



This MICCAI paper is the Open Access version, provided by the MICCAI Society. It is identical to the accepted version, except for the format and this watermark; the final published version is available on SpringerLink.

# Iterative Online Image Synthesis via Diffusion Model for Imbalanced Classification

Shuhan Li<sup>†</sup>, Yi Lin<sup>†</sup>, Hao Chen, and Kwang-Ting Cheng

The Hong Kong University of Science and Technology, Hong Kong, China  
slidm@connect.ust.hk

**Abstract.** Accurate and robust classification of diseases is important for proper diagnosis and treatment. However, medical datasets often face challenges related to limited sample sizes and inherent imbalanced distributions, due to difficulties in data collection and variations in disease prevalence across different types. In this paper, we introduce an Iterative Online Image Synthesis (IOIS) framework to address the class imbalance problem in medical image classification. Our framework incorporates two key modules, namely Online Image Synthesis (OIS) and Accuracy Adaptive Sampling (AAS), which collectively target the imbalance classification issue at both the instance level and the class level. The OIS module alleviates the data insufficiency problem by generating representative samples tailored for online training of the classifier. On the other hand, the AAS module dynamically balances the synthesized samples among various classes, targeting those with low training accuracy. To evaluate the effectiveness of our proposed method in addressing imbalanced classification, we conduct experiments on the HAM10000 and APTOS datasets. The results obtained demonstrate the superiority of our approach over state-of-the-art methods as well as the effectiveness of each component. The source code is available at [https://github.com/ustlsh/IOIS\\_imbalance](https://github.com/ustlsh/IOIS_imbalance).

**Keywords:** Image synthesis · Imbalanced classification · Diffusion model.

## 1 Introduction

Image classification is an essential task for medical image analysis and has wide applications in medical datasets, such as distinguishing benign or malignant tumors [4, 16], grading specific diseases [26, 27], and diagnosing various diseases [1, 13]. However, medical datasets often encounter challenges related to insufficient sample sizes or inherent imbalanced distributions, which can be attributed to difficulties in data collection and variations in disease prevalence across different types [2, 15]. These challenges can lead to poor generalization and introduce biases, compromising the performance of deep learning models.

To address the challenges of class imbalance, existing solutions can be categorized into three main groups: re-weighting, re-sampling, and data synthesis.

---

<sup>†</sup> Equal contribution;  corresponding author.

Re-weighting methods [5, 14, 22, 23] aim to balance the loss between the majority and minority classes. An example is Focal Loss [14], which adjusts hyperparameters to increase the loss for minority classes while decreasing it for majority classes. This approach encourages the model to focus more on challenging samples. On the other hand, re-sampling and data synthesis methods directly modify the distributions of the original dataset. Re-sampling methods [8, 17, 19] involve oversampling or undersampling techniques. Oversampling increases the sampling probability of minority classes, whereas undersampling decreases the probability of majority classes. Data synthesis methods [3, 18, 21, 28] utilize generative models such as generative adversarial networks (GANs) to generate new images for minority classes. These generated samples help balance the sample size across classes, thereby addressing the class imbalance issue.

Recently, the diffusion model (DM) has emerged as a potent generative model that exhibits superior performance compared to GANs [6, 20]. Building on the success of DM, various methods have demonstrated the potential of DM in improving medical image classification by synthesizing training samples [29, 30], which is particularly useful for medical image classification tasks with limited training data. However, existing methods typically generate images independently before commencing the training process, keeping the training images unchanged throughout. This separated pipeline may lead to overfitting due to the lack of synthetic diversity or model collapse problem [10]. Moreover, the portion of synthetic images for each class is often determined manually in these methods, which may not align with the dynamic requirements of the classifier for each class during the training process [24].

