

# Fully-Convolutional Siamese Networks for Object Tracking

박태우

# Object Tracking이란?

- 영상에서 bounding box가 그려진 특정 object를 계속해서 추적하는 기술.
- 이전에는 컴퓨터 비전 기술로 구현.
- 최근에는 딥러닝을 접목시켜 더욱 더 높은 성능을 보여줌.



# 동작 알고리즘

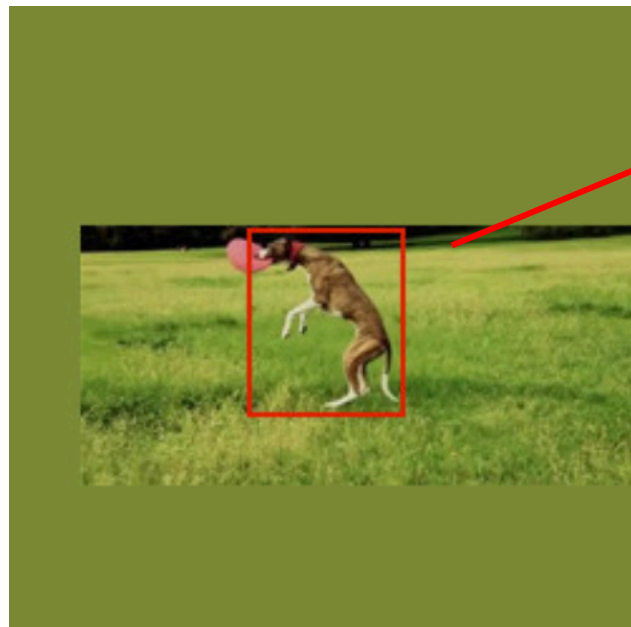
본 논문에서는 특정 프레임  $t$ 에서 object를 이후의 프레임에서 찾는 것을 목표로 한다.

기준 object



$t$

찾으려는 목표

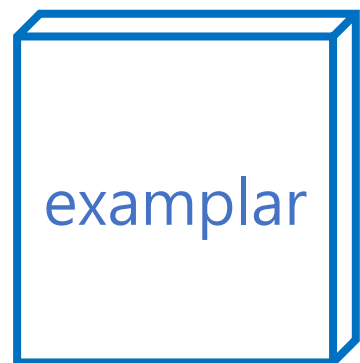


$t+n$

**딥러닝 모델에서 Siamese network를 사용.**

두 개의 입력에 대해 독립적으로 두 개의 합성곱 신경망을 실행한 뒤 비교하는 아이디어를 삼 네트워크라고 합니다.

