Wearable Technology & Mobile Perception for Monitoring Stability in IADLs in Older Adults with MCI

Zhenhao Zhao[1] Victoria Pham[2] Leslie Davidson[2] Chung Hyuk Park[1*] and Keith Cole[2*]

***Abstract*— Mild cognitive impairment (MCI) is associated with a higher risk of falls in older adults, which can result in severe injuries and significant financial costs. Cognitive-motor dual-task tests show promise in predicting MCI-related falls, but their clinical applicability is limited. This study aims to develop an intelligent mobile perception system for assistive robots or wearable devices, using deep learning techniques to analyze kinematic data from daily activities. We will recruit 20 MCI subjects and 20 cognitively normal subjects to perform grocery shopping in a simulated environment, which requires both cognitive and motor skills. Inertial measurement units and cameras will be used to collect kinematic data, and feature extraction methods will identify performance parameters for gait and non-gait based activities. The goal of this research is to create a framework using kinematic data and advanced deep learning algorithms to assess older adults' ability to perform naturalistic movements, ultimately reducing their risk of falls.**

I. INTRODUCTION

Mild cognitive impairment (MCI) is a condition characterized by a decline in cognitive function greater than what would be expected for an individual's age and level of education [7]. Although MCI does not significantly impair daily activities, it increases the risk of falling. Recent studies have shown that people with MCI have a significantly higher risk of falling, with an odds ratio (OR) of 1.98 and a 95% confidence interval of 1.11-3.53 [4], [11]. Falls are a serious concern as they can lead to debilitating injuries, such as broken bones and head injuries [1], [15]. They can also have a significant financial impact, with an average cost of \$62K$\sim$\$64K for a fall or fall with any injury [5].

Cognitive-motor dual-task tests have shown promise in predicting MCI and related falls [2]. However, there are concerns that tests performed in a clinical environment do not translate to function in a real-world daily environment[13]. To address this issue, we aim to build an intelligent mobile perception that can be applied to assistive robots or wearable devices for performing the dual-task tests in real-world environments. We will accomplish this through using state-of-the-art deep learning (DL) techniques to recognize and analyze the kinematic data of instrumental activities of daily living (IADL) collected from inertial measurement units (IMUs) and cameras. Specifically, we will recruit 20 MCI subjects and 20 cognitively normal subjects to perform grocery shopping in a simulated grocery store. Grocery shopping requires both cognitive abilities, such as finding the specific shopping items, and motor skills, such as walking with a basket. Finally, we will use feature extraction methods to identify specific kinematic performance parameters for each gait and non-gait-based activities. Our aim of this research is to use both kinematic data and advanced deep learning algorithms to develop a framework for recognizing and determining differences in naturalistic movements for older cognitively healthy adults and those with MCI.

[1]*Assistive Robotics and Tele-Medicine (ART-Med) Lab., Department of Biomedical Engineering, The George Washington University,* Washington, DC, 20052, USA

[2]*Department of Health, Human Function, and Rehabilitation Sciences, School of Medicine and Health Sciences,, The George Washington University,* Washington, DC, 20052, USA

[*]*Corresponding authors*

II. METHOD

A. Data Collection

A simulated grocery store was constructed in order to collect IADL kinematic data for the human subjects. Table I shows the IADL in grocery shopping. Each subject will have 4 IMUs on their body (see *A-1)*) and an action camera (for single person point of view), and two depth cameras (Azure Kinect and ZED-2) will be setup in the environment. Each subject will perform a structured grocery shopping task of navigating the simulated store to find items on a list. Each subject will perform five separate lists of shopping. To further investigate the relationship between cognition and the IADL kinematic data, we also perform cognitive testing using the Clinical Dementia Rating Scale and the NIH Toolbox Cognition Battery.

