

**Air traffic controllers' mental fatigue recognition:
A multi-sensor information fusion-based deep
learning approach**

Xiaoqing Yu, Chun-Hsien Chen, Haohan Yang *
School of Mechanical and Aerospace Engineering,
Nanyang Technological University 639798, Singapore

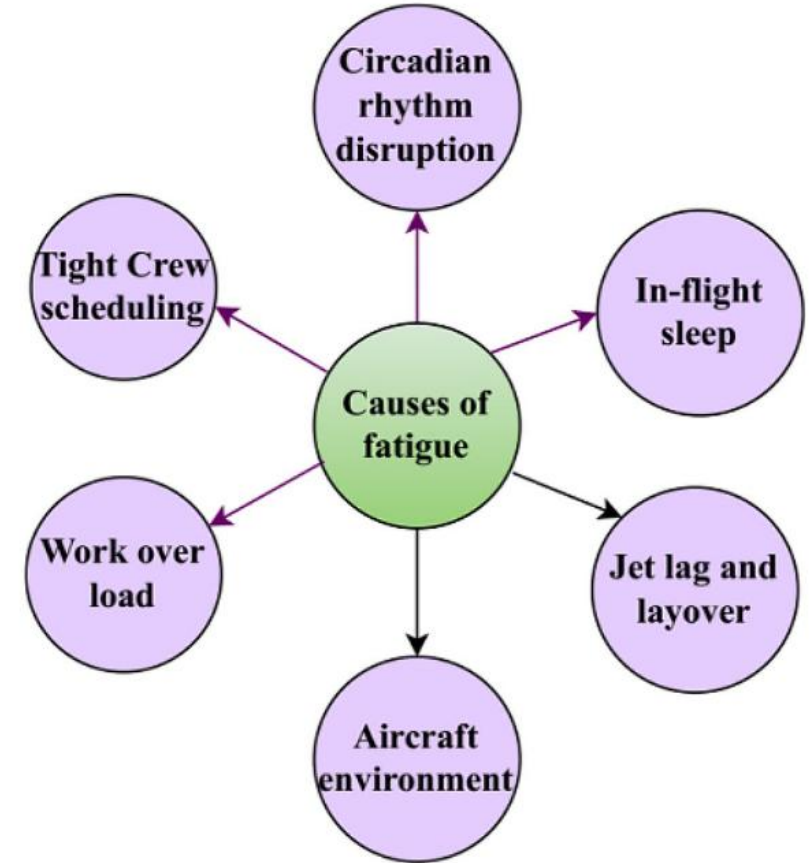
Advanced Engineering Informatics

Introduction

- Air traffic controllers: monitoring radar systems, communicating with pilots, and determining aircraft's location, altitude, and speed, among other tasks.
- Mental fatigue: drastically reduce operators' situation awareness and slow down reaction time

Previous Methods

- rated by subjective measures
- monitoring the operators' eye movements, electroencephalogram (EEG), heart rate, among others.
- RNN and LSTM



Introduction

- EEG: 脑电图
- EMG: 肌电图
- ECG: 心电图
- EDA: 皮肤电反应
- Respiration: 呼吸频率
- GSR: 皮肤电反应
- SKT: 皮肤温度