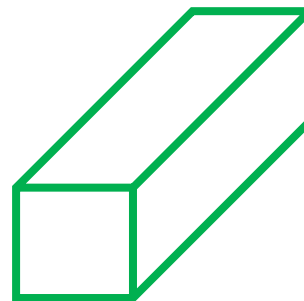
# 동작 알고리즘



127 x 127 x 3



Neural Network



6 x 6 x 128

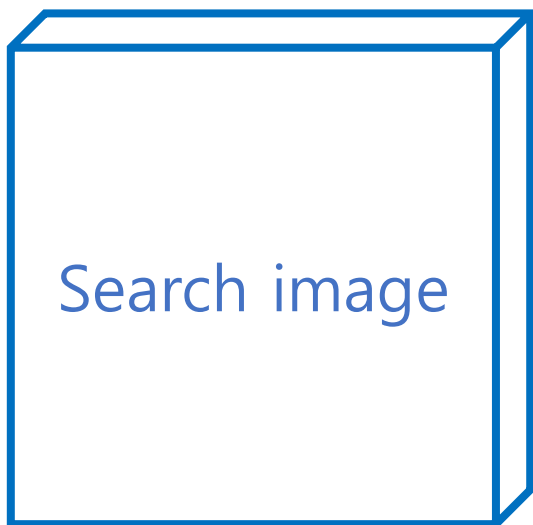


\*



Score map

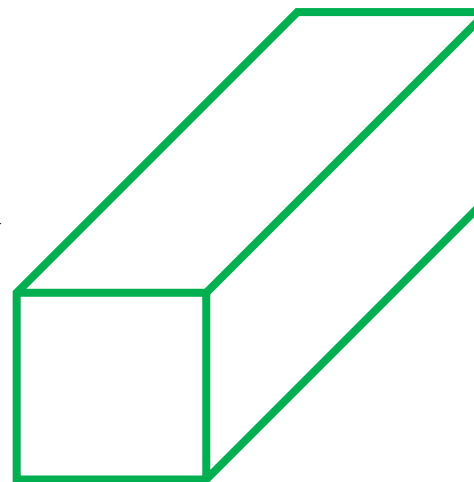
17 x 17 x 1



255 x 255 x 3



Neural Network



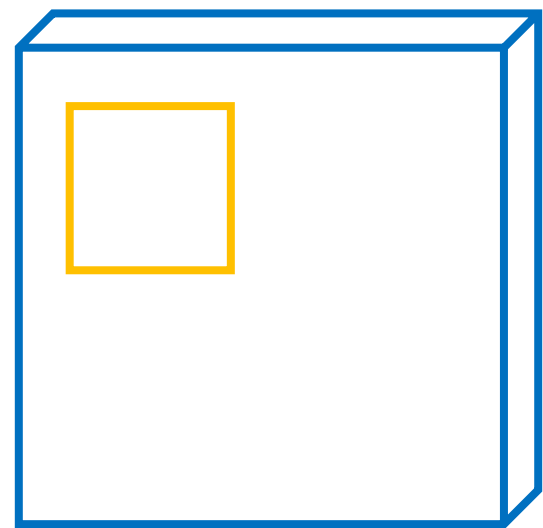
22 x 22 x 128



# 동작 알고리즘

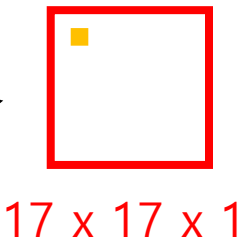
Search image에서 sub-window를  
convolutional 했을 때 위치에 score를 저장한다.

Score map 각 점에서 convolutional 하기 전  
크기가 sub-window가 된다.



