Webcam-based eye tracking

Information on the technology and accuracy

By Tess den Uyl and Amogh Gudi
Research Consultants, Vicarious Perception Technologies BV

A solution by Noldus Information Technology

WEBCAM-BASED EYE TRACKING TECHNOLOGY

FaceReader offers more than just facial expression analysis. It also provides eye tracking analysis using a regular webcam. Our eyes are our windows to the world, and our gaze reveals what we are looking at, often indicating what we find important. Eye gaze can be stimulus-driven, revealing what implicitly draws attention, but it can also be guided by explicit motivation.

Eye gaze can be stimulus-driven, revealing what implicitly draws attention, but it can also be guided by explicit motivation.

Eye tracking is therefore used across a wide range of fields, such as psychology, marketing, and usability research. Traditionally, eye tracking technology involves dedicated hardware, such as an infrared camera, attached to a computer, with varying price ranges and accuracy depending on requirements. In contrast, using a simple RGB camera enables eye tracking on any computer or laptop with a webcam, greatly increasing flexibility.

This white paper discusses the methodology and accuracy behind the technology. It also provides a validation example for a specific application and explains how to start using it in your own project.

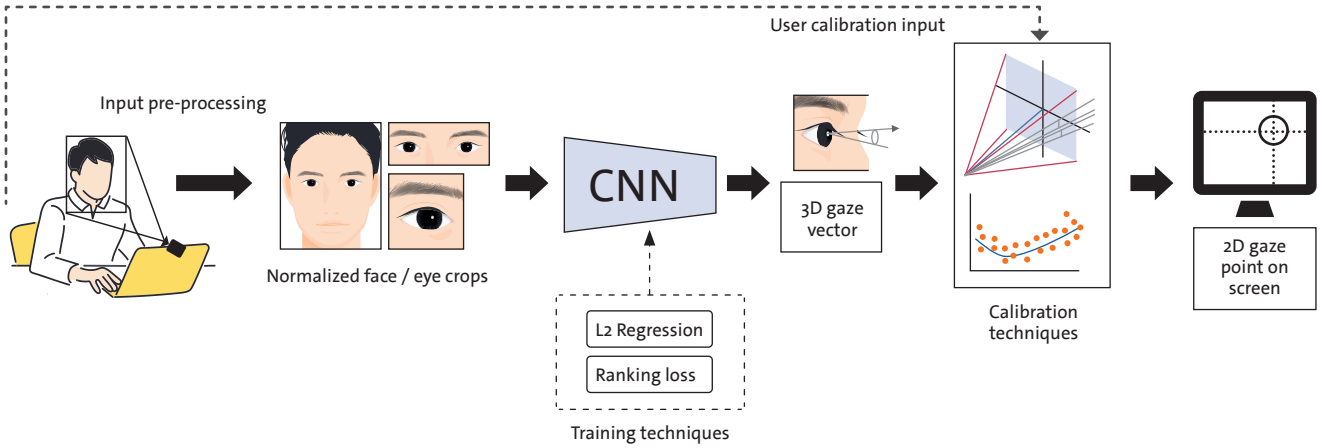
HOW IT WORKS

The eye tracking algorithm [1] estimates gaze direction and maps it to an image on the screen. The first step is face analysis to locate key points in the face (such as the eyes and nose) and estimate the head pose.

The eye image and head pose information are then processed by a pre-trained convolutional neural network (CNN) to produce a 3D gaze vector (a line segment in three dimensions) representing the direction of the person's eye gaze. The neural network is trained on large, labelled datasets containing screen locations and video recordings, enabling it to learn the relationship between the appearance of the eyes and the corresponding gaze vector.

To map the 3D gaze model to a specific location on the screen, the system must know exactly where the screen is. This is achieved through a calibration task, in which the participant follows dots on the screen. The 3D vector is then mathematically converted into 2D gaze coordinates (x, y) on the screen.

Illustration of processing steps of the eye tracking algorithm. [1]



1. Based on Gudi, Amogh, Xin Li, and Jan van Gemert. *Efficiency in real-time webcam gaze tracking.* European Conference on Computer Vision. Springer International Publishing, 2020.

