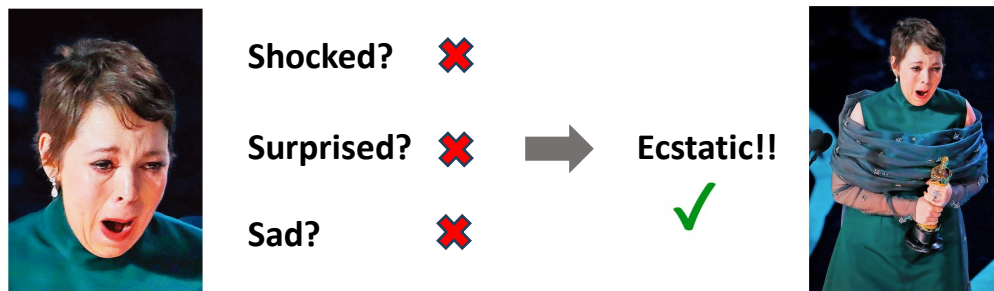# BEAT: Berkeley Emotion and Affect Tracking Dataset

Zhihang Ren    Jefferson Ortega    Yunhui Guo    Stella X. Yu    David Whitney
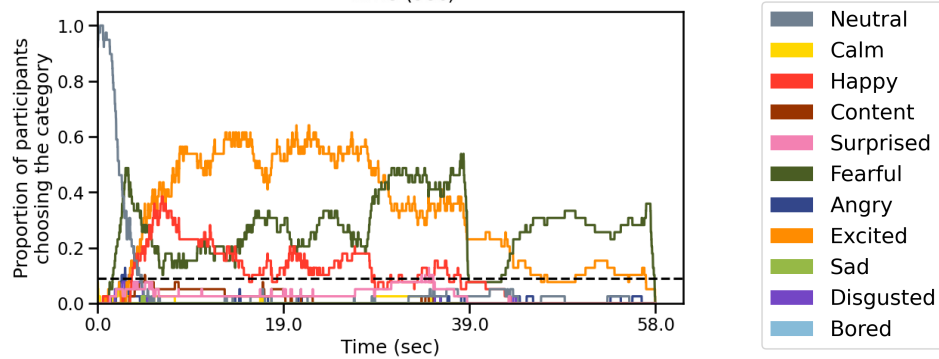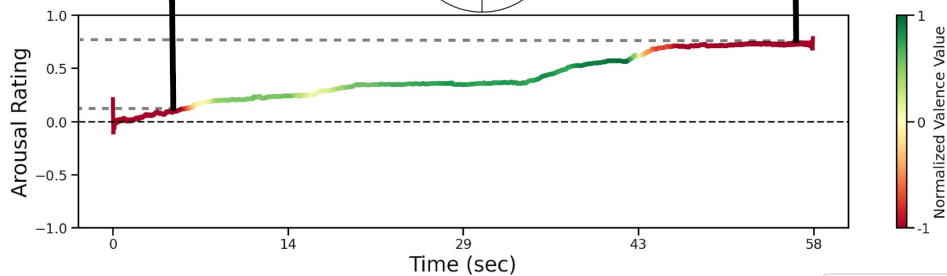
CVPR JUNE 17-21, 2024
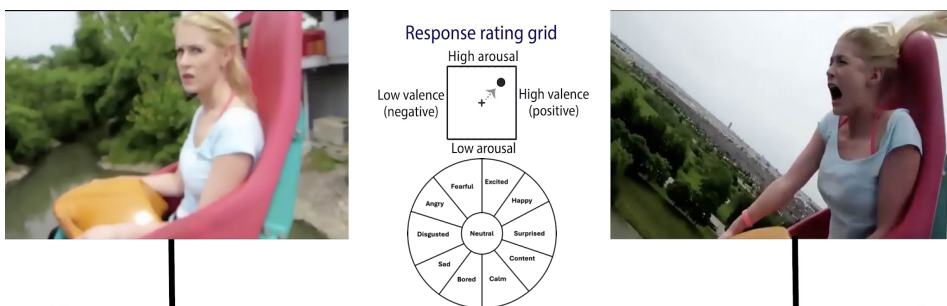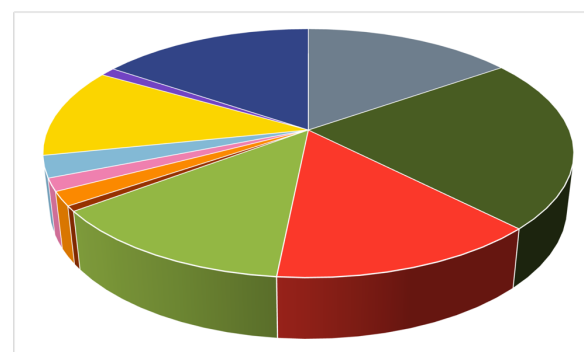
## Context is important in emotion recognition



