
Applications of Deep Learning for Computer Vision

2021.03.05

Data Mining and Quality Analytics Lab

박진혁

- **Introduction**
- **What is Computer Vision?**
- **Applications of deep learning for Computer Vision**
- **Conclusion**

Introduction

❖ 발표자 소개



- 박진혁
- Data Mining & Quality Analytics Lab(김성범 교수님)
- 석.박사 통합과정 4학기 재학 중(2019.8 ~)

✓ 관심 분야

- Computer Vision
- Image Captioning
- Object Detection

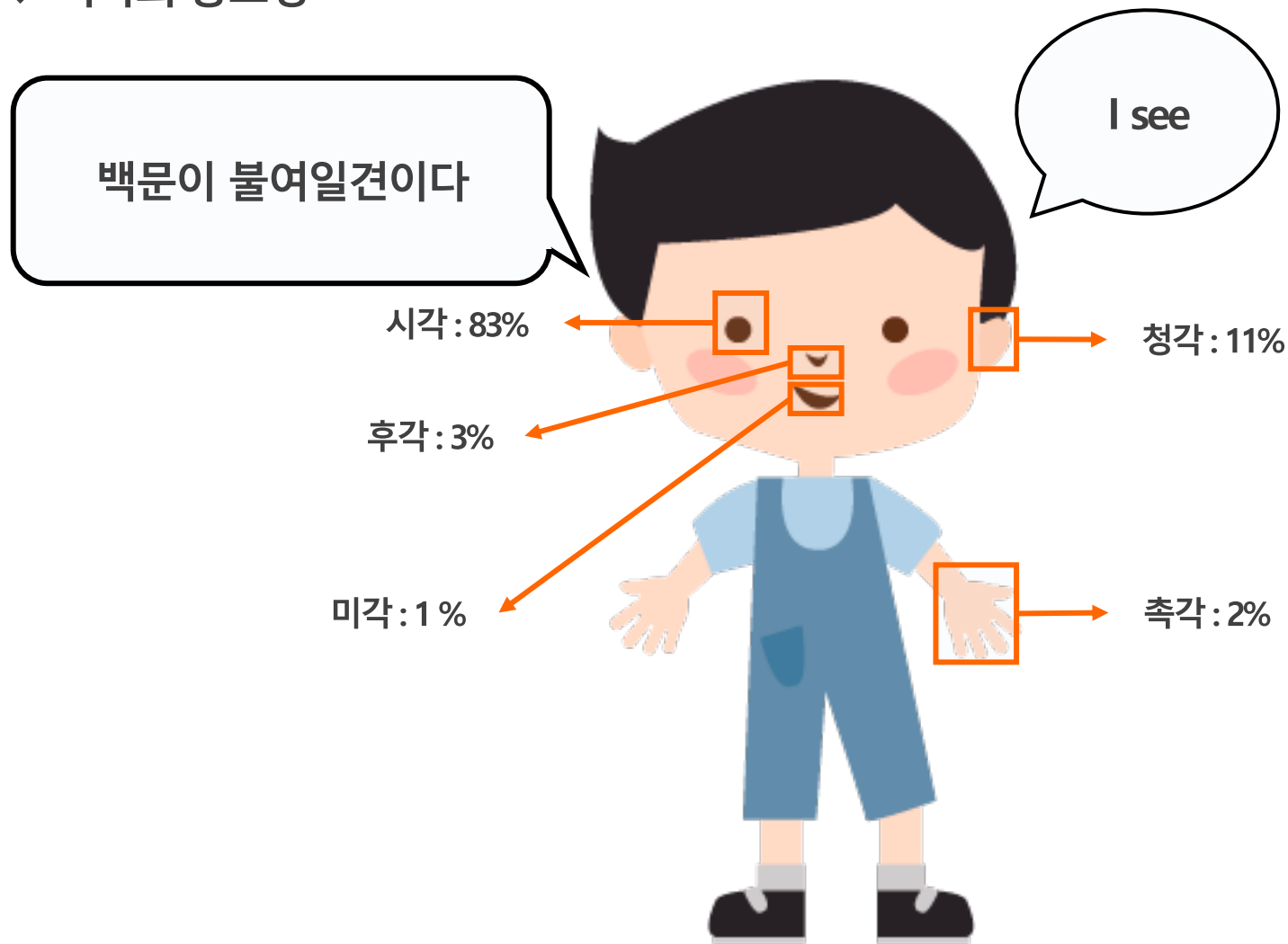
• 박진혁

• Data Mining & Quality Analytics Lab(김성범 교수님)

• 석.박사 통합과정 4학기 재학 중(2019.8 ~)

Introduction

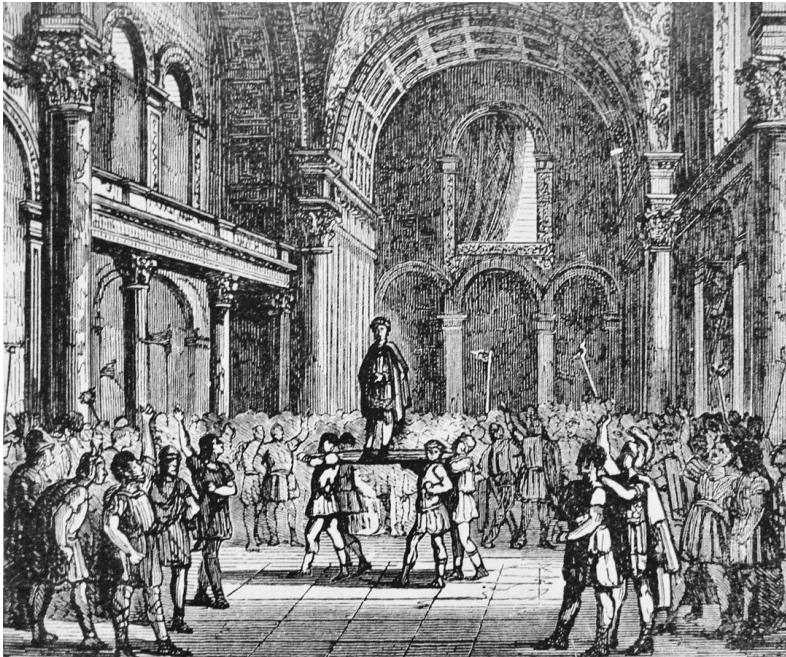
❖ 시각의 중요성



Introduction

❖ 시각의 중요성

- 중세 시대(암송, 경청) → 르네상스 시대 (독서)



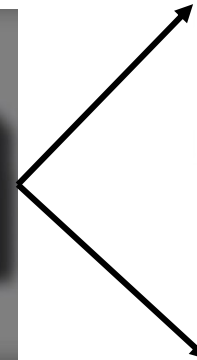
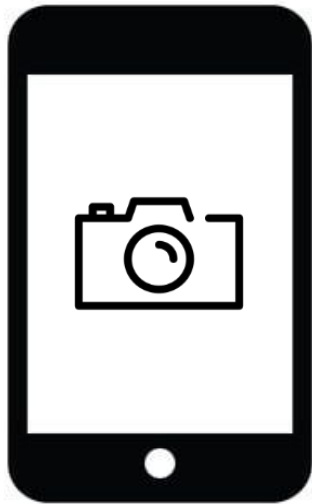
중세 시대



르네상스 시대

Introduction

❖ 시각의 중요성



What is Computer Vision

❖ Computer Vision(CV)

