CS60050 : MACHINE LEARNING TERM PROJECT

Artistic rendering of images

ArtificiallyArtistic

Team Members:
Ashrujit Ghoshal(14CS10060)
Sayan Ghosh(14CS10061)
Arundhati Banerjee(14CS30043)
Sayan Mandal(14CS30032)
Mousam Roy(14CS30019)
Sourav Pal(14CS10062)
Sohan Patro(14CS30044)
Projjal Chanda(14CS10057)
Pradeep Dogga(14CS10013)
Aniket Suri(14CS10004)

Primary papers :

1.*A Neural Algorithm of Artistic Style:* By:Leon A. Gatys,Alexander S. Ecker,Matthias Bethge.

This paper introduces an artificial system based on a Deep Neural Network that creates artistic images of high perceptual quality. The system uses neural representations to separate and recombine content and style of images, providing a neural algorithm for the artistic rendering of such images in a variety of artistic styles that the network has been trained in.

2.*Separating Style and Content on a Nonlinear Manifold* By:Ahmed Elgammal and Chan-Su Lee

This paper addresses how to separate style and content on manifolds representing dynamic objects. A decomposable generative model that explicitly decomposes the intrinsic body configuration (content) as a function of time from the appearance (style) of the person performing the action as time-invariant parameter is presented in the paper.

3.*Semantic Style Transfer and Turning Two-Bit Doodles into Fine Artwork*:By Alex J. Champandard

This paper introduces a novel concept to augment CNNs with semantic annotations, either by manually authoring pixel labels or using existing solutions for semantic segmentation. The result is a content-aware generative algorithm that offers meaningful control over the outcome. Applications include semantic style transfer and turning doodles with few colors into masterful paintings.

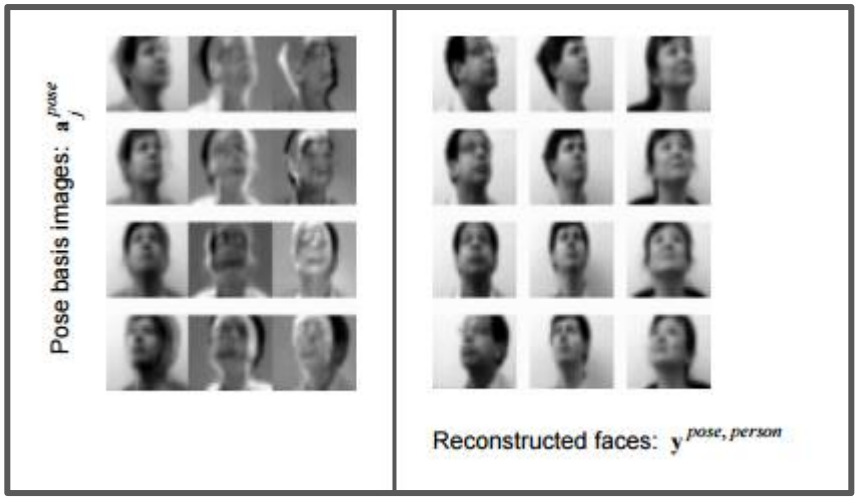
4.*Artistic style transfer for videos*:By Manuel Ruder, Alexey Dosovitskiy, Thomas Brox