To address aforementioned issues, in this paper, we propose a novel **Iterative Online Image Synthesis** framework, named **IOIS**, to address the class imbalance problem in medical image classification. The main contributions of our work are summarized as follows: 1) At the instance level, we introduce an Online Image Synthesis (OIS) module to alleviate the data insufficient problem. For each epoch, we employ the gradient of the training classifier as guidance for the diffusion model to generate samples tailored for online training. As the classifier develops during training, the synthesized images become more representative of their respective classes, which in turn benefits the online training process iteratively. 2) At the class level, we propose an Accuracy Adaptive Sampling (AAS) module to dynamically balance the synthesized samples among various classes. AAS aims to generate more samples for classes with low training accuracy, addressing the class imbalance issue in a targeted manner. 3) We conduct experiments on two public medical image datasets to evaluate the effectiveness of our method. The results demonstrate the superiority of our approach compared to state-of-the-art methods, as well as the effectiveness of each component.

## 2 Methodology

As depicted in Fig. 1, the proposed Iterative Online Image Synthesis (IOIS) framework comprises three key components: a classifier, the Online Image Syn-

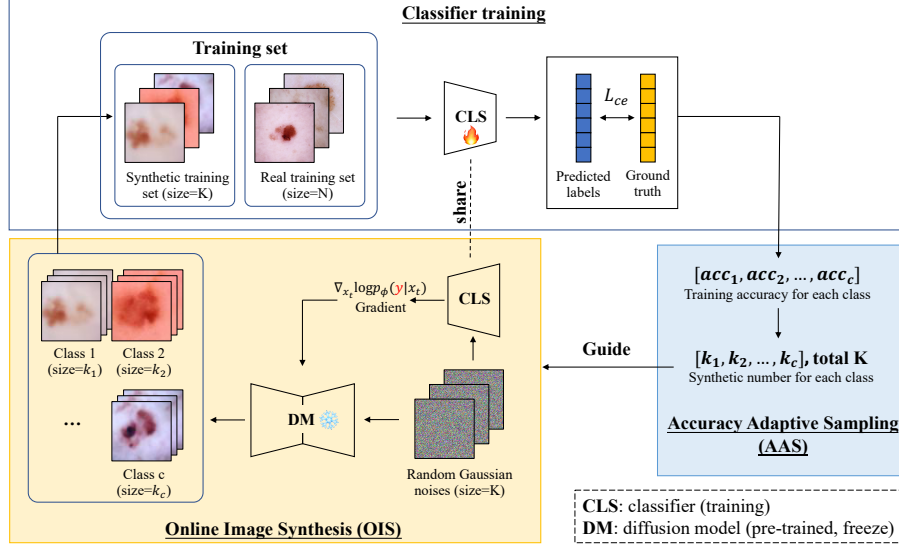


Fig. 1: The framework of the proposed IOIS. It comprises three components, which are iteratively performed: training classifier, determining class distribution of synthetic images via AAS, and synthesizing images via OIS.

thesis (OIS) module, and the Accuracy Adaptive Sampling (AAS) module. Firstly, a conventional classifier training process is conducted, where we adopt ResNet-50 as the backbone architecture. Then, the OIS module is introduced to generate augmented data tailored to the classifier during the training process. Additionally, the AAS module is proposed to update the distribution of synthesized images in the diffusion model based on the feedback from the classifier’s performance. In the following, we will detail each component.

## 2.1 Diffusion Model

Diffusion model (DM) [7] is a probabilistic generative model that aims to learn the probability distribution of a given dataset. It defines the *forward* process to gradually add Gaussian noise to a clean image and the *reverse* process to recover the noised image step by step.