IADL Tasks involved in Grocery Shopping
Static Stance
Static Stance with head turns (searching)
Walking without a basket
Walking with a basket containing 5 pounds
Walking with a basket containing 10 pounds
Turning around a corner
Reaching down to pick an item on a low shelf
Reaching up for an item on a high shelf

TABLE I

LIST OF DISCRETE BEHAVIORS TO MONITOR IN IADL.

1) IMU Data:

The study participants will be outfitted with four inertial measurement units (IMUs) to collect movement data, with each IMU housing three sensors: accelerometers, gyroscopes, and magnetometers. Specifically, one IMU will be placed on the head, one on the lumbar region, and one on each foot. The collected IMU data will be meticulously labeled for each movement and subjected to Fourier Transform-based denoising techniques to prepare it for the next step of deep learning training.

2) Video Data: Each subject will be monitored using a tri-camera setup, comprising of a first-person point-of-view action camera and two third-person point-of-view depth cameras. The first-person point-of-view action camera, which is equipped with an anti-shake function, will be mounted on the subject's chest to capture ground truth actions during the IADL tasks. Additionally, two third-person point-of-view depth cameras equipped with depth sensors will be used to capture 3D information of the subject's movements, which can be used for deep learning model training and to assist with movement feature extraction.

B. Deep Learning IADL recognition Model

1) Video Data Processing for IADL Recognition: For the temporal recognition of IADL in video-based continuous data streams, common approaches include utilizing recognition algorithms such as Long Short-Term Memory (LSTM) [8] or transformer-based action recognition models [3]. Alternatively, Convolutional Neural Network (CNN)-based models, such as the fast-slow ResNet [6], can also be employed. At this stage, the optimal model for our application remains undetermined. The model's output will be a one-dimensional (1D) feature vector, which will be prepared for subsequent decoding or feature fusion processes. Fig. 1 shows the workflow of this processing.

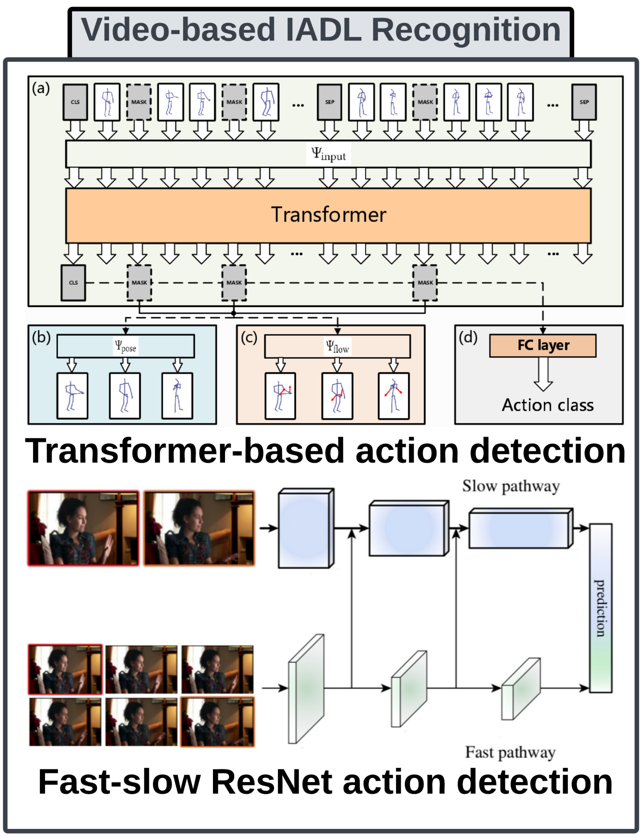

Fig. 1. Video Data Processing for IADL Recognition.

2) IMU Data Processing for IADL Recognition: Alongside video data, we will employ kinematic data gathered via IMU to facilitate the recognition of IADL. For IMU data processing, we will first apply Fourier transformation and other denoising techniques to preprocess the raw data, preparing it for input into neural networks. Subsequently, we will employ either a CNN-based deep learning model [12] or a transformer-based deep learning model [14], [10]. This step will produce a feature matrix, which will be prepared for decoding or further processing. Fig. 2 shows the workflow of this processing.

