

SENSORBASIERTE BEWEGUNGSANALYSE IM FAHRZEUG TEIL 2: OBJEKTKLASSIFIKATION MIT MASCHINELLEM LERNEN

Tristan Langenberg, Oliver Jokisch und Tilo Strutz

Institut für Kommunikationstechnik, Hochschule für Telekommunikation Leipzig (HfTL)

tristan.langenberg@hftl.de, jokisch@hftl.de, strutz@hftl.de

Kurzfassung: Der vorliegende Beitrag diskutiert ein sensorbasiertes Bewegungsanalysekonzept im Fahrzeugkontext und beschränkt sich auf Signalmerkmale von dreidimensionalen Beschleunigungs- und Gyroskopdaten im bewegten Fahrzeug. Forschungsgegenstand ist die Fahrbahnanalyse zur Entwicklung von Fahrassistenzsystemen in den Bereichen Sicherheit und Komfort.

Die Untersuchung beruht auf einem früheren Beitrag [1] zum gleichen Forschungsgegenstand, der smartphonebasierte Messdaten betrachtete und erste Erfahrungen mit der Merkmalsanalyse und -klassifikation von Beschleunigungssignalen zusammenfasste. Der hier vorgestellte Teil 2 greift das Anwendungsszenario der Straßencharakterisierung auf und stellt ein System zur Objektklassifikation unter Verwendung eines Algorithmus zum maschinellen Lernen vor. Es wird ein Neuronales Netz (NN) als Methode des überwachten Lernens, speziell ein Multi-Layer-Perzeptron (MLP), zur Klassifikation verwendet.

Zum Anlernen des neuronalen Netzes wird eine selbst aufgezeichnete Datenbank eingesetzt, die aus fünf Objekttypen (Temposchwelle, Bremsschwelle, Kanaldeckel, Brückenfuge, idealer Straßenuntergrund) mit zusätzlichen Dimensionsinformationen (rechts, links, frontal) besteht. Die insgesamt 833 Datensätze mit zugrunde liegenden Geschwindigkeiten zwischen 5 und 75 km/h wurden mit zwei verschiedenen Fahrzeugen und einem selbst konstruierten Messgerät¹ aufgenommen. Einflüsse von variierenden Fahrzeuggeschwindigkeiten und -typen werden bei der Datenanalyse und Datenaufbereitung kompensiert.

Im Ergebnis werden unter den genannten Randbedingungen Objekterkennungsraten von 75 bis 85 % erzielt, die von den jeweils trainierten Datensätzen und den extrahierten Signalmerkmalen, z. B. den Ortsspektren der Messdaten, beeinflusst sind.

1 Einleitung

Im Themengebiet der Objektklassifikation mit sensorbasierten Daten, die in Fahrzeugen aufgezeichnet werden, gibt es relevante Voruntersuchungen z. B. von Wang, Kodagoda und Khusaba [3]. Dieser Ansatz klassifiziert den Fahrbahnuntergrund im Hinblick auf die Objekttypen Wiese, Schotter und Asphalt mit einer Support Vector Machine (SVM). Dabei wird der Fokus auf die Geschwindigkeitsunabhängigkeit gelegt, indem feste Geschwindigkeiten von 20, 30 und 40 $\frac{km}{h}$ vorgegeben und alle Messdaten in den Frequenzbereich transformiert werden. Die Arbeit erreicht Erkennungsraten von bis zu 90 %. Eine weitere Untersuchung von Gonzalez, Martinez und Carlos [2] klassifiziert drei konkrete Objekte: Schlaglöcher, Temposchwellen und Metallschwellen. Die Autoren nutzten als Klassifikationssystem ein neuronales Netz mit einer

¹Das Messgerät basiert auf einem Raspberry PI 2, ausgerüstet mit einem 9D-Sensor von adafruit .

Erkennungsraten von bis zu 86 %. Im Vergleich zu diesen beiden Untersuchungen beinhaltet die vorliegende Untersuchung eine variable Fahrzeuggeschwindigkeit von 5 bis $75 \frac{km}{h}$, fünf unterschiedliche Objekttypen, zwei Fahrzeuge zur Datenaufzeichnung und eine höhere Abtastfrequenz der Beschleunigungsdaten von 800 Hz anstelle von 100 Hz (smartphonebasierter Sensor) bzw. von 300 Hz (eigene Messplattform namens CRUISE²). Außerdem wird in der vorliegenden Analyse eine Kreuzvalidierung vorgenommen.

