

DATA MINING

MILESTONE-3

SUMMARY

Deep Extreme Cut (DEXTR), which derives an object segmentation from its four extreme points—the leftmost, rightmost, topmost, and bottommost pixels.

DEXTR outperforms other methods using extreme points or object proposals (PASCAL), and provides a better input to video object segmentation. DEXTR can also be used to obtain dense annotations to train supervised techniques. DEXTR is significantly more efficient than training from the ground truth for a given target quality.

Annotation:

The common annotation pipeline for segmentation can also be assisted by DEXTR. In this framework, instead of detailed polygon labels, the workload of the annotator is reduced to only providing the extreme points of an object, and DEXTR produces the desired segmentation. In this pipeline, the labelling cost is reduced by a factor of 10 (from 79 seconds needed for a mask, to 7.2 seconds needed for the extreme clicks). We show that our method produces very accurate masks and the results trained on them are on par with those trained on the ground-truth annotations in terms of quality, with much less annotation budget.

Training and testing details:

DEXTR is trained on PASCAL 2012 Segmentation for 100 epochs or on COCO 2014 training set for 10 epochs. The learning rate is set to 10^{-8} , with momentum of 0.9 and weight decay of 5×10^{-4} . A mini-batch of 5 objects is used for PASCAL, whereas for COCO, due to the large size of the database, we train on 4 GPUs with an effective batch size of 20. Training on PASCAL takes approximately 20 hours on a Nvidia TitanX GPU, and 5 days on COCO. Testing the network is fast, requiring only 80 milliseconds.