HuStar AI Course: Computer Vision

GAN & StyleTransfer

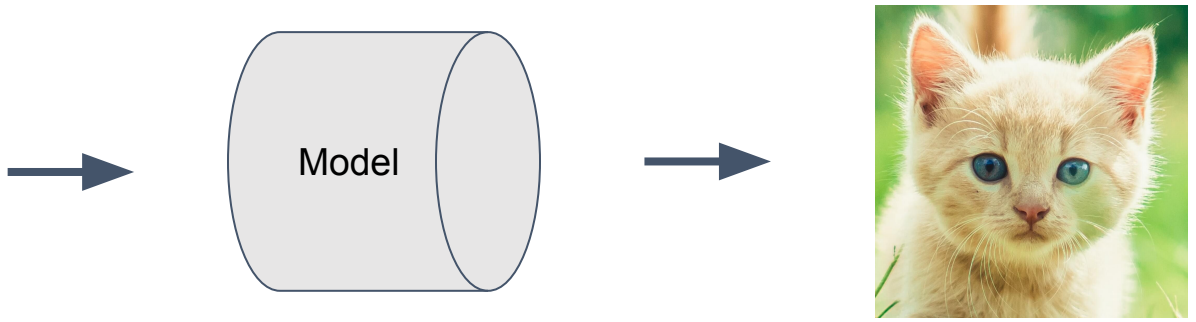
Janghun Jo Geonung Kim

Computer Graphics Lab.

POSTECH

Generative Model

[Cat Face Dataset]



[New Cat Face]

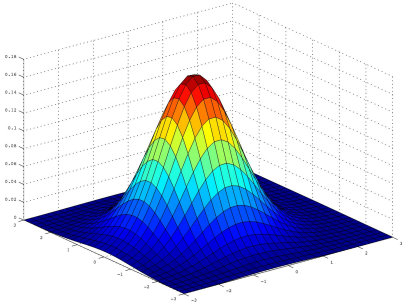
- Generative Model
  - **GAN(Generative Adversarial Networks)**
  - VAE(Variational AutoEncoder)
  - Normalizing Flow
  - DDPM(Denoising Diffusion Probabilistic Model)
  - ...

https://nvlabs.github.io/alias-free-gan/

Generative Adversarial Networks

https://nvlabs.github.io/alias-free-gan/

Generative Adversarial Networks



Normal Distribution

Sampling z

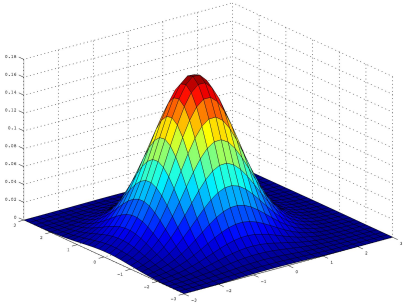
Generator

[Real Image]

Discriminator

Real/Fake
or
real인 확률

Generative Adversarial Networks



Normal Distribution

Sampling z

Generator

$$\nabla_{\theta_d} \frac{1}{m} \sum_{i=1}^{m} \left[\log D\left(\boldsymbol{x}^{(i)}\right) + \log\left(1 - D\left(G\left(\boldsymbol{z}^{(i)}\right)\right)\right) \right].$$

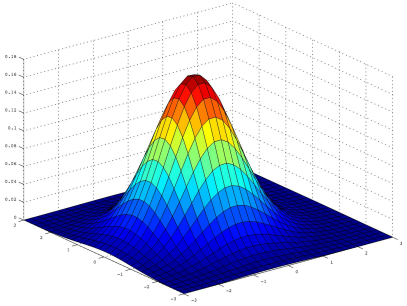
[Discriminator Training]

[Real Image]

Discriminator

Real/Fake
or
real인 확률

Generative Adversarial Networks



Normal Distribution

Sampling z

Generator

$$\nabla_{\theta_g} \frac{1}{m} \sum_{i=1}^{m} \log \left(1 - D\left(G\left(\boldsymbol{z}^{(i)}\right)\right)\right)$$

[Generator Training]

[Real Image]

Discriminator

Real/Fake
or
real인 확률

Style Transfer

[Style Image]

Image Style Transfer Using Convolutional Neural Networks

Leon A. Gatys
Centre for Integrative Neuroscience, University of Tübingen, Germany
Bernstein Center for Computational Neuroscience, Tübingen, Germany
Graduate School of Neural Information Processing, University of Tübingen, Germany
leon.gatys@bethgelab.org

Alexander S. Ecker
Centre for Integrative Neuroscience, University of Tübingen, Germany
Bernstein Center for Computational Neuroscience, Tübingen, Germany
Max Planck Institute for Biological Cybernetics, Tübingen, Germany
Baylor College of Medicine, Houston, TX, USA

