Perceptual test on StyleGAN3

Beozzo Emanuele
Xompero Leonardo

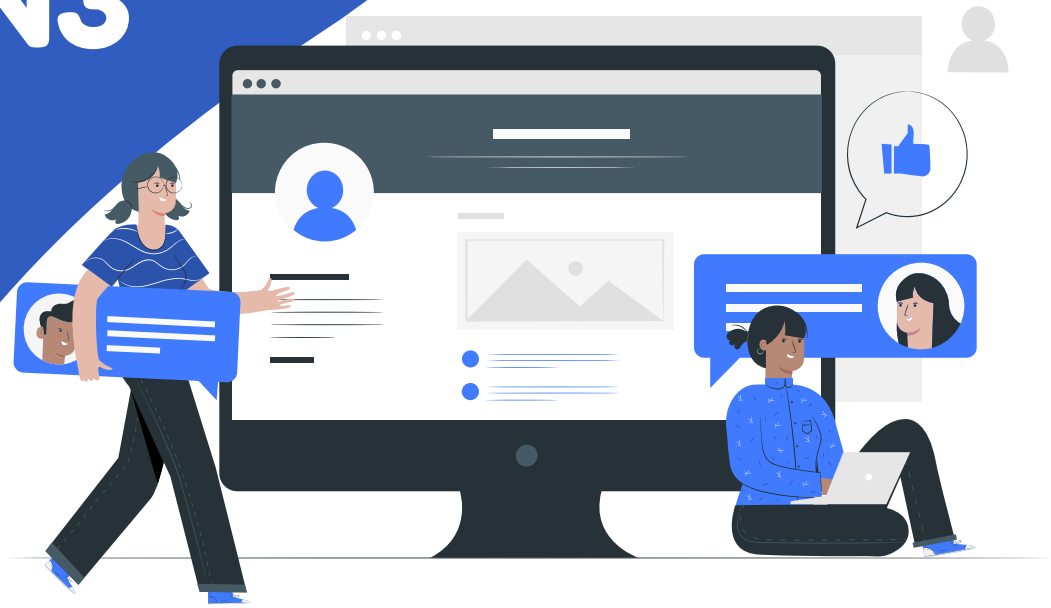
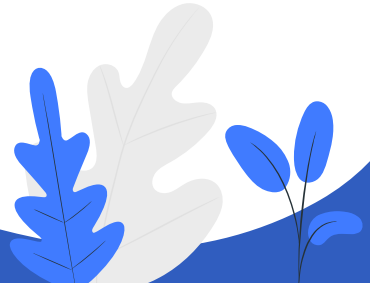


Table of contents

1 Introduction

Study of the previous paper and StyleGAN3 material



Previous experiment

- “More Real than Real: A Study on Human Visual Perception of Synthetic Faces”
- Provides quantitative evidence on how the quality and realism of face images generated with cutting-edge AIs makes it hard for human viewers to recognize them as synthetic
- Three GAN networks were tested and compared: PGGAN (AI-17), StyleGAN (AI-18) and StyleGAN2 (AI-19)

StyleGAN3

Generates state of the art results for un-aligned datasets and looks much more natural in motion (good for video and animation)

https://lambdalabs.com/blog/stylegan-3/

2 Dataset

Selection of image for the dataset

Datasets

FFHQ
(Flickr-Faces-HQ)

The dataset consists of 70,000 high-quality PNG images at 1024×1024 resolution and contains considerable variation in terms of age, ethnicity and image background.

StyleGAN3 - R

Example of images obtained from StyleGAN3 with config R (translation and rotation equiv.)

Selection criteria

We selected the images from each dataset with the following criteria:

- 30 images from FFHQ and 30 images from StyleGAN3-R
  - 15 females and 15 males each
- Caucasian (since most datasets are still unbalanced toward this ethnicity)
- Age in the range 20-50 years
- High-quality images
  - Not blurry images
  - No contrasts or strange illumination
  - Frontal looks
  - Without reflections on the eyes

