

Aktivitätserkennung mit Wearables

Abstract—Aktivitätserkennung und Wearable Computing sind Forschungsgebiete in den Bereichen Pervasive und Ubiquitous Computing. Durch die steigende Anzahl von Smartphones und Wearables, welche verschiedene Sensoren enthalten, wird die Aktivitätserkennung mit Wearables möglich. Neben dem alltäglichen Gebrauch wird die Aktivitätserkennung in Bereichen wie Gesundheitswesen, Militär, Gefahrenabwehr und Arbeitssicherheit angewandt. In dieser Arbeit werden die Mechanismen der Aktivitätserkennung durch Wearables und die Aktivitätserkennungssysteme präsentiert. Maschinelles Lernen und seine Anwendung in der Aktivitätserkennung wird auf einem Beispiel mit Hidden-Markow-Modellen näher untersucht. Daneben werden auch die verwendeten Sensoren und die Taxonomie der Aktivitätserkennung betrachtet.

Schlüsselwörter—Aktivitätserkennung; Wearables; Maschinelles Lernen; Aktivitätserkennungssysteme; Hidden-Markow-Modell

I. EINFÜHRUNG

Heutzutage gibt es 2.8 Mrd. Smartphones in der Welt und es wird erwartet, dass es bis 2021 noch 925 Mio. anderen Arten von Wearables geben wird [1, 2]. Diese Geräte werden immer leistungstärker und haben viele Sensoren mit welchen sie Daten sammeln können [3]. Auf dem Markt gibt es Wearables wie Fitness-Tracker und Smartwatches, die von den Herstellern wie Fitbit¹, Apple² und Samsung³ produziert sind. Sie sind auf die Messung von Daten ausgerichtet, die für das Training und gesundes Leben wichtig sind. Beispielsweise können sie die Schrittzahl und den Kalorienverbrauch des Benutzers messen. Außerdem gibt es auch API von Google⁴. Sie ermöglicht den Anwendungen auf den Android Systemen auf die Aktivitäten des Benutzers zu reagieren. Die Schwierigkeit der Aktivitätserkennung zeigt sich darin, dass diese API zurzeit nur zwischen acht Aktivitäten unterscheiden kann: im Fahrzeug, auf dem Fahrrad, zu Fuß, Rennen, Spazierengehen, Kippen, Unbekannt und Bewegungslos.

Aktivitätserkennung hat ebenfalls viele Anwendungen und kann beispielsweise in den Bereichen wie Altenpflege und Rehabilitation oder im Militärbereich und Gefahrenabwehr angewandt werden. Durch die Anwendung in der Altenpflege und Rehabilitation kann die Qualität des Lebens von den älteren Menschen verbessert werden [4]. Die Aktivitätserkennungssysteme in diesem Anwendungsgebiet können beispielsweise erkennen, wenn eine Person stürzt, und dann die Pfleger benachrichtigen [5]. Im Unterschied zu den anderen Anwendungsbereichen ist im Gefahrenabwehr und Militärbereich

Aufdringlichkeit der verwendeten Systeme weniger wichtig. Das Militärpersonal ist daran gewöhnt, mit viel Ausrüstung zu arbeiten, und deswegen ist die Verwendung aufdringlicheren Systemen in diesem Anwendungsbereich möglich. Andererseits müssen die verwendeten Sensoren wegen der besonderen Umstände widerstandsfähiger sein [6]. Sie müssen z. B. waserdicht und stoßfest sein. Aktivitätserkennungssysteme für die Gefahrenabwehr wurden im Rahmen der europäischen Forschungsprojekte wie SAFESENS [7] und Euripides Intex [8] untersucht und entwickelt. Im Rahmen von Euripides Intex wurde eine intelligente Jacke entwickelt, welche die Aktivitäten von Feuerwehrleuten erkennt. Falls die Feuerwehrleute in Gefahr sind, aktiviert das System den Alarm in der Feuerwehrzentrale. Des Weiteren, dass es nicht unbedingt um die Erkennung von menschlichen Aktivitäten gehen muss, zeigen auch die Forschungsprojekte, in denen die Wearables eingesetzt wurden, um Tiere zu verfolgen und so ihre Gewohnheiten besser zu erkennen [9, 10].

In dieser Arbeit beschäftigen wir uns mit der Aktivitätserkennung mit Wearables. Im Kapitel II wird die Taxonomie der Aktivitätserkennung betrachtet. Im Kapitel III werden die Sensoren sowie die Unterteilungen der Aktivitäten dargestellt. Im Kapitel IV werden Mechanismen zur Aktivitätserkennung betrachtet, insbesondere im Bezug auf das maschinelle Lernen, welches hier näher betrachtet wird. Verschiedene Strategien des maschinellen Lernens werden ebenso angesehen. Kapitel V betrachtet das Verfahren der Aktivitätserkennung mit maschinellem Lernen näher. Kapitel VI befasst sich mit dem FE-AT Ansatz, welcher zwei existierende Ansätze kombiniert. Im Kapitel VII werden einige Aktivitätserkennungssysteme dargestellt und miteinander verglichen. Am Ende wird im Fazit alles zusammengefasst.

II. TAXONOMIE DER AKTIVITÄTSERKENNUNG

Zuerst wird nach den Arten der Sensoren, welche für die Erkennung der Aktivitäten verwendet werden, zwischen (1) sensorbasierter und (2) visionbasierter Aktivitätserkennung unterschieden [11]. Die visionbasierte Aktivitätserkennung verwendet Videokameras als Sensoren und erkennt die Aktivitäten aus den Videos mit Hilfe von Methoden aus dem Bereich der Bilderkennung. Bei dieser Art der Aktivitätserkennung sind aber immer die Bedenken hinsichtlich der Privatsphäre vorhanden [12]. Ferner, problematisch ist es auch, dass die Methoden für Videoverarbeitung viel Rechenleistung brauchen.

Sensorbasierte Aktivitätserkennung benutzt auf der anderen Seite die Sensoren, welche die Daten in einer Zeitreihe darstellen (z. B. Beschleunigungssensor misst die Beschleunigung in x-, y- und z-Achse mit einer bestimmten Abtastgeschwindigkeit). Sie wird dann in die Externe und die Aktivitätserkennung mit Wearables untergliedert [13]. In der externen Aktivitätserkennung werden die Sensoren auf Objekte befestigt. Die Akti-

¹<https://www.fitbit.com/de/home>

²<https://www.apple.com/lae/watch/>

³<http://www.samsung.com/us/mobile/wearables/>

⁴<https://developers.google.com/android/reference/com/google/android/gms/location/DetectedActivity>

vitäten, welche im Zusammenhang mit diesen Objekten sind, werden dann erkannt. Diese Art der Aktivitätserkennung kann deswegen auch die komplexeren Aktivitäten erkennen, z. B. wenn die Duschkabine die Sensoren enthalten würde, könnte die Aktivität 'Duschen' einfacher erkannt werden. Ein weiteres Beispiel hierfür sind die Aktivitätsüberwachungssysteme, die in einem Haus installiert werden können [14]. Das Problem dieser Art der Erkennung ist, dass die Aktivitäten nur in der Reichweite von Sensoren erkannt werden können. Die Aktivitätserkennung mit Wearables löst dieses Problem jedoch, weil die Sensoren dann immer mit der Person sind. Allerdings sind die objektabhängigen Aktivitäten schwieriger mit Wearables zu erkennen (z. B. ist es schwer zu erkennen, ob jemand einen Kühlschrank öffnet, wenn alle Sensoren an der Person selbst – in Form von Wearables – angebracht sind). Schließlich existiert noch ein hybrider Ansatz, der die externe Aktivitätserkennung mit der gleichzeitigen Nutzung von Wearables kombiniert [13].

