CPSL
Cyber-Physical Systems Laboratory

Detecting Mental Disorders with Wearables: A Large Cohort Study

Ruixuan Dai[1], Thomas Kannampallil[1,2], Seunghwan Kim[2], Vera Thornton[3], Laura Bierut[3], Chenyang Lu[1,4,5]

[1]Departmet of Computer Science & Engineering

[2]Departmet of Anesthesiology

[3]Departmet of Psychiatry

[4]Departmet of Medicine

[5]AIM Institute – AI and IoT for Medicine

Washington University in St. Louis

Mental Health Crisis

- Depressive and anxiety disorders are prevalent.
  - **~3.8%** of the population (i.e., **280 million**) experience depression (WHO).
- Over **50%** of patients are not recognized or treated.
- Clinical visit is time-consuming and expensive.
  - Hindering timely diagnosis and intervention

- Detect mental disorders with wearables devices?
  - Unobtrusive, multi-modal sensing
  - Activities, heart rate, and sleep are associated with mental health
  - More than **500 millions** sold in 2021 globally

Need: Large and Diverse Cohorts

- Wearable mental health studies: **small** cohorts with **limited diversity**
  - 18 pregnant women [UbiComp'19]
  - 48-652 college students [UbiComp'14, UbiComp'19, J. Biomed. Inform. '22]
  - 1,002 healthy subjects [NPJ Digit. Med. '18]

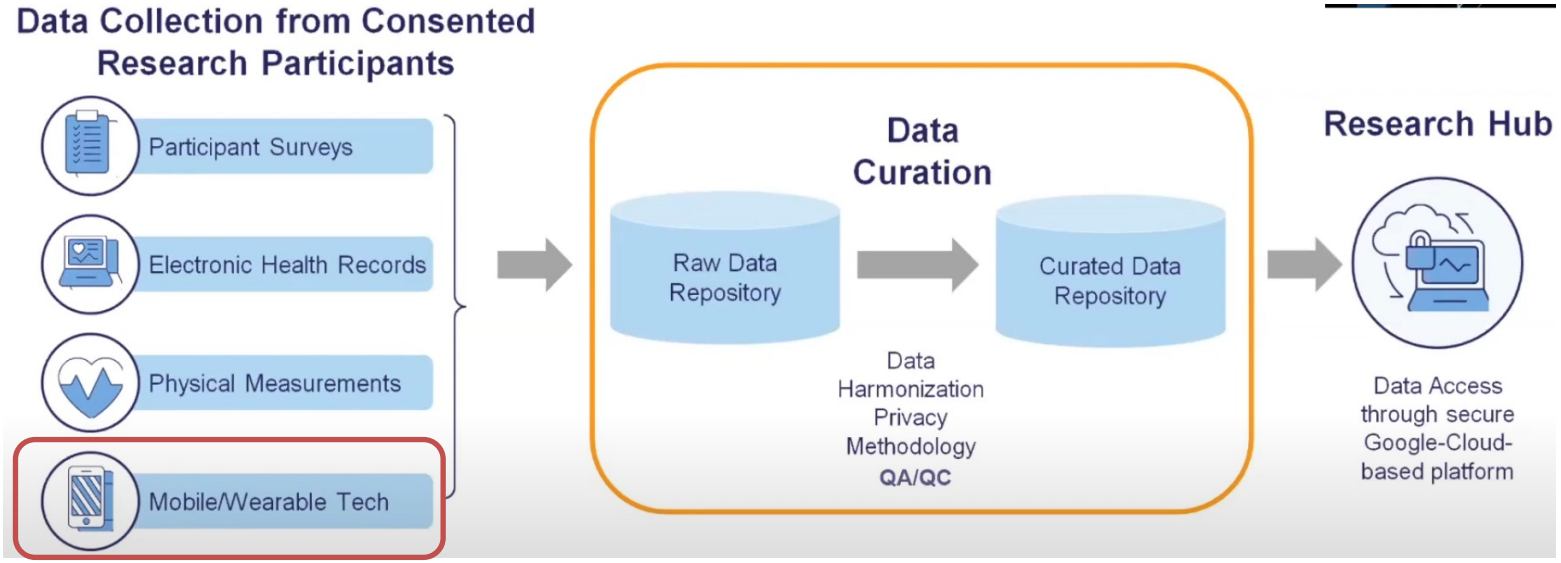
- Machine learning approaches
  - Shallow machine learning models
    - Rely on ad hoc feature engineering
    - Limited predictive power
  - Deep models with questionable generality