Machine Learning (ML)-Algorithmen werden, im Gegensatz zu anderen Verfahren aus der multivarianten Statistik, zur Datenanalyse vermehrt eingesetzt, da sie eine höhere Aussagekraft in der Untersuchung von dynamischen Daten erzielen. Typische ML-Algorithmen sind z. B. Clusteranalysen, Prädiktionsanalysen oder Klassifikationen.

In dieser Untersuchung, die ein geeignetes Klassifikationssystem sucht, kommen mobile Sensordaten von Fahrzeugen zum Einsatz, die mit einem neuronalen Netz verarbeitet werden. Die Messdaten entstammen einer eigens dafür erzeugten Datenbank, die mit einer für diese Untersuchung neu entwickelten Messplattform aufgezeichnet wurde.

Ziel ist es, ein Klassifikationssystem zu entwickeln, das verschiedene Fahrbahnobjekte und deren Eigenschaften automatisch erkennen kann. Die verwendeten Messdaten werden durch Fahrstil, Fahrzeugtyp, Fahrbahnuntergrund sowie Fahrzeuggeschwindigkeit beeinflusst, womit veränderliche Merkmale von Fahrbahnobjekten unterstellt werden. Deshalb wird das Klassifikationssystem so konzeptioniert, dass die Ergebnisse möglichst unabhängig vom verwendeten Fahrzeugfahrwerk sowie von der gefahrenen Geschwindigkeit ermittelt werden.

2 Messdaten

2.1 Messplattform und Datenerfassung

Die Basis der Messplattform bildet ein Raspberry PI Modell 2 (PI) [4]. Der PI ist ein leistungsstarker und kostengünstiger Einplatinencomputer, der mit einem 900 MHz Quad-Core Prozessor und 1 GB Arbeitsspeicher ausgestattet ist. Er hat einen geringen Stromverbrauch und ist sehr klein, weshalb er für mobile Einsatzszenarien qualifiziert erscheint.

Der PI wird mit einem LSM9DS0 Sensor von adafruit, welcher ein 3-Achsen-Beschleunigungssensor, einen Gyroskopsensor und einen Magnetometer inklusive Thermometer besitzt, via I²C Bus³ kombiniert. Des Weiteren werden eine Raspberry PI-Kamera und weitere Peripheriekomponenten angeschlossen. Die Kamera dient zur Unterstützung der Messdatennotation und die Peripheriekomponenten werden zur Steuerung der Messplattform und zur Markierung der Objekte benötigt.

Die Datenerfassung erfolgt mit einer selbst programmierten Routine. Dabei werden die Daten bzw. Steuerungsregister des LSM9DS0 Sensors über den PI ausgelesen und gesetzt, vgl. [5]. Es werden Beschleunigungsdaten mit einer Abtastfrequenz f_a von 800 Hz und einem Messbereich von $\pm 4g$ und Gyroskopdaten mit 400 Hz Abtastfrequenz und einem Wertebereich von $245 \frac{deg}{s}$ erfasst. Die Gyroskopdaten werden zur nachfolgenden Datenbereinigung benötigt, um das Koordinatensystem des Sensors im Bezugssystem des Fahrzeuges abzubilden. Die Messgrößen werden jeweils in X-, Y-, und Z-Richtung aufgezeichnet und zusammen mit ihren diskreten Abtastzeitpunkten kT_a [s] mit $k = 1, 2, 3, \dots, K$ gespeichert. Durch die hohe Abtastfrequenz kann eine örtliche Auflösung Δx von 0,05 m nach dem Nyquist-Shannon-Abtasttheorem garantiert werden, wenn eine Geschwindigkeit von $\Delta x \cdot \frac{f_a}{2} = 20 \frac{m}{s}$ ($72 \frac{km}{h}$) nicht überschritten wird.

