Programm zur Innovationsförderung des BMEL

Bewertung der Tiergerechtheit und des Tierwohls in der landwirtschaftlichen Nutztierhaltung unter Einsatz geeigneter Indikatoren

Autowohl

Automatisierte Erfassung von Tierwohlindikatoren bei Geflügel

Abschlussbericht

Arbeitspakete 2.2 - 2.4

Ansprechpartner: Jenny Stracke, Nicole Kemper, Birgit Spindler

Institut für Tierhygiene, Tierschutz und Nutztierethologie (ITTN) der Stiftung Tierärztliche Hochschule Hannover

Bischofsholer Damm 15 (Gebäude 116)

30173 Hannover

T: +49 (0)511 856 8960

F: +49 (0)511 856 8998

Mail: jenny.stracke@tiho-hannover.de

Inhaltsverzeichnis

I. Kurzdarstellung 4
I.1 Aufgabenstellung 4
I.2 Voraussetzungen, unter denen das Vorhaben durchgeführt wurde 4
I.3 Planung und Ablauf des Vorhabens 5
I.4 Wissenschaftlicher und technischer Stand an den angeknüpft wurde 7
I.5 Zusammenarbeit mit anderen Stellen 8
II. Eingehende Darstellung 10
II.1 Verwendung der Zuwendungen und des erzielten Ergebnisses einschließlich Vergleich mit den vorgegebenen Projektzielen 10
II.1.1 AP 2.2 Fußballenveränderungen (FPD) am Schlachthof bei Puten 10
II.1.1.1 Beobachterabgleich 10
II.1.1.2 Validierung: Vergleich Kamerasystem vs. adspektorische Begutachtung 11
II.1.1.3 Validierung: Vergleich Kamerasystem vs. visuelle Beurteilung auf Herdenbasis (am Band) 12
II.1.1.4 Optimierung: Systematische Bewertung der Fehler und Optimierung des Algorithmus 13
II.1.1.5 Validierung: Histologie 15
II.1.1.6 Standardisierung 16
II.1.2 Arbeitspaket 2.3 Brusthautveränderungen bei der Pute 16
II.1.2.1 Beobachterabgleich 16
II.1.2.2 Validierung: Erhebung der Prävalenz und Abgleich der automatischen Erkennung mit einer visuelle Beurteilung (am Band) 17
II.1.2.3 Optimierung: Systematische Bewertung der Fehler und Optimierung des Algorithmus 18
II.1.2.4 Alternative Erkennung Brustblasen Probelauf NIR Kamerasystem 19
II.1.3 AP 2.4 Kratzer am Schlachthof bei Puten 20
II.1.3.1 Beobachterabgleich 20
II.1.3.2 Prävalenzerhebung 21
II.1.3.3 Kamerabilder 21
II.2 Wichtigste Positionen des zahlenmäßigen Nachweises 22
II.3 Notwendigkeit und Angemessenheit der geleisteten Arbeit 22
II.4 Voraussichtlicher Nutzen, insbesondere der Verwertbarkeit des Ergebnisses im Sinne 22
des fortgeschriebenen Verwertungsplans 22
II.4 Darstellung des während der Durchführung des Vorhabens dem ZE bekannt gewordenen Fortschritts auf dem Gebiet des Vorhabens bei anderen Stellen 23
II.6 Darstellung der erfolgten oder geplanten Veröffentlichungen des Ergebnisses nach Nr. 6 23
III. Erfolgskontrollbericht 24
III.1 Beitrag des Ergebnisses zu den förderpolitischen Zielen 24
III.2 Wissenschaftlich-technische Ergebnisse, erreichte Nebenergebnisse, gesammelte 24
wesentliche Erfahrungen 24

III.3 Fortschreibung des Verwertungsplans 25
III.4 Arbeiten, die zu keiner Lösung geführt haben 26
III.5 Präsentationsmöglichkeiten für mögliche Nutzer 26
III.6 Einhaltung der Kosten- und Zeitplanung 26
IV. Kurzfassung 27
Verwendete Literatur: 28
Anhang 1: Merkblatt zum Beurteilerabgleich 29
Anhang 2: Rahmenpapier allgemeine Vorgehensweise für Veröffentlichungen 30
Anhang 3: Handlungsempfehlung „Automatisierte Erfassung Fußballenveränderung“ 31
Anhang 5: Veröffentlichungen der Stiftung Tierärztliche Hochschule im Rahmen des Projektes 32

I.Kurzdarstellung

I.1 Aufgabenstellung

Dieses Verbundprojekt diente der Entwicklung verbesserter und neuer Messsysteme zur automatisierten bildgestützten Erfassung von Tierwohlindikatoren am Schlachthof und im Bestand bei Legehennen, Masthühnern und Puten und wurde im Rahmen der Innovationsförderung des BMEL gefördert. Beteiligt waren vier wissenschaftlichen Partner (Stiftung Tierärztliche Hochschule Hannover (TiHo), Ludwig-Maximilians-Universität München (LMU), Hochschule Osnabrück (HS OS), Universität Kassel) sowie vier Unternehmen der vor- und nachgelagerten Produktion (CLK GmbH Bildverarbeitung und Robotik, Barnstorfer Geflügelschlachterei Dietrich Menge GmbH, Heidemark GmbH und Brüterei Süd ZN der BWE-Brüterei Weser-Ems GmbH & Co. KG). Aufgeteilt in acht Arbeitspaketen (AP), sollten bereits bestehende Messsysteme standardisiert, optimiert und validiert, sowie neue Systeme entwickelt werden. Ziel des Vorhabens war es, das Potential einer besseren Standardisierung und objektiven Befunderhebung mit dem Vorteil einer zeiteffizienten Datenerhebung weiter zu erschließen, um routinemäßige Tierwohlbeurteilungen am Schlachthof und im Geflügelbetrieb verstärkt und valide einsetzen zu können.

