

(12) NACH DEM VERTRAG ÜBER DIE INTERNATIONALE ZUSAMMENARBEIT AUF DEM GEBIET DES
PATENTWESENS (PCT) VERÖFFENTLICHTE INTERNATIONALE ANMELDUNG

(19) Weltorganisation für geistiges Eigentum
Internationales Büro

(43) Internationales Veröffentlichungsdatum
26. August 2021 (26.08.2021)



(10) Internationale Veröffentlichungsnummer
WO 2021/165077 A1

(51) Internationale Patentklassifikation:
G06K 9/00 (2006.01) *G06K 9/62* (2006.01)

(21) Internationales Aktenzeichen: PCT/EP2021/052931

(22) Internationales Anmeldedatum:
08. Februar 2021 (08.02.2021)

(25) Einreichungssprache: Deutsch

(26) Veröffentlichungssprache: Deutsch

(30) Angaben zur Priorität:
10 2020 201 939.8
17. Februar 2020 (17.02.2020) DE

(71) Anmelder: **ROBERT BOSCH GMBH** [DE/DE]; Post-
fach 30 02 20, 70442 Stuttgart (DE).

(72) Erfinder: **RITTEL, Michael**; Bahnhofstr. 33, 71706
Markgroeningen (DE). **OEHLERKING, Jens**; Grossbee-
renstrasse 18, 70499 Stuttgart (DE). **GLADISCH, Chris-
toph**; Calwerstr. 27/1, 71272 Renningen (DE). **GROH,
Konrad**; Klopstockstrasse 21, 70193 Stuttgart (DE). **WIL-
LERS, Oliver**; Talstr. 6, 71404 Korb (DE). **SUDHOLT,
Sebastian**; Erfurter Str. 9/1, 74196 Neuenstadt Am Kocher
(DE). **WOEHRLE, Matthias**; Erwin-Baelz-Str. 32, 74321
Bietigheim-Bissingen (DE). **HEINZEMANN, Christian**;
Jommellistr. 19, 71640 Ludwigsburg (DE).

(81) Bestimmungsstaaten (soweit nicht anders angegeben, für
jede verfügbare nationale Schutzrechtsart): AE, AG, AL,
AM, AO, AT, AU, AZ, BA, BB, BG, BH, BN, BR, BW, BY,
BZ, CA, CH, CL, CN, CO, CR, CU, CZ, DE, DJ, DK, DM,
DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT,
HN, HR, HU, ID, IL, IN, IR, IS, IT, JO, JP, KE, KG, KH,
KN, KP, KR, KW, KZ, LA, LC, LK, LR, LS, LU, LY, MA,

(54) Title: METHOD AND DEVICE FOR EVALUATING IMAGE CLASSIFIERS

(54) Bezeichnung: VERFAHREN UND VORRICHTUNG ZUR BEWERTUNG EINES BILDKLASSIFIKATORS

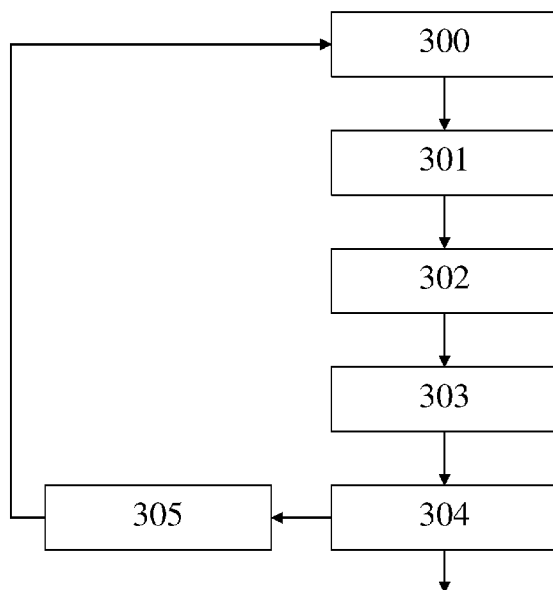


Fig. 1

(57) **Abstract:** The invention relates to a computer-implemented method for evaluating image classifiers (60), wherein a classifier output (y) of the image classifier (60) is used to actuate an at least partly autonomous robot (100, 220). The evaluation method has the following steps: • ascertaining (300) a first data set containing image data which is assigned annotations, said annotations containing information on the scene displayed in the respective image and/or on the image regions to be classified and/or on movement information of the robot (100, 220); • ascertaining (301) scene regions (212), which can be reached by the robot (100, 220), on the basis of the annotations; • ascertaining (302) relevance values for image regions to be classified by the image classifier; • classifying (303) the image data of the first image data set by means of the image classifier; and • evaluating (304) the image classifier on the basis of image regions correctly classified by the image classifier (60), image regions incorrectly classified by the image classifier, and the calculated relevance values of the corresponding image regions.

(57) **Zusammenfassung:** Computerimplementiertes Verfahren zur Bewertung eines Bildklassifikators (60), wobei eine Klassifikatorausgabe (y) des Bildklassifikators (60) zur Ansteuerung eines zumindest teilweise autonomen Roboters (100, 220) verwendet wird, wobei das Verfahren zur Bewertung folgende Schritte umfasst: • Ermittlung (300) eines ersten Datensatzes, wobei der erste Datensatz Bilddaten enthält, wobei den Bilddaten Annotationen zugewiesen sind, wobei die Annotationen Informationen über die im jeweiligen Bild abgebildete Szene und/oder über zu klassifizierende Bildbereiche und/oder über Bewegungsinformationen des Roboters (100, 220) enthalten; • Ermittlung (301) von durch den Roboter (100,



WO 2021/165077 A1

MD, ME, MG, MK, MN, MW, MX, MY, MZ, NA, NG, NI,
NO, NZ, OM, PA, PE, PG, PH, PL, PT, QA, RO, RS, RU,
RW, SA, SC, SD, SE, SG, SK, SL, ST, SV, SY, TH, TJ, TM,
TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, WS, ZA, ZM,
ZW.

- (84) Bestimmungsstaaten** (soweit nicht anders angegeben, für jede verfügbare regionale Schutzrechtsart): ARIPO (BW, GH, GM, KE, LR, LS, MW, MZ, NA, RW, SD, SL, ST, SZ, TZ, UG, ZM, ZW), eurasisches (AM, AZ, BY, KG, KZ, RU, TJ, TM), europäisches (AL, AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HR, HU, IE, IS, IT, LT, LU, LV, MC, MK, MT, NL, NO, PL, PT, RO, RS, SE, SI, SK, SM, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, KM, ML, MR, NE, SN, TD, TG).

Veröffentlicht:

- mit internationalem Recherchenbericht (Artikel 21 Absatz 3)

220) erreichbaren Bereichen (212) der Szenen basierend auf den Annotationen; • Ermittlung (302) von Relevanzwerten für vom Bildklassifikator zu klassifizierenden Bildbereiche; • Klassifizierung (303) der Bilddaten des ersten Bilddatensatzes mittels des Bildklassifikators; • Bewertung (304) des Bildklassifikators basierend auf durch den Bildklassifikator (60) korrekt klassifizierten Bildbereichen und falsch klassifizierten Bildbereichen, sowie den berechneten Relevanzwerten der entsprechenden Bildbereiche.

5 BeschreibungTitelVerfahren und Vorrichtung zur Bewertung eines Bildklassifikators

10 Die Erfindung betrifft ein Verfahren zum Bewerten eines Bildklassifikators, ein Verfahren zum Trainieren eines Bildklassifikators, ein Verfahren zum Betreiben eines Bildklassifikators, eine Trainingsvorrichtung, ein Computerprogramm, ein Ansteuersystem und ein maschinenlesbares Speichermedium.

15 Stand der Technik

"Reachability Analysis and its Application to the Safety Assessment of Autonomous Cars", Matthias Althoff, Dissertation, Technische Universität München, 2010 offenbart ein Verfahren zum Bestimmen von erreichbaren Bereichen eines autonomen Fahrzeugs.

20 Vorteile der Erfindung

25 Bildklassifikatoren stellen eine Schlüsseltechnologie zum Betreiben von zumindest teilweise autonomen und/oder mobilen Robotern da. Es hat sich gezeigt, dass aus Daten gelernte Bildklassifikatoren, insbesondere Neuronale Netze, derzeit die besten Klassifikationsleistungen erbringen.

30 Die Untersuchung dieser machine-learning-basierten Bildklassifikatoren gestaltet sich jedoch schwierig, da oftmals nicht offensichtlich ist, wie ein Bildklassifikator zu seiner Klassifikation gelangt. Insbesondere bei sicherheitskritischen Anwendungen macht es diese Eigenschaft schwer mit Sicherheit zu bestimmen, dass ein Produkt, welches einen machine-learning-basierten Bildklassifikator enthält, in seiner Umgebung sicher agiert.

35

Der Vorteil des Verfahrens mit Merkmalen gemäß dem unabhängigen Anspruch 1 liegt darin, einen Einblick in die Funktionsweise eines Bildklassifikators zu erhalten. Insbesondere ermöglicht das Verfahren die Ermittlung von unter Sicherheitsaspekten relevanten Elementen eines Bildes, die der Bildklassifikator erkennen soll. Dies erlaubt einen Einblick in die Genauigkeit der Klassifikationen des Klassifikators. Im Umkehrschluss kann das Verfahren verwendet werden, um bestimmen zu können, ob ein mobiler Roboter, der seine Navigation basierend auf Ausgaben eines Bildklassifikators vornimmt, sicher genug ist, um ihn betreiben zu können.

