# CycleMix: A Holistic Strategy for Medical Image Segmentation from Scribble Supervision

## Supplementary Material

Table 1. Application to different mixup methods.

| Methods | LV | MYO | RV | Avg |
|---|---|---|---|---|
| w/o CycleMix | | | | |
| MixUp [7] | .804±.152 | .715±.129 | .774±.158 | .765 |
| CutMix [5] | .767±.237 | .726±.176 | .784±.197 | .759 |
| Puzzle Mix [3] | .821±.158 | .715±.125 | .806±.162 | .781 |
| Co-mixup [2] | .709±.232 | .689±.153 | .803±.097 | .734 |
| w/ CycleMix | | | | |
| MixUp [7] | .849±.139 | .761±.102 | .802±.138 | .804 |
| CutMix [5] | .835±.183 | .771±.166 | .716±.206 | .774 |
| Puzzle Mix [3] | **.880±.115** | **.825±.072** | **.860±.089** | **.855** |
| Co-mixup [2] | .797±.158 | .735±.109 | .775±.171 | .769 |

Table 2. Comparison with baselines on entire ACDC training dataset of 70 subjects. For weakly-supervised segmentation, **bold** denotes the best results, underline denotes the second best performance.

| Methods | Data | ACDC | | | |
|---|---|---|---|---|---|
| | | LV | MYO | RV | Avg |
| 70 scribbles | | | | | |
| UNet$^{+}_{pce}$ [4] | scribbles | .808±.161 | .749±.099 | .779±.133 | .779 |
| MixUp [7] | scribbles | .804±.152 | .715±.129 | .774±.158 | .765 |
| Cutout [1] | scribbles | .815±.172 | .758±.134 | .817±.123 | .797 |
| CutMix [5] | scribbles | .767±.237 | .726±.176 | .784±.197 | .759 |
| Puzzle Mix [3] | scribbles | .821±.158 | .715±.125 | .806±.162 | .781 |
| Co-mixup [2] | scribbles | .709±.232 | .689±.153 | .803±.097 | .734 |
| CycleMix(ours) | scribbles | **.880±.115** | **.825±.072** | **.860±.089** | **.855** |
| 70 masks | | | | | |
| UNet$^{+}_{F}$ | masks | .883±.130 | .831±.093 | .870±.096 | .862 |
| Puzzle Mix$_F$ [2] | masks | .912±.082 | .842±.081 | .887±.066 | .880 |

## A. Extra experiments

We conducted the extensive experiments on the entire training set of ACDC, which includes 70 subjects and corresponding scribble annotations. Firstly, we proved the applicability of CycleMix to various mixup methods. Then, we compared CycleMix to different mixup methods on the entire training dataset of ACDC.

### A.1. Application to different mixup methods.

Table. 1 presents the performance of different mixup augmentation methods with and without CycleMix on ACDC dataset. CycleMix consistently improved the segmentation results of all listed mixup methods by large margins range from 1.5% to 7.4%. For MixUp, CutMix and Co-mixup, CycleMix obtained an obvious performance gain of 3.9%, 1.5%, and 3.5%, respectively, demonstrating the applicability of the proposed CycleMix to different mixup augmentation methods. The benefit was even more evident on the Puzzle Mix, where CycleMix boosted the Dice Score to 85.5% with an improvement of 7.4%.

### A.2. Comparison on the entire ACDC training set

To further validate the effectiveness of proposed CycleMix, we trained CycleMix and mixup baselines on the entire ACDC training dataset of 70 objects. The experiment results are reported in Table. 2. As one can observed, our CycleMix significantly outperformed all baselines trained with scribble annotations. Compared to $2_{nd}$ best average Dice Score, CycleMix obtained an remarkable performance gain of 5.8% (85.5% vs 79.7%), demonstrating the effectiveness of our proposed method. As Table. 2 shows, CycleMix exhibited performance matching the fully-supervised methods.

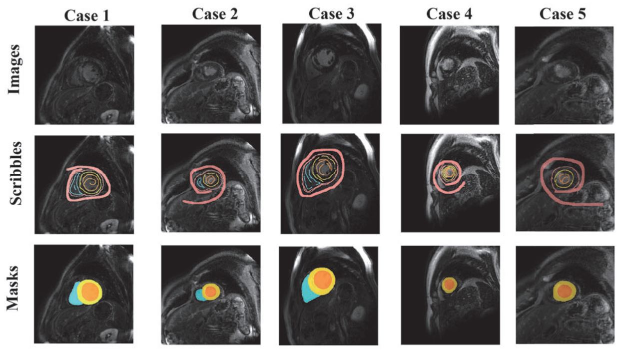


Figure 1. Scribble annotations of MSCMR dataset. The first three cases show the complete heart, including LV, MYO, and RV. The last two cases present scribble annotations when only LV and MYO are visible.

## B. Scribble annotations of MSCMR

On MSCMR LGE dataset, we use ITK-SNAP [6] to manually draw scribbles within the available segmentation masks of MSCMR dataset. The RV, LV and MYO are annotated separately. Considering that the LGE CMR segmentation *per se* is more complex than images from ACDC, we additionally draw a scribble of background to outline the heart. More scribble annotations are visualized in Figure. 1.

## References


[1] Terrance DeVries and Graham W Taylor. Improved regularization of convolutional neural networks with cutout. *arXiv preprint arXiv:1708.04552*, 2017. 1

[2] JangHyun Kim, Wonho Choo, Hosan Jeong, and Hyun Oh Song. Co-mixup: Saliency guided joint mixup with supermodular diversity. In *International Conference on Learning Representations*, 2021. 1

[3] Jang-Hyun Kim, Wonho Choo, and Hyun Oh Song. Puzzle mix: Exploiting saliency and local statistics for optimal mixup. In *International Conference on Machine Learning (ICML)*, 2020. 1

[4] Meng Tang, Abdelaziz Djelouah, Federico Perazzi, Yuri Boykov, and Christopher Schroers. Normalized cut loss for weakly-supervised cnn segmentation. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pages 1818–1827, 2018. 1

[5] Sangdoo Yun, Dongyoon Han, Seong Joon Oh, Sanghyuk Chun, Junsuk Choe, and Youngjoon Yoo. Cutmix: Regularization strategy to train strong classifiers with localizable features. In *International Conference on Computer Vision (ICCV)*, 2019. 1

[6] Paul A Yushkevich, Joseph Piven, Heather Cody Hazlett, Rachel Gimpel Smith, Sean Ho, James C Gee, and Guido Gerig. User-guided 3d active contour segmentation of anatomical structures: significantly improved efficiency and reliability. *Neuroimage*, 31(3):1116–1128, 2006. 1

[7] Hongyi Zhang, Moustapha Cisse, Yann N. Dauphin, and David Lopez-Paz. mixup: Beyond empirical risk minimization. *International Conference on Learning Representations*, 2018. 1