Semantic Segmentation of PHT based on Improved DeeplabV3+

Haiquan Fang¹

¹Zhejiang University of Technology

March 07, 2024

Abstract

This work aimed to address the two shortcomings of printed and handwritten texts (PHT) classification. The classification accuracy of FCN and U-net, which are used for PHT pixel-level classification, still has room to improve. PHT public datasets have small sample sizes, and the generalization ability of the models are not good. In this paper, first, a pixel-level sample making method for PHT identification was proposed, and a PHT dataset 2021 (PHTD 2021), containing 3,000 samples, was constructed. Second, because there is a large number of words but the contours are small in documents, the DeeplabV3+ model was improved. The network layer number and pooling times were reduced, and the convolution kernel and dilated rate were increased. In the experiment, the improved DeeplabV3+ model had a classification accuracy of 95.06% on the test samples from PHTD 2021 data set. The improved DeeplabV3+ model has a higher recognition accuracy than the FCN and DeeplabV3+ models. Finally, after the classification of PHT, applications of handwritten texts removal and handwritten texts extraction are provided.

Hosted file

1-Semantic Segmentation of PHT based on Improved DeeplabV3+-202202010.docx available at https://authorea.com/users/740978/articles/713522-semantic-segmentation-of-pht-based-on-improved-deeplabv3