

Democratizing Ski Safety: Real-Time Turn Segmentation with Smartphone IMU and Causal LSTM Networks

Michał Szymocha, Piotr Kacprzak, Jakub Robak and Wojciech Turek

AGH University of Krakow, 30-059 Krakow, Poland
jakubrobak@agh.edu.pl

Abstract

Anterior cruciate ligament (ACL) injury is one of the most common and serious injuries in sports, particularly among recreational skiers. Research shows that structured technique awareness and continuous feedback can significantly reduce the risk of such injuries, yet access to professional instructors is limited to wealthy athletes who can afford continuous private coaching, creating a harmful inequity in injury prevention. This gap can be mitigated by automating the real-time analysis of skiing techniques available to the wider recreational skiing community. The approach relies exclusively on inertial sensors embedded in standard smartphones, eliminating the need for specialized equipment and enabling broad social scalability. To support immediate feedback, the system operates causally, producing predictions based solely on past observations. The work is conducted in cooperation with professional ski instructors, ensuring that problem formulation, data annotation, and result evaluation reflect real-world coaching practices and injury prevention needs. The model is evaluated using Leave-One-Subject-Out validation on a public, in-the-wild dataset, demonstrating robust generalization across skiers, achieving an average directional accuracy of 89.8%, while maintaining extremely low inference latency suitable for on-device mobile deployment. This work outlines a practical pathway to democratizing injury prevention in recreational sports.

1 Introduction

Alpine skiing is one of the most popular forms of winter recreation. The latest industry reports estimate the number of slope visits at more than 366 million per year [Vanat, 2025]. Its popularity remains unimpeded by a high risk of serious injury. Extensive studies of competitive skiing [Pujol *et al.*, 2007; Jordan *et al.*, 2017] have shown a high incidence of anterior cruciate ligament (ACL) ruptures, highlighting the inherent biomechanical risks of this sport. In recreational skiing, these risks are further exacerbated by a lack of continuous technical supervision and awareness of movements,

often leading to an unsafe loss of ski control [Ruedl *et al.*, 2023]. Beyond severe injuries, this constitutes a broad public safety challenge that affects millions of recreational skiers each season [Davey *et al.*, 2019]. Crucially, fundamental research has shown that awareness of structured technique and ongoing correction can significantly reduce the risk of serious knee sprains [Ettlinger *et al.*, 1995].

Unlike professional athletes who train in structures with regular coaching support, recreational skiers typically lack access to systematic feedback on technique. Research on sport participation reveals persistent socioeconomic inequalities. Individuals of higher socioeconomic status are significantly more likely to participate in organized physical activity than those from lower income groups, indicating that financial resources influence access to sport [Richard *et al.*, 2023]. Furthermore, research indicates that the costs associated with participating in sports are perceived as a significant barrier. Many families limit participation in activities due to their inability to purchase equipment, memberships, or training sessions [O’Kane *et al.*, 2025]. In addition, literature reviews highlight that equipment expenses, fees, and other costs are among the most common practical barriers to participation in sports across various age groups [Somerset and Hoare, 2018]. This creates a significant gap in safety support and movement education among the broader recreational ski community. Widely available smartphone-based mobile systems offering automated real-time monitoring could help alleviate this inequality by providing feedback without the recurring costs associated with traditional coaching.

Existing smartphone apps for skiers are usually based on GNSS data, which do not offer the time resolution required for detailed technique evaluation [Petrella *et al.*, 2025]. There are advanced commercial systems, such as smart insoles with pressure sensors (e.g., Carv systems [Enoiu *et al.*, 2023]), that provide granular feedback. However, these are expensive and require specialized equipment, limiting their accessibility. In contrast, inertial sensors embedded in standard smartphones offer a widely available sensing modality for movement analysis, enabling socially scalable technique assessment [Azadi *et al.*, 2022].

In order to become an efficient “digital safety assistant”, a smartphone must first detect and understand the basic unit of a skier’s movement, i.e. a single turn. Precise temporal segmentation, as shown in the literature [Martínez *et al.*, 2021],

is a necessary condition for further biomechanical analysis, which is crucial for safety assessment. Synchronization errors at this stage lead to false risk assessments, making a robust segmentation algorithm the foundation of the entire system.

Efficient operation in uncontrolled real-world conditions (in-the-wild) encounters a significant technological barrier resulting from the need to reconcile algorithmic robustness with real-time requirements. Physics-based methods that use rigid thresholds achieve high precision in laboratory settings but fail to generalize to uncontrolled conditions [Yamagiwa *et al.*, 2014; Martínez *et al.*, 2021]. In contrast, advanced deep learning architectures often rely on non-causal or bidirectional processing, requiring access to future context [Chen *et al.*, 2022] and thus precluding their use for real-time injury prevention. Therefore, there is an urgent need to develop an architecture that retains the robustness of deep learning while operating causally, eliminating the delays that result from analyzing future context.

This study is conducted in collaboration with professional ski instructors, ensuring that problem formulation, data annotation, and evaluation criteria reflect real-world coaching, movement education, and safety practices rather than purely algorithmic assumptions. As a result, this work provides a foundational building block for future smartphone-based digital safety assistants by enabling a reliable real-time interpretation of ski maneuvers.

The contributions of this work are threefold.

(i) We formulate the problem of ski turn segmentation as a causal frame-level recognition task suitable for real-time processing on mobile devices.

(ii) We propose a lightweight, learning-based approach for turn segmentation using only inertial sensors embedded in standard smartphones, designed to operate robustly in real-world, in-the-wild conditions.

