YOLACT: Real-time Instance Segmentation

Nguyễn Tấn Phúc

University of Information Technology
HCMC, Vietnam

What ?

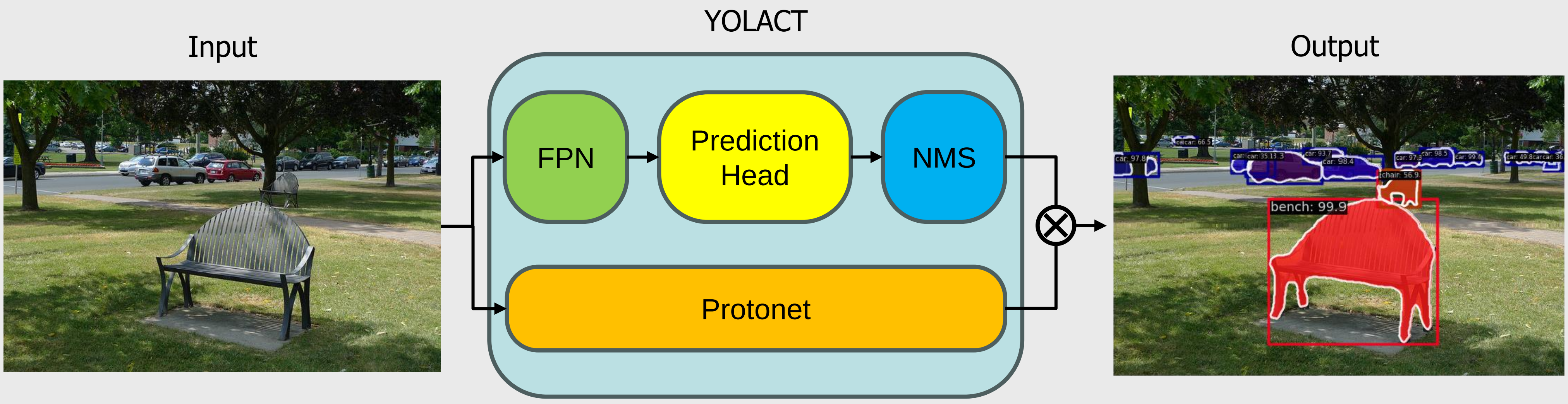
We introduce an instance segmentation algorithm, in which we have:

- Proposed a innovative one-stage instance segmentation algorithm to delineate and identify objects within an image.
- Proposed Fast NMS to trade a little performance off for speed.
- Compare our model with competitive SOTA methods.

Why ?

- Instance segmentation is necessary as it provides detailed information about individual objects within an image. This fine-grained segmentation is essential for various applications, such as object tracking, scene understanding, and even medical imaging.
- Prevailing two-stage methods have focus primarily on **performance** over **speed**. One-stage methods still have been far from **real-time**.

Overview



Description

1. Instance Segmentation

- YOLACT breaks up instance segmentation (see Fig. 1) into two parallel tasks:
  - (1) generating a dictionary of non-local prototype masks over the entire image.
  - (2) predicting a set of linear combination coefficients per instance.
- Then producing a full-image instance segmentation from these two components: for each instance, linearly combine the prototypes using the corresponding predicted coefficients and then crop with a predicted bounding box.

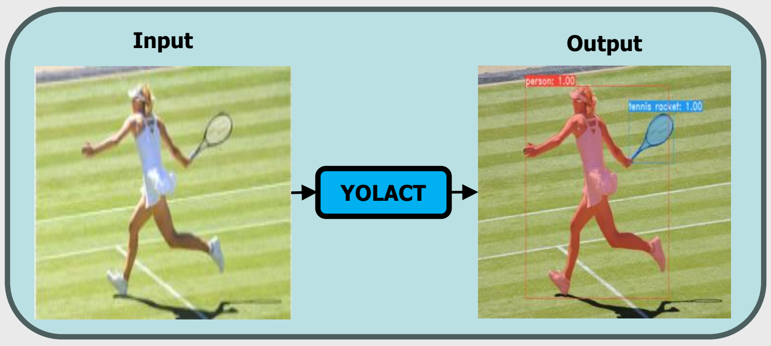


Figure 1. Input & Output illustration

2. Prototype Generation

- Uses an Fully Convolution Network to produce a set of image-sized "prototype masks" that do not depend on any one instance (see Fig. 2).

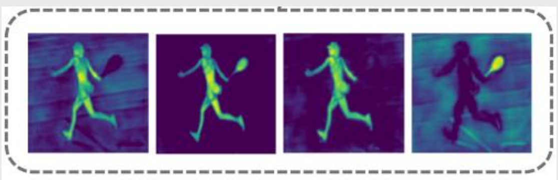


Figure 2. Prototype Generation

3. Mask Coefficients

- Adds an extra head to the object detection branch (RetinaNet backbone with ResNet-101 and FPN) to predict a vector of "mask coefficients" for each anchor that encode an instance's representation in the prototype space (see Fig. 3)

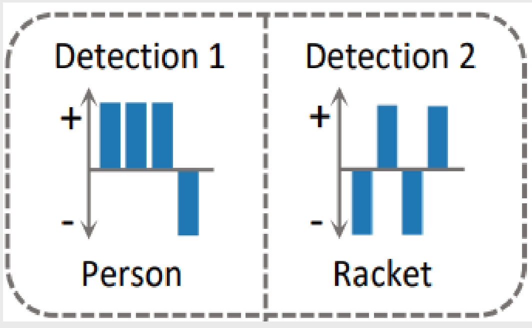


Figure 3. Mask coefficients

4. Mask Assembly

- For each instance that survives NMS, we construct a mask for that instance by linearly combining the work of those two stages above (see Fig. 4).

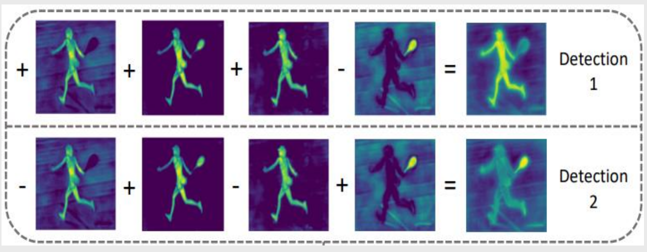


Figure 4. Mask assembly

NII

Nguyễn Tấn Phúc – Trường Đại học Công nghệ Thông tin
TEL : 0941959027 Email : phucnt.17@uit.edu.vn