- **Small datasets limit the rigor and generality of results!**

Our Work

- A **large** and **diverse** dataset
  - **8,996** participants of the "All of Us" program
  - **Longitudinal wearable data** and **mental health diagnosis**

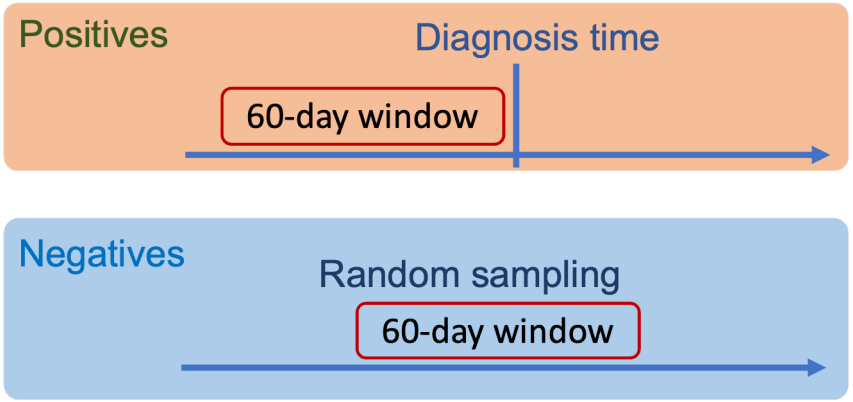
- **WearNet**: deep model for detecting mental disorders with wearables



Source: All of Us program

Outcomes and Inputs

- Mental disorders
  - 20 depressive/anxiety disorder diagnoses identified by clinical experts
  - Positive label: a participant with any of the diagnoses

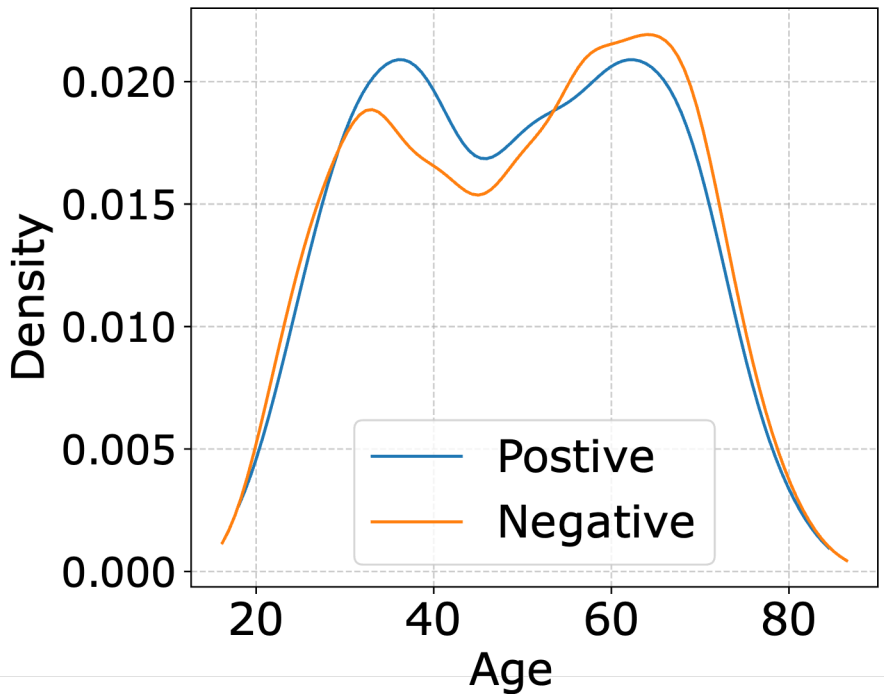


- Input features
  - Wearable data in a 60-day window
    - 10 types of daily summaries from Fitbit cloud
    - Derived based on step and heart rate time series
    - e.g., total steps, average heart rate, sedentary minutes
  - Static data from electrical health records and surveys
    - Age, race, ethnicity, gender, education, smoke history, alcohol history

