Aalborg Universitet

Artificial Intelligence for Detecting Indoor Visual Discomfort from Facial Analysis of Building Occupants

Johra, Hicham; Gade, Rikke; Poulsen, Mathias Østergaard; Christensen, Albert Daugbjerg; Sarey Khanie, Mandana; Moeslund, Thomas B.; Jensen, Rasmus Lund

PAPER • OPEN ACCESS

Artificial Intelligence for Detecting Indoor Visual Discomfort from Facial Analysis of Building Occupants

View the article online for updates and enhancements.

Artificial Intelligence for Detecting Indoor Visual Discomfort from Facial Analysis of Building Occupants

Hicham Johra[1]*, Rikke Gade[2], Mathias Østergaard Poulsen[2], Albert Daugbjerg Christensen[2], Mandana Sarey Khanie[3], Thomas Moeslund[2] and Rasmus Lund Jensen[1]

[1] Department of the Built Environment, Aalborg University, Thomas Manns Vej 23, DK-9220, Aalborg Øst, Denmark
[2] Department of Architecture, Design and Media Technology, Aalborg University, Rendsburggade 14, DK-9000, Aalborg, Denmark
[3] Department of Civil Engineering, Technical University of Denmark, Brovej 118, DK-2800, Kongens Lyngby, Denmark
* Corresponding author: hj@build.aau.dk

Abstract. Glare is a common local visual discomfort that is difficult to identify with conventional light sensors. This article presents an artificial intelligence algorithm that detects subjective local glare discomfort from the image analysis of the video footage of an office occupant's face. The occupant's face is directly used as a visual comfort sensor. Results show that it can recognize glare discomfort with around 90% accuracy. This algorithm can thus be at the basis of an efficient feedback control system to regulate shading devices in an office building.

1. Introduction

Enhancing the indoor daylight conditions is a key issue to address as it greatly impacts building occupants comfort and satisfaction, particularly in a working environment. The illumination conditions in a daylit room can include situations that reduce visibility and create visual discomfort. One of them is known as glare. It is usually marked by occupants' behaviours such as blinking [1], looking away [2] or shielding the eyes when encountering extreme contrasts. If persistent, it can result in visual fatigue [3][4] and is one of the main drivers for building users' interaction with the façade settings [5], which consequently can change the building's performance. Optimized integration of glare-free daylight solutions thus proves to be crucial for a sustainable building design.

Office workplaces are highly prone to be affected by glare discomfort as a result of non-optimized lighting. An automated shading system that could prevent glare in dynamic weather changes, locally [6] and at an individual level [7], while providing adequate daylight penetration proves to be an effective solution. However, local indoor visual comfort is highly dependent on the location and orientation of the occupants and windows, the furniture layout, the position of the sun in the sky, the cloud cover, and the surfaces blocking or reflecting sunlight. It is thus very complicated to assess the effective local visual comfort of occupants based on fixed sensors placed in a room. In addition, visual comfort can drastically change within few seconds in windy countries with rapid cloud cover variations.

Despite glare being one of the main drivers for building occupants to activate solar shadings, the current commercial systems do not support accurate visual comfort assessment. The existing visual comfort metrics rely on several photometric values to be measured in the room, and can only partially predict actual occupants' perception of visual comfort with high uncertainties. Consequently, building users are

complaining and their unpredictable and non-optimized interactions with shading devices and electric lighting induce a higher energy demand.

While simulation-assisted data-driven methods can be employed for glare control strategies with automated shading systems [8], advances in artificial intelligence (AI) could open up new opportunities to enhance the visual comfort conditions at the local and individual level. In response to glare discomfort, humans reshape their orbital rims and eyes to dim the collected light. This anatomic adjustment reflex to the lighting conditions is very distinctive and visible on the human's face. Moreover, a beam of light striking the eyes of a person will leave clear over-exposed light patches on his/her face. These are very convenient as they can be identified by means of video-based face analysis. Indeed, machine learning methods for AI have made considerable progress in the field of facial recognition and analysis over the last few years. This technology should now be mature enough to decipher the subjective local indoor visual comfort perception of building occupants from an image of his/her face.

The aim of the project presented in this article is to develop a proof of concept prototype to demonstrate that the subjective visual comfort of an office occupant can be accurately assessed by a computer algorithm analysing the video footage of the human's face. The output of such an algorithm can be used as feedback signal to control dimmable artificial lighting and motorized solar shading devices to optimize the effective local visual comfort of the occupants.

