

Einzeltieridentifikation über biometrische Bildanalyse mittels künstlicher Intelligenz

Franziska Deißing¹, 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. Im Rahmen zweier Machbarkeitsstudien wurde untersucht, inwiefern eine eindeutige Einzeltieridentifikation über Bildanalyse prinzipiell möglich ist, welche Tierindividuellen Merkmale Unterschiede zwischen Individuen zeigen und wie sich der Forschungsstand darstellt, sowie die aktuelle Verfügbarkeit von technischen Möglichkeiten zur Tieridentifikation. Weiterhin konnten mögliche Lösungsansätze zum grundsätzlichen Vorgehen im Bereich optischer Tieridentifikation entwickelt werden.

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 ein ganzes Leben lang. Kennzeichnung und Registrierung von Rindern ist nach der Viehverkehrsverordnung (ViehVerkV vom 06.07.2007, zuletzt geändert 03.05.2016) des Bundesministeriums für Ernährung und Landwirtschaft in Deutschland genau definiert.

Diese Art von Kennzeichnung bringt neben dem Nutzen aber auch einige Nachteile mit sich. Die für die obligatorische Kennzeichnung nach ViehVerkV zugelassenen elektronischen Ohrmarken arbeiten im Low Frequenzbereich entsprechend der DIN ISO 11785. Für eine zuverlässige Tieridentifikation müssen die Tiere vereinzelt den Reader in einem Abstand von 40-80 cm passieren. Aufgrund der relativ langen Lesezyklen sollten die Tiere langsam, mit nicht mehr als 3 m/s am Reader vorbeigeführt werden. Zur

¹ Landesamt für Umwelt, Landwirtschaft und Geologie, Abteilung 7 Landwirtschaft, Am Park 3, 04886 Arzberg, franziska.deissing@smekul.sachsen.de, <https://www.landwirtschaft.sachsen.de/digitalisierung-tierhaltung-51670.html>

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

Vermeidung der Lesung benachbarter Tiere sind die Readerstationen entsprechend abzuschirmen und vor der Einwirkung anderer Magnetfelder zu schützen.

Nach dem sich in den letzten Jahrzehnten die Biometrie als Wissenschaft zur Messung und Auswertung von Merkmalen an Lebewesen etablieren konnte, gewinnen diese lang bekannten Methoden heute durch das maschinelle Lernen und die verfügbaren Rechenkapazitäten in einer Cloud oder offline rasant an Bedeutung. Wie alle Lebewesen besitzen Rinder individuelle Merkmale, die für eine biometrische Identifizierung herangezogen werden können, sofern deren Ausprägung eine ausreichende Varianz und damit Differenzierung erlaubt. Als Beispiele könnten hier die Musterung des Felles, die Aderung von Körperregionen, Körperproportionen, Haarwirbel und Farbe genannt werden. Die Aufgabe für die Zukunft liegt in der Entwicklung eines geeigneten Erkennungssystems, das den Umweltbedingungen in der Tierhaltung, der Tiergesundheit und der Mobilität der Tiere Rechnung trägt. Zielstellung des Vorhabens ist das Aufzeigen der prinzipiellen Möglichkeit der eindeutigen Einzeltieridentifizierung über biometrische Bildgebungsverfahren und die Verfahrensbeschreibung aus aktuellen Forschungsinitiativen.

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, das Projekt Cow-body-scan (Pache et al) und das Projektpaper zum Themenkomplex optische Tieranalyse (Klück, Deissing 2019) zentrale Ansätze. Alle drei Voruntersuchungen beschäftigen sich mit bildgebenden Verfahren in der Rinderhaltung und mit der Möglichkeit einzelne Individuen von anderen zu unterscheiden. Herauszustellen ist dabei das nicht nur die Unterscheidung der Einzeltiere im Fokus der diskutierten Fragestellung steht, sondern die tatsächliche eindeutige Identifikation der jeweiligen Individuen.