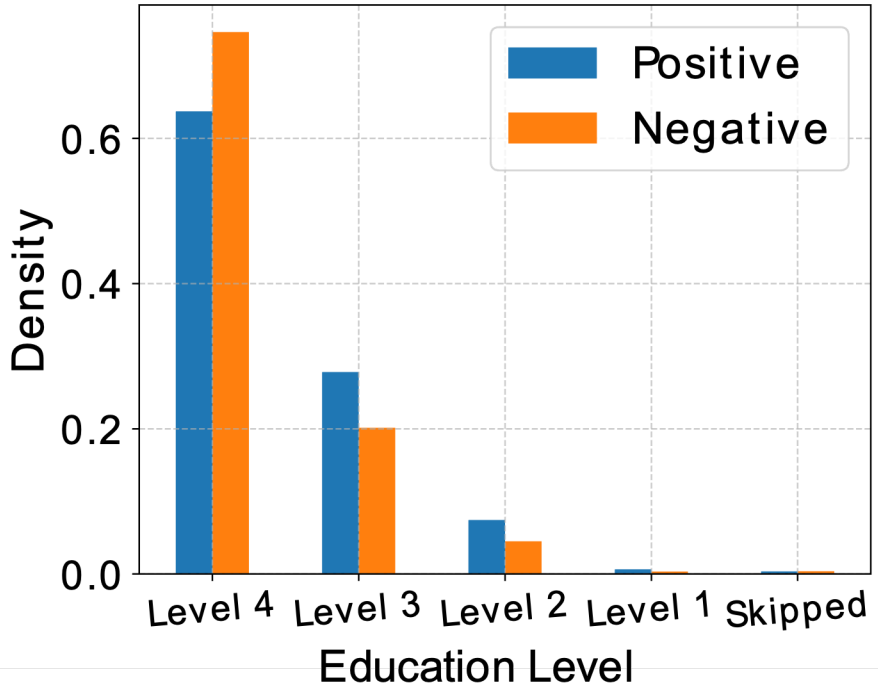
Large and Diverse Cohort

- **8,996** participants including **1,247** with mental disorders

Age: mean 48.6 (s.d. 15.9)