III. AKTIVITÄTEN UND SENSOREN

A. Aktivitäten

In der Literatur findet man verschiedene Unterteilungen der Aktivitäten. O. D. Lara und M. A. Labrador unterteilen die Aktivitäten in sieben Aktivitätstypen: Bewegungsaktivitäten, Transportation, Handynutzung, tägliche Aktivitäten, Militär-, Oberkörper- und Übungsaktivitäten [12]. Auf der anderen Seite, M. Cornacchia et al. [15] unterteilen die Aktivitäten in (1) die Aktivitäten mit globaler Bewegung und (2) die Aktivitäten mit lokaler Interaktion. Die Aktivitäten mit globaler Bewegung sind diejenigen, bei denen sich der ganze Körper bewegt. Bus fahren und Spazierengehen sind Beispiele dafür. Bei den Aktivitäten mit lokaler Interaktion hingegen interagiert man mit einigen Objekten, wie z.B. beim Zähneputzen. Diese Unterteilungen sind nützlich, weil sie uns beispielsweise erlauben Sensortypen oder Positionen der Sensoren in den Zusammenhang mit Aktivitätstypen zu bringen. Bestimmte Sensortypen sind für eine spezifische Art von Aktivitäten besser geeignet (z. B. eignen sich GPS-Sensoren mehr für (1) und Beschleunigungssensoren für (2)). Sensorpositionen werden ebenfalls in der Abhängigkeit von den erkannten Aktivitäten gewählt (z. B. eignet sich die Gürtellinie für (1) und das Handgelenk für (2)). Letztlich bauen sie eine Struktur auf, in welche wir die Aktivitätserkennungssysteme einordnen können.

B. Sensoren

In der Aktivitätserkennung benutzte Sensoren sind zahlreich. Einer der am häufigsten benutzten Sensortypen ist der Beschleunigungssensor [16]. Er misst die Beschleunigung in x-, y- und z-Achse. In einer Studie in [17, 41] wurde gezeigt, dass die Obergrenze für die Genauigkeit der Erkennung von zehn unterschiedlichen Aktivitäten nur mit den Beschleunigungssensoren 97,3 % ist. Daneben wurde bemerkt, dass die Addition vom Barometer zu den Beschleunigungssensoren die Genauigkeit für die Erkennung der Aktivitäten 'Treppenabsteigen' und 'Treppensteigen' um 20 % vergrößert [17, 18, 41].

Neben dem Beschleunigungssensor werden oft Gyroskop und Magnetometer eingesetzt [19]. Gyroskop misst die Win-

kelgeschwindigkeit. Es verbessert die Genauigkeit der Winkelbestimmung und axialer Beschleunigung. Magnetometer misst die Magnetfelder und kann deswegen wie Kompass funktionieren. Im Vergleich zum Beschleunigungssensor erkennt das Gyroskop die Aktivitäten 'Treppensteigen' und 'Treppenabsteigen' besser, während der Beschleunigungssensor 'Stehen' und 'Sitzen' besser erkennt [40].

Andere häufig verwendete Sensoren sind Umgebungssensoren und der GPS-Sensor. Zu den Umgebungssensoren gehören Mikrophone, Lichtsensoren und Thermometer. GPS in der Kombination mit den Beschleunigungssensoren ermöglicht die Erkennung von globalen Aktivitäten sowie die Erkennung vom Transporttyp [13, 42]. Oft findet man ebenfalls die Sensoren, welche die physiologische Signale erkennen, wie beispielsweise die Sensoren, welche die Herzfrequenz oder Atemfrequenz erkennen können. Es wurde jedoch gezeigt, dass diese Sensoren einige Aktivitäten nur mit Verzögerung erkennen können, weil beispielsweise die Herzfrequenz noch einige Zeit nach dem 'Rennen' erhöht bleibt [34].

IV. MECHANISMEN ZUR AKTIVITÄT SERKENNUNG

Lara und Labrador definieren formal das Aktivitätserkennungsproblem wie folgt [13]:

Definition 1: Gegeben ist die Menge $M = \{M_0, \dots, M_k\}$, wobei M_i einem Attribut entsprechende Zeitreihe darstellt, welche in einem Zeitintervall $I = [t_\alpha, t_\omega]$ definiert ist. Gesucht ist eine temporale Partition $\langle I_0, \dots, I_{r-1} \rangle$ von I . Sie ist basiert auf den Daten in M und auf den Beschriftungen (in einer Menge), die in jedem Intervall I_j ausgeführte Aktivität darstellen. Die Zeitintervalle I_j sind konsekutiv, nicht leer, nicht überlappend und es gilt $\bigcup_{j=0}^{r-1} I_j = I$.

Dieses Problem ist deterministisch nicht lösbar, weil die Anzahl der möglichen Kombinationen von Attributen und Aktivitäten sehr groß werden kann [13]. Deshalb wird die Aktivitätserkennung mit Wearables meistens mit Hilfe vom maschinellen Lernen durchgeführt. Maschinelles Lernen ist ein Teilgebiet der künstlichen Intelligenz, das mit den anderen Bereichen wie insbesondere Statistik, aber auch Informatik, Mathematik und Physik verbunden ist, und sich mit der Konstruktion und Untersuchung von lernfähigen Systemen und Algorithmen, welche aus Daten lernen können, befasst [20, 21].

Maschinelles Lernen wird nach den Lernstrategien [13], welche verwendet werden, in das überwachte, unüberwachte und teilüberwachte Lernen unterteilt. Im überwachten Lernen bestehen die Trainingsdaten aus den Features und entsprechenden Aktivitäten. Im unüberwachten Lernen werden in den Trainingsdaten nur die Features enthalten und im teilüberwachten Lernen ist nur eine begrenzte Anzahl von Features mit ihnen entsprechenden Aktivitäten beschriftet. Da es sehr schwer ist, die Aktivitäten ohne wenigstens einer kleinen Anzahl von beschrifteten Features zu erkennen, benutzen die Aktivitätserkennungssysteme das überwachte und teilüberwachte Lernen [13].

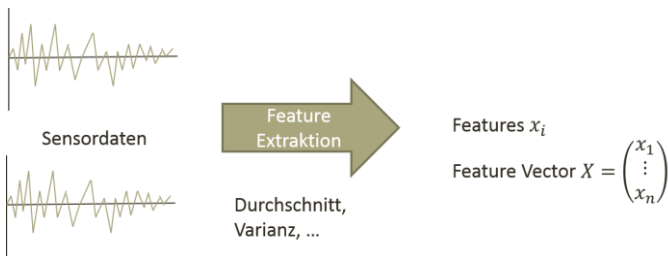


Abb. 1. Funktionsweise der Feature-Extraktion.

In der Aktivitätserkennung benutzte Ansätze werden in die wissensbasierten und datengesteuerten Ansätze unterteilt [11]. Da die datengesteuerten Ansätze auf die Sensordaten konzentriert sind, werden sie auch bottom-up Ansätze genannt. Zu den datengesteuerten Ansätzen gehören beispielsweise die generativen Methoden, die ein Wahrscheinlichkeitsmodell aus den Daten aufbauen und versuchen die Aktivitäten mit Hilfe der Wahrscheinlichkeiten im Modell zu erkennen. Die bekanntesten generativen Ansätze sind diejenigen, die Hidden-Markov-Modelle benutzen. Andererseits werden die wissensbasierte Ansätze top-down Ansätze genannt, weil sie schon existierendes Wissen in einer Domäne benutzen und damit das Modell top-down aufbauen [11, 22]. Sie brauchen jedoch die Experten, welche die Daten annotieren können. Ein weiterer Nachteil von diesen Modellen ist, dass sie statisch sind und sich daher mühsam an unterschiedliche Situationen anpassen lassen.

