Training and testing labelled image and video datasets of human faces for different indoor visual comfort and glare visual discomfort situations

Johra, Hicham; Veit, Martin; Poulsen, Mathias Østergaard; Christensen, Albert Daugbjerg; Gade, Rikke; Moeslund, Thomas B.; Jensen, Rasmus Lund

DOI (link to publication from Publisher):
10.54337/aau542153983

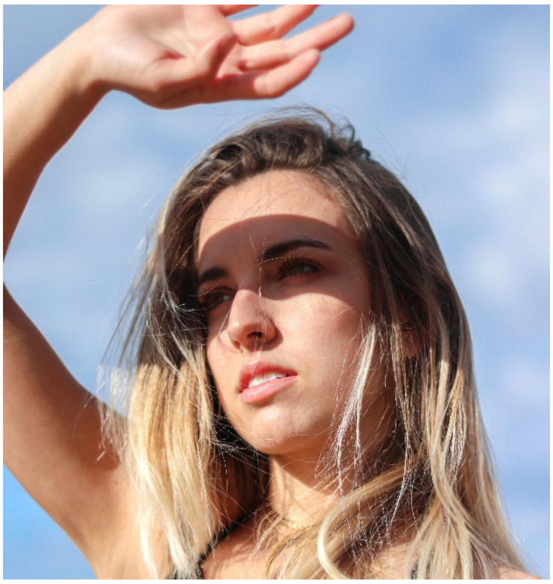
Publication date:
2023

Document Version
Publisher's PDF, also known as Version of record

DEPARTMENT OF THE BUILT ENVIRONMENT
AALBORG UNIVERSITY

Training and testing labelled image and video datasets of human faces for different indoor visual comfort and glare visual discomfort situations

Hicham Johra
Martin Veit
Mathias Østergaard Poulsen
Albert Daugbjerg Christensen
Rikke Gade
Thomas Moeslund
Rasmus Lund Jensen



ISSN 1901-726X
Technical Report No. 316

Aalborg University
Department of the Built Environment
Division of Sustainability, Energy & Indoor Environment

Technical Report No. 316

Training and testing labelled image and video datasets of human faces for different indoor visual comfort and glare visual discomfort situations

by

Hicham Johra
Martin Veit
Mathias Østergaard Poulsen
Albert Daugbjerg Christensen
Rikke Gade
Thomas Moeslund
Rasmus Lund Jensen

July 2023

Scientific Publications at the Department of the Built Environment

Technical Reports are published for timely dissemination of research results and scientific work carried out at the Department of the Built Environment at Aalborg University. This medium allows publication of more detailed explanations and results than typically allowed in scientific journals.

Technical Memoranda are produced to enable the preliminary dissemination of scientific work by the personnel of the Department of the Built Environment where such release is deemed to be appropriate. Documents of this kind may be incomplete or temporary versions of papers—or part of continuing work. This should be kept in mind when references are given to publications of this kind.

Contract Reports are produced to report scientific work carried out under contract. Publications of this kind contain confidential matter and are reserved for the sponsors and the Department of the Built Environment. Therefore, Contract Reports are generally not available for public circulation.

Lecture Notes contain material produced by the lecturers at the Department of the Built Environment for educational purposes. This may be scientific notes, lecture books, example problems or manuals for laboratory work, or computer programs developed at the Department of the Built Environment.

Theses are monograms or collections of papers published to report the scientific work carried out at the Department of the Built Environment to obtain a degree as either PhD or Doctor of Technology. The thesis is publicly available after the defence of the degree.

Latest News is published to enable rapid communication of information about scientific work carried out at the Department of the Built Environment. This includes the status of research projects, developments in the laboratories, information about collaborative work and recent research results.

Published 2023 by
Aalborg University
Department of the Built Environment
Thomas Manns Vej 23
DK-9220 Aalborg Ø, Denmark

Printed in Aalborg at Aalborg University

ISSN 1901-726X
Technical Report No. 316

Foreword

The aim of this technical report is to provide a description and access to labelled image and video datasets of human faces that have been generated for different indoor visual comfort and glare visual discomfort situations. These datasets have been used to train and test a computer-vision artificial neural network detecting glare discomfort from images of human faces [1].

