


Einzeltieridentifikation über biometrische Bildanalyse mittels künstlicher Intelligenz

Franziska Deißing ¹, Tobias Heinzmann², Dorothee Heyde, Dr. Steffen Pache,
Dr. Maximilian Deutsch

Abstract: Wie alle Lebewesen besitzen Rinder individuelle Merkmale, die für eine biometrische Identifizierung herangezogen werden können, sofern deren Ausprägung mit Methoden des maschinellen Lernens eine reproduzierbare Klassifizierung und eine eindeutige Differenzierung erlaubt. Auf Basis von zwei Machbarkeitsstudien wird gezeigt, inwiefern eine eindeutige Einzeltieridentifikation über Bildanalyse prinzipiell möglich ist, welche tierindividuellen Merkmale zwischen Individuen Unterschiede zeigen, wie sich der Forschungsstand darstellt und welche technischen Möglichkeiten zur biometrischen Tieridentifikation aktuell verfügbar sind. Daneben werden mögliche Lösungsansätze zum Forschungsbedarf im Bereich der optischen Tieridentifikation aufgezeigt.

Keywords: Biometrische Identifizierung, Maschinelles Lernen, Datasets

1 Einleitung

Elektronische Tierkennzeichnung ist die Voraussetzung der digitalen Datenerhebung im Stall und wird seit vielen Jahren in der Praxis eingesetzt. Die Kennzeichnungshardware als Halsbandtransponder oder elektronische Ohrmarke begleitet das Tier während seiner gesamten Lebenszeit. Die administrative Kennzeichnung und Registrierung von Rindern ist nach der Viehverkehrsverordnung [VVVO] genau definiert.

Diese Art von Kennzeichnung bringt neben dem Nutzen aber auch einige Nachteile mit sich. Die für die Kennzeichnung nach VVVO zugelassenen elektronischen Ohrmarken arbeiten im Low Frequenzbereich entsprechend der DIN ISO 11785. Für eine zuverlässige elektronische Tieridentifikation müssen die Tiere langsam, mit nicht mehr als 3 m/s, und vereinzelt in einem Abstand von 40-80 cm am Lesegerät vorbeilaufen. Zur Vermeidung der Lesung benachbarter Tiere sind die Lesestationen entsprechend abzuschirmen und vor der Einwirkung anderer Magnetfelder zu schützen. Nachdem sich in den letzten Jahrzehnten die Biometrie als Wissenschaft zur Klassifizierung und Identifizierung von Merkmalen an Lebewesen etablieren konnte, gewinnt diese Methode des maschinellen Lernens durch die verfügbaren Rechenkapazitäten auf vernetzten Servern zunehmend an Bedeutung. Wie alle Lebewesen besitzen Rinder individuelle Merkmale, die für eine biometrische Identifizierung herangezogen werden können, sofern deren Ausprägungen

¹ Landesamt für Umwelt, Landwirtschaft und Geologie, Abteilung 7 Landwirtschaft, Am Park 3, 04886

Arzberg, franziska.deissing@smekul.sachsen.de,  <https://orcid.org/0000-0000-0000-0000>

² Softwareentwickler und IT-Berater, Martinstraße 18, 04318 Leipzig, mail@tobias-heinzmann.de

genügend Varianz aufweist und damit Differenzierung erlaubt. Als Beispiele könnten hier die Musterung des Felles, die Aderung von Körperregionen, Körperproportionen, Haarwirbel und Farbschattierungen genannt werden. Die Aufgabe für die Zukunft liegt in der Entwicklung eines geeigneten, stressarmen Tieridentifikationssystems, das den Umweltbedingungen in der Tierhaltung, der Tiergesundheit und der Mobilität der Tiere Rechnung trägt. Auf der Grundlage der Recherche aktueller Forschungsinitiativen werden nachfolgend prinzipielle Möglichkeiten der Einzeltieridentifizierung durch Klassifizierung biometrischer Merkmalsausprägungen mittels künstlicher Intelligenz aufgezeigt.