Matthias Bethge
Centre for Integrative Neuroscience, University of Tübingen, Germany
Bernstein Center for Computational Neuroscience, Tübingen, Germany
Max Planck Institute for Biological Cybernetics, Tübingen, Germany

[Content Image]

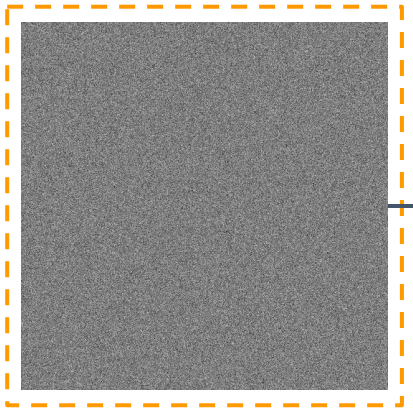
+

[Results]

Style Transfer

[Style Image]

Optimization Target



Back propagation

VGG Model
(pretrained)

[Result Image]

[Content Image]

I_s : style image
I_c : content image
I_r : result image

$$L_{total} = L_{style} + L_{content}$$

$$L_{style} = ||f_{style}(I_s) - f_{style}(I_r)||$$

$$L_{content} = ||f_{content}(I_c) - f_{content}(I_r)||$$

VGG Network

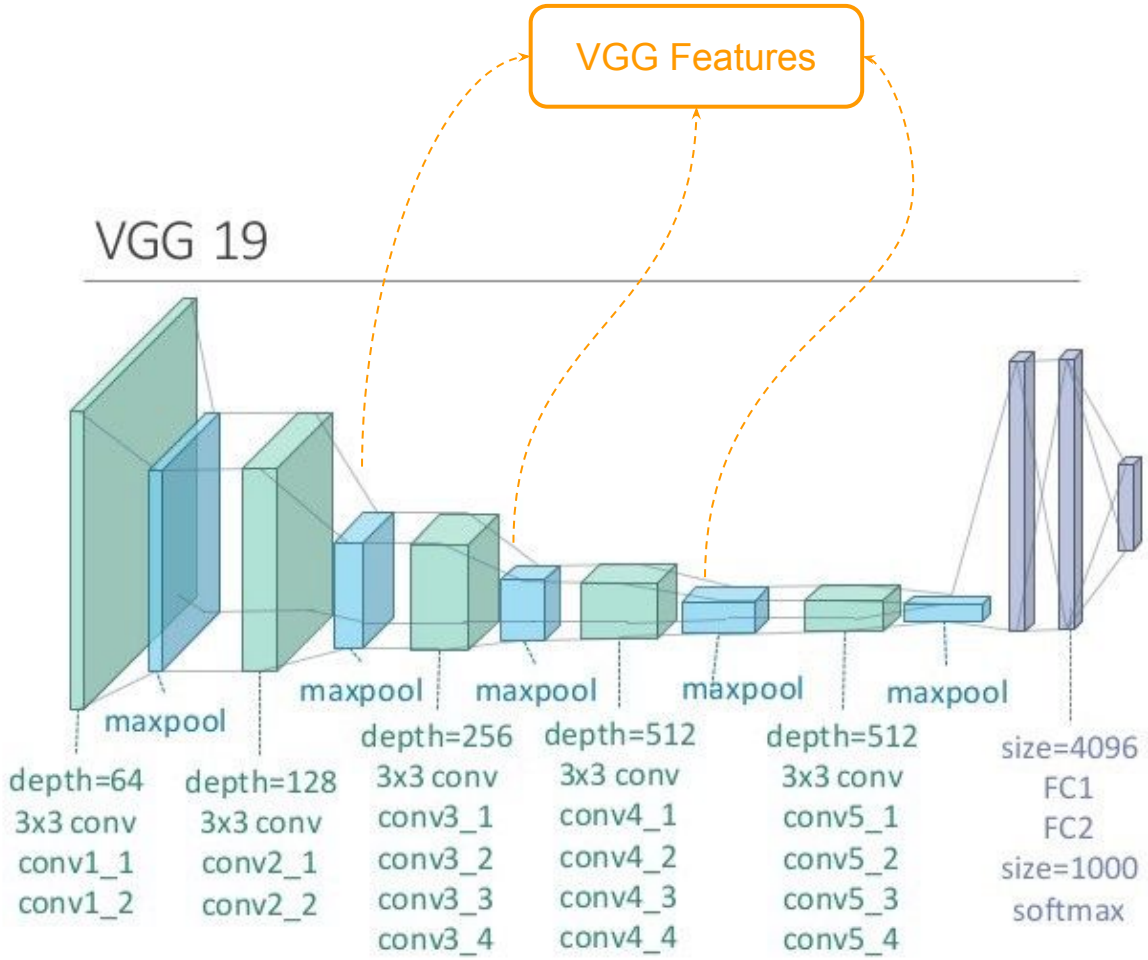
Published as a conference paper at ICLR 2015

