ハイパースペクトル画像におけるピーマンの正確なセグメンテーションのための LPF 誘起ミックスアップの研究

1275120 Jin Yuzhou 【 画像情報工学研究室 】

LPF-Induced Mixup for Accurate Segmentation of Peppers in Hyperspectral Images

1275120 Jin Yuzhou 【 Image Processing and Informatics Lab. 】

1 Introduction

Table 1 Mean F1 scores and variance of different data augmentation methods for 6-fold cross validation Pixel-Level classification experiment.

	Average F1	Std
Without DA	0.660	0.043
With LPF	0.786	0.117
Mixup(α = 0.2)	0.590	0.086
Mixup(α = 0.4)	0.582	0.085
Random select	0.615	0.007
Random select with LPF	0.644	0.010
Mix amplitude	0.632	0.014
Mix amplitude with LPF	0.561	0.128
Mixup by LPF(our)	0.819	0.092

In modern agriculture, automated harvesting is key to improving efficiency and reducing labor costs. Traditional manual identification is accurate but inefficient and not suitable for mass production. It is difficult to have a high accuracy rate by relying on methods such as color thresholding. Therefore deep learning is beginning to be applied to the field of agriculture.

To enhance segmentation accuracy and robustness, this study proposes a green pepper segmentation method based on hyperspectral images, combined with a data augmentation approach that integrates spatial-domain low-pass filtering (LPF) and Mixup.

The experimental results show that the proposed data augmentation method improved the average F1 score by 15.9% in the MLP classification task and by 2.1% in the CNN segmentation task, while reducing the standard deviation by 0.002, significantly enhancing model stability and generalization ability.

Additionally, this study investigates the sensitivity of multiscale targets to convolution kernel size and compares the performance of different data augmentation methods. Unlike traditional approaches, the study uses random pixel fusion (LPF-Mixup), which performs best in the spatial domain.

2 Method And Materials

This chapter presents the technical methods for addressing the challenges of hyperspectral image segmentation in the agricultural domain, specifically focusing on green pepper segmentation. To tackle the overconfidence issue [1] in deep neural networks, the Mixup data augmentation technique is introduced.By linearly interpolating between samples and their labels [2], Mixup smooths decision boundaries and enhances the model's generalization and robustness. Additionally, Gaussian low-pass filtering (LPF) is employed to generate mixed pixels, further improving the distinction between the target and background.

2.1 Data Augmentation

The Mixup technique generates new training data by blending randomly selected samples, enabling the model to produce more calibrated predictions under uncertainty. The enhanced LPF-Mixup method extends low-pass filtering to both input data and labels, creating more realistic target pixel distributions and optimizing segmentation performance.Show as Fig.1.

2.2 Network Structure

Image Segmentation: A convolutional neural network (CNN) based on the U-Net architecture is utilized. Pixel Classification: A multilayer perceptron (MLP) with a structure of 512-256-128-64-1 is implemented.

3 Experimentals And Experimental Result

3.1 Results Of Pixel-Wise Classification Experiment

Table 3 shows the average F1 scores as well as the standard deviation of different data augmentation approaches on the test set for four images. It can be seen that the proposed party has a 15.9% improvement, so our proposed method has great potential in the field of hyperspectral data augmentation.

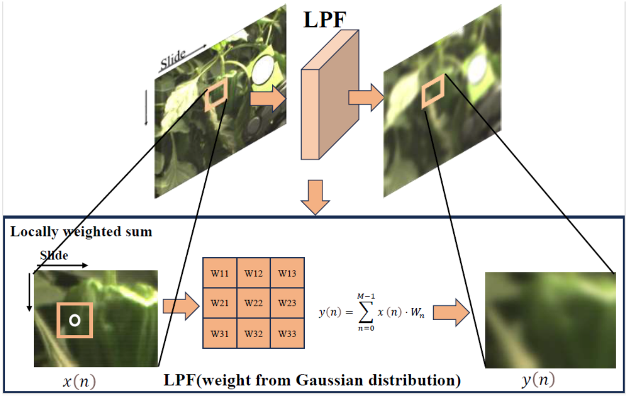

Fig 1 LPF causes the target spectrum to be mixed with surrounding pixels.

Table 2 Mean F1 scores and variance of different data augmentation methods for 6-fold cross validation Pixel-Level classification experiment.

