New Global Pooling
For Generating Score Map

20150560 CSE Dongwon Kim

20150384 CSE Eunyoung Hyung

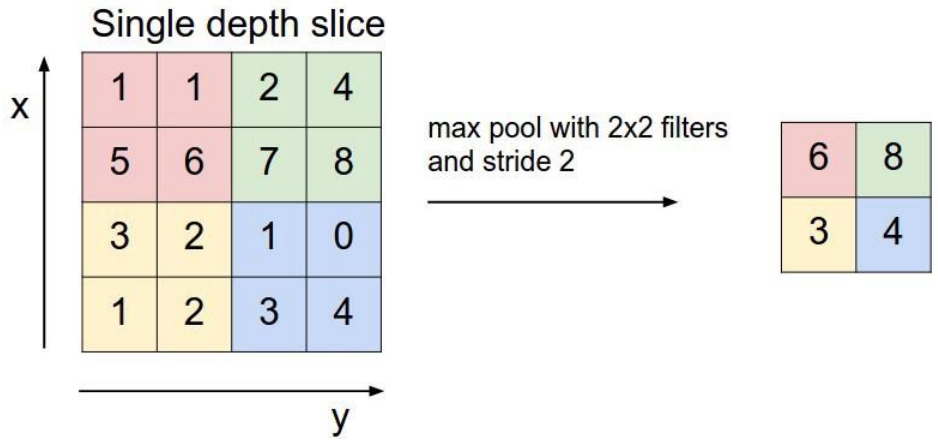
POSTECH

Introduction

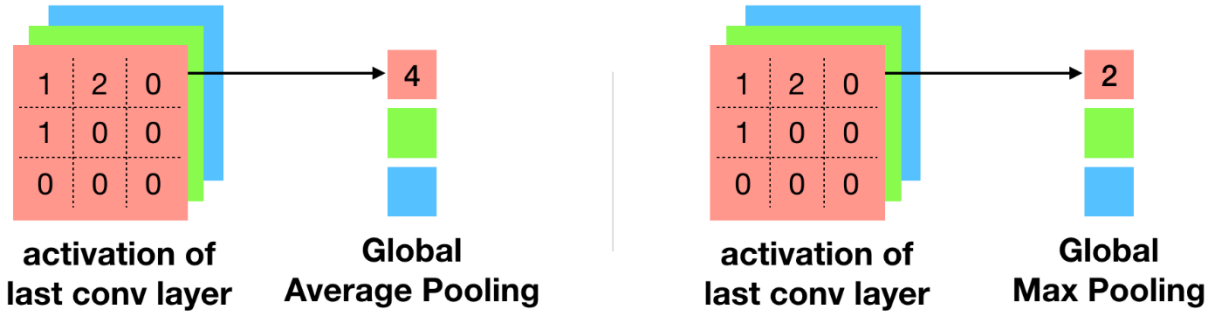
Global Pooling?

