TATC: Predicting Alzheimer's Disease with Actigraphy Data

Jia Li, Yu Rong, Helen Meng, Zhihui Lu, Timothy Kwok, Hong Cheng

The Chinese University of Hong Kong
Tencent AI Lab

Overview

Background

The world is witnessing a dramatic increase of elderly population. Alzheimer's Disease, is the most common cause of dementia among the elderly.

- AD will double in the next 20 years.
- Timely and accurate diagonosis is vital.
- Current diagnosis relies on clinical test and doctors' experience.

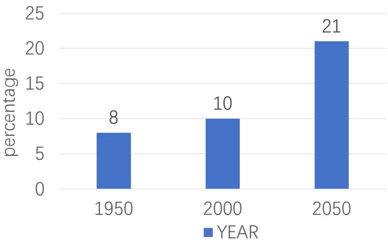


Figure: Proportion of population 60 years or older:world,1950-2050 by UN

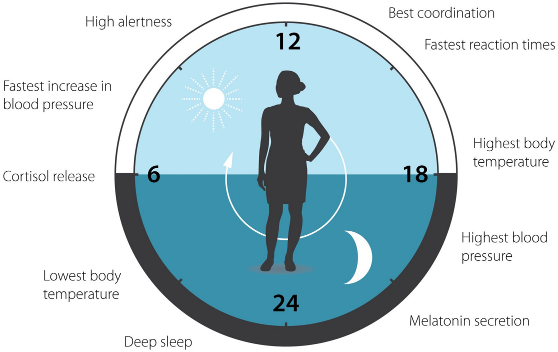
Background

- Recent studies identified physical activity as one risk factor for AD.
- Wrist-worn devices can be used to monitor physical activities.

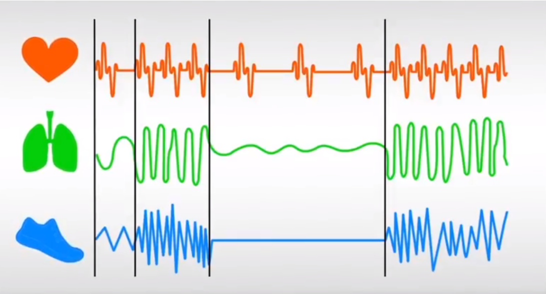
Can actigraphy data be used to predict Alzheimer's Disease?

Motivation

The Nobel Prize in Medicine 2017 was awarded to J. Hall, M. Rosbash and M. Young for their finding on how biological rhythms govern human life.



Motivation



Human daily activities can be understood as various states, e.g., sleeping, sedentary activity, exercising.

Contributions

- Take a data mining approach to predicting AD from actigraphy data in daily living environment.
- Design TATC, an attention-based model for multivariate time series classification with meaningful interpretation.
- Report experience and insights from this cohort study, particularly on data collection and practical value.

Participants

Prince of Wales Hospital launched the project of Hong Kong Alzheimer's Disease Study in 2016, in which 560 Chinese men and 500 Chinese women aged 65 years and older were recruited.

- AD subjects were invited in memory clinics and took AD drugs for at least 3 months.
- Cognitive status was evaluated by HK-MoCA and clinical doctors.
- Subjects were categorized into three groups: normal control (NC), mild cognitive impairment (MCI) and Alzheimer's Disease (AD).

Participants

Group	No. of subjects	Age	MASCH	BMI	MHDIAB	MHMI	GRIPAM
NC	441 (M/F: 287/154)	82.4	7.7	23.5	Y/N: 89/352	Y/N: 50/391	24.9
MCI	103 (M/F: 57/46)	83.3	4.3	23.4	Y/N: 21/82	Y/N: 9/94	18.4
AD	185 (M/F: 68/117)	80.6	6.5	23.5	Y/N: 50/135	Y/N: 20/165	14.2

Table: Statistics of personal information of subjects in NC, MCI and AD groups

Type	Notation	Meaning
Personal information	Gender	M=male, F=female
	Age	65 years and older
	MASCH	years of education
	BMI	body mass index
Clinical history	MHDIAB	medical history of diabetes: Y=Yes, N=No
	MHMI	medical history of heart disease: Y=Yes, N=No
Physical test	GRIPAM	maximum grip strength (kg)

Table: Features of personal particulars

Physical Activity Records

Participants were invited to wear an actigraph GT3X on their non-dominant wrist for 7 consecutive days. The device contains a 3-axis accelerometer and a light sensor.

