0.828(0.043)	0.783(0.072)	0.782(0.068)	0.828(0.09)	0.933(0.133)	**1(0)**	0.791(0.056)	0.934(0.022)
	Epilepsy	0.72(0.058)	0.731(0.083)	0.757(0.089)	0.685(0.091)	0.695(0.107)	**0.822(0.093)**	0.764(0.051)	0.692(0.12)
	NATOPS	0.659(0.026)	0.657(0.044)	0.708(0.034)	0.684(0.03)	**0.82(0.017)**	0.785(0.06)	0.716(0.014)	0.695(0.044)
	EthanolConcentration	0.195(0.042)	0.139(0.019)	0.195(0.041)	0.124(0.034)	0.119(0.029)	**0.233(0.014)**	0.103(0.014)	0.211(0.037)
	FaceDetection	0.546(0.016)	0.541(0.016)	**0.555(0.019)**	0.541(0.015)	0.525(0.013)	0.501(0.013)	0.53(0.006)	0.543(0.014)
	FingerMovements	0.511(0.021)	0.494(0.072)	**0.584(0.053)**	0.524(0.017)	0.348(0.02)	0.427(0.028)	0.546(0.034)	0.519(0.038)
	Avw_F1	0.646(0.039)	0.657(0.061)	0.722(0.052)	0.681(0.056)	0.537(0.081)	0.685(0.043)	0.708(0.037)	**0.758(0.043)**
	Av_Rank	5.31	5.85	3.15	4.54	6.46	4.23	3.77	**2.69**
40%	ArrowHead	0.597(0.072)	0.578(0.124)	0.724(0.063)	0.725(0.045)	0.171(0.032)	0.445(0.122)	0.729(0.045)	**0.748(0.077)**
	CBF	0.697(0.023)	0.767(0.046)	0.822(0.024)	0.646(0.052)	0.553(0.003)	**0.927(0.026)**	0.745(0.046)	0.836(0.037)
	FaceFour	0.537(0.072)	0.455(0.116)	0.601(0.098)	0.557(0.031)	0.245(0.081)	0.577(0.058)	0.707(0.045)	**0.82(0.088)**
	MelbournePedestrian	0.64(0.027)	0.838(0.009)	**0.858(0.011)**	0.847(0.016)	0.711(0.018)	0.672(0.012)	0.834(0.011)	0.81(0.034)
	OSULeaf	0.543(0.027)	0.576(0.061)	0.706(0.039)	0.498(0.119)	0.147(0.061)	0.481(0.108)	0.665(0.067)	**0.753(0.064)**
	Plane	0.523(0.08)	0.627(0.09)	0.652(0.101)	0.671(0.133)	0.769(0.218)	**0.928(0.097)**	0.6(0.117)	0.765(0.116)
	Symbols	0.618(0.043)	0.806(0.036)	0.901(0.026)	0.957(0.019)	0.672(0.161)	0.906(0.035)	**0.984(0.008)**	0.789(0.068)
	Trace	0.656(0.071)	0.685(0.065)	0.776(0.075)	0.665(0.101)	0.738(0.356)	**1(0)**	0.69(0.088)	0.818(0.092)
	Epilepsy	0.576(0.107)	0.586(0.086)	0.608(0.106)	0.622(0.06)	0.649(0.206)	**0.748(0.08)**	0.7(0.066)	0.657(0.076)
	NATOPS	0.575(0.056)	0.643(0.028)	0.631(0.056)	0.592(0.057)	**0.768(0.062)**	0.717(0.048)	0.626(0.017)	0.591(0.041)
	EthanolConcentration	0.186(0.03)	0.198(0.038)	0.188(0.03)	0.136(0.042)	0.146(0.035)	**0.217(0.036)**	0.1(0.003)	0.213(0.03)
	FaceDetection	**0.521(0.02)**	0.516(0.015)	0.518(0.009)	0.516(0.011)	0.398(0.079)	0.493(0.011)	0.514(0.012)	0.513(0.012)
	FingerMovements	0.482(0.023)	0.484(0.039)	0.501(0.021)	0.518(0.051)	0.382(0.048)	0.46(0.043)	**0.52(0.019)**	0.485(0.035)
	Avw_F1	0.55(0.05)	0.597(0.058)	0.653(0.051)	0.611(0.057)	0.488(0.104)	0.659(0.052)	0.647(0.042)	**0.677(0.059)**
	Av_Rank	6.31	4.77	3.38	4.62	6.00	3.77	3.85	**3.31**

Table 7: The average of weighted F1-score (standard deviation) of methods on 13 benchmark datasets without noisy labels. The best results are in **bold**.

Datasets	Vanilla	SIGUA	Co-teaching	Mixup-BMM	Dividemix	Sel-CL	SREA	CTW
ArrowHead	0.913(0.038)	0.899(0.036)	0.898(0.033)	0.771(0.042)	0.209(0.128)	0.61(0.063)	**0.913(0.044)**	0.87(0.029)
CBF	**1(0)**	**1(0)**	**1(0)**	0.771(0.039)	0.417(0.151)	0.923(0.037)	**1(0)**	**1(0)**
FaceFour	0.982(0.022)	0.964(0.034)	0.973(0.037)	0.945(0.034)	0.337(0.041)	0.704(0.095)	**0.991(0.028)**	0.955(0.018)
MelbournePedestrian	0.92(0.009)	0.925(0.006)	0.93(0.009)	0.846(0.037)	0.41(0.039)	0.692(0.01)	**0.933(0.011)**	0.898(0.011)
OSULeaf	0.915(0.04)	0.874(0.061)	0.916(0.031)	0.708(0.064)	0.382(0.091)	0.681(0.083)	**0.957(0.032)**	0.953(0.014)
Plane	0.99(0.019)	0.99(0.012)	**0.995(0.01)**	0.913(0.089)	0.531(0.178)	0.99(0.02)	0.985(0.01)	**0.995(0.02)**
Symbols	0.977(0.005)	0.975(0.01)	0.983(0.009)	0.769(0.002)	0.4(0.068)	0.927(0.028)	**0.993(0.005)**	0.988(0.005)
Trace	**1(0)**	**1(0)**	**1(0)**	**1(0)**	0.391(0.115)	**1(0)**	**1(0)**	**1(0)**
Epilepsy	0.96(0.018)	0.971(0.019)	0.978(0.021)	0.967(0.027)	0.531(0.057)	0.905(0.034)	0.978(0.02)	**0.982(0.021)**
NATOPS	0.851(0.021)	0.864(0.021)	0.876(0.017)	0.856(0.011)	0.638(0.153)	**0.893(0.029)**	0.88(0.021)	0.887(0.014)
EthanolConcentration	0.202(0.03)	**0.235(0.039)**	0.213(0.043)	0.122(0.032)	0.122(0.022)	0.235(0.025)	0.099(0.053)	0.195(0.001)
FaceDetection	0.669(0.013)	0.667(0.021)	**0.683(0.019)**	0.65(0.01)	0.373(0.079)	0.529(0)	0.62(0.013)	0.62(0.025)
FingerMovements	0.532(0.025)	0.513(0.035)	0.521(0.035)	**0.553(0.027)**	0.492(0.092)	0.53(0.061)	0.53(0.032)	0.535(0.006)
Avw_F1	0.839(0.018)	0.837(0.023)	**0.844(0.02)**	0.759(0.032)	0.402(0.093)	0.74(0.04)	0.837(0.02)	0.837(0.013)
Av_Rank	3.73	3.88	**3.12**	5.46	7.92	5.38	3.23	3.27

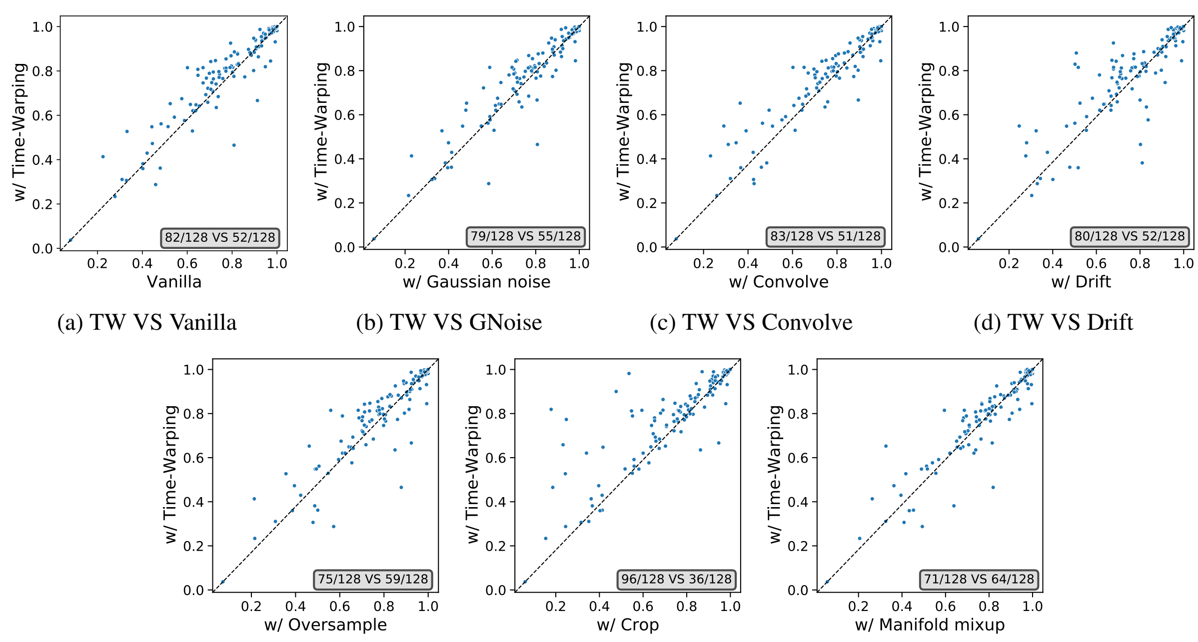


(a) TimeWarp: AAug vs CAug (b) GaussNoise: AAug vs CAug (c) Convolve: AAug vs CAug

(d) Drift: AAug vs CAug (e) Oversample: AAug vs CAug (f) Crop: AAug vs CAug

(g) MF_Mixup: AAug vs CAug

Figure 4: The t-SNE visualizations on the Trace dataset. *AAug*: Augment all instances. *CAug*: Augment confident instances. Augmenting on confident instances helps model learn more compact representations from each class.



(a) TW VS Vanilla (b) TW VS GNoise (c) TW VS Convolve (d) TW VS Drift

(e) TW VS Oversample (f) TW VS Crop (g) TW VS MF_Mixup

Figure 5: Comparation of augmentation methods on 128 UCR datasets without noisy labels. In each subplot, the vertical coordinate of each blue point indicates the weighted F1 score obtained by TW on the dataset, and the horizontal coordinate indicates the weighted F1 score obtained by the other method. The gray box like "x/128 VS y/128" means Time-Warping performs better on x datasets among 128 UCR datasets, and another method performs better on y datasets (if the difference between the F1 scores of the two methods on the dataset is less than 1^{-4}, then we consider them to have the same result. It means that $x + y$ may be bigger than 128). *TW*: Time-Warping. *GNoise*: Gaussian Noise. *MF_Mixup*: Manifold Mixup. It can be found that Time-Warping performs better than other augmentation methods.

Table 8: Comparison of methods on 128 UCR datasets, dealing with 30% symmetric noise. The average of weighted F1-score is reported. The best results are in **bold**.