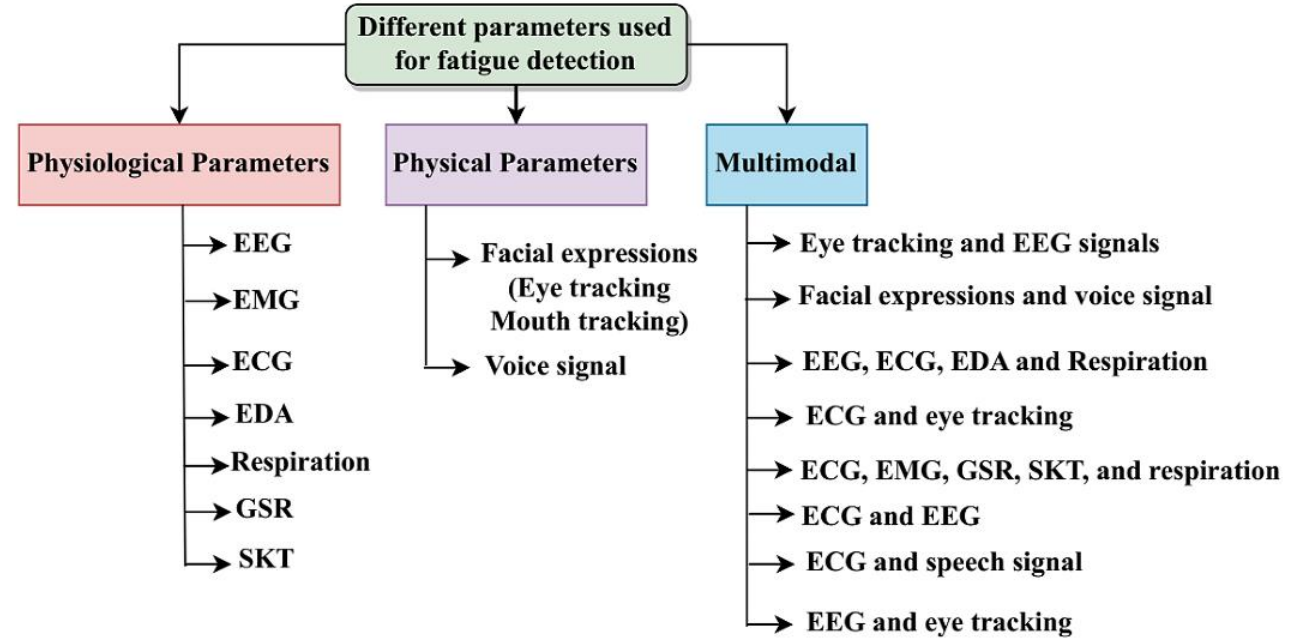
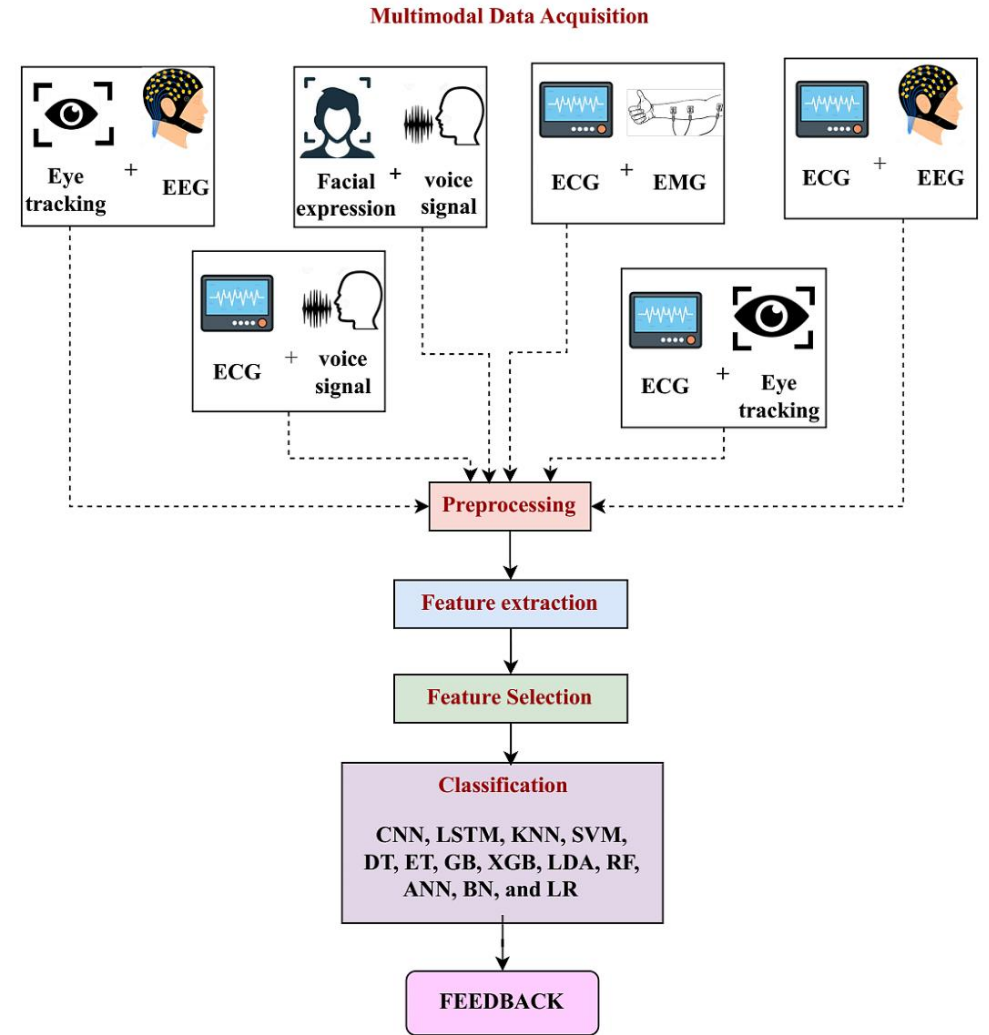
