

Exploring Ski Proficiency Estimation from Foot Pressure Patterns in a Simulator-Based Environment

Shigeharu Ono

s2120010@jaist.ac.jp

Japan Advanced Institute of Science and Technology

Ryosuke Atsumi

Sony Payment Services Inc.

Hideaki Kanai

Japan Advanced Institute of Science and Technology

Hideki Koike

Institute of Science Tokyo

Article

Keywords: Ski simulator, deep learning, multi-class classification, image classification, feature visualization

Posted Date: July 22nd, 2026

DOI: <https://doi.org/10.21203/rs.3.rs-10312009/v1>

License:   This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

Additional Declarations: No competing interests reported.

Version of Record: A version of this preprint was published at Scientific Reports on October 16th, 2026. See the published version at <https://doi.org/10.1038/s41598-026-76341-y>.

Exploring Ski Proficiency Estimation from Foot Pressure Patterns in a Simulator-Based Environment

Shigeharu Ono^{1*}, Ryosuke Atsumi², Hideaki Kanai¹,
Hideki Koike³

^{1*}Japan Advanced Institute of Science and Technology, Asahidai, Nomi,
923-1292, Ishikawa, Japan.

²Sony Payment Services Inc., Takanawa, Minato, 108-0074, Tokyo,
Japan.

³Institute of Science Tokyo, Ōokayama, Meguro, 152-8550, Tokyo, Japan.

*Corresponding author(s). E-mail(s): s2120010@jaist.ac.jp;

Abstract

This study investigated whether deep learning can classify skiing proficiency from plantar pressure distributions measured during simulator-based training and whether the resulting predictions reflect longitudinal changes in skill acquisition. Weight-shift images obtained from 27 participants were classified into beginner, novice, and intermediate proficiency levels using a YOLO11-based image classifier under a leave-one-participant-out evaluation. Longitudinal changes in classification outcomes were then analyzed using linear and spline-based mixed-effects models. To interpret the learned features, Eigen-CAM visualizations and quantitative analyses of attention regions were performed. The proposed classifier demonstrated reliable discrimination between beginner and novice pressure patterns but showed limited generalization to the minority intermediate class. Longitudinal analyses suggested that the probability of non-beginner classifications generally increased during practice, while the spline-based mixed-effects model indicated a nonlinear learning trajectory characterized by rapid early improvement followed by a plateau. Furthermore, repeated-measures correlation demonstrated a significant association between forefoot-focused attention and the probability of non-beginner classification, suggesting that the model captured biomechanically meaningful movement characteristics. These findings indicate that deep learning can provide an objective measure of skill acquisition in simulator-based ski training while offering interpretable visual explanations of the learned movement features.

Keywords: Ski simulator, deep learning, multi-class classification, image classification, feature visualization.

1 Introduction

Virtual reality (VR)-based sports training systems provide repeatable training conditions and objective performance measurements[1, 2]. Also in skiing, several ski simulators (e.g., PRO SKI SIMULATOR, SkyTechSport, and Ski Machine) and VR-based ski training systems[3, 4] have been developed to support off-season training and skill acquisition.

However, learning to ski remains challenging, especially for beginners who require continuous feedback on weight shifting and balance. Automatic proficiency estimation enables personalized feedback without constant instructor supervision. Deep learning has recently been applied to sports skill indicators to identify movement characteristics associated with different proficiency levels [5, 6].

However, existing ski simulators have limitations, including fixed foot positions, limited body tilt, and simplified ground conditions. These constraints may reduce the usefulness of simulator data for extracting proficiency-related features[7, 8]. Consequently, it remains unclear whether biomechanical information measured in simulator environments contains sufficient information for reliable proficiency estimation.

Among sensing modalities for skiing, plantar pressure is particularly suitable because weight shifting and balance control are fundamental skiing skills. Previous studies have used foot-pressure sensors for performance monitoring and real-time feedback in both real and simulator-based environments [4, 9, 10]. Deep learning has also been applied to plantar pressure maps for foot-type classification, pressure estimation, gait recognition, and biomechanical analysis [11–14]. However, these studies primarily addressed walking or clinical applications, leaving the use of simulator-derived pressure maps for estimating skiing proficiency unexplored.

Accordingly, this study investigates the estimation of skiing proficiency directly from simulator-derived plantar pressure maps. Unlike previous approaches based on posture estimation, inertial sensors, or RGB images collected in real skiing environments [15–17], we use only plantar pressure distributions acquired from a low-cost ski simulator. We further employ Eigen-CAM [18] to visualize the learned spatial features and examine how they change during repeated simulator training. This study aimed to answer the following research questions:

- Can foot-pressure patterns be used to estimate skiing proficiency levels?
- Can the proposed classifier capture longitudinal changes in weight-shifting patterns during repeated simulator training?
- What pressure-distribution characteristics are associated with different proficiency levels?

First, we created a YOLO11[19]-based classifier based on weight-shifting data from 27 participants using a ski simulator (Pro Ski-Simulator: POWER SKI SIMULATOR).

Second, we created feature heatmaps and compared them among three proficiency levels using Eigen-CAM[18]. Third, we conducted a training experiment using VizSki[3] with 6 non-experienced participants.

2 Related Research

2.1 Ski Skill Indicators and Motion Analysis

In actual skiing, IMUs are widely used to measure posture. They have been applied to estimate 3D orientation in ski jumping[20–22], landing momentum[23], and ski-turning skills such as edge angle, symmetry, and body balance[15, 16]. GNSSs have also been used to measure skier trajectories in cross-country skiing[24], ski jumping[25], and alpine skiing[26]. Together, IMUs and GNSSs provide information such as turn depth, turn length, and posture.

However, in simulator-based environments, the data from these sensors is constrained because the user’s feet are fixed in a parallel orientation to the front, tilt is not considered, and the environment is indoor. Moreover, the ski simulator used in this study restricts the user’s forward movement. Given these limitations, this study focused on foot pressure data from pressure sensors.

2.2 Foot Pressure Measurement for Ski Learning

Although IMUs are a commonly used solution for measurement in skiing, some studies have measured foot pressure data in both actual and virtual ski environments. Michahelles and Schiele[9] implemented a skiing monitoring system that incorporated several sensing devices, including pressure sensors. Hasegawa et al.[10] developed a system that provides real-time auditory sonification of a user’s center of gravity. In virtual ski training, Ono et al.[4] provided visual feedback on foot pressure.