2 Stand des Wissens

2.1 Status quo

Eigene Voruntersuchungen zur beschriebenen Problemstellung fanden bereits in verschiedenen Bereichen statt. Hierbei sind die Projekte on-Farm-Recording [Pa13], Cow-body-scan [De19] und ein eigener Vorbericht zum Themenkomplex Optische Tieranalyse aus 2020 zentrale Ansätze. Alle drei beschäftigen sich mit bildgebenden Verfahren in der Rinderhaltung und mit der Möglichkeit einzelne Individuen von anderen zu unterscheiden. Im Fazit wurde deutlich, dass die tatsächliche eindeutige Identifikation von Rindern mittels biometrischer Bildanalyse noch weit von praxistauglichen Anwendungen entfernt ist.

2.2 Regional unterschiedliche Motivationen

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. Insbesondere der globale Handel mit Nutztieren ist an einer eindeutigen Zuordnung von tierspezifischen Gesundheitsdaten und Herkunftszuordnungen interessiert. Im asiatischen Raum geht die Motivation von der Versicherungsbranche aus. Versicherungsbetrug mit gefälschten Tierdaten sind dort ein weit verbreitetes Problem. Projekte im europäischen Raum werden vornehmlich in Großbritannien bearbeitet und sind im Bereich Systemanbindung im Herdenmanagement angesiedelt. Bemerkenswert ist, dass eher kleinere und mittelständige IT-Unternehmen im Bereich der optischen Tieridentifikation forschen. Diese Teams vertrauen trotz der etablierten, elektronischen Kennzeichnungsmethoden auf die Anwendung dieser

neuartigen, biometrischen Tieridentifikation und schätzen das Potential für Folgeinnovationen hoch ein.

3 Material und Methoden

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

3.1 Patent- und Marktrecherche

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 zunächst eine automatisierte Breitensuche durchgeführt. Um aus den über 1.700 Fundstellen die relevanten Patente herauszufinden, erfolgte im Anschluss eine teil-manuelle Sichtung von ca. 500 Dokumenten.

Das Ziel der Marktrecherche war neben der Darlegung fertiger Produktlösungen das Aufzeigen von Projekten, die an diesem Thema arbeiten, unabhängig ihrer kommerziellen Verwertung, oder des Reifegrades der Technologie. Die Suche wurde zweistufig durchgeführt. Zunächst wurden alle den Autoren bekannten Unternehmen und Verbände via Email kontaktiert, darauf folgte eine intensive Internetrecherche und Kontaktaufnahme zu den identifizierten Vorhaben.

3.2 Technologiebeschreibung

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]. Der am häufigsten von unterschiedlichen Projekten angewandte Ansatz, sind Konvolute Neuronale Netzwerke, die mit Bilddaten von Tiergesichtern erstellt wurden. Auf der Suche nach Lösungsansätzen für die biometrische Identifizierung von Rindern wurden diese Modelle des maschinellen Lernens intensiver untersucht.

4 Ergebnisse

4.1 Patent- und Marktrecherche

Die aufgezeigten Recherchemethoden und die daraus resultierten Erkenntnisse werden im Folgenden diskutiert. In der Patentrecherche wurden insgesamt 25 relevante Ansätze

detektiert, vornehmlich aus China (21), Schweden (2), Dänemark und USA (je 1). Die Bereitschaft relevante Kontakte weiter zu vermitteln war in der Gruppe der kontaktierten Personen generell hoch, sodass eine gute Durchdringung der aktuellen Arbeiten erreicht werden konnte. Schlussendlich wurden ca. 95 Akteure kontaktiert. Ein weiterführender Informationsaustausch fand mit zehn Institutionen statt. Am Ende der Patent- und Marktrecherche ergaben sich 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 war. Aber gerade in China fehlen in vielen Patenten detaillierte Einblicke in Modellstrukturen, die jedoch zentraler Gegenstand zur Beurteilung der aktuellen, chinesischen Produktentwicklungen sind. 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 am 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 Einzeltieridentifikation durchführen kann. 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 beim Recherchezeitpunkt bei ca. 800 Tieren.