Offenbarung der Erfindung

In einem ersten Aspekt beschäftigt sich die Erfindung mit einem computerimplementierten Verfahren zur Bewertung eines Bildklassifikators, wobei eine Klassifikatorausgabe des Bildklassifikators zur Ansteuerung eines zumindest teilweise autonomen Roboters (100, 220) bereitgestellt wird, wobei das Verfahren zur Bewertung folgende Schritte umfasst:

- Ermittlung (300) eines ersten Datensatzes, wobei der erste Datensatz Bilder enthält, wobei den Bildern Annotationen zugewiesen sind, wobei die Annotationen Informationen über die im jeweiligen Bild abgebildete Szene und/oder über zu klassifizierende Bildbereiche und/oder über Bewegungsinformationen des Roboters (100, 220) enthalten;
- Ermittlung (301) von durch den Roboter (100, 220) erreichbaren Bereichen (212) der Szenen basierend auf den Annotationen;
- Ermittlung (302) von Relevanzwerten für vom Bildklassifikator zu klassifizierenden Bildbereiche;
- Klassifizierung (303) der Bilder des ersten Bilddatensatzes mittels des Bildklassifikators;
- Bewertung (304) des Bildklassifikators basierend auf durch den Bildklassifikator (60) korrekt klassifizierten Bildbereichen und falsch klassifizierten Bildbereichen, sowie den berechneten Relevanzwerten der entsprechenden Bildbereiche.

Unter einem Bildklassifikator kann eine Vorrichtung verstanden werden, die dahingehend ausgeprägt ist, dass sie Bilder (auch: Bilddaten) entgegennehmen kann und eine Klassifikationsausgabe erzeugen kann, die die Bilddaten oder Teile davon charakterisiert. Zum Beispiel kann ein Bildklassifikator verwendet werden, um zu bestimmen, in welchen Teilen eines Eingabebildes sich Objekte befinden. Bei einem mobilen Roboter, wie beispielsweise einem autonomen Fahrzeug, kann ein Bildklassifikator so verwendet werden, dass er andere Straßenteilnehmer detektiert. Die entsprechende Klassifikatorausgabe kann dann verwendet werden, um den Roboter anzusteuern. Zum Beispiel kann mit Hilfe der Klassifikatorausgabe eine Trajektorie bestimmt werden, auf der der Roboter kollisionsfrei durch seine Umwelt bewegt. D.h. vorzugsweise zeigen die Bilddaten eine Umgebung des Roboters.

Neben der Objektdetektion kann ein Bildklassifikator auch für andere Klassifikationsaufgaben verwendet werden, zum Beispiel für semantische Segmentierung. Hierbei klassifiziert der Bildklassifikator jeden gewünschten Punkt in einem Eingabebild, zum Beispiel jeden Pixel eines Kamerabildes, in eine gewünschte Klasse. Dies kann zum Beispiel dafür verwendet werden, dass ein mobiler Roboter basierend auf einem Eingabebild die Grenzen des fahrbaren Bereichs der Umgebung erkennt und basierend darauf eine Trajektorie plant.

Ein Bildklassifikator kann ein Modell aus dem Bereich des Machine Learning, wie etwa ein neuronales Netz, enthalten. Das Modell kann verwendet werden, um die Eingabe des Bildklassifikators zu klassifizieren. Des Weiteren kann der Bildklassifikator Vor- und/oder Nachverarbeitungsverfahren. Im Falle einer Objektdetektion kann ein Nachverarbeitungsverfahren zum Beispiel eine Non-Maximum Suppression sein, die verwendet werden kann, um verschiedene Bounding Boxen gleicher Objekte zu fusionieren.

Für einen Bildklassifikator können unterschiedliche Arten von Bildern als Eingabedaten verwendet werden, insbesondere Sensordaten, beispielsweise von einem Kamerasensor, einem Radarsensor, einem LIDAR-Sensor, einem Ultraschallsensor oder einem Infrarotkamerasensor. Auch Audioaufnahmen von Mikrofonen können als Bilddaten dargestellt und als Eingabe für einen Bildklassifikator verwendet werden, zum Beispiel in Form von Spektralbildern. Es ist weiterhin

vorstellbar, dass mehrere Arten von Sensordaten kombiniert werden können, um ein Eingabedatum für den Bildklassifikator zu erhalten.

Alternativ können Bilddaten mit Hilfe von computergestützten Maßnahmen synthetisch erzeugt werden. Zum Beispiel können Bilder basierend auf physikalischen Modellen berechnet bzw. gerendert werden.

Die Bilder, die zur Eingabe für den Bildklassifikator verwendet werden, können von einem Sensor direkt aufgenommen und an den Bildklassifikator weitergegeben werden. Alternativ können Bilddaten vor der Klassifikation aufgenommen bzw. erzeugt werden und anschließend vorzugsweise auf einem Speichermedium zwischengespeichert werden, bevor sie an den Bildklassifikator weitergegeben werden. Als Speichermedium können hierbei insbesondere Festplatten, Flash Drives oder Solid State Disks verwendet werden. Auch können Bilddaten in einem dynamischen Speicher vorgehalten werden.

Basierend auf der Ausgabe des Bildklassifikators kann ein Ansteuersignal ermittelt werden, mit dem ein zumindest teilweise autonomer Roboter angesteuert werden kann. Unter einem zumindest teilweise autonomen Roboter kann ein Roboter verstanden werden, der zumindest zeitweise ohne Steuerung eines Menschen eine Aufgabe selbständig durchführt. Hierfür kann er zum Beispiel Sensoren und Aktuatoren verwenden. Ein teilweise autonomer Roboter kann zum Beispiel ein autonom fahrendes Fahrzeug, ein Rasenmäherroboter, ein Saugroboter oder eine Drohne sein. Im Folgenden wird unter dem Begriff Roboter ein zumindest teilweise autonomer Roboter verstanden.

Das Verfahren benötigt einen ersten Bilddatensatz, mit dessen Hilfe die Bewertung durchgeführt werden kann. Unter einem Bilddatensatz kann eine Menge von Bilddaten verstanden werden, wobei den Bilddaten spezifische Informationen in Form von Annotationen zugeordnet sein können. Unter Annotation eines Bilddatums kann hierbei eine Menge von Informationen verstanden werden, die das Bilddatum oder Teile davon beschreiben oder weitere zusätzliche Informationen über das Bild beinhalten.

Bilddaten können Szenen abbilden, wobei Szenen Objekte enthalten können. Im Falle von mit einem oder mehreren Sensoren aufgenommenen Bilddaten kann unter einer Szene eine Situation der realen Welt, insbesondere der Umgebung des Roboters, verstanden werden. Zum Beispiel kann eine Szene eine Menge von Objekten in einer Straßensituation darstellen. Unter Objekten können in diesem Fall zum Beispiel andere Straßenteilnehmer verstanden werden.

Im Falle von synthetisch erzeugten Bilddaten kann unter einer Szene die virtuelle Welt verstanden werden, auf deren Basis ein Bilddatum synthetisiert wurde. In diesem Fall können die Objekte als virtuelle Elemente der Szene verstanden werden.

Bilddaten können Annotationen zugeordnet werden, wobei Annotationen Informationen über die jeweils abgebildete Szene und/oder Bildbereiche umfassen können. Zum Beispiel kann eine Annotation eine Menge von Bounding Boxen enthalten, die die Position vom im Bilddatum abgebildeten Objekten beschreiben. Alternativ oder zusätzlich ist vorstellbar, dass die Annotation pixelgenaue Informationen bezüglich der Klasse eines Pixels (d.h. eine semantische Segmentierung) des Bilddatums beinhaltet.

Alternativ oder zusätzlich ist vorstellbar, dass eine Annotation Informationen von Wetter- und/oder Umwelteinflüssen beinhaltet, die bei der Aufnahme des spezifischen Bilddatums vorlagen, z.B. Regen, Sonneneinstrahlung, Tageszeit oder Bodenbeschaffenheit.

Alternativ oder zusätzlich ist vorstellbar, dass eine Annotation Informationen über die Szene enthält, in der das Bild aufgenommen wurde. Für den Fall einer realen Szene, kann die Annotation zum Beispiel Informationen über die relative Position des Sensors im Verhältnis zu anderen Objekten der Szene enthalten. Diese Information kann später beispielsweise genutzt werden, um die 3-dimensionale Position eines Objekts zu bestimmen, das in einem Bilddatum (z.B. einem Kamerabild) 2-dimensional abgebildet ist. Alternativ ist auch denkbar, dass die 3-dimensionalen Positionsinformationen von Objekten in der Szene direkt in der Annotation enthalten sind, zum Beispiel in Form eines relativen Vektors vom Sensor zum Objekt.

Der erste Bilddatensatz kann für das Verfahren vorzugsweise so gewählt werden, dass zur Aufzeichnung ein System verwendet wird, das dem späteren Roboter entspricht oder ähnelt. Zum Beispiel kann zur Bewertung des Bildklassifikators, der in einem autonomen Fahrzeug verwendet werden soll, der erste Bilddatensatz so aufgenommen werden, dass ein Testfahrer das Fahrzeug derart steuert, dass gewünschte Bilddaten von den Sensoren des Fahrzeugs aufgezeichnet werden können. Alternativ ist vorstellbar, dass der erste Datensatz von einem Fahrzeug aufgenommen wird, das von der Sensorik baugleich zu dem Fahrzeug ist, für den der Bildklassifikator bewertet werden soll.

