

# 꼼꼼한 딥러닝 논문 리뷰와 코드 실습

Deep Learning Paper Review and Code Practice

나동빈([dongbinna@postech.ac.kr](mailto:dongbinna@postech.ac.kr))

Pohang University of Science and Technology

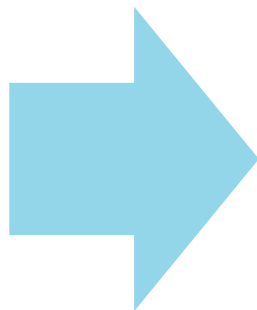
# Image Style Transfer Using Convolutional Neural Networks (CVPR 2016)

- 본 논문에서는 CNN을 활용해 이미지의 스타일(style)을 전송하는 방법을 제안합니다.



Content

Style Transfer 결과



Style



Style

# Image Style Transfer Using Convolutional Neural Networks (CVPR 2016)

- 스타일 전송(style transfer)을 위해 사전 학습된(Pre-trained) CNN 모델을 사용합니다.
- 네트워크의 가중치는 고정한 뒤에 이미지를 변경시키는 방법을 사용합니다.
  - 이미지를 학습(업데이트)한다고 이해할 수 있습니다.
- 이미지를 최적화(optimization)한다는 것은 어떤 의미일까요?



*target*

이미지 변수  $x$  값이 변환되는 과정



$\vec{x}_1$



$\vec{x}_2$



$\vec{x}_3$

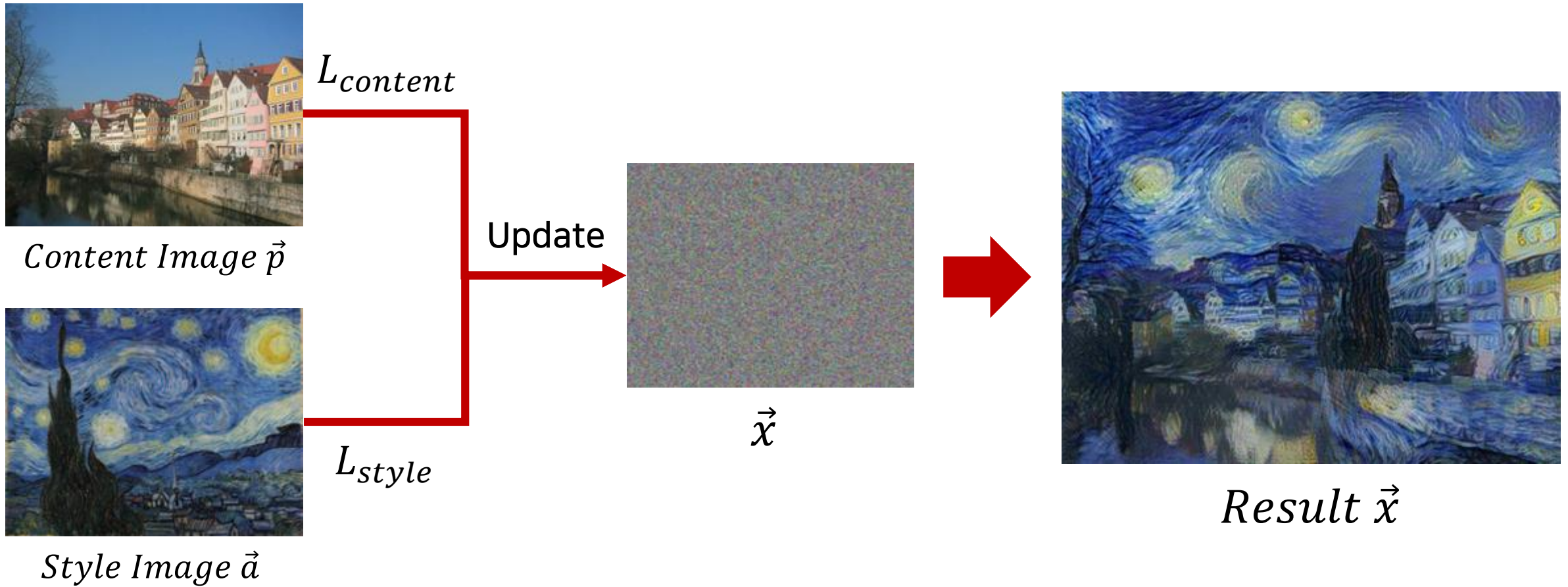
.....