However, existing studies primarily use pressure data for monitoring or feedback and do not investigate whether pressure patterns can be used directly to estimate skier proficiency. At least in actual ski situations, Wojtyczek et al.[27], and Simsek et al.[28] indicated that balance performance was related to improved skiing performance. Therefore, we hypothesized that weight-shifting skills could indicate a user’s experience level, even when using a ski simulator.

2.3 Deep Learning for Sports Skill Classification

Several studies have applied deep learning to evaluate sports skills rather than merely recognize activities. Tabrizi et al.[5] employed deep learning models to classify table-tennis forehand strokes according to proficiency level, demonstrating that movement patterns contain discriminative information related to skill acquisition. Similarly, AI Golf [6] analyzed golf swings and provided self-training feedback based on extracted movement features. These studies suggest that deep learning can support skill indicators by identifying subtle differences in movement characteristics among athletes of different proficiency levels. Ronao and Cho[29] utilized a Deep Convolutional

139 Neural Network for human activity recognition using sensor data from a smart-
140 phone gyroscope. Similarly, recognition systems have been implemented for beach
141 volleyball[30].

142 Although these examples focused on sensor-based recognition, image-based recog-
143 nition systems have also been developed[17, 31, 32]. Yoon et al. [31] introduced an
144 object-tracking system for basketball that identifies each player and the ball holder
145 in video frames and constructs a passing network from the tracking data. Lv et
146 al.[32] developed a key-frame extraction system for weightlifting that uses a fully
147 convolutional network (FCN) and a CNN to estimate and classify a user’s posture.
148 Furthermore, Ludwig et al. [17] designed a performance estimation system for ski
149 jumping that employed a CNN for posture estimation and principal component anal-
150 ysis to estimate jumping distances from these estimates. These systems classify data
151 by tracking user bodies and posture in real environments.

152

153 2.4 Pressure-Map-Based Learning Approaches

154

155 Recent studies have demonstrated that plantar pressure maps can be effectively
156 analyzed using deep learning models for gait analysis, foot-type classification, and
157 plantar pressure estimation[11–14, 33]. Pressure maps preserve the spatial distribu-
158 tion of plantar loading, allowing convolutional neural networks to automatically learn
159 discriminative features from pressure images.

160 Singh et al.[33] transformed sensor data into image representations for deep
161 learning, demonstrating that spatial sensor patterns can be effectively learned using
162 image-based neural networks. Similarly, Mun et al.[12] estimated plantar pressure dis-
163 tributions using deep learning, while Zhao et al.[11] showed that plantar pressure
164 images can be directly classified using convolutional neural networks for foot-type
165 recognition.

166 More recently, pressure-map learning has been extended to gait recognition and
167 biomechanical indicators. Wu et al.[13] proposed a CNN-based framework for gait
168 recognition using plantar pressure maps. Moreover, Dindorf et al.[14] summarized
169 the rapid growth of deep learning in plantar pressure analysis for rehabilitation, gait
170 indicators, and clinical diagnosis. These studies suggest that pressure maps contain
171 rich spatial information that can be effectively exploited by modern deep learning
172 methods.

173 However, existing pressure-map studies have primarily focused on walking, reha-
174 bilitation, or medical diagnosis [12–14]. In contrast, skiing is substantially different
175 from the movement patterns observed during walking. Consequently, it remains
176 unclear whether the success of pressure-map learning in gait analysis can be directly
177 transferred to simulator-based ski training.

178

179 2.5 Research Gap and Positioning of This Study

180

181 Previous studies have demonstrated the effectiveness of wearable sensors for evaluating
182 skiing performance [9, 10, 15–17] and deep learning for sports skill indicators [5, 6].
183 However, most have focused on real-world skiing or cross-sectional skill classification

184

[5, 6, 15–17], providing limited evidence that AI-based evaluations can objectively measure skill acquisition during simulator-based training.

To address these gaps, this study investigates whether plantar-pressure images acquired using a ski simulator can serve as an objective indicator of skiing proficiency and learning progress. We develop a multi-class deep learning classifier, evaluate longitudinal changes in classification outcomes using mixed-effects models, and quantitatively analyze Eigen-CAM attention patterns to examine whether the learned features correspond to meaningful changes in weight-shifting behavior. By integrating classification accuracy, longitudinal validation, and model interpretability, this study provides a comprehensive framework for objectively evaluating skill acquisition in simulator-based ski training.

3 System Components

3.1 Data Collection Environment

Fig. 1(a) illustrates the implemented system. Foot pressure data were collected in a simulator-based ski training system consisting of VizSki[3], VIVE Pro, and Pro Ski-Simulator: POWER SKI SIMULATOR as its foundational environments. The POWER SKI SIMULATOR offers cost-effective and more space-efficient advantages than do other simulators. However, as mentioned above, the simulator has limitations, such as the user’s feet being fixed in a parallel position, a lack of consideration for tilt, and the ground surface’s status.

3.2 Pressure Sensors

Since weight shifting is a fundamental skill in skiing, and previous studies have reported improvements in balance through ski training, foot-pressure data were selected as the primary modality for estimating proficiency. During sessions with VizSki, the system captured the weight-shifting data of the user using pressure sensor modules (Fig. 1(b)).

This study used pressure sensor modules with 32 pressure sensors (16 per foot). Although the sensor positions are almost based on those of the foot part in Fig. 2, position misalignment could occur in each participant because each sensor was independent and needed to be replaced. The pressure sensor module includes a pressure sensor (FSR402, 13 mm circle $\times$ 56 mm, force sensitivity range of 0.2–20 N), an Arduino Mega 2560, and an XBee WiFi (S6B) module (DIGI-XB2B-WFWT-001) to establish a wireless connection. Obtained weight shift data from the pressure sensor modules were exported as comma-separated values (CSV)-formatted files as 10-bit analog input data. The plantar pressure dataset used in this study has been made publicly available through IEEE DataPort [34].