Alternativ ist vorstellbar, dass der erste Bilddatensatz synthetisch mit Hilfe eines computergestützten Modells erzeugt wird. Hierbei kann das Modell vorzugsweise so gewählt werden, dass es dem Roboter in Form, physikalischen Eigenschaften und Sensorik zumindest ähnelt. Die Szene kann in diesem Fall als die Kombination von Anordnung und Eigenschaften von virtuellen Objekten verstanden werden, mit deren Hilfe synthetische Bilddaten erzeugt werden können.

Alternativ kann der erste Bilddatensatz auch aus bestehenden Quellen gewonnen werden. Zum Beispiel existiert eine Reihe von frei zugänglichen Datensätzen im Internet, die zum Zwecke der Beurteilung eines Bildklassifikators benutzt werden können.

Die für das Verfahren benötigten Annotationen können entweder manuell oder zumindest teilautomatisiert für die verschiedenen Bilddaten erzeugt werden. Vorzugsweise enthalten die Annotationen relative Informationen bezüglich der zu klassifizierenden Element eines Bilddatums und dem System, was zur Aufzeichnung verwendet wird/wurde. Zum Beispiel kann ein Fahrzeug derart ausgeprägt sein, dass es einen kamerabasierten Bilddatensatz aufnehmen kann, der im Anschluss zur Bewertung eines Bildklassifikators verwendet werden kann, der im Fahrzeug oder einem baugleichen Fahrzeug später verwendet werden soll. Die Annotationen der Bilddaten können in diesem Fall zum Beispiel Bounding Boxen von in der Umgebung des Fahrzeugs zu detektierenden Objekten beinhalten. Zusätzlich können sie Informationen darüber enthalten, welche Position die in ei-

nem spezifischen Bilddatum zu erkennenden Objekte im Verhältnis zum Fahrzeug haben. Diese Informationen können später genutzt werden, um den Relevanzwert eines Objekts zu bestimmen.

5 Alternativ können bei synthetisch erzeugten Bilddaten die Modelldaten des synthetischen Modells direkt als Informationen in die Annotationen aufgenommen werden. Zum Beispiel können die im vorherigen Absatz beschriebenen Daten mit Hilfe eines Computers entsprechende simuliert werden. Hierfür wird ein virtuelles Modell des Sensors benötigt, sowie seine Position in der simulierten Szene.

10 Diese Position und/oder Positionen von simulierten Objekten, die später vom Bildklassifikator erkannt werden sollen, können in diesem Fall direkt in die Annotation mit aufgenommen werden.

15 Vorzugsweise enthalten die Annotationen darüber hinaus Informationen bezüglich z.B. der Geschwindigkeit des Roboters, der Beschleunigung, der Lenkwinkel, der Antriebsstellungen oder der geplanten Trajektorie, wobei jede dieser Informationen Werte abbildet, die zum Zeitpunkt der Aufnahme eines Bilddatums vorliegen oder vorlagen. Diese Informationen sind vorzugsweise, soweit sinnvoll, auch für die Objekte der Szene in der Annotation enthalten.

20 Mit Hilfe der Bewegungsinformationen können dann Bereiche einer Szene ermittelt werden, die der Roboter zum Zeitpunkt der Aufnahme in einer bestimmten Zeit erreichen hätte können. Diese Bereiche können z.B. mit Hilfe der Time-To-Collision mit anderen Objekten und/oder der Time-To-React und/oder Time-To-Brake und/oder Time-To-Steer und/oder Time-To-Kickdown ermittelt werden. Die Bereiche können als sicherheitskritische Bereich verstanden werden, in denen der Roboter andere Objekte mit hoher Genauigkeit und Zuverlässigkeit erkennen können muss, um zum Beispiel eine sichere Trajektorie zu planen.

30 Ein zu klassifizierender Bildbereich kann als zumindest Teil eines Bilddatums verstanden werden, für den der Bildklassifikator ein bestimmtes Objekt oder eine bestimmte Klasse präzisieren soll, was durch den Bildbereich abgebildet ist. Im Falle einer Objektdetektion können die zu klassifizierenden Bereiche als Abbildungen der Objekte verstanden werden, die vom Bildklassifikator detektiert werden sollen. Im Falle einer semantischen Segmentierung können die Bildbereiche

35

als die Pixel eines Bildes verstanden werden, wobei jeder Pixel einem Objekt zugeordnet werden kann.

Den zu klassifizierenden Bildbereichen können dann Relevanzwerte zugeordnet werden. Vorzugsweise erhält jeder zu klassifizierende Bereich einen Relevanzwert. Unter Relevanzwert kann hier ein Wert verstanden werden, der angibt, in wie weit eine Fehlklassifikation des Bildklassifikators für diesen Bildbereich kritisch für das Verhalten des Roboters werden kann, der den Bildklassifikator verwendet. Zum Beispiel können Bildbereichen, die sehr weit vom Roboter entfernte Objekte abbilden, kleine Relevanzwerte zugeordnet werden. Bildbereichen, die dem Roboter nahegelegene Objekte abbilden, können hingegen hohen Relevanzwert erhalten, da eine Fehlklassifikation von ihnen größeren Einfluss auf den Roboter hätte.

Relevanzwerte können durch skalare Werte abgebildet werden. Dabei kann ein Relevanzwert binär oder reellwertig sein.

Für alle oder Teile der zu klassifizierenden Bildbereiche kann dann bestimmt werden, ob sie vom Bildklassifikator korrekt klassifiziert wurden. Die Bewertung des Bildklassifikators kann dann vorzugsweise auf Basis der Relevanzwerte der falsch klassifizierten Bildbereiche erfolgen. Zum Beispiel kann die Bewertung in Form einer Summe oder eines Durchschnitts der Relevanzwerte der falsch klassifizierten Bereiche erfolgen.

Der Vorteil dieser Herangehensweise ist, dass ein numerischer und objektiver Wert bestimmt werden kann, der verwendet werden kann, um eine Entscheidung zu treffen, ob der Bildklassifikator als Teil der Ansteuerung des Roboters verwendet werden kann. Dies ermöglicht einen differenzierten Einblick in die Funktionsweise eines wie oben beschriebenen Bildklassifikators. Dies ist eine erhebliche Verbesserung, da insbesondere machine-learning-basierte Bildklassifikatoren Black Box Verhalten aufweisen, das anderweitig nicht befriedigend einsehbar ist.

Im ersten Aspekt der Erfindung ist weiterhin vorstellbar, dass die zu klassifizierenden Bildbereiche jeweils einem Objekt zugeordnet sind.

Der Vorteil dieser Herangehensweise ist, dass die Relevanz eines Bildbereichs die Relevanz des Objekts widerspiegeln kann. Dies erlaubt die Bewertung der Relevanz von Objekten in einer Szene basierend auf den Bildbereichen. Im Gegenzug erlaubt dies einen detaillierten Einblick in das Verhalten des Bildklassifikators zur Erkennung von Objekten in einer Szene.

Im ersten Aspekt der Erfindung ist weiterhin vorstellbar, dass der Schritt der Ermittlung der durch den Roboter erreichbaren Bereiche auf Bewegungsinformationen des Roboters basiert.

Die Bewegungsinformationen des Roboters, wie etwa Geschwindigkeit oder Beschleunigung, erlauben eine Bestimmung, in welche Bereiche der Szene der Roboter sich überhaupt plausiblerweise bewegen kann. Objekte in diesen Bereichen sollten daher mit hoher Genauigkeit vorhergesagt werden können, um den Roboter so anzusteuern, dass er zum Beispiel nicht mit den Objekten kollidiert. Die Bewegungsinformationen können, falls vorhanden, aus der Annotation extrahiert werden oder mit Hilfe der Bilddaten geschätzt werden. Zum Beispiel können mehrere aufeinanderfolgende Bilder des ersten Bilddatensatzes verwendet werden, um die Geschwindigkeit des Roboters zu schätzen. Im Falle von beispielsweise stationären Fertigungsroboter können Informationen über die erreichbaren Bereiche des Roboters auch aus Datenblättern des Roboters gewonnen werden.

Der Vorteil dieser Herangehensweise ist daher, dass Bereiche bestimmt werden können, in denen der Bildklassifikator Objekte zuverlässig erkennen können sollte. Da andere Bereiche der Szene weniger oder nicht relevant sein können, erlaubt dieses Verfahren daher einen detaillierten und zielgerichteten Einblick in die Arbeitsweise des Bildklassifikators und die Bewertung erlaubt eine bessere Abschätzung der Erkennungsleistung (auch: Performanz) des Bildklassifikators. Dies liegt daran, da in bekannten Verfahren die Erkennungsleistung eines Bildklassifikators auf allen Bildbereichen eines Bildes geschätzt werden. Über die ermittelten erreichbaren Bereiche lässt sich daher die Erkennungsleistung des Roboters in Bezug auf einen sicheren und fehlerfreien Betrieb viel besser bewerten.

Im ersten Aspekten der Erfindung ist weiterhin vorstellbar, dass der Schritt der Ermittlung der Relevanzwerte folgende Schritte umfasst:

- Ermittlung von Tiefeninformationen der Objekte;
- Bestimmung eines Verhältnisses von Objektpositionen zu erreichbaren Bereichen basierend auf den ermittelten Tiefeninformationen;
- Ermittlung der Relevanzwerte basierend auf dem Verhältnis.

5

Der Vorteil dieser Herangehensweise ist, dass Bildbereichen beispielsweise nur dann ein Relevanzwert ungleich Null zugewiesen werden kann, wenn die entsprechenden Objekte tatsächlich mit dem Roboter sicherheitskritisch interagieren können. Beispielsweise ist die Erkennung eines Fußgängers für die Trajektorienplanung eines Roboters unter sicherheitskritischen Maßstäben irrelevant, falls der Roboter zum Beispiel nicht schneller als 30 km/h fahren kann und der zu erkennende Fußgänger zum Beispiel mehr als 500m entfernt ist. Im Umkehrschluss kann das Verhalten des Bildklassifikators in Bezug zur Sicherheit des Roboters durch das vorgestellte Verfahren viel genauer bewertet werden.

