



FindingEmo: An Image Dataset for Emotion Recognition in the Wild

Laurent Mertens

NeurIPS, 2024

Co-Authors: Elahe' Yargholi, Hans Op de Beeck,
Jan Van den Stock, Joost Vennekens

What is FindingEmo?



A dataset comprised of 25,869 annotations for as many images.

Target: Emotion Recognition

Dimensions: Valence, Arousal, Emotion (Plutchik's Wheel of Emotions)

Table 1: Comparison of relevant datasets. “V/A/D” indicates which of the Valence, Arousal and Dominance dimensions were annotated.

Name	Nb. images	Image source	Annotation target	V/A/D	Emotions scheme	Reference
EMOTIC	23,571	COCO + Ade20k + internet	Single person	V/A/D	26 custom emotion categories	[24]
CAER-S	70,000	TV Shows	Single person (face visible)	–	Ekman’s 6 + neutral	[32]
HECO	9,385	HICO-DET + V-COCO + film + internet	Single person	V/A/D	Ekman’s 6 + Excitement and Peace	[27]
FindingEmo	25,869	Internet	Whole image	V/A	Plutchik’s Wheel of Emotions	This paper



Photo by David Shankbone, CC 3.0 (Wikimedia)

Gathering images

Scrape images using Duckduckgo, by generating an army of specific queries

	A	B	C
1	Emotion	Subject	Setting
2	accepting	children	anniversary
3	adoring	youth	beach
4	affectionate	adults	birthday
5	amused	seniors	circus
6	appreciative	people	competition
7	carefree	students	conference

Query = "Emotion subject setting"
e.g., "amused seniors circus"

Consider first 1000 results → apply some filtering → download remaining images.



#ComeToTheDuckSide

www.duckduckgo.com

Custom annotation interface



⑦ Keep/reject image?:

- Keep
- Reject

⑦ Tags:

- Bad quality photo
- Copyright
- Watermark
- No Interaction
- No people
- Text
- Not Applicable

⑦ Age group:

- Children
- Youth
- Young Adults
- Adults
- Seniors

⑦ Deciding factor(s) for emotion:

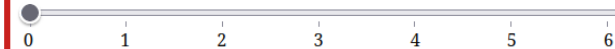
- Neutral
- Body language
- Conflict context vs. person
- Context
- Facial expression
- Staging

Menu

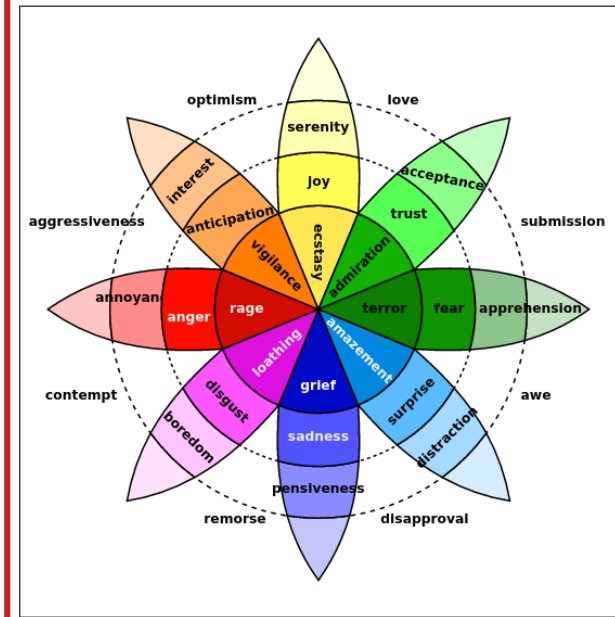
⑦ Negative/Positive:



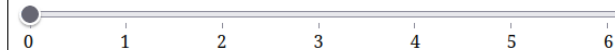
⑦ Intensity:



⑦ Main emotion:



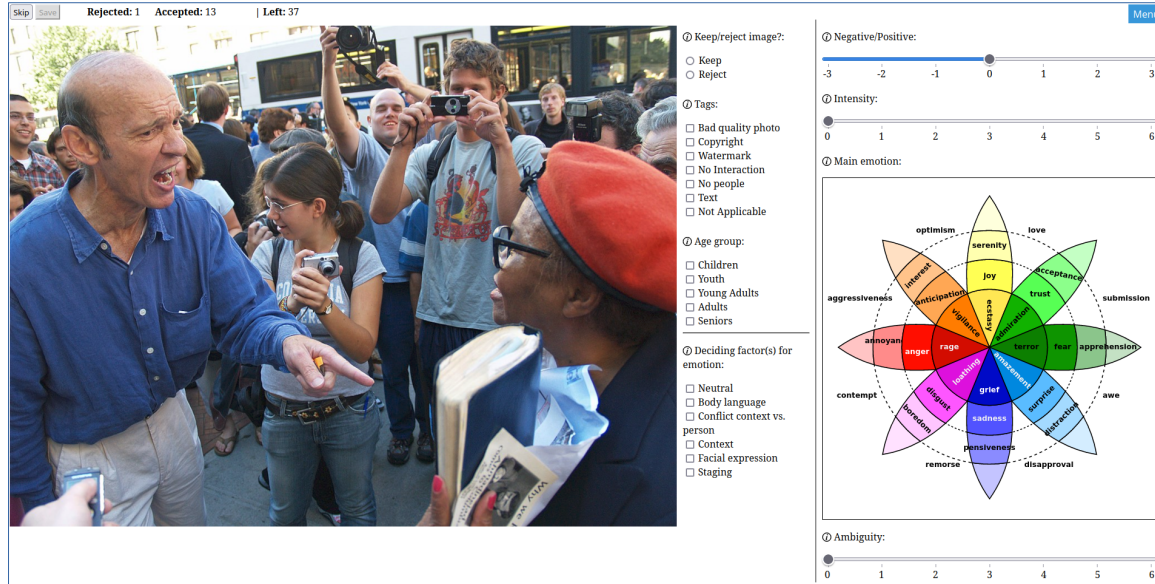
⑦ Ambiguity:



Main annotation dimensions



Custom annotation interface



The screenshot shows a web-based annotation interface. On the left, there is a photo of a man in a blue shirt talking to a woman in a red beret. The interface includes a top bar with 'Skip', 'Rejected: 1', 'Accepted: 13', and 'Left: 37'. Below the photo are several sections of controls:

- Keep/reject image?:** Radio buttons for 'Keep' and 'Reject'.
- Tags:** A list of checkboxes: 'Bad quality photo', 'Copyright', 'Watermark', 'No Interaction', 'No people', 'Text', and 'Not Applicable'.
- Age group:** Checkboxes for 'Children', 'Youth', 'Adults', and 'Seniors'.
- Deciding factor(s) for emotion:** Checkboxes for 'Neutral', 'Body language', 'Conflict context vs. person', 'Context', 'Facial expression', and 'Staging'.

On the right side of the interface, there are three sliders and a central emotion wheel:

- Negative/Positive:** A slider ranging from -3 to 3.
- Intensity:** A slider ranging from 0 to 6.
- Main emotion:** A circular emotion wheel with 16 segments, each labeled with an emotion: optimism, serenity, love, joy, acceptance, submission, trust, fear, apprehension, awe, indignation, disapproval, remorse, sadness, pensiveness, and grief.
- Ambiguity:** A slider ranging from 0 to 6.