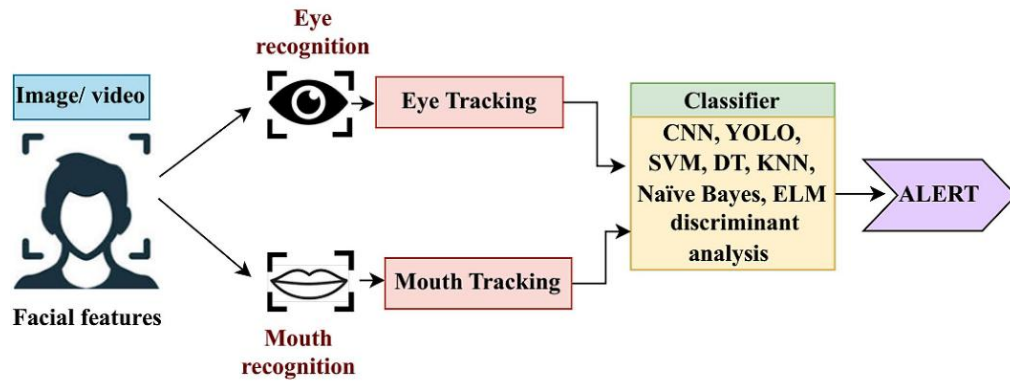
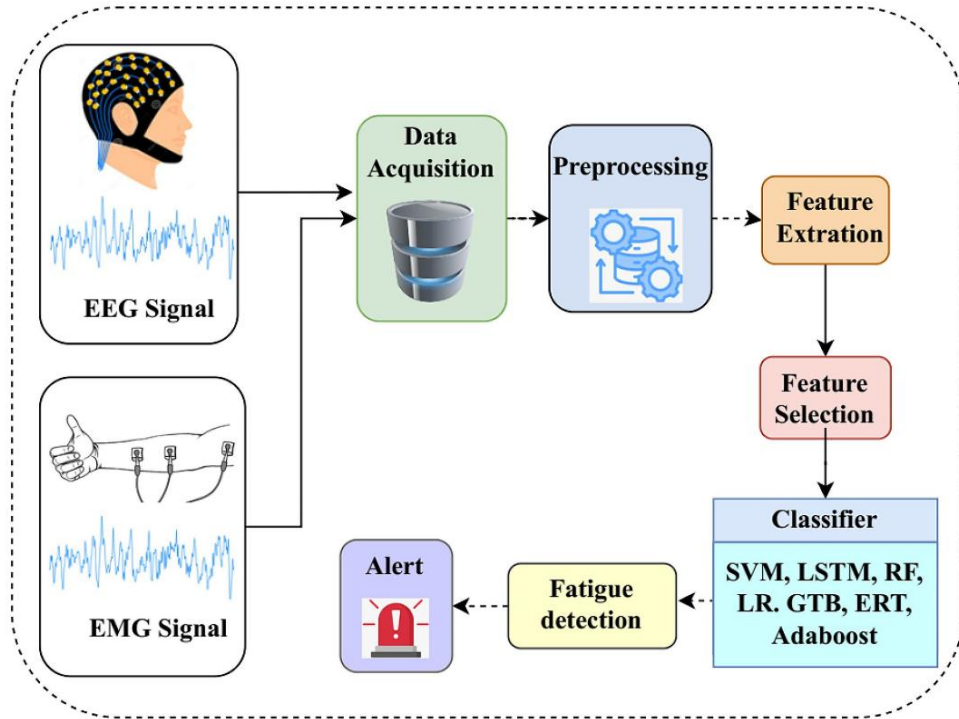