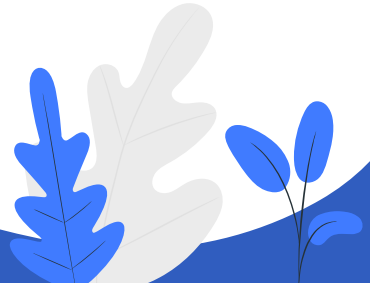
Examples

FFHQ

StyleGAN3

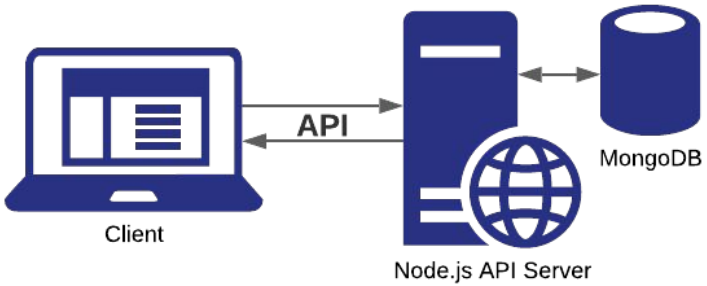
3 Interface

Implementation of the web page

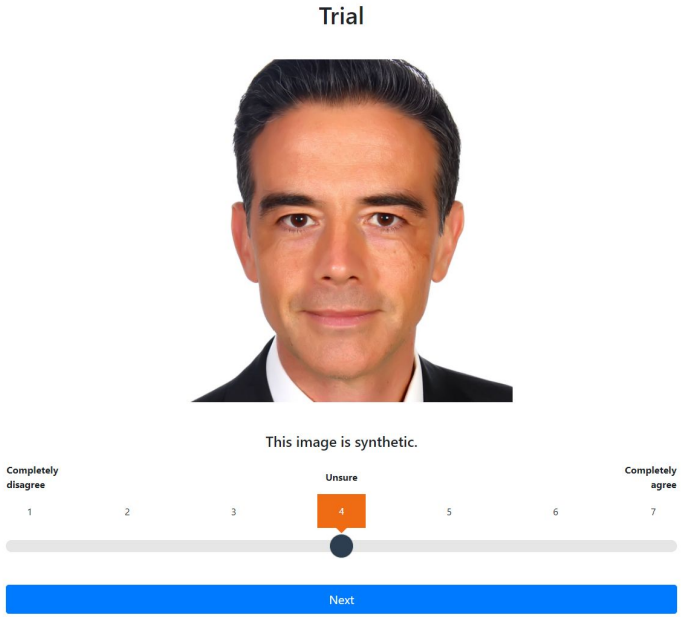


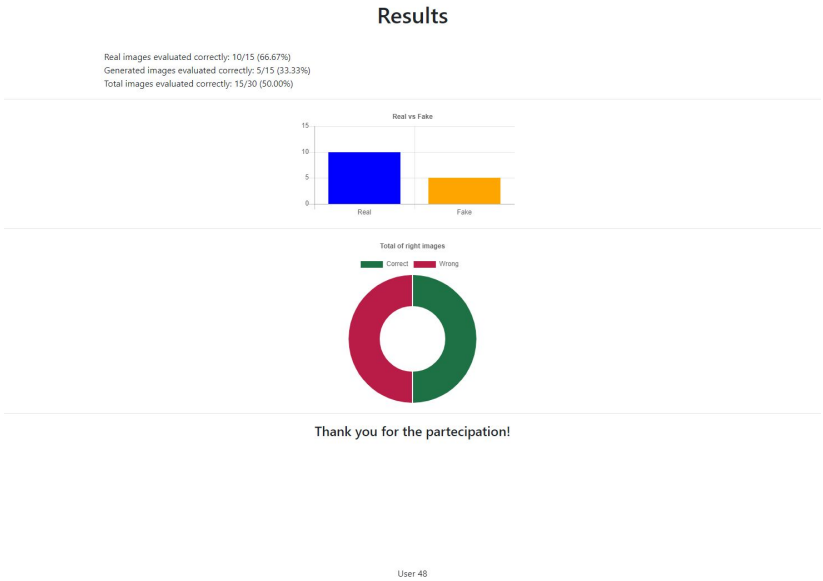
How the application works

- Familiarization with one known image
- Retrieval of the images list for the trial
  - Creation of the user
  - Creation of the list of 30 images (15 real, 15 generated) based on the number of the user
- Load of each image
  - Displayed for 3 seconds
  - Evaluation (from 1 to 7) sent to the server and stored in the database where 1 is a real image and 7 a synthetic image
- Results calculation and display



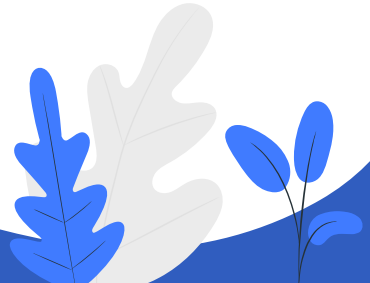
Interface of the application





4 Experiment

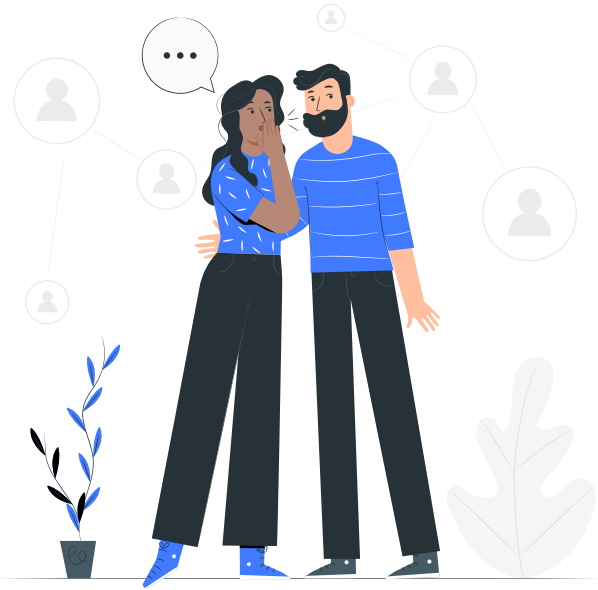
How the experiment was conducted



Test

- Offline experiment (face to face)
- Setup:
  - External monitor
  - 50/60 cm of distance
  - Images of 15x15 cm
- Briefly explanation of the goal and how the system works
- Supervision of evaluation process