4.2 Technologiebeschreibung

Die Ergebnisse der Patentrecherche zeigen, dass sich 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 Modelle erwiesen, wenn es um das Klassifizieren von Bilddaten geht. Sie werden intensiv beim Erstellen von Modellen für die menschliche Gesichtserkennung eingesetzt. Die so gewonnenen Erfahrungen sind grundlegend für die optische Klassifizierung und Identifizieren von Tieren. 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 wie z.B. unterschiedliche Aufnahmewinkel, oder variierende Lichtverhältnisse. 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 Menschen und Maschinen leicht zu erkennen und somit leichter verwertbar aufzunehmen als beispielsweise ein definierter Bereich der Flanke, oder eines Euters, der z.B. beim Venenscannen notwendig ist. Die Auflösung der Bilder ist dennoch nicht irrelevant. Je

größer die Pixelanzahl der 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 wichtiger Faktor, der aufgrund des geringen Alters der Technologie bislang noch keinen Eingang in die Literatur gefunden hat, ist dabei die Standardisierung. Soll eine rein optische Identifikation zukünftig Marktreife erreichen, so muss man sich mit der Möglichkeit auseinandersetzen, 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 basiert der Fortschritt maßgeblich auf der Verfügbarkeit von Daten. Es existieren vielfältige Datensets zu allen möglichen Themen. Datensets bestehen neben den Bildern aus Labels, anhand derer das trainierte Modell lernt. Hierfür müssen die Bilder von Tieren 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.

5 Diskussion

Bezugnehmend auf die Technologiebeschreibung ist es wichtig zu verstehen, dass leistungsfähige Modelle sukzessiv 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 Datensets. Bestehende Modelle werden weiter verbessert, sodass immer mehr Tiere damit identifiziert werden können. Dies ist ein stetiger Prozess. Nach mehreren Rechercheergebnissen 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 zu identifizierenden 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, könnte somit die Reduzierung der zu verwaltenden Tiermenge durch die Einführung der Kategorie *Betrieb* im ID-Schema bestehend aus Land-Bundesland-*Betrieb*-Tier sein. Hierdurch kann die Menge der optisch zu identifizierenden Tiere deutlich unter der Leistungsgrenze künftig absehbarer CNN-Modelle geteilt werden. Darüber hinaus muss die Genauigkeit, mit der ein einzelnes Tier identifizierbar ist, festgelegt werden. Hierzu ist eine rechtliche Grundlage notwendig, die für alle Systeme einzuhalten ist und für den Anwendungsfall der Tieridentifikation mit ausschließlich optisch-biometrischen Mitteln eine Genauigkeit festlegt, die ein System erfüllen muss. Rechtssicherheit bieten Werte nur bei einer Genauigkeit von über 99%. Grundlegend sind für das Training von Modellen große Datensets notwendig. Diese sind

ausschlaggebend für die Verbesserung der Identifizierung. Datasets bestehen nicht nur aus Bilddateien, sondern auch aus den jeweiligen Labels, d.h. den zu jedem Bild vorhandenen Informationen, die es dem Modell schlussendlich ermöglichen maschinell 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 dabei durch die Ausschöpfung von Parallelität und Synergieeffekten erreicht.

6 Schlussfolgerung

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

7 Literaturverzeichnis

- [BABR01] Banko, M.; Brill, E.: and Eric Brill: Scaling to Very Very Large Corpora for Natural Language Disambiguation. Microsoft Research, 2001, dl.acm.org/doi/pdf/10.3115/1073012.1073017.
- [De19] Deißing, F.; Dorsch, U.; Huhne, T.; Schneider, P.; Pache, S.; Fischer, R.: With 3D image analysis system "cow body scan" - keeping animal welfare of dairy cows under control. Animal Health Report November/2019, 30-31, 2019.
- [Pa13] Pache, S.; Hölscher, R.; Steinhöfel, I.; Fischer, R.: Automatische Vermessung der Körperkondition von Milchkühen mit 3-D-Bildverarbeitungstechnik. In (KTBL): 11. Tagung: Bau, Technik und Umwelt 2013 in der landwirtschaftlichen Nutztierhaltung. KTBL, Darmstadt, 275-280, 2013.