Von der TiHo Hannover wurden folgende Arbeitspakete bearbeitet:

- **AP 2.2 Fußballenveränderungen (FPD) am Schlachthof bei Puten:**
  In diesem AP sollte ein bereits bestehendes (automatisiertes) Bonitursystem validiert, standardisiert und weiterentwickelt werden, sowie über histologische Untersuchungen Rückschlüsse auf die Tiefe der Fußballenveränderungen getroffen werden.
- **AP 2.3 Brusthautveränderungen am Schlachthof bei Puten:**
  In diesem AP sollte ein bereits bestehendes automatisiertes System validiert, standardisiert und weiterentwickelt werden.
- **AP 2.4 Kratzer am Schlachthof bei Puten:**
  In diesem AP sollte ein automatisiertes System zur Erfassung von Kratzern und Pickverletzungen am Körper von Puten neu entwickelt werden, um in einem zweiten Schritt standardisiert und validiert zu werden.

I.2 Voraussetzungen, unter denen das Vorhaben durchgeführt wurde

Das Institut für Tierhygiene, Tierschutz und Nutztierethologie (ITTN) der Stiftung Tierärztliche Hochschule Hannover (TiHo) unter der Leitung von Frau Prof. Dr. Kemper bearbeitet grundlegende Fragestellungen zur Tierhaltung, zum Tierschutz und zur Ethologie von Nutztieren. Die Arbeitsgruppe Tierschutz im Institut unter der Leitung von Frau Dr. Spindler bearbeitet tierschutzrelevante Fragestellungen aus dem Bereich der Haltung mit der Erhebung und Bewertung tier- und ressourcenbezogener Indikatoren. Ein wesentlicher Bestandteil aller Projekte ist die Erhebung von Indikatoren zur Beurteilung der Tiergesundheit und des Tierwohls (z.B. Bonitur von Gefiederschäden und Hautverletzungen und entzündlichen Hautschäden sowie der Fußballen-Veränderungen) im Stall

und am Schlachthof. Das Projekt wurde wissenschaftlich begleitet von Frau Dr. Jenny Stracke. Sie ist als PostDoc am ITTN angestellt und bearbeitet eine Reihe von Projekten zum Thema Tierwohl, mit Schwerpunkt auf der Erhebung von Tierwohlindikatoren und Verhaltensbeobachtungen.

Alle Arbeitspakete wurden in enger Zusammenarbeit mit der Firma CLK GmbH (CLK) bearbeitet. CLK entwickelt und vertreibt Bildverarbeitungssysteme zur Fußballenkontrolle, Hockburn-Erkennung, Kratzspur-Erkennung sowie zur Brusthautkontrolle bei Masthühnern, Mastputen und Enten. Solche Systeme sind in diversen großen Geflügelschlachthöfen in Deutschland installiert. Das Know-How liegt insbesondere in der Kameratechnik (2D, 3D, Farbe) und der Entwicklung stabiler Algorithmen zur automatischen Bewertung.
Die Untersuchungen wurden in Zusammenarbeit mit dem Unternehmen Heidemark GmbH (Ahlhorn) durchgeführt. Heidemark ist einer der größten Anbieter von Geflügelprodukten (Pute). Zum Thema Tierwohl wurde u.a. das Projekt „Tierwohl geprüft" ins Leben gerufen, um Tierwohl anhand festgelegter Kriterien messbar zu machen. Im Schlachthof der Firma Heidemark GmbH sind zwei Systeme der Firma CLK verbaut: ein Kamerasystem zur Detektion von Fußballenveränderungen (Turkey Check 2.0) und ein System zur Detektion von Veränderungen der Brusthaut (kombiniertes 3D-RGB-System).

I.3 Planung und Ablauf des Vorhabens

Der geplante zeitliche Ablauf der Arbeitspakete wich bereits durch einen abweichenden Projektstart (geplant November 2016, bewilligt zum 19.04.2017) von der ursprünglichen Vorhabenbeschreibung ab, die Erreichung der im Projektplan beschriebenen Meilensteine wurde, in Abstimmung mit dem Konsortium, bei einem ersten Projekttreffen entsprechend angepasst. Unterschiedliche Meilensteine waren im Projektplan festgelegt und sollten für die unterschiedlichen Arbeitspakete gleichermaßen realisiert werden.

