CV

AI명예학회

SKHU

Classification

CAT

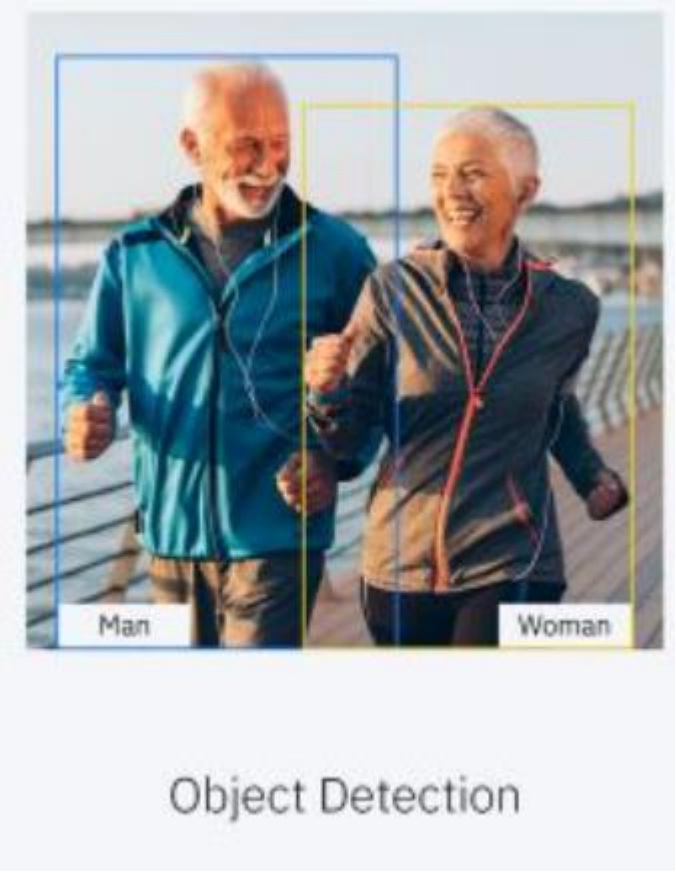
No spatial extent

우리가 했던 분류

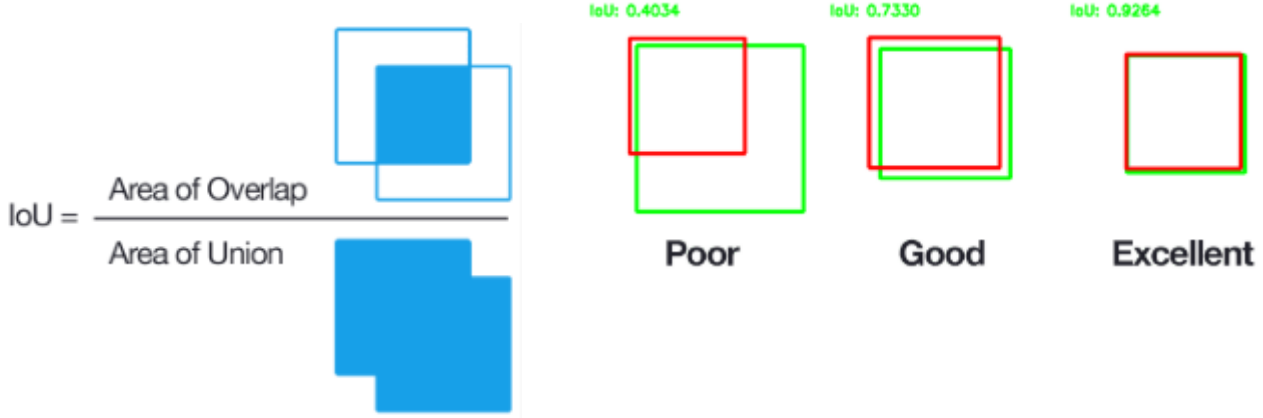