Instead of relying on multiple illuminance sensors in a room and identifying the complex correlation between the sensors' measurements and the actual visual comfort of the occupant, the current concept uses directly the occupant's face as a sensor for subjective and localized visual comfort (which is what actually matters to the occupant). This smart control system based on facial analysis of occupants' footage should be able to react promptly to the rapid changes of outdoor conditions which can induce intermittent glare. This is particularly frequent in Nordic countries like Denmark where the cloud cover can drastically change in a matter of seconds, and the low position of the sun in the sky typically causes glare for a significant share of the day, especially during wintertime.

This paper presents the different steps for the development of a glare detection algorithm based on facial image analysis. After discussing the results of the prototype, the article closes with preliminary conclusions and information about the future work that will be soon conducted.

2. Study case description

The scope of the current project is restricted to a single occupant in an office room working at a computer station placed next to a large window. The glare discomfort is generated by spotlights emulating the sun on the other side of the window (outside the test room) and directed directly at the office's occupant. The height of the spotlights is similar to that of the sun when it is low on the horizon at sunrise or sunset. (see Figure 1).
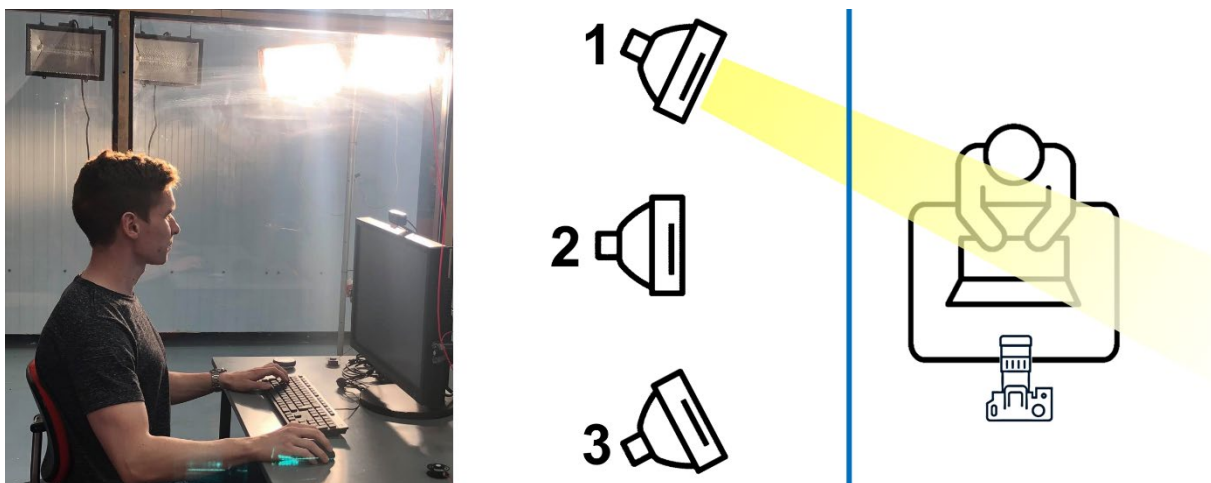
Figure 1. Experimental setup for the acquisition of the training data.

The source of glare (spotlight) can have 3 different horizontal orientations with regards to the occupant: 1) the spotlight is slightly behind the head of the occupant and the light strikes the occupant's face on the side and from behind 2) the spotlight is on the side of the occupant, perpendicular to his/her line of sight; the light strikes the occupant's face on the side 3) the spotlight is slightly in front of the occupant and the light strikes the occupant's face on the side and the front (see Figure 1). Video footage of the occupant's face is recorded with a single webcam mounted on the computer screen.

3.   Training data acquisition

The first step to develop the visual discomfort detection algorithm is to generate a labelled set of training data: video footage of the face of different humans with indications about whether or not they are experiencing visual discomfort. To that matter, a test room with a controlled environment is set up in a laboratory (see the previous section). The training data acquisition campaign consisted of 17 experimental tests conducted with 15 different participants. Although this is a small number of participants, there is a certain variety of people's characteristics: age, gender, face feature, glasses, etc. Each test takes place as follows: the participant sits down in front of the computer and reads a short introductory text on a graphical interface developed for this experiment. The initial phase of the experiment has a neutral indoor illuminance of around 500 lux (adjusted by artificial dimmable light). The participant goes through the different phases of the experiment by pressing the "next" button on the interface. At each phase, the participant is asked to read a short text, and then to answer 3 questions related to that text. The participant then indicates if he/she is currently experiencing thermal discomfort, visual discomfort or acoustic discomfort, and the next phase begins. At the beginning of a new phase, the light condition is changed: one of the 3 glare sources is switched on (see Figure 1) to potentially induce glare discomfort to the participant, or all spotlights are switched off and the indoor illuminance is kept at 500 lux (neutral situation without glare discomfort). There are 12 phases: 6 neutral scenes with no glare sources, and 2 repetitions of each of the 3 glare source positions. Each phase usually lasts around 1 minute and 30 seconds.

