

Facial attribute editing using neural networks

Evgeny Zholkovskiy

Research advisor: Ivan Oseledets

MSc Data Science

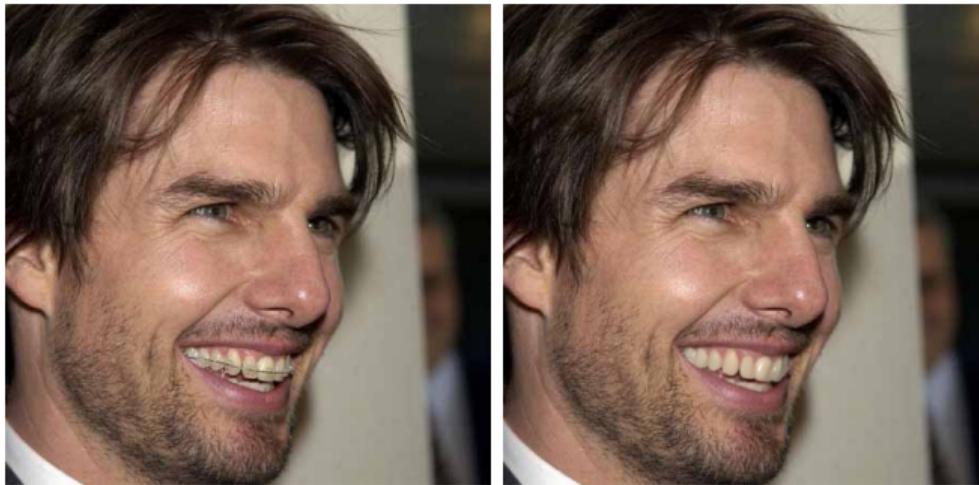
May, 2021

Introduction

- ▶ Photo editing is now widely used in entertainment, media, medicine and many other areas.
- ▶ Automated image editing with desired properties can simplify routine work and make editing available for everyone.
- ▶ Face editing is challenging because even small defects affects human perception very much.

Problem

Facial attributes editing, in particular dental braces removing from a photo is considered in this work.