3.3 Input Images

The obtained pressure data were processed using Savitzky-Golay smoothing and Z-score standardization before creating heatmaps. After that, the data were normalized to 8-bit values using min-max normalization to reduce differences in features due to

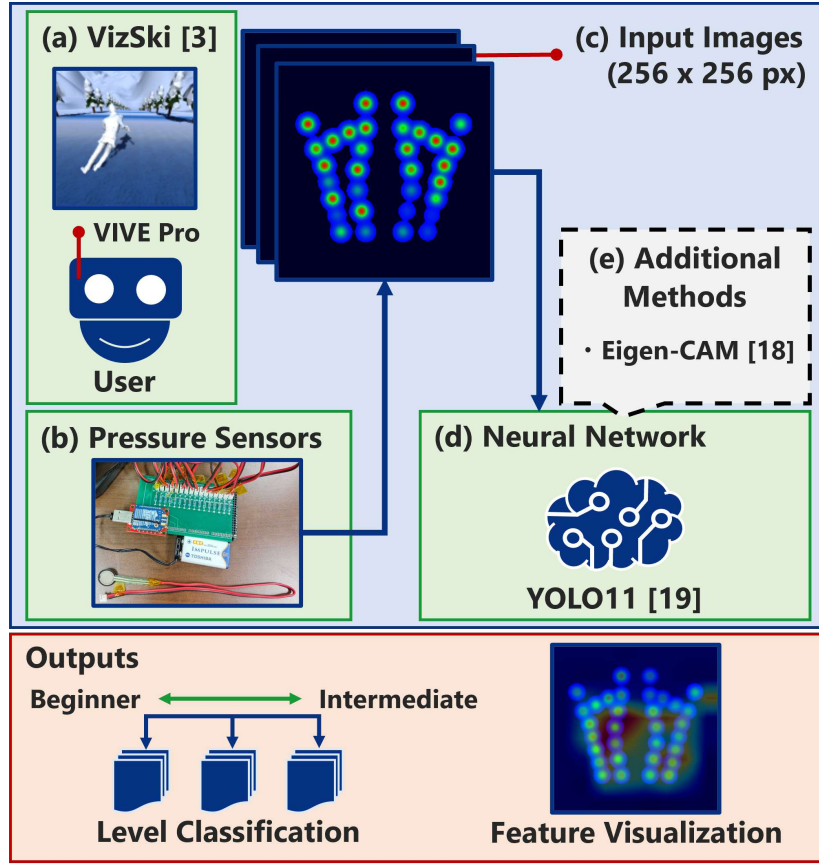


Fig. 1 System components

participant weight. After normalization, the data were plotted in the positions shown in Fig. 2 to create 256×256 size portable network graphics (PNG) heatmap images. Each heatmap image was used as input to the feature-extraction network (Fig. 1(c)).

When a user plays VizSki, they must wear a head-mounted display and hold the simulator's handrail for safety. Because of this limitation, the user's posture while playing VizSki may differ from that during actual skiing. This limitation may also affect the user's weight-shifting patterns. The features obtained may not always align with the weight-shifting patterns observed in actual skiing situations. However, some studies[27, 28] indicate that ski training improves balance performance. Therefore, we believe that classifying and visualizing weight-shifting features by proficiency level is helpful for learning and teaching skiing.

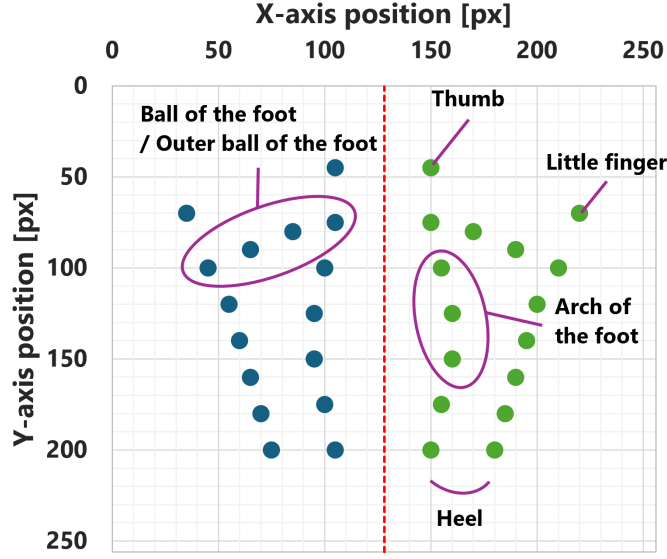


Fig. 2 Positions of each pressure sensor in the input image (256×256 px)

3.4 Neural Network

This study used the YOLO11 classification model (yolo11n-cls) as the feature-extraction network (Fig. 1(d)). YOLO11 adopts spatial pyramid pooling-fast (SPPF) and cross-stage partial with spatial attention (C2PSA) blocks in its backbone network[19]. We selected YOLO11 because it provides a lightweight convolutional architecture with strong image-classification performance while preserving spatial feature representations. Since the objective of this study was not only to classify foot-pressure heatmaps but also to interpret the learned spatial features using Eigen-CAM[18], a convolutional neural network was preferable, as the spatial correspondence between activation maps and input images is more straightforward than with transformer-based architectures. Furthermore, the lightweight design of YOLO11 makes it suitable for future real-time deployment in simulator-based training systems. The model was trained to classify foot-pressure heatmaps into three proficiency levels (Beginner, Novice, Intermediate).

3.5 Explainability Analysis

Although if the classifier achieved high classification performance, classification accuracy alone does not reveal which regions of the foot-pressure image contributed to the decision. To verify whether the model relied on meaningful weight-shifting characteristics rather than participant-specific artifacts, an explainability analysis was conducted using Eigen-CAM (Fig. 1(e)). Eigen-CAM visualizes the spatial regions that contribute most strongly to the network’s classification decision [18]. Unlike Grad-CAM[35],

Eigen-CAM does not rely on gradient information and instead extracts the principal component of the activation tensor using singular value decomposition (SVD). The resulting activation map highlights image regions that strongly influence the classifier output.

To improve robustness, Eigen-CAM maps were generated for all validation images and averaged within each participant-session pair. The averaged activation maps were then overlaid on the corresponding mean pressure images to visualize the regions consistently attended to by the classifier. Changes in the distribution of attention across multiple training sessions may reveal how participants modified their weight-shifting strategies. Therefore, Eigen-CAM can be used not only to explain classification results but also to visualize potential learning-related changes.

4 Classification Test

4.1 Procedure

We collected weight-shifting data from 27 participants (20 males and 7 females). Table 1 details the attributes of the participants and their corresponding class assignments based on these experiences. In this phase of the experiment, participants were classified into three experience levels based on the number of days they had skied: 0 d (beginner), 1–5 d (novice), and 6 d or more (intermediate). The intermediate group contained only four participants, resulting in class imbalance. This imbalance was mainly due to recruiting snowplow skiers specifically to minimize the impact of variations in turning techniques, not the parallel skiers.