$\vec{x}_{iters}$

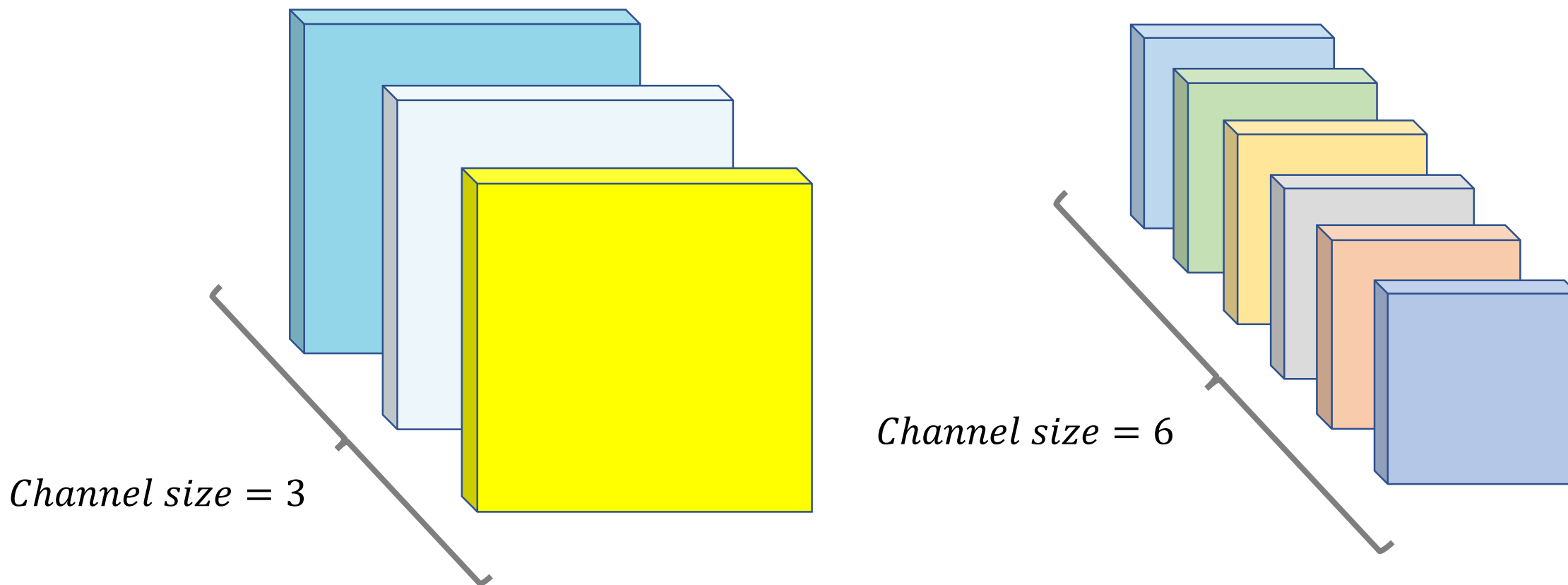
# Image Style Transfer Using Convolutional Neural Networks (CVPR 2016)

$$L_{total}(\vec{p}, \vec{a}, \vec{x}) = \alpha L_{content}(\vec{p}, \vec{x}) + \beta L_{style}(\vec{a}, \vec{x})$$



