

Machbarkeitsrecherche Optische Tieridentifikation

Schriftenreihe, Heft 3/2023



Machbarkeitsrecherche Optische Tieridentifikation

Dipl. Inf. (FH) Tobias Heinzmann

(Softwareentwicklung und IT-Beratung)

Dorothee Heyde, Dr. Steffen Pache, Franziska Deißing, Dr. Maximilian Deutsch

(Landesamt für Umwelt, Landwirtschaft und Geologie)

Vorwort zur Einordnung in die Arbeit des LfULG

Die aktuellen Ansprüche hinsichtlich Klima- und Umweltschutz, Tierwohl und Gesundheitsvorsorge, Transparenz, Rückverfolgbarkeit, Verbraucherschutz und Qualitätssicherung zwingen den Landwirt zur Steigerung der Ressourceneffizienz. Herausgestellt zu betrachten sind vor allem die Arbeitseffizienz und der Mitteleinsatz. Produktionsverfahren ändern sich und einzeltierbezogene valide Prognose- und Steuer-elemente bilden die Grundlage für die ganzheitliche digitale Verfahrensgestaltung bereits mecha-nisierter und autonomer Produktionsabschnitte.

Um die rasante Entwicklung der Digitalisierung in unserer Gesellschaft zu begleiten und die Landwirtschaft u. a. bei der digitalen Transformation zu unterstützen, schaffte der Freistaat Sachsen mit der Initiative simul+InnovationHub (SIH) die Möglichkeit, unterschiedliche Anwendungen in der Praxis in regionalen Ex-perimentierfeldern zu testen. Die Inhalte dieses Beitrages sind ursprünglich als Teil des Themenfeldes 1 „Experimentierfeld 5G in Land- und Forstwirtschaft“ aus dem SIMUL+SIH-Vorhaben geplant worden. Im laufenden Projekt wird das Vorgehen am Bedarf der landwirtschaftlichen Praxisbetriebe in Sachsen aus-gerichtet. Durch die aktive Recherche der Fachreferate zu aktuellen Forschungen im Bereich Digi-talisierung, eigener Forschungsinitiativen, Rückkopplung der Bedarfe aus der Praxis durch Gespräche und Veranstaltungen im Jahresverlauf mit Anwendern, Gesprächen mit Anbietern und Recherche vorhandener Technologien am Markt konnten einzelne Bereiche wichtiger Digitalisierungsinitiativen herausgearbeitet werden und wurden als Arbeitspakete für die verschiedenen Facetten der Landwirtschaft fixiert. Im ersten von vier Arbeitspaketen im Bereich der Digitalisierung in der Tierhaltung werden optische Sensoren zur Tieridentifikation und als Tierwohllindikator erprobt und bewertet. Mit dem Ziel eines Lösungsansatz für die Bildanalyse zum Zweck der optische Tieridentifikation wurde eine Machbarkeits-recherche initiiert und in Auftrag gegeben. Das Vorgehen und die Ergebnisse sind im folgenden beschrieben.

Inhaltsverzeichnis

Vorwort zur Einordnung in die Arbeit des LfULG	3
1 Einleitung	10
2 Patentrecherche.....	11
2.1 Ablauf der Patentrecherche	11
2.1.1 Automatisierte Suche mittels Open Patent Service	11
2.1.1.1 Automatisierte Breitensuche	11
2.1.1.2 Kategorien	12
2.1.1.3 Ergebnisse der Breitensuche	12
2.1.2 Ausschluss nicht relevanter Dokumente	12
2.1.3 Suche im deutschen Sprachraum	12
2.2 Relevante Patentedokumente.....	13
2.2.1 Vollständige Lösung mittels Gesichtserkennung und Referenzdatenabgleich.....	13
2.2.1.1 Verteilung nach Ländern.....	15
2.2.2 Ansätze um Daten und/ oder Modell aktuell zu halten	15
2.2.2.1 Verteilung nach Ländern.....	16
2.2.3 Teillösung Gesichtserkennung, ohne Referenzdatenabgleich.....	16
2.2.3.1 Verteilung nach Ländern.....	17
2.2.4 Alternative Ansätze.....	18
2.2.4.1 Verteilung nach Ländern.....	19
2.2.5 Verteilung nach Ländern alle Ansätze gesamt	20
2.3 Zusammenfassung Patentrecherche	20
2.3.1 Patentrecht	20
2.3.2 Motivation Patente anzumelden.....	20
2.3.3 China	21
2.3.4 Alter der Patente.....	21
2.3.5 Qualität und Quantität der Daten	21
2.3.6 Auswertung Patentrecherche.....	22
2.3.6.1 Gesichtserkennung mit CNN.....	22
2.3.6.2 Anforderung an Kameras.....	22
2.3.6.3 Auflösung der Aufnahmen	22
2.3.6.4 Standardisierbarkeit.....	23
2.3.6.5 Leseempfehlung	23
3 Marktrecherche.....	24
3.1 Suchvorgang	24
3.2 Motivationen, Möglichkeiten und Verortung	24
3.2.1 Tierortung	24
3.2.2 Systemanbieter.....	25
3.2.3 Biosecurity und Healthcare	25
3.2.4 Schutz vor Versicherungsbetrug	25
3.3 Aktive Projekte.....	26
3.3.1 Cattleeye	26

3.3.2	Cainthus	26
3.3.3	Cattletracks.....	27
3.3.4	University of Kentucky	27
3.3.5	Stellapps.....	28
3.4	Forschungseinrichtungen Deutschland	28
3.4.1	Universität Kassel	28
3.4.2	Leibniz-Institut für Nutztierbiologie (FBN).....	29
3.4.3	Cattlehub	29
3.5	Forschungseinrichtungen UK.....	30
3.5.1	Agricultural Engineering Precision Innovation Centre (Agri-EPI).....	30
3.5.2	Rothamsted Research	30
3.6	Inaktive Projekte	30
3.6.1	Aktivstall	30
3.6.2	Agsenze	31
3.6.3	Milkalyser.....	31
3.7	Ungeklärte Projekte	31
3.7.1	Techvantage.....	31
3.7.2	Mooofarm	32
3.8	Bewertung der Marktsituation	32
3.8.1	Beurteilungsfähigkeit der Modelle	32
3.8.2	Komplexität und Vergleich mit Ohrmarkensystem.....	32
3.8.3	Prognose Marktreife	33
3.8.4	Zusammenfassung der Marktsituation	33
3.8.5	Expertengruppen	34
4	Technologie	35
4.1	Künstliche Intelligenz	35
4.1.1	Maschinelles Lernen	35
4.1.2	Leistungsfähigkeit von Modellen	36
4.1.3	Tiefes maschinelles Lernen	37
4.2	Künstliche Neuronale Netze.....	37
4.2.1	Neuronen.....	39
4.2.2	Lernen	40
4.3	Convolutional Neural Networks.....	40
4.3.1	Convolutional Layer	41
4.3.2	Filter	44
4.3.3	Pooling Layer.....	46
4.3.4	Feature Maps und Feature-Vektoren	47
4.3.5	Modelle trainieren	48
4.3.6	ImageNet Large Scale Visual Recognition Challenge	49
4.3.7	Transfer learning.....	49
4.4	Daten.....	49
4.5	Erkennung	51
4.5.1	Object recognition.....	51
4.5.2	Object detection.....	51
4.5.3	Face detection	51

4.5.4	Face recognition	51
4.5.5	Facial Keypoint Detection	51
4.5.6	Holistischer Ansatz	52
4.5.7	Alternative Ansätze.....	52
4.5.8	2D und 3D	52
4.6	Systemaufbau.....	52
4.6.1	Hardwareanforderungen	52
4.6.2	Image preprocessing	53
4.6.3	Modell.....	54
4.6.4	Folgeprozesse	54
5	Handlungsempfehlung	55
5.1	Vorhandene Möglichkeiten zur optischen Tieridentifikation am Markt.....	55
5.2	Darstellung geeigneter Lösungsansätze	55
5.2.1	Reduzierung der zu identifizierenden Tiermenge	55
5.2.2	Festlegung der Genauigkeit.....	55
5.3	Formulierung der Schritte bis zur Anwendungsreife	56
5.4	Aufzeigen von Forschungsansätzen noch bestehender Lücken um Anwendungsreife zu erreichen.....	56
5.4.1	Kommunikation und Zusammenarbeit.....	56
5.4.2	Daten.....	57
5.4.3	Initiieren eigener Forschungsprojekte	58
6	Zusammenfassung	59
7	Schlussfolgerung	62
8	Quellen	63
8.1	Web-Seiten.....	63
8.2	Papers	64
8.3	Bücher	64

Abbildungsverzeichnis

Abbildung 1: Länderverteilung Patentfamilien vollständige Lösung.....	15
Abbildung 2: Länderverteilung Patentfamilien Daten aktualisieren.....	16
Abbildung 3: Länderverteilung Patentfamilien Teillösung Gesichtserkennung	17
Abbildung 4: Länderverteilung Patentfamilien alternative Ansätze.....	19
Abbildung 5: Länderverteilung Patentfamilien gesamt	20
Abbildung 6: kleines neuronales Beispielnetz, bestehend aus acht Neuronen.....	38
Abbildung 7: Beispielnetz mit einer Eingabeschicht, drei versteckten Schichten und einer Ausgabeschicht Quelle der Darstellung: [NI15].....	38
Abbildung 8: Aufbau eines Neurons	39
Abbildung 9: Beispiel Wahrnehmungsfelder	41
Abbildung 10: Verbindungen des ersten Schrittes: ein 25-Neuronen großer Kernel wird auf ein Neuron der nächsten Schicht abgebildet.....	43
Abbildung 11: Verbindungen des zweiten Schrittes: ein 25-Neuronen großer Kernel wird auf das zweite Neuron der nächsten Schicht abgebildet. Die zweite Schicht entsteht so durch die vollständige Iteration über alle Kernel.....	43
Abbildung 12: Beispielmuster Horizontalfilter.....	44
Abbildung 13: Beispielmuster Vertikalfilter.....	44
Abbildung 14: Nach Anwendung eines Horizontalfilters.....	45
Abbildung 15: Originalbild.....	45
Abbildung 16: Nach Anwendung eines Vertikalfilters.....	45
Abbildung 17: Max-Pooling mit einem Pooling-Kernel der Größe $2*2=4$ und einer Schrittweite die der Kantenlänge des Kernels (also 2) entspricht. Anhand der Färbung kann erkannt werden, welche Areale von Schicht 1 der Kernel bei einem Schritt sieht. Er ermittelt den größten Wert dieses Areals und gibt ihn an das Neuron der nächsten Schicht weiter. Das Resultat dieses Prozesses ist dann Schicht 2.	46
Abbildung 18: Auswirkung der Methoden Max-Pooling und Min-Pooling. Neben Max-Pooling können auch andere Pooling-Verfahren Anwendung finden. Max-Pooling ist aber das am häufigsten verwendete.	47
Abbildung 19: Vollständiges CNN. Subsampling ist ein Synonym für Pooling.....	48
Abbildung 20: Beispiel bounding box.....	50
Abbildung 21: Beispiel Object segmentation masks	50

Tabellenverzeichnis

Tabelle 1: Rechenbeispiel	42
---------------------------------	----

Abkürzungsverzeichnis

AB	Arbeitsbreite
IT	Informationstechnik
LfULG	Sächsisches Landesamt für Umwelt, Landwirtschaft und Geologie
UK	United Kingdom
USD	United States dollar

1 Einleitung

Das Interesse der Gesellschaft am Herstellungsprozess der Lebensmittel befindet sich in einem stetig wachsenden Prozess. Das Bedürfnis nach Transparenz und lückenloser Rückverfolgbarkeit ist bei den Verbrauchern größer als je zuvor. Diesen Bedarf greift auch die neue »Ampel-Koalition« im geplanten »Umbau der Landwirtschaft« auf. Eine lückenlose Datenerfassung im Stall ist demnach nicht mehr nur hinsichtlich der Produktions- und Wirtschaftlichkeitsaspekte notwendig. Sie hat auch stetig zunehmende Relevanz für die gesellschaftliche Akzeptanz.

Mit dieser Entwicklung verbunden ist das seit Jahren anhaltende Wachstum des Dokumentationsdrucks. Die Digitalisierung hat sich in diesem Bereich bewährt und stellt einen Lösungsansatz für die Tierhalter da. Grundvoraussetzung für die digitale Datenerhebung ist eine elektronische Kennzeichnung der Tiere. In der deutschen Rinderhaltung ist die Kennzeichnung und Registrierung nach der Viehverkehrsverordnung (ViehVerkV vom 06.07.2007, zuletzt geändert 03.05.2016) durch das Bundesministerium für Ernährung und Landwirtschaft festgelegt. Die verschiedenen Optionen für eine elektronische Kennzeichnung beim Rind z.B. in Form einer Ohrmarke oder als Transponder am Halsband inkludieren jedoch auch Nachteile. Die Anbringung stellt für beide Seiten, Mensch und Tier, eine Belastung dar. Beim Verlust der Hardware ist das Nachrüsten mit Stress verbunden und die Zeit bis zur Wiederherstellung der automatisierten Datenerhebung führt zu einer Lücke in der Dokumentation. Supplementär weisen die Technologien im Low Frequenzbereich Schwachstellen in der Zuverlässigkeit der Readertechnologien auf, die durch viele Faktoren im Stall und während des Produktionsablaufes beeinflusst werden kann.

In den letzten Jahrzehnten konnte die Biometrie als Wissenschaft zur Analyse der Ausprägung sichtbarer Merkmale in vielen Bereichen unserer Gesellschaft zur Identifizierung von Lebewesen genutzt werden. Durch den technischen Fortschritt im Bereich der Rechenkapazität und des maschinellen Lernens gewinnt dieses Verfahren immer mehr an Bedeutung im Alltag. Die für die biometrische Identifizierung benötigten Merkmale besitzen auch Rinder. Die Musterung des Fells, die Aderung an bestimmten Körperteilen, Körperproportionen, Haarwirbel und Fellfarbe sind Beispiel für Merkmale, die in ihrer Ausprägung die benötigte Varianz aufweisen könnten. Bildgebende Verfahren im zwei- oder dreidimensionalen Bereich kombiniert mit einer nachgelagerten Datenanalyse könnten eine Alternative zur Identifizierung von Einzeltieren im Rinderbereich darstellen und eine elektronische Tiererkennung mit ihren Nachteilen ablösen.

Für das strukturierte Vorgehen zur Bearbeitung dieses vielschichtigen Themas wurde eine umfassende Recherchearbeit beauftragt. Dabei wurde sowohl der Stand der Technik als auch der aktuelle Umsetzungsfortschritt im Themenbereich abgebildet. Zur Zielmotivation gehört die Herausstellung prinzipieller Möglichkeiten einer tierindividuellen Erkennung, ohne eine separate Tierkennzeichnung vornehmen zu müssen. Dabei soll der aktuelle Stand der technischen Entwicklung abgebildet werden.

2 Patentrecherche

Die Patentrecherche ist eine gute Methode um mit einer Technologierecherche zu beginnen. Es lässt sich sehr gut die Evolution von technischen Lösungen nachvollziehen und erkennen, wo Kompetenzen im recherchierten Themenbereich verortet sind. Beachtet werden muss allerdings, dass die international unterschiedlichen Patentrechtsgebungen dazu führen, dass Patente in unterschiedlichen Ländern und Themenbereichen auch unterschiedlich intensiv eingesetzt werden.

2.1 Ablauf der Patentrecherche

Die durchgeführte Patentrecherche kann grob in zwei Arbeitsschritte unterteilt werden. Um möglichst alle Patente zu finden die einen Beitrag zur optischen Tieridentifikation liefern können, wurde mit einer automatisierten Breitensuche begonnen. Diese lieferte sehr viele Dokumente als Ergebnis und natürlich befanden sich viele Patente darunter, die keine Relevanz für das vorliegende Thema haben. Um aus den Ergebnissen der automatisierten Suche die relevanten herauszufinden, wurde anschließend eine teil-manuelle Sichtung durchgeführt.

2.1.1 Automatisierte Suche mittels Open Patent Service

Es wurde der OPS (Open Patent Service) des Europäischen Patentamts für die automatisierte Patentsuche verwendet. Dies ist eine Schnittstelle zum automatisierten Zugriff auf die Patentdatenbank des Europäischen Patentamts. Eine Beschreibung dieser Schnittstelle ist zu finden unter: <https://www.epo.org/searching-for-patents/data/web-services/ops.html>

Mit OPS werden neben der Datenbank des Europäischen Patentamts auch alle an das Europäische Patentamt angeschlossenen Patentdatenbanken durchsucht. Dies sind Datenbanken aus über 80 verschiedenen Ländern: <https://www.epo.org/searching-for-patents/data/coverage.html>

Es kann somit davon ausgegangen werden, dass eine Patentsuche mittels OPS alle weltweit relevanten Erfindungen abdeckt.

2.1.1.1 Automatisierte Breitensuche

Die Breitensuche diente dem Finden aller möglicherweise relevanten Patentedokumente. Ziel war es, möglichst viele Treffer zu erhalten um keine Patente mit etwaig für das Thema der optischen Tieridentifikation interessanten Inhalt unberücksichtigt zu lassen.

Patentedokumente bestehen aus mehreren Feldern. Unter anderem enthalten sind Textfelder, die zusammen das im klassischen Sinne verstandene Patentedokument bilden:

- Titel
- Abstract
- Patentanspruch
- Beschreibung

Mit der automatisierten Suche via OPS können Titel und Abstract in englischer Sprache aller Dokumente durchsucht werden.

Für die Suche im aktuellen Projekt wurden Wortkombinationen bestehend aus zwei Worten gebildet. Das erste Wort ist eines der Synonyme für eine Kuh und das Zweite bezog sich auf die Identifikation, bzw. für die Identifikation relevanten Eigenschaften.

Suchworte erster Stelle sind:

animal, bovine, cattle, cow, livestock

Suchworte zweiter Stelle sind:

artificial, biometric, camera, deep, detection, identification, identity, image, individual, intelligence, learning, optical, pattern, recognition, training, visual

Sodass beispielsweise die erste Suchanfrage »animal artificial«, die zweite »animal biometric«, usf., bis zur letzten »livestock visual« gebildet wurden. Die Suchworte waren UND-verknüpft und bezogen sich auf Titel und Abstract aller Patente in den relevanten Kategorien. Alle diese Kombinationen wurden abgefragt und die Ergebnisse zur späteren Verarbeitung gespeichert.

2.1.1.2 Kategorien

Patente sind Kategorien zugeordnet, ein Patent kann mehreren Kategorien zugewiesen sein. Eine Kategorie definiert den Bereich auf den sich die Erfindung bezieht, oder die Technologie mit der die Erfindung umgesetzt wurde. Für die vorliegende Suche wurden Patente dieser Kategorien untersucht:

- A01K11/006: »Automatic identification systems for animals, ...«
- A01K29/00: »Monitoring or measuring activity, e.g. detecting heat or mating«
- A01K2227/101: »Bovine«
- G06K9/00885: »Biometric patterns ...«
- G07C9/37: »Using biometric data, e.g. fingerprints, iris scans or voice recognition«
- G06N3/02: »Computer systems based on biological models, using neural network models«

Näheres zu Kategorien unter: https://worldwide.espacenet.com/classification?locale=en_EP

2.1.1.3 Ergebnisse der Breitensuche

Nach Bereinigung der Duplikate lieferte die Breitensuche über 1.700 Dokumente als Ergebnis.

2.1.2 Ausschluss nicht relevanter Dokumente

Zur weiteren Eingrenzung wurden nur Dokumente ab dem Jahr 2015 verwertet. Es wird vorausgesetzt, dass eine mögliche Lösung realistisch nur unter Verwendung von bildverarbeitenden, künstliche Lernmethoden möglich ist. Diese existierten zwar schon vor 2015, aber wirkliche Fortschritte wurden hier erstens hauptsächlich nur in der Humanidentifikation generiert und zudem kann davon ausgegangen werden, dass wenn es eine ältere, fundierte Methode zur optischen Tieridentifikation geben würde, sie am Markt mittlerweile einschlägig bekannt geworden und dem sächsischen Landesamt für Umwelt, Landwirtschaft und Geologie nicht entgangen wäre.

Somit wurden für die Jahre 2015 bis 2020 viele Abstracts manuell gesichtet und alle unpassenden Dokumente anhand von Negativ-Schlüsselwörtern, oder falls dies aufgrund der allgemeinen Formulierung des Abstracts nicht möglich war, anhand der Patentfamilie, ausgeschlossen. Es wurden somit über 500 Patentdokumente teilweise oder in Gänze gesichtet.

2.1.3 Suche im deutschen Sprachraum

Da die Suche mit der OPS-Schnittstelle nur für Dokumente in englischer Sprache funktioniert, musste eine Suche ebenfalls im deutschen Sprachraum durchgeführt werden. Es ist davon auszugehen, dass Patente mittlerweile nicht mehr nur in Deutschland sondern immer auch parallel mindestens im europäischen Raum angemeldet werden. Die manuelle Suche lieferte keine Treffer, die nicht auch schon in der internationalen, automatisierten Suche enthalten waren.

2.2 Relevante Patentedokumente

Es werden Patente folgend unterschieden in

- **Vollständige Lösung:** Patente die Systeme beschreiben, mit denen die einzelnen Tiere erkannt und identifiziert werden können. Zur Identifikation wird das zu identifizierende Tier mit Referenzdaten abgeglichen.
- **Ansätze um Daten und Model aktuell zu halten:** diese Patente bearbeiten die Problematik, dass es nicht genügt, ein Tier zu identifizieren – es muss auch sichergestellt sein, dass das System stabil gegen Veränderung ist und dass die Daten von einem Tier auch aktuell gehalten werden.
- Eines der beiden zentralen Probleme welches ein System der optischen Tieridentifikation bewältigen muss, ist die Erkennung: es muss erkannt werden, ob und wo in einem Bild sich ein passendes, zu identifizierendes Tier (-Gesicht) befindet.

Die zweite große Schwierigkeit neben dem Erkennen ist der Abgleich mit der Referenzdatenbank. Da sich alle Bilder nur bestenfalls ähnlich sind, kann es keine 100%ige Zuordnung von erkanntem Tier auf dem Bild und denen in der Referenzdatenbank befindlichen Merkmalen geben. Patente denen der zweite **Schritt der Identifikation fehlt, die aber eine Erkennung leisten**, sind in der dritten Gruppe enthalten.

- **Alternative Ansätze:** Patente sind so definiert, dass sie neue Erfindungen beschreiben müssen. Es kann natürlich nicht ausgeschlossen werden, dass auch andere Technologien einen Beitrag zur optischen Tieridentifikation liefern können. Diese Gruppe an Patenten soll jene Ansätze enthalten.

2.2.1 Vollständige Lösung mittels Gesichtserkennung und Referenzdatenabgleich

- Patentfamilie 71904404

»Animal biological feature recognition method and device, computer equipment and storage medium«

Gesichtserkennung mittels CNN, Transfer Learning von Netzen für Humanerkennung und ImageNet, unterteilt in obere und untere Gesichtshälfte.

- 2020 08 11: [CN111523479](#), Agricultural Information Inst Caas

- Patentfamilie 71208820

»Cow face detection and recognition method based on deep learning«

Gesichtserkennung frontal, links und rechts. Verwendung Residual Network 101 (ResNet101) und Mask Region-CNN (Mask R-CNN).

