

CAM

(Class Activation Mapping)

※ Zhou, Bolei, et al. "Learning deep features for discriminative localization." Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. 2016.

ISL

안재원

CONTENTS

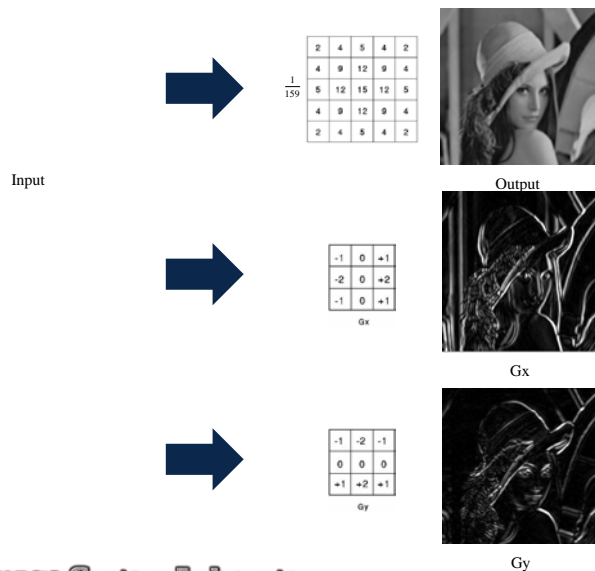
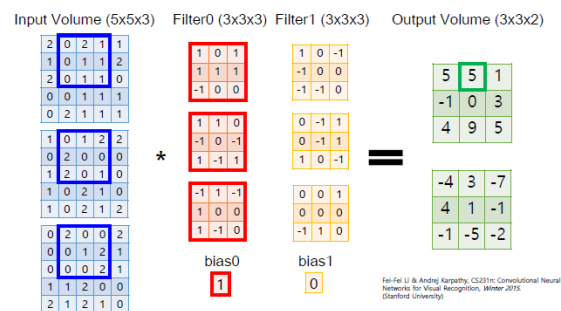
- Intro
- Class Activation Mapping
- Result

Convolution Neural Network

1. Convolution

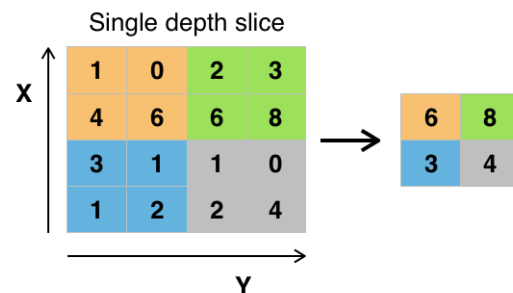
- 신호처리에서 사용됨.
- 마스크 연산
- 신호의 특징을 검출할 때 사용.

$$(f * g)(t) = \int_{-\infty}^{\infty} f(\tau)g(t-\tau)d\tau$$



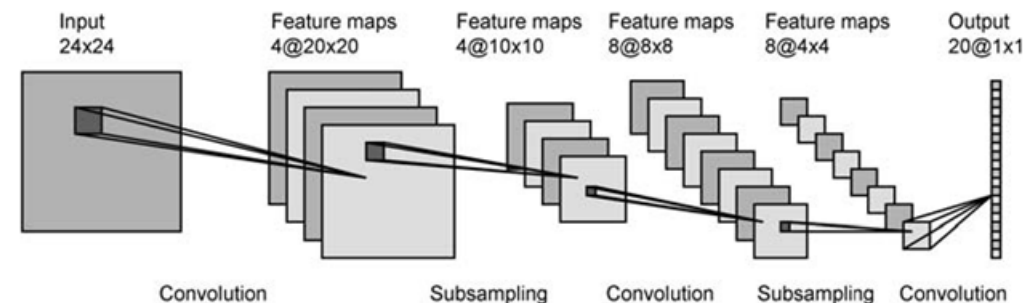
2. Pooling

- Down Sampling
- 보통 Max pooling을 사용한다.



