Long Short-Term Memory Network for Elderly Movement Classification

Yunqing Zhao[*]
University of Washington
yzhao73@uw.edu

Shangming Zhuo[*]
University of Washington
oiviauw@uw.edu

Abstract

With the global aging population on the rise, ensuring the safety of elderly individuals living alone has become a key concern. Many elderly people face mobility issues, making them susceptible to falls or prolonged inactivity. To address this, we developed an LSTM-based posture recognition system that detects 11 activity types and provides real-time alerts for safety monitoring. Our goal is to offer a non-intrusive solution that enhances elderly independence while providing reassurance to family members.

Our code is available:

https://github.com/oliviashangming/513_Group8_datasheet-code

1 Introduction

1.1 Context and Importance of the Problem

With the global aging population increasing, ensuring the safety and well-being of elderly individuals, especially those living alone or frequently going out, has become an urgent social concern. Many elderly people experience reduced mobility, balance impairments, or underlying health conditions, making them more susceptible to falls or prolonged periods of inactivity. In such cases, delays in intervention can lead to serious consequences, including severe injuries or even life-threatening situations.

Traditional monitoring solutions include manual check-ins, surveillance cameras, and wearable emergency alert devices. Manual check-ins require regular in-person visits, but due to limited human resources, timely detection of emergencies is challenging. Surveillance cameras can provide video monitoring but pose privacy risks and require continuous observation, lacking automatic anomaly detection capabilities. Wearable emergency alert devices rely on user activation; however, in sudden emergencies, elderly individuals may be unable to use these devices promptly. Additionally, some regions have introduced smart wristbands equipped with emergency call, location tracking, and health monitoring functions, but issues such as false alarms, missed detections, and low willingness of elderly individuals to wear them persist.

Therefore, a more intelligent, automated, and non-intrusive monitoring system is essential to enhance elderly safety while preserving their independence and dignity.

1.2 Problem Statement

This project aims to develop an elderly activity recognition system based on machine learning to address these challenges. The system is designed to identify 11 distinct activities, including walking, falling, sitting, standing up, and lying down, among others. By applying machine learning techniques to posture data, the system can classify different activities in real-time and detect critical situations, such as falls or prolonged inactivity, enabling timely alerts for caregivers and family members.

Unlike traditional monitoring methods, this system leverages advanced machine learning models and real-time data processing to provide accurate and automated activity recognition. The objective is to create a precise, efficient, and non-intrusive solution that enhances elderly safety, reduces family members' concerns, and ultimately improves the quality of life and independence of elderly individuals.

1.3 Technical Foundation

The system utilizes publicly available Kaggle datasets containing labeled time-series coordinate data. Through machine learning methods, we establish correlations between limb kinematics and activity states. To reduce computational complexity while maintaining diagnostic precision, we focus on four critical skeletal points, enabling real-time analysis on edge devices without relying on

[*]Authors contributed equally to this work

cameras, thereby safeguarding the privacy of the elderly.

1.4 System Architecture

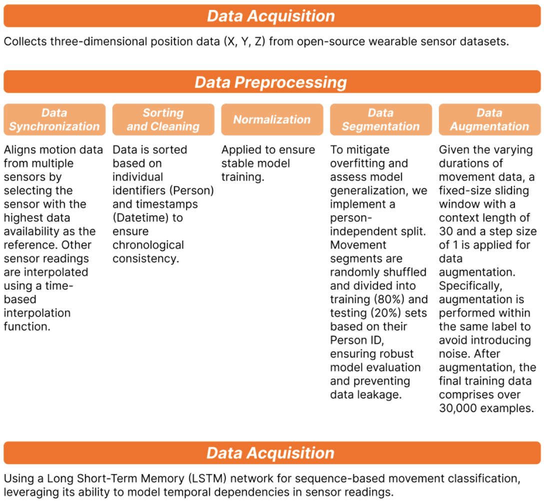

Figure 1: Overall schematic.