- 사물을 인지하여 의미 있는 정보를 수학적 알고리즘을 통해 추출하는 학문
- 지도학습의 한 분야로 사진/동영상에 대한 정답(Label)이 존재

Object Detection



Thor, Captain America, Car

Image Segmentation



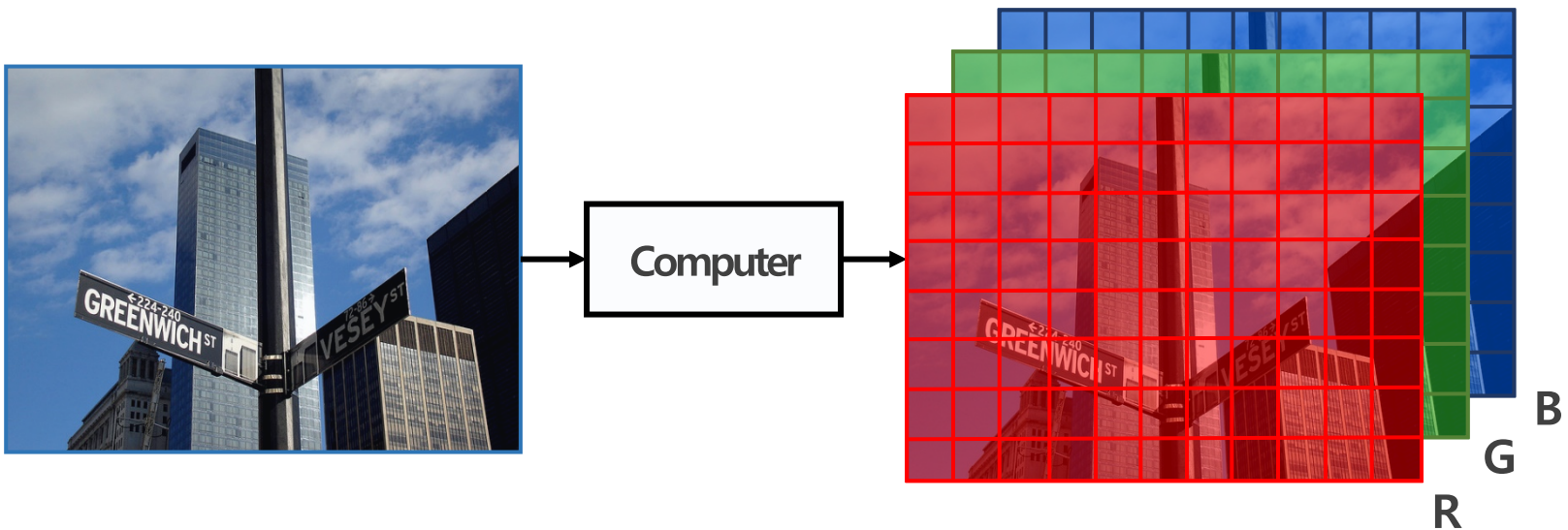
Thor, Captain America, Car



What is Computer Vision

❖ Computer Vision(CV)

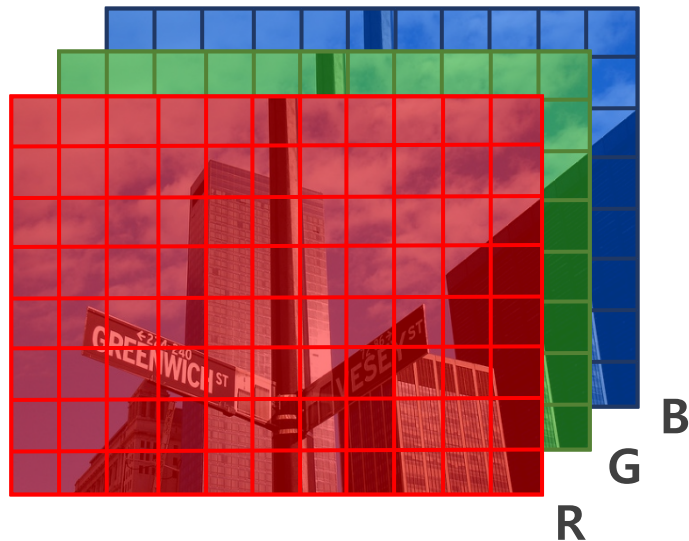
- 픽셀 : 이미지를 구성하는 가장 기본적인 단위
- RGB channel : 각 픽셀을 red, green, blue에 해당하는 3개의 실수로 표현한 channel



What is Computer Vision

❖ Computer Vision(CV)

- 픽셀 : 이미지를 구성하는 가장 기본적인 단위
- RGB channel : 각 픽셀을 red, green, blue에 해당하는 3개의 실수로 표현한 channel



2	3	3	8
6	7	2	8
5	4	3	3
3	2	7	6

<Feature Map>

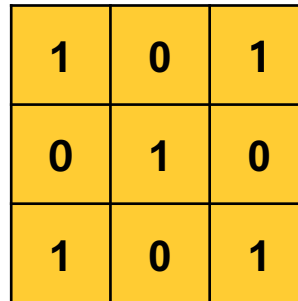
What is Computer Vision

❖ Convolutional Neural Networks (CNN)

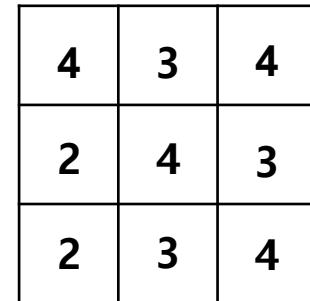
- Convolutional Neural Networks : 이미지의 특징을 추출하는 알고리즘
- Filter : 이미지 특징을 찾아내기 위한 파라미터



