

# Image Style Transfer Using Convolutional Neural Networks

윤세환

# 목차

- style transfer의 정의
- 기존 방식의 한계점
- 본 논문에서 제시한 방법
- 세부 구현
- 결과
- 의의

# Style Transfer의 정의

- 기반이 되는 콘텐츠 이미지에 스타일 이미지의 화풍을 추출해 새로운 이미지를 생성하는 기법
- Texture Transfer의 한 일종으로 본 논문에서는 취급하고 있으며, Texture Transfer는 input image와 어떠한 texture(패턴)이 있을 때 이 texture를 input image와 구조가 같은 이미지를 만드는 기법을 의미
- CNN 기반의 Style transfer가 있기 전, 알고리즘을 기반으로 한 방식들이 연구되어왔으며, Image Quilting 알고리즘을 사용한 경우가 많았음
  - Image Quilting은 작은 텍스처 이미지를 이와 유사한 더 큰 텍스처 이미지로 확장하는 방법

# 용어 정리

- 스타일을 추출하는 이미지를 Source Image
- 스타일을 적용하고자 하는 이미지를 Target Image



source image

+



target image

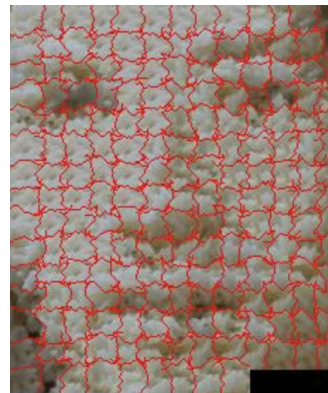


# 기존 방식의 한계점

- Image Quilting를 사용한 기존 방식은 **source image**와 **target** 이미지간 유사한 부분끼리 매칭시켜서 이미지를 생성하는 기법을 사용했는데, 이 방식을 사용할 경우, **target** 이미지의 **low level** 특징만을 사용할 수 밖에 없다.
- 즉, 사물을 인식하고 해당 사물에 스타일을 입히는 방식이라기보다는, **target**과 **source** 이미지를 여러 조각으로 나눈 후 유사한 이미지끼리 매칭시는 방식



