One-stage Progressive Dichotomous Segmentation

Jing Zhu    Karim Ahmed    Wenbo Li    Yilin Shen    Hongxia Jin
Samsung Research America AI Center

Problem

Current existing methods can be classified into two categories. Multi-stage methods have the potential to generate superior results, but they often come with time and memory expenses. On the contrary, one-stage methods offer a more straightforward approach with lower computation costs. However, they typically yield inferior performance because many effective high-complexity networks (e.g., transformer), cannot directly handle high-resolution images due to resource limit.

Approach

• We propose a one-stage framework with an efficient yet effective convolutional attention module that could directly work on high-resolution images for dichotomous segmentation.

• We design the progressive prediction schema into the decoder of the model which enables the gradual refinement of the segmentation map level by level. A multi-scale supervision loss is introduced to enhance the model’s training capabilities. An overview of the multi-scale convolution attention is depicted at the upper left, where $D_0$-$D_n$ denote depth-wise convolutions.

Contribution

• Novel Point 1: a one-stage framework with an efficient yet effective multi-scale convolutional attention feature extractor, enabling direct processing of high-resolution images for dichotomous segmentation.

• Novel Point 2: a progressive decoder with a specifically designed progressive prediction mechanism, which generates an initial segmentation map using the lowest-resolution features and progressively refines the map’s resolution level by level, leveraging the extracted multi-scale features from the feature extractor.

Experimental Results

Table 1: Comparison of dichotomous segmentation on four DISK [28] testing subsets with alternative approaches. Overall performance is computed by taking the mean value of the scores from the four subsets. The best performance has been marked in blue and the second-best result is marked in red. Higher $F^+$, $F^*$, S, E scores and lower M, HCE values indicate the better performance. Our method outperforms all the the single-stage methods with fewer model parameters and computational operations (FLOPs).

Table 2: The ablation studies on DISK-1 to verify the effectiveness of each module. We can observe the progressive scheme has a greater effect on the model’s performance than the multi-scale supervision, while they together yield the best performance.