Global Pooling

- Pooling

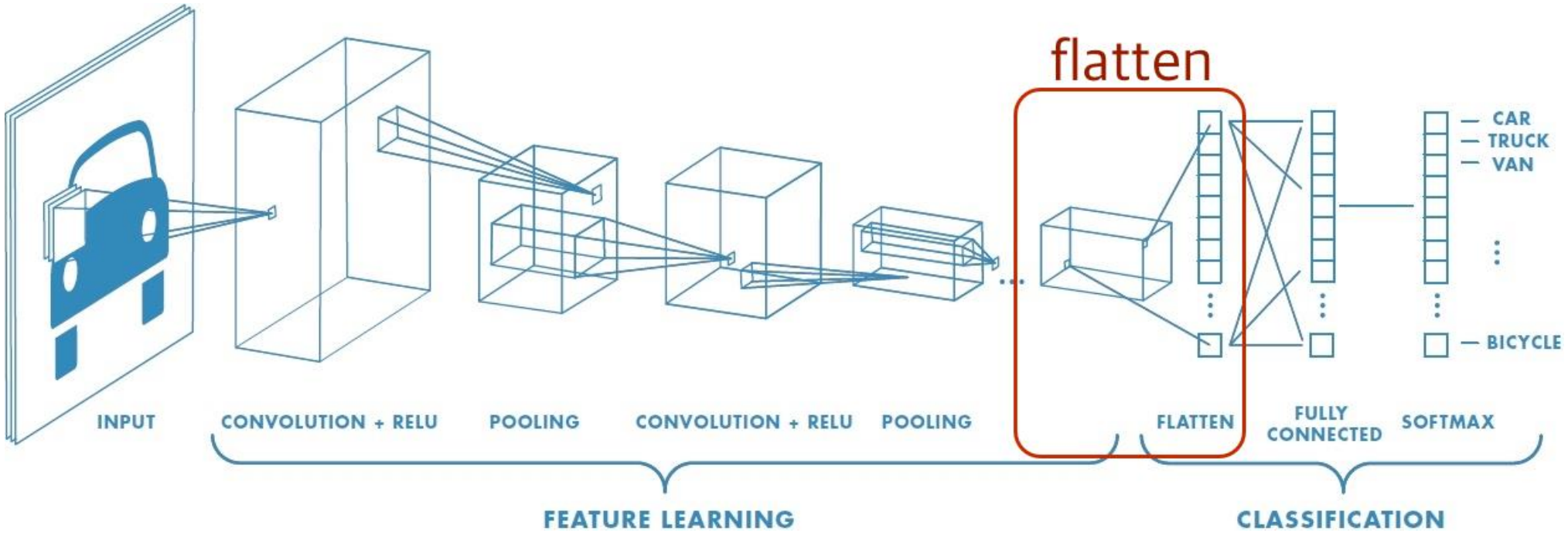


- Global pooling



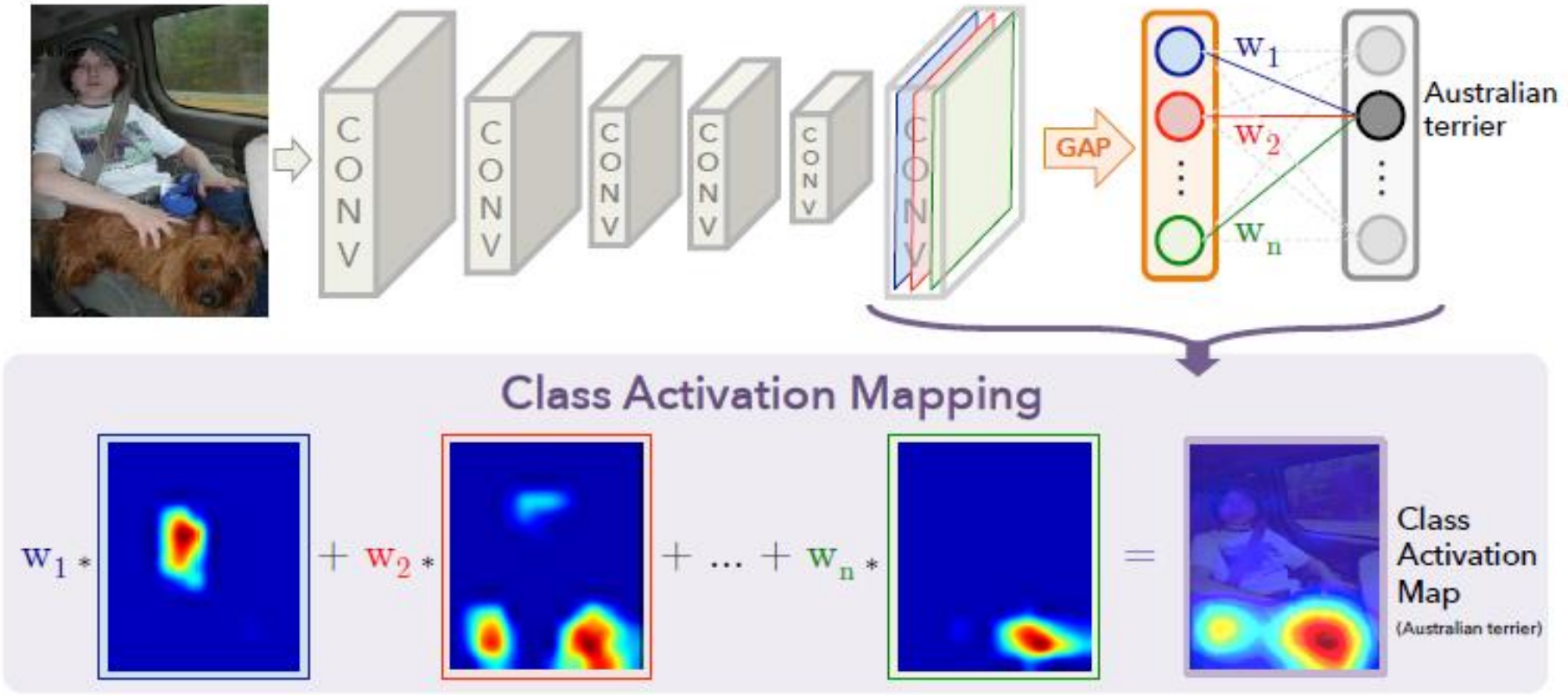
Global Pooling

- CNN for classification



Global Pooling

- Instead, global pooling can be used to estimate the location of object



Related Works

That gave us the idea

Related Works

- Global Max Pooling(GMP)
- Global Average Pooling(GAP)
- Log-sum-exponential Pooling(LSE)
- Global Weighted Rank Pooling(GWRP)

Global Max Pooling(GMP)

- Hypothesizes the location of the object in the image at the position with the maximum score.
- Therefore, GMP often underestimates the sizes of objects.

Figure 4: **Illustration of the weakly-supervised learning procedure.** At training time, given an input image with an aeroplane label

Global Average Pooling(GAP)

- The global average pooling(GAP) outputs the spatial average of the feature map.
- GAP, also used frequently, encourages all responses to be high so it often overestimates them.