V. VERFAHREN DER AKTIVITÄTSEKKNUNG MIT MASCHINELLEM LERNEN

Das Verfahren der Aktivitätserkennung mit dem maschinellen Lernen besteht aus einer Trainings- und einer Testphase [13, 17]. Am Anfang der beiden Phasen wird das Verfahren der Feature-Extraktion vorgenommen.

A. Feature-Extraktion

In der Feature-Extraktion werden die Features aus den rohen Sensordaten extrahiert. Man kann es als eine Art von der Dimensionsreduktion ansehen. Features sind die Eigenschaften der Daten, mit welchen in der Trainingsphase das Modell trainiert wird [44]. Sie werden mit den Feature-Extraktion-Methoden wie Durchschnitt und Varianz extrahiert (Abb. 1). Die Sensoren arbeiten mit einer Abtastfrequenz und sammeln die Daten in den Zeitfenstern. Aus den Daten in diesen Zeitfenstern werden mit den gewählten Feature-Extraktion-Methoden die Features erstellt. Normalerweise gibt es mehrere Features pro Zeitfenster, deren Anzahl und Qualität von der Art und Anzahl der verwendeten Sensoren und Feature-Extraktion-Methoden abhängen. Sie werden dann mit Hilfe von einem Feature-Vektor dargestellt, der zu einem Zeitfenster entspricht. Wenn man beispielsweise den Durchschnitt als Feature-Extraktion-Methode wählt und diese auf Daten des Beschleunigungssensors anwendet, dann enthält aus einem Zeitfenster extrahierter Feature-Vektor die Werte für durchschnittliche Beschleunigung in der x-, y- und z-Achse in diesem Zeitfenster.

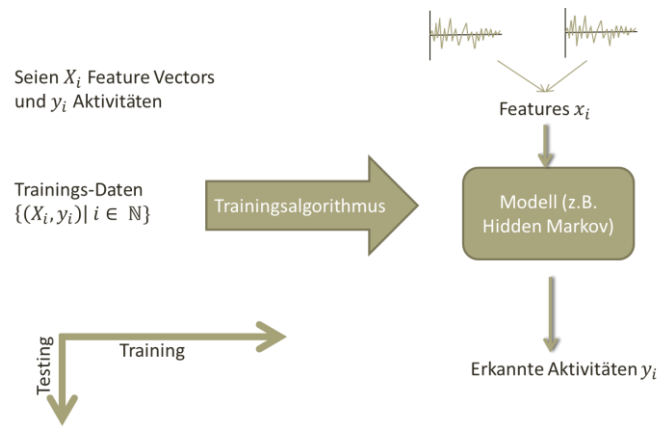


Abb. 2. Funktionsweise der Trainings- und Testphasen.

Ein wichtiger Punkt bei der Feature-Extraktion ist die Auswahl der Feature-Extraktion-Methode. Wang et al. zeigen, dass diese Auswahl die Genauigkeit der Aktivitätserkennung stark beeinflussen kann [17, 23]. Sie haben gezeigt, dass bei Anwendung der Standardabweichung als Feature-Extraktion-Methode auf die Beschleunigungssensordaten mit einer Genauigkeit von 99,4 % zwischen 'Bewegung' und 'nicht Bewegung' unterschieden werden kann. Des Weiteren können in [24] und [25] auch deutliche Unterschiede in Abhängigkeit von der Auswahl der Features bemerkt werden. Aus diesem Grund wird in der Aktivitätserkennung für die Auswahl der Features normalerweise die MRMR-Methode (engl. Minimum Redundancy and Maximum Relevance) verwendet. Sie hat als Kriterium für die maximale Relevanz die Größe der Unterschiede zwischen den Features und als Kriterium für die minimale Redundanz, dass die Features nicht die gleichen Informationen enthalten [13, 26]. Beispielsweise, wenn man die Varianz als Feature-Extraktion-Methode benutzt, bringt daneben noch die Nutzung der Standardabweichung keine neuen Informationen. Außerdem, wenn weniger Features extrahiert werden müssen, wird auch der benötigte Rechenbedarf verringert [13]. Die empirischen Ergebnisse können letztendlich auch widersprüchlich sein. Huynh und Schiele haben festgestellt, dass die Varianz angewandt auf die Daten der Beschleunigungssensoren höhere Genauigkeit als Durchschnitt liefert [25], während Pirttikangas, Fujinami und Nakajima entgegengesetzte Ergebnisse bekommen haben [24].

Bezüglich der Frequenz, mit welcher die Sensoren arbeiten, wurde erkannt, dass diese in der Abhängigkeit von den erkannten Aktivitäten gewählt werden soll. Zum Beispiel brauchen die komplexeren Aktivitäten längere Zeitfenster. In [27] wurde herausgefunden, dass die Bewegungsaktivitäten mit den Beschleunigungssensoren mit einer Frequenz von 15 bis 20 Hz am besten erkannt werden können und dass die Vergrößerung der Frequenz keine Vorteile bringt, sondern nur den Energieverbrauch negativ beeinflusst.

B. Trainingsphase

Nach der Feature-Extraktion werden die extrahierten Features zusammen mit den entsprechenden Aktivitäten als Train-

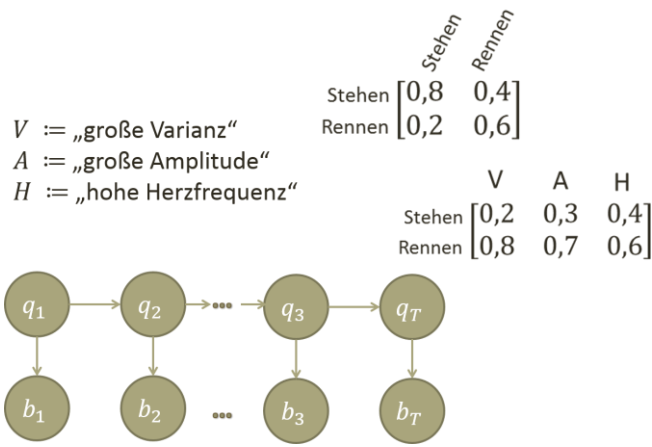


Abb. 3. Graphische Darstellung eines Hidden-Markow- Modells

ningsdaten zur Verfügung gestellt [13]. Die Trainingsdaten, die in der Trainingsphase verwendet werden, hängen von der verwendeten Strategie des maschinellen Lernens ab. Sie enthalten nur Feature-Vektoren (im unüberwachten Lernen) oder Feature-Vektoren und entsprechende Aktivitäten (im überwachten Lernen). An die Trainingsdaten wird ein Trainingsalgorithmus angewandt, mit welchem das Modell aus den Daten kreiert wird (Abb. 2). Der verwendete Trainingsalgorithmus und das erzeugte Modell hängen vom Typ des Klassifikationsalgorithmus ab, welches bei der Aktivitätserkennung benutzt wird.

C. Testphase

In der Abb. 2 wird die Testphase dargestellt. Zuerst werden die Feature-Vektoren mit Feature-Extraktion extrahiert [13]. Danach wird mit einem Algorithmus und dem in Trainingsphase trainierten Modell berechnet, welche Aktivität am wahrscheinlichsten dem Feature-Vektor entspricht. Diese Aktivität wird als erkannte Aktivität zurückgegeben.