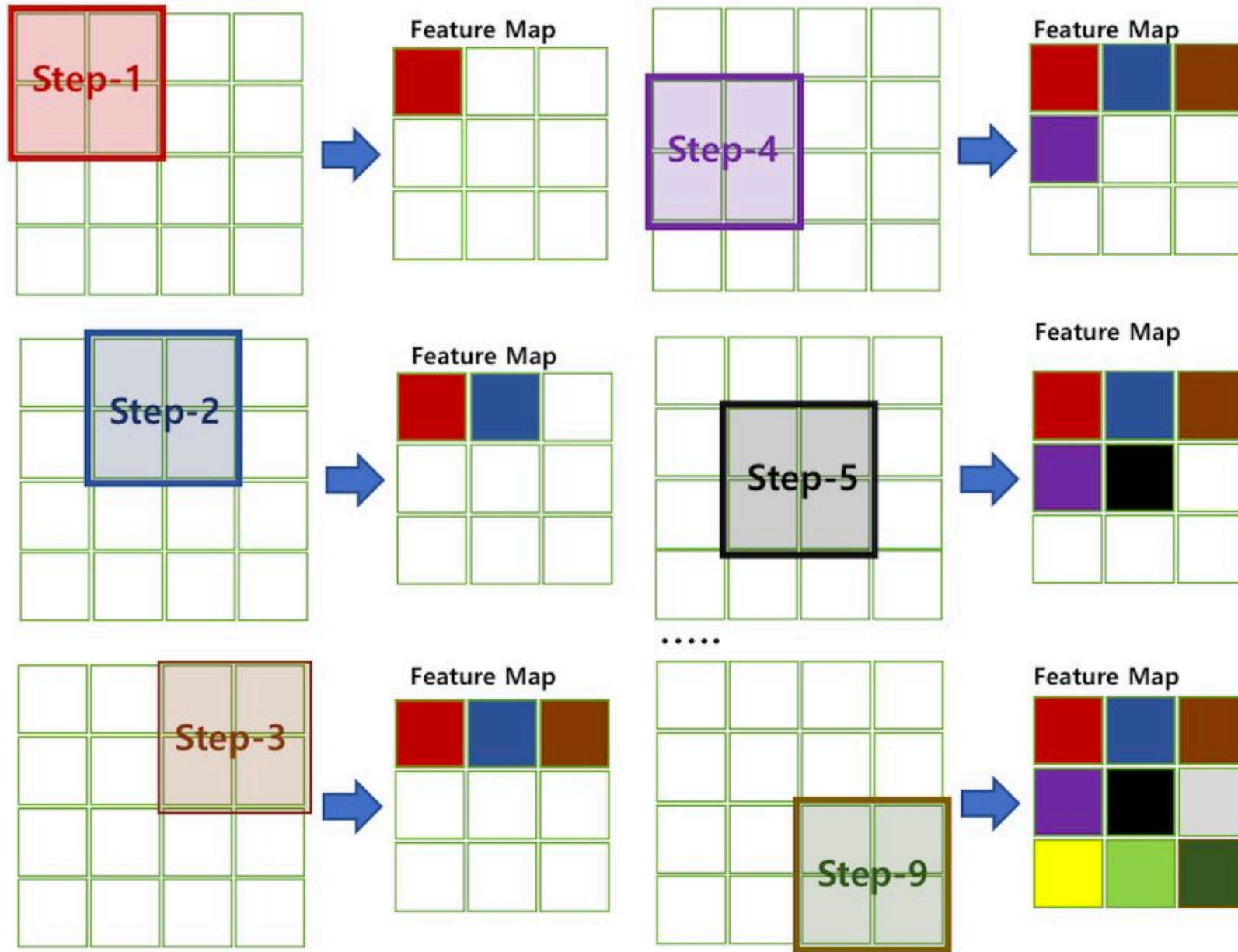
255 x 255 x 3

fully-convolutional  
with stride k

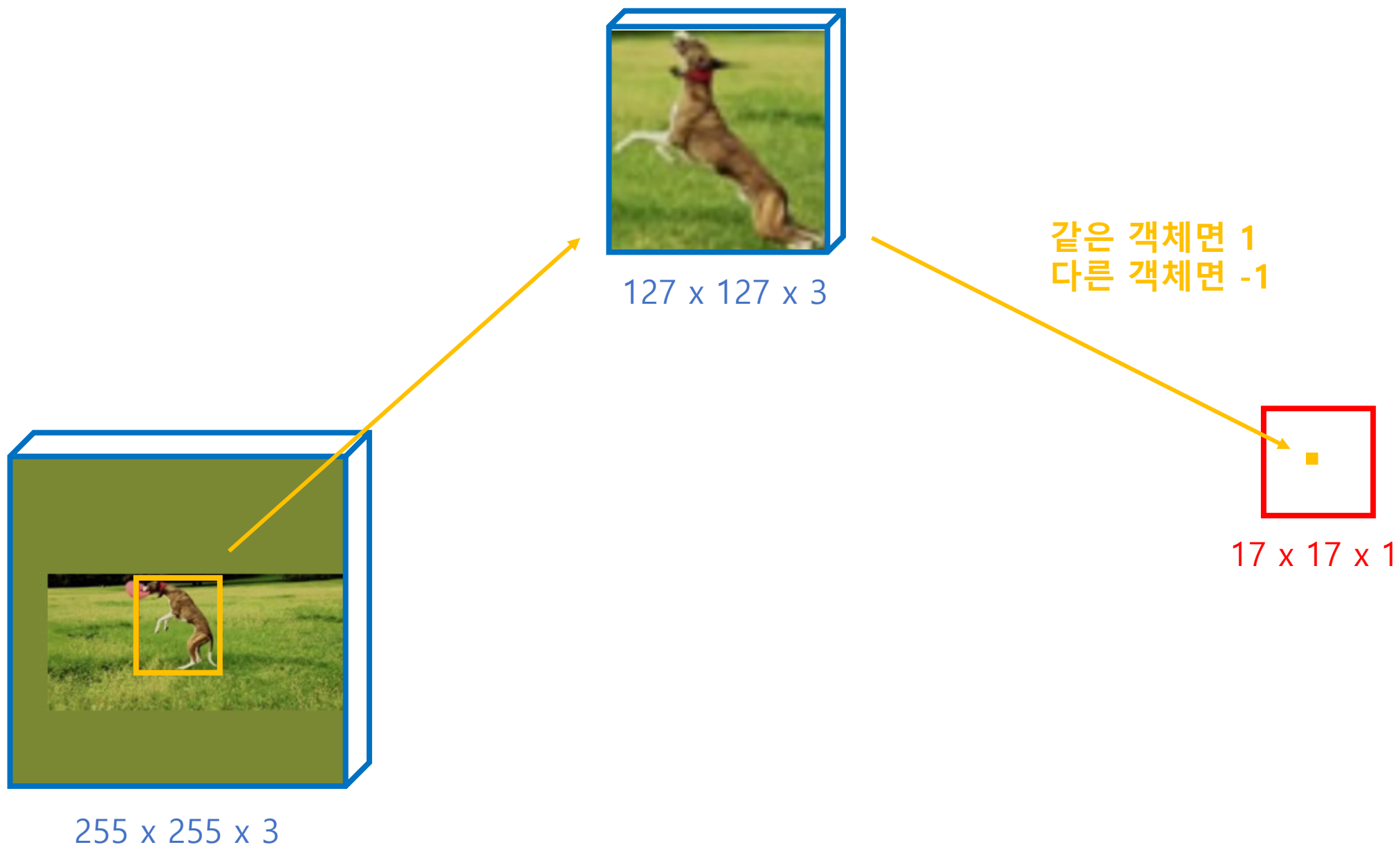


17 x 17 x 1

# 동작 알고리즘



# 동작 알고리즘





# Training Loss

한 점에서의 logistic loss

$$\ell(y, v) = \log(1 + \exp(-yv))$$

Score map 전체의 평균

$$L(y, v) = \frac{1}{|\mathcal{D}|} \sum_{u \in \mathcal{D}} \ell(y[u], v[u])$$

SGD 알고리즘을 이용해  
loss의 최소값을 위한  $\theta$ 를 찾음

$$\arg \min_{\theta} \mathbb{E}_{(z, x, y)} L(y, f(z, x; \theta))$$

# 정리

## 특징

- 전체 이미지에서 object 1개를 계속해서 추적하는 tracker
- 정확도가 높지는 않지만 속도는 상당히 빠름.
- Siamese network를 사용했음.

# 최근 tracking

- 최근 딥러닝을 이용한 tracking의 경우 Multi Object Tracking을 수행함.
- Object가 사라졌다가 다시 등장했을 때 재인식하는 Re-Identification 기술 사용.
- MOTChallenge에서 다양한 dataset 제공.