ICASSP 2023 - 2023 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP) | 978-1-7281-6327-7/23/$31.00 ©2023 IEEE | DOI: 10.1109/ICASSP49357.2023.10094893

BOOSTING SIGNAL MODULATION FEW-SHOT LEARNING WITH PRE-TRANSFORMATION

Peng Sun[*,1] *Jie Su*[*,2] *Zhenyu Wen*[1] *Yejian Zhou*[1] *Zhen Hong*[1] *Shanqing Yu*[1] *Huaji Zhou*[3]

[1] Zhejiang University of Technology
[2] Newcastle University
[3] Xidian University

ABSTRACT

The recent flourish of deep learning on various tasks is largely accredited to the rich and high-quality labeled data. Nonetheless, collecting sufficient labeled samples is not very practical for many real applications. *Few-shot Learning* (FSL) provides a promising solution that allows a model to learn the concept of novel classes with a few labeled samples. However, many existing FSL methods are only designed for computer vision tasks and are not suitable for radio signal recognition. This paper calls for a radically different approach to FSL: in contrast to developing a new FSL model, we should focus on transforming the radio signal to be better processed by the state-of-the-art (SOTA) FSL model. We propose Modulated Signal Pre-transformation (MSP), a parameterized radio signal transformation framework that encourages the signals having the same semantics to have similar representations. MSP currently adapts to various SOTA FSL models for signal modulation recognition and can support the mainstream deep learning backbone. Evaluation results show that MSP improves the performance gains for many SOTA FSL models while maintaining flexibility.

***Index Terms*—** Signal Processing, Machine Learning, Few Shot Learning

1. INTRODUCTION

Deep learning-based methods have shown advantages in radio signal modulation recognition [1]. However, the availability of labeled datasets for the radio signal is very limited. Unlike image data, to label these types of data, human annotators needs to have knowledge of radio signal processing.

FSL was proposed to solve such label-scarce problems by learning a model on a set of base classes (i.e., base set) and studying its adaptability to novel classes with only a few samples (i.e., support set). Many classical FSL methods have been proposed in computer vision, and prior works can be roughly cast into three categories. Metric-based methods like Matching network [2], Relation Network [3], Prototypical Network [4] and RENet [5]) aim to learn a set of project functions (embedding functions) and metrics to measure the similarity between samples. Meta-based methods (MAML [6], ProtoMAML [7]) aim to use a model agnostic meta-learner to train a good basic model on a variety of training tasks, such that given a new task with only few training samples, a small number of gradient steps is sufficient to produce a good generalization model. Augmentation-based methods [8, 9] aim to design different sample generation strategies for novel classes to encourage representation learning. Benefiting from the success of FSL methods in computer vision, a few improved variations of these methods [10] have recently been proposed to perform few-shot recognition on modulated signals. For example, ARN [11] provides an improved metric-based method by utilizing the channel and spatial attention in the feature extractor.

However, the FSL methods in modulated signals mainly focus on improving the learning approach. When adapting those methods to the radio signals data, they may suffer from a 'poor generalization' problem, where the learned semantic/constant information can not be effectively generalized to

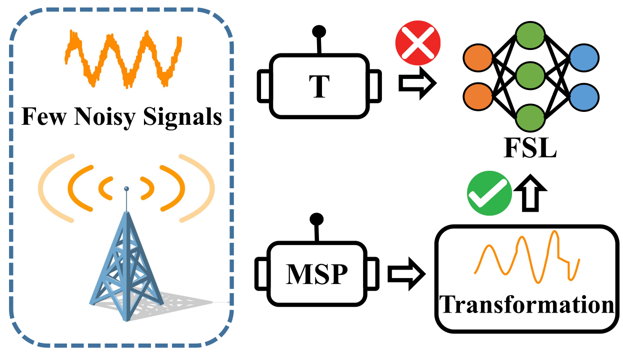

Fig. 1: The received modulated signals often suffer from noise interference during the open environment transmission, and the quantity of collected signals is usually less. Our proposed MSP framework presents a novel transformation strategy, taking the noisy problem into account, and boosting the performance of the FSL recognition.

* indicates equal contribution. Correspondence to: Yejian Zhou.

the query/novel classes. This may be caused by the following reasons: 1) The received modulated signals often suffer from **noise interference** during the open environment transmission, making representation learning more difficult. 2) Radio signal data has the **unique characteristics** of periodicity, symmetry, etc., which are not easily learned by deep learning models with insufficient samples.