(iii) We validate the proposed approach under subject-independent Leave-One-Subject-Out evaluation on a public, expert-annotated dataset, demonstrating strong generalization across skiers while meeting the latency constraints required for on-device deployment.

2 Related Work

Traditional turn detection methods in alpine skiing are based on physical models. A comprehensive review of published approaches [Martínez *et al.*, 2021] shows that existing methods achieve high accuracy in the laboratory, but field validation revealed reduced performance for non-carving styles (e.g., drifted: 0.833, snowplow: 0.538 for the Ratio score) and variability dependent on snow conditions [Martínez *et al.*, 2019]. So far, machine learning has focused mainly on classifying skiers' entire runs, not on the precise segmentation of turn boundaries [Jones *et al.*, 2016].

A recent study [Robak and Turek, 2025] addressed the problem of detecting ski turns using smartphone IMU sensors by developing a gradient descent-based algorithm on a public dataset. The method achieved an F1 score of 0.943 in detecting complete turns (turn-level) and 89.7% accuracy in mapping the turn direction (left/right) for each signal time

point. However, this approach requires offline processing of entire runs (post-hoc optimization), operates at the level of global transition points instead of frame-level prediction, and is evaluated using randomized data splitting without subject hold-out validation.

In related domains, LSTM networks have shown promising results. LSTM models for gait phase estimation have achieved errors below 100 ms on IMU sensors [Ding *et al.*, 2018; Tang *et al.*, 2024], and recent studies using a single IMU on a smartphone reported a median error of around 70–80 ms with 92% accuracy of events [Larsen *et al.*, 2024]. Human activity recognition (HAR) using LSTM and CNN achieves 95–99% accuracy [Xia *et al.*, 2020], and video temporal action segmentation uses bidirectional models achieving high precision in boundary localization [Ding *et al.*, 2023].

Despite these advances, existing approaches fundamentally differ from the requirements of ski turn segmentation in real-world recreational conditions. Most HAR and temporal segmentation systems operate in time windows of a few seconds rather than single-sample predictions, utilize bidirectional architectures requiring access to future data, and are predominantly validated in controlled laboratory settings. These assumptions limit the applicability of such methods to real-time, safety-oriented systems running on mobile devices in uncontrolled skiing conditions.

Furthermore, compact LSTM models can perform real-time causal inference on mobile devices [Chen *et al.*, 2022], but their application to temporal event detection on raw IMU data from ski turns in real-world terrain (in-the-wild) remains unexplored.

Based on this review, four key research gaps can be identified:

(i) The absence of causal methods for real-time ski turn segmentation using mobile inertial sensors.

(ii) Limited generalization between skiers, skill levels, and riding styles in real uncontrolled slope conditions, where existing algorithms exhibit reduced performance.

(iii) The lack of efficient deployment of advanced sequence models, such as LSTM networks, on resource-constrained smartphones for real-time temporal event detection.

(iv) The absence of systematic comparisons between simple rule-based approaches and modern deep learning architectures for the ski turn segmentation task.

This work addresses these gaps by proposing a causal LSTM architecture for real-time ski turn segmentation on raw IMU data from standard smartphones in real-world, in-the-wild conditions, based on an extended dataset with additional attributes (skill level, skiing style) to improve cross-skier generalization.

3 Dataset and Expert Domain Knowledge

The effectiveness of AI systems for recreational safety support depends on the reliability of training data and its connection with the biomechanics of movement. This section describes a set of IMU data recorded in real-world conditions (in-the-wild) that was analyzed and annotated by an expert ski instructor. By combining smartphone-based sens-

ing with instructor-provided annotations and contextual metadata, the dataset supports safety-oriented applications such as risk-aware feedback and technique assessment in recreational sports. Importantly, relying on commodity mobile devices reduces economic barriers and enables inclusive access to these systems, regardless of financial means or access to professional instruction.

3.1 Data Source

The experimental basis of this work is a publicly available IMU dataset recorded during real-world skiing runs (in-the-wild) [Robak and Turek, 2025]. The dataset is annotated by a certified ski instructor using synchronized video recordings that serve as a reference for temporal alignment and the definition of the turn boundary. In accordance with the annotation protocol, each timestamp is assigned a binary label corresponding to the current turn direction (left or right), as described in professional curricula, such as the PSIA Alpine Technical Manual. In this setting, straight or no-turn segments are treated as brief transition periods rather than separate target states, since our goal is to analyze the active turning phases most relevant for turn-shape assessment and safety feedback. The collection comprises 105 runs performed by 11 different skiers: 7 males and 4 females, with 9 participants aged 18 to 30 years and 2 aged 40 to 55 years, resulting in 1,781 recorded turns. Each ski run is recorded as a continuous multivariate IMU time series sampled at 10 Hz using inertial sensors embedded in standard smartphones. Following prior work on smartphone-based ski motion analysis, the device is mounted on the skier’s calf, providing stable coupling with lower-limb kinematics while remaining practical for real-world use. The participants represented a full spectrum of skill levels, from beginner to certified ski instructor, making the data highly representative of the diverse recreational skier population. An important element of the realism of the data is the fact that the skiers moved freely, without an imposed route or gates, which introduces natural variability in the trajectory and movement dynamics typical of ski tourism. The data was recorded using inertial sensors built into standard smartphones, which guarantees high social scalability of the solution.

3.2 Expert-Led Metadata Enrichment

A key value of the collection is the rigorous process of determining the “ground truth” conducted by a ski instructor. Instead of relying on simplistic physical thresholds, the limits of each turn were defined on the basis of biomechanical analysis of synchronized video recordings. This high labeling precision is essential for ACL injury prevention systems, where fractions of a second in motion phase detection can strongly affect the accuracy of biomechanical risk assessment in dynamically changing slope conditions [Spörri *et al.*, 2016].

