

Image to image synthesis of face sketches to realistic images of faces

Jason Li, Joe Suzuki, Anthony Sokolov

Description:

Our algorithmic project will aim to train a model capable of synthesizing realistic faces from outline sketches. Such a model will be useful in applications of generative animations and rendering, and can also be used to augment and expand existing datasets by creating new training instances. We will explore several different approaches to this problem, guided by the state of the art in research. Our approaches will focus on various GAN architectures due to the fact that this type of model has seen success in the task of image translation in recent years, and is the most common algorithm we saw when reviewing current research.

Related work / sources:

Image-to-Image Translation with Conditional Adversarial Networks

This paper pioneers the use of conditional GANs in image translation tasks. We plan to explore the architecture outlined in this paper and tune the hyperparameters to maximize performance on sketch to face translations. We will also use the FCN scores from this paper as baselines for our performance.

Unpaired Image-to-Image Translation using Cycle-Consistent Adversarial Networks

We will use this paper as a reference for our semi-supervised approach. This paper outlines a GAN that can be used for image translation without explicit pairs in the training set. Our training data for such an approach would consist of a set of sketches and a set of faces, but there would be no pair mapping between the sets.

Scribbler: Controlling Deep Image Synthesis with Sketch and Color

While there are many image-to-image translation methods out there, Scribbler allows user input of colors that gives users more control over the output. We plan to build on this approach and explore options of automatically choosing color rather than prompting the user.

Model implementations

We plan to use sci-kit learn, TensorFlow, and Keras to implement our algorithms.

We plan to use OpenCV for preprocessing operations on our dataset

Github link

<https://github.com/AnthonySokolov/sketch-to-face-translation>

Enumerate the CSCI 353/795 related topics (questions and solution approaches)

each individual will contribute to the project.

Generative/discriminative learning - The GAN models we will explore feature generative and discriminative networks trained in tandem

Neural networks - We will explore GAN models that are based on artificial neural nets

Supervised / semi supervised learning - We will explore various approaches using both paired and unpaired images for training data

Data cleaning and augmentation - We will need to process our dataset to generate facial outlines. Our models can also be used to augment data and create training instances for image classifiers.

Team member's role and contribution

Anthony Sokolov:

Use the canny edge detection algorithm to generate facial outlines from the dataset of facial images.

Download the weights for the image classifier to be used for FCN loss

Implement model based on conditional GAN architecture (supervised)

Implement model based on cycle consistent GAN architecture (semi supervised)

Joe Suzuki:

Implement PatchGAN discriminator model and U-Net encoder-decoder generator model

Implement L1 and L1 + conditional adversarial loss for evaluation + explore combination of loss functions

Jason Li:

Work on PatchGAN discriminator for conditional GAN model

Implement model based on scribbler architecture

Implement model based on cycle consistent GAN paper

Dataset:

We will use a compiled dataset of facial images. We have identified multiple data sources that we plan to compile into one large dataset. We will train our network to be able to generate images of the target class (facial images). We will then test the model by inputting sketches and applying a loss function based on how well the output image fits the target class. We wish to generate pairs of facial sketches and the images that they correspond with. To do this we will apply an edge detection algorithm to our dataset of facial images to create corresponding facial outlines.

Links to data sources:

<http://mmlab.ie.cuhk.edu.hk/projects/CelebA.html>

<https://www.kaggle.com/kpvisionlab/tufts-face-database>

Timeline:**Week of 11/2:**

- Download and compile dataset of facial images
- Generate dataset of facial outlines to be used for model training
- Download weights for the model that will be used for FCN score
- Begin work on conditional GAN model

Week of 11/9:

- Continue work on conditional GAN model
- identify the optimal architecture and hyperparameters of cGAN model
- Save weights of optimal model and record FCN score metrics
- Begin work on scribbler based model

Week of 11/16:

- Continue working on second model
- Experiment with approaches to train a network to add color

Week of 11/23:

- Begin work on cycle consistent GAN model

Week of 11/30:

- Prepare the deliverables, start working on presentation
- Finish evaluations of approach
- Work on documentation

Week of 12/7:

- Complete documentation, presentation and demonstration

Presentation Day (12/15)

Compile all work together and present

Demonstration

We plan to demo our model performing on various unseen examples. We will focus on our best performing model but also show how the various models we worked on differ. We will also record a video of our model's performance on various examples in addition to a paper and small poster. Our deliverables will be the weights for all of our models, as well as the code and notebooks used in their training.

Evaluation

We will compare all of our models using the same FCN score metric. We will test how well a pre-trained facial image classifier performs on the images synthesized by our models. Because the evaluation of generative models is still an open problem, we will also take our own perception of the model's realism into account.

Citations

Phillip Isola, Jun-Yan Zhu, Tinghui Zhou, Alexei A. Efros "Image-to-Image Translation with Conditional Adversarial Networks" CVPR, 2017

Jun-Yan Zhu, Taesung Park, Phillip Isola, and Alexei A. Efros. "Unpaired Image-to-Image Translation using Cycle-Consistent Adversarial Networks", in IEEE International Conference on Computer Vision (ICCV), 2017.

Patsorn Sangkloy, Jingwan Lu, Chen Fang, Fisher Yu, James Hays "Scribbler: Controlling Deep Image Synthesis with Sketch and Color" CVPR 2017