

Maschinelles Lernen für die Situationserkennung einer aktiven Kniegelenkorthese

Machine Learning for Situation Detection of an Active Knee-Ankle-Foot-Orthosis

Philipp Mattfeldt, Bastian Latsch, Jürgen Hielscher, Sascha Desch, Niklas Schäfer, Romal Chadda, Roland Werthschützky, André Seyfarth, Mario Kupnik

Technische Universität Darmstadt, Fachgebiet Mess- und Sensortechnik, 64283 Darmstadt, Deutschland, bastian.latsch@tu-darmstadt.de

Kurzfassung

Um den Menschen bedarfsgerecht und ergonomisch unterstützen zu können, müssen Assistenzsysteme die Bewegungen und Bewegungsabsichten des Nutzers frühzeitig erkennen. Die Individualität der menschlichen Bewegung ist dabei eine besondere Herausforderung. Dieser Beitrag zeigt, dass maschinelles Lernen geeignet ist, um aus den Daten ausgewählter, am Körper getragener Sensoren sieben alltägliche Bewegungssituationen zu unterscheiden. Ein Convolutional Neural Network erkennt Merkmale in den Bewegungsdaten, welche nicht nur auf die aktuelle Situation schließen lassen, sondern auch als Frühindikatoren dienen, aus denen eine kommende Bewegung antizipiert werden kann. Dazu ermittelt das Netzwerk für jede vordefinierte Situation eine Plausibilität im laufenden Betrieb basierend auf den Sensorsignalen des Assistenzsystems. Das hier beschriebene Modell wurde mit Bewegungsdaten von sechs Personen trainiert, wodurch eine Erkennungsgenauigkeit von 99,1 % erreicht wurde. Die Robustheit der Erkennung ist mit der Einführung einer Konfidenz gegeben und somit ein Einsatz der Situationserkennung an einer aktiven Kniegelenkorthese möglich.

Abstract

In order to support humans ergonomically and in line with their needs, assistance systems have to recognize the user's movements and movement intentions at an early stage. Here, the individuality of human movement is a key challenge. This paper shows that machine learning is suitable to distinguish seven common movement situations from the data of selected sensors attached to the body. A Convolutional Neural Network detects features in the motion data that not only indicate the current situation but also provide indicators from which an upcoming motion can be anticipated. For this purpose, the network determines a plausibility for each predefined situation in the ongoing operation based on the measured values of the assistance systems. The model described was trained with the motion data of six participants resulting in a recognition accuracy of 99,1 %. The robustness of the recognition is given with the aid of a confidence, and thus an application of the situation recognition in an active knee orthosis is possible.

1 Einleitung

Der demografische Wandel und der damit einhergehende steigende Anteil älterer Menschen stellt die Gesellschaft vor eine große Herausforderung. Vor allem ältere oder körperlich eingeschränkte Menschen verfügen oft nur über eine eingeschränkte Mobilität [1]. Dabei können einfache Bewegungen wie Treppensteigen oder das Aufstehen von einem Stuhl zur Belastung werden. In einem solchen Fall helfen Assistenzsysteme [2].

In vorheriger Arbeit [3] wurde eine aktive Kniegelenkorthese entwickelt, die das Aufstehen und Treppensteigen unterstützt (Abb. 1a). Durch die Orthese wird gerade der aufgrund der Muskelschwäche fehlende Teil des Drehmoments im Knie kompensiert. Je nach Bewegungssituation und Bedürfnis des Anwenders sind allerdings unterschiedliche Unterstützungsfaktoren notwendig. Um das Zuschalten der Unterstützung im richtigen Zeitpunkt und mit angepasstem Drehmoment zu gewährleisten, ist die Erkennung der aktuellen Bewegungssituation im laufenden Betrieb essenziell. Für die Erkennung und Analyse von Bewegungsabläufen

werden in der Sportwissenschaft und Biomechanik mehrere bewegungsspezifische Messgrößen am menschlichen Körper zeitkontinuierlich erfasst. Hierbei wird zwischen kinematischen Messgrößen, wie Segmentbeschleunigung und Gelenkwinkel, physiologischen Messgrößen, wie der Muskelaktivität, und dynamometrischen Messgrößen, wie den Bodenreaktionskräften und dem Körperschwerpunkt (center of pressure, CoP), unterschieden. Mithilfe solcher Messgrößen ist eine Klassifizierung der aktuellen Bewegungssituation möglich [4].

