Instance-Level Salient Object Segmentation

Aniket Joshi (20161166)
Prakyath Madadi (20161236)
Vashist Madiraju (20161222)

Objective

The main objective of this project is to implement a salient instance segmentation method that produces a saliency mask with distinct object instance labels for an input image.

Motivation

The main motivation behind this project is that most of the current instance segmentation methods are unable to find the object instances from the detected saliency regions and this method. These image object instances are important for object recognition and improving the vision based pipelines.

Overview

What is Instance-Level Salient Object Detection?

- ➔ Salient object detection attempts to locate the most noticeable and eye-attracting object regions in images.
- ➔ It is a fundamental problem in computer vision and serves as a pre-processing step to facilitate a wide range of vision applications.
- ➔ This project tackles a more challenging task, instance-level salient object segmentation, which aims to identify individual object instances in the detected salient regions.

Steps Involved

1) Estimating binary saliency map

In this sub-task, a pixel-level saliency mask is predicted,indicating salient regions in the input image.

2) Detecting salient object contours

In this sub-task, we perform contour detection for individual salient object instances.

3) Identify salient object instances

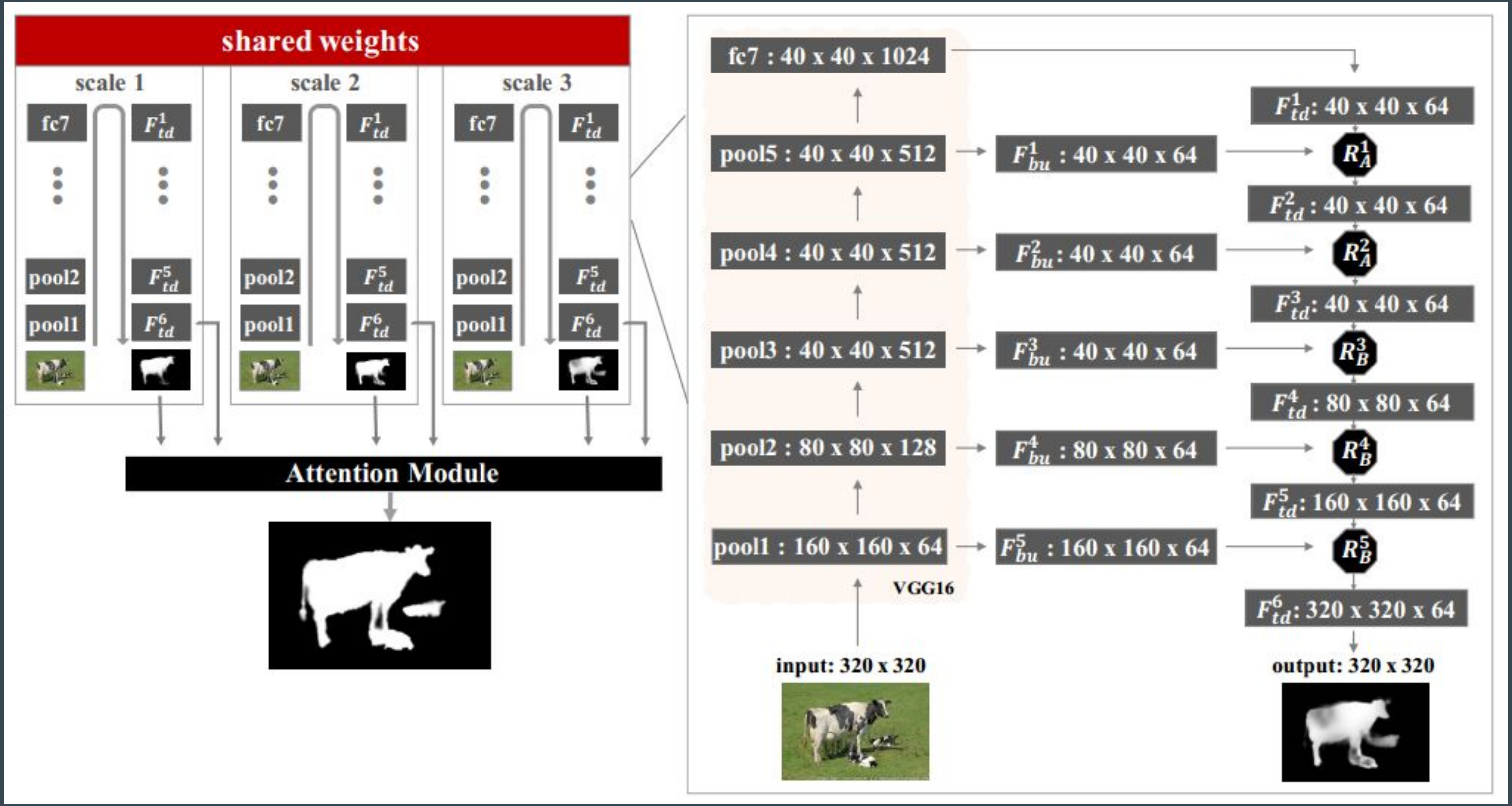
In this sub-task, salient object proposals are generated, and a subset of salient objects proposals are selected to cover the salient regions.

