



CNN VISUALIZATION

Hey CNN! Why did you say that?



Why need to visualize CNN?

□ 坦克問題



□ 確保 CNN 不要學壞!



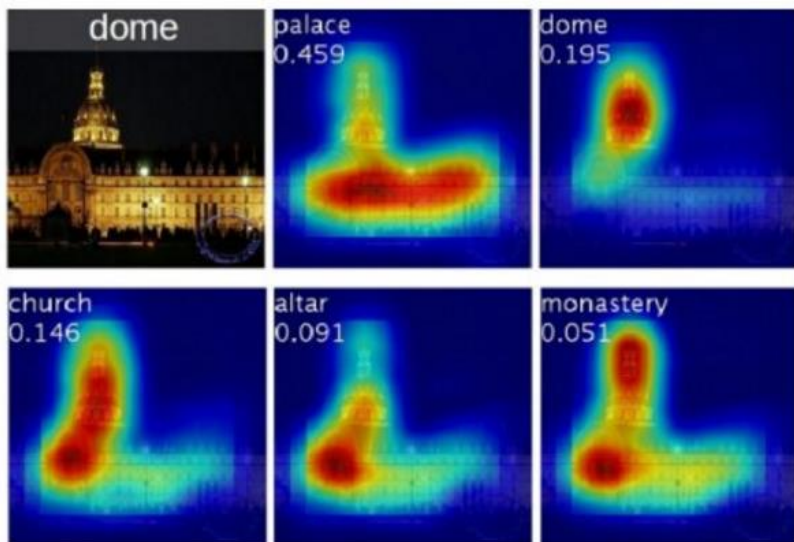
Papers

- ❑ 2015, CVPR, MIT, Learning Deep Features for Discriminative Localization
 - ❑ Class Activation Map (CAM)
- ❑ 2017, CVPR, Virginia Tech, Georgia Institute of Technology, Grad-CAM: Visual Explanations from Deep Networks via Gradient-based Localization
 - ❑ Gradient Class Activation Map (Grad-CAM)
- ❑ <https://github.com/mrgloom/CNN-heatmap>

Learning Deep Features for Discriminative Localization



Learning Deep Features for Discriminative Localization



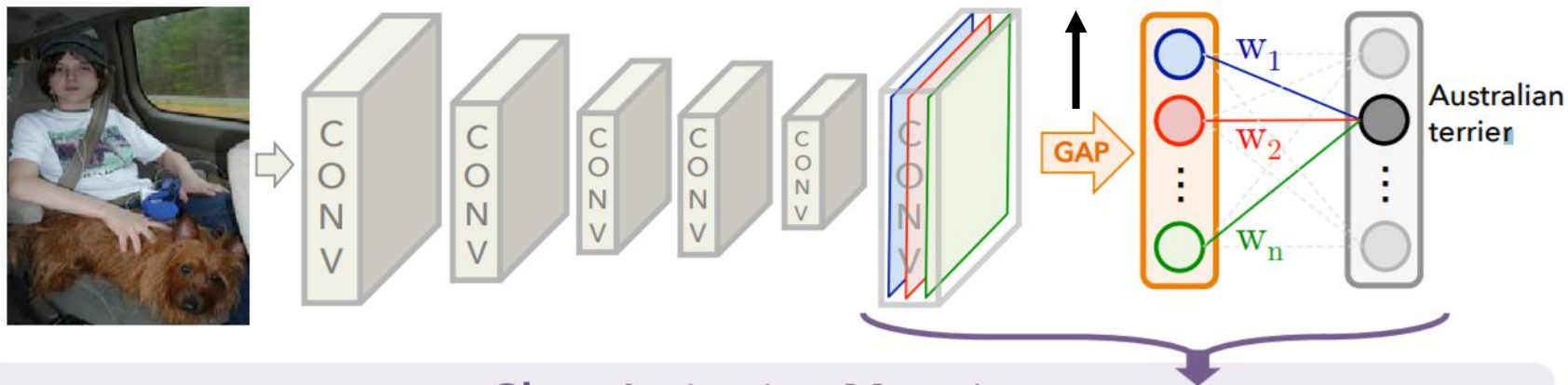
Class activation maps of top 5 predictions



