



Nicht nur Müdigkeit verursacht viele Unfälle, auch das Telefonieren mit dem Smartphone oder der Verzehr von Lebensmitteln während der Fahrt stellen ein hohes Risiko dar. ©

Dusan Petkovic / Shutterstock, puhha / Shutterstock

DEEP LEARNING BEI FAHRASSISTENZSYSTEMEN

Bildbasiertes CNN-System

ARRK Engineering kann Ablenkungen durch Mobiltelefone oder Nahrungsaufnahme automatisch erkennen und kategorisieren. Dazu wurden Bilder mit Infrarotkameras aufgenommen und für das maschinelle Anlernen von unterschiedlichen Convolutional Neural Network-Systemen (CNN) verwendet. So können zuverlässig verschiedene Szenarien am Steuer erkannt werden.

Seit einigen Jahren kommen in der Automobilindustrie bereits Systeme zum Einsatz, die bei auftretender Müdigkeit warnen. Dafür analysieren diese Fahrassistenten beispielsweise die Blickrichtung des Fahrers und erkennen automatisch Abweichungen vom üblichen Verhalten am Steuer. Existierende Warnsysteme können bisher nur bestimmte Gefahrensituationen richtig erfassen. „Bei einigen Tätigkeiten wie Essen, Trinken oder Telefonieren wird die Kopfhaltung kaum verändert, da der Fahrer weiter auf die Straße vor sich schaut,“ berichtet Benjamin Wagner, Senior Consultant für Fahrassistentensysteme bei ARRK Engineering. Deswegen hat ARRK Engineering eine Versuchsreihe zur Identifikation von unterschiedlichen Körperhaltungen durchgeführt, um die Nutzung von Mobiltelefonen und den Verzehr von Lebensmitteln automatisch zu erkennen. Damit alle Arten von visueller, aber auch manueller und kogniti-

ver Ablenkung korrekt erfasst werden, testete ARRK verschiedene Deep Learning Modelle und trainierte sie mit den ermittelten Daten.

Anlernen der Systeme

Für den Versuchsaufbau (Bild 1) wurden in einem Testwagen zwei Kameras mit aktiver Infrarotbeleuchtung jeweils links und rechts des Fahrers an der A-Säule installiert. Beide Kameras verfügten

über eine Frequenz von 30 Hz und lieferten 8-bit-Graustufenbilder mit einer Auflösung von 1280 x 1024 Pixeln. Die Kameras waren außerdem mit einem IR-Langpassfilter ausgestattet, um den größten Teil des Lichts aus dem sichtbaren Spektrum mit einer Wellenlänge unter 780 nm zu blockieren. „Dadurch stellten wir sicher, dass das empfangene Licht hauptsächlich von den IR-Strahlern stammt und deren volle Funktionalität sowohl bei Tag als auch bei Nacht ge-

»Existierende Warnsysteme können bisher nur bestimmte Gefahrensituationen richtig erfassen. Denn bei einigen Tätigkeiten wie Essen, Trinken oder Telefonieren wird die Kopfhaltung kaum verändert, da der Fahrer weiter auf die Straße vor sich schaut.«

Benjamin Wagner, Senior Consultant für Fahrassistentensysteme bei ARRK Engineering.

währleistet ist," so Wagner. Darüber hinaus vermied die Blockade des sichtbaren Tageslichts auch mögliche Schatteneffekte in der Fahrerkabine, die sonst zu Fehlern bei der Gesichtserkennung führen können. Damit die Bilder auf beiden Seiten zur gleichen Zeit aufgenommen wurden, wurde ein Raspberry Pi 3 Model B+ verwendet, das ein Triggersignal an beide Kameras im Moment der Aufnahme sendet.

Mit diesem Aufbau wurden die Bilder der Körperhaltungen von 16 Probanden in einem stehenden Auto aufgenommen. Um möglichst vielfältige Daten zu erzeugen, unterschieden sich die Probanden beispielsweise in Geschlecht, Alter oder Kopfbedeckung, aber auch verschiedene Mobiltelefonmodelle, Lebensmittel und Getränke wurden genutzt. „Für die jeweiligen Ablenkungsarten erstellten wir fünf Kategorien, in die später die Körperhaltungen einsortiert werden konnten. Dabei handelte es sich um: ‚keine sichtbare Ablenkung‘, ‚Telefonieren am Smartphone‘, ‚manuelle Bedienung des Smartphones‘, ‚Essen sowie Trinken‘, aber auch das ‚Halten von Lebensmitteln oder Getränken‘“, erläutert Wagner. Nach der Aufnahme wurden die Bilder der beiden Kameras entsprechend kategorisiert und danach für das maschinelle Anlernen des Systems genutzt.