Overview

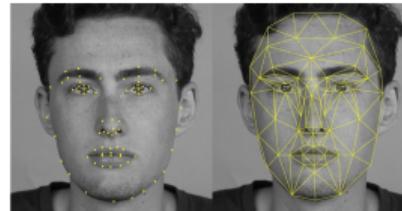
Generative networks



Segmentation



Facial landmarks



Approaches

1. Segmentation + Inpainting

- ▶ Pros: Less artifacts
- ▶ Cons: Expensive dataset

2. Generative Models

- ▶ Pros: Dataset can be obtained automatically
- ▶ Cons: Image artifacts

Objectives

1. Obtain a dataset for the problem
2. Build an image processing pipeline

Dataset



Dataset

Semi-automatic cleaning process

- ▶ Parse retrieval systems for braces
- ▶ Process Flickr-Faces-HQ Dataset

Obtained dataset

- ▶ 14k images without braces
- ▶ 1.5k images with braces

Attention-Guided GAN

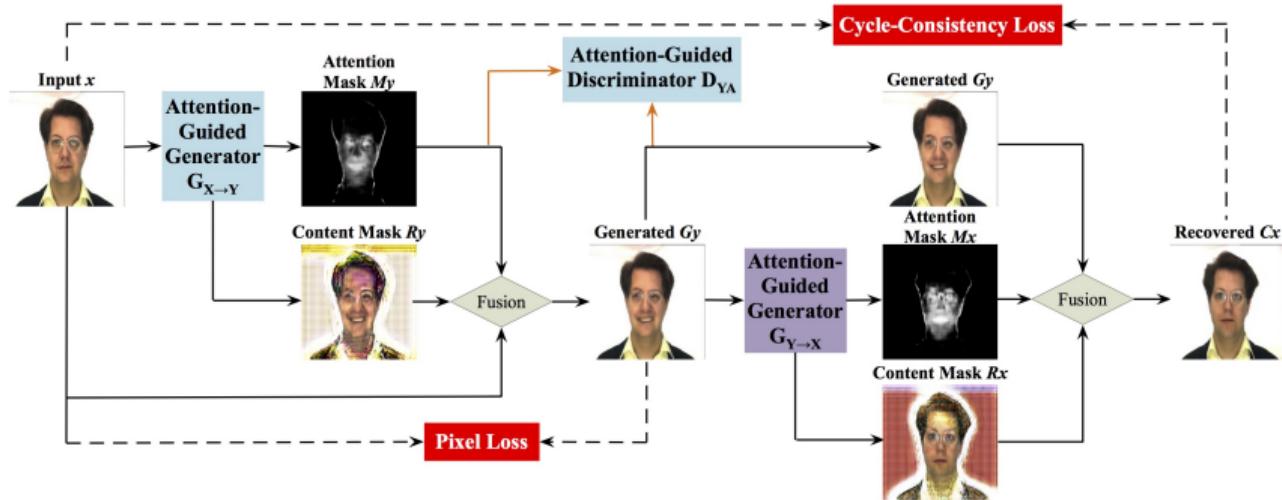
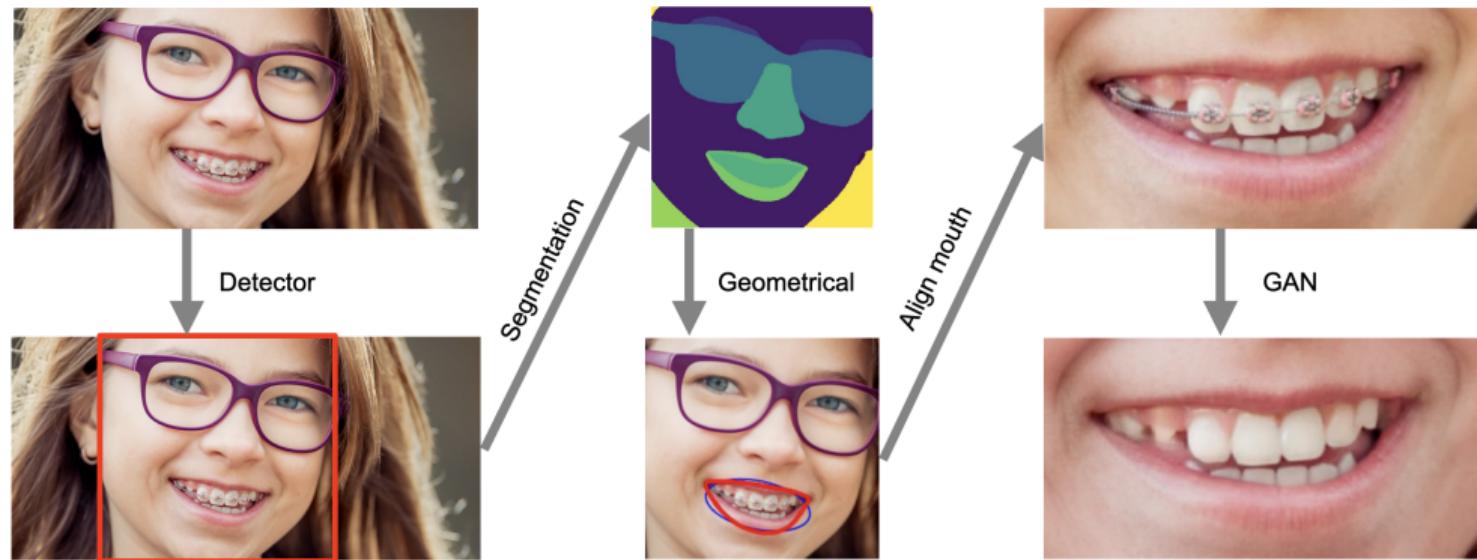