- Meilenstein 1: Vorarbeiten, Installation der Technik, Festlegung eines Beobachtersystems, Durchführung der Beobachterabgleiche (geplanter Abschluss bis November 2017)
- Meilenstein 2: automatisierte Erfassung und Bonitur, histopathologische Untersuchungen, Optimierung der automatischen Systeme (geplanter Abschluss Januar 2018)
- Meilenstein 3: Validierung der Ergebnisse, erste Auswertung (geplanter Abschluss Mai 2019)
- Meilensteine 4 und 5: Erarbeitung von Anleitungen zur Anwendung bzw. Schulungsunterlagen, Publikation der Ergebnisse (geplanter Abschluss bis Dezember 2019)

Die Erreichung der einzelnen Meilensteine verschob sich im Laufe der Projektbearbeitung. Insbesondere die Neuentwicklung einer Technik zur Erkennung von Kratzern und die Weiterentwicklung der Automatik zur Erkennung von Brustblasen erwiesen sich als komplexer als ursprünglich gedacht.

Für AP 2.2 (Fußballenveränderungen) wurden alle Meilensteine erreicht, Im Laufe des Projektes verschob sich allerdings die Erreichung des Meilensteins 3 deutlich nach hinten, da für die Optimierung und Validierung des Systems ein iterativer Prozess zum Einsatz kam. Bei der Entwicklung der automatischen Erkennung von Brusthautveränderungen (AP2.3) wurden die Ziele nur teilweise erreicht. Auch hier kam ein iterativer Prozess zum Einsatz, allerdings wurde nach wiederholter Validierung und Optimierung der Schluss gezogen, dass das verwendete System keine verwertbaren reliablen Ergebnisse zur Erkennung von Brustblasen liefern kann. Parallel wurde an Alternativmethoden gearbeitet. Daher wurde der Meilenstein 4 nur teilweise umgesetzt (ursprünglich geplante Schulungsunterlagen konnten nicht erstellt werden). Für AP 2.4 (Erkennung von Kratzern) wurden die Meilensteine 1-3 und 5 inhaltlich bearbeitet, das Ziel, ein validiertes System in die Anwendung zu überführen und entsprechende Schulungsunterlagen zu erarbeiten, wurde für diesen Punkt nicht erfüllt, da die Entwicklung dieser neuen Technik in der Projektlaufzeit nicht abgeschlossen werden konnte.

Näheres zu den erreichten Zielen inklusive einer Gegenüberstellung mit den ursprünglich im Projektplan vorgesehenen Zielen sind unter Punkt II zu finden.

Des Weiteren war im Projektplan die Durchführung verschiedener Wissenschaftlertreffen (6x) vorgesehen. Diese wurden über die Jahre verteilt durchgeführt.

- Wissenschaftlertreffen 1 (30.08.2017) in Ahlhorn mit der Hochschule Osnabrück und der Heidemark GmbH: Besprechung des Scoringsystems für die Erhebung der Fußballenveränderungen und der bekannten Problematiken bei der automatischen Erfassung.
- Wissenschaftlertreffen 2 (20.09.2017) zwischen der CLK GmbH und der Heidemark GmbH in Ahlhorn: Präzision der Zielsetzung und des Vorgehens.
- Wissenschaftlertreffen 3 (20.11.2017) an der LMU München: Besprechung der Fußballenproblematik und Durchführung des Beobachterabgleichs im Arbeitspaket 2.1.
- Wissenschaftlertreffen 4 (23.08.2018) mit der Universität Kassel in Kassel: Besprechung des generellen Vorgehens beim Beobachterabgleich und der Autorenschaft verschiedenen Veröffentlichungen. Für beides entstanden Merkblätter die in Anlage 1 und 2 aufgeführt sind.
- Wissenschaftlertreffen 5 (28.09.2018) zwischen der TiHo und der CLK in Altenberge: Diskussion des Vorgehens in den einzelnen Arbeitspaketen des Zeitplans.
- Wissenschaftlertreffen 6 (09.07.2019) mit der CLK in Altenberge: Besprechung erster Ergebnisse und des weiteren Vorgehens bei der Entwicklung der NIR-Technik.

Abweichend vom ursprünglichen Projektplan, in dem zwei Konsortium-Treffen und drei Workshops zum Wissenstransfer in die Praxis angesetzt waren, wurden (entsprechend der gestellten und bewilligten Änderungsanträge) drei Konsortium-Treffen und zwei Workshops durchgeführt.

Die Konsortium-Treffen fanden am 26.06.2017, am 25.05.2018 und am 25.06.2019 statt. Die Ergebnisse aller Arbeitspakete wurden in zwei Workshops zusammengeführt und einer ausgewählten

Zuhörerschaft aus verschiedensten Bereichen präsentiert. Die Workshops fanden am 13.02.2019 und am 21.11.2019 statt.