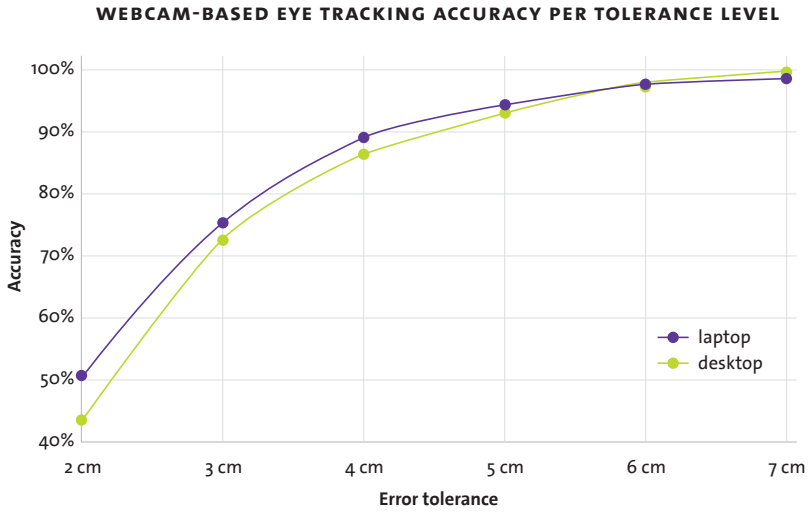
ACCURACY OF WEBCAM-BASED EYE TRACKING TECHNOLOGY

We tested the system on our own validation dataset under natural lighting conditions, using data from 17 participants (12 males, 5 females; 9 aged 18–59). Participants completed a dot following task and several psychological tasks. Data was collected with both the webcam-based eye tracking algorithm and a dedicated hardware-based eye tracker (Tobii Nano Pro) in parallel. Participants were seated approximately 75 cm from a 1920×1080p 23-inch screen. They also performed a dot following task on a 15-inch laptop (without the hardware eye tracker).

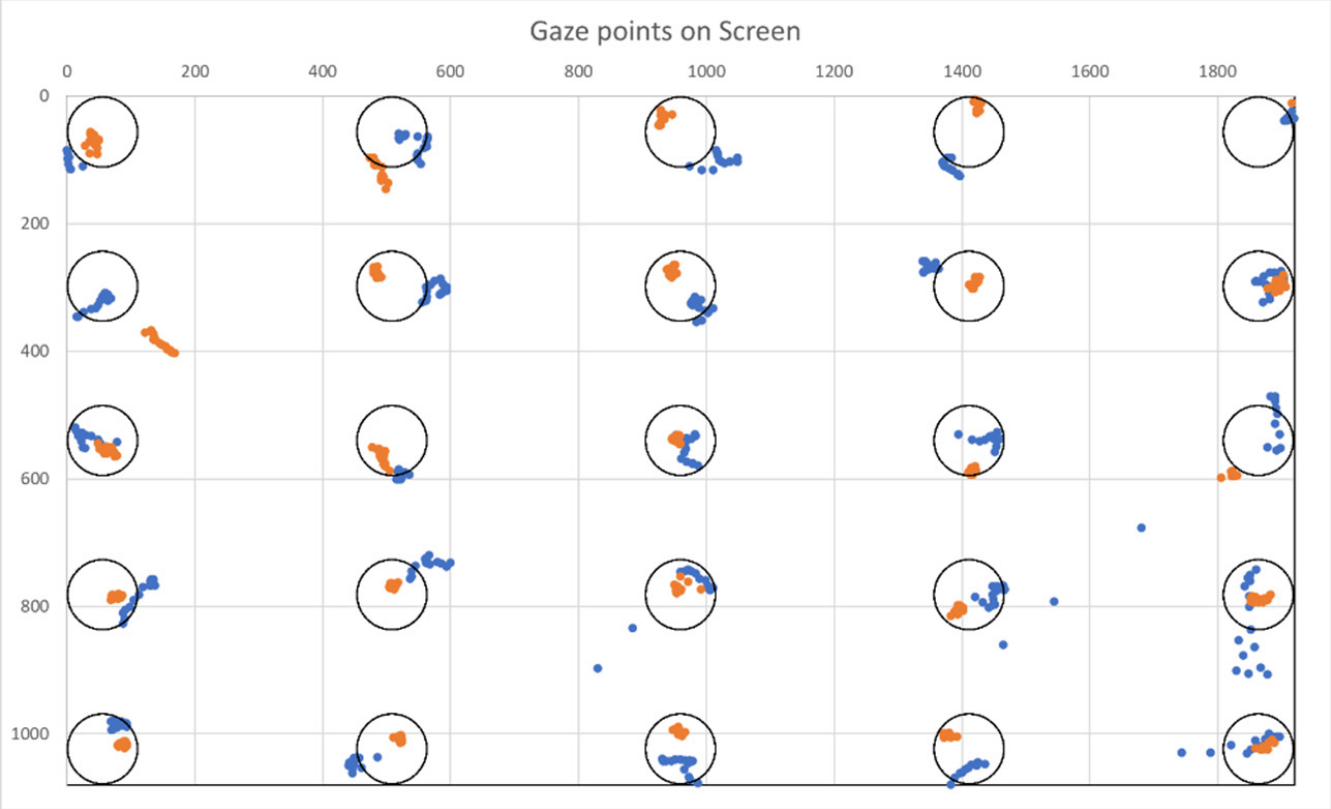
	EYE TRACKING ERROR RATE	
	FaceReader (webcam-based eye tracking)	Tobii Nano Pro (hardware-based eye tracking)
Desktop [23" screen]	2.4 ±1.5 cm | 1.8 ±1.2 °	0.63 ±0.3 cm | 0.51 ±0.2 °
Laptop [15" screen]	2.3 ±2.0 cm | 1.8 ±1.5 °	N/A

