ICANN19 Object-Level Salience Detection By Progressively Enhanced Network

28th International Conference on Artificial Neural Networks

Wang Yuan[1*], Haichuan Song[1*], Xin Tan[2], Chengwei Chen[1], Shouhong Ding[3], Lizhuang Ma[1,2]

1 East China Normal University, Shanghai, China

2 Shanghai Jiao Tong University, Shanghai, China

3 Tencent, Shanghai, China

Abstract

Saliency detection plays an important role in computer vision area. However, most of the previous works focus on detecting the salient regions, rather than the objects, which is more reasonable in many practical applications. In this paper, a framework is proposed for detecting the salient objects in input im-ages. This framework is composed of two main components: 1). progressive-ly enhanced network (PEN) for amplifying the specified layers of the net-work and merging the global context simultaneously; 2). object-level bound-ary extraction module (OBEM) for extracting the complete boundary of the salient object. Experiments and comparisons show that the proposed frame-work achieves state-of-the-art results. Especially on many challenging da-tasets, our method performs much better than other methods.

Introduction

Given an input image and extracting the most salient object are vital in computer vision. This is also known as saliency detection. Because most of the previous work aim to extract the most salient regions, rather than salient objects. This gets worse when the difference between foreground and background are inconspicuous (the target person head and the background in Fig.1.), while the target object consists of several parts with significant differences (the head and the white shirt of the target person in Fig.1.) According to previous methods, it is easy to detect the white shirt is salient, but it's tricky to distinguish the head from the background, and group the head and shirt as a whole. So their maps are always incomplete. Therefore, object-level detection is more reasonable and important in saliency detection.

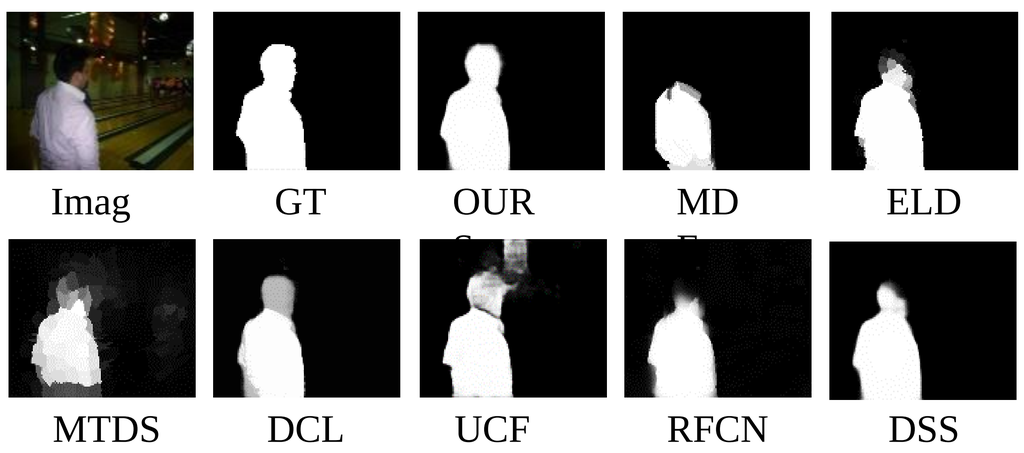


Fig.1. Result maps of different methods. The results of previous methods are either blurred or incomplete.

