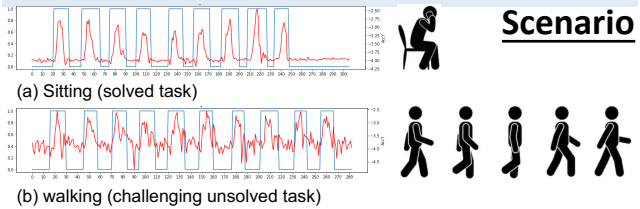
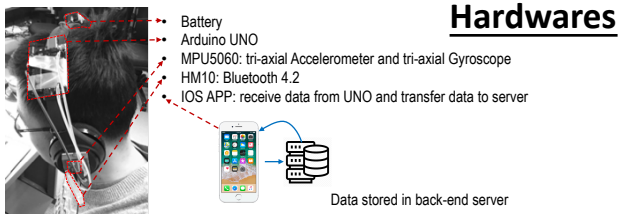
Conditional-UNet: A Condition-aware Deep Model for Coherent Human Activity Recognition From Wearables

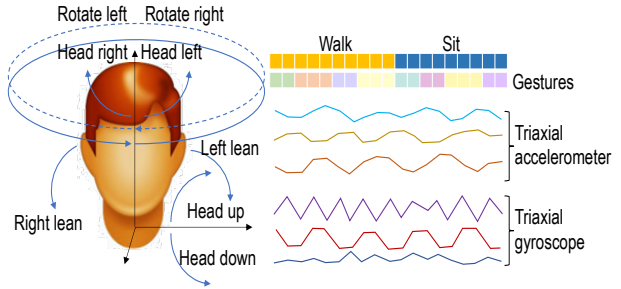
Liming Zhang, George Mason University; Wenbin Zhang, University of Maryland; Baltimore CountyNathalie Japkowicz, American University.12/10/2020

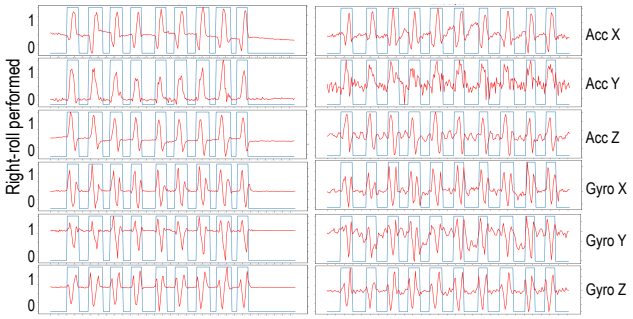
Motivations: Many conditions are hard when body is moving.





Coherent Human Activity Recognition (Co-HAR) with ***single-location*** sensors.





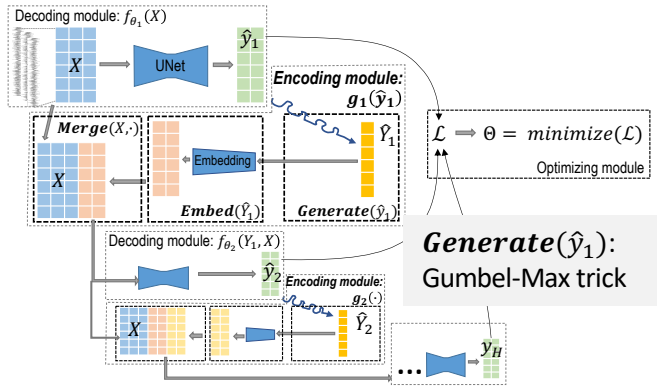
(a) Right-roll under sit (b) Right-roll under walk

A novel condition-aware deep model called "Conditional-UNet"

Challenges

(1) The single location of sensors has mutual impact on signals.
(2) The imbalanced domination of different activities could fade away the signals of the other activities.
(3) The multi-label window problem for activities of various duration.

Deep architectures for a new encoding module



$Generate(\hat{y}_1)$: Gumbel-Max trick

Limitations:
Real-life scenarios are more complicated. Need to include more deep learning methods to compare.
We run deep models on desktop GPU, but computation power is constrained in real-world wearables.

Future works
Is such trained model transferred for real-world scenario? Or need re-training?
In the data likelihood loss, hierarchical labels can be considered or imbalanced class problem can be studied in the future.

A more general likelihood estimation

Raw data likelihood formula

$$p(Y_1, \ldots, Y_H|X) = p_{\theta_1}(Y_1|X)p_{\theta_2}(Y_2|Y_1, X)\ldots p_{\theta_H}(Y_H|P_{H-1}, \ldots, Y_1, X) \quad (1)$$

Y_i: different labels
X: sensors data

Our approach: Conditional data likelihood factorization as a more general framework

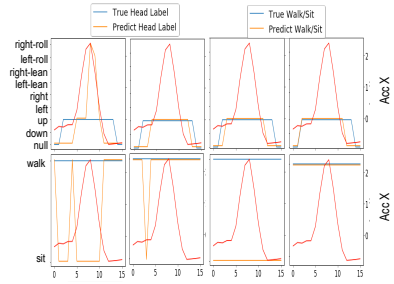
$$\mathcal{L} = log(p(Y_1, \ldots, Y_H|X)) = \sum_t^T (log(p_{\theta_1}(Y_{1,t}|X)) + \cdots + log(p_{\theta_H}(Y_{H,t}|P_{H-1,t}, \ldots, Y_{1,t}, X))) \quad (2)$$

Existing approaches: multi-label classification assuming conditional independences

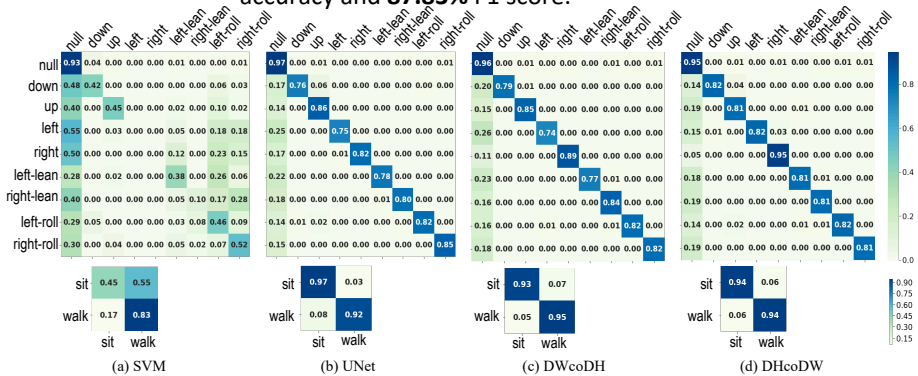
$$p(Y_1, \ldots, Y_H|X) = p_{\theta_1}(Y_1|X)p_{\theta_2}(Y_2|X)\ldots p_{\theta_H}(Y_H|X)$$

Experiment results

Baseline models:
SVM, UNet
Two alternatives of Conditional-Unet:
(1) DWcoDH, Walking conditioned on Head
(2) DHcoDW, Head conditioned on Walking
achieves up to **92.06%** accuracy and **87.83%** F1 score.



(a) SVM (b) UNet (c) DWcoDH (d) DHcoDW



(a) SVM (b) UNet (c) DWcoDH (d) DHcoDW