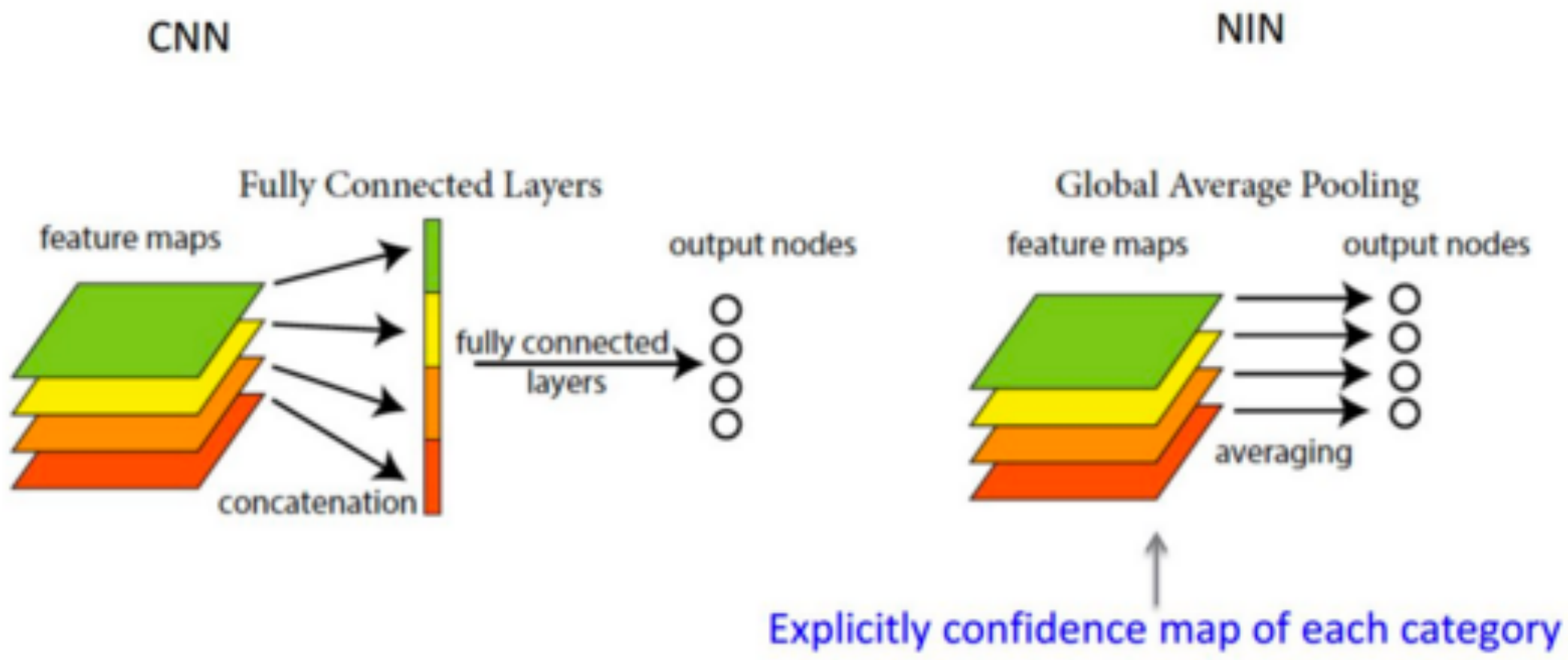
Learning Deep Features for Discriminative Localization

