

Everyday Mobile Visual Attention (EMVA) Dataset End User License Agreement (06/2020)

ACADEMIC OR NON-PROFIT ORGANIZATION NONCOMMERCIAL RESEARCH USE ONLY

By using or downloading the dataset and any of the associated software, You are agreeing to the terms of this license agreement. If You do not agree with these terms, You may not use nor download the dataset.

This is a license agreement ("Agreement") between your academic institution or non-profit organization or self (called "Licensee" or "You" in this Agreement) and ETH Zürich (called "Licensor" in this Agreement). All rights not specifically granted to You in this Agreement are reserved to the Licensor.

MATERIAL UNDER LICENSE

The material under License (hereinafter: "Dataset") is a database described in detail in Appendix 1. The Dataset is composed of videos collected from the front-facing camera of mobile devices as well as associated sensor data, usage logs, interactions events, and location data. The Dataset additionally contains annotations for the eye contact detection task. This dataset was collected as a result of a compensated user study as approved by the Ethics Committee of ETH Zürich (EK 2018-N-118). All data contained within the dataset have been collected and processed in accordance with the laws applicable in Switzerland.

COPYRIGHT

The Dataset is the sole property of the Licensor and is protected by copyright. The Licensor reserves all rights to use and distribute the Dataset in any way. The Dataset shall remain under the exclusive property of the Licensor.

ACCESS

You may only use this Dataset after this Agreement has been signed and returned to the Licensor. You have to return the signed agreement by email, in PDF format, to the following e-mail address: emva.dataset@gmail.com.

PERMITTED USES

The Dataset may only be used for your own non-commercial internal research purposes.

The use of screenshots, images, or videos of the study subjects in your own work, including in publications is permitted. However, the use of images that contain bystanders are not permitted (see below). In case of doubt, please send an email to emva.dataset@gmail.com.

USES NOT PERMITTED

You may not distribute, copy or use the Dataset except as explicitly permitted herein. In particular, any commercial use of the Dataset is strictly prohibited.

The use of screenshots, images, or videos of bystanders, i.e. other people who have been captured by the camera (e.g. in the background but not only) and who have not given their explicit consent, is explicitly forbidden without the Licensor's consent.

DISTRIBUTION

You shall not, without prior authorization from the Licensor, transfer in any way, permanently or temporarily, distribute or broadcast all or parts of the Dataset to third parties.

DERIVATIVES

You may create derivatives of or make modifications to the Dataset, however, you agree that all and any such derivatives and modifications will be owned by the Licensor and become a part of the Dataset licensed to You under this Agreement. In case of modifications or derivatives, you must provide to the Licensor one copy of your changes in a recognized electronic format via email sent to emva.dataset@gmail.com within thirty (30) days of the publication date of any publication that relates to any such derivatives or modifications. You understand that the Licensor is not obligated to distribute or otherwise make available any derivatives or modifications provided by you.

SUPPORT

The Dataset is provided as-is. No additional software support or training by the Licensor is provided as part of this Agreement.

AMENDMENTS

The Licensor is allowed to amend this License at any time without the prior consent of the Licensee. You will be informed about changes and given the choice to opt out of this License within thirty (30) days at the following email address: emva.dataset@gmail.com. If the amendment is not substantial, the Licensor may modify this Agreement without any notice and it will be fully applicable to the Licensee within thirty (30) days from its publication on www.emva-dataset.org.

PUBLICATIONS

Any work or publication that uses the Dataset, or results obtained with it, has to reference the following publication:

Mihai Bâce, Sander Staal, and Andreas Bulling. 2020. Quantification of Users' Visual Attention During Everyday Mobile Device Interactions. In Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems (CHI '20). Association for Computing Machinery, New York, NY, USA, 1–14.

DOI:<https://doi.org/10.1145/3313831.3376449>

EXCLUSIVE REMEDY AND LIMITATION OF LIABILITY

To the maximum extent permitted under the applicable law, the Licensor shall not be liable for direct, indirect, special, incidental, or consequential damages or lost profits related to the Licensee's use of and/or inability to use the Dataset, even if the Licensor is advised of the possibility of such damage. The Licensor shall not be held responsible for any illegal or criminal use of the Dataset by the Licensee.

GOVERNING LAW

This Agreement shall be construed and enforced in accordance with the Swiss law.

WARRANTIES

By signing this Agreement, you agree to have read, understood, and to strictly respect the conditions set forth herein and to respect all the laws applicable in Switzerland with regards to the data contained within the Dataset collected and processed by the Licensor.