Class activation maps for one object class

CAM structure

$13 \times 13 \times 1024 \rightarrow 1 \times 1 \times 1024$

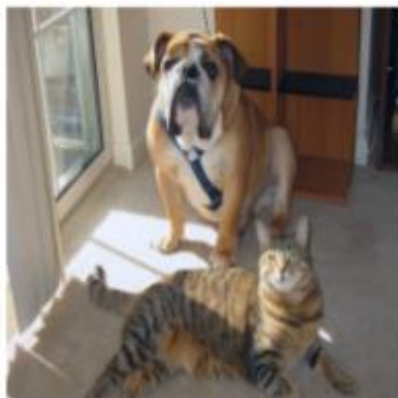


Class Activation Mapping

$$W_1 * \text{Heatmap}_1 + W_2 * \text{Heatmap}_2 + \dots + W_n * \text{Heatmap}_n = \text{Class Activation Map (Australian terrier)}$$

The equation shows the weighted sum of individual activation heatmaps. Each heatmap is a 13×13 grid representing the spatial distribution of activations for a specific class. The weights W_1, W_2, \dots, W_n are the output of the GAP layer. The final result is the Class Activation Map for the "Australian terrier" class, which highlights the regions of the input image that are most responsible for the classification.

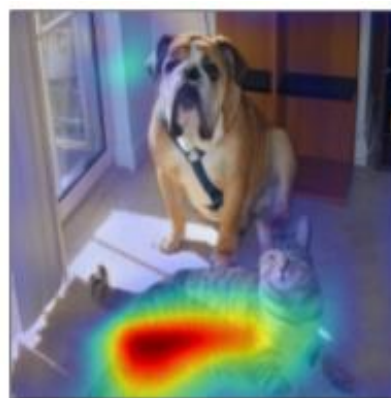
Grad-CAM



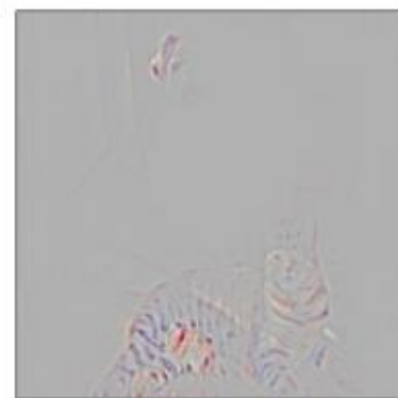
(a) Original Image



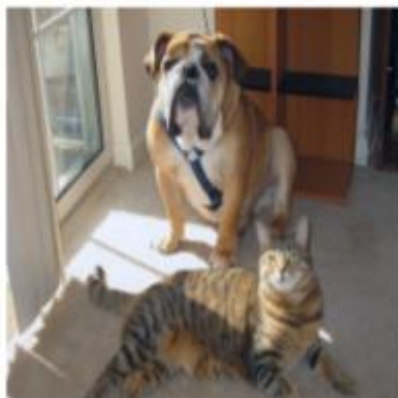
(b) Guided Backprop 'Cat'



(c) Grad-CAM for 'Cat'



(d) Guided Grad-CAM 'Cat'



(g) Original Image



(h) Guided Backprop 'Dog'