Der Anspruch dabei ist eine individuell auf den Nutzer zugeschnittene und optimierte Verarbeitung. Dadurch wird eine über die Nutzungsdauer andauernde Flexibilität erreicht, die auf Änderungen im Bewegungsverhalten konstruktiv eingehen kann. Ein verbreiteter Ansatz zur Klassifizierung, der die genannten Punkte erfüllt, ist Machine Learning. Hierbei werden mittlerweile vermehrt Ansätze des Deep Learning für die Erkennung menschlicher Aktivitäten (human activity recognition) eingesetzt [5]. In diesem Zusammenhang kommt besonders häufig die Klasse der Convolutional Neural Networks (CNNs) zum Einsatz [5]. Beispiels-

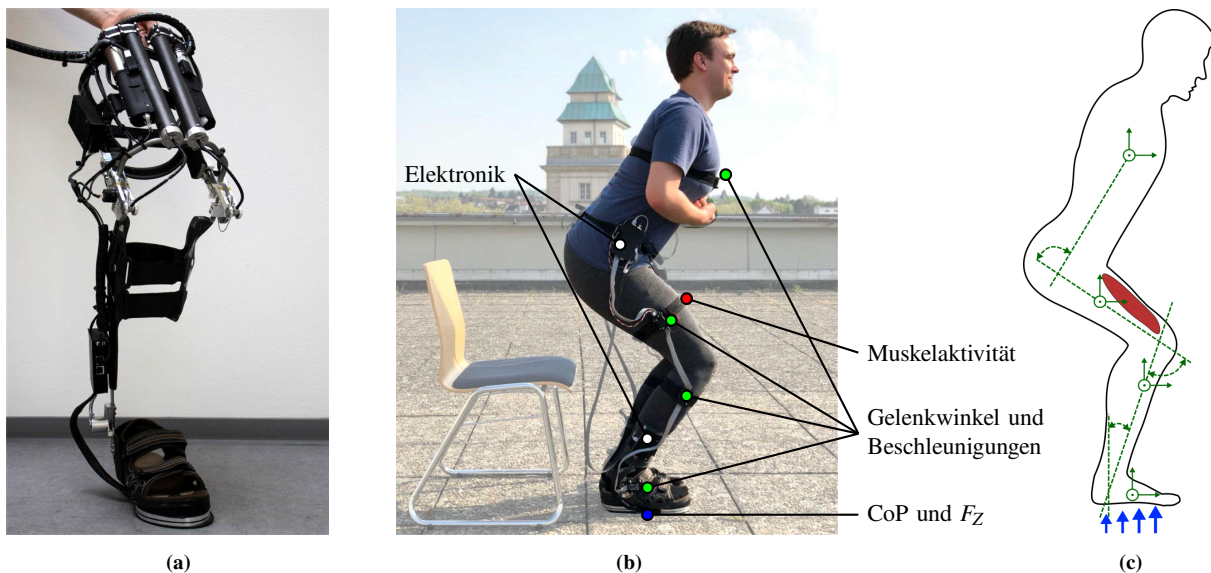


Abbildung 1 (a) Labormuster der aktiven Kniegelenkorthese nach [3] mit Kraftsensoren im Fußteil und Erfassung der Gelenkwinkel in Knie und Sprunggelenk. Die Aktorik ist am Oberschenkelteil befestigt. (b) Die Umsetzung der Sensorik-Orthese beinhaltet die zu erfassenden kinematischen Größen Gelenkwinkel sowie lineare Beschleunigungen und Rotationsgeschwindigkeiten (grün), die physiologische Messgröße Muskelaktivität der Muskelbündel am Oberschenkel (rot) sowie die dynamometrischen Messgrößen Körperschwerpunkt (CoP) und Bodenreaktionskräfte F_Z (blau). (c) Schematische Skizze der erfassten Messgrößen an der Sensorik-Orthese.

weise werden basierend auf den Gyroskop- und Beschleunigungsdaten eines Smartphones sechs verschiedene Aktivitäten (Gehen, Treppauf, Treppab, Sitzen, Stehen, Liegen) klassifiziert [6]. Viele dieser Ansätze berichten bereits hohe Erkennungsgenauigkeiten. Allerdings sind sie oftmals nicht für den Einsatz im laufenden Betrieb geeignet oder werden nicht für die Erkennung von Übergangssituationen, wie Aufstehen und Hinsetzen, ausgelegt. Beide Aspekte sind für die Anwendung bei einer aktiven Kniegelenkorthese jedoch essenziell. Ein weiterer wesentlicher Aspekt ist, dass unter realen Bedingungen auch unbekannte Situationen auftreten, welche dann oftmals trotzdem eindeutig einer bekannten Situation zugeordnet werden. Dabei gibt das Modell in der Regel keinen Aufschluss darüber, in welchem Maß der Erkennung vertraut werden kann.

In dieser Arbeit wird ein Modell zur Situationserkennung für die Anwendung an einer aktiven Kniegelenkorthese vorgestellt, das die zuvor genannten Herausforderungen adressiert. Hierzu wird für die Generierung von Trainings- und Testdaten zunächst eine passive Orthese mit Sensorik ausgestattet (Sensorik-Orthese). Das trainierte mehrschichtige CNN mit abschließendem Fully-Connected-Layer wird evaluiert. Hierbei wird untersucht, wie die Berücksichtigung personalisierter Trainingsdaten die Modellgüte beeinflusst. Darüber hinaus wird ein Gütemaß eingeführt, welches das Vertrauen quantifiziert, das einer Erkennung des Modells entgegengebracht werden kann.