In the *forward* process, we gradually add Gaussian noise to an image sampled from a data distribution  $\mathbf{x}_0 \sim q(\mathbf{x})$  in  $T$  steps to obtain noisy samples  $\{\mathbf{x}_1, \dots, \mathbf{x}_t, \dots, \mathbf{x}_T\}$ , where  $t \in \{1, \dots, T\}$ . We use  $q(\mathbf{x}_t|\mathbf{x}_{t-1})$  to represent forward process at  $t$ -th step,

$$q(\mathbf{x}_t|\mathbf{x}_{t-1}) = \mathcal{N}(\mathbf{x}_t; \sqrt{1 - \beta_t}\mathbf{x}_{t-1}, \beta_t\mathbf{I}), \quad (1)$$

where  $\beta_t \in (0, 1)$  denotes the variance which controls the step size for  $t$ -th step, and  $\mathbf{I}$  is the identity matrix. In the *reverse* process, we train a model  $\epsilon_\theta$  to

predict the noise added on  $\mathbf{x}_t$ . Following [7], the simplified training objective is to minimize the mean square error loss:

$$L_{diff} = E_{t, \mathbf{x}_0, \epsilon} [\|\epsilon - \epsilon_\theta(\mathbf{x}_t, t)\|^2]. \quad (2)$$

In the inference stage, we start from a randomly sampled pure Gaussian noise  $\mathbf{x}_T$  and remove one step noise by

$$\mathbf{x}_{t-1} = \frac{1}{\sqrt{\alpha_t}} \left( \mathbf{x}_t - \frac{1 - \alpha_t}{\sqrt{1 - \bar{\alpha}_t}} \epsilon_\theta(\mathbf{x}_t, t) \right) + \sigma_t \mathbf{z}, \mathbf{z} \sim \mathcal{N}(0, \mathbf{I}). \quad (3)$$

After  $T$  steps calculation, we can finally obtain  $\mathbf{x}_0$  from  $\mathbf{x}_T$ .

This label-free pre-training approach presents promising prospects, as it allows us to harness the vast reserves of unlabeled data and facilitates the utilization of a wide array of pre-trained diffusion models, thereby enhancing the scalability of our approach. Once the diffusion model is pre-trained, we freeze the parameters of the DM and only perform inference to generate augmented data for the classifier training. The class distribution of the synthetic images is dynamically determined by the proposed OIS and AAS modules, which we will describe in the following sections.

## 2.2 Online Image Synthesis with Classifier Guidance

In contrast to existing methods that solely rely on a pre-trained diffusion model to generate augmented data for training the classifier, we propose a novel module called online image synthesis (OIS). OIS addresses the adaptivity limitations in the classifier by producing augmented data specifically tailored to the classifier during the training process. The key idea of the OIS module is to utilize the development of the classifier to synthesize more representative images for each class. Specifically, during the classifier training process, instead of solely using the real images to guide the DM inference as in Eq. (3), we further incorporate the classifier’s feedback by using the classifier’s gradient of  $\mathbf{x}_t$  from the output of diffusion model:

$$\hat{\epsilon} = \epsilon_\theta(\mathbf{x}_t, t) - s \cdot \nabla_{\mathbf{x}_t} \log p_\phi(y|\mathbf{x}_t). \quad (4)$$

Then,  $\mathbf{x}_{t-1}$  is derived by replace the  $\epsilon_\theta$  with the updated value  $\hat{\epsilon}$  in Eq. (3). With the iteration of several steps, the intermediate image  $\mathbf{x}_t$  becomes close to the distribution of class  $y$  as the decreasing of the noise level. In the end, the recovered image  $\mathbf{x}_0$  belongs to class  $y$ .

## 2.3 Accuracy Adaptive Sampling for Imbalanced Data

To achieve a better balance of samples among different classes, we propose an Accuracy Adaptive Sampling (AAS) module designed to dynamically determine the class distributions of synthetic images based on the accuracy of each class.

