

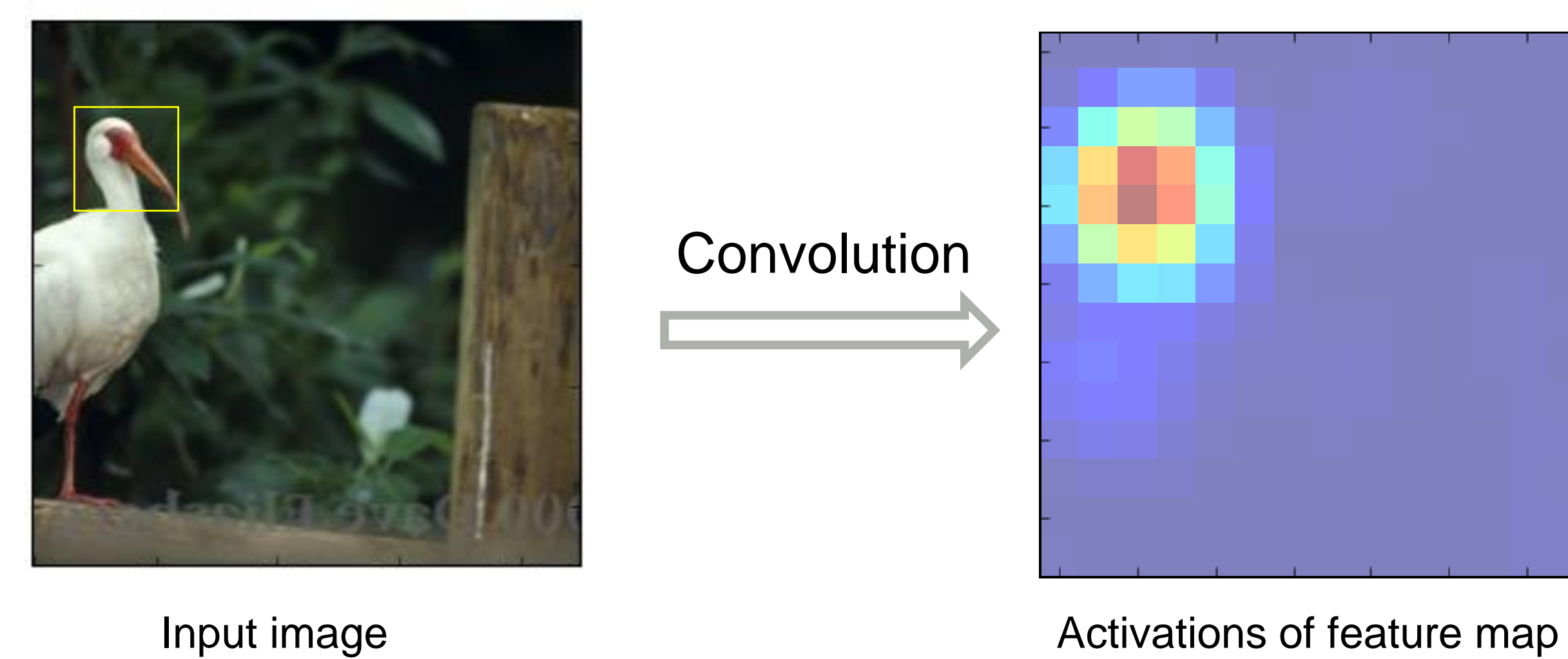
# WEAKLY SUPERVISED OBJECT LOCALIZATION WITH DEEP CONVOLUTIONAL NEURAL NETWORK BASED ON SPATIAL PYRAMID SALIENCY MAP