10

15

Im ersten Aspekt der Erfindung ist weiterhin vorstellbar, dass der Schritt der Bewertung des Bildklassifikators das Bestimmen eines Bewertungsmaßes umfasst, weiter wobei das Verfahren zur Bewertung des Bildklassifikators folgenden zusätzlichen Schritt enthält:

20

- Nachtrainieren des Bildklassifikators basierend auf einem zweiten Bilddatensatz, falls das Bewertungsmaß schlechter ist als ein vorgegebener Bewertungsschwellenwert.

25

Beispielsweise kann das Bewertungsmaß derart gewählt sein, dass es umso größer ist, je schlechter die Performanz des Bildklassifikators ist. In diesem Fall würde also nachtrainiert werden, wenn das Bewertungsmaß größer als der Bewertungsschwellenwert ist.

30

Das Bewertungsmaß kann, wie oben beschrieben, beispielsweise die Summe oder der Durchschnitt der Relevanzwerte aller fehlklassifizierten Bildbereiche sein.

Unter nachtrainieren des Bildklassifikators kann ein Verfahren verstanden werden, das die Parameter des Bildklassifikators mit Hilfe des zweiten Bilddatensatzes so anpasst, dass die Erkennungsgenauigkeit des Bildklassifikators mit Hilfe

des zweiten Bilddatensatz weiter verbessert wird. Hierzu kann der zweite Bilddatensatz wiederum Annotationen enthalten, die zur Anpassung der Parameter mit Hilfe eines überwachten Lernverfahrens verwendet werden können. Der zweite Bilddatensatz kann hierbei mit den gleichen Verfahren wie der erste Bilddatensatz ermittelt werden. Weiterhin ist denkbar, dass der zweite Bilddatensatz zumindest Teile der Bilddaten und/oder Annotationen des ersten Bilddatensatzes enthält.

Der Vorteil dieser Herangehensweise ist, dass durch die verbesserte Erkennungsgenauigkeit die Sicherheit und damit die Erkennungsleistung des Roboters wesentlich gesteigert werden kann. Dies bedingt ein besseres Gesamtsystem, das in der realen Welt besser und zuverlässiger funktioniert.

Im ersten Aspekt der Erfindung ist weiterhin vorstellbar, dass die beschriebenen Schritte iterativ wiederholt werden, bis eine vorgegebene Anzahl von Iterationen durchgeführt wurde und/oder der vorgegebene Bewertungsschwellenwert unterschritten wurde.

Der Vorteil dieser Herangehensweise ist, dass die Erkennungsleistung des Bildklassifikators so lange verbessert werden kann, bis die Erkennungsleistung ausreichend ist, um ihn in einem realen Produkt betreiben zu können. Zusätzlich bietet diese Herangehensweise auch den Vorteil, dass ein gerade genügendes Bewertungsergebnis weiter verbessert werden kann und so ein gewisser Sicherheitspuffer bezüglich der Erkennungsgenauigkeit des Bildklassifikators erzielt werden kann. Weiterhin ist denkbar, dass der Bildklassifikator mit zwischen den Iterationen unterschiedlichen zweiten Bilddatensätzen nachtrainiert wird, um die Erkennungsleistung weiter zu steigern.

Bei dem vorgestellten iterativen Vorgehen können Bilddaten des ersten Bilddatensatzes für zumindest einen Teil des zweiten Bilddatensatz verwendet werden. Vorzugsweise können in jeder Iteration Bilder aus dem ersten Bilddatensatz entfernt oder ausgetauscht werden. Des Weiteren können in jeder Iteration Bilder aus dem zweiten Bilddatensatz entfernt oder ausgetauscht werden. Findet ein

Austausch von Bilddaten beim ersten und/oder zweiten Bilddatensatz statt, können vorzugsweise die Annotationen so abgeändert werden, dass sie Informationen über die neuen Bilddaten enthalten.

5 Im ersten Aspekt der Erfindung ist weiterhin vorstellbar, dass das Nachtrainieren des Bildklassifikators basierend auf Relevanzwerten von Bildbereichen des zweiten Bilddatensatzes durchgeführt wird.

10 Der Vorteil dieser Herangehensweise ist, dass aus Bewertungssicht weniger oder nicht-relevante Bildbereiche im Training derart gewichtet werden können, dass sie einen geringen oder keinen Einfluss auf das Training des Bildklassifikators haben. Dies führt zu einem einfacheren Training des Bildklassifikators, was im Umkehrschluss die Erkennungsleistung des Bildklassifikators erheblich steigert. Dies führt zu einer Steigerung der Performanz des Gesamtsystems, wenn
15 ein so trainierter Bildklassifikator als Teil der Steuerung eines Roboters verwendet wird.

Nachfolgend werden Ausführungsformen der Erfindung unter Bezugnahme auf die beiliegenden Zeichnungen näher erläutert. In den Zeichnungen zeigen:

20

Figur 1 schematisch das Ablaufdiagramm des Verfahrens einer Bewertung eines Bildklassifikators;

Figur 2 schematisch ein Steuersystem;

25

Figur 3 schematisch ein autonomes Fahrzeug, das vom Steuersystem gesteuert wird;

Figur 4 schematisch einen Fertigungsroboter, der vom Steuersystem gesteuert wird.
30

Beschreibung der Ausführungsbeispiele

Figur 1 zeigt ein Ablaufdiagramm eines Verfahrens zum Bewerten eines Bildklassifikators (60). In diesem Ausführungsbeispiel ist der Bildklassifikator (60) derart

ausgeprägt, dass er Fahrzeuge in Autobahnsituationen erkennen kann, wobei die Klassifikatorausgabe (y) Bounding Boxes beinhaltet.

5 In einem ersten Schritt (300) wird ein Bilddatensatz ermittelt. Dies kann zum Beispiel unter Zuhilfenahme eines Testfahrzeugs geschehen, in dem eine Kamera installiert ist, die zur Aufnahme von Bilddaten geeignet ist. In diesem Ausführungsbeispiel zeigt der Bilddatensatz Bilddaten von Autobahnsituationen, auf denen Fahrzeuge zu erkennen sind. Weiter werden in diesem Ausführungsbeispiel während der Aufnahme des Bilddatensatzes den Bilddaten jeweils Fahrzeugdaten wie Geschwindigkeit und Lenkwinkel zugeordnet, die zum jeweiligen Aufnahmezeitpunkt eines Bildes vorliegen. Alternativ können diese Fahrzeugdaten auch nach der Aufnahme aus den aufgenommenen Bilddaten geschätzt werden.

15 Der so ermittelte Bilddatensatz wird anschließend manuell von einem Menschen mit Annotationen versehen. Alternativ kann anstelle der manuellen Annotation hier auch eine halbautomatische Annotation mit Hilfe eines zweiten Bildklassifikators durchgeführt werden. In diesem Fall kann der zweite Bildklassifikator Annotationen vorschlagen, die von einem Menschen überprüft und gegebenenfalls abgeändert werden können. Alternativ ist weiterhin denkbar, dass der zweite Bildklassifikator die Annotation vollautomatisiert durchführt, indem die Vorschläge des zweiten Bildklassifikators direkt als Annotationen verwendet werden.

25 Die erzeugten Annotationen enthalten in diesem Ausführungsbeispiel Informationen bezüglich der anderen Fahrzeuge in den jeweils aufgenommenen Bilddaten, Bounding Boxen der Fahrzeuge im Bild, sowie Einbauposition und Orientierung des Kamerasensors. In weiteren Ausführungsbeispielen ist vorstellbar, dass die Annotationen zusätzlich 3-dimensionale Informationen, wie Position, Orientierung, Geschwindigkeit und/oder Fahrtrichtung der entsprechend zu detektierenden Fahrzeuge enthalten.

30 In einem zweiten Schritt (301) wird dann für die Bilder des Bilddatensatzes jeweils bestimmt, welche Bereiche das Testfahrzeug in einer festgelegten Zeit zum Zeitpunkt der Aufnahme des Bildes hätte erreichen können. Als Zeit kann hier beispielsweise die Time-To-React verwendet werden. Alternativ ist vorstellbar, dass anstelle der Time-To-React die Time-To-Collision, Time-To-Brake, Time-

35

To-Steer oder Time-To-Kickdown oder Kombinationen dieser Zeiten verwendet werden können.

Die Berechnung der erreichbaren Bereiche geschieht unter Zuhilfenahme der Geschwindigkeitsinformationen in den Annotationen, sowie Informationen über die Position des Fahrzeugs. Das Ergebnis ist eine Information darüber, welche Bereiche das Fahrzeug zum Zeitpunkt der Aufnahme eines Bilddatums in der Szene, in der das Bilddatum aufgenommen wurde, in einer bestimmten Zeit hätte erreichen können.

In einem dritten Schritt (302) wird für die anderen Fahrzeuge der Bilddaten ein Relevanzwert bestimmt. Hierfür werden auf Basis der annotierten Bounding Boxes und der Einbauposition des Kamerasensors die 3-dimensionale Position der anderen Fahrzeuge ermittelt. Alternativ können diese Informationen auch direkt aus der Annotation extrahiert werden, falls sie vorhanden sind.

