

Automatische Langlauf Sub- Technik Klassifikation

Exposé zur Bachelorarbeit

Verfasser:

Louis Feuillet
Herisauer Strasse 86
9015 St.Gallen
Louis.feuillet@stud.fhgr.ch

Referent/in:

Prof. Dr. Martin Bünner

Korreferent/in:

Stefano Balestra

Modul:

Bachelor Thesis (cds-904) FS25

Organisation:

FHGR - Fachhochschule Graubünden

Eingereicht am: 13.06.2025

Inhaltsverzeichnis

Abbildungs-, Tabellen-, Abkürzungsverzeichnis	1—II
1 Einleitung	1
1.1 Forschungsfrage	1
2 Methodik	2
2.1 Forschungsstand	2
2.2 Datensatz	3
2.3 Datenvorverarbeitung	3
2.4 Modellwahl	4
2.5 Metriken	4
3 Vorläufige Gliederung	5
4 Zeitplan	5
5 Verwendete Literatur	6
6 Weiterführende Literatur	7

Abbildungs-, Tabellen-, Abkürzungsverzeichnis

- IMU** – Inertial Measurement Unit
- CNN** – Convolutional Neural Network
- LSTM** – Long Short-Term Memory
- GNSS** – Global Navigation Satellite System
- GPS** – Global Positioning System (Ortungssystem)
- Hz** – Hertz (Frequenzeinheit)
- ENU** – East-North-Up (geodätisches Koordinatensystem)

1 Einleitung

Skilanglauf ist eine Wintersportart, die sowohl physische Ausdauer als auch technische Präzision erfordert. Mit der zunehmenden Professionalisierung und Technisierung des Sports wächst der Anspruch, Bewegungsabläufe detailliert zu analysieren und gezielt zu optimieren. Einerseits müssen Athleten den Streckenverlauf genau kennen, um ihren Energieeinsatz strategisch anzupassen (Smith 1992). Andererseits ist eine präzise Kommunikation zwischen Trainer und Athlet notwendig, um im Training und Wettkampf gezielt auf bestimmte Bewegungsphasen reagieren zu können (Smith 1992).

Vor diesem Hintergrund wurden im Skilanglauf sogenannte Subtechniken eingeführt, welche standardisierte Fortbewegungsformen sind, die je nach Gelände, Geschwindigkeit und Technikstil variieren und auch als "Gears" bezeichnet werden (Pellegrini et al., 2021). Die korrekte Identifikation dieser Subtechniken bildet die Grundlage für technische Analysen und leistungsspezifisches Feedback.

Dank Fortschritten in der Sensortechnologie ist es mittlerweile möglich, Bewegungen präzise zu erfassen. Inertialmesseinheiten (IMU) (Rindal, et al. 2017) oder Gyroskope (Jang, et al. 2018) werden an Körperstellen wie Brust, Armen, Beinen oder Skiern angebracht und liefern hochfrequente Daten zur Bewegungserkennung in Echtzeit.

Zusätzlich ermöglicht maschinelles Lernen die automatische Klassifikation solcher Bewegungsmuster. Trainierte Modelle können anhand physikalischer Merkmale vorhersagen, welche Subtechnik aktuell ausgeführt wird (Pousibet-Garrido, et al. 2024). Dabei spielen sowohl die Wahl der Modellarchitektur als auch Art, Position und Qualität der Sensorik eine zentrale Rolle für die Klassifikationsgüte.

Im Gegensatz zu vielen bisherigen Studien, die auf IMUs mit 100–200 Hz oder Motion-Capture-Systeme mit hoher räumlich-zeitlicher Auflösung zurückgreifen, verwendet diese Arbeit ausschliesslich GPS-Daten mit einer Abtastrate von lediglich 10 Hz. Diese deutlich reduzierte zeitliche Auflösung stellt eine besondere methodische Herausforderung dar, da feinmotorische Bewegungsdetails im Rohsignal nicht direkt erkennbar sind. Um dennoch relevante Bewegungsmerkmale extrahieren zu können, wird eine mehrstufige Vorverarbeitung durchgeführt, bei der aus den GPS-Positionsdaten approximierte Beschleunigungen und zyklische Muster abgeleitet werden. Ziel ist es, trotz der limitierten Datenqualität ein robustes Modelltraining zur Subtechnikklassifikation zu ermöglichen.