²Cas Research Ute for Intelligence, Safty and Exploration

³Inter-Integrated Circuit

Objekttyp	Anzahl Pkw 1	Anzahl Pkw 2	Gesamtanzahl
Temposchwelle	111	110	221
Bremsschwelle	62	76	138
Kanaldeckel	108	78	186
Musterbeispiel	111	97	208
Brückenfuge	36	44	80
Gesamt	424	409	833

Tabelle 1 - Datenbasis: Anzahl von Events je Objekttyp und Fahrzeug.

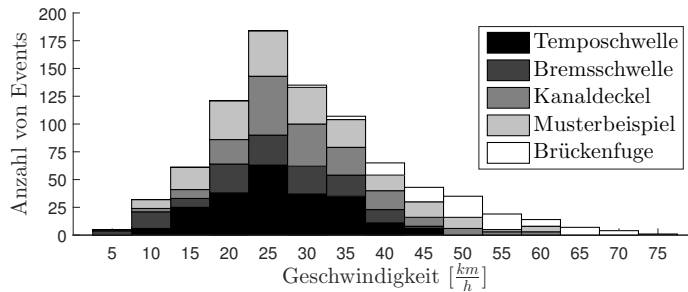


Abbildung 1 - Histogramm der Anzahl von Events gegenüber ihren Geschwindigkeiten.

2.2 Datenbasis

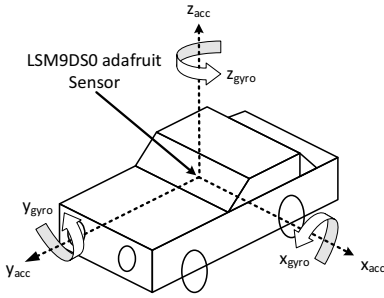
Die Datenbasis besteht aus fünf Objekttypen, die für jeweils zwei Fahrzeuge angelegt sind. Das erste Messfahrzeug ist ein alter Volkswagen Bora 1.6, Baujahr 2001 mit 77 kW und 1215 kg Leergewicht (Pkw 1). Das zweite Fahrzeug ist ein moderner Volkswagen Golf 1.2 TSI, Baujahr 2014 mit 81 kW und 1210 kg Leergewicht (Pkw 2). Die Anzahl von Objekten pro Typ wurde auf zehn Objekte (manuell ausgewählt und als repräsentativ für den jeweiligen Objekttyp definiert) begrenzt. Jede Überquerung eines Objektes erzeugt ein sogenanntes Event. Tabelle 1 fasst die Anzahl der Events pro Objekttyp und Fahrzeug zusammen. Insgesamt stehen 833 Events für die Analyse zur Verfügung, die sich auf beide Fahrzeuge pro Objekttyp gleich verteilen.

Die Dauer eines Events liegt zwischen 0,45 s und 3,96 s, und die mittlere Dauer beträgt 1,26 s mit einer Abweichung von -0,4 s bzw. +0,5 s für die beiden Fahrzeuge. Die mittlere Geschwindigkeit während der Aufnahme von Events beträgt $30,1 \frac{m}{s^2}$ mit einer Abweichung von $-1,4 \frac{m}{s^2}$ und $+1,6 \frac{m}{s^2}$ für die Fahrzeuge Pkw 1 und Pkw 2.

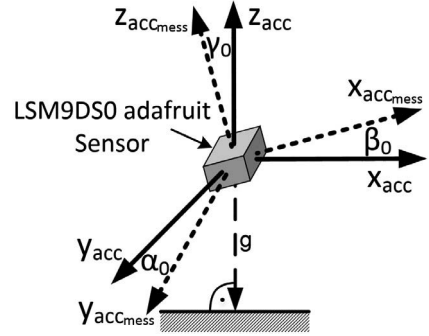
Die Datensätze beinhalten eine mittlere Geschwindigkeit von $28,7 \frac{km}{h}$ (Pkw 1) und $31,7 \frac{km}{h}$ (Pkw 2). Der Geschwindigkeitsbereich von $22,5$ bis $32,5 \frac{km}{h}$ ist auch der Bereich, der die höchste Anzahl von Datensätzen (320 Events), siehe Abbildung 1, aufweist. Abbildung 1 gibt Aufschluss über die variable Fahrzeuggeschwindigkeit während der Datenaufzeichnung und stellt dies in Form eines Histogramms mit einer Intervallbreite von $5 \frac{km}{h}$ dar.