	Average F1	Std
Without DA	0.660	0.043
With LPF	0.786	0.117
Mixup($\alpha = 0.2$)	0.590	0.086
Mixup($\alpha = 0.4$)	0.582	0.085
Mixup by LPF	0.819	0.092

3.2 Results Of Image-Wise Classification Experiment

Table 2 lists the results of the experiment, including the test evaluation scores for each data augmentation method. From the results, we can observe that the proposed Mixup

Table 3 Mean F1 scores and variance of different data augmentation methods on segmentation experiment.

	Average F1	Std
Without DA	0.942	0.00275
Flip and crop	0.955	0.00330
With LPF	0.958	0.02295
Mixup by LPF	0.963	0.00082

by LPF data augmentation method outperforms the methods not augmented with data and other augmentation methods, achieving the highest average test F1 score of 0.963. Additionally, it exhibits a lower standard deviation (0.00082) compared to the baseline and same-class augmentation, indicating that the model's performance is more stable and consistent across different folds of the cross-validation.

3.3 Results of Pixel-wise experiments for multi-scale targets

The Fig .2 show the average F1 scores for small and large targets on the test set.The following conclusions can be drawn from this set of data regarding small and large targets: When the kernel size was set to 5, the accuracy for small targets dropped significantly to 0.775, and the accuracy for large targets also decreased to 0.857. This suggests that smaller kernel sizes have a more pronounced negative effect on the segmentation of small targets. As the kernel size increased to 15 and 25, the accuracy for small targets improved to 0.948 and 0.951, while for large targets, it reached 0.944 and 0.931, respectively. This indicates that larger kernel sizes can effectively enhance the recognition performance for small targets while maintaining high accuracy for large targets。 As the kernel size increased, performance for both small and large targets began to decline, indicating that excessive smoothing from larger kernels negatively impacts recognition accuracy.

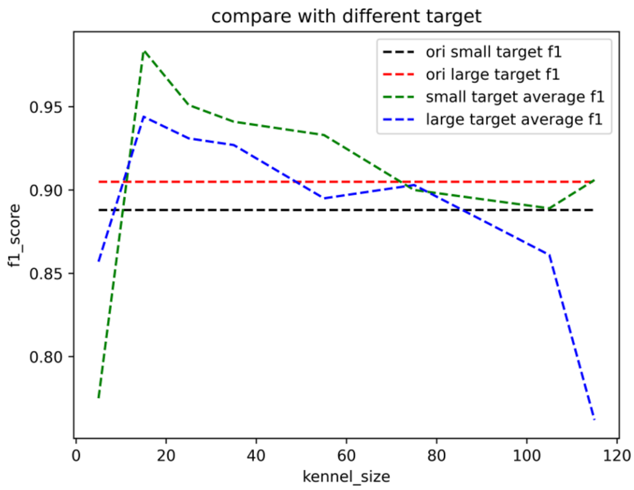

Fig 2 Comparison of f1 scores for Multi-scale objectives.

4 Conclusion

This study proposes an intra-image data augmentation method based on Mixup, using low-pass filtering (LPF) to blend green peppers with surrounding objects and reduce sensor noise in low-light areas, significantly improving segmentation accuracy.

In green pepper segmentation, convolution kernel performance varies with kernel size, especially for targets with size discrepancies. Multi-scale classification experiments helped identify suitable kernel sizes for small and large targets. LPF intensity also impacts segmentation performance by causing different blurring effects.

Due to limited hyperspectral datasets, the study could not fully explore their potential in semantic segmentation.

The high-dimensional nature of hyperspectral data leads to information loss during feature extraction, a challenge for future research. We aim to develop more efficient solutions for enhancing hyperspectral image data.

References

[1] Chuan Guo, Geoff Pleiss, Yu Sun, Kilian Q. Weinberger, "On calibration of modern neural networks," International Conference on Machine Learning, pp. 1321-1330, 2017.

[2] Sunil Thulasidasan, Gopinath Chennupati, Jeff Bilmes, Tanmoy Bhattacharya, Sarah Michalak, "On Mixup Training: Improved Calibration and Predictive Uncertainty for Deep Neural Networks," arXiv:1905.11001, 2020, `https://arxiv.org/abs/1905.11001`.