Detecting mental fatigue from eye-tracking data gathered while watching video:

Evaluation in younger and older adults.

Presented by: Gentry Atkinson

Paper & Authors

- Yasunorid Yamada, Masatomo Kobayashi
- IBM Research- Tokyo
- Accepted: June 26, 2018
- Journal: Artificial Intelligence in Medicine
- Journal's H-Index: 74 (compare to CVPR's H-Index of 96)

Introduction

- "Health monitoring technology in everyday situations is expected to improve quality of life and **support aging populations**."
- Many fatigue studies using biometric data have only gathered data from **younger populations**.
- Fatigue should also be **measurable during natural actions** with unobtrusive technology.
- A **novel feature** set is proposed to assess fatigue from eye tracking data and is evaluated in 2 experiments.

Impact & Contribution

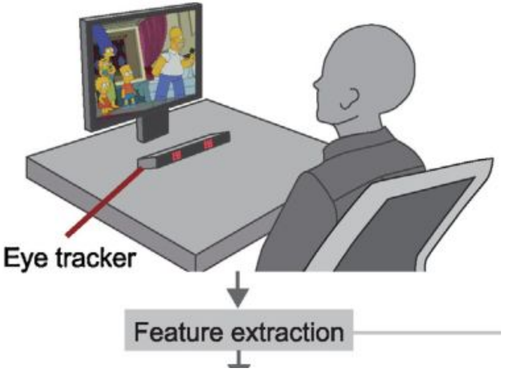
- "Mental fatigue has been suggested as one of the most frequent causes of accidents and **errors in the workplace**." [*A multifaceted investigation of the link between mental fatigue and task disengagement*, 2015]
- "Fatigue-related accidents and errors in the US may reach as a high as **$31.1 billion**." [*The associations of insomnia with costly workplace accidents and errors: results from the American insomnia survey*, 2012]

- A model to detect mental fatigue in young and old adults.
- Eye tracking data from adults watching video clips before and after cognitive tasks.
- Demonstration that this model can detect mental fatigue from cognitive tasks despite age differences.

Definitions

- **Mental Fatigue**: "...the feeling that people might experience during or after cognitive activities."
- **Saliency**: Describes a thing as being particularly noticeable.
- **Support Vector Machine:** A classifier which non-linearly separates learned groups by mapping their inputs to a higher-dimension feature space.
- **Fixation and Saccade**: I sincerely hope that these are familiar to you.

Eye Tracking in Natural Viewing



- Eye tracking data was gathered while participants were **viewing video clips**.
- Participants were told to **watch videos naturally** and were not explicitly informed about eye tracking.
- Videos were shown on a 20 inch screen at 30Hz with a resolution of 1600x1200. Participants were **85cm from the screen**.
- Eye movement and pupil data was collected with an **EMR ACTUS** device at **60Hz sample rate**. Calibrated with 9 points for every recording phase. (~17ms/sample)

Fatigue Detecting Model

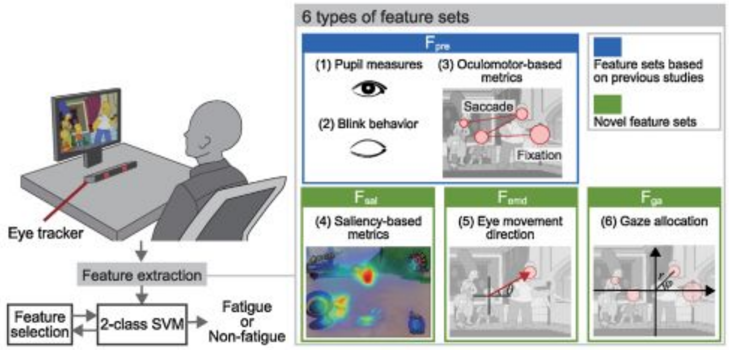


Fig. 1. Overview of our fatigue-detection model. Our model first extracts six types of feature sets from eye-tracking data collected while participants watch video. Using a subset of the features selected by a feature selection method, a two-class classifier using support vector machine (SVM) model estimates whether that person is fatigued or not.