Diese Arbeit untersucht daher, wie zuverlässig verschiedene Klassifikationsmodelle in der Lage sind, Subtechniken im Skilanglauf allein auf Basis von GPS-Daten vorherzusagen, die in der Form von geschätzten Inner-Body-Beschleunigungen vorliegen.

1.1 Forschungsfrage

«Welches der untersuchten Klassifikationsmodelle kann auf Basis von aus 10 Hz GPS-Daten abgeleiteten Inner-Body-Beschleunigungen die Subtechniken von Skilangläufer am zuverlässigsten vorhersagen? »

2 Methodik

2.1 Forschungsstand

Dies Literaturanalyse zeigt Ansätze zur Lösung der Forschungsfrage. Allerdings unterscheiden sich die Eckpunkte wie Sensorenart, Anzahl Probanden, Versuchsaufbau, Datenaufbereitung sowie die Modellwahl.

Gløersen & Gilgien (2021) untersuchten die Klassifikation von Skating-Subtechniken mittels GNSS-Sensordaten (50 Hz), welcher auf dem Helm angebracht wurde. Dabei wurden acht professionelle Skilangläufer analysiert, sechs auf einem Laufband zur Modellentwicklung und zwei auf einer Outdoor-Roller-Ski-Strecke zur Validierung. Es wurden fünf Sub-Techniken klassifiziert. Die Klassifikation erfolgte über ein Feed-Forward-Neuronales Netz sowie Entscheidungsregeln. Besonders G2–G4 konnten mit 92–97 % Genauigkeit erkannt werden; G5 hingegen nur mit 32 %. Die Studie nutzte Fourier-Analysen zur Merkmalerkennung.

Stöggl et al. (2014) verwendeten Smartphones mit integriertem Accelerometer (80 Hz), um Skating-Subtechniken von 11 professionellen Athleten auf einem Laufband zu klassifizieren. Es wurden fünf Subtechniken unterschieden. Die Klassifikation erfolgte über Markov-Ketten mit multivariaten Gaussverteilungen. Trainiert wurde das Modell sowohl für individuelle Probanden als auch kollektiv. Die Gesamtgenauigkeit lag bei 96 %, wobei Übergangsfehler zwischen den Sub-Techniken am häufigsten bei G3 und G4 auftraten.

Pousibet-Garrido et al. (2024) setzten IMUs (100–200 Hz) direkt auf den Skiern ein, um drei Skating-Subtechniken an einem künstlichen Anstieg von zwei professionelle Athleten zu klassifizieren. Die Klassifikation erfolgte mittels einer Kombination von Convolutional Neural Network (CNN) und Long Short-Term Memory- Network (LSTM), unter Verwendung eines Adam-Optimierers. Als Merkmale dienten Zeitfenster sowie Segmente aus den drei Achsen. Die Genauigkeit betrug 98 %, mit 90 % in der Einzelbewertung.

Uda et al. (2024) untersuchten vier Subtechniken Mithilfe eines kinematischen GNSS-Sensors (100 Hz) positioniert am Kopf, wobei zwei professionelle Athleten auf einer echten Langlauf Strecke analysiert wurden. Die Erkennung basierte auf der Auswertung von Wellenformen. Zur Korrektur von Höhen- und Kurveneinflüssen wurde ein gleitender Mittelwert sowie die Umwandlung in ENU-Koordinaten verwendet. Die Klassifikation erreichte eine Genauigkeit von 97,7%.

Meland (2017) entwickelte ein automatisiertes System zur Erkennung und Klassifikation von Bewegungszyklen im Langlauf unter Nutzung von IMUs (50–100 Hz) positioniert am Rücken. Drei Athleten wurden bei einem realen Lauf analysiert. Acht Subtechniken wurden mit Template-Matching, Cosine Similarity und Cluster-Analyse erkannt. Die Merkmalsextraktion erfolgte über Low-Pass-Filterung, Peakerkennung und Resampling. Die Klassifikationsgenauigkeit lag bei 94 %.

Takeda et al. (2019) verwendeten einen GNSS-Sensor (100 Hz), um klassische Sub-Techniken bei einem professionellen Athleten im Outdoor-Szenario zu klassifizieren. Zur

Signalbereinigung wurde ein gleitender Mittelwert der Höhendaten genutzt. Die Klassifikationsgenauigkeit betrug 98,6 %.