D. Funktionsweise auf einem Beispiel mit Hidden-Markow-Modell

Hidden-Markow-Modell (HMM) besteht aus den Zuständen q_t und den Beobachtungen b_t (zur Zeit t) sowie den Übergangs- und Beobachtungsmatrizen [33]. Wir definieren in (1) zuerst die kürzere Schreibweise für die Wahrscheinlichkeiten. Sei Z_t die Zufallsvariable für einen Zustand zur Zeit t und sei B_t die Zufallsvariable für eine Beobachtung zur Zeit t .

$$P(Z_t = q_t \wedge B_t = b_t) = P(q_t \wedge b_t) \quad (1)$$

Hidden-Markow-Modell heißt "hidden" (dt. versteckt), weil die Zustände nicht bekannt sind. Es ist ein Markow-Modell, weil es die Markow-Eigenschaft erfüllt [43]. Die Markow-Eigenschaft (2) besagt, dass jeder Zustand nur von dem vorherigen Zustand abhängt. Die früheren Zustände haben keinen Einfluss auf den jetzigen Zustand.

$$P(q_t | q_1 \wedge \dots \wedge q_{t-1}) = P(q_t | q_{t-1}) \quad (2)$$

Daneben sind auch die Beobachtungen b_t voneinander unabhängig (3).

$$P(b_t | q_1 \wedge \dots \wedge q_t \wedge \dots \wedge q_T \wedge b_1 \wedge \dots \wedge b_t \wedge \dots, b_T) = P(b_t | q_t) \quad (3)$$

Eigenschaften (2) und (3) eines HMM kann man auch aus seiner graphischen Darstellung in der Abb. 3 auslesen. Die Pfeile zwischen den Zuständen und Beobachtungen stellen die Abhängigkeiten zwischen ihnen dar. Nach der Feature-Extraktion stehen die Trainingsdaten zur Verfügung. Hier wird angenommen, dass die Feature-Vektoren mit den entsprechenden Aktivitäten beschriftet sind. An diese Daten wird der Baum-Welch-Algorithmus angewandt, der als Trainingsalgorithmus im Fall vom HMM verwendet wird. Baum-Welch-Algorithmus ist eine besondere Art vom Expectation-Maximization-Algorithmus, der aus den Trainingsdaten das HMM aufbaut [32]. Expectation-Maximization-Algorithmus besteht aus dem E(xpectation)-Schritt und dem M(aximization)-Schritt. Im E-Schritt schließt er von den gegebenen Parametern auf die unbekannten Werte. Im Anschluss werden im M-Schritt die Parameter optimiert [45]. In jedem E-Schritt dieses Algorithmus werden mit der Forward-Backward-Inferenz die Wahrscheinlichkeiten für die unbekannten Zustände unter der Bedingung, dass einige Beobachtungen beobachtet wurden, gefunden. Das heißt, die $P(q_t | B)$ und $P(q_t \wedge q_{t+1} | B)$ werden bestimmt. Im Folgenden wird die Forward-Backward-Inferenz beschrieben.

Forward-Backward-Inferenz [28]: Sei q_t definiert als der Zustand q zur Zeit t und sei b_t definiert als die Beobachtung b zur Zeit t . Sei $Q = q_1 \wedge q_2 \wedge \dots \wedge q_T$ und sei $B = b_1 \wedge b_2 \wedge \dots \wedge b_T$. Die graphische Darstellung vom HMM in der Abb. 3 stellt dann die multivariate Verteilung $P(Q \wedge B)$ über alle Zustände und Beobachtungen dar. Die Pfeile repräsentieren dabei die bedingten Wahrscheinlichkeiten in (4).

$$P(Q \wedge B) = P(q_1) \left(\prod_{t=1}^{T-1} P(q_{t+1} | q_t) \right) \left(\prod_{t=1}^T P(b_t | q_t) \right) \quad (4)$$

Man berechnet zuerst den Forward Teil der Forward-Backward-Inferenz $\alpha_t = P(q_t \wedge b_1 \wedge \dots \wedge b_t)$. Nach der Definition und den Abhängigkeiten, die aus dem Graph abgelesen werden können, wissen wir, dass

$$\alpha_1 = P(q_1 \wedge b_1) = P(q_1)P(b_1 | q_1) \quad (5)$$

$$\begin{aligned} \alpha_2 &= P(q_2 \wedge b_1 \wedge b_2) \\ &= \sum_{q_1} P(q_1 \wedge b_1)P(q_2 | q_1)P(b_2 | q_2) \end{aligned}$$

$$= \sum_{q_1} \alpha_1 P(q_2|q_1) P(b_2|q_2) \quad (6)$$

Dabei steht Σ_{q_t} für die Summe über allen möglichen unterschiedlichen Zuständen von q_t und die Formel (6) bekommt man, wenn man $\alpha_1 = P(q_1 \wedge b_1)$ aus (5) verwendet. Aus (5) und (6) kann man dann rekursiv die Forward-Inferenz in (7) definieren.

$$\begin{aligned} \alpha_{t+1} &= P(q_{t+1} \wedge b_1 \wedge \dots \wedge b_{t+1}) \\ &= \sum_{q_t} \alpha_t P(q_{t+1}|q_t) P(b_{t+1}|q_{t+1}) \end{aligned} \quad (7)$$

Der Backward Teil der Forward-Backward-Inferenz $\beta_t = P(b_{t+1} \wedge \dots \wedge b_T|q_t)$ wird danach bestimmt. Hierbei geht man von β_T abwärts vor. Zuerst wird β_T in (8) definiert.

$$\beta_T = P(\emptyset|q_T) = 1 \quad (8)$$

Und β_{t-1} wird nun rekursiv berechnet:

$$\begin{aligned} \beta_{t-1} &= P(b_t \wedge \dots \wedge b_T|q_{t-1}) \\ &= \sum_{q_t} P(q_t|q_{t-1}) P(b_t \wedge b_{t+1} \wedge \dots \wedge b_T|q_t) \\ &= \sum_{q_t} P(q_t|q_{t-1}) P(b_t|q_t) P(b_{t+1}, \dots, b_T|q_t) \\ &= \sum_{q_t} P(q_t|q_{t-1}) P(b_t|q_t) \beta_t \end{aligned} \quad (9)$$

Wenn man α_t und β_t in (10) kombiniert, bekommt man die Wahrscheinlichkeit, dass q_t der unbekannte Zustand zur Zeit t ist und dass die Beobachtungen B beobachtet wurden.

$$P(q_t \wedge B) = P(q_t, b_1, \dots, b_t) P(b_{t+1}, \dots, b_T|q_t) = \alpha_t \beta_t \quad (10)$$

$P(q_t \wedge B)$ ist proportional zur $P(q_t|B)$, was man durch Normalisierung von $P(q_t \wedge B)$ bekommen kann. Neben $P(q_t|B)$ wird in der Forward-Backward Inferenz auch $P(q_t \wedge q_{t+1}|B)$ berechnet. Bei der Ableitung von $P(q_t \wedge q_{t+1}|B)$ in (11) wird im Nenner der Satz der totalen Wahrscheinlichkeit verwendet. Im Zähler wird die Unabhängigkeit der Zufallsvariablen ausgenutzt (siehe Graph in der Abb. 3).

$$\begin{aligned} P(q_t \wedge q_{t+1}|B) &= \frac{P(q_t \wedge q_{t+1} \wedge b_1 \dots b_t \wedge b_{t+1} \wedge b_{t+2}, \dots b_T)}{P(B)} \\ &= \frac{P(q_t \wedge b_1 \dots b_t) P(q_{t+1}|q_t) P(b_{t+2} \wedge \dots b_T|q_{t+1}) P(b_{t+1}|q_{t+1})}{\sum_{t=1}^{t=T} P(q_t \wedge B)} \\ &= \frac{\alpha_t P(q_{t+1}|q_t) \beta_{t+1} P(b_{t+1}|q_{t+1})}{\sum_{t=1}^{t=T} \alpha_t} \end{aligned} \quad (11)$$

Auf diese Weise kann man mit dem Forward-Backward-Algorithmus auf die unbekannten Zustände schließen. In jedem Erwartungsschritt vom Baum-Welch-Algorithmus wird die Forward-Backward-Inferenz benutzt, und so werden die Wahrscheinlichkeiten in der Übergangsmatrix und Beobachtungsmatrix berechnet.