(a) Input (b) Subsampling

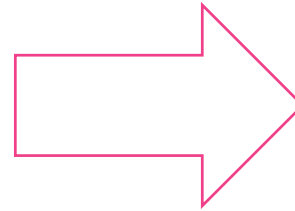
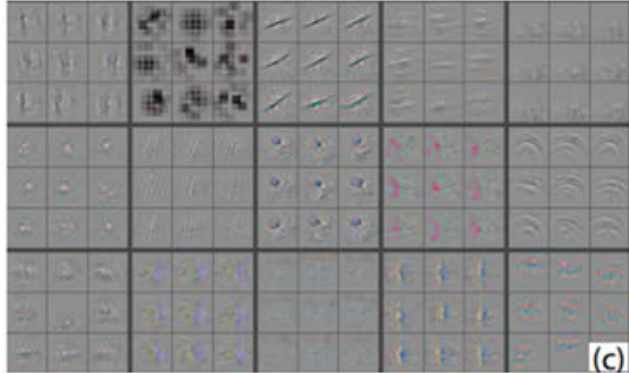
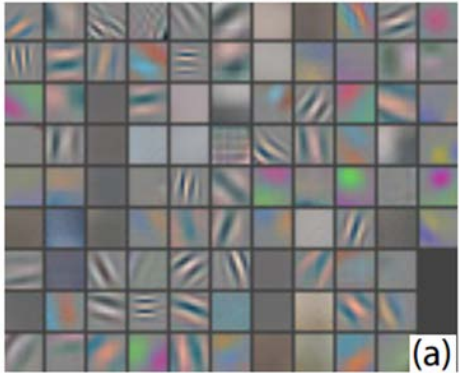
3. CNN



- 학습을 위한 특징 검출 과정.
- 연산량과 변수를 줄이기 위한 과정.
- MLP에 비하여 줄었지만, 하드웨어의 발전과 맞물려 사용 가능한 수준에 도달했다.
- 아직 오랜 연산 시간과 많은 변수가 필요하다.

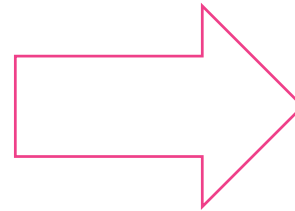
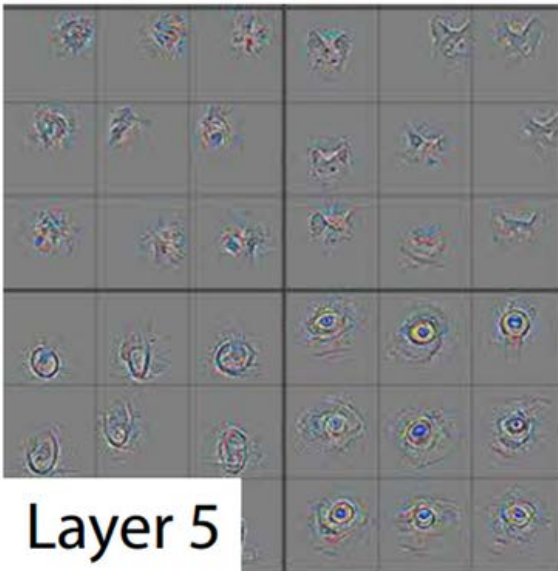
Convolution Neural Network

1. 얕은 layer



Hand made filter : 사람이 생각할 수 있는 영역

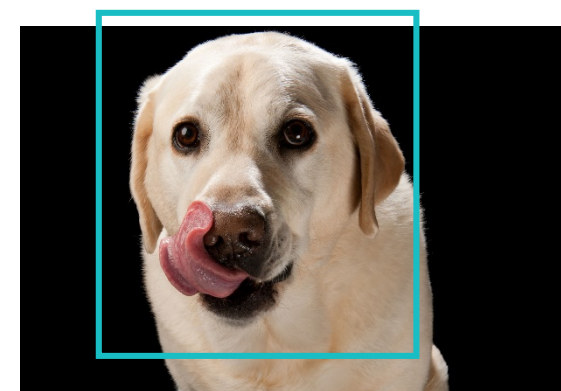
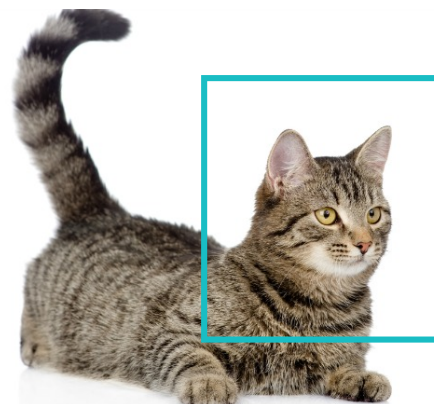
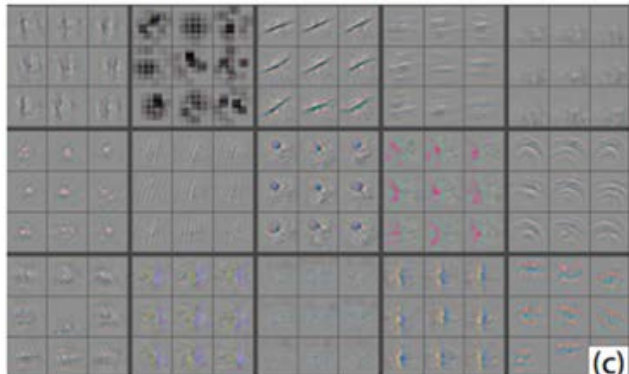
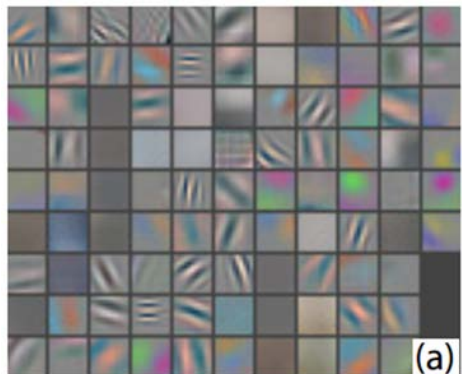
2. 깊은 layer