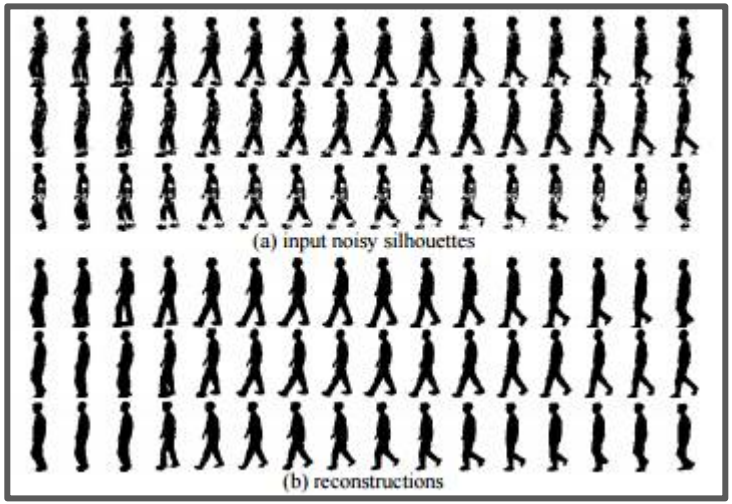
This paper makes use of recent advances in style transfer in still images and proposes new initializations and loss functions applicable to videos

Current practice :

There have recently been a lot of interesting contributions to the issue of style transfer using deep neural networks. Gatys et al. proposed a novel approach using neural networks to capture the style of artistic images and transfer it to real world photographs. Their approach uses high-level feature representations of the images from hidden layers of the VGG convolutional network to separate and reassemble content and style.


Interpolated smiles for four different people
Interpolated smiles at intermediate (new) people styles.


Pose basis images: $\mathbf{a}_j^{pose}$
Reconstructed faces: $\mathbf{y}^{pose,\,person}$


(a) input noisy silhouettes
(b) reconstructions

Convolutional Neural Networks (CNN)

It is a type of feed forward ANN in which the connectivity between its neurons is inspired by the organization of the animal visual cortex, whose individual neurons are arranged in such a way that they respond to overlapping regions tiling the visual field.

When used for image recognition and image processing CNNs consist of multiple layers of small neuron collections which process a small part of the image.After processing each regions the output from all these layers are merged or processed together which gives us a better representation of the image.

Advantages : 3d arrangement of neurons , local connectivity , use of shared weight in convolutional layers, thus using the same filter for each pixel thus optimizing both space and time .

Use of CNN in our work :

Following the work of Gatys we will be using the VGG network , a CNN network, which uses 16 convolutional and 5 pooling layers.(The VGG-19 concept ranked 2nd in the ImageNet 2014 object recognition challenge)

The input image is encoded in each layer of the CNN by the filter responses to that image. To visualize the image information at each level we use gradient descent on a white noise image to find another image that matches the feature of this image . We then define the squared error loss between the images from which we calculate the gradient with respect to image using standard techniques of back propagation.

Tools needed :

Language : Python

Libraries :

1. Theano
2. Scikit-learn
3. TensorFlow

Dataset source :

These are collections of paintings by different artists which could be used as training data for artificial image rendering:

1. WikiArt

2. Art Resource

3. Princeton University Datasets

Roadmap :

1. September : We plan to prepare two implementations of the VGG convNet model, one using Google's TensorFlow, and the other using Theano+Scikit-learn. The model will be trained to be able to separate content from style in images, and apply a chosen style(or combination of styles) to a provided image. The output will be the stylised image.

   Note: We intend to extend the problem of style transfer in images to doodles as well.

2. October: After analysing the performance and efficiency of these implementations with regard to stylisation of images, the models will be extended to incorporate style transfer in videos. In this regard, the added constraint of optical and temporal coherence will be taken care of.

Stretch goal

- November: We shall try to extend style transfer techniques to apply it to to construct 3D images from the artistically rendered 2D images, using techniques of space filling.

Reference:

StereoBrush: Interactive 2D to 3D Conversion Using Discontinuous Warps

By: O. Wang, M. Lang, M. Frei, A. Hornung, A. Smolic, M. Gross

PROGRESS MADE

- Reviewed the primary papers
- Consulted existing open-source implementations of the separation of style and context technique proposed by Gatys
    - Discussed about the different tools that can be used - Theano, ScikitLearn and TensorFlow among team members.
    - Discussed about the various CNN model implementations available - especially the Caffe models
- Started implementation of convolutional neural network in python