Salient object detection of images with various object properties

M.A.M.K. Weerasinghe[*], M.K.S. Madushika, & D.W.C.P. Kumari

Department of Computer Science, University of Ruhuna, Sri Lanka

Corresponding author: *weerasinghe10294@usci.ruh.ac.lk*

Abstract

Convolutional Neural Networks (CNNs) have recently shown robust feature representation capabilities and have substantially aided the development of Salient Object Detection (SOD). Most of the SOD models are used encoder-decoder architecture. The encoder is often composed of a pre-trained classification model. In the decoder, extracted features of the encoder are combined to generate saliency maps. Here we propose a method called SOD-VOP (Salient Object Detection of Images with Various Object Properties) for the decoder with two networks based on outperforming network. The first network called the Decomposition network, generates body, detail, and skeleton maps simultaneously and further combines maps in the completion network. We use the traditional image processing technique called Distance Transformation (DT) to decouple the original saliency label to the detail map and body map. Detail maps mainly pay attention to edge pixels more than traditional edge detection. The body map discards edge pixels and only pays attention to center areas. After composing these maps with skeleton maps are utilized to fill flaws and suppress noises and can give a better prediction. Another specialty of our model is it performed well with various object properties like multiple, large, small, and moderate objects in simple, clean, and complex backgrounds. Experiments also conducted on five benchmark datasets, which demonstrated that the proposed model outperformed existing networks. After considering our results and comparing them with the others, we can conclude giving more attention to pixels near the edge helps to get better results.

Keywords: *Encoder-decoder architecture, Label decoupling, Saliency completion, Salient object detection*