To further adapt the data to the requirements of real-world injury prevention, the original dataset was augmented with expert metadata. The Enrichment process was conducted by the same certified ski instructor, who performed a detailed analysis of synchronized video recordings and raw inertial sensor signals. As a result, each of the 105 runs was assigned

precise contextual attributes: skier skill level (skier level) distributed as 2 beginners, 5 intermediates, 3 advanced, and 1 expert; skiing style (skier style); and a unique user identifier (skier ID). In accordance with the annotation protocol, each timestamp is assigned a binary label corresponding to the current turn direction (left or right), without introducing an explicit straight-skiing or no-turn class.

The addition of these parameters allows for rigorous validation of the model’s ability to generalize across different user technical profiles. From the perspective of ACL injury prevention, this approach is essential for creating a system that can accurately interpret movement dynamics in individuals with a wide range of technical skills, from beginners to experts. Integrating domain knowledge in the feature set definition stage lays the foundation for the democratization of professional safety systems, making them effective for every smartphone user.

4 Methodology

This section presents a causal approach to ski turn recognition from smartphone IMU data. The goal is accurate frame-level segmentation in real-world conditions under mobile and real-time constraints. We use a lightweight recurrent model that relies only on past observations.

4.1 Problem Formulation

We address the problem of online ski turn recognition from multivariate inertial sensor data. Given a continuous IMU time series recorded during a skiing run, the task is to predict the current turn direction at each time step, resulting in a frame-level binary labeling of left and right turns.

Formally, let $x_t \in \mathbb{R}^d$ denote the multichannel IMU measurement at time step t . The objective is to learn a model f_θ that produces an online prediction of the current turn direction based solely on past and present observations in Equation 1.

$$\hat{y}_t = f_\theta(x_{1:t}). \quad (1)$$

The problem is challenging due to substantial noise in consumer-grade IMU signals, pronounced inter-subject variability in skiing technique, and changing environmental conditions encountered in real-world skiing. Moreover, enforcing causality introduces an inherent trade-off between temporal precision and latency, as reliable detection of turn transitions may require the accumulation of sufficient historical context without ever accessing future data, which is a strict constraint for both the unidirectional LSTM and preprocessing imposed by real-time mobile deployment rather than accuracy optimization.

4.2 Data Preprocessing

The preprocessing pipeline is designed to stabilize raw IMU signals for recurrent modeling while preserving causality and minimizing additional latency, which is critical for real-time operation on mobile devices.

To handle phase discontinuities in rotational signals, we apply phase unwrapping to the orientation components to obtain continuous angular representations suitable for temporal modeling [Robak and Turek, 2025]. Without unwrapping,

angular representations exhibit artificial discontinuities that introduce spurious high-magnitude jumps, which can significantly degrade temporal modeling in recurrent networks.

To reduce high-frequency measurement noise while preserving short-term motion dynamics, each signal component is filtered online using an exponential moving average (EMA). This filter is strictly causal, as each output value depends solely on the current observation and the previous filter state, allowing it to be applied in real time without access to future samples. EMA filtering is performed independently for each run to prevent information leakage between sessions.

Then, feature normalization is subsequently applied using statistics computed exclusively on the training data within each fold.

To prepare data for temporal modeling, we generate fixed-length sequences using a sliding window approach with a unit stride, labeled according to the ground-truth class of the final timestamp, which is consistent with the causal setting adopted for online prediction. This solution provides sufficient time context to model skiing dynamics while limiting the sequence length for stable optimization and effective training.

In addition to raw sensor signals, we derive a small set of physics-informed features to capture motion characteristics relevant to skiing behavior. For accelerometer and gyroscope signals, we compute magnitude vectors using the ℓ_2 norm to represent overall motion intensity. Jerk is computed as the temporal derivative of acceleration, providing a measure of movement smoothness [Hogan and Sternad, 2009]. These features offer significant physical context with negligible computational overhead. By capturing motion intensity and smoothness through the ℓ_2 norm and jerk, we provide the model with interpretable indicators of skier stability and dynamics that are not immediately accessible from raw signals alone. Most importantly, we calculate and include the first and second derivatives of yaw. Finally, local rolling statistics are computed to capture short-term temporal patterns without violating causality.

Overall, the proposed data preprocessing pipeline provides noise-robust, strictly causal input representations suitable for real-time recursive modeling in uncontrolled, real-world skiing conditions.

4.3 Data Augmentation

The data augmentation strategy is introduced to mitigate limited subject diversity and variability in execution speed, which are particularly pronounced under subject-independent (LOSO) evaluation and can hinder generalization in real-world deployment. Notably, the dataset is relatively small, and models tend to overfit quickly on it.

To increase temporal variability while preserving the semantic structure of individual turns, we apply data augmentation using segment-wise time warping. Unlike standard window warping applied to entire sequences, we perform warping independently within each annotated turn segment of a ski run. This design preserves behavioral transitions while allowing realistic variations within turns.

Window warping is applied using four temporal scale factors: 0.5, 0.75, 1.25, and 1.5, following established practice in

inertial time-series augmentation [Rashid and Louis, 2019]. The window ratio parameter is set to 1.0, allowing warping to span the full duration of each turn segment. This procedure produces four augmented variants for each original run.

The warping operation relies on temporal interpolation, which has the additional side effect of mildly smoothing high-frequency sensor noise while preserving low-frequency motion dynamics. As a result, the augmented signals lead to more stable optimization during training.