Questionnaire



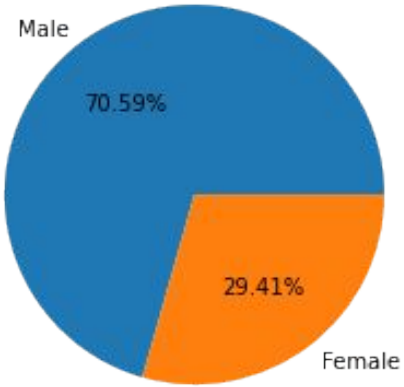
- After the test
- Separate computer
- Through Google Form
- We filled it for them
- Questions about:
  - Demographic information (gender, age, country)
  - Use of glasses
  - Knowledge about Deep-fake
  - Strategies used during the experiment

5 Results

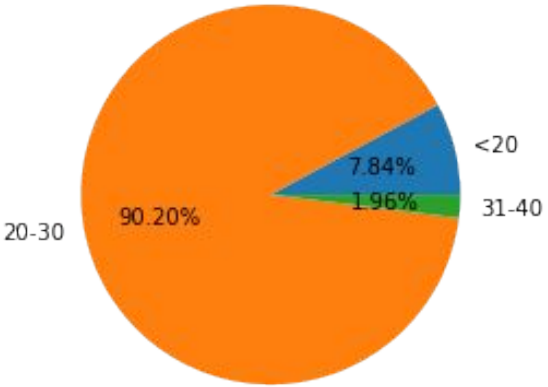
Analysis of collected data

Our participants

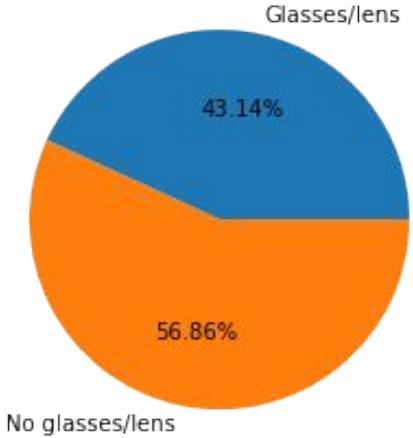
Gender



Age



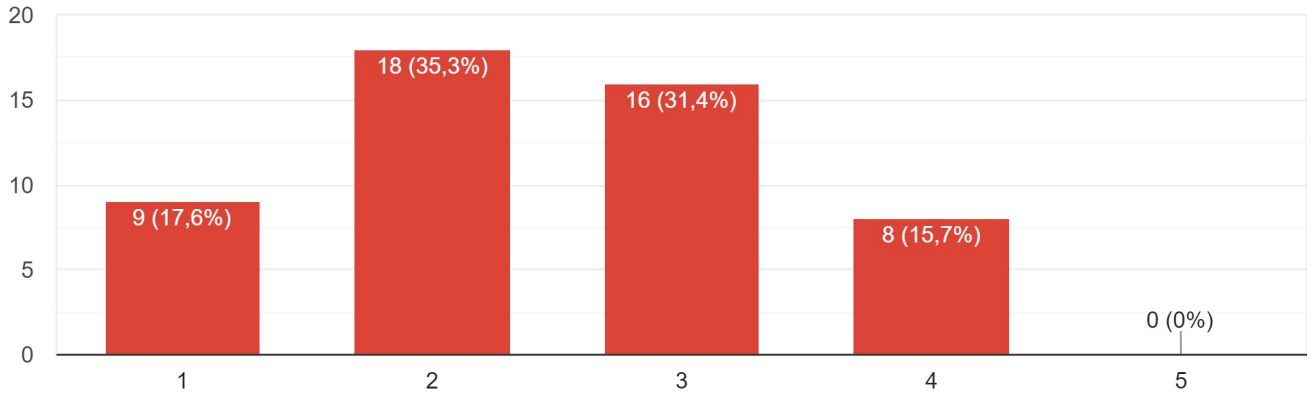
Glasses



51 participants, most of them were Italian

Our participants

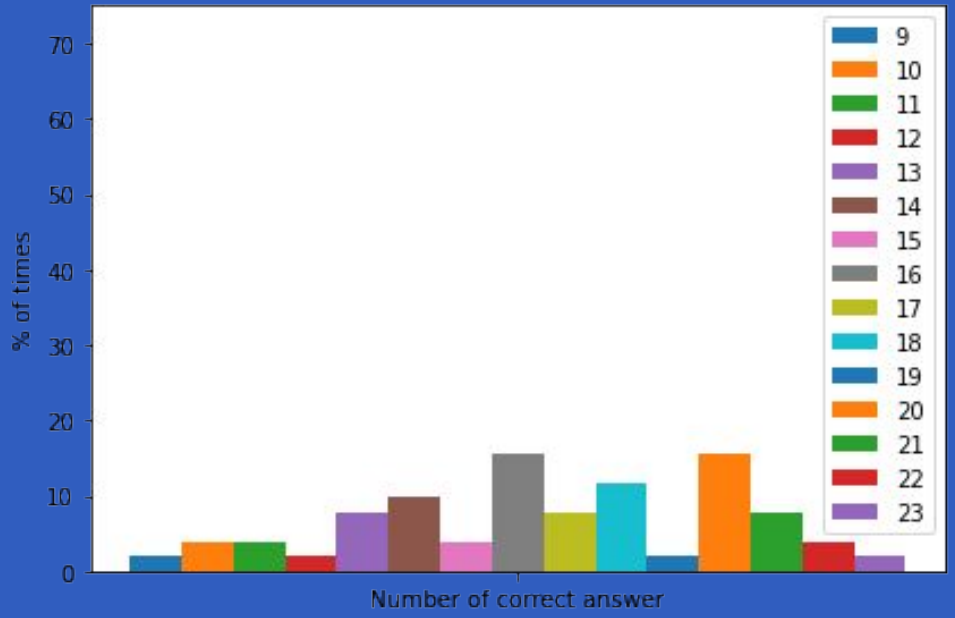