## CNN 모델의 특징 맵(Feature Map)

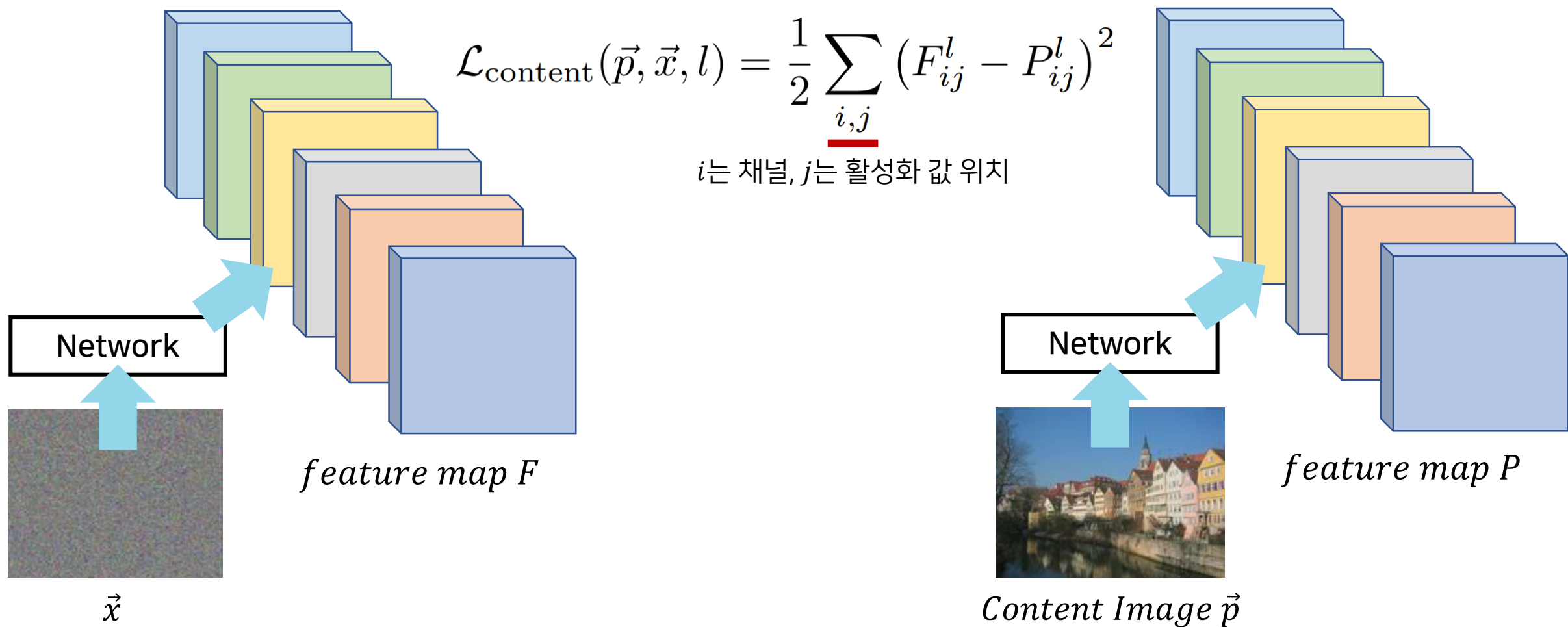
- 일반적으로 CNN에서 레이어가 깊어질수록 채널의 수가 많아지고 너비와 높이는 줄어듭니다.
- 컨볼루션 레이어의 서로 다른 필터들은 각각 적절한 특징(feature) 값을 추출하도록 학습됩니다.