2 Sensorik-Orthese

Das Ziel der Sensorik-Orthese ist, eine umfangreiche Erfassung verschiedener Messgrößen zu ermöglichen, ohne die Probanden in ihrer Bewegungsfreiheit einzuschrän-

ken. Hierfür wird bei der Entwicklung auf die Verwendung möglichst leichter und flexibler Materialien geachtet. So werden beispielsweise die Sprung-, Knie- und Hüftgelenkwinkel berührungslos aus den Orientierungen inertialer Messeinheiten (IMUs) ermittelt, welche jeweils mit elastischen Manschetten an Oberkörper, Oberschenkel, Unterschenkel und am Fuß appliziert sind (Abb. 1). Zusätzlich werden über die IMUs (BNO055, Bosch Sensortec, Reutlingen, Deutschland) die linearen Beschleunigungen sowie die Rotationsgeschwindigkeiten dieser Segmente hinsichtlich der drei Raumrichtungen erfasst. Die Detektion der Muskelaktivität des Kniestreckers erfolgt über Elektromyografie (EMG) mit Sensoren (MyoWare Muscle Sensor, SparkFun Electronics, Boulder, Colorado, USA) an den Oberschenkeln und Biegesensoren (Flex-Sensor 4.5", Spectra Symbol, Salt Lake City, Utah, USA), die in der Oberschenkel-Manschette integriert sind und so die Kontraktion der Oberschenkel-Muskulatur detektieren. Die Bodenreaktionskräfte sowie der Körperschwerpunkt werden mittels acht Wägezellen (FX1091, Measurement Specialties Inc., Hampton, Virginia, USA) in den Schuhen ermittelt. Die Sensorik-Orthese übermittelt die sich so ergebenden 64 Messgrößen mit einer Abtastrate von 1 kHz an einen Rechner zur weiteren Verarbeitung.

Mit der Sensorik-Orthese führten sechs Probanden unterschiedlichen Alters (39 ± 15 Jahre), Geschlechts (vier männlich, zwei weiblich), Größe ($1,78 \text{ m} \pm 0,09 \text{ m}$) und Gewichts ($76 \text{ kg} \pm 12 \text{ kg}$) jeweils 30 Minuten lang Bewegungen aus den sieben Alltagsbewegungen durch. Dabei wurde den Probanden keine Einschränkung hinsichtlich Reihenfolge oder auch Ausführungsart gegeben, um den Datensatz mit Diversität zu füllen und den Anteil der einzelnen Situationen (Tab. 1) nicht vorherzubestimmen.

Tabelle 1 Die Verteilung der repräsentierten Situationen im gesamten dreistündigen Datensatz berechnet sich aus ihrer empirischen Beobachtung und dem zeitlichen Anteil.

Situation	Anteil h_i	Umfang im Datensatz
Gehen \rightarrow	25 %	ca. 4500 Schritte
Stehen \uparrow	15 %	ca. 40 Minuten
Sitzen \downarrow	15 %	ca. 40 Minuten
Treppauf \uparrow	12,5 %	} jeweils ca. 2100 Schritte
Treppab \downarrow	12,5 %	
Aufstehen \nearrow	10 %	} jeweils ca. 700 Vorgänge
Hinsetzen \searrow	10 %	

Für den hier verwendeten Ansatz mit überwachtem Lernen ist es essenziell, dass zu jedem Zeitpunkt im Datensatz das Label (sog. Ground Truth), also die korrekte Situation zu den erfassten Messgrößen, vorliegt. Das Labeln der Daten erfolgt in der Literatur meist manuell während der Erfassung oder im Nachhinein manuell unter Zuhilfenahme synchronisierter Videoaufnahmen [7] und stellt damit einen erheblichen Arbeitsaufwand dar. Zur Erleichterung dieses Prozesses wurde basierend auf einem Zustandsautomaten ein „Auto-Labeler“ entwickelt, der einen mehrere Sekunden umfassenden Zeitraum überblicken kann. Durch dieses zusätzlich verfügbare Wissen gegenüber einer Echtzeit-Verarbeitung kann robust mittels boolescher Logik zwischen den sieben Alltagssituationen unterschieden werden. Die Daten können so nach der Aufnahme automatisiert, optional mit geringfügig manueller Anpassung, für die spätere Nutzung aufbereitet werden. Abschließend werden 80 % der gelabelten Daten eines Probanden dem Trainingsdatensatz und die restlichen 20 % dem Testdatensatz zugeteilt.

3 Netzarchitektur der Situationserkennung

Für das Modell zur Situationserkennung wird ein künstliches neuronales Netz mit mehreren Schichten so genannter Convolutional-Layer und zwei finalen Fully-Connected-Layern gewählt. Die Inputs sind zeitliche Fenster T_F von zwei Sekunden Länge (Abb. 2). Die Convolutional-Layer extrahieren dabei zeitliche und signal-übergreifende Merkmale (Features) aus den Sensorsignalen. Die abschließenden Fully-Connected-Layer lernen schließlich aus der Kombination der jeweiligen Merkmale die Plausibilitäten für jede Situation. Entgegen vieler bisheriger Forschungen [6], [8] wird der Output-Layer nicht via Softmax-Funktion als One-Hot-Encoded gestaltet, sondern über eine lineare Aktivierungsfunktion als Regression. Dem aktuellen Trend der „explainable AI“ [9] folgend wird dieser Ansatz im Kontext der Orthese zur Steigerung der Nutzersicherheit gewählt. Er ermöglicht es, durch dem Modell nachgeschaltete Operationen wie die Bewertung einer Konfidenz, in die Entscheidung des Modells einzugreifen.