I.4 Wissenschaftlicher und technischer Stand an den angeknüpft wurde

Gesundheitsbezogene Tierwohlindikatoren bei Hühnern (Legehennen und Masthühner) und Puten, die am Schlachthof (potentiell) automatisiert erhoben werden können, wie entzündliche Hautveränderungen an den Füßen und ggf. Brust- sowie Hautverletzungen und Brustbeinschäden sind in allen einschlägigen Vorschlägen zu wichtigen Tierwohlindikatoren enthalten (z.B. Welfare Quality® 2009, KTBL 2014, 2015). Eine automatisierte, bildgestützte Datenerfassung bietet sich vor allem am „Flaschenhals“ Schlachthof an. Die Überwachung derartiger Tierwohlparameter kann maßgeblich dazu beitragen, die Haltung landwirtschaftlicher Nutztiere langfristig zu verbessern. Im Falle der Fußballenbewertung von Puten am Schlachthof wurde ein Bewertungssystem im Rahmen des Gesundheitskontrollprogrammes von Puten (VDP 2013) als Benchmarksystem etabliert, um so die Tierhalter für die Problematik zu sensibilisieren, das Management zu verbessern und somit eine schrittweise Verbesserung der Fußballengesundheit zu erreichen. Die Erkennung von FPD wird in den größeren Schlachtbetrieben in Deutschland bereits automatisiert vorgenommen. Im QS Leitfaden „Befunddaten in der Geflügelschlachtung“ (QS Fachgesellschaft Geflügel GmbH, Version 01.01.2019) wird die kameragestützte Bewertung von Fußballenveränderungen ab einer gewissen Schlachtkapazität (mehr als 500 Tiere/Stunde) sogar gefordert. In Deutschland ist hierfür in der Regel Technik der CLK GmbH verbaut, aber auch andere Firmen produzieren entsprechende Systeme (z.B. Meyn Food Processing Technology BV, Niederlande). Anhand zweidimensionaler Bilderfassung wird hier die Größe der veränderten Fußballenhaut des Mittelfußballens (Metatarsalballen) ermittelt (s. Abbildung 1). Die Bewertung jedes erfassten Fußes wird in einer Datenbank gespeichert, wodurch automatisch eine Statistik für den Mäster erstellt wird.

Die Bewertung des Schweregrades solcher Veränderungen wird in der Regel über Boniturnoten kategorisiert. Um Scoring-Systeme am Schlachthof einsetzen zu können, müssen diese schnell und einfach anwendbar sein; die zugrundeliegende Problematik soll dennoch sinnhaft und in mess- und bewertbaren Kategorien widergespiegelt werden. Damit solche Systeme auch dem wirtschaftlichen Wettbewerb standhalten können, sind Objektivität und Validität entscheidende zusätzliche Kriterien (Lund et al., 2017; van Harn und de Jong, 2017). Für die Verwendung zur betrieblichen Eigenkontrolle, zu Zertifizierungszwecken, durch Kontrollbehörden oder für ein nationales Monitoring, aber auch für die Nutzung zu wissenschaftlichen Zwecken, ist eine Vergleichbarkeit der erhobenen Daten ausschlaggebend. Eine Studie von Vanderhasselt et al. (2013) wies auf Schwierigkeiten in der Übereinstimmung zwischen einem automatischen System zur Erfassung von Fußballenveränderungen und der manuellen Bonitur bei Masthühnern (Firma Meyn) hin; vorherige Ergebnisse von de Jong et al. (2008) wurden dadurch bestätigt. Wissenschaftliche Ergebnisse zu den in Deutschland verwendeten Systemen der Firma CLK bei Puten fehlen bisher. Auch Fragen zur Reliabilität der Erhebung und Bewertung der Befunde sind unbeantwortet.

Abbildung 1: Erfassungsbereich/Referenzbereich (Mittelfußballen) in grün; durch FPD verfärbter Bereich in rot (© CLK GmbH)

Als ein weiterer Tierschutzindikator bei Puten können Brustblasen und Breast Buttons herangezogen werden. Allain et al. (2013) zeigten eine Prävalenz von Brusthautveränderungen von 30 % bei männlichen Tieren zum Zeitpunkt der Schlachtung. Brusthautveränderungen sind somit aus der Sicht des Tierschutzes, aber auch im Hinblick auf mögliche wirtschaftliche Einbußen von großem Interesse. Auch hierzu gibt es Bemühungen der standardisierten Erfassung. Seit Januar 2020 sind QS-zertifizierte Betriebe dazu angehalten, Brusthautveränderungen in drei Klassen zu erheben und den Anteil der betroffenen Tiere je nach Klasse in die entsprechende Datenbank einzugeben (QS, 2019). Bisher werden diese Daten allerdings manuell erhoben.

Kratz- und Pickverletzungen bei der Pute werden bisher noch nicht standardmäßig erfasst. Auch gibt es bisher nur wenig wissenschaftliche Belege über die Höhe der Prävalenzen in den Beständen. Eine Studie von Allain et al. (2013) wies nur geringe Prävalenzen (unter 2%) nach allerdings wurden in dieser Studie ausschließlich männliche Tiere untersucht. In der Praxis mehren sich Hinweise, dass es bei der Art der Verletzungen Geschlechtsunterschiede geben könnte.

I.5 Zusammenarbeit mit anderen Stellen

Das Projekt wurde in enger Zusammenarbeit mit allen beteiligten Projektpartnern durchgeführt. In den bearbeiteten Arbeitspaketen wurde in erster Linie mit der Firma CLK (Erarbeitung, Optimierung der Algorithmen und der Technik) und der Heidemark GmbH (Bereitstellung der Schlachttiere, fachliche Unterstützung) zusammengearbeitet.