The test participants are purposely not instructed about the exact goal of the experiment to avoid disturbances and biases in the results. The texts to be read, the questions related to the latter, and the query about thermal comfort and acoustic comfort are only intended to blur the focus of the experiment.

4.   Development of the glare detection algorithm

After synchronization of the video footage with the answers of the participants, the training data is used to develop the facial analysis algorithm. The glare detection task is defined as a binary classification problem: comfort or discomfort. It must be done using only the video feed of the participant's face.

Firstly, the video is resampled at 3 images per second. Each image is processed individually, hence the algorithm classifies the visual comfort of individual images. The first part of the algorithm must identify the face in the image, and crop it down to the face only. This task is performed by a pre-trained face classifier based on the *Haar Feature-based Cascade Classifier* method (proposed by Viola and Jones [9]) and using the *OpenCV* implementation. This algorithm is pre-trained, meaning that no further training with the current project's data is needed to perform the task.

The facial analysis for classifying comfort and discomfort is based on a Convolutional Neural Network (CNN). These networks are designed for analyzing the content of images. The main challenge with them in general is the number of parameters that needs to be trained. It requires a large amount of labelled training data, which can be very time-consuming. However, large publicly available image datasets can be used for teaching the network basic image recognition. Then a smaller dataset of images from the relevant task can be used to train the last layer of the network, which will decide the classification output. For this work, a *VGG-16* network [10] with the *TensorFlow Keras* implementation is used. The *VGG-16* has 16 weight layers (convolutional layers and fully connected layers). The inputs to the network are RGB images rescaled to 224x224 pixels. The CNN is pre-trained on the *ImageNet* dataset which consists of more than 1.2 million images labelled with 1000 object classes [11]. The transfer learning concept is then applied to train the fully connected classification layer (last layer of the network) with the comfort/discomfort-labelled images from the training data acquired for this project (see Figure 2).
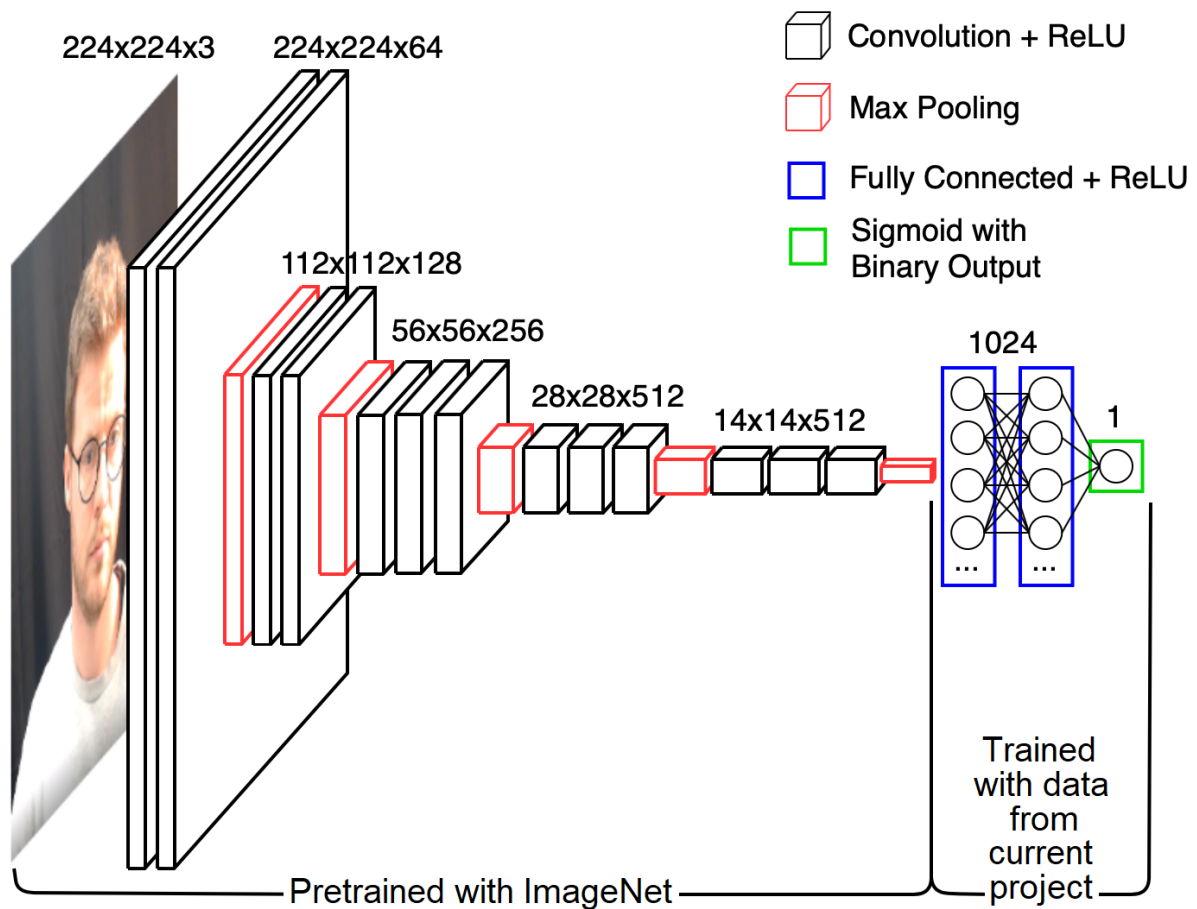
Figure 2. The VGG16 network with its components. The rectified linear unit (ReLU) function is a piecewise linear function that outputs the input value if that value is positive, and outputs zero for negative input values. ReLU is used as the activation function for the convolutional layers and the fully connected layers. Max pooling with size 2x2 and stride 2x2 outputs the maximum value in the given 2x2 area: it is used to reduce the computational cost of the network.