Datasets	Vanilla	SIGUA	Co-teaching	Mixup-BMM	DivideMix	Sel-CL	SREA	CTW
ACSF1	0.545(0.056)	0.609(0.076)	**0.663(0.082)**	0.311(0.089)	0.195(0.028)	0.185(0.056)	0.629(0.075)	0.351(0.073)
Adiac	0.586(0.062)	0.631(0.065)	**0.645(0.068)**	0.209(0.013)	0.081(0.019)	0.13(0.034)	0.292(0.071)	0.534(0.089)
AllGestureWiimoteX	0.512(0.03)	0.595(0.02)	0.591(0.038)	0.543(0.014)	0.585(0.033)	0.613(0.021)	0.568(0.015)	**0.714(0.024)**
AllGestureWiimoteY	0.506(0.056)	0.639(0.042)	0.621(0.033)	0.56(0.039)	0.625(0.066)	0.61(0.02)	0.604(0.024)	**0.753(0.036)**
AllGestureWiimoteZ	0.403(0.027)	0.491(0.027)	0.476(0.046)	0.418(0.04)	0.551(0.015)	0.431(0.029)	0.485(0.014)	**0.657(0.028)**
ArrowHead	0.779(0.047)	0.773(0.105)	**0.869(0.039)**	0.741(0.1)	0.162(0.031)	0.573(0.037)	0.813(0.029)	**0.869(0.048)**
BME	0.841(0.056)	0.898(0.048)	0.948(0.058)	0.863(0.086)	0.973(0.043)	**1(0)**	0.889(0.049)	0.965(0.011)
Beef	0.514(0.144)	0.513(0.131)	**0.542(0.18)**	0.186(0.037)	0.246(0.111)	0.372(0.087)	0.519(0.091)	0.537(0.19)
BeetleFly	0.402(0.107)	0.413(0.16)	0.536(0.231)	0.458(0.183)	0.333(0)	0.554(0.204)	0.518(0.193)	**0.669(0.107)**
BirdChicken	0.531(0.198)	0.69(0.159)	**0.715(0.12)**	0.557(0.246)	0.333(0)	0.476(0.111)	0.458(0.183)	0.669(0.124)
CBF	0.789(0.036)	0.844(0.048)	0.921(0.003)	0.721(0.021)	0.55(0.003)	0.931(0.024)	0.884(0.029)	**0.965(0.013)**
Car	0.569(0.128)	0.348(0.131)	0.63(0.202)	0.321(0.049)	0.1(0)	0.363(0.024)	**0.774(0.017)**	0.745(0.227)
Chinatown	0.767(0.026)	0.862(0.084)	0.908(0.022)	0.941(0.025)	0.899(0.147)	0.913(0.048)	**0.949(0.018)**	0.941(0.022)
ChlorineConcentration	0.844(0.019)	0.922(0.005)	**0.932(0.006)**	0.657(0.051)	0.397(0.023)	0.371(0.032)	0.859(0.029)	0.807(0.027)
CinCECGtorso	0.678(0.056)	0.847(0.05)	0.903(0.031)	0.813(0.03)	0.385(0.219)	0.504(0.085)	**0.954(0.024)**	0.928(0.029)
Coffee	0.848(0.138)	0.925(0.15)	0.982(0.036)	0.877(0.246)	0.314(0.04)	0.763(0.263)	0.908(0.082)	**1(0.043)**
Computers	0.568(0.042)	0.563(0.05)	0.583(0.028)	0.5(0.059)	0.333(0)	0.626(0.038)	0.583(0.069)	**0.702(0.038)**
CricketX	0.487(0.028)	0.47(0.03)	0.547(0.043)	0.531(0.047)	0.388(0.099)	0.554(0.04)	0.574(0.027)	**0.667(0.018)**
CricketY	0.463(0.038)	0.446(0.055)	0.533(0.009)	0.494(0.047)	0.596(0.068)	0.522(0.022)	0.572(0.015)	**0.671(0.014)**
CricketZ	0.457(0.053)	0.432(0.043)	0.49(0.023)	0.467(0.027)	0.444(0.104)	0.557(0.024)	0.562(0.017)	**0.678(0.025)**
Crop	0.695(0.08)	0.744(0.003)	**0.764(0.007)**	0.738(0.006)	0.63(0.005)	0.587(0.005)	0.753(0.005)	0.761(0.006)
DiatomSizeReduction	0.953(0.023)	0.956(0.049)	0.974(0.013)	0.981(0.016)	0.145(0.005)	0.509(0.092)	**0.987(0.012)**	0.922(0.097)
DistalPhalanxOutlineAgeGroup	0.699(0.06)	0.728(0.063)	0.741(0.077)	0.741(0.106)	0.743(0.116)	0.779(0.136)	0.768(0.102)	**0.803(0.127)**
DistalPhalanxOutlineCorrect	0.647(0.05)	0.707(0.05)	0.714(0.072)	0.486(0.157)	0.439(0.06)	0.382(0.122)	0.705(0.057)	**0.752(0.071)**
DistalPhalanxTW	0.645(0.011)	0.685(0.034)	0.705(0.015)	0.658(0.035)	0.597(0.006)	0.703(0.025)	0.715(0.023)	**0.722(0.009)**
DodgerLoopDay	0.23(0.07)	0.171(0.055)	0.312(0.049)	0.266(0.075)	0.218(0.062)	0.272(0.067)	0.344(0.027)	**0.433(0.064)**
DodgerLoopGame	0.706(0.071)	0.643(0.053)	0.718(0.041)	0.704(0.072)	0.341(0.024)	**0.865(0.068)**	0.753(0.072)	0.831(0.045)
DodgerLoopWeekend	0.695(0.086)	0.715(0.052)	0.738(0.053)	0.797(0.105)	0.588(0.015)	0.736(0.225)	0.828(0.073)	**0.89(0.075)**
ECG200	0.684(0.102)	0.701(0.119)	0.757(0.135)	0.8(0.037)	0.531(0.016)	0.765(0.058)	0.732(0.072)	**0.807(0.043)**
ECG5000	0.886(0.008)	0.914(0.016)	0.926(0.023)	0.92(0.012)	0.915(0.021)	0.823(0.103)	0.874(0.026)	**0.929(0.02)**
ECGFiveDays	0.762(0.072)	0.883(0.04)	0.907(0.014)	**1(0)**	0.333(0.003)	0.61(0.152)	0.985(0.014)	0.962(0.021)
EOGHorizontalSignal	0.433(0.074)	0.482(0.087)	0.501(0.067)	0.349(0.072)	0.382(0.136)	0.505(0.028)	**0.557(0.065)**	0.53(0.075)
EOGVerticalSignal	0.331(0.077)	0.404(0.059)	0.436(0.062)	0.368(0.051)	0.361(0.132)	0.486(0.059)	**0.493(0.061)**	0.467(0.033)
Earthquakes	0.686(0.017)	0.674(0.023)	0.691(0.023)	0.68(0.032)	0.709(0.007)	0.498(0.106)	0.611(0.057)	**0.73(0.053)**
ElectricDevices	0.637(0.036)	0.691(0.035)	0.724(0.041)	0.74(0.056)	0.569(0.034)	0.648(0.077)	**0.761(0.043)**	0.759(0.064)
EthanolLevel	0.155(0.031)	0.13(0.033)	0.16(0.053)	0.101(0.002)	0.101(0.002)	0.141(0.052)	0.1(0.001)	**0.172(0.017)**
FaceAll	0.747(0.038)	0.876(0.038)	0.901(0.026)	**0.947(0.018)**	0.919(0.023)	0.682(0.102)	0.91(0.029)	0.931(0.023)
FaceFour	0.772(0.069)	0.63(0.09)	0.875(0.053)	0.817(0.079)	0.361(0.06)	0.7(0.099)	**0.921(0.061)**	0.896(0.022)
FacesUCR	0.747(0.012)	0.871(0.026)	0.919(0.02)	**0.951(0.009)**	0.928(0.021)	0.738(0.027)	0.925(0.017)	0.929(0.012)
FiftyWords	0.422(0.034)	0.477(0.038)	0.539(0.021)	0.338(0.021)	0.317(0.05)	0.424(0.018)	0.57(0.015)	**0.584(0.015)**
Fish	0.631(0.044)	0.76(0.081)	0.839(0.051)	0.344(0.099)	0.11(0.053)	0.12(0.033)	**0.923(0.043)**	0.877(0.016)
FordA	0.669(0.058)	0.778(0.079)	0.822(0.017)	**0.875(0.014)**	0.349(0.001)	0.544(0.033)	0.735(0.013)	0.804(0.083)
FordB	0.669(0.043)	0.757(0.076)	0.818(0.063)	**0.891(0.051)**	0.343(0)	0.478(0.046)	0.711(0.039)	0.763(0.08)
FreezerRegularTrain	0.825(0.066)	0.918(0.02)	0.949(0.023)	0.789(0.178)	0.593(0.099)	0.768(0.011)	**0.994(0.004)**	0.993(0.014)
FreezerSmallTrain	0.814(0.073)	0.961(0.024)	0.961(0.015)	0.793(0.148)	0.668(0.127)	0.877(0.024)	**0.995(0.005)**	0.994(0.004)
Fungi	0.755(0.056)	0.9(0.017)	0.909(0.038)	0.936(0.026)	0.636(0.167)	0.803(0.105)	**0.99(0.021)**	0.899(0.02)
GestureMidAirD1	0.434(0.057)	0.456(0.092)	0.499(0.032)	0.387(0.052)	0.495(0.052)	0.511(0.062)	0.506(0.054)	**0.546(0.089)**
GestureMidAirD2	0.422(0.021)	0.418(0.036)	0.449(0.075)	0.29(0.057)	0.439(0.027)	0.397(0.083)	0.421(0.056)	**0.462(0.054)**
GestureMidAirD3	0.24(0.023)	0.238(0.075)	0.238(0.064)	0.211(0.049)	0.231(0.1)	0.157(0.04)	0.246(0.053)	**0.258(0.075)**
GesturePebbleZ1	0.589(0.038)	0.642(0.081)	0.697(0.043)	0.673(0.093)	0.585(0.127)	**0.744(0.06)**	0.675(0.095)	0.724(0.048)
GesturePebbleZ2	0.542(0.068)	0.522(0.038)	0.654(0.122)	0.633(0.061)	0.482(0.196)	**0.761(0.041)**	0.613(0.027)	0.691(0.061)
GunPoint	0.718(0.072)	0.868(0.062)	0.95(0.023)	0.908(0.089)	0.601(0.068)	0.632(0.169)	0.9(0.032)	**0.97(0.012)**
GunPointAgeSpan	0.729(0.048)	0.831(0.062)	0.886(0.057)	0.881(0.044)	0.332(0.008)	0.809(0.08)	0.918(0.032)	**0.964(0.019)**
GunPointMaleVersusFemale	0.716(0.023)	0.898(0.058)	0.953(0.022)	0.912(0.065)	0.362(0.005)	**0.978(0.016)**	0.915(0.071)	0.971(0.025)
GunPointOldVersusYoung	0.737(0.061)	0.795(0.025)	0.875(0.022)	0.783(0.071)	0.457(0.142)	0.528(0.026)	0.898(0.04)	**0.929(0.044)**
Ham	0.68(0.068)	0.702(0.075)	0.692(0.04)	**0.71(0.045)**	0.334(0.023)	0.579(0.101)	0.65(0.05)	0.669(0.062)
HandOutlines	0.329(0.16)	0.263(0.155)	**0.544(0.194)**	0.192(0)	0.192(0)	0.186(0.005)	0.498(0)	0.473(0.101)
Haptics	0.294(0.068)	**0.351(0.046)**	0.324(0.073)	0.333(0.046)	0.096(0.028)	0.194(0.049)	0.253(0.046)	0.302(0.087)
Herring	0.506(0.093)	0.478(0.043)	0.517(0.081)	0.452(0.017)	0.452(0.017)	0.498(0.052)	0.452(0.017)	**0.534(0.033)**
HouseTwenty	0.773(0.048)	0.835(0.076)	0.912(0.1)	0.753(0.122)	0.297(0.054)	0.854(0.052)	0.886(0.066)	**0.918(0.025)**
InlineSkate	0.337(0.094)	0.43(0.086)	0.476(0.087)	0.256(0.061)	0.113(0.037)	0.238(0.025)	**0.488(0.026)**	0.355(0.089)
InsectEPGRegularTrain	0.791(0.033)	0.779(0.073)	0.899(0.038)	0.883(0.064)	0.38(0.234)	**0.963(0.034)**	0.843(0.022)	0.922(0.045)
InsectEPGSmallTrain	0.751(0.091)	0.833(0.075)	0.846(0.054)	0.855(0.108)	0.188(0.002)	**0.931(0.041)**	0.833(0.035)	0.889(0.054)
InsectWingbeatSound	0.458(0.021)	0.552(0.018)	0.556(0.025)	**0.624(0.019)**	0.408(0.013)	0.341(0.016)	0.579(0.021)	0.603(0.03)
ItalyPowerDemand	0.729(0.018)	0.891(0.061)	0.918(0.023)	**0.965(0.007)**	0.334(0.002)	0.817(0.056)	0.957(0.009)	0.96(0.015)