공간 개념 존재 X

AI명예학회
SKHU

Object Detection

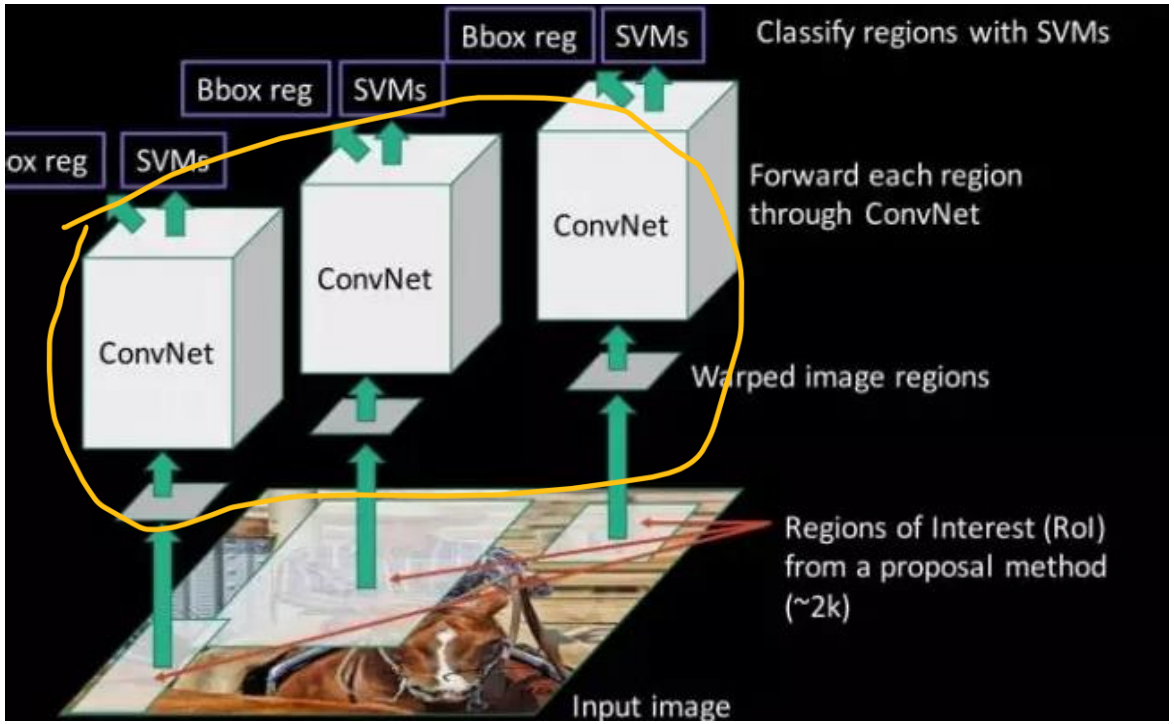


1. 내가 이미지에서 라벨링한 구역 == 실측값
2. 모델이 예측한 구역 == 예측값
3. 이 두개의 IOU값.



