Machine Learning for Movement Pattern Changes during Kinect-Based Mixed Reality Exercise Programs in Women with Possible Sarcopenia: Pilot Study

Yunho Sung a,1, Ji-won Seo a, Byunggul Lim a,b, Shu Jiang a, Xinxing Li a, Parivash Jamrasi a, So Young Ahn a, Seohyun Ahn a, Yuseon Kang a, Hyejung Shin a, Donghyun Kim a, Dong Hyun Yoon c,d, and Wook Song a,c,e,*

*a Department of Physical Education, Seoul National University, Seoul, Korea
bResearch Institute, Dr.Exsol.Inc, Director of Research, Seoul, Korea
cInstitute on Aging, Seoul National University, Seoul, Korea
dDepartment of Rehabilitation Medicine, Seoul Metropolitan Government Boramae Medical Center,
Seoul, Republic of Korea
eInstitute of Sport Science, Seoul National University, Seoul, Korea

Corresponding author

Wook Song

Department of Physical Education, Seoul National University, Seoul, Korea. Institute on Aging, Seoul National University, Seoul, Korea. Institute of Sport Science, Seoul National University, Seoul, Korea.

E-mail: songw3@snu.ac.kr

1 Corresponding Wook Song, Department of Physical Education, Seoul National University, Seoul, Korea. Email: songw3@snu.ac.kr
ORCID ID Numbers
Yunho Sung : 0000-0002-4493-250X
Ji-won Seo : 0000-0003-0279-7562
Byunggul Lim : 0000-0002-4819-7370
Shu Jiang : 0000-0001-6790-300X
Xinxing Li : 0000-0002-1221-428X
Parivash Jamrasi : 0000-0002-5262-8655
So Young Ahn : 0000-0002-0075-7500
Seohyun Ahn : 0009-0001-4590-2214
Yuseon Kang : 0009-0007-3894-9819
Hyejung Shin : 0000-0002-1409-8833
Donghyun Kim : 0000-0001-6862-5566
Dong Hyun Yoon : 0000-0002-3997-2670
Wook Song : 0000-0002-8825-6259

Running title
ML for Sarcopenic Women's Movement in KMR

Conflict of Interest
The authors declare that they have no conflicts of interest regarding the publication of this article
Abstract

Background: Sarcopenia is a muscle wasting condition that affects elderly individuals. It can lead to changes in movement patterns, which can increase the risk of falls and other injuries. Methods: Elderly women participants aged ≥65 years who could walk independently were recruited and classified into two groups based on knee extension strength (KES). Participants with low KES scores were assigned to the possible sarcopenia group (PSG, n=7) and an 8-week exercise intervention was implemented. Healthy seniors with high KES scores were classified as the reference group (RG, n=4), and a 3-week exercise intervention was conducted. Kinematic movement data were recorded during the intervention period. All participants' exercise repetitions were used in the data analysis (number of data points =1,128).

Results: The PSG showed significantly larger movement patterns in knee rotation during wide squats compared to the RG, attributed to weakened lower limb strength. The voting classifier, trained on the movement patterns from wide squats, determined that significant differences in overall movement patterns between the two groups persisted until the end of the exercise intervention. However, after the exercise intervention, significant improvements in lower limb strength in the PSG resulted in reduced knee rotation ROM and Max, thereby stabilizing movements and eliminating significant differences with the RG.

Conclusions: This study suggests that exercise interventions can modify the movement patterns in elderly individuals with possible sarcopenia. These findings provide fundamental data for developing an exercise management system that remotely tracks and monitors the movement patterns of older adults during exercise activities.

Keywords: Exercise training, Sarcopenia, Machine Learning, Movement, Mixed Reality

1. Introduction

Sarcopenia, an age-related muscle disorder in older adults, is characterized by decreased muscle mass, strength, and function (1). Those affected often struggle to maintain functional activities of daily life, leading to reduced physical activity, diminished quality of life, and
increased risk of falls and fractures. These changes may advance to a vicious cycle, leading to long-term care facility admission and mortality, thereby emphasizing the need for early diagnosis and exercise prescription (2–4). Patients with severe sarcopenia face a fourfold higher risk of mortality within 2.6 years owing to various complex reasons (5).