Das Modell lernt während des Trainings die Plausibilität einer jeden Situation als Funktion der Sensorsignale. Das Ler-

nen dieser Funktion entspricht einer Optimierung der Gewichte, welche die freien Parameter der einzelnen Schichten sind. Im Falle der Convolutional-Layer entsprechen diese Parameter den Filtern, mit denen das Eingangssignal gefaltet wird, um das Ausgangssignal zu berechnen. Bei dem Fully-Connected-Layer sind die Gewichte die Einträge in einer Matrix, mit der das Eingangssignal multipliziert wird, um die Ausgänge zu erhalten. Um mit diesem Aufbau, der im Wesentlichen aus zwei linearen Operationen (Faltung und Matrix-Multiplikation) besteht, nichtlineare Funktionen wie die Situationserkennung der Orthese zu erlernen, wird zwischen den einzelnen Schichten eine Nichtlinearität eingebracht. Hierfür wird die ReLU-Funktion (Rectified Linear Unit) aufgrund ihrer schnellen Lernkonvergenz in Mehrschichtmodellen und der weiten Verbreitung in der Klassifizierung [10] genutzt.

4 F_1 -Maß und Konfidenz zur Evaluierung der Modellgüte

Zur Evaluierung von Machine-Learning-Klassifizierern wird in der Regel eine Bewertung nach dem F_1 -Maß mit k -fold cross-validation (hier $k = 3$) durchgeführt [11]. Das F_1 -Maß ist ein Gütemaß, das die zwei Gütekriterien *Präzision* und *Trefferquote* gleichmäßig gewichtet über alle N Klassen (hier $N = 7$) kombiniert. Die Präzision einer Klasse gibt an, ob eine positive Entscheidung des Modells korrekt war, wenn eine Klasse als positiv gelabelt war. Die Präzision ist

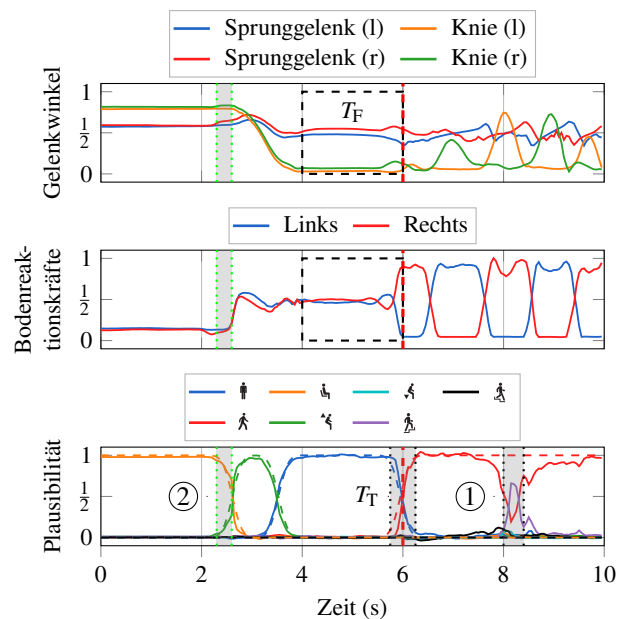


Abbildung 2 Die Sensordaten (Fensterbreite $T_F = 2$ s) fließen zum aktuellsten Zeitpunkt (rot gestrichelt) in die Prognose des Modells ein. Aufgrund der natürlichen Unschärfe beim Situationsübergang wird ein Toleranzbereich $T_T = \pm 250$ ms definiert, in dem eine Transition des Modells als richtig bewertet wird. ① Dabei kann es auch zu einer Fehleinschätzung mit anschließender Korrektur zwischen Plausibilität (durchgezogen) und Ground Truth (gestrichelt) kommen. ② Das Modell erkennt Frühindikatoren für einen sich ankündigenden Situationswechsel.

somit das Verhältnis aus den Wahr-Positiv-Entscheidungen zur Summe aus Wahr-Positiv (WP) und Falsch-Positiv (FP). Die Trefferquote einer Klasse gibt an, ob das Modell eine Klasse als positiv erkannt hat, wenn sie dies in Wahrheit war. Die Trefferquote ist also das Verhältnis aus Wahr-Positiv-Entscheidungen zur Summe aus Wahr-Positiv und Falsch-Negativ (FN).