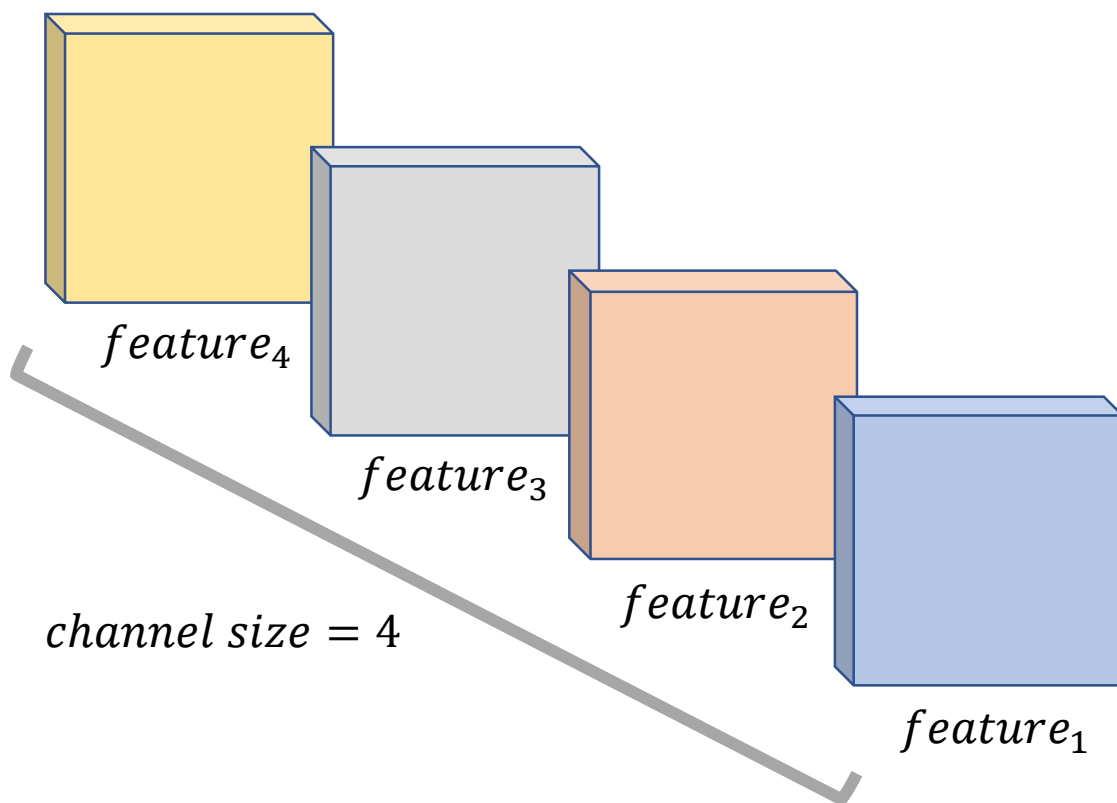
## 콘텐츠 손실(Content Loss)

- 콘텐츠 손실(content loss)은 두 이미지 특징(feature)의 activation 값이 동일하도록 만듭니다.



# 스타일 손실(Style Loss)

- 스타일(Style)은 서로 다른 특징(feature)간의 상관관계(correlation)를 의미합니다.
- $G_{ij}$  = 특징  $i$ 와 특징  $j$ 의 상관관계(correlation)



$$G_{ij} = \sum_k F_{ik} F_{jk} \quad (\text{feature } i \text{와 feature } j \text{의 내적})$$

