FALL RISK DETECTION BY CLASSIFYING OBSTACLE INTERACTION USING MACHINE LEARNING

Yejin Lee[1,2], Jeongmin Kim[1,2], Kang-Ho Lee[1], Hak Yi[2], Keonyoung Oh[2], Jang-Hyuk Cho[3], Kyoung-Tae Kim[3], Yongkoo Lee[1]
[1]Korea Institute of Machinery and Materials, Daegu, Republic of Korea
[2]School of Mechanical Engineering, Kyungpook National University, Daegu, Republic of Korea
[3]Department of Rehabilitation Medicine, Keimyung University Dong-San Hospital, Daegu, Republic of Korea
Corresponding Author Email: ylee2012@kimm.re.kr

Introduction

Falls during gait primarily occur due to tripping over obstacles or losing balance. There is a need for technologies that can detect and respond to such events at an early stage. The aim of this study is to accurately classify obstacle trip events using gait data from 28 healthy participants and to analyze the minimum proportion and combination required to detect obstacle contact after the initiation of the swing phase.

Methods

During the swing phase, interactions with obstacles were classified into three classes based on the level of risk [1]. All datasets were resampled to 100 data points using cubic interpolation. Additionally, the dataset was divided into 10 segments with a 10% ratio. This study aims to identify the minimum data ratio required to achieve optimal fall detection accuracy and to evaluate performance differences based on the combination of hip, knee, and ankle joint angles. The algorithms used included tree-based methods such as Random Forest (gini), Extra Trees (gini, entropy), LightGBM, XGBoost, and CatBoost.

Results and Discussion

Compared to the Swing Phase, the improvement in accuracy with an increasing data ratio was relatively smaller in the post-Stance Phase region. However, accuracy generally tended to improve as the data ratio increased. Notably, models achieved near-optimal performance when data from the post-trip Stance Phase (40–60%) was included. As shown in Figure 1, accuracy significantly improved within the 40–60% data ratio range, whereas additional data in the 70–100% range contributed minimally to learning performance, indicating a data convergence effect. While similar trends were observed across models, the LightGBM model demonstrated relatively superior performance.

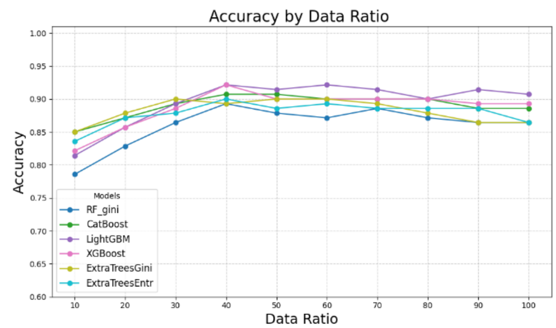

Figure 1: Accuracy by Data Ratio

Conclusions

The performance of intention recognition algorithms depends not only on the characteristics of the data but also on the selection and combination of joint locations and channels. Performance differences based on the choice of location and dimension have a significant impact on the practical application of the algorithm in real-world scenarios. In this study, the optimization of these factors aims to maximize the efficiency of fall prevention systems, ultimately ensuring the safety of individuals during gait.

Acknowledgments

This research was supported by the NK254H projects (Development of core mechanical technology implementing immersive extended reality for overcoming severe motor dysfunction) of the Korea Institute Machinery & Materials

References

[1] JM. Kim, et al. (2025). *J Mech Sci Tech*, **39(3)**.
[2] Liu Y et al. (2023). Comput Biol Med, 163: 107192.

Table 1: The best-performing combination for each joint angle

Part	Data Ratio	Model	Accuracy	Sensitivity (HIGH)	Sensitivity (LOW)
Hip	50	ExtraTreesEntr	0.792	0.90	0.625
Knee	40	CatBoost	0.907	0.95	1.000
Ankle	60	ExtraTreesGini	0.871	0.95	0.812
Hip+Knee	50	LightGBM	0.921	0.95	0.937
Knee+Ankle	60	XGBoost	0.914	0.96	0.937
Hip+Ankle	30	ExtraTreesGini	0.878	0.95	0.875
Hip+Knee+Ankle	40	LightGBM	0.921	0.95	1.000