Eye tracking error rates in terms of average deviations on the screen (in cm) as well as average angular deviation in degrees.

The results on the dot following task (13 dots are used for calibration and 21 for testing/validation) on a desktop computer show that the algorithm predicts gaze points on screen with an average deviation of 2.4±1.5 cm (or an average of 93 pixels, approximately 4.2% deviation of the screen size). This corresponds to an approximate angular deviation of 1.8±1.2°.



Plot of the webcam-based eye tracking accuracy for different tolerance levels - deviations within tolerance are counted as correct.



Example data from one participant: gaze points plotted on the screen from the webcam-based eye tracker (blue dots) vs the hardware eye tracker (orange dots) with respect to the instructed dot locations (open dots).

The results on a laptop with a smaller screen also achieve a similar accuracy (using its low-resolution built-in webcam). In addition, the accuracy obtained using a dedicated hardware eye tracker (Tobii Nano Pro) is also shown for comparison (only available for the desktop).

In terms of accuracy, FaceReader's webcam-based eye tracking achieved 85–90% accuracy within a 4 cm tolerance, and about 75% within a 3 cm tolerance. In practice, this means it is well-suited for applications where gaze targets are spaced 3–4 cm or more apart. For finer targets (1–2 cm), hardware eye tracking can be preferable as webcam-based eye tracking accuracy in this range drops below 50%.

These results are comparable to, or better than, other webcam-based eye tracking systems on the market, although direct comparisons can be difficult due to differences in screen sizes (errors sometimes reported in pixels) and testing setups.