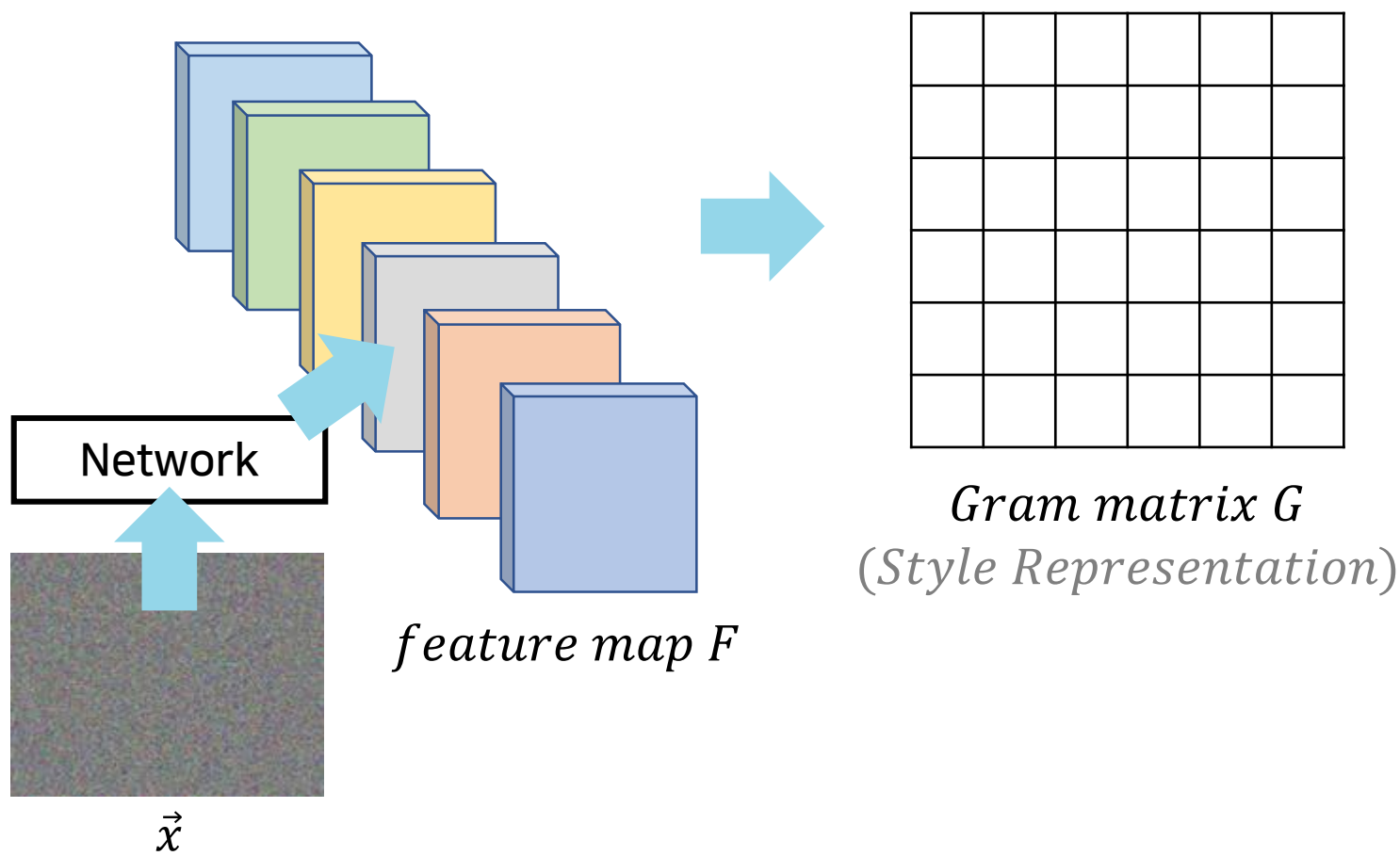
$k$ 는 활성화 값 위치

$G_{11}$	$G_{12}$	$G_{13}$	$G_{14}$
$G_{21}$	$G_{22}$	$G_{23}$	$G_{24}$
$G_{31}$	$G_{32}$	$G_{33}$	$G_{34}$
$G_{41}$	$G_{42}$	$G_{43}$	$G_{44}$

Gram Matrix (= Style Representation)

# 스타일 손실(Style Loss)

- 스타일 손실(style loss)은 **두 이미지의 특징 상관관계를 유사하게** 만들어줍니다.



두 이미지의 Gram Matrix가 유사하도록

$$E_l = \frac{1}{4N_l^2 M_l^2} \sum_{i,j} \overbrace{(G_{ij}^l - A_{ij}^l)^2}^{\uparrow}$$

$$\mathcal{L}_{\text{style}}(\vec{a}, \vec{x}) = \sum_{l=0}^L \underbrace{w_l E_l}_{\downarrow}$$

어떤 레이어에 많은 가중치를 둘 지



# Style Reconstruction과 Content Reconstruction 살펴보기

## [ Style Reconstructions ]

- Gram Matrix는 채널의 크기만큼 커지게 됩니다.

(a) conv1\_1

(b) conv1\_1, conv2\_1

(c) conv1\_1, conv2\_1, conv3\_1

(d) conv1\_1, conv2\_1, conv3\_1, conv4\_1,

(e) conv1\_1, conv2\_1, conv3\_1, conv4\_1, conv5\_1

## [ Content Reconstructions ]

- 깊어질수록 구체적인 픽셀 정보는 소실됩니다.