Image Processing Pipeline



Color Correction

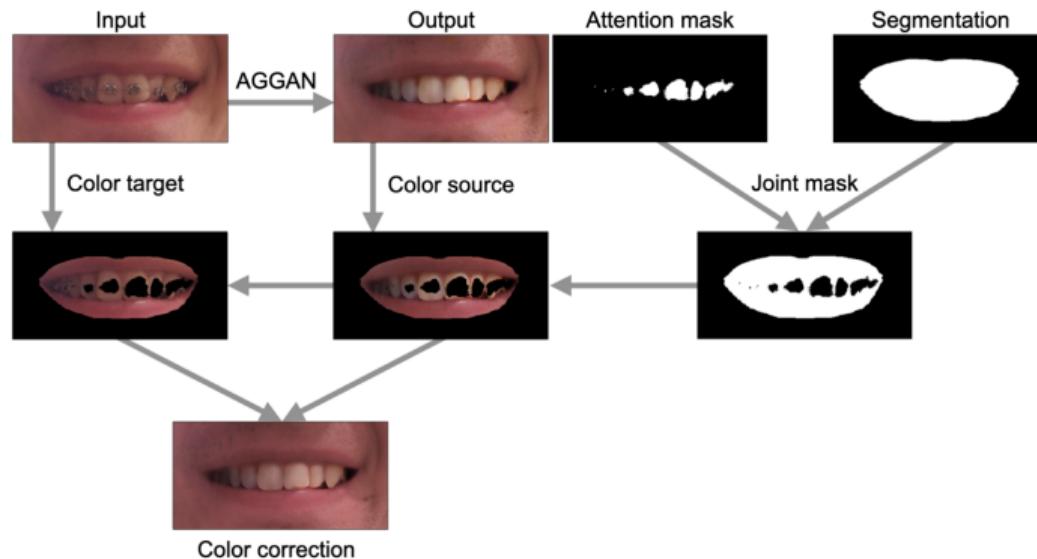


Image processing pipeline

1. Face detection
 - ▶ MTCNN, BCE + MSE
 - ▶ FDDB dataset (3k images, bboxes)
2. Face Segmentation
 - ▶ BiSeNet, BCE
 - ▶ CelebAMask-HQ dataset (30k images, segmentation)
3. Mouth align
4. Removing braces
 - ▶ Attention-Guided GAN, Adversarial, Cycle, Consistency
5. Wrap result

Inference time

	Core i7, 2.6 GHz	GeForce RTX 2080 Ti
Face detection	204 ms	138 ms
Face segmentation	670 ms	65 ms
Mouth alignment		165 ms
GAN	1070 ms	23 ms
Colour correction		18 ms
Wrap result		31 ms
Total	2158 ms	440 ms

Results - Examples



Results - Examples



Results - Examples



Results - Examples



Evgeny Zhokovskiy
Facial attribute editing using neural networks

Results - Closer look



Results - Problems

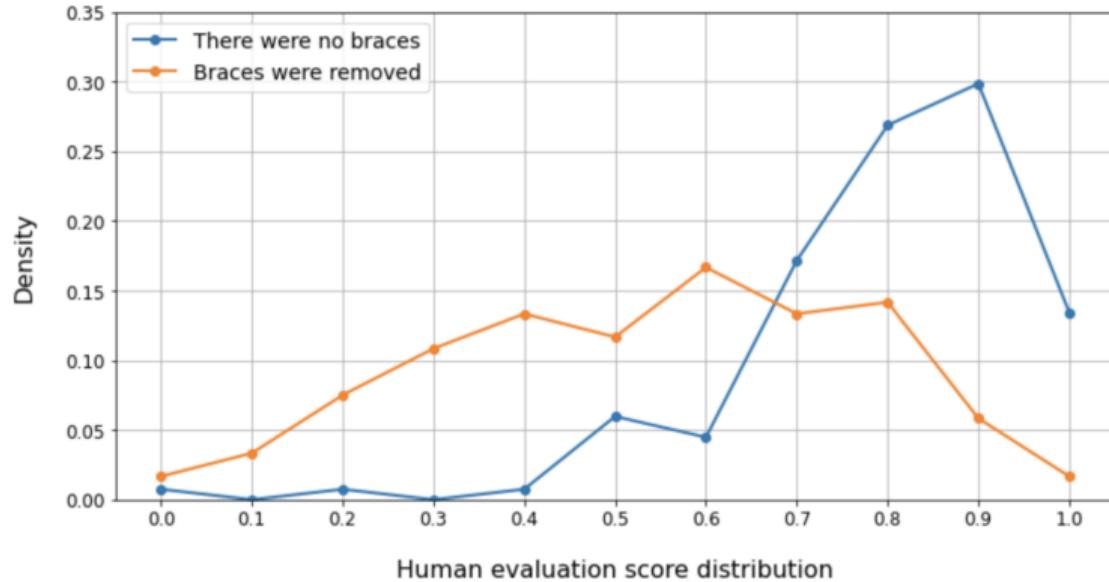
- ▶ Noise and artefacts
- ▶ Incomplete braces removal
- ▶ Image tone change