- 2020 07 03: [CN111368766](#), Yunnan Anhua Disaster Prevention and Mitigation Tech Co Ltd Li Taoyong

- Patentfamilie 71029994

»Dairy cow individual identification system based on deep learning and identification method thereof«

CNN, Image-Größe 256*256, Feature-Vektor wird mit nicht näher beschriebener Feature-Vektor-Bibliothek abgeglichen.

- 2020 06 16: [CN111291683](#), Univ Inner Mongolia

- Patentfamilie 68281908

»Animal Identification Using Vision Techniques«

Allgemeine Beschreibung der Identifikation von Tieren anhand biometrischer Merkmale. Beispiele sind Gesichtserkennung und künstliche Intelligenz.

◦ 2020 04 16: [wo2020076225](#), Delaval Holding Ab

■ Patentfamilie 68890247

»Neural network-based cattle face recognition method, device, equipment and storage medium«

NN mit max, sum, oder average pooling (also CNN), automatisches Hinzufügen neuer Bilder in Datenbank.

◦ 2019 12 24: [CN110610125](#), Ping an Tech Shenzhen Co Ltd

■ Patentfamilie 66072883

»Intelligent feeding system and method for livestock«

Zwar für Schweine, vermutlich aber auch ebenso auf Kühe anwendbar. Enthält u.a. Gesichtserkennung mittels CNN. Yolo, MTCNN, MobileNet, OpenCV.

◦ 2019 04 16: [CN109618961](#), Beijing Jd Finance Tech Holding Co Ltd

◦ 2020 06 18: [WO2020119184](#), Jingdong Digits Tech Holding Co Ltd

■ Patentfamilie 64418078

»Animal identity identifying method and device, computer device and storage medium«

Gesichtserkennung für Tiere und im speziellen Kühe mittels FaceNet.

◦ 2018 11 30: [CN108921026](#), Ping an Tech Shenzhen Co Ltd

◦ 2019 12 05: [WO2019227616](#), Ping an Tech Shenzhen Co Ltd

■ Patentfamilie 62926752

»Livestock identification method and apparatus«

Bild des zu identifizierenden Tieres wird über ein 3D-Model gelegt, welches dann mittels eines CNN klassifiziert wird.

◦ 2018 07 27: [CN108335366](#), Deng Changshun

■ Patentfamilie 60105588

»Cow face identification method based on convolutional neural network and classifier model«

Sehr detailliert beschriebenes Identifikationsverfahren mittels CNN und einem angeschlossenen Classifier. Jede Schicht wurde beschrieben.

◦ 2017 10 24: [CN107292298](#), Univ Beifang Nationalities

2.2.1.1 Verteilung nach Ländern

Länderverteilung Patentfamilien vollständige Lösung

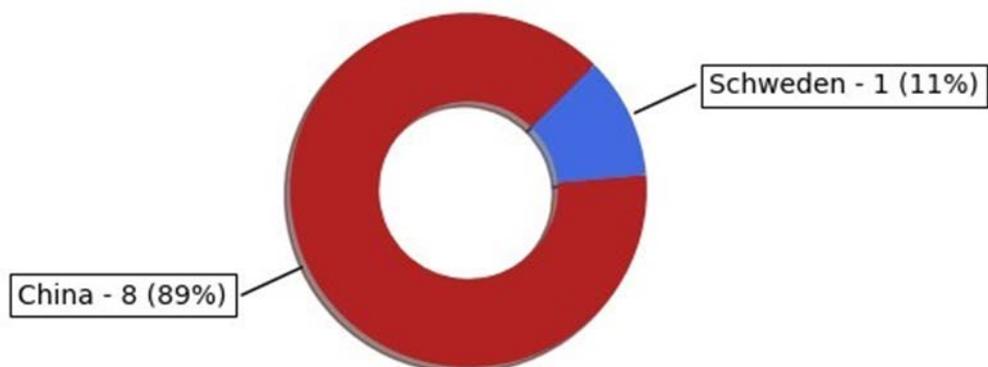


Abbildung 1: Länderverteilung Patentfamilien vollständige Lösung

2.2.2 Ansätze um Daten und/ oder Modell aktuell zu halten

■ Patentfamilie: 67386428

»A self-learning animal identification algorithm«

Ansatz wie das System stabil gehalten wird gegen Veränderung, wie bspw. durch Veränderung der Gesichtspportionen anhand von Wachstum.

◦ 2019 07 03: SE1800191, Delaval Holding Ab

■ Patentfamilie 65927456

»A method for livestock identity non-intervention automatic registration and identification«

Kontinuierliches Erheben aktueller Bilddaten und Hinzufügen zur Datenbank.

◦ 2019 04 05: CN109583400, Chengdu Muyun Huishi Tech Co Ltd

2.2.2.1 Verteilung nach Ländern

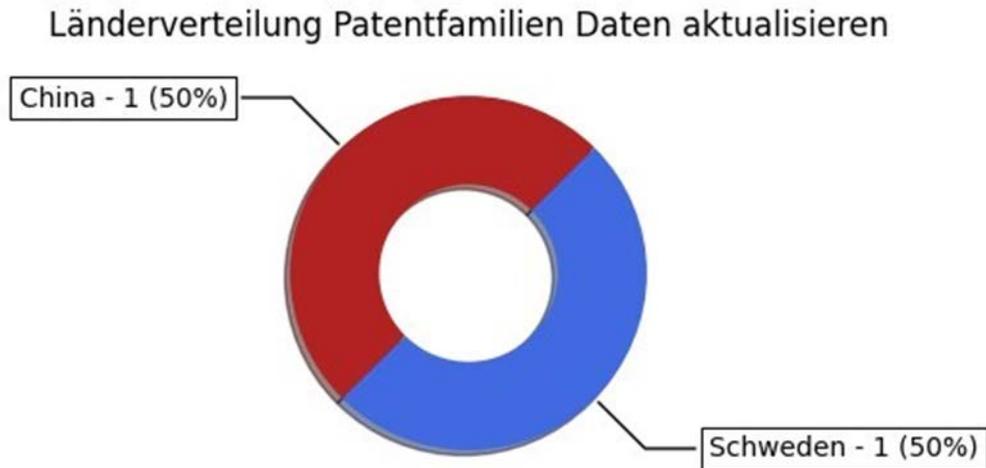


Abbildung 2: Länderverteilung Patentfamilien Daten aktualisieren

2.2.3 Teillösung Gesichtserkennung, ohne Referenzdatenabgleich

■ Patentfamilie 71204308

»Cattle face recognition method and device«

Netze nicht näher beschrieben. Mehrere Bilder eines Gesichts werden zusammengeführt (engl. merged).

◦ 2020 07 03: [CN111368657](#), Jd Digits Holding Co Ltd

■ Patentfamilie 68360529

»Cow face training method and device«

Gesichtserkennung Kuh mittels beschriebenerem CNN, Image-Größe 224*224*3.

◦ 2019 11 05: [CN110414369](#), Inst Animal Husbandry & Veterinary

Medicine Anhui Academy Agricultural Sciences

■ Patentfamilie 68028680

»Cow face and cow face key point detection method based on Mask-RCNN«

Erstellen und Trainieren eines Mask-RCNN mit Eingangs-Image von 1024*1024*3, VGGNet und FPN.

◦ 2019 10 01: [CN110298291](#), Univ Jilin

■ Patentfamilie 67568386

»Animal face recognition method and device, computer equipment and storage medium«

Bilinear-CNN, structural similarity index (SSIM) um die Ähnlichkeit zweier Images zu vergleichen.

◦ 2019 08 16: [CN110135231](#), Hangzhou Huimv Tech Co Ltd

■ Patentfamilie 64600562

»A method and system for cow face alignment based on depth neural network«

Gesichtserkennung speziell für Kühe mittels CNN. Respektiert unterschiedliche Aufnahmewinkel und Veränderung durch Alterung.

◦ 2018 12 14: [CN109002769](#), Shenyuan Hengji Tech Co Ltd

■ Patentfamilie 64094327

»Method and apparatus for extracting livestock feature vectors, computer device and storage medium«

Generelle Gesichtserkennung bei Tieren mittels VGG19, CNN .

◦ 2018 11 13: [CN108805137](#), Ping an Tech Shenzhen Co Ltd

◦ 2019 10 24: [WO2019200735](#), Ping an Tech Shenzhen Co Ltd

■ Patentfamilie 64004219

»Method and system for facial identification of animals under small sample conditions«

Gesichtserkennung von Tieren, mit Bildern aus verschiedenen Winkeln aufgenommen. Berücksichtigt neben gängigen Schlüsselpunkten wie Augen, Ohren, Nase auch die Textur.

◦ 2018 11 06: [CN108764159](#), Beijing Res Ct Information Tech Agriculture

■ Patentfamilie 61175362

»Detection algorithm of animal face key point«

Grundlegende Beschreibung wie Gesichtserkennung bei Tieren mit Bildern mit verschobenen Winkeln umgesetzt werden kann, CNN.

◦ 2018 02 16: [CN107704817](#), Chengdu Pinguo Technology Co Ltd

2.2.3.1 Verteilung nach Ländern

Länderverteilung Patentfamilien Teillösung Gesichtserkennung

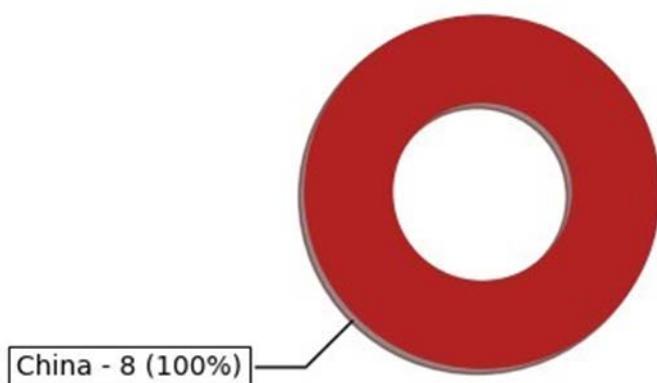


Abbildung 3: Länderverteilung Patentfamilien Teillösung Gesichtserkennung

2.2.4 Alternative Ansätze

■ Patentfamilie 68711128

»Non-contact animal identity vein recognition scanner and recognition method«

Mit LED-Laser werden Venenmuster abgetastet. Vermutlich deutlich schwerer umzusetzen als auf Gesichtserkennung basierender Ansatz, da 1) Fokussierung schwieriger als konventionelle Kamera, 2) keine automatische Erkennung der zuvor verwendeten Körperregion, dadurch kann Körperregion teilvariieren, somit auch das aufgenommene Muster.

◦ 2019 12 06: [CN110543852](#), Hongyi Tech Development Tianjin Co Ltd

■ Patentfamilie 64012134

»Livestock identification method and device, and storage medium«

Nebst klassischer Gesichtserkennung werden hier zusätzlich Klauen, Ohren und Hinterteil mit dem Gesicht zu einem Bild zusammengefügt und dieses durch ein CNN klassifiziert, welches detailliert beschrieben wird ([0092] und [0093]).

◦ 2018 11 06: [CN108764045](#), Ping an Tech Shenzhen Co Ltd

◦ 2019 10 31: [WO2019205377](#), Ping an Tech Shenzhen Co Ltd

■ Patentfamilie 60009746

»Methods and Systems for Biometric Identification of Dairy Animals Using Vein Pattern Recognition«

Identifikation durch Venenmuster am Euter. Das gesamte Herdenmanagementsystem ist beschrieben. Der Vergleichsalgorithmus ist leider nicht Bestandteil der Erfindung.

◦ 2018 03 22: [US2018082016](#), Src Inc

◦ 2018 03 29: [WO2018057840](#), Src Inc

■ Patentfamilie 58951879

»Dairy cow individual recognition method based on deep convolutional neural network«

Anstelle des Gesichts wird die Textur der Seitenansicht zur Identifikation verwendet. Detaillierte Beschreibung des Netzes, CNN, Caffe Framework.

◦ 2017 05 31: [CN106778902](#), Univ Hebei Technology

■ Patentfamilie 55100431

»Dairy cow identification method and system«

Verwendet Bilder von der Seite und des Rückens von Kühen. Kein Neuronales Netz, sondern ein Perceptual Hash Algorithmus, der das Bild zunächst stark verkleinert um anschließend einen Hash zu berechnen, welcher mit bereits in der Datenbank befindlichen abgeglichen wird.

◦ 2016 01 20: [CN105260750](#), Beijing Res Ct Inf Tech Agri

■ Patentfamilie 53524612

»System and Method for Identification of Individual Animals Based on Images of the Back«

Es wird ein topographisches Abbild des Rückens der Tiere erstellt und mit Referenzdaten verglichen.

- 2017 01 05: [CA2989258](#), Viking Genetics Fmba
- 2017 01 05: [WO2017001538](#), Viking Genetics Fmba
- 2018 01 18: [AU2016286295](#), Viking Genetics Fmba
- 2018 03 20: [CN107820616](#), Viking Genetics Fmba
- 2018 06 29: [EA201890099](#), Викинг Генетикс Фмба
- 2018 08 15: [MX2017016878](#), Viking Genetics Fmba
- 2019 06 28: [NZ738150](#), Viking Genetics Fmba
- 2020 01 23: [AU2019283978](#), Viking Genetics Fmba
- 2020 05 07: [US2020143157](#), Viking Genetics Fmba

2.2.4.1 Verteilung nach Ländern

Länderverteilung Patentfamilien alternative Ansätze

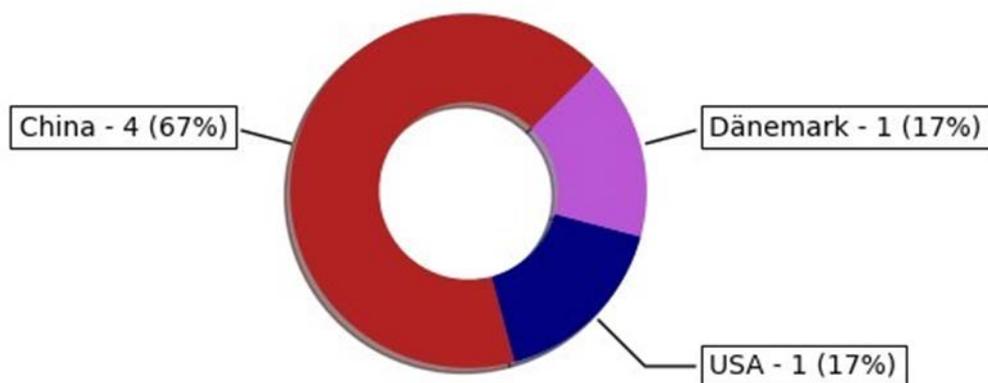


Abbildung 4: Länderverteilung Patentfamilien alternative Ansätze

2.2.5 Verteilung nach Ländern alle Ansätze gesamt

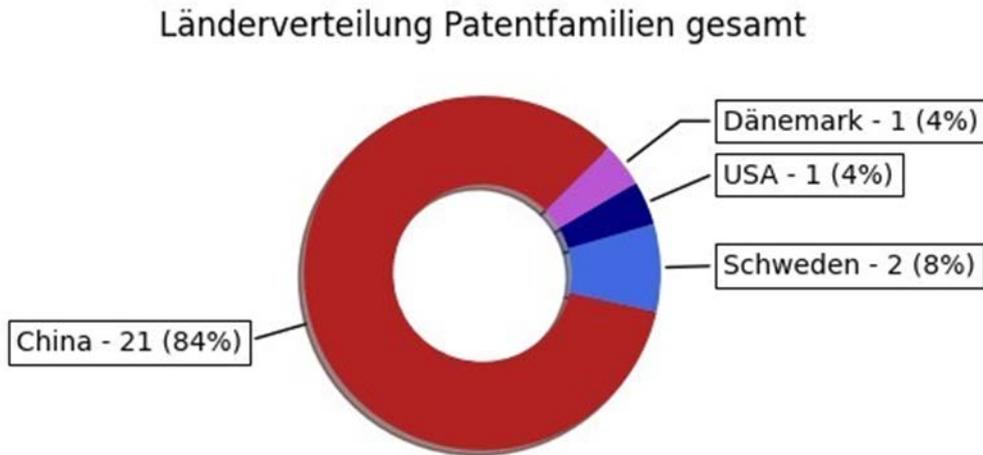


Abbildung 5: Länderverteilung Patentfamilien gesamt

2.3 Zusammenfassung Patentrecherche

2.3.1 Patentrecht

Das deutsche Patentrecht deckt sich weitestgehend mit dem europäischen. Hierzulande und in Europa sind Patente ein Mechanismus um hauptsächlich Erzeugnisse des sekundären Sektors zu schützen: eine angemeldete Erfindung muss mindestens einen real-stofflichen Anteil haben und mindestens im Teil ein real-stoffliches, technisches Problem lösen. Reine Softwarepatente sind in der EU nicht üblich, vielmehr werden Softwareerzeugnisse als Form geistiger Schöpfung durch das Urheberrecht geschützt.

Auch wenn die Schlüsseltechnologie der Lösung des Problems der optischen Tiererkennung klar im Softwarebereich zu suchen ist, so muss ein System welches die Lösung umsetzt, immer auch mit technischen Mitteln einen Bezug zu real-stofflichen Tieren herstellen und hat rückwirkend wiederum Einfluss auf agrarwirtschaftliche Prozesse durch technische Mittel. Deshalb wird davon ausgegangen, dass eine Lösung sehr wohl in Europa patentierbar ist.

2.3.2 Motivation Patente anzumelden

Patente können aus unterschiedlichen Motivationen heraus angemeldet werden. Zunächst natürlich um eine Erfindung zu schützen. Aber eben auch um ein juristisches Mittel gegen Konkurrenten zur Verfügung zu haben, sollten diese im Patent beschriebenen Bereich tätig werden. Manche Patente sind so allgemein gehalten, dass im engeren Sinne nicht mehr von einer konkreten Erfindung die Rede sein kann, denn eher von einer Idee. Eine Hilfe beim Abschätzen inwiefern ein Patent tatsächlich erfolgreich umgesetzt wurde, ist die Betrachtung des Detailgrads mit dem die Erfindung beschrieben wird.

2.3.3 China

Es wurden hauptsächlich chinesische Patente gefunden. Den Dokumenten ist zu entnehmen, dass dort derzeit eine sehr starke Automatisierung im Agrarsektor stattfindet. Weiterhin kann davon ausgegangen werden, dass in China Methoden künstlicher Intelligenz extrem massiv auf sehr breiter Basis in vielen Bereichen eingesetzt werden, sodass das Vordringen dieser Technologien in alle Domänen zu erwarten ist, so wie dies im Agrarsektor derzeit auch geschieht.

Aus Sicht der angemeldeten Patente mit Thema der optischen Tieridentifikation mittels Methoden des maschinellen Lernens ist China absolut führend.

2.3.4 Alter der Patente

Im Bereich des maschinellen Lernens findet zurzeit global eine intensive Weiterentwicklung statt. Dies hat zur Folge, dass Methoden laufend verbessert werden und ständig leistungsfähigere Modelle entstehen. Es kann also die Faustregel aufgestellt werden, dass jüngere Ansätze, sofern sie mit ausreichendem Detailgrad beschrieben wurden, bessere Ergebnisse liefern als ältere.

2.3.5 Qualität und Quantität der Daten

Was sich anhand von Patenten nicht abschätzen lässt, aber für Modelle künstlichen Lernens von sehr großer Bedeutung ist, sind die Menge und die Qualität der Daten mit denen ein Model trainiert wurde. Die Daten werden selten umfänglich beschrieben und sind auch nicht Bestandteil eines Patents. Jedoch sind die Erfolge, wie sie derzeit in der Humanidentifikation verzeichnet werden hauptsächlich auf das Vorhandensein sehr großer Mengen an qualifizierten Bildern mit Menschen zurückzuführen¹. Diese Daten existieren für Tiere nicht im selben Umfang. Um diesem Mangel entgegenzuwirken, verwendet beispielsweise ein Patent (71904404) die Methode des Transfer learnings, wo Modelle Verwendung finden, die ursprünglich mit Humandaten trainiert wurden. Prinzipiell kann es gut funktionieren, da sich Teile eines neuronalen Netzes durchaus zum Wiederverwenden eignen. Auf lange Sicht wäre natürlich das Trainieren mit solchen Daten besser, die auch später zur Klassifizierung Anwendung finden sollen. Da es aber auf absehbare Zeit keine so umfangreichen Datensätze speziell für Bovine geben wird, sodass ein neuronales Netz ausschließlich damit trainiert werden könnte, ist transfer learning eine gute Methode um mit diesem Mangel umzugehen.

Auch was die Leistungsfähigkeit der einzelnen Modelle angeht wird geschwiegen. Diese ist relativ zur Herdengröße und der Heterogenität der Tiere. Eine erfolgreiche Identifikation ist in einer Herde bestehend aus fünf sehr unterschiedlichen Tieren wesentlich einfacher, als wenn die Herde eine Größe von beispielsweise 10.000 Tieren hat und aus sich sehr ähnlichen Tieren besteht.

Wie erfolgreich ein Modell also ist, kann anhand einer Patentschrift nicht abschließend bewertet werden.

¹ Im kommerziellen Kontext kann von Datensets mit der Größe von 10-100Mio Bildern ausgegangen werden.

2.3.6 Auswertung Patentrecherche

2.3.6.1 Gesichtserkennung mit CNN

Es kann Zusammengefasst werden, dass sich deutlich der Ansatz der Gesichtserkennung mittels Convolutional Neuronal Networks (CNN) als Lösung in den meisten Patenten durchsetzt. CNNs haben sich in den letzten Jahren als leistungsfähige Methode erwiesen, wenn es um das Klassifizieren von Bilddaten geht. Sie werden intensiv eingesetzt und die so entstandene Erfahrung beim Erstellen von Modellen für die menschliche Gesichtserkennung wird nun auch beim Erkennen und Identifizieren von Tieren angewandt.

Grundlegend geht es bei CNNs darum, dass es verschiedene Schichten mit unterschiedlichen Aufgaben gibt. Diese können an die folgende Schicht Daten weitergeben, die auf ein bestimmtes Merkmal reduziert sind, wie beispielsweise ein eingegangenes Bild auf seine Kanten reduzieren. Insgesamt können so im Vergleich zu Netzen bei denen alle Schichten gleich groß und voll miteinander verknüpft sind, deutlich mehr Daten mit der selben Anzahl an Operationen verarbeitet werden. Dies lässt dann deutlich leistungsfähigere Modelle zu als unter Verwendung von klassischen tiefen neuronalen Netzen. Näheres hierzu im Abschnitt Convolutional Neural Networks.

Ein Bild ist der Input eines Modells. Der Output ist eine Liste von Werten die das Ergebnis der Berechnungen auf dem Bild sind, auch Feature-Vektor genannt. Umso besser das Modell ist, umso ähnlicher sind sich die Feature-Vektoren von Bildern eines Tieres unabhängig anderer Eigenschaften des Bildes. Umso weniger wirken sich z.B. unterschiedliche Aufnahmewinkel, oder variierende Lichtverhältnisse auf den Inhalt des Feature-Vektors aus. Verbessert wird dies unter anderem dadurch, dass der Teil des Bildes in dem sich das Gesicht des Tieres befindet erkannt, ausgeschnitten und normalisiert wird. Das bedeutet dass Gesichter auf unterschiedlichen Bildern möglichst so aufbereitet werden, dass Varianz von vorn herein reduziert wird, bevor diese zur Identifikation Verwendung finden.