To overcome the above-mentioned challenges, in this paper, we shift the attention and propose, MSP, a parameterized radio signal transformation framework. MSP is to obtain the constants with high-density (e.g., semantic information) while filtering the no constants such as added noises, providing better intra-class concentrated [12] constants for downstream FSL methods.

To the best of our knowledge, MSP is the first framework that solves the radio signal modulation recognition FSL problem by transforming radio signals to be better processed by SOTA FSL models. We perform extensive experiments and analysis, and make the following key observations:

- We design MSP as an end-to-end joint pre-training framework for learning effective representation from modulated signal data.
- The MSP framework consists of two designed modules, and we demonstrate its effectiveness under high and low noise conditions.
- The existing SOTA FSL methods obtain decent performance gain on radio signal modulation FSL task by attaching the proposed MSP framework.

2. SIGNAL PRE-TRANSFORMATION

To obtain the constant semantic information from the signals, we aim to maximize the mutual information between two signal segments x_1 and x_2 that have similar semantics, which can be formalized as:

$$I(x_1, x_2) = \sum\nolimits_{x_1,x_2} p(x_1, x_2) \log(\frac{p(x_1, x_2)}{p(x_1) \cdot p(x_2)}) \quad (1)$$

To this end, the proposed MSP framework is designed to minimize an InfoNCE loss (see Eq. 9), which is equivalent to maximizing Eq. 1. Figure 2 demonstrates the process of the proposed MSP framework. Specifically, we first utilize an *Adaptive Noise Filtering* module to reduce the interference of semantically irrelevant information in input signal (i.e., inherent Gaussian noise). Moreover, we design a constraint loss (see Eq. 7) to make it more approximate an optimum signal-filter. Then, we use a *Info-preserved Augmentation* module to augment (but remain modulation type unchanged) signal to support the optimization of InfoNCE loss. At last, we exploit a shared weight parametric model (e.g., CNN) to convert the signals into high-level representations (i.e., capture the semantically relevant information).

2.1. Info-preserved Augmentation

[13] discuss that data augmentation can greatly improve the performance of extracting semantic information by InfoNCE. These improvements come from two aspects: 1) data augmentation can increase the size of training data; 2) data augmentation increases the number of data that have similar semantics to maximize the mutual information. We designed two novel types of signal augmentation to generate information-preserved data samples.

Interception. For a quadrature modulated signal segment x in original length L, the interception operation can be written as:

$$g_{inter}(x, a, L') = [x(a), x(a+1), ..., x(a+L')] \quad (2)$$

where $x_{a,L'}$ is the intercepted signal and L' represents the predefined interception length. The intercept start point a should be selected from $[0, L-L']$ interval.

Rotation. In order to facilitate the illustration of rotation, we will introduce a background knowledge of modulated signals. A quadrature modulated signal/complex signal segment x is formed by the signal pairs $\{x_R, x_I\}$ where x_R and x_I represents the real and imaginary part of signals, respectively. The relationship between these signals is given by:

$$x = x_R + j \cdot x_I \quad (3)$$

where j is a imaginary number, that is, $j = \sqrt{-1}$. We apply the Euler formula [14] to rotate the original signals. For a quadrature modulated signal, the rotation formulation could be written as:

$$\begin{aligned} g_{rotate}(x, \theta) &= x \cdot e^{\theta j} \\ &= (x_R + j \cdot x_I)(cos(\theta) + j \cdot sin(\theta)) \end{aligned} \quad (4)$$

where $g_{rotate}(x, \theta)$ represents the rotation augmentation operation and $\theta \sim U(0, 2\pi)$ denotes the rotation angle drawn from the uniform distribution.

2.2. Adaptive Noise Filtering

In order to filter signals adaptively, we utilize a parameterized Gaussian filter [15] to filter noisy signals adaptively, which can be formulated as:

$$f_\theta(x) = x * G \quad s.t. \quad G(n) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{n^2}{2\sigma^2}} \quad (5)$$

and $G(n)$ is the n-index variable of the Gaussian filter kernel. The parameter σ can be optimized by learning to adapt to different signal types. Moreover, an optimum filter $f_\theta^*(x)$ should then be equivariant [16] to the augmentations $g(x)$, i.e., swapping the order of augmentation and filtering should yield the same result:

$$g(f_\theta^*(x)) = f_\theta^*(g(x)) \quad (6)$$

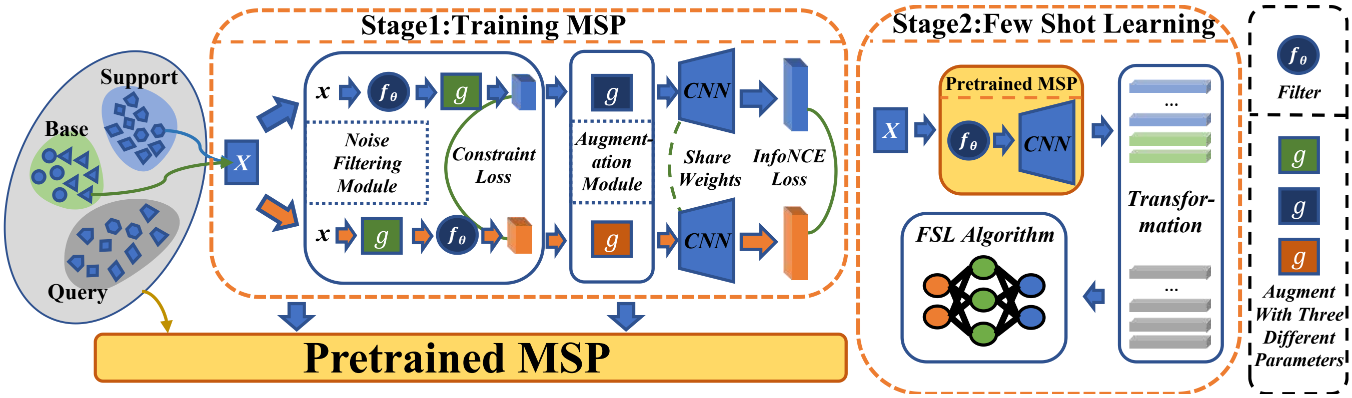

Fig. 2: The structure of the proposed Modulated Signal Pre-transformation (MSP) framework. The base set classes are disjoint with support and query sets. Further, the query set shares the same classes with the support set and is only used for testing, which will not contribute to the training procedure.

Therefore, under this assumption, we utilize the parameterized filter f_θ to approximate the optimum filter by minimizing the constraint loss:

$$\mathcal{L}_G = \|g(f_\theta(x)) - f_\theta(g(x))\|_2^2 \tag{7}$$

2.3. Optimization Objective

The MSP framework is trained simultaneously under the optimization of the constraint loss $\mathcal{L}_G$ and InfoNCE loss $\mathcal{L}_C$, which can be formulated as:

$$\mathcal{L} = \mathcal{L}_C + \mathcal{L}_G \tag{8}$$

Specifically, the InfoNCE loss could be formulated as:

$$\mathcal{L}_C = \mathcal{E}_{x_1^i, x_2^i \sim p(x_1, x_2)}(\log \frac{\sum_{j \neq i}^{N} h(x_1^i, x_2^j)}{h(x_1^i, x_2^i)}) \tag{9}$$

where $h(x_1, x_2) = \exp\{sim(g_\theta(x_1), g_\theta(x_2))/\alpha\}$, g_θ represents the MSP framework and α represents a temperature parameter. $sim(u, v) = u^{\mathrm{T}}v / \|u\| \|v\|$ denotes the dot product between l_2 normalized u and v (i.e. cosine similarity), which is used to measure the similarity between u and v. We regard the augmented samples from the same signal as positives x_1^i, x_2^i, and augmented samples from different signals as negatives x_1^i, x_2^j.

3. EXPERIMENTS

In this section, we evaluate the performance of MSP by comparing the SOTA FSL algorithms and these algorithms assembled with our framework.

3.1. Datasets

We evaluate our proposed MSP framework on three benchmark datasets, namely, signal-128 [1], signal-512, and signal-1024 [17], where the numbers represent the length of the signals.

Signal-128 is a public radio dataset that contains 11 modulations. Each modulation type has 20 different Signal-to-Noise Ratios (SNRs) with 1000 samples. SNR is a measurement of signal noise level as denoted $Signal/Noise$, where the higher SNR, the less noise is included in a signal.

Signal-512 is a private dataset that considers several non-ideal effects of communication systems, including carrier phase, pulse shaping, frequency offsets, and noise. Each data sample contains 64 symbols, and the oversampling rate is 8, so the number of sampling points for each sample is 512.