2.3 Datenaufbereitung

Die aufgezeichneten Daten werden durch die Ausrichtung und Positionierung des Sensors beeinflusst. Die Abweichung des Koordinatensystems des Messgeräts zum Bezugssystem im Fahrzeug ist entscheidend für die exakte Datenerfassung. Diese Herausforderung ist jedoch in der Praxis nicht optimal lösbar, vgl. auch [1], so dass eine Nachberechnung zum Bestimmen der Abweichungen vom Fahrzeugbezugssystem notwendig wird. Abbildung 2 veranschaulicht die Abweichung des Messgerätekoordinatensystems ($x_{acc_{mess}}, y_{acc_{mess}}, z_{acc_{mess}}$) zum Fahrzeugbezugssystem ($x_{acc}, y_{acc}, z_{acc}$).



(a) Ausrichtung des Sensors im Bezugssystem



(b) Positionierung des Sensors im Fahrzeug

Abbildung 2 - Ausrichtung des Sensors (links) im Fahrzeug sowie (rechts) im Bezugssystem.

Die Fahrzeugbeschleunigung in X-, Y- und Z-Richtung wird abhängig von der Fahrzeugrotation im Raum mit der Erdbeschleunigung g überlagert. Um die Erdbeschleunigung herauszufiltern, werden die Gyroskopdaten und die Winkel β_{start} und γ_{start} , welche die Abweichung zu Messfahrtbeginn repräsentieren, verwendet. Damit wird die zeitliche Fahrzeugrotation und die jeweils wirkende Erdbeschleunigung beschrieben.

$$x_a(kT_a) = x_{acc_mess}(kT_a) - g \cdot \sin(|\beta_{start} + y_{gyro}(kT_a)|) \cdot \cos(\gamma_{start} - x_{gyro}(kT_a)) \quad (1)$$

$$y_a(kT_a) = y_{acc_mess}(kT_a) - g \cdot \cos(\beta_{start} + y_{gyro}(kT_a)) \cdot \sin(|\gamma_{start} - x_{gyro}(kT_a)|) \quad (2)$$

$$z_a(kT_a) = z_{acc_mess}(kT_a) - \sqrt{g^2 - x_{acc_neu}(kT_a)^2 - y_{acc_neu}(kT_a)^2} \quad (3)$$

Mit den Gleichungen (1) bis (3) wird die überlagerte Erdbeschleunigung aus allen drei Dimensionen der Beschleunigungsdaten ($x_{acc}(kT_a)$, $y_{acc}(kT_a)$, $z_{acc}(kT_a)$) entfernt. Zusätzlich werden mit Hilfe der Winkel α_0 , β_0 und γ_0 , siehe Abbildung 2, die Amplitudenverhältnisse der Beschleunigungsdaten in allen Dimensionen zueinander korrigiert. Dies geschieht unter Verwendung einer Rotationsmatrix.

$$\begin{pmatrix} x_{acc}(kT_a) \\ y_{acc}(kT_a) \\ z_{acc}(kT_a) \end{pmatrix} = \begin{pmatrix} \frac{1}{\cos(\alpha_0) \cdot \cos(\beta_0)} & -\sin(\alpha_0) \cdot \cos(\beta_0) & -\cos(\alpha_0) \cdot \sin(\beta_0) \\ -\sin(\alpha_0) \cdot \cos(\gamma_0) & \frac{1}{\cos(\alpha_0) \cdot \cos(\gamma_0)} & -\cos(\alpha_0) \cdot \sin(\gamma_0) \\ -\sin(\beta_0) \cdot \cos(\gamma_0) & -\cos(\beta_0) \cdot \sin(\gamma_0) & \frac{1}{\cos(\beta_0) \cdot \cos(\gamma_0)} \end{pmatrix} \cdot \begin{pmatrix} x_a(kT_a) \\ y_a(kT_a) \\ z_a(kT_a) \end{pmatrix}$$

3 Analyse

3.1 Methodik

Mit den aufgenommenen und aufbereiteten Messdaten soll ein neuronales Netz trainiert werden, das die fünf Objekttypen korrekt klassifizieren kann. Die Suche nach einem geeignetem Klassifikationssystem wird testweise mit drei Varianten der Eingabedaten in Z-Richtung durchgeführt:

- Variante A: Verwendung der bereinigten Beschleunigungsdaten $z_{acc}(kT_a)$,

- Variante B: Verwendung der Ortsspektren der Beschleunigungsdaten
 $|Z_{acc}(e^{j2\pi f/f_a})| = |DFT(z_{acc}(kX_a))|$ mit $X_a \bullet \rightarrow T_a$,
- Variante C: Verwendung der Profildaten
 $z_{mov}(kX_a) = \sum_{k=1}^K \sum_{k=1}^K z_{acc}(kT_a) + z_{acc}((k+1)T_a)$.