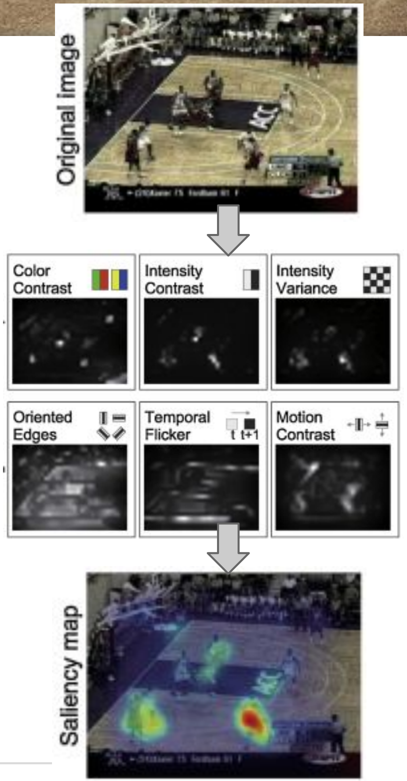
- Eye tracking features are gathered in 6 sets: 3 established and 3 novel.
- Features are selected from the total set by **Recursive Feature Elimination** to remove age dependent features.
- The selected features are used as inputs to train an SVM which classifies samples as "fatigued" or "not fatigued".

Established Feature Extraction

- **Set 1 Oculomotor-based**: saccade amplitude, saccade duration, saccade rate, fixation duration, etc. **9 total features**.
- **Set 2 Blinks**: blink rate, blink duration, etc. **7 total features.**
- **Set 3 Pupil Measures**: pupil diameter, constriction velocity, etc. **6 total features.**

Novel Feature Extraction

- **Set 4 Gaze Allocation**: radial coordinates were binned for fixations and for all eye movements. Probabilities were computed for each bin and used as features. **72 total features.**
- **Set 5 Eye-Movement Direction:** directions were binned in a similar fashion to gaze allocation. **50 total features.**
- **Set 6 Saliency:** a composite measure of saliency is developed from 6 individual measures. The Area Under the Curve the salliency for saccade endpoints is then computed for a participant.



Experiments

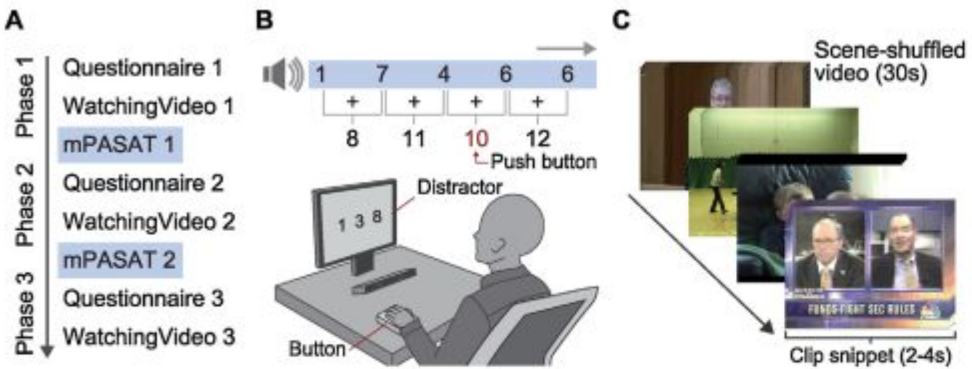


Fig. 3. Experimental setup: (A) overall procedure, (B) mental calculation task called mPASAT, and (C) examples of scene-shuffled video clips.

