Development of creative AI
Neural Style Transfer (NST)

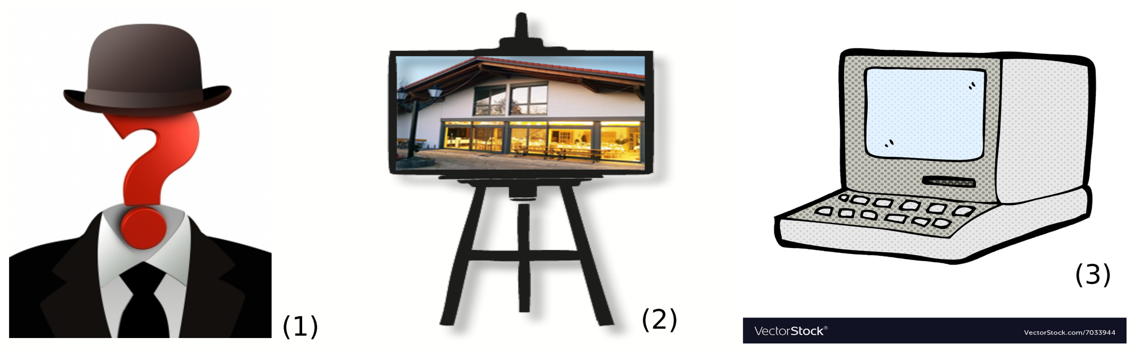
Schenk Torsten Michael, Rahimian Mashhadi Omid
s63550@beuth-hochschule.de, s74304@beuth-hochschule.de

BEUTH HOCHSCHULE FÜR TECHNIK BERLIN
University of Applied Sciences

Learning from Images WS2019
Prof. Dr. Hildebrand

Last resort creative AI

In this project we want to present the development from **deep dream** towards **NST**. Creative AI is one of the wonderful byproducts of deep neural networks. We show cost functions and visualize intermediate steps during reconstruction of the result image. In recent past the measure of intelligent behavior was based on turing tests or playing board games. Basically skipping the act of creating art.



Rumors that human intelligence is unique, still prevail today. Though this might change soon!

Motivation

The psychedelic images produced by deep dream catched our interest. Two main ap-proaches exist to merge the activation of high layers with a content image. Deep dream just uses the activation from a pretrained network and dependent on this network the results are somewhat preprogrammed. An InceptionV3 model with weights from imagenet has a over-representation of dogs and cats and is therefore prone to merge eyes and hairs into images. On NST the images can be choosen freely.

Preparation of Data

As input we use 3 channel RGB 8 bit color images. The content images show familiar known objects like houses, humans, food or outdoor scenes. The style images for NST should contain strong pattern, colors, painting styles or abstract art for best results. In case of controlled dreams a content image can be used as guide as optimization objective.

Content (4)

Style (5)