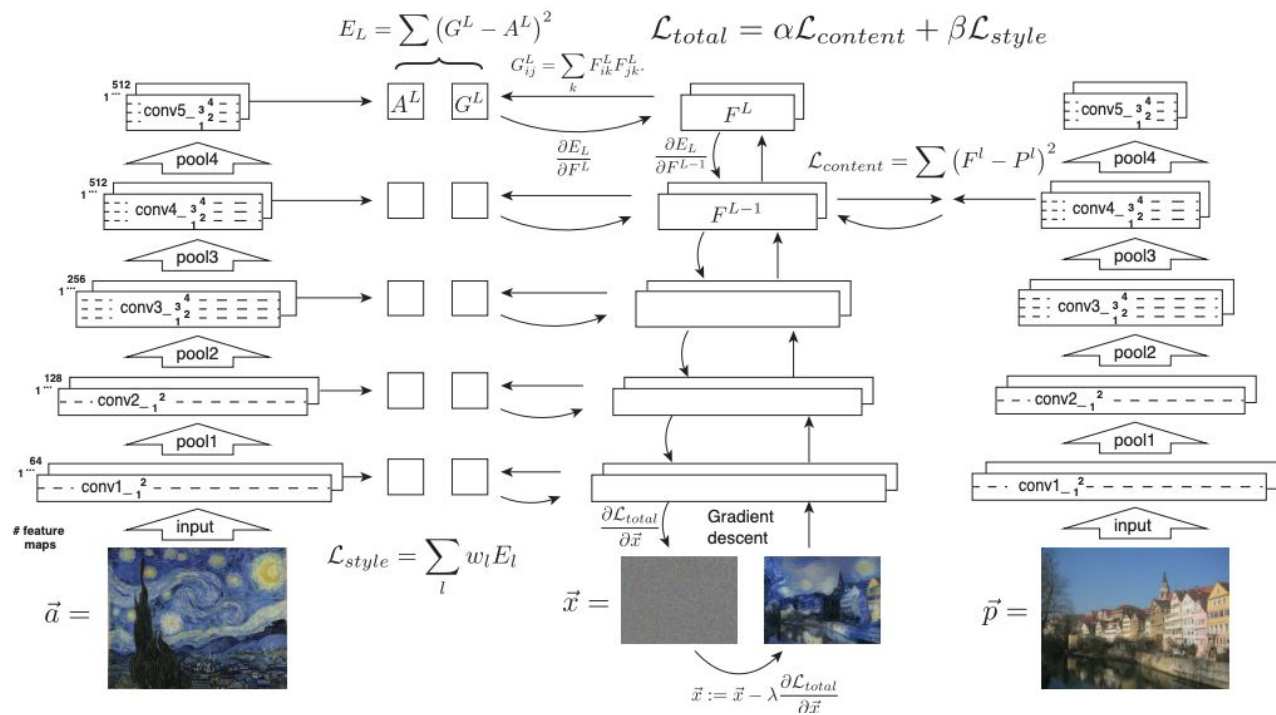
+



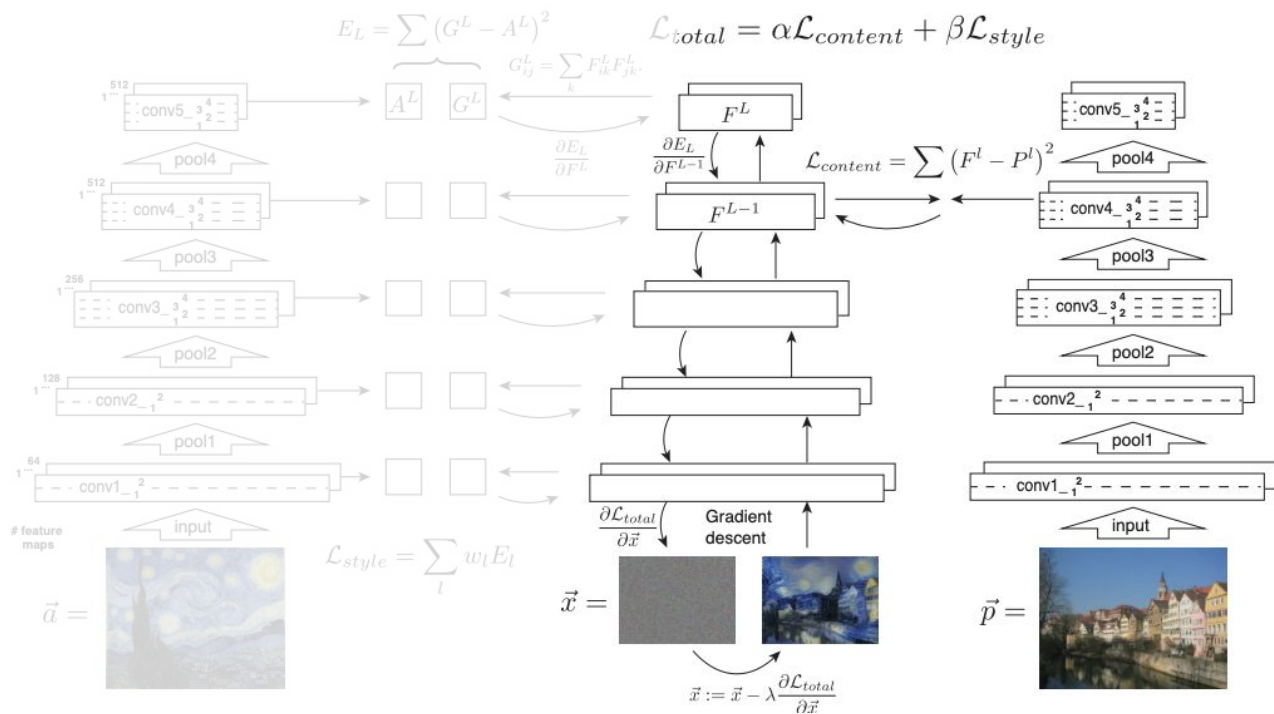
## 본 논문에서 제시한 방법

- **target image**에서 이미지내 **content**(사물같은거)를 추출하고, **source image**에서 스타일을 추출하여 추출한 **representation**을 사용하여 이미지를 생성
- 이미지의 특징을 추출할 때 **CNN**기반 모델을 사용 (**VGG-19**)
- **source image**의 **style**을 추출할 때는 비교적 앞에 있는 레이어의 결과값을 사용
- **target image**의 **content**를 추출할 때는 뒷단 레이어의 결과값을 사용
- 추출한 **style**, **content representation**값과 모두 유사한 이미지를 생성

