ISSN 2616-5775 Vol. 7, Issue 2: 12-18, DOI: 10.25236/AJCIS.2024.070202

A Mixup-based Margin Aware and Calibration Model for Imbalanced in Soil Classification

Xin Zheng[1,a,*], Xin Bai[1,b]

[1]*School of Information Science and Engineering, Chongqing Jiaotong University, Chongqing, China*
[a]*823495617@qq.com*, [b]*230326401@qq.com*
[*]*Corresponding author*

Abstract: *In the process of soil image classification, the issue of class imbalance occurs, which leads to a decline in the generalization performance of the classifier due to the lack of data from minority classes. We investigated the effectiveness of Mixup through margin statistical analysis and successfully improved the deep imbalanced classification with uneven margins. Additionally, we investigated the relationship between margins and logits, and empirically discovered that uncalibrated margins exhibit a positive correlation with logits. Based on this revelation, we propose a Mixup-based margin-aware and calibration model to address the challenge of handling imbalanced soil image classification data. We conducted experiments using the Soil dataset and additionally tested the generalization capabilities of our method on the CIFAR10-LT, CIFAR100-LT, and ImageNet-LT datasets. The experimental results indicate that our approach achieved impressive results.*

Keywords: *Imbalanced Classification; Soil Classification; Mixup; Margin*

1. Introduction

This paper collects six types of soil samples from a region in southeastern China, including clay, silty clay, muddy clay, muddy silty clay, silty, and gravel. After the laboratory testing personnel conducted physical and chemical analysis on the soil samples, they named and classified the samples to create a soil image classification dataset. This dataset suffers from an imbalance problem, as the number of samples for each class differs significantly. If the dataset used by the model is imbalanced, it tends to bias towards the majority class and often has poor generalization ability when identifying minority classes.

Long-tailed class imbalance is a common problem in image classification tasks, which often limits the practicality of recognition models based on deep networks in real-world applications. This is because they tend to bias towards the dominant classes and perform poorly on the tail classes. To address the long-tailed class imbalance problem, there are several strategies, which usually involve re-sampling or re-weighting techniques.

Motivated by LDAM[1], We examined the effectiveness of Mixup[2] through margin statistical analysis, and LDAM has successfully addressed the depth-imbalanced classification with uneven margins. Experimental results indicate that the gap between the majority and minority classes is loosely correlated with the accuracy of depth-imbalanced classification. We found that Mixup implicitly reduces the margin gap, providing a theoretical basis for its effectiveness. Building on this theoretical foundation, we discover that during the first stage of standard training, the margin gap can be more explicitly tuned by implementing margin-aware Mixup. The Margin-Aware Mixup (MAM) approach we have developed sets a new benchmark in performance on established tests for imbalanced classification, overtaking both Mixup and LDAM with considerable margins, especially in environments marked by severe data imbalance.

Furthermore, we delved into the correlation between margin and logit values, undertaking numerous experiments which revealed that both are associated with the quantity of images per class. Specifically, prior to any calibration efforts, it was observed that classes featuring a higher volume of images exhibited larger margins and logits. Conversely, classes with fewer images displayed reduced margins and logits[3]. Moreover, we found that uncalibrated margins and logits will substantially detract from image classification performance. To address this issue, we integrated margin calibration into our experimental setup to achieve logits that are more evenly balanced. Following the extraction of representations and classifier heads through standard training processes, we implemented a straightforward class-specific model for margin calibration.

In summary, our primary contributions are encapsulated in the following points:

1) We undertook a comprehensive investigation into uncalibrated margins from the standpoint of margin analysis.

2) Drawing inspiration from Mixup, we formulated a Margin-Aware Mixup (MAM) model.

3) We unveiled a specialized Margin Calibration (MC) model designed for meticulous margin adjustment.

4) By integrating MAM and MC at various stages of training, we developed an innovative framework named MAMMC.