In der Abb. 3 kann man dann sehen, wie ein trainiertes Hidden-Markow-Modell im Bereich der Aktivitätserkennung aussehen kann. Dabei repräsentieren die Buchstaben V, A und H die Features, welche aus Sensordaten extrahiert wurden. Die Matrizen stellen die Übergangs- und Beobachtungsmatrix dar und in der Übergangsmatrix sind die Wahrscheinlichkeiten für den jetzigen Zustand in Abhängigkeit vom letzten Zustand dargestellt. Zum Beispiel, wenn der letzte Zustand 'Stehen' war, dann ist die Wahrscheinlichkeit, dass der jetzige Zustand wieder 'Stehen' ist 80 % und dass er 'Rennen' ist 20 %. Die Beobachtungsmatrix enthält die Wahrscheinlichkeiten für den jetzigen Zustand in Abhängigkeit von der beobachteten Beobachtung. Beispielsweise, falls V beobachtet wurde, ist die Wahrscheinlichkeit, dass der jetzige Zustand 'Stehen' ist 20 % und dass er 'Rennen' ist 80 %.

Wenn das Hidden-Markow-Modell in der Aktivitätserkennung benutzt wird, werden meistens so viele Modelle trainiert, wie es Aktivitäten zu erkennen gibt. Beispielsweise in [31] wird zwischen sechs Aktivitäten mit sechs HMM unterschieden. Einzelne Modelle wurden so trainiert, dass sie erkennen, ob die entsprechende Aktivität ausgeführt wird oder nicht. Bei der Eingabe einer Sequenz aus den Feature-Vektoren wird die Aktivität erkannt, deren Modell die höchste Wahrscheinlichkeit ausgibt.

VI. FE-AT ANSATZ

Der Ansatz, der bei begrenzten Trainingsdaten Vorteile bringt, ist der FE-AT Ansatz [29]. Aus diesem Grund wird er vor allem eingesetzt, wenn ein bestehendes Aktivitätserkennungssystem erweitert werden soll. Der Ansatz wird FE-AT genannt, weil er aus der Fusion der featurebasierten (FE) und attributbasierten (AT) Ansätzen entstanden ist. Featurebasierter Ansatz ist der Ansatz, welcher in der Aktivitätserkennung üblicherweise benutzt wird und welchen wir bereits auf dem Beispiel der Aktivitätserkennung mit HMM gesehen haben.

Bei dem attributbasierten Ansatz bestehen die Aktivitäten aus den Attributen [29]. Attribute sind bestimmte Positionen des Körpers oder Eigenschaften der Aktivität (z. B. 'Schwingende Arme' oder 'Ruhige Beine'). In diesem Ansatz werden mit einem klassischen featurebasierten Ansatz (siehe Abb. 4, Attribut-Detektoren) die Attribute erkannt. Der wesentliche Teil von diesem Ansatz ist die Matrix, welche die Attribute und

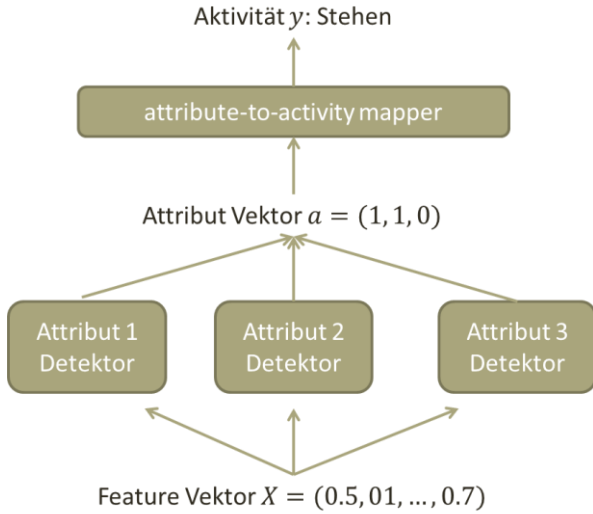


Abb. 4. Funktionsweise des FE-AT Ansatzes [29].

		0	0	1	0
Rennen		0	1	0	0
Fahrad fahren		1	1	0	1
Sitzen		1	1	0	1
Liegen		1	1	0	1
		Ruhige Körper	Körper ändert den Winkel nicht	Arme schwingen	Ruhige Beine

Aktivitäten (links) **Attribute** (unten)

Abb. 5. Beispiel eines Attribute-Aktivität Matrix [29].

Aktivitäten in die Beziehung bringt. In der Abb. 5 kann man auf einem einfachen Beispiel sehen, wie dieser Matrix angewandt auf die Erkennung von Aktivitäten aussehen könnte. In der Abb. 4 kann man sehen, wie der FE-AT Ansatz funktioniert. Zuerst wird der Feature-Vektor mit der Feature-Extraktion extrahiert. Danach wird mit Hilfe von Attribute-Detektoren erkannt, welche Attribute sich in diesem Feature-Vektor befinden. Die erkannten Attribute werden mit einem Attributvektor repräsentiert, welcher dann zum ‚attribute-to-activity mapper‘ übergeben wird. Im ‚attribute-to-activity mapper‘ wird der Attributvektor mit der Matrix aus der Abb. 4 verglichen und damit wird die Aktivität erkannt (beispielsweise kann die Matrix mit dem Vektor multipliziert werden und die resultierenden Zeilen die 1 enthalten würden, würden den erkannten Aktivitäten entsprechen).

In der Abb. 5 ist auch das größte Problem des attributbasierten Ansatz zu sehen. Wenn zwei oder mehr Aktivitäten die gleichen Attribute haben (wie ‚Sitzen‘ und ‚Liegen‘ in der Abb. 5), kann mit dem ‚attribute-to-activity mapper‘ die Aktivität nicht eindeutig bestimmt werden. Dieses Problem kann durch das Hinzufügen der weiteren Attribute gelöst werden.

Die Fusion der beiden Ansätze wird mit der Scoring-Funktion beschrieben:

$$f_{FE-AT}(y|x) = \begin{cases} P_{FE}(y|x) + P_{AT}(y|x) & \text{falls } y = y_{ne} \\ P_{FE}(y|x) & \text{sonst} \end{cases} \quad (12)$$

Im Vergleich mit den anderen Fusion-Ansätzen [29, 30] wie (13) und (14) ist die Besonderheit des FE-AT Ansatzes, dass aus dem Attribut-basierten Ansatz bekommene Wahrscheinlichkeit für eine Aktivität nur verwendet wird, falls es um die neue Aktivität geht.

$$f_{MAX}(y|x) = \max\{P_{FE}(y|x), P_{AT}(y|x)\} \quad (13)$$

$$f_{SUM}(y|x) = P_{FE}(y|x) + P_{AT}(y|x) \quad (14)$$

Evaluation des FE-AT Ansatzes [29]: Nguyen et al. haben den FE-AT Ansatz mit den drei öffentlichen Datensätzen bewertet und mit dem featurebasierten, attributbasierten und den anderen Fusionsansätzen verglichen (siehe Abb. 6). Sie haben zwischen der Erkennung der Aktivitäten, welche dem System hinzugefügt wurden und den schon existierenden Aktivitäten unterschieden. Im Fall der existierenden Aktivitäten hat der FE-AT Ansatz keine Verbesserung der Genauigkeit gegenüber dem featurebasierten Ansatz gezeigt. Andererseits wurde aber die Erkennung der hinzugefügten Aktivitäten deutlich besser als bei dem featurebasierten Ansatz oder den anderen Fusionsansätzen. In beiden Fällen war die Genauigkeit des attributbasierten Ansatzes schlechter als bei den anderen Ansätzen. Das liegt daran, dass einige Aktivitäten mit den gleichen Attributvektoren beschrieben wurden. Schließlich zeigte diese Studie auch, dass die Vorteile vom FE-AT Ansatz mit der wachsenden Anzahl der Trainingsdaten für die hinzugefügte Aktivität im Vergleich zum featurebasierten Ansatz verschwinden (siehe Abb. 7).