Der Relevanzwert kann für alle Fahrzeuge als 1 definiert werden, die sich in einem der im vorherigen Schritt bestimmten Bereiche befinden, und andernfalls als 0. Alternativ ist möglich, dass den Fahrzeugen ein Wert zwischen 0 und 1 zugewiesen wird, falls sie sich außerhalb eines im vorherigen Schritt bestimmten Bereiches befinden. Alternativ ist weiterhin vorstellbar, dass auch Fahrzeuge in einem der im vorherigen Schritt bestimmten Bereiche einen Wert zwischen 0 und 1 zugewiesen bekommen. Weiterhin ist denkbar, dass der Relevanzwert eines Objekts auch von der Geschwindigkeit und Trajektorie des Objekts abhängt. Beispielsweise können auch Objekte außerhalb der erreichbaren Bereiche einen Relevanzwert größer 0 erhalten, falls sie sich zum Beispiel auf die entsprechenden erreichbaren Bereiche zubewegen.

In einem vierten Schritt (303) werden die Bilddaten des ersten Bilddatensatzes durch den Bildklassifikator (60) klassifiziert. Bei der Klassifikation kann ein Fahrzeug in einem Bilddatum entweder detektiert, also korrekt klassifiziert, werden oder nicht.

In einem fünften Schritt (304) wird die Erkennungsgenauigkeit des Bildklassifikators (60) bewertet. Als Bewertungsmaß kann eine Summe der Relevanzwerte

verwendet werden, die zu Fahrzeugen gehören, die nicht detektiert wurden. Alternativ ist vorstellbar, dass auch der Mittelwert oder der Median der Relevanzwerte als Bewertungsmaß verwendet werden kann.

5 Falls das Bewertungsmaß schlechter ist als ein vorgegebener Bewertungsschwellenwert, kann in einem sechsten Schritt (306) der Bildklassifikator (60) unter Zuhilfenahme eines zweiten Bilddatensatzes nachtrainiert werden. In diesem Ausführungsbeispiel ist beispielsweise vorstellbar, dass der Bewertungsschwellenwert als 0 definiert wird. Dies ist gleichbedeutend mit der Aussage, dass alle
10 Fahrzeuge mit einer Relevanz größer 0 erkannt werden müssen. Für den Fall, dass dies nicht Eintritt wird der Bildklassifikator nachtrainiert. Der zweite Bilddatensatz kann in diesem Fall nach einem der Verfahren ermittelt werden, die auch zur Ermittlung des ersten Bilddatensatzes verwendet werden können. Falls das Bewertungsmaß dem Bewertungsschwellenwert genügt, kann der Bildklassifikator (60) freigegeben werden.
15

In weiteren Ausführungsbeispielen ist vorstellbar, dass die Schritte der Bewertung des Bildklassifikators (300, 301, 302, 303, 304, 305) so lange iterativ wiederholt werden, bis das Bewertungsmaß dem Schwellenwert genügt.
20

In weiteren Ausführungsbeispielen ist vorstellbar, dass die Schritte der Bewertung des Bildklassifikators (300, 301, 302, 303, 304, 305) so lange iterativ wiederholt werden, bis eine vorher definierte Anzahl von Iterationen absolviert wurde.
25

Figur 2 zeigt einen Aktor (10) in seiner Umgebung (20) in Interaktion mit einem Steuerungssystem (40). In vorzugsweise regelmäßigen zeitlichen Abständen wird die Umgebung (20) mit einem Sensor (30), insbesondere einem bildgebenden Sensor wie einem Videosensor, erfasst, der auch durch eine Mehrzahl von Sensoren gegeben sein kann, beispielsweise eine Stereokamera. Das Sensorsignal (S) – bzw. im Fall mehrerer Sensoren je ein Sensorsignal (S) – des Sensors (30) wird an das Steuerungssystem (40) übermittelt. Das Steuerungssystem (40) empfängt somit eine Folge von Sensorsignalen (S). Das Steuerungssystem (40) ermittelt hieraus Ansteuersignale (A), welche an den Aktor (10) übertragen werden.
30
35

Das Steuerungssystem (40) empfängt die Folge von Sensorsignalen (S) des Sensors (30) in einer optionalen Empfangseinheit (50), die die Folge von Sensorsignalen (S) in eine Folge von Eingangsbildern (x) umwandelt (alternativ kann auch unmittelbar je das Sensorsignal (S) als Eingangsbild (x) übernommen werden). Das Eingangsbild (x) kann beispielsweise ein Ausschnitt oder eine Weiterverarbeitung des Sensorsignals (S) sein. Das Eingangsbild (x) umfasst einzelne Frames einer Videoaufzeichnung. Mit anderen Worten wird Eingangsbild (x) abhängig von Sensorsignal (S) ermittelt. Die Folge von Eingangsbildern (x) wird einem Bildklassifikator (60) zugeführt, der zum Beispiel wie im ersten Ausführungsbeispiel bewertet wurde und dessen Bewertungsmaß unter dem Bewertungsschwellenwert lag.

Der Bildklassifikator (60) wird vorzugsweise parametrierbar durch Parameter (ϕ), die in einem Parameterspeicher (P) hinterlegt sind und von diesem bereitgestellt werden.

Der Bildklassifikator (60) ermittelt aus den Eingangsbildern (x) Klassifikatorausgaben (y). Die Klassifikatorausgaben (y) werden einer optionalen Umformeinheit (80) zugeführt, die hieraus Ansteuersignale (A) ermittelt, welche dem Aktor (10) zugeführt werden, um den Aktor (10) entsprechend anzusteuern. Die Klassifikatorausgabe (y) umfasst Informationen über Objekte, die der Sensor (30) erfasst hat.

Der Aktor (10) empfängt die Ansteuersignale (A), wird entsprechend angesteuert und führt eine entsprechende Aktion aus. Der Aktor (10) kann hierbei eine (nicht notwendigerweise baulich integrierte) Ansteuerlogik umfassen, welche aus dem Ansteuersignal (A) ein zweites Ansteuersignal ermittelt, mit dem dann der Aktor (10) angesteuert wird.

In weiteren Ausführungsformen umfasst das Steuerungssystem (40) den Sensor (30). In noch weiteren Ausführungsformen umfasst das Steuerungssystem (40) alternativ oder zusätzlich auch den Aktor (10).

In weiteren bevorzugten Ausführungsformen umfasst das Steuerungssystem (40) einen oder eine Mehrzahl von Prozessoren (45) und wenigstens ein maschinenlesbares Speichermedium (46), auf dem Anweisungen gespeichert sind, die dann, wenn sie auf den Prozessoren (45) ausgeführt werden, das Steuerungssystem (40) veranlassen, das erfindungsgemäße Verfahren auszuführen.

In alternativen Ausführungsformen ist alternativ oder zusätzlich zum Aktor (10) eine Anzeigeeinheit (10a) vorgesehen.

Figur 3 zeigt, wie das Steuerungssystem (40) zur Steuerung eines wenigstens teilweise autonomen Roboters, hier eines wenigstens teilautonomen Kraftfahrzeugs (100), eingesetzt werden kann.

Bei dem Sensor (30) kann es sich beispielsweise um einen vorzugsweise im Kraftfahrzeug (100) angeordneten Videosensor handeln.

Der Bildklassifikator (60) ist eingerichtet, aus den Eingangsbildern (x) Objekte zu identifizieren.

Bei dem vorzugsweise im Kraftfahrzeug (100) angeordneten Aktor (10) kann es sich beispielsweise um eine Bremse, einen Antrieb oder eine Lenkung des Kraftfahrzeugs (100) handeln. Das Ansteuersignal (A) kann dann derart ermittelt werden, dass der Aktor oder die Aktoren (10) derart angesteuert wird, dass das Kraftfahrzeug (100) beispielsweise eine Kollision mit den vom Bildklassifikator (60) identifizierten Objekte verhindert, insbesondere, wenn es sich um Objekte bestimmter Klassen, z.B. um Fußgänger, handelt.

Alternativ kann es sich bei dem wenigstens teilautonomen Roboter auch um einen anderen mobilen Roboter (nicht abgebildet) handeln, beispielsweise um einen solchen, der sich durch Fliegen, Schwimmen, Tauchen oder Schreiten fortbewegt. Bei dem mobilen Roboter kann es sich beispielsweise auch um einen wenigstens teilautonomen Rasenmäher oder einen wenigstens teilautonomen Putzroboter handeln. Auch in diesen Fällen kann das Ansteuersignal (A) derart ermittelt werden, dass Antrieb und/oder Lenkung des mobilen Roboters derart

angesteuert werden, dass der wenigstens teilautonome Roboter beispielsweise eine Kollision mit vom Bildklassifikator (60) identifizierten Objekten verhindert.

Alternativ oder zusätzlich kann mit dem Ansteuersignal (A) eine Anzeigeeinheit (10a) angesteuert werden, und beispielsweise die ermittelten sicheren Bereiche dargestellt werden. Auch ist es beispielsweise bei einem Kraftfahrzeug (100) mit nicht automatisierter Lenkung möglich, dass die Anzeigeeinheit (10a) mit dem Ansteuersignal (A) derart angesteuert wird, dass sie ein optisches oder akustisches Warnsignal ausgibt, wenn ermittelt wird, dass das Kraftfahrzeug (100) droht, mit einem der vom Bildklassifikator (60) identifizierten Objekte zu kollidieren.

