2020 19th ACM/IEEE International Conference on Information Processing in Sensor Networks (IPSN)

Poster Abstract: Combating Transceiver Layout Variation in Device-Free WiFi Sensing using Convolutional Autoencoder

Isura Nirmal[1],Abdelwahed Khamis[1],Wen Hu[1], Mahbub Hassan[1],Xiaoqing Zhu[2]
[1] University of New South Wales
[2] Cisco Systems, Inc
{b.isuranirmal,a.khamiss,wen.hu,mahbub.hassan}@unsw.edu.au
xiaoqzhu@cisco.com

ABSTRACT

Sensitivity of WiFi channel measurements to the transceiver placement is a major limitation for on-demand deployment of device-free WiFi sensing in environments where the transmitting/receiving devices may move. Using publicly available datasets, we show that even slight deviations of transmitter/receiver placements from the reference values can degrade device-free gesture recognition accuracy significantly. We design a convolutional autoencoder to translate WiFi spectrograms from arbitrary receiver placements to a reference placement configuration in a given area of interest with minimal human effort. Our experiments with the public datasets reveal that the proposed autoencoder can successfully reduce WiFi measurement variability caused by transmitter/receiver movement, which ultimately increases gesture recognition accuracy by up to 58%.

KEYWORDS

Activity Recognition,Wireless Sensing, Deep Learning

1 INTRODUCTION

Device free wireless sensing has emerged as a promising alternative to the wearable paradigms. In spite of tremendous research effort in localization, activity recognition and healthcare applications, current models are still far from real world implementations. WiFi enabled IoT device's locomotion can negatively affect model's performance.

Sensitivity of WiFi channel measurements to the deployment configuration (i.e. position of sender and receiver w.r.t the user) is known as one of the key limitation of current WiFi device-free sensing. As a result, most WiFi sensing systems requires operating in the same configuration all the time [5].Had the sender and/or receiver configurations been altered or the user followed a different path, the captured features (i.e. spectrogram) will be different. The reason behind this is that spectrograms are mainly capturing the rate of change in wireless paths lengths [3]. Limited by this, systems trained on these features are required to either assume fixed configuration (such as fixed sender and receiver setup in a corridor environment [5]) or do re-training for every new configuration which is not practical.

2 MOTIVATION & PROPOSAL

In the literature, two generic approaches were followed to resolve this problem. The first one is to learn a *configuration-independent model* that able to work across various configuration . Examples include EI [1] that uses training data from multiple domains to learn configuration independent representations. This requires collecting data from numerous configurational changes (≥ 22 [1]) before the model can generalize. The second approach is to *perform translation* to/from a reference configuration. In this direction, we find CrossSense [6] that translates (or "roams") the model itself using data from new configuration and transfer learning techniques requiring retraining, albeit with low overhead, in new configurations. Another way is to translate the actual wireless measurements directly.WiAg[4] that constructs virtual samples in all possible target configurations given the actual samples and the gesture shape in the reference configuration. Consequently, classifiers in all possible configuration can be built and trained on virtual samples however this requires the support of wearable IMU to capture data.

Inspired by this and in line with translation techniques, our approach to counter the limitations is developing a system that answers the questions of " What will spectrogram look like when we change the configuration from X to Y?". At the core of this system is automatic translation of the spectrogram to the target configuration using deep convolutional autoencoder. Figure 1 illustrates the concept of our proposal in one example application scenario in which user controls IoT WiFi-enabled devices in smart home using touchless gestures. A translation model would enable a reference classifier to be used for newly installed devices with a new configuration (a) or tolerate configuration changes of an existing devices (b).

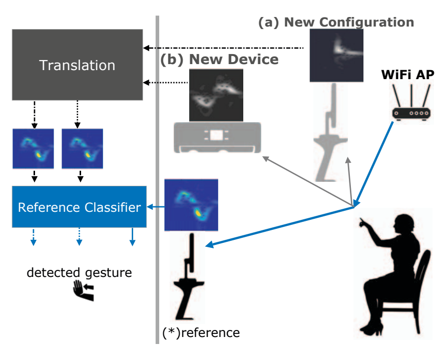

Figure 1: Proposed sensing and inference pipeline

3 EXPERIMENTS

We are using a publicly available dataset provided by [7] and selected 4 activities (Push&Pull, Sweep, Clap and Swipe) performed by multiple users in an enclosed environment. The Channel State