AI명예학회
SKHU

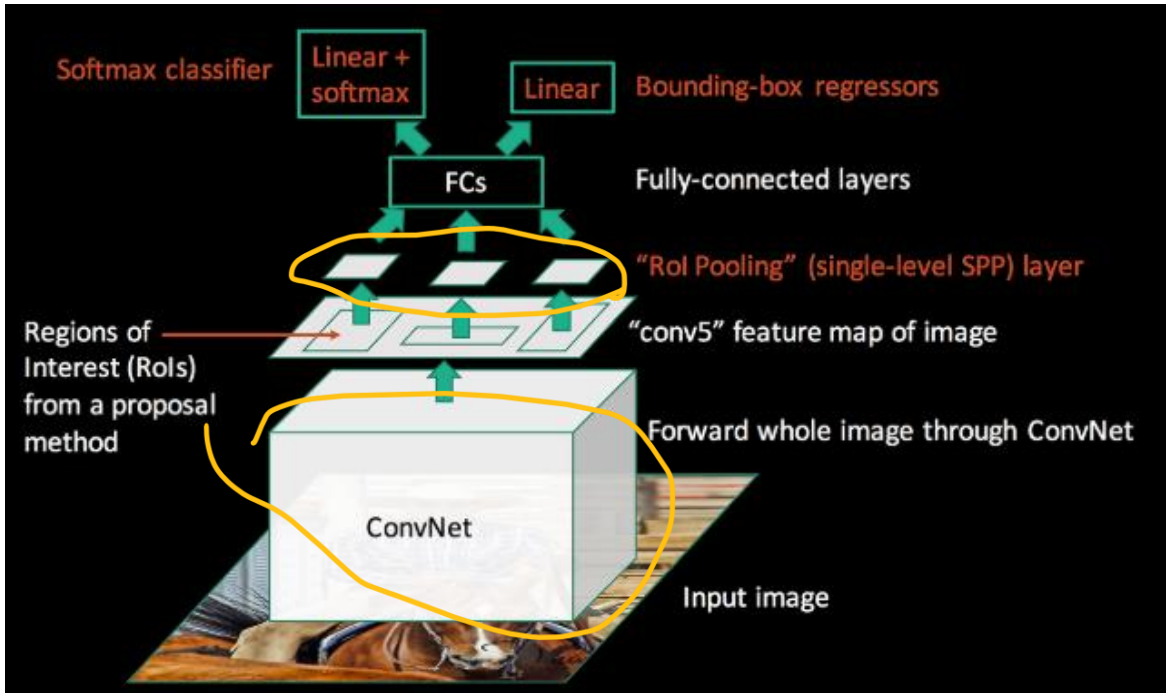
Object Detection: R-CNN



1. Input Image
2. RoI(Region of interest) : Selective Search
3. RoI 동일 사이즈로 Resizing
4. 각 독립적 ConvNet 통과(질감, edge 등)
5. SVM 분류 통한 클래스 예측
6. Bounding-Box Regressor(새로운 물체 위치 좌표값)

AI명예학회
SKHU

Object Detection: Fast R-CNN



1. Input Image
2. 한방에 ConvNet ➔ Feature map 출력
3. RoI
4. RoI Pooling
5. FC Layer

RoI Pooling:
RoI(관심 영역)를 일정한
그리드로 나누고, 그 후에
Max Pooling

AI명예학회
SKHU

Object Detection: YOLO

One Stage vs Two Stage

R-CNN: Region Proposal & CNN ➔ Two stage

YOLO: 이미지 전체 한방에 ConvNet,
Bounding Box 좌표값, 클래스 확률, 신뢰도 동시 출력 ➔ One Stage

RoI 빠져서 빠른 속도,

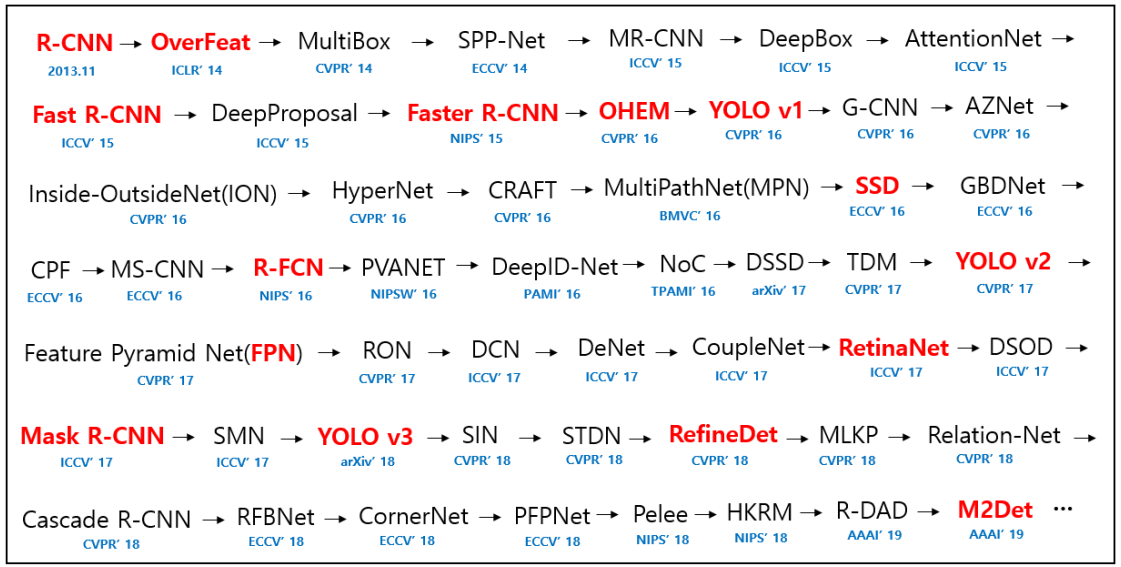
but, 물체 작다하면 R-CNN 계열 유리. ➔ 정밀도

https://pjreddie.com/darknet/yolo/

We use a totally different approach. We apply a single neural network to the full image. This network divides the image into regions and predicts bounding boxes and probabilities for each region. These bounding boxes are weighted by the predicted probabilities.