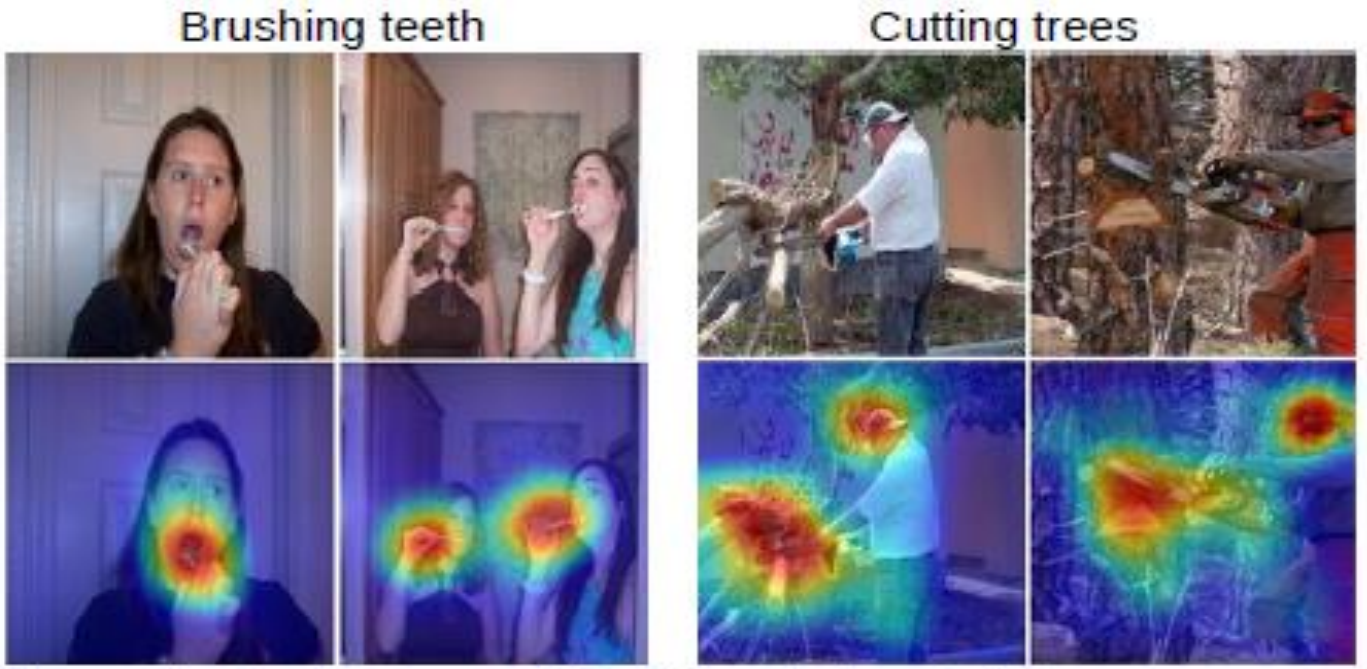


Figure 1. A simple modification of the global average pooling layer combined with our class activation mapping (CAM)

Log-sum-exponential Pooling(LSE)

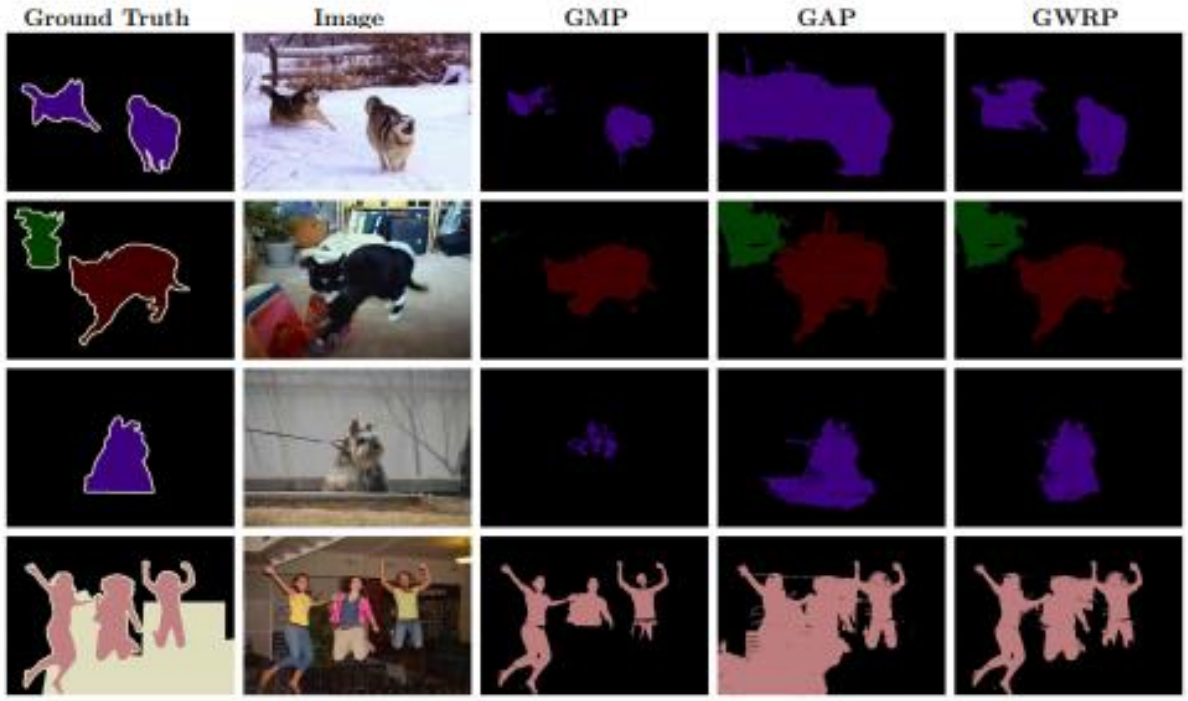
- LSE takes advantages of both GMP and GAP by using a parameter r which makes pooling become between GAP and GMP
  - The function's output is close to GAP when r is small and GMP when r is large.