Table 4 Comparison of different parameters employed for fatigue recognition among aviation professionals

Parameters	Cost	Simplicity	Efficiency	Detection speed	Real-Time Deployment	References
EEG	High	Difficult	High	Fast	Low	[62–67]
EMG	High	Difficult	High	Fast	Low	[71]
Speech signal	Low	Easy	Medium	Slow	High	[78–80]
Facial expressions	Low	Easy	Medium	Slow	High	[82–84]
Eye tracking + EEG	Medium	Medium	High	Medium	Medium	[92, 93]
Facial expressions + voice signal	Low	Easy	Medium	Slow	High	[94]
EEG + EMG + GSR + SKT + respiration	Medium	Difficult	High	Fast	Low	[89]
ECG + EEG + EDA + respiration	High	Difficult	High	Fast	Low	[90]
Voice signal + ECG	Low	Medium	High	Medium	Medium	[91]
Eye tracking + ECG	Medium	Medium	High	Medium	Medium	[17]

Abbreviations: EEG: electroencephalogram; EMG: electromyography; GSR: galvanic skin response; SKT: skin temperature; ECG: electrocardiogram; EDA: electrodermal activity

Introduction



Method

- **Signal preprocessing**
 - EEG Signals: different frequency bands (delta, theta, alpha, beta)
 - Eye Movements: gaze positions and speed
- **1D CNN module**
 - extracting features
- **LSTM module**
 - processes the temporal dependencies
- **Attention mechanism**
 - focus on the more fatigue-related signals
- **Fatigue classification**
 - alert, moderate fatigue or Intensive fatigue state