$$\text{Präzision} = \frac{WP}{WP + FP}, \quad \text{Trefferquote} = \frac{WP}{WP + FN} \quad (1)$$

$$F_1 = \sum_i^N 2 \cdot h_i \cdot \frac{\text{Präzision}_i \cdot \text{Trefferquote}_i}{\text{Präzision}_i + \text{Trefferquote}_i} \quad (2)$$

Zur Evaluierung der Modelle im laufenden Betrieb wird der Testdatensatz kontinuierlich (alle 10 ms) hinsichtlich der Situationserkennung ausgewertet. Für die Berechnung des F_1 -Maßes wird an jedem dieser Zeitpunkte die dominante Situation als Argmax-Funktion der Plausibilitäten bestimmt. Eine Herausforderung stellt die Definition des exakten Übergangzeitpunkts zwischen zwei Situationen, so dass nicht immer eine scharfe Trennung möglich ist. Dies ist auch beim manuellen Labeln schwierig, da teilweise bei den Situationswechseln keine eindeutige Trennlinie existiert. Um einer Verfälschung der Modellgüte durch diese Unschärfe vorzubeugen, muss das Modell innerhalb eines Toleranzfensters $T_T = \pm 250\text{ms}$ den Situationswechsel erkennen, damit diese Entscheidung als richtig gewertet wird (Abb. 2).

Zur Steigerung der Nutzersicherheit ist das Überprüfen der getroffenen Entscheidung erforderlich. Hierzu wird das Modell um eine Konfidenz erweitert, die dem neuronalen Netz nachgeschaltet ist. Diese stellt ein Gütemaß dar, welche anhand eigens definierter Kriterien die Plausibilitäten des Modells in den vergangenen 200 ms bewertet. Zu diesen Kriterien zählen die relativen Differenzen aller Plausibilitäten und das Rauschen der Plausibilitäten. Ersteres bedeutet im Falle einer unsicheren Entscheidung des Modells, dass mehrere Situationen eine Plausibilität größer als Null besitzen. Letzteres bedeutet im Falle einer unsicheren Entscheidung des Modells, dass die Plausibilitäten stärker rauschen und die Transitionen zwischen zwei Situationen sehr steil sind. Beide Kriterien werden über Gewichtungsfaktoren miteinander kombiniert und stellen die Konfidenz dar. Unterschreitet die Konfidenz einen definierten Schwellenwert (z. B. 50 %), kann dies als Trigger für einen übergeordneten Sicherheitsmodus genutzt werden.

5 Experimente

Das vorgestellte Modell wird mit den Trainingsdaten trainiert und anhand von Testdaten, die das Modell zuvor nicht gesehen hat, evaluiert. Dabei ist für den Einsatz solcher Situationserkennung vor allem die Robustheit des Modells relevant und wird daher in zwei Fällen untersucht. Im ersten Fall ist die universelle Einsetzbarkeit des Modells von Interesse. Dies wird durch eine allgemeine und breite Anwendung der Kniegelenkorthese motiviert und erfordert, dass das Modell auf die Daten eines unbekanntem Anwenders adaptierbar ist. Hierzu werden in einer Versuchsreihe meh-

rere Modellvarianten mit gleichen Parametern, jedoch mit den Daten variierender Probanden, trainiert. Für die Evaluierung wird jeder Proband gegen alle Modelle getestet, so dass folgende Szenarien entstehen:

- **Personalisiert:** Test eines jeden Probanden mit dem Modell aus seinen individuellen Trainingsdaten;
- **Exklusiv:** Test eines jeden Probanden mit Modellen, die ausschließlich mit Daten anderer Probanden trainiert wurden, wobei die Anzahl der Probanden im Training erhöht wird;
- **Alle:** Test eines jeden Probanden mit einem universellen Modell, welches mit den Trainingsdaten aller Probanden trainiert wurde.

Im zweiten Fall zur Untersuchung der Robustheit wird das Verhalten des Modells bei unbekanntem Situationen geprüft. Dies ist ein wesentlicher Sicherheitsaspekt für den Nutzer der Kniegelenkorthese. Im Alltag können neben den sieben hier behandelten Situationen weitere Situationen und besondere Bewegungen, wie beispielsweise das Stürzen des Nutzers, vorkommen. Auch in diesen Fällen ist die Robustheit der Situationserkennung für die Nutzersicherheit unabdingbar. Dafür wird das Modell Situationen ausgesetzt, die während des Trainings in dieser Form nie aufgetreten sind.