All augmented samples are generated exclusively from the training subset within each LOSO fold, and models are trained using the augmented data. It improves the generalization ability of a model, thanks to the mild smoothing side effect of augmentation. Importantly, the original (non-augmented) training data is *not* used during optimization. Validation and test subsets remain completely unmodified, ensuring that hyperparameter tuning and final evaluation reflect performance on realistic, non-distorted sensor dynamics.

In general, this augmentation strategy effectively quadruples the size of the training set while maintaining physically plausible temporal relationships within behavioral segments. It also reflects realistic differences in how quickly a skier executes a turn without altering the underlying motion pattern.

4.4 Training Strategy

A critical failure mode in kinematics-based Human Activity Recognition (HAR) is the “identity bias”, where models overfit biometric characteristics (e.g., unique sensor mounting angles or specific posture habits) rather than learning universal movement mechanics. Standard randomized k -fold cross-validation exacerbates this by allowing temporally correlated windows from the same subject to appear in both training and testing sets, leading to inflated performance estimates. [Tello *et al.*, 2024]

To enforce the learning of generalizable subject-invariant physical features, we employ a strict Leave-One-Subject-Out (LOSO) training strategy. The training process is iterative: for each unique subject with at least 5 descents $k \in \{1, \dots, N\}$, we create a test set containing exclusively the runs performed by subject k . The model is trained from scratch on the remaining pool $S \setminus \{k\}$. This strategy simulates a “cold-start” deployment scenario, ensuring that the optimization process never observes the kinematic characteristics of the target user, thereby testing the model’s ability to apply learned biomechanics to unseen individuals.

Additionally, within each training iteration, we enforce strict data hygiene to prevent spatiotemporal leakage. The training pool is further divided into train and validation subsets. Crucially, this split is performed at the entire run rather than at the window level. Splitting by window would result in training and validation samples that are milliseconds apart, trivially correlating in phase and amplitude. Our file-level splitting ensures that the validation set represents distinct events.

Finally, to stabilize the training of the recurrent architecture, we adopt a Dense Supervision objective. Rather than optimizing a single prediction at the end of the sequence, which forces error gradients to backpropagate throughout the temporal depth, often leading to vanishing gradients, we su-

pervise the network at every time step t . This approach enhances our model’s causal capabilities. The objective function is Cross-Entropy Loss.

Synthesizing strict LOSO validation with dense supervision and run-level hygiene yields a robust optimization pipeline that effectively eliminates data leakage while stabilizing the learning of subject-invariant biomechanics.

5 Proposed Architecture

To address the challenges of alpine skiing recognition, specifically the variability in sensor placement, the multi-frequency nature of motion, and long-term temporal dependencies—we propose a hybrid Residual Recurrent Network. Our architecture transforms raw inertial signals into a rich latent representation through three distinct stages, as illustrated in Figure 1.

Learnable Sensor Calibration and Alignment Sensor misalignment remains a bottleneck in wearable-based motion analysis, where physical orientation often varies between subjects. Rather than relying on imprecise manual alignment, we incorporate an entry-level Calibration Layer designed to normalize the input manifold before temporal processing. This 1×1 convolutional block serves as a learnable spatial transformer, optimizing a projection matrix W_{cal} that aligns disparate sensor frames into a canonical feature representation. This data-driven realignment reduces the model’s sensitivity to physical setup, allowing it to focus on the underlying motion dynamics rather than subject-specific sensor bias.

Multi-Scale Spatiotemporal Feature Extraction Skiing dynamics exhibit features at different temporal resolutions: high-frequency vibrations indicating ski-snow interaction (edge chatter) and low-frequency distinct phases (weight shift, steering). A single convolutional kernel size fails to capture this spectrum simultaneously. We employ a Multi-Scale Convolutional Block comprising four parallel branches with kernel sizes $k \in \{1, 3, 5, 11\}$.

- The $k = 1$ branch captures instantaneous global trends.
- The $k = 3, 5$ branches extract short-term interactions and rapid orientation changes.
- The $k = 11$ branch captures long-duration maneuvers and trajectory shapes.

The outputs are concatenated to form a unified feature map. To explicitly model the interdependencies between physical channels (e.g., the correlation between roll rate and lateral acceleration), we integrate a Squeeze-and-Excitation (SE) block [Hu *et al.*, 2018]. This mechanism recalibrates channel-wise feature responses by explicitly modeling dependencies, effectively acting as a soft attention mechanism that suppresses noise and amplifies informative physical signals.

Residual Recurrent Encoding We model the dynamics of each turn using a stacked LSTM architecture. To improve training stability, we incorporate residual skip connections ($h_t^l = \text{LSTM}(h_t^{l-1}) + h_t^{l-1}$), which allow gradients to flow more freely and enable the model to bypass layers that do not contribute to the representation. To ensure these representations remain robust, we use Locked Dropout. Unlike standard

dropout, this approach fixes the dropout mask across the entire sequence, preventing the model from relying on localized noise and encouraging the learning of stable temporal patterns [Merity *et al.*, 2017].