LargeKitchenAppliances	0.654(0.045)	0.746(0.057)	0.761(0.067)	0.601(0.037)	0.231(0.129)	0.762(0.081)	0.697(0.028)	**0.853(0.043)**
Lightning2	0.626(0.072)	0.617(0.064)	0.653(0.047)	0.656(0.048)	0.45(0.019)	0.551(0.082)	0.653(0.109)	**0.803(0.084)**
Lightning7	0.639(0.087)	0.58(0.088)	0.667(0.031)	0.602(0.042)	0.446(0.119)	**0.787(0.084)**	0.663(0.09)	0.759(0.094)
Mallat	0.566(0.142)	0.836(0.045)	0.829(0.098)	0.965(0.009)	0.459(0.082)	0.949(0.007)	**0.976(0.003)**	0.858(0.033)
Meat	0.767(0.13)	0.595(0.09)	**0.898(0.058)**	0.169(0.004)	0.167(0)	0.167(0)	0.746(0.138)	0.846(0.166)
MedicalImages	0.626(0.029)	0.666(0.048)	0.693(0.025)	0.555(0.051)	0.446(0.115)	0.552(0.016)	0.636(0.035)	**0.73(0.025)**
MelbournePedestrian	0.75(0.012)	0.864(0.005)	0.875(0.009)	0.889(0.01)	0.59(0.032)	0.676(0.017)	0.888(0.005)	**0.894(0.011)**
MiddlePhalanxOutlineAgeGroup	0.613(0.066)	0.688(0.065)	0.691(0.072)	0.719(0.126)	0.67(0.099)	**0.732(0.103)**	0.686(0.115)	0.724(0.106)
MiddlePhalanxOutlineCorrect	0.643(0.041)	**0.706(0.063)**	0.663(0.089)	0.423(0.109)	0.477(0.003)	0.402(0.166)	0.528(0.059)	**0.706(0.112)**
MiddlePhalanxTW	0.48(0.043)	0.534(0.016)	0.537(0.018)	0.482(0.04)	0.478(0.004)	**0.562(0.023)**	0.464(0.028)	0.541(0.033)
MixedShapes	0.614(0.096)	0.836(0.017)	0.887(0.041)	0.45(0.085)	0.37(0.084)	0.88(0.015)	**0.97(0.01)**	0.935(0.036)
MixedShapesSmallTrain	0.726(0.041)	0.849(0.046)	0.912(0.014)	0.541(0.097)	0.542(0.244)	0.852(0.023)	**0.954(0.011)**	0.916(0.033)
MoteStrain	0.724(0.034)	0.819(0.026)	0.863(0.009)	0.898(0.017)	0.291(0.001)	0.854(0.027)	0.85(0.02)	**0.903(0.01)**
NonInvasiveFetalECGThorax1	0.57(0.036)	0.731(0.022)	0.777(0.031)	0.82(0.023)	0.641(0.014)	0.089(0.013)	**0.832(0.017)**	0.7(0.022)
NonInvasiveFetalECGThorax2	0.625(0.025)	0.833(0.021)	**0.861(0.012)**	0.791(0.015)	0.743(0.032)	0.204(0.047)	0.847(0.026)	0.816(0.043)
OSULeaf	0.697(0.03)	0.675(0.036)	0.766(0.039)	0.624(0.09)	0.348(0.049)	0.544(0.1)	0.728(0.05)	**0.846(0.056)**
OliveOil	**0.478(0.153)**	0.248(0.006)	0.245(0)	0.185(0.055)	0.229(0.031)	0.159(0.106)	0.216(0.058)	0.433(0.065)
PLAID	0.56(0.023)	0.639(0.022)	0.693(0.014)	0.588(0.046)	0.537(0.102)	0.67(0.031)	**0.764(0.026)**	0.759(0.026)
PhalangesOutlinesCorrect	0.649(0.036)	**0.695(0.046)**	0.678(0.095)	0.553(0.094)	0.477(0.041)	0.503(0.117)	0.585(0.071)	0.591(0.1)
Phoneme	0.134(0.031)	0.17(0.019)	0.165(0.035)	0.215(0.031)	0.234(0.058)	0.231(0.014)	0.238(0.022)	**0.24(0.021)**
PickupGestureWiimoteZ	0.417(0.166)	0.367(0.071)	0.423(0.077)	0.43(0.103)	0.445(0.145)	0.681(0.07)	0.559(0.076)	**0.718(0.13)**
PigAirwayPressure	**0.047(0.015)**	0.027(0.019)	0.024(0.019)	0.003(0.005)	0.04(0.028)	0.036(0.019)	0.028(0.018)	0.031(0.022)
PigArtPressure	0.229(0.054)	0.17(0.036)	0.287(0.082)	0.148(0.059)	0.124(0.049)	0.157(0.026)	**0.697(0.016)**	0.163(0.034)
PigCVP	0.16(0.046)	0.171(0.056)	0.165(0.058)	0.022(0.022)	0.04(0.013)	0.109(0.02)	**0.304(0.059)**	0.282(0.048)
Plane	0.793(0.059)	0.895(0.026)	0.951(0.035)	0.981(0.018)	0.805(0.217)	**0.981(0.018)**	0.89(0.019)	0.963(0)
PowerCons	0.724(0.055)	0.724(0.053)	0.724(0.045)	0.721(0.057)	0.333(0)	**0.807(0.044)**	0.739(0.016)	0.761(0.035)
ProximalPhalanxOutlineAgeGroup	0.675(0.059)	0.797(0.055)	0.791(0.044)	0.815(0.063)	0.725(0.062)	**0.822(0.089)**	0.8(0.108)	0.821(0.058)
ProximalPhalanxOutlineCorrect	0.683(0.037)	0.769(0.065)	**0.777(0.029)**	0.55(0.243)	0.549(0.002)	0.582(0.259)	0.454(0.158)	0.575(0.246)
ProximalPhalanxTW	0.686(0.021)	0.78(0.027)	0.778(0.017)	0.744(0.015)	0.532(0.064)	0.758(0.016)	0.754(0.032)	**0.781(0.019)**
RefrigerationDevices	0.424(0.046)	0.41(0.042)	0.453(0.075)	0.376(0.066)	0.167(0)	0.471(0.103)	0.474(0.072)	**0.481(0.076)**
Rock	0.32(0.193)	0.208(0.14)	0.164(0.085)	0.37(0.093)	0.188(0.12)	0.595(0.126)	**0.741(0.12)**	0.385(0.213)
ScreenType	0.438(0.043)	0.406(0.038)	0.432(0.025)	0.394(0.036)	0.167(0)	0.415(0.05)	0.425(0.016)	**0.466(0.086)**
SemgHandGenderCh2	0.624(0.024)	0.649(0.043)	0.648(0.041)	0.609(0.044)	0.45(0)	0.466(0.134)	0.623(0.028)	**0.686(0.096)**
SemgHandMovementCh2	0.265(0.046)	0.317(0.071)	0.285(0.021)	0.311(0.087)	0.078(0.061)	0.343(0.046)	0.347(0.038)	**0.441(0.081)**
SemgHandSubjectCh2	0.306(0.094)	0.272(0.146)	0.296(0.069)	0.438(0.075)	0.067(0)	0.321(0.06)	**0.602(0.067)**	0.41(0.077)
ShakeGestureWiimoteZ	0.615(0.076)	0.589(0.143)	0.648(0.085)	0.658(0.068)	0.745(0.105)	**0.858(0.046)**	0.657(0.07)	0.803(0.06)
ShapeletSim	0.899(0.054)	0.863(0.083)	0.914(0.026)	0.778(0.062)	0.3(0.067)	**1(0)**	0.769(0.07)	0.995(0.01)
ShapesAll	0.447(0.057)	0.682(0.055)	0.633(0.053)	0.289(0.035)	0.442(0.125)	0.363(0.032)	**0.748(0.029)**	0.616(0.035)
SmallKitchenAppliances	0.577(0.048)	0.562(0.05)	0.593(0.04)	0.563(0.038)	0.167(0)	0.615(0.058)	0.612(0.053)	**0.695(0.062)**
SmoothSubspace	0.745(0.068)	0.78(0.029)	0.838(0.054)	0.866(0.053)	0.761(0.251)	**0.97(0.012)**	0.783(0.053)	0.897(0.026)
SonyAIBORobotSurface1	0.815(0.042)	0.884(0.035)	0.923(0.033)	**0.97(0.022)**	0.404(0.004)	0.968(0.017)	0.807(0.025)	0.953(0.02)
SonyAIBORobotSurface2	0.764(0.02)	0.882(0.026)	0.891(0.024)	**0.946(0.02)**	0.617(0.12)	0.804(0.014)	0.863(0.029)	0.887(0.018)
StarlightCurves	0.749(0.03)	0.722(0.06)	0.701(0.145)	0.955(0.02)	0.327(0.003)	0.884(0.028)	**0.962(0.01)**	0.723(0.116)
Strawberry	0.803(0.056)	0.9(0.053)	**0.927(0.043)**	0.858(0.058)	0.424(0.122)	0.616(0.076)	0.842(0.045)	0.876(0.043)
SwedishLeaf	0.721(0.01)	0.875(0.028)	0.854(0.016)	0.864(0.017)	0.716(0.046)	0.71(0.056)	0.875(0.005)	**0.905(0.015)**
Symbols	0.806(0.058)	0.942(0.014)	0.947(0.023)	0.965(0.026)	0.757(0.03)	0.913(0.019)	**0.986(0.004)**	0.968(0.005)
SyntheticControl	0.796(0.051)	0.876(0.037)	0.89(0.039)	0.857(0.038)	**0.979(0.018)**	0.967(0.019)	0.88(0.02)	0.929(0.028)
ToeSegmentation1	0.755(0.092)	0.734(0.079)	0.825(0.073)	0.765(0.048)	0.601(0.241)	**0.879(0.074)**	0.805(0.032)	0.858(0.021)
ToeSegmentation2	0.795(0.074)	0.695(0.034)	0.805(0.077)	0.809(0.091)	0.634(0.024)	0.74(0.064)	0.748(0.056)	**0.853(0.065)**
Trace	0.736(0.062)	0.745(0.084)	0.898(0.033)	0.929(0.06)	0.875(0.25)	0.891(0.124)	0.889(0.03)	**0.96(0.01)**
TwoLeadECG	0.795(0.013)	0.96(0.018)	0.968(0.015)	0.993(0.005)	0.415(0.068)	0.601(0.043)	**0.999(0.002)**	0.996(0.003)
TwoPatterns	0.864(0.024)	0.924(0.012)	0.974(0.01)	0.983(0.017)	**1(0)**	**1(0)**	0.894(0.013)	0.983(0.004)
UMD	0.727(0.11)	0.649(0.15)	0.784(0.144)	0.829(0.066)	0.609(0.107)	**0.978(0.021)**	0.745(0.081)	0.899(0.097)
UWaveGestureLibraryAll	0.554(0.043)	0.581(0.111)	0.729(0.044)	0.854(0.022)	0.737(0.047)	0.747(0.012)	0.817(0.02)	**0.897(0.036)**
UWaveGestureLibraryX	0.564(0.031)	0.644(0.023)	0.695(0.026)	0.74(0.021)	0.673(0.041)	0.735(0.021)	0.77(0.011)	**0.771(0.026)**
UWaveGestureLibraryY	0.518(0.021)	0.572(0.026)	0.632(0.013)	0.644(0.023)	0.544(0.029)	0.588(0.008)	0.655(0.014)	**0.708(0.025)**
UWaveGestureLibraryZ	0.534(0.013)	0.609(0.039)	0.643(0.008)	0.701(0.014)	0.584(0.017)	0.666(0.007)	0.724(0.013)	**0.725(0.015)**
Wafer	0.818(0.004)	0.932(0.009)	0.953(0.006)	0.95(0.009)	0.843(0.001)	0.883(0.024)	**0.971(0.012)**	0.826(0.007)
Wine	0.395(0.148)	0.349(0.021)	0.349(0.021)	0.328(0.024)	0.328(0.024)	0.335(0.027)	0.349(0.021)	**0.527(0.113)**
WordSynonyms	0.492(0.038)	0.523(0.032)	0.552(0.031)	0.351(0.051)	0.291(0.022)	0.373(0.041)	0.551(0.055)	**0.619(0.032)**
Worms	0.515(0.072)	0.454(0.042)	0.5(0.056)	0.543(0.078)	0.297(0.066)	0.435(0.034)	0.523(0.054)	**0.565(0.044)**
WormsTwoClass	0.639(0.069)	**0.655(0.078)**	0.617(0.109)	0.59(0.069)	0.397(0.145)	0.542(0.056)	0.627(0.092)	0.642(0.04)
Yoga	0.69(0.019)	0.727(0.036)	0.825(0.02)	0.528(0.057)	0.342(0.04)	0.56(0.022)	0.825(0.023)	**0.827(0.044)**
Avw_F1(Av_std)	0.608(0.057)	0.652(0.055)	0.696(0.05)	0.63(0.055)	0.449(0.058)	0.601(0.056)	0.7(0.043)	**0.733(0.052)**
#Best	2	4	12	10	2	19	27	**55**
p-value	6.7E-32	4.5E-16	4.5E-07	1.8E-14	1.7E-31	9.2E-16	5.0E-04	-