Abstract

Glare is a common local and subjective visual discomfort that is difficult to identify with conventional light sensors. However, computer vision technology can be used to detect it from the analysis of human facial expressions. The current dataset consists of labelled and categorized facial images and video footage samples from 23 different building occupants who are exposed (or not) to glare visual discomfort. The occupants sit at a desktop table while reading a text and answering questions on a computer screen during a controlled laboratory experiment. Spotlights placed at various positions with different angles relative to the occupants are used to generate glare visual discomfort. The dataset includes face extractions, which allow to directly use the occupant's face as a glare visual discomfort sensor. The dataset is structured into a training data subset and a testing data subset. Each subset is further divided into data corresponding to glare visual discomfort experienced by the occupant and data corresponding to no glare discomfort. The data release contains two versions of the dataset: one which has the raw video footage samples and the other one which has images resampled from the video footage samples. This dataset is suited for, but not limited to, machine learning methods, and the training, testing and benchmarking of computer vision algorithms. Although containing only 23 different participants, this dataset was proven useful and was used to successfully train a Convolutional Neural Network (CNN) classifier that can detect glare visual discomfort from face images with an out-of-sample assessment accuracy of around 90% [1]. Such algorithms can be the basis for an efficient feedback control system to regulate shading devices to eliminate the local glare discomfort in an indoor environment.

Keywords

Glare visual discomfort detection, experimental ground truth, human face analysis, artificial intelligence algorithm training, computer vision.

Specifications table

Subject	Computer Vision and Pattern Recognition, Applied Machine Learning, Control and Systems Engineering, Architecture
Specific subject area	Detection of glare visual discomfort from image analysis of human faces. Training and testing datasets for machine learning methods and computer vision algorithms.
Type of data	Images Videos
How the data were acquired	The data acquisition consists in recording the video footage of the face of the participants while they were sitting in a lined-up office environment. During the experiment, the participants were sitting at a desktop table and asked to read a series of short texts on a computer screen and answer a few questions concerning these texts (the questions are irrelevant to the experiment itself, they are only intended to keep the participants performing a typical office task at a desktop). The illuminance inside the test room is maintained at around 500 lux by artificial dimmable light. Spotlights placed on the other side of the test room window (outside the test room, in the laboratory environment) generate glare visual discomfort, similar to that produced by the sun when it is low on the horizon at sunrise or sunset. Throughout the different phases of the experiment, the different sources of glare (spotlight) are alternatively switched on and off to generate (or not) visual discomfort (or absence of visual discomfort) at different horizontal orientations with regard to the occupant sitting at the desktop. At each phase of the experiment, a new visual comfort/discomfort situation is set and the occupant is asked to read a small text and answer a few questions. After each series of questions, the occupant is asked to indicate the perceived thermal, visual and acoustic discomfort. The video footage of the face of the occupant is continuously recorded with a webcam (Logitech Webcam C930e: Full HD 1080p / 30 FPS resolution) placed above the computer screen on which participants are reading and answering questions. The video footages are labelled as “glare” (presence of glare visual discomfort) and “noglare” (no glare visual discomfort experienced by the occupant) based on the occupants’ statements about their visual discomfort at each phase of the experiment. The test participants are not instructed about the exact goal of the experiment.

	The dataset is gathered from 23 participants and includes a certain variety of characteristics: age, gender, face feature, glasses, etc. In order to generate the labelled image dataset, the labelled video footages are resampled at a 0.5 FPS rate, and the YOLOFace extraction algorithm [2] is used to crop the image down to the face of the occupants only. Detailed information on the acquisition method for this dataset can be found in Johra et al., 2021 [1].
Data format	Raw Filtered Classified
Description of data collection	The video and image datasets are generated from experimental tests on 23 different participants and includes a certain variety in characteristics: age, gender, face feature, glasses, etc. The original dataset was generated from experimental tests on 24 different participants but all the data of one of the participants was entirely excluded as this participant did not want his/her data to be publically available. After synchronization of the video footages with the answers of the participants regarding their visual discomfort at the different phases of the experiment to label the video dataset, the video footages are resampled at a 0.5 FPS rate, and the YOLOFace extraction algorithm [2] is used to crop the image down to the face of the occupants only and generate the labelled image dataset.
Data source location	Institution: Aalborg University, Department of the Built Environment City/Town/Region: Aalborg Country: Denmark
Data accessibility	Repository name: zenodo.org Data identification name of the image dataset: Training and testing labelled dataset of human faces for different indoor visual comfort and glare visual discomfort situations – images: dataset_images.zip Direct URL to the image dataset: https://doi.org/10.5281/zenodo.6895080 [3]