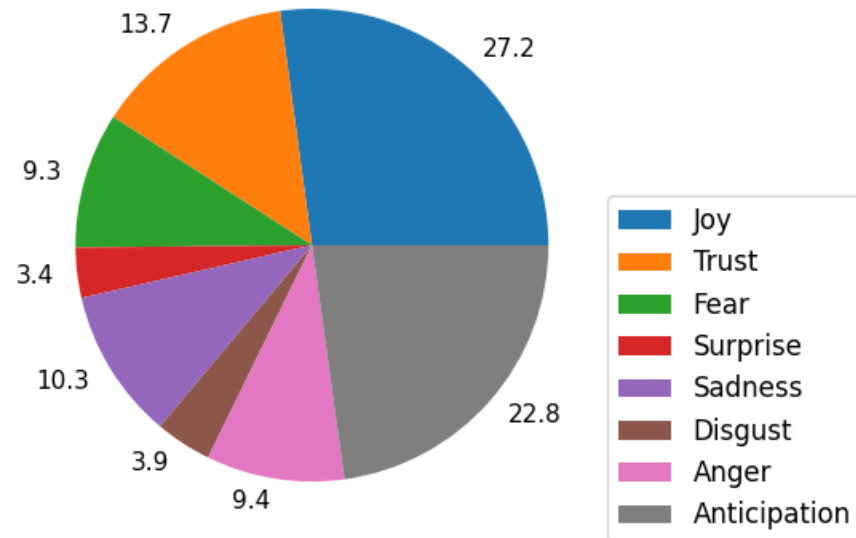
Annotator recruitment

www.prolific.com



655 annotators contributed

Emotion distribution



Paper sneak peak!

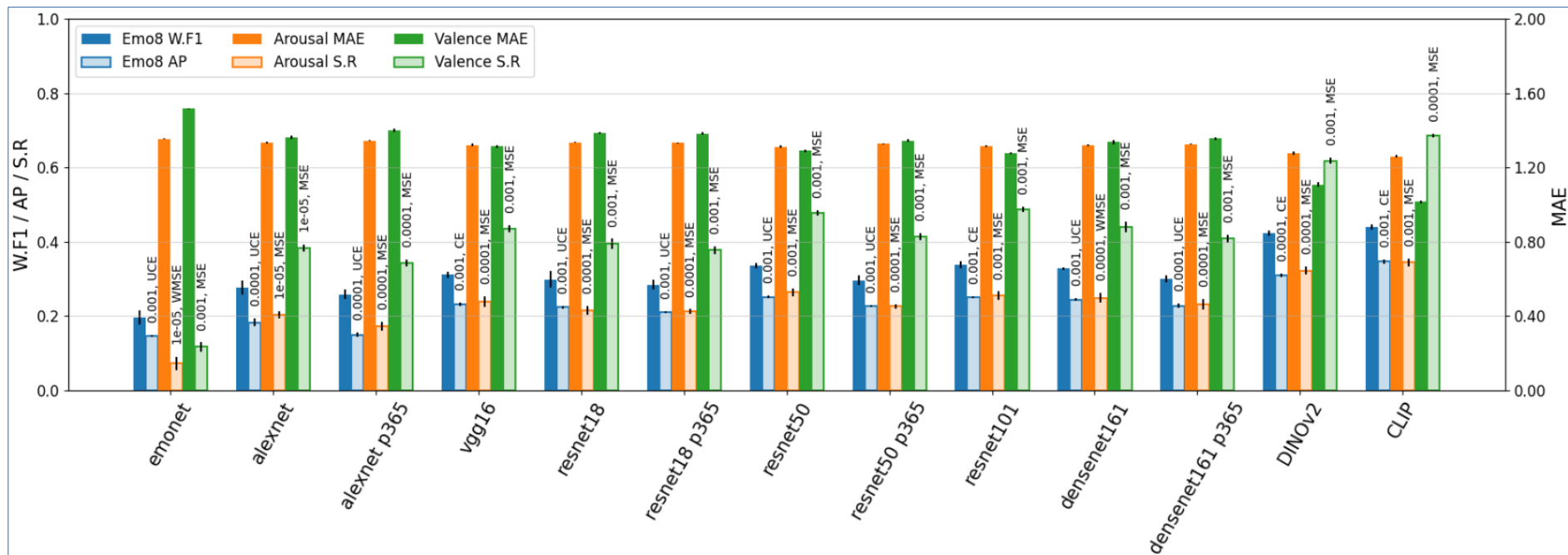


Figure 5: Test data baseline performance on the Emo8 classification and Arousal and Valence regression tasks. Metrics are: Weighted F1 (W.F1) and Average Precision (AP) for classification, and Mean Absolute Error (MAE) and Spearman R correlation coefficient (S.R) for regression. The starting learning rate and loss corresponding to each model are displayed above the training bars. (U)CE = (Unbalanced)CrossEntropyLoss, (W)MSE = (Weighted)MeanSquaredError loss, p365 = original model trained on Places365 dataset.



FindingEmo

<https://gitlab.com/EAVISE/lme/findingemo>



FindingEmo

Thanks!

<https://gitlab.com/EAVISE/lme/findingemo>