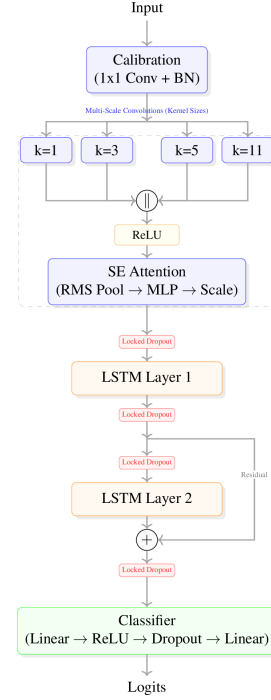


Figure 1: SkiC-LSTM architecture scheme

Layer	Output Dim.	Configuration / Parameters
1. Pre-processing		
Permute	(C_{in}, T)	Transpose $(B, T, C_{in}) \rightarrow (B, C_{in}, T)$
Calibration	(C_{in}, T)	Conv1d $(C_{in} \rightarrow C_{in}, k=1)$, bias=False
BatchNorm	(C_{in}, T)	BN1d $(C_{in} \text{ features})$, $\epsilon = 1e^{-5}$
2. Feature Extraction		
Multi-Scale Conv	$(64, T)$	4 Parallel Branches ($C_{out} = 16$ each): 1. Conv1d $(k=3, p=1)$ 2. Conv1d $(k=5, p=2)$ 3. Conv1d $(k=11, p=5)$ 4. Conv1d $(k=1, p=0)$
SE Block	$(64, T)$	Reduction $r=4$, RMS Pooling
Activation	$(64, T)$	MLP: $64 \rightarrow 16 \xrightarrow{\text{ReLU}} 64 \xrightarrow{\text{Sigmoid}}$ ReLU
3. Recurrent Layers		
Permute	$(T, 64)$	Transpose to (B, T, C_{feat})
Dropout	$(T, 64)$	Locked Dropout ($p=0.30$)
LSTM Layer 1	$(T, 128)$	Input: 64, Hidden: 128, Unidirectional
LSTM Layer 2	$(T, 128)$	Input: 128, Hidden: 128, Unidirectional Residual Connection: $x + \text{LSTM}(x)$
4. Classification Head		
Slice	(128)	Select last time step $t=T$
FC Layer 1	(128)	Linear $128 \rightarrow 128$, ReLU
Dropout	(128)	Standard Dropout ($p=0.30$)
FC Layer 2	(2)	Linear $128 \rightarrow 2$ (Logits)

Table 1: Detailed architecture of the proposed SkiC-LSTM model. The input sequence has dimensions $B \times T \times C_{in}$, where T represents the window size and C_{in} represents the number of input features. Model parameters were tuned using grid search over a candidate set of hyperparameters.

6 Experiments

Multiple experiments were conducted to examine the performance of our architecture in comparison to existing approaches under standard evaluation protocols and the generalization across different skiers, skier levels, and styles. Comparisons are performed under identical training and testing conditions against strong baselines.

6.1 Compared Baselines

We compare the proposed method against a set of classical and deep-learning baselines commonly used in time-series Human Activity Recognition (HAR), covering both feature-based and end-to-end learning paradigms.

Concretely, we include **Random Forest** and **XGBoost** as representative feature-based baselines, as tree ensembles are routinely reported in HAR benchmarks and comparative studies [Syed *et al.*, 2020; Zhang *et al.*, 2019].

For end-to-end learning, we use **LSTM** and **ResNet1D**-style architectures, which are commonly adopted deep baselines for wearable/smartphone time-series HAR across standard datasets [Hammerla *et al.*, 2016; Ordóñez and Roggen, 2016; Kim and Kim, 2025]. While bidirectional models are feasible, we do not include them because they require future context and therefore fall outside the real-time causal setting considered here.

6.2 Evaluation Metrics

Performance is evaluated using Accuracy and Intersection over Union. Each time step is treated as an independent prediction, measuring the model’s ability to maintain consistent semantic interpretation throughout the duration of a turn, rather than only predicting the dominant label of a window. Accuracy quantifies the proportion of timestamps for which the model correctly predicts the direction of a turn. The Intersection over Union metric is calculated by taking the average of the best possible *IoU* for each turn matched with the most overlapping continuous prediction.

We compute accuracy metrics both globally and per individual skier. This dual-level evaluation allows us to verify the overall efficacy of the dataset while simultaneously ensuring the model remains robust against inter-subject variability in style and technique. To avoid the high uncertainty associated with skiers who have very few timestamps, we reduced the evaluation pool to valid skiers with at least 5 descents. This resulted in a final test dataset consisting of 99 runs, 9 different skiers, and 1656 turns.

The inclusion of *IoU* is particularly important for this domain. While accuracy provides a general performance overview, it can be misleadingly high in datasets dominated by a background class (e.g., long periods of straight skiing). *IoU*, conversely, measures the precise temporal overlap between the predicted and ground-truth events. It penalizes both over-prediction (false positives) and under-prediction (missed duration) equally, ensuring that a high score reflects not just the detection of a turn but the accurate recovery of its boundaries.

This evaluation protocol is deliberately strict: even small temporal misalignments at turn boundaries are penalized at

every affected timestamp. Consequently, the reported metrics reflect both classification correctness and temporal precision, which are critical for downstream applications such as real-time feedback and fine-grained skill assessment.

6.3 Implementation Details

All models operate on sliding windows of fixed length $W = 12$ with a stride of one. This dense overlapping scheme ensures high temporal resolution and maximizes data utilization while preserving local temporal continuity.

All available IMU channels are used as input features; however, some of them are transformed, allowing the models to exploit the full inertial signal space captured by the device.

Neural network models are implemented in PyTorch [Paszke *et al.*, 2019]. The ResNet1D baseline follows a standard residual design consisting of an initial convolutional stem (kernel size 3), followed by six residual blocks operating at a constant channel width of 128. Each block uses kernel size 3 and dropout rate of 0.1.