Tabelle 1: Beobachterabgleiche in den einzelnen Arbeitspaketen

Arbeitspaket	**Hauptverantwortliche Einrichtung**	**Beobachterabgleich**
Fußballen- und Fersenhöcker – Veränderungen am Schlachthof bei Masthühnern	LMU München	Tierärztliche Hochschule Hannover

Fußballenveränderungen am Schlachthof bei Puten	Tierärztliche Hochschule Hannover	Hochschule Osnabrück
Brusthautveränderungen am Schlachthof bei Puten	Tierärztliche Hochschule Hannover	Heidemark GmbH
Kratzer am Schlachthof bei Puten	Tierärztliche Hochschule Hannover	Universität Kassel
Gefiederschäden Pute im Bestand	Universität Kassel	Tierärztliche Hochschule Hannover

Aber auch zwischen den Arbeitspaketen fand ein reger Austausch statt. So wurden die durchzuführenden Beobachterabgleiche in Zusammenarbeit mit allen Beteiligten wissenschaftlichen Einrichtungen durchgeführt (s. Tabelle 1 für die entsprechenden Kooperationen der TiHo).

II. Eingehende Darstellung

Die diesem Teilprojekt zugeordneten Arbeitspakete werden im Folgenden näher erläutert. Alle Zuwendungen wurden entsprechend des Antrags bzw. der gestellten und bewilligten Änderungsanträge im Rahmen der Projektarbeiten verwendet (Einzelheiten siehe Verwendungsnachweis).

II.1 Verwendung der Zuwendungen und des erzielten Ergebnisses einschließlich Vergleich mit den vorgegebenen Projektzielen

II.1.1 AP 2.2 Fußballenveränderungen (FPD) am Schlachthof bei Puten

Als Grundlage für die Bewertung von FPD bei Puten wurde ein 5-stufiger Score (nach Hocking et al. 2008) genutzt (Abbildung 2). Dies ist ein visuelles Bewertungssystem, das den Schweregrad einer Veränderung anhand der Größe der farblich veränderten Areale auf dem Mittelfußballen (Metatarsalballen) bewertet. Dies Bewertungssystem wird auch in der Praxis üblicherweise als Scoring-System herangezogen.

Abbildung 2.: Fußballenveränderungen nach HOCKING et al. (2008): Score 0 = intakter Fußballen, Score 1 = Veränderungen betreffen <10 % des Fußballens, Score 2 = Veränderungen betreffen <25 % des Fußballens, Score 3 = Veränderungen betreffen <50 % des Fußballens, Score 4 = Veränderungen betreffen >50 % des Fußballens (©Spindler)

II.1.1.1 Beobachterabgleich

Der Beobachterabgleich zur Beurteilung der Fußballen der Puten am Schlachthof wurde in Zusammenarbeit mit der Hochschule Osnabrück und der Firma Heidemark GmbH durchgeführt. Eine Durchführung am laufenden Schlachtband war nicht möglich, da sich eine Erfassung derselben Füße im Tempo des laufenden Betriebes als nicht praktikabel erwies. Daher wurden die Füße (n = 400 Fußpaare (200 Hennen/ 200 Hähne) am Ende der Schlachtbahn vom Haken entnommen. Dort wurden sie von vier Beobachtern (2x Hochschule Osnabrück; 2x Tierärztliche Hochschule) bewertet. Zusätzlich wurde von jedem der Fußpaare ein digitales Foto erstellt. Auch die Klassifizierung anhand der digitalen Fotos wurde über einen Beobachterabgleich (3 Personen, TiHo) abgeglichen.

Als statistisches Maß wurde der Prevalence-Adjusted Bias-Adjusted Kappa (PABAK) berechnet (s. Anlage 1).

Ergebnisse:

Für den Abgleich der Bewertung der FPD anhand der entnommenen Füße ergab sich ein Wert von 0,7 ein Abgleich der Bewertung der digitalen Bilder resultierte in einem PABAK von 0,74. Somit wurde bei beiden Verfahren eine gute Übereinstimmung zwischen den Beobachtern festgestellt.

II.1.1.2 Validierung: Vergleich Kamerasystem vs. adspektorische Begutachtung

Für einen Abgleich der Klassifizierung der Füße durch das Kamerasystem und der Einteilung durch einen visuellen Beobachter wurden Füße am Anfang des Schlachtbandes mit Kabelbinder versehen. Die markierten Füße (n = 1993) durchliefen dann das Kamerasystem und wurden am Ende des Schlachtbandes entnommen. Dort wurden die einzelnen Füße adspektorisch begutachtet, zusätzlich wurden digitale Fotos erstellt.

Ergebnisse:

Von 1993 bewerteten Füßen wurde bei 50,6 % der Füße eine perfekte Übereinstimmung erreicht (gleiche Scorestufe). Bei 43,3 % wich die Beurteilung um eine Note voneinander ab (akzeptable Übereinstimmung). Bei 6,1 % der Füße wurde eine Abweichung von mehr als einer Note festgestellt. Das automatische System bewertete die Füße insgesamt besser als ein manueller Beobachter (s. Abbildung 3; Tabelle 2).