2.3.6.2 Anforderung an Kameras

Ein großer Vorteil dieses Ansatzes ist es, dass an die Fotos der zu identifizierenden Tiere relativ geringe Ansprüche gestellt werden. Es werden keine Spezialkameras benötigt und zudem ist das Gesicht eines Tieres von Mensch und Maschine leicht zu erkennen und somit leichter verwertbar aufzunehmen, als beispielsweise ein definierter Bereich der Flanke, oder eines Euters, wie es beim Venenscannen notwendig ist. Handelsübliche Kameras wie sie in den allermeisten Mobiltelefonen verbaut sind, können bei ausreichenden Lichtverhältnissen Bilder erzeugen, die für die Identifikation durch Gesichtserkennung verwertbar sind. Mit entsprechender Zoom-Optik könnte dies auch auf größere Distanzen geschehen.

2.3.6.3 Auflösung der Aufnahmen

Die Auflösung der Bilder ist dennoch nicht irrelevant. Je größer die Bilder sind, die vom Modell verarbeitet werden können, um so mehr Daten stehen pro Aufnahme zur Verfügung. Gerade bei sehr großen Herden mit vielen sehr ähnlichen Tieren sind geringe Unterschiede entscheidend für die erfolgreiche Identifikation. Viele Modelle verwenden Bilder in der Größe von ungefähr 220*220 bis 256*256. Sehr wahrscheinlich ist dies zu wenig um in großen Herden eine gute Identifikation gewährleisten zu können. Hier gilt: umso größer ein Bild ist umso besser für die Identifikation. Da aber die Größe von Modellen maßgeblich durch Hardware begrenzt wird und einfache Kameras keine guten hochauflösenden Bilder bei normalen Lichtverhältnissen erzeugen können, ist die Auflösung neben der Quantität und Qualität der Daten eine Messlatte, die erst mit der Zeit langsam erhöht werden kann.

2.3.6.4 Standardisierbarkeit

Ohrmarken verbleiben am Tier.

Ein Punkt der aufgrund des geringen Alters der Technologie noch nirgends erwähnt wurde, ist die Standardisierbarkeit. Soll eine rein optische Identifikation irgendwann marktreif sein, so muss sich mit der Möglichkeit auseinander gesetzt werden, wie die Datensätze die mit einem Tier assoziiert sind (also die Sammlung an Feature-Vektoren eines Tieres) von einem Besitzer zum nächsten übermittelt werden können.

Eine Möglichkeit dies generell zu vermeiden ist das Erstellen eines Zentralregisters. Dies dürfte aber aufgrund von Datenschutzbestimmungen und international unterschiedlichen Rechtsgrundlagen und Begehrlichkeiten schwierig sein. Zudem müssten die Datensätze regelmäßig aktualisiert werden, was innerhalb eines Betriebs mit einer Softwarelösung einfach, aber für ein Zentralregister erheblich schwieriger ist.

Eine andere mögliche Lösung wäre somit das Übermitteln der Datensätze mit dem Tier, sobald es den Besitzer wechselt. Dies allerdings setzt einheitliche Modelle voraus, da Feature-Vektoren ja das Ergebnis der Berechnung eines Modells sind und sich (noch) nicht mit Feature-Vektoren andersartiger Modelle vergleichen lassen. Um dem Entstehen mehrerer unterschiedlicher, proprietärer und inkompatibler Formate entgegenzuwirken, wie sie in vielen anderen Branchen leider gängig sind, müsste dieses Thema früh bedacht werden.

2.3.6.5 Leseempfehlung

Die vom Verfasser empfohlenen, exemplarischen Dokumente sind:

- Komplettlösung, sehr kleines Dataset mit Foto von links, rechts und frontal
 - Patentfamilie: 71208820
 - Patent: Cow face detection and recognition method based on deep learning
- Nur Gesichtserkennung, aber sehr detailliert und mit hoher Auflösung der Bilder
 - Patentfamilie: 68028680
 - Patent: Cow face and cow face key point detection method based onMask-RCNN

3 Marktrecherche

Nach der Patentrecherche wurde eine Marktrecherche durchgeführt. Das Ziel hierbei war nicht nur das Ausfindigmachen von fertigen Produkten, sondern ebenfalls ein Bild darüber zu erzeugen, welche Projekte am Thema arbeiten, unabhängig von etwaiger kommerzieller Verwertung, oder des Reifegrades der Technologie. Insofern wird im Folgenden der Begriff Projekt verwendet um zu verdeutlichen, dass nicht zwingend ein marktwirtschaftlich agierendes Unternehmen, sondern auch eine nichtkommerzielle Forschungseinrichtungen das Thema bearbeiteten kann.

3.1 Suchvorgang

Die Suche wurde zweistufig durchgeführt. Zunächst wurden alle dem LfULG bekannten Unternehmen und Verbände via Email kontaktiert, bei denen die Hoffnung bestand, dass Bezug zum Thema vorhanden ist. Des Weiteren wurden Projekte durch Internetrecherche gefunden, welche dann ebenfalls angeschrieben wurden.

Aus den Antworten dieser Anschreiben konnte ein grobes Gefühl dafür entwickelt werden, wie verbreitet die Forschung am Thema ist. Sehr hilfreich waren die Hinweise und Einsichten die sich aus den anschließenden Kommunikationen ergaben. Manche Projekte haben wenig Medienpräsenz und sind deswegen nur schwer mittels ausschließlicher Onlinesuche zu finden. Die Bereitschaft relevante Kontakte weiter zu vermitteln war in der Gruppe der kontaktierten Personen generell hoch, sodass eine gute Durchdringung erreicht werden konnte. Schlussendlich wurden ca. 95 Akteure insgesamt kontaktiert und mit ungefähr zehn fand weiterführender Informationsaustausch statt. Am Ende der Marktrecherche ergaben sich zusätzlich einige Expertengespräche, die das entstandene Bild abrundeten und bestätigten.

Zu beachten ist, dass natürlich kein absolut vollständiger, Überblick über alle Projekte weltweit generiert werden kann. Einerseits, da die Suche auf den deutschen und englischen Sprachraum beschränkt ist und gerade in China viel Aktivität erwartet wird. Zum anderen sind Veröffentlichungen zu Projekten im universitären Umfeld nicht so umfangreich vor ihrem Abschluss, wie Unternehmen die Aufmerksamkeit für ihre Lösungen erzeugen wollen. Da jedoch durchweg alle Experten die Suchergebnisse bestätigten, wird davon ausgegangen, dass der erstellte Überblick zum aktuellen Zeitpunkt sehr gut ist.

3.2 Motivationen, Möglichkeiten und Verortung

Bevor auf die einzelnen Projekte eingegangen wird, ein paar Worte zu den unterschiedlichen vorgefundenen Motivationen. Diese vermitteln auch einen Blick darüber, welches Potential in der Technologie von den jeweiligen Akteuren vermutet wird.

3.2.1 Tierortung

Um den Aufwand der Tierortung bei großen Weideflächen in den USA zu reduzieren, forscht ein Projekt an der University of Kentucky an Tierortung und Identifikation mittels Drohnen.

Da durch den Verlust von Tieren durch Gesundheitsprobleme dem USamerikanischen Agrarsektor jährlich ca. 1Mrd USD verloren gehen, investiert hier auch das US-amerikanische Landwirtschaftsministerium.

3.2.2 Systemanbieter

Die in Europa gefundenen Projekte haben die Motivation der Produktentwicklung als Systemanbieter im Herdemanagement. Die aktiven europäischen Projekte sind derzeit ausschließlich im UK angesiedelt.

Große Systemanbieter scheinen kein Interesse an der Forschung im thematisierten Bereich zu haben, es handelt sich bei den gefundenen Projekten um kleinere oder kleine mittelständische Unternehmen.

Dadurch dass Ohrmarken zwar mit Nachteilen behaftet aber überall vorhanden sind, ist die Motivation diese nicht zu verwenden und ein eigenes Identifikationssystem zu entwickeln nur moderat. Das einzige in Europa gefundene aktive Projekt ist Cattleeye. Die Entwicklung der erforderlichen Modelle ist doch umfangreicher, als einige Hersteller zunächst annahmen, weshalb die Technologie bisher noch keinen Zugang zu marktreifen Produkten fand. Die Projekte Cainthus, Aktivstall und Agsenze forschten zwar erfolgreich an der Technologie, der Aufwand welcher zur Marktreife notwendig gewesen wäre, überstieg jedoch den ökonomischen Nutzen.

3.2.3 Biosecurity und Healthcare

Um den Viehbestand in den USA vor der Verbreitung ausländischer Krankheiten zu schützen, entwickelt das Projekt Cattletracks, unter Verwendung von Gesichtserkennung bei Rindern ein Rapid Disease Traceability System (System zur schnellen Seuchenverfolgung). Ein gesteigertes Interesse an diesem Ansatz seitens der Fleischindustrie konnte bei der Recherche identifiziert werden.

Des Weiteren besteht generell globales Interesse an der Zurverfügungstellung des Werdegangs von Tieren, sowie deren Gesundheitsakte inklusive Einblick in durchgeführte veterinärmedizinische Maßnahmen. Diese Möglichkeiten werden oft mit der Identifikation von Tieren anhand optischer Merkmale verknüpft.

[KSN20], [BHF20], Stellapps, Cattletracks

3.2.4 Schutz vor Versicherungsbetrug

Wie chinesischen Patentschriften zu entnehmen ist, scheint die Hauptmotivation in China der Schutz vor Versicherungsbetrug zu sein. Selbiges trifft auch auf indische Projekte zu, die in der Marktrecherche gefunden wurden.

Ohrmarken lassen sich entfernen und austauschen und bieten somit keinen sicheren Schutz vor intendierter Fehlidentifikation der Tiere. Diese Fälle scheinen zumindest so zahlreich zu sein, dass Versicherungsunternehmen als Hauptinteressenten für die Entwicklung alternativer, fälschungssicherer Methoden mittels einfacher, auf optischem Wege zu erfassender Merkmale der Tiere im asiatischen Raum benannt werden können.

Die Region und Motivation ist aus mehreren Gründen bemerkenswert:

- Das Ziel, Ohrmarken abzulösen, ist hier keine interessante Eigenschaft die ein Produkt erweitern könnte, wie bei europäischen Systemanbietern der Fall, sondern es ist die Hauptmotivation.
- Das intensive Vorhandensein von Knowhow im KI-Bereich in den beiden Ländern vereinfacht die Entwicklung.
- Das der Forschung und Entwicklung zur Verfügung stehende Kapitalvolumen ist hier als nicht unerheblich einzuschätzen, da es dem Schaden durch Versicherungsbetrug mehr oder minder direkt gegenübersteht. Zudem sind Versicherungsunternehmen finanzkräftiger als kleine und mittelständische Unternehmen im europäischen Agrarsektor.

Insofern ist zu erwarten, dass chinesische und indische Unternehmen die besten Voraussetzungen haben um eine marktreife Lösung zu entwickeln. Es lässt sich derzeit allerdings nicht bewerten, wann diese im europäischen Markt ankommen wird und welcher Reifegrad zu erwarten ist.

3.3 Aktive Projekte

Viele Projekte arbeiten mit am Tier angebrachten Sensoren und Transmittern. Da die Fragestellung dieser Studie klar die Verwendung von am Tier verbrachten Geräten ausschließt, wurden Projekte deren Technologie solche Methoden verwendet, in der Recherche nicht berücksichtigt. Ebenso sind hier keine Projekte enthalten bei denen die Funktionsweise nicht geklärt und somit ein Auslesen von Transmittern und/ oder Ohrmarken zu Identifikationszwecken nicht ausgeschlossen werden konnte.

Im Folgenden sind die aktiven und für das Thema relevanten Projekte, absteigend nach Relevanz sortiert dargestellt.

3.3.1 Cattleeye

Sitz: Belfast, Nordirland, UK

Website: cattleeye.com

Beschreibung und Einschätzung:

Cattleeye wurde im Januar 2019 gegründet und bestehend aus einem kleinen aber versierten Team von Spezialisten. Das Unternehmen ist Systemanbieter im Bereich Monitoring von Gesundheit und Leistungsfähigkeit von Milchkühen. Die Überwachung wird u.a. mit einfachen Kameras an Ein- und Ausgang am Melkstand umgesetzt. Die Ergebnisse sind als erfolgversprechend einzustufen.

Das Unternehmen ist der einzige gefundene Anbieter aus Europa mit einem System das Individualidentifikation beinhaltet. Sollte eine Erprobung der Technologie mittels Teststellung angestrebt werden, ist Cattleeye das Unternehmen welches hierfür derzeit am ehesten in Frage kommt.

3.3.2 Cainthus

Sitz: Dublin, Irland

Website: cainthus.com

Beschreibung und Einschätzung:

2014 als IT-Unternehmen mit Fokus Computer-Vision im Agrarsektor gegründet.

Cainthus ist das Unternehmen mit der meisten Erfahrung, wenn es um optische Tieridentifikation im bovinen Bereich geht.

Derzeit besteht die Zielgruppe aus Großfarmen im US-amerikanischen Markt.

Es werden Tiere auf Stallebene (~400 Tiere) unterschieden, allerdings nicht einzeln identifiziert. Dies entspricht auch den Arbeitsabläufen der meisten Kunden. Erforderliche Eingriffe erfolgen pro Stall, was ökonomisch und tierwohlbezogen sinnvoller ist, als reaktive, tierindividuelle Maßnahmen nach einem Problem.

[FNAC20-1], [FNAC20-2], [BPDHI18]

3.3.3 Cattletracks

Sitz: Leavenworth, Kansas, USA

Website: blackhereford.com; cattletracs.com

Beschreibung und Einschätzung:

Das Projekt startete 2019. Die Idee ist, ein »National Animal Disease Traceability System« zu entwickeln.

Das System wird in Kooperation des Agrarunternehmens J&N Ranch, LLC mit der Kansas State University umgesetzt. Zunächst wurde begonnen, Kurzvideos von Tieren zu sammeln um eine Datenbasis zu begründen und erste Modellberechnungen durchzuführen. Schließlich wurde eine App für Apple-Handys entwickelt, mit der Benutzer Fotos von Rindern hochladen können. Eine App für Android-Geräte ist mittlerweile ebenso verfügbar. Die Aufnahmen sorgen für ein kontinuierliches Wachsen der Trainingsdaten und ein verbessern des Modells.

Positiv bewertet wird bei diesem Ansatz, dass:

- der Zugang für Kunden und Teilnehmer sehr niederschwellig ist: die App kann einfach installiert werden und ist kostenlos.
- die Teilnahme freiwillig ist. In den USA kann eine drohende Zwangsteilnahme durch behördlichen Druck schnell zur Ablehnung führen.
- in Qualität und Diversität die App genau die Trainingsdaten liefert, die später, wenn das System weiter fortgeschritten sein wird, auch hauptteilig verarbeitet werden.
- das System kontinuierlich mit dem Trainingsdatenvolumen wachsen kann.

Vermutlich wird ein 3D-Modell des Tiergesichts erstellt, da die App mehrere Bilder aus verschiedenen Winkeln aufnimmt.

Angeblich besteht großes Interesse seitens der Fleischindustrie an dem Projekt, mit dem Wunsch eine Nachverfolgbarkeit der Tiere herzustellen.

Es bleibt jedoch abzuwarten, wie erfolgreich die Forschung am Model ist. Umfangreiche Daten bedeuten ja nicht, dass die Forschung schneller geht, sie wird erst durch deren Vorhandensein ermöglicht.[KSN20], [BHF20]

3.3.4 University of Kentucky

Sitz: Lexington, Kentucky, USA

Website: Keine eigene Seite, News-Artikel hauptsächlich der Universität Kentucky wo das Projekt angesiedelt ist: [UKY19], [UKY20], [BBCUKY]

Beschreibung und Einschätzung:

Gerade die Tierortung sei bei großen Weideflächen und geringer zur Verfügung stehender Arbeitszeit eine Belastung für die hiesigen Rinderproduzenten. [UKY19]

Um die Entwicklung eines Modells zu ermöglichen, wurden Züchter aufgerufen, Bilder von Rinden via Email einzusenden.

Der zweite Artikel erschien ein Jahr später und beschreibt gesteigertes Interesse am Projekt. Weitere Fachkräfte, speziell auch für die Entwicklung geeigneter Drohnen welche die Tiere wenig irritieren, wurden hinzugezogen. [UKY20]

Das US-amerikanische Landwirtschaftsministerium unterstützt das Projekt seit Anfang 2018 mit 900.000USD. [UKY20], [BBCUKY]

Hierbei wird ein 3D-Modell des Tiergesichts für die Gesichtserkennungsalgorithmen angewandt.

Leider blieben Kontaktversuche erfolglos.

3.3.5 Stellapps

Sitz: Bangalore, Indien

Website: stellapps.com

Beschreibung und Einschätzung:

2010 gegründet hat Stellapps derzeit rund 400 Mitarbeiter, bezeichnet sich allerdings selbst noch als Startup. An der Tieridentifikation arbeitet ein relativ kleines Team, mit teilweise wechselnden Teilnehmern.

Die beiden Motivationen für die Entwicklung eines Tieridentifikationssystems sind:

- Schutz vor Versicherungsbetrug
- präventive Gesundheitsfürsorge

Es wurde eine App für Android entwickelt, mit der auf betriebsebene Tiere registriert werden können. Dies reduziert die Population innerhalb derer Tiere zuverlässig identifizierbar sein müssen auf die Anzahl der Tiere pro Betrieb.

Es bestehen derzeit drei Testinstallationen in Indien und eine in Pakistan.

Stellapps ist auf der Suche nach Partnern, mit denen das Projekt weiter vorangetrieben werden kann.

Das Model verwendet Gesichtserkennung. Über die Erkennungsrate und die Größe der Datenbasis konnten ohne Vertraulichkeitsvereinbarung keine Informationen erlangt werden.

[INNW10]

3.4 Forschungseinrichtungen Deutschland

Folgend die im Agrarbereich agierenden Forschungseinrichtungen in Deutschland mit vermutetem starken Bezug zum Thema, auch wenn hier aktuell keine aktive Forschung im thematisierten Bereich betrieben wird.

3.4.1 Universität Kassel

Sitz: Witzenhausen, Deutschland

Website: Agrartechnik Universität Kassel: <https://www.uni-kassel.de/fb11agrар/fachgebiete/-/einrichtungen/agrartechnik/startseite>, Dr. Abozar Nasirahmadi: <https://www.uni-kassel.de/fb11agrар/fachgebiete/-/einrichtungen/agrartechnik/dr-abozar-nasirahmadi>

Beschreibung und Einschätzung:

Dr. Abozar Nasirahmadi Head of Machine Learning and Smart Systems der Universität Kassel, Fachbereich Ökologische Agrarwissenschaften im Fachgebiet Agrartechnik.

Dr. Nasirahmadi forscht seit vielen Jahren an Computervision und künstlichen Neuronalen Netzen im Bereich der Agrarwissenschaften und hat somit einen sehr guten Überblick über Akteure und Methoden. Ein weiteres Bearbeiten des Themas sollte in jedem Fall seine Beteiligung anstreben.

3.4.2 Leibniz-Institut für Nutztierbiologie (FBN)

Sitz: Dummerstorf, Deutschland

Website: FBN allgemein: fbn-dummerstorf.de.

Beschreibung und Einschätzung:

Das Leibniz-Institut für Nutztierbiologie (FBN) in Dummerstorf bei Rostock ist ein staatlich grundfinanziertes, außeruniversitäres Forschungsinstitut in Form einer Stiftung des Öffentlichen Rechts des Landes Mecklenburg-Vorpommern.

Derzeit findet zum Thema der optischen Tieridentifikation für Kühe keine Forschung am FBN statt, allerdings besteht dort generelles Interesse an solchen Methoden mindestens als Werkzeug für weitere Forschungen im Nutztierbereich.

Aufgrund der umfassenden thematischen Verortung des FBN ist davon auszugehen, dass das Erscheinen künftiger Lösungen im Bereich der optischen Tieridentifikation am ehesten hier erfasst und bewertet werden. Da die vorliegende Recherchestudie ein Bild der aktuellen Situation darstellt und keine Dauerbeobachtung ist, sollte eine perspektivische Zusammenarbeit mit dem FBN generell angestrebt werden. Ansprechpartner ist u. a. Dr. Christian Manteuffel.

3.4.3 Cattlehub

Sitz: Bonn, Deutschland

Website: cattlehub.de

Beschreibung und Einschätzung:

Ein durch das Bundesministerium für Ernährung und Landwirtschaft geförderter Zusammenschluss der Universität Jena, des Thünen-Institut für Agrartechnologie Braunschweig, der Universität Bonn, des sächsischen Landesamtes für Umwelt, Landwirtschaft und Geologie, der Friedrich-Schiller-Universität Jena, der Technischen Universität Dresden, der Landwirtschaftskammer Nordrhein-Westfalen und der Technischen Universität Chemnitz.

Selbstbeschreibung:

Im Experimentierfeld »CattleHub« geht es um den Einsatz digitaler Technik in der Rinderhaltung. Dabei soll vor allem die Etablierung digitaler Assistenzsysteme in der Rinderhaltung unterstützt und aktiv gestärkt werden. Ziel ist, dass Landwirte und deren Tiere durch verbesserte Entscheidungsempfehlungen profitieren.

3.5 Forschungseinrichtungen UK

Im Folgenden sind die im Agrarbereich agierenden Forschungseinrichtungen im Vereinten Königreich mit vermutetem starken Bezug zum Thema aufgeführt, auch wenn hier aktuell keine aktive Forschung im thematisierten Bereich betrieben wird.

3.5.1 Agricultural Engineering Precision Innovation Centre (Agri-EPI)

Sitz: Easter Bush, Roslin, Schottland, UK

Website: agri-epicentre.com

Beschreibung und Einschätzung:

Das Agricultural Engineering Precision Innovation Centre ist eine gut vernetzte, staatlich grundfinanzierte Forschungseinrichtung.

Derzeit werden hier keine Forschungsprojekte mit Themennähe zur optischen Tieridentifikation betrieben.

3.5.2 Rothamsted Research

Sitz: Harpenden, England, UK

Website: rothamsted.ac.uk

Beschreibung und Einschätzung:

Rothamsted Research ist eine staatlich grundfinanzierte Forschungseinrichtung mit eigener Forschungsfarm.

Kontakt zu Dr. Khalid Mahmood wurde hergestellt. Es besteht seinerseits Interesse an Kooperation und Forschungsaufträgen. Er ist ehemaliges Teammitglied des Milkalyser-Projekts und für Anfragen und Informationsaustausch offen.

Derzeit werden hier keine Forschungsprojekte mit Themennähe zur optischen Tieridentifikation betrieben.

3.6 Inaktive Projekte

Folgend alle gefundenen Projekte die vormals am Thema gearbeitet haben und aus verschiedenen Gründen nicht mehr im Themenbereich der rein optischen Tieridentifikation mittels biometrischer Daten aktiv sind

3.6.1 Aktivstall

Sitz: Harpenden, England, UK

Website: rothamsted.ac.uk

Beschreibung und Einschätzung:

Aktivstall ist ein Pionier in der Gesichtserkennung für Tiere in Deutschland. Zwar nicht für Kühe, sondern für Pferde, aber der Ansatz der 2012/ 2013 versucht wurde, entspricht im Wesentlichen dem der heute angewandten Methoden.