Therefore, early diagnosis, exercise prescription, and self-management are essential for sarcopenia prevention and treatment (6). The Kinect device (Microsoft, Redmond, WA, USA) enables exercise, early diagnosis, and self-management. The Kinect V2 captures and digitizes the movements of 25 body joints in real time without requiring markers, providing a foundation for comparative movement analysis (7). Kinect-based mixed reality (KMR) offers consistent feedback per standard and analyzes user movements during exercise. In the minimal-contact era, KMR enables remote, no-contact exercise interventions using stored data for analysis. Furthermore, previous studies on unsupervised exercise interventions using KMR have reported significant improvements in lower-body strength (8). Therefore, this device is useful for both training and diagnosing abnormal movements.

This study investigated the movement patterns of older adults with possible sarcopenia and examined how exercise interventions altered these patterns. Previous research has predominantly conducted cross-sectional analyses of movement patterns in frail older adults, often noting issues such as insufficient muscle contraction in the calves during walking (9) and a leaning trunk when rising from a chair (10). Although machine-learning models have been employed to automatically classify frailty levels, few studies have explored the effects of exercise interventions on movement patterns changes (11). Therefore, we aimed to determine whether the movement patterns of older adults with sarcopenia can be distinguished from those of healthy individuals, and whether exercise interventions can align their movement patterns more closely with healthy standards. Our findings may contribute significantly to the development of diagnostic and treatment methods for specific conditions, ultimately enhancing the quality of life of older adults by observing both cross-sectional differences and long-term changes through exercise interventions.
2. MATERIALS AND METHODS

The Institutional Review Board of Seoul National University reviewed and approved this study (IRB No. 2206/001-015). The trial was conducted in accordance with the principles of the Declaration of Helsinki.

2.1. Study Participants

This prospective study recruited women aged ≥65 years who were capable of walking and exercising independently. Individuals who had undergone surgery within the last 6 months, which could affect joint movement, were excluded. Among the 14 enrolled participants, 11 completed the study. The recruited participants were divided into a possible sarcopenia group (PSG) and a reference group (RG) based on their lower-limb strength. The Asian Working Group for Sarcopenia recommends assessing grip strength or sit-to-stand time to identify “possible sarcopenia” in primary care settings (12, 13). While handgrip strength is a convenient tool for measuring upper-body strength, lower-body muscles have the highest distribution in the human body and significantly impact daily activities and motions. Therefore, to analyze the movement patterns, we used knee-extension strength (KES) to classify sarcopenia. We assessed KES using thresholds of <18 kg for men and <16 kg for women (14–20). The overall flow of this study is shown in Figure 1.

Figure 1. CONSORT flow diagram of the study participants.
Study design for an 8-week exercise intervention program for individuals in the possible sarcopenia group and a 3-week exercise program in the reference group.

2.2. Study Procedure

Both the PSG and RG underwent a pre-assessment, with the PSG undergoing an additional post-assessment and an 8-week exercise intervention. During the intervention period, the first and second weeks served as adaptation periods for the participants to familiarize themselves with proper exercise techniques and KMR use. The movement pattern analysis utilized PSG data from weeks 3–8. Therefore, the RG underwent a 3-week exercise intervention to align with the PSG data. We compared exercise data from the PSG with that from week 3 of healthy older adults in the RG.

The KMR-based exercise intervention was conducted twice a week for approximately 30 min. The participants performed the exercises at a rate of perceived exertion level of 12–15. The exercise intervention consisted of six movements: an overhead-side bend, wide squat, shoulder press, straight leg deadlift, two-arm dumbbell row, and dumbbell floor press.
2.3. Measurements

2.3.1. Demographic characteristics and anthropometric measures

Information on participants’ ages was recorded. Height was measured using a stadiometer, while weight, body mass index (BMI), and skeletal muscle mass were measured using bioelectrical impedance analysis (InBody 320; Inbody, Seoul, South Korea). All measurements were made by the same researcher using the same device at the same location and time, with participants wearing lightweight clothing and excluding metal components.