Figur 4 zeigt, wie das Steuersystem (40) zur Steuerung eines Fertigungsroboters (220), wie zum Beispiel eines PUMA Roboters, verwendet werden kann, wobei der Arbeitsraum (212) des Fertigungsroboters (220) auch von Personen (210) betreten werden kann. Es ist vorstellbar, dass in diesem Ausführungsbeispiel das Steuersystem (40) Bilddaten von einem Kamerasensor (30) erhält, auf deren Basis es einen Aktuator (10) ansteuert, wobei der Aktuator (10) die Bewegung des Fertigungsroboters (220) sowie einen Greifer am Ende des Arms des Fertigungsroboters (220) antreibt, mit dem Werkstücke (211a, 211b) erfasst werden können.

Über die Bilddaten des Kamerasensors (30) kann das Steuersystem (40) mit Hilfe des beinhalteten Bildklassifikators (60) außerdem Personen (210) erkennen, die sich im Arbeitsraum (212) des Fertigungsroboters (220) befinden. Für den Fall, dass eine oder mehrere Personen (210) im Arbeitsraum (212) erkannt wurden, kann die Bewegung des Fertigungsroboters (220) derart durch das Steuersystem (40) angepasst werden, dass die Person oder die Personen (210) nicht vom Fertigungsroboter (220) berührt oder verletzt wird/werden. Optional ist auch vorstellbar, dass die Bewegung des Fertigungsroboters (220) derart gewählt wird, dass der Arm des Fertigungsroboters (220) einen gewissen Mindestabstand zu der Person oder den Personen (210) im Arbeitsraum (212) einhält.

Für dieses Ausführungsbeispiel ist es möglich, dass der Bildklassifikator (60) mit Bildern von Personen (210) in oder um den Arbeitsraum (212) des Fertigungsroboters (220) trainiert wurde. Zur Bewertung, ob der Fertigungsroboter (220) sicher betrieben werden kann, kann ein erster Bilddatensatz aufgenommen werden, wobei die Bilder des ersten Bilddatensatzes ebenfalls Personen (210) in oder um den Arbeitsraum des Fertigungsroboters (220) zeigen können. Die Bilder des ersten Bilddatensatzes können zur Bewertung mit Annotationen in Form von Bounding Boxen für die Personen (210) auf den entsprechenden Bildern versehen werden, wobei jeder Bounding Box außerdem ein Relevanzwert zugewiesen wird. Dieser Relevanzwert kann als 1 definiert werden, falls die entsprechende Bounding Box eine Person (210) zeigt, die sich im Arbeitsraum (212) des Fertigungsroboters (220) befindet, und andernfalls als 0 definiert werden.

Zur Bewertung kann anschließend festgelegt werden, dass die Summe der Relevanzwerte der vom Bildklassifikator (60) nicht erkannten Bounding Boxes des ersten Datensatzes 0 sein muss. Dies ist gleichbedeutend mit der Aussage, dass der Bildklassifikator (60) keine Person (210) innerhalb des Arbeitsraums (212) des Fertigungsroboters (220) fehldetektieren darf, während dies bei Personen außerhalb des Arbeitsraums nicht gefordert ist. Alternativ ist vorstellbar, dass Personen außerhalb des Arbeitsraums (212) höhere Relevanzwerte erhalten, je näher sie am Arbeitsraum (212) stehen. Es ist weiterhin vorstellbar, dass in diesem Fall die Summe der Relevanzwerte größer als 0 sein darf, um den Bildklassifikator (60) als ausreichend sicher für den Einsatz zu bewerten.

5 Ansprüche

1. Computerimplementiertes Verfahren zur Bewertung eines Bildklassifikators (60), wobei eine Klassifikatorausgabe (y) des Bildklassifikators (60) zur Ansteuerung eines zumindest teilweise autonomen Roboters (100, 220) bereitgestellt wird, wobei das Verfahren zur Bewertung folgende Schritte umfasst:
- Ermittlung (300) eines ersten Datensatzes, wobei der erste Datensatz Bilder enthält, wobei den Bildern Annotationen zugewiesen sind, wobei die Annotationen Informationen über die im jeweiligen Bild abgebildete Szene und/oder über zu klassifizierende Bildbereiche und/oder über Bewegungsinformationen des Roboters (100, 220) enthalten;
 - Ermittlung (301) von durch den Roboter (100, 220) erreichbaren Bereichen (212) der Szenen basierend auf den Annotationen;
 - Ermittlung (302) von Relevanzwerten für vom Bildklassifikator zu klassifizierenden Bildbereiche;
 - Klassifizierung (303) der Bilddaten des ersten Bilddatensatzes mittels des Bildklassifikators;
 - Bewertung (304) des Bildklassifikators basierend auf durch den Bildklassifikator (60) korrekt klassifizierten Bildbereichen und falsch klassifizierten Bildbereichen, sowie den berechneten Relevanzwerten der entsprechenden Bildbereiche.
2. Verfahren nach Anspruch 1, wobei die zu klassifizierenden Bildbereiche jeweils einem Objekt (210) zugeordnet sind.
3. Verfahren nach Anspruch 1 oder 2, wobei die Ermittlung (301) der durch den Roboter (100, 220) erreichbaren Bereiche auf Bewegungsinformationen des Roboters (100, 220) basiert.

4. Verfahren nach einem der Ansprüche 1 bis 3, wobei der Schritt der Ermittlung (302) der Relevanzwerte folgende Schritte umfasst:
 - Ermittlung von Tiefeninformationen der Objekte (210);
 - Bestimmung einer Relation von Objektpositionen zu erreichbaren Bereichen (212) basierend auf den Tiefeninformationen;
 - Ermittlung der Relevanzwerte basierend auf der Relation.
5. Verfahren nach einem der Ansprüche 1 bis 4, wobei der Schritt der Bewertung (304) des Bildklassifikators das Bestimmen eines Bewertungsmaßes umfasst, weiter wobei das Verfahren zur Bewertung des Bildklassifikators (60) folgenden zusätzlichen Schritt enthält:
 - Nachtrainieren (305) des Bildklassifikators (60) basierend auf einem zweiten Bilddatensatz, falls das Bewertungsmaß schlechter ist als ein vorgegebener Bewertungsschwellenwert.
6. Verfahren nach Anspruch 5, wobei die Schritte iterativ wiederholt werden, bis eine vorgegebene Anzahl von Iterationen durchgeführt wurde und/oder der vorgegebene Bewertungsschwellenwert unterschritten wurde.
7. Verfahren nach Anspruch 5 oder 6, wobei das Nachtrainieren des Bildklassifikators (60) basierend auf Relevanzwerten von Bildbereichen des zweiten Bilddatensatzes durchgeführt wird.
8. Computerimplementiertes Verfahren zum Betreiben eines Bildklassifikators (60), wobei der Bildklassifikator (60) nach einem der Ansprüche 1 bis 5 bewertet oder nach einem der Ansprüche 5 bis 7 nachtrainiert wurde.
9. Steuersystem (40) zur Ansteuerung eines zumindest teilweise autonomen Roboters, wobei das Ansteuersystem (40) einen Bildklassifikator (60) nach einem der Ansprüche 1 bis 7 beinhaltet, des Weiteren wobei das Steuersystem (40) Bilddaten (x) basierend auf Sensorsignalen (S) eines Sensors (30) an den Bildklassifikator (60) übergibt und/oder ein Ansteuersignal (A) basierend auf der Klassifikatorausgabe (y) ermittelt wird, wobei das Ansteuersignal (A) zum Ansteuern eines Aktuators (10) des Roboters verwendet wird.