Leider zeigten die Forschungsergebnisse damals, dass eine Identifikation zwar durchaus machbar, aber die Produktreife im angestrebten Marktsegment ökonomisch nicht sinnvoll umsetzbar ist.

3.6.2 Agsenze

Sitz: Lancaster, England, UK

Website: agsenze.com

Beschreibung und Einschätzung:

Agsenze ist ein Tier-Monitoring-Systemhersteller aus England, der maschinelles Lernen einsetzt um eine Gesundheitsbewertung von Tieren durchzuführen.

In der Vergangenheit wurde an optischer Tieridentifikation geforscht, allerdings fand diese Technologie keinen Zugang zu aktuellen Produkten. Es besteht aber mittel- bis langfristig der Wunsch, dieses Thema wieder aufzugreifen.

3.6.3 Milkalyser

Sitz: Exeter, England, UK

Beschreibung und Einschätzung:

Das Projekt existiert nicht mehr.

Ehemals Zusammenschluss des Unternehmens Milkalyser (mittlerweile Lely), der Universität von Exeter und Rothamsted Research.

Das Projekt wird von Lely nicht fortgeführt und die ehemaligen Wissensträger sind teilweise nicht für Lely tätig.

Ziel war die Identifikation mit optischen Methoden anhand biometrischer Daten.

3.7 Ungeklärte Projekte

Folgend Projekte von denen angenommen wird, dass aktuell Forschung und Entwicklung zum Thema betrieben wird, allerdings aufgrund der mangelnden Informationslage keine weiteren Aussagen getroffen werden können.

3.7.1 Techvantage

Sitz: Kerala, Indien

Website: techvantagesystems.com/moo-id

Beschreibung und Einschätzung:

Techvantage ist ein großer Dienstleister im Bereich Datascience und künstliche Intelligenz. Es werden Lösungen für diverse Branchen angeboten.

Das Unternehmen hat vermutlich eine App entwickelt, mit der Rinder registriert werden können. Der Registrator authentifiziert sich mittels seiner Aadha a r id.

Die Motivation ist Schutz vor Versicherungsbetrug.

Ohne Vertraulichkeitsvereinbarung war das Unternehmen nicht zur Kommunikation bereit.

3.7.2 Moofarm

Sitz: Gurugram, Indien

Website: <https://mooo.farm/>

Beschreibung und Einschätzung:

Laut des Artikels des Innovation News Network, arbeitet moofarm an einer Lösung zur optischen Tieridentifikation. Die Erkennungsrate liegt nach eigenen Angaben bei 95%, allerdings gibt es weder Angaben zur Größe noch zur Qualität des Datensets und die Kontaktaufnahme blieb erfolglos. [INNW10]

3.8 Bewertung der Marktsituation

3.8.1 Beurteilungsfähigkeit der Modelle

Da die Entwicklung der Modelle zur optischen Tieridentifikation sehr aufwändig ist und bei der Suche nach den passenden Parametern Insiderwissen viel helfen kann, geben Hersteller verständlicherweise nur begrenzt Informationen über die Herzstücke ihrer Systeme preis. Offengelegt wird meist nur die Erkennungsrate und eventuell eine vage Beschreibung der Daten. Allerdings ist das Bewerten ohne genauere Kenntnis der zum Erstellen des Modells verwendeten Datensätze, insbesondere deren Qualität und Umfang, nur sehr begrenzt möglich.

Wenn beispielsweise ein großes Dataset aus vielen hochauflösenden und detailreichen Aufnahmen sehr heterogener Tiere besteht, ist eine Erkennungsrate von 95% weniger gut, als wenn die selbe Erkennungsrate erreicht wird mit Daten von der selben Anzahl an gleichartigen Tieren und mit heterogenen Aufnahmebedingungen (Licht, Winkel, Detailgrad, Auflösung, usw.).

3.8.2 Komplexität und Vergleich mit Ohrmarkensystem

Es existiert zu diesem Zeitpunkt (Anfang 2021) kein System, das auch nur ansatzweise in der Lage ist, Ohrmarken abzulösen. Dies liegt daran, dass die Komplexität der Modelle mit der Anzahl der zu unterscheidenden Tiere enorm steigt. Die kleinste Menge an Tieren, die derzeit mit Ohrmarken ein Identitätsmerkmal erhalten, ist die eines Bundeslandes. [BMEL19] Diese Zahl ist viel zu groß, als dass ein existierendes KI-Modell alle Tiere hiervon auch nur ansatzweise zuverlässig unterscheiden könnte.

Eine mögliche Eingrenzung dieses Problems könnte mit der Verkleinerung der Menge der zu unterscheidenden Tiere erreicht werden. Es könnte die Identifikationsmerkmalshierarchie um eine weitere, kleinere Untergruppe erweitert werden, beispielsweise die der Besitzer. Damit würden nicht maximal $99.999.999^2$ Tiere unterscheidbar sein müssen.

Aktuelle Modelle arbeiten mit weniger als 1.000 Tieren.

² 99.999.999 ist die maximale Anzahl an Möglichkeiten mit 8-stelligen Ohrmarkenziffern, ohne 0. Es wird hier davon ausgegangen, dass es pro Besitzer nicht so viele Tiere gibt.

3.8.3 Prognose Marktreife

Es wird hier nicht davon ausgegangen, dass innerhalb der nächsten fünf Jahre die Marke der 10.000 voneinander zu unterscheidenden Tiere erreicht wird. Diese Einschätzung geht davon aus, dass die Entwicklung der optischen Tieridentifikation ungefähr weiterhin so verläuft wie dies bisher der Fall war.

Es muss auch explizit darauf hingewiesen werden, dass es derzeit generell sehr viele Forschungsprojekte im Bereich des maschinellen Lernens gibt, vollkommen unabhängig vom Gegenstand der untersuchten Thematik. Meist schaffen es sehr wenige dieser Ansätze zur Marktreife. Es werden zwar künftig immer mehr Produkte mit KI-Technologien auf den Markt kommen, bis diese jedoch auf breiter Basis im klassischen Mittelstand wirklich ankommen dürfte es noch einige Jahre dauern.

Umstände welche die Verfügbarkeit von Produkten mit optischer Tieridentifikation beschleunigen können sind:

- Es arbeiten im englisch- und deutschsprachigen Raum sehr wenige Projekte aktiv am Thema. Aufgrund der Sprachbarriere kann nur sehr schwer abgeschätzt werden, wie viele Produkte in China existieren. Bei der Fülle an gefundenen Patenten ist jedoch davon auszugehen, dass chinesische Hersteller weiter sind, als europäische, oder amerikanische. Es ist allerdings fraglich wie stark diese zunächst auf den europäischen Markt drängen, die kulturelle und sprachliche Hürde dürfte hier als recht hoch einzustufen zu sein und die Gewinnmargen im heimischen, asiatischen Markt veranlassen wohl kaum zum Blick nach Europa, da eine enorme Automatisierung dort eine sehr starke Nachfrage vermuten lässt.
- Die KI-Welt ist hochdynamisch, regelmäßige Innovation im Bereich des maschinellen Lernens können die Zuverlässigkeit von Modellen und deren Fähigkeit zu generalisieren erhöhen. Neuere Ansätze sind meist komplexer als ältere und es kann somit etwas dauern, bis diese von Projekten adaptiert werden. Dennoch muss jederzeit mit Leistungssteigerungen durch Innovation gerechnet werden.
- Das Entstehen von hochwertigen Trainingsdatenbibliotheken könnte das ganze Thema erheblich fördern: dass ein Projekt zunächst viel Aufwand betreiben muss, um überhaupt Trainingsdaten zu erhalten, bindet viel Energie und wirkt sicherlich stark hemmend. Würde der Zugang zu Trainingsdaten erheblich erleichtert, würden mehr Projekte am Thema tätig werden.

3.8.4 Zusammenfassung der Marktsituation

Die europäischen Projekte Agsenze, Aktivstall und Milkalyser haben an der Technologie der optischen Tiererkennung geforscht und sie hauptsächlich aufgrund ökonomischer Entscheidungen nicht zur Marktreife vorangetrieben.

Cainthus hat die meiste Erfahrung, allerdings ebenfalls entschieden, dass die Marktreife ökonomisch noch nicht sinnvoll ist. Das Unternehmen wird allerdings begrenzte Individualidentifikationseigenschaften in sein neues Produkt »ALUS Health« integrieren, das Ende 2021 verfügbar sein soll. Die Zielgruppe von Cainthus sind allerdings hauptsächlich Größtfarmen.

Das US-amerikanische Cattletracks wird als geschlossenes System wahrgenommen, das von dessen Nutzern ausschließlich mit einer App auf einem mobilen Endgerät bedient wird. Unabhängig von der Qualität der Erkennungsrate arbeitet das System nicht wie ein Herdemanagementsystem im klassischen Sinne: die automatisierte Erkennung und Re-Identifikation im Stall scheint derzeit kein verfolgtes Anwendungsziel zu sein. Allerdings befindet sich dieses Projekt in der Entwicklungsphase und sammelt

Daten um das Modell zu verbessern. Welche Möglichkeiten sich hier bieten und ob sie umgesetzt werden, sollte beobachtet werden.

Derzeit ist Cattleeye das einzige Projekt mit verfügbarer Teststellung in Europa. Das Projekt ist jung, aber vielversprechend. Die vom Projekt angegebene Erkennungsrate ist die beste hier bekannte. Sollte eine Erprobung der Technologie durch eine bezahlte Teststellung angestrebt werden, so ist Cattleeye derzeit das Unternehmen das hierfür am ehesten in Frage kommt.

Das indische Unternehmen Stellapps arbeitet an einer Lösung die ebenfalls mit einem mobilen Endgerät bedient wird. Auch hier besteht die Möglichkeit einer bezahlten Teststellung.

Von den beiden Projekten mooofarm und Techvantage kann nur gemutmaßt werden, wie gut die Erkennungsrate ist und welche Gestalt für das Produkt angedacht wird, da hier keine weiterführenden Informationen zu erlangen waren.

3.8.5 Expertengruppen

Für alle Folgeaktivitäten sollten mindestens die drei Expertengruppen Cattlehub, das Leibniz-Institut für Nutztierbiologie (FBN) und die Universität Kassel und mit einbezogen werden.

4 Technologie

Mit diesem Text wird einerseits versucht, eine für Laien geeignete Hinführung zum Thema zu erreichen und andererseits auf aktuelle Forschung in den für die optische Tieridentifikation relevanten Bereichen des maschinellen Lernens einzugehen. Hierzu wird top-down vorgegangen, d.h. vom Allgemeinen zum Speziellen. Viele Konzepte müssen stark vereinfacht dargestellt werden.

Des Weiteren wird versucht so weit wie möglich in deutscher Sprache zu bleiben. Allerdings gibt es für manche Konzepte keine gängige Übersetzung. Hier werden die englischen Begriffe verwendet.

4.1 Künstliche Intelligenz

Künstliche Intelligenz (KI) ist eine sehr allgemeine Definition, hierunter versammeln sich viele Konzepte und Technologien, wie auch das derzeit so populäre maschinelle Lernen. KI ist ein junger³, interdisziplinärer, wissenschaftlicher Fachbereich, mit vielen Verbindungen in diverse andere Fachbereiche⁴: »Die KI beinhaltet momentan eine große Vielzahl von Unterbereichen, von allgemeinen Bereichen (wie beispielsweise Lernen und Wahrnehmen) bis hin zu speziellen Bereichen wie beispielsweise Schach spielen, das Beweisen mathematischer Theoreme, das Schreiben von Gedichten, das Führen eines Fahrzeugs auf einer belebten Straße oder die Krankheitsdiagnose. KI ist für jede intellektuelle Aufgabe relevant. Es handelt sich um ein wirklich universelles Gebiet.« [RUNO12_Einf]

Für Unbeteiligte mag das Thema der KI etwas ungreifbar erscheinen, hauptsächlich weil so viele verschiedenartige Anwendungsfälle hierdurch bearbeitet werden und nicht auf den ersten Blick klar wird, wie eine Technologie so vielfältig eingesetzt werden kann. Hierzu: es handelt sich nicht um eine Technologie, sondern um vielfältige Konzepte von denen manche schon lange existieren, wohingegen andere noch sehr jung sind. Dies zu wissen erleichtert zwar nicht unbedingt das Verständnis von KI, allerdings wird klar, dass für verschiedene Anwendungsfälle und Anforderungen sehr verschiedene Methoden eingesetzt werden und KI keineswegs ein einheitlicher Fachbereich ist. Ein tieferer und differenzierter Blick in die verschiedenen Themenbereiche ist zum Verständnis notwendig.

Die aktuellen Erfolge und somit auch die Popularität des Themas wurden erst durch das Vorhandensein ausreichender Rechenleistung und großer Datenmengen ermöglicht. Das weitere Ansteigen der Verfügbarkeit dieser beiden Ressourcen lässt erwarten, dass auch weiterhin Fortschritte und Innovationen aus dem Bereich der KI zu erwarten sind.

4.1.1 Maschinelles Lernen

Ein derzeit sehr populärer Unterbereich der KI, ist das maschinelle Lernen (ML). Im Allgemeinen spricht man von ML, wenn ein Programm »... seine Leistung für zukünftige Aufgaben verbessert, nachdem [es] Beobachtungen über die Welt gemacht hat.« [RUNO12_Lern] »Das heißt, es werden nicht einfach die Beispiele auswendig gelernt, sondern Muster und Gesetzmäßigkeiten in den Lerndaten erkannt.« [WIKIML21]

³ Seit Ende des zweiten Weltkriegs, der Name künstliche Intelligenz wurde 1956 geprägt. [RUNO12_Einf]

⁴ Berührte Fachgebiete sind u.a.: Philosophie (Logik, formale Regeln, Entstehen von Wissen, ...), Mathematik (Logik, gültige Schlüsse, Berechenbarkeit, ...), Wirtschafts- und Sozialwissenschaften (Renditenmaximierung, Zukunftsvorhersagen, ...),

Neurowissenschaften (Wie verarbeitet das Gehirn Informationen?), Psychologie (Wie denken und handeln Menschen und Tiere?), Technische Informatik (Aufbau effizienterer Computer), Regelungstheorie und Kybernetik, Linguistik... [RUNO12_Einf]

Ein weiterer Punkt, der ML für Laien oft ungreifbar macht, ist das Lernen an sich: also die sich verändernde Struktur eines Computerprogramms. Klassische Programme können zwar hochkomplex werden, aber sie verändern sich nicht mit den von ihnen verarbeiteten Daten, sie generieren neue Daten, oder ändern vorhandene. Die Struktur von lernenden Algorithmen (auch Modelle genannt), verändert sich ebenso wenig wie bei einem klassischen Computerprogramm. Wenn bei einem klassischen Computerprogramm im Allgemeinen die Regeln, wie Daten zu verarbeiten sind, zum Start definiert sein müssen, macht lernende Algorithmen aus, dass die Regeln zwar ebenso festgelegt sind, aber die Gewichtung der Regeln sich im Laufe der Arbeitsschritte an die verarbeiteten Daten anpassen. Sie passen sich also über die Iterationen an die Daten an, die sie als Eingaben erhalten, was klassische Computerprogramme nicht machen. Der Typ und die Form der Daten die von einem solchen Modell verarbeitet werden können, lassen sich aber nachträglich nicht ändern, die Struktur des Modells ist hierauf ebenso festgelegt wie bei klassischen Programmen. Es ist beispielsweise nicht möglich, mit einem Modell das für das Verarbeiten von allgemeinen Bilddaten erstellt wurde, andersförmige Daten wie sie zum Beispiel in der Sprachverarbeitung verwendet werden, oder zeitreihenbasierte Messwerte einer Wetterstation, zu verarbeiten.

Es existieren diverse Ansätze wie ML umgesetzt werden kann, hier die zwei grundlegendsten:

- **Nicht überwachtetes Lernen:** ein Modell findet in Daten selbstständig Muster und gruppiert die Einzelwerte entsprechend. Verfahren des nicht überwachten Lernens dienen meist der Exploration von Daten und sind im Kontext dieses Textes nicht weiter relevant.
- **Überwachtetes Lernen:** ein Modell lernt Muster anhand von bereits bewerteten Daten. D.h. es wird mit Daten trainiert, die zuvor meist von Menschen entsprechend der Zielvorgabe beschriftet (gelabelt) wurden.

Dies könnte beispielsweise so aussehen, dass die Trainingsdaten dergestalt sind, dass zu einer Vielzahl von Fotos von verschiedenen Objekten die Informationen, welches Objekt auf dem jeweiligen Foto dargestellt ist, enthalten sind. Diese Benennungen der Objekte nennt man Labels. Beim Trainieren des Modells versucht dieses dann, seine Ausgabe immer weiter so zu verändern, dass sie zu den Labels passt.

Die Vorhersage, welches Objekt in einem Bild enthalten ist, wird mit der Anzahl der Iterationen besser. Sind die Trainingsdaten vollkommen verwendet worden, wird die Lernphase beendet und das Modell eingefroren, es verändert sich dann nicht mehr. Die Hoffnung ist nun, dass die Treffergenauigkeit der Aussagen des Modells einen entsprechend hohen Grad erreicht, sodass es in der Lage ist, auch neue Bilder die nicht gelabelt sind, anhand der gelernten Muster in die richtige Kategorie einzusortieren.

4.1.2 Leistungsfähigkeit von Modellen

Das Vorhandensein immer größerer Datenmengen ermöglicht das Entstehen immer präziser und allgemeiner arbeitender Modelle. Denn der Umfang an Daten, mit denen Modelle trainiert werden, ist ebenso wichtig, wie die Wahl der passenden Modell-Architektur und die Konfiguration dieser. [BABR01]. Es sind also zwei Faktoren relevant für das erfolgreiche Erstellen eines guten Modells. Zum einen muss natürlich ein der Komplexität der Aufgabe entsprechendes Modell gefunden werden, was einer nicht zu unterschätzenden Forschungsleistung bedarf. Aber ohne umfangreiche und gut für diese Forschung

vorbereiteten Datasets können solche Modelle nicht trainiert werden. Das Erstellen solcher Datasets benötigt einen ebenso nicht zu unterschätzenden manuellen Aufwand.

Die wachsende Breite des Anwendungsspektrums von Modellen innerhalb einer Domäne (wie gut ein Modell generalisiert, also auf einen möglichst weiten Bereich innerhalb der Domäne anwendbar ist), kann als Leistungssteigerung so verstanden werden: für einen Algorithmus kann es einfach sein, Objekte innerhalb zweier Objektklassen zu unterscheiden. Beispielsweise zu erkennen, ob auf einem Bild entweder eine Instanz der Objektklasse Haus abgebildet ist oder eine Instanz der Objektklasse Baum.

Instanzen dieser beiden Klassen stellen sich sehr verschieden dar in Form, Textur und Farbe und sind somit von einem Modell recht gut zu unterscheiden. Soll ein Modell nun jedoch zwischen allen gängigen mitteleuropäischen Baumarten unterscheiden können, setzt dies voraus, dass es mit vielen Instanzen jeder einzelner dieser Arten trainiert worden ist. Umfangreiche Datasets sind also essentiell für das Erzeugen solcher spezifischer Anwendungsfälle. Ist das Ziel, ein Modell zu erstellen, das zwischen 1.000 Klassen vielfältigster Objekte unterscheiden kann, werden sehr große Datasets für das Trainieren dieses Modells benötigt. Eine solch große Datenbank ist beispielsweise das Projekt ImageNet, mit derzeit rund 14Mio Bildern.

Natürlich spielt auch die Beschaffenheit der Modelle eine wesentliche Rolle für ihre Leistungsfähigkeit. Je höher auflösend, also größer die Bilder sind, die ein Modell verarbeiten kann, umso mehr Details können herangezogen werden zum Unterscheiden einzelner Instanzen und somit zum Feststellen der entsprechenden Klasse.

Generell ist es so, dass nie eine 100%ige Klassifizierungsleistung erreicht wird. Man spricht in diesem Zusammenhang auch von der Präzision bzw. accuracy eines Modells. Je nach Anwendungsfall genügen accuracies von um die 95%. Es darf allerdings niemals ignoriert werden, dass ein Unterschied besteht zwischen einem eindeutigen, sicheren Ergebnis und einem zu 95% sicheren. Eine accuracy von 95% ist bereits sehr gut.

4.1.3 Tiefes maschinelles Lernen

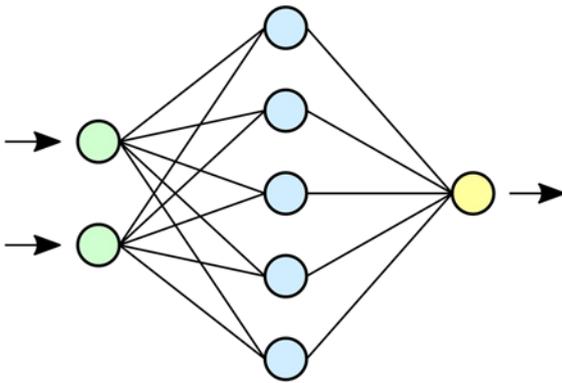
Innerhalb des maschinellen Lernens gibt es die Spezialisierung des tiefen maschinellen Lernens, das sich auf das Lernen mittels tiefer neuronaler Netze bezieht.

4.2 Künstliche Neuronale Netze

Die Idee, die zur Entstehung künstlicher neuronaler Netze führte, war, zu »... versuchen die Arbeitsweise des menschlichen Gehirns nachzubilden.«[LÄCL08]. Ein menschliches Gehirn besteht aus ca. 86 Mrd. Neuronen [SHH09], die miteinander zu einem gigantischen Netzwerk verbunden sind. Generell bestehen Netzwerke aus Knoten und Verbindungen zwischen den Knoten. Im Fall neuronaler Netze werden die Knoten Neuronen genannt.

Folgend wird ausschließlich auf die Funktionsweise künstlicher Neuronen und künstlicher neuronaler Netze (Artificial Neural Networks, ANN, auch kurz NN) eingegangen, also solchen wie sie in Form von Algorithmen in Computern umgesetzt werden. Da Rechner ausschließlich mit Zahlen arbeiten, müssen alle Daten, so sie mit einem NN verarbeitet werden sollen, in eine numerische Repräsentation überführt werden.

Lernen wird durch das Verstärken, oder Abschwächen der Weiterleitungsfähigkeit der Verbindungen zwischen Neuronen umgesetzt. Die Veränderung der Weiterleitungsfähigkeit führt dann dazu, dass nachfolgende Neuronen unterschiedlich stark aktiviert werden.



Quelle der Darstellung: [WIKIKNN21]

Abbildung 6: kleines neuronales Beispielnetz, bestehend aus acht Neuronen

Wie im Beispielnetz zu sehen, werden Neuronen in Schichten angeordnet. Daher kommt auch das Adjektiv »tief« in tiefem Lernen: Jede Schicht nach der ersten wird als versteckte Schicht bezeichnet (hier blau). Sobald ein Netz mehrere versteckte Schichten hat, spricht man von einem tiefen Netz und somit von tiefem Lernen. Das kleine Beispielnetz ist also noch kein tiefes Netz. Grundlegend können Netze sehr viele versteckte Schichten haben. Die erste (hier grüne) Schicht ist die Eingabeschicht, sie nimmt die zu verarbeitenden Daten lediglich an und verändert diese nicht. Die Zweite (blau) ist eine versteckte Schicht, die mit der an sie anschließenden Ausgabeschicht, welche hier nur aus einem einzigen Neuron (gelb) besteht, verbunden ist.