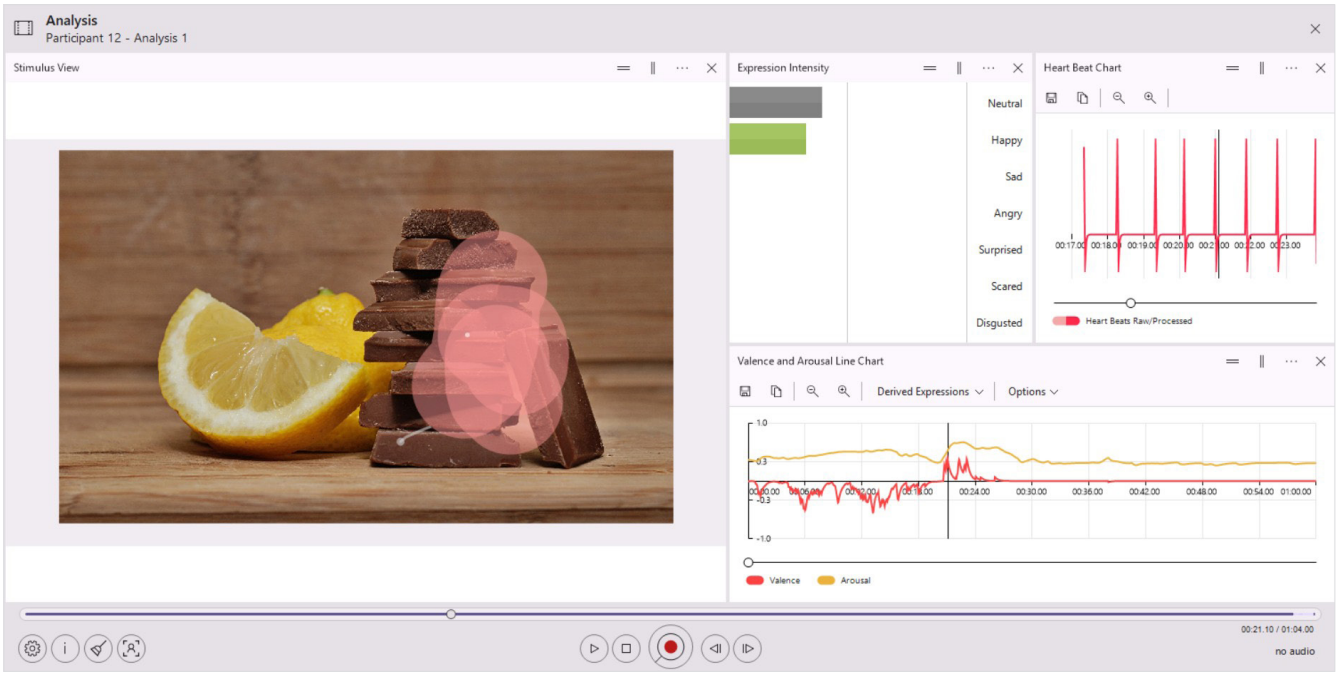
HOW CAN WEBCAM-BASED EYE TRACKING BE USED

An important but simple finding within eye tracking research is that people look longer at what they find interesting. We let the participants perform two simple choice tasks, where they viewed a few images of items (e.g. cookies, chocolate, etc.) and were allowed to choose one afterwards.

A comparison of the webcam-based eye tracking gaze estimations with the hardware-based eye tracker estimations shows a high correlation between the X and Y coordinate estimation (X: Mean r = 0.87, Y: Mean r = 0.77).

Gaze estimations can be classified into being a saccade (fast movements with less visual information processing) or fixation (slower eye focus relevant for information processing) based on the speed of the gaze movements. For each category that was presented total duration of fixation time was measured.

Eye tracking shows us what grabs attention. In this screenshot, the participant's gaze is clearly drawn to the chocolate. Emotional responses are shown in the charts, with measures of expression intensity, heart rate, and valence/arousal.