6 Ergebnisse & Diskussion

Das Experiment zum universellen Einsatz des Modells zeigt, dass die Genauigkeit mit steigender Anzahl an Probanden stetig ansteigt (Abb. 3). Probandenanzahl „1“ bis „5“ bilden die Gruppe *Exklusiv*. Eine Probandenanzahl von eins bedeutet, dass jeder Proband gegen die personalisierten Modelle jedes anderen Probanden getestet wurde. Bei sechs verfügbaren Probanden, wird also jeder Proband gegen die fünf personalisierten Modelle der anderen Probanden getestet. Bei einem Modell der Probandenanzahl „2“ enthalten die Modelle entsprechend alle möglichen Kombinationen von jeweils zwei Probanden, jedoch immer exklusive des zu testenden Probanden. Eine Probandenanzahl von sechs entspricht somit *Alle*. Neben der steigenden Genauigkeit sinkt gleichzeitig die Streuung zwischen den einzelnen Modellen einer Probandenanzahl. Dies wird vor allem beim Testen gegen die personalisierten Modelle der anderen Probanden „1“ deutlich. Hier weist das F_1 -Maß teilweise eine starke Streuung von bis zu 17 % auf (Proband P3 und P6). Es wird allerdings auch deutlich, dass vereinzelte Probanden bereits gut durch personalisierte Modelle anderer Probanden repräsentierbar sind (Probanden P1 und P5). Dies weist darauf hin, dass sich die Bewegungsmuster dieser Probanden ähneln. Dennoch lassen sich personalisierte Modelle für einen universellen Einsatz optimieren. Wird die Anzahl der Probanden im Trainingsdatensatz erhöht, steigt die durchschnittliche Genauigkeit von 93,9 % bei einer Probandenanzahl von eins auf 98,2 % bei einem Modell mit fünf Probanden an. Werden zusätzlich Daten des zu testenden Probanden in den Trainingsdatensatz mit aufgenommen, steigt die Genauigkeit nochmals um 0,9 Prozentpunkte auf 99,1 %

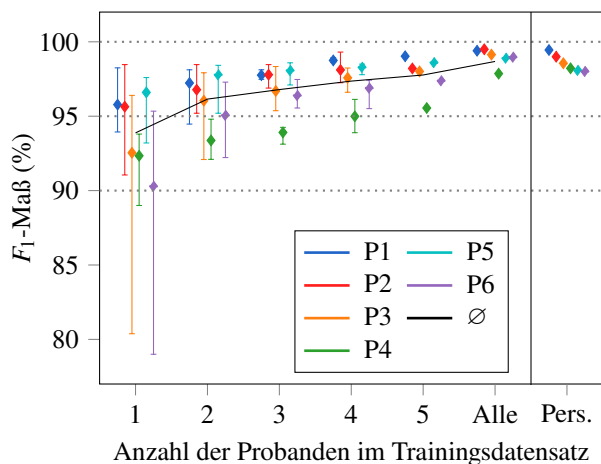


Abbildung 3 Beim Vergleich der Modellgenauigkeit mit unterschiedlicher Anzahl an Probanden im Trainingsdatensatz wird gegen mehrere Modelle bestehend aus den Daten von insgesamt sechs Probanden getestet. Ein farbiger Punkt markiert den Mittelwert eines Probanden (P1 bis P6) für diese Anzahl an Probanden, mit denen das Modell trainiert wurde. Der zugehörige Fehlerbalken beschreibt den Bereich der Modellgenauigkeit, der beim Testen gegen die Modelle dieser Probandenanzahl erreicht wird. Sowohl die Modellgenauigkeit als auch die Streuung der Genauigkeit verbessern sich mit steigender Probandenanzahl und übersteigen teilweise sogar die des personalisierten Modells (Pers.).

an (Alle). Dieses universelle Modell mit den Daten aller verfügbaren Probanden ist selbst im Vergleich zu den personalisierten Modellen eines jeden Probanden mit ihrer durchschnittlichen Genauigkeit von 98,6 % um 0,5 Prozentpunkte genauer. Dabei erfolgt die Erkennung mit einer durchschnittlichen Konfidenz von 99,4 % und einer Laufzeit von 5 ms pro Berechnung (GeForce GT440, NVidia, Santa Clara, USA). In der Konfusionsmatrix (Tab. 2) wird deutlich, dass nur in wenigen bestimmten Kombinationen die vom universellen Modell ermittelte dominante Plausibilität von der Ground Truth abweicht. Eine nähere Betrachtung der Abweichungen zeigt, dass in nahezu allen Fällen eine Abweichung aus der Unschärfe der Ground Truth resultiert. Diese existierende Unschärfe ist jedoch grundsätzlich nicht negativ, sondern bietet dem Modell die Möglichkeit, Frühindikatoren für einen kommenden Situationswechsel in Signalverläufen zu erkennen. So interpretiert beispielsweise das universelle Modell das Heranziehen der Füße und das Vorbeugen des Oberkörpers im ¶ als Vorankündigung für ¶. Die Plausibilität steigt hier bereits auf ca. 30-40 % an (Abb. 2, ②). Durch diese Frühindikation des Modells ist die Orthese schneller bereit dem Anwender das unterstützende Drehmoment hinzuschalten, sobald dieser mit dem Belasten der Füße das ausschlaggebende Signal zur ¶ Bewegung gibt.

Bei unbekanntem Situationswechsel wie beispielsweise dem *Einbeinstand mit Knie- und Fußwackeln* (Abb. 4, ①) erkennt das universelle Modell dies als wechselndes ¶. Obwohl in diesem Fall keine eindeutige Ground Truth besteht, ist die Erkennung des universellen Modells nachvollziehbar und verständlich. Das initiale In-die-Knie-Gehen beim *In-der-*

Tabelle 2 Der Konfusionsmatrix des universellen Modells mit einem F_1 -Maß von 99,1 % und einer Konfidenz von 99,4 % zufolge werden alle Situationen bis auf wenige Kombinationen richtig erkannt. Abweichungen sind in so gut wie allen Fällen auf die Unschärfe der Ground Truth zurückzuführen.