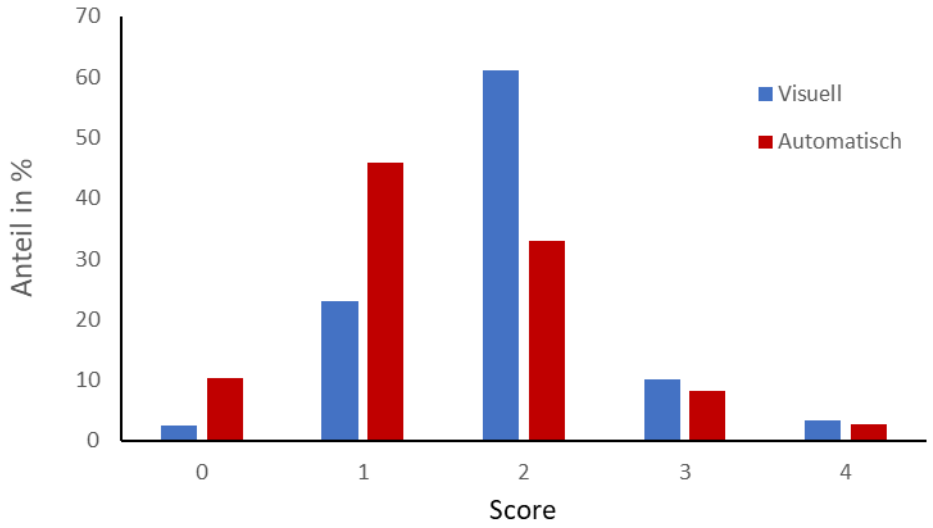


Abbildung 3: Vergleich automatische vs. visuelle Bewertung; blau: visuelle Bewertung: rot: automatische Bewertung: n=1993 Füße

Tabelle 2: Kreuztabelle der Übereinstimmung zwischen automatischer und manueller/visueller Klassifizierung der Fußballenveränderungen. In grün dargestellt die Anzahl der perfekten (dunkelgrün) und der akzeptablen (hellgrün;

Abweichung um nur eine Stufe) Übereinstimmungen, in rot dargestellt Anzahl der Abweichungen um mehr als eine Note.

Projektstart (2017)	Visuell					
Automatisch	Score 0	Score 1	Score 2	Score 3	Score 4	Gesamt
Score 0	33	96	72	3	1	205
Score 1	18	331	535	27	3	914
Score 2	0	30	534	81	11	656
Score 3	0	1	74	75	14	164
Score 4	0	2	2	14	36	54
Gesamt	51	460	1217	200	65	**1993**

II.1.1.3 Validierung: Vergleich Kamerasystem vs. visuelle Beurteilung auf Herdenbasis (am Band)

Im Falle eines Kameraausfalls, ist eine Vergleichbarkeit einer manuellen Beurteilung am Schlachtband mit einem automatisierten System auf Herdenbasis notwendig.

Um dies zu untersuchen, wurden verteilt über die Schlachtpartie linke und rechte Füße / Herde am Band visuell bonitiert. Ziel war es jeweils 200 Fußpaare verteilt auf Anfang, Mitte und Ende der Schlachtpartie zu beurteilen, aufgrund der Geschwindigkeit des Schlachtbandes und der unterschiedlichen Herdengrößen, konnte diese Zielstellung allerdings nicht präzise eingehalten werden. Zeitgleich wurde jeder Fuß von dem automatischen System erfasst. Der Algorithmus bewertete zu diesem Zeitpunkt jeweils den linken Fuß / Fußpaar. Konnte dieser nicht detektiert werden, wurde der jeweils rechte Fuß verwendet. Insgesamt wurden 44 Herden (520.020 automatisch und 18.620 manuell bewertete Füße) in die Auswertung eingeschlossen.

Ergebnisse:

Die automatische Bewertung der Füße fiel positiver aus, als die visuelle Bewertung. Es wurden deutlich mehr Füße in die Scorestufen 0 und 1 einsortiert, wogegen die visuelle Bewertung mehr Füße in die Scorestufen 2, 3 und 4 einstufte (Abbildung 4).

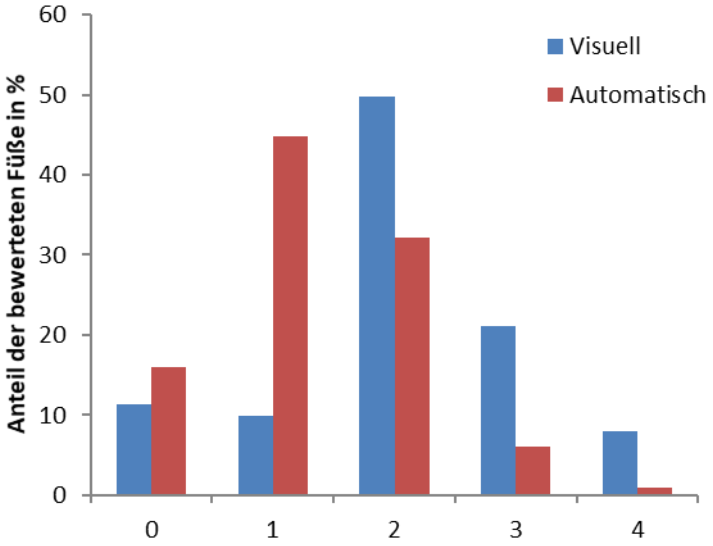


Abbildung 4: Vergleich automatische vs. visuelle Bewertung; blau: visuelle Bewertung: n=18.620 Füße; rot: automatische Bewertung: n: 520.020 Füße