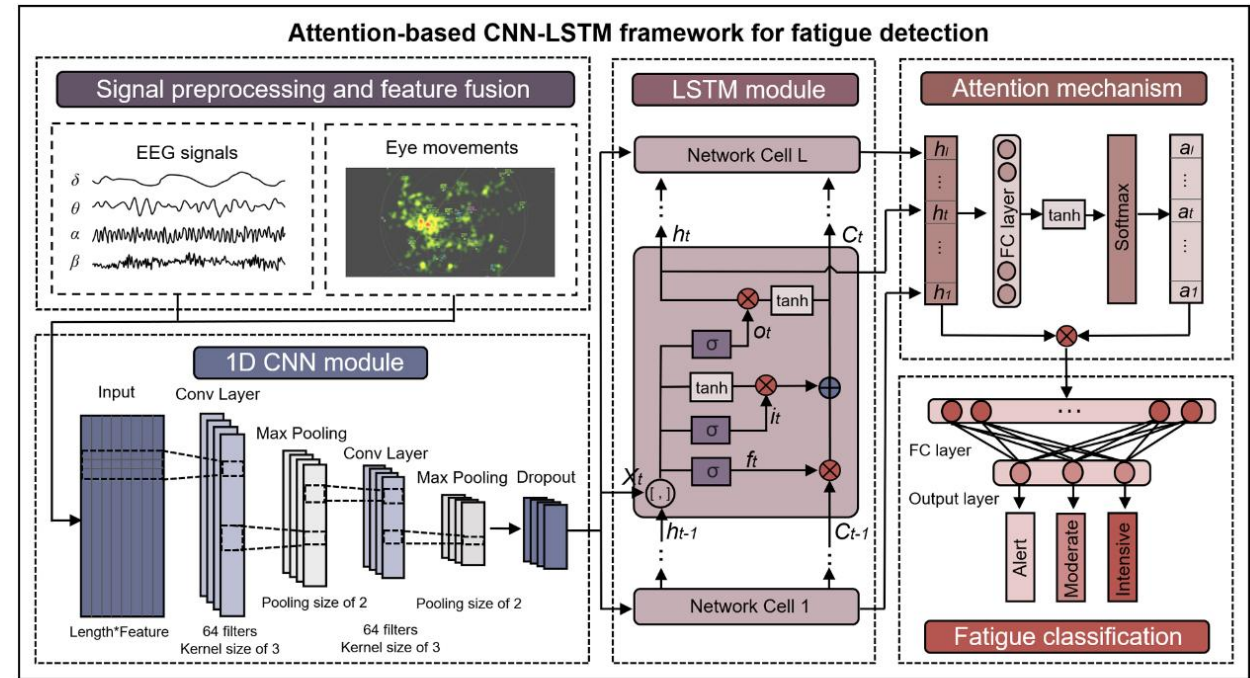
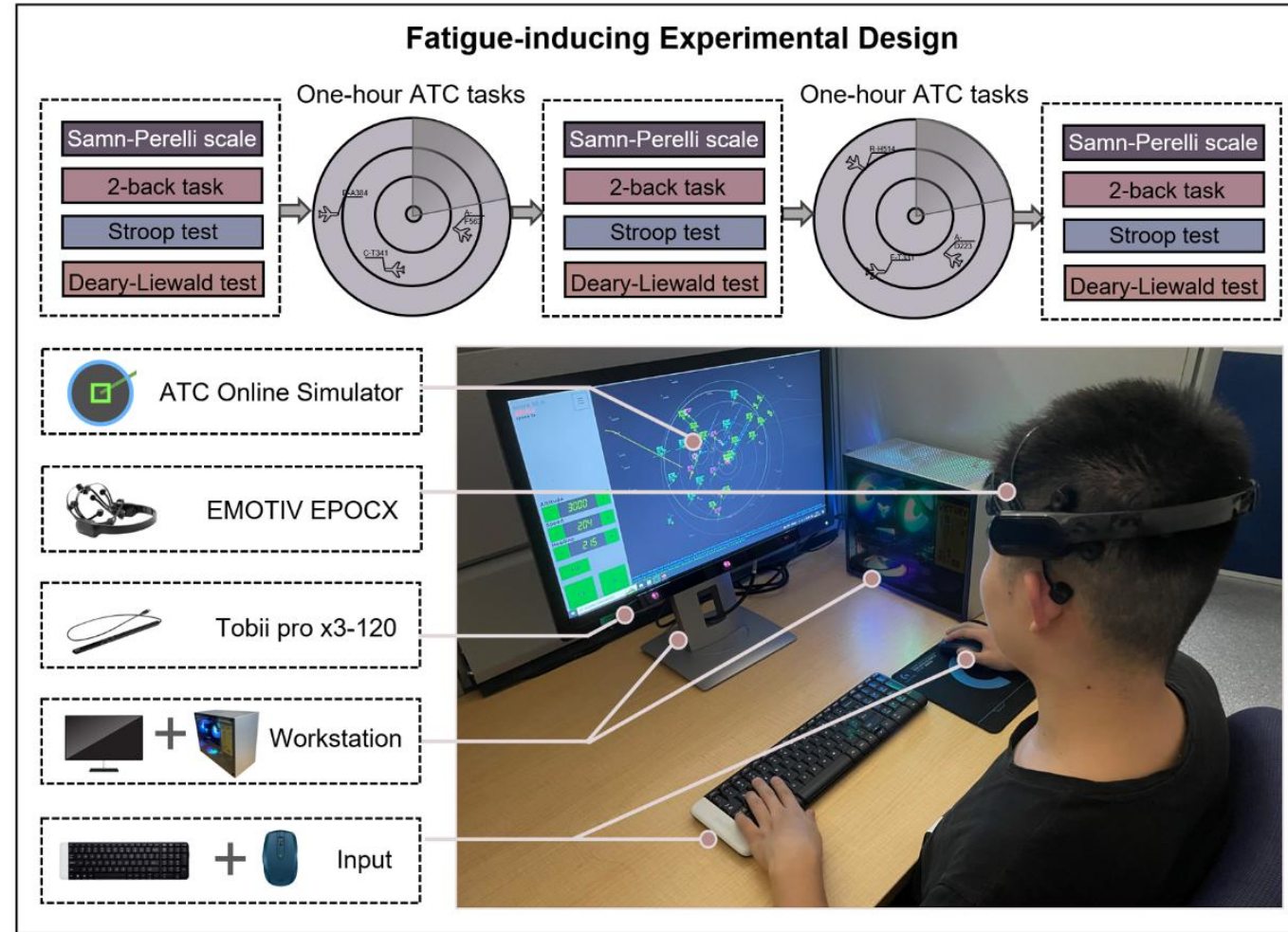


Fig. 1. Framework of the proposed multimodal fusion-based MF recognition scheme.

