

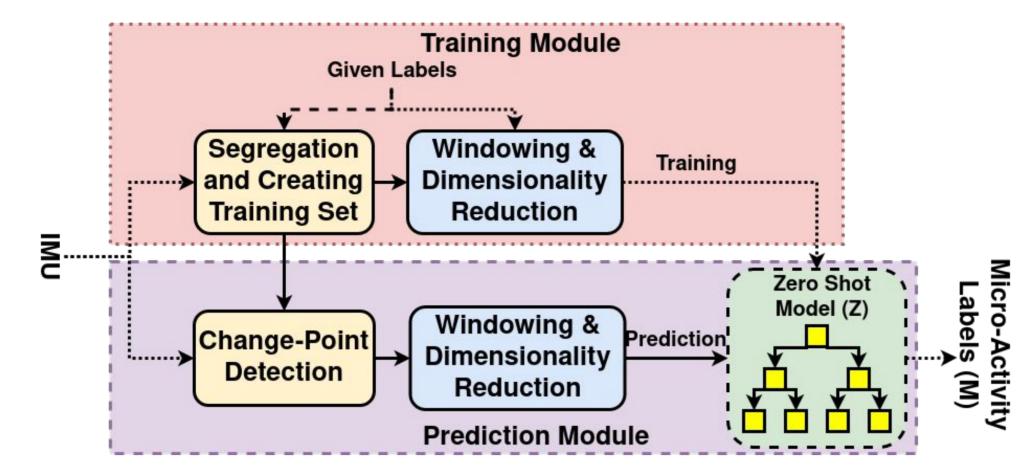
Demo: Automated Micro-Activity Annotations for Human Activity Recognition with Inertial Sensing



Soumyajit Chatterjee, Bivas Mitra, Sandip Chakraborty

Department of Computer Science and Engineering, Indian Institute of Technology Kharagpur

Motivation and Broad Idea



1. Automating the complicated tasks of getting annotations for micro-activities

Generation of Micro-Activity Annotations

Algorithm 2 Predicting Micro-Activities

Input Accelerometer data $\mathcal{I}_{t+\tau}^t$ for activities where $\tau \geq 10$ s and the trained zero-shot model \mathbb{Z} .

Output Predicted micro-activities $\mathbb{M}_{t+\tau}^t$, where $\mathbb{M}_{t+\tau}^t = \{m_1, m_2, \dots, m_n\}$, where *n* is the total number of micro-activities performed by the subject in the time duration $[t, t+\tau]$.

1: $\mathbb{W}_{t+\tau}^t = change_point(\mathcal{I}_{t+\tau}^t)$ {Here, $\mathbb{W}_{t+\tau}^t$ is the set of change-point windows.} 2: $\mathbb{M}_{t+\tau}^t = \{\}$

3: for each change window ω in $\mathbb{W}_{t+\tau}^t$ do

- 2. Challenges: Number and characteristics of the micro-activities may not be known apriori
- 3. Primary Idea: Using the short-duration macro-activity labels along with zero-shot learning
- 4. Experiments done on Kitchen dataset [1] containing complex activities of daily living

Zero-Shot Learning and Verb Attribute Embeddings

٢	Transitivity			Aspect	Motion	Time	Social	Bodyparts						Effect on Arguments										
	Transitivity		Arms					Head	Legs	Torso	Other		Effect on Arguments											
Iſ	1	1	1	3	3	1	0	1	0	0	0	0	0	1	0	0	0	1	1	0	0	1	1	0

Sample verb-attributes [2] for the activity verb "spray" from the Kitchen dataset