Random Forest and XGBoost baselines are trained using standard fitting procedures from scikit-learn, while their hyperparameters were tuned via grid search across commonly used configuration ranges. All baselines share the same windowing strategy and are evaluated under identical data partitions to ensure a fair comparison.

SkiC-LSTM follows the specification in Table 1. We optimize it with AdamW using learning rate $1 \cdot 10^{-5}$ and weight decay $1 \cdot 10^{-4}$. To improve generalization, we apply a dropout rate of 0.3. Training runs for a maximum of 40 epochs, utilizing an early stopping mechanism with a patience of 5 epochs.

A characteristic of frame-level classifiers is that the raw output exhibits high-frequency jitter. While these short duration anomalies do not significantly impact global accuracy metrics, they affect the visual smoothness required for user-facing applications. To mitigate this, we implemented a median filter as a post-processing step. Although this introduces a 0.5s latency, resulting in a near-real-time rather than a strictly real-time system, it effectively eliminates artifacts. Crucially, this lightweight post-processing, combined with the SkiC-LSTM’s efficient inference, ensures the pipeline remains computationally inexpensive, making it highly suitable for direct deployment on mobile edge devices where resource constraints are paramount. For comparison purposes, we applied the same post-processing to every baseline to ensure that all of the results are comparable.

All preprocessing and data-dependent transformations are computed per training fold, as described in Training Strategy section, with no test-set statistics used. All baselines are evaluated under the same causal input protocol as our model (tree models use an unrolled representation). Experiments use the dataset from section 3.

7 Results

The proposed method is compared with representative baselines under identical training and testing splits, and performance is reported using evaluation metrics defined in subsection 6.2. The quantitative results are presented first, followed by ablation studies and analyzes to better understand the contributions of individual components.

7.1 Quantitative Results

Method	Per-skier Acc	Per-run Acc	Per-run IoU
ResNet1D	86.45	87.18	75.36
Standard LSTM	86.84	87.30	75.64
Random Forest	87.18	87.73	75.01
XGBoost	88.50	88.66	78.51
SkiC-LSTM (Ours)	89.77	89.85	79.45

Table 2: Segmentation performance (%) on the Test Set.

SkiC-LSTM achieves the best results under identical splits, with the highest per-skier accuracy (89.77), per-run accuracy (89.85), and per-run IoU (79.45) in Table 2. The IoU gain indicates more temporally coherent segments and reduced boundary jitter.

Method \ Style	Carving	Quick	Skidded	Snowplow
ResNet1D	87.83	78.89	87.89	85.38
LSTM	89.08	84.44	87.97	82.57
RandomForest	89.40	82.52	87.80	85.46
XGBoost	90.94	88.18	89.30	84.74
SkiC-LSTM (Ours)	92.10	87.60	89.70	88.24

Table 3: Comparison of Accuracy (%) across different skiing styles.

Method \ Level	0	1	2	3
ResNet1D	83.58	88.91	89.33	80.58
LSTM	82.00	89.68	90.32	81.02
RandomForest	85.12	88.97	89.10	84.01
XGBoost	84.45	90.45	90.82	86.31
SkiC-LSTM (Ours)	87.07	90.99	91.53	88.11

Table 4: Comparison of Accuracy (%) across different skier skill levels (0: Beginner, 1: Intermediate, 2: Advanced, 3: Expert).

SkierID	ResNet1D	LSTM	RF	XGB	SkiC-LSTM
1	89.13	91.54	91.11	90.89	91.44
2	82.00	78.25	83.67	81.63	85.58
4	91.36	92.06	91.32	91.71	92.31
5	85.09	85.60	86.50	87.14	88.49
6	88.31	87.95	87.17	89.79	89.55
8	83.35	85.18	82.55	88.20	89.22
9	89.04	90.11	88.69	89.19	92.15
10	89.18	89.90	89.63	91.68	91.06
11	80.58	81.02	84.01	86.31	88.11
Mean	86.45	86.84	87.18	88.50	89.77
σ	3.78	4.78	3.26	3.21	2.20

Table 5: Comparison of Accuracy (%) for each skier with at least 5 runs.

We further stratify the results by skiing style, skill level, and individual skiers to assess generalization and subject-specific variability.

7.2 Robustness Analysis

Visual inspection confirms the architecture’s strong capacity for preserving the morphological shape of motion sequences. Quantitative error analysis reveals that accuracy losses are primarily attributable to temporal phase shifts rather than event misclassifications; the model remains highly reliable in detecting turn occurrence and duration. These phase shifts typically manifest as consistent differences of n frames, a phenomenon inherent to the causal nature of SkiC-LSTM processing historical context.

Regarding style generalization, the model demonstrated the expected superior predictive performance on clean carving turns. For ”snowplow” techniques, data were scarce: only 10 descents, mostly from one subject during LOSO training. Despite this, the performance degradation was acceptable, with a directional accuracy drop of only 2% Table 3. In particular, our model demonstrated better generalization compared to recurrent baselines, which failed to converge effectively in low-data scenarios.

Performance stratification by skier level followed an expected trend: advanced skiers yielded the highest accuracy, followed by the intermediate group, which accounted for the majority of the dataset. Consistent with our expectations, Table 4 shows that the model’s performance was lowest for beginners. This drop is attributable to a combination of factors: the inherently high variability of beginner movements and the limited sample size, as only two beginner skiers were included in the study.

Finally, the IoU gains in Table 2 indicate improved segment overlap with the ground truth, reflecting better boundary alignment and reduced temporal fragmentation.

Our model achieved an inference time of around 0.1 ms per frame (on a single core 3 GHz CPU) leaving several orders of magnitude of compute headroom for deployment on resource-constrained devices, while matching the performance of much heavier offline models [Robak and Turek, 2025], which are not suitable for real-time feedback.