2. Related Work

2.1. Method Based on Re-Sampling

Traditional deep network training is based on random sampling of small batches for gradient descent, which neglects the imbalance problem in long-tailed learning. Common methods for addressing imbalanced classification involve under-sampling the majority classes or over-sampling the minority classes. Kang et al.[4] suggested that Decoupling, in conjunction with square root sampling and progressive balancing sampling, enhances training strategies for addressing long-tailed recognition challenges. Nonetheless, the implementation of these strategies necessitates prior knowledge regarding the frequency of training samples across various categories. To tackle this challenge, Feng et al.[5] introduced an innovative adaptive sampling strategy termed LOCE. This method tracks the model's training performance across various classes by leveraging average class prediction scores and employs these scores to adjust the sampling rates for different classes accordingly. Furthermore, Zang et al.[6] put forward a meta-learning-oriented approach named FASA. This strategy employs the model's classification loss on a balanced meta-validation set as a metric to fine-tune the sampling rates for various classes. Consequently, it increases the sampling frequency from underrepresented tail classes to address imbalance more effectively.

2.2. Method Based on Re-Weighting

To mitigate the challenges posed by class imbalance, re-weighting techniques modify the training loss values for different classes through multiplication with unique weights tailored to each class[7]. The most straightforward approach is to directly use the label frequency of training samples for loss reweighting, thereby rebalancing the uneven class gradients. Cui et al.[8] introduced a Class-Balanced (CB) loss, which recalibrates the weight of each class based on its effective number of instances. This method amplifies the significance of minority instances within the loss function, thereby averting the underrepresentation of less populated classes. Ren et al.[9] proposed the Balanced Softmax technique, which alleviates the class imbalance bias by adjusting the predicted logits using the label frequencies before calculating the final loss. Lin et al.[10] put forward the concept of Focal Loss, which diverges from using training label frequency and instead focuses on recalibrating weights based on the difficulty of predicting each class. To address the issue of overfitting, Cao et al.[1] developed a theoretically grounded approach known as Label Distribution-Aware Margin (LDAM) loss, aimed at reducing the generalization bound through margin optimization. Additionally, they proposed a Delayed Reweighting (DRW) technique to further improve model performance.

2.3. Method Based on Re-Margining

In response to class imbalance, re-margining endeavors to mitigate this challenge by fine-tuning the loss function through the subtraction of unique margin factors tailored to each class. This methodology enables the calculation of distinct minimum margins between features and classifiers for every category, effectively addressing disparity among classes. Cao et al.[11] introduced an innovative frequency metric grounded in inter-class feature compactness, subsequently employing this metric to re-margining the feature space of the tail domain. Nevertheless, while this re-margining technique aimed at enlarging the margin for tail classes might effectively address certain issues, it also carries the risk of diminishing feature learning capabilities within head classes. Wu et al.[12] employed a scale-invariant classifier and re-balanced the data by utilizing margin engineering during the training phase and margin adjustment

ISSN 2616-5775 Vol. 7, Issue 2: 12-18, DOI: 10.25236/AJCIS.2024.070202

during the inference phase to enhance the adversarial robustness under the long-tailed distribution. The decision margin adjustment technique refines the classifier's head in a learnable manner post-standard training, leveraging strategies such as maximum likelihood estimation among others. Nonetheless, prevailing methods of decision margin adjustment overlook the crucial aspect of margin calibration. Our objective is centered on fine-tuning the margin to achieve equitable prediction outcomes.

3. Method

We propose MAMMC, A Mixup-based Margin Aware and Calibration Model, specifically designed to tackle imbalance issues prevalent in soil image datasets. Drawing inspiration from Mixup, we infuse the concept of variable margins into this framework through our Margin-Aware-Mixup (MAM). Subsequently, to refine these margins post-training, we apply an Margin Calibration (MC) during the latter training phase for precise margin adjustment.

3.1. Margin-Aware-Mixup

We propose to integrate the concept of uneven margins into the Mixup-based data augmentation technique. Given a training sample x and its corresponding label y, the margin of (x,y) is defined as follows:

$$\gamma(x,y) = f(x)_y - \max_{j \neq y} f(x)_j \tag{1}$$

LDAM proposes a loss function that encourages class-dependent margins to address the class imbalance problem. Theoretically, it has been derived that the ideal margin for each class is proportional to the ratio of positive to negative samples in that class.

$$\sigma_j = C / n_j^{1/4} \tag{2}$$

where C is a constant. Theoretically, this loss function encourages larger margins for the minority class. Assuming (x_i,y_i) and (x_j,y_j) are two samples from different classes, the distance between the decision margin between class i and class j for sample xi is defined as follows:

$$\kappa_i = 1/n_i^{\omega}; \kappa_j = 1/n_j^{\omega} \tag{3}$$

where ω is a hyper-parameter that needs to be tuned to achieve the optimal balance in the proposed margin-aware mixing technique. Experimental analysis has been conducted on this hyper-parameter.