2.2 Datensatz

Der Datensatz umfasst GPS-Daten (10 Hz) von acht professionellen Langläuferinnen und Langläufern, die am Wettbewerb in Goms (Wallis) 2024 teilgenommen haben. Dabei trugen sie jeweils ein GPS-Messgerät auf dem Rücken. Die Gruppe bestand aus vier Männern und vier Frauen. Insgesamt wurden 13 Durchgänge aufgezeichnet. Einige dieser Durchgänge sind jedoch nicht vollständig gelabelt.

Beim Labeln der Zielvariable wurden für diesen Datensatz folgende Sub-Techniken ("Gears") definiert:

- Gear 2 Skating (**Asymmetrischer Schlittschuhschritt**): Hier erfolgt ein asymmetrischer Doppelstockeinsatz bei jedem zweiten Beinschritt, wobei ein Arm als Führungsarm agiert und der andere unterstützt.
- Gear 3 Skating (**1:1 Schlittschuhschritt**): In dieser Technik wird bei jedem Beinschritt ein symmetrischer Doppelstockeinsatz ausgeführt.
- Gear 4 Skating (**1:2 Schlittschuhschritt**): Diese Technik beinhaltet einen symmetrischen Doppelstockeinsatz bei jedem zweiten Beinschritt.
- Gear 5 Skating (**Schlittschuhschritt ohne Stockeinsatz**): Es werden ausschliesslich Beinschritte ausgeführt, ohne Einsatz der Stöcke.
- Gear 6 Skating (**Hocke**): Der Körper befindet sich in einer tiefen Position ohne Arm- und Beinbewegungen zur Reduktion des Luftwiderstands.
- Gear 4 Classic (**Doppelstockschub**): Diese klassische Technik nutzt einen symmetrischen Doppelstockeinsatz, während das Gleiten ohne aktiven Beinschub erfolgt.

Es ist zu beachten, dass nicht alle in Pellegrini et al. (2021) aufgeführten Subtechniken im vorliegenden Datensatz verwendet wurden. Dies lässt sich durch die spezifische Charakteristik der untersuchten Strecke erklären.

2.3 Datenvorverarbeitung

Zur Vorbereitung der zyklusbasierten Klassifikation von Skilanglauf-Subtechniken wurde eine mehrstufige Vorverarbeitung der GPS-Rohdaten entwickelt. Ein Bewegungszyklus entspricht dabei einer vollständigen Fortbewegungsphase im Skating-Stil, typischerweise zwischen zwei vergleichbaren biomechanischen Zuständen.

Die GPS-Koordinaten (x, y, z) werden zunächst in ein ebenes Koordinatensystem überführt und mithilfe eines Savitzky–Golay-Filters geglättet um die Schwerpunktbahn (x_{sp}, y_{sp}) zu erhalten. Geschwindigkeit $(v_{x_{sp}}, v_{y_{sp}})$ und Beschleunigung $(a_{x_{sp}}, a_{y_{sp}})$ dieser Bahn werden anschliessend über numerische Ableitungen mit einem tieferen Filterfenster berechnet. Zur Isolierung Inner-body Bewegungen (x_{ib}, y_{ib}) wird zusätzlich ein Savitzky–Golay-Filter direkt auf die Rohkoordinaten angewendet. Die Differenz zwischen diesen „inner-body“- und Schwerpunkt-Beschleunigungen ergibt ein Residualsignal $(a_{x_{res}}, a_{y_{res}})$, das feinere Bewegungsdynamiken abbildet. Die Richtungskomponenten wurden anschliessend mithilfe eines lokal berechneten Tangentenwinkels φ in eine vorwärts- und seitwärtsorientierte

Achse rotiert ($a_{forward}$, $a_{lateral}$). Für die Höheninformationen (z) wird auch die vertikale Beschleunigung ($a_{vertical}$) berechnet, allerdings ohne Rotation.

Die Segmentierung der Bewegung in Einzelzyklen erfolgt über die Detektion lokaler Maxima im lateralen Signal. Um zeitlich konsistente Eingaben für das Modelltraining zu ermöglichen, wurden alle gültigen Zyklen auf eine einheitliche Punktzahl interpoliert.

Zur Validierung der Datenvorverarbeitung wurden zusätzlich synthetische Testdaten erzeugt (Surrogate-Daten).