VERY DEEP CONVOLUTIONAL NETWORKS FOR LARGE-SCALE IMAGE RECOGNITION

Karen Simonyan* **& Andrew Zisserman+**
Visual Geometry Group, Department of Engineering Science, University of Oxford
{karen,az}@robots.ox.ac.uk

ABSTRACT

In this work we investigate the effect of the convolutional network depth on its accuracy in the large-scale image recognition setting. Our main contribution is a thorough evaluation of networks of increasing depth using an architecture with very small (3×3) convolution filters, which shows that a significant improvement on the prior-art configurations can be achieved by pushing the depth to 16–19 weight layers. These findings were the basis of our ImageNet Challenge 2014 submission, where our team secured the first and the second places in the localisation and classification tracks respectively. We also show that our representations generalise well to other datasets, where they achieve state-of-the-art results. We have made our two best-performing ConvNet models publicly available to facilitate further research on the use of deep visual representations in computer vision.



VGG Network

Pixel Comparison

Input

Conv 1-1 Conv 1-2 Pooing

Conv 2-1 Conv 2-2 Pooing

Conv 3-1 Conv 3-2 Conv 3-3 Pooing

Conv 4-1 Conv 4-2 Conv 4-3 Pooing

Conv 5-1 Conv 5-2 Conv 5-3 Pooing

Feature Comparison

Input

Conv 1-1 Conv 1-2 Pooing

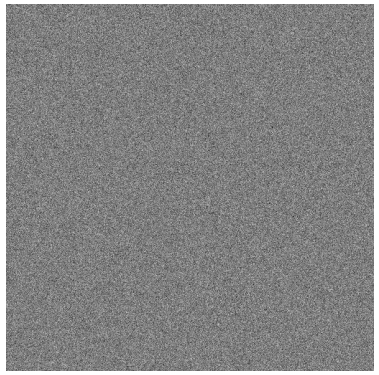
Conv 2-1 Conv 2-2 Pooing

Conv 3-1 Conv 3-2 Conv 3-3 Pooing

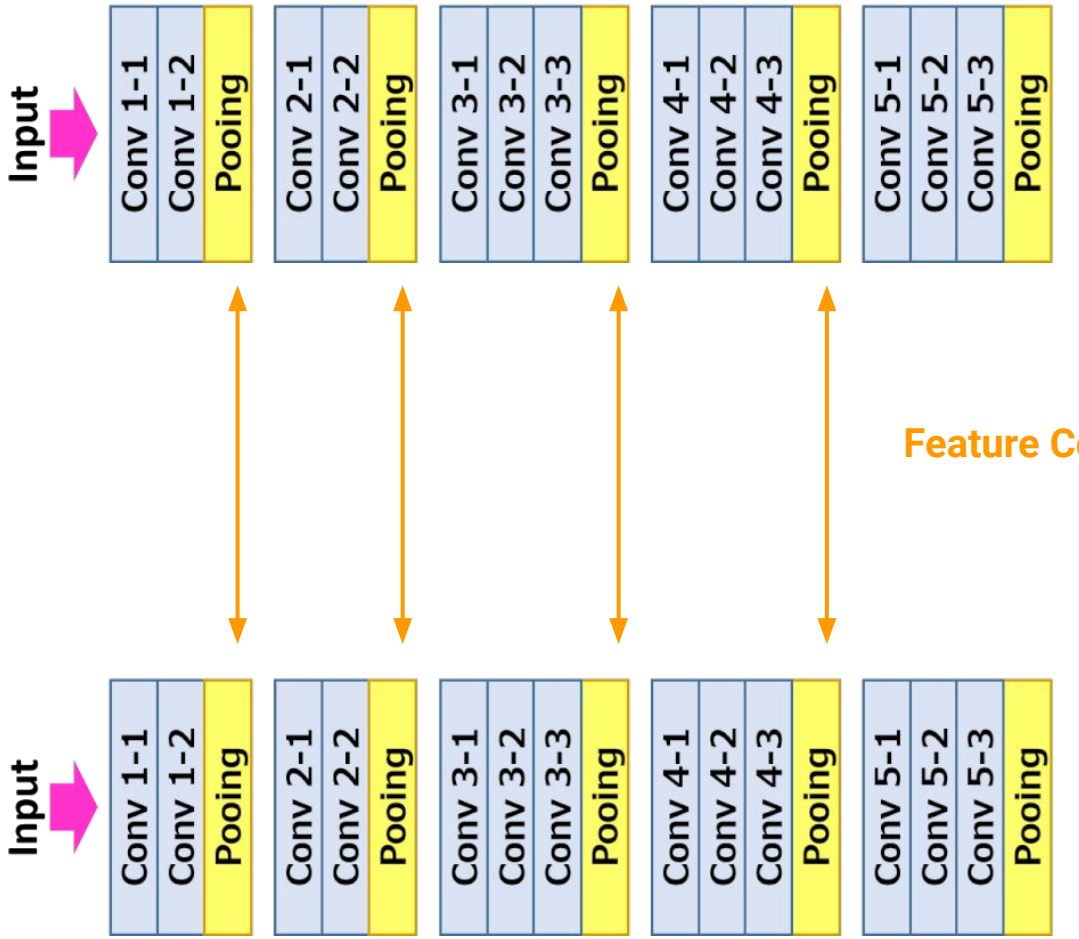
Conv 4-1 Conv 4-2 Conv 4-3 Pooing

Conv 5-1 Conv 5-2 Conv 5-3 Pooing

VGG Network

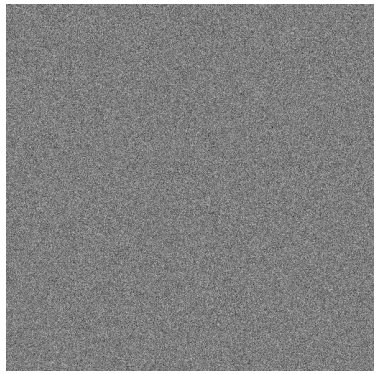


Content Loss

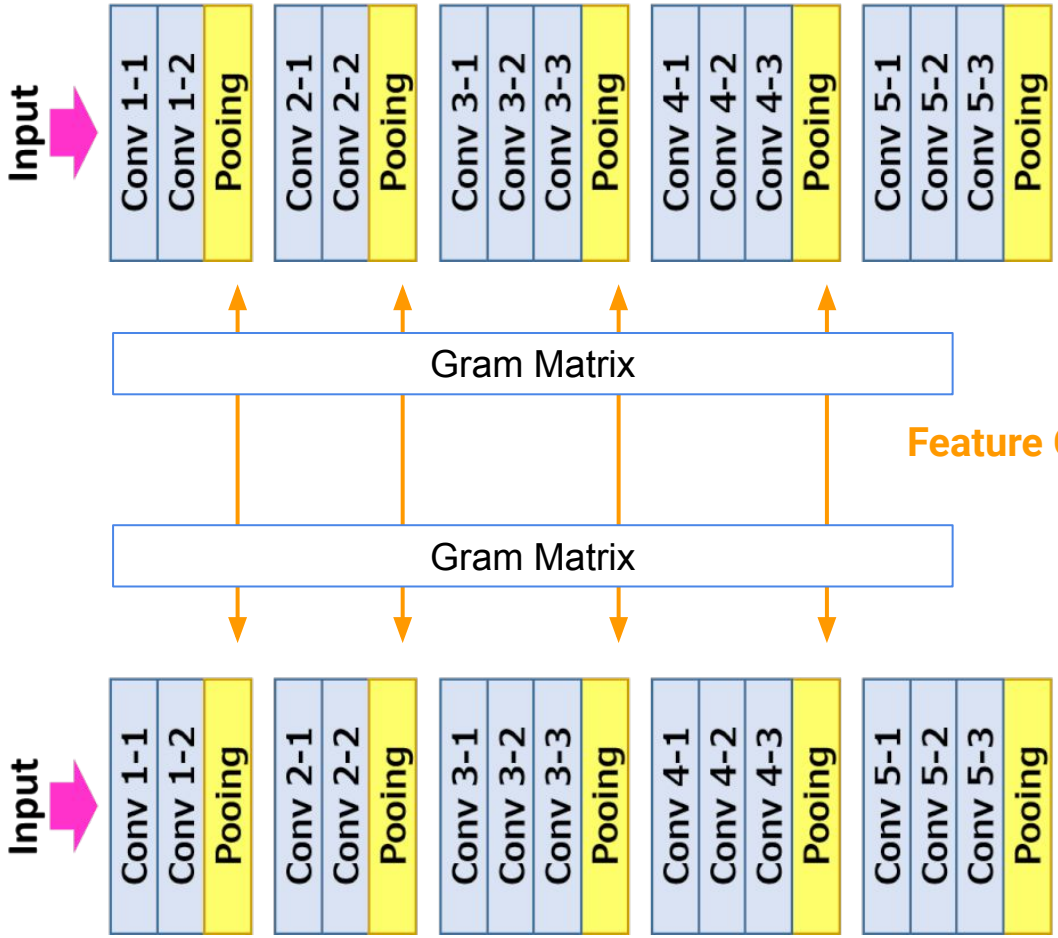


Feature Comparison

VGG Network



Style Loss



Feature Gram Matrix Comparison

Style Transfer