VII. AKTIVITÄTSERKENNUNGSSYSTEME

Eine herkömmliche Architektur der Aktivitätserkennungssysteme besteht aus den Sensoren, dem Integrationsgerät und dem Server, mit welchem über Internet kommuniziert wird [13]. Übliche Integrationsgeräte sind Smartphones oder Laptops. In besonderen Fällen können alle drei Komponenten der Erkennungssysteme in einem einzigen Gerät enthalten sein. Dies ist beispielsweise der Fall, wenn Smartphone als Integrationsgerät verwendet wird und ausschließlich die im Smartphone

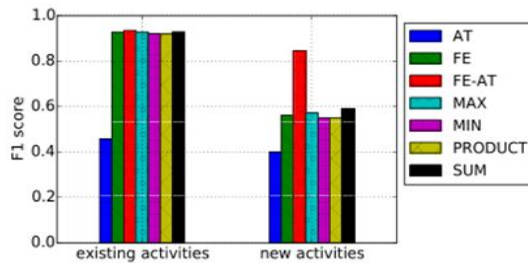


Abb. 6. FE-AT Ansatz im Vergleich mit den anderen Ansätzen [29].

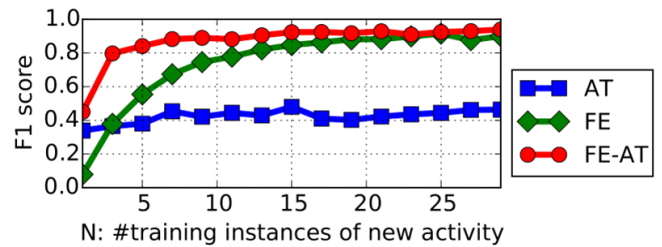


Abb. 7. FE-AT Ansatz verglichen mit Feature- und Attribut-basierten Ansätzen in Abhängigkeit von der Anzahl der Trainingsdaten der hinzugefügten Aktivität [29].

ne enthaltenen Sensoren genutzt werden. Die gesammelten Daten können potenziell ohne Kommunikation mit dem Server auf dem Smartphone gespeichert und verarbeitet werden.

Aktivitätserkennungssysteme werden je nach der Reaktionszeit in Online- und Offline-Systeme unterteilt [13]. Online-Systeme reagieren und erkennen die Aktivitäten sofort und können deswegen in den Echtzeitsystemen angewandt werden. Die Offline-Systeme brauchen eine bestimmte Zeit für die Erkennung von Aktivitäten. Der Grund dafür können beispielsweise eine hohe Verarbeitungszeit oder die Notwendigkeit der Kommunikation mit dem Server sein. Insbesondere in der Trainingsphase unterscheiden sich diese Systeme so, dass die Daten in den Online-Systemen sofort beschriftet werden.

Betrachtet werden die Systeme, die in der Geschichte der Aktivitätserkennung wichtig sind [37]; Systeme, die besondere Wearables verwenden [35, 40]; ein System, das Aktivitäten von Tieren erkennt [9]; ein System, welches die Aktivitäten nur anhand der Augenblinkfrequenz erkennt [39]. Ferner unterscheiden sich die vorgestellten Aktivitätserkennungssysteme in der Anzahl und Art der erkannten Aktivitäten. Es ist bemerkenswert, dass sich die Anzahl der erkannten Aktivitäten negativ auf die Genauigkeit auswirkt. Andererseits ist bei den Sensoren die Position und Art des Sensors wichtiger als ihre Anzahl. Im Folgenden werden die Aktivitätserkennungssysteme in der Tabelle 1 näher betrachtet.

A. Online-Systeme

Tapia et al. [34]: Ihr System erkennt 30 Übungsaktivitäten, wie z.B. 'Gewichte heben', 'Rudern' oder 'Liegestützen'. Sie benutzen 5 Beschleunigungssensoren am Arm und am Handgelenk, am Oberschenkel und am Knöchel sowie an den Hüften. Als Integrationsgerät wird ein Laptop verwendet, und wegen der großen Anzahl von Sensoren ist das System sehr störend. Wenn das System in der Trainingsphase an den Benutzer angepasst wird, zeigt es deutlich höhere Genauigkeit (86 %). Ansonsten verringert sich die Genauigkeit auf 56 %.

Kao et al. [35]: Die Besonderheit des Systems von Kao et al. ist neben der Erkennung der globalen Bewegungsaktivitäten auch die Möglichkeit, die Aktivitäten mit lokaler Interaktion zu erkennen. Das wird von den gewählten Sensorpositionen ermöglicht. Dieses System erkennt sieben verschiedene tägliche und Bewegungsaktivitäten wie 'Zähneputzen', 'Schlagen', 'Klopfen' und 'Rennen'. Außerdem ist das System nicht aufdringlich und unkompliziert, weil nur ein Sensor auf dem

Handgelenk benutzt wird. Ein eingebettetes System, das getragen werden muss, wird als Integrationsgerät verwendet. Schließlich zeichnet sich das System von Kao et al. mit einer niedrigen durchschnittlichen Reaktionszeit (10 ms) aus.

COSAR [36]: Dieses System erkennt zehn tägliche und Bewegungsaktivitäten wie 'Zähneputzen' und 'Schreiben an der Tafel'. Es werden zwei Beschleunigungssensoren benutzt, einer am Handgelenk und einer im Handy, der auch als Integrationsgerät verwendet wird. Bei der Aktivitätserkennung werden die zuvor ausgeführten Aktivitäten und der Standort des Benutzers in Betracht gezogen. Der Standort wird mit dem GPS-Sensor im Handy bestimmt. Da dieses System die Aktivitäten mit den Standorten in einer Matrix in Zusammenhang bringt, kann es beispielsweise erkennen, dass der Benutzer nicht joggt, wenn er sich im Wohnzimmer befindet.

B. Offline-Systeme

Bao und Intille [37]: Dieses System wurde in 2004 veröffentlicht und stellte damals einen bedeutenden Schritt vorwärts dar. Es verwendet 5 Beschleunigungssensoren, die an Knie, Knöchel, Arm, Ellbogen und an den Hüften platziert sind. Die Autoren stellten jedoch fest, dass die Genauigkeit nicht wesentlich beeinträchtigt wird (bis zu 5 % niedriger), falls nur Beschleunigungssensoren an den Hüften und am Handgelenk verwendet werden. Zwanzig tägliche und Bewegungsaktivitäten wie 'Fernsehen', 'Staubsaugen' und 'Fahrrad fahren' können erkannt werden. Für Speicherung der Daten wurden Hoarder-Boards benutzt, welche die Benutzer in vier Handybeutel getragen haben. Die geringste Genauigkeit erreichte das System bei 'Dehnung' (41,42 %) und 'Fahren mit dem Aufzug' (43,58 %), was oft mit der 'Benutzung der Rolltreppen' (70,56 %) verwechselt wurde. Zur weiteren Verbesserung der Genauigkeit der Aktivitätserkennung haben die Autoren die Verwendung zusätzlicher Sensoren wie GPS, Herzfrequenz und die Berücksichtigung der Standortinformationen (was später in COSAR umgesetzt wurde) vorgeschlagen.