2.3.2. Muscle strength

We measured the muscle strengths of the upper and lower extremities. Upper-extremity strength was measured using a digital handgrip dynamometer (TKK5401; Takei, Niigata, Japan). The average values of two repetitions on the left and right sides were recorded. KES and knee-flexion strength (KFS) were measured using a handheld dynamometer (model 01163; Lafayette Instrument Company, Lafayette, IN, USA). The participants sat on a chair with their hips and knees bent at 90°. The assessor placed the dynamometer on the talotibial joint and lateral malleolus to assess the isometric voluntary contractions of the knee extensor and flexor muscle strengths (21). The participants were instructed to perform two maximum contractions for each muscle group on their right leg, which were held for 5 s each. The average of two trials was recorded.

2.3.3. Lower-limb muscle function

We assessed lower-limb muscle function using the five-times sit-to-stand test (5xSTS) and recorded using KMR. The 5xSTS measures the time taken to stand up and sit down five times from a sitting to a standing position as quickly as possible with arms folded across the chest (22).
2.3.4. *Questionnaires*

We used the Recommended Food Score (RFS) to evaluate the overall dietary quality of the participants, assessing the consumption of foods consistent with existing dietary guidelines (23). We applied the Korean Global Physical Activity Questionnaire (GPAQ) to assess physical activity levels (24), and activities of daily living (ADL) to assess the participants’ ability to perform essential daily tasks independently, such as eating, dressing, and walking (25).

2.4. *Data Storage during the Exercise Intervention*

Owing to the challenges in participant recruitment during the coronavirus disease 2019 (COVID-19) pandemic, we conducted a pilot study with 11 participants. Previous studies conducted gait analyses using machine learning with repeated measurements from <10 participants (26, 27). Similarly, in the present study, we collected data on all instances of physical activity performed by the participants for the machine learning analysis. During the 8-week exercise intervention period, the x, y, and z values of all movements of the 25 joints of the participants were collected for the entire exercise period between weeks 3 and 8. As we observed dynamic changes in KES compared to grip strength before and after the exercise intervention, the analysis included only the wide squat exercise, which was related to changes in lower-limb strength.

2.5. *Machine Learning*

We used the exercise movement data collected through KMR during the exercise intervention period to create a machine-learning model to classify the two groups. This model examined weekly changes in the PSG. The importance of the variables used to classify the two groups was extracted to assess individual changes in the squat variable.
2.6. Data Collection and Smoothing

The raw data of the 25 joints collected through KMR contained noise that significantly affected the accuracy of the learning process. To improve the learning accuracy, it was necessary to reduce noise in the raw data and accurately separate the data based on the patient’s movement cycle (11). In this study, a moving average filter was applied to smooth the raw data, and the z-coordinates of the patient’s head were used to track the stand-up and sit-down points, allowing the data to be segmented based on the movement cycles.

2.7. Feature Extraction

The 25 joints of the patients provided a multitude of data as they moved in different directions. However, meaningful results were challenging to obtain from raw, unprocessed three-dimensional (3D) coordinates. Feature extraction involved the conversion of meaningless raw data into meaningful data. In this study, joint angles were calculated based on the sagittal plane for flexion and the transverse plane for rotation, enabling the calculation of the angles between the joints. To quantify the movement of the trunk, knee, and ankle, flexion and rotation for each joint were defined as the range of motion (ROM), maximum, and minimum, resulting in 18 features.

2.8. Feature Selection

Feature selection is the process of selecting the most useful features for training a model. In this study, we applied recursive feature elimination with cross-validation using a support vector machine (RFECV-SVM) to identify the optimal set of features (28). We divided the entire dataset into a 7:3 ratio and used only the training dataset for the RFECV-SVM. Starting with all features, the feature with the lowest importance was removed at each step, and the RFECV-SVM was trained with fivefold cross-validation using accuracy scoring (28). This process was repeated until the optimal number of features yielding the best results was
obtained. In addition, to extract feature importance rankings, we used an optimal set of features to train random forest (RF), gradient boosting, extreme gradient boosting (XGBoost), and adaptive boosting (AdaBoost) models and determined the ranks using the importance value.

2.9. Model Training and Evaluation

Kinect-related research has used various machine-learning models. Among these, support vector machines (SVMs) have demonstrated high accuracy in analyzing skeletal muscle posture (29, 30). The k-nearest neighbors (KNN) algorithm has also yielded promising results in person-identification studies by combining static and dynamic gait features (31). RF is known for its low computational cost and high performance, which effectively mitigates data overfitting (32). Gradient boosting (GB) has demonstrated high accuracy in age classification studies using the Kinect (33). These findings indicate that machine learning models can be effectively utilized for model training using Kinect data.