Hier ein etwas größeres Beispielnetz. Es ist deutlich zu sehen, wie die Anzahl an Verbindungen mit der Anzahl an Neuronen exponentiell zunimmt. Hieran wird auch sichtbar, weshalb NN so rechenintensiv sind. Netze mit mehreren Schichten aus vielen hundert Neuronen sind keine Seltenheit.

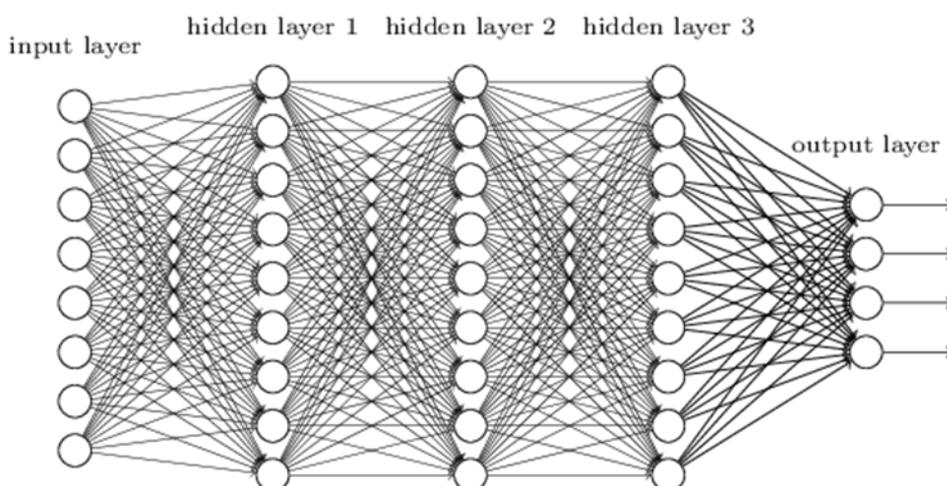


Abbildung 7: Beispielnetz mit einer Eingabeschicht, drei versteckten Schichten und einer Ausgabeschicht Quelle der Darstellung: [NI15]

4.2.1 Neuronen

Ein Neuron besteht hauptsächlich aus den folgenden Elementen:

- **Eingänge:** Ein Neuron kann beliebig viele Eingänge haben. Über die Eingänge werden Daten in numerischer Form an das Neuron übermittelt. Bis auf Neuronen der Eingabeschicht sind die Eingänge mit allen Ausgängen der Neuronen voriger Schichten verbunden.
- **Gewichtungen:** Sehr wichtig ist, dass ein Neuron selbst keinerlei Lernfähigkeit hat. Es sind die Verbindungen zwischen den Neuronen, die mit Gewichtungen gedämpft oder verstärkt werden. Wird beim Lernen beispielsweise festgestellt, dass eine Verbindung zu intensiv ist und häufig zu starke Werte an das folgende Neuron liefert, so wird dieses Gewicht reduziert. Andersherum findet ein Stärken von Verbindungen statt, wenn beim Lernen festgestellt wird, dass diese zu schwach sind.

Die Verbindungen zwischen den Ein- und Ausgängen der Neuronen sind also gewichtet und es sind auch diese Gewichte, die die eigentliche Lerneigenschaft umsetzen.

- **Aktivierungsfunktion:** Das Neuron muss irgendwie von den Eingabedaten zu seinen Ausgabedaten kommen. Meist werden die Eingangsdaten mit den jeweiligen Gewichten multipliziert und die Ergebnisse mit einer Übertragungsfunktion summiert. Dieses Resultat wird mit einer Aktivierungsfunktion verrechnet und über den Ausgang an die Neuronen der nächsten Schicht als Aktivierung dieser übermittelt.
- **Ausgänge:** Das Ergebnis der Berechnung wird über die Ausgänge ausgegeben. Ein Neuron kann beliebig viele Ausgänge haben und bis auf die letzte Schicht – der Ausgabeschicht – sind die Ausgänge mit den Eingängen der Neuronen der folgenden Schicht verbunden.

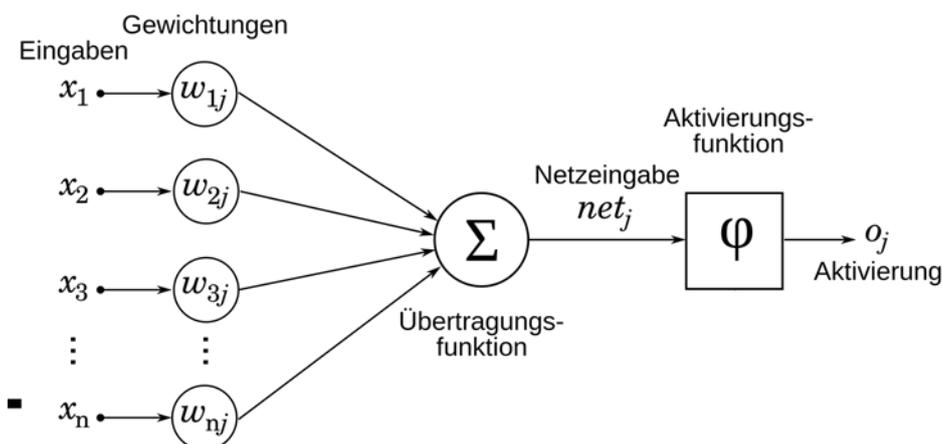


Abbildung 8: Aufbau eines Neurons

Quelle der Darstellung: [WIKIKNN21]

Ein einzelnes Neuron j mit mehreren Eingaben (x_1 - x_n) und deren Gewichten (w_{1j} - w_{nj}). Die Eingaben und entsprechende gewichte werden multipliziert und anschließend in der Übertragungsfunktion summiert. Das Ergebnis ist die Netzeingabe des Neurons j . Über die Aktivierungsfunktion wird die Ausgabe (o_j) errechnet und als Aktivierung an Neuronen folgender Schichten weitergegeben.

4.2.2 Lernen

Bisher wurde der Aufbau und die Komponenten eines NN skizziert. Nun fehlt noch zu klären, wie das Lernen eigentlich funktioniert, also wie mit den Trainingsdaten die Gewichte des Netzes so angepasst werden, dass die Ausgabe des Netzes möglichst deckungsgleich mit den Labels ist. Hierzu muss festgestellt werden, ob es eine Differenz zwischen Ausgabe des Netzes und dem Label gibt und wenn ja, wie groß diese ist.

Um das Beispiel mit den gelabelten Bilddaten zu konkretisieren: ein Netz soll trainiert werden, um zu erkennen, ob sich auf einem Bild ein Objekt der vier Objektklassen Äpfel, Fahrräder, Häuser oder Kühe befindet. Da es nur numerische Werte ausgeben kann, erhält jede Klasse eine numerische Entsprechung, also bspw. Äpfel: 0, Fahrräder: 1, Häuser: 2, Kühe: 3. Digitale Bilddaten sind nichts anderes als Zahlen, die üblicherweise die Farbintensität der drei Farbkanäle Rot, Grün und Blau eines jeden Pixels wiedergeben. Diese Pixelwerte können problemlos einem NN über dessen Eingänge zur Verarbeitung eingegeben werden. Sobald also Muster in diesen Zahlenwerten vorkommen, die auf das Vorhandensein eines Hauses schließen lassen, soll die Ausgabe des Netzes eine 2 ausgeben.

Zu Beginn des Trainingsprozesses, wenn die Gewichte des NN noch undefiniert sind, kommt an der Ausgabeschicht ein Zufallswert heraus. Ziemlich sicher entspricht die erste Ausgabe somit nicht der gesuchten 2 für ein Haus, das auf dem Bild war, welches wir dem Netz als erste Eingabe gegeben hatten. Die Differenz des Ausgabewerts (wir nehmen hier einfach eine zufällige 1 an) zum Zielwert 2, ist der Ausgabefehler. Hier also wäre der Ausgabefehler $1 (2-1=1)$.

Blickt man nun auf die letzte Schicht vor der Ausgabeschicht, lässt sich erkennen, welche der Ausgabewerte stärker zum Fehler beigetragen haben und welche Neuronen eher einen geringen Anteil am Fehler hatten. Entsprechend werden nun die Gewichte angepasst: die Verbindungen mit geringem Fehleranteil werden nur mild, oder gar nicht verändert, während Verbindungen mit hohem Fehleranteil eine stärkere Gewichtsangpassung erfahren. Sobald die letzte Schicht vor der Ausgabeschicht berechnet und aktualisiert wurde, kann mit der Schicht davor ebenso verfahren werden. Bis schließlich das komplette Netz rückwärts berechnet und die fehlerintensiven Gewichte korrigiert wurden. Diese Gewichtungsadaption wird nun bei jedem Bild auf dem ganzen Netz durchgeführt, sodass die Gewichte immer geringere Fehler produzieren sollen.

Dieses Verfahren nennt sich Backpropagating und ist die Schlüsselidee für den Erfolg künstlicher neuronaler Netze. Sie wurde 1986 von David E. Rumelhart et al. mit dem Paper [RU86] publiziert. Wichtig für Backpropagating ist eine weitere Methode, das Gradientenverfahren, oder auch Gradientenabstieg, ein Minimierungsverfahren welches es ermöglicht die Gewichte so anzupassen, dass der Ausgabefehler minimiert wird. Die genauen Funktionsweisen dieser beider Methoden können im Umfang dieses Textes nicht durchleuchtet werden, aber aufgrund ihrer Wichtigkeit für das tiefe Lernen sind sie omnipräsent und in jeder Literatur oder Seite zu finden die sich mit tiefen Lernen beschäftigen.

4.3 Convolutional Neural Networks

Die Ergebnisse der Patentrecherche lassen erkennen, dass für das Klassifizieren von Bilddaten für die Identifikation von Bovinen hauptsächlich eine Spezialform neuronaler Netze, die Convolutional Neural Networks (CNN) angewandt werden. Dies ist nicht weiter verwunderlich, haben sich CNNs als sehr leistungsfähig beim Klassifizieren von Bilddaten erwiesen und sind sehr weit verbreitet.

Auch für CNNs war wieder die Biologie Ideengeber: sie sind dem visuellen Cortex von Säugetieren nachempfunden. [GÉ19] Hier ist es so, dass es Wahrnehmungsfelder gibt, die bestimmte Muster erkennen. Es sind also immer Gruppen von Neuronen mit der nächsten Schicht verbunden. Im Gegensatz zu NN wie sie bisher in diesem Text beschrieben wurden, sind also nicht alle Neuronen mit allen Neuronen der nächsten Schicht verbunden sondern nur bestimmte Felder einer Schicht mit Neuronen der Folgeschicht.

Diese Felder (auch Kernel) sind flächige Teilbereiche einer Schicht, die mit nur wenigen Neuronen der nächsten Schicht verbunden sind, wodurch eine Lokalität entsteht.

Das Konzept der convolutional neural networks erreichte ab 1998 seinen Durchbruch mit der Veröffentlichung dieses mittlerweile berühmten Papers: [LECU98]. Ziel war damals das automatische Erkennen von handgeschriebenen Ziffern auf Bankchecks. Der französische Informatiker Yann LeCun erstellte im Zuge seiner Arbeit auch das MNIST-Datenset, ein Dataset mit gelabelten, handgeschriebenen Ziffern der Zahlen 0-9. Ein neuronales Netz mit MNIST zu trainieren ist das »Hello World« eines jeden Interessierten, beim Lernen tiefen maschinellen Lernens. [MNIST]

Konventionelle, also voll-verbundene NN, arbeiten zwar auch sehr gut mit Bildern, allerdings steigt ihre Schichten-, aber vor allem Verbindungsanzahl extrem mit Größe und Komplexität der zu verarbeitenden Bilder, sodass man hier allein schon durch die reale Berechenbarkeit schnell an die Grenzen des Machbaren stößt. Da CNNs sehr viel weniger Verbindungen benötigen als voll-verbundene NNs, sind sie deutlich sparsamer zu berechnen, bzw. erlauben es, auf der selben Hardware viel leistungsfähigere Netze zu berechnen.

Mit folgenden drei Konzepten lassen sich CNNs verstehen: convolutional Layers und den dadurch ermöglichten Filtern und dem Pooling.

4.3.1 Convolutional Layer

Die bereits erwähnten Wahrnehmungsfelder sind essentiell für CNNs. Zum besseren Verständnis zunächst eine Abbildung. Diese wurde dem freien Online-Buch neuralnetworksanddeeplearning.com entnommen, ein guter Einsteigerkurs zum Thema tiefes Lernen mit neuronalen Netzen.

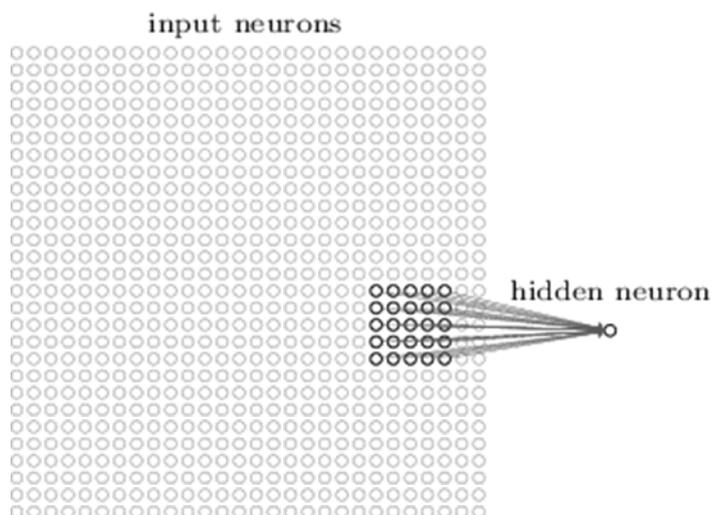


Abbildung 9: Beispiel Wahrnehmungsfelder

Quelle der Darstellung: [NI15]

In diesem einfachen Beispiel sind $5 \times 5 = 25$ Neuronen der Eingabeschicht zu einem Kernel (=Wahrnehmungsfeld) gruppiert. Die Ausgänge dieser bilden den Eingang eines Neurons in der nächsten, der ersten versteckten Schicht.

Verglichen dazu: in einem voll-verbundenen NN wäre jedes Neuron der Eingabeschicht mit jedem Neuron der ersten versteckten Schicht verbunden.

Ein kleines Rechenbeispiel zu diesem kleinen Netz:

Tabelle: Rechenbeispiel

Voll-verbundenes neuronales Netz	CNN
Kantenlängen der Eingabeschicht: X = 28; Y = 28 Neuronenanzahl Eingabeschicht: $28 * 28 = 784$ Anzahl Verbindungen von der Eingabeschicht zur ersten versteckten Schicht bei identischer Größe beider Schichten: $784 * 784 = 614.656$	Kantenlängen der Eingabeschicht: X = 28; Y = 28 Neuronenanzahl Eingabeschicht: $28 * 28 = 784$ Anzahl Verbindungen von der Eingabeschicht zur ersten versteckten Schicht bei identischer Größe beider Schichten: $784 * 25 = 19.600$

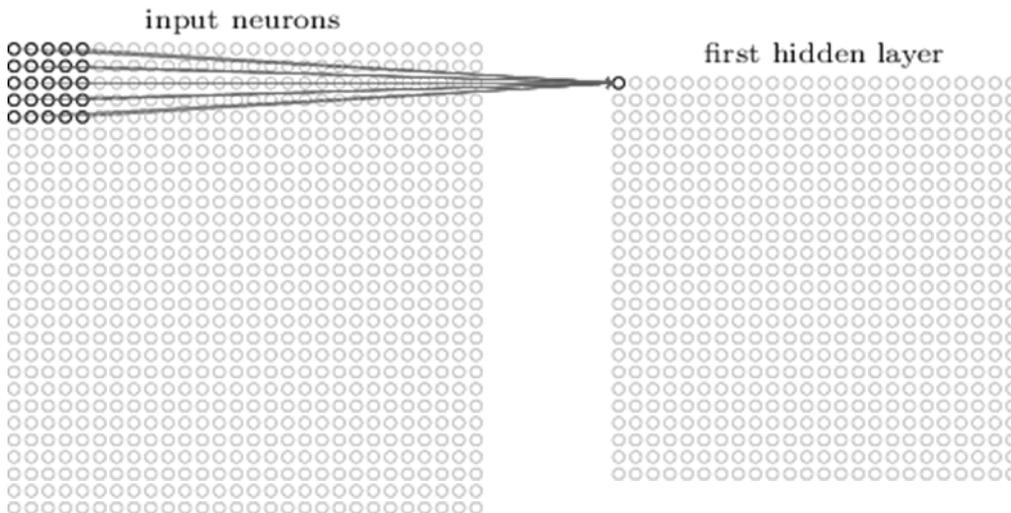
Ein Bild mit einer Kantenlänge von 28×28 Pixel ist nun nicht wirklich groß.

Zum Vergleich: würde man ein Schwarzweißbild einer 1 Megapixel-Kamera verwenden, hätte es, wenn es quadratisch wäre, eine Kantenlänge von 1.000 Pixel. Es gäbe also 1×10^{12} (1.000.000.000.000, eine Trillion) Verbindungen von der Eingabeschicht zur ersten versteckten Schicht bei einem voll verbundenen Netz. Da ein tiefes Netz viele Schichten benötigt um komplexere Muster verarbeiten zu können, lässt sich schnell erkennen, dass dies kein praktikabler Ansatz ist. Aber um fair zu bleiben muss allerdings gesagt werden, dass eine Kantenlänge von 1000 Pixel auf derzeit gängiger Hardware auch mit CNNs absolut nicht berechenbar ist. Die Grenzen des in angemessener Zeit Berechenbaren wird durch sie nicht aufgehoben, nur verschoben.

Die Größe der Bilder die aktuelle Netze verarbeiten liegt bei ImageNetModellen beispielsweise bei 224×224 Pixeln.

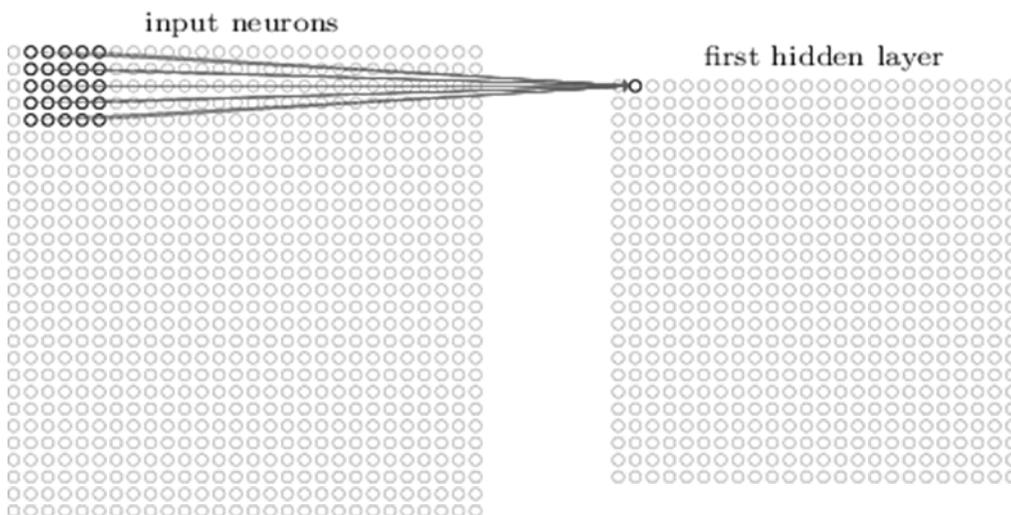
Convolutional kernel sind oft kleiner als die 25 aus dem obigen Beispiel, meist $3 \times 3 = 9$, oder 16 Neuronen groß. Das Patent CN111291683 beschreibt den Aufbau des Modells unter Abschnitt [0019] folgend: »... the size of the convolution kernel of the convolution layer is 3×3 , ... the size of the normalized image is 256×256 ...«. Dieser Aufbau kann als repräsentativ für leistungsfähige, aktuelle Modelle angesehen werden.

Durch das Verschieben des Kernels um immer ein Neuron entstehen so die Verbindungen zur Folgeschicht:



Quelle der Darstellung: [NI15]

Abbildung 10: Verbindungen des ersten Schrittes: ein 25-Neuronen großer Kernel wird auf ein Neuron der nächsten Schicht abgebildet.



Quelle der Darstellung: [NI15]

Abbildung 11: Verbindungen des zweiten Schrittes: ein 25-Neuronen großer Kernel wird auf das zweite Neuron der nächsten Schicht abgebildet. Die zweite Schicht entsteht so durch die vollständige Iteration über alle Kernel

Die zweite Schicht ist somit etwas kleiner als die Eingabeschicht. Ein weiterer Aspekt ist zudem die Lokalität, die dadurch entsteht, dass nur ein Teil der Neuronen mit einem Neuron der Folgeschicht verbunden sind: Neuronen der Folgeschicht sehen nicht das komplette Bild, sondern nur einen kleinen Ausschnitt.

4.3.2 Filter

Diese Lokalität ermöglicht eine faszinierende Eigenschaft: Filter. Filter funktionieren wie in der herkömmlichen Bildbearbeitung auch: je nach Beschaffenheit werden bestimmte Eigenschaften des Bildes stärker zum Vorschein gebracht und andere gemindert. Zwei gängige Filter bei CNNs sind beispielsweise Horizontal- und Vertikalfilter. Diese sind so beschaffen, dass manche Gewichte des convolutional kernels 0 sind und manche 1. Zur Erinnerung: die Verbindungen der Neuronen des Kernels werden mit den Gewichten multipliziert bevor sie dem nächsten Neuron als Eingabe dienen.

Horizontalfilter

Wird angenommen der Kernel ist 3*3 Neuronen groß, so lässt sich mit den Gewichten zum Beispiel dieses Muster abbilden:

0	0	0
1	1	1
0	0	0

Abbildung 12: Beispielmuster Horizontalfilter

Dieses Gewichtemuster sorgt nun dafür, dass lediglich die drei Neuronen der Eingangsschicht, die auf der Ebene mit den Einsen liegen, ihre Werte voll an das Neuron der Folgeschicht weitergeben können. Die andern drei darüber und die drei darunter werden mit 0 multipliziert, was eine Weitergabe der Werte komplett unterbindet. Da der Kernel jedoch immer nur ein Neuron weiterbewegt wird, und in der nächsten Reihe tiefer ebenso verfahren wird, entsteht ein Horizontalfilter. Das heißt, dass alle horizontale Linien auf dem Bild stärker zum Vorschein treten, als alle schrägen, oder vertikalen.