Image



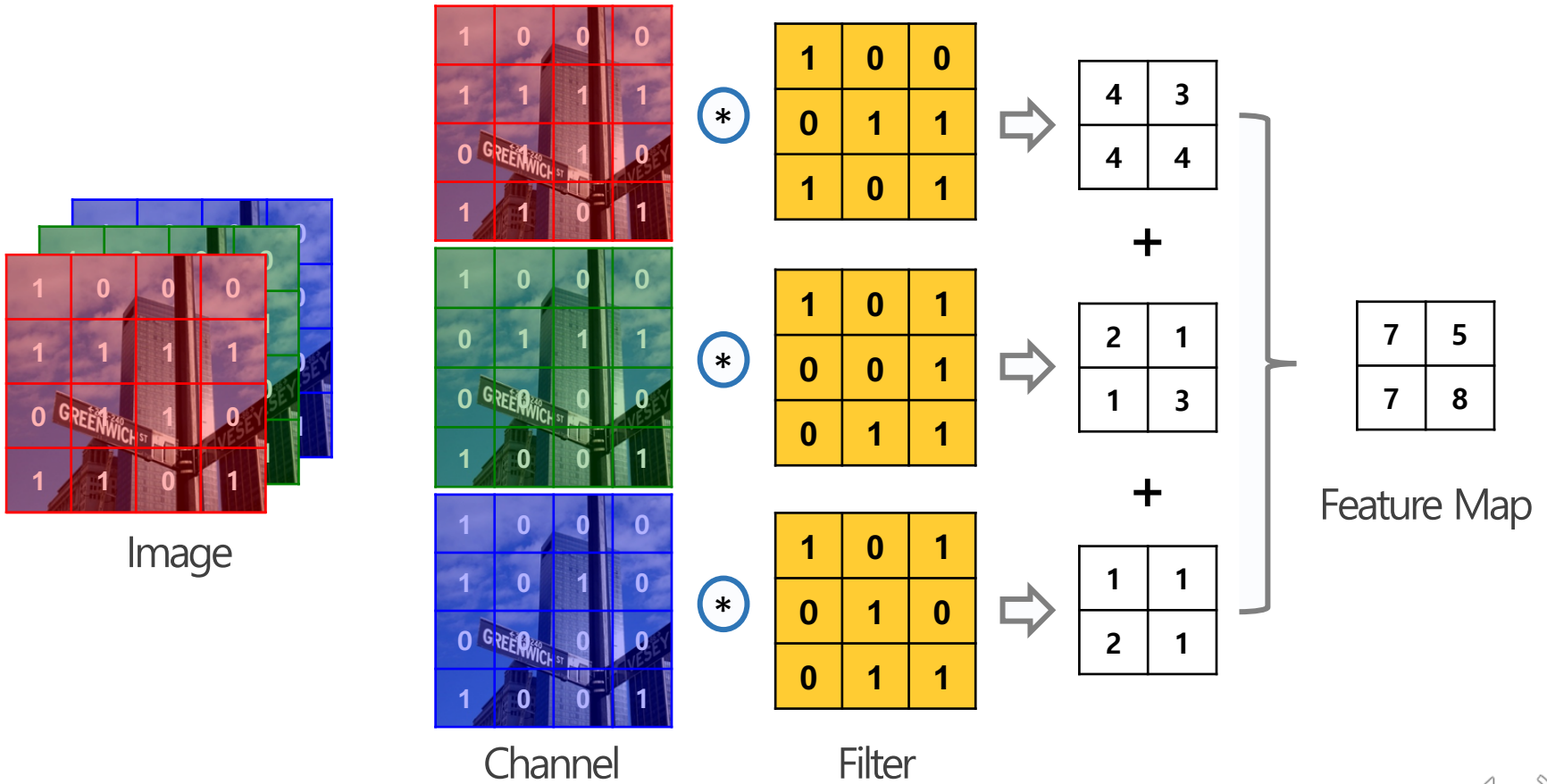
Filter



Feature Map

What is Computer Vision

❖ CNN



Applications of Deep Learning for Computer Vision

❖ Image Classification and Localization

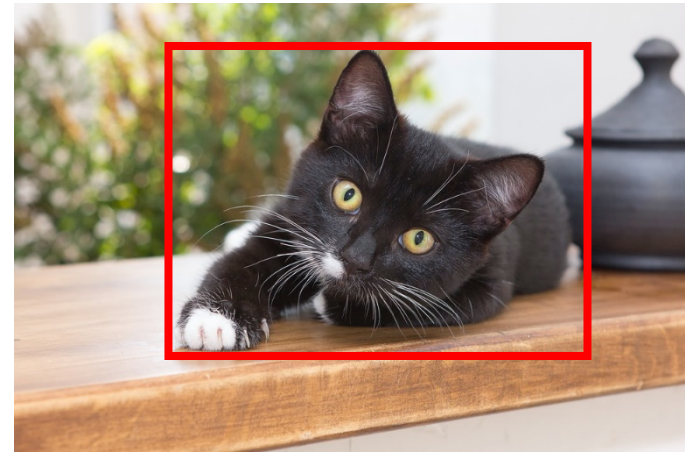
- Image Classification : 이미지가 사전에 정의한 범주 중 어떤 범주에 속하는지 분류하는 문제
- Image Localization : 이미지 내 사전에 정의한 범주 중 어떤 범주에 속하며 어디에 존재하는 예측

<Classification>



CAT

<Localization>



CAT

[1] Krizhevsky, A., Sutskever, I. & Hinton, G. E. (2012). Imagenet classification with deep convolutional neural networks. *Advances in neural information processing systems*, 25, 1097-1105.

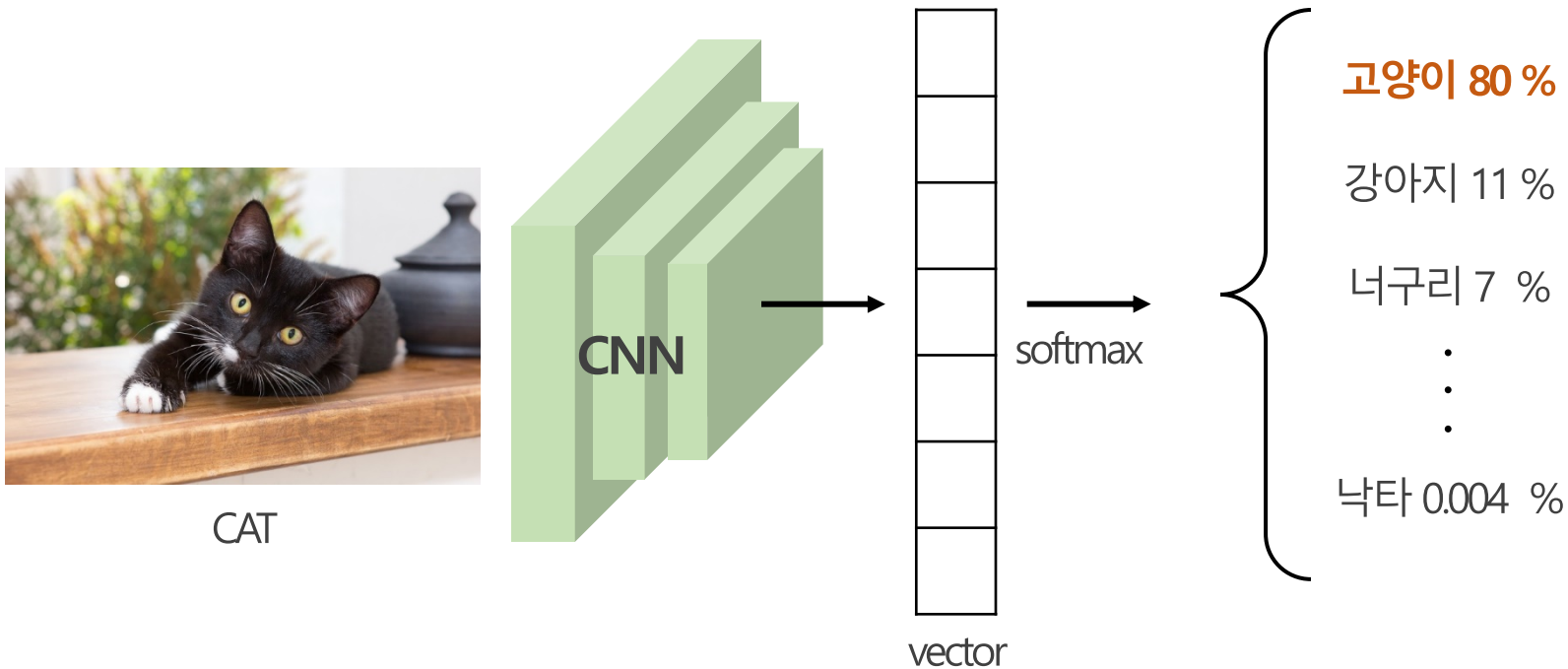
[2] He, K., Zhang, X., Ren, S. & Sun, J. (2016). Deep residual learning for image recognition. In *Proceedings of the IEEE conference on computer vision and pattern recognition* (pp. 770-778).