Ladha et al. [9]: Dieses System erkennt nicht die menschlichen Aktivitäten, sondern siebzehn Hundeaktivitäten wie das 'Bellen', 'Essen', 'Schnüffeln' und 'Springen'. Die Daten wurden nur mit einem wasserdichten Beschleunigungssensor im Hundehalsband gesammelt. Abtastrate wurde auf 30 Hz gesetzt und die Daten wurden auf ein 4 Gb Flash-Speicher gespeichert. Das Experiment wurde mit achtzehn Hunden aus dreizehn verschiedenen Hunderassen ausgeführt. Die Hunde wurden 20-40

	Aktivitäten (#)	Sensoren (#)	Integrations- gerät	Experiment	Aufdringlichkeit	Genauigkeit
Online						
Tapia [34]	Übung (30)	Beschleunigungs- (5), Herzfrequenz	Laptop	LAB	Hoch	86%
Kao [35]	Bewegungs-, Tägliche (7)	Beschleunigungs- (1)	Custom	N/S	Niedrig	94,71%
COSAR [36]	Bewegungs-, Tägliche (10)	Beschleunigungs- (2)	Handy	NAT	Medium	93%
Offline						
Bao [37]	Bewegungs-, Tägliche (20)	Beschleunigungs- (5)	/	NAT	Hoch	84%
Ladha [9]	Hundeaktivitäten (17)	Beschleunigungs- (1, Hals)	/	NAT	Niedrig	68,6%
Khan [38]	Bewegungs-, Übergänge (15)	Beschleunigungs- (1, Brust)	Computer	NAT	Medium	97,9%
Ishimaru [39]	Tägliche (5)	Beschleunigungs-, Umgebungs- (4)	Google Glass	LAB	Niedrig	82%

Tabelle 1. Betrachtete Aktivitätserkennungssysteme [13].

N/S – nicht spezifiziert
LAB – Experiment im Labor
NAT – Naturalistisches Experiment

Minuten gefilmt, und anschließend wurden die Daten von einem Experten mit Hilfe der Videoaufnahmen annotiert. Als Klassifikationsalgorithmus wurde KNN verwendet. Die Aktivitäten der größeren Hunde wurden besser als der kleineren Hunde erkannt und das System hat am besten zwischen 'Essen' und 'Trinken' unterschieden.

Khan et al. [38]: Dieses System unterscheidet zwischen fünfzehn Aktivitäten und den Übergängen zwischen ihnen. Es erkennt z. B. statische Aktivitäten wie 'Liegen', Übergänge wie 'vom Liegen zum Sitzen' und dynamische Aktivitäten wie 'Rennen'. Ein einziger Sensor auf der Brust wurde benutzt. Als Integrationsgerät wurde Computer verwendet und der Sensor hat die Daten mit Bluetooth nach Computer geschickt, wo sie verarbeitet und gespeichert wurden.

Ishimaru et al. [39]: Das System von Ishimaru unterscheidet zwischen fünf täglichen Aktivitäten ('Lesen', 'Reden', 'Fernsehen', 'Lösen von mathematischen Problemen' und 'Sägen'). Als Integrationsgerät wurde Google Glass benutzt und nur die in Google Glass eingebaute Sensoren wurden verwendet (Kamera, Mikrophone, inertielle Messeinheit und Näherungssensor). Für die Erkennung von Aktivitäten wurden die Augenblinkfrequenz und Kopfbewegung betrachtet. Ferner wurde der Näherungssensor für die Erkennung der Augenblinkfrequenz verwendet. Im Experiment haben acht Probanden mit Google Glass jede Aktivität fünf Minuten ausgeführt. Einerseits war die Genauigkeit der Aktivitätserkennung bei ausschließlicher Verwendung der Augenblinkfrequenz durchschnittlich bei 67 % (abhängig von der Probandin, variierte sie im Bereich von 52 % bis 82 %). Wurde dagegen nur die Kopfbewegung verwendet, lag sie im Durchschnitt bei 63 % und reichte von 50 % bis 83 %. Bei der Kombination beider Ansätze

ze schließlich betrug die Genauigkeit im Durchschnitt 82 % und lag in einem Bereich von 74 % bis 92 %.

VIII. FAZIT

In dieser Arbeit wurden die Aktivitätserkennungssysteme mit Wearables betrachtet. Heutige Aktivitätserkennungssysteme können zwischen wenigen Aktivitäten mit hoher Genauigkeit unterscheiden. Am besten werden die Bewegungsaktivitäten erkannt. Die Kombination der Ansätze bietet sich für die Zukunft an. Interessant für die Zukunft sind auch Systeme, wie das System von Ishimaru, das aus kleinen Aktionen der Benutzer die Aktivitäten erkennen können. Weitere Herausforderungen für die Zukunft sind die Erkennung der komplexeren Aktivitäten (wie z. B. Fußballspielen) und die Entwicklung der Systeme, die zwischen den Aktivitäten von mehreren Typen unterscheiden könnten, sowie die Entwicklung der standardisierten Datensätze, mit welchen die Aktivitätserkennungssysteme besser miteinander verglichen werden könnten.

REFERENZEN

- [1] Number of smartphone users worldwide from 2014 to 2020. In Statista - The Statistics Portal
- [2] Number of connected wearable devices worldwide from 2016 to 2021. In Statista - The Statistics Portal.
- [3] Want, R. (2009). When cell phones become computers. *IEEE Pervasive Computing*, 8(2), 2-5.
- [4] Patel, S., Park, H., Bonato, P., Chan, L., & Rodgers, M. (2012). A review of wearable sensors and systems with application in rehabilitation. *Journal of neuroengineering and rehabilitation*, 9(1), 21
- [5] Hong, Y. J., Kim, I. J., Ahn, S. C., & Kim, H. G. (2008, December). Activity recognition using wearable sensors for elder care. In *2008 second international conference on future generation communication and networking* (Vol. 2, pp. 302-305). IEEE.