Vertikalfilter

Mit ebenjenem Verfahren lässt sich sehr einfach auch ein solcher Filter erzeugen:

0	1	0
0	1	0
0	1	0

Abbildung 13: Beispielmuster Vertikalfilter

Dieser Vertikalfilter begünstigt alle vertikalen Linien im Bild. Da diese Erklärung aufgrund ihrer Knappheit etwas abstrakt wirkt, hier für jeden Filter ein Beispiel:

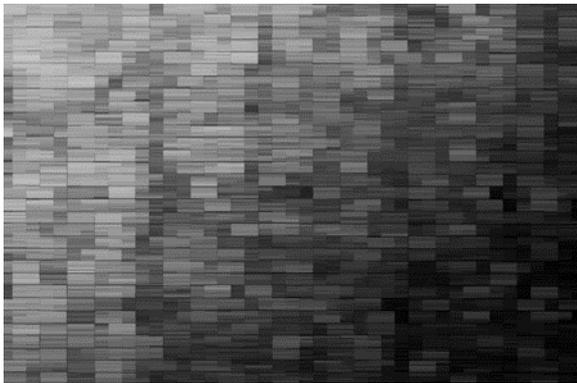


Abbildung 14: Nach Anwendung eines Horizontalfilters

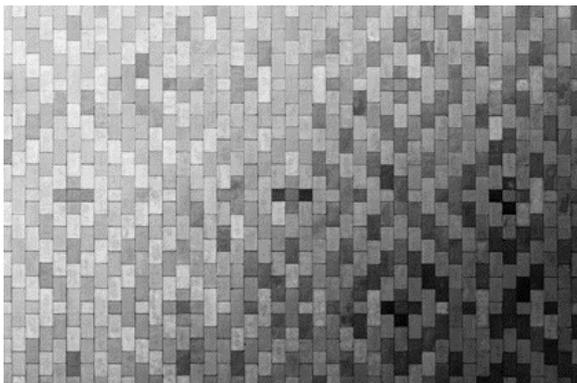


Abbildung 15: Originalbild

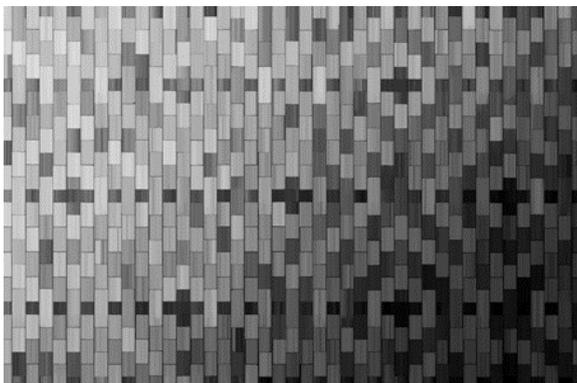


Abbildung 16: Nach Anwendung eines Vertikalfilters

Was Filter wirklich faszinierend macht, ist die Tatsache, dass sie von alleine durch das Lernen entstehen. [GÉ19] (S.450) Horizontal- und Vertikalfilter sind sehr einfache, anschauliche Beispiele, tatsächlich jedoch existieren vielfältige, diverse Muster, die sich in convolutional kernels bilden können, während ein Modell trainiert wird.

Zur Erinnerung: die Gewichte bilden den Lerneffekt ab, da sie entsprechend der Fehlerintensität bei jedem einzelnen der Bilder, die durch das Netz »geschoben« werden unterschiedlich stark angepasst werden. Somit lässt sich sehr plastisch vorstellen, wie sich das Netz immer feiner an die Gesamtheit der Bilder anpasst. Dadurch wird auch klar, weshalb Modelle große Datenmengen benötigen, um gut ausformuliert zu werden, wie die Datenmenge zum Trainieren eigentlich nie groß genug sein kann.

Analog zum biologischen, visuellen Kortex bilden sich also auch hier Muster mit denen das Klassifizieren von Objekten ermöglicht wird.

4.3.3 Pooling Layer

Ein weiterer Mechanismus, der CNNs von voll-verbundenen NN unterscheidet ist Pooling. Auch hier kommen Kernel zum Einsatz.

Pooling wird verwendet, um Stabilität gegenüber Variation zu erhalten, also um sich auf das Wesentliche zu konzentrieren und unempfindlich gegenüber Eigenschaften in einem Bild zu werden, die nicht zur Klassifikation beitragen. Hier wird der Tatsache Rechnung getragen, dass sich intensive Ausgabewerte von Neuronen stärker auf das Ergebnis des Netzes auswirken als niedrige Werte¹⁴. Dies kann so verstanden werden, dass immer die hohen, intensiven Ausgaben von Neuronen am ehesten zur Erkennung beitragen.

Ein weiterer positiver Nebeneffekt ist das erneute Reduzieren der Schichtengröße.

Im Gegensatz zu convolutional kernel, passen sich Pooling-Kernel nicht an die Daten an, sie lernen nicht und haben keine Gewichte sondern ermitteln den größten Wert innerhalb ihres Wahrnehmungsfeldes und leiten ausschließlich diesen weiter.

Hier kann sogar so vorgegangen werden, dass die Schrittweite der Kantenlänge des Kernels entspricht, was bei einem $2 \times 2 = 4$ großen Kernel zu einem Reduzieren der nächsten Schichtgröße um 75% führt. Es wird also jedes Areal des Wahrnehmungsfeldes nur einmal abgefragt und somit schafft es nur der größte Wert, die der Pooling-Kernel betrachtet, in die nächste Schicht:

1	3	5	1	3	1
0	2	2	2	4	7
0	5	9	5	1	0
0	1	9	3	1	6
0	7	0	5	2	0
1	7	8	1	4	5

Schicht 2

3	5	7
5	9	6
7	8	5

Schicht 1

Abbildung 17: Max-Pooling mit einem Pooling-Kernel der Größe $2 \times 2 = 4$ und einer Schrittweite die der Kantenlänge des Kernels (also 2) entspricht. Anhand der Färbung kann erkannt werden, welche Areale von Schicht 1 der Kernel bei einem Schritt sieht. Er ermittelt den größten Wert dieses Areals und gibt ihn an das Neuron der nächsten Schicht weiter. Das Resultat dieses Prozesses ist dann Schicht 2.

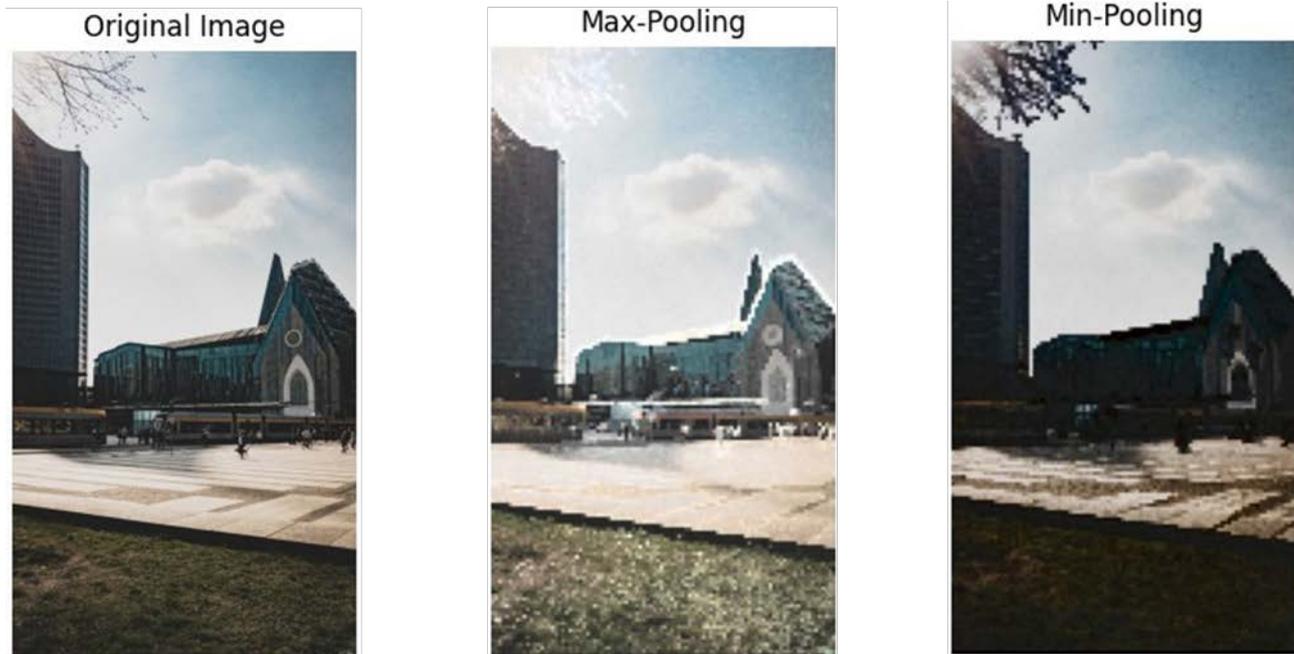


Abbildung 18: Auswirkung der Methoden Max-Pooling und Min-Pooling. Neben Max-Pooling können auch andere Pooling-Verfahren Anwendung finden. Max-Pooling ist aber das am häufigsten verwendete.

Es muss allerdings erwähnt werden, dass Pooling zwar im Kontext von CNNs sehr relevant ist. Es ist aber auch ein destruktives Verfahren wodurch Daten verloren gehen. Bei der Identifikation von Tieren ist es wichtig, möglichst wenige Informationen zu verlieren. Insofern wird davon ausgegangen, dass Pooling eher weniger stark in CNNs Anwendung findet, die zur Identifikation erstellt wurden.

4.3.4 Feature Maps und Feature-Vektoren

Eine Schicht die durch convolutional kernel entstanden ist, wird auch feature map genannt. Generell ist es so, dass in CNNs mehrere convolutional kernel pro Schicht angewandt werden, sodass mehrere feature maps entstehen. Näher darauf einzugehen ist hier nicht notwendig, da nur ein grobes Grundverständnis von CNNs vermittelt werden soll und dem Leser durch diesen Text nicht ermöglicht wird, eigene Modelle zu erstellen.

Es lässt sich jedoch zusammenfassen, dass mit Filtern Muster erkannt werden und dadurch, dass pro Schicht viele convolutional kernel angewandt werden können, auch ebenso viele Muster vom Netz registriert werden können. Die Ergebnisschicht einer dieser Erkennungen wird dann feature map genannt.

Zusammen mit der Reduktion der Größe in tieferen Schichten, enthalten feature maps immer komplexere Muster, bis sie schließlich gegen Ende des Netzes vollständige Objekte erkennen können. Um in den letzten Schichten des Netzes nun von diesen umfangreichen Mustern zu der Entscheidung zu kommen, zu welcher Klasse das Objekt auf dem Bild¹⁵ gehört, gibt es am Ende mindestens eine vollverbundene Schicht. Diese ist eindimensional und somit ein Vektor und wird Feature-Vektor genannt.

Angenommen der Feature-Vektor wäre 1.000 Neuronen groß, so würde das Netz 1.000 Zahlen am Ende pro Bild ausgeben, in welchen die Klassifikationsleistung des Netzes enthalten ist.

Ein Tier würde also durch 1.000 Zahlen repräsentiert. Je ähnlicher die Tiere auf unterschiedlichen Bildern sich sind, umso ähnlicher sollten sich dann auch die Feature-Vektoren sein. 100%ige identisch werden sich die Feature-Vektoren nie sein, allein schon deshalb, weil verschiedene Bilder ein und desselben Tieres mindestens leicht unterschiedlich sind. Allerdings sollten die Unterschiede groß genug sein um Tiere voneinander unterscheiden zu können.

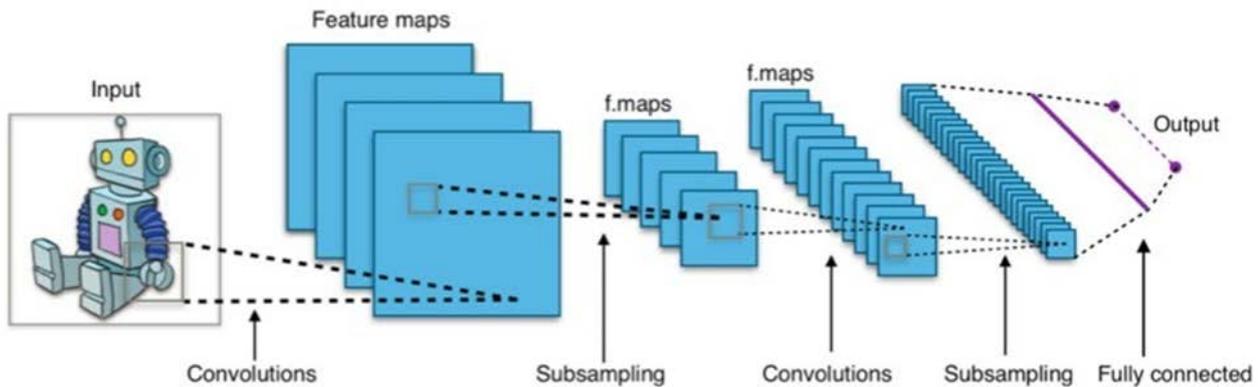


Abbildung 19: Vollständiges CNN. Subsampling ist ein Synonym für Pooling.

Quelle der Darstellung: [WIKICNN21]

4.3.5 Modelle trainieren

Beim maschinellen Lernen spricht man von Parametern und Hyperparametern. Parameter sind solche Variablen, die durch den Lernprozess angepasst werden, also zum Beispiel die Gewichte eines NN.

Hyperparameter sind alle jene Variablen, die das Lernverhalten beeinflussen. Also im Kontext von NN beispielsweise die Anzahl und Größe der Schichten oder Art der Aktivierungsfunktion. Im Kontext von CNNs gibt es zusätzlich noch: die Kernelgröße, die Anzahl der Filter, die Schrittweite, die Entscheidung, wann eine Pooling-Schicht eingefügt wird usw..

Das Herausfinden, welche Architektur¹⁶ gute oder gar die besten Ergebnisse für eine Aufgabe liefert, ist keine triviale Aufgabe. Trainingsdurchläufe sind zeitaufwändig und es ist nicht so, dass klar ist, warum ein Netz reagiert, wie es reagiert. NN werden deswegen zu den Black Box Verfahren des maschinellen Lernens gezählt. Sie sind sehr leistungsfähig, aber es ist nicht einfach nachzuvollziehen wie sie zu ihren Entscheidungen kommen. Ein relativ junger Forschungszweig beschäftigt sich mit dem Thema der Darstellbarkeit der Interna von NN: Feature Visualization. Hier ein sehr umfangreicher Artikel zum Thema: [FEAVIS17]

Die Darstellung dessen, was ein NN sieht, bzw. wie es sieht und entscheidet ist nicht eindeutig und recht zeitaufwändig. Möchte man ein NN trainieren, ist Geduld und Gefühl für die Daten von Nöten.

Das Entwerfen einer passenden Architektur und das Trainieren eines neuronalen Netzes ist kein deterministischer, oder geradlinig planbarer Prozess, vielmehr eine Forschungsaufgabe mit viel Ausprobieren über viele Iterationen hinweg.

Der Prozess, eine gute Architektur und somit auch die richtigen Hyperparameter zu finden, wird auch hyperparameter optimization, bzw. Hyperparameteroptimierung genannt.

4.3.6 ImageNet Large Scale Visual Recognition Challenge

Die Netz-Architekturen werden immer komplexer und ausgefeilter. Was den Wettbewerb um die beste image-classification angeht ist die »ImageNet Large Scale Visual Recognition Challenge« [ILSVRC] die Königsklasse der Wettbewerbe. Hier messen sich die größten Netze, die größten Namen und Konzerne und es entstehen große Innovationen. Die Gewinner-Architekturen versprechen immer größere Leistungsfähigkeit und finden dadurch schnell Zugang zur praktischen Anwendung als Werkzeug für vielfältigste Forschungsbereiche. Einen sehr guten Überblick bietet [INWIN].

2012 läutete Alex Krizhevsky mit dem Gewinnen der ImageNet Competition mit seinem CNN [KRIZ12] die CNN-Ära in der image-classification ein. Große Namen wie Google und Microsoft ließen mit eigenen Modellarchitekturen nicht lange auf sich warten.

4.3.7 Transfer learning

Oft ist es so, dass für eine Aufgabe nur begrenzte Datensätze zur Verfügung stehen. Um ein CNN vollständig zu trainieren werden jedoch sehr umfangreiche Datensätze benötigt. Zu kleine Datensätze führen zu keinem guten Ergebnis. Da die unteren Schichten nahe an der Eingangsschicht generell bei CNNs nur sehr allgemeine Strukturen und Muster erkennen, lassen sich diese sehr gut wiederverwenden, sodass deren Gewichte eingefroren werden können. Dieser Mechanismus wird transfer learning bezeichnet. Es werden nur die oberen Schichten neu trainiert, die bei komplexeren Mustern bis hin zu vollständigen Objektklassen aktiviert werden. »The more training data you have, the more layers you can unfreeze.«[GÉ19] Generell wird transfer learning häufig in Projekten eingesetzt, die nicht ausreichend Daten haben, oder aber deren Klassifikationsziel nicht allzu weit von einem bereits verfügbarem Modell entfernt ist.

4.4 Daten

In allen Bereichen des maschinellen Lernens ist der Fortschritt hungrig nach Daten. Es existieren vielfältige Datasets zu allen möglichen Themen:

[WIKIDAS21]. Neben dem bereits mehrfach erwähnten ImageNet ist das Open Image Dataset von Google [OPENIMA21] sehr gut geeignet um zu veranschaulichen, was ein großes Dataset ausmacht.

Datasets bestehen neben den Bildern natürlich auch aus Labels anhand derer das trainierte Modell lernt. Hierfür müssen die Bilder von Menschen manuell gelabelt, oder zumindest qualifiziert werden. Oft kommen hier mittlerweile auch schon Modelle zum Einsatz, die Labels erstellen. Diese Labels müssen dann noch vom Menschen geprüft werden da nicht ausgeschlossen werden kann, dass Fehler enthalten sind. Fehler in Trainingsdaten sind fatal, da sie Fehlerhafte Modelle zu Folge hätten. Bei mehreren Millionen Bildern und mehreren möglichen Objekten pro Bild ist der Aufwand gute Labels zu erstellen nicht zu unterschätzen. Näheres zum Erstellen eines Datasets unter [KUZOPIM20].

Labels sind in separaten Dateien enthalten und werden über eine Referenz mit der Bilddatei assoziiert. Leider unterscheiden sich die Projekte in der Art, wie die Labels zur Verfügung gestellt werden. Grundsätzlich kann aber unterschieden werden in folgende Kategorien von Labels:

- Image-level labels: Pro Bild wird angegeben, welche Objekte sich auf der Abbildung befinden, aber nicht wo auf dem Bild.
- Object bounding boxes: eine bounding box wird so gezeichnet, dass sie möglichst das vollständige Objekt beinhaltet und so wenig wie möglich Fläche, die nicht zum Objekt gehört:

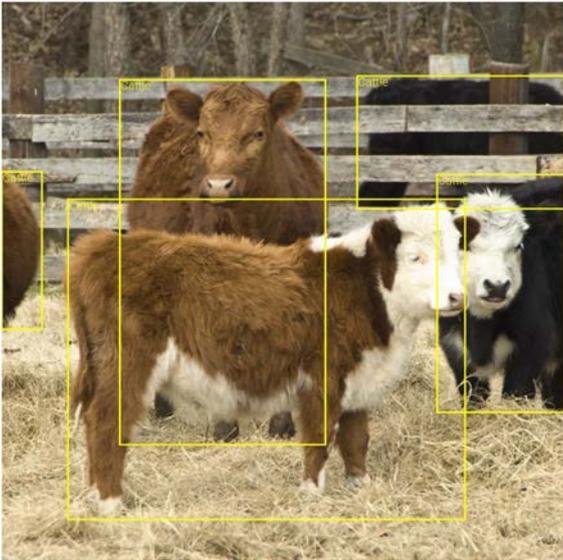


Abbildung 20: Beispiel bounding box

Quelle der Darstellung: [OPENIMA21]

- Object segmentation masks: die Fläche die ein Objekt ausfüllt, wird durch das Zeichnen einer Umrandung gekennzeichnet



Abbildung 21: Beispiel Object segmentation masks

Quelle der Darstellung: [OPENIMA21]

Universelle Datasets wie Open Image o.ä., enthalten mehrere hundert Kategorien an Objekten auf mehreren Millionen Bildern und sind somit gut um ein Netz grundlegend zu trainieren. Ein Projekt das aber Kühe identifizieren möchte benötigt zusätzlich sehr viele gelabelte Daten und zwar in der Form, dass nicht nur Kühe erkannt werden, sondern auch unterschieden (identifiziert) werden können. Das bedeutet, dass von jedem einzelnen Tier viele Aufnahmen vorhanden sein müssen, die gelabelt und entsprechend mit diesem einen Tier assoziiert sind.

4.5 Erkennung

Verschiedene Termini müssen unterschieden werden, im Kontext der Identifikation von individuellen Objekten und Tieren. Um nicht noch weitere, zusätzliche Begriffe und Interpretationen zu verwenden werden hier anstatt der deutschen Begriffe die englischen Originalbegriffe verwendet.

4.5.1 Object recognition

Object recognition ist der generelle Überbegriff für image classification und object localisation.

Image classification bezeichnet den Prozess festzustellen, ob in einem Bild eine Instanz einer bestimmten Objektklasse enthalten ist.

Object localization findet die Stelle im Bild, wo das Objekt enthalten ist. Diese wird mit einer bounding box gekennzeichnet. Eine erfolgreiche image classification ist hierfür Voraussetzung.

4.5.2 Object detection

Object detection beinhaltet Image classification und object localization für multiple Objekte in einem Bild: das Klassifizieren und die Lokalisation mehrerer Objekte auf einem Bild.

Diese Aufgabe wird mittlerweile sehr gut von NN erledigt. Die Anzahl der erkennbaren Objektklassen nimmt stetig ebenso zu, wie die Fehlerrate der Klassifikation abnimmt. Große Datasets sind darauf ausgelegt um das Trainieren von Modelle genau hierfür zu ermöglichen.

Ein Objekt in einem Bild zu finden und zu klassifizieren funktioniert sehr gut mit aktuellen Modellen und wird auch weithin kommerziell eingesetzt und lässt sich einfach in bestehende Softwareprodukte integrieren.

4.5.3 Face detection

Analog zur object detection bezeichnet face detection das Erkennen des Vorhandenseins und Lokalisieren von Gesichtern auf Bildern.

4.5.4 Face recognition

Es scheint gemeinhin weniger Eindeutigkeit zu bestehen, was den Begriff face recognition angeht, er wird oft synonym verwendet zu face detection, was unpräzise aber dennoch weit verbreitet ist. Je nach Kontext lohnt also der genauere Blick, was der jeweilige Autor tatsächlich bezeichnet.

Im Kontext der Identitätsfeststellung eines Individuums wird face recognition gleichgesetzt mit Identifizieren und Verifizieren von Gesichtern. Also ist hier der Abgleich, ob ein gefundenes Gesicht auf einem Bild zu einer bekannten Person gehört, mit enthalten. Dieser Abgleich findet in der Regel mit Datenbankeinträgen statt, sodass diese Art von System ein 1:N-Problem löst: eine Identität wird aus vielen möglichen ermittelt und bestätigt.