Für diese drei Varianten wird jeweils mit zwei unterschiedlichen neuronalen Paradigmen, Multilayer Perceptron (MLP) bzw. Cascading Multilayer Perceptron (C-MLP), mit jeweils 50 verschiedenen Netzstrukturen (und -größen) trainiert. Es wird die erfolgreichste Netzstruktur der besten Variante und des besten NN-Paradigmas als Klassifikationssystem gesucht. Die Netzstrukturen ergeben sich aus der systematischen und schrittweisen Abarbeitung der gelisteten Randbedingungen:

- maximale Anzahl von Neuronen je verdeckte Netzschicht: 50 Neuronen,
- maximale Anzahl von verdeckten Netzschichten: 2 verdeckte Schichten,
- Schrittweite zur Veränderung der Anzahl von Neuronen je Schicht: 5 Neuronen.

3.2 Vorbereitung

Die Eingangssignale für die gewählten neuronalen Netze müssen eine feste Länge haben. Deshalb werden die Eingabedaten der Variante A auf 400 Werte⁴, die der Variante B auf 257 Werte⁵ und einen Frequenzbereich von (0 Hz; 50 Hz] und die der Variante C auf 400 Werte und eine Fahrbahnprofillänge von 20 m linear interpoliert. Außerdem werden die Amplituden der Eingabedaten auf das Intervall [0; 1] relativ zueinander normiert.

Alle neuronalen Netzstrukturen beruhen auf den gleichen Netzkonfigurationen, und die praktische Umsetzung erfolgt unter Verwendung der MATLAB Neuronal Network Toolbox von MathWorks Inc. Als Aktivierungsfunktion aller Neuronen wird die Sigmoid-Funktion verwendet. Als Fehlerfunktion wird die Methode der kleinsten Fehlerquadrate (MSE) eingesetzt und das Lernverfahren wird mit dem Gradientenabstiegsverfahren (SCG)⁶ realisiert.

Die NN-Ausgabeschicht umfasst jeweils fünf Elemente (für jedes mögliche Event). Die Zielvektoren in der Lernphase bestehen aus einem 1-Element für die richtige Klasse und vier 0-Elementen.

3.3 Durchführung

Die Analyse wird mit allen 833 verfügbaren Events durchgeführt. Das Training jedes neuronalen Netzes wird durch drei Phasen bestimmt: Trainingsphase, Validierungsphase und Testphase. Die interessanteste ist die Testphase, denn in dieser wird das trainierte NN im Hinblick auf die Aspekte der Generalisierung und Überanpassung mit den Testdaten bewertet. Die Datenverteilung auf die drei Phasen erfolgt in der Proportion 80 % Trainingsdaten, 10 % Validierungsdaten und 10 % Testdaten. Eine gesicherte Aussage bezüglich der Testdatenergebnisse für die verschiedenen Varianten und Netzstrukturen wird durch eine Kreuzvalidierung erreicht. Dazu werden zehn Wiederholungen pro Dateneingabevariante, Netzarchitektur sowie NN-Paradigma durchgeführt. Jeder Testlauf verwendet 10 % Testdaten, wobei durch eine gleichmäßige Datenverteilung sichergestellt wird, dass sich die Testdaten für die zehn Wiederholungen nicht überschneiden.

⁴400 Werte entsprechen einer Eventdauer von 0,5 s und die mittlere Eventdauer der Datenbasis beträgt 0,45 s.

⁵Die 257 Werte des Frequenzspektrums ergeben sich aus der nächsten Zweierpotenz von 400 zu $2^9/2+1 = 257$.

⁶Scaled Conjugated Gradient.