CRF based refinement method is applied to improve the spatial coherence.

Binary Saliency Map and Object Contouring

- We will use a deep multi-scale refinement network (MSRNet), which can generate very accurate results for both salient region detection and object contour detection.
- The deep network consists of three parallel streams processing scaled versions of the same input image and a learned attention model to fuse results at different scales from the three streams. The three streams share the same network architecture, a refined VGG network, and its associated parameters. This refined VGG network is designed to integrate the bottom-up and top-down information in the original network.
- MSRNet can not only integrate bottom-up and top-down information for saliency inference but also attentionally determine the pixel-level weight of each salient map by looking at different scaled versions of the same image.

MSRNet Architecture



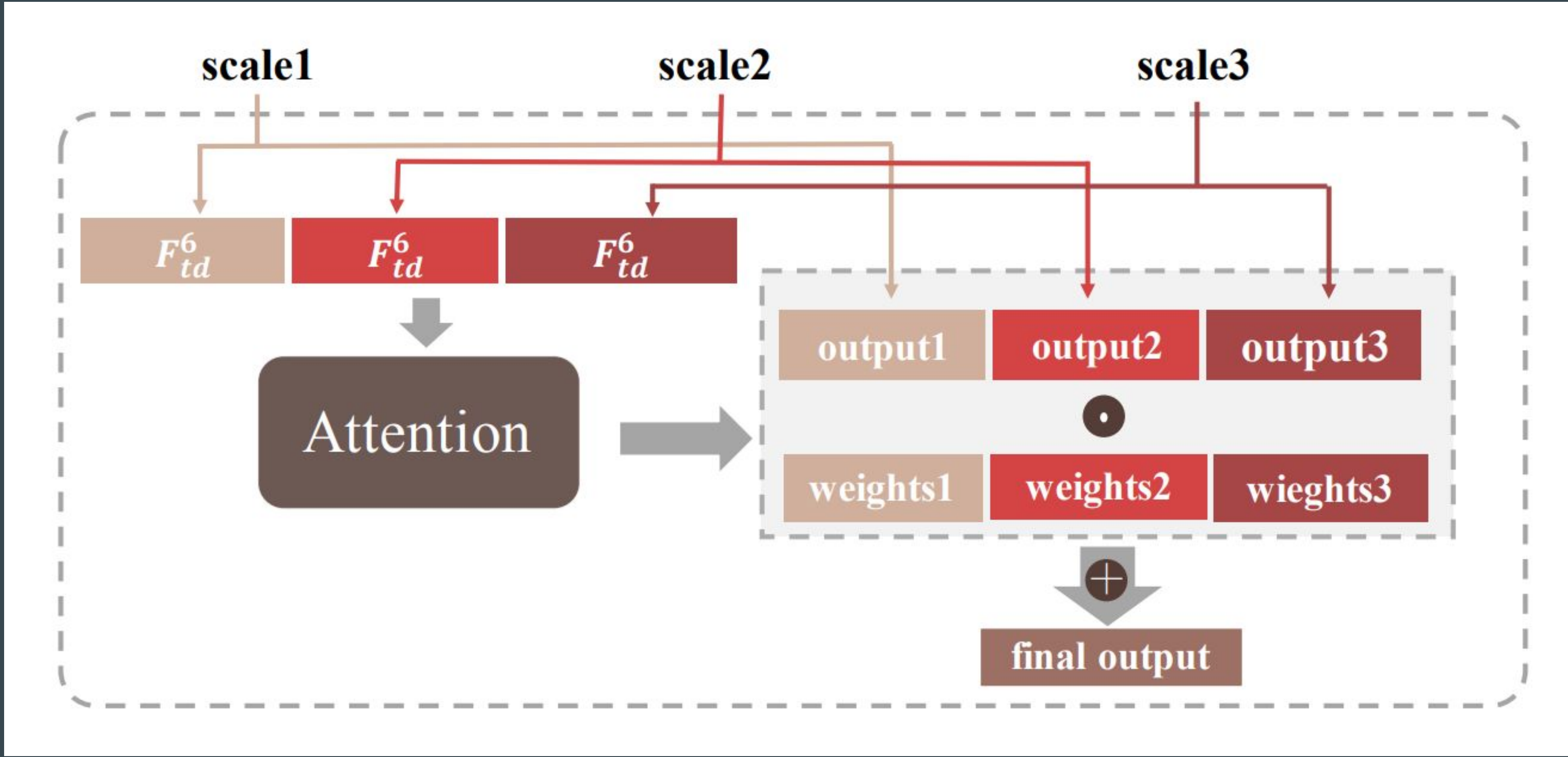
MSRNet Features

- Bottom-Up Network :
  - We use a modified VGG network. It takes low level features (pixels) as input, such as colours and texture, and propagates it up through the layers.
  - Information from an input image needs to be passed from the bottom layers up in a deep network before being transformed into high-level semantic information.
- Top-Down Network :
  - We use a top down model to use and incorporate high level semantic information.
  - It is propagated from the top layers further down, and is integrated with low-level information obtained from the intermediate stages of the Bottom-up network.
  - This integration of high level information with low level features, results in high precision contour detection results.