$$\lambda_y = \begin{cases} 1 - \dfrac{1-\lambda_x}{2\kappa_i/(\kappa_i+\kappa_j)}, \lambda_x \geq \lambda_{xbest} \\ \dfrac{\lambda_x}{2\kappa_j/(\kappa_i+\kappa_j)}, \lambda_x < \lambda_{xbest} \end{cases} \tag{4}$$

where λ_x and the Mixup-selected pairs (x_i,y_i) and (x_j,y_j) is similar to the original Mixup, and $\lambda_y \in [0,1]$, $\lambda_{xbest} = \kappa_j/(\kappa_i+\kappa_j)$. If the mixing factor $\lambda_x = \kappa_j/(\kappa_i+\kappa_j)$, The probability of outputting the synthesized example for both class i and class j should be 0.5.

Finally, we propose the Margin-Aware-Mixup formula as follows:

$$\begin{aligned} \widetilde{x} &= \lambda_x x_i + (1-\lambda_x) x_j \\ \widetilde{y} &= \lambda_y y_i + (1-\lambda_y) y_j \end{aligned} \tag{5}$$

In the original Mixup, the mixing factor for synthesizing x and y are the same, that is $\lambda_x = \lambda_y$. The core idea of Mixup for addressing class imbalance is to obtain different y. We propose a method that achieves this idea by combining margin-aware.

3.2. Margin Calibration

To obtain the calibrated logits, MC calibrates the margins after standard training. Specifically, a simple class margin calibration model is used during training, with the original margins remaining fixed:

ISSN 2616-5775 Vol. 7, Issue 2: 12-18, DOI: 10.25236/AJCIS.2024.070202

$$\hat{d}_j = \alpha_j \cdot d_j + \beta_j \tag{6}$$

where α_j and β_j is the learnable parameters of class j. In the experiment, the calibration is extended from class j to other classes, and finally to the entire dataset. Therefore, the calibrated logit is calculated as:

$$\|W_j\|\hat{d}_j = \|W_j\|(\alpha_j \cdot d_j + \beta_j) = \alpha_j \cdot \eta_j + \beta_j \cdot \|W_j\| \tag{7}$$

Where η_j are the initial fixed logits. Then, we can obtain the calibrated prediction distribution:

$$p(y = y_i \mid x_i; \theta_r, \theta_c) = \frac{\exp(\alpha_{y_i} \cdot \eta_{y_i} + \beta_{y_i} \cdot \|W_{y_i}\|)}{\sum_{j=1}^{K} \exp(\alpha_j \cdot \eta_j + \beta_j \cdot \|W_j\|)} \tag{8}$$

Moreover, to obtain a more balanced gradient during the training process, we re-weighting the loss as in previous studies. The weight of the yi class is calculated as follows:

$$U_{y_i} = K \cdot \frac{(1/n_{y_i})^{\gamma}}{\sum_{j=1}^{K} (1/n_j)^{\gamma}} \tag{9}$$

Whereγis the scale hyper-parameter. Whenγ=0, the weights of all classer are set to 1, indicating that there is no need for re-weighting. Similar to previous studies, we propose a new loss function for class re-weighting, and the loss for the margin calibration model is:

$$L(x_i, y_i; \widetilde{\theta}_r, \widetilde{\theta}_c, \alpha, \beta) = -U_{y_i} \cdot \log\left(\frac{\exp(\alpha_{y_i} \cdot \eta_{y_i} + \beta_{y_i} \cdot \|W_{y_i}\|)}{\sum_{j=1}^{K} \exp(\alpha_j \cdot \eta_j + \beta_j \cdot \|W_j\|)} \right) \tag{10}$$

Where $\widetilde{\theta}_r$ and $\widetilde{\theta}_c$ parameter is frozen in the training process. The entire training process consists of two stages: standard training and margin calibration model training.