In this study, we created a voting model combining SVM, KNN, RF, and GB to achieve more accurate classification. Voting models aggregate predictions made by multiple classifiers based on the predicted class probabilities, leading to improved classification accuracy (11).

The evaluation of machine-learning models is a critical step in measuring and enhancing their performance. We used accuracy, precision, recall, F1 score, and area under the receiver operating characteristic curve (AUROC) as metrics for model evaluation.

2.10. Statistical Analysis

We conducted statistical analyses using IBM SPSS Statistics for Windows, version 26.0 (IBM Corp., Armonk, NY, USA) to assess the exercise-related effects of the intervention. The Mann–Whitney U test was used to compare differences in pre-to post-measurements and exercise intervention data between the PSG and RG. The Wilcoxon test was used to compare differences within the PSG. The significance level for all statistics was set at $P<.05$. Descriptive statistics,
including the mean (M) and standard deviation (SD), were calculated for all variables. We plotted data obtained from the statistical analyses using GraphPad Prism version 9.4.1 (GraphPad Software, San Diego, CA, USA). Real-time 3D coordinate data of the 25 joints extracted from KMR, along with data preprocessing and the learning process for the machine learning models, were analyzed using Python (version 3.9.12; Python Software Foundation, Wilmington, DE, USA).

3. Results

3.1. Participants

Among 14 older adult women recruited to participate in this study, we allocated 8 and 6 to the PSG and RG, respectively. However, owing to COVID-19, two participants withdrew from the RG, and one participant from the PSG was excluded because of <80% adherence. Therefore, 11 older women completed the study. Their general characteristics are presented in (Table 1). Age (PSG: 80.00±6.48 years, n=7; RG: 72.75±7.93 years, n=4), weight (PSG: 51.67±4.56 kg; RG: 60.05±6.87 kg), and BMI (PSG: 18.53±2.89; RG: 21.05±5.76) did not differ significantly between the groups. However, appendicular skeletal-muscle composition differed significantly between the groups (PSG: 5.30±0.23 kg/m²; RG: 5.93±0.46 kg/m², P=.024).

Table 1. Body composition and muscle strength changes before and after the exercise interventions

<table>
<thead>
<tr>
<th>Variable</th>
<th>PSG (n=7)</th>
<th>RG (n=4)</th>
<th>Pre</th>
<th>Post</th>
<th>P-valuea</th>
<th>P-valueb</th>
<th>P-valuec</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age</td>
<td>80.00±6.48</td>
<td>72.75±7.93</td>
<td>-</td>
<td>.109</td>
<td>-</td>
<td>.073</td>
<td>.042</td>
</tr>
<tr>
<td>Body composition</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Weight (kg)</td>
<td>51.67±4.56</td>
<td>51.91±4.49</td>
<td>.866</td>
<td>60.05±6.87</td>
<td>.073</td>
<td>.024</td>
<td>.073</td>
</tr>
<tr>
<td>BMI (m²/kg)</td>
<td>18.53±2.89</td>
<td>17.86±3.14</td>
<td>.672</td>
<td>21.05±5.76</td>
<td>.412</td>
<td>.412</td>
<td>.412</td>
</tr>
<tr>
<td>ASM (kg)</td>
<td>5.30±0.23</td>
<td>5.33±0.25</td>
<td>.866</td>
<td>5.93±0.46</td>
<td>.024</td>
<td>.073</td>
<td>.073</td>
</tr>
<tr>
<td>Muscle strength</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Grip strength (kg)</td>
<td>20.13±2.48</td>
<td>20.70±3.35</td>
<td>.398</td>
<td>20.88±3.46</td>
<td>.927</td>
<td>.648</td>
<td>.648</td>
</tr>
<tr>
<td>KES (kg)</td>
<td>12.76±2.61</td>
<td>16.12±3.47</td>
<td>.018*</td>
<td>19.09±1.83</td>
<td>.006**</td>
<td>.315</td>
<td>.315</td>
</tr>
<tr>
<td>KFS (kg)</td>
<td>9.15±2.90</td>
<td>10.31±3.21</td>
<td>.237</td>
<td>13.73±3.28</td>
<td>.042*</td>
<td>.109</td>
<td>.109</td>
</tr>
<tr>
<td>Lower limb muscle function</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4-meter gait test (s)</td>
<td>3.98±0.73</td>
<td>3.74±0.39</td>
<td>.352</td>
<td>3.38±0.37</td>
<td>.230</td>
<td>.230</td>
<td>.230</td>
</tr>
<tr>
<td>5×STS</td>
<td>12.36±3.59</td>
<td>10.33±2.14</td>
<td>.128</td>
<td>10.35±2.22</td>
<td>.527</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>Questionnaire</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Values are presented as mean±standard deviation. PSG, possible sarcopenia group; RG, Reference group; BMI, body mass index; ASM, appendicular skeletal muscle; KES, knee-extension strength; KFS, knee flexion strength; 5xSTS, five times sit-to-stand; GPAQ, global physical activity questionnaire; RFS, recommended food score; ADL, activities of daily living; aP-value for the comparison within PSG pre-test and post-test; bP-value for the comparison between PSG pre-test and RG; cP-value for the comparison between PSG post-test and RG; *P<.05. **P<.01