Applications of Deep Learning for Computer Vision

❖ Image Classification and Localization

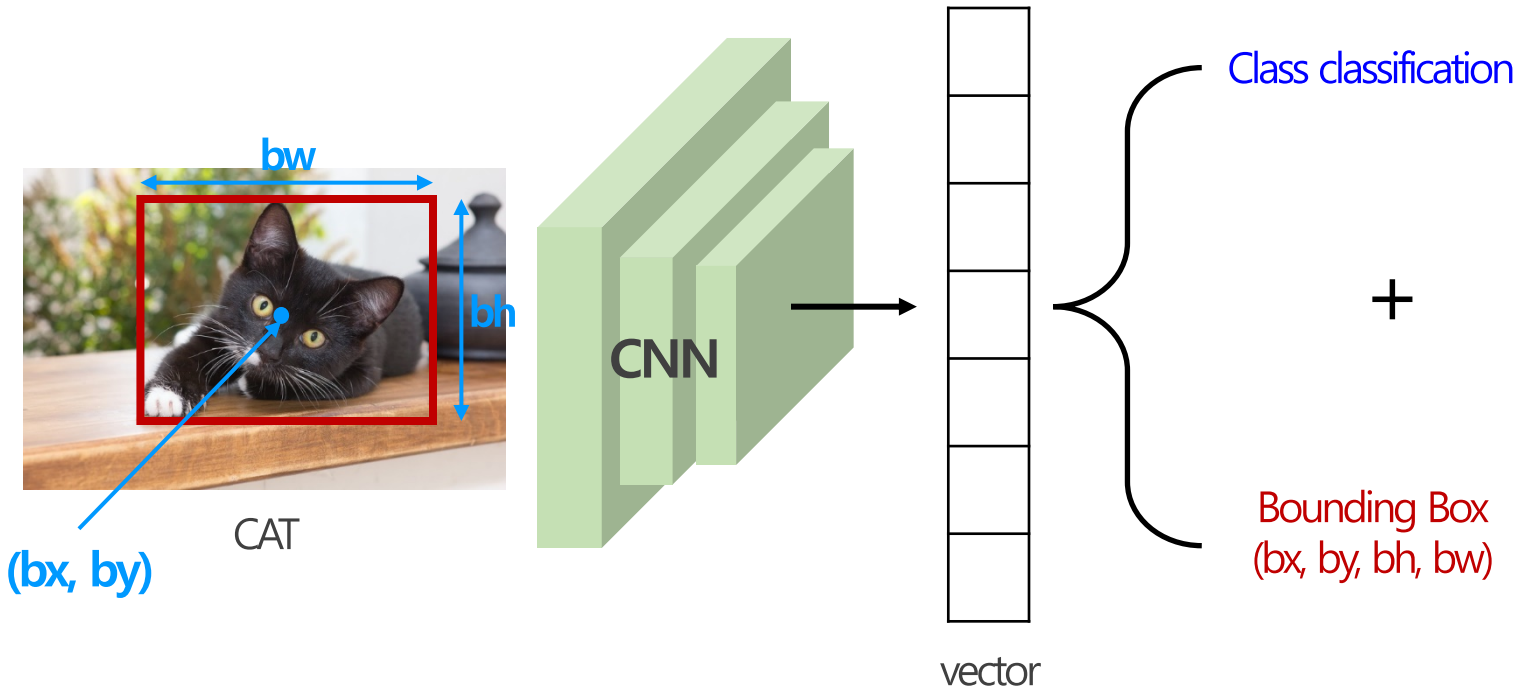
- **Image Classification** : 이미지가 사전에 정의한 범주 중 어떤 범주에 속하는지 분류하는 문제
- Image Localization : 이미지 내 사전에 정의한 범주 중 어떤 범주에 속하며 어디에 존재하는 예측



Applications of Deep Learning for Computer Vision

❖ Image Classification and Localization

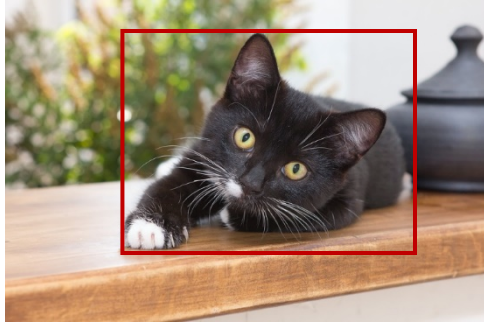
- Image Classification : 이미지가 사전에 정의한 범주 중 어떤 범주에 속하는지 분류하는 문제
- **Image Localization** : 이미지 내 사전에 정의한 범주 중 어떤 범주에 속하며 어디에 존재하는 예측



Applications of Deep Learning for Computer Vision

❖ Image Classification and Localization

- Image Classification : 이미지가 사전에 정의한 범주 중 어떤 범주에 속하는지 분류하는 문제
- **Image Localization** : 이미지 내 사전에 정의한 범주 중 어떤 범주에 속하며 어디에 존재하는 예측



CAT

$$\rightarrow \mathbf{y} = \begin{bmatrix} P_c \\ b_x \\ b_y \\ b_h \\ b_w \\ c_1 \\ c_2 \\ c_3 \end{bmatrix}$$

P_c : $P_c = 1$ (Object 존재), $P_c = 0$ (Object가 background)

b_x, b_y, b_h, b_w : Bounding Box의 좌표

c_1, c_2, c_3 : Class1, Class2, Class3



Applications of Deep Learning for Computer Vision

❖ Object Detection

- Object Detection : 여러가지 물체에 대한 Classification + 물체의 위치정보를 파악하는 Localization

Object Detection = Multi-Labeled **Classification**
+ Bounding Box Regression(**Localization**)