4. Experiments

4.1. Datasets

Our experiments utilize three public datasets, including CIFAR-10-LT, CIFAR-100-LT, ImageNet-LT. The imbalance ratio used in the CIFAR dataset is defined as Nmax/Nmin, where Nmax is the number of samples in the largest class, and Nmin is the number of samples in the smallest class. For the CIFAR dataset, experiments were conducted with two imbalance ratios, 100 and 200. For the ImageNet-LT dataset, the classes are further divided into three groups based on the number of images: Many-shot (over 100 images), Medium-shot (20 to 100 images), and Few-shot (fewer than 20 images). Ultimately, we deploy our proposed model on the Soil dataset for evaluation.

4.2. Implementation Details

For the CIFAR-10-LT, CIFAR-100-LT, and Soil datasets, we used ResNet32 as the backbone network. The standard training phase lasted for 200 epochs, and the margin calibration phase consisted of 10 epochs. We applied a cosine learning rate scheduler with an initial learning rate that gradually decreased from 0.05 to 0, and the batch size was 256. For the ImageNet-LT dataset, we selected ResNet50 as the backbone network, and the standard training phase consisted of 100 epochs. The margin calibration phase also consisted of 10 epochs, using a cosine learning rate scheduler with an initial learning rate that gradually decreased from 0.1 to 0. The batch size was 128. During the standard training phase, we optimized the model using stochastic gradient descent (SGD) with a momentum of 0.9, weight decay of 2e-4, and an initial learning rate of 0.1. Theγwas set to 1.2. GPU using NVIDIA GeForce RTX 3090 with 24G memory, PyTorch 1.9.0 , Python 3.8, and CUDA 11.1.

4.3. Ablation study

Table 1 presents the experimental results on the CIFAR-10-LT, CIFAR-100-LT, and Soil datasets, with the imbalance ratio set to 100 for CIFAR-10-LT and CIFAR-100-LT. MAMMC improves the accuracy by 6.7%, 7.0%, and 8.2% respectively compared to standalone MAM. Additionally, MAMMC outperforms MAM-DRW, which combines MAM with delayed reweighting, by 1.4%, 4.4%, and 1.5% respectively in terms of accuracy. The experimental results indicate that the fusion embedding of MAMMC during different training stages is superior to other standalone training models, as well as the fusion model with delayed re-weighting.

Table 1: Accuracy on CIFAR-LT and Soil with imbalance ratio of 100.

Method	Top-1 Accuracy (%)		
	CIFAR-10-LT	CIFAR-100-LT	Soil
ResNet32	75.1	42.4	70.7
ResNet32+MAM	76.6	42.8	72.3
ResNet32+MC	81.5	45.2	77.9
ResNet32+MAM-DRW	81.9	46.3	79.0
ResNet32+MAMMC	**83.3**	**50.7**	**80.5**

4.4. Comparisons with State-of-the-Art Methods

4.4.1. Experiments on CIFAR-LT

Table 2 presents the experimental results on the CIFAR-10-LT and CIFAR-100-LT datasets, with imbalance ratio set to 100 and 200, respectively. Compared to other re-sampling, re-weighting, and re-margining methods, our proposed method MAMMC demonstrates superior performance than other methods in the table. The accuracy of MAMMC on CIFAR-10-LT(100), CIFAR-10-LT(200), CIFAR-100-LT(100), and CIFAR-100-LT(200) surpasses Balanced Softmax by 0.2%, 0.3%, 0.4%, and 0.6%, respectively, with the numbers in parentheses indicating the imbalance ratios.

Table 2: Accuracy on CIFAR-LT with imbalance ratio of 100 and 200.

Method	Top-1 Accuracy (%)			
	CIFAR-10-LT		CIFAR-100-LT	
	100	200	100	200
Focal Loss[10]	77.1	71.8	43.8	40.2
Mixup-DRW[2]	82.0	78.3	47.5	43.5
Class Balanced Loss[8]	78.2	72.6	44.6	39.9
LDAM-DRW[1]	77.5	73.8	41.3	37.1
Decouple-cRT[4]	82.0	76.6	50.0	44.5
Equalization Loss[13]	78.5	74.6	47.4	43.3
Balanced Softmax[9]	83.1	79.0	50.3	45.9
BBN[14]	79.8	-	42.6	-
Remix-DRW[15]	82.3	77.8	46.0	42.8
LADE[16]	81.8	76.9	45.4	43.6
DisAlign[17]	78.0	71.2	49.1	43.6
Hybrid-SC[18]	81.4	-	46.7	-
RISDA[19]	79.9	74.0	50.1	44.7
LOBM[20]	78.7	-	46.2	-
MAMMC	**83.3**	**79.3**	**50.7**	**46.5**