The KES, KFS, and GPAQ scores differed significantly between the PSG and RG at baseline. Regarding muscle strength, the grip strength did not differ significantly between the groups after the exercise intervention. Following the 8-week exercise intervention, the PSG showed a significant improvement in KES over time, as well as a significant difference compared with the RG. The KFS and GPAQ scores differed significantly between the groups at baseline but did not differ significantly in the PSG over time. Regarding physical function, the 5xSTS, RFS, and ADL scores did not differ significantly between the groups at baseline. Additionally, although the time taken to complete the 5xSTS did not differ significantly, the motion data collected using the Kinect revealed significant differences between the two groups at baseline (Table 2).

Table 2. Movement differences between the two groups in the five-times sit-to-stand test

<table>
<thead>
<tr>
<th>Variable</th>
<th>PSG Pre (Data=35)</th>
<th>Post (Data=35)</th>
<th>P-valuea</th>
<th>RG Pre (Data=20)</th>
<th>P-valueb</th>
<th>P-valuec</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Rotation</strong></td>
<td>Arm_ROM (°)</td>
<td>0.17±0.10</td>
<td>0.11±0.04</td>
<td>.003**</td>
<td>0.10±0.04</td>
<td>0.002**</td>
</tr>
<tr>
<td>Forearm_ROM (°)</td>
<td>0.23±0.12</td>
<td>0.17±0.09</td>
<td>.053</td>
<td>0.19±0.11</td>
<td>.162</td>
<td>.588</td>
</tr>
<tr>
<td>KES (°)</td>
<td>0.07±0.03</td>
<td>0.04±0.02</td>
<td>&lt;.001**</td>
<td>0.06±0.03</td>
<td>.875</td>
<td>&lt;.001**</td>
</tr>
<tr>
<td>Ankle_ROM (°)</td>
<td>0.14±0.11</td>
<td>0.07±0.03</td>
<td>.006**</td>
<td>0.09±0.04</td>
<td>.793</td>
<td>.100</td>
</tr>
<tr>
<td>Trunk_ROM (°)</td>
<td>0.04±0.04</td>
<td>0.01±0.012</td>
<td>&lt;.001**</td>
<td>0.02±0.01</td>
<td>.066</td>
<td>.037*</td>
</tr>
<tr>
<td><strong>Flexion</strong></td>
<td>Arm_ROM (°)</td>
<td>27.69±10.30</td>
<td>33.14±9.96</td>
<td>.008**</td>
<td>28.74±12.59</td>
<td>.813</td>
</tr>
<tr>
<td>Forearm_ROM (°)</td>
<td>29.33±21.33</td>
<td>36.09±22.59</td>
<td>.318</td>
<td>33.30±22.18</td>
<td>.540</td>
<td>.588</td>
</tr>
<tr>
<td>Knee_ROM (°)</td>
<td>65.31±23.35</td>
<td>66.31±23.16</td>
<td>.857</td>
<td>66.22±20.29</td>
<td>.060</td>
<td>.452</td>
</tr>
</tbody>
</table>
Values are presented as mean±standard deviation. PSG, possible sarcopenia group; RG, reference group; ROM, range of motion; aP-value for the comparison within PSG pre-test and post-test; bP-value for the comparison between PSG pre-test and the RG; cP-value for the comparison between PSG post-test and the RG; *P<.05. **P<.01

After the 8-week exercise intervention, the ROM of arm rotation in the PSG showed significant improvement, leading to non-significant differences compared to that in the RG.