8 Conclusion

The developed hybrid LSTM architecture successfully segments IMU data into individual skiing maneuvers with high efficiency. By achieving precise turn segmentation even in the presence of ”in-the-wild” noise, the model proves to be more resilient than traditional physics-based approaches, which often struggle with the irregularities of real-world recreational skiing.

The causal architecture and very low inference time support real-time processing on standard smartphone hardware. By enabling automated data structuring without the need for professional oversight, this work offers a practical path to mitigate harmful inequity in access to coaching and injury prevention tools. Ultimately, this research provides the essential framework for a future real-time smartphone assistant, where reliable maneuver segmentation serves as the critical first step toward comprehensive technique evaluation and skier safety¹.

¹Source code, pretrained weights, and curated dataset are available at <https://github.com/mszymocha/SkiC-LSTM>.

References

- [Azadi *et al.*, 2022] Behrooz Azadi, Michael Haslgrübler, Bernhard Anzengruber-Tanase, Stefan Grünberger, and Alois Ferscha. Alpine skiing activity recognition using smartphone’s imu. *Sensors*, 22(15):5922, 2022.
- [Chen *et al.*, 2022] Zhe Chen, Hugh T Blair, and Jason Cong. Energy-efficient lstm inference accelerator for real-time causal prediction. *ACM Transactions on Design Automation of Electronic Systems (TODAES)*, 27(5):1–19, 2022.
- [Davey *et al.*, 2019] Annabelle Davey, Nathan K Endres, Robert J Johnson, and Jasper E Shealy. Alpine skiing injuries. *Sports health*, 11(1):18–26, 2019.
- [Ding *et al.*, 2018] Zhen Ding, Chifu Yang, Kai Xing, Xueyan Ma, Kai Yang, Hao Guo, ChunZhi Yi, and Feng Jiang. The real time gait phase detection based on long short-term memory. In *2018 IEEE Third International Conference on Data Science in Cyberspace (DSC)*, pages 33–38. IEEE, 2018.
- [Ding *et al.*, 2023] Guodong Ding, Fadime Sener, and Angela Yao. Temporal action segmentation: An analysis of modern techniques. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 46(2):1011–1030, 2023.
- [Enoiu *et al.*, 2023] Sandu Razvan Enoiu, Denisa Iulia Brus, and Veronica Mindrescu. New technology in education on performance analysis. wearable sensors utility on alpine skiing. *Revista Romaneasca pentru Educatie Multidimensionala*, 15(4):279–296, 2023.
- [Ettlinger *et al.*, 1995] Carl F Ettlinger, Robert J Johnson, and Jasper E Shealy. A method to help reduce the risk of serious knee sprains incurred in alpine skiing. *The American journal of sports medicine*, 23(5):531–537, 1995.
- [Hammerla *et al.*, 2016] Nils Y. Hammerla, Shane Halloran, and Thomas Plötz. Deep, convolutional, and recurrent models for human activity recognition using wearables. In *Proceedings of the Twenty-Fifth International Joint Conference on Artificial Intelligence (IJCAI-16)*, pages 1533–1540, 2016.
- [Hogan and Sternad, 2009] Neville Hogan and Dagmar Sternad. Sensitivity of smoothness measures to movement duration, amplitude, and arrests. *Journal of motor behavior*, 41(6):529–534, 2009.
- [Hu *et al.*, 2018] Jie Hu, Li Shen, and Gang Sun. Squeeze-and-excitation networks. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 7132–7141. IEEE/CVF, 2018.
- [Jones *et al.*, 2016] Michael Jones, Casey Walker, Zann Anderson, and Lawrence Thatcher. Automatic detection of alpine ski turns in sensor data. In *Proceedings of the 2016 ACM International Joint Conference on Pervasive and Ubiquitous Computing: Adjunct*, pages 856–860, 2016.
- [Jordan *et al.*, 2017] Matthew J Jordan, Per Aagaard, and Walter Herzog. Anterior cruciate ligament injury/reinjury in alpine ski racing: a narrative review. *Open access journal of sports medicine*, pages 71–83, 2017.
- [Kim and Kim, 2025] Sang-Un Kim and Joo-Yong Kim. Improving human activity recognition through 1D-ResNet: A wearable wristband for 14 workout movements. *Processes*, 13(1):207, 2025.
- [Larsen *et al.*, 2024] Aske G Larsen, Line Ø Sadolin, Trine R Thomsen, and Anderson S Oliveira. Accurate detection of gait events using neural networks and imu data mimicking real-world smartphone usage. *Computer Methods in Biomechanics and Biomedical Engineering*, pages 1–11, 2024.
- [Martínez *et al.*, 2019] Aaron Martínez, Rüdiger Jahnel, Michael Buchecker, Cory Snyder, Richard Brunauer, and Thomas Stöggel. Development of an automatic alpine skiing turn detection algorithm based on a simple sensor setup. *Sensors*, 19(4):902, 2019.
- [Martínez *et al.*, 2021] Aaron Martínez, Cory Snyder, Stephanie R Moore, and Thomas Stöggel. A comprehensive comparison and validation of published methods to detect turn switch during alpine skiing. *Sensors*, 21(7):2573, 2021.
- [Merity *et al.*, 2017] Stephen Merity, Nitish Shirish Keskar, and Richard Socher. Regularizing and optimizing lstm language models. *arXiv preprint arXiv:1708.02182*, 2017.
- [Ordóñez and Roggen, 2016] Francisco Javier Ordóñez and Daniel Roggen. Deep convolutional and LSTM recurrent neural networks for multimodal wearable activity recognition. *Sensors*, 16(1):115, 2016.
- [O’Kane *et al.*, 2025] S Maria O’Kane, Roisin McCafferty, Alison M Gallagher, Marie H Murphy, Ciara Fitzpatrick, and Angela Carlin. “sport is not a luxury.” the perceived impact of the cost-of-living crisis on sport and physical activity opportunities for children and adolescents in northern ireland (ni): A cross-sectional survey. *BMC Public Health*, 25(1):1675, 2025.
- [Paszke *et al.*, 2019] Adam Paszke, Sam Gross, Francisco Massa, Adam Lerer, James Bradbury, Gregory Chanan, Trevor Killeen, Zeming Lin, Natalia Gimelshein, Luca Antiga, Alban Desmaison, Andreas Kopf, Edward Yang, Zachary DeVito, Martin Raison, Alykhan Tejani, Sasank Chilamkurthy, Benoit Steiner, Lu Fang, Junjie Bai, and Soumith Chintala. Pytorch: An imperative style, high-performance deep learning library. In *Advances in Neural Information Processing Systems 32*, pages 8024–8035. Curran Associates, Inc., 2019.
- [Petrella *et al.*, 2025] Davide Petrella, Lynn Ellenberger, and Matthias Gilgien. Accuracy evaluation of smartphone-based gnss position and speed tracking for ski-slope and safety management. *PLoS One*, 20(8):e0327896, 2025.
- [Pujol *et al.*, 2007] Nicolas Pujol, Marie Philippe Rousseaux Blanchi, and Pierre Chambat. The incidence of anterior cruciate ligament injuries among competitive alpine skiers: a 25-year investigation. *The American journal of sports medicine*, 35(7):1070–1074, 2007.
- [Rashid and Louis, 2019] Khandakar M. Rashid and Joseph Louis. Window-warping: A time series data augmenta-