---

**Algorithm 1:** Framework of IOIS, given a diffusion model  $\epsilon_\theta$ , and gradient scale  $s$ .

---

**Input:** Real training images  $\mathbf{x}_{re} \sim D_{real}$ , random Gaussian noise  $\mathbf{x}_T \sim \mathcal{N}(0, \mathbf{I})$   
**Output:** Labels of images in the test set  
 Initialize the parameters of classifier  $\phi$  with pre-trained ImageNet weights;  
**for** each epoch **do**  
     Update  $\phi$ ;  
     Compute accuracy for each class on the training set  $[acc_1, \dots, acc_c]$  ;  
     Compute the synthetic numbers for each class  $[k_1, \dots, k_c]$  by Eq. (5);  
     Sample  $K$  images with diffusion model  $\epsilon_\theta$  and the classifier  $\phi$ ;  
     Merge  $K$  synthesized images and the real training images as the training set for the next epoch.

---

Specifically, for each epoch of classifier training, we compute the training accuracy for each class, resulting in a vector  $[acc_1, acc_2, \dots, acc_c]$ , where  $c$  represents the number of classes. Based on the training accuracy, we calculate the synthetic number, denoted as  $k_i$ , for each class  $i$ . The value of  $k_i$  is determined by the following equation:

$$[k_1, k_2, \dots, k_c] = \text{Softmax}([1 - acc_1, 1 - acc_2, \dots, 1 - acc_c]) * K, \quad (5)$$

where  $K$  denotes the total size of synthetic images for all the classes. The obtained number  $k_i$  for class  $i$  reflects the performance, that is the lower accuracy of the class, the more samples will be generated in the third step. As a result, the classes with lower accuracy obtain more training samples in the next epoch, which can improve the following performance. Algorithm 1 outlines the whole pipeline of our proposed approach.

### 3 Experiments

#### 3.1 Datasets and Evaluation Metrics

We evaluate our method on two public medical image datasets with imbalanced class distributions. The **HAM10000** dataset [25] consists of 10,015 dermatoscopic images from 7 skin lesion categories. We follow the data splitting settings in [11], where the proportions of the training set, the validation set, and the test set are 70%, 10%, and 20%, respectively. The **APTOS** dataset [9] includes 3,662 retinal fundus images which are divided into five categories. Similarly, we split 70%, 10%, and 20% of the dataset as training, validation, and test set. The imbalanced ratios [12] of HAM10000 and APTOS are 59 and 9, respectively. We leverage three evaluation metrics for imbalanced classification which are Macro-F1, Balanced Accuracy (B-ACC), and Matthew’s correlation coefficient (MCC).

Table 1: Comparisons with state-of-the-art methods on HAM10000 and APTOS datasets. Results of other methods are re-implemented under the same setting.

| Method            | HAM10000 Dataset |              |              | APTOS Dataset |              |              |
|-------------------|------------------|--------------|--------------|---------------|--------------|--------------|
|                   | Macro-F1         | B-ACC        | MCC          | Macro-F1      | B-ACC        | MCC          |
| CE Loss           | 77.78            | 76.78        | 76.11        | 69.05         | 67.91        | 76.79        |
| Focal Loss [14]   | 77.72            | 77.54        | 75.04        | 70.22         | 67.82        | 77.47        |
| CB-Focal [5]      | 77.96            | 79.90        | 75.31        | 70.54         | 70.36        | 77.19        |
| Sqrt-RS [19]      | 78.55            | 79.44        | 76.13        | 70.02         | 68.61        | 77.68        |
| PG-RS [8]         | 79.09            | 78.45        | 77.78        | 70.77         | 69.77        | 76.98        |
| Cell-GAN [21]     | 78.89            | 79.43        | 77.74        | 71.10         | 70.26        | 78.86        |
| StyleGAN2-ADA [3] | 80.03            | 80.14        | 78.82        | 72.48         | 71.05        | 79.26        |
| Ours (+offline)   | 80.09            | 79.97        | 78.79        | 72.79         | 70.86        | 79.95        |
| Ours (+OIS)       | 80.89            | 80.20        | 79.03        | 72.72         | 71.08        | 80.17        |
| Ours (+OIS, AAS)  | <b>81.97</b>     | <b>81.50</b> | <b>80.64</b> | <b>73.02</b>  | <b>72.27</b> | <b>80.76</b> |