		Modellentscheidung						
		Stehen	Gehen	Sitzen	Aufstehen	Hinsetzen	Treppauf	Treppab
Ground Truth	Stehen	98,89	1,11	0	0	0	0	0
	Gehen	0,55	98,84	0	0	0,08	0,21	0,32
	Sitzen	0	0	99,81	0,19	0	0	0
	Aufstehen	0	0,27	0	99,73	0	0	0
	Hinsetzen	0	0	0	0	100	0	0
	Treppauf	0	0,88	0	0	0	99,12	0
	Treppab	0	0,6	0	0	0	0	99,4

Hocke-Wippen ② wird souverän als ¶ verbucht. Nachfolgend wechselt die Prognose wieder zurück zu ¶. Somit erkennt das universelle Modell nachvollziehbar diese Pose eher als Stehen statt Sitzen. Beim *Auf-Zehenspitzen-Stehen* ③ werden keine merklichen Impulse ausgelöst, von der Entscheidung ¶ abzuweichen. Die letzte Testsituation, das *dreimalige Springen* ④, ist hingegen eine Herausforderung für das universelle Modell. Das initiale In-die-Knie-Gehen wird noch verständlicherweise als ¶ erkannt. Die Landung resultiert jedoch in konfusen Plausibilitäten, aus denen keine dominante Situation abgeleitet werden kann. In diesem Fall greift die absichernde Konfidenz, sodass das Modell in den übergeordneten Sicherheitsmodus gelangt und die Sicherheit für den Nutzer weiterhin gegeben ist.

7 Schlussfolgerung

Der vorgestellte Ansatz zeigt, dass die Situationserkennung mittels neuronaler Netze basierend auf der Messung von Bewegungsdaten möglich ist. Dabei hat sich eine größere Menge an Trainingsdaten als Vorteil gezeigt, um ein universelles Modell zu erstellen, welches durchschnittlich besser abschneidet als ein personalisiertes Modell. Ein Grund dafür sind Bewegungen, die ein Proband anders als üblich ausführt. Eine korrekte Erkennung wird in solchen Fällen durch die Diversität aller Probanden in einem universellen Modell möglich. Mit Einführung der Konfidenz sind unbekannte Situationen erkennbar und somit erhöht sich die Nutzersicherheit. In weiteren Untersuchungen wird ein größerer Trainingsdatensatz erhoben, um die Anwendung des universellen Modells bei einer erhöhten Probandenanzahl zu bestätigen. Damit lässt sich die aktive Kniegelenkorthese um ein universelles Modell zur Situationserkennung erweitern, um den Nutzer dem Bedarfsfall entsprechend zu unterstützen.

Danksagung

Diese Forschung wurde unterstützt von der Deutschen Forschungsgemeinschaft unter den Projektnummern 164569068 und 402740893.