- 22 -

10. Trainingsvorrichtung, welche eingerichtet ist, das Verfahren nach einem der Ansprüche 5 bis 7 auszuführen.

5

11. Computerprogramm, welches eingerichtet ist, das Verfahren nach einem der Ansprüche 1 bis 7 auszuführen.

12. Maschinenlesbares Speichermedium (46, 146), auf dem das Computerprogramm nach Anspruch 11 gespeichert ist.

10

1/4

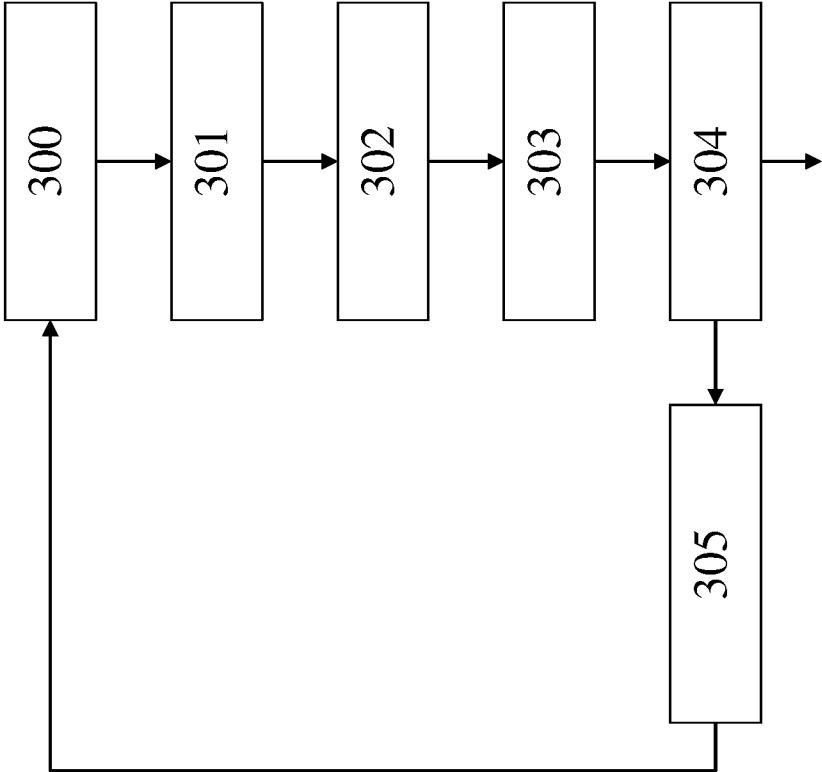


Fig. 1

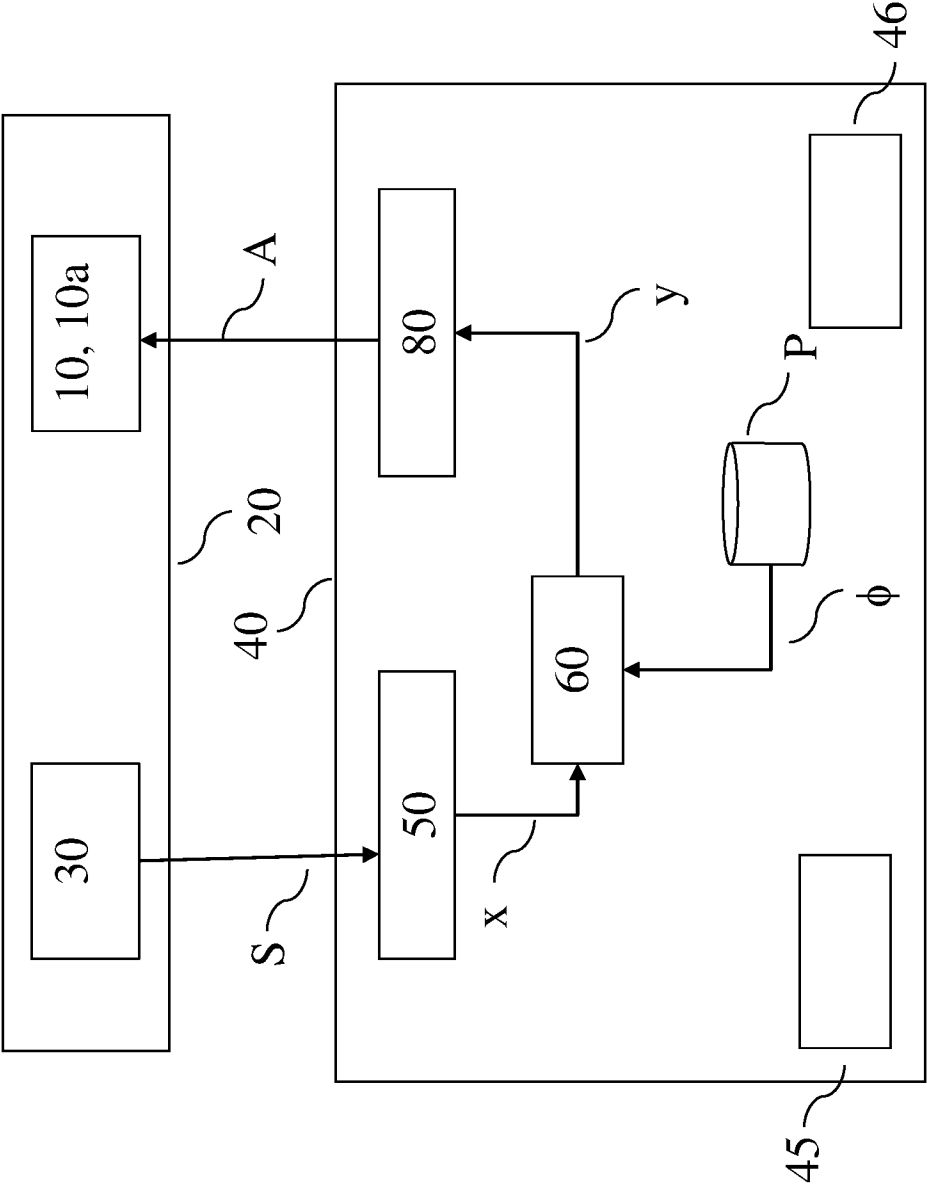


Fig. 2

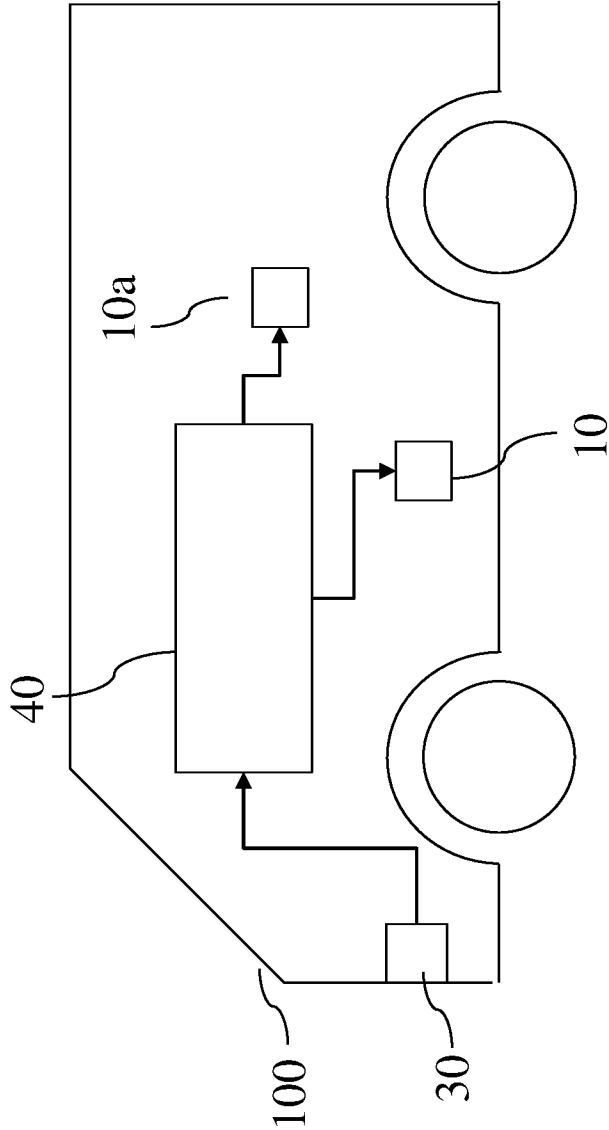


Fig. 3

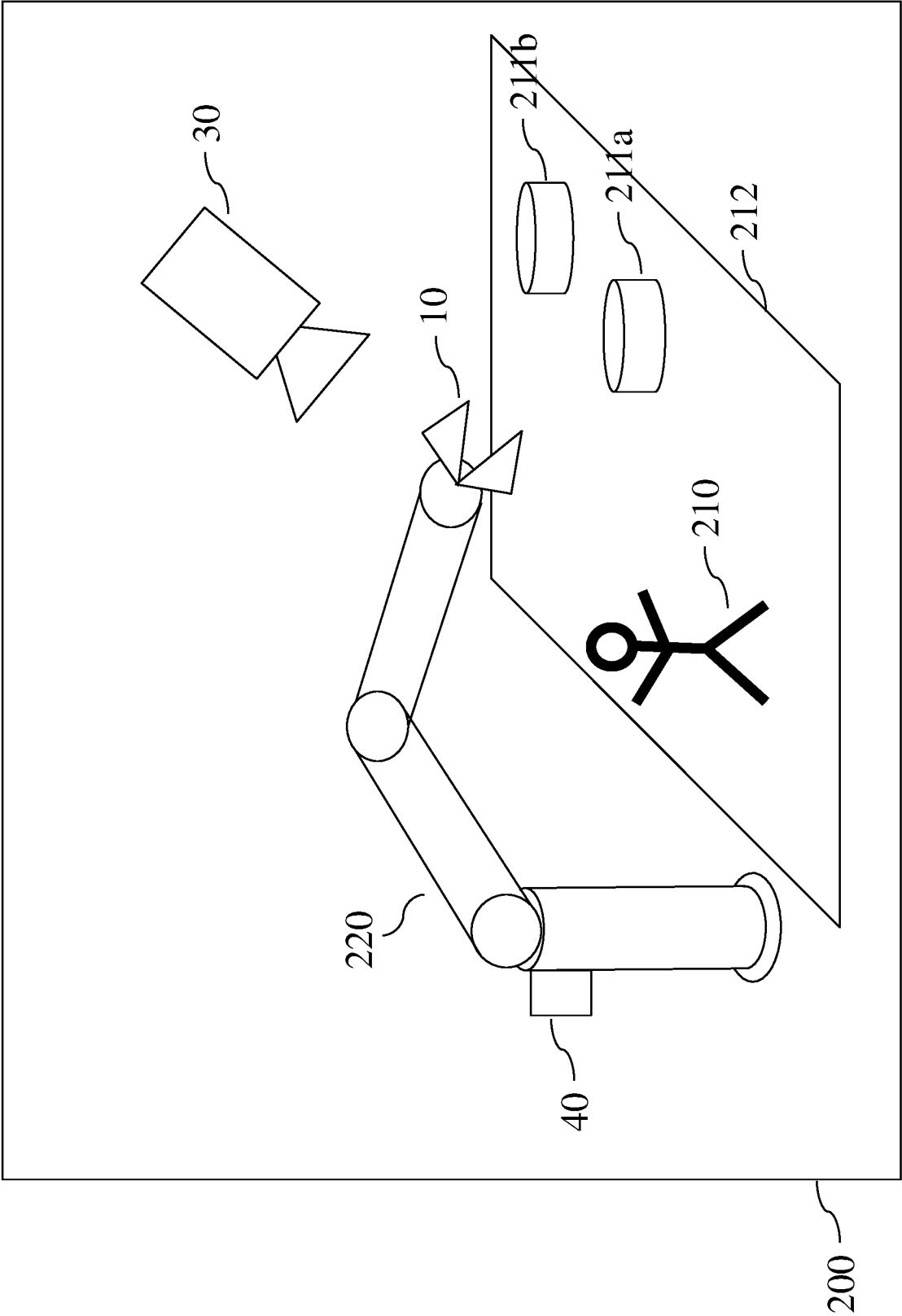


Fig. 4

INTERNATIONAL SEARCH REPORT

International application No.