4.4.2. Experiments on ImageNet-LT

Table 3 presents the experimental results on the ImageNet-LT dataset, where the classes are divided into three groups based on the number of images: Many-shot (over 100 images), Medium-shot (20 to 100 images), and Few-shot (fewer than 20 images). MAMMC demonstrates superior performance compared to all re-weighting methods. When using all category images, MAMMC improves the accuracy by 0.2% compared to MARC. In the Many-shot group, it performs lower than methods such as Seesaw; in the Medium-shot group, it improves the accuracy by 0.3% compared to DisAlign; in the Few-shot group, it improves the accuracy by 0.3% compared to MARC.

ISSN 2616-5775 Vol. 7, Issue 2: 12-18, DOI: 10.25236/AJCIS.2024.070202

Table 3: Accuracy on ImageNet-LT with different groups.

Method	Top-1 Accuracy (%)			
	All	Many	Medium	Few
Focal Loss[10]	43.7	64.3	37.1	8.2
LDAM-DRW[1]	49.8	60.4	46.9	30.7
Decouple-cRT[4]	47.3	58.8	44.0	26.1
Balanced Softmax[9]	51.4	62.2	48.8	29.8
LADE[16]	51.9	62.3	49.3	31.2
DisAlign[17]	52.2	60.8	50.4	34.7
Seesaw[21]	50.4	**67.1**	45.2	21.4
MARC[3]	52.3	60.4	50.3	36.6
MAMMC	**52.5**	62.5	**50.7**	**36.9**

4.5. Hyper-Parameter Analysis

To analyze the role of the scale hyper-parameter γ in the margin calibration model, experiments were conducted on the CIFAR-10-LT and CIFAR-100-LT datasets with an imbalance ratio of 100. The experimental results are shown in Figure 1. The horizontal axis represents the value of the scale hyper-parameter γ, and the vertical axis represents the Top-1 Accuracy. From the figure, it can be observed that when γ=0, the accuracy of CIFAR-10-LT and CIFAR-100-LT is 80.1% and 48.2%, respectively. At this point, no loss re-weighting techniques are used during training, and the model remains effective. When γ=1.2, the accuracy of CIFAR-10-LT and CIFAR-100-LT is 83.3% and 50.7%, respectively. Therefore, when training other datasets, we directly use the setting with γ=1.2.

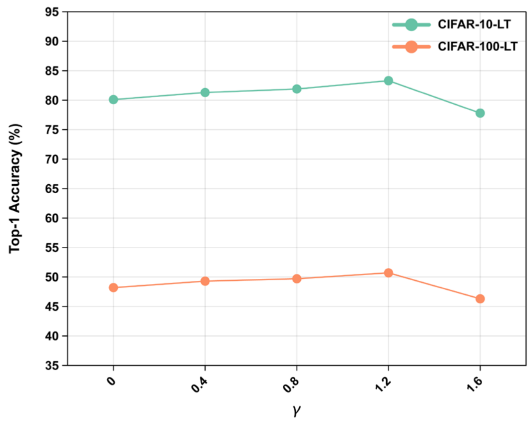

Figure 1: Accuracy on CIFAR-LT with different scale hyper-parameterγ.

5. Conclusions

We utilize margin statistical approaches to assess the model's proficiency in acquiring precise representations from a margin perspective within a class-imbalanced learning setting. Introducing the Margin-Aware Mixup (MAM), we harness Mixup technology to establish variable edges deliberately. In typical imbalanced classification contexts, the original Mixup technique inadvertently results in uneven margin. By explicitly manipulating the extent of this edge imbalance, we fine-tune the model’s performance. Furthermore, we employ a MC margin calibration model to obtain balanced logits in long-tailed visual recognition, achieving good results without altering the model's representation. Ultimately, our study amalgamates MAM and MC across varying phases of training, thereby introducing a groundbreaking model framework dubbed MAMMC. Through this margin-aware and calibration approach based on Mixup, the problem of soil class imbalance can be effectively addressed, thereby improving the model's performance in practical applications. Moreover, this method can be applied to other domains with imbalanced classification issues, providing a new perspective for solving such problems.