Signal-1024 is a public over-the-air radio signal which provides 24 types of digital and analog modulations. Each modulation type includes 26 diverse SNRs, and each SNR with 4096 samples.

3.2. Experiment setup

Evaluation Models. The proposed MSP framework was implemented in Pytorch [18] and trained on a Tesla V100. We study the performance of our proposed MSP framework using six existing FSL models, namely, MAML, MatchNet, ProtoNet, RelatNet, TIM-GD, TIM-ADM [19] and REnet.

Experiment setting. Following previous studies [20], we conduct our experiments in 5-way-5-shot and 5-way-1-shot settings, i.e., five novel classes, each of them only has 5 and 1 samples, respectively. For each dataset, two different SNR conditions are studied: -4 dB, and 18 dB.

Training setting. Our framework is optimized by Adam optimizer [21] with a learning rate of 0.001. The maximum training epoch is set to 50. The input batch size is set to 70 and 42 for the 5-way-5-shot and 5-way-1-shot, respectively.

3.3. Experimental results

Table 1 demonstrates that our proposed MSP framework can effectively boost the performance of the current SOTA FSL methods on both 5-way-1-shot and 5-way-5-shot modulation recognition tasks with different backbone structures. Furthermore, we observe that the RelatNet with our MSP framework

Backbone	Method	*Signal-128*				*Signal-512*				*Signal-1024*			
		1-shot		5-shot		1-shot		5-shot		1-shot		5-shot	
		Origin	+MSP	Origin	+MSP	Origin	+MSP	Origin	+MSP	Origin	+MSP	Origin	+MSP
CNN	MAML	40.67%	+ 21.33%	61.93%	+ 13.57%	25.27%	+ 2.31%	27.52%	+ 2.39%	69.45%	+ 4.93%	80.43%	+ 2.93%
	MatchNet	70.95%	+ 9.54%	73.97%	+ 9.46%	41.45%	+ 0.53%	42.33%	+ 2.43%	84.33%	+ 4.16%	87.41%	+ 5.36%
	ProtoNet	71.61%	+ 10.23%	76.59%	+ 10.45%	39.09%	- 0.28%	41.92%	+ 3.54%	82.17%	+ 6.73%	85.75%	+ 4.58%
	RelatNet	67.73%	+ 13.76%	73.42%	+ 8.79%	37.76%	+ 0.83%	39.63%	+ 3.00%	80.11%	+ 7.35%	80.89%	+ 8.28%
	Tim-GD	72.23%	+ 6.45%	78.47%	+ 6.56%	39.79%	- 0.19%	41.68%	+ 3.99%	83.22%	+ 3.61%	87.61%	+ 5.08%
	Tim-ADM	73.49%	+ 6.96%	79.36%	+ 7.21%	42.23%	+ 0.10%	42.81%	+ 2.73%	82.19%	+ 4.29%	88.15%	+ 4.74%
	Renet	69.77%	+ 4.28%	73.58%	+ 3.93%	37.58%	+ 0.33%	40.26%	+ 0.64%	81.75%	+ 2.46%	88.19%	+ 5.49%
ResNet	MAML	38.39%	+ 15.48%	63.41%	+ 3.45%	27.45%	+ 0.76%	29.92%	+ 9.38%	67.00%	+ 6.35%	79.98%	+ 4.31%
	MatchNet	69.55%	+ 12.89%	74.14%	+ 8.34%	40.75%	+ 1.39%	42.55%	+ 3.74%	80.92%	+ 6.31%	83.34%	+ 8.77%
	ProtoNet	71.27%	+ 8.58%	77.67%	+ 7.26%	40.27%	+ 0.52%	42.79%	+ 3.06%	81.05%	+ 5.92%	82.12%	+ 11.70%
	RelatNet	62.47%	+ 16.15%	70.39%	+ 14.93%	39.23%	+ 0.41%	37.47%	+ 3.67%	70.15%	+ 12.97%	80.45%	+ 7.78%
	Tim-GD	72.66%	+ 13.34%	78.49%	+ 6.35%	41.73%	+ 0.87%	43.29%	+ 4.69%	86.38%	+ 3.78%	87.13%	+ 7.50%
	Tim-ADM	73.58%	+ 12.82%	79.44%	+ 7.25%	42.21%	+ 0.28%	43.72%	+ 3.75%	86.95%	+ 4.37%	88.37%	+ 6.89%
	Renet	68.74%	+ 2.53%	73.21%	+ 2.35%	38.65%	+ 0.32%	42.45%	+ 1.55%	83.85%	+ 1.67%	87.69%	+ 5.19%

Table 1: 5-way few-shot classification accuracy on signal-128, signal-512 and signal-1024 dataset. “Origin” represents the few-shot learning result without the MSP framework, while “+MSP” denotes the few-shot learning result with the MSP framework. Note: ‘+’ and ‘-’indicates the performance gain and drop respectively. The top performance gain is in ‘_’ underline format.

achieves six top performance gains with an average growth of 8.16% in 12 tasks.

