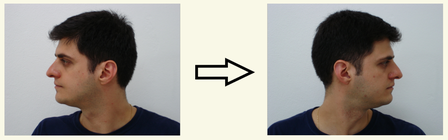
Pose Manipulation

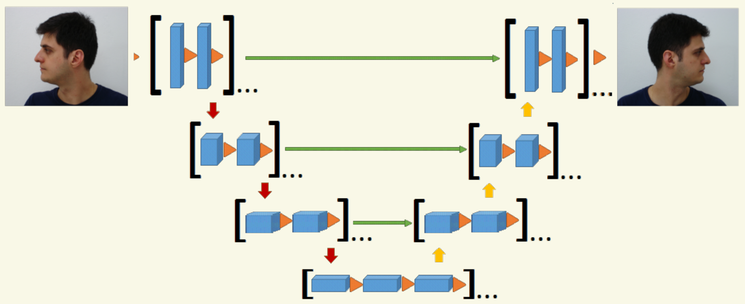
Overview

- Change style or pose of object.



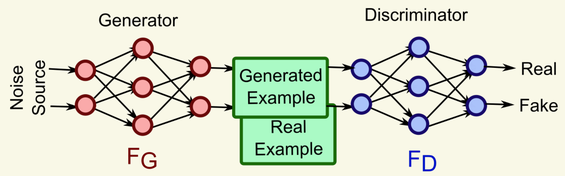
- Paper
  1. Pose Guided Person Image Generation
     https://arxiv.org/abs/1705.09368

U-Network



- Fully convolutional (3x3 filters)
- Increase/Decrease channel size $\Longleftrightarrow$ Decrease/Increase image size
- Upsamling
  - Transposed convolution
  - Interpolation

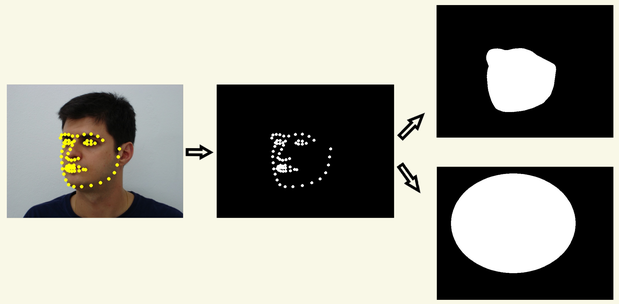
Generative Adversarial Networks



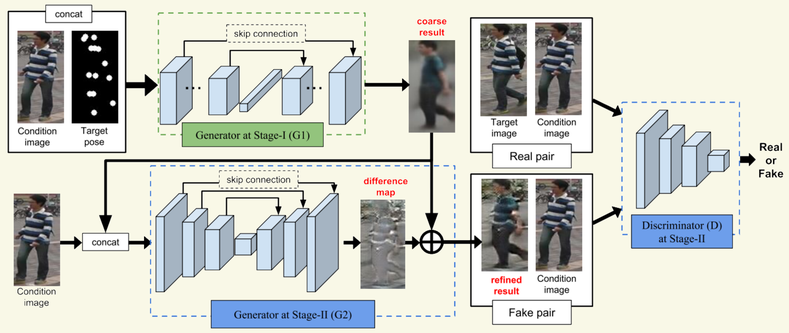
- Generator
  $\Longrightarrow$ Try to trick Discriminator
- Discriminator
  $\Longrightarrow$ Try to distinguish

Pose Guided Person Image Generation

- Pose Mask and Embedding



- Architecture



- Generator G1
  - Goal: Try to generate coarse image with new pose
  - Input: Person (RGB image), Landmarks heatmap(68 channel image)
  - Output: Coarse pose changed person (RGB image)
  - Loss:

$$L_{G1} = ||(G1(I_S, P_T) - I_T) \odot (1 + M_T)||_1$$

- Generator 2
  - Goal: Try to refine image of G1
  - Input: Person (RGB image)
  - Output: Difference map (RGB image)
  - Loss:

$$L_{G2} = L_{BCE}(D(I_S, I_{G2}), 1) + \lambda||(I_{G2} - I_T) \odot (1 + M_T)||_1$$

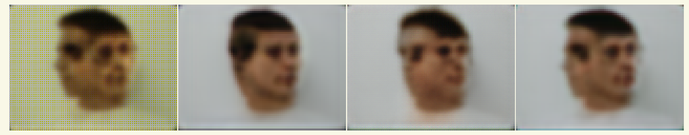
- Discriminator
  - Goal: Tries to distinguish between real and fake pairs
  - Input: Person (RGB image), Person with target pose (RGB image)
  - Output: 1 or 0
  - Loss:

$$L_D = L_{BCE}(D(I_S, I_T), 1) + L_{BCE}(D(I_S, I_{G2}), 0)$$

- Training
  - Train only G1
  - Fix G1
  - Train G2 and D

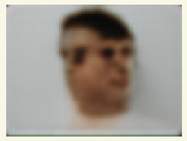
Decisions

- Testing of Upsampling and Masking

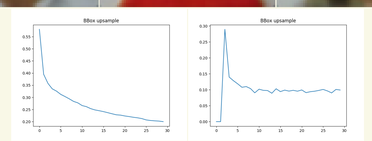


$\Longrightarrow$ BBox and Interpolation.

- Excluded borders from image I_{G2}.
- Appending also source landmarks due to:



Final Results



SSIM(I_S, I_T)	SSIM(I_T, I_{G1})	SSIM(I_T, I_{G2})	RGB-L1-Distance
1.0	0.882	0.877	0.024
0.552	0.730	0.721	0.046

Jan Sieber
uni@email-master.de
Ruprecht-Karls-University Heidelberg