Table 9: Comparison of methods on 128 UCR datasets, dealing with 40% asymmetric noise. The average of weighted F1-score is reported. The best results are in **bold**.

Datasets	Vanilla	SIGUA	Co-teaching	Mixup-BMM	DivideMix	Sel-CL	SREA	CTW
ACSF1	0.395(0.07)	0.428(0.082)	0.414(0.073)	0.267(0.081)	0.141(0.054)	0.195(0.045)	**0.498(0.02)**	0.343(0.056)
Adiac	0.432(0.033)	0.442(0.031)	**0.442(0.027)**	0.151(0.027)	0.07(0.028)	0.114(0.036)	0.218(0.026)	0.421(0.086)
AllGestureWiimoteX	0.409(0.044)	0.495(0.021)	0.482(0.028)	0.422(0.044)	0.504(0.085)	0.516(0.027)	0.457(0.023)	**0.542(0.025)**
AllGestureWiimoteY	0.444(0.026)	0.501(0.024)	0.501(0.035)	0.433(0.05)	0.529(0.048)	0.553(0.053)	0.485(0.027)	**0.616(0.029)**
AllGestureWiimoteZ	0.371(0.024)	0.403(0.018)	0.395(0.023)	0.357(0.035)	0.397(0.065)	0.392(0.073)	0.445(0.039)	**0.523(0.042)**
ArrowHead	0.526(0.049)	0.614(0.025)	0.74(0.071)	0.653(0.108)	0.171(0.032)	0.389(0.098)	0.61(0.185)	**0.778(0.059)**
BME	0.594(0.065)	0.63(0.137)	0.764(0.081)	0.795(0.131)	0.784(0.14)	**0.977(0.175)**	0.655(0.046)	0.887(0.076)
Beef	0.302(0.112)	0.264(0.117)	0.307(0.092)	0.211(0.112)	0.093(0.091)	0.259(0.101)	**0.399(0.128)**	0.375(0.09)
BeetleFly	0.511(0.198)	0.488(0.177)	0.367(0.101)	0.594(0.306)	0.333(0)	0.619(0.16)	0.61(0.218)	**0.657(0.083)**
BirdChicken	0.399(0.209)	0.307(0.053)	0.413(0.16)	0.467(0.223)	0.333(0)	**0.645(0.16)**	0.432(0.169)	0.502(0.22)
CBF	0.615(0.031)	0.712(0.047)	0.742(0.027)	0.62(0.031)	0.482(0.091)	**0.938(0)**	0.645(0.02)	0.786(0.034)
Car	0.376(0.074)	0.273(0.133)	0.393(0.154)	0.218(0.097)	0.1(0)	0.229(0.069)	0.444(0.073)	**0.544(0.194)**
Chinatown	0.645(0.03)	0.794(0.099)	**0.895(0.02)**	0.886(0.046)	0.737(0.177)	0.892(0.146)	0.893(0.06)	0.865(0.055)
ChlorineConcentration	0.671(0.022)	0.674(0.044)	0.699(0.022)	0.561(0.053)	0.257(0.143)	0.391(0.018)	**0.756(0.023)**	0.641(0.029)
CinCECGtorso	0.561(0.047)	0.606(0.042)	0.79(0.036)	0.634(0.038)	0.181(0.161)	0.403(0.046)	**0.878(0.046)**	0.805(0.022)
Coffee	0.549(0.071)	0.418(0.169)	0.554(0.298)	0.355(0.041)	0.334(0.045)	0.43(0.02)	0.368(0.204)	**0.747(0.08)**
Computers	0.529(0.069)	0.561(0.043)	0.564(0.09)	0.564(0.024)	0.333(0)	0.614(0.117)	0.566(0.055)	**0.638(0.047)**
CricketX	0.382(0.03)	0.335(0.02)	0.369(0.006)	0.362(0.065)	0.265(0.037)	0.525(0.031)	0.452(0.043)	**0.541(0.033)**
CricketY	0.381(0.033)	0.352(0.05)	0.412(0.017)	0.365(0.052)	0.172(0.024)	0.467(0.06)	0.452(0.04)	**0.573(0.019)**
CricketZ	0.376(0.014)	0.349(0.049)	0.394(0.031)	0.329(0.031)	0.265(0.087)	0.485(0.086)	0.407(0.047)	**0.497(0.016)**
Crop	0.566(0.016)	0.613(0.016)	0.691(0.007)	0.692(0.008)	0.595(0.016)	0.574(0.012)	**0.698(0.008)**	0.644(0.003)
DiatomSizeReduction	0.726(0.039)	0.885(0.117)	0.959(0.034)	0.885(0.161)	0.13(0.011)	0.279(0.124)	**0.984(0.038)**	0.947(0.01)
DistalPhalanxOutlineAgeGroup	0.579(0.032)	0.675(0.061)	0.658(0.026)	0.656(0.122)	0.485(0.077)	0.675(0.098)	**0.75(0.252)**	0.719(0.05)
DistalPhalanxOutlineCorrect	0.594(0.064)	0.65(0.063)	**0.689(0.085)**	0.353(0.111)	0.415(0.113)	0.471(0.103)	0.643(0.226)	0.449(0.05)
DistalPhalanxTW	0.572(0.04)	0.634(0.071)	**0.689(0.028)**	0.608(0.044)	0.599(0.006)	0.678(0.013)	0.675(0.02)	0.611(0.02)
DodgerLoopDay	0.183(0.074)	0.171(0.048)	0.271(0.135)	0.252(0.048)	0.16(0.042)	0.256(0.109)	0.351(0.07)	**0.361(0.049)**
DodgerLoopGame	0.564(0.058)	0.551(0.07)	0.553(0.088)	0.588(0.089)	0.334(0.025)	0.591(0.027)	0.586(0.124)	**0.618(0.051)**
DodgerLoopWeekend	0.637(0.058)	0.715(0.109)	0.618(0.034)	0.613(0.115)	0.588(0.015)	0.743(0.203)	0.613(0.24)	**0.818(0.163)**
ECG200	0.605(0.114)	0.636(0.051)	0.691(0.09)	0.628(0.138)	0.546(0.038)	**0.747(0.056)**	0.579(0.064)	0.698(0.077)
ECG5000	0.71(0.028)	0.784(0.037)	0.889(0.01)	0.685(0.141)	**0.89(0.024)**	0.801(0.019)	0.71(0.09)	0.721(0.017)
ECGFiveDays	0.637(0.02)	0.816(0.02)	0.865(0.025)	**0.926(0.025)**	0.333(0.003)	0.572(0.003)	0.924(0.114)	0.878(0.059)
EOGHorizontalSignal	0.356(0.062)	0.344(0.094)	0.442(0.111)	0.336(0.08)	0.314(0.091)	0.42(0.043)	**0.47(0.045)**	0.405(0.088)
EOGVerticalSignal	0.328(0.046)	0.263(0.028)	0.356(0.038)	0.291(0.054)	0.275(0.084)	0.402(0.045)	**0.432(0.057)**	0.383(0.024)
Earthquakes	0.587(0.078)	0.565(0.068)	0.631(0.065)	0.586(0.104)	**0.709(0.007)**	0.403(0.242)	0.573(0.085)	0.703(0.029)
ElectricDevices	0.508(0.046)	0.588(0.048)	0.671(0.041)	0.655(0.049)	0.553(0.039)	0.65(0.063)	0.67(0.077)	**0.679(0.056)**
EthanolLevel	0.21(0.156)	0.128(0.052)	0.179(0.092)	0.1(0.001)	0.1(0.001)	0.164(0.002)	0.101(0.055)	**0.253(0.002)**
FaceAll	0.628(0.018)	0.762(0.043)	0.802(0.014)	0.821(0.028)	**0.835(0.064)**	0.594(0.075)	0.748(0.055)	0.822(0.042)
FaceFour	0.576(0.17)	0.547(0.081)	0.596(0.075)	0.583(0.127)	0.341(0.048)	0.511(0.122)	0.636(0.091)	**0.772(0.122)**
FacesUCR	0.607(0.027)	0.74(0.013)	0.805(0.008)	0.823(0.044)	**0.893(0.024)**	0.71(0.043)	0.756(0.051)	0.827(0.015)
FiftyWords	0.355(0.022)	0.38(0.031)	0.42(0.049)	0.241(0.061)	0.273(0.034)	0.311(0.014)	0.475(0.036)	**0.484(0.059)**
Fish	0.535(0.091)	0.504(0.122)	0.68(0.06)	0.384(0.103)	0.036(0)	0.13(0.063)	0.734(0.073)	**0.793(0.014)**
FordA	0.582(0.024)	0.669(0.108)	**0.791(0.026)**	0.697(0.057)	0.349(0.001)	0.508(0.031)	0.649(0.043)	0.673(0.021)
FordB	0.59(0.063)	0.632(0.038)	0.721(0.052)	**0.816(0.063)**	0.343(0)	0.501(0.091)	0.603(0.065)	0.579(0.021)
FreezerRegularTrain	0.74(0.024)	0.918(0.023)	0.896(0.033)	0.748(0.077)	0.752(0.056)	0.759(0.236)	0.916(0.012)	**0.921(0.064)**
FreezerSmallTrain	0.672(0.025)	0.874(0.041)	0.882(0.028)	0.622(0.044)	0.716(0.089)	0.821(0.139)	**0.899(0.014)**	0.877(0.082)
Fungi	0.646(0.07)	0.685(0.057)	0.683(0.045)	0.675(0.057)	0.424(0.141)	0.505(0.025)	0.724(0.119)	**0.796(0.056)**
GestureMidAirD1	0.345(0.057)	0.338(0.048)	0.357(0.031)	0.287(0.044)	0.271(0.086)	0.333(0.035)	0.347(0.06)	**0.435(0.037)**
GestureMidAirD2	0.34(0.043)	0.36(0.079)	0.344(0.059)	0.217(0.042)	0.287(0.064)	0.327(0.029)	0.386(0.072)	**0.411(0.044)**
GestureMidAirD3	0.185(0.052)	0.144(0.043)	0.162(0.05)	0.185(0.053)	0.139(0.056)	0.132(0.054)	**0.209(0.016)**	0.192(0.034)
GesturePebbleZ1	0.451(0.066)	0.479(0.069)	0.528(0.062)	**0.552(0.103)**	0.367(0.102)	0.483(0.102)	0.492(0.11)	0.52(0.097)
GesturePebbleZ2	0.45(0.07)	0.466(0.118)	0.493(0.108)	0.491(0.122)	0.342(0.078)	0.453(0.18)	0.498(0.092)	**0.559(0.111)**
GunPoint	0.649(0.069)	0.711(0.04)	0.791(0.06)	0.754(0.136)	0.385(0.103)	0.486(0.195)	0.784(0.069)	**0.854(0.058)**
GunPointAgeSpan	0.628(0.06)	0.715(0.087)	0.74(0.071)	0.73(0.134)	0.385(0.118)	0.692(0.222)	0.741(0.145)	**0.862(0.056)**
GunPointMaleVersusFemale	0.66(0.065)	0.781(0.101)	0.804(0.094)	0.814(0.077)	0.482(0.237)	**0.964(0.299)**	0.841(0.045)	0.874(0.044)
GunPointOldVersusYoung	0.607(0.066)	0.698(0.043)	0.775(0.057)	0.597(0.122)	0.451(0.114)	0.519(0.068)	0.817(0.077)	**0.838(0.076)**
Ham	**0.557(0.117)**	0.502(0.089)	0.434(0.065)	0.526(0.06)	0.323(0.021)	0.442(0.106)	0.553(0.101)	0.526(0.041)
HandOutlines	0.4(0.138)	**0.526(0.057)**	0.503(0.007)	0.191(0.001)	0.375(0.15)	0.19(0.15)	0.498(0.003)	0.364(0)
Haptics	0.249(0.041)	0.269(0.047)	0.262(0.037)	0.24(0.031)	0.126(0.087)	0.144(0.098)	0.2(0.055)	**0.29(0.064)**
Herring	**0.539(0.08)**	0.484(0.074)	0.475(0.04)	0.363(0.111)	0.408(0.091)	0.361(0.111)	0.363(0.105)	0.51(0.111)
HouseTwenty	0.502(0.066)	0.555(0.123)	0.737(0.068)	0.492(0.082)	0.269(0.006)	**0.748(0.233)**	0.599(0.085)	0.673(0.083)
InlineSkate	0.379(0.034)	0.355(0.03)	0.362(0.023)	0.297(0.054)	0.147(0.048)	0.253(0.044)	**0.433(0.027)**	0.354(0.028)
InsectEPGRegularTrain	0.672(0.078)	0.751(0.067)	0.828(0.057)	0.637(0.137)	0.545(0.256)	**0.946(0.103)**	0.689(0.023)	0.844(0.07)
InsectEPGSmallTrain	0.633(0.094)	0.654(0.092)	0.743(0.06)	0.613(0.143)	0.376(0.212)	**0.904(0.162)**	0.669(0.058)	0.806(0.076)
InsectWingbeatSound	0.357(0.031)	0.444(0.037)	0.47(0.018)	**0.553(0.02)**	0.364(0.054)	0.331(0.05)	0.44(0.013)	0.466(0.02)