AI명예학회
SKHU

Object Detection



https://roboflow.com/

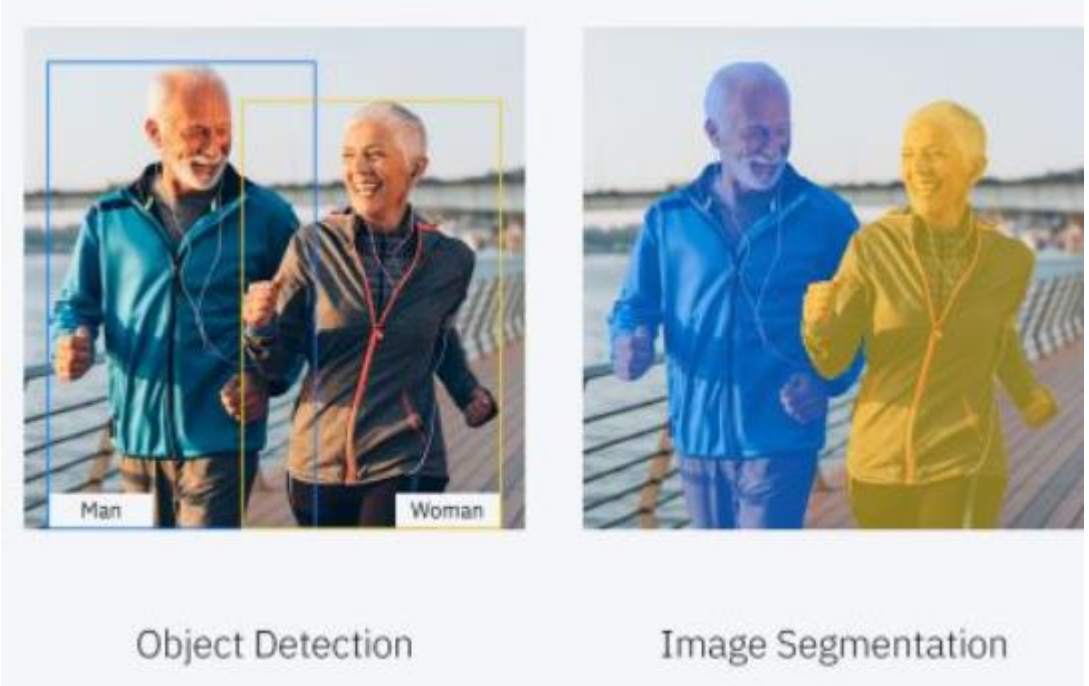
https://github.com/HumanSignal/labelImg/releases

https://github.com/tensorflow/models/tree/master/research/object_detection → tensorflow