# 세부 구현



# 세부 구현 - Content Representation





# 세부 구현 - Content Representation

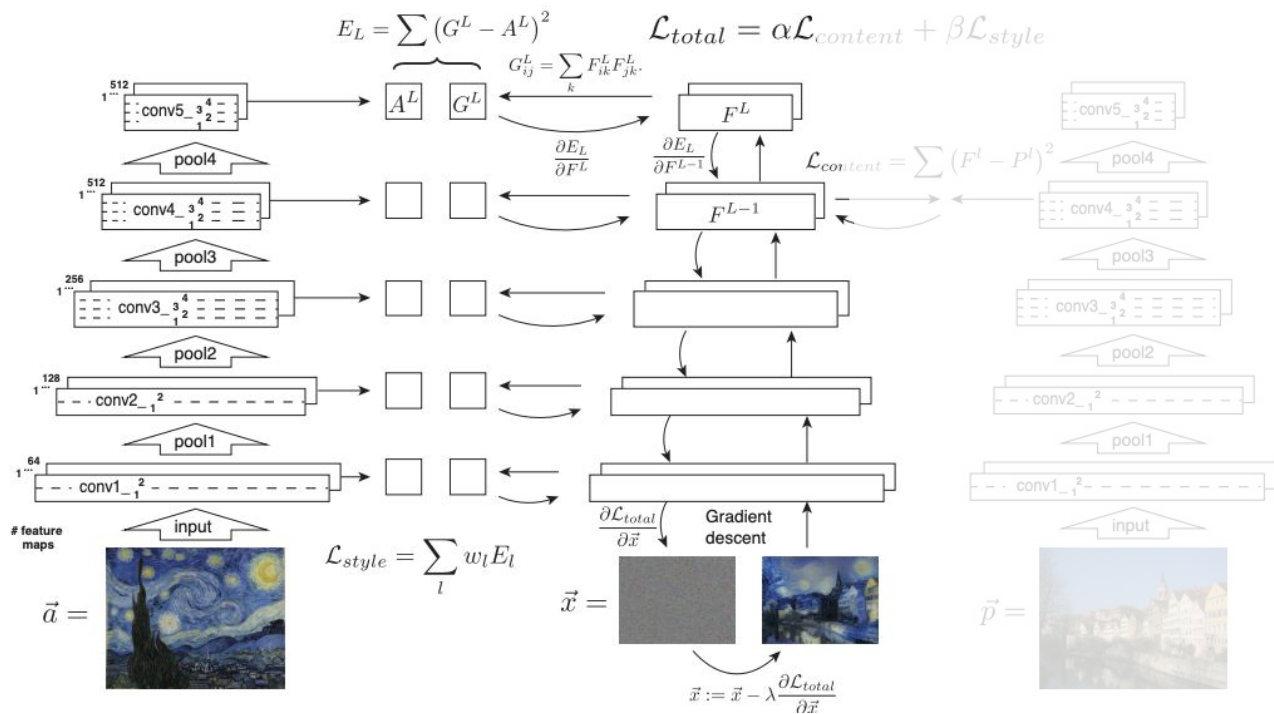
Content를 추출할 layer의 깊이에 따른 결과

- Conv2\_2의 결과를 보면, 저차원의 특성 (모서리 등)이 너무 선명해서 단순히 색깔을 입혀놓은 것 같은 느낌이 듭니다

- Conv4\_2의 결과를 보면 고차원의 특징 (사물)은 유지되면서 저차원의 특징은 어느 정도 제거된 모습을 보임



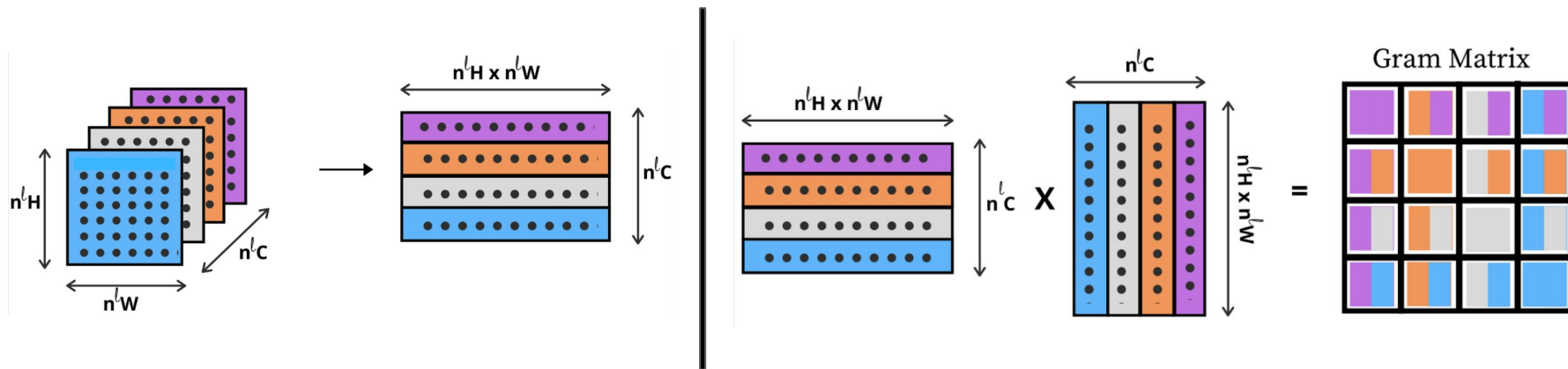
# 세부 구현 - Style Representation



# 세부 구현 - Style Representation

## Gram Matrix

- source image로부터 얻은 low level feature값을 채널마다 flatten한 후, 행렬곱을 수행하여 얻은 matrix

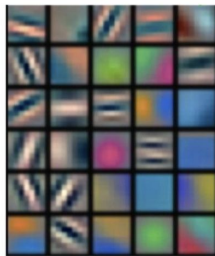


## 세부 구현 - Style Representation

단순 **feature** 값을 그대로 사용하지 않고 **gram matrix**를 사용하는 이유

- **low level feature**는 스타일뿐만 아니라 모서리, 가로/세로 직선과 같은 패턴을 인식하는 기능도 수행하기 때문에 스타일 외 다른 정보들도 가지고 있다.