CAR



DOG



Applications of Deep Learning for Computer Vision

❖ Object Detection

- Sliding window : 이미지에서 모든 영역을 다양한 크기의 window로 sliding하며 탐색하는 방법

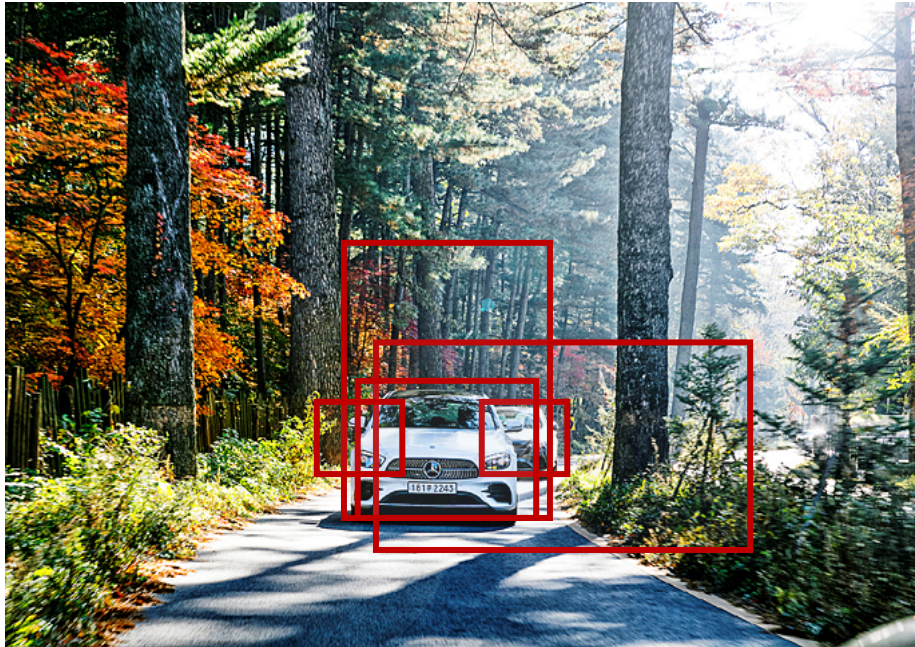
모든 영역을 보기 때문에 연산량이 많아짐



Applications of Deep Learning for Computer Vision

❖ Object Detection

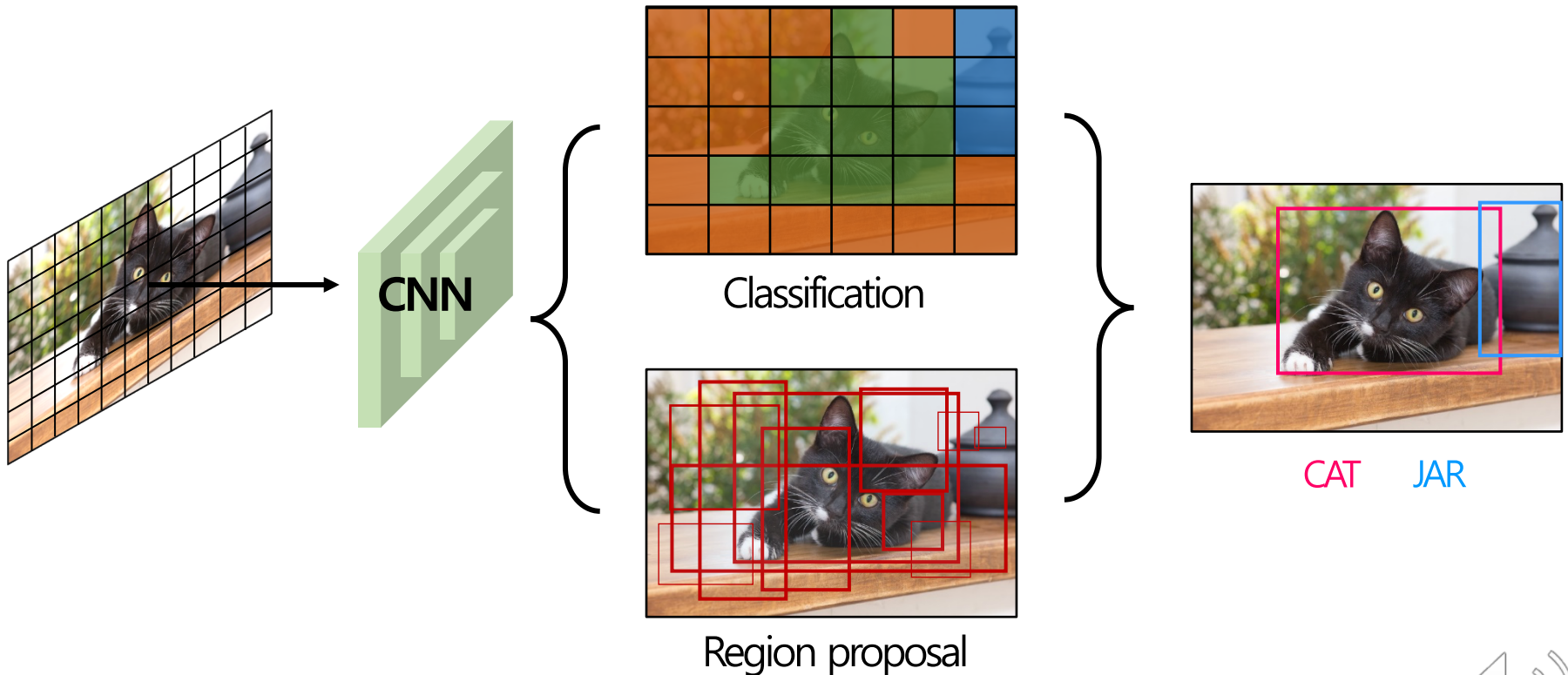
- Region proposal : 이미지 안에서 '물체가 있을 법한 영역'을 찾아내는 방식
 - **Selective Search, Edge boxes, ...**



Applications of Deep Learning for Computer Vision

❖ Object Detection

- 1-Stage Detector : Regional proposal과 Classification이 동시에 이루어짐(속도 ↑)
- 2-Stage Detector : Regional proposal과 Classification이 순차적으로 이루어짐 (정확도 ↑)

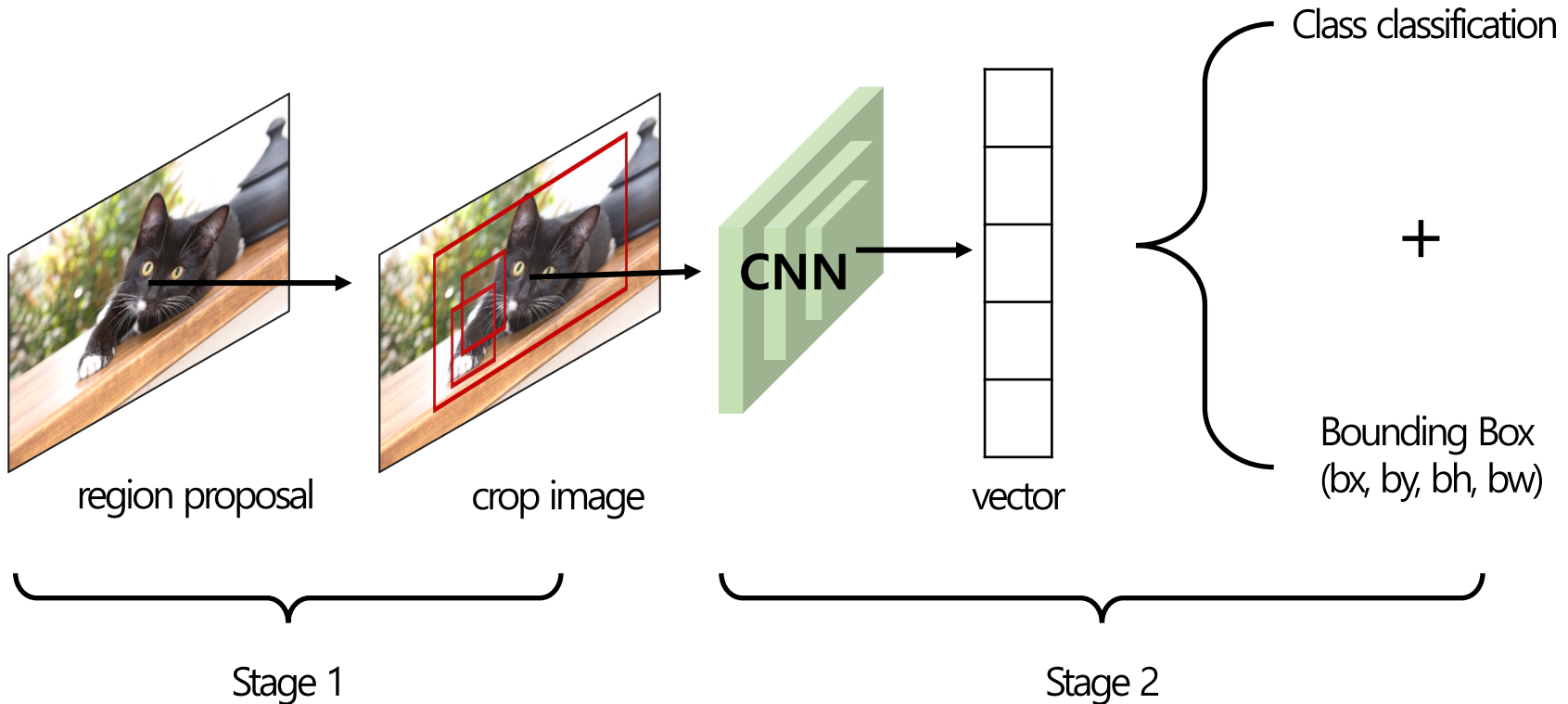


[1] Redmon, J., Divvala, S., Girshick, R., & Farhadi, A. (2016). You only look once: Unified, real-time object detection. In Proceedings of the IEEE conference on computer vision and pattern recognition (pp. 779-788).