Organization Name (Institute, University):	
Organization address:	
Full Name:	
Position:	
Email address:	
Date:	
Signature:	

--	--

**APPENDIX 1:
A Short Description of the Dataset**

The EMVA dataset contains video snippets from 33 participants collected using the front-facing camera of their own mobile devices over more than two weeks in-situ, during everyday routine.

Table 1 describes the data sources in detail. Some data sources might not be present for all participants (e.g. many mobile phones did not have an ambient temperature sensor).

For more details on how the Dataset was collected, please refer to the following publication:

Mihai Bâce, Sander Staal, and Andreas Bulling. 2020. Quantification of Users’ Visual Attention During Everyday Mobile Device Interactions. In Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems (CHI ’20). Association for Computing Machinery, New York, NY, USA, 1–14.

DOI:<https://doi.org/10.1145/3313831.3376449>

Table 1: Detailed description of the data sources

Data	Description
Participant ID	Each participant was assigned a 128-bit account UUID (Universally Unique Identifier), which has been randomly created on the first start of the app and is used to identify each user account.
Timestamps	Various timestamps are stored. For instance, the time when a new session begins, the time a notification has been displayed, or the time of a sensor reading. A timestamp is built by saving the current date in nanoseconds since 1970. An example value is "1486395620249".
Camera Data	

Video from the front-facing camera	When users unlocked their mobile device, the app recorded a video from the front-facing camera. The recording lasted until the the device went back into standby.
Device Integrated Sensors	
Accelerometer	Measures the applied acceleration force (in m/s^2) to the device on all three physical axes (x, y, and z).
Gyroscope	Measures a device's rate of rotation in rad/s around each of the three physical axes (x, y, and z).
Magnetic field	Measures the ambient geomagnetic field for all three physical axes (x, y, z) in μT .
Proximity	Measures the proximity of an object in cm relative to the display of the device. This sensor is typically intended for finding out whether the device is currently held up against your ear.
Ambient temperature	Measures the ambient room temperature in degrees Celsius ($^{\circ}C$).
Step counter	Counts the number of steps you have taken since the last reboot of the device.
Rotation data	Measures the orientation of a device by computing the three axes of the device's rotation vector.
Light	Measures the ambient light level (illumination) in lx.
Location Data	
Place	<p>If users enabled GPS on their device, the data collection application logged the latitude, longitude and the accuracy of the last GPS measurement while the phone was unlocked. For anonymisation, this data was converted into general place types.</p> <p>This conversion was done using the Google Nearby Places Search Request API. For example, the values "47.37642" (latitude) and "8.54767" (longitude), which are the coordinates for the main building at ETH Zürich, would be stored as "library", "point of interest" and "establishment".</p> <p>Furthermore, some participants opted to use Private Locations.</p>

	No location data was gathered when users were within 100m from a private location.
Other Collected Data	
Activity	The user's activity while using the device was stored. It was accessed via the Google API for activity recognition and the possible values are: "IN_VEHICLE", "ON_BICYCLE", "ON_FOOT", "RUNNING", "STILL", "TILTING", "UNKNOWN", and "WALKING". Each activity type has a confidence value associated with it.
Notifications	The dataset contains events concerning mobile notifications: When a new notification appeared or when it was removed from the device status bar, together with metadata, like the package name of the application which created it. None of the actual content of the notification was stored.
Touch events	The timestamp and the event itself. The location of the touch event was not stored (in agreement with Android's security policy).
Calendar	If users allowed calendar access, the application stored a flag (0 or 1) indicating whether there was an appointment in the device calendar or not. If so, the device also stored the start and end time of the appointment. No information about the appointment's content was stored.
Nearby Bluetooth devices	The application logged the RSSI (Received Signal Strength Indicator) value and Bluetooth MAC address for nearby Bluetooth devices. To ensure the privacy of other devices, each Bluetooth address was appended with a secret pepper and afterwards hashed with a non-invertible hash function. Nearby devices are scanned once, for 10 seconds, every time the user unlocked the device.
Application	Stored the package name of the application which ran in the foreground (i.e. the application used by the user).
Device Events	
Ring mode	Stored the current ring mode of the device. Possible values are "Normal", "Silent" or "Vibrate".

Out of memory	Whenever the device had less than 100 MB of free storage, the data collection service was forced to stop and this event was logged.
Device lock/unlock	Each lock and unlock event of the device screen was logged.
Battery percentage	Each battery level change (i.e., percentage change) was stored.
Charging state	The timestamp of the event when the device was charging or stopped charging was logged. The dataset also includes whether the charging was tethered or not.
Brightness and brightness mode	Stored the screen level brightness and the brightness mode of the device (manual or automatic - the device adjusted the screen brightness based on the ambient illumination).
Screen orientation	Orientation of the device. Possible values are: 0, 90, 180 or 270 degrees.