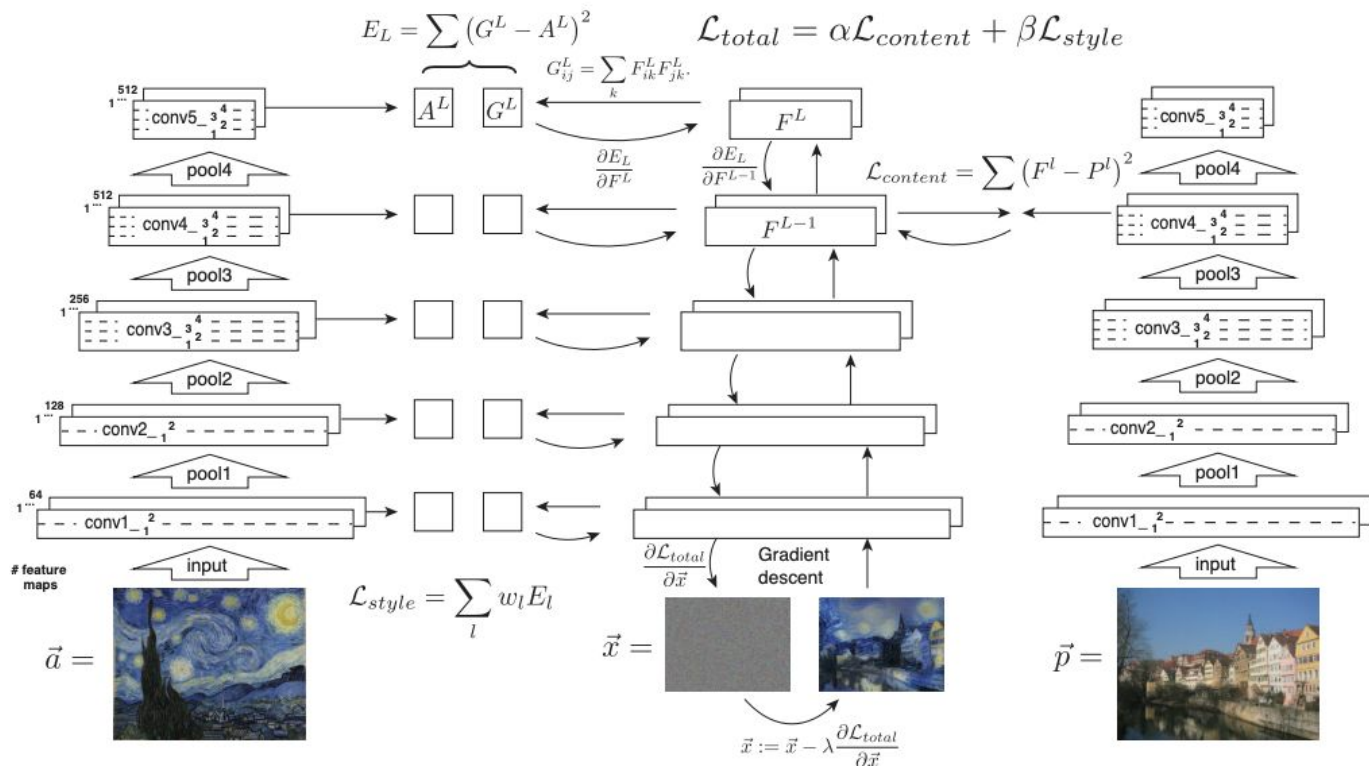
low-level features



- 다만, 모서리, 가로, 세로와 같은 사물을 판단하는데 사용하는 정보와는 달리, 스타일에 대한 정보는 모든 채널에 약간씩 존재할 것이라 판단하였고, 채널마다 존재하는 스타일을 추출하기 위해 상관계수를 사용했다.

(gram matrix의 형태가 공분산 행렬의 형태와 매우 유사하기에, gram matrix를 이루는 각 요소들을 상관계수로 보고 있음)

# 세부 구현



# 결과





## 한계점

- 이미지 1개를 합성할 때마다 학습을 수행하므로 시간, 자원이 많이 소요됨
- 이미지의 해상도가 커짐에 따라 합성 속도가 많이 느려짐
- 생성된 이미지를 보면, 낮은 수준의 **noise**가 발생하는 경우가 있음



## 참고 링크

- Gram Matrix 정리 : <https://aigong.tistory.com/360>
- Image Quilting을 사용한 Texture Transfer 정리 :  
<https://www.codeproject.com/Articles/24172/Texture-Transfer-using-Efros-Freeman-s-Image-Quilt>