https://github.com/facebookresearch/detectron2 → Pytorch

AI명예학회
SKHU

Image Segmentation

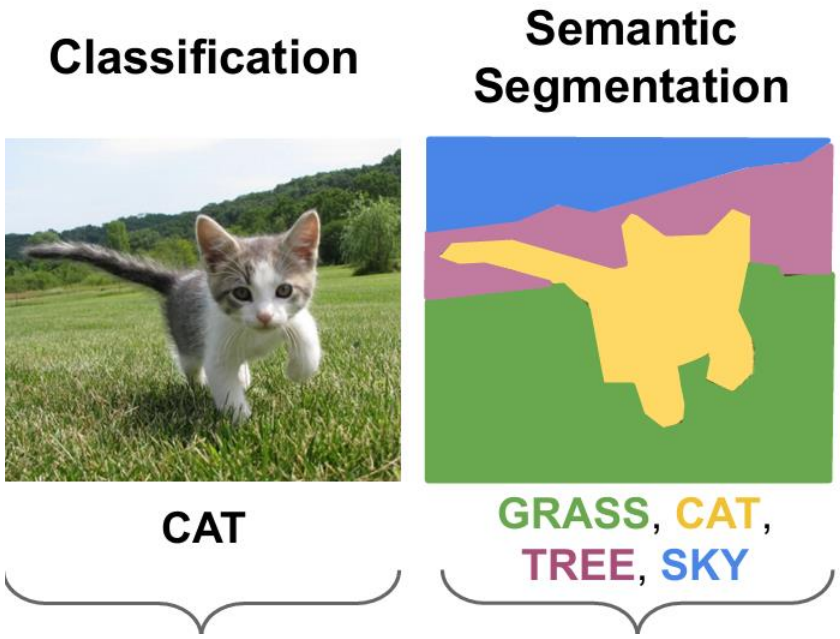


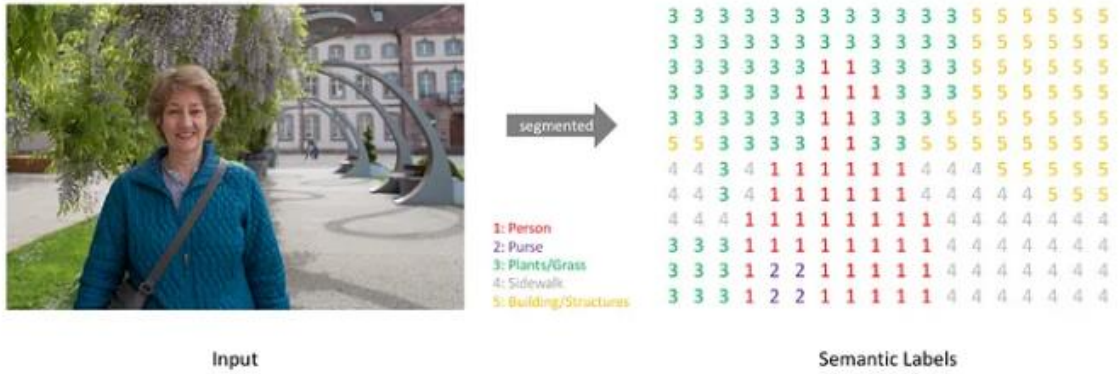
물체 위치 Bounding Box?

물체 실루엣?

AI명예학회
SKHU

Image Segmentation: Semantic Segmentation





단순 Pixel 변화에 따른 것.

이 친구가 사람인지 고양이인지
하늘인지 등등.