	Data identification name of the video dataset: Training and testing labelled dataset of human faces for different indoor visual comfort and glare visual discomfort situations – videos: dataset_videos.zip Direct URL to the video dataset: https://doi.org/10.5281/zenodo.6895054 [4] A copy of the GDPR (General Data Protection Regulation in the European Union) consent form that has been signed by each participant of this data collection campaign is available as supplementary materials to that dataset. The access to the datasets is controlled. To access the datasets and download them, request access directly on the zenodo website (click the “request access” button). This dataset is available for research and scientific purposes only. Proper reference and citation to that dataset are required. To use this dataset for other purposes, such as commercial applications, please contact Hicham Johra (hj@build.aau.dk) or Aalborg University, Department of the Built Environment (build@build.aau.dk).
Related research article	H. Johra, R. Gade, M.Ø. Poulsen, A.D. Christensen, M.S. Khanie, T. Moeslund, R.L. Jensen, Artificial Intelligence for Detecting Indoor Visual Discomfort from Facial Analysis of Building Occupants, Journal of Physics: Conference Series 2042 (2021), 012008. https://doi.org/10.1088/1742-6596/2042/1/012008

Value of the data

- Glare is a common local and subjective visual discomfort that is difficult to identify with conventional light sensors but can be detected with computer vision methods by analysing human facial expressions.
- This dataset consists of labelled and categorized facial images and video footage samples from 23 different building occupants who are exposed (or not) to glare visual discomfort.
- This dataset is suited for, but not limited to, machine learning methods, and the training, testing and benchmarking of computer vision algorithms for the detection of glare discomfort.
- The facial expression analysis allows using the occupant's face as a subjective glare visual discomfort sensor.
- The feedback from a computer vision-based glare detection algorithm can be used to control a shading device that eliminates the local visual discomfort in the indoor environment.

Objective

This dataset was generated to address the challenge of identifying glare discomfort with computer vision algorithms, which is a common visual discomfort that is difficult to detect using conventional sensors and assessment methods. The dataset consists of labelled and categorized facial images and video footage samples from 23 different building occupants who were exposed (or not) to glare visual discomfort. This dataset is suited for, but not limited to, machine learning methods and the training, testing and benchmarking of computer vision algorithms. Although limited in the number of different subjects, this dataset was proven useful as it was used to successfully train a Convolutional Neural Network (CNN) classifier detecting glare discomfort from face images with an out-of-sample assessment accuracy of around 90%. This CNN classifier was presented at a peer-reviewed international scientific conference [1]. The current data article adds value to the published research article [1] by describing the acquisition method for the dataset in a synthetic manner and making the dataset available for research, development and other scientific purposes.

Data description

The dataset contains two subsets: one being the labelled sampled images, the other being the labelled videos from the experiments.

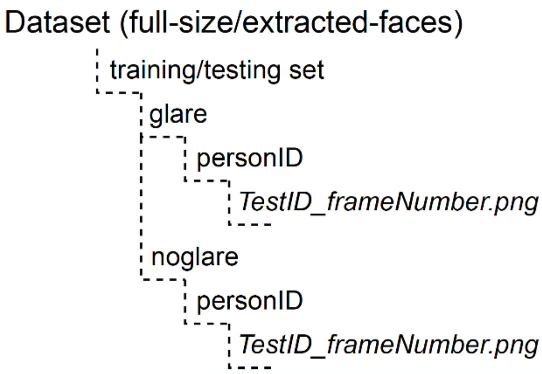
Although containing only 23 different participants, this dataset is believed to be very useful because of its uniqueness. There is, at the moment, no similar dataset for the detection of glare discomfort in the indoor environment. This dataset can thus seed the development of new datasets for detecting glare discomfort from images and videos of human faces. Moreover, the relevance of this dataset has been demonstrated in a peer-reviewed article [1]. This dataset was sufficient to successfully train and test a Convolutional Neural Network (CNN) classifier that can detect glare visual discomfort from face images with an out-of-sample assessment accuracy of around 90%. This glare detection classifier algorithm was developed with off-the-shelf pre-trained neural networks that can readily be used for image analysis and object detection [1].