ItalyPowerDemand	0.594(0.053)	0.872(0.022)	0.88(0.024)	0.944(0.026)	0.359(0.051)	0.749(0.198)	**0.947(0.053)**	0.888(0.017)
LargeKitchenAppliances	0.538(0.079)	0.61(0.051)	0.641(0.04)	0.51(0.047)	0.268(0.138)	0.672(0.024)	0.579(0.067)	**0.778(0.033)**
Lightning2	0.505(0.121)	0.557(0.114)	0.541(0.112)	0.442(0.103)	0.423(0.106)	0.578(0.094)	0.578(0.06)	**0.623(0.074)**
Lightning7	0.368(0.081)	0.389(0.102)	0.423(0.06)	0.448(0.057)	0.382(0.21)	**0.709(0.122)**	0.536(0.124)	0.574(0.066)
Mallat	0.533(0.062)	0.777(0.119)	0.882(0.032)	0.866(0.078)	0.668(0.05)	0.948(0.054)	**0.977(0.008)**	0.794(0.01)
Meat	0.499(0.204)	0.384(0.163)	0.369(0.17)	0.167(0)	0.167(0)	0.167(0)	0.481(0)	**0.623(0.15)**
MedicalImages	0.552(0.026)	0.595(0.041)	0.587(0.031)	0.514(0.024)	0.457(0.04)	0.401(0.059)	0.597(0.034)	**0.613(0.037)**
MelbournePedestrian	0.593(0.024)	0.73(0.041)	**0.847(0.012)**	0.841(0.023)	0.653(0.018)	0.683(0.041)	0.804(0.018)	0.797(0.015)
MiddlePhalanxOutlineAgeGroup	0.519(0.06)	0.574(0.077)	0.633(0.064)	0.651(0.098)	0.497(0.175)	0.548(0.217)	0.65(0.219)	**0.662(0.1)**
MiddlePhalanxOutlineCorrect	0.571(0.057)	**0.672(0.058)**	0.663(0.093)	0.463(0.028)	0.37(0.132)	0.494(0.107)	0.477(0.135)	0.456(0.003)
MiddlePhalanxTW	0.444(0.027)	0.488(0.02)	**0.539(0.015)**	0.469(0.041)	0.479(0.004)	0.537(0.026)	0.49(0.024)	0.472(0.018)
MixedShapes	0.556(0.062)	0.751(0.083)	0.876(0.032)	0.642(0.177)	0.456(0.144)	0.859(0.084)	**0.895(0.046)**	0.782(0.016)
MixedShapesSmallTrain	0.549(0.035)	0.733(0.059)	0.834(0.047)	0.555(0.209)	0.154(0.079)	**0.847(0.179)**	0.843(0.013)	0.817(0.04)
MoteStrain	0.631(0.014)	0.785(0.045)	0.799(0.037)	**0.856(0.024)**	0.474(0.195)	0.824(0.026)	0.798(0.031)	0.81(0.016)
NonInvasiveFetalECGThorax1	0.51(0.027)	0.585(0.027)	0.633(0.03)	0.604(0.036)	0.451(0.11)	0.062(0.042)	**0.693(0.019)**	0.605(0.023)
NonInvasiveFetalECGThorax2	0.512(0.029)	0.648(0.025)	0.717(0.038)	0.596(0.061)	0.315(0.136)	0.184(0.063)	**0.744(0.036)**	0.7(0.04)
OSULeaf	0.531(0.067)	0.483(0.104)	0.589(0.053)	0.509(0.087)	0.17(0.109)	0.448(0.09)	0.604(0.063)	**0.724(0.067)**
OliveOil	**0.316(0.259)**	0.254(0.226)	0.166(0.097)	0.166(0.097)	0.15(0.089)	0.075(0.084)	0.206(0.054)	0.137(0.079)
PLAID	0.519(0.031)	0.537(0.046)	0.583(0.04)	0.494(0.031)	0.605(0.028)	0.536(0.098)	**0.749(0.05)**	0.623(0.065)
PhalangesOutlinesCorrect	0.585(0.027)	0.592(0.057)	**0.609(0.064)**	0.508(0.141)	0.498(0)	0.449(0.124)	0.491(0.132)	0.469(0.136)
Phoneme	0.121(0.021)	0.146(0.009)	0.154(0.021)	0.195(0.013)	0.166(0.022)	0.186(0.02)	0.204(0.021)	**0.233(0.018)**
PickupGestureWiimoteZ	0.315(0.133)	0.292(0.084)	0.282(0.06)	0.288(0.075)	0.351(0.159)	**0.456(0.077)**	0.377(0.041)	0.386(0.074)
PigAirwayPressure	**0.039(0.034)**	0.014(0.009)	0.013(0.01)	0.007(0.007)	0.025(0.031)	0.03(0.013)	0.031(0.015)	0.023(0.025)
PigArtPressure	0.129(0.037)	0.099(0.044)	0.159(0.058)	0.1(0.051)	0.083(0.029)	0.123(0.035)	**0.507(0.058)**	0.133(0.031)
PigCVP	0.114(0.068)	0.12(0.077)	0.112(0.029)	0.064(0.025)	0.019(0.006)	0.062(0.026)	**0.297(0.012)**	0.129(0.05)
Plane	0.627(0.038)	0.696(0.063)	0.763(0.03)	0.865(0.086)	0.402(0.11)	0.774(0.112)	0.647(0.099)	**0.894(0.065)**
PowerCons	0.615(0.026)	0.668(0.08)	0.647(0.042)	0.604(0.054)	0.333(0)	**0.755(0.049)**	0.621(0.063)	0.696(0.042)
ProximalPhalanxOutlineAgeGroup	0.572(0.044)	0.745(0.059)	0.717(0.062)	0.727(0.054)	0.629(0.173)	0.734(0.06)	0.727(0.053)	**0.784(0.059)**
ProximalPhalanxOutlineCorrect	0.59(0.035)	0.705(0.053)	**0.72(0.049)**	0.566(0.238)	0.549(0.002)	0.585(0.193)	0.456(0.236)	0.572(0.27)
ProximalPhalanxTW	0.605(0.037)	0.721(0.019)	0.721(0.036)	0.669(0.018)	0.464(0.025)	**0.747(0.082)**	0.69(0.021)	0.631(0.023)
RefrigerationDevices	0.328(0.043)	0.353(0.042)	0.395(0.056)	0.287(0.081)	0.167(0)	0.394(0.032)	**0.432(0.108)**	0.406(0.025)
Rock	0.096(0.05)	0.045(0.016)	0.105(0.044)	0.176(0.16)	0.06(0.02)	**0.45(0.085)**	0.444(0.071)	0.196(0.214)
ScreenType	0.405(0.061)	0.374(0.064)	0.373(0.062)	0.356(0.072)	0.164(0.006)	0.391(0.062)	0.431(0.081)	**0.46(0.039)**
SemgHandGenderCh2	0.549(0.026)	0.588(0.039)	**0.629(0.05)**	0.57(0.044)	0.45(0)	0.601(0.102)	0.548(0.053)	0.59(0.044)
SemgHandMovementCh2	0.228(0.02)	0.218(0.032)	0.236(0.051)	0.271(0.027)	0.077(0.039)	0.257(0.041)	0.317(0.045)	**0.347(0.022)**
SemgHandSubjectCh2	0.314(0.047)	0.347(0.042)	0.455(0.044)	0.38(0.043)	0.141(0.149)	0.329(0.067)	0.47(0.086)	**0.521(0.029)**
ShakeGestureWiimoteZ	0.419(0.068)	0.457(0.111)	0.453(0.052)	0.45(0.164)	0.47(0.124)	**0.563(0.111)**	0.511(0.132)	0.534(0.063)
ShapeletSim	0.549(0.056)	0.524(0.15)	0.596(0.052)	0.604(0.078)	0.333(0)	**1(0)**	0.551(0)	0.775(0.071)
ShapesAll	0.396(0.02)	0.444(0.038)	0.424(0.056)	0.241(0.032)	0.246(0.045)	0.257(0.058)	**0.587(0.018)**	0.464(0.073)
SmallKitchenAppliances	0.43(0.037)	0.489(0.048)	0.509(0.039)	0.479(0.039)	0.167(0)	**0.591(0.055)**	0.507(0.013)	0.568(0.03)
SmoothSubspace	0.65(0.062)	0.66(0.03)	0.673(0.069)	0.697(0.081)	**0.906(0.098)**	0.906(0.042)	0.657(0.146)	0.736(0.063)
SonyAIBORobotSurface1	0.707(0.044)	0.713(0.058)	0.827(0.049)	0.842(0.089)	0.434(0.144)	**0.966(0.233)**	0.687(0.013)	0.834(0.033)
SonyAIBORobotSurface2	0.625(0.029)	0.669(0.079)	**0.795(0.027)**	0.78(0.068)	0.503(0.066)	0.791(0.118)	0.735(0.031)	0.791(0.02)
StarlightCurves	0.499(0.13)	0.534(0.072)	0.721(0.064)	0.625(0.191)	0.316(0.026)	0.882(0.076)	**0.926(0.018)**	0.582(0.068)
Strawberry	0.659(0.086)	0.68(0.095)	**0.819(0.025)**	0.532(0.061)	0.377(0.155)	0.569(0.126)	0.455(0.103)	0.752(0.222)
SwedishLeaf	0.546(0.028)	0.663(0.031)	0.698(0.033)	0.715(0.078)	0.653(0.077)	0.637(0.029)	0.676(0.047)	**0.8(0.035)**
Symbols	0.626(0.043)	0.679(0.067)	0.915(0.038)	0.779(0.145)	0.823(0.118)	0.906(0.034)	**0.991(0.029)**	0.928(0.004)
SyntheticControl	0.632(0.051)	0.768(0.033)	0.778(0.063)	0.664(0.058)	**0.962(0.023)**	0.958(0.182)	0.681(0.028)	0.825(0.035)
ToeSegmentation1	0.589(0.095)	0.623(0.13)	0.694(0.057)	0.689(0.087)	0.346(0.022)	0.676(0.079)	0.61(0.081)	**0.722(0.049)**
ToeSegmentation2	0.64(0.032)	**0.757(0.066)**	0.685(0.066)	0.696(0.11)	0.639(0.018)	0.6(0.077)	0.654(0.12)	0.685(0.019)
Trace	0.567(0.092)	0.652(0.038)	0.714(0.083)	0.581(0.136)	0.807(0.231)	0.78(0.223)	0.657(0.077)	**0.965(0.052)**
TwoLeadECG	0.631(0.034)	0.797(0.051)	0.851(0.013)	0.822(0.09)	0.334(0.001)	0.599(0.135)	0.886(0.055)	**0.901(0.106)**
TwoPatterns	0.657(0.027)	0.855(0.023)	0.901(0.012)	0.868(0.08)	**1(0)**	**1(0)**	0.683(0)	0.876(0.012)
UMD	0.577(0.099)	0.551(0.174)	0.69(0.088)	0.676(0.121)	0.524(0.2)	**0.88(0.171)**	0.569(0.126)	0.764(0.057)
UWaveGestureLibraryAll	0.514(0.052)	0.603(0.076)	0.641(0.051)	**0.81(0.024)**	0.695(0.05)	0.758(0.029)	0.612(0.019)	0.696(0.013)
UWaveGestureLibraryX	0.46(0.04)	0.514(0.036)	0.624(0.021)	0.72(0.023)	0.684(0.024)	**0.735(0.031)**	0.672(0.02)	0.625(0.005)
UWaveGestureLibraryY	0.447(0.027)	0.475(0.03)	0.548(0.029)	0.549(0.017)	0.404(0.092)	0.568(0.032)	0.532(0.013)	**0.571(0.009)**
UWaveGestureLibraryZ	0.483(0.021)	0.549(0.028)	0.567(0.029)	0.613(0.012)	0.42(0.027)	**0.642(0.031)**	0.612(0.023)	0.631(0.017)
Wafer	0.734(0.016)	0.9(0.012)	0.895(0.011)	0.887(0.016)	0.843(0.001)	0.77(0.001)	**0.916(0.02)**	0.782(0.03)
Wine	0.349(0.021)	0.349(0.021)	0.349(0.021)	0.319(0.02)	0.328(0.024)	**0.39(0.024)**	0.339(0.113)	0.369(0.025)
WordSynonyms	0.419(0.044)	0.442(0.036)	0.485(0.034)	0.296(0.043)	0.264(0.047)	0.294(0.051)	**0.527(0.042)**	0.506(0.047)
Worms	0.466(0.059)	0.41(0.133)	**0.476(0.053)**	0.474(0.046)	0.297(0.126)	0.439(0.052)	0.453(0.053)	0.469(0.069)
WormsTwoClass	0.536(0.081)	0.489(0.12)	**0.603(0.057)**	0.589(0.064)	0.281(0.066)	0.541(0.107)	0.596(0.091)	0.589(0.044)
Yoga	0.583(0.027)	0.584(0.067)	**0.724(0.035)**	0.475(0.05)	0.358(0.032)	0.525(0.064)	0.686(0.026)	0.659(0.016)
Avw_F1(Av_std)	0.499(0.058)	0.544(0.064)	0.59(0.052)	0.53(0.074)	0.399(0.068)	0.542(0.084)	0.589(0.066)	**0.621(0.054)**
#Best	4	3	15	6	7	22	26	**45**
p-value	2.5E-27	4.8E-15	1.8E-04	1.0E-15	8.5E-26	5.6E-07	2.5E-03	-