Moreover, we observe the performance improvement is very slight in the 5-way-1-shot task with the signal-512 dataset, even experiencing about 0.28% performance drop for ProtoNet with CNN backbone. One possible conjecture is that large intra-class gaps in the signal-512 dataset make statistical properties learning of different categories from a single sample more challenging. This conjecture is further proved that when we increase the number of samples (i.e., 5-way-5-shot), the performance gain on the signal-512 dataset is improved.

3.4. Ablation study on MSP framework

To study the effectiveness of the two components in the MSP framework, we conduct experiments on every single component and their combinations. Figure 3 reports the detailed results of the ablation studies, and we can see that each component/module in the MSP framework contributes positively to the final results. Moreover, the combination of noise filtering and augmentation module always performs the best compared with the single component. Besides, we also observe that the contribution of the noise filtering module is marginal under high SNR scenarios. That could result from the clean signals provided in high SNR scenarios, where filtering is no more required.

4. CONCLUSIONS

We propose MSP, a novel radio signal pre-processing framework, which can be attached to various SOTA FSL models for the modulation recognition task. Specifically, the MSP framework utilizes online Info-preserved augmentations to generate diverse signal segments and remove the carried noises. Finally, a feature enhancement module simplifies the signal representations to encourage representation learning. Extensive experiments result presents the effectiveness of the proposed MSP framework.

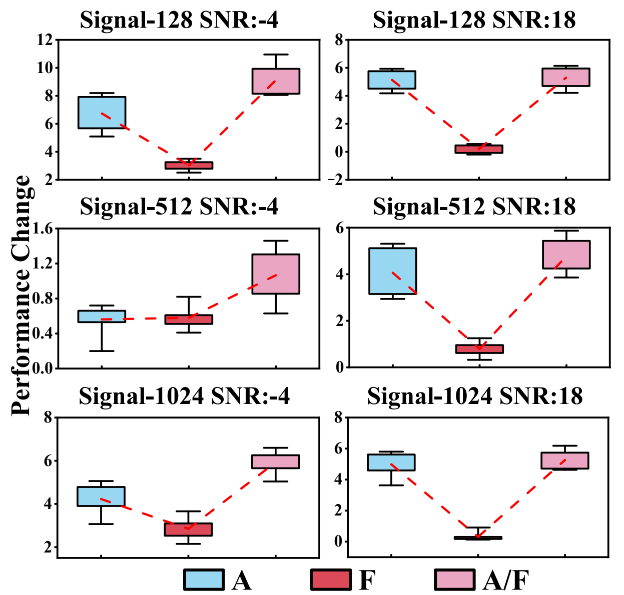

Fig. 3: Ablation study on 5-way-5-shot task for three datasets under two different SNR scenarios. ‘A’, and ‘F’ represents augmentation and noise filtering, respectively. The ‘A/F’ indicates the combination of augmentation and noise filtering. The dotted line indicates the average performance change

Acknowledgement

This work was supported in part by the National Natural Sciences Foundation of China under Grant No. 62101494 and No. 62072408.

5. REFERENCES

[1] Timothy J O’Shea, Johnathan Corgan, and T Charles Clancy, “Convolutional radio modulation recognition networks,” in *International conference on engineer-*

ing applications of neural networks. Springer, 2016, pp. 213–226.

[2] Oriol Vinyals, Charles Blundell, Timothy Lillicrap, Daan Wierstra, et al., "Matching networks for one shot learning," *Advances in neural information processing systems*, vol. 29, 2016.

[3] F. Sung, Y. Yang, Z. Li, X. Tao, and T. M. Hospedales, "Learning to compare: Relation network for few-shot learning," in *2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 2018.

[4] Jake Snell, Kevin Swersky, and Richard Zemel, "Prototypical networks for few-shot learning," *Advances in neural information processing systems*, vol. 30, 2017.