Average pooling Max pooling LSE pooling

Figure 5. Heat map for class *train* generated by proposal network trained with average pooling, max pooling and LSE pooling respectively.

Global Weighted Rank Pooling(GWRP)

- Also generalization of GMP and GAP
  - Regards high score of pixel as more important one
  - But also takes the low score of pixel into account
- It encourages objects to occupy a certain fraction of an image, but, unlike GAP, is less prone to overestimating object sizes.

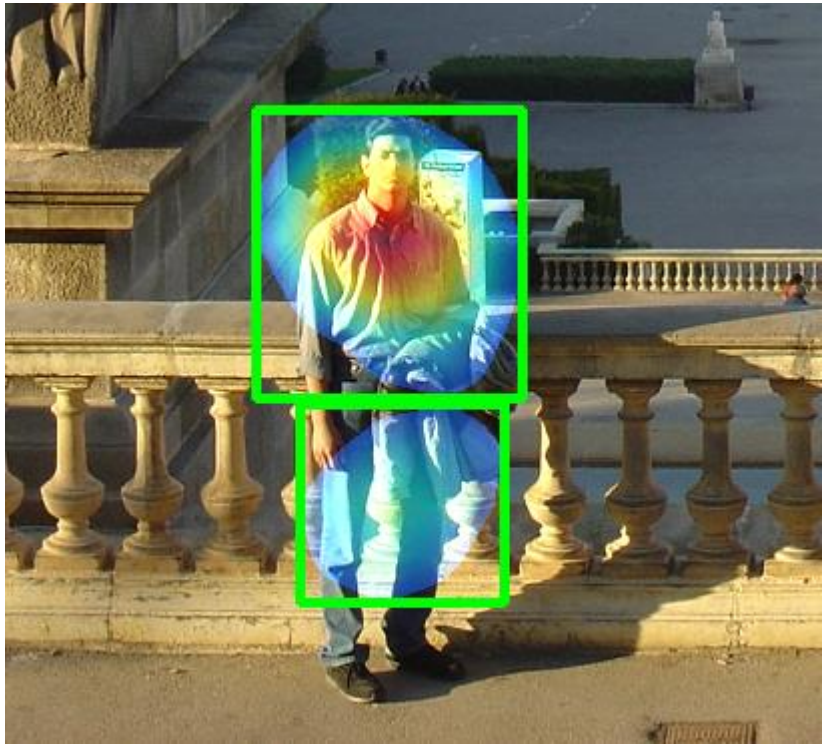


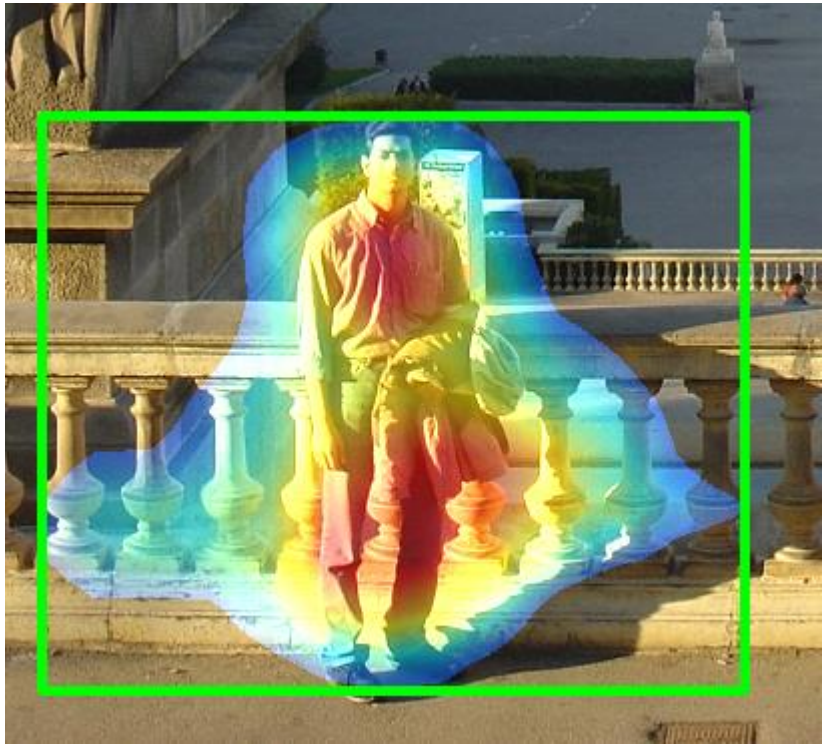
Our Work

New Global Pooling

Problem of GMP and GAP

- GMP assumes smaller size, GAP assumes larger size
- Can we make pooling to assume appropriate size of object?
  - -> Weighted average





Idea

- Consider every pixels in feature map in global pooling
  - But, Focus on more important pixels