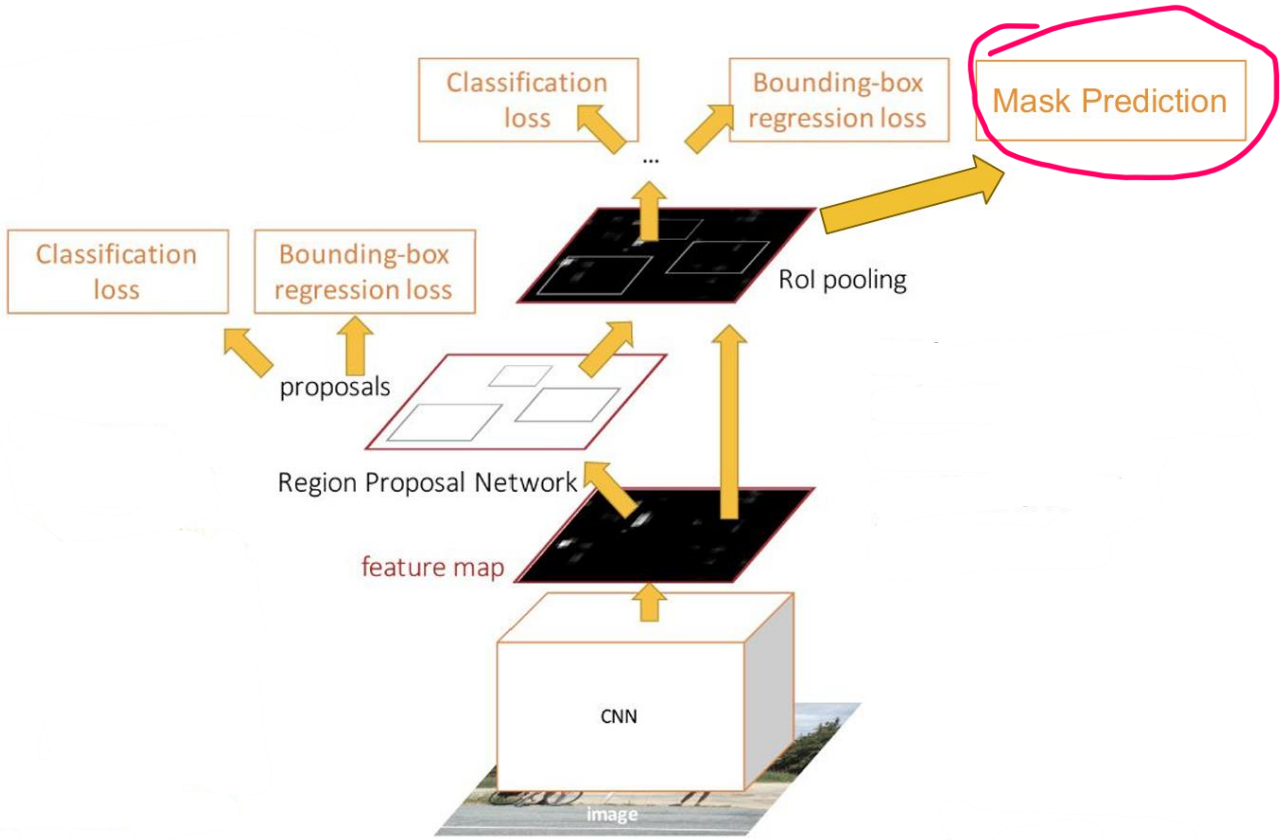
자세한건 다음시간에...

AI명예학회
SKHU

Image Segmentation: Instance Segmentation

AI명예학회
SKHU

Instance Segmentation: Mask R-CNN

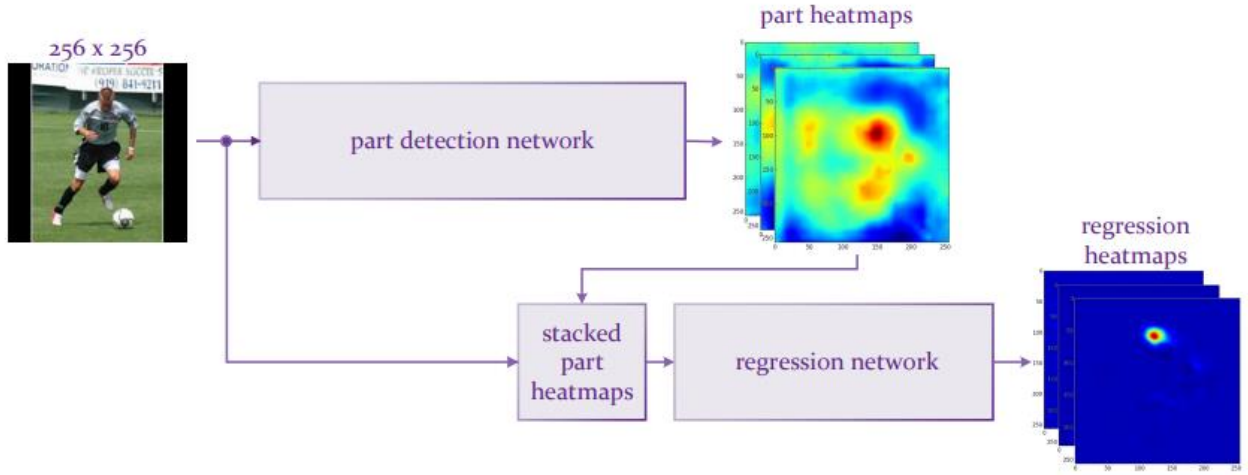


1. Input Image
2. CNN
3. RPN
4. Region Proposal: 물체 후보 영역 좌표 & 신뢰도 출력 값
5. RoI Pooling
6. 최종출력: 물체 클래스, Bounding Box 좌표값, 마스크 (픽셀단위)

RPN: Selective Search의 진보된 버전.