3.2. Optimal set of features

The raw data on wide squats collected through the KMR were smoothed and divided into cycles, generating 18 features. To reduce the model complexity, we conducted procedures to prevent overfitting and improve the generalization performance using a feature-selection process. RFECV-SVM was used on the training dataset to determine the appropriate number of features for model training (Supplementary Figure S1). The model achieved the best performance in classifying the two groups when nine features were used.

The results of the SVM, KNN, RF, GB, and voting models in the training and test datasets are presented in Table 3.

Table 3. Evaluating the sarcopenia classification model based on movement patterns during wide squats

<table>
<thead>
<tr>
<th>Model</th>
<th>Accuracy (%)</th>
<th>Recall (%)</th>
<th>Precision (%)</th>
<th>F1 score (%)</th>
<th>AUROC (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Train dataset</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SVM</td>
<td>92.41±2.61</td>
<td>95.87±1.76</td>
<td>92.72±2.38</td>
<td>94.26±1.93</td>
<td>95.38±4.43</td>
</tr>
<tr>
<td>KNN</td>
<td>91.96±2.89</td>
<td>93.47±2.28</td>
<td>94.14±2.55</td>
<td>93.80±2.20</td>
<td>95.82±3.76</td>
</tr>
<tr>
<td>RF</td>
<td>90.17±2.71</td>
<td>96.21±1.70</td>
<td>91.16±1.99</td>
<td>92.27±1.48</td>
<td>96.45±1.35</td>
</tr>
<tr>
<td>GB</td>
<td>89.72±3.15</td>
<td>93.47±2.28</td>
<td>90.84±3.95</td>
<td>92.23±2.29</td>
<td>95.42±1.62</td>
</tr>
<tr>
<td>Voting</td>
<td>92.85±3.07</td>
<td>95.87±3.20</td>
<td>92.75±2.16</td>
<td>94.25±2.17</td>
<td>97.00±3.07</td>
</tr>
</tbody>
</table>
Values are presented as mean ± standard deviation. SVM, support vector machine; KNN, K-nearest neighbor; RF, random forest; GB, gradient boosting. The training data set was evaluated using fivefold cross-validation.

In the training dataset, the voting classifier achieved an accuracy of 92.85% and AUROC of 97%, showing an overall superior performance compared with the other models. Similarly, in the test dataset, the voting classifier demonstrated an accuracy of 92.75% and AUROC of 97.27%.

The results of the investigation into the effect of the exercise intervention in the PSG on the probability of classifying possible sarcopenia using our model, which classifies possible sarcopenia based on a single wide squat movement, are shown in Figure 2.

**Figure 2.** Temporal weekly variations of the probability of sarcopenia diagnosis
sarcopenia by evaluating weekly data in the possible sarcopenia group. W: weeks of the exercise intervention in the possible sarcopenia group, RG: reference group. *P < .01: significant difference between W3 and other weeks, as well as the RG. †P < .01: significant difference between the RG and other weeks.

During the transition from weeks 3 to 4, the probability of being classified as having sarcopenia decreased significantly in the PSG. However, even at week 8, the PSG still differed significantly from the RG. Evaluation of all components of the squat movement revealed that the PSG continued to differ significantly from the RG.

The variable importance values for each feature in the RF, XGBoost, AdaBoost, and GB models, as well as the average variable importance values across the four models, are depicted in Figure 3.

**Figure 3.** Ranked feature importance of squat variables from machine learning models

Each classifier calculates variable importance using the importance value, which is displayed on the x-axis. The variable importance rankings are shown on the y-axis. XGBoost: extreme gradient boosting; AdaBoost: adaptive boosting; ROM: range of motion.

3.3. Changes in the values of variables associated with specific movements in wide squats

The changes in the individual features of wide squats ranked in order are presented in Table 4.
Table 4. Group differences in individual variables when performing wide squats.

