Exploring the Impact of Optimizers on CSI-Based Sign Language Recognition: A Comparison of AdaHessian and First-Order Methods

Joy Saha

February 22, 2025

1 Project Overview

Title: Optimizer Comparison for CSI-Based Sign Language Detection

Objective: This project aims to investigate the performance of different optimization algorithms in training a sign language detection system based on Channel State Information (CSI). The primary focus is on evaluating the AdaHessian optimizer[2], a second-order optimization technique, to determine if it leads to improved convergence and performance compared to traditional first-order optimizers like Adam and SGD. The goal is to understand how leveraging second-order methods like AdaHessian affects the training dynamics, accuracy, and robustness when applied to CSI data for sign language recognition.

2 Background and Motivation

The use of CSI data for human activity recognition, including sign language detection, has gained significant attention due to its potential in providing rich, detailed signals in wireless communication environments. AdaHessian introduces a second-order optimizer that incorporates curvature information, potentially offering a more refined approach to training models. The motivation behind this project is to explore whether second-order methods like AdaHessian provide any tangible benefits over first-order optimizers in this context, particularly in terms of convergence speed, noise robustness, and overall model performance. The specific advantages under investigation are:

- **Convergence Speed:** By incorporating curvature estimates, AdaHessian may reduce the number of training iterations required for convergence compared to first-order methods.
- **Robustness to Noisy Data:** CSI data is prone to noise. AdaHessian's features such as spatial averaging and momentum might help mitigate the impact of noise during training.
- **Performance Improvement:** We aim to assess if AdaHessian's adaptive learning rate adjustment for each parameter, due to second-order information, improves accuracy in detecting sign language gestures.

3 Project Objectives

1. **Data Acquisition:**
   - Collect high-quality CSI data for a variety of sign language gestures.
   - Pre-process the data with noise filtering and normalization techniques to prepare it for model training.
2. **Model Development:**
   - Develop a convolutional neural network (CNN) model to process the preprocessed CSI data.

- Compare different optimizers, including the second-order AdaHessian and first-order methods such as Adam and SGD, to evaluate their impact on training dynamics.
  - AdaHessian approximates the Hessian diagonal using techniques such as Hutchinson's method.
  - It incorporates spatial averaging to smooth out local curvature noise and uses momentum for stability.
  - These features enable AdaHessian to adjust learning rates adaptively, potentially resulting in faster convergence and more stable training.

3. **Performance Evaluation:**

- Evaluate the performance of the model using standard metrics (accuracy, precision, recall, F1-score).
- Benchmark the robustness of AdaHessian against first-order optimizers (e.g., Adam, SGD, AdamW) under varying noise conditions.
- Investigate how well the system performs under different user setups and environmental conditions to gauge its generalization capabilities.

4 CSI Dataset Overview

The sign gesture is performed in front of a WiFi Station (STA) that communicates with a nearby WiFi Access Point (AP), from which CSI data is collected. The CSI amplitude and phase, each originally having a size of $(3, 30, 200)$, are combined and reshaped by the input layer into a tensor of size $(200, 60, 3)$. Here, 200 represents the number of samples per antenna, 3 corresponds to the number of antennas, and 60 denotes the total number of subcarriers, where 30 are allocated for amplitude and 30 for phase. The CSI dataset has been collected in two different settings: Home and Lab[1]. Additionally, a multi-user setting has been used to capture diverse sign language data. The dataset details are summarized in Table 1.

Table 1: CSI Dataset Collection Overview

Dataset Setting	**Number of Signs**	**Repetitions per Sign**	**Total Samples**
Home Dataset	276	10	2,760
Lab Dataset	276	20	5,520
Multi-User Setting	150	10 per user (5 users)	7,500

5 Conclusion

This project investigates the effectiveness of different optimizers, with a primary focus on AdaHessian, for CSI-based sign language detection. By comparing AdaHessian, a second-order optimizer, to first-order optimizers such as Adam and SGD, the project aims to determine if second-order methods lead to better convergence speed, improved robustness to noisy data, and higher overall performance. The findings will provide insights into the utility of second-order optimization techniques in the context of CSI-based human activity recognition.

References

[1] Yongsen Ma, Gang Zhou, and Shuangquan Wang. WiFi Sensing with Channel State Information: A Survey. *ACM Comput. Surv.*, 52(3), Article 46 (June 2019), 36 pages. https://doi.org/10.1145/3310194.

[2] Zhewei Yao, Amir Gholami, Sheng Shen, Kurt Keutzer, and Michael W. Mahoney. ADAHESSIAN: An Adaptive Second Order Optimizer for Machine Learning. *CoRR*, abs/2006.00719, 2020. https://arxiv.org/abs/2006.00719.