Table 1 Participants of The Classification Test

Class	Skiing Experience	Participants (Male/Female)	Mean Experience
Beginner	None	12 (8/4)	0
Novice	1–5 days	11 (9/2)	2.5 days
Intermediate	above 6 days	4 (3/1)	19 days

Participants completed an informed consent form before participating in the experiment. In the experiment, the participants engaged in measurement sessions after the practice session. In the measurement session, participants played the VizSki on a ski simulator for 1 min. For safety reasons, they were required to hold onto the handrails attached to the simulator. Consequently, the participants were advised to move their bodies without exerting force to maintain a natural weight shift. Participants were required to complete the measurement session three times. Among them, the second and third measurements were used for classification. To evaluate whether weight-shifting patterns can be classified within participants, images from the first 40 s of each session were used for training, and those from the final 20 s for validation. Experiments in this section were conducted as part of a study approved by the institutional ethical committee (JAIST Life Science Committee, ID: Human 04-056). This procedure was in accordance with the committee’s guidelines and regulations.

Because training and validation data were obtained from the same participants and sessions, the evaluation primarily measures participant-non-independent classification performance rather than generalization to unseen participants. However, to evaluate classifier performance, this study created both participant-dependent and participant-independent models. In the participant-independent model, the Leave-One-Participant-Out (LOPO) method was used. In contrast to the participant-dependent model, images from all 60 s of each session were used for training (26 participants) and validation (one participant) per fold.

4.2 Result

We acquired 59,292 input images from 54 sessions with 27 participants.

In the participant-dependent model, we used 10 classifiers (seed 0–9) and averaged their classification accuracies to assess robustness. 39,528 (beginner: 17,568; novice: 16,104; intermediate: 5,856) images captured during the initial 40 s of the experiment were designated as training data, and the remaining 19,764 (beginner: 8,784; novice: 8,052; intermediate: 2,928) images from the latter 20 s served as validation data.

In the participant-independent model, we used 27 classifiers based on the Leave-One-Participant-Out (LOPO) method and calculated the average classification accuracy. 57,096 images from 26 participants were designated as training data, and the remaining 2,196 images from one participant were used as validation data for each model.

Both models used YOLO11 (yolo11n-cls). The AdamW optimization algorithm was used to train these classifiers with a learning rate of $1e-4$. The batch size was set to 128 to ensure the minority class (intermediate) images were included.

4.2.1 Level Classification

Fig. 3 shows the confusion matrices for proficiency-level classification in both models. The label counts are the sums of the classification results across all folds. Table 2 summarizes the classification performance obtained using both participant-dependent and participant-independent (LOPO) evaluations. The participant-independent (LOPO) model showed moderate performance for the Beginner and Novice classes ($F1 = 0.681$ and 0.674 , respectively), performance for the Intermediate class remained poor ($F1 = 0.028$), indicating limited generalization to unseen participants, particularly for minority classes. The poor LOPO performance for the intermediate class was likely due to severe class imbalance (only 4 participants) and inter-participant variability. The participant-dependent evaluation demonstrated that the proposed architecture can reliably distinguish pressure patterns within individuals.

4.2.2 Attention Analysis of Foot Pressure Patterns

Figure 4 presents the class-wise mean Eigen-CAM visualizations. The intermediate group exhibited a slightly broader attention region extending toward the forefoot compared with the beginner group, although the overall activation patterns remained largely similar across proficiency levels. However, the differences between novice and intermediate participants and between beginner and novice participants were less

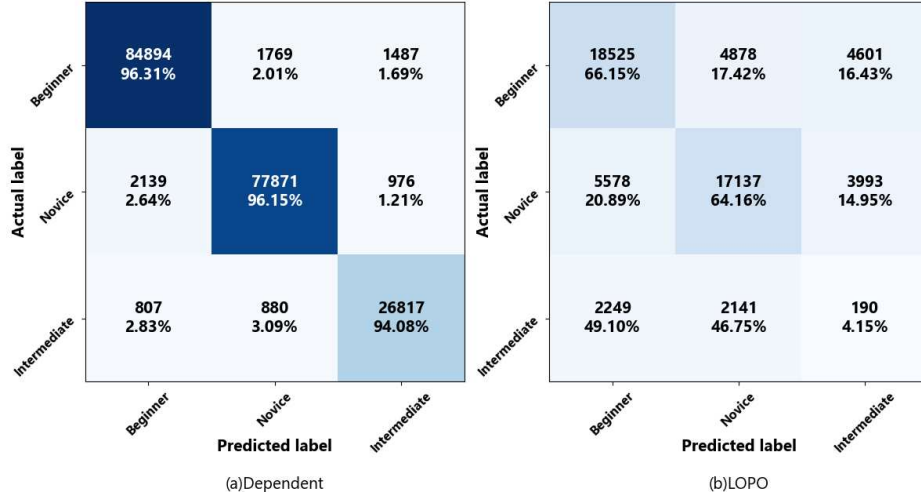


Fig. 3 Confusion matrix for proficiency level classification

Table 2 Evaluation metrics for each class

Class	Dependent (Accuracy: 0.959)			LOPO (Accuracy: 0.604)		
	Precision	Recall	F1-Score	Precision	Recall	F1-Score
Beginner	0.963	0.966	0.965	0.661	0.703	0.681
Novice	0.962	0.967	0.964	0.642	0.709	0.674
Intermediate	0.941	0.916	0.928	0.041	0.022	0.028

apparent. This observation is consistent with the lower separability of these two classes, particularly in the participant-independent evaluation. Therefore, while the feature visualization suggests that weight distribution is relevant to proficiency estimation, additional data from more experienced skiers are required to identify robust proficiency-specific features.

5 Exploratory Analysis of Training Progress

The classification experiment showed that the proposed classifier could distinguish weight-shift patterns associated with the predefined proficiency groups. Based on this result, we investigated whether changes in classifier outputs correspond to changes in participants' weight-shift patterns during repeated simulator training. Rather than evaluating skiing proficiency directly, this analysis examines how the learned feature representations respond to longitudinal changes in foot-pressure patterns.