[5] Dahyun Kang, Heeseung Kwon, Juhong Min, and Minsu Cho, "Relational embedding for few-shot classification," in *Proceedings of the IEEE/CVF International Conference on Computer Vision*, 2021, pp. 8822–8833.

[6] Chelsea Finn, Pieter Abbeel, and Sergey Levine, "Model-agnostic meta-learning for fast adaptation of deep networks," in *International conference on machine learning*. PMLR, 2017, pp. 1126–1135.

[7] Eleni Triantafillou, Tyler Zhu, Vincent Dumoulin, Pascal Lamblin, Utku Evci, Kelvin Xu, Ross Goroshin, Carles Gelada, Kevin Swersky, Pierre-Antoine Manzagol, et al., "Meta-dataset: A dataset of datasets for learning to learn from few examples," *arXiv preprint arXiv:1903.03096*, 2019.

[8] Kai Li, Yulun Zhang, Kunpeng Li, and Yun Fu, "Adversarial feature hallucination networks for few-shot learning," in *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 2020, pp. 13470–13479.

[9] Eli Schwartz, Leonid Karlinsky, Joseph Shtok, Sivan Harary, Mattias Marder, Abhishek Kumar, Rogerio Feris, Raja Giryes, and Alex Bronstein, "Delta-encoder: an effective sample synthesis method for few-shot object recognition," *Advances in Neural Information Processing Systems*, vol. 31, 2018.

[10] Quan Zhou, Ronghui Zhang, Junsheng Mu, Hongming Zhang, Fangpei Zhang, and Xiaojun Jing, "Amcrn: Few-shot learning for automatic modulation classification," *IEEE Communications Letters*, vol. 26, no. 3, pp. 542–546, 2021.

[11] Zilin Zhang, Yan Li, and Meiguo Gao, "Few-shot learning of signal modulation recognition based on attention relation network," in *2020 28th European Signal Processing Conference (EUSIPCO)*. IEEE, 2021, pp. 1372–1376.

[12] Jovana Mitrovic, Brian McWilliams, and Melanie Rey, "Less can be more in contrastive learning," 2020.

[13] Ting Chen, Simon Kornblith, Mohammad Norouzi, and Geoffrey Hinton, "A simple framework for contrastive learning of visual representations," in *International conference on machine learning*. PMLR, 2020, pp. 1597–1607.

[14] JeanFrançois Dufourd, "Polyhedra genus theorem and euler formula: A hypermap-formalized intuitionistic proof," *Theoretical Computer Science*, vol. 403, no. 2, pp. 133–159, 2008.

[15] Jan Issac, Manuel Wüthrich, Cristina Garcia Cifuentes, Jeannette Bohg, Sebastian Trimpe, and Stefan Schaal, "Depth-based object tracking using a robust gaussian filter," in *2016 IEEE International Conference on Robotics and Automation (ICRA)*. IEEE, 2016, pp. 608–615.

[16] Dongdong Chen, Julián Tachella, and Mike E Davies, "Equivariant imaging: Learning beyond the range space," in *Proceedings of the IEEE/CVF International Conference on Computer Vision*, 2021, pp. 4379–4388.

[17] Timothy J O'Shea, T. Roy, and T. C. Clancy, "Over the air deep learning based radio signal classification," *IEEE Journal of Selected Topics in Signal Processing*, vol. PP, no. 99, pp. 1–1, 2017.

[18] Adam Paszke, Sam Gross, Francisco Massa, Adam Lerer, James Bradbury, Gregory Chanan, Trevor Killeen, Zeming Lin, Natalia Gimelshein, Luca Antiga, et al., "Pytorch: An imperative style, high-performance deep learning library," *Advances in neural information processing systems*, vol. 32, 2019.

[19] Malik Boudiaf, Imtiaz Ziko, Jérôme Rony, José Dolz, Pablo Piantanida, and Ismail Ben Ayed, "Information maximization for few-shot learning," *Advances in Neural Information Processing Systems*, vol. 33, pp. 2445–2457, 2020.

[20] Kai Li, Yulun Zhang, Kunpeng Li, and Yun Fu, "Adversarial feature hallucination networks for few-shot learning," in *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 2020, pp. 13470–13479.

[21] Zijun Zhang, "Improved adam optimizer for deep neural networks," in *2018 IEEE/ACM 26th International Symposium on Quality of Service (IWQoS)*. IEEE, 2018, pp. 1–2.