Hongming Shan
shanh@rpi.edu

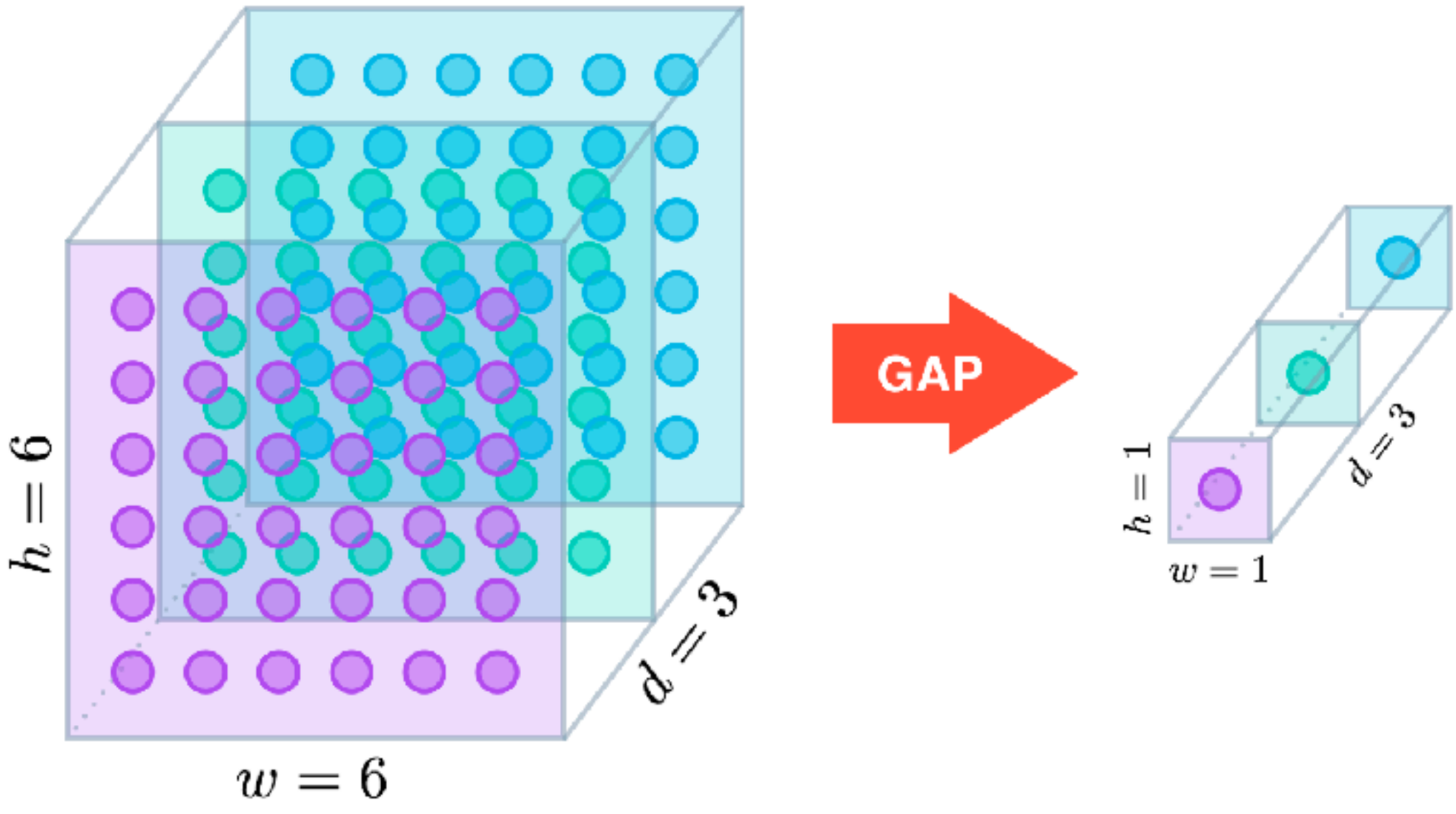
Review : Global Average Pooling in Network in Network

- Qing led a discussion about Network in Network last week.
  - Linear convolution layer -> mplconv layer
  - Fully Connected layer -> **Global average pooling**



Review: Global Average Pooling

- GAP layers reduce each ***h*x*w*** feature map to a single number by simply taking the average of all ***hw*** values.