Training with Short-Duration Macro-Activities

Algorithm 1 Training the Zero-Shot Model

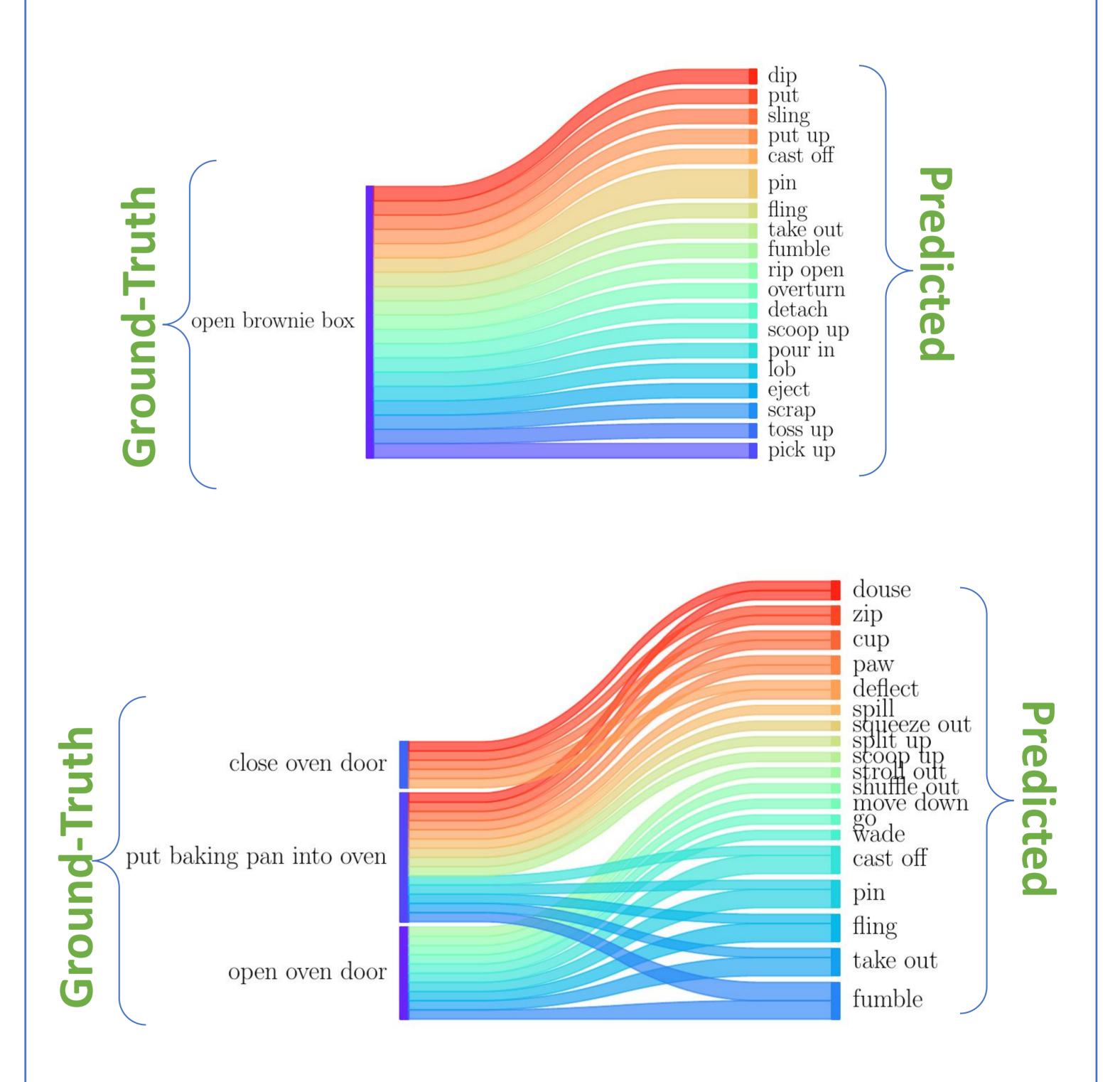
- **Input:** Training Set $\mathbb{V} = \{(u, a) : u \in \mathcal{I}, a \in \mathcal{A}\}$, where \mathcal{I} is the accelerometer data and \mathcal{A} is the set of short-duration macro-activity labels. **Output** Trained zero-shot model \mathbb{Z}
- 1: $(\mathbb{S}, \mathcal{A}) = red_dimen_labelwise$ (\mathbb{V}, d) {We fix d = 2. Here, \mathbb{S} , which is the transformed accelerometer data with dimensions reduced to d.}
- 2: $\mathbb{Z} = []$
- 3: for i = 1 to \mathbb{N} do