1.5 Related Work

Existing elderly activity monitoring solutions can be broadly classified into three main categories: wearable sensor-based systems, vision-based systems, and ambient sensor-based systems. Wearable Sensor-Based Systems: Many elderly monitoring solutions utilize wearable devices such as smartwatches or inertial measurement units (IMUs) to track movement patterns and detect falls. These systems typically determine falling actions by detecting rapid changes in specific axis values. However, they provide limited assistance in recognizing and identifying daily activities beyond fall detection. Vision-Based Systems: Camera-based monitoring systems analyze video feeds to detect falls and inactivity. While these systems achieve high accuracy, they present significant privacy concerns and environmental dependencies. If an elderly individual leaves the monitored area, the system becomes ineffective. Furthermore, their performance is influenced by lighting conditions, obstructions, and camera placement. Ambient Sensor-Based Systems: These solutions employ pressure sensors, infrared sensors, or radar technologies to monitor movement patterns within a home environment. While non-intrusive, they often require extensive installation and may struggle to accurately differentiate between various postures and activities.

Although these monitoring systems have made progress, existing solutions often fail to strike a balance between accuracy, usability, and privacy. Traditional wearable motion recognition devices are limited to detecting specific actions, vision-based systems raise privacy concerns, and ambient sensors may lack precision in posture recognition. Our approach improves upon these existing solutions by leveraging an LSTM-based posture recognition model to process multi-sensor data collected from IMUs. By integrating multiple sensors and employing deep learning techniques, our system achieves high accuracy while preserving user privacy and comfort.

2 Methods

Our method is composed of two parts, 1. the time series data preprocessing, 2. model training and evaluation.

2.1 Preprocessing

We collect open-sourced location wearable sensors data (X,Y,Z). In detail, we collect tabular data with each entry a 3D location sampled from the sensor along with its time stamp. We first apply Data Synchronization: Motion data from multiple sensors are aligned by selecting the sensor with the highest data availability as the reference. Other sensor readings are interpolated using a time-based interpolation function. After that, we apply Sorting and Cleaning: Data is sorted based on individual identifiers (Person) and timestamps (Datetime) to ensure chronological consistency.We also use normalization to ensure a stable training. To mitigate overfitting and assess model generalization, we implement a person-independent split. Movement segments are randomly shuffled and divided into training (80%) and testing (20%) sets based on their Person ID. This ensures a robust model evaluation, preventing data leakage. After aligning the timestamp for all sensors and interpolation, our data has 12 features coming from the four 3D location sensors on one person.

As the movement data varies in their duration, we apply a fixed-size sliding window of context length 30 and step size 1 for data augmentation. Specifically, we augment the data within the same label to avoid introducing noise. After augmentation, our final training data have more than 30,000 examples.

2.2 Model training

To classify a time series data, we propose to use the Long Short-Term Memory (LSTM) network for sequence-based movement classification, given its ability to model temporal dependencies in sensor readings.

3 Experiments

3.1 Experiment Setup

Our model architecture consists of three LSTM layers with 128, 64, and 32 hidden units. We also add two dense layers with a softmax head to output a probability over labels. Furthermore, we apply dropout with dropout ratio 0.3 for all LSTM layers to avoid overfitting. We use LeakyReLU as the activation. We train our model using Adam optimizer with an initial learning rate 0.01. We set the batch size to 32 and train the model for 50 epochs. We also apply an exponentially decaying learning rate. Finally, we apply early stopping with patience 5. We evaluate the model's accuracy and loss using the hold-out validation set. We conduct our experiments on a personal Laptop with Apple M2 CPU.

3.2 Baseline

We propose to use the Gated Recurrent Unit (GRU) as the baseline method given its computational efficiency.

4 Experimental Results

We first show the results for our model in Figure 2. We observe that both the training and validation loss decreases during training. Correspondingly, the accuracy increases and finally reaches 0.9563 on the validation set. This results highlight the effectiveness of our approach.