- [6] Wyss, T., & Mäder, U. (2010). Recognition of military-specific physical activities with body-fixed sensors. *Military medicine*, 175(11), 858-864.
- [7] Scheurer, S., Tedesco, S., Brown, K. N., & O'Flynn, B. (2017, May). Human activity recognition for emergency first responders via body-worn inertial sensors. In *2017 IEEE 14th International Conference on Wearable and Implantable Body Sensor Networks (BSN)* (pp. 5-8). IEEE.
- [8] Soukup, R., Blecha, T., Hamacek, A., & Reboun, J. (2014, September). Smart textile-based protective system for firefighters. In *Proceedings of the 5th Electronics System-integration Technology Conference (ESTC)* (pp. 1-5). IEEE.
- [9] Ladha, C., Hammerla, N., Hughes, E., Olivier, P., & Ploetz, T. (2013, September). Dog's life: wearable activity recognition for dogs. In *Proceedings of the 2013 ACM international joint conference on Pervasive and ubiquitous computing* (pp. 415-418).
- [10] Venkatraman, S., Long, J. D., Pister, K. S., & Carmenta, J. M. (2007, August). Wireless inertial sensors for monitoring animal behavior. In *2007 29th Annual International Conference of the IEEE Engineering in Medicine and Biology Society* (pp. 378-381). IEEE.
- [11] Chen, L., Hoey, J., Nugent, C. D., Cook, D. J., & Yu, Z. (2012). Sensor-based activity recognition. *IEEE Transactions on Systems, Man, and Cybernetics, Part C (Applications and Reviews)*, 42(6), 790-808.
- [12] Caine, K. E., Fisk, A. D., & Rogers, W. A. (2006, October). Benefits and privacy concerns of a home equipped with a visual sensing system: A perspective from older adults. In *Proceedings of the human factors and ergonomics society annual meeting* (Vol. 50, No. 2, pp. 180-184). Sage CA: Los Angeles, CA: SAGE Publications.
- [13] Lara, O. D., & Labrador, M. A. (2012). A survey on human activity recognition using wearable sensors. *IEEE communications surveys & tutorials*, 15(3), 1192-1209.
- [14] Van Kasteren, T. L. M., Englebienne, G., & Kröse, B. J. (2010). An activity monitoring system for elderly care using generative and discriminative models. *Personal and ubiquitous computing*, 14(6), 489-498.
- [15] Cornacchia, M., Ozcan, K., Zheng, Y., & Velipasalar, S. (2017). A survey on activity detection and classification using wearable sensors. *IEEE Sensors Journal*, 17(2), 386-403.
- [16] Mukhopadhyay, S. C. (2014). Wearable sensors for human activity monitoring: A review. *IEEE sensors journal*, 15(3), 1321-1330.
- [17] Morales, J., & Akopian, D. (2017). Physical activity recognition by smartphones, a survey. *Biocybernetics and Biomedical Engineering*, 37(3), 388-400.
- [18] Moncada-Torres, A., Leuenberger, K., Gonzenbach, R., Luft, A., & Gassert, R. (2014). Activity classification based on inertial and barometric pressure sensors at different anatomical locations. *Physiological measurement*, 35(7), 1245.
- [19] Shoaib, M., Scholten, H., & Havinga, P. J. (2013, December). Towards physical activity recognition using smartphone sensors. In *2013 IEEE 10th international conference on ubiquitous intelligence and computing and 2013 IEEE 10th international conference on autonomic and trusted computing* (pp. 80-87). IEEE.
- [20] Michalski, R. S., Carbonell, J. G., & Mitchell, T. M. (Eds.). (2013). *Machine learning: An artificial intelligence approach*. Springer Science & Business Media.
- [21] Schapire, R.E., & Wang, C. (2008). Cos 511: Theoretical Machine Learning.
- [22] Chen, L., Nugent, C. D., & Wang, H. (2011). A knowledge-driven approach to activity recognition in smart homes. *IEEE Transactions on Knowledge and Data Engineering*, 24(6), 961-974.
- [23] Wang, Y., Lin, J., Annavaram, M., Jacobson, Q. A., Hong, J., Krishnamachari, B., & Sadeh, N. (2009, June). A framework of energy efficient mobile sensing for automatic user state recognition. In *Proceedings of the 7th international conference on Mobile systems, applications, and services* (pp. 179-192).
- [24] Pirttikangas, S., Fujinami, K., & Nakajima, T. (2006, October). Feature selection and activity recognition from wearable sensors. In *International symposium on ubiquitous computing systems* (pp. 516-527). Springer, Berlin, Heidelberg.
- [25] Huynh, T., & Schiele, B. (2005, October). Analyzing features for activity recognition. In *Proceedings of the 2005 joint conference on Smart objects and ambient intelligence: innovative context-aware services: usages and technologies* (pp. 159-163).
- [26] P Peng, H., Long, F., & Ding, C. (2005). Feature selection based on mutual information criteria of max-dependency, max-relevance, and min-redundancy. *IEEE Transactions on pattern analysis and machine intelligence*, 27(8), 1226-1238.
- [27] Maurer, U., Smailagic, A., Siewiorek, D. P., & Deisher, M. (2006, April). Activity recognition and monitoring using multiple sensors on different body positions. In *International Workshop on Wearable and Implantable Body Sensor Networks (BSN'06)* (pp. 4-pp). IEEE.
- [28] Huang, B., CS5824/CS4824/ECE5424/ECE4424: Machine Learning, Fall 2015, Virginia Tech
- [29] Nguyen, L. T., Zeng, M., Tague, P., & Zhang, J. (2015, September). Recognizing new activities with limited training data. In *Proceedings of the 2015 ACM International Symposium on Wearable Computers* (pp. 67-74). ACM.
- [30] Mangai, U. G., Samanta, S., Das, S., & Chowdhury, P. R. (2010). A survey of decision fusion and feature fusion strategies for pattern classification. *IETE Technical review*, 27(4), 293-307.
- [31] San-Segundo, R., Montero, J. M., Moreno-Pimentel, J., & Pardo, J. M. (2016). HMM adaptation for improving a human activity recognition system. *Algorithms*, 9(3), 60.
- [32] Rabiner, L., & Juang, B. (1986). An introduction to hidden Markov models. *IEEE ASSP Magazine*, 3(1), 4-16.
- [33] Jurafsky, D., Martin, J.H., *Speech and Language Processing* (3rd ed. draft)
- [34] Tapia, E. M., Intille, S. S., Haskell, W., Larson, K., Wright, J., King, A., & Friedman, R. (2007, October). Real-time recognition of physical activities and their intensities using wireless accelerometers and a heart rate monitor. In *2007 11th IEEE international symposium on wearable computers* (pp. 37-40). IEEE.
- [35] Kao, T. P., Lin, C. W., & Wang, J. S. (2009, July). Development of a portable activity detector for daily activity recognition. In *2009 IEEE international symposium on industrial electronics* (pp. 115-120). IEEE.
- [36] Riboni, D., & Bettini, C. (2011). COSAR: hybrid reasoning for context-aware activity recognition. *Personal and Ubiquitous Computing*, 15(3), 271-289.
- [37] Bao, L., & Intille, S. S. (2004, April). Activity recognition from user-annotated acceleration data. In *International conference on pervasive computing* (pp. 1-17). Springer, Berlin, Heidelberg.
- [38] Khan, A. M., Lee, Y. K., Lee, S. Y., & Kim, T. S. (2010). A triaxial accelerometer-based physical-activity recognition via augmented-signal features and a hierarchical recognizer. *IEEE transactions on information technology in biomedicine*, 14(5), 1166-1172.
- [39] Ishimaru, S., Kunze, K., Kise, K., Weppner, J., Dengel, A., Lukowicz, P., & Bulling, A. (2014, March). In the blink of an eye: combining head motion and eye blink frequency for activity recognition with google glass. In *Proceedings of the 5th augmented human international conference* (pp. 1-4).
- [40] Shoaib, M., Bosch, S., Incel, O. D., Scholten, H., & Havinga, P. J. (2014). Fusion of smartphone motion sensors for physical activity recognition. *Sensors*, 14(6), 10146-10176.
- [41] Berchtold, M., Budde, M., Gordon, D., Schmidtke, H. R., & Beigl, M. (2010, October). Actiserv: Activity recognition service for mobile phones. In *International symposium on wearable computers (ISWC) 2010* (pp. 1-8). IEEE.
- [42] Reddy, S., Mun, M., Burke, J., Estrin, D., Hansen, M., & Srivastava, M. (2010). Using mobile phones to determine transportation modes. *ACM Transactions on Sensor Networks (TOSN)*, 6(2), 1-27.
- [43] Ghahramani, Z. (2001). An introduction to hidden Markov models and Bayesian networks. In *Hidden Markov models: applications in computer vision* (pp. 9-41).
- [44] Ng, A. (2011). CS229 Notes.
- [45] Murphy, K. P. (2012). *Machine learning: a probabilistic perspective*. MIT press.