AI명예학회
SKHU

Pose Estimation



1. Input Image
2. 관절 위치 detection(픽셀 단위의 세부 단위)
3. Heatmap 추출(어깨, 무릎 위치 등) ➔ 해당 관절 위치 가능성 높은거 밝은 색.
4. 각 관절 정보 결합
5. 회귀 Network 거쳐서 관절 정확한 위치 추출.

➔ Detection 결과를 Regression이 정밀하게 조정.

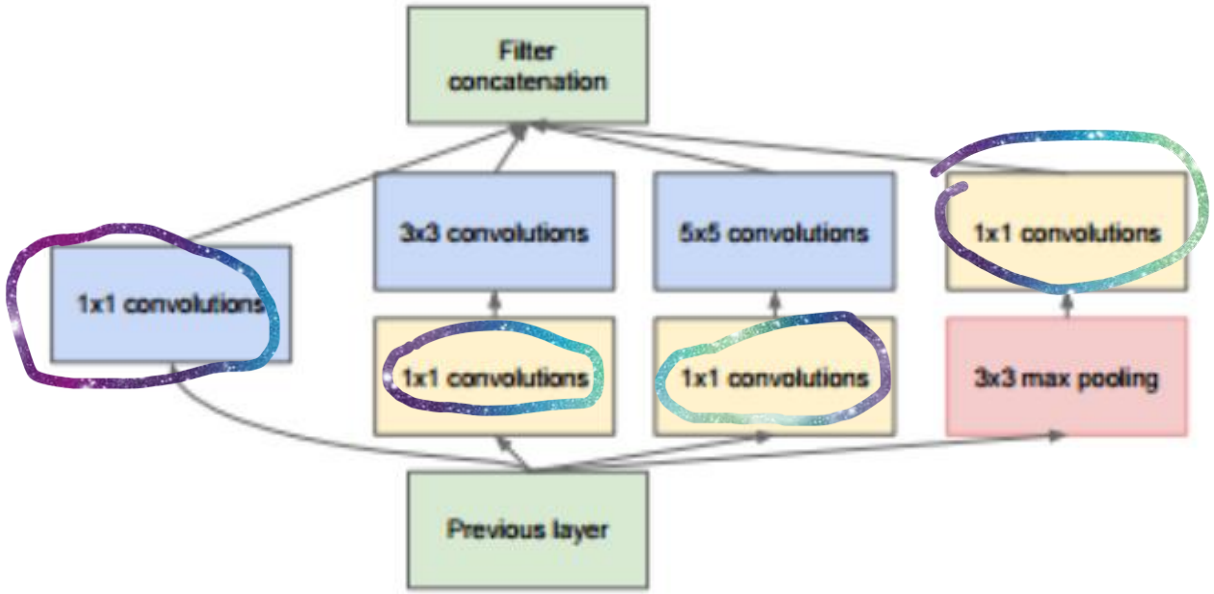
AI명예학회
SKHU

대표적 Pose Estimation 오픈소스 라이브러리

- OpenPose
- MediaPipe
- DensePose

AI명예학회
SKHU

지금까지의 CNN 계층 이해



(b) Inception module with dimensionality reduction

GoogleNet

Input_data = 32 x 32 x 3
3x3 필터 수 = 64,
5 x 5 필터 수 = 32

3x3 conv 연산량:
Input_data * (3 * 3) * 64 = 1769472

5 x 5 conv 연산량:
Input_data * (5 * 5) * 32 = 2457600

Total: 4227072

AI명예학회
SKHU

지금까지의 CNN 계층 이해

1 x 1 사용시:

1x1 필터 수: 16

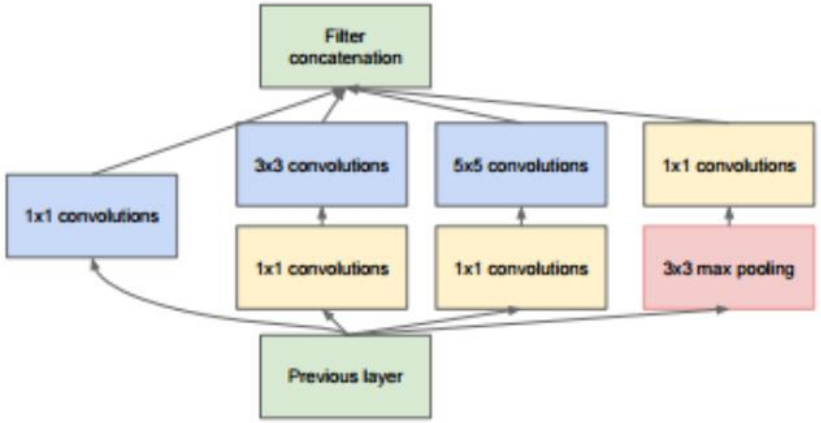
Input_data = 32 x 32 x 3

1x1 conv 연산량:
input_data * (1* 1) * 16 = 49152

1x1 적용 후 3x3 conv 연산량:
32 * 32 * 16 * 9 * 64 = 943718

1x1 적용 후 5x5 conv 연산량:
32 * 32 * 16 * 25 * 32 = 1310720

Total: 2303590



(b) Inception module with dimensionality reduction

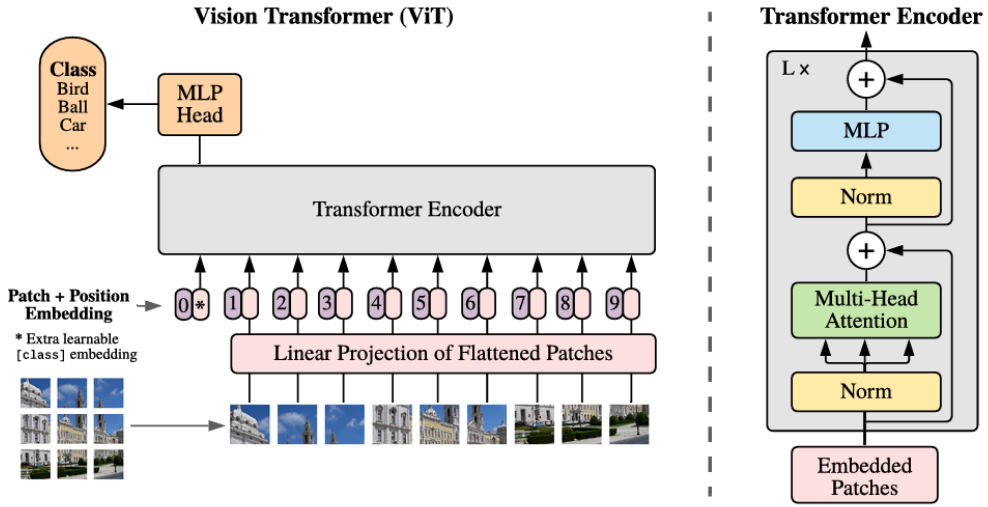
획기적 CNN 모델.

but, 연산량, 계산량 줄이는 것.

이런건 이미 얼추 정복된 분야

AI명예학회
SKHU

Vision Transformer(ViT)



1. Input Image
2. 작은 조각들로 쪼개기(Patch)
3. 선형 변환(동일 차원의 벡터)
4. 각각의 위치정보 추가 & Token 부여
5. Transformer Encoder
6. Class token -> MLP Head
7. 최종 클래스 예측

AI명예학회
SKHU

Assignment

1. 수업 내용 정리

2. 분류 모델 만들기
   2.1 이미지 데이터셋 구축
   2.2 관절 추출(분류 정확도 최대한 높일 수 있는 좌표값)
   2.3 CSV 파일 추출
   2.3 머신러닝 분류기 사용 후 정확도 까지 볼 것.
모든 좌표는 쓰지 말 것.

3. 제출 후, Pr 코멘트에 데이터셋은 무엇인지, 추출한 좌표, 이 좌표를 추출한 이유는 무엇인지.

AI명예학회
SKHU