5. Results and discussions

After examining the training data, 3 persons were discarded due to synchronization errors and substantial problems with automatic face detection. Hence, 12 test participants are used for training and testing the prototype of the glare detection algorithm. 2 different strategies have been investigated to split the experimental dataset into training and test data. In experiment 1, the data is split based on individual people, such that images of the same person cannot be represented in both training and test data. Hence, this experiment shows the performance of the algorithm when presented to unseen faces. 4-fold cross-validations are run, in which each iteration uses 3 persons as test data and the remaining 9 for training the classifier. Therefore, all 12 persons are alternatively used for training and testing.

In experiment 2, the effect of including the same person in training and test data is studied. All videos are cut into shorter sub-sequences consisting of a single experimental phase. 5-fold cross-validations are then run, in which each iteration takes 80% of the pooled sub-sequences as training data and 20% as test data. All participants are thus included in both training and test sets, but only with data from different phases of the experiment.

Table 1. Algorithm classification accuracy.

Iteration number	1	2	3	4	5	Mean accuracy	Std. dev.
Experiment 1	71.7%	99.0%	92.2%	97.3%	-	**90.01%**	10.89%
Experiment 2	85.3%	91.8%	82.9%	90.6%	86.6%	**87.44%**	3.31%

With a classification accuracy of around 90%, one can see that the glare detection algorithm performs very well in both experiments (see Table 1). Experiment 1 shows the highest mean accuracy but also the highest standard deviation. This indicates that some persons may be harder to classify due to appearance variations, different facial expressions, distinct personal tolerance to glare discomfort, or disruptive elements on the image such as glasses. In experiment 2, all test persons are mixed, which explains why the results are more even.

Table 2. Examples of classification results. Columns 2 and 4 are correctly classified, while columns 3 and 5 are incorrectly classified by the algorithm.
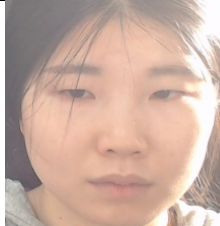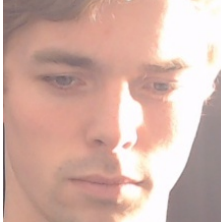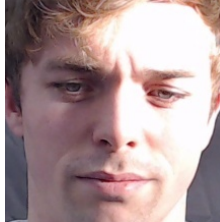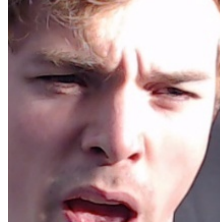
True label	Comfort	Comfort	Discomfort	Discomfort
Algorithm classification	Comfort	Discomfort	Discomfort	Comfort
Example 1				
Example 2				

