ON THE USE OF CONCORDANCE CORRELATION COEFFICIENT FOR EVALUATING FIRST SHOT ANOMALOUS SOUND DETECTION

Technical Report

*Bagus Tris Atmaja**

National Institute of Advanced
Industrial Science and Technology
Tsukuba, Japan
b-atmaja@aist.go.jp

Akira Sasou

National Institute of Advanced
Industrial Science and Technology
Tsukuba, Japan
a-sasou@aist.go.jp

ABSTRACT

The choice of the loss function is a critical aspect of machine/deep learning. In this study, we investigate the use of the concordance correlation coefficient (CCC) as a loss function for first-shot anomaly sound detection. We compare the performance of CCC with the commonly used loss function, mean squared error (MSE). Furthermore, we benchmark CCC, MSE, and selective Malahanobis distance equally. The results show that CCC outperforms MSE and Selective Mahalanobis in terms of the harmonic mean of pAUC scores. We repeated the experiments of our method with CCC five times, and we obtained similar results across four runs showing the stability of our method.

Index Terms— anomalous sound detection, condition-based monitoring, concordance correlation coefficient, loss function

1. SYSTEM DESCRIPTION

The description of the system can be read in the original papers [3, 4, 2] along with their implementations [1]. The dataset for the 2023 challenge is based on the datasets provided in the previous years [5, 6]. The only change we made is to change the loss function (`loss_fn` function in the original implementation) from MSE to the concordance correlation coefficient.

CCC loss (CCCL) is formulated as

$$CCC = \frac{2\rho_{xy}\sigma_x\sigma_y}{\sigma_x^2 + \sigma_y^2 + (\mu_x - \mu_y)^2}, \tag{1}$$

$$CCCL = 1 - CCC, \tag{2}$$

where μ_x and μ_y are the means of the predicted and ground truth values, respectively. σ_x and σ_y are the standard deviations of the predicted and ground truth values, respectively, and ρ_{xy} is the Pearson correlation between the predicted and ground truth values. CCC loss is arguably more effective than other error-based loss functions, especially when the metric is CCC. CCC is more effective than other correlation functions since it not only accounts for the relation of the two variables but also for the exact difference in values [7].

Listing 1 shows our results with CCC loss and original MSE baseline, while Table 1 shows its results. It is shown from four experiments that the CCC loss is arguably better for obtaining the average pAUC scores than MSE and selective Mahalanobis distance methods.

*Thanks to NEDO Japan (project JPNP20006) for funding.

2. REFERENCES

[1] https://github.com/nttcslab/dase2023_task2_baseline_ae

[2] https://dcase.community/challenge2023/task-first-shot-unsupervised-anomalous-sound-detection-for-machine-condition-monitoring

[3] N. Harada, D. Niizumi, Y. Ohishi, D. Takeuchi, and M. Yasuda, “First-shot anomaly sound detection for machine condition monitoring: A domain generalization baseline,” Mar. 2023, [Online]. Available: http://arxiv.org/abs/2303.00455.

[4] Kota Dohi, Keisuke Imoto, Noboru Harada, Daisuke Niizumi, Yuma Koizumi, Tomoya Nishida, Harsh Purohit, Ryo Tanabe, Takashi Endo, and Yohei Kawaguchi, ‘Description and discussion on dcase 2023 challenge task 2: first-shot unsupervised anomalous sound detection for machine condition monitoring,” In arXiv e-prints: 2305.07828, 2023.

[5] Noboru Harada, Daisuke Niizumi, Daiki Takeuchi, Yasunori Ohishi, Masahiro Yasuda, and Shoichiro Saito. ToyADMOS2: another dataset of miniature-machine operating sounds for anomalous sound detection under domain shift conditions. In Proceedings of the Detection and Classification of Acoustic Scenes and Events Workshop (DCASE), 1–5. Barcelona, Spain, November 2021.

