Investigating the Feasibility and Efficacy of Data Augmentation Strategies for Context Recognition In-the-Wild

Ha Le
Northeastern University
le.ha1@northeastern.edu

1 Introduction and motivation

Recently, deep learning frameworks have been shown to perform well in time series analysis and signal processing. However, multi-class deep learning models owe their power to the large training sets, with many data points in each class. Training deep neural networks for context recognition on free-living condition can be challenging because of the lack of complete, balanced, and labeled data sets. Data augmentation (DA) is a well-known technique used primarily in computer vision to expand the training set by performing certain mutations on the existing data. Due to the characteristics of the data, many traditional data augmentation techniques cannot be applied to sensor data, which limits the usage of DA techniques. For the final project, I would like to investigate the feasibility and efficiency of using different data augmentation techniques to perform context recognition on multi-sensor, multi-class real-world data.

2 Description and objectives

During six weeks, I plan to investigate the feasibility and efficiency of different data augmentation techniques on deep learning models in context recognition in the wild task. Some possible questions that I want to answer are:

1. *Does the use of one or a combination of DA techniques can increase the robustness of models?*
2. *Do the effects of DA vary between different labels?*
3. *Can we find an automatic strategy to choose the best DA strategies for different context labels?*

For the project, I will be using the *extraSensory* dataset (Vaizman et al., 2017), which contains the data for a multi-modal experiment annotated with rich context labels. The data was collected using the participants' smartphones and companion smartwatches, in their natural environment for a week. There are more than 20 different self-reported context labels available in the published dataset.

3 Related Work

In the past, few research have been done to explore the benefits of data augmentation on sensor data. (Um et al., 2017) has shown that using a combination of data augmentation techniques (cropping, permutation, and time warping), they were able to increase the performance of a Convolutional Neural Network (CNN) model by 10% to classify different motor state in people with Parkinson disease. (Saeed et al., 2020) showed augmentation strategies can improve the robustness of CNN on 7-sensor datasets with 4 complex tasks, resulting in a 1.5 to 10 F1 score increase from the baseline. (Kraft et al., 2020) used two augmentation techniques on multiple deep learning models (LSTM, ResNet, and CNN), shifting and axis-rotation, and was able to achieve an accuracy score of 97% for fall detection tasks across 5 different datasets. However, most related studies have only used augmentation techniques on very controlled-setting data sets, predicting very limited and specialized tasks.

Other methods have also been explored to tackle the lack of quality-labeled sensor data in real-world settings. (Vaizman et al., 2018) used transfer learning (using a pre-trained model to recognize new label) to improve the performance of a Multi-Layer Perceptron (MLP) on the *extraSensory* dataset. Recently, multiple efforts have been made to use self-supervised learning to extract features from multi-sensor data, since time-synchronous data from different sensors can act as natural transformations of others (Jain et al., 2022; Deldari et al., 2022). However, most efforts have been focused on the model side, which involves complicated mathematics background and heavy tuning and implementation. Many data augmentation techniques are easy to implement and have been shown to improve performance without increasing the model's complexity.

4 Expected Timeline

The final project is expected to last for 6 weeks. The first week will be dedicated to getting familiar with the dataset, and performing appropriate data preprocessing and research on different DA techniques. The third week will be used to choose a set of diverse DA methods (time-domain, magnitude-domain, and frequency domain) and one or several deep learning models for implementation. The fourth and fifth weeks will be used to train, tune and evaluate the performance of the model, as well as find the best combination of DA. The original paper of the *extraSensory* dataset will be used as the baseline for evaluation. The write-up and preparation for the presentation will be done in the sixth week. However, I am aware that the plan is subjected to modification in case of emergency and based on the number of people working on it.

5 Resources

All the data processing and model training processes will be done using my laptop. Additional computing resources might be requested from the cluster if needed. In the case, if multiple students working on the project, we can use Northeastern's Teams and Github to facilitate communication and collaboration.

References

Shohreh Deldari, Hao Xue, Aaqib Saeed, Daniel V. Smith, and Flora D. Salim. 2022. Cocoa: Cross modality contrastive learning for sensor data. *Proc. ACM Interact. Mob. Wearable Ubiquitous Technol.*, 6(3).

Yash Jain, Chi Ian Tang, Chulhong Min, Fahim Kawsar, and Akhil Mathur. 2022. Collossl: Collaborative self-supervised learning for human activity recognition. *Proc. ACM Interact. Mob. Wearable Ubiquitous Technol.*, 6(1).

Dimitri Kraft, Karthik Srinivasan, and Gerald Bieber. 2020. Wrist-worn accelerometer based fall detection for embedded systems and iot devices using deep learning algorithms. PETRA '20, New York, NY, USA. Association for Computing Machinery.

Aaqib Saeed, Ye Li, Tanir Ozcelebi, and Johan Lukkien. 2020. Multi-sensor data augmentation for robust sensing. In *2020 International Conference on Omni-layer Intelligent Systems (COINS)*, pages 1–7.

Terry T. Um, Franz M. J. Pfister, Daniel Pichler, Satoshi Endo, Muriel Lang, Sandra Hirche, Urban Fietzek, and Dana Kulić. 2017. Data augmentation of wearable sensor data for parkinson's disease monitoring using convolutional neural networks. In *Proceedings of the 19th ACM International Conference on Multimodal Interaction*, ICMI '17, page 216–220, New York, NY, USA. Association for Computing Machinery.

Yonatan Vaizman, Katherine Ellis, and Gert Lanckriet. 2017. Recognizing detailed human context in the wild from smartphones and smartwatches. *IEEE Pervasive Computing*, 16(4):62–74.

Yonatan Vaizman, Nadir Weibel, and Gert Lanckriet. 2018. Context recognition in-the-wild: Unified model for multi-modal sensors and multi-label classification. *Proc. ACM Interact. Mob. Wearable Ubiquitous Technol.*, 1(4).