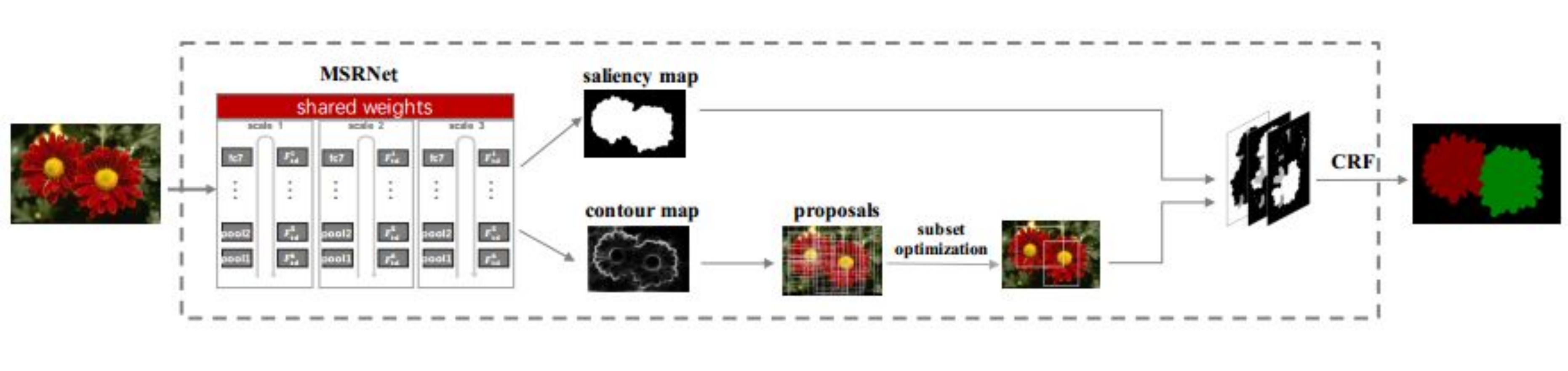
MSRNet Features

- MSRNet :
    - As we can conclude from the properties of the above networks, a network should consider both bottom-up and top-down information propagation and output a label map.
    - We want the saliency map and contour size to have the same resolution as the input image. Thus the output should be a label map of same resolution as the input image.

Attention Module Architecture



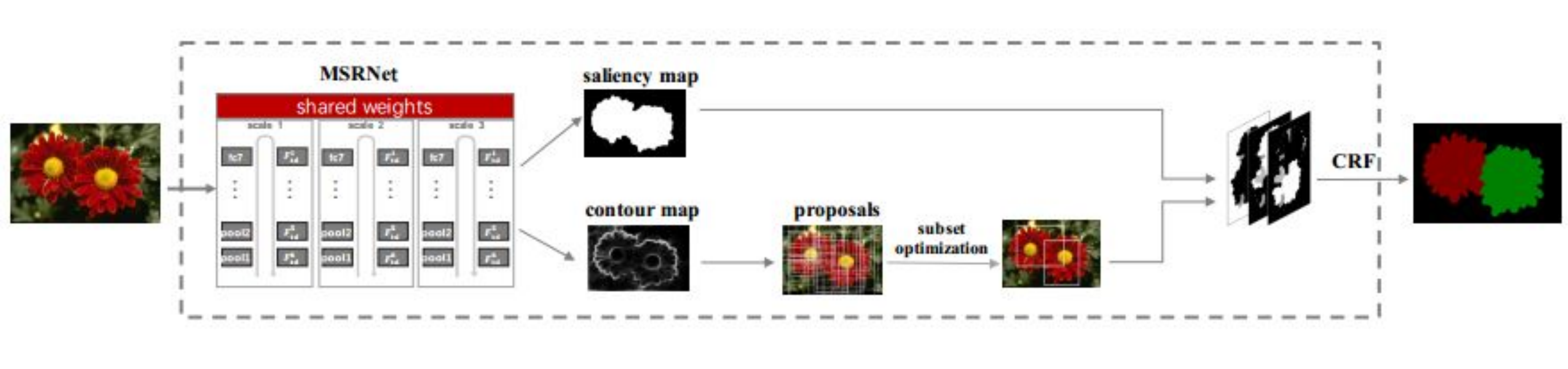
Algorithm Pipeline



Identifying Salient Object Instances

- Given the detected contours of salient object instances, we apply multiscale combinatorial grouping (MCG) to generate a number of salient object proposals.
- Though the generated object proposals are of high quality, they are still noisy and tend to have severe overlap.
- We apply MAP based subset optimisation to optimize the number and locations of the detected windows from the set of object proposals.
- We further filter out noisy or overlapping proposals and produce a compact set of segmented salient object instances. Finally, a fully connected CRF model is employed to improve spatial coherence and contour localization in the initial salient instance segmentation.