### 3.2 Impelementation Details

We utilize ResNet-50 as the backbone architecture and initialize the parameters with the pre-trained model on the ImageNet dataset. To train the model, we set the batch size to 32, and resize the input images to 224px  $\times$  224px. We apply the standard augmentations including random cropping, random rotation (-30 degrees to +30 degrees), and horizontal flipping. We choose the SGD optimizer with a 0.0125 initial learning rate for 200 epochs of training. We decay the learning rate by 0.1 times at 60, 120, and 180 epochs. To prevent overfitting, we select the test model based on the maximum Macro-F1 on the validation set.

### 3.3 Comparisons with State-of-the-art Methods

We conducted a comprehensive comparison of our proposed method with several state-of-the-art approaches on the HAM10000 and APTOS datasets. To ensure a fair evaluation, we re-implemented the competing methods using the same backbone architecture and data augmentation strategies. The comparison results are summarized in Table 1. The compared methods are specifically designed to address the issue of imbalanced classification, including re-weighting techniques (Focal Loss [14] and CB-Focal [5]), re-sampling approaches (Sqrt-RS [19] and PG-RS [8]), and GAN-based synthetic methods (Cell-GAN [21] and StyleGAN2-ADA [3]). Compared to the baseline method (CE Loss), which employs cross-entropy loss for training, other methods demonstrate improvements across all three metrics in general. However, the re-weighting and re-sampling techniques achieve these enhancements for minority classes at the expense of accuracy for majority classes. For instance, on the HAM10000 dataset, the CB-Focal method obtains a 3.12% improvement in B-ACC while experiencing a 0.80% decline in MCC relative to CE Loss. In contrast, the GAN-based synthetic methods generally enhance accuracy for both majority and minority classes. Nonetheless, the

improvements are insufficient for certain minority classes due to the inadequate number of synthetic samples generated for these categories. As evidenced by the experimental results in Table 1, the re-weighting and re-sampling methods are characterized by lower Macro-F1 and MCC scores, while the GAN-based synthetic methods exhibit reduced B-ACC values. This observation suggests that the predictions generated by these methods tend to yield higher rates of false positives or false negatives. Our proposed method, which iteratively incorporates online image synthesis and accuracy adaptive sampling, shows significant improvements for minority classes while maintaining high accuracy for majority classes. Specifically, our approach outperforms competing methods across all evaluation metrics. Specifically, our method achieves a 4.53% and 3.97% increase in MCC compared to CE Loss on HAM10000 and APTOS, respectively.

### 3.4 Ablation Study

To evaluate the effectiveness of the proposed online image synthesis (OIS) and accuracy adaptive sampling (AAS), we conducted an ablation study on HAM10000 and APTOS datasets. We compared three models, one with offline image synthesis strategy, one with OIS module, and one with both OIS and AAS. The results of these models are presented in the last three rows of Table 1. For the experiment with the offline setting, we employed the pre-trained diffusion model and classifier to generate images for each class in advance and trained the model with the real training set and the same synthetic images under the same settings. Compared to the results of the baseline method (CE Loss), our model with offline synthesis method obtains improvements of 2.31%, 3.19%, and 2.68% in Macro-F1, B-ACC, and MCC on the HAM10000 dataset, respectively. Moreover, it achieves a little enhancement than Cell-GAN and StyleGAN2-ADA, which indicates the outperformance of the diffusion models over GANs. For the experiments with the OIS strategy only, we sampled images per epoch with the same size for each class. Compared to the offline method, the online module provides 0.24% and 0.22% enhancement of MCC for HAM10000 and APTOS datasets. To further assess the effectiveness of the AAS strategy, we applied it and observed 1.61% and 0.59% improvement of MCC for two datasets, respectively.