II.1.1.4 Optimierung: Systematische Bewertung der Fehler und Optimierung des Algorithmus

Um die Ursachen der Unterschiede zwischen automatischer Erfassung und visueller Bonitur systematisch zu erfassen, wurden die Kamerabilder nach folgenden Kriterien bewertet:

- Referenzbereich: (gut / zu groß oder zu klein (Abweichungen um mehr als ½ Zehenglied / ganz falsch erkannt (Abweichungen um mehr als 1 Zehenglied) / verschoben)
- Veränderungen (erkannt / zu viel erkannt / zu wenig erkannt)
- Andere mögliche Fehlerquellen (Fuß schief / Verschmutzung / etc.)

Nach einer zunächst subjektiven Identifikation der Hauptfehlerquellen wurde der Algorithmus durch die CLK GmbH modifiziert. Nach einer ersten Modifikation (Mod1) wurden 3110 Bilder systematisch auf oben beschriebene Fehlerquellen untersucht. Ausgehend von diesen Ergebnissen wurde der Algorithmus ein zweites Mal angepasst (Mod2). Zusätzlich wurde der Algorithmus für beide Füße eines Fußpaares ausgegeben, um eine zeitgleiche Erfassung linker und rechter Füße zu gewährleisten und hier systematische Unterschiede zu untersuchen. Diese Ergebnisse wurden erneut anhand von 2788 Bildern validiert.

Ergebnisse:

In den meisten Fällen war die Fehlerquelle ein falsch detektierter Referenzbereich. Als zweite Ursache stellte sich heraus, dass teilweise die Veränderungen nicht oder falsch (zu klein / zu groß) erkannt wurden, insbesondere in der Scorestufe 1 (punktuelle Veränderungen, < 10 % des Fußballens betroffen) wurden Veränderungen teilweise nicht detektiert. Hier war auffällig, dass viele dieser Veränderungen farblich leicht von denen in den höheren Scorestufen abwichen (hellere Braunfärbung). Diese Fehlerquelle konnte aber in Mod2 durch Einstellungen der Farberkennung

minimiert werden. Eine weitere Fehlerquelle mit der Folge einer falschen Erkennung war die Aufhängung der Füße. Teilweise hingen die Füße schief in der Aufhängung und wurden dadurch perspektivisch falsch erfasst. Im untersuchten System war hier vorwiegend der linke Fuß betroffen. Dies liegt vermutlich an der Schnitttechnik während des Schlachtprozesses zum Zeitpunkt der Abtrennung des Tierkörpers von den beiden noch in den Haken hängenden Füßen. Die Ergebnisse für Mod1 sind in Tabelle 3 zu finden, die Ergebnisse für Mod2 sind in Tabelle 4 erfasst.

Tabelle 3: Automatisierte Erfassung der Fußballen und der veränderten Flächen nach Optimierung des Algorithmus (Mod1); n=3110 Füße, Ausfallrate: 14,7 %, mehrere Fehlerquellen/Fuß möglich.

Keine Fehler	26,8 %
Fehler	58,5 %
Referenzbereich falsch detektiert	48,2 %
Veränderung falsch detektiert	28,3 %
Referenz + Veränderung falsch detektiert	21,4 %
Schiefe Einhängung	44,7 %

Tabelle 4: Automatisierte Erfassung der Fußballen und der veränderten Flächen nach Optimierung des Algorithmus (Mod2); n=2788 Füße (1394 linke Füße, 1394 rechte Füße), Ausfallrate 10,5 % (davon: 32,2 % rechte Füße, 67,8 % linke Füße, mehrere Fehlerquellen/Fuß möglich).

Keine Fehler	**31,7 %**	Linker Fuß	26,5 %
		Rechter Fuß	73,4 %
Fehler	57,8%	Linker Fuß	59,6 %
		Rechter Fuß	40,4 %
Referenzbereich falsch detektiert	46,6 %	Linker Fuß	60,8 %
		Rechter Fuß	39,2 %
Veränderung falsch detektiert	24,4 %	Linker Fuß	60,1 %
		Rechter Fuß	39,9 %
Referenz + Veränderung falsch detektiert	17,1 %	Linker Fuß	70,6 %
		Rechter Fuß	29,4 %
Schiefe Einhängung	41,6 %	Linker Fuß	73,1 %
		Rechter Fuß	26,9 %

Die Fehlerrate wurde deutlich minimiert, allerdings besteht noch weiterer Optimierungsbedarf, insbesondere bei der Erkennung des Referenzbereiches. Dieser wird in der Regel als zu groß angenommen (44,8 % der bewerteten Füße).