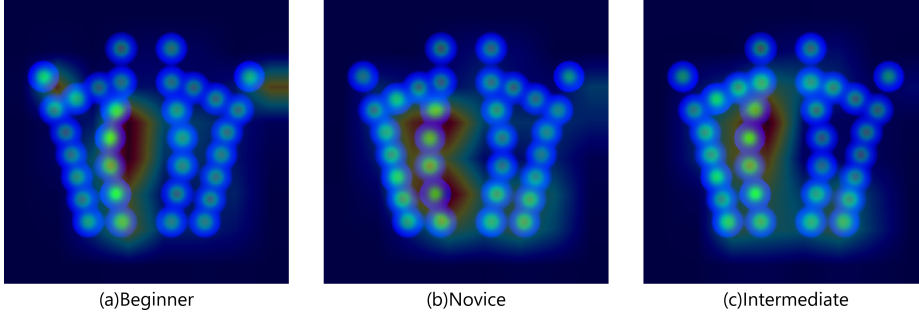


Fig. 4 Class-wise mean Eigen-CAM visualization results

5.1 procedure

To investigate longitudinal changes in weight-shift patterns, six beginner participants (three males and three females) from the classification experiment were measured repeatedly during simulator training. Participants completed an informed consent form before participating in the experiment. As in the experiment in Section 4, the participants engaged in measurement sessions after the practice session. In the measurement session, participants played the VizSki on a ski simulator for 1 min. Participants were required to complete the measurement session three times. Among them, the second and third measurements were used for classification. To evaluate whether weight-shifting patterns can be classified within participants, images from the first 40 s of each session were used for training, and those from the final 20 s for validation. We conducted these measurements two or three times per week, for a total of eight measurements (note: we suspended the experiment for subject 6 midway due to injuries outside the experiment). Experiments in this section were conducted as part of a study approved by the institutional ethical committee (JAIST Life Science Committee, ID: Human 04-056). This procedure was in accordance with the committee’s guidelines and regulations.

5.2 Results

5.2.1 Longitudinal Changes in Classification Outputs

We obtained 101,016 images from 92 sessions involving 6 participants. Of these, 33,672 images from the latter 20s of the experiment served as the validation data. We employed 10 classifiers in a classification test and calculated the average classification accuracy to assess robustness.

Figure 5(a) illustrates the temporal changes in classification outputs for each participant. In the initial training/validation sessions in Section 4, most samples were classified as Beginner, consistent with the participants’ original labels. As training progressed, several participants (Subjects 1, 2, 4, and 5) exhibited progressively higher proportions of samples classified as Novice or Intermediate continuously or temporarily.

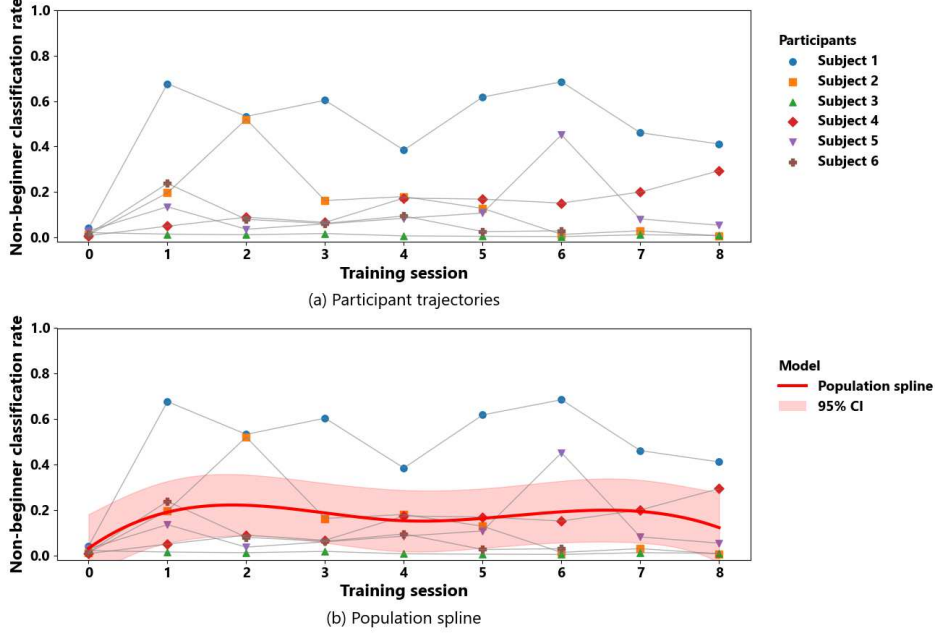


Fig. 5 Training status based on the classification result

To summarize the overall improvement across participants, we compared the non-beginner classification rate between the origin and final training sessions. The mean non-beginner classification rate increased by 0.113 (95% bootstrap CI: [0.002, 0.240]). Although the average classification rate increased, the Wilcoxon signed-rank test ($V = 3$, $p = 0.156$) did not reach statistical significance. The effect size was moderate (Hedges' $g = 0.56$, 95% CI: [0.117, 1.186]), suggesting that some participants exhibited meaningful improvements, although substantial inter-individual variability was observed.

5.2.2 Statistical Analysis of Longitudinal Trends

While the paired comparison summarized the overall change from the beginning to the end of training, it did not characterize the trajectory across intermediate sessions. Therefore, we fitted linear and spline-based mixed-effects models to evaluate the longitudinal trend throughout the training period. We first fitted a linear mixed-effects model with session as a fixed effect and participant as a random intercept:

$$y_{ij} = \beta_0 + \beta_1 s_{ij} + b_i + \varepsilon_{ij} \quad (1)$$

Session was treated as a fixed effect, while participant was included as a random intercept to account for repeated measurements. In addition, because visual inspection suggested that learning trajectories might be nonlinear, we additionally fitted a spline-based mixed-effects model using cubic B-splines (4 degrees of freedom):

$$y_{ij} = \beta_0 + \sum_{k=1}^K \beta_k B_k(s_{ij}) + b_i + \varepsilon_{ij} \quad (2)$$

B_k denotes the cubic B-spline basis functions and b_i represents the participant-specific random intercept. The statistical significance of the session effect was assessed using Wald tests on the fixed-effect coefficient. Models were estimated using maximum likelihood, and their goodness of fit was compared using the Akaike Information Criterion (AIC).

Table 3 shows the comparison result of the linear and spline mixed-effects models. The non-beginner classification rate showed no significant linear increase across sessions ($\beta = 0.006$, $p = 0.432$). However, the likelihood ratio test showed the spline mixed-effects model provided a better fit than the linear model (LR = 9.36, $p = 0.025$). It indicates that the longitudinal change was nonlinear rather than monotonic. The spline model also increased the marginal R^2 and the conditional R^2 , suggesting that nonlinear session effects explained additional variation beyond participant-specific differences.

Table 3 Comparison of the linear and spline mixed-effects models

Model	Fixed effect	Random effect	LogLik	AIC	Marginal R^2	Conditional R^2
Linear	Session	Participant	25.34	-42.67	0.005	0.569
Spline	Cubic B-spline (4 df)	Participant	30.02	-46.04	0.076	0.650

The estimated trajectory by the spline-based mixed-effects model in Fig. 5(b) exhibited three characteristic patterns: an initial increase in the non-beginner classification rate in the early training sessions, a relatively stable plateau in the middle sessions, and fluctuations in the later sessions. These findings suggest that participants tended to improve rapidly during the initial stage of training, then stabilize, after which individual differences became more pronounced.