### 3.5 Discussions

**Class-wise comparison.** To evaluate the performance of different methods on each class within the HAM10000 dataset, we computed the relative variations in accuracy for all seven classes in comparison to the CE Loss. The methods considered for comparison are Focal Loss, Sqrt-RS, StyleGAN2-ADA, and our proposed approach. We arranged the seven categories in descending order based on their class sizes, and the results are depicted in Fig. 2. Compared to the CE Loss, Focal Loss exhibits a substantial decline in accuracy for major classes (MEL, BKL, and BCC), despite achieving significant improvements for some minor classes (AK and DR). Furthermore, the Sqrt-RS and StyleGAN2-ADA methods demonstrate suboptimal performance for the BKL and BCC classes. In

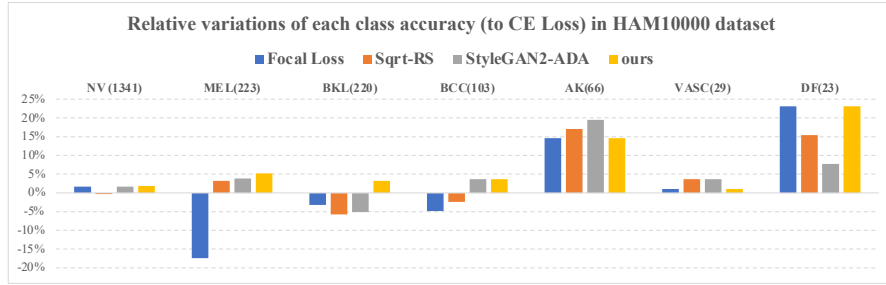


Fig. 2: Comparisons of per class accuracy of different methods on HAM10000 dataset. The relative variations of each class accuracy compared to the CE loss are computed. The numbers in brackets following the class name denote the size of samples in that class.

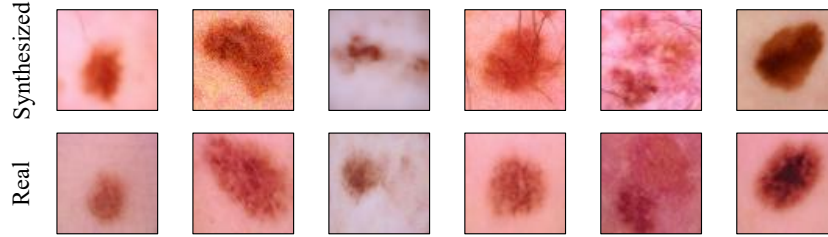


Fig. 3: Visualization examples of synthesized images and real images.

contrast, our proposed approach consistently exhibits enhancements in accuracy for both major and minor classes when compared to the other methods.

**Visualization of synthetic images.** To validate the quality of the synthesized images, we present several visual examples of the generated images in Fig. 3 and identify the most similar images within the real dataset for each synthetic image. Upon examination, it is evident that the generated samples exhibit diversity while maintaining similarity to the real images. For instance, the first column of images displays a consistent shape, but with varying background illumination levels. The third column of images features a comparable texture, yet with diverse lesion shapes.

## 4 Conclusion

In conclusion, we present a novel approach to tackle the issue of imbalanced classification in medical image analysis. Building on the superior performance of the diffusion model in synthesizing high-quality images, we integrate online

image synthesis and accuracy adaptive sampling to iteratively augment training samples that are tailored to the classifier. Our method demonstrates superior performance when compared to several state-of-the-art approaches on the HAM10000 and APTOS datasets. Future research may explore the application of our approach to other medical imaging tasks and the development of more advanced techniques for image synthesis and sampling.

**Acknowledgments.** This research was supported by National Natural Science Foundation of China/HKSAR Research Grants Council Joint Research Scheme under Grant N\_HKUST627/20 and by Nanhai People’s Government of Foshan under Project FSNH 22EG05/RG203.

**Disclosure of Interests.** The authors have no competing interests to declare that are relevant to the content of this article.

