GMIS generate pixel affinities based on CNN for semantic segmentation.

GMIS splits the task of instance segmentation into two sequential steps. The first step utilizes CNN, which specifically is Deeplabv3, to obtain class information and pixel affinity of the input image, while the second step applies the graph merge algorithm on those results to generate the pixel-level masks for each instance, which is a novel proposal-free instance segmentation scheme, where semantic information and pixel affinity information are both used to derive in- stance segmentation results.

The method demonstrated that even with a simple graph merge algorithm, we can outperform other methods, including proposal-based ones. It clearly shows that proposal- free methods can have comparable or even better performance than proposal- based methods.

The method also shows that semantic segmentation network is reasonably suitable for pixel affinity prediction with only the meaning of the output changed.

Results notably improves the performance and achieves 27.3 AP on Cityscapes test set, which outperforms Mask RCNN trained with only Cityscapes train data by 1.1 points (4.2% relatively).