- Use weighted average to bring this 'focus' in global pooling
  - Score of each pixels becomes weight
  - No additional hyperparameter
  - No additional learning parameter

- Global Squared Average Pooling(GSAP)
  - $\sum_{i,j} \frac{f_k(i,j)^2}{\sum_{i,j} f_k(i,j)}$

Experiments

Setup

- Keras with Theano background
- VGG16 without last fully connected layer as pre-trained model

- Learning rate 0.001
- Trained for 40 epoch
- Train last 4 convolution layers of VGG16 + fully connected layer

- Dataset: INRIA Person dataset
  - Consisted with 2 category: img with people, and without people
  - Only image and label was trained

- Bounding box
  - It is drawn for regions with value > 20% of the max value of the CAM.

Evaluation

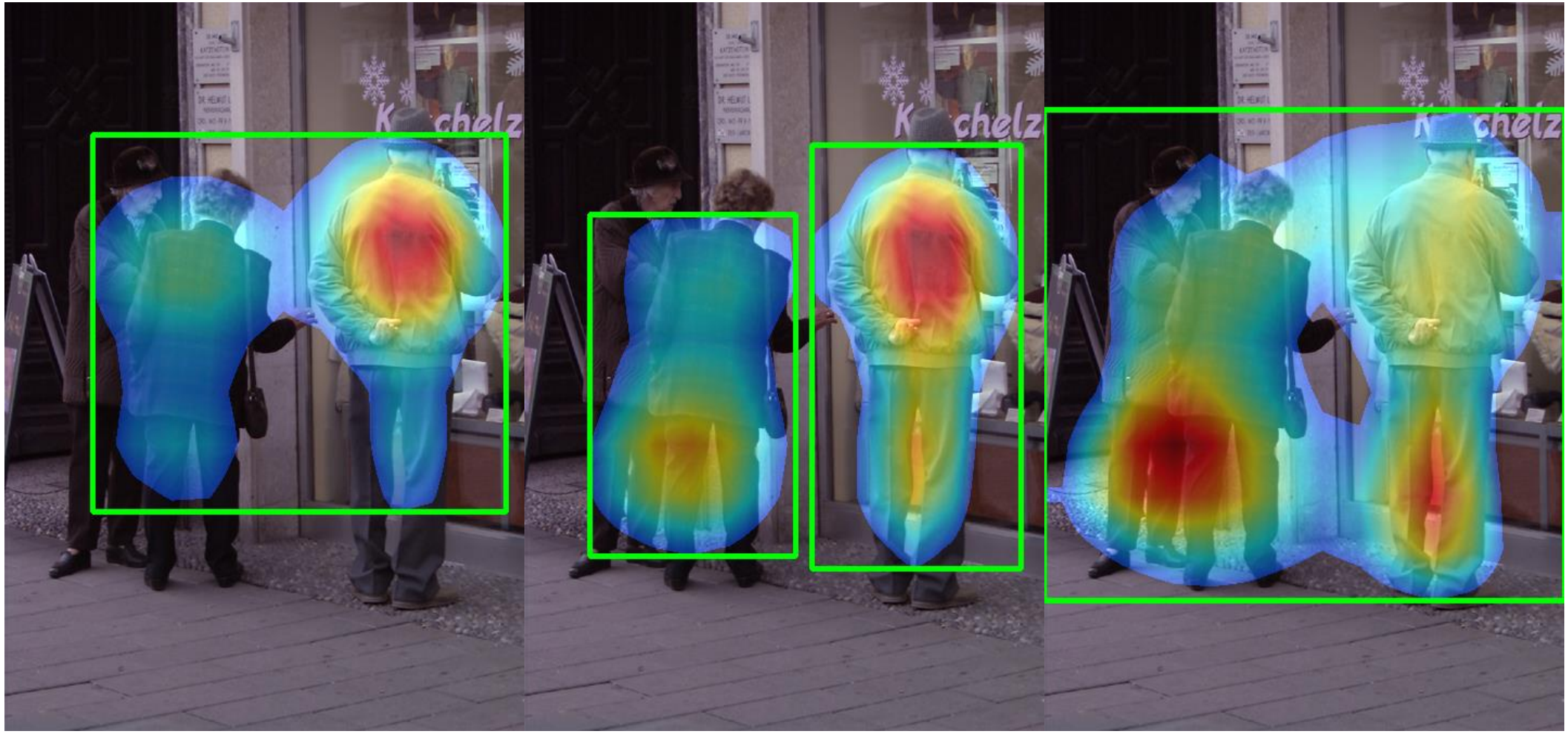
- Classification error
  - Test set of INRIA person dataset(588 images)