Learning Deep Features for Discriminative Localization

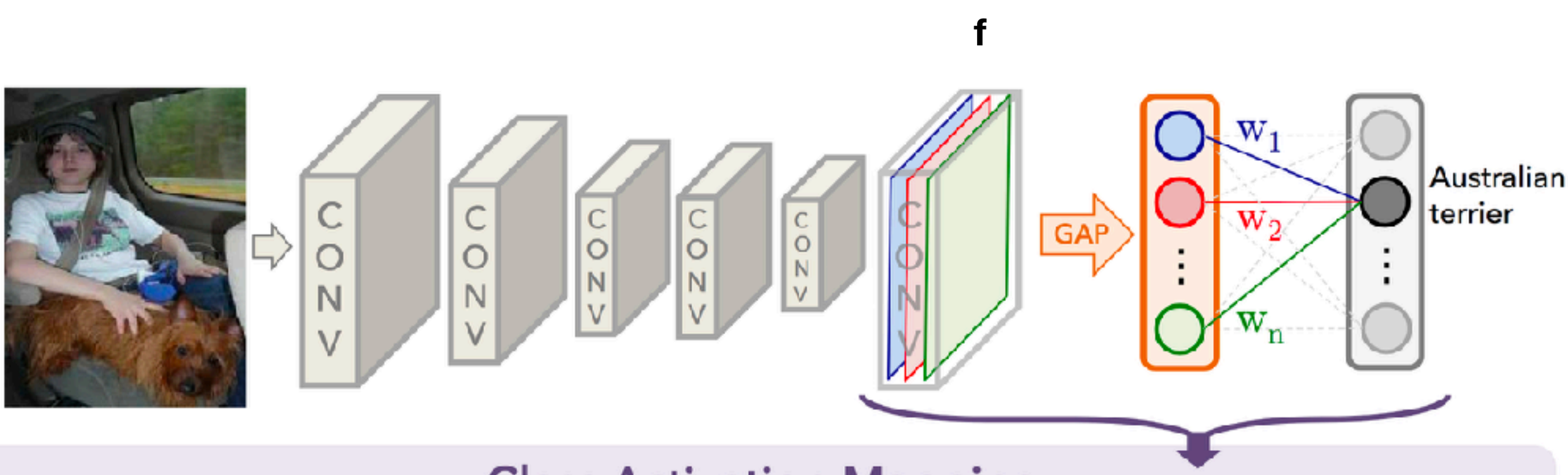
Contribution:

GAP layer can enable CNN to have **remarkable localization ability** despite being trained on **image classification**.

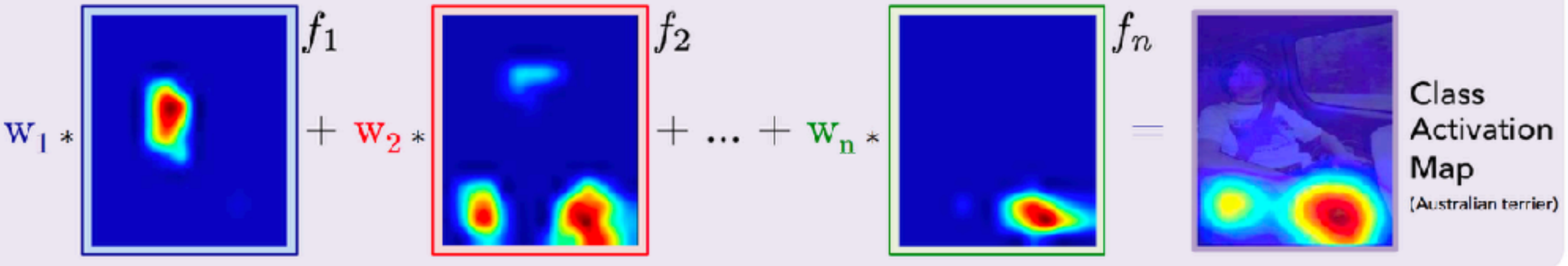
- What object is contained in the image (image classification)

- Where the object is in the image (object localization)

GAP-CNN



Class Activation Mapping



CAMs

Figure 3. The CAMs of two classes from ILSVRC [21]. The maps highlight the discriminative image regions used for image classification, the head of the animal for *briard* and the plates in *barbell*.