Knowledge of deep fake



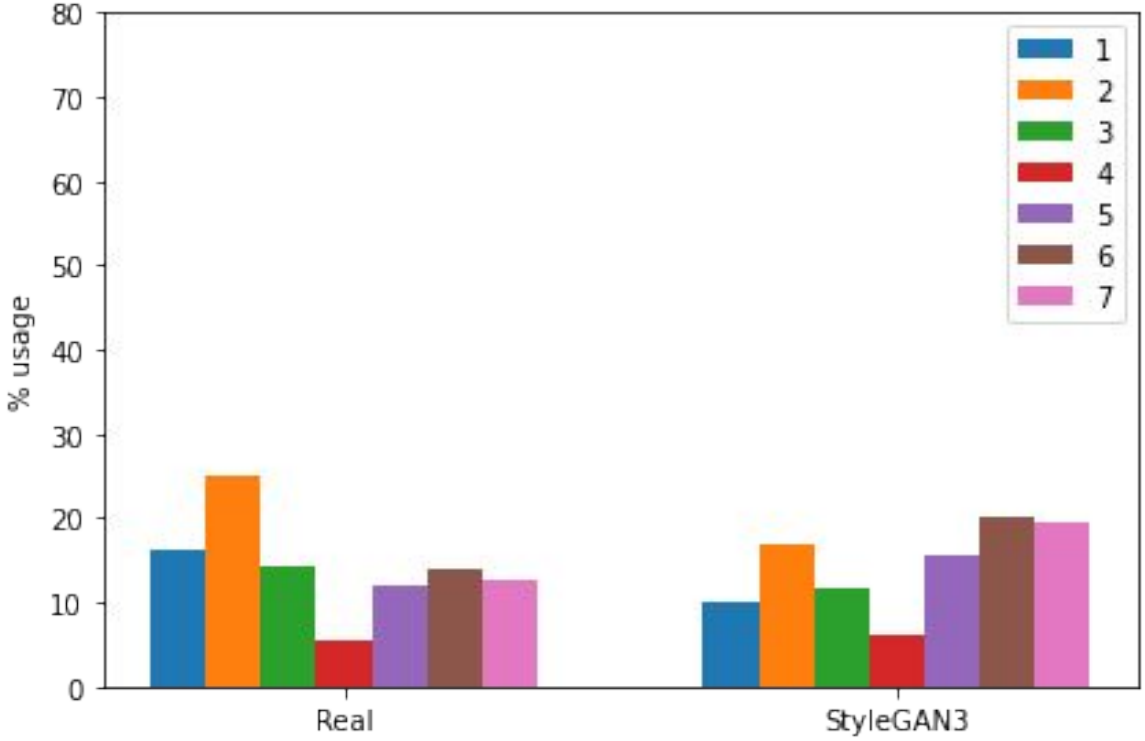
Low

High

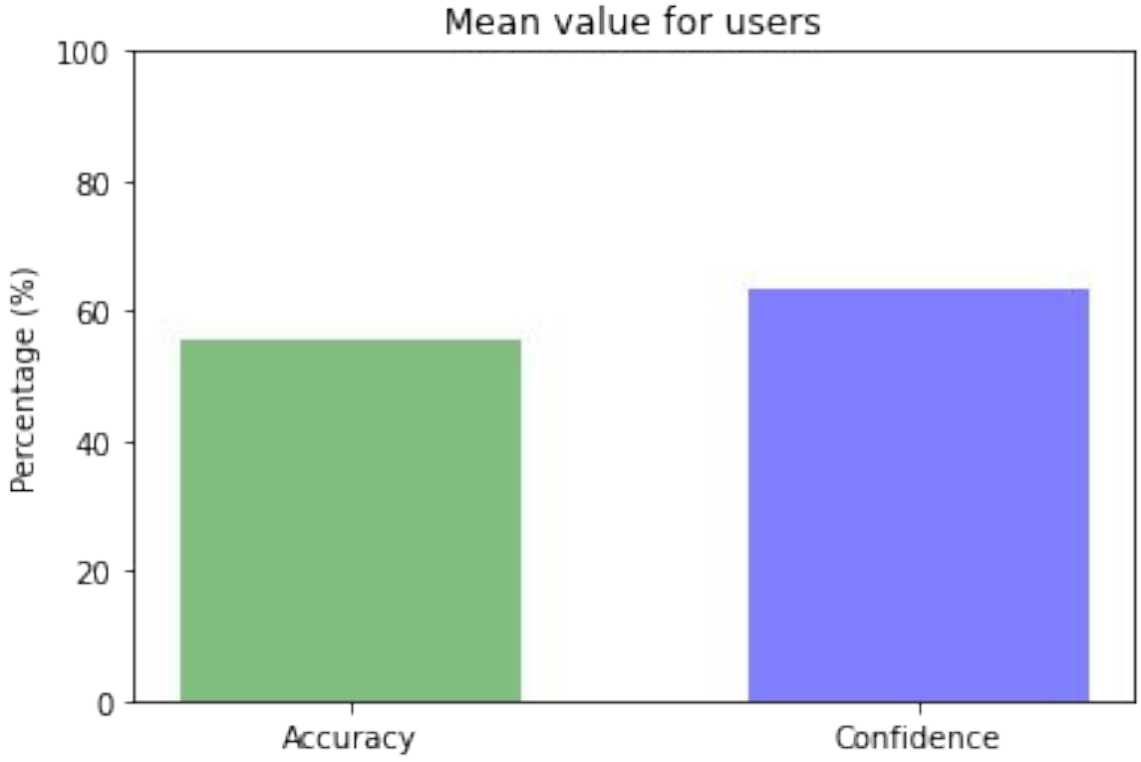
Distribution of results



Distribution of evaluations

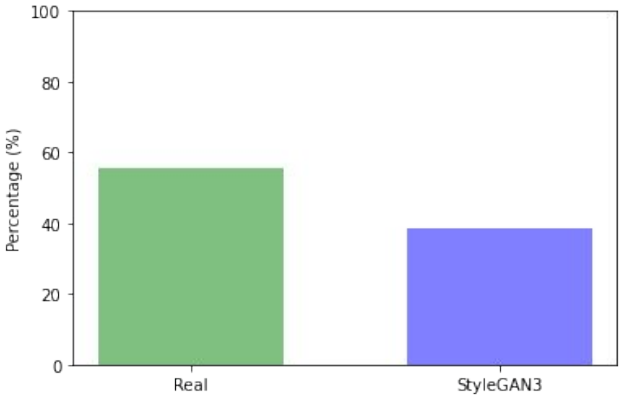


Metrics per users

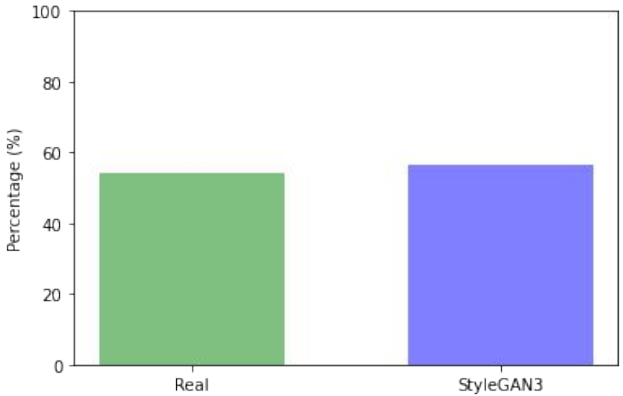


Metrics per images