4.5.5 Facial Keypoint Detection

Facial keypoint detection beschreibt eine Methode, bei der markante Punkte, wie Augen, Nasenansatz, Mund, etc. gefunden und deren Abstände zueinander als eindeutige Merkmale zu einem face-print zusammengefasst werden. Diese lassen sich vergleichen und ermöglichen somit die Aussage, ob die Abbildungen von Personen, oder Tieren auf unterschiedlichen Bildern zur selben Person gehören, oder nicht. Im Humanbereich wird eine Vielzahl dieser facial keypoints verwendet. Apple benutzt laut eigenen Angaben über 30.000 solcher Punkte um eine 3D Abbildung des Gesichts anzufertigen. Im Gegensatz dazu wurden in den meisten Patenten aus dem Agrarsektor unter 10 verwendet um Kühe zu identifizieren. Es wird vermutet, dass dies zu wenige sind um mit einer größeren Menge zuverlässig zu arbeiten.

4.5.6 Holistischer Ansatz

Beim holistischen Ansatz werden anstatt keypoints zu definieren und auf Bildern zu finden, CNNs »einfach« mit einer Vielzahl an Bildern mit ausschließlich Gesichtern trainiert. Hierzu können klassische image classifier Architekturen verwendet werden. Anstatt nun am Ende des Netzes eine Klassifikation durchzuführen, wird die softmax-Funktion, die normalerweise aus dem letzten Feature-Vektor die konkrete Klasse ableitet, entfernt und die Werte des Feature-Vektoren als face print verwendet. Anschließend kommt eine weitere Trainingsmethode (Triplet-Loss) zum Einsatz, um die Gewichte für die Identifikation zu verfeinern. Aktuelle Feature-Vektor-Größen für menschliche Gesichter liegt bei ~2048. Diese können nun in einer Datenbank gespeichert werden. Feature-Vektoren von neu eingelesenen Bildern können nun mit denen in der Datenbank verglichen werden.

4.5.7 Alternative Ansätze

Die meisten Ansätze verwenden Aufnahmen der Tiergesichter, insofern ist hier eindeutig ein Mainstream zu erkennen. Dies soll allerdings nicht bedeuten, dass andere Ansätze nicht existieren, oder nicht erfolgreich sein können. Da gerade bei einfarbigen Tieren die Identifikation anhand von Felltexturen wegfällt, ist der Gedanke weitere Körpermerkmale wie komplette Seitenaufnahmen oder Bewegungsmuster zur Identifikation hinzuzunehmen durchaus nachvollziehbar. Damit einher geht natürlich auch, dass die Anforderungen an die Aufnahmen, respektive an die Aufnahmemethode, steigen. Denn meist hat dies zur Folge, dass Tiere nur in bestimmten Situationen aufgenommen werden können, wie beispielsweise beim Gehen, oder vollständig von der Seite. Dies schränkt die Anwendung der Identifikationsmethode stark ein, da zum Beispiel ruhende Tiere nicht identifiziert werden können.

Welcher Ansatz sich durchsetzt kann nur abgewartet werden. Die größte Anwendungsbreite ermöglicht natürlich die reine Gesichtserkennung, allerdings stellt diese höhere Ansprüche an das Modell. Zuversichtlich stimmt, dass die meisten Projekte diesen Ansatz wählen.

4.5.8 2D und 3D

Bei der Humanidentifikation kann davon ausgegangen werden, dass meist eine 3D-Repräsentation des Gesichts erstellt wird. Da Räumlichkeit im Vergleich zu 2D Aufnahmen eine nicht zu unterschätzende zusätzliche Informationsebene bietet, sind 3D-Repräsentationen besser geeignet.

Allerdings werden zur Anfertigung eines 3D-Modells mehrere 2D-Aufnahmen benötigt wie sie mit herkömmlichen Kameras angefertigt werden können. Dies erhöht wiederum die Anforderungen an die Aufnahmemethodik, wenn auch nicht so stark wie bei Seitenaufnahmen der Tiere. Mit genug räumlichem Versatz könnten 3D-Aufnahmen auch mit mehreren Kameras simultan angefertigt werden.

4.6 Systemaufbau

Um eine Vorstellung zu ermöglichen, wie ein Identifikationssystem mittels 2DAufnahmen aufgebaut sein könnte, wird hier ein Beispielentwurf skizziert.

4.6.1 Hardwareanforderungen

Der Erfolg des maschinellen Lernens im Bereich der Computer Vision ist auch darauf zurückzuführen, dass Bilder die mit herkömmlichen Kameras erstellt werden, als Datenmaterial vollkommen ausreichend sind. Vielmehr ist es sogar so, dass aktuelle Handykameras Bilder mit weit größeren Auflösungen produzieren, als sie von CNNs verarbeitet werden können.

Alle Bilder die mit demselben CNN verarbeitet werden, müssen das selbe Format haben, also dieselbe Anzahl Pixel. Aktuelle sind dies meist $224 \times 224 \times 3$, also ein quadratisches Bild mit 224 Pixel Kantenlänge und drei Farbkanälen. Aufnahmen dieses Formats sind mit jeder einfachen Kamera herzustellen.

Im Vergleich zu anderen biometrischen Identifikationsverfahren wie zum Beispiel Nahaufnahmen von Nasenmuster, Venenmuster mittels Laserabtastung, oder topografischen Rückenreliefs, stellen Aufnahmen vom Gesicht der Tiere weit geringere Ansprüche an die Aufnahmetechnik. Da hierfür einfache Kameras genügen, erweitern sich die Anwendungsmöglichkeiten entsprechend.

Nicht zu vernachlässigen ist jedoch die Beleuchtung. Kameras benötigen eine exakte Lichtmenge für Aufnahmen. Bei zu viel Licht wird die Aufnahmezeit verringert, bei zu wenig Licht muss die Aufnahmezeit verlängert werden. Dies ist so lange umsetzbar, bis Bewegung des Motivs zu einem Verwackeln führt.

Am besten arbeiten CNNs, wenn sich die Bilder idealerweise lediglich in den Eigenschaften unterscheiden, die auch zur Identifikation der einzelnen Tiere verwendet werden. Eine perfekte Umgebung wäre also ein komplett einfarbiger Hintergrund, der auf allen Aufnahmen identisch ist und somit die Mustererkennung nicht beeinflusst. Die Farbe und der Kontrast verändern sich auch mit der Lichtmenge: nimmt die Lichtmenge ab, werden die Farben dunkler und weniger differenziert. In einem idealen Aufbau wäre somit auch die Lichtmenge bei allen Aufnahmen identisch. Ein Ideal ist immer ein Gedankenspiel das sich nicht umsetzen lässt, allerdings hilft es beim Erkennen welche Eigenschaften Probleme erzeugen können und somit, welche Umgebungsparameter es zu minimieren, oder maximieren gilt. Im Stall muss somit sicherlich mittels zusätzlicher Beleuchtung für eine ausreichende Lichtmenge gesorgt werden. Im Freien kann dies abhängig von der Witterung und Tageszeit ebenfalls der Fall sein. Denkbar ist natürlich auch, eine variable Lichtmenge beim Erstellen des Datasets zu berücksichtigen, was aber den Aufwand der Erstellung des Dataset erhöht.

Je nach Geschwindigkeit der Tieridentifikation, so sie erstmals erfolgreich umgesetzt wurde, wird dann auch das Identifizieren individueller Tiere im Videostream ermöglicht. Und der nächste Schritt hiernach wäre das Identifizieren mehrerer Tiere gleichzeitig in einem Videostream. Die benötigte Rechenleistung macht die Installation eines leistungsfähigen Rechners notwendig. Allerdings kann ein Videostream leicht mit entsprechender Netzwerkverbindung an den, den Stream prozessierenden Rechner geschickt werden: ein Stream mit einer Auflösung von 1280×720 (HD) ist für eine Identifikation vollkommen ausreichend und benötigt mit entsprechend guter Komprimierung (zb. H.264 high quality) keine 2Mb/s Netzbandbreite.

4.6.2 Image preprocessing

Auf Bildern die von Kuhgesichtern mit einer Kamera angefertigt werden befindet sich das Gesicht nicht immer an derselben Stelle und sehr wahrscheinlich variiert auch der Aufnahmewinkel, sowie die Helligkeit etwas. Um die Varianz so weit wie möglich zu reduzieren ist der erste Schritt, das Gesicht zu lokalisieren und auszuschneiden. Dies würde mit einem separaten CNN umgesetzt werden, das darauf trainiert ist, Kuhgesichter in Bildern zu detektieren. Das Gesicht würde anhand der bounding box ausgeschnitten werden, sodass ein Bild entsteht welches neben dem Gesicht so wenig wie möglich anderen Inhalt hat. Ausschließlich dieses neue Teilbild wird dann weiter verarbeitet.

Um die Varianz zwischen verschiedenen Aufnahmen entsprechend des Kontrastbereichs zu reduzieren, werden die Bilder normalisiert. Ein Pixel kann für jeden der drei Farbkanäle Werte zwischen 0 und 255

annehmen, wobei 0 bedeutet, dass die entsprechende Farbe in dem entsprechenden Pixel nicht enthalten ist und 255 die maximale Farbintensität bedeutet. Die meisten Bilder nutzen dieses Spektrum nicht vollständig aus, sie erscheinen sie etwas stumpf, oder bedeckt. Beim Normalisieren eines Bildes wird das Spektrum so umgerechnet, dass die das Bild darstellenden Pixel dieses Spektrum voll ausschöpfen. Hierdurch werden die enthaltenen Muster deutlicher erkennbar, da die Unterschiede in Farbintensität, welche die Muster ja ausmachen, verstärkt werden.

Bevor ein Bild schließlich für das NN verwendbar ist, muss es standardisiert werden. Da NN am besten arbeiten, wenn alle Variablen denselben Wertebereich haben, werden die Variablen so umgerechnet, dass der Wertebereich zwischen 0 und 1 liegt. Im Fall der Farbkanäle wird das mit der Division jeden Wertes durch den Maximalwert 255 umgesetzt.

Diese Schritte beschreiben das normale Vorverarbeiten von Bildern, wenn sie mit NN klassifiziert werden sollen. Oft wird auch nur mit Graustufenbildern gearbeitet, da diese schon sehr viel Information enthalten und deutlich kleinere Modelle benötigen. Wie jedoch die Bilder vollständig vorbereitet werden hängt natürlich vom Modell ab, das dann folgend zur face recognition verwendet wird.

4.6.3 Modell

Derzeit erscheint der Ansatz der Gesichtserkennung unter Verwendung von transfer-learning am erfolgversprechendsten zu sein. Also ein leistungsfähiges Modell zu verwenden, das mit einem großen imageclassification-Dataset trainiert wurde und dessen eingangsnahen Schichten wiederverwendet werden. Die späteren, komplexeren Muster erkennenden Schichten würden dann mit einem extra hierfür angefertigten Dataset für Kuhgesichter trainiert werden.

Ob 2D, oder 3D und ob keypoints oder der ganzheitliche Ansatz sich durchsetzen, kann derzeit nicht bestimmt werden.

Das Entwickeln eines passenden Modells ist aktueller Forschungsgegenstand und somit kann hier keine Aussage darüber getroffen werden, wie ein erfolgreiches Modell aussehen wird.

4.6.4 Folgeprozesse

Die Identifikation von Tieren mittels optischer Methoden ermöglicht diverse neue Anwendungsfälle. Zusammen mit der Nachverfolgung (tracking) von individuellen Tieren könnte eine lückenlose Überwachung der Tiere erreicht werden. Angereichert mit den Ergebnissen der visuellen Verhaltens- und Tierwohlanalyse könnte jedes Tier derart überwacht werden, dass frühzeitig und automatisch erkannt und reagiert werden kann, sobald die überwachten Parameter einen definierten Grenzbereich verlassen.

5 Handlungsempfehlung

Die Leistungsbeschreibung dieser Recherchestudie definiert die Aufgabe so, dass es Ziel ist herauszufinden, ob sich Ohrmarken absehbar ablösen lassen durch ausschließliche Verwendung biometrischer Merkmale und optischer Aufnahmemethoden. Falls eine solche Technologie nicht existiert, soll ein Weg aufgezeigt werden, wie sie erreicht werden kann.

Ein Modell, das leistungsfähig genug ist um eine Menge von über 1.000 Tieren zuverlässig voneinander zu unterscheiden und zu identifizieren, existiert aktuell nicht bzw. konnte durch die beschriebenen Methoden nicht gefunden werden. Der Autor schätzt, dass es auch in den nächsten fünf Jahren kein System geben wird, welches 10.000 Tiere zu identifizieren vermag, vorausgesetzt die Forschung und Entwicklung in diesem Bereich verläuft in etwa mit der Geschwindigkeit wie bisher. Zur Erinnerung: das Ohrmarkensystem versieht 100.000.000 Tiere mit einem Identifikationsmerkmal innerhalb eines Bundeslandes.

5.1 Vorhandene Möglichkeiten zur optischen Tieridentifikation am Markt

Sollte der Wunsch bestehen ein kommerzielles Produkt zu Testzwecken zu evaluieren, so wird empfohlen das Unternehmen Cattleeye hierfür zu kontaktieren. Derzeit scheint dieses Projekt am aktivsten in Europa an einer Lösung zu arbeiten und auch das leistungsfähigste Modell entwickelt zu haben. Zudem bietet Cattleeye einen Aufbau mittels Teststellung an, was zum Beispiel Cattletracks nicht anbietet.

Das recht bekannte Unternehmen Cainthus hat langjährige Erfahrung im Bereich der optischen Tieridentifikation, allerdings derzeit keine Individualerkennungseigenschaften in seinem Produkt.

5.2 Darstellung geeigneter Lösungsansätze

Unabhängig des Entwicklungsstandes etwaiger Projekte, müssen folgende beiden Themen bearbeitet werden, bevor es einen Lösungsansatz geben kann der in der Lage wäre, Ohrmarken zu ersetzen.

5.2.1 Reduzierung der zu identifizierenden Tiermenge

Das Ablösen von Ohrmarken ist in der Form nicht möglich, da die Menge der Tiere die durch das Ohrmarkenschema mit einem Identifikationsmerkmal ausgestattet werden zu groß ist, als dass sie realistisch mit Mitteln des maschinellen Lernens umsetzbar wäre.

Es muss also die Menge der Tiere drastisch reduziert werden, die durch ein solches System identifiziert werden sollen.

Umgesetzt werden könnte dies beispielsweise durch das Einfügen einer zusätzlichen Hierarchieebene im Tier-ID-System. Also zum Beispiel zu Nationalstaat und Bundesland eine zusätzliche Kategorie, wie die des Betriebes. So ließe sich die Menge der zu identifizierenden Tiere deutlich näher an eine Mengengröße bringen, welche sich mit Methoden des maschinellen Lernens absehbar bearbeiten lässt.

5.2.2 Festlegung der Genauigkeit

Zusätzlich besteht weiterhin das Problem, dass die Genauigkeit mit der ein einzelnes Tier identifizierbar ist, mit der Anzahl der Tiere, die ein System zu verarbeiten hat, abnimmt. Es muss hier also für den Anwendungsfall der Tieridentifikation mit ausschließlich optisch-biometrischen Mitteln eine Genauigkeit festgelegt werden, die ein System erfüllen können muss. Solange ein Tier lediglich für betriebsinterne

Prozesse optisch-biometrisch identifiziert wird, sollte eine Genauigkeit von ~95% ausreichend sein. Rechtssicherheit hingegen dürfte nur mit Werten um, oder über 99% erreicht werden. Dass solche Systeme absehbar im Agrarbereich entstehen wird hier als unwahrscheinlich bewertet. Dennoch muss eine Richtlinie festgelegt werden, anhand derer sich Projekte, die an der Technologie forschen, orientieren können.

5.3 Formulierung der Schritte bis zur Anwendungsreife

Eine Anforderung dieser Recherchestudie ist das Aufzeigen notwendiger Schritte bis die Technologie einen Reifegrad erreicht hat, der die Anwendung im praktischen Betrieb ermöglicht.

Bezugnehmend auf die Technologiebeschreibung des vorangegangenen Kapitels ist es wichtig zu verstehen, dass leistungsfähige Modelle sukzessive entwickelt werden. Methoden des maschinellen Lernens werden laufend erweitert und verbessert, die Erfolge gehen Hand in Hand mit steigender Rechenleistung und immer größeren zum Training geeignete Datasets. Es ist nicht so, dass das eine Modell entwickelt wird, welches dann schlussendlich die Lösung für das angegebene Problem darstellt. Vielmehr werden bestehende Modelle immer weiter verbessert, sodass immer mehr Tiere damit identifiziert werden können. Es gibt also keine definierte Grenze, ab wann ein Modell »fertig« ist, vielmehr wird die Anzahl der Tiere erhöht, die ein Modell zuverlässig unterscheiden kann.

Das Entwickeln solcher Modelle ist Gegenstand aktueller Forschungsarbeit und der Weg bis zum Ziel kann nicht einfach definiert, sondern er muss erforscht werden.

Generell ist es so, dass je größer die Menge der Tiere ist, die eindeutig voneinander unterscheidbar und somit identifizierbar sein sollen, umso größer ist auch das zu lösende Problem. 400-500 Tiere zuverlässig zu unterscheiden scheint derzeit nach Herstellerangaben stabil umsetzbar. Je nach Heterogenität der Tiere können aktuell auch in Mengen von bis zu 800 Tieren Individuen identifiziert werden. Die Komplexität des Problems steigt aber nicht linear, sondern exponentiell mit der Anzahl der Tiere.

Der Technologie-Reifegrad ist also direkt abhängig von der Anzahl der Tiere. Bestünde die Menge aus 100 Tieren, wären aktuelle Modelle durchaus leistungsfähig genug direkt daraus ein Produkt zur Marktreife zu erstellen. Besteht die Menge aus 1.000 Tieren, befindet sich die Technologie noch in der Grundlagenforschung. Und es ist nicht absehbar, dass Ohrmarken in ihrer aktuellen Form abgelöst werden können.

5.4 Aufzeigen von Forschungsansätzen noch bestehender Lücken um Anwendungsreife zu erreichen

Eine Forderung der Ausschreibung ist das »Aufzeigen von Forschungsansätzen noch bestehender Lücken um Anwendungsreife zu erreichen«. Folgend sollen ein paar Möglichkeiten aufgezeigt werden, welche Maßnahmen als zielführend erachtet werden, um die Forschung und Entwicklung leistungsfähigerer Modellen günstig zu beeinflussen.

5.4.1 Kommunikation und Zusammenarbeit

Während den Recherchen zum Thema fiel auf, dass die Technologie zur optischen Tieridentifikation bisher keine eigene Plattform hat und die Projekte untereinander spärlich vernetzt erscheinen. Dies liegt sehr wahrscheinlich an der geringen Anzahl der Projekte die hier tätig sind, sowie am Pioniercharakter

dieser. Trotz der Konkurrenz der kommerziellen Projekte wurde aber eine hohe Kommunikationsbereitschaft festgestellt.

Ohne großen Aufwand könnte eine kleine Fachkonferenz, oder Workshop veranstaltet werden. Durch die Corona-Pandemie hat sich digitale Videokommunikation durchgesetzt, was ein immenser Vorteil für das Veranstalten kleiner Konferenzen ist. Es können internationale Akteure sehr bequem ohne zeit- und kostenintensive Anreise teilnehmen.

Einerseits wäre dies eine Möglichkeit um weitere Einblicke in die aktuelle Forschung und Produktentwicklung zu erhalten, andererseits könnten Probleme, Bedürfnisse und Kapazitäten kommuniziert und so etwaige Synergien erkennbar werden.

Da aktive Forschung bezüglich des Themas in Deutschland nicht zu existieren scheint, müsste die Konferenz in englischer Sprache abgehalten werden. Neben denen am Markt gefundenen Unternehmen, sollten auch Vertreter der Expertengruppen eingeladen werden.

5.4.2 Daten

Das Erstellen von geeigneten Datasets ist sehr aufwändig. Für Identifikationssysteme im Humanbereich wird angenommen, dass die Datasets die zum Trainieren notwendig waren, aus 10Mio bis 100Mio Bildern bestehen. Solche Systeme sind in China bereits erfolgreich im Einsatz und die Unternehmen die diese Systeme erstellt haben, erzeugten die Trainingsdaten selbst. Leider sind diese somit nicht öffentlich verfügbar.

Datasets bestehen nicht nur aus Bilddateien, sondern auch aus den jeweiligen Labels. Also den zu jedem Bild vorhandenen Informationen die es dem Modell schlussendlich ermöglichen zu lernen. Je nach Modellentwurf unterscheiden sich die Labels voneinander, sodass das Erstellen von Datasets nicht unabhängig vom Modell durchführbar ist.

Könnten sich die Teilnehmer in der Fachtagung auf eine Definition einigen die festlegt, welcher Gestalt ein einheitlich nutzbares Dataset sein müsste, so wäre es möglich ein öffentlich verfügbares, oder zumindest gemeinsam nutzbares, Dataset zu erstellen. Mit diesem könnten dann auch kleinere Projekte, zum Beispiel an Universitäten an geeigneten Modellen forschen. Die größte Hürde für das Entstehen vieler kleiner Forschungsprojekte wäre somit behoben.

Parallelität ist ein sehr wichtiges Kriterium in der Forschung. Je mehr Projekte es gibt, die an einer Thematik forschen, umso größer der Austausch, die Synergieeffekte und schlussendlich die Innovation. Oft entstehen neue Modelle dadurch, dass bestehende Ansätze angepasst und verbessert werden. Das eifrige Veröffentlichen neuer Ideen und Entdeckungen durch die KI-Community ist ein wichtiger Faktor für den Erfolg dieser Technologie.

Ein niederschwellig verfügbares Dataset würde die Entwicklung neuer Modelle stark begünstigen.

5.4.3 Initiieren eigener Forschungsprojekte

Sollte der Wunsch bestehen, eigene Forschungsaktivitäten anzustrengen, wären folgende Punkte zu berücksichtigen:

- Ein schneller Erfolg ist nicht zu erwarten. Im kommerziellen Umfeld wurden biometrische Individual-Identifikationseigenschaften bisher bei Bovinen nicht umgesetzt, weil der Aufwand den Nutzen überstieg. Dies bedeutet nicht, dass die Umsetzung nicht möglich ist, aber sehr wohl, dass sie nicht einfach umsetzbar ist. Auch haben Projekte mehrere Jahre an Forschung investiert, ohne ein fertiges Produkt erstellen zu können. Es muss also davon ausgegangen werden, dass mehrere Jahre notwendig sind, um einen Prototypen zu entwickeln, der zuverlässig mehr als 1000 homogene Tiere unterscheiden können soll.
- Es wird entsprechende IT-Infrastruktur benötigt, sowie Fachpersonal um diese aufzubauen und zu pflegen. Dies sind Server mit leistungsstarken Grafikkarten, schnelle Dateiserver die mit sehr großen Datenmengen performant umgehen können und ein entsprechend leistungsfähiges Netzwerk. Selbst wenn in etwaigen Forschungseinrichtungen diese Infrastruktur schon vorhanden sein sollte, muss der Mehraufwand durch die zusätzliche Nutzung berücksichtigt werden. Natürlich verlangsamt konkurrierender Zugriff auf gemeinsam genutzte Ressourcen den Forschungs- und Entwicklungsprozess teilweise erheblich.
- Sollte für das Forschungsprojekt kein Dataset bestehen, muss dieses erstellt werden. Die Gestalt des Datasets muss vom gesamten Team festgelegt werden. Es werden auch hierfür Fachkräfte benötigt. Mit neuen Erkenntnissen aus der Forschung werden wahrscheinlich neue Anforderungen an die Gestalt des Datasets herangetragen, was wiederum zu großem Aufwand führen kann. Der Aufwand für das Erstellen guter Trainingsdaten kann nicht unterschätzt werden. Für das Erstellen der Labels könnten aber auch zusätzliche studentische Hilfskräfte engagiert werden. Es wird Zugang zu einer großen Menge an Kühen oder Rindern benötigt, um diese geeignet und mehrfach fotografieren, bzw. filmen zu können. Idealerweise über einen längeren Zeitraum um Alterungseffekte auf das Lernverhalten des Modells untersuchen zu können. Wichtig ist natürlich, dass die Identität kontinuierlich dokumentiert wird.
- Ein interdisziplinäres Forscherteam, das mindestens Kenntnisse im agrartechnischen Bereich hat, sowie natürlich mindestens ein dauerhafter Forscher im maschinellen Lernen der keinerlei Nebentätigkeiten wahrnehmen muss. Unterstützt werden könnte dieser von Doktoranden, falls die Organisationsstruktur dies zulässt.
- Aufgrund der Vorkenntnisse und Synergien wird empfohlen, ein entsprechendes Forschungsprojekt an bereits bestehenden Instituten der Agrarinformatik anzusiedeln.