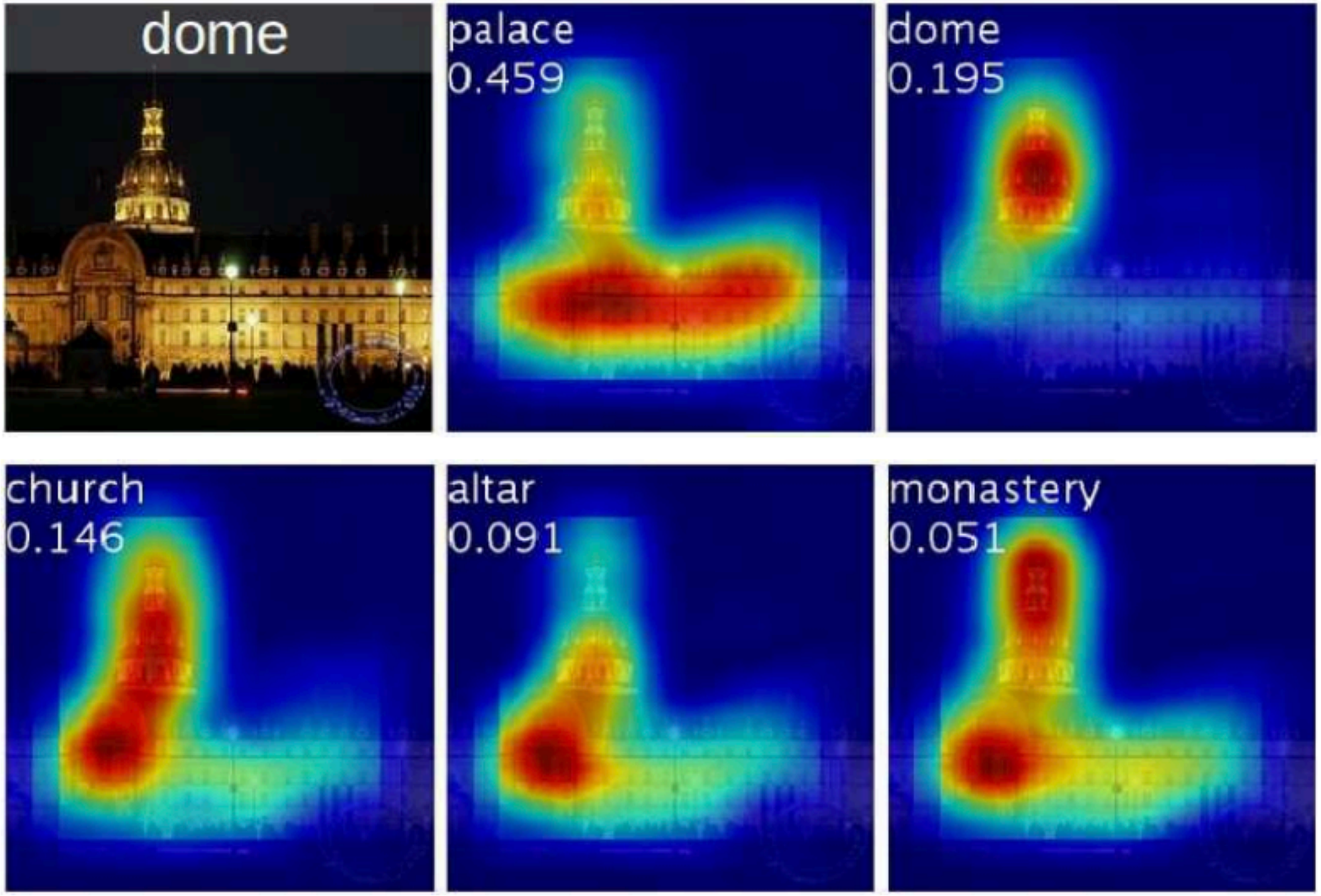


Figure 4. Examples of the CAMs generated from the top 5 predicted categories for the given image with ground-truth as dome. The predicted class and its score are shown above each class activation map. We observe that the highlighted regions vary across predicted classes e.g., *dome* activates the upper round part while *palace* activates the lower flat part of the compound.

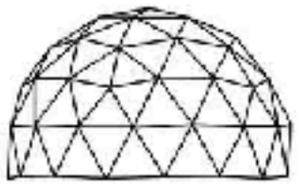
Dome

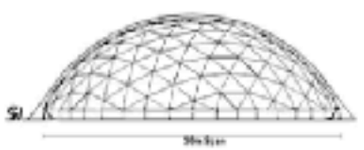






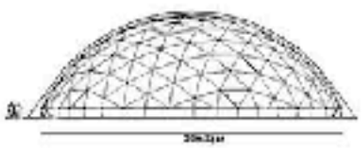
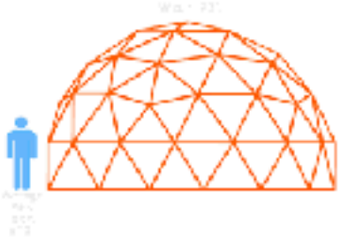