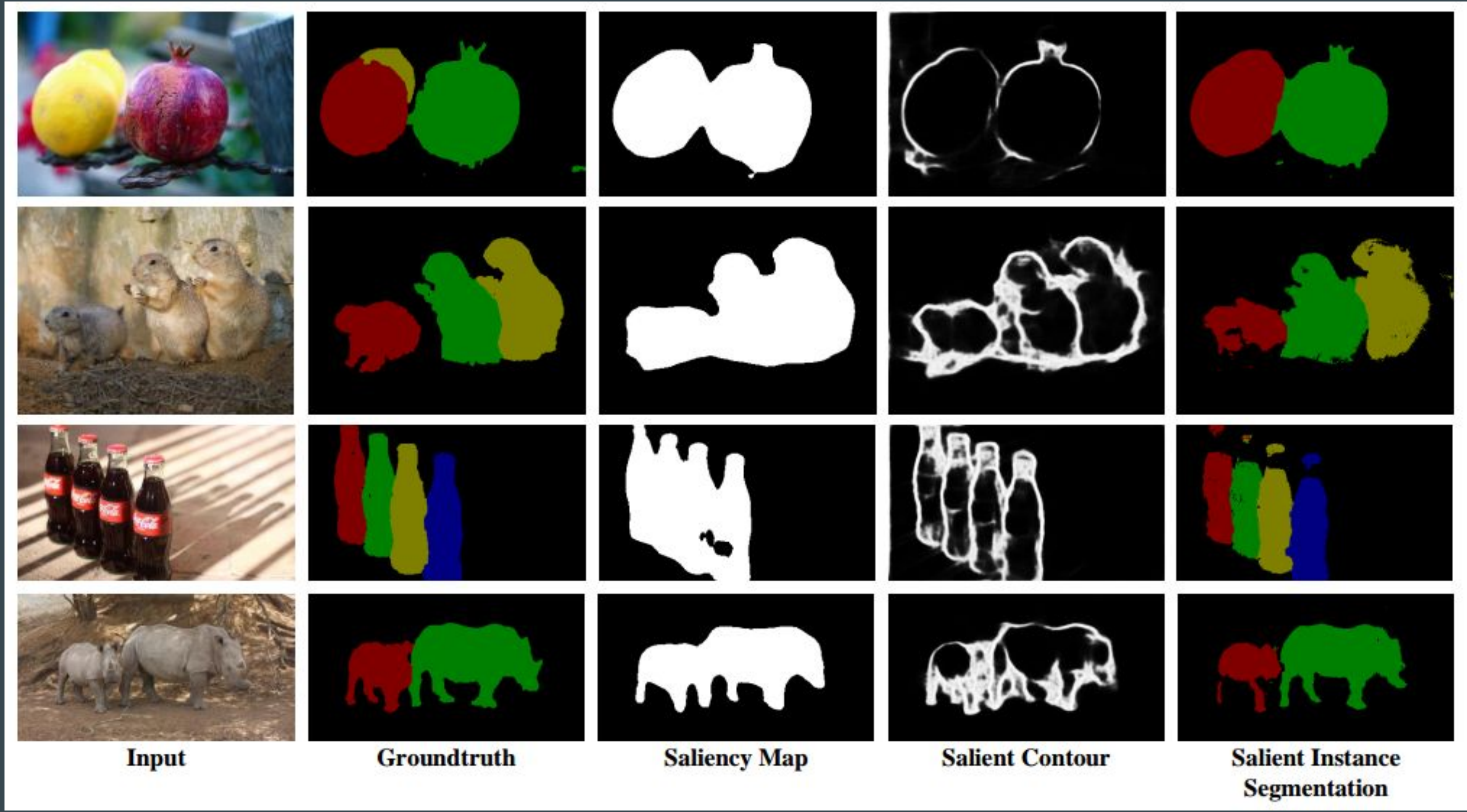
Algorithm Pipeline



Milestones

What we did	What's left
● Implemented MSRNet in Keras Framework. ● Obtained Dataset.	● Implement MCG algorithm. ● Implement CRF module.

Expected Results



Thank You