3) Feature Fusion and Multi-Modal Recognition: Following the completion of the previous steps, we will have a 1D vector representing visual information and a matrix representing IMU kinematic information. We will fuse these data sources using deep learning techniques. Initially, we will attempt matrix concatenation, inputting the fused matrix into a fully connected network to obtain preliminary results, which will serve as our baseline. Subsequently, we will explore more advanced feature fusion techniques, such as transformers [17], [18], to integrate the features and perform the final IADL recognition.

C. Features Extraction

Once we recognize each IADL for the subject's movement, we can extract features from the healthy and MCI subjects' movements. Temporal and frequency domain features of gait and non-gait movements will be analyzed. Statistical analysis will determine differences between groups for key features. Significant features may be fed back into the deep learning algorithm, though We will make appropriate adjustments as necessary. The workflow of this study is illustrated in Fig. 3.

III. Initial Result

A. Grocery Store Setting up

We constructed a simulated grocery store in the lab. The layout of the simulated grocery store imitates the real grocery store as much as possible to give the subject a realistic shopping experience. Perishable items such as fruit and bread, were replaced with one-to-one simulation models, and filled with sand to simulate the weight of real products. Other products with a longer shelf life, such as candy and paper towels, use real products directly. We also made labels with bar codes and corresponding prices for each product to better simulate the real shopping environment. At the same time, we also partitioned the entire grocery store area with screens and hidden the camera position as much as possible to give the subjects an immersive shopping experience. The Fig. 4 shows the setting of the grocery store.

B. Data Structure of the Subjects Data

We have currently collected data on 19 subjects and expect to have data on 21 more subjects to be collected. All subject data are stored anonymously, and the data are only marked by subject number and date of data collection. The general idea of the data storage structure is *Data Type* $\rightarrow$ *Subject* $\rightarrow$ *Round* $\rightarrow$ *Video/CSV file*.

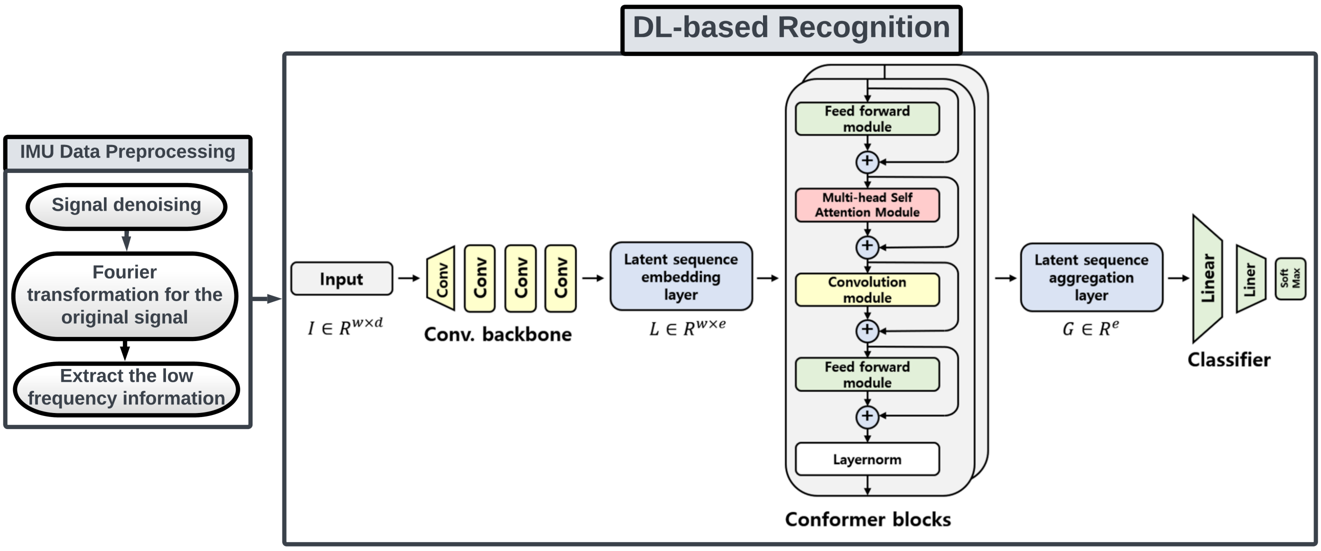

Fig. 2. IMU Data Processing for IADL Recognition.