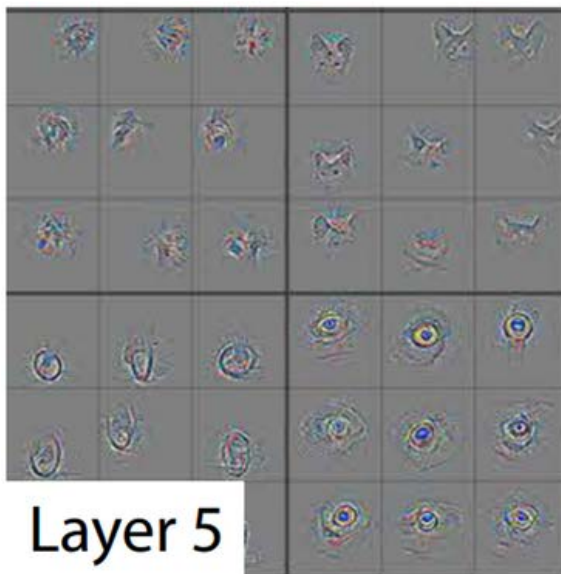
사람이 생각하기 힘든 영역

Convolution Neural Network

1. 얕은 layer



2. 깊은 layer



- Supervise learning
- Bounding box
- Label

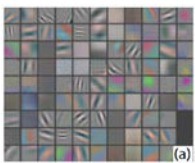


- Weakly-supervised object localization(Label only)
- Global Average pooling



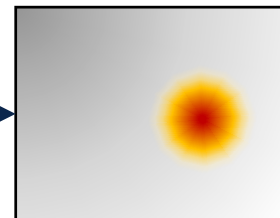
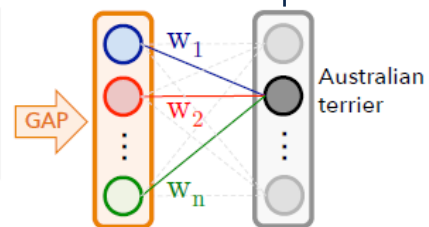
Class Activation Mapping