4.1 Baseline Comparison

We compare our model with the baseline, a GRU-based model in Figure 3. Specifically, the GRU-based method achieves a validation accuracy of 0.9263. This results shows that our LSTM-based model outperforms the GRU based model, further validating the superiority of the LSTM model.

4.2 Ablation Study

We study the effect of learning rate on the model performance. In this comparison, we set the learning rate to 0.1. We plot the training and validation metrics in Figure 4. We show that the model does not convergent and thus is sensitive to the learning rate.

Conclusion and Limitations

Conclusion: In this work, we use a LSTM model for time series data classification. Starting with raw unaligned 3D location data and through progressive data preprocessing and augmentation, we curate a high quality movement data for classification. We conduct experiments to validate the model performance and compare it with the baseline model. Through ablation study on the learning rate, we show the model is sensitive to this hyperparameter.

Limitation: Also LSTM achieves a high accuracy, we argue that a finer grained sensor data and a more powerful model is needed for practical deployment. Also, our current model is trained and tested on a fixed context window, a dynamic context window might be needed for further usage.

Division of work: This project was a fully collaborative effort between Shangming and Yunqing. Throughout the development process, both members actively participated in all major steps. Rather than strictly dividing tasks, we worked together in each phase, discussing multiple approaches, testing different strategies, and continuously refining our methodology by selecting the most effective solutions from our combined efforts to ensure the best possible outcome. Additionally, both members contributed equally to documentation and final report writing.

References

Lin, Hsin-Chang, et al. "Fall Recognition Based on an IMU Wearable Device and Fall Verification through a Smart Speaker and the IoT." Sensors, vol. 23, no. 12, 2023, p. 5472, doi:10.3390/s23125472.

Yu, Xiaoqun, et al. "A Large-Scale Open Motion Dataset (KFall) and Benchmark Algorithms for Detecting Pre-Impact Fall of the Elderly Using Wearable Inertial Sensors." *Frontiers in Aging Neuroscience*, vol. 13, 15 July 2021, [doi:10.3389/fnagi.2021.692865]

UCI Machine Learning and Collaborator. Localization Data for Posture Reconstruction. Kaggle, Updated 8 years ago, https://www.kaggle.com/datasets/uciml/posture-reconstruction. Use ChatGPT to assist with technical and grammar checking.

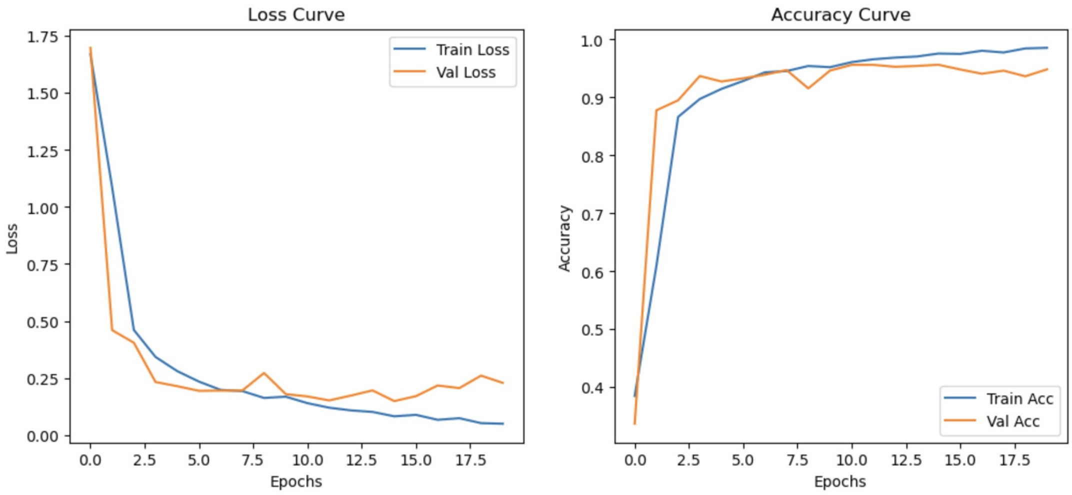

Figure 2: The training and validation metrics during training.