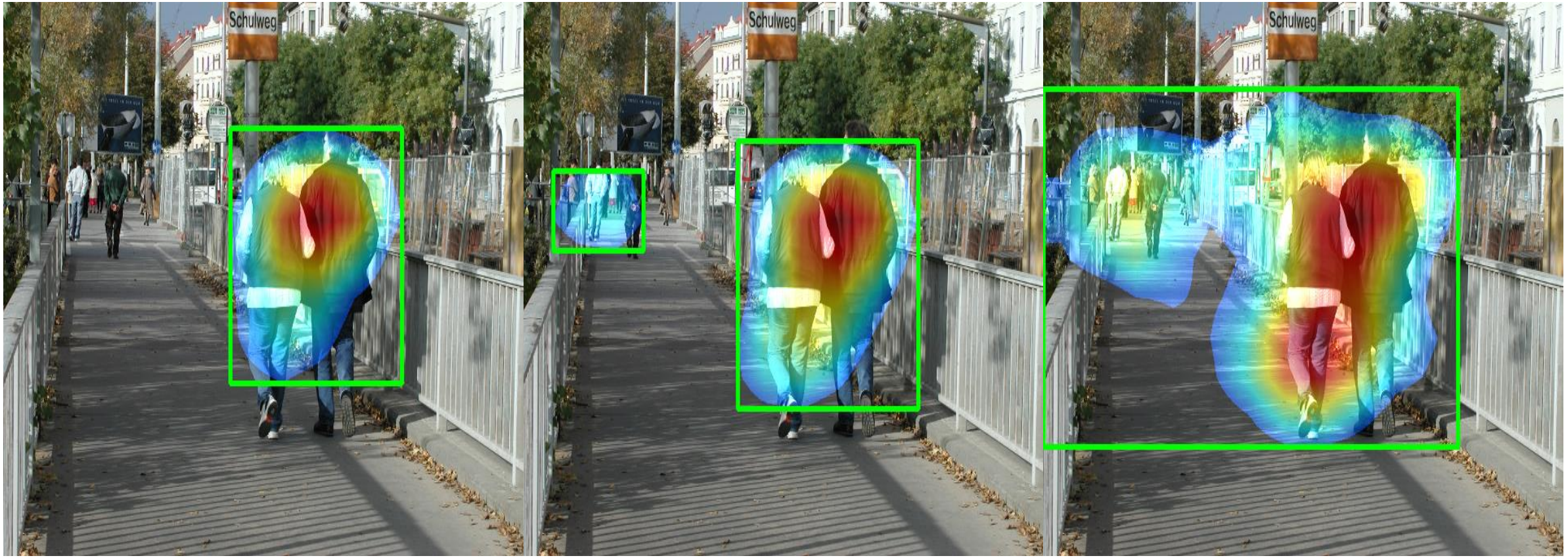
Method	Total loss	Accuracy
GMP	0.58	0.962
GAP	0.07	0.979
GSAP	0.08	0.967

