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 County
Nathalie Japkowicz, American University

12/10/2020
Motivations

(1) Sensor-embedded wearables are more and more popular. etc.

(2) Head gesture recognition with wearables are a trending research.

(3) Real world conditions can be complicated when users are moving. None of current works tried to solve this.

Scenario

(a) Sitting (solved task)

(b) walking (challenging unsolved task)

Hardwares

- Battery
- Arduino UNO
- MPU5060: tri-axial Accelerometer and tri-axial Gyroscope
- HM10: Bluetooth 4.2
- IOS APP: receive data from UNO and transfer data to server

Data stored in back-end server
Challenges

A new problem: *Coherent* Human Activity Recognition (Co-HAR) with *single-location* sensors

Specifically, there are modeling challenges as follows:

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
Collected experiment data

(a) Right-roll under sit

(b) Right-roll under walk
A novel condition-aware deep model called “Conditional-UNet”

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

\[
\mathcal{L} = \log(p(Y_1, \ldots, Y_H | X)) = \sum_{t} \left( \log(p_{\theta_1}(Y_{1,t} | X)) + \cdots + \log(p_{\theta_H}(Y_{H,t} | P_{H-1,t}, \ldots, Y_{1,t}, X)) \right) \quad (2)
\]

Existing approaches: multi-label classification assuming conditional independences

\[
p(Y_1, \ldots, Y_H | X) = p_{\theta_1}(Y_1 | X)p_{\theta_2}(Y_2 | X) \ldots p_{\theta_H}(Y_H | X)
\]
A novel condition-aware deep model called “Conditional-UNet”

Deep architectures

Unet used a 32 hidden neuron for the first upsampling and last downsampling layers, and is very efficient for computing.
A novel condition-aware deep model called “Conditional-UNet”

Compared Naïve-Max and Gumbel-Max trick for $\text{Generate}(\hat{y}_1)$ sampling operation

(a) Unstable Naïve-Max trick

(b) Stable Gumbel-Max trick

(c) True logits
Experiment results

Baseline models:
SVM, UNet

Two alternative models of Conditional-Unet:
1) DWcoDH, Walking conditioned on Head
2) DHcoDW, Head conditioned on Walking

Conditional-Unet outperforms existing state-of-the-art UNet model, and achieves up to 92.06% of accuracy and 87.83% of F1 score. Also, DHcoDW’s performance is good for all gesture types, not for some gestures like other baselines.
Contributions, Limitations and Future works

Contributions:
Addressed a challenging problem Co-HAR for which a new dataset was collected.
Proposed a novel condition-aware deep model called “Conditional-UNet”.

Limitations:
• It is still not real-life scenario.
• 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.

Code and data: https://github.com/tongjiyiming/Conditional-UNet