- 2 experiments were conducted: 1 showing that **the model can detect mental fatigue** the other showing that mental fatigue is induced by the tests **not just by watching videos**.
- The cognitive test used is the modified paced auditory serial attention test (mPASAT). Participants listen to a series of numbers and press a button when two consecutive numbers sum to 10/

Experiment 1 Design

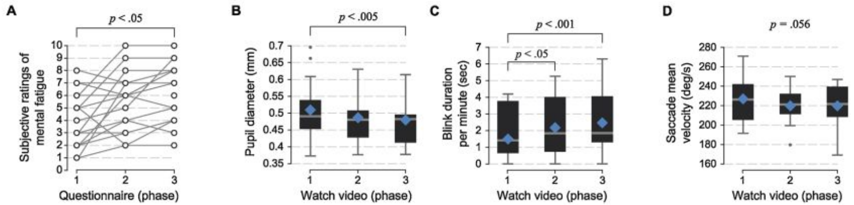
- Data is collected from 20 participants. 8 and 12 male. Mean age: 47.5 years with a SD of 20.5 years. 9 participants > 50 years of age. 2 were excluded for calibration errors. **N=18**.
- **H_0**: a classifier will be no more likely to detect fatigue after the test as before.
- **H_1**: a classifier will be more likely to detect fatigue after the mPASAT test.
- Participant data is collected during 3 30 second videos divided by 2 17-minute cognitive tests. **Independent Variable:** test is administered 0, 1, or 2 times.
- Results are calculated **within** the participant population.

Experiment 2 Design

- Additional data was collected from 11 participants. 2 female and 9 male. Mean age 29.7 years with a SD of 9.8 years. **N=11.**
- **H_0:** the new group will experience the same levels of fatigue as the first.
- **H_1:** the new group will have a measurably different level of fatigue from the first.
- Participants in this group watched 3 30-second video clips and eye data was gathered. **Independent variable**: cognitive tests are given or not.
- Results are calculated on comparisons **between** groups 1 and 2.

Experiment 1 Results

- Dunn multiple comparison post-hoc test shows increased fatigue from phase 1 to 3 with **$p<0.05$.**
- Post-hoc analysis shows decreased pupil diameter from phase 1 to 3 **$p<0.005$.**
- Post-hoc analysis of blink duration shows an increase from P1 to P3 with **$p<0.001$.**
- "Our hypothesis was that reflexive eye movement guided by bottom-up attention **increases with mental fatigue**." ANOVA with post hoc Bonferroni multiple comparison supports this with $p<0.001$.
- Overall SVM classifier accuracy increases from 77.1% for established features to 91.0% with novel features.



Model	Detection performance (%)			
	Accuracy	Precision	Recall	F-measure
F_{pre}	77.1	78.6	72.9	75.6
$F_{pre} + F_{sal}$	80.7	79.4	83.0	81.0
$F_{pre} + F_{smd}$	82.9	83.2	82.4	82.7
$F_{pre} + F_{ga}$	84.7	84.6	84.9	84.7
$F_{pre} + F_{sal} + F_{smd} + F_{ga}$	**91.0**	**91.4**	**90.3**	**90.8**