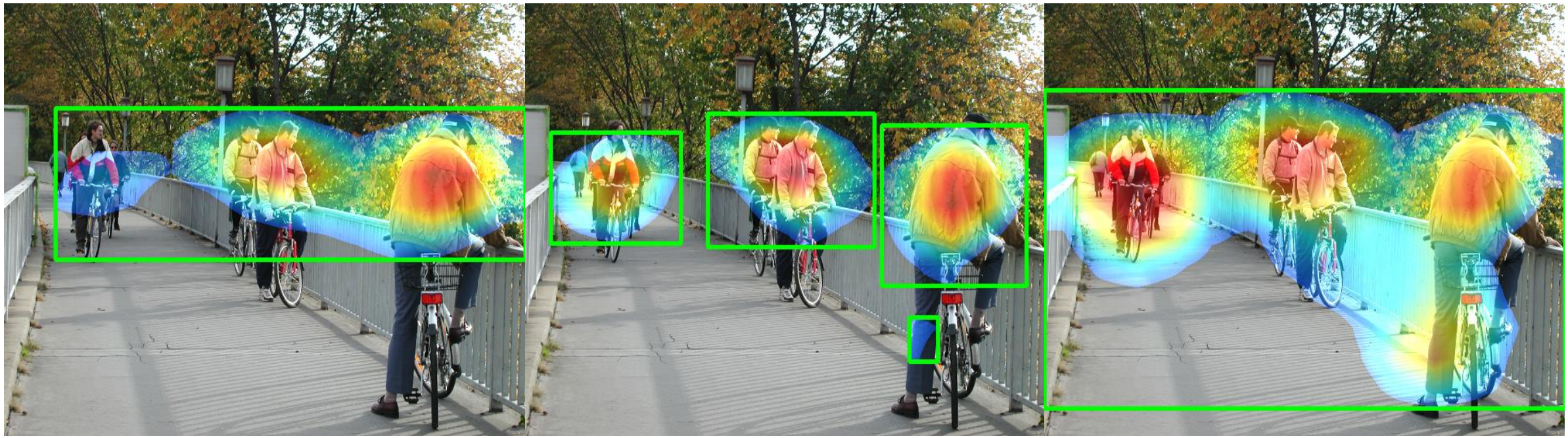
  - No significant accuracy drop compared to GMP and GAP

Evaluation

- Detection evaluation
  - Because of the time limitation, we haven't perform qualitative evaluation.
  - But there are many examples shows GSAP's advantage, especially in multi object detection. (GMP-GSAP-GAP)




Schulweg
Schulweg
Schulweg



Conclusion

- GAP and GMP has problem when estimating object's size.
- GSAP, can solve this issue. It focus more on important region.
- Classification accuracy is higher than GMP, almost similar with GAP.
- GSAP especially performs better in multi-object-detection problem.
  - GMP tends to predict scores for only top1 object.
    - Because it train/predict only top score pixel.
  - GAP tends to predict high scores for interval between object.
    - It focus on every pixel equally, therefore predicted objects becomes larger and m erged.