Fatigue-inducing experiment design

- **Familiarization Phase**
 - participants achieve 90% accuracy in ATC tasks
- **Fatigue Measurement**
 - Samn-Perelli scale to assess their subjective fatigue levels
- **Cognitive Tasks**
 - evaluate the impact of fatigue on cognitive performance (reaction time and accuracy rate)
 - 2-back test (working memory)
 - Stroop test (thinking)
 - Deary-Liewald test (visual concentration)
- **Task Execution Phase**
 - perform air traffic control tasks one hour (monitoring radar screens and guiding aircraft in and out of the airport)
 - monotonous and repetitive



Experiments

- **Evaluation metrics and baselines**
- accuracy (Acc), precision (Pr), recall (Re), F1 score
- Logistic Regression
- Random Forest
- SVM
- LSTM
- CNN-LSTM
- RecMF
- 5-fold cross-validation
- different window sizes (1s, 3s, 5s)

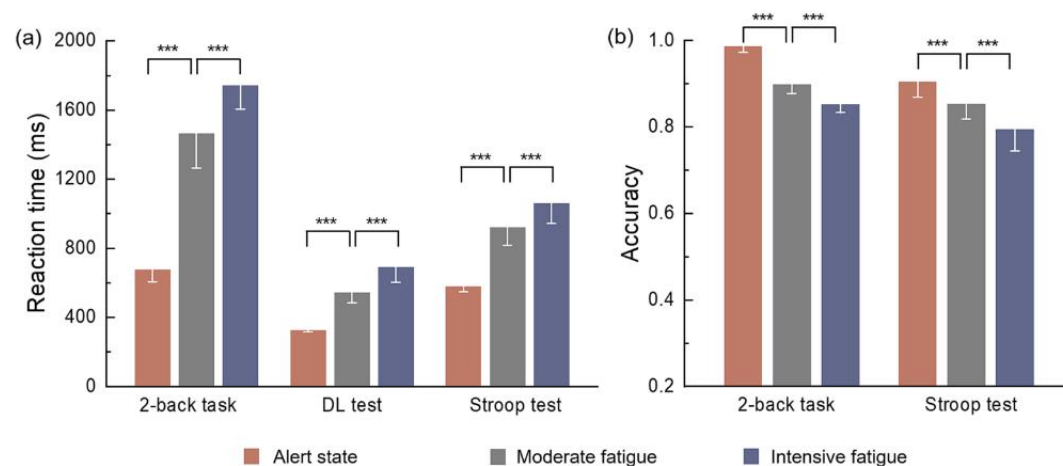


Fig. 3. Repeated measures ANOVA results of the effects of MF on cognitive performance during typical cognitive tasks, (a) reaction time, (b) accuracy. The error bar represents σ_s and the mark *** means $p < 0.001$.