References

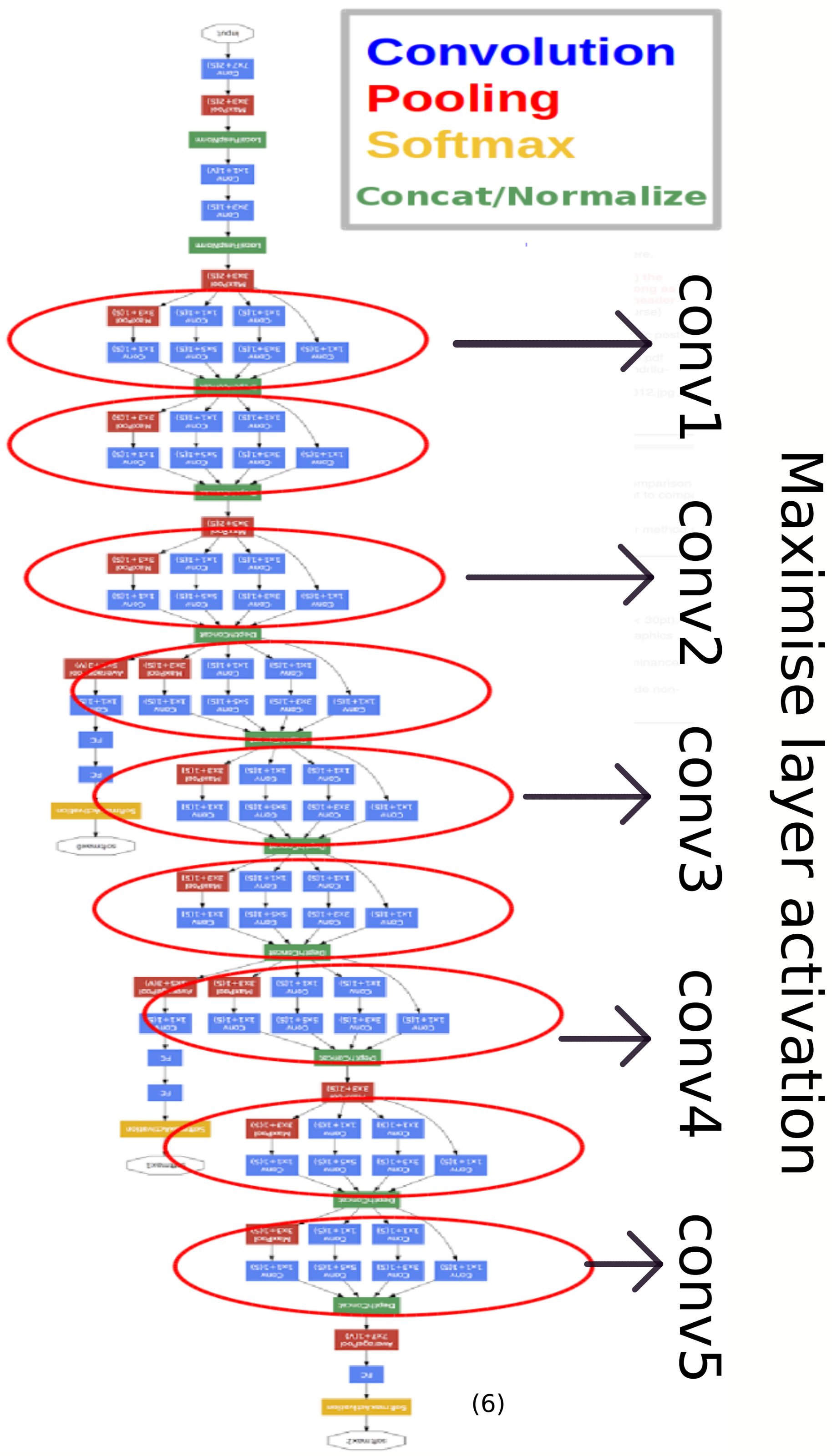
- o Deep Learning mit Python and Keras: Francois Chollet, mitp
- o Python Machine Learning, Sebastian Raschka & Vahid Mirjalili, Packt
- o Image Style Transfer Using Convolutional Neural Networks, Leon A.Gatys e.Al.
- o https://github.com/google/deepdream
- o Neural Style Transfer: A Review, Yongcheng Jing, e.Al.
- o Going deeper with convolutions, Christian Szeedy, e.Al (GoogleLeNet)

(1) www.selfavenue.com, (2) www.pinclipart.com, (3) www.vectorstock.com, (4) www.suedkurier.de, (5) www.amazon.de, (6) Going Deeper with convolutions: Christian Szegedy, (7) www.cs.toronto.edu, (8) Image Style Transfer Using Convolutional Neural Networks: Leon A. Gatys

