
Supplementary for: A realistic approach to generate masked faces applied on two novel masked face recognition data sets

Tudor Mare
SecurifAI
tudor.mare@securifai.ro

Georgian Duta
SecurifAI
georgian.duta@securifai.ro

Mariana-Iuliana Georgescu
SecurifAI
University of Bucharest
georgescu_lily@yahoo.com

Adrian Sandru
SecurifAI
adrian.sandru@securifai.ro

Bogdan Alexe
SecurifAI
University of Bucharest
bogdan.alexefmi.unibuc.ro

Marius Popescu
SecurifAI
University of Bucharest
popescunmarius@gmail.com

Radu Tudor Ionescu
SecurifAI
University of Bucharest
raducu.ionescu@gmail.com

1 Data Set Documentation

The COVID-19 pandemic raises the problem of adapting face recognition systems to the new reality, where people may wear surgical masks to cover their noses and mouths. Traditional data sets for training these systems were released before the pandemic, being unsuited due to the lack of examples of people wearing masks. To this end, we introduce synthesized masks for two existing large data sets, CelebA and CASIA-WebFace. The images are generated using a series of tools and algorithms (SparkAR Studio, Dlib-ml, ColorMatcher, custom verification scripts). The released data consists of images containing generated mask overlays that can be applied to a set of existing data sets: CelebA and CASIA-WebFace. The released images do not contain confidential or private information, as they contain only mask overlays. We note that our data sets can also be used to train mask detectors and even facial landmark detectors under mask occlusion.

2 Hosting and Maintenance Plan

The project is hosted and version-tracked via Github. It will be permanently available under the link https://github.com/securifai/masked_faces. The download links of the mask image data sets can be found in the README file included in the Github repository.

We are committed to maintain and actively develop the data sets for a minimum of 3 years since the publication date. We plan to generate additional mask images for other face recognition data sets. For feedback and suggestions, we recommend contacting one of the authors by e-mail.

3 Licensing and Distribution

The masked face data sets are released by SecurifAI, under the CC BY-NC-SA 4.0 (<https://creativecommons.org/licenses/by-nc-sa/4.0/>) license agreement, which permits use and sharing for noncommercial purposes only.

4 Author Statement

We, the authors, will bear all responsibility in case of violation of rights.