2.2 Motivationstreiber der technischen Umsetzung

Die treibende Entwicklungsarbeit im Bereich der optischen Tieridentifikation ist länderspezifisch individuell ausgeprägt. Ein Projekt der University of Kentucky, USA, forscht an Tierortung und Identifikation mittels Drohnenbildgebung. 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 der Gesichtserkennung 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 Hauptmotivation in der Anwendung von KI zur Fälschungsprävention. Diese Initiative geht aus der Versicherungsbranche hervor. Der Versicherungsbetrug und gefälschte Tierdaten sind weit verbreitete Problemstellungen. Projekte im europäischen Raum setzen

bei der Produktentwicklung auf den Bereich Systemanbindung im Herdenmanagement. Dabei sind aktive Initiativen hauptsächlich in Großbritannien angesiedelt. Das augenscheinlich größte Potential hat dabei das Projekt Cattleeye. Herausragend ist zudem, dass eher kleinere und mittelständige Unternehmen im Bereich Tieridentifikation forschen. Die gute Verfügbarkeit und weltweit anerkannte Tierkennzeichnung durch Ohrmarken scheint keinen Konkurrenzdruck zu erzeugen.

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 speziellen Themenbereich abgebildet.

3.1 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. Um aus den Ergebnissen der automatisierten Suche die relevanten Ergebnisse herauszufinden, wurde im Anschluss eine teil-manuelle Sichtung durchgeführt.

3.2 Marktrecherche

Das Ziel der Marktrecherche war nicht nur das Aufzeigen fertiger Produktlösungen, es kann aufgezeigt werden, welche Projekte am Thema arbeiten, unabhängig ihrer kommerziellen Verwertung, oder des Reifegrades der Technologie. Die Suche wurde zweistufig durchgeführt. Zunächst wurden alle dem LfULG bekannten Unternehmen und Verbände via Email kontaktiert, darauf folgte eine intensive Internetrecherche und Kontaktaufnahme zu den identifizierten Vorhaben.

3.3 Technologiebeschreibung

Der Begriff Künstliche Intelligenz beschreibt keine Technologie, sondern vielfältige Konzepte. Für verschiedene Anwendungsfälle und Anforderungen werden unterschiedlichste Methoden eingesetzt. 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)

4 Ergebnisse

Die aufgezeigten Recherchemethoden und die daraus resultierten Erkenntnisse können im Folgenden näher betrachtet werden. Hierbei werden vor allem die wichtigsten Eckpunkte festgehalten, um diese Schwerpunkte im Folgenden näher zu vertiefen.

4.1 Patent- und Marktrecherche

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. 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 und gerade in China weiterhin eine gesteigerte Aktivität erwartet wird. Zum anderen veröffentlichten Projekte im universitären Umfeld nicht umfangreich vor Abschluss der Forschungsarbeit, wie am Markt agierende Unternehmen. 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. Eine Unterscheidung ist allerdings vorerst bei 800 Tieren in der Praxisanwendung erfolgt.

4.2 Technologiebeschreibung

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: [WIKIDAS21]. 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.

5 Diskussion

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 immer weiter verbessert, sodass immer mehr Tiere damit identifiziert werden können. Dies ist ein stetiger Prozess, um die identifizierbare Tieranzahl fortlaufend steigern zu können. 400-500 Tiere zuverlässig zu unterscheiden scheint derzeit nach Herstellerangaben stabil umsetzbar 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. Aktuell kann kein Modell zuverlässig 1.000 Tiere identifizieren. Die Tieridentifikation mittels Ohrmarke ermöglicht es 10 000 000 Tiere zweifelsfrei zuzuordnen. Die Möglichkeiten die optische Identifizierung zu nutzen kann durch eine Reduzierung der Tiermenge erfolgen. 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 bearbeitet werden können. 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. Ein niederschwellig verfügbares Dataset würde die Entwicklung neuer Modelle stark begünstigen. Davon würde die Forschung in diesem Bereich Elementar profitieren und Hemmnisse für kleine Forschungsgruppen beseitigt werden. Die Nutzung von Parallelität und Synergieeffekten fördern aktiv die Innovation.

6 Schlussfolgerung

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 kann die Möglichkeit der Tieridentifikation zur Verknüpfung mit Technologien bei der Weideüberwachung genutzt werden, wie beispielsweise die Erkennung von Einzeltieren über Drohnenflug. Dabei ist die größte Chance zur Einsatzfähigkeit und einwandfreien Zuordnung optischer Daten durch immer höhere Verarbeitungskapazität in kürzester Zeit zu erreichen. Ein weiterer positiver Innovationstreiber ist die kontinuierlich verbesserte Kameratechnik, die zur Verfügung steht. Die Kamera sieht mehr als der Mensch sich vorstellen kann, somit verbessern sich maßgeblich die Chancen der Tierkontrolle in allen Bereichen.