NN-Paradigma	Variante A	Variante B	Variante C	Gesamt
MLP	67,8 %	73,3 %	61,2 %	67,4 %
C-MLP	61,6 %	71,0 %	56,3 %	63,0 %

Tabelle 2 - Zusammenfassung der durchschnittlichen Treffsicherheiten für die Objektklassifikation.

Das Training einer neuronalen Netzstruktur wird abgebrochen, wenn mindestens eine der drei Bedingungen erfüllt wird:

- Anzahl von Epochen $E = 1000$,
- Anzahl Validationchecks = 25,
- Validierungsfehler $\varepsilon_{val} \leq 10^{-3}$.

Die Anzahl von Validationchecks ist ein vorzeitiges Abbruchkriterium, das erkennt, wenn sich der Gesamtfehler der Netzschätzung nicht mehr verringert. Es wird abgebrochen, wenn der Gesamtfehler der Validierungsdaten 25 Epochen in Folge konstant bleibt oder sogar steigt.

Zur Bewertung der Ergebnisse der verschiedenen Netzstrukturen und Paradigmen werden drei Kriterien, die sich u. a. aus der zehnfachen Kreuzvalidierung ergeben, definiert:

- mittlerer Validierungsfehler $\bar{\varepsilon}_{val} = \frac{1}{10} \sum_{i=1}^{10} \varepsilon_{val_i}$,
- mittlere Erkennungsrate in der Konfusionsmatrix $\bar{c}_{test} = \frac{1}{10} \sum_{i=1}^{10} \frac{RP_i + RN_i}{RP_i + RN_i + FP_i + FN_i}$ in [%],
- Spannweite der Erkennungsrate $\Delta c_{test} = \max(\{c_{test}\}) - \min(\{c_{test}\})$ in [%]

4 Ergebnisse

Allgemein ist anhand des Kriteriums der durchschnittlichen Erkennungsrate (für alle 50 Netzstrukturen) festzustellen, dass die Variante B mit 73,3 % bzw. 71,0 % die besten Klassifikationsergebnisse erzielt. Im allgemeinen Vergleich zwischen den Paradigmen liegt das MLP mit über 5 % Abstand vor dem C-MLP, siehe Tabelle 2. Somit konzentriert sich die weitere Auswahl eines Klassifikationssystems auf eine MLP-Netzarchitektur der Variante B.

Speziell erweist sich der Testlauf der Variante B mit einem 257-50-40-5 Netz als der Erfolgreichste, denn die 257-50-40-5 MLP Netzstruktur erreicht nach der zehnmaligen Kreuzvalidierung die besten Werte in allen Kriterien: kleinster mittlerer Validierungsfehler mit $\bar{\varepsilon}_{val} = 0,066$, höchste mittlere Erkennungsrate bei $\bar{c}_{test} = 80,8$ % und geringste Spannweite in der Erkennungsrate Δc_{test} von 8,9 %. Damit vereint diese neuronale Netzstruktur neben den besten Testergebnissen auch die stabilsten Ergebnisse.

Anzumerken ist, dass mit einer höheren Neuronenzahl in den verdeckten Schichten eine größere Erkennungsrate erreicht wird und eine zweite verdeckte Schicht nicht zwangsläufig mit einer größeren Erkennungsrate einhergeht. Jedoch sorgt die Kombination mit einer zweiten verdeckten Schicht für verlässlichere Ergebnisse im Hinblick auf Δc_{test} . Ein geringer Δc_{test} -Wert deutet auf eine gelungene Kreuzvalidierung und damit auf ein generalisierendes Netz hin.

Um die Ergebnisse des Klassifikationssystems zur Erkennung der Objekttypen besser einzuschätzen, wird ein vom Fahrzeugfahrwerk unabhängiger Test vorgenommen. Dieser überprüft, ob das gefundene Klassifikationssystem durch die Fahrzeugfahrwerke von Pkw 1 und Pkw 2 beeinflusst wird und ob ein Fehler in der Datenverteilung vorliegt. Dazu wird das gefundene 257-50-40-5 MLP-Netz der Variante B mit den 424 Pkw 1-Events ohne weitere Abbruchbedingungen einmal 250 Epochen lang trainiert und anschließend gegen die 409 Pkw 2-Events

