Group Equivariant Convolutional Neural Networks for Color Fundus Images Super-Resolution

Juntao Jiang juntaojiang@zju.edu.cn and **Yong Liu** yongliu@iipc.zju.edu.cn
38 Zheda Road, West Lake District, Zhejiang University, Yuquan Campus

Abstract

High-resolution (HR) color fundus images can provide finer details and help with a more accurate diagnosis, while deep learning-based super-resolution methods usually require a large amount of data. This paper integrated Group Equivariant Convolution Neural Networks (G-CNNs) in Modified SRResNet (MSRResNet) for fundus images super-resolution. Experiments showed that it leads to some improvement compared with regular CNNs.

Keywords: Fundus images, Super-resolution, Group Equivariant Convolution Neural Networks, Modified SRResNet

1. Introduction

Fundus imaging is widely used to screen and diagnose ophthalmic diseases. Compared with low-resolution (LR) photography, high-resolution (HR) color fundus images can provide finer details. However, due to the limitations in image acquisition time or hardware (Li et al. (2021)), sometimes it is hard to get images with sufficient spatial resolution.

Super-resolution (SR) is an approach to obtaining HR images from LR observations (Gerchberg (1974)). Single-image SR (SISR) is often considered an ill-posed inverse problem, as different HR images can be degraded to the same LR image. Recently, many deep-learning-based methods have been proposed to address the SISR problem, mainly designed for natural images but usually also work on medical images.

However, deep-learning-based methods usually require significant training data to enhance the model's generalization ability and avoid overfitting. Meanwhile, data augmentation methods are considered sub-optimal regarding sample complexity and computationally more expensive (Mei et al. (2021)).

The Group Equivariant Convolutional Neural Network (G-CNN) (Cohen and Welling (2016)) was proposed to improve the statistical efficiency of deep learning methods by exploiting symmetries. Without increasing the number of parameters, it increases the expressive capacity of the network. On the classification task, G-CNN achieved state-of-the-art results on MNIST-rot (Larochelle et al. (2007)) and CIFAR10 (Krizhevsky et al. (2009)). G-CNN has been widely used in medical image analysis. However, the exploration of G-CNNs' applications in medical image super-resolution is not much.

To better exploit symmetries in Fundus images, this paper proposed a G-MSRResNet, which simply integrated G-Convolution in the Modified SRResNet (MSRResNet) (Ledig et al. (2017),Wang et al. (2018)). We trained the G-MSRResNet on the DRIVE image dataset (Staal et al. (2004)). Experiments showed that the G-MSRResNet performed better than MSRResNet with regular CNN.

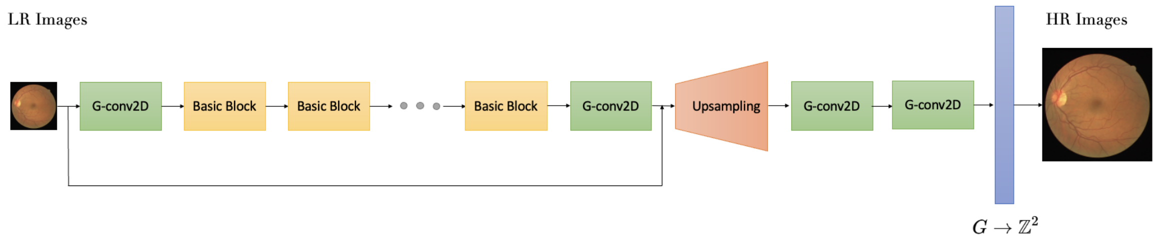

Figure 1: The architecture of Group Equivariant MSRResNet

2. Related Works

The Group and Group Convolution A group $(G; \circ)$ consists of a set G and a binary composition operator $\circ$ and satisfies the properties of closure, associativity, identity, and invertibility. Given a group G and a map $\Phi : X \to Y$ between two G-sets X and Y, Φ is equivariant i.f.f. $\Phi(g \cdot x) = g \cdot \Phi(x), \forall x \in X, \forall g \in G$ (Dey et al. (2020)). Our work applied two types of plane symmetry groups $p4$ (all compositions of 90-degree rotations and translations) and $p4m$ (all compositions of 90-degree rotations, reflections, and translations) (Schattschneider (1978)).

Group convolution(G-Convolution) in the first layer of a G-CNN can be represented as $[f \star \psi](g) = \sum_{y \in \mathbb{Z}^2} \sum_k f_k(y) \psi_k \left(g^{-1} y\right)$, where the input image f and the filter ψ are functions of the plane $\mathbb{Z}^2$ while the feature map $f \star \psi$ is a function on the discrete group G. In the layers after the first layer, it can be represented as $[f \star \psi](g) = \sum_{h \in G} \sum_k f_k(h) \psi_k \left(g^{-1} h\right)$ (Cohen and Welling (2016)). Finally, for the super-resolution tasks, the feature map for each filter over the set of transformations can be pooled to the domain of feature maps from G back to $\mathbb{Z}^2$.

Modified SRResNet The Modified SRResNet (MSRResNet) consists of 16 residual modules and a global skip connection, which is an improved version of SRResNet (Ledig et al. (2017)) without batch normalization layer according to improvement in Wang et al. (2018).