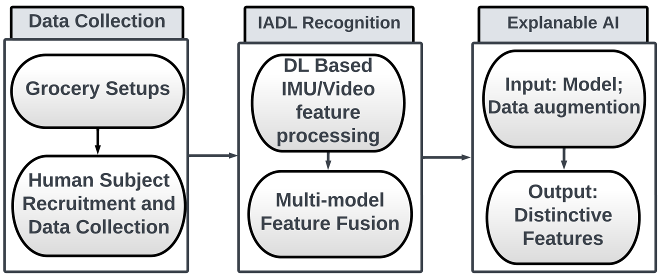

Fig. 3. Data analysis and deep learning process flow.

C. Deep Learning Model and Initial Data Analysis

Currently, our main focus is on subject recruitment and data collection, preparation, and analysis. In order to perform the required data analysis. We are assembling a range of actions of daily living (ADL) and instrumental activities of daily living (IADL) datasets [16], [9]. We plan to use these datasets for our initial training and then apply transfer learning to our self-collected data.

We will also apply signal processing algorithms (e.g., Fourier transformation) to extract gait cycles as well as non-gait tasks (e.g. bending) during the tasks from the IMU data. Another aspect of our work involves tracking the skeleton of the human subjects, thereby laying the groundwork for future visual analysis. As demonstrated in Fig. 5, we provide an example of preliminary analysis of our gait and behavior analysis.

IV. Conclusion

In summary, our study aims to introduce a deep-learning-based multi-modal fusion recognition algorithm. This algorithm, incorporating both visual and IMU kinematic perceptions, will be trained using self-collected data from simulated grocery shopping tasks. Additionally, we will employ