Konfusionsmatrix - Testdaten: GOLF

IST-Ausgabe \hat{y}	1	103 25.2%	18 4.4%	0 0.0%	0 0.0%	0 0.0%	85.1% 14.9%
	2	12 2.9%	38 9.3%	4 1.0%	1 0.2%	0 0.0%	69.1% 30.9%
	3	0 0.0%	0 0.0%	50 12.2%	11 2.7%	2 0.5%	79.4% 20.6%
	4	0 0.0%	2 0.5%	33 8.1%	88 21.5%	0 0.0%	71.5% 28.5%
	5	0 0.0%	2 0.5%	8 2.0%	3 0.7%	34 8.3%	72.3% 27.7%
			89.6% 10.4%	63.3% 36.7%	52.6% 47.4%	85.4% 14.6%	94.4% 5.6%
		1	2	3	4	5	
		SOLL-Ausgabe y					

(a) mit Pkw 1 Trainingsdaten

Konfusionsmatrix - Testdaten: BORA

IST-Ausgabe \hat{y}	1	97 22.9%	19 4.5%	0 0.0%	0 0.0%	0 0.0%	83.6% 16.4%
	2	8 1.9%	52 12.3%	0 0.0%	1 0.2%	0 0.0%	85.2% 14.8%
	3	1 0.2%	5 1.2%	47 11.1%	9 2.1%	10 2.4%	65.3% 34.7%
	4	0 0.0%	0 0.0%	40 9.4%	94 22.2%	2 0.5%	69.1% 30.9%
	5	0 0.0%	2 0.5%	4 0.9%	1 0.2%	32 7.5%	82.1% 17.9%
			91.5% 8.5%	66.7% 33.3%	51.6% 48.4%	89.5% 10.5%	72.7% 27.3%
		1	2	3	4	5	
		SOLL-Ausgabe y					

(b) mit Pkw 2 Trainingsdaten

Abbildung 3 - Konfusionsmatrizen der Fahrzeugkreuzvalidierung für das 257-50-40-5 MPL-Netz. Ausgaben: Temposchwelle (1), Bremsschwelle (2), Kanaldeckel (3), Musterbeispiel (4), Brückenfuge (5).

getestet. Danach wird der Test umgedreht mit den Pkw 2-Events als Trainingsdaten und den Pkw 1-Events als Testdaten wiederholt.

Die Ergebnisse der Fahrzeugkreuzvalidierung sind in Form von Konfusionsmatrizen in Abbildung 3 dargestellt. Beide Konfusionsmatrizen weisen im Vergleich zum gemischten Training der Variante B Erkennungsraten von 76,5 % und 75,9 %, die ca. 5 % geringer sind, aus. Den Konfusionsmatrizen ist zu entnehmen, dass es in beiden Matrizen die gleichen Objekttypen sind, die von FP(falsch-positiv)- bzw. FN(falsch-negativ)-klassifizierten Events betroffen sind. Zum einen sind dies die Temposchwellen (1) und die Bremsschwelle (2), die in Summe 30 (a) und 27 (b) falsch klassifizierte Events hervorbringen. Andererseits sind die Events der Kanaldeckel (3) und Musterbeispiele (4) untereinander von insgesamt 44 (a) und 49 (b) falsch zugewiesenen Events betroffen.

Insgesamt belegen die konstanten Erkennungsraten von über 75 % zur Klassifikation der Objekttypen, dass es möglich ist, ein Klassifikationssystem zu erstellen, das vom Fahrzeugfahrwerk und von der Geschwindigkeit unabhängig ist. Des Weiteren bestätigt die geringe Differenz der Erkennungsrate von 0,6 % eine Fahrzeugunabhängigkeit für die eingesetzte Datenbasis. Dies bedeutet, dass ein erfolgreich trainiertes und generalisierendes Netz als Klassifikationssystem gefunden wurde.

5 Diskussion

Die Suche eines geeigneten Klassifikationssystems ergibt, dass genau ein MLP-Netz und die Eingabevariante B (Ortsspektren der Beschleunigungsdaten) die besten Klassifikationsergebnisse liefern.