## References

1. Baccouche, A., Garcia-Zapirain, B., Castillo Olea, C., Elmaghraby, A.: Ensemble deep learning models for heart disease classification: A case study from mexico. *Information* **11**(4), 207 (2020)
2. Banik, D., Bhattacharjee, D.: Mitigating data imbalance issues in medical image analysis. In: *Data preprocessing, active learning, and cost perceptive approaches for resolving data imbalance*, pp. 66–89. IGI Global (2021)
3. Carrasco Limeros, S., Majchrowska, S., Zoubi, M.K., Rosén, A., Suvilehto, J., Sjöblom, L., Kjellberg, M.: Assessing gan-based generative modeling on skin lesions images. In: *Machine Intelligence and Digital Interaction Conference*. pp. 93–102. Springer Nature Switzerland Cham (2022)
4. Cui, W., Peng, Y., Yuan, G., Cao, W., Cao, Y., Lu, Z., Ni, X., Yan, Z., Zheng, J.: Fmrnet: A fused network of multiple tumoral regions for breast tumor classification with ultrasound images. *Medical Physics* **49**(1), 144–157 (2022)
5. Cui, Y., Jia, M., Lin, T.Y., Song, Y., Belongie, S.: Class-balanced loss based on effective number of samples. In: *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*. pp. 9268–9277 (2019)
6. Dhariwal, P., Nichol, A.: Diffusion models beat gans on image synthesis. *Advances in neural information processing systems* **34**, 8780–8794 (2021)
7. Ho, J., Jain, A., Abbeel, P.: Denoising diffusion probabilistic models. *Advances in Neural Information Processing Systems* **33**, 6840–6851 (2020)
8. Kang, B., Xie, S., Rohrbach, M., Yan, Z., Gordo, A., Feng, J., Kalantidis, Y.: Decoupling representation and classifier for long-tailed recognition. *arXiv preprint arXiv:1910.09217* (2019)
9. Karthik, Maggie, S.D.: Aptos 2019 blindness detection (2019), <https://kaggle.com/competitions/aptos2019-blindness-detection>
10. Kebaili, A., Lapuyade-Lahorgue, J., Ruan, S.: Deep learning approaches for data augmentation in medical imaging: a review. *Journal of Imaging* **9**(4), 81 (2023)
11. Li, J., Cao, H., Wang, J., Liu, F., Dou, Q., Chen, G., Heng, P.A.: Learning robust classifier for imbalanced medical image dataset with noisy labels by minimizing invariant risk. In: *International Conference on Medical Image Computing and Computer-Assisted Intervention*. pp. 306–316. Springer (2023)