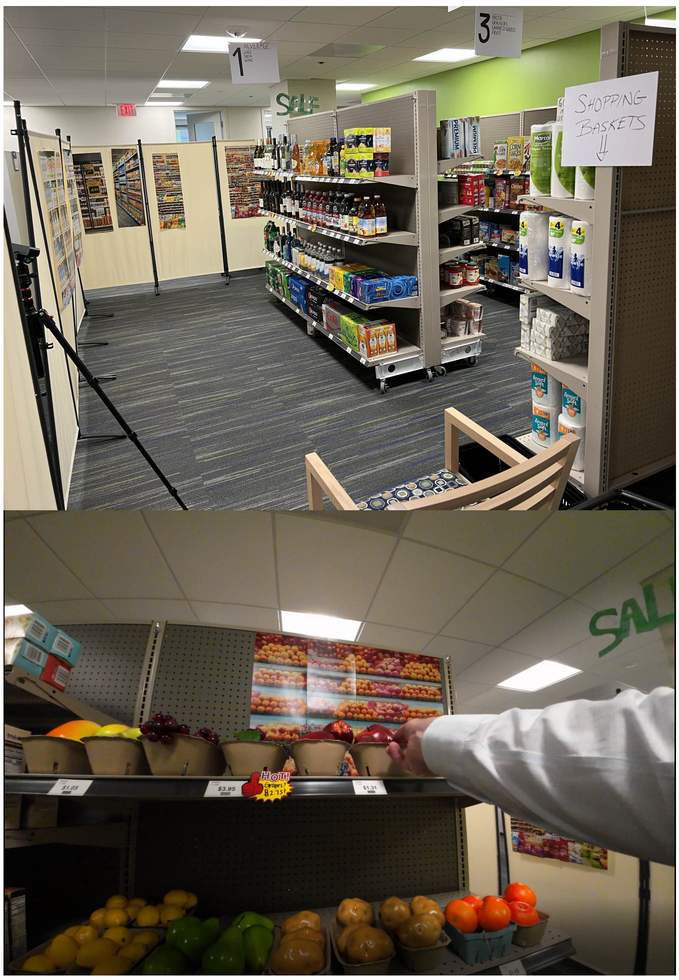

Fig. 4. Top: Third-person point-of-view camera view for the simulated grocery; Bottom: First-person point-of-view camera view of the subjects.

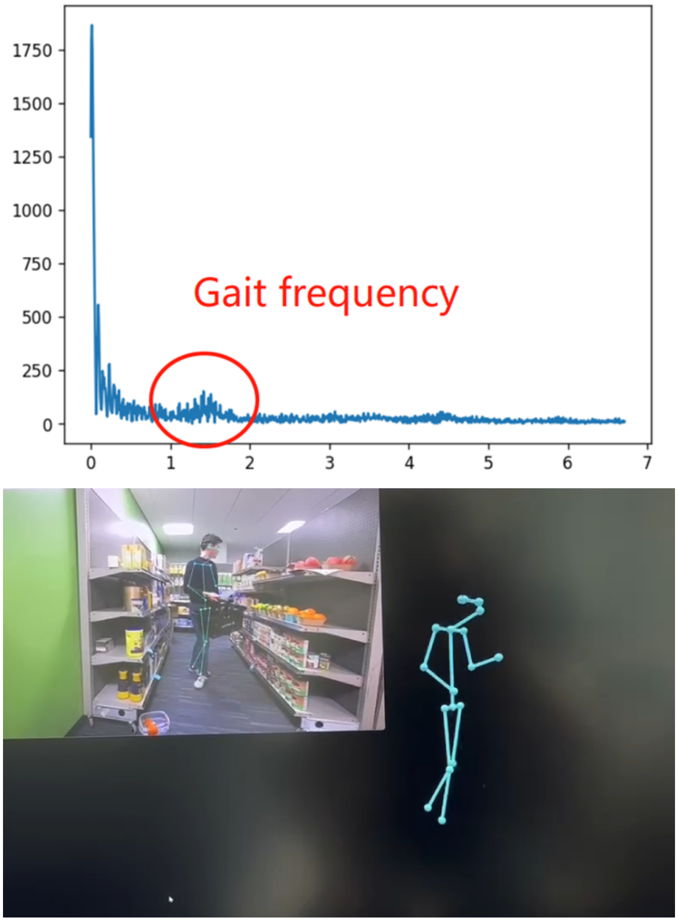

Fig. 5. Top: Fourier transformation for the initial gait analysis; Bottom: Skeleton tracking for the human subjects.

feature extraction techniques to discern specific performance parameters for both gait and non-gait based activities. The culmination of these efforts will result in a comprehensive framework designed to evaluate the capability of older adults to execute daily naturalistic activity. The ultimate goal of this endeavor is to mitigate their risk of falls, thereby enhancing their overall safety and well-being.

References

[1] B. H. Alexander, F. P. Rivara, and M. E. Wolf. The cost and frequency of hospitalization for fall-related injuries in older adults. *American journal of public health*, 82(7):1020–1023, Jul 1992.

[2] A. Bishnoi and M. E. Hernandez. Dual task walking costs in older adults with mild cognitive impairment: a systematic review and meta-analysis. *Aging & Mental Health*, 25(9):1618–1629, 2021. PMID: 32757759.

[3] Y.-B. Cheng, X. Chen, D. Zhang, and L. Lin. Motion-transformer: self-supervised pre-training for skeleton-based action recognition. *Proceedings of the 2nd ACM International Conference on Multimedia in Asia*, 2021.