- tion of imu data for construction equipment activity identification. In *Proceedings of the 36th International Symposium on Automation and Robotics in Construction (ISARC 2019)*, volume 36, pages 651–657. International Association for Automation and Robotics in Construction (IAARC), 2019.
- [Richard *et al.*, 2023] Viviane Richard, Giovanni Piumatti, Nick Pullen, Elsa Lorthé, Idris Guessous, Nicola Cantoreggi, and Silvia Stringhini. Socioeconomic inequalities in sport participation: pattern per sport and time trends—a repeated cross-sectional study. *BMC public health*, 23(1):785, 2023.
- [Robak and Turek, 2025] Jakub Robak and Wojciech Turek. Turn detection in alpine skiing using smartphone sensors. In *International Conference on Computational Science*, pages 86–94. Springer, 2025.
- [Ruedl *et al.*, 2023] Gerhard Ruedl, Markus Posch, Elena Pocecco, Katja Tecklenburg, Birgit Schliernzauer, Michael D Kennedy, Martin Faulhaber, and Martin Burtcher. Association of personal and equipment-related factors on acl injury risk in alpine skiers with cautious or risk-taking behaviour: A case-control study. *AIMS public health*, 10(2):348, 2023.
- [Somerset and Hoare, 2018] Sarah Somerset and Derek J Hoare. Barriers to voluntary participation in sport for children: a systematic review. *BMC pediatrics*, 18(1):47, 2018.
- [Spörri *et al.*, 2016] Jörg Spörri, Josef Kröll, Matthias Gilgien, and Erich Müller. Sidecut radius and the mechanics of turning—equipment designed to reduce risk of severe traumatic knee injuries in alpine giant slalom ski racing. *British journal of sports medicine*, 50(1):14–19, 2016.
- [Syed *et al.*, 2020] Abbas Shah Syed, Zafi Sherhan Syed, Muhammad Shehram Shah, and Salahuddin Saddar. Using wearable sensors for human activity recognition in logistics: A comparison of different feature sets and machine learning algorithms. *International Journal of Advanced Computer Science and Applications (IJACSA)*, 11(9):644–649, 2020.
- [Tang *et al.*, 2024] Lyndon Tang, Mohammad Shushtari, and Arash Arami. Imu-based real-time estimation of gait phase using multi-resolution neural networks. *Sensors*, 24(8):2390, 2024.
- [Tello *et al.*, 2024] Andrés Tello, Victoria Degeler, and Alexander Lazovik. Too good to be true: accuracy overestimation in (re)current practices for human activity recognition. In *2024 IEEE International Conference on Pervasive Computing and Communications Workshops and Other Affiliated Events (PerCom Workshops)*, pages 511–517. IEEE, 2024.
- [Vanat, 2025] Laurent Vanat. 2025 International Report on Snow & Mountain Tourism. Laurent Vanat Consulting SARL, Geneva, 2025. Available at <https://www.vanat.ch/>.
- [Xia *et al.*, 2020] Kun Xia, Jianguang Huang, and Hanyu Wang. Lstm-cnn architecture for human activity recognition. *Ieee Access*, 8:56855–56866, 2020.
- [Yamagiwa *et al.*, 2014] Shinichi Yamagiwa, Hiroyuki Ohshima, and Kazuki Shirakawa. Development of skill scoring system for ski and snowboard. In *International Congress on Sports Science Research and Technology Support*, pages 1–15. Springer, 2014.
- [Zhang *et al.*, 2019] Wenting Zhang, Xiaohui Zhao, and Zan Li. A comprehensive study of smartphone-based indoor activity recognition via XGBoost. *IEEE Access*, 7:80027–80042, 2019.