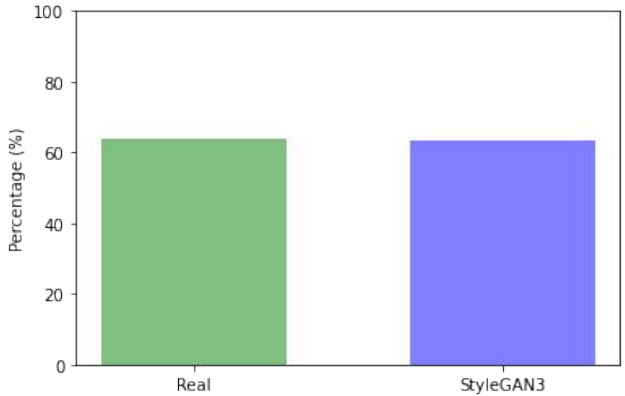
Mean realism rate



Mean accuracy

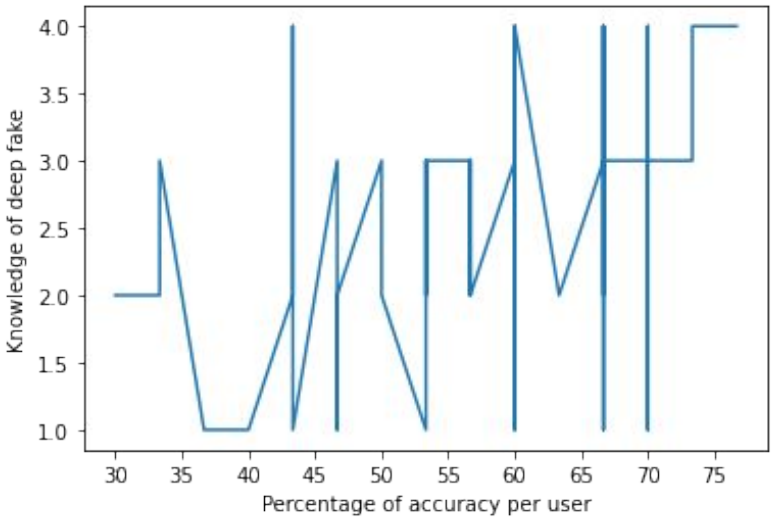


Mean confidence

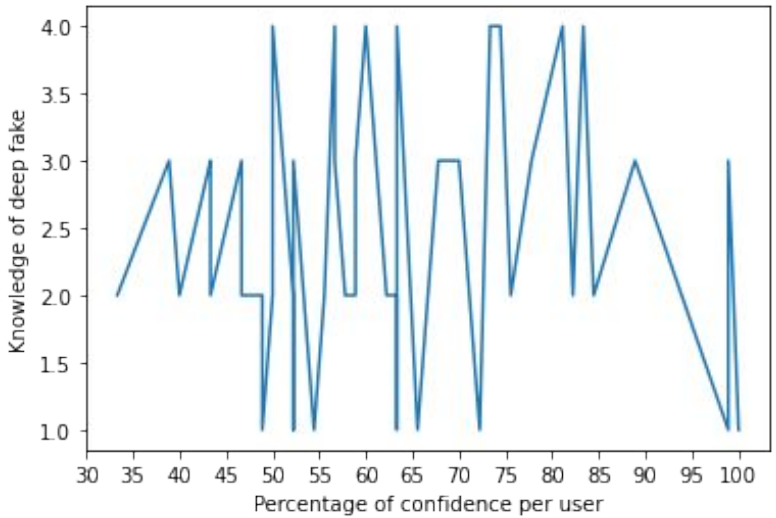


Relation between knowledge and users metrics

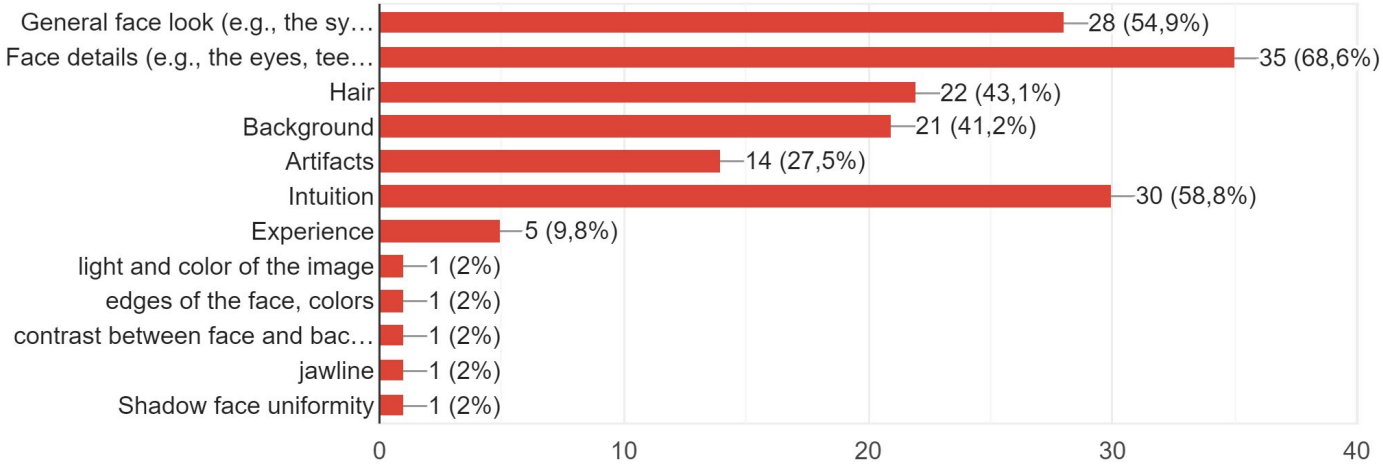
Relationship between accuracy and knowledge