Table 10: The average of weighted F1-score last ten epochs of methods on 13 benchmark datasets dealing with symmetric noise. The best results are in **bold**

Noise	Datasets	Vanilla	SIGUA	Co-teaching	Mixup-BMM	Dividemix	Sel-CL	SREA	CTW
15%	ArrowHead	0.878	0.850	**0.920**	0.809	0.322	0.586	0.903	0.890
	CBF	0.936	0.955	0.977	0.773	0.599	0.930	0.979	**0.998**
	FaceFour	0.929	0.894	0.969	0.931	0.435	0.680	0.972	**0.974**
	MelbournePedestrian	0.850	0.890	**0.904**	0.895	0.467	0.696	0.903	0.901
	OSULeaf	0.739	0.713	0.840	0.738	0.454	0.582	0.873	**0.924**
	Plane	0.870	0.952	0.968	0.984	0.583	0.993	0.913	**0.997**
	Symbols	0.876	0.946	0.967	0.977	0.670	0.939	0.989	**0.989**
	Trace	0.948	0.975	0.994	0.935	0.800	0.950	0.978	**1.000**
	Epilepsy	0.864	0.865	0.930	0.872	0.476	0.869	0.918	**0.936**
	NATOPS	0.756	0.797	**0.840**	0.785	0.700	0.828	0.832	0.839
	EthanolConcentration	0.225	0.222	0.221	**0.253**	0.250	0.237	0.213	0.214
	FaceDetection	0.621	0.617	**0.643**	0.607	0.505	0.504	0.569	0.589
	FingerMovements	0.519	0.520	0.520	0.533	0.504	0.502	**0.550**	0.502
	Avw_F1	0.770	0.784	0.822	0.776	0.520	0.715	0.815	**0.827**
	Av_Rank	6.462	5.923	3.462	5.077	9.077	7.192	4.077	**3.346**
30%	ArrowHead	0.787	0.823	**0.879**	0.802	0.308	0.592	0.836	0.867
	CBF	0.801	0.870	0.915	0.706	0.666	0.941	0.866	**0.952**
	FaceFour	0.811	0.734	0.929	0.813	0.457	0.721	0.936	**0.970**
	MelbournePedestrian	0.751	0.857	**0.891**	0.891	0.481	0.684	0.884	0.881
	OSULeaf	0.654	0.642	0.801	0.705	0.400	0.539	0.778	**0.833**
	Plane	0.784	0.920	0.936	0.981	0.655	0.981	0.848	**0.995**
	Symbols	0.792	0.918	0.956	0.954	0.672	0.917	**0.988**	0.982
	Trace	0.659	0.755	0.869	0.913	0.750	0.852	0.858	**0.990**
	Epilepsy	0.778	0.778	0.841	0.753	0.382	0.828	**0.870**	0.799
	NATOPS	0.688	0.699	0.771	0.699	0.687	**0.814**	0.754	0.752
	EthanolConcentration	0.237	**0.247**	0.212	0.245	0.246	0.242	0.231	0.227
	FaceDetection	0.553	0.552	**0.567**	0.555	0.509	0.508	0.525	0.539
	FingerMovements	0.531	0.504	0.533	0.503	0.529	0.537	0.509	**0.543**
	Avw_F1	0.679	0.715	0.777	0.732	0.518	0.704	0.760	**0.795**
	Av_Rank	6.923	6.000	3.615	5.154	8.846	5.846	4.615	**3.385**
45%	ArrowHead	0.510	0.582	0.720	0.708	0.327	0.587	0.625	**0.779**
	CBF	0.600	0.668	0.756	0.585	0.663	**0.935**	0.621	0.812
	FaceFour	0.587	0.472	0.680	0.598	0.345	0.541	0.735	**0.794**
	MelbournePedestrian	0.627	0.829	**0.868**	0.865	0.518	0.667	0.853	0.827
	OSULeaf	0.475	0.519	**0.674**	0.628	0.315	0.382	0.579	0.673
	Plane	0.619	0.812	0.884	0.871	0.705	**0.978**	0.707	0.921
	Symbols	0.648	0.864	0.934	0.953	0.672	0.896	**0.984**	0.905
	Trace	0.563	0.674	0.842	0.689	0.750	0.769	0.732	**0.882**
	Epilepsy	0.640	0.610	0.732	0.652	0.427	**0.752**	0.689	0.668
	NATOPS	0.509	0.580	0.648	0.516	0.689	**0.729**	0.577	0.572
	EthanolConcentration	0.229	0.231	0.233	0.255	0.253	0.245	**0.269**	0.226
	FaceDetection	0.508	0.511	**0.518**	0.512	0.502	0.504	0.511	0.511
	FingerMovements	0.495	0.517	0.510	0.486	0.515	0.494	**0.524**	0.500
	Avw_F1	0.539	0.605	0.692	0.640	0.514	0.652	0.647	**0.698**
	Av_Rank	8.308	6.538	**3.385**	5.231	7.731	5.462	4.154	4.462
60%	ArrowHead	0.391	0.367	0.366	0.407	0.308	0.358	0.384	**0.450**
	CBF	0.415	0.489	0.476	0.407	0.395	**0.887**	0.442	0.533
	FaceFour	0.453	0.359	0.443	0.443	0.285	0.406	0.477	**0.615**
	MelbournePedestrian	0.462	0.790	**0.836**	0.822	0.511	0.633	0.824	0.717
	OSULeaf	0.397	0.388	0.538	0.556	0.293	0.387	0.414	**0.557**
	Plane	0.473	0.694	0.762	0.688	0.543	**0.938**	0.545	0.738
	Symbols	0.465	0.789	0.915	0.918	0.634	0.877	**0.980**	0.749
	Trace	0.368	0.356	0.498	0.482	0.639	**0.713**	0.403	0.631
	Epilepsy	0.409	0.387	0.567	0.428	0.263	**0.583**	0.468	0.453
	NATOPS	0.379	0.367	0.489	0.393	**0.698**	0.639	0.426	0.403
	EthanolConcentration	**0.275**	0.242	0.237	0.254	0.251	0.257	0.213	0.272
	FaceDetection	0.476	0.471	0.487	0.471	0.497	**0.499**	0.496	0.483
	FingerMovements	0.514	**0.514**	0.483	0.458	0.510	0.489	0.474	0.478
	Avw_F1	0.421	0.478	0.546	0.518	0.448	**0.590**	0.503	0.544
	Av_Rank	6.769	7.000	4.923	5.615	7.231	**4.231**	5.308	4.462

Table 11: The average of weighted F1-score last ten epochs of methods on 13 benchmark datasets dealing with asymmetric noise. The best results are in **bold**