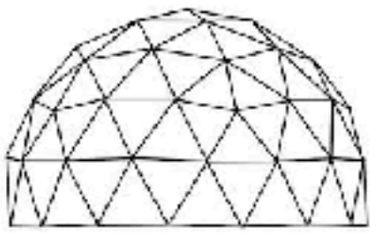


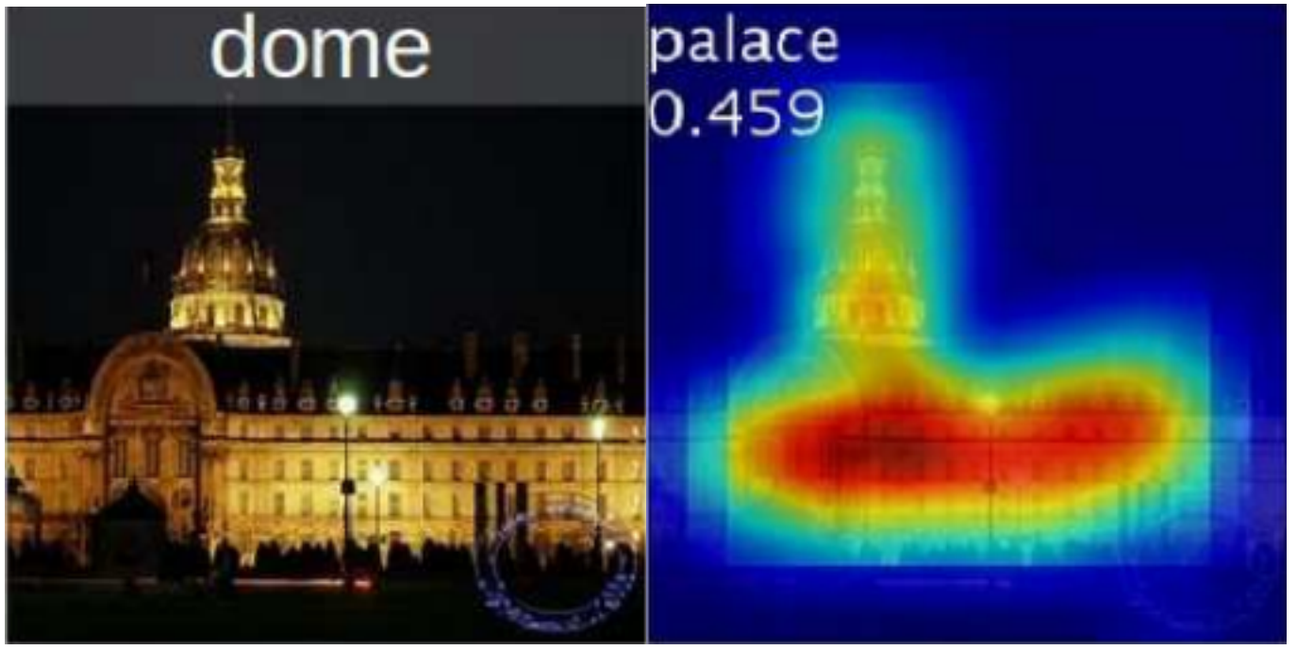




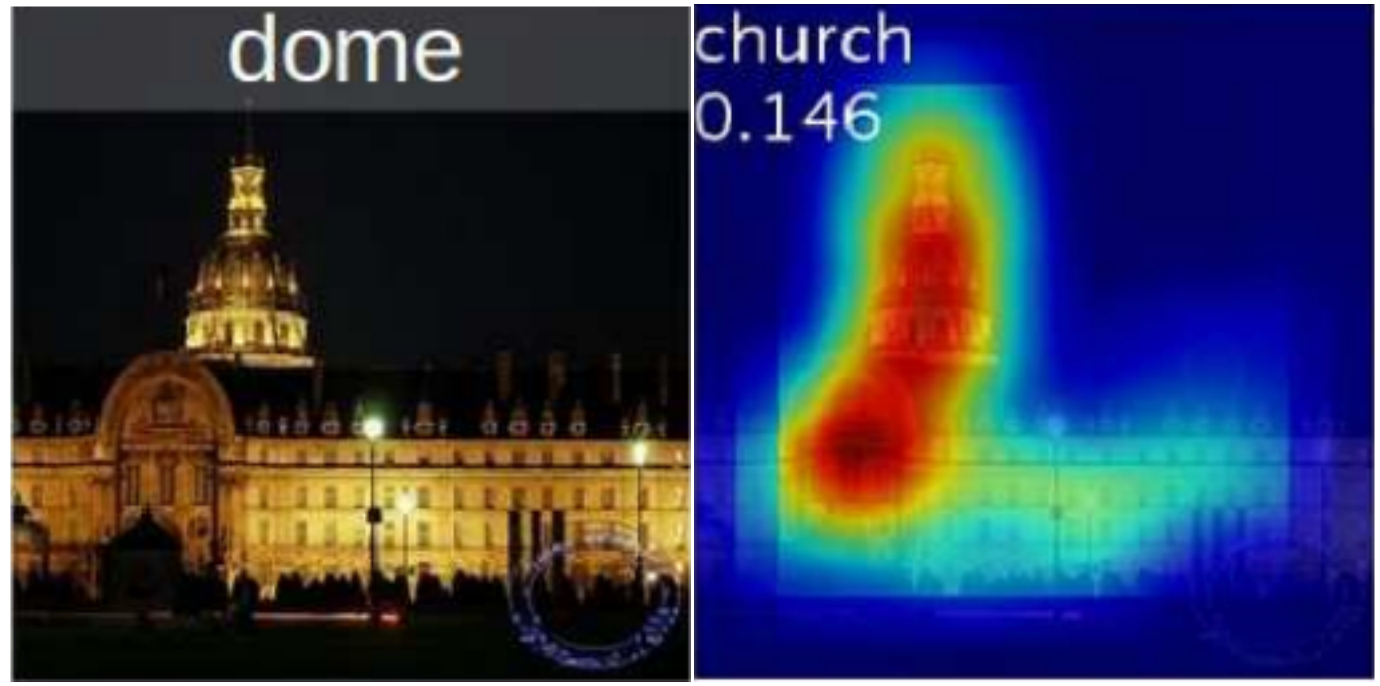






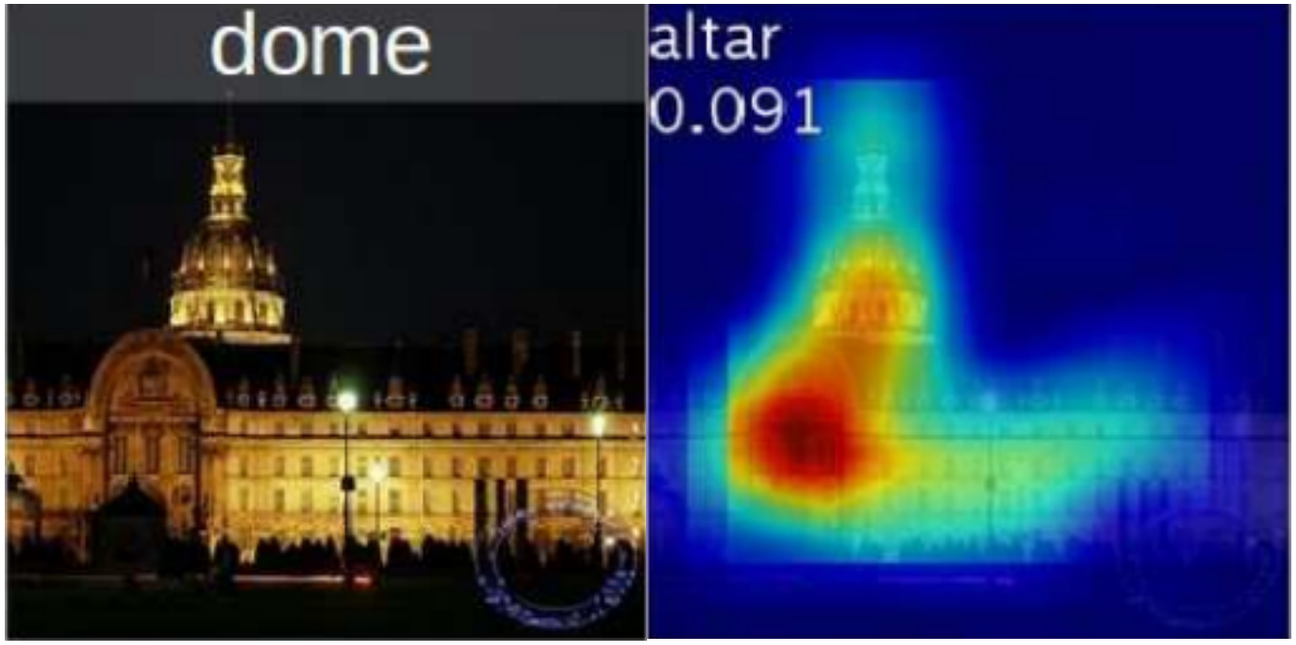

Palace



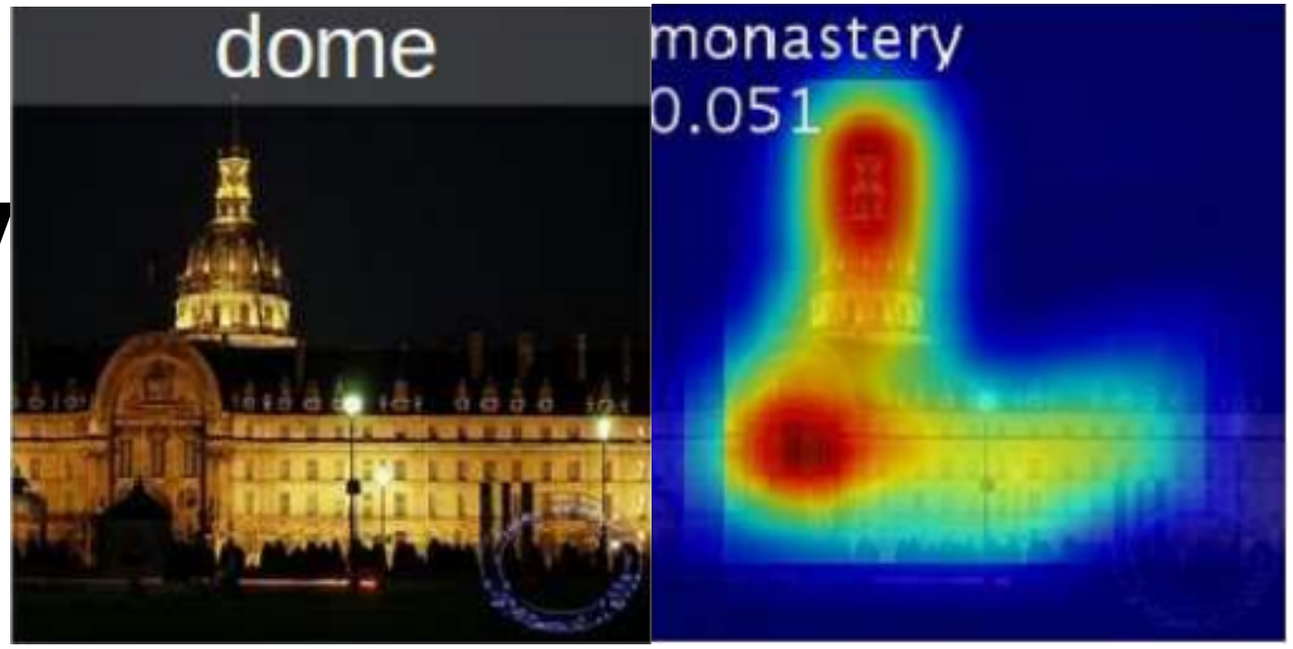
Church



Altar



Monastery



What the CNN is looking and how it shifts the attention in the video.

Here we apply the class activation mapping to a video, to visualize what the CNN is looking and how CNN shifts its attention over time. The word on top-left is the top-1 predicted object label, the heatmap is the class activation map, highlighting the importance of the image region to the prediction.

YouTube Video

Thanks!