Zhiqiang Wan and Haibo He  
University of Rhode Island, RI, USA, 02881

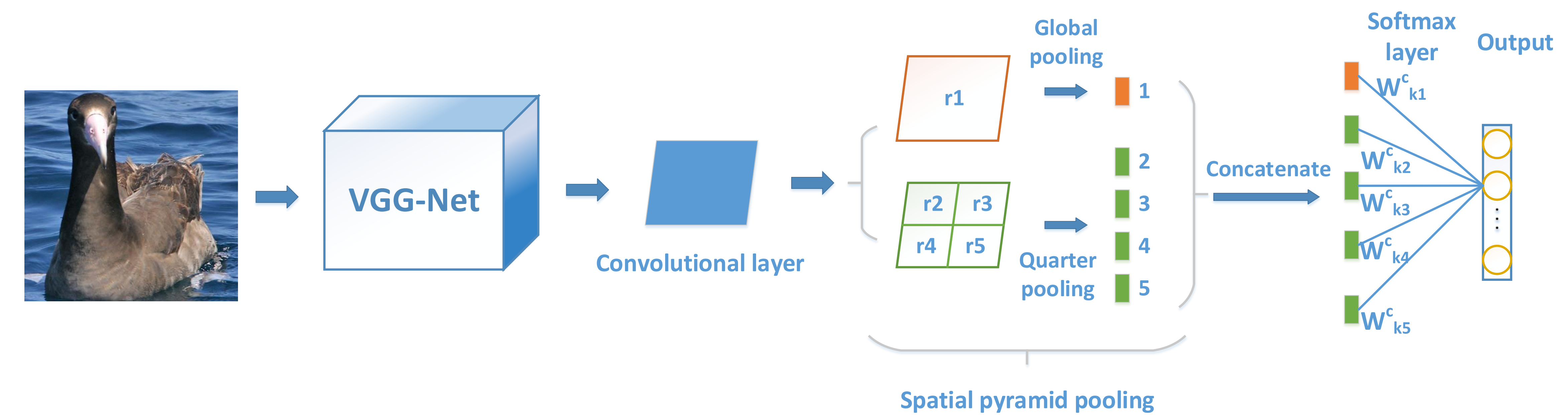
## Outline:

- Background for weakly supervised object localization;
- Key idea: exploiting the localization ability of the convolutional layer;
- Experimental results and analysis;
- Conclusions and discussions;

## Localization ability of the convolutional layer:



## Architecture of the proposed model:



### Spatial Pyramid Pooling

Global pooling:

$$z_{k1} = \frac{1}{N_1} \sum_{x,y \in r1} f_k(x,y)$$

Quarter pooling:

$$z_{k2} = \frac{1}{N_2} \sum_{x,y \in r2} f_k(x,y)$$

### Softmax Layer

Input:

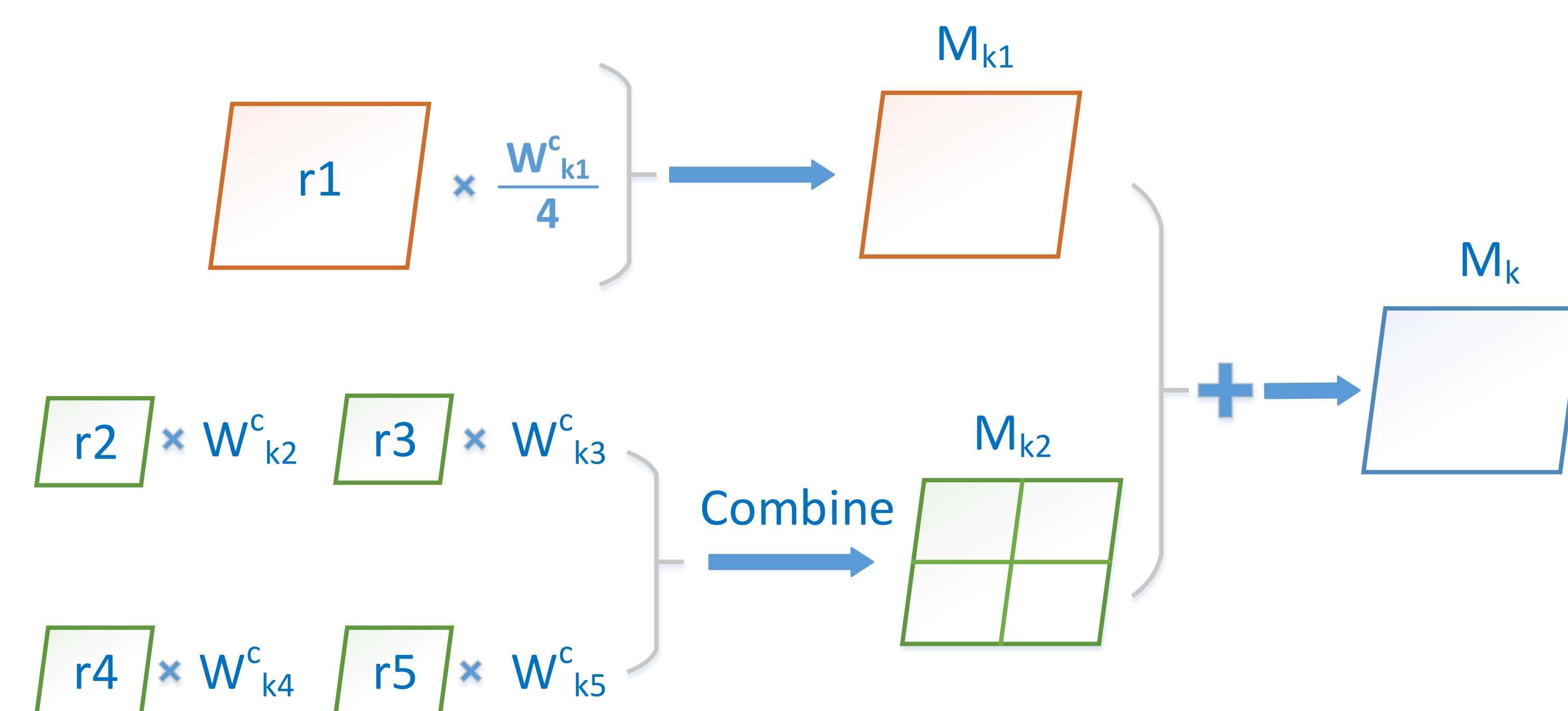
$$S_c = \sum_k w_{k1}^c z_{k1} + \sum_k w_{k2}^c z_{k2} + \sum_k w_{k3}^c z_{k3} + \sum_k w_{k4}^c z_{k4} + \sum_k w_{k5}^c z_{k5}$$

Output:

$$P_c = \frac{\exp(S_c)}{\sum_c \exp(S_c)}$$

## Object localization with spatial pyramid saliency map:

### Importance of each feature map:



### Spatial pyramid saliency map construction:

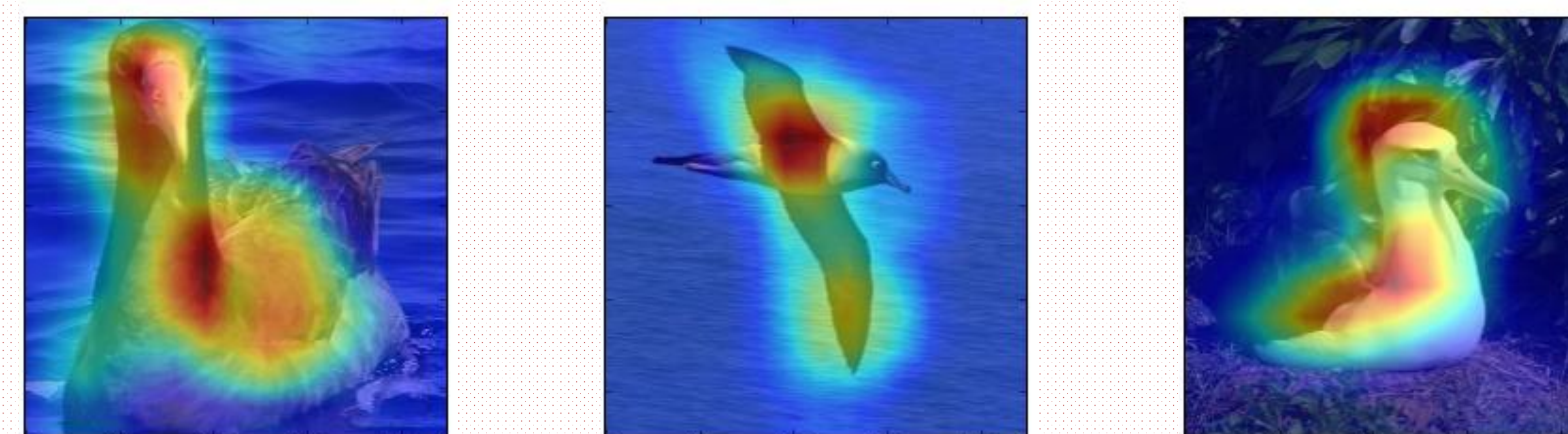
$$I_c(x,y) = \sum_k M_k(x,y)$$

### Object localization:

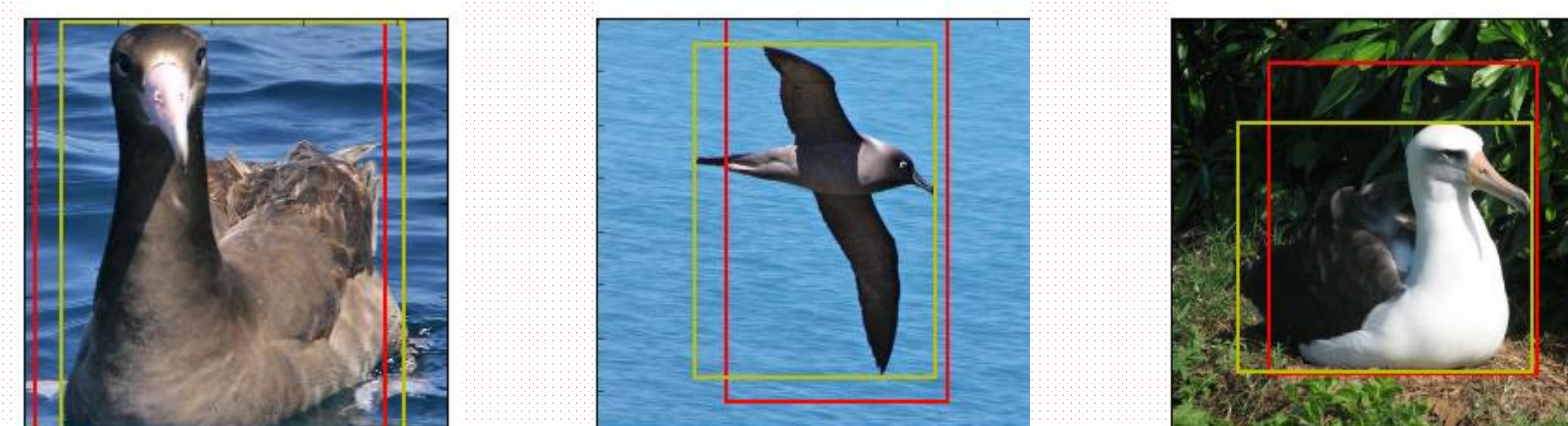
- Up-sample the saliency map into the size of the input image.
- Highlight the object with this saliency map.