<table>
<thead>
<tr>
<th>Rank</th>
<th>Variable</th>
<th>PSG Week 3 (Data=416)</th>
<th>PSG Week 8 (Data=487)</th>
<th>PSG P-value</th>
<th>RG Week 3 (Data=225)</th>
<th>RG P-valuea</th>
<th>RG P-valueb</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Knee flexion_ROM (°)</td>
<td>53.50±12.56</td>
<td>54.01±11.71</td>
<td>.408</td>
<td>59.21±6.71</td>
<td>&lt;.001**</td>
<td>&lt;.001**</td>
</tr>
<tr>
<td>2</td>
<td>Knee rotation_ROM (°)</td>
<td>0.14±0.10</td>
<td>0.14±0.06</td>
<td>&lt;.001**</td>
<td>0.14±0.05</td>
<td>&lt;.001**</td>
<td>.521</td>
</tr>
<tr>
<td>3</td>
<td>Ankle flexion_ROM (°)</td>
<td>16.63±4.74</td>
<td>16.95±5.11</td>
<td>.415</td>
<td>15.51±3.46</td>
<td>.001**</td>
<td>&lt;.001**</td>
</tr>
<tr>
<td>4</td>
<td>Trunk flexion_ROM (°)</td>
<td>18.06±6.41</td>
<td>19.40±5.85</td>
<td>&lt;.001**</td>
<td>22.80±6.58</td>
<td>&lt;.001**</td>
<td>&lt;.001**</td>
</tr>
<tr>
<td>5</td>
<td>Knee rotation_Max (°)</td>
<td>0.10±0.08</td>
<td>0.11±0.05</td>
<td>&lt;.001**</td>
<td>0.11±0.04</td>
<td>&lt;.001**</td>
<td>.714</td>
</tr>
<tr>
<td>6</td>
<td>Trunk flexion_Min (°)</td>
<td>-16.69±6.84</td>
<td>-18.18±6.15</td>
<td>.004**</td>
<td>-21.44±6.70</td>
<td>&lt;.001**</td>
<td>&lt;.001**</td>
</tr>
<tr>
<td>7</td>
<td>Ankle flexion_Max (°)</td>
<td>3.11±3.21</td>
<td>2.64±2.78</td>
<td>.247</td>
<td>1.48±2.04</td>
<td>&lt;.001**</td>
<td>&lt;.001**</td>
</tr>
<tr>
<td>8</td>
<td>Knee flexion_Min (°)</td>
<td>-1.68±3.67</td>
<td>-1.91±3.52</td>
<td>.043*</td>
<td>-2.29±5.55</td>
<td>.113</td>
<td>.942</td>
</tr>
<tr>
<td>9</td>
<td>Trunk flexion_Max (°)</td>
<td>1.37±1.50</td>
<td>1.23±1.27</td>
<td>.643</td>
<td>1.37±1.67</td>
<td>.711</td>
<td>.995</td>
</tr>
</tbody>
</table>

Values are presented as mean ± standard deviation. PSG, possible sarcopenia group; RG, reference group; ROM, range of motion; Max, maximum; Min, minimum; *P-value for the comparison between PSG week 3 and 8; †P-value for the comparison between PSG week 3 and RG; ‡P-value for the comparison between PSG week 8 and RG; *P<.05, **P<.01

The second- and fifth-ranked variables (knee rotation ROM and knee rotation maximum, respectively) differed significantly between the PSG and RG at 3 weeks. However, after the exercise intervention, the PSG values were similar to those of the RG and the difference was not significant, while the differences in other factors between the PSG and RG remained significant.

4. Discussion

This study analyzed data from pre- and post-test assessments using the 5xSTS. The results demonstrated that while the PSG and RG showed different movement patterns owing to differences in muscle strength, improving muscle strength in the PSG resulted in close approximations of the movement patterns of the RG. During the 5xSTS, the PSG displayed more...
dynamic arm movements compared to the RG. The 5xSTS primarily relies on lower limb strength, and individuals typically compensate for weak lower limb strength with increased dynamic upper-body movements, such as enhanced trunk flexion during the forward-acceleration phase (10, 34, 35). However, following the exercise intervention in our study, the dynamic arm movements of the PSG decreased and did not differ significantly from those in the RG. These observations suggest that the significant improvement in lower limb strength in the PSG group reduced the need for compensatory arm movements, indicating that although the PSG and RG differed in compensatory actions owing to differences in lower limb strength, improvements in muscle strength stabilized the movement patterns of the PSG.