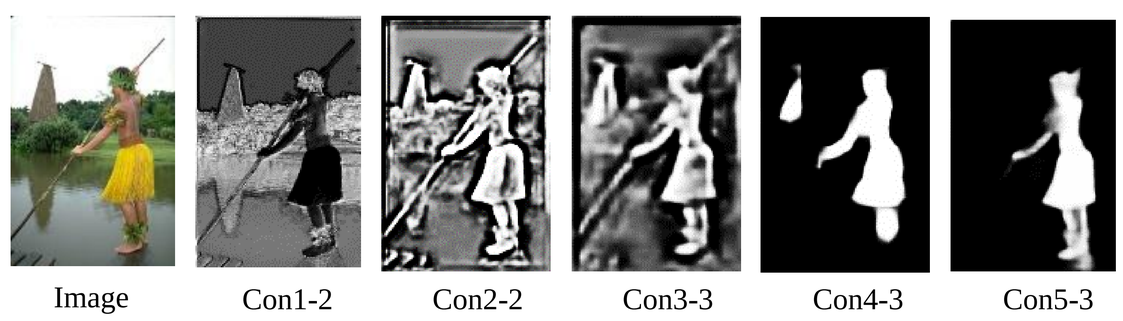
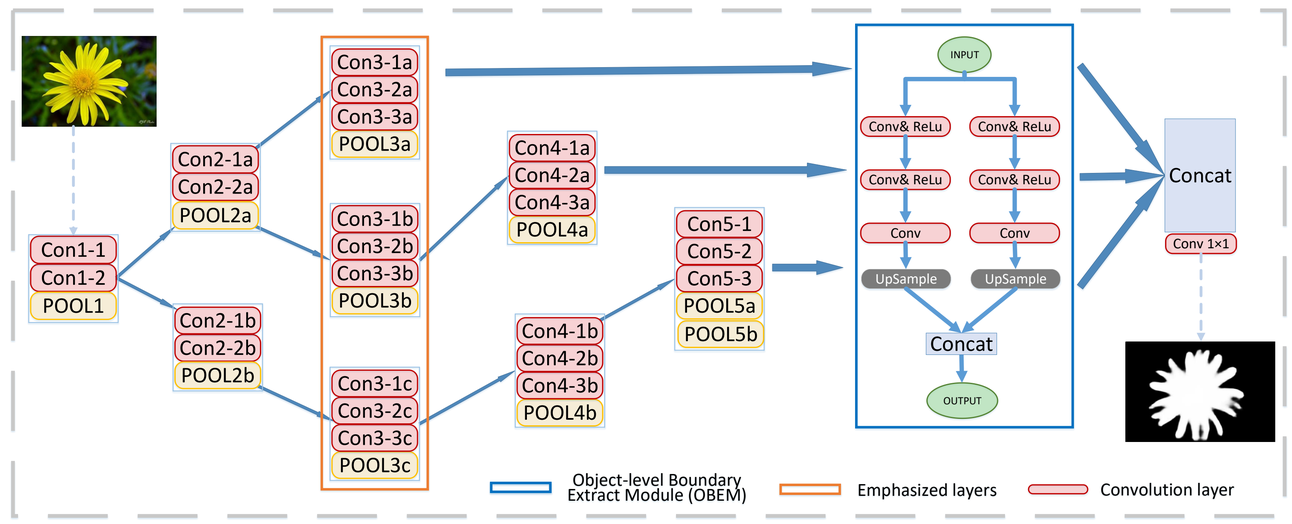


Fig.2. Function of different layers. From left to right: the input image, the output maps of the convolution layers 1-2, 2-2, 3-3, 4-3 and 5-3 of PEN showed in Fig.3.



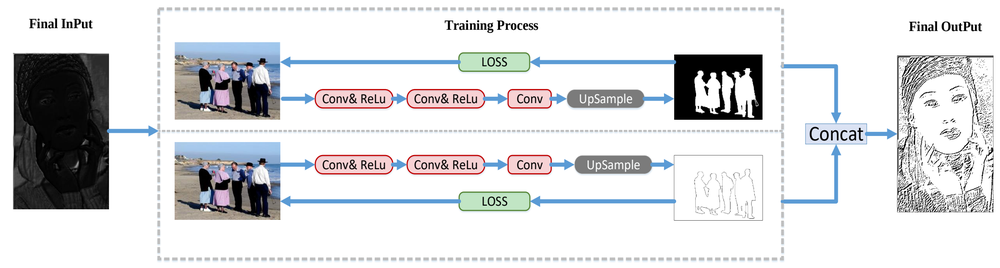


Fig.3. The main architecture of progressively enhanced network (PEN) with object-level boundary extraction module (OBEM) embedded. OBEM marked with blue box will be detailed in sec.3.2. The others will be proposed in sec 3.1. And the architecture of OBEM and its effect in the final result.

Conclusion

In order to obtain a better detection of the salient objects, rather than the regions, a novel framework is proposed in this paper. It consists of two main components: PEN and OBEM. By amplifying the specified layers from FCNs, and merging the global context at the same time, PEN is able to separate the objects from the background effectively, even in complex cases. OBEM on the other hand is able to detect the objects with more precise boundaries. Experiments results show that the complete framework with PEN and OBEM have better performance than other methods, especially when the raw image is complicated on semantics.

ManuscriptID: 239

Other contact Information:
51184501076@stu.ecnu.edu.cn,
hcsong@sei.ecnu.edu.cn