Experiments

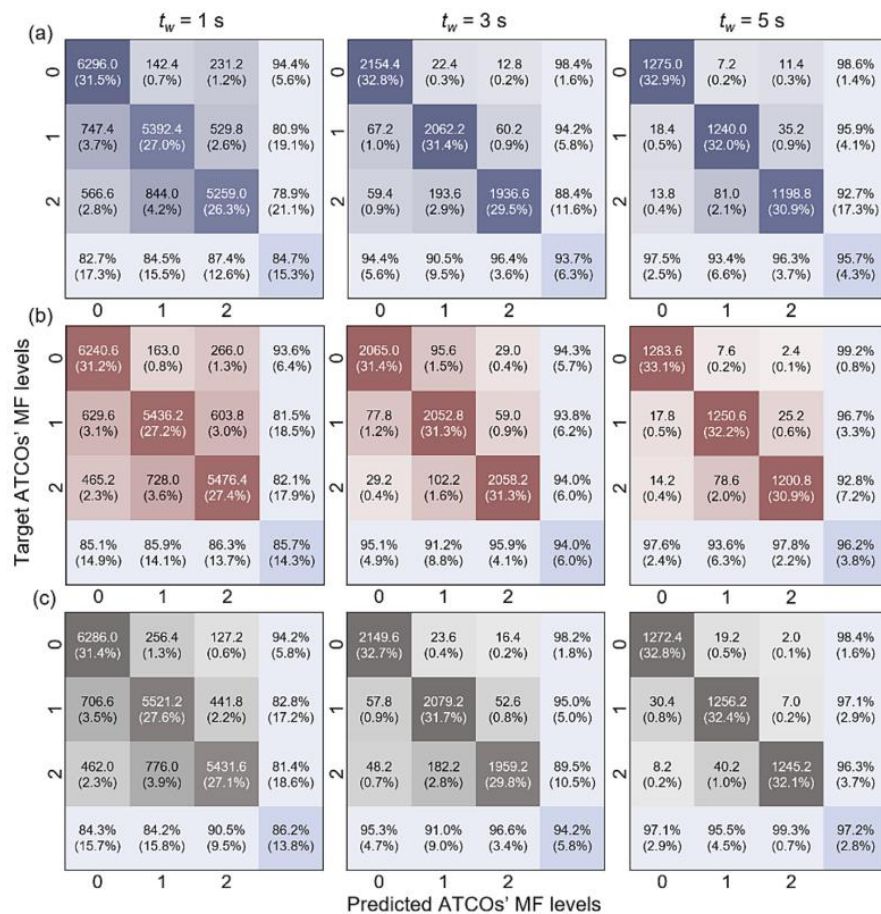


Fig. 5. Confusion matrices of MF recognition in different historical horizons during typical cognitive tasks, including (a) 2-back task, (b) DL test, and (c) Stroop test. The labels 0, 1, and 2 represent the alert state, moderate fatigue, and intensive fatigue, respectively.

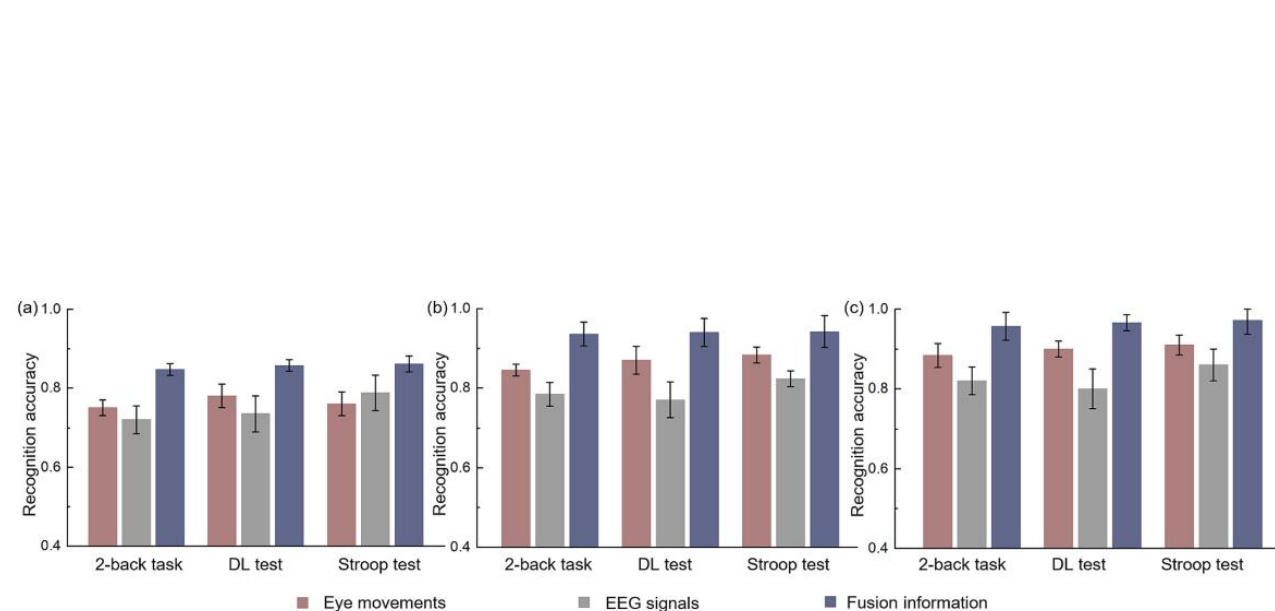


Fig. 4. Recognition accuracy of the RecMF model with temporal information across three cognitive tasks. The subfigures (a), (b), and (c) display the historical horizons of 1 s, 3 s, and 5 s, respectively. The error bar (set at 5σ) is used to provide a clearer comparison.