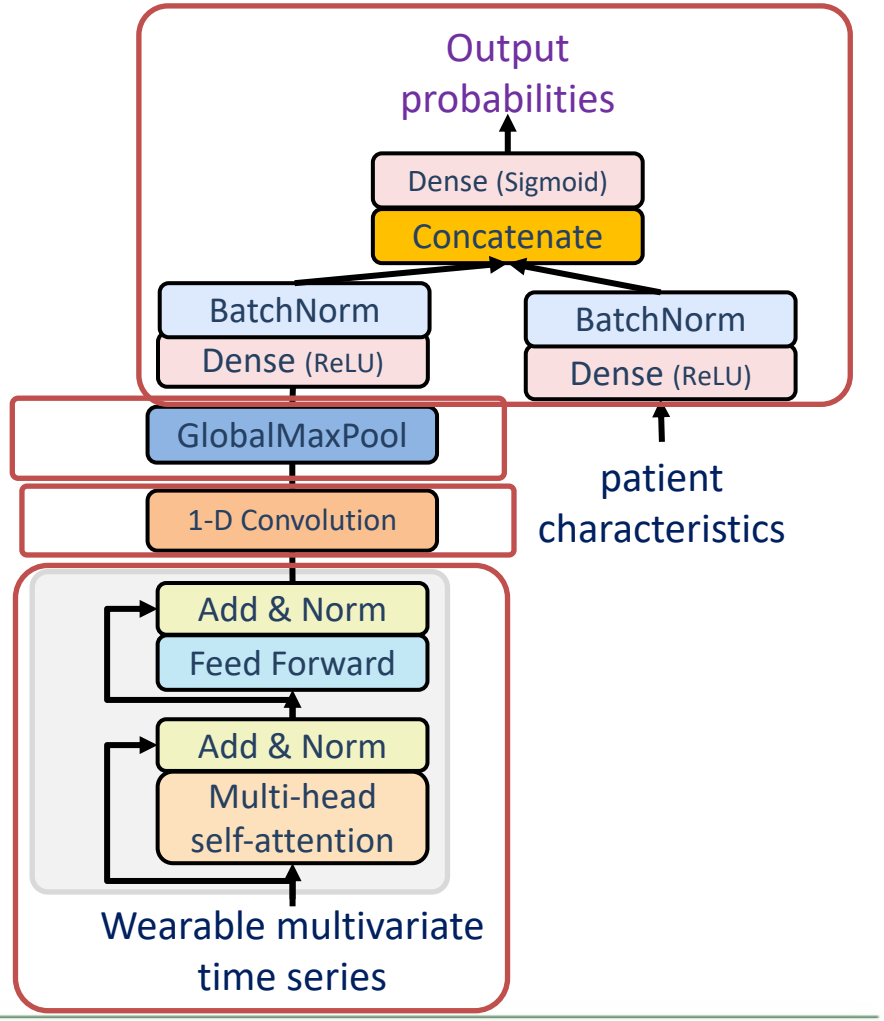
Education



Level1: below high school
Level2: twelve Or GED
Level3: College One to Three
Level4: College graduate or above

WearNet: Deep Model for Detecting Mental Disorders

- Combining transformer encoder and convolutional neural network
- Transformer encoder (multi-head self-attention)
  - Identifies patterns across multiple timestamps
- Convolutional neural network (1-d)
  - Integrates neighborhood patterns
- Global max pooling
  - Identifies one global pattern for robustness
- Integrate wearable and static data at the top
  - Captures the underlying characteristics



Detection Performance

- **AUROC**: Area Under the Receiver Operating Characteristic Curve
- **AUPRC:** Area Under the Precision-Recall Curve

Category	Model	AUROC	AUPRC	
Shallow Models	LR	0.701(0.000)	0.351(0.000)	
	SVM	0.592(0.000)	0.290(0.000)	
	RF	0.661(0.005)	0.349(0.007)	
	GBDT	0.685(0.001)	0.365(0.000)	
Deep Models	bi-LSTM	0.702(0.015)	0.464(0.011)	**RNN and TCN Models**
	BRITS	0.693(0.012)	0.445(0.011)	
	CrossNet	0.682(0.021)	0.429(0.014)	
	TCN	0.629(0.021)	0.235(0.024)	
	Informer	0.705(0.008)	0.428(0.011)	**Transformer Models**
	WearNet	0.717(0.009)	0.487(0.008)	

- **Shallow models** with feature engineering are less predictive.
- **RNN and TCN models** underperformed Transformer models.
- **WearNet** achieved the best predictive performance.

Ablation Study

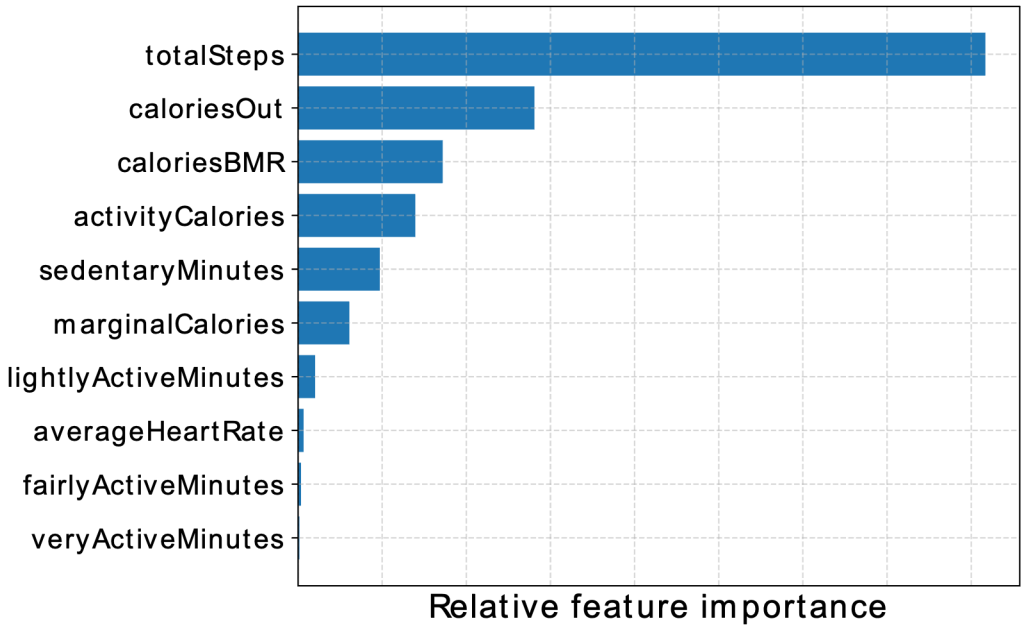
Model	AUROC	AUPRC
WearNet	0.717(0.009)	0.487(0.008)
WearNet(without convolutional layer)	0.673(0.008)	0.251(0.008)
WearNet(without static characteristics)	0.702(0.007)	0.456(0.007)
WearNet(without wearable data)*	0.650(0.003)	0.222(0.003)