Image dataset

The image dataset consists of two main folders. One includes the full-size images of the participants in the experiment. The other includes the cropped faces extracted from the full-size images.

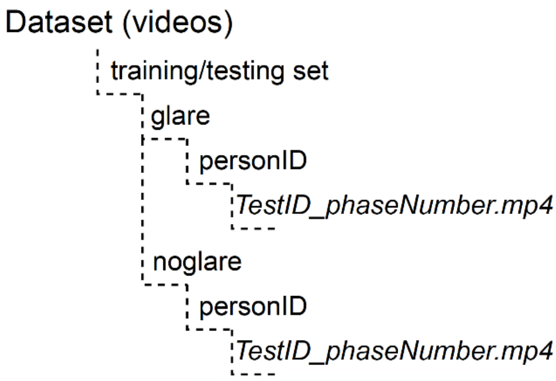
Both main folders contain a total of 14,133 samples. These are divided into a training set with 18 participants and 11,096 samples and a testing set with 5 participants and 3,037 samples. The testing and training datasets are formed to represent, as much as possible, the general dataset characteristics with regards to age, gender, glasses, beards distribution and a representative amount of “glare” and “noglare” samples. The participants are fully split between the two sets, i.e., none of the participants is present in both the training and the testing set. Apart from the training and testing set, the dataset is divided into “glare” and “noglare” folders. These consist of a subfolder for each participant in the set. This subfolder is given a person ID and contains all the samples for the given person. The name of a sample (image file) is the test ID and the frame number, which allows identifying all independent samples in the video. The folder structure is given in the figure below.

Training		**Testing**	
Share	78.5 %	Share	21.5 %
N_{total}	11,096	N_{total}	3,037
N_{glare}	3,738	N_{glare}	928
$N_{noglare}$	7,358	$N_{noglare}$	2,109



Video dataset

The structure of the video dataset is similar to the image dataset. The videos are split into a training and a testing set with the same 18 and 5 participants, respectively. Each video is cut into 12 phases, which have been divided into “glare” and “noglare” folders according to the perceived visual discomfort. These consist of a subfolder for each participant in the set. This subfolder is given a person ID and contains all the videos for the given person. The name of a sample (video file) is the test ID and the phase number. The folder structure is given in the figure below.



Experimental design, materials and methods

The data acquisition consisted in recording the video footage of the face of the participants while they were sitting in a lined-up office environment. During the experiment, the participants were sitting at a desktop table placed next to a large window through which glare visual discomfort could occur. The participants were asked to read a series of short texts on a computer screen and answer a few questions concerning these texts. These questions are irrelevant to the experiment itself; they are only intended to keep the participants performing a typical office task at a desktop and blur the exact goal of the experimental test. The test participants are purposely not instructed about the exact goal of the experiment to avoid disturbances and biases in the results.

The illuminance inside the test room is maintained at around 500 lux by artificial dimmable light. Spotlights placed on the other side of the test room window (outside the test room, in the laboratory environment) and directed at the office's occupant generate glare visual discomfort, similar to that produced by the sun when it is low on the horizon at sunrise or sunset. The sources of glare (spotlight) have 3 different horizontal orientations with regard to the occupant: slightly behind the head of the occupant; on the side of the occupant, perpendicular to his/her line of sight; slightly in front of the occupant (see **Figure 1**).

