Exploring Contextual Engagement For Trauma Recovery

Svati Dhamija, Terrance Boult
University of Colorado Colorado Springs
{sdhamija,tboult}@vast.uccs.edu

Vision And Security Technology
University of Colorado Colorado Springs

The 1st International Workshop on Deep Affective Learning and Context Modeling in conjuction with CVPR 2017
CVPR July 21-26 2017 HONOLULU

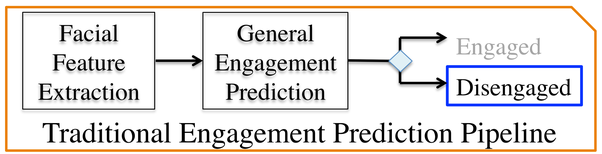
Why Engagement for Trauma Recovery?

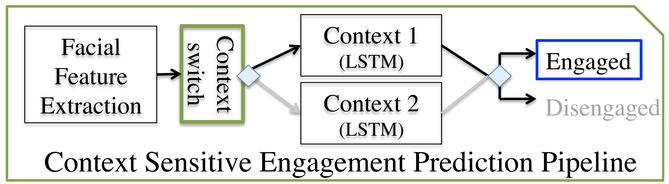
1. Each year, over 3 million people in the United States are affected by post-traumatic stress, a serious mental condition.
2. Personalization and automated adaption in self-help websites can positively aid people with mental health issues.
3. Self-reported user engagement with self-help websites has been found, in many psychology studies, to be highly correlated with outcomes.
4. Vision and learning based methods provide a proactive, scalable and cost-effective web-based treatment for trauma recovery by analyzing webcam feeds.
5. The website and task can adapt to enhance or maintain engagement and recovery based on reliable and quantitative engagement measurement.
6. Context specific modeling is more accurate than a generic user engagement model.

Contributions

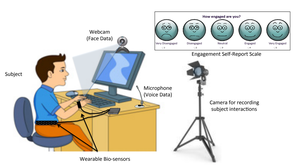
1. Exploration of engagement in two contextually different tasks within the recovery regime: "Relaxation" and "Triggers".
2. Developing automated engagement prediction methods based on automatically computed AUs using LSTMs. We train/test this on both subtasks within **trauma recovery**.
3. We build context-specific, cross-context and mixed prediction models and show the importance of context in predicting engagement from facial expressions.
4. Exploring relationship of subject's mood as an initialization parameter for engagement estimation.

Contextual Engagement Pipeline



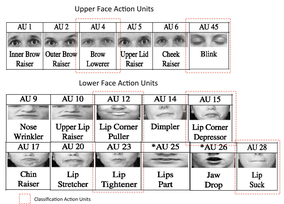


EASE Dataset



	SESSION 1							
	MODULE 1				MODULE 2			
	Task	Number of Videos	Number of Frames	Number of Self Reports	Task	Number of Videos	Number of Frames	Number of Self-Reports
Trigger Task followed by Relaxation Task	Trigger	52	806855	166	Relaxation	52	1579927	122
Relaxation Task followed by Trigger Task	Relaxation	43	1391803	91	Trigger	43	590953	98

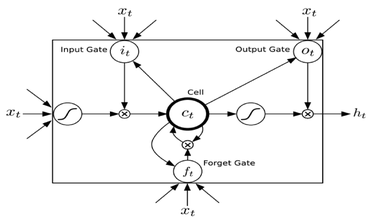
	SESSION 2							
	MODULE 1				MODULE 2			
	Task	Number of Videos	Number of Frames	Number of Self Reports	Task	Number of Videos	Number of Frames	Number of Self Reports
Trigger Task followed by Relaxation Task	Trigger	33	454510	94	Relaxation	33	1553409	63
Relaxation Task followed by Trigger Task	Relaxation	47	1139996	105	Trigger	47	544298	149



Dataset Details

1. Web-intervention used for trauma-recovery : ***http://ease.vast.uccs.edu/***.
2. Participants consisted of total 110 subjects with 88 Female, 17 Male, 5 did not specify in the age group of 18-79 years, with 80% being under the age of 46.
3. Study comprised of 3 Sessions using six Modules of trauma-recovery: Relaxation, Triggers, Social-Support, Self-Talk, Professional-Help, Unhelpful-Coping.
4. Face data, audio and sensory data (skin conductance, respiration, ECG) was captured from subjects while they were interacting with website while performing self-regulation exercises.
5. Self-Reports were collected from subjects about their engagement level, mood (Very Short Profile of Mood States (POMS) questionnaire) and self-efficacy measures.

Long Short-Term Memory

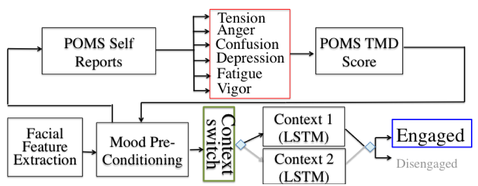


1. LSTM is a recurrent neural network that allows data to flow both forwards and backwards with explicit memory cells for storing information.
2. We model the engagement prediction problem as a sequence learning problem. Each sequence corresponds to an engagement self-report label provided by trauma subjects.

Mood-Aware Contextual Engagement

1. Does using current mood as a pre-conditioner improve engagement prediction for a given task?
2. Our very short POMS measure has first 24 questions from POMS-SF. Clustered into 5 negative sentiments: tension (5), depression (6), anger (5), fatigue (2), confusion (2) and 1 positive sentiment: vigor (4).

$$\text{POMStmd} = \frac{1}{21.1} \sum_{x \in \text{neg. sent.}} n(x) - \sum_{x \in \text{pos. sent.}} p(x)$$



Results and Conclusion

	RX	TR
Chance	32.8%	38.3%
SVC	31.4 ± 7%	35.2 ± 9%
LSTM	39.1 ± 8.8%	50.7 ± 11%

	RX - Test	TR - Test
RX - Train	39.1 ± 8.8 %	38.1 ± 4.4%
TR - Train	36.7 ± 3.3%	50.7 ± 11%
(RX + TR) - Train	39.5 ± 6.1%	49.1 ± 7.6%

	RX	TR
LSTM Baseline	0.9989	0.7653
POMS-aware LSTM	0.9493	0.6786

Experiment Details

1. **Features:** 20 Facial Action Units (AUs) are extracted using OpenFace Framework each from 30 sec video segment (900 frames/segment).
2. **Algorithms:** Support Vector Classification and LSTM based classification for Engagement Prediction
3. **Learning Protocol for Contextual Engagement Models:**
   - Training on Relaxation Module, Testing on Trigger Module & Relaxation Module
   - Train on Trigger Module, Testing on Trigger Module & Relaxation Module
   - Training on mixed (Relaxation & Trigger), testing on Trigger Module & Relaxation Module
4. **Mood-Aware Engagement Prediction**: Leave-One-Subject-Out methodology: Training on Relaxation Module & testing on Relaxation Module, Training of Trigger Module & Testing on Trigger Module

Conclusions

- User engagement is highly contextual and LSTM-based models can be used to study task specific facial behaviors.
- This work is the first step towards task-specific engagement.
- Adding subject's mood to the AU data demonstrates a clear improvement in engagement prediction performance.

Acknowledgements: The work support in part by NSF Research Grant SCH-INT 1418520