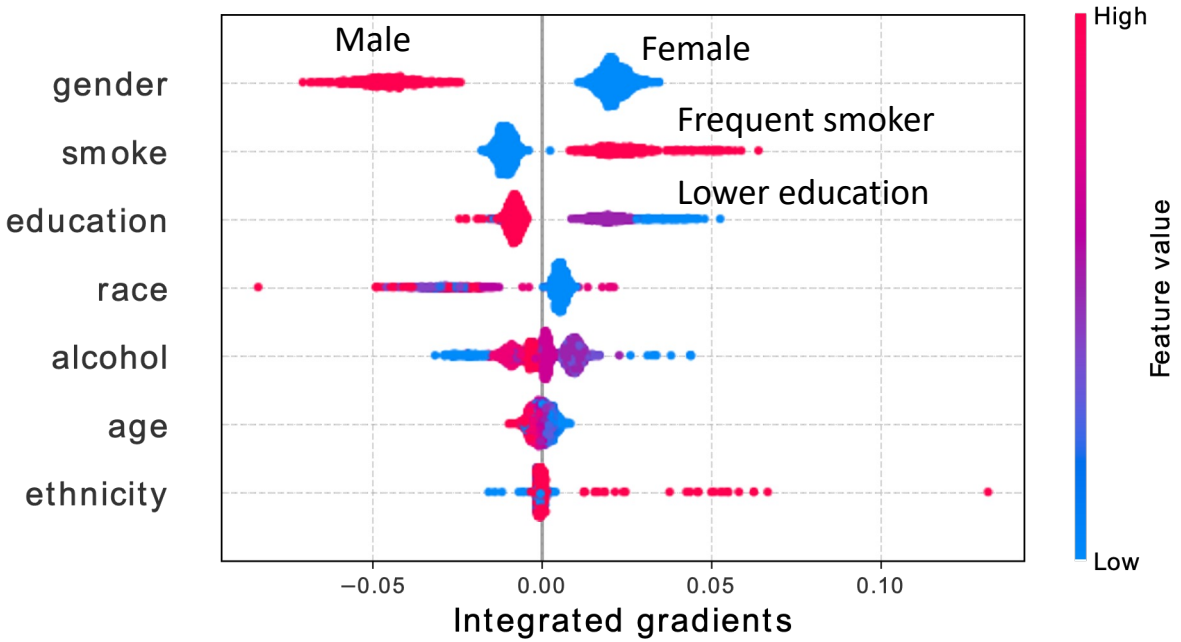
*The model reduces to a single-layer perceptron if we drop components for wearable data

- **Integrating** neighborhood patterns with **convolutional layer** helped.
- **Wearable time-series** had larger impacts than than static characteristics.
- Wearable data and static characteristics are **complementary** to each other.

Model Explanation

- Assign importance scores to features by approximating the integral of gradients





- **Total step** is the most important wearable feature.
- **Women** and **frequent smokers** are more likely to be diagnosed.

→ **Matching the literature**

Conclusion

- Mental disorders are **prevalent** but significantly **underdiagnosed**.
- A **large** and **diverse** cohort with 8,996 participants from All of Us.
- **WearNet**: a deep model for detecting mental disorders with wearables.
- An **unobtrusive** approach to detect mental disorders in the community.