Noise	Datasets	Vanilla	SIGUA	Co-teaching	Mixup-BMM	Dividemix	Sel-CL	SREA	CTW
10%	ArrowHead	0.846	0.852	0.887	0.787	0.329	0.585	0.863	**0.898**
	CBF	0.961	0.964	0.988	0.776	0.593	0.935	0.986	**0.998**
	FaceFour	0.956	0.930	0.937	0.919	0.534	0.737	0.939	**0.967**
	MelbournePedestrian	0.873	0.898	0.913	0.898	0.428	0.699	**0.916**	0.903
	OSULeaf	0.805	0.792	0.871	0.762	0.353	0.599	0.895	**0.951**
	Plane	0.915	0.960	0.976	0.990	0.599	0.993	0.945	**0.994**
	Symbols	0.908	0.952	0.961	0.952	0.667	0.929	**0.990**	0.990
	Trace	0.986	0.977	0.992	0.950	0.850	**1.000**	0.977	**1.000**
	Epilepsy	0.895	0.909	0.937	0.898	0.641	0.870	0.913	**0.954**
	NATOPS	0.803	0.801	0.839	0.820	0.749	0.866	0.865	**0.870**
	EthanolConcentration	0.218	0.210	0.241	0.237	**0.253**	0.248	0.241	0.226
	FaceDetection	0.633	0.634	**0.654**	0.611	0.508	0.514	0.583	0.600
	FingerMovements	0.524	0.514	0.516	0.531	0.499	**0.561**	0.488	0.521
	Avw_F1	0.794	0.799	0.824	0.779	0.539	0.734	0.815	**0.836**
	Av_Rank	6.077	5.962	3.615	5.846	9.308	6.115	4.654	**2.731**
20%	ArrowHead	0.817	0.813	**0.888**	0.808	0.308	0.575	0.836	0.883
	CBF	0.895	0.904	0.968	0.747	0.601	0.950	0.935	**0.992**
	FaceFour	0.814	0.814	0.838	0.824	0.450	0.645	0.876	**0.974**
	MelbournePedestrian	0.796	0.872	**0.899**	0.897	0.428	0.699	0.896	0.898
	OSULeaf	0.728	0.704	0.817	0.759	0.338	0.567	0.809	**0.917**
	Plane	0.836	0.941	0.967	0.984	0.568	0.984	0.871	**0.993**
	Symbols	0.858	0.931	0.955	0.975	0.672	0.929	0.984	**0.988**
	Trace	0.846	0.853	0.940	0.975	0.850	0.993	0.902	**1.000**
	Epilepsy	0.806	0.830	0.873	0.794	0.584	0.806	0.868	**0.920**
	NATOPS	0.753	0.761	0.794	0.771	0.680	**0.822**	0.802	0.817
	EthanolConcentration	0.225	0.229	0.214	0.242	0.247	0.222	**0.262**	0.217
	FaceDetection	0.602	0.607	**0.621**	0.583	0.504	0.508	0.554	0.580
	FingerMovements	0.516	**0.529**	0.512	0.524	0.492	0.527	0.488	0.520
	Avw_F1	0.730	0.753	0.791	0.760	0.517	0.710	0.776	**0.823**
	Av_Rank	6.846	5.769	4.192	5.000	9.231	6.462	4.769	**2.885**
30%	ArrowHead	0.760	0.758	0.845	0.770	0.351	0.535	0.770	**0.854**
	CBF	0.803	0.864	0.921	0.698	0.633	**0.938**	0.839	0.933
	FaceFour	0.736	0.676	0.802	0.687	0.396	0.642	0.841	**0.906**
	MelbournePedestrian	0.698	0.830	0.881	0.874	0.436	0.689	0.866	**0.881**
	OSULeaf	0.635	0.636	0.759	0.734	0.327	0.513	0.693	**0.819**
	Plane	0.726	0.851	0.876	0.977	0.540	**0.981**	0.818	0.932
	Symbols	0.755	0.917	0.956	0.951	0.673	0.933	**0.993**	0.978
	Trace	0.691	0.775	0.836	0.860	0.850	**0.987**	0.829	0.977
	Epilepsy	0.728	0.784	0.791	0.765	0.634	0.799	0.796	**0.820**
	NATOPS	0.690	0.721	0.758	0.706	0.750	**0.788**	0.738	0.737
	EthanolConcentration	0.202	0.204	0.216	0.247	**0.249**	0.202	0.213	0.247
	FaceDetection	0.552	0.539	**0.565**	0.543	0.501	0.510	0.525	0.555
	FingerMovements	0.496	0.491	0.510	0.500	0.489	0.475	**0.539**	0.489
	Avw_F1	0.652	0.696	0.747	0.716	0.525	0.692	0.728	**0.779**
	Av_Rank	7.538	6.692	3.923	5.077	8.500	6.000	4.692	**2.885**
40%	ArrowHead	0.605	0.622	0.718	0.684	0.337	0.413	0.675	**0.768**
	CBF	0.657	0.714	0.763	0.615	0.622	**0.949**	0.665	0.776
	FaceFour	0.472	0.444	0.544	0.601	0.493	0.552	0.616	**0.708**
	MelbournePedestrian	0.592	0.773	**0.849**	0.828	0.603	0.683	0.811	0.802
	OSULeaf	0.540	0.536	0.670	0.638	0.264	0.452	0.600	**0.737**
	Plane	0.623	0.761	0.784	0.828	0.429	0.823	0.666	**0.919**
	Symbols	0.655	0.790	0.928	0.841	0.731	0.911	**0.985**	0.915
	Trace	0.596	0.688	0.664	0.703	0.700	0.822	0.618	**0.953**
	Epilepsy	0.588	0.587	0.654	0.604	0.575	**0.671**	0.652	0.637
	NATOPS	0.577	0.607	0.642	0.600	**0.784**	0.746	0.617	0.656
	EthanolConcentration	0.230	0.226	0.229	0.243	0.244	0.226	**0.245**	0.204
	FaceDetection	0.518	0.513	**0.519**	0.515	0.497	0.510	0.512	0.507
	FingerMovements	0.508	0.499	0.495	**0.529**	0.500	0.504	0.507	0.506
	Avw_F1	0.551	0.597	0.651	0.633	0.522	0.636	0.628	**0.699**
	Av_Rank	7.154	6.846	4.385	4.538	7.615	5.385	4.538	**3.769**

Table 12: Full results of CAug equipped with different augmentation methods when dealing with 30% symmetric noise. The best results are in **bold** (comparison with different augmentation methods).

Datasets	TimeWarp	GaussNoise	Convolve	Drift	Oversample	Crop	MF_Mixup
ArrowHead	**0.891**	0.876	0.869	0.858	0.857	0.861	0.851
CBF	**0.955**	0.894	0.911	0.877	0.911	0.933	0.920
FaceFour	0.929	0.874	0.891	0.832	0.825	**0.935**	0.874
MelbournePedestrian	0.890	0.869	0.873	0.860	0.878	0.882	**0.904**
OSULeaf	0.849	0.752	0.764	**0.856**	0.785	0.839	0.787
Plane	**0.971**	0.906	0.928	0.854	0.878	0.933	0.899
Symbols	0.970	**0.984**	0.983	0.981	0.975	0.974	0.976
Trace	**0.980**	0.936	0.932	0.884	0.944	0.936	0.949
Epilepsy	0.812	0.818	0.745	0.828	0.820	0.739	**0.833**
NATOPS	**0.788**	0.699	0.717	0.673	0.694	0.735	0.676
EthanolConcentration	0.176	0.194	0.190	0.169	**0.211**	0.190	0.174
FaceDetection	0.539	**0.564**	0.561	0.559	0.555	0.549	0.557
FingerMovements	0.540	0.533	**0.540**	0.518	0.523	0.526	0.529
Avw_F1	**0.792**	0.761	0.762	0.750	0.758	0.772	0.764
Av_Rank	**2.85**	3.77	3.69	5.31	4.62	3.81	3.96

Table 13: Full results of AAug equipped with different augmentation methods when dealing with 30% symmetric noise. The best results are in **bold** (comparison with different augmentation methods) and nderlined results are better than CAug equipped with the same augmentation method.

Datasets	TimeWarp	GaussNoise	Convolve	Drift	Oversample	Crop	MF_Mixup
ArrowHead	0.871	0.847	0.866	.859	0.836	.885	0.893
CBF	0.962	0.890	0.840	.898	0.873	.944	0.855
FaceFour	0.954	0.861	.898	.852	.890	0.928	0.831
MelbournePedestrian	0.885	0.863	0.871	.872	0.860	0.879	**0.885**
OSULeaf	0.837	.778	.770	0.813	0.771	0.839	.790
Plane	**0.958**	.917	0.904	.862	.889	0.928	0.884
Symbols	0.979	0.971	0.958	0.968	0.967	0.972	0.967
Trace	**0.953**	0.883	0.873	.908	0.871	0.925	0.914
Epilepsy	0.788	0.761	0.790	0.781	0.779	0.737	0.782
NATOPS	0.731	0.693	0.675	0.668	0.686	.750	0.754
EthanolConcentration	.183	0.191	0.210	.210	0.174	.195	.181
FaceDetection	0.568	0.562	.568	.561	.555	.554	0.549
FingerMovements	0.528	0.507	0.558	.552	0.511	0.507	0.503
Avw_F1	**0.784**	0.748	0.752	.754	0.743	.772	0.753
Av_Rank	**2.00**	4.62	4.15	4.23	5.58	3.00	4.42

Table 14: Full results of CTW with reconstructing augmented instances when dealing with 30% symmetric noise. The best results are in **bold** (comparison with different augmentation methods) and nderlined results are better than CAug equipped with the same augmentation method.

Datasets	TimeWarp	GaussNoise	Convolve	Drift	Oversample	Crop	MF_Mixup
ArrowHead	0.856	0.878	0.866	.876	.874	.872	0.841
CBF	0.957	0.890	0.873	.886	.926	0.930	.934
FaceFour	0.962	0.834	0.861	.871	.858	0.899	.891
MelbournePedestrian	.895	.871	.878	0.858	0.872	.886	**0.899**
OSULeaf	**0.847**	.769	.791	0.788	0.755	0.810	.812
Plane	**0.946**	0.874	0.923	.873	0.874	0.919	.903
Symbols	.976	**0.979**	0.977	0.979	.977	.977	.978
Trace	**0.980**	.945	0.932	0.882	0.940	.945	0.923
Epilepsy	.831	0.817	.814	.843	.821	0.723	0.849
NATOPS	0.745	0.696	0.701	.719	.698	0.755	.718
EthanolConcentration	0.172	0.173	0.174	.187	0.205	.197	0.215
FaceDetection	.553	0.562	0.569	0.557	.561	.555	0.550
FingerMovements	0.468	0.523	0.511	0.546	.527	0.522	0.517
Avw_F1	**0.784**	0.755	0.759	.759	.761	0.768	.772
Av_Rank	**3.46**	4.38	4.62	4.15	4.38	**3.46**	3.54

Table 15: Full results of sample selection analysis on benchmark datasets when dealing with 60% symmetric noise. *Rec.*: Sample reconstruction. *Sel.*: Sample selection. *EPS*: Eliminating Selection Preference. *R*: The proportion of all losses that are less than ℓ_{thred}. The same proportion *R* is used for each class. The method with ***** selects instances from the whole dataset instead of each batch. The best results are in **bold**

Datasets	Vanilla + Rec. + Sel.	Vanilla + Rec. + Sel. (with EPS)	CTW	CTW (Sel. with GMM class by class)	CTW (Sel. with R class by class)	CTW*	CTW (Sel. with GMM class by class)*
ArrowHead	0.331	0.389	0.441	0.444	0.443	**0.478**	0.400
CBF	0.410	0.426	**0.537**	0.466	0.465	0.531	0.536
FaceFour	0.468	0.371	0.532	**0.568**	0.529	0.554	0.480
MelbournePedestrian	0.591	0.608	0.713	0.735	0.727	0.740	**0.757**
OSULeaf	0.468	0.456	0.544	0.570	0.544	0.566	**0.673**
Plane	0.543	0.520	0.722	0.658	0.731	0.846	**0.952**
Symbols	0.667	0.691	0.738	0.742	0.715	0.790	**0.972**
Trace	0.385	0.438	0.559	0.589	0.663	**0.678**	0.599
Epilepsy	0.422	0.412	0.428	**0.435**	0.359	0.402	0.406
NATOPS	0.400	0.359	0.399	0.394	0.377	**0.484**	0.413
EthanolConcentration	0.170	0.222	0.219	0.241	0.219	0.235	**0.246**
FaceDetection	0.477	0.474	0.480	0.480	**0.490**	0.487	0.487
FingerMovements	**0.489**	0.480	0.478	0.477	0.466	0.472	0.456
Ave_F1	0.448	0.450	0.522	0.523	0.518	0.559	**0.567**
Av_Rank	5.62	5.77	3.81	3.12	4.31	**2.62**	2.77