We trained our model on squat movements related to lower limb strength, to effectively identify significant differences during the exercise intervention period. This model successfully differentiated between the PSG and the RG, with an accuracy and AUROC of 92.75% and 97.27%, respectively. These results surpass the average accuracy of 80–85% reported in previous studies (11) that employed KNN, SVM, multilayer perceptron, bagging classifiers, and voting classifiers to categorize the movements of older adults, as assessed by gait, 30-s sit-to-stand, 2-min step, and 30-s arm curl movements. Consequently, the high accuracy of the machine learning model in the present study underscores its capability to successfully distinguish between older adults with and without possible sarcopenia based on their movement patterns.

Additionally, our model revealed a decreasing likelihood of diagnosing possible sarcopenia in the PSG from the 4th week of the exercise intervention. However, despite significant improvements in muscle strength after the 8-week exercise program in the PSG, persistent differences in exercise patterns between the PSG and RG were observed based on the classification model. This persistence may be attributed to the multi-joint movements involved in squat exercises, including hip, knee, and ankle flexion and extension (36, 37). We observed no significant improvements in
muscle groups other than the KES, which may have led to the continued differences in the overall movement patterns between the PSG and RG.

Therefore, rather than comprehensively comparing the entire squat movement using our classification model, we focused on distinguishing the movements of the individual joint components to identify their importance and observe the changes in individual functions. The initial examination of the individual elements of squat movement revealed significant differences in knee rotation, knee flexion, trunk flexion, and ankle flexion. These findings align with those of previous studies based on key indicators such as knee and hip joint-angle displacements, lateral sway, and trunk flexion (10), and are consistent with prior research suggesting that weak individuals may exhibit dynamic movements of the trunk, knee, and ankle during squats (38, 39). However, in our study, we observed significantly decreased post-intervention values only for knee rotation ROM and maximum values in the PSG, approaching values similar to those in the RG. During squatting, internal and external rotations of the femur accompany knee flexion and extension (40), often resulting in dynamic knee valgus, typically due to weak quadriceps, gluteal, and hamstring muscles (41, 42). The reduced knee rotation in the PSG in the present study can be attributed to KES improvements. Therefore, the significant difference in knee rotation between the PSG and RG can be explained by the initial lower quadriceps strength in the PSG and the stabilization of knee rotation movements with improving KES.

Our study, to the best of our knowledge, is the first to implement exercise interventions for individuals with possible sarcopenia and apply machine-learning models to observe changes in movement patterns. Our findings demonstrated distinct differences in movement patterns between healthy older adult women and those with possible sarcopenia, predominantly owing to muscle weakness associated with possible sarcopenia. Furthermore, the exercise intervention led to significant improvements in muscle strength, which, in turn, influenced changes in movement
patterns. Additionally, the classification model developed from the Kinect exercise data successfully differentiated between older adults with possible sarcopenia and healthy individuals based on their movement patterns, which enabled the tracking of changes in movement patterns. Therefore, this study provides foundational evidence for the development of an exercise management system to remotely monitor and track the condition of older adults.

This study had some limitations. First, this study served as a pilot study for the main experiment. Owing to the challenges in participant recruitment during the COVID-19 pandemic, the study proceeded with a minimal number of participants. Consequently, we predominantly recruited relatively easily accessible older adult women, which led to a limited sample size and potential sex bias. Second, there may be data quality issues owing to inherent noise in the Kinect data. To address this issue, we refined the raw data and discarded data with unclear cycle identification during the segmentation process. Moreover, the persistent significant differences in squat movements between the PSG and RG, even after the intervention, may be attributed to the relatively short intervention period to improve the strength of the various muscle groups. Therefore, future studies should include a larger number of participants with long-term interventions to investigate whether strength improvements lead to changes in movement patterns.

Author contributions

Acknowledgments
The authors thank all the participants for their participation in this study.

**Funding**

This work was supported by MyBenefit Co. (Seoul, S. Korea).

**Disclosure statement**

No potential conflicts of interest were reported by the author(s).