Experiments

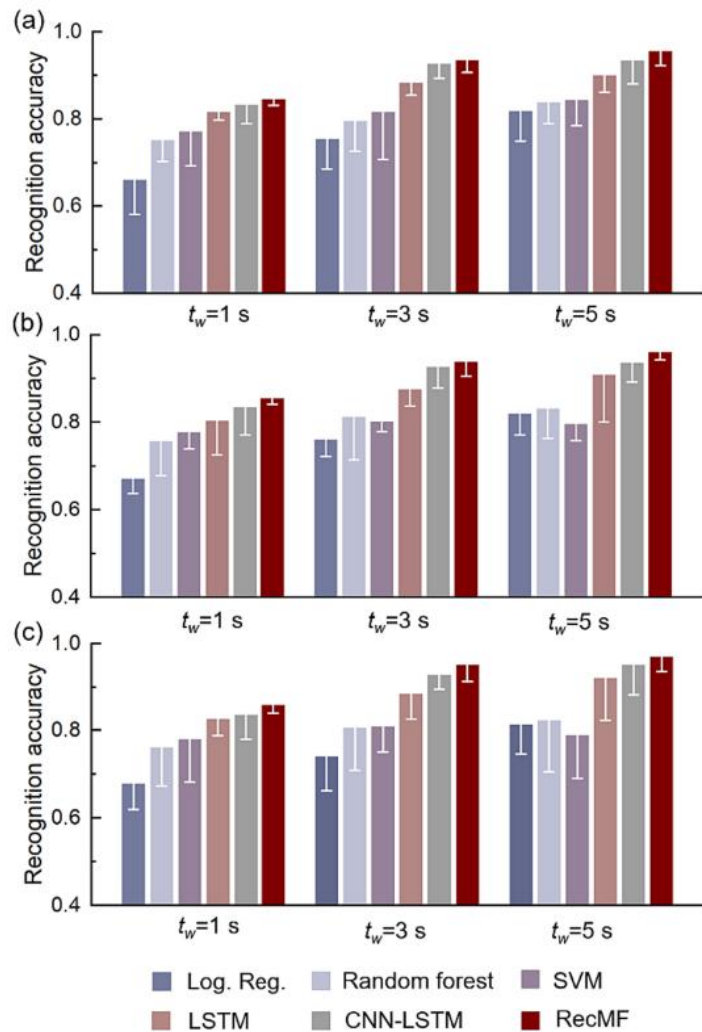


Fig. 6. Comparison of RecMF and other learning-based models in MF recognition accuracy during typical cognitive tasks of (a) 2-back task (b) DL test (c) Stroop test. The error bar is set at $5\sigma_x$.

Table 2

Comparison of models' recognition performance (F1 score) for MF in three cognitive tasks.

Methods	$t_w = 1$ s	$t_w = 3$ s	$t_w = 5$ s
2-back task			
Log. Reg.	0.661 (↓21.9%)	0.755 (↓19.3%)	0.819 (↓14.4%)
Random Forest	0.751 (↓11.2%)	0.796 (↓15.0%)	0.839 (↓12.3%)
SVM	0.772 (↓8.7%)	0.817 (↓12.7%)	0.844 (↓11.8%)
LSTM	0.817 (↓3.4%)	0.884 (↓5.6%)	0.901 (↓5.9%)
CNN-LSTM	0.834 (↓1.4%)	0.928 (↓0.9%)	0.935 (↓2.3%)
RecMF	0.846	0.936	0.957
DL test			
Log. Reg.	0.673 (↓21.4%)	0.762 (↓18.9%)	0.821 (↓14.7%)
Random forest	0.758 (↓11.4%)	0.814 (↓13.4%)	0.833 (↓13.4%)
SVM	0.779 (↓9.0%)	0.803 (↓14.6%)	0.798 (↓17.0%)
LSTM	0.805 (↓6.0%)	0.877 (↓6.7%)	0.910 (↓5.4%)
CNN-LSTM	0.836 (↓2.3%)	0.928 (↓1.3%)	0.937 (↓2.6%)
RecMF	0.856	0.940	0.962
Stroop test			
Log. Reg.	0.680 (↓21.1%)	0.742 (↓21.2%)	0.816 (↓16.0%)
Random forest	0.763 (↓11.5%)	0.809 (↓14.1%)	0.825 (↓15.1%)
SVM	0.782 (↓9.3%)	0.811 (↓13.9%)	0.791 (↓18.6%)
LSTM	0.828 (↓3.9%)	0.887 (↓5.8%)	0.923 (↓5.0%)
CNN-LSTM	0.840 (↓2.6%)	0.931 (↓1.2%)	0.954 (↓1.9%)
RecMF	0.862	0.942	0.972

Thanks