Evaluation - Baseline



Human Evaluation



Human Evaluation

- ▶ In 54% of cases, people failed to determine that braces were on the original image.
- ▶ In 20% of cases images that were not edited were considered as edited.

Braces Generation - Examples



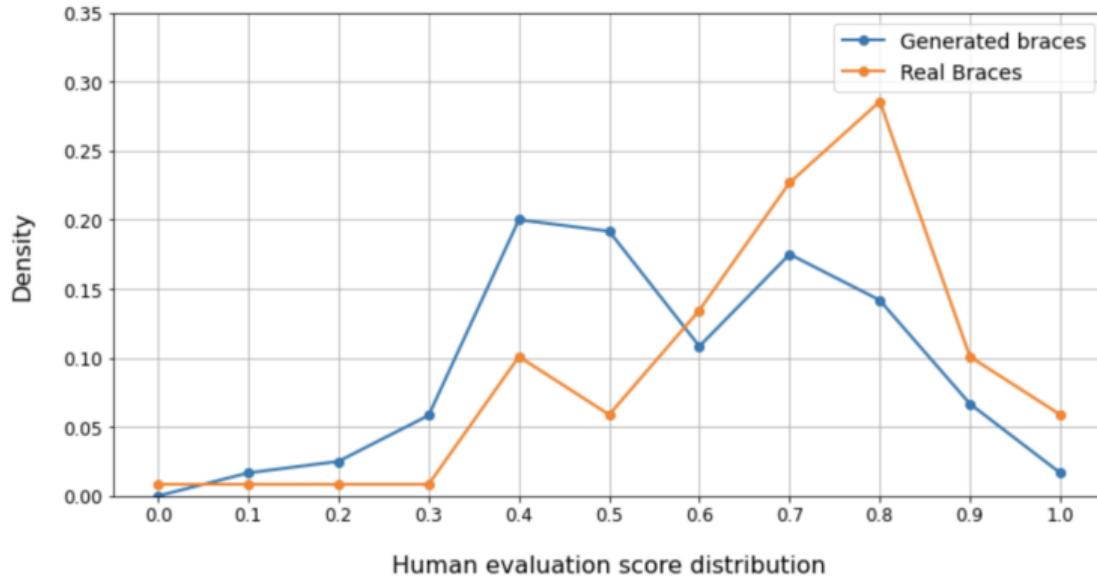
Braces Generation - Examples



Braces Generation - Examples



Human Evaluation - Generation



Human Evaluation - Generation

- ▶ In 58% of cases, people determined that braces were generated.
- ▶ In 31% of cases real braces were considered as generated.

Outcomes and Applications

- ▶ The new large dataset was obtained
- ▶ The pipeline that was built allows to easily edit facial attributes on photos
- ▶ The service is available as a Telegram bot

Applications:

- ▶ Photo editing
- ▶ Medicine and cosmetology
- ▶ Entertainment

Conclusion

- ▶ Having a good dataset as much important as training a machine learning model. Obtaining appropriate dataset is a challenging task.
- ▶ A pipeline of computer vision models for face detection, alignment and photo editing was built.

Outlook and applications

1. Obtain more data and improve quality
2. Framework may be adopted for another tasks
3. Service is planed to be tested publicly

Thank you!