Relationship between confidence and knowledge



Strategy used

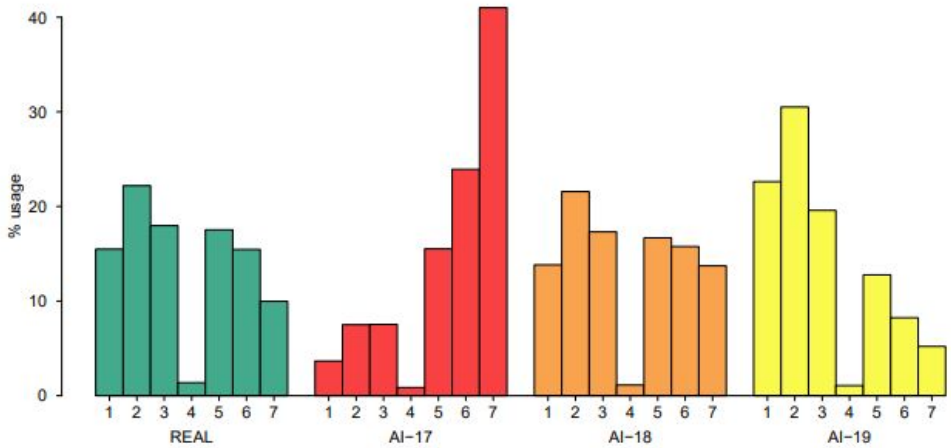


6 Conclusions

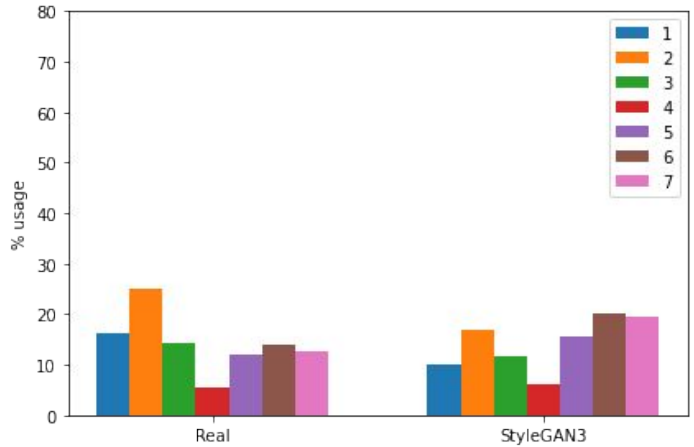
Final words

Comparison of results: evaluations' distribution

Previous experiment

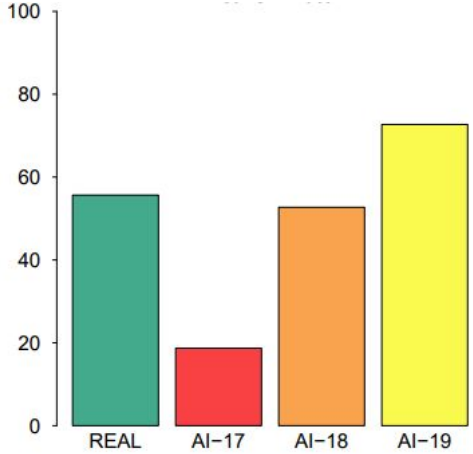


Our experiment

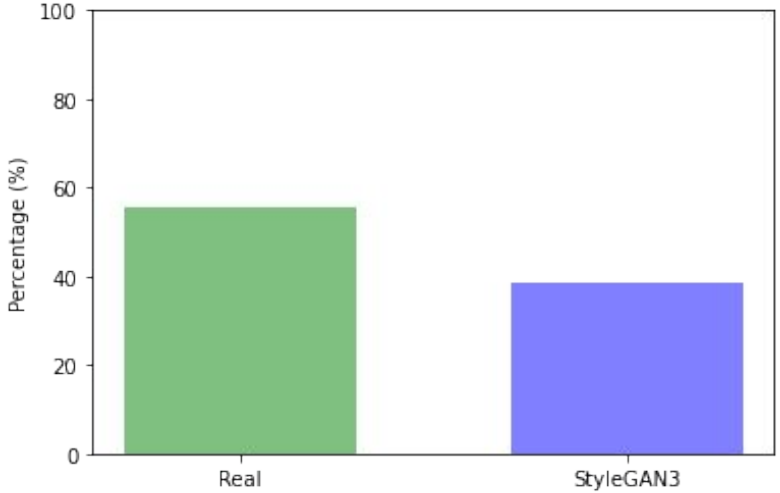


Comparison of results: realism rate

Previous experiment

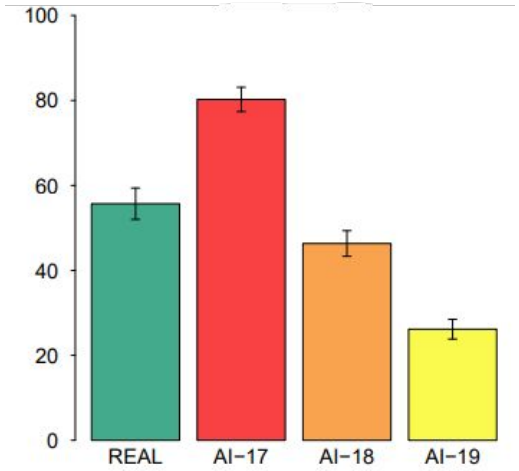


Our experiment

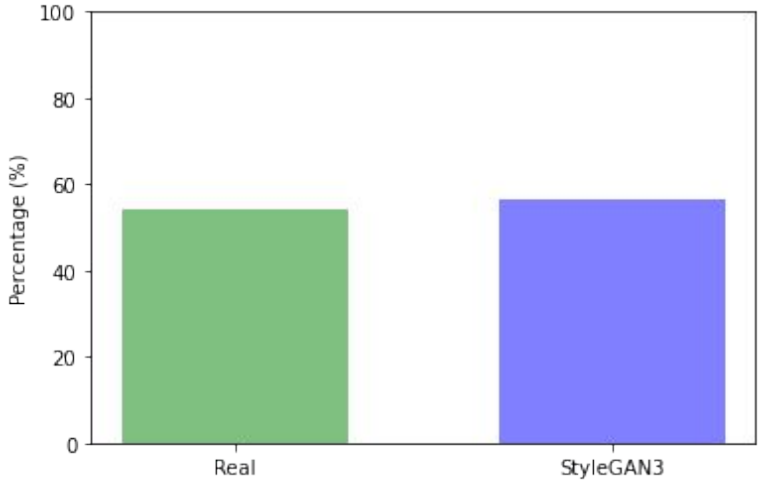


Comparison of results: accuracy

Previous experiment