SubjectID	**Date**	**Time**	X_{acc}	Y_{acc}	Z_{acc}	Lux
HK0001	20161001	23:59	23	8	5	0
HK0001	20161002	00:00	32	15	9	4
HK0001	20161002	00:01	7	5	2	0

Table: A multivariate time series example generated by actigraph GT3X

Observations

- MCI is very similar to NC, while AD can be separated from NC and MCI.
- The difference between groups varies over time.

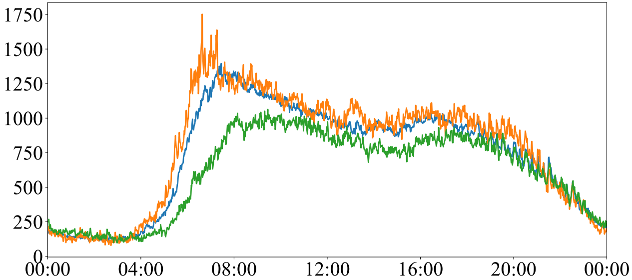


Figure: Average circadian activity by X_{acc} of NC (blue), MCI (orange), and AD (green).

Framework

TATC is a multivariate time series classification model, with composite TICC and CNN to extract meaningful features, and time-aware attention to model the effect of circadian rhythm.

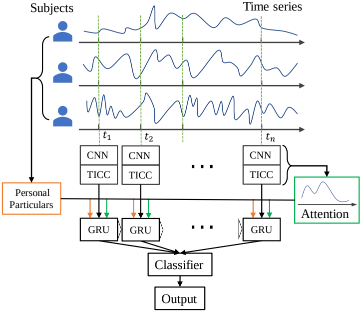


Figure: Framework of TATC.

TICC representation

TICC by Hallac et al. is the unsupervised component that infers hidden states, e.g., sleeping and exercising, from time series.

$$\min \sum_{k=1}^{\kappa} \sum_{x_j \in P_k} -\zeta(x_j, \Theta_k) + \beta \mathbb{1}\{x_{j-1} \notin P_k\}. \quad (1)$$

κ is the number of hidden states.
$\mathbb{1}\{x_{j-1} \notin P_k\}$ is the indicator function indicating that the previous hidden state is different from the current hidden state.
β is the penalty parameter.
$\zeta(x_j, \Theta_k)$ is the log likelihood that x_j belongs to P_k.

CNN representation

CNN is the supervised component to learn discriminative temporal features from the time series.

$$CNN_j = \mathrm{ReLU}(b + \sum_{k=1}^{4\gamma}[\sum_{l=1}^{L} \boldsymbol{F}_k(l)\bar{\boldsymbol{c}}_k(j-l)]). \tag{2}$$

b is the bias.
L is the length of filter.
$\boldsymbol{F}_k$ is the filter working on dimension k.

Time-aware Attention

Different time intervals in a day have different degree of importance to differentiate the subjects, which can be realized by the time-aware attention.

$$S_j = \boldsymbol{E}^\top \boldsymbol{tc}_j + b_0. \tag{3}$$

$\boldsymbol{E} \in \mathbb{R}^{\kappa+\beta}$ is the shared weight parameter of all time intervals.
$\boldsymbol{tc} \in \mathbb{R}^{(\kappa+\beta)}$ is the composite representation.

$$\alpha_j = \boldsymbol{W}_j^\top \boldsymbol{S}. \tag{4}$$

$\boldsymbol{W}_j \in \mathbb{R}^n$ is an n-dimensional vector of parameters.
The attention weights $\boldsymbol{\alpha} = \{\alpha_j\}_{j=1}^n$ is combined as the input for GRU.

Cluster Number

Use BIC to decide the optimal number of clusters in TICC, which is set to 5.

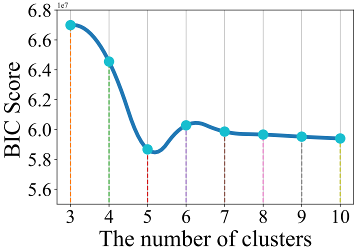


Figure: BIC score corresponding to different number of clusters.

Baselines

- DTW: implemented as a sum of squared DTW distances in each dimension;
- BOSS: train a base classifier on each univariate time series and build an ensemble of four base classifiers;
- SMTS: state-of-the-art multivariate time series classification model.

Imbalanced data is oversampled by SMOTE for all the methods.

NC vs AD Results

TATC achieves the best performance in AUC with a good balance between sensitivity and specificity.

Approach	**Sensitivity**	**Specificity**	**AUC**
DTW	90.3%	47.5%	68.9%
BOSS	38.7%	91.3%	76.1%
SMTS	45.2%	92.5%	84.5%
TATC	80.6%	86.3%	86.2%

Table: Quantitative comparison of different classifiers to predict AD

NC vs MCI Results

The result is not as good as predicting AD. MCI is further categorized into stable MCI (sMCI) and progressive MCI (pMCI). sMCI is physically as active as NC.