Table 16: Full results of sample selection analysis on benchmark datasets when dealing with 40% asymmetric noise. *Rec.*: Sample reconstruction. *Sel.*: Sample selection. *EPS*: Eliminating Selection Preference. *R*: The proportion of all losses that are less than ℓ_{thred}. The method with ***** selects instances from the whole dataset instead of each batch. The same proportion *R* is used for each class. The best results are in **bold**

Datasets	Vanilla + Rec. + Sel.	Vanilla + Rec. + Sel. (with EPS)	CTW	CTW (Sel. with GMM class by class)	CTW (Sel. with R class by class)	CTW*	CTW (Sel. with GMM class by class)*
ArrowHead	0.666	0.682	0.741	0.725	**0.795**	0.744	0.740
CBF	0.645	0.680	0.767	0.750	0.726	0.795	**0.979**
FaceFour	0.558	0.621	**0.754**	0.691	0.748	0.728	0.641
MelbournePedestrian	0.695	0.765	0.796	0.792	0.784	**0.845**	0.765
OSULeaf	0.560	0.632	0.722	0.632	0.657	**0.780**	0.776
Plane	0.709	0.690	0.906	0.857	0.800	**0.965**	0.950
Symbols	0.858	0.875	0.911	0.903	0.816	0.971	**0.984**
Trace	0.685	0.658	0.950	0.960	0.965	0.914	**1.000**
Epilepsy	0.554	0.577	0.612	0.632	0.644	**0.666**	0.606
NATOPS	0.609	0.579	0.655	0.596	0.605	**0.682**	0.645
EthanolConcentration	0.216	0.204	0.183	0.190	0.214	**0.219**	0.218
FaceDetection	**0.516**	0.516	0.505	0.503	0.513	0.515	0.509
FingerMovements	0.499	0.495	0.495	0.491	0.486	0.493	**0.519**
Ave_F1	0.598	0.613	0.692	0.671	0.673	0.717	**0.718**
Av_Rank	5.38	5.46	3.46	4.69	4.00	**2.15**	2.85

Table 17: Ablation analysis on benchmark datasets dealing with 30% symmetric noise. *Sel.*: Sample selection. *EPS*: Eliminating Selection Preference. *Rec.*: Sample reconstruction.

Datasets	CTW	w/o Sel.	w/o EPS	w/o TimeWarp	w/o Rec.
ArrowHead	0.851	0.840	0.850	**0.875**	0.860
CBF	**0.958**	0.913	0.943	0.888	0.943
FaceFour	**0.965**	0.850	0.955	0.853	0.927
MelbournePedestrian	**0.880**	0.822	0.879	0.872	0.879
OSULeaf	**0.851**	0.727	0.840	0.782	0.826
Plane	**0.995**	0.882	0.990	0.885	**0.995**
Symbols	0.978	0.831	0.953	**0.980**	0.965
Trace	**0.990**	0.918	**0.990**	0.921	0.985
Epilepsy	0.751	0.736	0.772	**0.828**	0.793
NATOPS	**0.747**	0.683	0.721	0.703	0.727
EthanolConcentration	0.202	0.198	0.191	0.196	**0.211**
FaceDetection	0.536	0.520	0.538	**0.563**	0.534
FingerMovements	0.519	0.535	0.525	**0.546**	0.498
Avw_F1	**0.786**	0.727	0.781	0.761	0.780
Av_Rank	**2.00**	4.54	2.85	2.92	2.69

Table 18: The influence of the hyperparameter β on benchmark datasets with 30% symmetric noise.

$\beta =$	2	4	6	8	10	12
ArrowHead	0.840	0.879	0.869	0.871	0.865	**0.895**
CBF	0.888	0.946	0.954	0.937	0.945	**0.957**
FaceFour	0.938	0.937	**0.957**	0.937	0.946	0.938
MelbournePedestrian	0.872	0.883	**0.886**	0.886	0.884	0.884
OSULeaf	0.804	0.831	0.833	0.838	**0.852**	0.830
Plane	**1.000**	0.995	0.995	0.995	0.995	0.990
Symbols	0.966	**0.981**	0.973	0.976	0.978	0.968
Trace	**1.000**	0.995	**1.000**	**1.000**	**1.000**	0.990
Epilepsy	0.723	0.783	0.815	0.767	0.804	**0.825**
NATOPS	0.721	0.719	0.723	0.729	**0.741**	0.728
EthanolConcentration	0.183	0.188	0.174	0.193	**0.207**	0.187
FaceDetection	0.529	0.535	**0.541**	0.531	0.537	0.530
FingerMovements	0.512	0.504	0.515	0.527	**0.539**	0.523
Avw_F1	0.767	0.783	0.787	0.784	**0.792**	0.788
Av_Rank	4.92	3.88	2.92	3.23	**2.46**	3.58

Table 19: The influence of the hyperparameter γ on benchmark datasets with 30% symmetric noise.

$\gamma =$	0.15	0.3	0.45	0.6	0.75	0.9
ArrowHead	0.872	0.865	0.842	0.857	**0.880**	0.870
CBF	**0.945**	0.945	0.928	0.924	0.915	0.915
FaceFour	0.947	0.946	0.937	0.947	0.964	**0.973**
MelbournePedestrian	0.882	**0.884**	0.872	0.880	0.874	0.875
OSULeaf	0.850	**0.852**	0.843	0.827	0.851	0.830
Plane	**0.995**	**0.995**	0.990	**0.995**	0.990	0.990
Symbols	**0.985**	0.978	0.976	0.971	0.945	0.953
Trace	**1.000**	**1.000**	0.995	**1.000**	0.985	0.985
Epilepsy	**0.830**	0.804	0.820	0.777	0.813	0.777
NATOPS	0.731	0.741	**0.756**	0.732	0.755	0.735
EthanolConcentration	0.203	0.207	0.187	**0.209**	0.178	0.179
FaceDetection	0.531	0.537	0.535	**0.539**	0.529	0.531
FingerMovements	0.536	**0.539**	0.508	0.536	0.506	0.525
Avw_F1	**0.793**	0.792	0.784	0.784	0.784	0.780
Av_Rank	**2.62**	2.38	4.00	3.38	4.15	4.46

Table 20: The influence of the hyperparameter λ on benchmark datasets with 30% symmetric noise.

$\lambda =$	0.2	0.4	0.6	0.8	1	1.2
ArrowHead	0.869	0.880	**0.885**	0.865	0.865	0.866
CBF	0.936	0.945	0.937	0.936	0.945	**0.950**
FaceFour	0.927	0.964	0.955	0.963	0.946	**0.964**
MelbournePedestrian	0.887	0.882	**0.889**	0.885	0.884	0.874
OSULeaf	**0.867**	0.831	0.838	0.815	0.852	0.850
Plane	0.981	0.986	**0.995**	**0.995**	**0.995**	**0.995**
Symbols	**0.986**	0.975	0.978	0.978	0.978	0.983
Trace	0.948	0.995	0.985	0.990	**1.000**	0.995
Epilepsy	**0.839**	0.835	0.809	0.808	0.804	0.827
NATOPS	0.762	0.751	0.750	0.756	0.741	**0.786**
EthanolConcentration	0.176	0.200	0.198	0.185	**0.207**	0.189
FaceDetection	0.540	0.536	0.539	**0.542**	0.537	0.532
FingerMovements	0.526	0.522	0.512	0.512	**0.539**	0.531
Avw_F1	0.788	0.792	0.790	0.787	0.792	**0.796**
Av_Rank	3.42	3.58	3.54	3.96	3.58	**2.92**

Table 21: The influence of the hyperparameter μ on benchmark datasets with 30% symmetric noise.

$\mu =$	0.2	0.4	0.6	0.8	1	1.2
ArrowHead	0.851	0.844	**0.882**	0.856	0.865	0.867
CBF	0.944	0.950	**0.961**	0.920	0.945	0.949
FaceFour	0.965	0.964	0.965	**0.974**	0.946	0.937
MelbournePedestrian	**0.887**	0.879	0.883	0.884	0.884	0.880
OSULeaf	0.801	0.854	0.845	0.848	0.852	**0.864**
Plane	0.990	0.995	0.986	**1.000**	0.995	0.995
Symbols	0.978	**0.982**	0.979	0.981	0.978	0.976
Trace	0.990	**1.000**	**1.000**	**1.000**	**1.000**	**1.000**
Epilepsy	0.819	0.794	0.830	**0.845**	0.804	0.798
NATOPS	0.764	**0.767**	0.740	0.751	0.741	0.754
EthanolConcentration	0.194	0.212	0.174	0.203	0.207	**0.225**
FaceDetection	0.536	0.531	0.536	**0.541**	0.537	0.535
FingerMovements	0.510	0.539	0.516	0.488	**0.539**	0.502
Avw_F1	0.787	**0.793**	0.792	0.792	0.792	0.791
Av_Rank	4.00	3.38	3.54	**3.00**	3.38	3.69

References

[Arazo *et al.*, 2019] Eric Arazo, Diego Ortego, Paul Albert, Noel O'Connor, and Kevin McGuinness. Unsupervised label noise modeling and loss correction. In *International conference on machine learning*, pages 312–321. PMLR, 2019.

[Castellani *et al.*, 2021] Andrea Castellani, Sebastian Schmitt, and Barbara Hammer. Estimating the electrical power output of industrial devices with end-to-end time-series classification in the presence of label noise. In *Joint European Conference on Machine Learning and Knowledge Discovery in Databases*, pages 469–484. Springer, 2021.

[Han *et al.*, 2018] Bo Han, Quanming Yao, Xingrui Yu, Gang Niu, Miao Xu, Weihua Hu, Ivor Tsang, and Masashi Sugiyama. Co-teaching: Robust training of deep neural networks with extremely noisy labels. *Advances in neural information processing systems*, 31, 2018.

[Han *et al.*, 2020] Bo Han, Gang Niu, Xingrui Yu, Quanming Yao, Miao Xu, Ivor Tsang, and Masashi Sugiyama. Sigua: Forgetting may make learning with noisy labels more robust. In *International Conference on Machine Learning*, pages 4006–4016. PMLR, 2020.

[Huang *et al.*, 2022] Yingsong Huang, Bing Bai, Shengwei Zhao, Kun Bai, and Fei Wang. Uncertainty-aware learning against label noise on imbalanced datasets. 2022.

[Karim *et al.*, 2022] Nazmul Karim, Mamshad Nayeem Rizve, Nazanin Rahnavard, Ajmal Mian, and Mubarak Shah. Unicon: Combating label noise through uniform selection and contrastive learning. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pages 9676–9686, 2022.

[Li *et al.*, 2020] Junnan Li, Richard Socher, and Steven CH Hoi. Dividemix: Learning with noisy labels as semi-supervised learning. *arXiv preprint arXiv:2002.07394*, 2020.

[Li *et al.*, 2022] Shikun Li, Xiaobo Xia, Shiming Ge, and Tongliang Liu. Selective-supervised contrastive learning with noisy labels. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pages 316–325, 2022.

[Verma *et al.*, 2019] Vikas Verma, Alex Lamb, Christopher Beckham, Amir Najafi, Ioannis Mitliagkas, David Lopez-Paz, and Yoshua Bengio. Manifold mixup: Better representations by interpolating hidden states. In *International Conference on Machine Learning*, pages 6438–6447. PMLR, 2019.