## Experimental results:

### Object localization results on Cub-200 dataset:

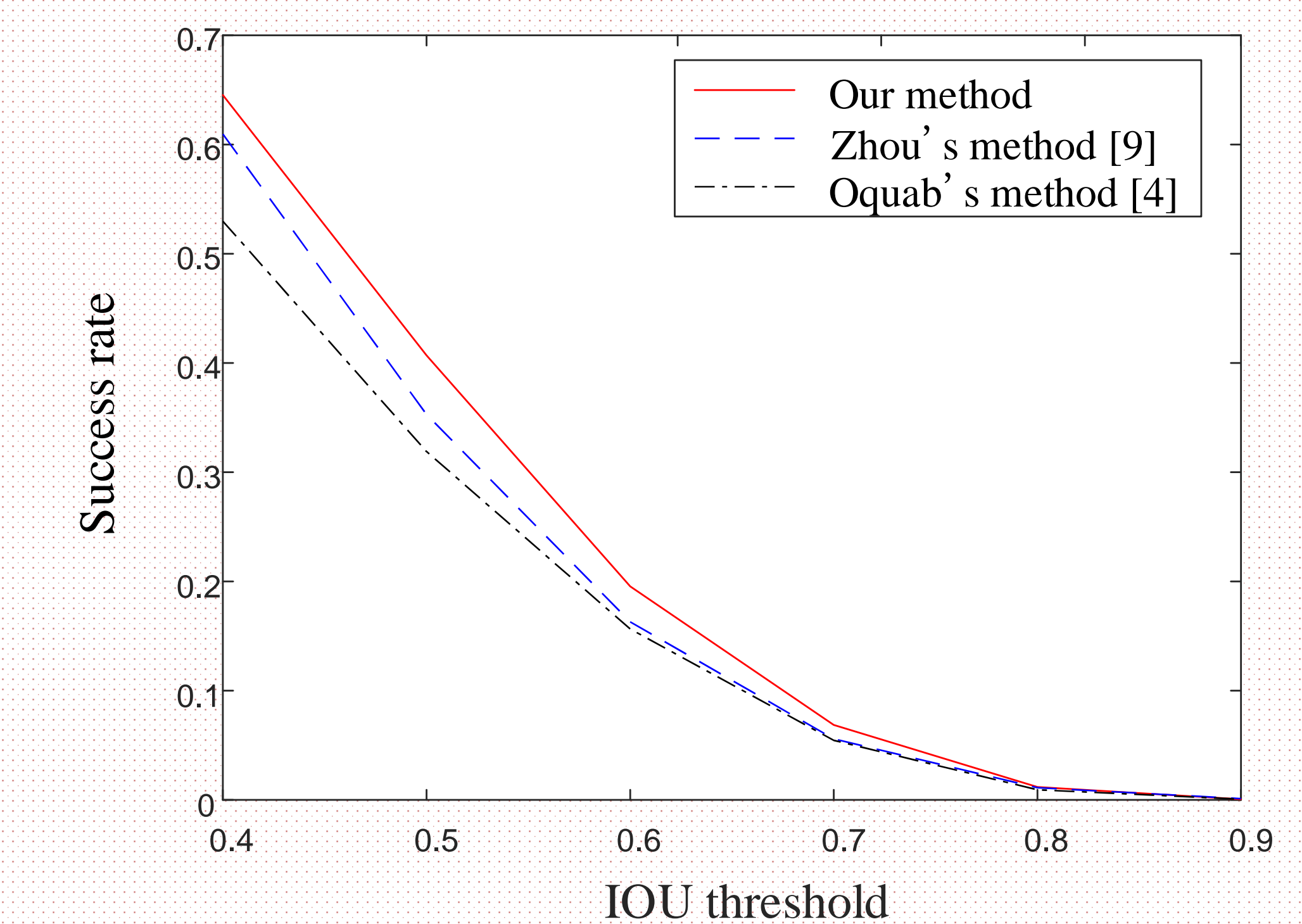


(a) Spatial pyramid saliency map



(b) Localization results (red) and ground truth (yellow)

### Performance comparison with benchmark methods:



Success rate under different IOU threshold

## Conclusions and Discussions:

- Propose a convolutional neural network based spatial pyramid saliency map to localize object.
- Experimental results on Cub-200 dataset verify the effectiveness of the proposed model.