12. Li, J., Chen, G., Mao, H., Deng, D., Li, D., Hao, J., Dou, Q., Heng, P.A.: Flat-aware cross-stage distilled framework for imbalanced medical image classification. In: International Conference on Medical Image Computing and Computer-Assisted Intervention. pp. 217–226. Springer (2022)
13. Li, S., Li, X., Xu, X., Cheng, K.T.: Dynamic subcluster-aware network for few-shot skin disease classification. *IEEE Transactions on Neural Networks and Learning Systems* (2023)
14. Lin, T.Y., Goyal, P., Girshick, R., He, K., Dollár, P.: Focal loss for dense object detection. In: Proceedings of the IEEE international conference on computer vision. pp. 2980–2988 (2017)
15. Lin, Y., Liu, L., Ma, K., Zheng, Y.: Seg4reg+: Consistency learning between spine segmentation and cobb angle regression. In: International conference on medical image computing and computer-assisted intervention. pp. 490–499. Springer (2021)
16. Lin, Y., Liu, Y., Chen, H., Yang, X., Ma, K., Zheng, Y., Cheng, K.T.: Lenas: Learning-based neural architecture search and ensemble for 3-d radiotherapy dose prediction. *IEEE Transactions on Cybernetics* (2024)
17. Lin, Y., Su, J., Wang, X., Li, X., Liu, J., Cheng, K.T., Yang, X.: Automated pulmonary embolism detection from ctpa images using an end-to-end convolutional neural network. In: International conference on medical image computing and computer-assisted intervention. pp. 280–288. Springer (2019)
18. Lin, Y., Wang, Z., Cheng, K.T., Chen, H.: Insmix: towards realistic generative data augmentation for nuclei instance segmentation. In: International Conference on Medical Image Computing and Computer-Assisted Intervention. pp. 140–149. Springer (2022)
19. Mahajan, D., Girshick, R., Ramanathan, V., He, K., Paluri, M., Li, Y., Bharambe, A., Van Der Maaten, L.: Exploring the limits of weakly supervised pretraining. In: Proceedings of the European conference on computer vision (ECCV). pp. 181–196 (2018)
20. Rombach, R., Blattmann, A., Lorenz, D., Esser, P., Ommer, B.: High-resolution image synthesis with latent diffusion models. In: Proceedings of the IEEE/CVF conference on computer vision and pattern recognition. pp. 10684–10695 (2022)
21. Shen, Z., Cao, M., Wang, S., Zhang, L., Wang, Q.: Cellgan: Conditional cervical cell synthesis for augmenting cytopathological image classification. In: International Conference on Medical Image Computing and Computer-Assisted Intervention. pp. 487–496. Springer (2023)
22. Tan, J., Lu, X., Zhang, G., Yin, C., Li, Q.: Equalization loss v2: A new gradient balance approach for long-tailed object detection. In: Proceedings of the IEEE/CVF conference on computer vision and pattern recognition. pp. 1685–1694 (2021)
23. Tan, J., Wang, C., Li, B., Li, Q., Ouyang, W., Yin, C., Yan, J.: Equalization loss for long-tailed object recognition. In: Proceedings of the IEEE/CVF conference on computer vision and pattern recognition. pp. 11662–11671 (2020)
24. Tang, Z., Gao, Y., Karlinsky, L., Sattigeri, P., Feris, R., Metaxas, D.: Onlineaug: Online data augmentation with less domain knowledge. In: Computer Vision–ECCV 2020: 16th European Conference, Glasgow, UK, August 23–28, 2020, Proceedings, Part VII 16. pp. 313–329. Springer (2020)
25. Tschandl, P., Rosendahl, C., Kittler, H.: The ham10000 dataset, a large collection of multi-source dermatoscopic images of common pigmented skin lesions. *Scientific data* **5**(1), 1–9 (2018)
26. Wang, X., Han, Y., Sun, G., Yang, F., Liu, W., Luo, J., Cao, X., Yin, P., Myers, F.L., Zhou, L.: Detection of the microvascular changes of diabetic retinopathy

- progression using optical coherence tomography angiography. *Translational vision science & technology* **10**(7), 31–31 (2021)
27. Yan, R., Ren, F., Li, J., Rao, X., Lv, Z., Zheng, C., Zhang, F.: Nuclei-guided network for breast cancer grading in he-stained pathological images. *Sensors* **22**(11), 4061 (2022)
  28. Yang, X., Lin, Y., Wang, Z., Li, X., Cheng, K.T.: Bi-modality medical image synthesis using semi-supervised sequential generative adversarial networks. *IEEE journal of biomedical and health informatics* **24**(3), 855–865 (2019)
  29. Ye, J., Ni, H., Jin, P., Huang, S.X., Xue, Y.: Synthetic augmentation with large-scale unconditional pre-training. In: *International Conference on Medical Image Computing and Computer-Assisted Intervention*. pp. 754–764. Springer (2023)
  30. Zhong, Y., Cui, S., Wang, J., Wang, X., Yin, Z., Wang, Y., Xiao, H., Huai, M., Wang, T., Ma, F.: Meddiffusion: Boosting health risk prediction via diffusion-based data augmentation. *arXiv preprint arXiv:2310.02520* (2023)