Approach	Sensitivity	Specificity	AUC
DTW	70.0%	48.3%	59.1%
BOSS	5.7%	95.5%	58.8%
SMTS	5.0%	91.0%	58.5%
TATC	42.3%	81.3%	61.7%

Table: Quantitative comparison of different classifiers to predict MCI

Measure of Hidden states

State	Interpretation	Measure	X_{acc}	Y_{acc}	Z_{acc}	Lux
*1	**good sleep**	PageRank	0	0	0	0
		Mean	0	0	0	0
*2	**sedentary activity**	PageRank	0.37	0.37	0.26	0
		Mean	704	746	917	0
*3	**light activity**	PageRank	0.22	0.34	0.20	0.24
		Mean	932	972	1240	23
*4	**moderate activity**	PageRank	0.31	0.30	0.25	0.14
		Mean	1293	1281	1600	118
*5	**exercising**	PageRank	0	0	0	0
		Mean	2563	2238	2270	628

Table: PageRank and mean values for five hidden states

Justification of interpretation

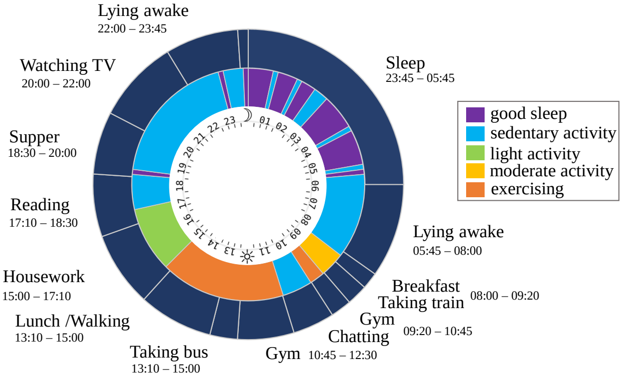
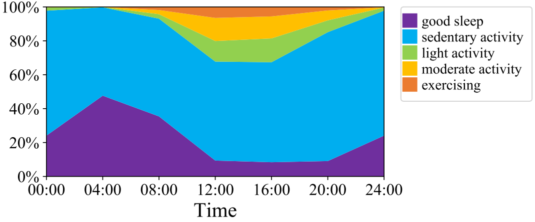
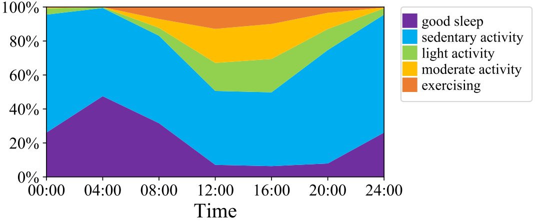


Figure: Self-report daily activities along with inferred hidden states. Inner circle represents hidden states inferred by TATC. Outer circle represents self-report activities.

NC vs AD comparison



(a) AD subjects' mean circadian activity



(b) NC subjects' mean circadian activity

Attention Mechanism

The largest attention weight appears at 4-8 AM, which is the usual wake-up time for the elderly.

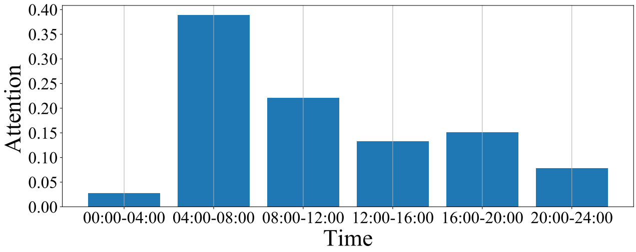


Figure: Average attention weight learned from NC and AD subjects.

Lessons

Data collection is challenging with respect to the aged cohort. Many subjects forgot to put on the device after bathing or swimming, or fill in the self-report questionnaire. The lesson is that data collection procedure should be made simple and bring little disturbance to the subjects' daily life.

Practical value

Traditional cognitive status diagnosis involves lots of clinical assessments and clinic visits, which bring much burden to the elderly and heavily rely on the domain knowledge of doctors.

TATC provides an automatic, low-cost solution for continuously monitoring the change of physical activities of subjects in daily living environment.

Conclusion

- TATC shows great potential and practical value in continuous monitoring of physical activities of subjects and in early detection of AD risk.
- For future work, we plan to explore the possibility of incorporating other measurements for predicting MCI.

Thank you.