Applications of Deep Learning for Computer Vision

❖ Object Detection

- 1-Stage Detector : Regional proposal과 Classification이 동시에 이루어짐(속도 ↑)
- **2-Stage Detector : Regional proposal과 Classification이 순차적으로 이루어짐 (정확도 ↑)**



[1] Ren, S., He, K., Girshick, R., & Sun, J. (2015). Faster r-cnn: Towards real-time object detection with region proposal networks. arXiv preprint arXiv:1506.01497.

Applications of Deep Learning for Computer Vision

❖ Object Segmentation

- Object Segmentation : 이미지 픽셀이 사전에 정의한 범주 중 어느 범주에 속하는지 예측
 - **Semantic segmentation**과 **Instance segmentation**으로 구성

Input Image



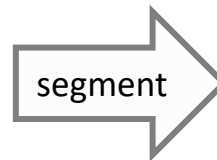
Applications of Deep Learning for Computer Vision

❖ Object Segmentation

- Object Segmentation : 이미지 픽셀이 사전에 정의한 범주 중 어느 범주에 속하는지 예측
 - **Semantic segmentation**과 **Instance segmentation**으로 구성



Image



3	3	3	3	3	3	3	3	3	3	3	3	3	5	5	5	5	5	5
3	3	3	3	3	3	3	3	3	3	3	3	3	5	5	5	5	5	5
3	3	3	3	3	3	1	1	3	3	3	3	3	5	5	5	5	5	5
3	3	3	3	3	1	1	1	1	3	3	3	5	5	5	5	5	5	5
3	3	3	3	3	3	1	1	3	3	3	5	5	5	5	5	5	5	5
5	5	3	3	3	3	1	1	3	3	5	5	5	5	5	5	5	5	5
4	4	3	4	1	1	1	1	1	1	4	4	4	5	5	5	5	5	5
4	4	3	4	1	1	1	1	1	1	4	4	4	4	4	5	5	5	5
4	4	4	1	1	1	1	1	1	1	4	4	4	4	4	4	4	4	4
3	3	3	1	1	1	1	1	1	1	4	4	4	4	4	4	4	4	4
3	3	3	1	2	2	1	1	1	1	4	4	4	4	4	4	4	4	4
3	3	3	1	2	2	1	1	1	1	4	4	4	4	4	4	4	4	4

semantic labels

- 1: Person
- 2: Purse
- 3: Plants/Grass
- 4: Sidewalk
- 5: Building/Structures

[1] He, K., Gkioxari, G., Dollár, P., & Girshick, R. (2017). Mask r-cnn. In Proceedings of the IEEE international conference on computer vision (pp. 2961-2969).

Applications of Deep Learning for Computer Vision

❖ Object Segmentation

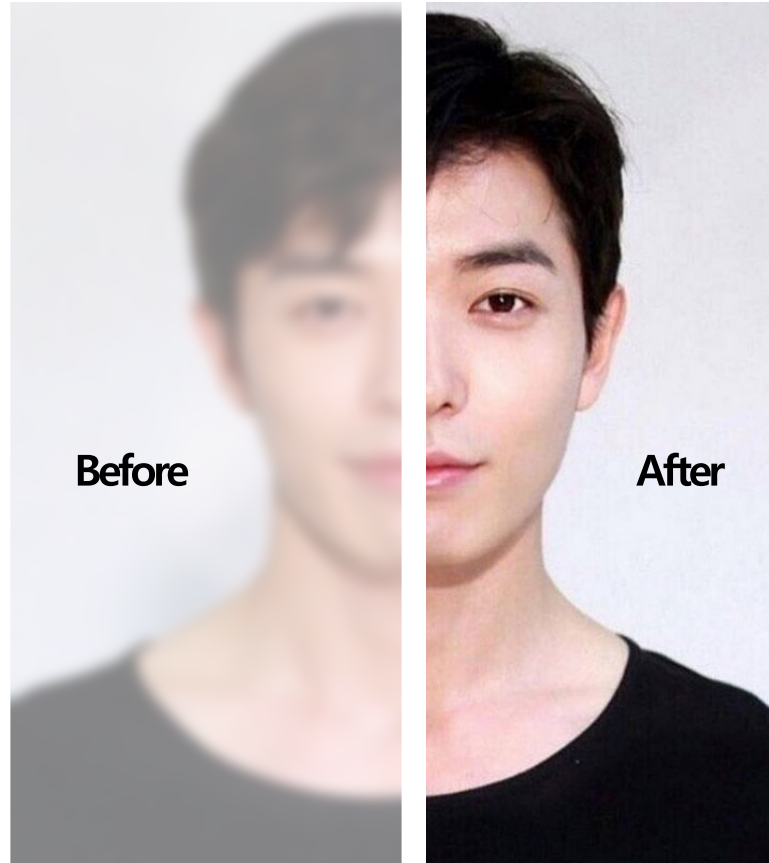
- Object Segmentation : 이미지 픽셀이 사전에 정의한 범주 중 어느 범주에 속하는지 예측
 - Semantic segmentation과 Instance segmentation으로 구성



Applications of Deep Learning for Computer Vision

❖ Image Super Resolution

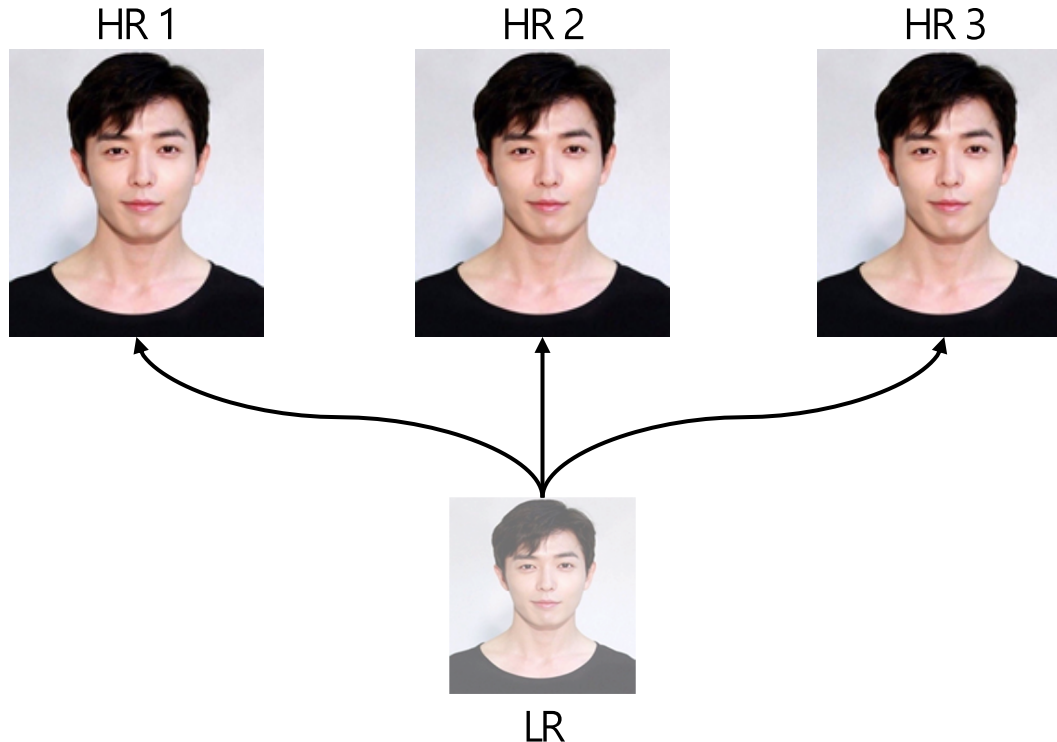
- Image Super Resolution : 저해상도(LR) 이미지를 고해상도(HR) 이미지로 변환하는 연구 분야
 - **Single Image Super Resolution**, Multi Image Super Resolution