Next, to better understand the biomechanical changes underlying the classifier-derived learning trajectory, we additionally analyzed representative foot-pressure variables using linear mixed-effects models. A linear mixed-effects model with participant as a random intercept was fitted for each biomechanical variable, and the statistical significance of the session effect was assessed using Wald tests on the fixed-effect coefficient.

Table 4 shows the summary of mixed-effects model analyses in foot-pressure. The reported p-values correspond to Wald tests of the session effect with the Benjamini-Hochberg method. Mean CoP shifted anteriorly over training, while CoP path length, forefoot pressure ratio, and lateral CoP variability increased. These findings suggest that participants progressively adopted larger and more active weight-transfer strategies rather than simply reducing postural sway.

Table 4 Summary of mixed-effects model analyses for biomechanical variables.

Metric	β	p	Interpretation
Mean CoP X position	+0.341	0.120	No significant change
Mean CoP Y position	-2.052	< 0.001	Forward CoP shift
SD of CoP (X)	+0.869	< 0.001	Increased lateral CoP variability
SD of CoP (Y)	-0.174	0.287	No significant change
CoP path length	+363.0	< 0.001	Longer CoP trajectory
Convex hull area	+122.2	0.020	Larger CoP sway area
95% ellipse area	+63.5	0.490	No significant change
Left-right symmetry error	+0.709	< 0.001	Increased left-right asymmetry
Left pressure ratio	-0.0075	< 0.001	Reduced left-foot loading
Right pressure ratio	+0.0075	< 0.001	Increased right-foot loading
Forefoot pressure ratio	+0.0186	< 0.001	Increased forefoot loading
Heel pressure ratio	-0.0246	< 0.001	Reduced heel loading
Mean total pressure	-197.0	< 0.001	Reduced overall pressure

6 Discussion

6.1 Interpretation of the results

The results presented in Fig. 3 suggest that plantar pressure distributions contain biomechanically meaningful information related to skiing experience. Because weight transfer between the left and right feet and between the forefoot and heel is essential for balance and turning, pressure distributions appear to capture important movement characteristics even without posture information.

The class-wise mean Eigen-CAM visualizations (Fig. 4) revealed largely similar attention patterns across proficiency levels, with only a modest expansion of attention toward the forefoot in the intermediate group. Although these visual differences were subtle, the accompanying mixed-effects analyses of biomechanical variables (Table 4) demonstrated systematic longitudinal changes in plantar-pressure characteristics, including anterior CoP displacement, increased forefoot loading, and changes in CoP trajectory. Together, these findings suggest that the classifier relied on distributed biomechanical characteristics rather than a single localized pressure region.

In the training experiment, several beginners showed a gradual increase in the proportion of samples classified as higher proficiency levels (Fig. 5). The mixed-effects analyses (Tables 3 and 4) further indicated that these changes were accompanied by systematic modifications of plantar-pressure behavior, including an anterior shift of the center of pressure, increased forefoot loading, and longer CoP trajectories. These findings suggest that participants progressively adopted more active weight-transfer strategies during simulator training rather than maintaining the initial pressure distribution. However, because the present study was conducted using a ski simulator, these biomechanical changes should be interpreted as simulator-specific movement adaptations rather than direct evidence of improved on-snow skiing performance.

The classification experiment used a participant-non-independent evaluation protocol, in which training and validation samples were obtained from the same participants but across different time segments. This design was selected because the number of

participants was limited and the primary objective of this study was to investigate whether proficiency-related patterns existed in foot-pressure distributions. Consequently, the reported classification accuracy should not be interpreted as an estimate of generalization performance for unseen users. Instead, it indicates that the classifier can consistently distinguish pressure patterns within the collected dataset. Although participant-independent evaluation is essential for developing a general-purpose proficiency estimator, participant-dependent evaluation remains useful for investigating whether meaningful pressure patterns exist within individuals and whether those patterns change throughout repeated training sessions.

6.2 Limitation

First, this study included no expert skiers, such as athletes or instructors, and proficiency levels were based on self-reported skiing experience. Consequently, future studies should recruit certified skiers to establish more reliable proficiency labels and improve the classifier’s robustness.

Second, the proficiency groups used in this study provided only coarse boundaries, particularly between the novice and intermediate classes. Future work should adopt standardized certification systems, such as the International Technical Skill Test, to define more objective proficiency criteria.

Third, because training and validation data included the same participants, the classifier may have learned participant-specific characteristics. Future work should evaluate the model using participant-independent protocols, such as leave-one-participant-out cross-validation, to assess generalizability to unseen users. In addition, temporal models such as LSTM were not used because the number of independent participants was limited, and consecutive pressure frames were highly correlated. To reduce model complexity and focus on interpretable spatial pressure patterns, this study adopted a YOLO11-based image classifier; temporal modeling is left for future work.

7 Conclusion

This study investigated the feasibility of estimating skiing proficiency from foot-pressure patterns acquired in a low-cost simulator-based training environment. Heatmap images generated from pressure-sensor data were classified using YOLO11, and the learned features were visualized with Eigen-CAM.

The results provide three principal findings. First, participant-dependent evaluation demonstrated that plantar pressure heatmaps contain sufficient information to distinguish predefined proficiency groups with high accuracy, indicating that weight-shifting patterns reflect skiing proficiency (RQ1). Second, longitudinal analyses suggested that the classifier outputs changed systematically throughout repeated simulator training, indicating that the learned representations captured changes in participants’ movement patterns over time (RQ2). Third, Eigen-CAM and quantitative attention analyses indicated that the classifier relied on biomechanically meaningful pressure-distribution characteristics rather than arbitrary image regions, supporting the interpretability of the proposed approach (RQ3).

691 However, this study was limited by the small sample size, self-reported profi-
692 ciency labels, and participant-dependent evaluation. Future work will collect data
693 from certified skiers, establish objective proficiency criteria, improve participant-
694 independent performance measures, and develop a real-time feedback system based
695 on extracted foot-pressure features. Overall, this study demonstrates the potential of
696 plantar-pressure-based deep learning for AI-assisted coaching in simulator-based ski
697 training.

698

699 **Statements and Declarations**

700

701 **Ethics approval and consent to participate**

702

703 Experiments in this research were conducted as part of a study approved by the JAIST
704 Life Science Committee (ID: Human 04-056) in accordance with the guidelines.

