UNIVERSITY of DENVER

AffectNet Database

AffectNet contains about 1M facial images collected from the Internet by querying three major search engines using 1250 emotion related keywords in six different languages. About half of the retrieved images (**~420K**) are manually annotated for the presence of seven discrete facial expressions (categorical model) and the intensity of valence and arousal (dimensional model). **The rest of the images (~550K) are automatically annotated using ResNext Neural Network trained on all manually annotated training set samples with average accuracy of 65%**. **AffectNet** is by far the largest existing database of facial expressions, valence, and arousal in the wild enabling research in automated facial expression recognition in two different emotion models.

AffectNet provides:

- Images of the faces
- Location of the faces in the images
- Location of the 68 facial landmarks
- Eleven emotion and non-emotion categorical labels (Neutral, Happiness, Sadness, Surprise, Fear, Disgust, Anger, Contempt, None, Uncertain, No-Face)
- Valence and arousal values of the facial expressions in continuous domain

We refer readers to our paper for details about the data collection, annotation, and model training.

Emotion categories:

Eleven annotated emotions are provided for images and indexed as follows: 0: Neutral, 1: Happiness, 2: Sadness, 3: Surprise, 4: Fear, 5: Disgust, 6: Anger, 7: Contempt, 8: None, 9: Uncertain, 10: No-Face

Number of manually annotated images in the training and validation set is shown in Table below:

Neutral	75,374
Happy	134,915
Sad	25,959
Surprise	14,590
Fear	6,878
Disgust	4,303
Anger	25,382
Contempt	4,250
None	33,588
Uncertain	12,145
Non-Face	82,915
Total	420,299

Valence & Arousal:

Valence and arousal values are provided as floating point numbers in the interval [-1,+1].

Images:

Provided images are result of 15% boundary expansion of OpenCV face detector. OpenCV face boundary boxes are provided in CSV files.

How to download the database:

AffectNet can be downloaded from here (the link changes every 30 days for security reasons).

Note that size of the database is about 122GB. It may take 1 or 2 days to download the entire database depending on your internet speed.

RAR files:

The images are compressed in RAR format (in both manually annotated and automatically annotated sets). Researchers should extract the compressed files after downloading the database. We suggest RAR/UNRAR packages for

Linux/Mac users and WinRAR application for Windows users for extracting the compressed files.

File lists:

Three file lists training.csv, validation.csv, and automatically_annotated.csv are provided. Instances in the training.csv and validation.csv refer to the images in the Manually_Annotated_compressed folder. Instances in the automatically_annotated.csv refer to images in Automatically_annotated_compressed folder.

Test set is not released at this moment. We are planning to organize a challenge on AffectNet in near future. Test set will be used to evaluate the participants.

The following tables give the results of our experiments on the validation set using our baseline methods (described in our paper) trained on the training set. We suggest researchers use the validation results as a baseline for comparison until the test set is released.

Categorical model:

	Down-Sampling	Up-Sampling	Weighted-Loss
ACCs	0.50	0.47	0.58
F1s	0.49	0.44	0.58
KAPPAs	0.42	0.38	0.51
ALPHAs	0.42	0.37	0.51
AUCPR	0.48	0.44	0.56
AUC	0.47	0.75	0.82

Dimensional model:

	Valence	Arousal
RMSE	0.37	0.41
CORR	0.66	0.54
SAGR	0.74	0.65
CCC	0.60	0.34

CSV structure:

The provided CSV files contain the following attributes:

- File path: subdirectory and file name of the image.
- Face_x: x coordination of the location of the face in the image.
- Face_y: y coordination of the location of the face in the image.
- Face_width: width of the detected face in the image.
- Face_height: height of the detected face in the image.
- Facial_landmarks: coordination (x and y) of the 68 detected facial landmarks. The x and y coordination are separated with a semi-colon (;) and have the following structure:
  x1;y1;x2;y2;x3;y3 …. x67;y67;x68;y68
- Expression: expression ID of the face (0: Neutral, 1: Happy, 2: Sad, 3: Surprise, 4: Fear, 5: Disgust, 6: Anger, 7: Contempt, 8: None, 9: Uncertain, 10: No-Face)
- Valence: valence value of the expression in interval [-1,+1] (for Uncertain and No-face categories the value is -2)
- Arousal: arousal value of the expression in interval [-1,+1] (for Uncertain and No-face categories the value is -2)

All papers that either uses AffectNet fully or partially or is based on the analysis / study of AffectNet database MUST cite the following paper:

Citation:

A. Mollahosseini; B. Hasani; M. H. Mahoor, "AffectNet: A Database for Facial Expression, Valence, and Arousal Computing in the Wild," in *IEEE Transactions on Affective Computing*, 2017.

BibTeX:
@ARTICLE{8013713,
author={A. Mollahosseini and B. Hasani and M. H. Mahoor},
journal={IEEE Transactions on Affective Computing},
title={AffectNet: A Database for Facial Expression, Valence, and Arousal

Computing in the Wild},
year={2017},
volume={PP},
number={99},
pages={1-1},}

The paper is in Arxiv: https://arxiv.org/abs/1708.03985