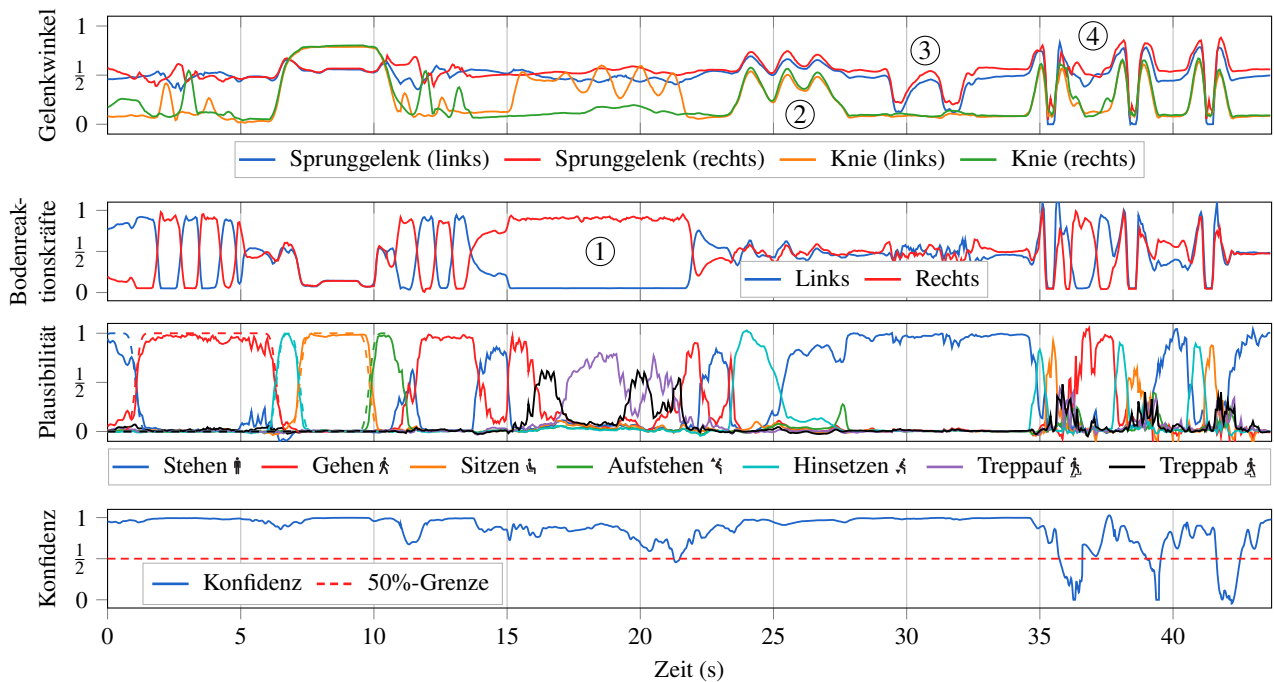


Abbildung 4 Aus den am Probanden gemessenen Sensorsignalen, wie Gelenkwinkel und Bodenreaktionskräfte, berechnet das Modell die Plausibilitäten für die aktuelle Bewegungssituation und gibt die Konfidenz der Berechnung an (normierte Verläufe). Bei angerlegten, also bekannten, Situationen (bis ca. 15 s) sind die Plausibilitäten der einzelnen Situationen (durchgezogen) nahezu gleich mit der vorliegenden Ground Truth (gestrichelt). Anhand von dem Modell unbekanntes lässt sich die Stärke unseres Ansatzes erkennen (ab 15 s), da in der Realität auch eine manuelle Einschätzung selten eindeutig ist. ① Beim Einbeinstand mit Knie- und Fußwackeln schwankt die Erkennung zwischen \uparrow und \downarrow , was nachvollziehbar der Situation am nächsten kommt. Da keine Plausibilität dominiert, sinkt die Konfidenz. ② Beim Wippen in der Hocke wird das In-die-Knie-Gehen korrekterweise als \downarrow erkannt. ③ Das Stehen auf Zehenspitzen besitzt wenig Einfluss auf die \uparrow Entscheidung. ④ Die Klassifizierung \downarrow beim Schwungholen zum dreimaligen Springen ist gefolgt von konfusen Plausibilitäten, woraufhin die Konfidenz absinkt.

Literatur

- [1] S. R. Lord, R. D. Clark und I. W. Webster, „Postural Stability and Associated Physiological Factors in a Population of Aged Persons,“ en, *Journal of Gerontology*, Jg. 46, Nr. 3, S. M69–M76, Mai 1991.
- [2] L. Zhang, G. Liu, B. Han u. a., „Assistive devices of human knee joint: A review,“ en, *Robotics and Autonomous Systems*, Jg. 125, S. 103–139, März 2020.
- [3] P. P. Pott, S. I. Wolf, J. Block u. a., „Knee-ankle-foot orthosis with powered knee for support in the elderly,“ en, *Proceedings of the Institution of Mechanical Engineers, Part H: Journal of Engineering in Medicine*, Jg. 231, Nr. 8, S. 715–727, Aug. 2017.
- [4] A. Gollhofer und E. Müller, Hrsg., *Handbuch Sportbiomechanik*, Ser. Beiträge zur Lehre und Forschung im Sport 171. Schorndorf: Hofmann-Verlag, 2009.
- [5] J. Wang, Y. Chen, S. Hao u. a., „Deep learning for sensor-based activity recognition: A survey,“ en, *Pattern Recognition Letters*, Jg. 119, S. 3–11, März 2019.
- [6] C. A. Ronao und S.-B. Cho, „Human activity recognition with smartphone sensors using deep learning neural networks,“ en, *Expert Systems with Applications*, Jg. 59, S. 235–244, Okt. 2016.
- [7] A. Bulling, U. Blanke und B. Schiele, „A tutorial on human activity recognition using body-worn inertial sensors,“ en, *ACM Computing Surveys*, Jg. 46, Nr. 3, S. 1–33, Jan. 2014.
- [8] B. Fang, Q. Zhou, F. Sun u. a., „Gait Neural Network for Human-Exoskeleton Interaction,“ *Frontiers in Neurobotics*, Jg. 14, S. 58, Okt. 2020.
- [9] A. Adadi und M. Berrada, „Peeking Inside the Black-Box: A Survey on Explainable Artificial Intelligence (XAI),“ *IEEE Access*, Jg. 6, S. 52 138–52 160, 2018.
- [10] K. Hara, D. Saito und H. Shouno, „Analysis of function of rectified linear unit used in deep learning,“ in *2015 International Joint Conference on Neural Networks (IJCNN)*, Killarney, Ireland: IEEE, Juli 2015, S. 1–8.
- [11] D. M. W. Powers, „Evaluation: from precision, recall and F-measure to ROC, informedness, markedness and correlation,“ *International Journal of Machine Learning Technology*, 2011.