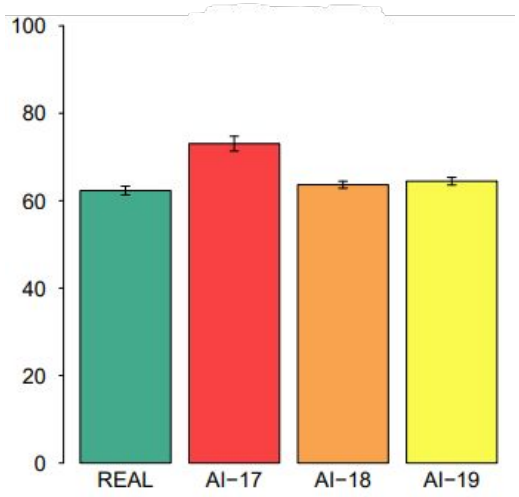


Our experiment

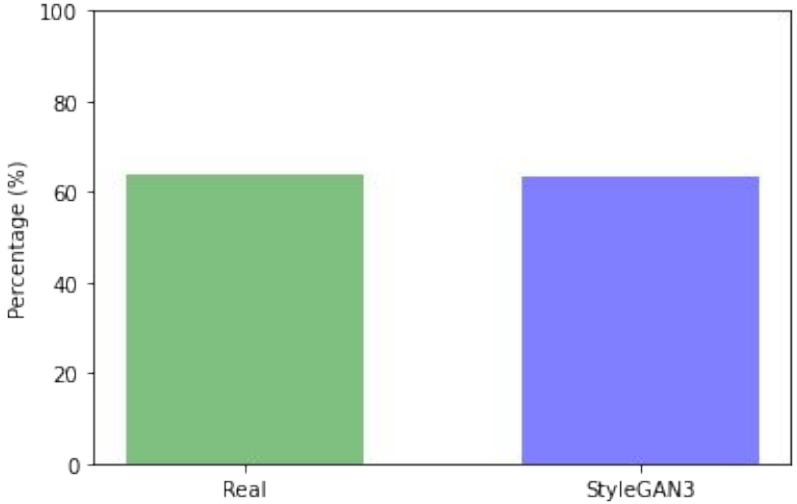


Comparison of results: confidence

Previous experiment



Our experiment

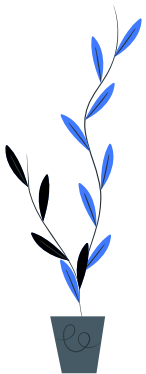


Comparison

StyleGAN-3 generates images that are more easily identifiable from the real ones

StyleGAN-2 produces better images

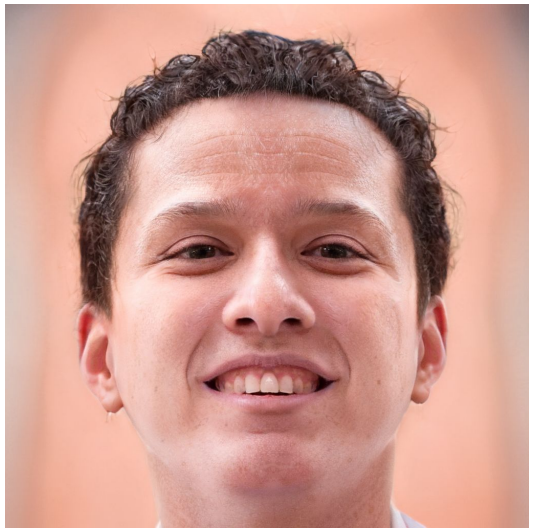
But StyleGAN-3 is better suited for video and animation (as stated in the paper)



Bonus

The best of the worst of StyleGAN3

StyleGAN is not infallible



StyleGAN is not infallible