Shocked? ✗
Surprised? ✗  →  Ecstatic!! ✓
Sad? ✗

## Annotation procedure for dynamic emotions



Response rating grid

High arousal

Low valence (negative) — High valence (positive)

Low arousal

## Comparison to other datasets

| Database | Annotation Type | Condition | # videos | Length of Videos | # Annotators | Context |
|---|---|---|---|---|---|---|
| RECOLA | Valence-Arousal (Continuous) | Controlled | 46 | 5min | 6 | ✕ |
| AFEW | 7 Basic Facial Expression | In-the-Wild | 1,809 | 0.3s - 5.4s | 3 | ✕ |
| Aff-Wild | Valence-Arousal (Continuous) | In-the-Wild | 298 | 6s - 14min28s | 8 | ✕ |
| Aff-Wild2 | Valence-Arousal (Continuous) | In-the-Wild | 260 | 4s - 15min4s | 4 | ✕ |
| CAER | 7 Emotion Categories | In-the-Wild | 13,201 | 1s - 5s | 6 | ✓ |
| EMOTIC | Mixed | In-the-Wild Image-based | \ | 18,316 images | 3-5 | ✓ |
| VEATIC | Valence-Arousal (Continuous) | In-the-Wild | 124 | 10s - 2min37s | 192 | ✓ |
| **BEAT (ours)** | **Mixed** | **In-the-Wild** | **124** | **10s - 2min37s** | **245** | ✓ |

## Distribution of Emotions



Annotator Feelings

Enjoyment (9)  4.98
Familiarity (5.0)  1.16



Density

Arousal ratings / Valence ratings

○ Hollywood movies
■ Documentaries
▲ Home videos
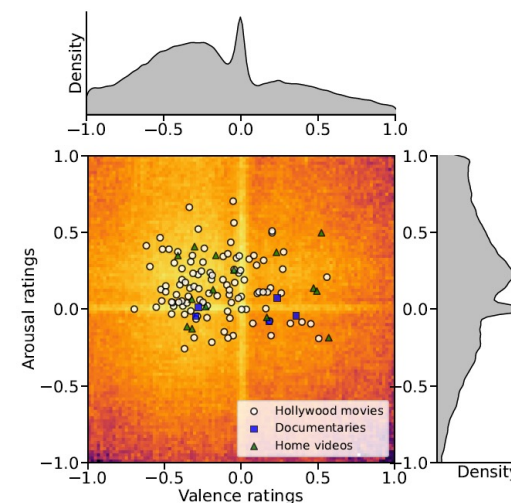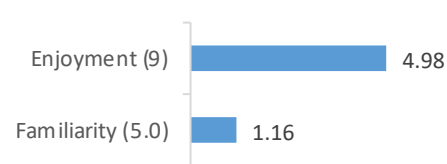
| Response Standard Deviation | |
|---|---|
| 5 annotators | all annotators |
| 0.094 | 0.028 |

## Dataset statistics

| Attribute | Description |
|---|---|
| No. of Frames | 257, 601 |
| No. of Videos | 124 |
| **Total No. of Annotators** | **245** |
| Avg. No. of Annotators per video | 56 |
| Length of Videos | 10 s - 2 min 37s |
| Mean Image Resolution | $854 \times 480$ |
| Hollywood movies | 104 |
| Documentaries | 5 |
| Home videos | 15 |

## Sample Videos with Annotations



Neutral   Happy   Sad   Surprised

Happy   Fearful

Neutral   Sad   Calm   Sad