Self-supervised learning of audio representations using angular contrastive loss

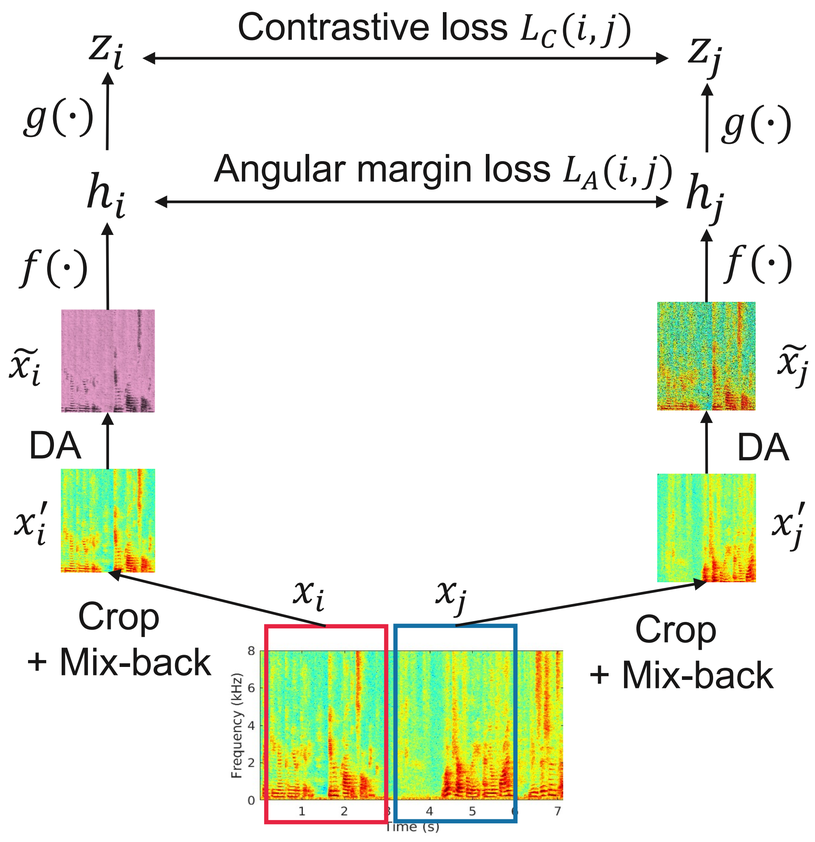
Shanshan Wang*, Soumya Tripathy, Annamaria Mesaros

Dept. of Information Technology and Communication Sciences, Tampere University

Introduction

- **Motivation**: Cross-Entropy (CE) loss does not explicitly optimize the feature embedding space
  - fails to learn clear clusters for classes -> affects feature quality
  - solution: auxiliary losses, e.g., angular margin loss (AMC)
- **Proposed**: Angular Contrastive Loss (ACL) - weighted sum of NT-Xent loss and angular margin loss
  - self-supervised learning (SSL)
  - NT-Xent is CE-based, expected to behave similarly with the addition of AMC

Method



$$ACL = \alpha * L_C + (1 - \alpha) * L_A \tag{1}$$

$$L_C(i,j) = -log \frac{exp(sim(z_i, z_j)/\tau)}{\sum_{k=1}^{2N} \mathbb{1}_{[k \neq i]} exp(sim(z_i, z_k)/\tau)} \tag{2}$$

$$L_A(i,j) = \begin{cases} (\cos^{-1}\langle h_i, h_j \rangle)^2 & \text{if } S_{ij} = 1 \\ max(0, m_g - \cos^{-1}\langle h_i, h_j \rangle)^2 & \text{if } S_{ij} = 0 \end{cases} \tag{3}$$

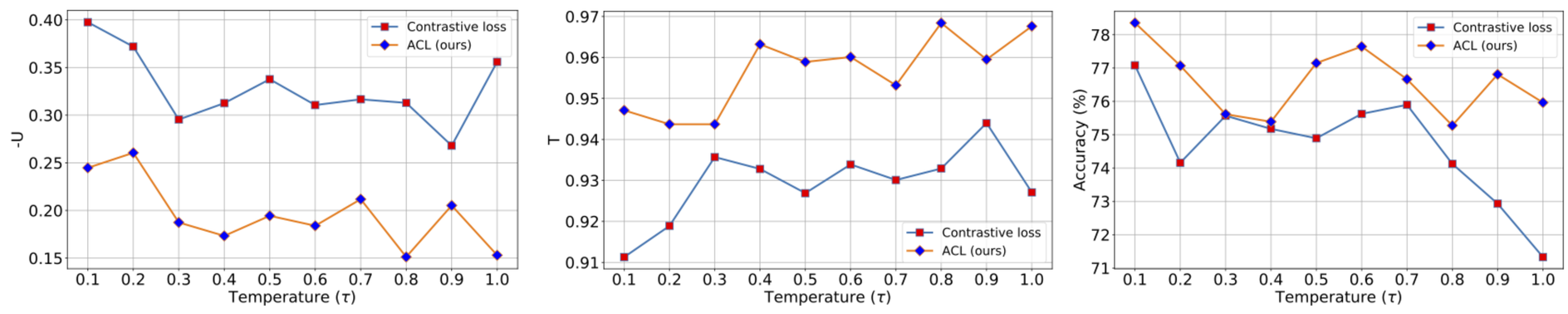
Experiments and Results

- **Experiments**: Train and test on FSDnoisy18K dataset
  - Supervised learning: CE loss + AMC
  - Self-supervised learning: ACL
- **Results and analysis**
  - Test accuracy

Training method	Best performance
Supervised baseline	70.1
Supervised with ACL	**73.6**
SSL baseline	74.2
SSL with ACL	**77.1**

- 1. Supervised learning: 70.1 $\rightarrow$ 73.6 (3.5 % boost)
- 2. SSL: 74.2 $\rightarrow$ 77.1 (2.9% boost)
- 3. **SSL performs better than supervised learning**

  - Feature analysis



  - ⋆ **ACL is always better** than NT-Xent loss irrespective of the temperature
  - ⋆ Slight increase in the feature tolerance value seems to be more beneficial
  - ⋆ The degradation in the uniformity does not necessarily harm the feature quality

Conclusion

- We proposed angular contrastive loss to improve the feature quality in the SSL.
- We validated the effectiveness of ACL for audio representation learning in both supervised learning and SSL.

Paper: `https://arxiv.org/pdf/2211.05442.pdf`
Code: `https://github.com/shanwangshan/Self_supervised_ACL`