[4] K. Delbaere, N. A. Kochan, J. C. Close, J. C. Menant, D. L. Sturnieks, H. Brodaty, P. S. Sachdev, and S. R. Lord. Mild cognitive impairment as a predictor of falls in community-dwelling older people. *The American Journal of Geriatric Psychiatry*, 20(10):845–853, 2012.

[5] P. C. Dykes, M. Curtin-Bowen, S. Lipsitz, C. Franz, J. Adelman, L. Adkison, M. Bogaisky, D. Carroll, E. Carter, L. Herlihy, M. E. Lindros, V. Ryan, M. Scanlan, M.-A. Walsh, M. Wien, and D. W. Bates. Cost of Inpatient Falls and Cost-Benefit Analysis of Implementation of an Evidence-Based Fall Prevention Program. *JAMA Health Forum*, 4(1):e225125–e225125, 01 2023.

[6] C. Feichtenhofer, H. Fan, J. Malik, and K. He. Slowfast networks for video recognition. In *Proceedings of the IEEE/CVF International Conference on Computer Vision (ICCV)*, October 2019.

[7] S. Gauthier, B. Reisberg, M. Zaudig, R. C. Petersen, K. Ritchie, K. Broich, S. Belleville, H. Brodaty, D. Bennett, H. Chertkow, J. L. Cummings, M. de Leon, H. Feldman, M. Ganguli, H. Hampel, P. Scheltens, M. C. Tierney, P. Whitehouse, and B. Winblad. Mild cognitive impairment. *The Lancet*, 367(9518):1262–1270, 2006.

[8] S. Hochreiter and J. Schmidhuber. Long short-term memory. 9(8):1735–1780, nov 1997.

[9] W. Kay, J. Carreira, K. Simonyan, B. Zhang, C. Hillier, S. Vijayanarasimhan, F. Viola, T. Green, T. Back, P. Natsev, M. Suleyman, and A. Zisserman. The kinetics human action video dataset, 2017.

[10] Y. Kim, W. Cho, K. Kim, and S. Lee. Inertial-measurement-unit-based novel human activity recognition algorithm using conformer. *Sensors (Basel)*, 22(10):3932, May 2022.

[11] T. Y. Liu-Ambrose, M. C. Ashe, P. Graf, B. L. Beattie, and K. M. Khan. Increased Risk of Falling in Older Community-Dwelling Women With Mild Cognitive Impairment. *Physical Therapy*, 88(12):1482–1491, 12 2008.

[12] R. Li, C. Fu, W. Yi, and X. Yi. Calib-net: Calibrating the low-cost imu via deep convolutional neural network. *Frontiers in Robotics and AI*, 8, 2022.

[13] J. Ren, E. Giannouli, O. Bock, S. Mellone, and W. Zijlstra. Mobility in old age: Capacity is not performance. *BioMed Research International*, 2016:3261567, 2016.

[14] Y. Shavit and I. Klein. Boosting inertial-based human activity recognition with transformers. *IEEE Access*, 9:53540–53547, 2021.

[15] D. A. Sterling, J. A. O'Connor, and J. Bonadies. Geriatric falls: injury severity is high and disproportionate to mechanism. *Journal of Trauma–Injury, Infection and Critical Care*, 50(1):116–119, 2001.

[16] UCI Machine Learning Repository. Activities of daily living (adls) recognition using binary sensors, 2023. Accessed: 2023-05-09.

[17] A. Vaswani, N. Shazeer, N. Parmar, J. Uszkoreit, L. Jones, A. N. Gomez, L. u. Kaiser, and I. Polosukhin. Attention is all you need. In I. Guyon, U. V. Luxburg, S. Bengio, H. Wallach, R. Fergus, S. Vishwanathan, and R. Garnett, editors, *Advances in Neural Information Processing Systems*, volume 30. Curran Associates, Inc., 2017.

[18] B. Xie, M. Sidulova, and C. H. Park. Robust multimodal emotion recognition from conversation with transformer-based crossmodality fusion. *Sensors*, 21(14), 2021.