Applications of Deep Learning for Computer Vision

❖ Image Super Resolution

- Image Super Resolution : 저해상도(LR) 이미지를 고해상도(HR) 이미지로 변환하는 연구 분야
 - **Single Image Super Resolution**, Multi Image Super Resolution

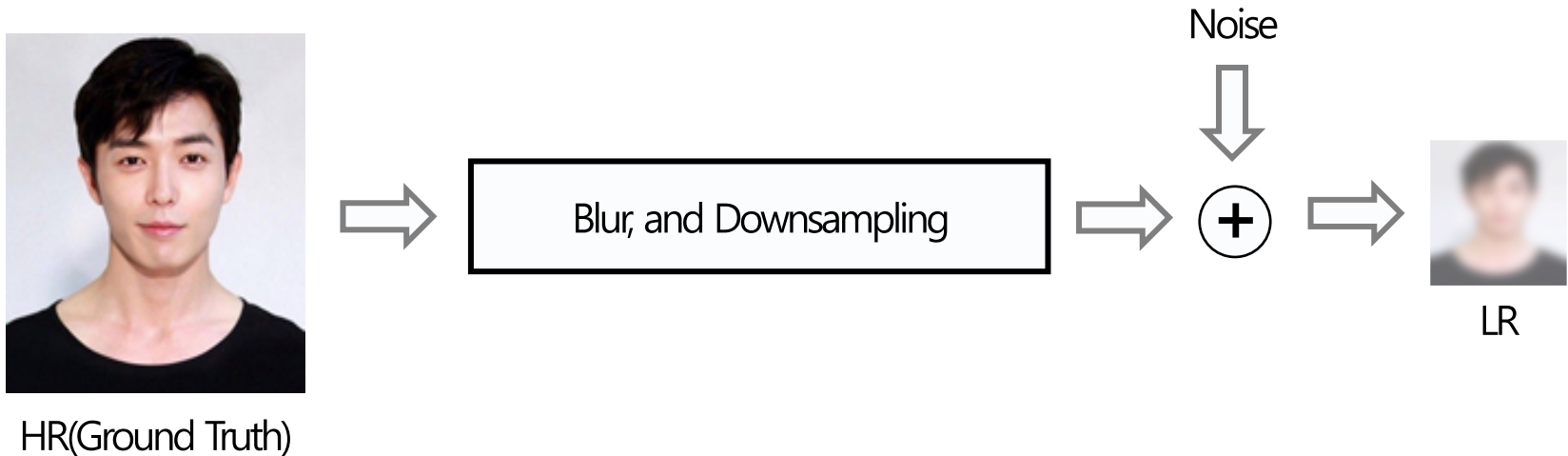


Ill-Posed Prob...

Applications of Deep Learning for Computer Vision

❖ Image Super Resolution

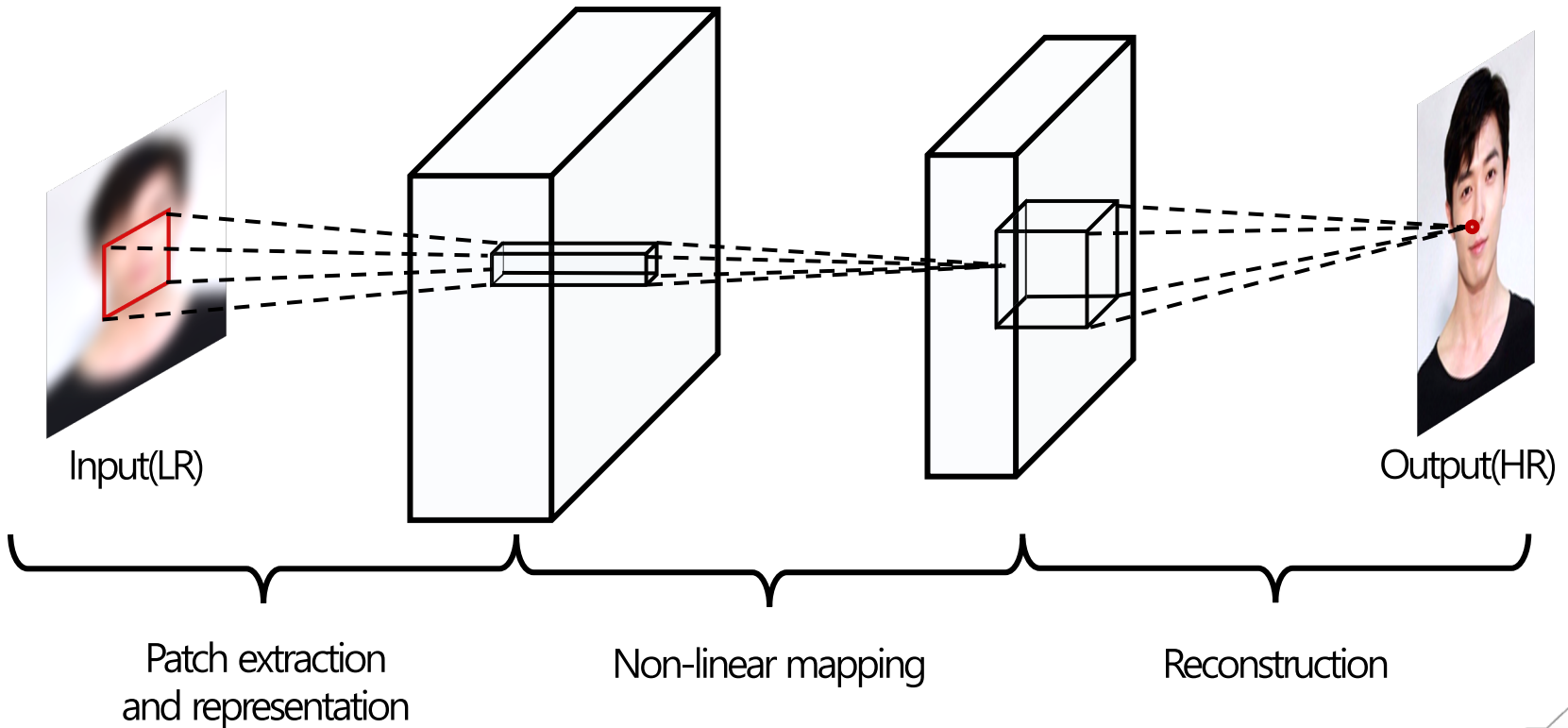
- Image Super Resolution : 저해상도(LR) 이미지를 고해상도(HR) 이미지로 변환하는 연구 분야
 - **Single Image Super Resolution**, Multi Image Super Resolution



Applications of Deep Learning for Computer Vision

❖ Image Super Resolution

- Image Super Resolution : 저해상도(LR) 이미지를 고해상도(HR) 이미지로 변환하는 연구 분야
 - **Single Image Super Resolution**, Multi Image Super Resolution



[1] Dong, C., Loy, C. C., He, K., & Tang, X. (2015). Image super-resolution using deep convolutional networks. *IEEE transactions on pattern analysis and machine intelligence*, 38(2), 295-307.