[6] Kota Dohi, Tomoya Nishida, Harsh Purohit, Ryo Tanabe, Takashi Endo, Masaaki Yamamoto, Yuki Nikaido, and Yohei Kawaguchi. Mimii dg: sound dataset for malfunctioning industrial machine investigation and inspection for domain generalization task. In Proceedings of the 7th Detection and Classification of Acoustic Scenes and Events 2022 Workshop (DCASE2022). Nancy, France, November 2022.

[7] B. T. Atmaja and A. Sasou, “Evaluating Variants of wav2vec 2.0 on Affective Vocal Burst Tasks,” in ICASSP 2023 - 2023 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), Jun. 2023, pp. 1–5, doi: 10.1109/ICASSP49357.2023.10096552.

Listing 1: CCC loss function in Python

```
def loss_fn(self, recon_x, x):
    """ CCC loss function """
    epsilon = 1e-8

    # Flatten the input tensors
    recon_x_flat = recon_x.view(recon_x.shape[0], -1)
    x_flat = x.view(x.shape[0], -1)

    # Calculate means
    recon_mean = torch.mean(recon_x_flat, dim=1, keepdim=True)
    x_mean = torch.mean(x_flat, dim=1, keepdim=True)

    # Center the tensors
    recon_centered = recon_x_flat - recon_mean
    x_centered = x_flat - x_mean

    # Calculate variances
    recon_var = torch.mean(recon_centered ** 2, dim=1, keepdim=True)
    x_var = torch.mean(x_centered ** 2, dim=1, keepdim=True)

    # Calculate covariance
    recon_cov = torch.mean(recon_centered * x_centered, dim=1, keepdim=True)

    # Calculate CCC
    ccc = 2 * recon_cov / (recon_var + x_var + epsilon)

    # Calculate CCC loss, 1 - CCC
    ccc_loss = 1 - ccc

    return ccc_loss
```

Table 1: Results of experiments with CCC loss

System	Metric	ToyCar	ToyTrain	fan	gearbox	bearing	slider	valve	Mean
Baseline MAHALA	AUC (source)	73.66	57.22	69.92	48.70	54.01	56.97	45.72	58.03
	AUC (target)	42.94	40.90	31.42	53.66	43.25	42.93	46.87	43.14
	pAUC (source, target)	49.00	48.32	50.61	50.18	49.87	48.45	49.37	49.40
Baseline MSE	AUC (source)	68.62	59.72	69.28	50.20	52.91	60.11	48.14	58.43
	AUC (target)	46.36	57.28	30.96	54.84	44.65	41.25	47.43	46.11
	pAUC (source, target)	50.42	48.47	50.53	50.63	49.79	50.26	49.03	49.88
CCC #1	AUC (source)	59.18	57.74	44.06	58.26	52.39	53.53	47.76	53.27
	AUC (target)	54.52	54.18	57.32	59.48	49.11	49.87	48.79	53.32
	pAUC (source, target)	48.68	50.16	50.79	50.97	50.29	50.45	50.03	50.20
CCC #2	AUC (source)	59.14	58.78	46.34	58.02	51.91	53.91	47.50	53.66
	AUC (target)	54.28	53.56	54.26	59.44	48.61	49.71	48.55	52.63
	pAUC (source, target)	48.63	49.89	52.00	50.82	50.45	50.82	49.97	50.37
CCC #3	AUC (source)	59.10	58.76	45.12	58.00	51.33	53.75	47.50	53.37
	AUC (target)	54.22	53.74	55.40	59.26	49.23	49.87	49.09	52.97
	pAUC (source, target)	48.89	50.05	51.18	50.55	50.63	50.79	49.97	50.29
CCC #4	AUC (source)	58.52	58.58	45.02	57.76	52.01	53.71	47.40	53.29
	AUC (target)	54.58	53.94	55.00	59.42	49.01	49.99	48.73	52.95
	pAUC (source, target)	48.79	50.16	51.37	50.87	50.71	50.71	49.97	50.37