3. Group Equivariant MSRResNet

The architecture of Group Equivariant MSRResNet is shown in Figure 1. The architectures of the basic block and upsampling block are shown in Figure 2. We used G-CNN to replace regular CNN in MSRResNet and used the nearest method for upsampling. And we projected the feature maps from G to $\mathbb{Z}^2$ space to get the final HR output after the final layer.

4. Experiments

We trained G-MSRResNets with $p4$ and $p4m$ groups on the DRIVE dataset and compared the results with MSRResNet with regular CNNs. We did super-resolution experiments with $\times 2$ and $\times 4$ scaling factors. The evaluation metric is PSNR and SSIM (Wang et al. (2004)). DRIVE is a dataset for retinal vessel segmentation with JPEG 40 color fundus images, which shows that the G-CNN can improve color fundus images' super-resolution. We split

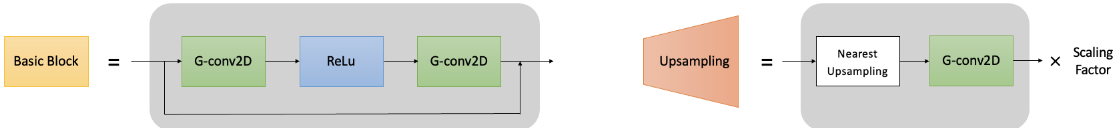

Figure 2: The architectures of the basic block (left) and upsampling block (right) in Group Equivariant MSRResNet

the original training set, from which 16 were for training and four were for validation. The original test set with 20 images is for testing. We downsampled these images to get LR images. The patch size in training is 128.

The initial learning rate is 1.0e-04, and a MultiStepLR scheduler is used. The total iterations for training are 20000. The learning rate is reduced to half of the original at 10000. The batch size is 16, and the optimizer is Adam. We validated and updated the best model every 1000 iterations.

The results of the experiments are shown in Table 1:

Table 1: Comparison Results on the DRIVE Dataset

Scaling Factor	**Convolution Type**	**PSNR**	**SSIM**
$\times 2$	Regular	43.79	0.9784
$\times 2$	$p4$ group	44.12	0.9795
$\times 2$	$p4m$ group	44.05	0.9793
$\times 4$	Regular	40.33	0.9382
$\times 4$	$p4$ group	40.79	0.9414
$\times 4$	$p4m$ group	40.84	0.9419

5. Conclusion

From the comparison experiments, G-MSRResNet performed better than regular MSRResNet on the DRIVE test. Further research will focus on comparing the equivariant group method and the data augmentation method for fundus images super-resolution, not only from the performance perspective but also from computational cost.

References

Taco Cohen and Max Welling. In Maria Florina Balcan and Kilian Q. Weinberger, editors, *Proceedings of The 33rd International Conference on Machine Learning*, volume 48 of *Proceedings of Machine Learning Research*, pages 2990–2999, New York, New York, USA, 20–22 Jun 2016. PMLR. URL https://proceedings.mlr.press/v48/cohenc16.html.

Neel Dey, Antong Chen, and Soheil Ghafurian. Group equivariant generative adversarial networks. *arXiv preprint arXiv:2005.01683*, 2020.

RW Gerchberg. Super-resolution through error energy reduction. *Optica Acta: International Journal of Optics*, 21(9):709–720, 1974.

Alex Krizhevsky, Geoffrey Hinton, et al. Learning multiple layers of features from tiny images. 2009.

Hugo Larochelle, Dumitru Erhan, Aaron Courville, James Bergstra, and Yoshua Bengio. An empirical evaluation of deep architectures on problems with many factors of variation. In *Proceedings of the 24th international conference on Machine learning*, pages 473–480, 2007.

Christian Ledig, Lucas Theis, Ferenc Huszár, Jose Caballero, Andrew Cunningham, Alejandro Acosta, Andrew Aitken, Alykhan Tejani, Johannes Totz, Zehan Wang, et al. Photo-realistic single image super-resolution using a generative adversarial network. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 4681–4690, 2017.

Y Li, Bruno Sixou, and F Peyrin. A review of the deep learning methods for medical images super resolution problems. *Irbm*, 42(2):120–133, 2021.

Song Mei, Theodor Misiakiewicz, and Andrea Montanari. Learning with invariances in random features and kernel models. *arXiv preprint arXiv:2102.13219*, 2021.

Doris Schattschneider. The plane symmetry groups: their recognition and notation. *The American Mathematical Monthly*, 85(6):439–450, 1978.

J. Staal, M.D. Abramoff, M. Niemeijer, M.A. Viergever, and B. van Ginneken. Ridge-based vessel segmentation in color images of the retina. *IEEE Transactions on Medical Imaging*, 23(4):501–509, 2004. doi: 10.1109/TMI.2004.825627.

Xintao Wang, Ke Yu, Shixiang Wu, Jinjin Gu, Yihao Liu, Chao Dong, Yu Qiao, and Chen Change Loy. Esrgan: Enhanced super-resolution generative adversarial networks. In *Proceedings of the European conference on computer vision (ECCV) workshops*, pages 0–0, 2018.

Zhou Wang, Alan C Bovik, Hamid R Sheikh, and Eero P Simoncelli. Image quality assessment: from error visibility to structural similarity. *IEEE transactions on image processing*, 13(4):600–612, 2004.