Applications of Deep Learning for Computer Vision

❖ Image Synthesis

- Image Synthesis : 새로운 이미지를 생성하거나 기존의 이미지를 합성하는 연구 분야



Bedrooms

[1] Radford, A, Metz, L, & Chintala S (2015). Unsupervised representation learning with deep convolutional generative adversarial networks. arXiv preprint arXiv:1511.06434.

[2] Oord, A.V.D., Kalchbrenner, N, Vinyals O, Espeholt, L, Graves, A, & Kavukcuoglu, K. (2016). Conditional image generation with pixelwise decoders. arXiv preprint arXiv:1606.05328.



Applications of Deep Learning for Computer Vision

❖ Image Synthesis

- Image Synthesis : 새로운 이미지를 생성하거나 기존의 이미지를 합성하는 연구



Zebras



Horses

[1] Radford, A, Metz, L, & Chintala, S (2015). Unsupervised representation learning with deep convolutional generative adversarial networks. arXiv preprint arXiv:1511.06434.

[2] Oord, A.V.D., Kalchbrenner, N, Vinyals, O, Espeholt, L, Graves, A, & Kavukcuoglu, K. (2016). Conditional image generation with pixelwise decoders. arXiv preprint arXiv:1606.05328.



Applications of Deep Learning for Computer Vision

❖ Image Reconstruction

- Image Reconstruction : 손상된 이미지를 원래 이미지처럼 복원하는 연구 분야

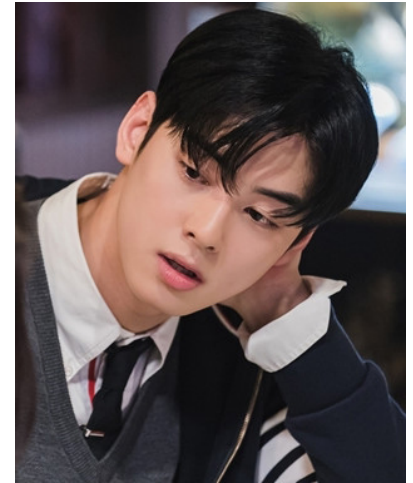
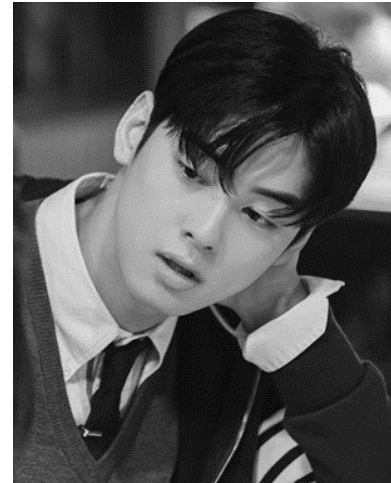


[1] Van Oord, A, Kaldobrenner, N, & Kavukcuoglu, K. (2016, June). Pixel recurrent neural networks. In International Conference on Machine Learning (pp. 1747-1756). PMLR

Applications of Deep Learning for Computer Vision

❖ Image Colorization

- Image Colorization : 손상된 이미지 색을 원래 이미지 색 처럼 복원하는 연구 분야



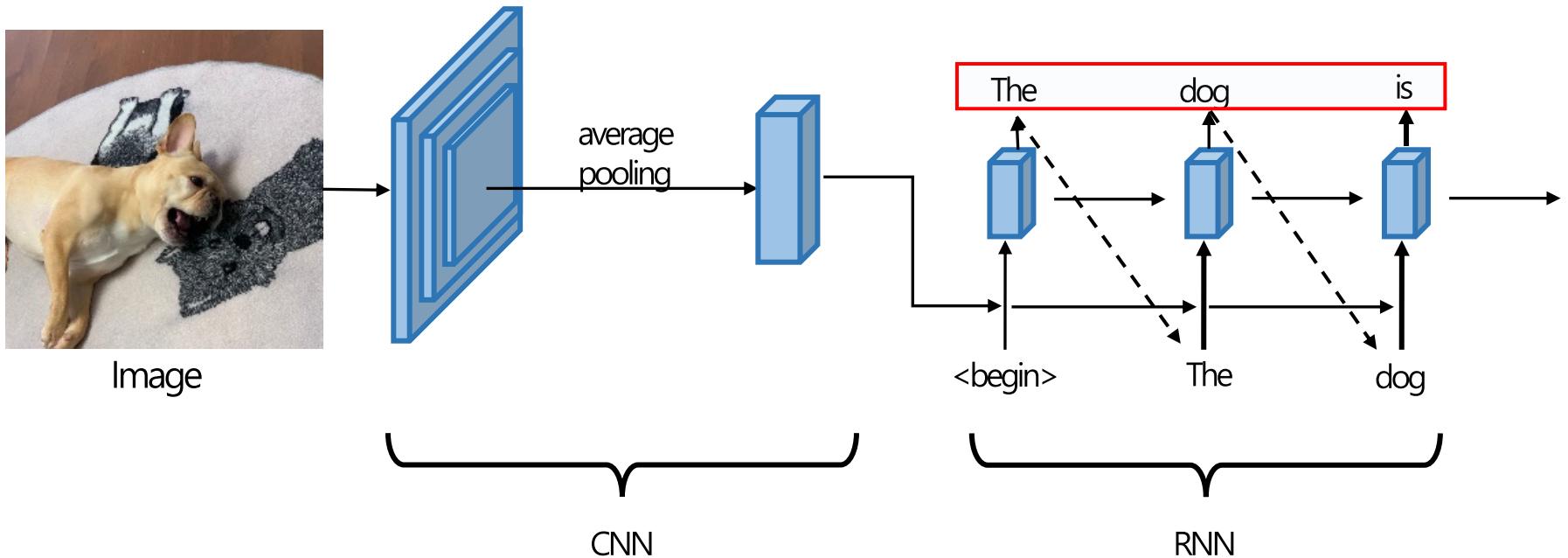
[1] Zhang, R, Isola, P, & Efros, A. A. (2016, October). Colorful image colorization. In European conference on computer vision (pp. 649-666). Springer, Cham.



Applications of Deep Learning for Computer Vision

❖ Computer Vision and NLP

- Image Captioning : 이미지를 설명하는 문장을 생성하는 연구 분야
 - 컴퓨터 비전과 자연어 처리를 연결하는 연구 분야



[1] Vinyals O, Toshev A, Bengio S, & Erhan D. (2015). Show and tell: A neural image caption generator. In Proceedings of the IEEE conference on computer vision and pattern recognition (pp. 3156-3164).

Applications of Deep Learning for Computer Vision

❖ Computer Vision and NLP

- Text to Image : Text를 가지고 실제 이미지와 같은 합성 이미지를 생성하는 연구 분야
 - 컴퓨터 비전과 자연어 처리를 연결하는 연구 분야

This small bird has a pink breast and crown, and black primaries and secondaries.



This magnificent fellow is almost all black with a red crest, and white cheek patch.



[1] Reed, S., Akata, Z., Yan, X., Logeswaran, L., Schiele, B. & Lee, H. (2016 June) Generative adversarial text-to-image synthesis. In International Conference on Machine Learning (pp. 1060-1069). PMLR



Conclusion

❖ 결론

- Computer Vision Tasks
 - Image Classification
 - Image Classification with Localization
 - Object Detection
 - Object Segmentation
 - Image Super Resolution
 - Image Synthesis
 - Image Reconstruction
 - Image Colorization
- Computer Vision and NLP
 - Image Captioning
 - Text to Image



감사합니다.