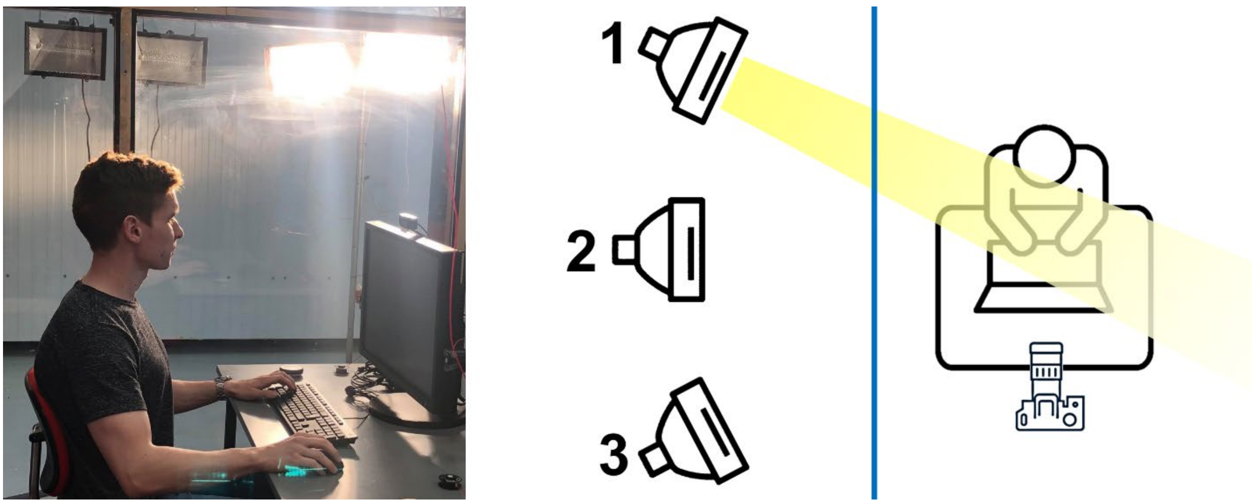


Figure 1: *View of the experimental setup for the acquisition of the dataset. Re-published with permission from Johra et al., 2021, "Artificial Intelligence for Detecting Indoor Visual Discomfort from Facial Analysis of Building Occupants", Journal of Physics: Conference Series 2042, 012008. https://doi.org/10.1088/1742-6596/2042/1/012008 [1].*

Throughout the 12 different phases of the experiment, the different spotlights are alternatively switched on and off to generate (or not) glare visual discomfort (or absence of visual discomfort) at different horizontal orientations with regard to the occupant sitting at the desktop. At each phase of the experiment, a new visual comfort/discomfort situation is set and the occupant is asked to read a small text and answer a few questions. After each series of questions, the occupant is asked to indicate the perceived thermal, visual and acoustic discomfort.

The video footage of the face of the occupant is continuously recorded with a webcam placed above the computer screen on which participants are reading and answering questions. The video footages are labelled as "glare" (presence of glare visual discomfort) and "noglare" (no glare visual discomfort experienced by the occupant) based on the occupants' statements about their visual discomfort at each phase of the experiment. The labelled videos are then resampled with 0.5 FPS (one image every two seconds), to generate an image dataset. The YOLOFace [2] face extractor is used to extract the face of the occupants.

The dataset is gathered from 23 participants and includes a certain variety of characteristics: age, gender, facial feature, glasses, etc.

Ethics statements

To conduct the experiments for the generation of the present dataset, the appropriate administrative body (Aalborg University – AAU Innovation – Grants & Contracts) has been contacted in order to verify the ethical soundness of the experiment and the necessary measures that had to be taken regarding the General Data Protection Regulation (GDPR). After informing this administrative body (Aalborg University – AAU Innovation – Grants & Contracts; protocol number: 2021-068-01292), all the participants in the experiment have been informed about the use of the collected data and a GDPR consent form has been sent to them. The authors hereby confirm that the relevant informed consent was obtained from all subjects who have participated in the generation of that dataset. A copy of the original consent form can be found in the dataset. Copies of the signed informed consent are retained by the authors.

No additional approval from institutional review boards or local ethics committees was necessary to conduct this experiment.

Credit author statement

Hicham Johra: Conceptualization, Methodology, Data acquisition, Data curation, Investigation, Supervision, Writing - Original draft preparation. **Martin Veit**: Methodology, Data acquisition, Data curation, Investigation. **Mathias Østergaard Poulsen**: Methodology, Data acquisition, Data curation, Investigation, Visualization. **Albert Daugbjerg Christensen**: Methodology, Data acquisition, Data curation, Investigation, Visualization. **Rikke Gade**: Methodology, Supervision. **Thomas B. Moeslund**: Methodology, Supervision, Resources. **Rasmus Lund Jensen**: Methodology, Supervision, Resources.