Es ist zu vermuten, dass das C-MLP unter der weiteren Einspeisung der Eingabedaten in die Ausgabeschicht leidet. Das Netz wird dadurch künstlich unterteilt, weshalb die verdeckten Schichten keinen bis einen geringen Einfluss auf die Ausgabedaten haben. Der Großteil der Klassifikation findet somit in der Ausgabeschicht statt, die mit fünf Neuronen keine korrekte

Klassifikation durchführen kann.

Die Variante B ist erfolgreicher, weil sie im Gegensatz zur Variante A auf den Ort substituiert ist. Im Vergleich zu den Varianten A und C betrachtet die Variante B die Erregerfrequenz der Objekte im Frequenzbereich anstelle der hervorgerufenen Beschleunigungen im Zeit- oder der Profildaten im Ortsbereich. Die Transformation in den Frequenzbereich macht die Variante B deshalb robuster gegenüber einer variierten Geschwindigkeit, gegenüber den Eigenschaften des Fahrzeugfahrwerks und gegenüber variierenden Objektanmessungen je Typ. Die Möglichkeit der Datentransformation wird bereits in vorangegangenen Untersuchungen benutzt, vgl. [3]. Die aktuelle Untersuchung bleibt jedoch auf die DFT beschränkt.

Bei der Betrachtung der Ergebnisse der Fahrzeugkreuzvalidierung für die 257-50-40-5 MLP-Netzstruktur der Variante B fällt auf, dass Tempo- und Bremsschwellen sowie Kanaldeckel und Musterbeispiele häufiger als andere Kombinationen vertauscht werden. Diese Ausprägungen sind darauf zurückzuführen, dass sich die genannten Objekttypen in ihren Abmessungen ähnlich sind. Ein Kanaldeckel ragt z. B. nur wenige Zentimeter aus der Fahrbahn, wodurch eine Verwechslung mit einem Musterbeispiel als idealer Fahrbahnuntergrund möglich ist.

In einem ersten Anwendungsszenario können Fahrzeughersteller einen Pkw-Prototypen mit der vorgestellten Messplattform ausstatten und diese an ihre vorhandene Kommunikationsschnittstelle anschließen. In einem zweiten Schritt lässt sich das prototypische, trainierte Klassifikationssystem im Echtzeitbetrieb auf einem PI testen. Die finale Klassifikationslösung lässt sich als Ergänzung der fahrzeugeigenen Sensorik einsetzen. Damit wäre ein Beitrag zum so genannten vernetzten Fahren in den Bereichen Sicherheit und Komfort möglich, vgl. ConnectedDrive von BMW [6]. Vorausfahrende können nachfolgende Fahrzeuge vor Straßenschäden oder spezifischen Fahrbahnobjekten warnen – Fahrzeuge, die ihrerseits dynamisch die Fahrwerkseigenschaften der veränderten Situation anpassen.

Literatur

- [1] LANGENBERG, T. UND JOKISCH, O.: *Smartphone-basierte Bewegungsanalyse im Fahrzeug*. Teil 1: Merkmalsanalyse von Beschleunigungsdaten, Tagungsband 26. Elektronische Sprachsignalverarbeitung Konferenz, S. 256 - 263, Eichstätt, 2015
- [2] GONZALEZ, L.C.; MARTINEZ, F. UND CARLOS, M.R.: *Identifying Roadway Surface Disruptions Based on Accelerometer Patterns*. In: IEEE Latin America Transactions, Vol. 12, No. 3, S. 455-460, South Brazil, 2014
- [3] WANG, S.; KODAGODA, S. UND KHUSHABA, R.: *Towards Speed-Independent Road-Type Classification*. In: IEEE 12th International Conference on Control, Automation, Robotics & Vision, S. 614-619, Guangzhou (China), 2015
- [4] Raspberry PI: *Raspberry Pi 2 Model B*. <https://www.raspberrypi.org/products/raspberry-pi-2-model-b/>, 06. Jan. 2016
- [5] STMicroelectronics: *LSM9DS0 - 3D accelerometer, 3D gyroscope, 3D magnetometer*. In: adafruit, Datenblatt, S. 41-53, 2013, <http://www.adafruit.com/datasheets/LSM9DS0.pdf>, 07. Okt. 2015
- [6] BMW AG: *ConnectedDrive Services*. <http://www.bmw.de/de/topics/faszination-bmw/connecteddrive/services-apps/connecteddrive-services.html>, 09. Nov. 2015