Feature Extraction

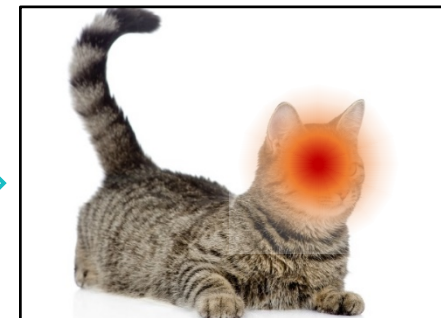


$$\text{Soft-max } S_c = \sum_k w_k^c \sum_{x,y} f_k(x,y)$$

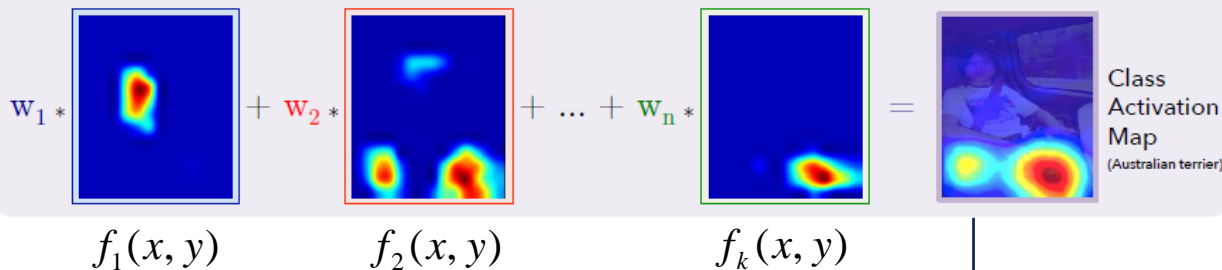
$$= \sum_k \sum_{x,y} w_k^c f_k(x,y)$$



Up-scaling

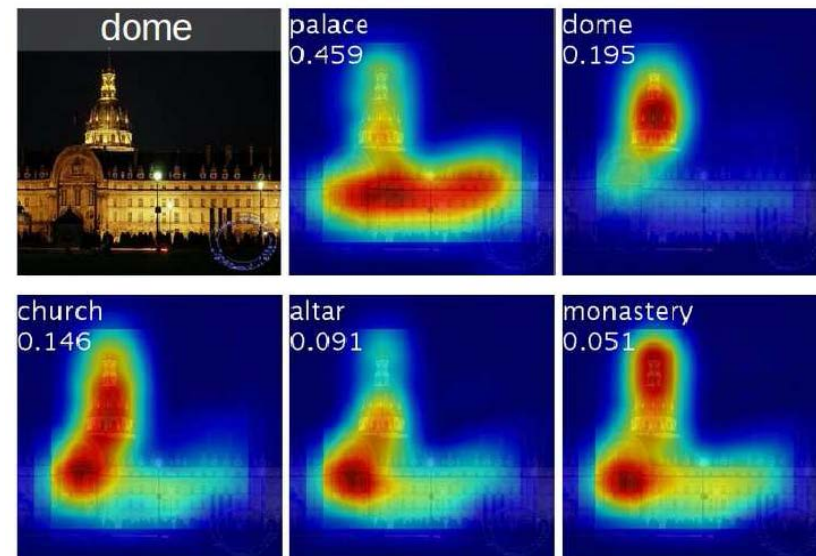


Class Activation Mapping

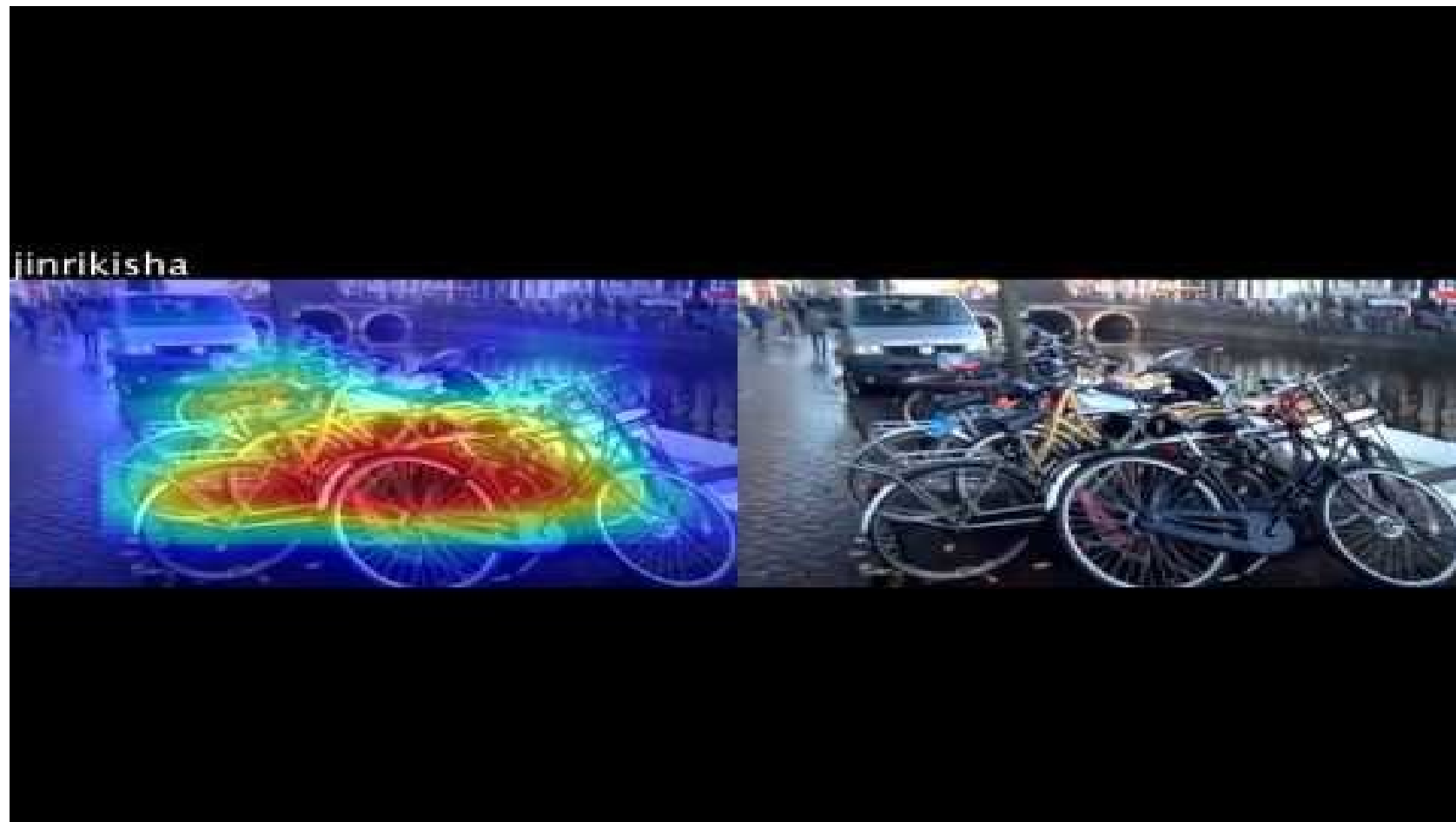