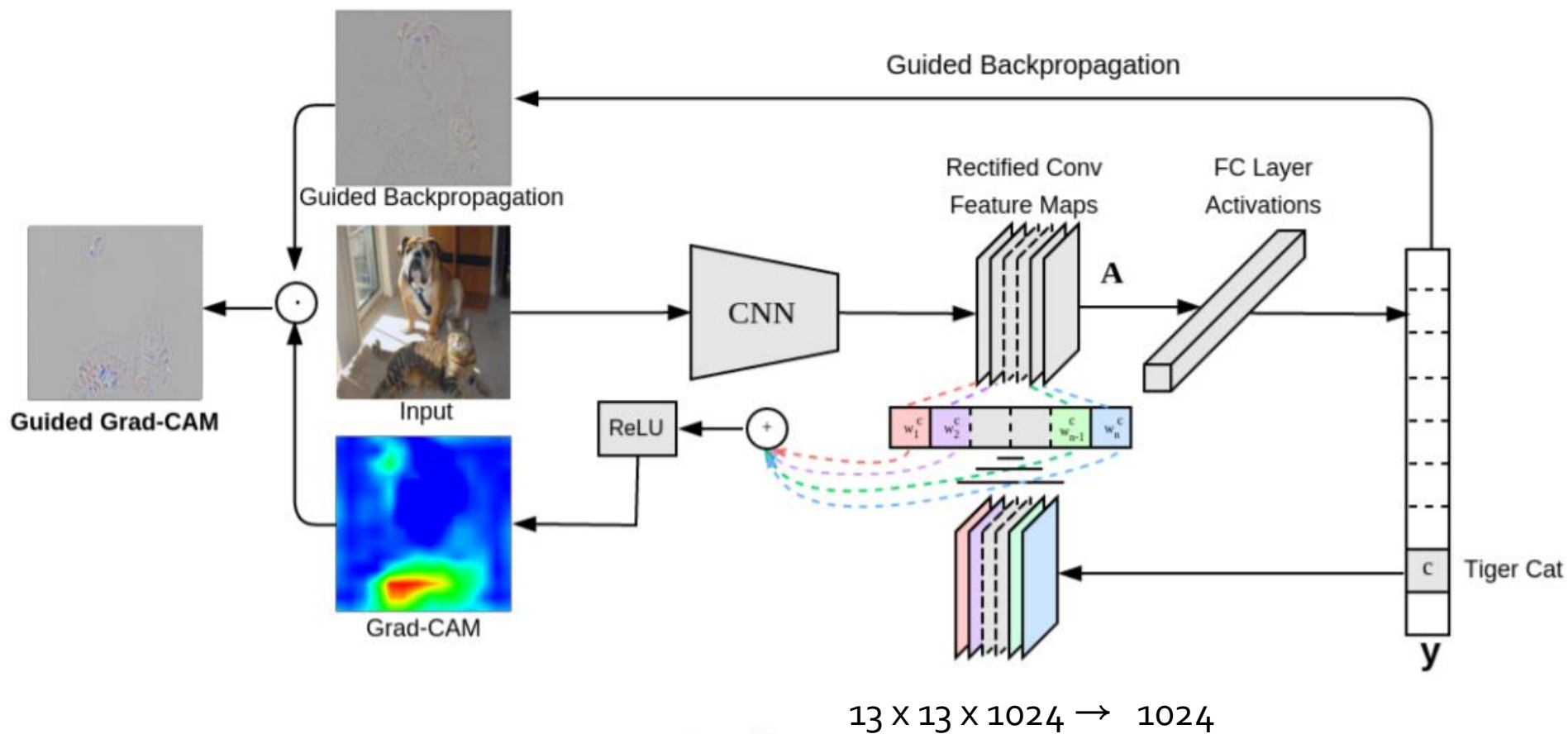


(i) Grad-CAM for 'Dog'



(j) Guided Grad-CAM 'Dog'

Grad-CAM structure





Ground truth: volcano

Ground truth: pineapple

Ground truth: volcano

Ground truth: polaroid camera

Ground truth: beaker

Ground truth: coil



Predicted: sandbar

Predicted: patio

Predicted: car mirror

Predicted: pencil sharpener

Predicted: syringe

Predicted: vine snake

(a)

(b)

(c)

(d)

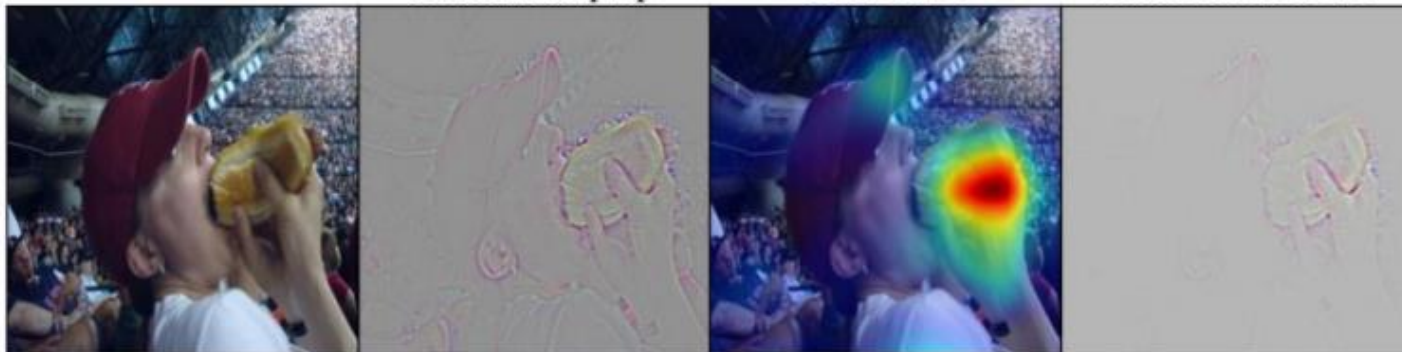
(e)