Experiment 2 Results

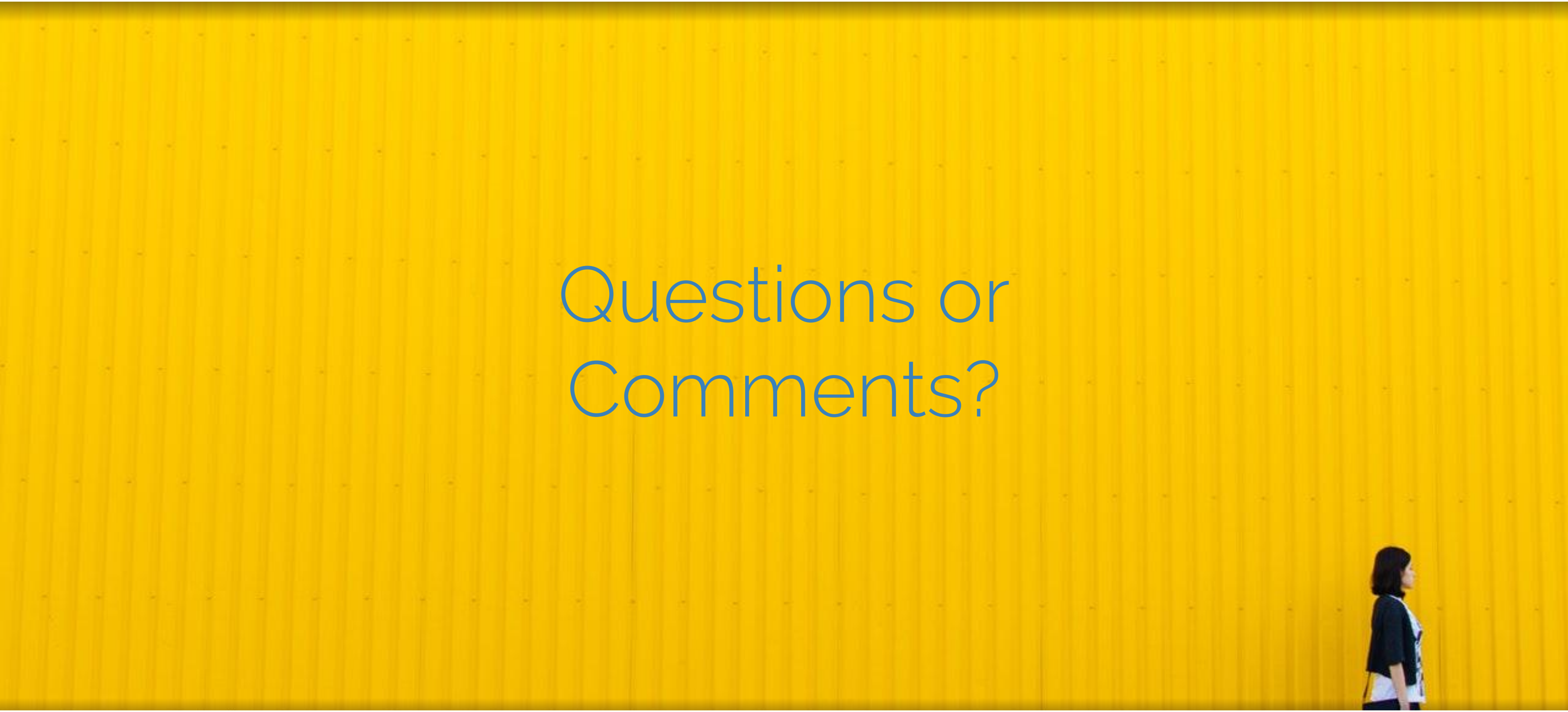
- Friedman Non-Parametric ANOVA was performed on subjective ratings for Group 2 and One-way Repeated Measures ANOVA was performed on eye data for Group 2. No significant difference was found **$p>0.05$.**
- Features were extracted from Phases 1 and 3. The fatigue detection model was then trained on Group 2 as described in Experiment 1. This model classified **91.9%** of the new samples as **"Not Fatigued".**
- This suggests that the model can specifically detect fatigue induced by mental tasks and not from the effects of watching videos.

Discussion & Conclusion

- We can show that the new features significantly improve fatigue detection based on 30 seconds of video watching.
- We have shown that the model can specifically distinguish between fatigue induced by cognitive work and by prolonged viewing.
- Age dependent features can be removed in an automated fashion to make a model robust across a range of ages.
- Only 14 out of 18 participants reported fatigue on the questionnaire even when the data indicated that they were fatigued.

Future Work

- This work included a small number of participants.
- Data was collected in a lab setting.
- The video clips used had limited content.
- Fatigue was treated as a binary condition. A model that rates a participants fatigue on a scale will be more useful for individual health.
- (All of these points are listed in the paper.)


Questions or
Comments?