Insgesamt werden die folgenden Features für das Trainieren der Modelle verwendet.

- zyklusweise segmentierte Inner-Body-Beschleunigungssignale (3D),
- zyklusweise Geschwindigkeit,
- Zykluslänge
- Steigung der Strecke pro Zyklus

2.4 Modellwahl

Für die Klassifikation der Subtechniken im Skilanglauf werden verschiedene Ansätze aus dem Bereich des maschinellen Lernens in Betracht gezogen. Die Auswahl orientiert sich an bestehenden Arbeiten, in denen sowohl klassische Verfahren wie Entscheidungsbäume und Hidden-Markov-Modelle (Stöggl, et al. 2014) als auch moderne Deep Learning Modelle wie CNNs oder LSTMs (Pousibet-Garrido, et al. 2024) erfolgreich angewendet wurden.

Da der Fokus dieser Arbeit auf der Auswertung zyklisch strukturierter GPS-basierter Bewegungsdaten mit begrenzter zeitlicher Auflösung liegt, werden insbesondere solche Modelle berücksichtigt, die robuste Ergebnisse bei tiefen Datenfrequenz (10 Hz) liefern können. Ziel ist es, Klassifikationsmodelle zu identifizieren, die sich durch eine gute Generalisierbarkeit, Effizienz und Stabilität gegenüber Störsignalen auszeichnen.

Die finale Auswahl der Modelle erfolgt im Rahmen einer empirischen Untersuchung, bei der unterschiedliche Ansätze anhand von Genauigkeit, Robustheit und Interpretierbarkeit miteinander verglichen werden.

2.5 Metriken

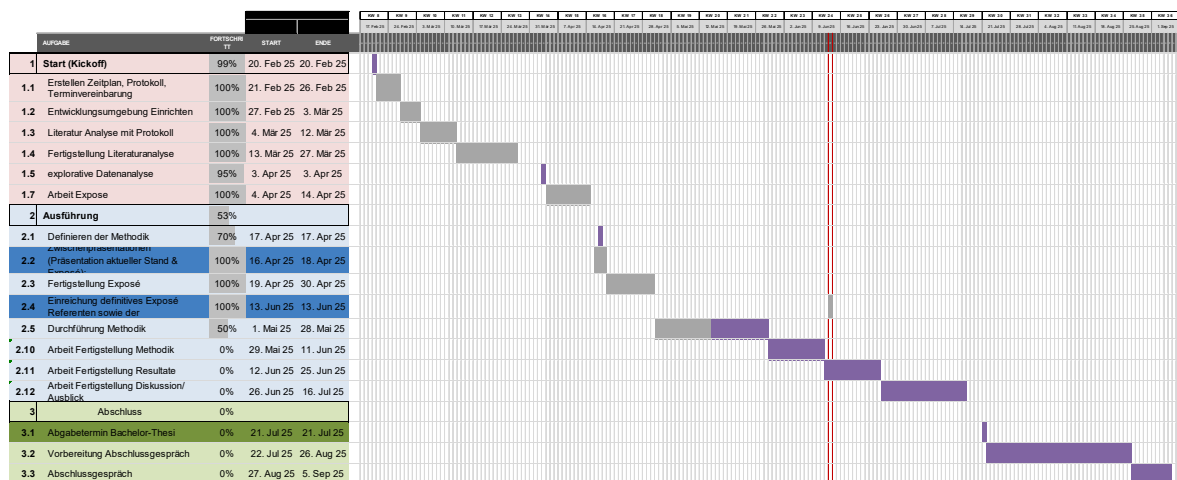
Zur Bewertung der Modellgüte wird eine Konfusionsmatrix verwendet, die die vorhergesagten Klassenzuordnungen des Modells den tatsächlichen Klassen gegenüberstellt. Jede Zeile der Matrix entspricht einer tatsächlichen Klasse, während die Spalten die vom Modell vorhergesagten Klassen darstellen. Korrekt klassifizierte Fälle befinden sich dabei auf der Diagonale (True Positives pro Klasse). Fehlklassifikationen sind außerhalb der Diagonale angeordnet. So lassen sich wichtigsten Kennwerte für jede Subtechnik lassen sich aus den Einträgen der Matrix direkt ableiten (Stöggl, et al. 2014).