(f)

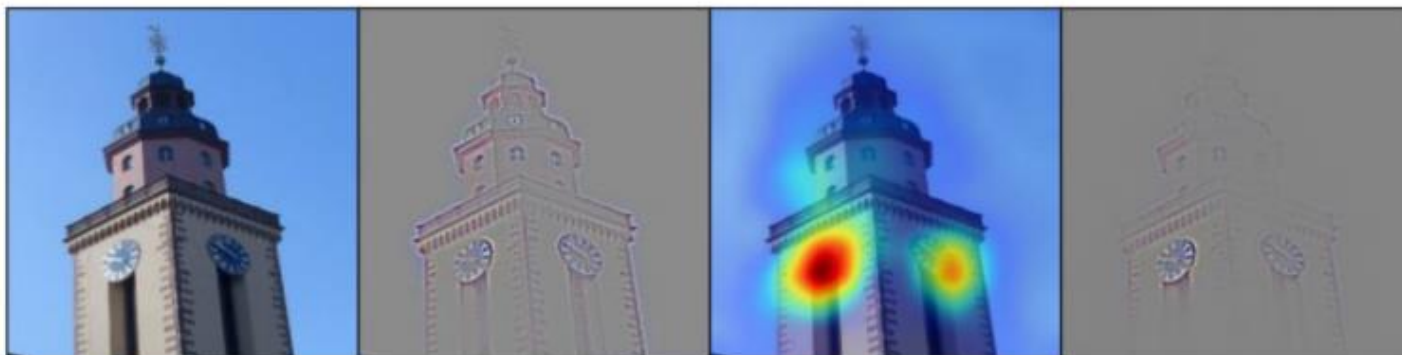
Guided Backprop

Grad-CAM

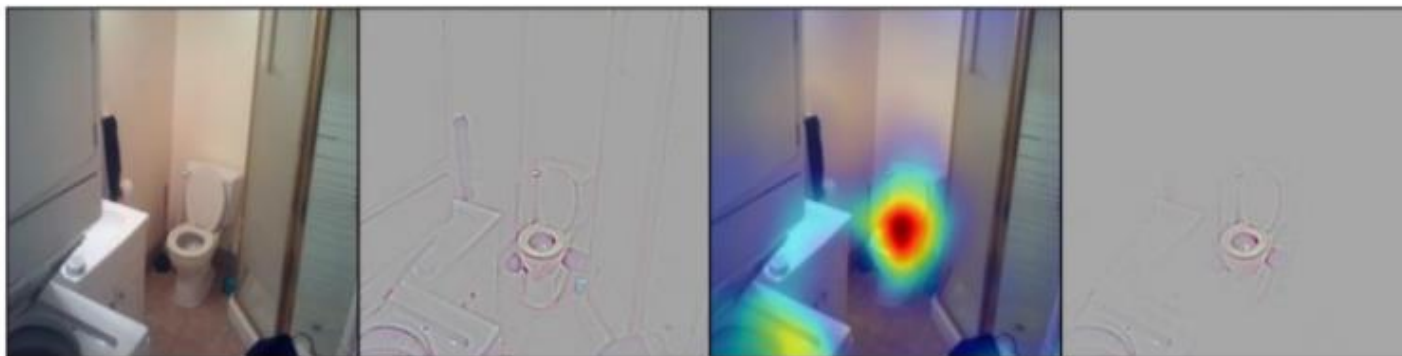
Guided Grad-CAM



A man is holding a hot dog in his hand



A large clock tower with a clock on the top of it



A bathroom with a toilet and a sink