Credit author statement

The authors would like to express their gratitude to all the subjects who have participated in the generation of that dataset. This work was supported by the AI for the People Center (https://www.ai.aau.dk) of Aalborg University, Denmark.

References

[1] H. Johra, R. Gade, M.Ø. Poulsen, A.D. Christensen, M.S. Khanie, T. Moeslund, R.L. Jensen, Artificial Intelligence for Detecting Indoor Visual Discomfort from Facial Analysis of Building Occupants, Journal of Physics: Conference Series 2042 (2021) 012008. https://doi.org/10.1088/1742-6596/2042/1/012008.

[2] T. Nguyen, YOLOFace, https://github.com/sthanhng/yoloface.

[3] H. Johra, M. Veit, Training and testing labelled dataset of human faces for different indoor visual comfort and glare visual discomfort situations - images. Zenodo, v1.0.0, 2022. https://doi.org/10.5281/zenodo.6895080.

[4] H. Johra, M. Veit, Training and testing labelled dataset of human faces for different indoor visual comfort and glare visual discomfort situations - videos. Zenodo, v1.0.0, 2022. https://doi.org/10.5281/zenodo.6895054.

Recent publications in the Technical Report Series

Hicham Johra. Long-Term Stability and Calibration of the Reference Thermometer ASL F200. DCE Technical Reports No. 266. Department of Civil Engineering, Aalborg University, 2019.

Hicham Johra, Olena K. Larsen, Chen Zhang, Ivan T. Nikolaisson, Simon P. Melgaard. Description of the Double Skin Façade Full-Scale Test Facilities of Aalborg University. DCE Technical Reports No. 287. Department of Civil Engineering, Aalborg University, 2019.

Hicham Johra. Overview of the Typical Domestic Hot Water Production Systems and Energy Sources in the Different Countries of the World. DCE Technical Report No. 288. Department of Civil Engineering, Aalborg University, 2019.

Hicham Johra. Thermal Properties of Building Materials - Review and Database. DCE Technical Report No. 289. Department of the Built Environment, Aalborg University, 2021.

Hicham Johra. Performance overview of caloric heat pumps: magnetocaloric, elastocaloric, electrocaloric and barocaloric systems. Technical Report No. 301. Department of the Built Environment, Aalborg University, 2022.

Martin Veit, Hicham Johra. Experimental Investigations of a Full-Scale Wall Element in a Large Guarded Hot Box Setup: Methodology Description. Technical Report No. 304. Department of the Built Environment, Aalborg University, 2022.

Hicham Johra. Datasets on the work habits of international building researchers. Technical Report No. 305. Department of the Built Environment, Aalborg University, 2022.

Hicham Johra. General study case description of TMV 23: A multi-storey office building and Living Lab in Denmark. Technical Report No. 306. Department of the Built Environment, Aalborg University, 2023.

Martin Veit, Hicham Johra. A comparative study of BSim and COMSOL Multiphysics for steady-state and dynamic simulation of transmission loss. Technical Reports No. 309. Department of the Built Environment, Aalborg University, 2023.

Hicham Johra, Mathilde Lenoël. Experimental setup description and raw data from the micro-climate measurement campaign of the outdoor air temperature around an office building in Denmark during summer. Technical Reports No. 313. Department of the Built Environment, Aalborg University, 2023.

Hicham Johra, Markus Schaffer, Gaurav Chaudhary, Hussain Syed Kazmi, Jérôme Le Dréau, Steffen Petersen. Coherent description of 48 metrics to compare, validate and assess accuracy of building energy models and indoor environment simulations. Technical Reports No. 314. Department of the Built Environment, Aalborg University, 2023.

Martin Veit, Christian Grau Sørensen, Michal Z. Pomianowski, Hicham Johra, Anna Marszal-Pomianowska. A practical approach to set up a simple database architecture using SQLite and Python. Technical Reports No. 315. Department of the Built Environment, Aalborg University, 2023.

ISSN 1901-726X
Technical Report No. 316