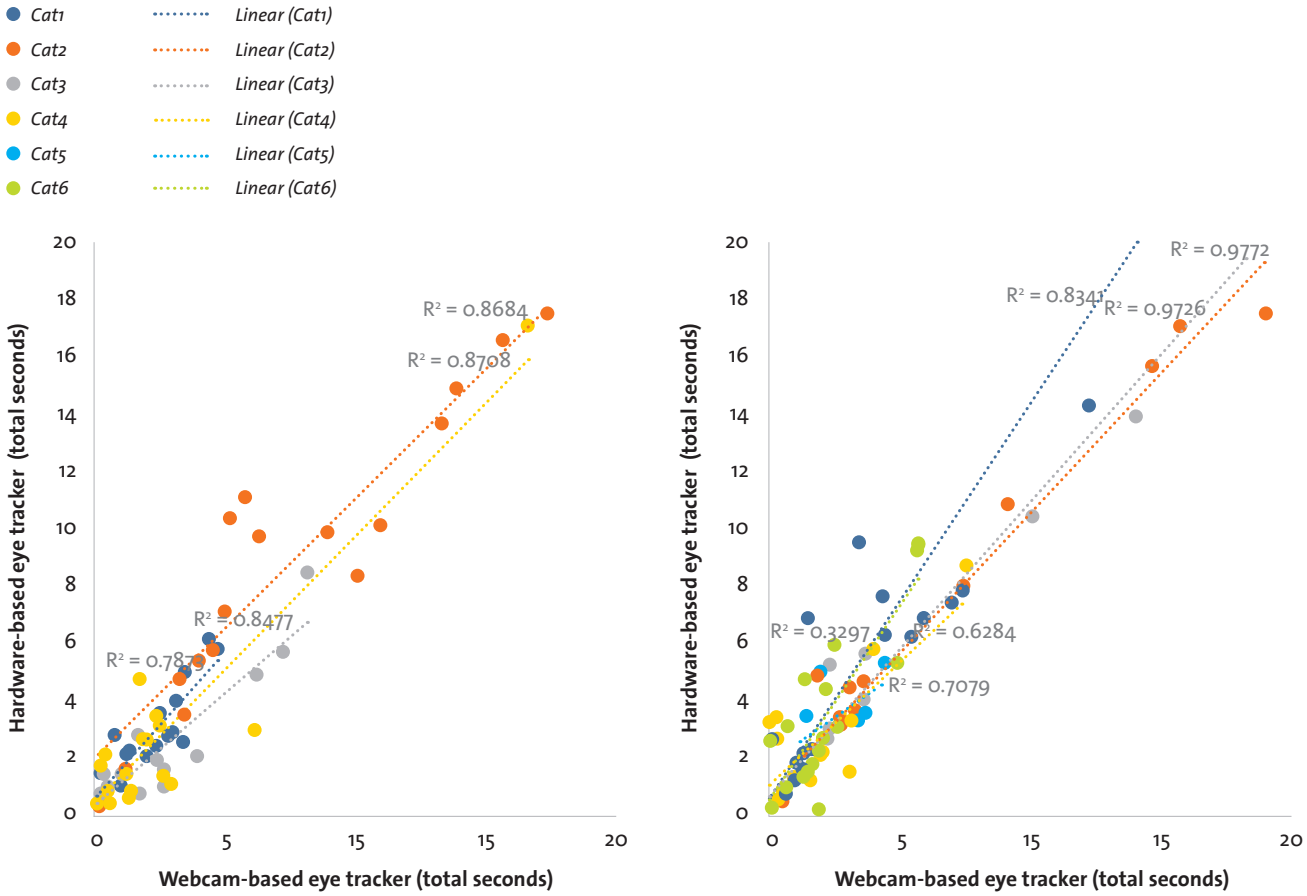


VALIDATION OF THIS APPLICATION

Both the webcam-based eye tracking algorithm and the hardware-based Tobii Nano Pro eye tracker produced the same overall interpretation of the results: 79% of participants looked longest at the category they ultimately preferred (which they then received as a reward). With the webcam-based eye tracker, participants spent on average 9.3 seconds fixating on the chosen category and 2.0 seconds on the non-chosen categories, compared to 11.1 and 2.6 seconds respectively with the Tobii Nano Pro. The slightly lower estimates obtained with the webcam-based eye tracker are most likely explained by a somewhat higher proportion of failed eye localizations.

In addition, the total fixation duration for each category indicated strong correlation between the two systems (mean r = 0.91, see figure).

Correlation within different categories in two tasks between fixation duration of hardware-based eye tracker and webcam-based eye tracker output.



GET STARTED NOW

> Webcam-based eye tracking is very suitable for researching behavior in realistic settings where you expect clear discernible effects.

The results show that eye tracking via the webcam can be a useful and easily applicable research tool. Under good conditions varying individual and screen characteristics do not cause large differences in performance.

Generally speaking, there is often a trade-off between ecological validity and measurement precision. If high precision is required, a lab setting is more appropriate where participants can also put their head in a headrest and have long calibration sessions. Webcam-based eye tracking is very suitable for researching behavior in realistic settings where you expect clear discernible effects.

You can use webcam-based eye tracking in a flexible and mobile lab environment by installing FaceReader on a laptop. In addition, the stimulus presentation tool should be used for the calibration step.

You can also use webcam-based eye tracking via the FaceReader Online platform for online testing. Simply turn gaze tracking on in your experiment design and a calibration will be added. Results can be visualized in heatmaps and are quantified with gaze statistics such as fixation duration.

The webcam-based eye tracking results can be visualised with a heatmap indicating hot spots of gaze focus per frame or over the whole presentation time.

INTERNATIONAL HEADQUARTERS
Noldus Information Technology bv
Wageningen, The Netherlands
Phone: +31-317-473300

E-mail: contact@noldus.com

NORTH AMERICAN HEADQUARTERS
Noldus Information Technology Inc.
Leesburg, VA, USA
Phone: +1-703-771-0440
Toll-free: 1-800-355-9541
E-mail: info@noldus.com

REPRESENTATION
We are also represented by a worldwide network of distributors and regional offices. Visit our website for contact information.

noldus.com