978-1-7281-5497-8/20/$31.00 ©2020 IEEE
DOI 10.1109/IPSN48710.2020.00-19

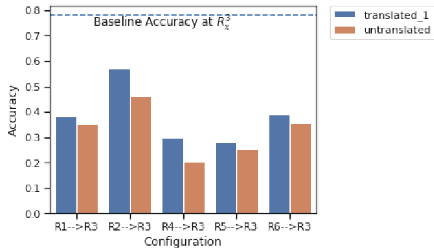

Figure 2: Classification Results

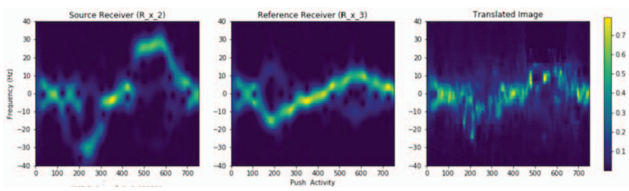

Figure 3: Translated Image

Information (CSI) is captured synchronously by 6 receivers placed in fixed positions in the environment. We use spectrogram image generated from the CSI data as the input for our deep model.

Among 6 receivers, we are taking one receiver, R_x^3 as a reference receiver where the considerable multi paths from wireless link are received. Our notation R_x^1−>R_x^3 shows the translation from source location R_x^1 to reference location R_x^3. Maximum $R_x - T_x$ distance is 2 m and receiver locations are apart 0.9 m to 2.8 m.

For the inference model,[CNN+GRU] based classification module proposed by [7] is used. Figure 2 shows the range of accuracies from each source receiver to the reference receiver and classification accuracy.

We have tested the classification model first with the untranslated images (untranslated) from source location resulting 20% to 45% accuracy for the classifier trained with images from reference location.

Then, the translator is trained with Push & Pull activity only (translated_1) preventing other activities to be unseen to the translator model. R_x^2 to R_x^3 translated images provide 58% of accuracy performing better than the untranslated data. This is a cost saving as only one activity is used for training. However, the baseline accuracy is high as 78% when the classifier is using R_x^3 samples itself. In Figure 3 , we have included the translated image obtained for the best performing transceiver pair (R_x^2 −> R_x^3 SSI = 0.482).

4 CONCLUSION AND FUTURE WORKS

In this work, we are trying to address the tremendous effort in model retraining and recollection of data in wireless sensing based large scale human activity recognition models with the support of deep learning based translation model.

To the best of our knowledge, our work is the first to address explicit configuration to configuration translation in wireless sensing using deep learning

Our preliminary results act as a proof-of-concept for the future works but not without the limitations. Our deep translation module requires the human intervention for activity labeling and pairing.This is to be improved with GAN architecture. Similar works including Mic2Mic[2] demonstrated the applicability in improved GAN architectures for non-paired domain to domain feature translations in other application areas.

ACKNOWLEDGMENTS

We would like to acknowledge the authors of Widar 3.0 [7] for making available their data set publicly. This work is partially funded by a CISCO Research Center University Grant.

REFERENCES

[1] Wenjun Jiang, Chenglin Miao, Fenglong Ma, Shuochao Yao, Yaqing Wang, Ye Yuan, Hongfei Xue, Chen Song, Xin Ma, Dimitrios Koutsonikolas, et al. 2018. Towards environment independent device free human activity recognition. In *Proceedings of the 24th Annual International Conference on Mobile Computing and Networking*. ACM, 289–304.

[2] Akhil Mathur, Anton Isopoussu, Fahim Kawsar, Nadia Berthouze, and Nicholas D. Lane. 2019. MIC2MIC: Using cycle-consistent generative adversarial networks to overcome microphone variability in speech systems. *IPSN 2019 - Proceedings of the 2019 Information Processing in Sensor Networks* (2019), 169–180. https://doi.org/10.1145/3302506.3310398

[3] Kun Qian, Chenshu Wu, Zheng Yang, Yunhao Liu, and Kyle Jamieson. 2017. Widar: Decimeter-level passive tracking via velocity monitoring with commodity Wi-Fi. In *Proceedings of the 18th ACM International Symposium on Mobile Ad Hoc Networking and Computing*. 1–10.

[4] Aditya Virmani and Muhammad Shahzad. 2017. Position and orientation agnostic gesture recognition using wifi. In *Proceedings of the 15th Annual International Conference on Mobile Systems, Applications, and Services*. 252–264.

[5] Wei Wang, Alex X Liu, and Muhammad Shahzad. 2016. Gait recognition using wifi signals. In *Proceedings of the 2016 ACM International Joint Conference on Pervasive and Ubiquitous Computing*. 363–373.

[6] Jie Zhang and Zheng Wang. [n.d.]. CrossSense : Towards Cross-Site and Large-Scale WiFi Sensing. ([n. d.]).

[7] Yue Zheng, Yi Zhang, Kun Qian, Guidong Zhang, Yunhao Liu, Chenshu Wu, and Zheng Yang. 2019. Zero-Effort Cross-Domain Gesture Recognition with Wi-Fi. (2019), 313–325. https://doi.org/10.1145/3307334.3326081