Da es sich um einen unausgeglichene (imbalanced) Datensatz handelt, wird zusätzlich der F1-Score ($\frac{2 * \text{True Positiv}}{2 * \text{True Positiv} + \text{False Positiv} + \text{False Negativ}}$) verwendet (Géron 2018).

3 Vorläufige Gliederung

- 1 Einleitung
 - 1.1 Forschungsstand
 - 1.2 Forschungsfrage
- 2 Methodik
 - 2.2 Sensor Spezifikation
 - 2.2 Datensatz
 - 2.3 Datenvorverarbeitung
 - 2.4 Beschrieb des verwendeten Klassifikators
 - 2.5 Metrik
- 3 Resultate
- 4 Diskussion
- 5 Fazit

4 Zeitplan



5 Verwendete Literatur

- Géron, Aurélien . *Praxiseinstieg Machine Learning mit Scikit-Learn und TensorFlow*. O'Reilly, 2018.
- Gløersen, Øyvind , und Matthias Gilgien. «Classification of Cross-Country Ski Skating Sub-Technique Can Be Automated Using Carrier-Phase Differential GNSS.» Norway: MDPI, 2021.
- Jang, Jihyeok, et al. «A Unified Deep-Learning Model for Classifying the Cross-Country Skiing Techniquea Using Wearable Gyroscope Sensors.» *Sensor*. Korea: Sensors, November 2018.
- Meland, Henrik J. «Automated detection and classification of movement cycles in cross-country skiing through analysis of inertial sensor data.» Oslo, Norway: Norwegian University of Science and Technology, 2017.
- Pellegrini, Barbara, et al. «Methodological Guidelines Designed to Improve the Quality of Research on Cross-Country Skiing.» *Journal of Science in Sport and Exercise*. springer, 2021.
- Pousibet-Garrido, Antonio , et al. «Gear Classification in Skating Cross-Country Skiing Using Inertial Sensors and Deep Learning.» *Sensor*. University of Granada: Sensors, 2024.
- Rindal, Ole Marius Hoel, Trine M. Seeberg, Johannes Tjønnås, Pål Haugnes, und Øyvind Sandbakk. «Automatic Classification of Sub-Techniques in classical Cross-Country Skiing using a Machine Learning Algorithm on Micro-Sensor Data.» *Sensor*. Norwegian University of Science and Technology: Sensors, 2017.
- Smith, Gerald A. «Biomechanical analysis of cross-country skiing techniques.» *Medicine and Science in Sports and Exercise*. Oregon State University: NIH, 1992.
- Stöggl, Thomas , et al. «Automatic Classification of the Sub-Techniques (Gears) Used in Cross-Country Ski Skating Employing a Mobile Phone.» University of Salzburg: Sensors, 2014.
- Takeda, Masaki, et al. «Cross-Country Skiing Analysis and Ski Technique Detection by High-Precision Kinematic Global Navigation Satellite System.» Kyoto, Japan: Sensors, 2019.
- Uda, Shunya, et al. «Cross-Country Ski Skating Style Sub-Technique Detection and Skiing Characteristic Analysis on Snow Using High-Precision GNSS.» Basel, Switzerland: Sensors, 2024.

6 Weiterführende Literatur

- Johansson, M., Korneliusson, M., & Lizbat Lawrence, N. (2019). Identifying cross country skiing techniques using power meters in ski poles. Gothenburg, Sweden.
- Marsland, F., Lyons, K., Anson, J., Waddington, G., Macintosh, C., & Chapman, D. (2012). Identification of Cross-Country Skiing Movement Patterns Using Micro-Sensors Using Micro-Sensors. University of Canberra, Australia: Sensors.
- Polo-Rodríguez, A., Escobedo, P., Martínez-Martí, F., Marcen-Cinca, N., Carvajal, M., Medina-Quero, J., & Martínez-García, M. (2025). A Comparative Study of Plantar Pressure and Inertial Sensors for Cross-Country Ski Classification Using Deep Learning. Granada, Spain: Sensors.
- Rassem, A., El-Beltagy, M., & Saleh, M. (2017). Cross-Country Skiing Gears Classification using Deep Learning. Cornell University .
- Seeberg, T., Tjønnås, J., Rindal, O., Haugnes, P., Dalgard, S., & Sandbakk, Ø. (2017). A multi-sensor system for automatic analysis of classical crosscountry skiing techniques. online: Springer.