705

706 **Availability of data and materials**

707

708 The raw dataset supporting this study is publicly available in IEEE DataPort at DOI:
709 10.21227/e3jv-7t50. The dataset includes plantar pressure measurements collected
710 from beginner to intermediate skiers using the PRO SKI SIMULATOR.

711 **Competing interests**

712

713 The authors have no conflict of interest to declare that is relevant to the content of
714 this article.

715

716 **Funding**

717

718 This work was supported by JST SPRING (JPMJSP2102), JST CREST
719 (JPMJCR17A3), JST Moonshot R&D Grant (JPMJMS2012), and JSPS KAKENHI
720 (JP22K12288).

721 **Authors' contributions**

722

- 723 • Conceptualization: Shigeharu Ono, Ryosuke Atsumi
- 724 • Data curation: Shigeharu Ono
- 725 • Formal analysis: Shigeharu Ono
- 726 • Funding acquisition: Shigeharu Ono, Hideaki Kanai, Hideki Koike
- 727 • Investigation: Shigeharu Ono, Ryosuke Atsumi
- 728 • Methodology: Shigeharu Ono, Ryosuke Atsumi
- 729 • Project administration: Shigeharu Ono
- 730 • Resources: Hideaki Kanai, Hideki Koike
- 731 • Software: Shigeharu Ono, Ryosuke Atsumi
- 732 • Supervision: Hideaki Kanai
- 733 • Validation: Shigeharu Ono
- 734 • Visualization: Shigeharu Ono
- 735 • Writing – original draft: Shigeharu Ono
- 736 • Writing — review&editing: Shigeharu Ono, Hideaki Kanai

References

- [1] Neumann, D.L., Moffitt, R.L., Thomas, P.R., Loveday, K., Watling, D.P., Lombard, C.L., Antonova, S., Tremeer, M.A.: A systematic review of the application of interactive virtual reality to sport. *Virtual Reality* **22**(3), 183–198 (2018) <https://doi.org/10.1007/s10055-017-0320-5> . Springer
- [2] Murray, E.G., Neumann, D.L., Moffitt, R.L., Thomas, P.R.: The effects of the presence of others during a rowing exercise in a virtual reality environment. *Psychology of Sport and Exercise* **22**, 328–336 (2016) <https://doi.org/10.1016/j.psychsport.2015.09.007> . Elsevier
- [3] Wu, E., Perteneder, F., Koike, H., Nozawa, T.: How to vizski: Visualizing captured skier motion in a vr ski training simulator. In: *The 17th International Conference on Virtual-Reality Continuum and Its Applications in Industry. VRCAI '19*, New York, NY, USA (2019). <https://doi.org/10.1145/3359997.3365698> . Association for Computing Machinery
- [4] Ono, S., Kanai, H., Atsumi, R., Koike, H., Nishimoto, K.: Learning support and evaluation of weight-shifting skills for novice skiers using virtual reality. In: *International Conference on Human-Computer Interaction*, pp. 226–237 (2022). https://doi.org/10.1007/978-3-031-05887-5_16 . Springer
- [5] Tabrizi, S.S., Pashazadeh, S., Javani, V.: Comparative study of table tennis fore-hand strokes classification using deep learning and svm. *IEEE Sensors Journal* **20**(22), 13552–13561 (2020) <https://doi.org/10.1109/JSEN.2020.3005443> . IEEE
- [6] Liao, C.-C., Hwang, D.-H., Koike, H.: Ai golf: Golf swing analysis tool for self-training. *IEEE Access* **10**, 106286–106295 (2022) <https://doi.org/10.1109/ACCESS.2022.3210261> . IEEE
- [7] Panizzolo, F.A., Marcolin, G., Petrone, N.: Comparative evaluation of two skiing simulators as functional training devices for recreational skiers. *Journal of sports science & medicine* **12**(1), 151 (2013). Dept. of Sports Medicine, Medical Faculty of Uludag University
- [8] Moon, J., Koo, D., Kim, K., Shin, I., Kim, H., Kim, J.: Effect of ski simulator training on kinematic and muscle activation of the lower extremities. *Journal of Physical Therapy Science* **27**(8), 2629–2632 (2015) <https://doi.org/10.1589/jpts.27.2629>
- [9] Michahelles, F., Schiele, B.: Sensing and monitoring professional skiers. *IEEE Pervasive Computing* **4**(3), 40–45 (2005) <https://doi.org/10.1109/MPRV.2005.66> . IEEE
- [10] Hasegawa, S., Ishijima, S., Kato, F., Mitake, H., Sato, M.: Realtime sonification of the center of gravity for skiing. In: *Proceedings of the 3rd Augmented*

- Human International Conference. AH '12, pp. 1–4 (2012). <https://doi.org/10.1145/2160125.2160136> . Association for Computing Machinery
- [11] Zhao, Y., Zhou, J., Qiu, F., Liao, X., Jiang, J., Chen, H., Lin, X., Hu, Y., He, J., Chen, J.: A deep learning method for foot-type classification using plantar pressure images. *Frontiers in Bioengineering and Biotechnology* **11**, 1239246 (2023) <https://doi.org/10.3389/fbioe.2023.1239246>
- [12] Mun, F., Choi, A.: Deep learning approach to estimate foot pressure distribution in walking with application for a cost-effective insole system. *Journal of NeuroEngineering and Rehabilitation* **19**(1), 4 (2022) <https://doi.org/10.1186/s12984-022-00987-8>
- [13] Wu, C.-C., Tsai, C.-W., Wu, F.-E., Chiang, C.-H., Chiou, J.-C.: Plantar pressure-based gait recognition with and without carried object by convolutional neural network-autoencoder architecture. *Biomimetics* **10**(2), 79 (2025) <https://doi.org/10.3390/biomimetics10020079>
- [14] Dindorf, C., Dully, J., Simon, S., Perchthaler, D., Becker, S., Ehmann, H., Diers, C., Garth, C., Fröhlich, M.: Toward automated plantar pressure analysis: machine learning-based segmentation and key point detection across multicenter data. *Frontiers in bioengineering and biotechnology* **13**, 1579072 (2025) <https://doi.org/10.3389/fbioe.2025.1579072>
- [15] Snyder, C., Martínez, A., Jahnel, R., Roe, J., Stöggl, T.: Connected skiing: Motion quality quantification in alpine skiing. *Sensors* **21**(11) (2021) <https://doi.org/10.3390/s21113779> . MDPI
- [16] Yamagiwa, S., Ohshima, H., Shirakawa, K.: Development of skill scoring system for ski and snowboard. In: Cabri, J., Barreiros, J., Pizarat Correia, P. (eds.) *Sports Science Research and Technology Support*, pp. 1–15 (2015). https://doi.org/10.1007/978-3-319-25249-0_1 . Springer
- [17] Ludwig, K., Einfalt, M., Lienhart, R.: Robust estimation of flight parameters for ski jumpers. In: 2020 IEEE International Conference on Multimedia & Expo Workshops (ICMEW), pp. 1–6 (2020). <https://doi.org/10.1109/ICMEW46912.2020.9105973> . IEEE
- [18] Muhammad, M.B., Yeasin, M.: Eigen-cam: Class activation map using principal components. In: 2020 International Joint Conference on Neural Networks (IJCNN), pp. 1–7 (2020). <https://doi.org/10.1109/IJCNN48605.2020.9206626> . IEEE
- [19] Khanam, R., Hussain, M.: Yolov11: An overview of the key architectural enhancements. *arXiv preprint arXiv:2410.17725* (2024) <https://doi.org/10.48550/arXiv.2410.17725>