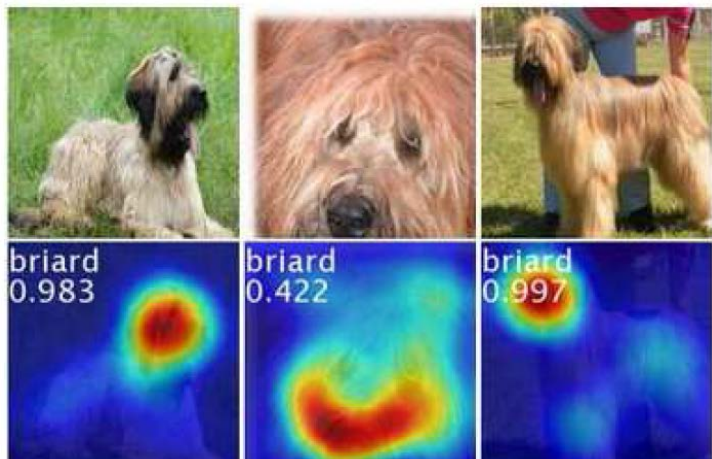


$$M_c(x,y) = \sum_k w_k^c f_k(x,y)$$

- 무엇을 보고 판단했는가?



Result



Result

Cleaning the floor



Cooking



Fixing a car



Polo



Rowing



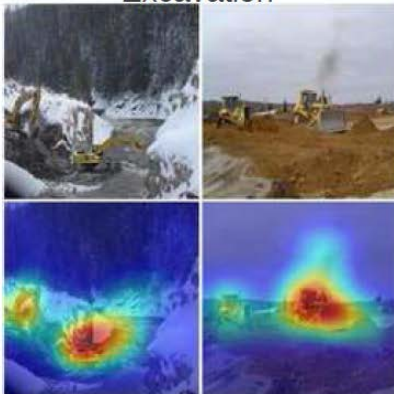
Croquet



Banquet hall



Excavation



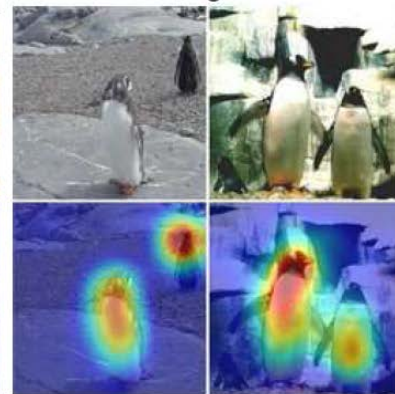
Playground



Mushroom



Penguin



Teapot



Q & A