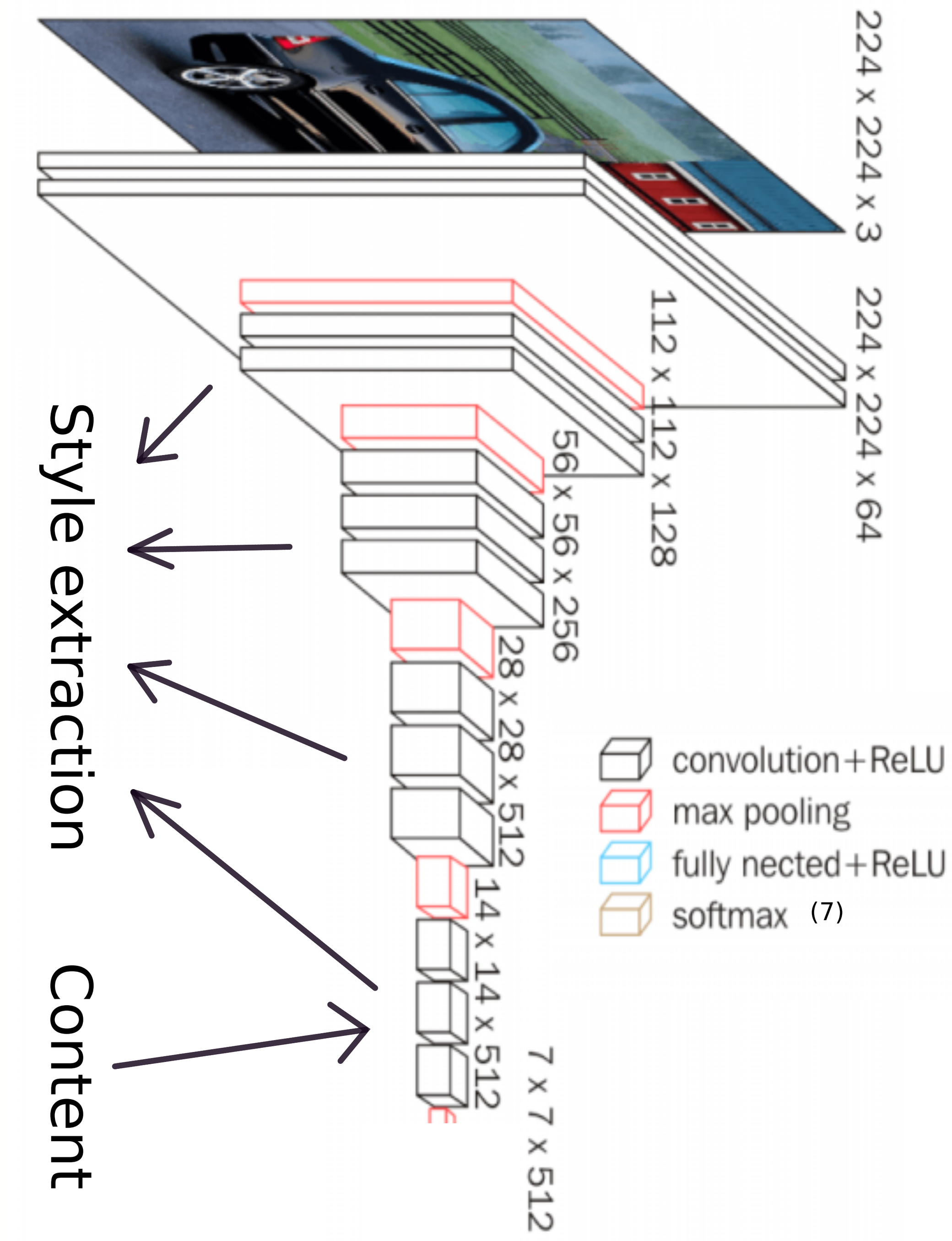
NN for deep dream and NST

Both networks used for creative AI are very deep neural networks. The deep dream algorithmus (e.g.GoogleLeNet) makes use of any layer, of which the activation is maximised. For NST the network has to be used for the style, content and artificial image to be created. A set of convolution layers is used to create gram matrices. Gram matrices reflect the similarity between features. Softmax or fully connected layers are not used.