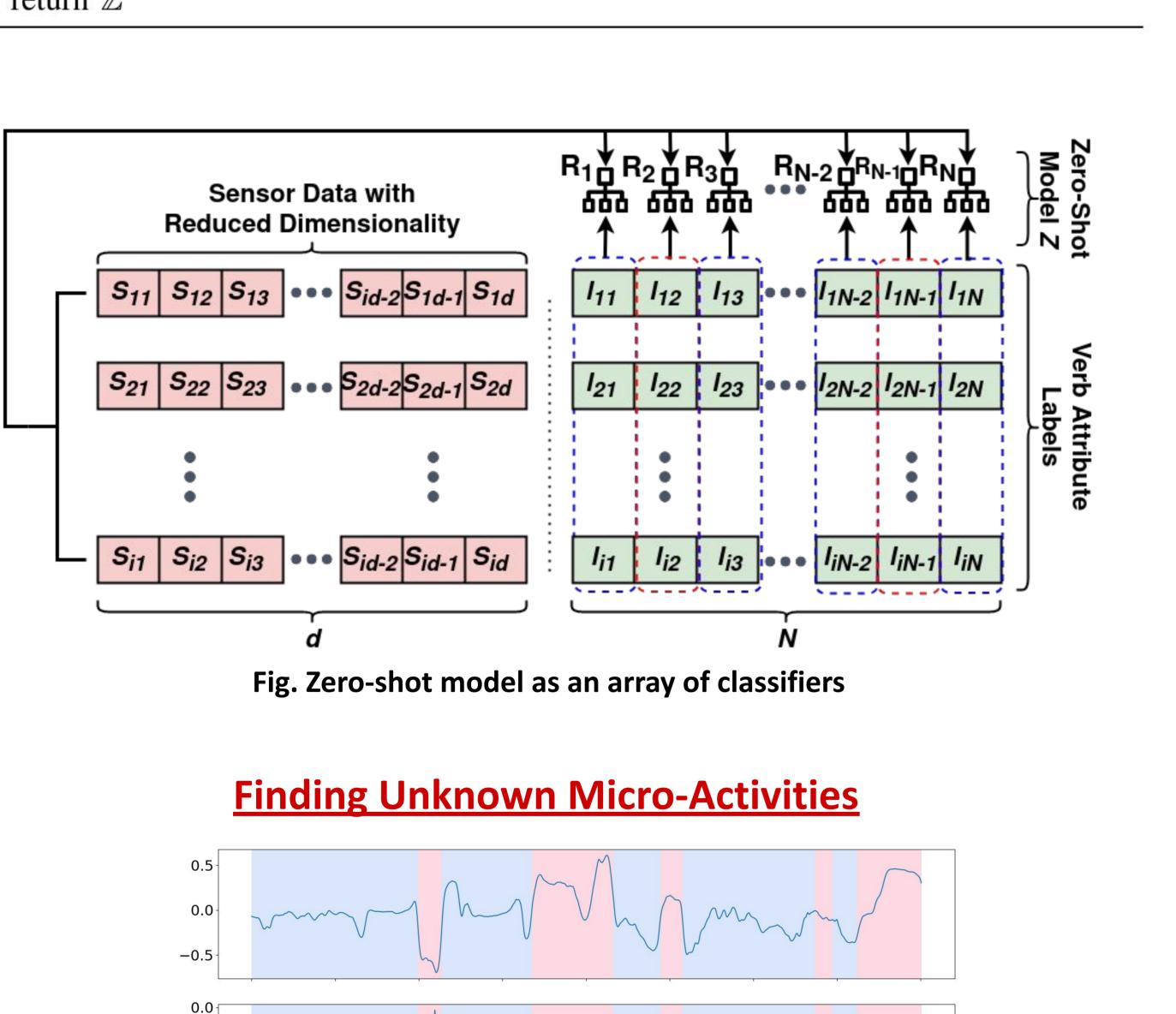
4:
$$\mathbb{Z}[i] = \text{train_model}(\mathcal{R}_i, \mathbb{S}, \text{verb_attribute}(\mathcal{A})[:,i])$$

- 5: end for
- 6: return \mathbb{Z}

- 4: J_ω = average accelerometer data fro the window ω
 5: J_ω = red_dimen (J_ω, d) {We fix d = 2.}
 6: m_ω = predict_micro_activities (ℤ, J_ω) {Here, m_ω is the micro-activity predicted for the change-window ω.}
 7: M^t_{t+τ} = M^t_{t+τ} ∪ {m_ω}
 8: end for
 9: return M^t_{t+τ}
- 1. Accelerometer across change-point windows is used as an input to the zero-shot model
- 2. Output is a set of micro-activities in the form of attribute embeddings
- 3. Micro-activities defined by observing the closest known verbs in the embedding space
- 4. Limitations: Unnecessary verbs may appear due to hubness

Demonstration of Labeling Performance





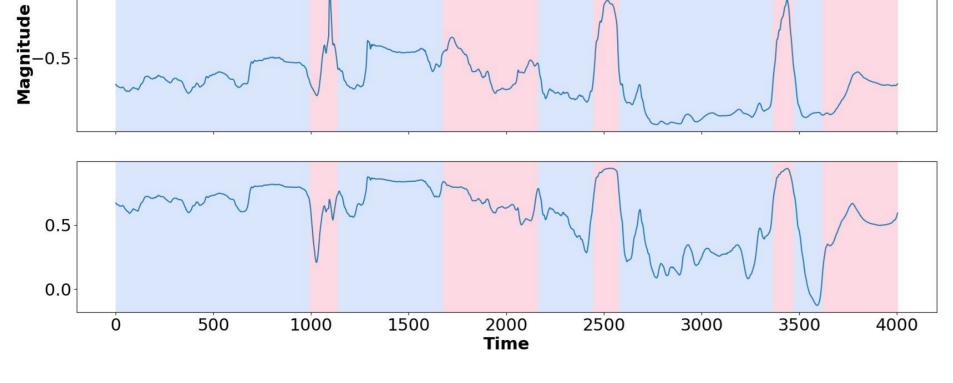


Fig. Change-points detected with penalty 50

- 1. Unsupervised change-point detection can help locate the activity boundaries
- 2. Each activity window can potentially represent one micro-activity
- 3. Use window-based change-point detection [3] with window-size = 100 and RBF kernel

<u>References</u>

[1] E. H. Spriggs, F. De La Torre, and M. Hebert, "Temporal segmentation and activity classification from first-person sensing," in IEEE CVPR Workshops. IEEE, 2009, pp. 17–24.
[2] R. Zellers and Y. Choi, "Zero-shot activity recognition with verb attribute induction," in EMNLP, 2017.

[3] C. Truong, L. Oudre, and N. Vayatis, "Selective review of offline change point detection methods," Signal Processing, 2020.

Contact

- Website: <u>https://sites.google.com/view/sjitiit/home</u>
- Twitter: <u>https://twitter.com/Soumyaj44848812</u>
- Email: <u>soumyachat@iitkgp.ac.in</u>
- Alternate Email: <u>sjituit@gmail.com</u>