Table 2 presents some examples of correct and incorrect classification results from the glare detection algorithm. Improper automatic face detection and extraction has been identified as a typical source of error and misclassification. A wrongly-cropped image is an insufficient input to the CNN classifier and impairs its performance (see Figure 3).
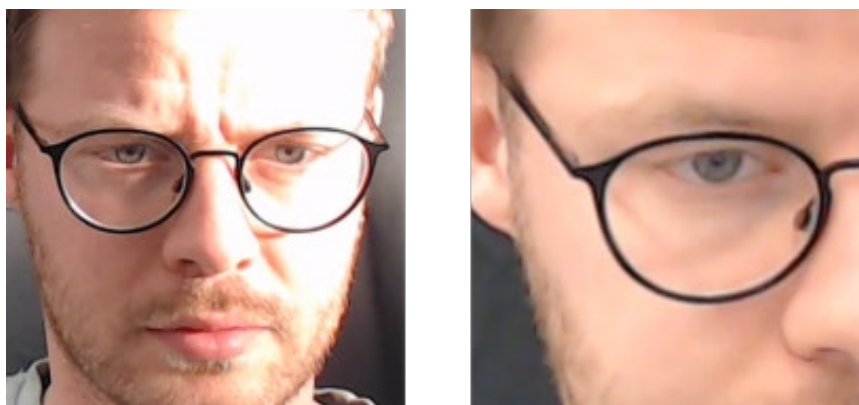
Figure 3. Good (left) and bad (right) examples of automatic face extraction.

6. Conclusions and future work

The results of this study show that it is possible to use AI to detect local glare discomfort from the facial analysis of the building's occupants. A glare detection algorithm has been developed with off-the-shelf pre-trained neural networks that can readily be used for various image analysis and object detection. It is thus possible to train the last layers of the network for specific applications with a fairly limited training dataset and minimum computation power. The neural network created for this project can detect

local glare discomfort with an accuracy of around 90% from the video feed of a webcam filming a person sitting at an office desk. This algorithm can thus be at the basis of an efficient feedback control system to regulate shading devices in a single-office building and eliminate local glare discomfort.

These first promising results motivate the continuation of the project to improve the capacities of the glare detection algorithm and integrate it into a smart dynamic façade system. The next steps for future work are as follows:

- Increase of the training dataset with more participants, different backgrounds and glare discomfort situations.
- Gain a deeper understanding of the artificial neural network operation when it performs the facial analysis for glare detection.
- Use the output of the glare detection algorithm to regulate the position of simple shading devices in an office such as Venetian blinds or shutters.
- Set up a reinforcement learning interface for self-optimizing glare detection.
- Use the algorithm to detect glare discomfort on multiple occupants in a room and control accordingly flexible dynamic façade that can shade local parts of the glazings without blocking the entire view to the outdoor.
- Study the integration of additional building information and sensors feedback to optimize indoor visual comfort (fish-eye view, location of glare sources, the geometry of the room and furniture layout, position of the occupants, etc).

Acknowledgements

References
[1] Doughty, M. J. (2014). Spontaneous eyeblink activity under different conditions of gaze (eye position) and visual glare, Graefe's Arch. Clin. Exp. Ophthalmol., 252(7), 1147–1153.

[2] Sarey Khanie, M., Stoll, J., Einhäuser, W., Wienold, J., & Andersen, M. (2017). Gaze and discomfort glare, Part 1: Development of a gaze-driven photometry, Light. Res. Technol., 49(7) 845–865.

[3] Boyce, P. R. (2014). Human factors in lighting, 3rd edition.

[4] Stone, P. (2009). A model for the explanation of discomfort and pain in the eye caused by light, Light. Res. Technol., 41(2), 109–121.

[5] Galasiu, A. D., & Veitch, J. A. (2006). Occupant preferences and satisfaction with the luminous environment and control systems in daylit offices: a literature review, Energy Build., 38, 728–742.

[6] Dubois, M. C., Gentile, N., Laike, T., Bournas, I., & Alenius, M. (2019). Daylighting and lighting under a nordic cloud, Studentlitteratur.

[7] Dubois, M. C., Demers, C. M. H., & Potvin, A. (2007). The influence of daylighting on occupants: comfort and diversity of luminous ambiences in architecture. Proceedings of the American Solar Energy Society (ASES), SOLAR 2007.

[8] Xie, J., & Sawyer, A. O. (2021). Simulation-assisted data-driven method for glare control with automated shading systems in office buildings, Build. Environ., 196, 107801.

[9] Viola, P., & Jones, M. J. (2001). Rapid object detection using a boosted cascade of simple features. The 2001 IEEE Computer Society Conference on Computer Vision and Pattern Recognition.

[10] Simonyan, K., & Zisserman, A. (2015). Very Deep Convolutional Networks for Large-Scale Image Recognition. International Conference on Learning Representations.

[11] ImageNet Large Scale Visual Recognition Challenge (ILSVRC) http://www.image-net.org/challenges/LSVRC/index.