6 Zusammenfassung

Die Nutzung biometrischer Merkmale zur Tiererkennung in der Rinderhaltung besitzt großes Potenzial und weckt weltweites Interesse. Die Motivationen welche den Arbeiten im Bereich der optischen Tieridentifikation zugrunde liegen, sind regional verschieden. Ein Projekt der University of Kentucky, USA, forscht an Tierortung und Identifikation mittels Drohnen. Durch die Ortstypischen großen Weideflächen bildet die Tierortung und Gesundheitskontrolle von Einzeltieren einen entscheidenden Kostenfaktor, der durch die Forschungsinitiative verringert werden soll. Ein weiterer Ansatz aus den USA ist die Nutzung zur Seuchenverfolgung und Krankheitsvorsorge. Gerade der globale Handel mit Nutztieren ist an einer eindeutigen Zuordnung von tierspezifischen Gesundheitsdaten und Herkunftszuordnungen interessiert. Im asiatischen Raum liegt die Motivation meist in der Fälschungsprävention. Diese Initiativen gehen aus der Versicherungsbranche hervor. Versicherungsbetrug und gefälschte Tierdaten sind dort weit verbreitete Probleme. Projekte im europäischen Raum jedoch stammen aus dem Bereich Systemanbindung im Herdenmanagement. Dabei sind aktive Initiativen hauptsächlich in Großbritannien angesiedelt. Bemerkenswert ist zudem, dass eher kleinere und mittelständige Unternehmen im Bereich Tieridentifikation forschen. Die Hersteller scheinen trotz langjährig etablierter Tierkennzeichnungs-methoden auf das Potential der neuartigen Tieridentifikation zu vertrauen und schätzen das Potential für Folgeinnovationen hoch ein.

Für das strukturierte Vorgehen zur Bearbeitung dieses vielschichtigen Themas wurde eine umfassende Rechercharbeit durchgeführt. Dabei wurde sowohl der Stand der Technik als auch der aktuelle Umsetzungsfortschritt im Themenbereich abgebildet.

Die Bereitschaft relevante Kontakte weiter zu vermitteln war in der Gruppe der angefragten Personen generell hoch, sodass eine gute Durchdringung erreicht werden konnte. Schlussendlich wurden ca. 95 Akteure insgesamt kontaktiert. Ein weiterführender Informationsaustausch fand mit zehn Institutionen statt. Am Ende der Marktrecherche ergaben sich zusätzlich einige Expertengespräche, die das entstandene Bild abrundeten und bestätigten. Zu beachten ist, dass kein absolut vollständiger Überblick über alle Projekte weltweit generiert werden kann. Einerseits, da die Suche auf den deutschen und englischen Sprachraum beschränkt ist. Aber gerade der Blick nach China lohnt, detaillierte Einblicke in Modellstrukturen lieferten zwar nur einige wenige Patente, der Grund dafür ist, dass gerade dieses Wissen zentraler Gegenstand aktueller Produktentwicklungen ist. Als aktive Initiative mit großem Potential für die praktische Umsetzung der Thematik wird das Unternehmen Cattleeye identifiziert. Cattleeye wurde im Januar 2019 gegründet und besteht aus einem kleinen aber versierten Team von Spezialisten. Das Unternehmen ist Systemanbieter im Bereich Monitoring von Gesundheit und Leistungsfähigkeit von Milchkühen. Die Überwachung wird u. a. mit einfachen Kameras an Ein- und Ausgang des Melkstandes umgesetzt. Die Ergebnisse sind als erfolversprechend einzustufen. Das Unternehmen ist der einzige ermittelte Anbieter aus Europa mit einem System das Individualidentifikation durchführt. Das Unternehmen gibt eine Identifikationsrate von 98 % Präzision an bei gemusterten Tieren. Die Rate bei einfarbigen Individuen ist nicht angegeben. Die maximale Herdengröße innerhalb derer eine zuverlässige Identifizierung durchgeführt werden konnte, lag zum Recherchezeitpunkt bei ca. 800 Tieren.

Es kann Zusammengefasst werden, dass sich deutlich der Ansatz der Gesichtserkennung mittels Convolutional Neuronal Networks (CNN) als Lösung in den meisten Patenten durchsetzt. CNNs haben sich in den letzten Jahren als leistungsfähige Methode erwiesen, wenn es um das Klassifizieren von Bilddaten geht. Sie werden intensiv eingesetzt und die so entstandene Erfahrung beim Erstellen von

Modellen für die menschliche Gesichtserkennung wird nun auch beim Erkennen und Identifizieren von Tieren angewandt. Ein Bild ist der Input eines Modells. Der Output ist eine Liste von Werten, die das Ergebnis der Berechnungen auf dem Bild sind, auch Feature-Vektor genannt. Umso besser das Modell ist, umso ähnlicher sind sich die Feature-Vektoren von Bildern eines Tieres unabhängig anderer Eigenschaften des Bildes. Umso weniger wirken sich z.B. unterschiedliche Aufnahmewinkel, oder variierende Lichtverhältnisse auf den Inhalt des Feature-Vektors aus. Ein großer Vorteil dieses Ansatzes ist es, dass an die Fotos, der zu identifizierenden Tiere, relativ geringe Ansprüche gestellt werden. Es werden keine Spezialkameras benötigt und zudem ist das Gesicht eines Tieres von Mensch und Maschine leicht zu erkennen und somit leichter verwertbar aufzunehmen, als beispielsweise ein definierter Bereich der Flanke, oder eines Euters, wie es beim Venenscannen notwendig ist. Die Auflösung der Bilder ist dennoch nicht irrelevant. Je größer die Bilder sind, die vom Modell verarbeitet werden können, umso mehr Daten stehen pro Aufnahme zur Verfügung. Gerade bei sehr großen Herden mit vielen sehr ähnlichen Tieren sind geringe Unterschiede entscheidend für die erfolgreiche Identifikation. Ein Punkt der aufgrund des geringen Alters der Technologie noch nirgends erwähnt wurde, ist die Standardisierung. Soll eine rein optische Identifikation irgendwann marktreif sein, so muss sich mit der Möglichkeit auseinandergesetzt werden, wie die Datensätze, die mit einem Tier assoziiert sind (also die Sammlung an Feature-Vektoren eines Tieres) von einem Besitzer zum nächsten übermittelt werden können. In allen Bereichen des maschinellen Lernens ist der Fortschritt hungrig nach Daten. Es existieren vielfältige Datasets zu allen möglichen Themen. Datasets bestehen neben den Bildern aus Labels anhand derer das trainierte Modell lernt. Hierfür müssen die Bilder von Menschen manuell gelabelt, oder zumindest qualifiziert werden. Oft kommen hier mittlerweile auch schon Modelle zum Einsatz, die Labels erstellen. Diese Labels müssen vom Menschen geprüft werden, da nicht ausgeschlossen werden kann, dass Fehler enthalten sind. Fehler in Trainingsdaten sind fatal, da sie Fehlerhafte Modelle zur Folge haben. Bei mehreren Millionen Bildern und mehreren möglichen Objekten pro Bild ist der Aufwand gute Labels zu erstellen nicht zu unterschätzen.

Bezugnehmend auf die Technologiebeschreibung ist es wichtig zu verstehen, dass leistungsfähige Modelle sukzessive entwickelt werden. Methoden des maschinellen Lernens werden ständig erweitert und verbessert, die Erfolge gehen Hand in Hand mit steigender Rechenleistung und immer größeren zum Training geeigneter Datasets. Bestehende Modelle werden weiter verbessert, sodass immer mehr Tiere damit identifiziert werden können. Dies ist ein stetiger Prozess, um die identifizierbare Tieranzahl fortlaufend steigern zu können. Nach mehreren Herstellerangaben scheint eine Unterscheidung von 400-500 Tieren derzeit zuverlässig zu funktionieren. Die Komplexität des Problems steigt aber nicht linear, sondern exponentiell mit der Anzahl der Tiere. Der Technologie-Reifegrad ist also direkt abhängig von der Anzahl der Tiere. Zum Recherchezeitpunkt konnte kein Modell zuverlässig 1.000 Tiere identifizieren. Die Tieridentifikation mittels Ohrmarke ermöglicht es 10 000 000 Tiere zweifelsfrei zuzuordnen. Eine Möglichkeit die optische Identifizierung nutzbar zu machen, kann die Reduzierung der zu verwaltenden Tiermenge sein. Durch die Einführung der Kategorie Betrieb im ID-Schema bestehend aus Land-Bundesland-Betrieb-Tier. So kann die Menge der zu identifizierenden Tiere deutlich näher an eine Mengengröße angenähert werden, die durch Methoden des maschinellen Lernens absehbar bearbeitbar sind. Des Weiteren muss die Genauigkeit mit der ein einzelnes Tier identifizierbar ist, festgelegt werden. Hierzu wird eine rechtliche Grundlage notwendig, die für alle Systeme einzuhalten ist. Für den Anwendungsfall der Tieridentifikation mit ausschließlich optisch-biometrischen Mitteln muss eine Genauigkeit festgelegt werden, die ein System erfüllen muss. Rechtssicherheit bieten Werte nur bei einer Genauigkeit von über 99%. Für das Training von Modellen sind große Datasets notwendig, diese

sind ausschlaggebend für die Verbesserung der Identifizierung. Datasets bestehen nicht nur aus Bilddateien, sondern auch aus den jeweiligen Labels. Also den zu jedem Bild vorhandenen Informationen, die es dem Modell schlussendlich ermöglichen zu lernen. Je nach Modellentwurf unterscheiden sich die Labels voneinander, sodass das Erstellen von Datasets nicht unabhängig vom Modell durchführbar ist. Die Vernetzung von Projekten stellt die Grundlage für ein einheitliches Dataset dar. Dabei kann eine Definition zur Gestalt eines öffentlich verfügbaren bzw. von Interessengruppen gemeinsam nutzbaren Datasets erstellt werden und dieses über einen Zentralen Zugang zur Verfügung stehen. Niederschwellig verfügbare Datasets würden die Entwicklung neuer Modelle stark begünstigen. Davon würde die Forschung in diesem Bereich elementar profitieren und initiale Hemmnisse kleinerer Forschungsgruppen reduzieren. Eine aktive Förderung der Innovation wird durch die Ausschöpfung von Parallelität und Synergieeffekten erreicht.

Zum aktuellen Zeitpunkt und den vorliegenden Ergebnissen zugrunde legend, wird geschlussfolgert, dass die optische Tieridentifikation die handelsübliche Ohrmarke nicht ablösen wird. Die Verarbeitung von Bilddaten ist innerhalb einzelner Betriebsstrukturen interessant für die Steuerung und Unterstützung von definierten Geschäftsprozessen. Weiterhin ist die optische Tieridentifikation klar als Befähiger-Technologie für vielfältige Folgeinnovationen zu betrachten, wie beispielsweise die Weideüberwachung mittels Drohnen. Dabei ist die größte Chance zur Einsatzfähigkeit und einwandfreien Zuordnung optischer Daten durch immer höhere Verarbeitungskapazität in kürzester Zeit zu erreichen.

7 Schlussfolgerung

Durch die Erkenntnisse der Machbarkeitsstudie wird deutlich, dass die technische Umsetzung eines Verfahrens zur Einzeltieridentifikation auf Grundlage bildgebender Verfahren zum jetzigen Zeitpunkt bereits möglich ist. Essenziell für die Entwicklung marktreifer Systeme ist die zugrundeliegende Datenbank für die Programmierung des lernfähigen Algorithmus. Je größer der Datenpool an einzelnen Individuen, um die Lerngrundlage eines Systems zu bilden anhand derer Tiere identifiziert werden, desto höher ist der Lernerfolg des Systems und desto fehlerfreier die Einzeltieridentifikation. Die Erstellung eines öffentlich zugängigen Datensets, barrierefrei, durch unabhängige Stellen befüllt und über Landesgrenzen hinweg verfügbar stellt eine wichtige Basis für weiterführende Forschungen in dem Bereich dar. Dies ist sowohl für die Datenvielfalt und Qualität vorteilhaft, als auch für die Gleichstellung aller Forschungsgruppen. Das Landesamt für Umwelt, Landwirtschaft und Geologie (LfULG) hat eine beratende Rolle für politische Entscheidungen und die Umsetzung von FuE-Projekten, sowie den Wissenstransfer in die Praxis inne. Die vorliegende Machbarkeitsstudie zeigt klar, dass die Thematik der optischen Tieridentifikation eine länderübergreifende, tiefgründige Initiative erfordert um die Grundlagenforschung zu beleben und ein vielgliedriges Dataset zu erzeugen. Mit den vorliegenden Ergebnissen kann das LfULG eine vermittelnde Rolle einnehmen und fachlich beratend zur Verfügung stehen. Es wird angestrebt eine internationale Arbeitsgruppe zu diesem Thema zu initiieren und fachlich begleitend zu unterstützen, die Thematik weiter im Blick zu behalten, die Entwicklungen zu beobachten, aber zum aktuellen Zeitpunkt keine weiteren Forschungsschwerpunkte im eigenen Haus zu setzen.

8 Quellen

8.1 Web-Seiten

- [KSN20] Kansas State University (2020): Internetpublikation zum Projekt cattletracs, ksre.k-state.edu/news/stories/2020/09/facial-recognition-for-cattle.html, 23.02.2021
- [BHF20] o.V (2020). Unternehmen, welches das Produkt Cattletracks entwickelt, blackhereford.com, 23.02.2021
- [UKY19] Kentucky State University (2019) News, <https://news.ca.uky.edu/article/ready-clarabelle-her-close-uk-scientists-need-pictures-cows-research-project>, 23.02.2021.
- [UKY20] Kentucky State University (2019) News, <https://news.ca.uky.edu/article/could-drones-save-cows-why-uk-research-team-thinks-so> , 23.02.2021
- [BBCUKY] BBC (2021): News zum Projekt der Kentucky State University [UKY20], <https://www.bbc.com/future/bespoke/follow-the-food/drones-finding-cattle-where-cowboys-cannot-reach.html>, 23.02.2021.
- [FNAC20-1] Feed Navigator (2020): News-Artikel von zu Cainthus, <https://www.feednavigator.com/Article/2020/02/04/Cainthus-launches-dairy-feeding-monitoring-technology>, 24.02.2021
- [FNAC20-2] Feed Navigator (2020): News-Artikel von zu Cainthus, <https://www.feednavigator.com/Article/2020/12/01/Cainthus-launches-new-technology-allowing-large-dairies-to-further-optimize-milk-production-and-animal-welfare>, 24.02.2021
- [BPDHI18] Cainthus (2018), Interview mit David Hunt, Co-Founder und CSO, <https://www.bangkokpost.com/business/1516646/agritech-evangelist>, 24.02.2021
- [INN10] o.V. (2020): News-Artikel, 03.03.2021, innovationnewsnetwork.com/cattle-facialrecognition-could-combat-agricultural-fraud/3014/
- [WIKIML21] Wikipedia (2020) Maschinelles Lernen, de.wikipedia.org/wiki/Maschinelles_Lernen, 15.03.2021
- [WIKIKNN21] Wikipedia (2020) https://de.wikipedia.org/wiki/K%C3%BCnstliches_neuronales_Netz, 18.03.2021
- [WIKICNN21] Wikipedia (2021): Convolutional Neural Networks, de.wikipedia.org/wiki/Convolutional_Neural_Network, 22.03.2021
- [WIKIDAS21] Wikipedia (2021) List of datasets for machine-learning research, https://en.wikipedia.org/wiki/List_of_datasets_for_machine-learning_research, 22.03.2021
- [ILSVRC] ImageNet Large Scale Visual Recognition Challenge (ILSVRC): www.image-net.org/challenges/LSVRC
- [FEAVIS17] o.V. (2017): Artikel zu: Feature Visualization in NN: distill.pub/2017/featurevisualization
https://hci.iwr.uni-heidelberg.de/system/files/private/downloads/1004425242/johannes_vogt_feature_visualisation_slides.pdf
- [INWIN] Timeline der Klassifizierungsleistung von Modellen mit dem ImageNet Dataset, <https://paperswithcode.com/dataset/imagenet>
- [ZFNET13] Zeiler, M., Fergus, R.(2013): Visualizing and Understanding Convolutional Networks, arxiv.org/pdf/1311.2901.pdf

[OPENIMA21] Google Open Images Dataset(2021): opensource.google/projects/open-images-dataset,
resp.: storage.googleapis.com/openimages/web/factsfigures.html

8.2 Papers

- [BMEL19] Bundesministerium für Ernährung und Landwirtschaft (2019): Kennzeichnung und Registrierung von Rindern,
bmel.de/SharedDocs/Downloads/DE/_Tiere/Nutztiere/KennzeichnungRinder.pdf, 05.03.21
- [SHH09] Herculano-Houzel, S. (2009): The Human Brain in Numbers: A Linearly Scaled-up Primate Brain. In: *Front Hum Neurosci*, Band 3, Nr. 31, November 2009, S. 1–11, PMC 2776484
- [BABR01] Michele Banko, M., Brill, E. (2001): Scaling to Very Very Large Corpora for Natural Language Disambiguation, Microsoft Research, dl.acm.org/doi/pdf/10.3115/1073012.1073017
- [HLARG21] Hasani, R., Lechner, M., Amini, A., Rus, D., Grosu, R. (2021): Liquid Time-constant Networks, MIT 2021, arxiv.org/pdf/2006.04439.pdf, MIT News-Artikel zum Thema:
news.mit.edu/2021/machine-learning-adapts-0128, 11.03.2021
- [RU86] Rumelhart, D., Hinton, G. and Williams, R. (1986): Learning Internal Representations by Error Propagation«, web.stanford.edu/class/psych209a/ReadingsByDate/02_06/PDPVollChapter8.pdf
- [HUWI58] David H. Hubel, D., Wiesel, T.(1958): Single Unit Activity In Striate Cortex Of Unrestrained Cats, www.ncbi.nlm.nih.gov/pmc/articles/PMC1357023/pdf/jphysiol01301-0020.pdf
- [HUWI59] David H. Hubel, D., Wiesel, T.(1959): Receptive Fields Of Single Neurons In The Cat's Striate Cortex«, www.ncbi.nlm.nih.gov/pmc/articles/PMC1363130/pdf/jphysiol01298-0128.pdf
- [LECU98] Yann LeCun, Y. (1998), Gradient-Based Learning Applied to Document Recognition,
yann.lecun.com/exdb/publis/pdf/lecun-01a.pdf
- [KRIZ12] Krizhevsky, A., Sutskever, I., Hinton, G. (2012): maheNet Classification with Deep Convolutional Neural Networks,
papers.nips.cc/paper/2012/file/c399862d3b9d6b76c8436e924a68c45bPaper.pdf
- [KUZOPIM20] Kuznetsova, A. (2020): The Open Images Dataset V4, arxiv.org/pdf/1811.00982.pdf

8.3 Bücher

- [RUNO12] Russel, S., Norvoig, P. (2012): Künstliche Intelligenz – Ein moderner Ansatz, 3. aktualisierte Auflage, Hallbergmoos: Pearson Deutschland GmbH
- [LÄCL08] Lämmel, U., Cleve, J. (2008): Künstliche Intelligenz, 3. neu bearbeitete Auflage, München: Carl Hanser Verlag München
- [GÉ17] Géron, A.(2017): Hands-On Machine Learning with Scikit-Learn & TensorFlow, O'Reilly Media Inc.: Newton
- [GÉ19] Géron, A.(2019): Hands-On Machine Learning with Scikit-Learn, Keras & TensorFlow, 2nd Edition, Newton: O'Reilly Media Inc.
- [BR01] Bräunl, T., Feyrer, S., Rapf, W., Reinhardt, M.(2001): Parallel Image Processing, Heidelberg: Springer-Verlag Berlin Heidelberg
- [GO16] Goodfellow, G., Bengio, Y., Courville, A.(2016): Deep Learning, First edition, Cambridge: MIT Press
- [NI15] Nielsen, M. (2015): Neural Networks and Deep Learning, Determination Press
neuralnetworksanddeeplearning.com/chap6.html

Herausgeber:

Sächsisches Landesamt für Umwelt, Landwirtschaft und
Geologie (LfULG)
Pillnitzer Platz 3, 01326 Dresden
Telefon: + 49 351 2612-0
Telefax: + 49 351 2612-1099
E- Mail: Poststelle.LfULG@smekul.sachsen.de
www.lfulg.sachsen.de

Autor:

Tobias Heinzmann
Softwareentwicklung und IT-Beratung
Martinstraße 18, 04318 Leipzig
Telefon: +49 176 8794 4520
E-Mail: mail@tobias-heinzmann.de

Redaktion:

Franziska Deißing
Abteilung 7, Referat 7St
Am Park 3, 04886 Köllitsch
Telefon: + 49 34222 46-2110
E-Mail: Franziska.Deissing@smekul.sachsen.de

Fotos:

Urheberangaben finden sich unter den betroffenen
Abbildungen und Fotos im Text.

Redaktionsschluss:

27.10.2022

ISSN:

1867-2868

Hinweis:

Die Broschüre steht nicht als Printmedium zur Verfügung,
kann aber als
PDF-Datei unter <https://publikationen.sachsen.de>
heruntergeladen werden.

Verteilerhinweis

Diese Informationsschrift wird von der Sächsischen
Staatsregierung im Rahmen ihrer verfassungsmäßigen
Verpflichtung zur Information der Öffentlichkeit
herausgegeben.

Sie darf weder von Parteien noch von deren Kandidaten
oder Helfern zum Zwecke der Wahlwerbung verwendet
werden. Dies gilt für alle Wahlen. Missbräuchlich ist
insbesondere die Verteilung auf Wahlveranstaltungen,
an Informationsständen der Parteien sowie das Einlegen,
Aufdrucken oder Aufkleben parteipolitischer Informationen
oder Werbemittel. Untersagt ist auch die Weitergabe an
Dritte zur Verwendung bei der Wahlwerbung.

*Täglich für
ein gutes Leben.*

www.lfulg.sachsen.de