Contribution ID: **220** Type: **Sectional reports**

Text segmentation on photorealistic images

Thursday 13 September 2018 13:30 (15 minutes)

The paper proposes an algorithm for segmentation of text, applied or presented in photorealistic images, characterized by a complex background. Because of its application, the exact location of image regions containing text is determined. The algorithm implements the method for semantic segmentation of images, while the text symbols serve as detectable objects. The original images are pre-processed and fed to the input of the pre-trained convolutional neural network. The paper proposes a network architecture for text segmentation, describes the procedure for the formation of the training set, and considers the algorithm for pre-processing images, reducing the amount of processed data and simplifying the segmentation of the object "background". The network architecture is a modification of well-known ResNet network and takes into account the specifics of text character images. The convolutional neural network is implemented using CUDA parallel computing technology at the GPU. The experimental results for evaluating quality of the text segmentation IoU (Intersection over Union) criterion have proved effectiveness of the proposed method.

Summary

We propose the algorithm for the segmentation of text regions in photorealistic images. It consists of a preprocessing step, a recognition step, and a localization step. The second step uses the modified convolutional ResNet network for recognition. Unlike the original network, the modified neural network saves the geometric structure of text characters into the feature maps. The third step determines the exact localization of recognized text characters and finds areas of the image containing text. Experimental results show effectiveness of the proposed algorithm. Quality of segmentation is evaluated using the IoU metric and reaches 78%, which is sufficient for further processing of the image text using OCR systems. Use of parallel processing technologies significantly reduces processing time of large series of images.

Primary author: Dr GRISHKIN, Valery (SPbGU)

Co-authors: Mr EBRAL, Aleksaner (SPbGU); Mr SENE, Jean (SPbGU); Dr STEPENKO, Nikolai (SPbGu)

Presenter: Dr GRISHKIN, Valery (SPbGU)

Session Classification: 11. Big data Analytics, Machine learning

Track Classification: 11. Big data Analytics, Machine learning