Note: Learning Deep Features for Discriminative Localization

sgc

May 28, 2022

1 Introduction

This is one of the most important and classic method in WSSS, it is published in 2016.

2 Method

A class activation map for a particular category indicates the discriminative image regions used by the CNN to identify the category. Global average pooling outputs the spatial average of the feature map of each unit at the last convolutional layer. A weighted sum of the feature maps of the last convolutional layer is our class activation maps.

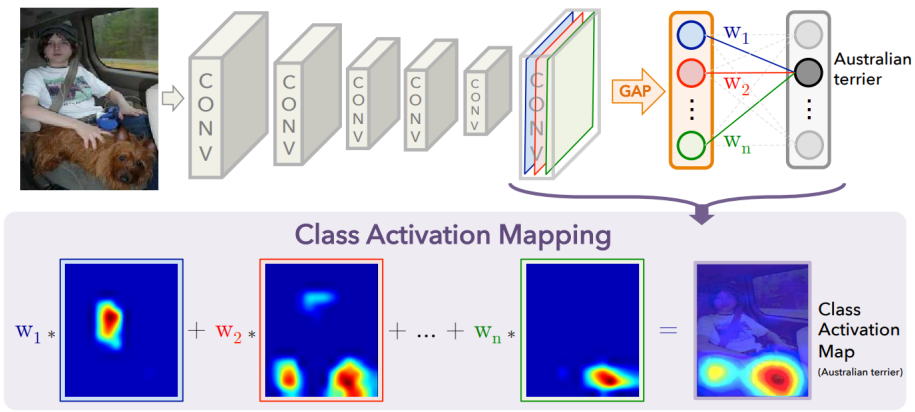

Figure 2. Class Activation Mapping: the predicted class score is mapped back to the previous convolutional layer to generate the class activation maps (CAMs). The CAM highlights the class-specific discriminative regions.

Detailed procedure: $f_k(x, y)$ is the activation of unit k in the last convolutional layer at spatial location(x,y).
ω_k^c indicates the importance of F_k for class c, it is the weight corresponding to class c for unit k.

First the class score

$$S_c = \sum_x \omega_k^c \sum_{x,y} f_k(x,y) = \sum_x \sum_{x,y} \omega_k^c f_k(x,y)$$

The CAM for class c M_c is generated by

$$M_c(x,y) = \sum_k \omega_k^c f_k(x,y)$$

At last project the CAM back to the size of original image.

References