PCT/EP2021/052931

A. CLASSIFICATION OF SUBJECT MATTER G06K 9/00(2006.01)i; G06K 9/62(2006.01)i According to International Patent Classification (IPC) or to both national classification and IPC		
B. FIELDS SEARCHED Minimum documentation searched (classification system followed by classification symbols) G06K Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched Electronic data base consulted during the international search (name of data base and, where practicable, search terms used) EPO-Internal, WPI Data		
C. DOCUMENTS CONSIDERED TO BE RELEVANT		
Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	OHN-BAR ESHED ET AL. "Are all objects equal? Deep spatio-temporal importance prediction in driving videos" <i>PATTERN RECOGNITION</i> , Vol. 64, 01 April 2017 (2017-04-01), pages 425-436 DOI: 10.1016/J.PATCOG.2016.08.029 ISSN: 0031-3203, XP029864359	1-3,5-12
A	abstract sections 1, 1.1, 4.1, 4.2, 5	4
A	WO 2019175012 A1 (BOSCH GMBH ROBERT [DE]) 19 September 2019 (2019-09-19) abstract page 10, lines 11-25 page 16, line 31 - page 17, line 32	1-12
A	JONATHON BYRD ET AL. "What is the Effect of Importance Weighting in Deep Learning?" <i>ARXIV.ORG, CORNELL UNIVERSITY LIBRARY, 201 OLIN LIBRARY CORNELL UNIVERSITY ITHACA, NY 14853</i> , 08 December 2018 (2018-12-08), XP081368966 abstract	1-12
☒ Further documents are listed in the continuation of Box C. ☒ See patent family annex.		
* Special categories of cited documents: "A" document defining the general state of the art which is not considered to be of particular relevance "E" earlier application or patent but published on or after the international filing date "L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified) "O" document referring to an oral disclosure, use, exhibition or other means "P" document published prior to the international filing date but later than the priority date claimed "T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention "X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone "Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art "&" document member of the same patent family		
Date of the actual completion of the international search 29 April 2021		Date of mailing of the international search report 06 May 2021
Name and mailing address of the ISA/EP European Patent Office p.b. 5818, Patentlaan 2, 2280 HV Rijswijk Netherlands Telephone No. (+31-70)340-2040 Facsimile No. (+31-70)340-3016		Authorized officer Philips, Petra Telephone No.

INTERNATIONAL SEARCH REPORT

International application No.

PCT/EP2021/052931**C. DOCUMENTS CONSIDERED TO BE RELEVANT**

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	M Althoff. "Reachability Analysis and its Application to the Safety Assessment of Autonomous Cars" 03 January 2018 (2018-01-03), Retrieved from the Internet: https://mediatum.ub.tum.de/doc/963752/642175.pdf XP055587085 cited in the application section 5.1	3

INTERNATIONAL SEARCH REPORT
Information on patent family members

International application No.

PCT/EP2021/052931

Patent document cited in search report			Publication date (day/month/year)	Patent family member(s)			Publication date (day/month/year)
WO	2019175012	A1	19 September 2019	CN	111868641	A	30 October 2020
				DE	102019202090	A1	19 September 2019
				EP	3765927	A1	20 January 2021
				US	2021078168	A1	18 March 2021
				WO	2019175012	A1	19 September 2019

A. KLASSIFIZIERUNG DES ANMELDUNGSGEGENSTANDES
 INV. G06K9/00 G06K9/62
 ADD.

Nach der Internationalen Patentklassifikation (IPC) oder nach der nationalen Klassifikation und der IPC

B. RECHERCHIERTE GEBIETE

Recherchierter Mindestprüfstoff (Klassifikationssystem und Klassifikationssymbole)
 G06K

Recherchierte, aber nicht zum Mindestprüfstoff gehörende Veröffentlichungen, soweit diese unter die recherchierten Gebiete fallen

Während der internationalen Recherche konsultierte elektronische Datenbank (Name der Datenbank und evtl. verwendete Suchbegriffe)

EPO-Internal, WPI Data

C. ALS WESENTLICH ANGESEHENE UNTERLAGEN

Kategorie*	Bezeichnung der Veröffentlichung, soweit erforderlich unter Angabe der in Betracht kommenden Teile	Betr. Anspruch Nr.
X	OHN-BAR ESHED ET AL: "Are all objects equal? Deep spatio-temporal importance prediction in driving videos", PATTERN RECOGNITION, Bd. 64, 1. April 2017 (2017-04-01), Seiten 425-436, XP029864359, ISSN: 0031-3203, DOI: 10.1016/J.PATCOG.2016.08.029	1-3,5-12
A	Zusammenfassung Abschnitte 1, 1.1, 4.1, 4.2, 5 -----	4
A	WO 2019/175012 A1 (BOSCH GMBH ROBERT [DE]) 19. September 2019 (2019-09-19) Zusammenfassung Seite 10, Zeilen 11-25 Seite 16, Zeile 31 - Seite 17, Zeile 32 ----- -/-	1-12



Weitere Veröffentlichungen sind der Fortsetzung von Feld C zu entnehmen



Siehe Anhang Patentfamilie

* Besondere Kategorien von angegebenen Veröffentlichungen :

"A" Veröffentlichung, die den allgemeinen Stand der Technik definiert, aber nicht als besonders bedeutsam anzusehen ist

"E" frühere Anmeldung oder Patent, die bzw. das jedoch erst am oder nach dem internationalen Anmeldedatum veröffentlicht worden ist

"L" Veröffentlichung, die geeignet ist, einen Prioritätsanspruch zweifelhaft erscheinen zu lassen, oder durch die das Veröffentlichungsdatum einer anderen im Recherchenbericht genannten Veröffentlichung belegt werden soll oder die aus einem anderen besonderen Grund angegeben ist (wie ausgeführt)

"O" Veröffentlichung, die sich auf eine mündliche Offenbarung, eine Benutzung, eine Ausstellung oder andere Maßnahmen bezieht

"P" Veröffentlichung, die vor dem internationalen Anmeldedatum, aber nach dem beanspruchten Prioritätsdatum veröffentlicht worden ist

"T" Spätere Veröffentlichung, die nach dem internationalen Anmeldedatum oder dem Prioritätsdatum veröffentlicht worden ist und mit der Anmeldung nicht kollidiert, sondern nur zum Verständnis des der Erfindung zugrundeliegenden Prinzips oder der ihr zugrundeliegenden Theorie angegeben ist

"X" Veröffentlichung von besonderer Bedeutung; die beanspruchte Erfindung kann allein aufgrund dieser Veröffentlichung nicht als neu oder auf erfinderischer Tätigkeit beruhend betrachtet werden

"Y" Veröffentlichung von besonderer Bedeutung; die beanspruchte Erfindung kann nicht als auf erfinderischer Tätigkeit beruhend betrachtet werden, wenn die Veröffentlichung mit einer oder mehreren Veröffentlichungen dieser Kategorie in Verbindung gebracht wird und diese Verbindung für einen Fachmann naheliegend ist

"&" Veröffentlichung, die Mitglied derselben Patentfamilie ist

Datum des Abschlusses der internationalen Recherche

29. April 2021

Absendedatum des internationalen Recherchenberichts

06/05/2021

Name und Postanschrift der Internationalen Recherchenbehörde
 Europäisches Patentamt, P.B. 5818 Patentlaan 2
 NL - 2280 HV Rijswijk
 Tel. (+31-70) 340-2040,
 Fax: (+31-70) 340-3016

Bevollmächtigter Bediensteter

Philips, Petra

C. (Fortsetzung) ALS WESENTLICH ANGESEHENE UNTERLAGEN		
Kategorie*	Bezeichnung der Veröffentlichung, soweit erforderlich unter Angabe der in Betracht kommenden Teile	Betr. Anspruch Nr.
A	JONATHON BYRD ET AL: "What is the Effect of Importance Weighting in Deep Learning?", ARXIV.ORG, CORNELL UNIVERSITY LIBRARY, 201 OLIN LIBRARY CORNELL UNIVERSITY ITHACA, NY 14853, 8. Dezember 2018 (2018-12-08), XP081368966, Zusammenfassung -----	1-12
A	M Althoff: "Reachability Analysis and its Application to the Safety Assessment of Autonomous Cars", 3. Januar 2018 (2018-01-03), XP055587085, Gefunden im Internet: URL:https://mediatum.ub.tum.de/doc/963752/642175.pdf in der Anmeldung erwähnt Abschnitt 5.1 -----	3

INTERNATIONALER RECHERCHENBERICHT

Angaben zu Veröffentlichungen, die zur selben Patentfamilie gehören

Internationales Aktenzeichen

PCT/EP2021/052931

Im Recherchenbericht angeführtes Patentdokument	Datum der Veröffentlichung	Mitglied(er) der Patentfamilie	Datum der Veröffentlichung
WO 2019175012 A1	19-09-2019	CN 111868641 A	30-10-2020
		DE 102019202090 A1	19-09-2019
		EP 3765927 A1	20-01-2021
		US 2021078168 A1	18-03-2021
		WO 2019175012 A1	19-09-2019