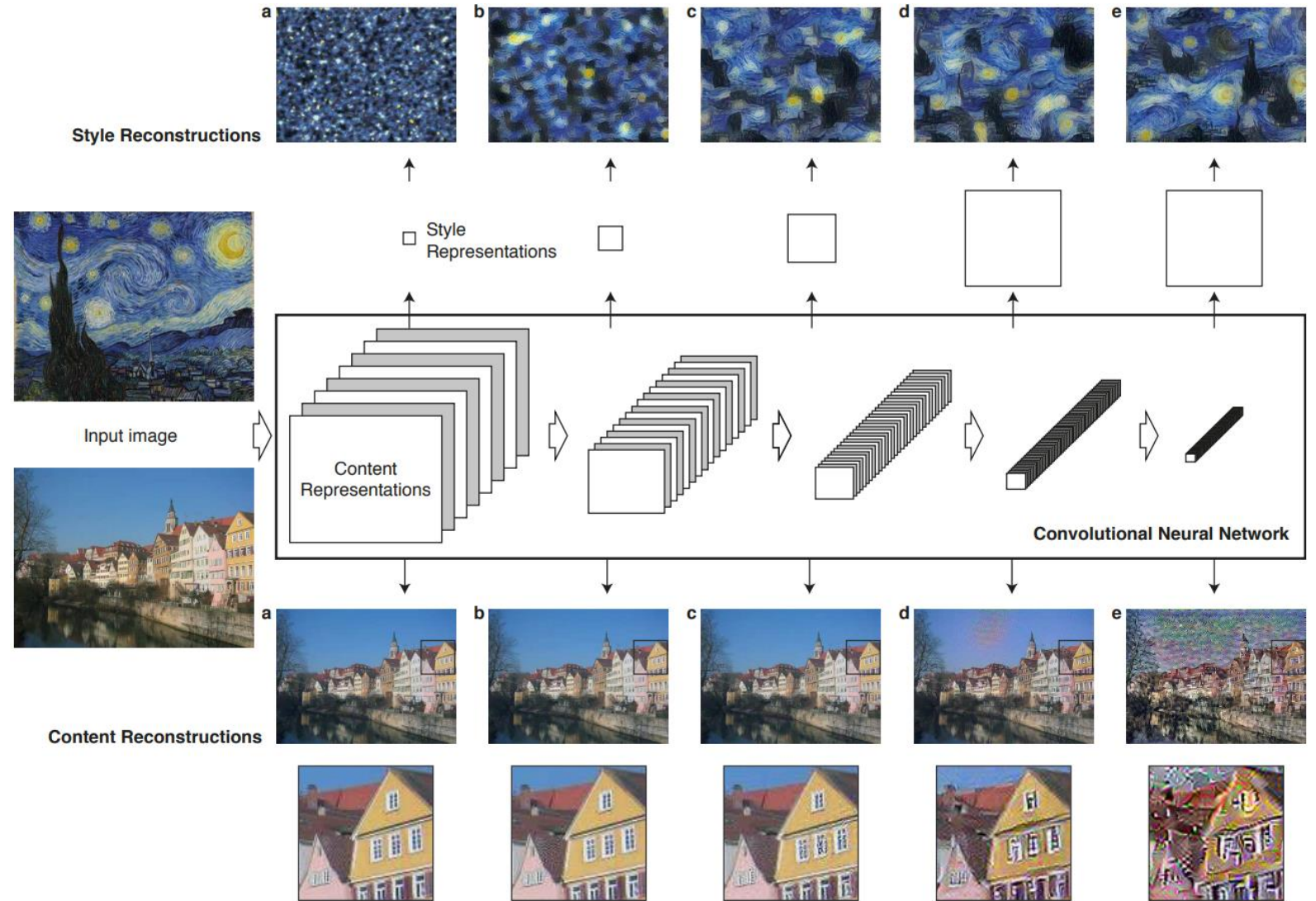
(a) conv1\_2

(b) conv2\_2

(c) conv3\_2

(d) conv4\_2

(e) conv5\_2



# Style Transfer 알고리즘