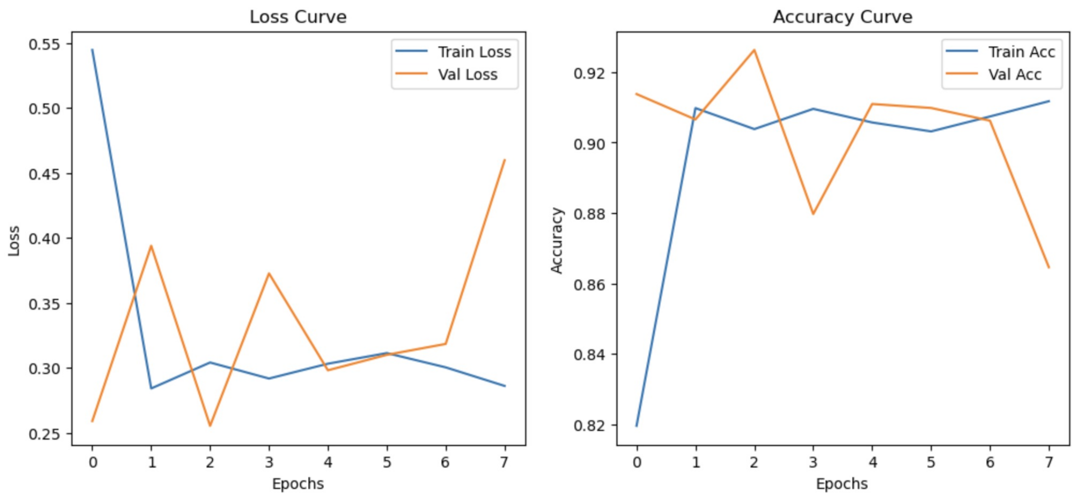

Figure 3: The training and validation metrics during training.

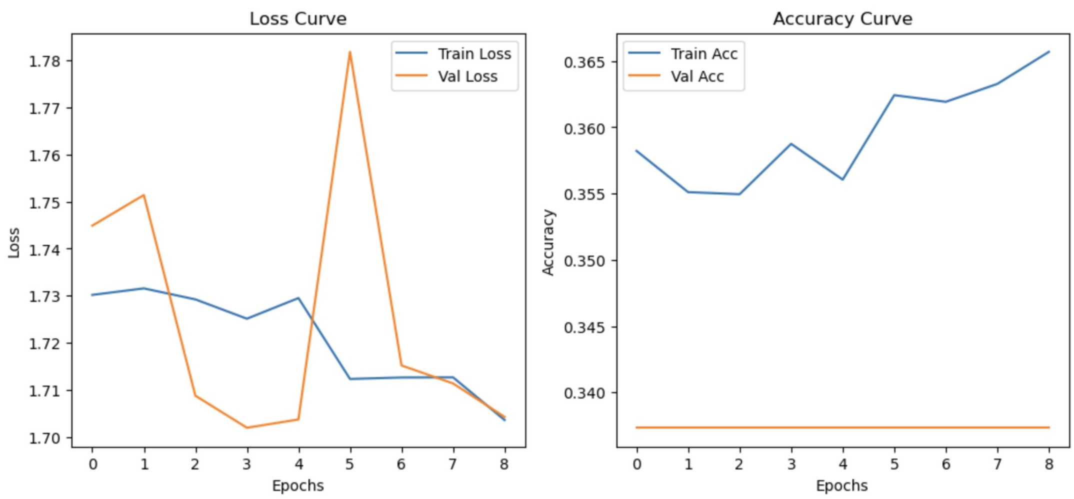

Figure 4: The training and validation metrics during training.

Cho, Kyunghyun, et al. "Learning phrase representations using RNN encoder-decoder for statistical machine translation." arXiv preprint arXiv:1406.1078 (2014).

Hochreiter, Sepp, and Jürgen Schmidhuber. "Long short-term memory." Neural computation 9.8 (1997): 1735-1780.