Training und Test der Bilderkennungssysteme

Zur Erkennung der Körperhaltungen wurden vier modifizierte CNN-Modelle verwendet: ResNeXt-34, ResNeXt-50, VGG-16 und VGG-19. Die beiden letztgenannten repräsentieren in der Praxis gängige Modelle wohingegen ResNeXt-34 und ResNeXt-50 eine dedizierte Struktur zur Verarbeitung von parallelen Pfaden enthalten. Für das Training der Systeme führte ARRK 50 Durchgänge mit dem Adam-Optimizer durch – einem Optimierungsalgorithmus mit adaptiver Lernrate. Dabei musste das CNN-Modell in jedem Durchgang die Körperhaltungen der Probanden in die vorher erstellten Kategorien einordnen. Mit jedem weiteren Schritt wurde diese Kategorisierung über ein Gradientenverfahren so angepasst, dass die Fehlerrate kontinuierlich sinkt. Zum Abschluss des Prozesses wurde ein dedizierter



Bild 1: Für den Versuchsaufbau wurden in einem Testwagen zwei Kameras mit aktiver Infrarotbeleuchtung jeweils links und rechts des Fahrers installiert.

© ARRK Engineering

Testdatensatz für die Berechnung der Wahrheitsmatrix verwendet, um die Fehlerquote pro Fahrerhaltung für jedes CNN-Modell zu analysieren. „Die Verwendung von zwei Kameras mit jeweils separat geschultem CNN-Modell ermöglicht eine optimierte Fallunterscheidung für die linke und die rechte Gesichtshälfte“, führt Wagner aus. „Dank dieses Vorgehens konnten wir das System mit der besten Performance ermitteln, um den Gebrauch von Mobiltelefonen und den Verzehr von Lebensmitteln für einen großen Bereich von Kopfwin-

keln zu erkennen.“ Die Gesamtauswertung ergab, dass die CNN-Modelle ResNeXt-34 und ResNeXt-50 die höchste Testgenauigkeit von 92,88 Prozent für die linke Kamera und 90,36 Prozent für die rechte Kamera erreichten – ein mit bisherigen Lösungen zur Erkennung von Müdigkeit absolut wettbewerbsfähiges Ergebnis.

Mit diesen Informationen erweiterte ARRK abschließend seine Trainingsdatenbank und kann beispielsweise bereits auf etwa 20.000 gelabelte Augendatensätze zugreifen. Darauf aufbauend ist die Entwicklung eines automatisierten, kamerabasierten Systems zur Validierung von Fahrerbeobachtungssystemen möglich. Für eine noch geringere Fehlerquote des Systems planen die Experten bereits einen Schritt weiter. „Um die Genauigkeit weiter zu verbessern, werden wir in einem nächsten Projekt andere CNN-Modelle einsetzen“, resümiert Wagner. Neben der Bewertung weiterer Klassifikationsmodelle werden sie dabei auch analysieren, ob die Integration von zugehörigen Objektpositionen aus dem Kamerabild zu zusätzlichen Verbesserungen führen kann. In diesem Zusammenhang werden Ansätze zu berücksichtigen sein, die auf der Erkennung von sogenannten Bounding Boxes und der semantischen Segmentierung basieren. Letztere ermöglichen neben einer Klassifizierung auch verschiedene Detailstufen hinsichtlich der Lokalisierung von Objekten. Auf diese Weise kann ARRK die Genauigkeit eines Fahrerassistenzsystems zur automatischen Erkennung von Ablenkungen am Steuer verbessern. ■ (oe)

www.ark-engineering.com

INFO

ARRK Engineering ist Teil des internationalen ARRK Firmenverbundes und spezialisiert auf sämtliche Leistungen rund um die Produktentwicklung. Mithilfe der Kompetenzen in Elektronik & Software, CAE, Material, Akustik, Composite, Karosserie, Antrieb, Fahrwerk, Interieur & Exterieur, Optische Systeme, Passive Sicherheit und Thermomanagement entwickelt das Unternehmen als langjähriger strategischer Entwicklungspartner für seine Kunden ganzheitlich und eigenverantwortlich Produkte. Zusammen mit den Schwesterunternehmen setzt ARRK Engineering Produktentwicklungen von der virtuellen Entwicklung bis hin zum Prototypen und der Produktion in Kleinserien um. Die Standorte der ARRK Engineering Division liegen in Deutschland, Rumänien, UK, Japan und China. Das Headquarter der Engineering Division ist die P+Z Engineering GmbH in Deutschland; ARRK Engineering beschäftigt über 1.200 Mitarbeiter.

Nach Unterlagen der P+Z Engineering GmbH, 80807 München