ISSN 2616-5775 Vol. 7, Issue 2: 12-18, DOI: 10.25236/AJCIS.2024.070202

References

[1] K. Cao, C. Wei, A. Gaidon et al. Learning imbalanced datasets with label-distribution-aware margin loss, Advances in neural information processing systems, vol. 32, 2019.
[2] H. Zhang, M. Cisse, Y. N. Dauphin et al. mixup: Beyond empirical risk minimization, arXiv preprint arXiv:1710.09412, 2017.
[3] Y. Wang, B. Zhang, W. Hou et al. Margin calibration for long-tailed visual recognition, in Asian Conference on Machine Learning, 2023: PMLR, pp. 1101-1116.
[4] B. Kang, S. Xie, M. Rohrbach et al. Decoupling representation and classifier for long-tailed recognition, arXiv preprint arXiv:1910.09217, 2019.
[5] C. Feng, Y. Zhong, and W. Huang Exploring classification equilibrium in long-tailed object detection, IEEE/CVF International conference on computer vision, 2021, pp. 3417-3426.
[6] Y. Zang, C. Huang, and C. C. Loy Fasa: Feature augmentation and sampling adaptation for long-tailed instance segmentation, IEEE/CVF International Conference on Computer Vision, 2021, pp. 3457-3466.
[7] Y. Zhang, B. Kang, B. Hooi et al. Deep long-tailed learning: A survey, IEEE Transactions on Pattern Analysis and Machine Intelligence, 2023.
[8] Y. Cui, M. Jia, T.-Y. Lin et al. Class-balanced loss based on effective number of samples, IEEE/CVF conference on computer vision and pattern recognition, 2019, pp. 9268-9277.
[9] J. Ren, C. Yu, X. Ma et al. Balanced meta-softmax for long-tailed visual recognition, Advances in neural information processing systems, vol. 33, pp. 4175-4186, 2020.
[10] T.-Y. Lin, P. Goyal, R. Girshick et al. Focal loss for dense object detection, IEEE international conference on computer vision, 2017, pp. 2980-2988.
[11] D. Cao, X. Zhu, X. Huang et al. Domain balancing: Face recognition on long-tailed domains, IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2020, pp. 5671-5679.
[12] T. Wu, Z. Liu, Q. Huang et al. Adversarial robustness under long-tailed distribution, IEEE/CVF conference on computer vision and pattern recognition, 2021, pp. 8659-8668.
[13] J. Tan, C. Wang, B. Li et al. Equalization loss for long-tailed object recognition, IEEE/CVF conference on computer vision and pattern recognition, 2020, pp. 11662-11671.
[14] B. Zhou, Q. Cui, X.-S. Wei et al. Bbn: Bilateral-branch network with cumulative learning for long-tailed visual recognition, IEEE/CVF conference on computer vision and pattern recognition, 2020, pp. 9719-9728.
[15] H.-P. Chou, S.-C. Chang, J.-Y. Pan et al. Remix: rebalanced mixup, in Computer Vision–ECCV 2020 Workshops: Glasgow, UK, August 23–28, 2020, Proceedings, Part VI 16, 2020: Springer, pp. 95-110.
[16] Y. Hong, S. Han, K. Choi et al. Disentangling label distribution for long-tailed visual recognition, IEEE/CVF conference on computer vision and pattern recognition, 2021, pp. 6626-6636.
[17] S. Zhang, Z. Li, S. Yan et al. Distribution alignment: A unified framework for long-tail visual recognition, IEEE/CVF conference on computer vision and pattern recognition, 2021, pp. 2361-2370.
[18] P. Wang, K. Han, X.-S. Wei et al. Contrastive learning based hybrid networks for long-tailed image classification, IEEE/CVF conference on computer vision and pattern recognition, 2021, pp. 943-952.
[19] X. Chen, Y. Zhou, D. Wu et al. Imagine by reasoning: A reasoning-based implicit semantic data augmentation for long-tailed classification, AAAI Conference on Artificial Intelligence, 2022, vol. 36, no. 1, pp. 356-364.
[20] S. Zhang, C. Chen, X. Zhang et al. Label-occurrence-balanced mixup for long-tailed recognition, in ICASSP 2022-2022 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), 2022: IEEE, pp. 3224-3228.
[21] J. Wang, W. Zhang, Y. Zang et al. Seesaw loss for long-tailed instance segmentation, IEEE/CVF conference on computer vision and pattern recognition, 2021, pp. 9695-9704.