- [20] Chardonens, J., Favre, J., Cuendet, F., Gremion, G., Aminian, K.: A system to measure the kinematics during the entire ski jump sequence using inertial sensors. *Journal of Biomechanics* **46**(1), 56–62 (2013) <https://doi.org/10.1016/j.jbiomech.2012.10.005> . Elsevier
- [21] Groh, B.H., Weeger, N., Warschun, F., Eskofier, B.M.: Simplified orientation determination in ski jumping using inertial sensor data. In: 2014 DGON Inertial Sensors and Systems (ISS), pp. 1–11 (2014). <https://doi.org/10.1109/InertialSensors.2014.7049482> . IEEE
- [22] Bessone, V., Petrat, J., Schwirtz, A.: Ski position during the flight and landing preparation phases in ski jumping detected with inertial sensors. *Sensors* **19**(11), 2575 (2019) <https://doi.org/10.3390/s19112575> . MDPI
- [23] Groh, B.H., Fritz, J., Deininger, M., Schwameder, H., Eskofier, B.M.: Unobtrusive and wearable landing momentum estimation in ski jumping with inertial-magnetic sensors. In: 2018 IEEE 15th International Conference on Wearable and Implantable Body Sensor Networks (BSN), pp. 102–105 (2018). <https://doi.org/10.1109/BSN.2018.8329669> . IEEE
- [24] Takeda, M., Miyamoto, N., Endo, T., Ohtonen, O., Lindinger, S., Linnamo, V., Stöggl, T.: Cross-country skiing analysis and ski technique detection by high-precision kinematic global navigation satellite system. *Sensors* **19**(22) (2019) <https://doi.org/10.3390/s19224947> . MDPI
- [25] Elfmark, O., Ettema, G., Groos, D., Ihlen, E.A., Velta, R., Haugen, P., Braaten, S., Gilgien, M.: Performance analysis in ski jumping with a differential global navigation satellite system and video-based pose estimation. *Sensors* **21**(16), 5318 (2021) <https://doi.org/10.3390/s21165318> . MDPI
- [26] Supej, M.: 3d measurements of alpine skiing with an inertial sensor motion capture suit and gnss rtk system. *Journal of sports sciences* **28**(7), 759–769 (2010) <https://doi.org/10.1080/02640411003716934> . Taylor & Francis
- [27] Wojtyczek, B., Paślawska, M., Raschner, C.: Changes in the balance performance of polish recreational skiers after seven days of alpine skiing. *Journal of human kinetics* **44**, 29–40 (2014) <https://doi.org/10.2478/hukin-2014-0108>
- [28] Simsek, E., Arslan, H., Polat, M., Koca, F.: The effect of alpine skiing training on balance performance. *African Educational Research Journal* **8**(2), 357–362 (2020). ERIC
- [29] Ronao, C.A., Cho, S.-B.: Human activity recognition with smartphone sensors using deep learning neural networks. *Expert Systems with Applications* **59**, 235–244 (2016) <https://doi.org/10.1016/j.eswa.2016.04.032>
- [30] Kautz, T., Groh, B.H., Hannink, J., Jensen, U., Strubberg, H., Eskofier, B.M.:

875 Activity recognition in beach volleyball using a deep convolutional neural net-
876 work: Leveraging the potential of deep learning in sports. *Data Mining and Knowl-*
877 *edge Discovery* **31**, 1678–1705 (2017) <https://doi.org/10.1007/s10618-017-0495-0>
878 . Springer
879

880 [31] Yoon, Y., Hwang, H., Choi, Y., Joo, M., Oh, H., Park, I., Lee, K.-H., Hwang, J.-
881 H.: Analyzing basketball movements and pass relationships using realtime object
882 tracking techniques based on deep learning. *IEEE Access* **7**, 56564–56576 (2019)
883 <https://doi.org/10.1109/ACCESS.2019.2913953> . IEEE
884

885 [32] Lv, C., Li, J., Tian, J.: Key frame extraction for sports training based on improved
886 deep learning. *Scientific Programming* **2021**(1), 1016574 (2021) [https://doi.org/](https://doi.org/10.1155/2021/1016574)
887 [10.1155/2021/1016574](https://doi.org/10.1155/2021/1016574)
888

889 [33] Singh, M.S., Pondenkandath, V., Zhou, B., Lukowicz, P., Liwickit, M.: Trans-
890 forming sensor data to the image domain for deep learning—an application to
891 footstep detection. In: 2017 International Joint Conference on Neural Networks
892 (IJCNN), pp. 2665–2672 (2017). <https://doi.org/10.1109/IJCNN.2017.7966182> .
893 IEEE
894

895 [34] Ono, S., Kanai, H.: Foot Pressure Dataset of Snowplow Skiers Collected Using
896 the PRO SKI SIMULATOR. <https://doi.org/10.21227/e3jv-7t50> . [https://dx.doi.](https://dx.doi.org/10.21227/e3jv-7t50)
897 [org/10.21227/e3jv-7t50](https://dx.doi.org/10.21227/e3jv-7t50)
898

899 [35] Selvaraju, R.R., Cogswell, M., Das, A., Vedantam, R., Parikh, D., Batra, D.:
900 Grad-cam: Visual explanations from deep networks via gradient-based localiza-
901 tion. In: Proceedings of the IEEE International Conference on Computer Vision,
902 pp. 618–626 (2017). <https://doi.org/10.1109/ICCV.2017.74> . IEEE
903
904
905
906
907
908
909
910
911
912
913
914
915
916
917
918
919
920