Inception, deep dream



VGG19 Net, NST

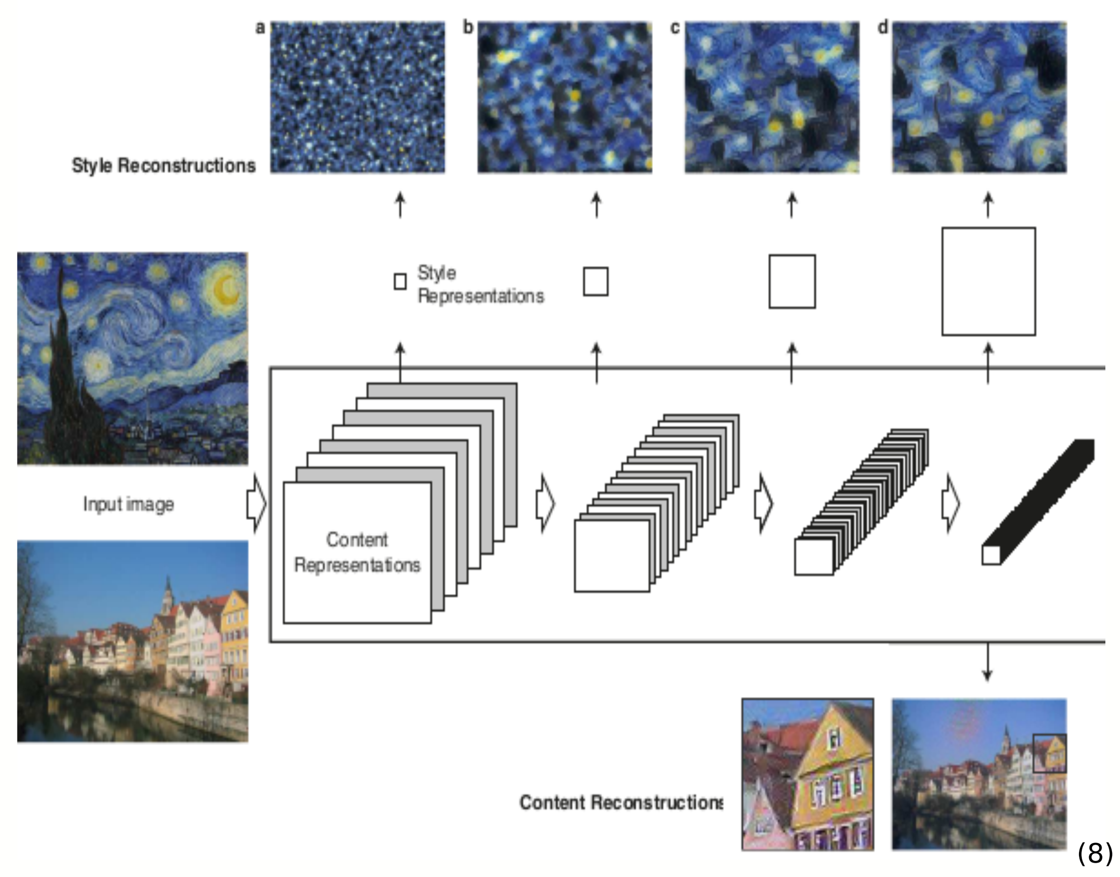


Layer activation

For creating impressive deepdreams lower layers show patterns, higher layers activation like dog, human or bird faces. All trained activations can be visualized.

Style & Content

For NST we will extract the style from various layers and minimise the sum of the style loss beta. The content is introduced in a high layer and only updated via the loss function alpha.



Loss functions for deep dream and NST

There are two loss functions, one is the content loss and other one is the style loss function. The content loss function ensures that the activations of the higher layers are similar between the content image and the generated image.

$$\mathcal{L}_{content} = \sum \left(F^l - P^l\right)^2$$

The style loss function makes sure that the correlation of activations in all the layers are similar between the style image and the generated image.

$$\mathcal{L}_{style} = \sum_l w_l E_l$$

The total loss function is:

$$\mathcal{L}_{total} = \alpha \mathcal{L}_{content} + \beta \mathcal{L}_{style}$$

where α and β are user-defined hyperparameters and are the weighting factor for content and style recunstruction. By controlling α and β you can control the amount of content and style injected to the generated image.