Der optimierte Algorithmus zur automatischen Bewertung der Füße wurde zudem erneut mit einer visuellen Bonitur verglichen. Auch nach der Optimierung fiel die automatische Bewertung der Füße positiver aus, als die visuelle Bewertung. Es wurden mehr Füße in die Scorestufen 0 und 1 einsortiert, wogegen die visuelle Bewertung mehr Füße in die Scorestufen 2 und 4 einstufte (Abbildung 5). Nichtsdestotrotz verbesserte sich die Übereinstimmung zwischen beiden Bewertungssystemen deutlich (s. Abbildung 5).

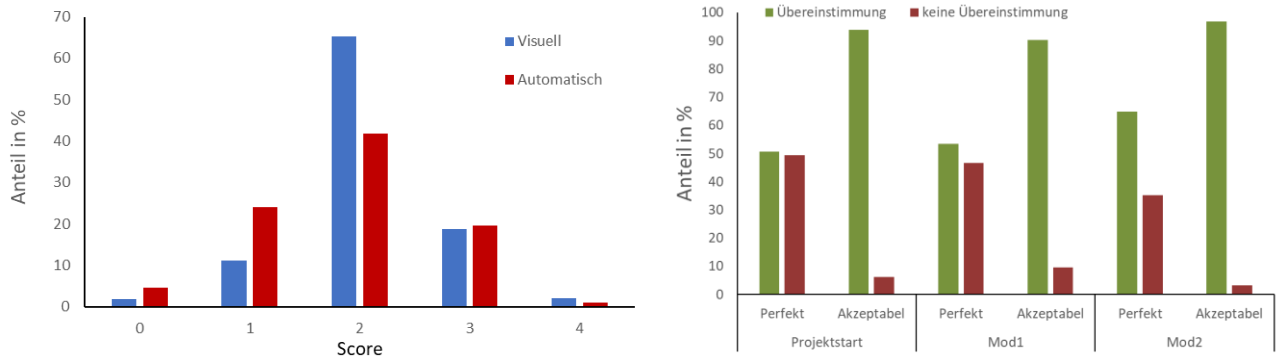


Abbildung 5: Vergleich automatische vs. visuelle Bewertung**; links:** Validierung nach Modifikation 2; blau: visuelle Bewertung; rot: automatische Bewertung: n: 2788 Füße; Anteil nicht bewerteter Füße (nicht sichtbar/verdreckt/etc.): visuell: 1,15%, automatisch: 9,07%; **rechts:** Übereinstimmung zwischen automatischer und visueller Klassifizierung in den unterschiedlichen Stadien des Optimierungsprozesses (perfekte Übereinstimmung = gleiche Bewertung durch visuelle und automatische Beurteilung, akzeptable Übereinstimmung = Abweichung um eine Boniturnote; in grün dargestellt die Übereinstimmungen in Prozent, in rot dargestellt die Abweichungen in Prozent; Projektstart = 2017, n=1993; Mod1 = Modifikation 1, Anfang 2019, n=3110; Mod2 = Modifikation 2, Projektende 2019, n=2788.

Nichtsdestotrotz wurden Abweichungen zwischen einer visuellen Bonitur und der automatischen Beurteilung deutlich, insbesondere auch bei Füßen bei denen keinerlei Fehler in der automatischen Erfassung auftraten, d.h. sowohl der Referenzbereich, als auch die Veränderung richtig detektiert wurden. Letzteres wird auf eine unterschiedliche (falsche) Wahrnehmung des visuellen Beobachters zurückgeführt.

II.1.1.5 Validierung: Histologie

Für histologische Untersuchungen wurde eine Stichprobe aus oben beschriebenen Füßen (II.1.1.2) verwendet, wobei jeweils der rechte Fuß beprobt wurde (n=100; 20 pro Scorestufe). Die Größen der jeweiligen Fußballen und der veränderten Fläche wurden mit einem Grafikprogramm (ImageJ) vermessen, die veränderte Fläche wurde dann in Relation zum Fußballen gesetzt. Anschließend wurden histologische Schnitte angefertigt, wobei mittig der Läsionen geschnitten wurde. Die Schnitte wurden in Formalin fixiert und anschließend zur weiteren Analyse an das Institut für Pathologie (Stiftung Tierärztliche Hochschule) gebracht. Dort wurden unterschiedliche Parameter erfasst (Hyperkeratose, Erosionen, Ulzerationen, Granulationsgewebe (Narbengewebe) und multifokale perivaskuläre lymphozytäre Pododermatits (Hinweis auf einen chronischen Entzündungsprozess (Abbildung 6). Alle Parameter wurden zunächst in ihrer binomialen Ausprägung erfasst (Auftreten ja / nein), zudem wurden für die Parameter Hyperkeratose, Ulzeration und multifokale, lymphozytäre, perivaskuläre Pododermatitis der Schweregrad der Ausprägung bestimmt (geringgradig / mittelgradig / hochgradig). Als Zusammenhangsmaß wurde der Korrelationskoeffizient nach Spearman berechnet.

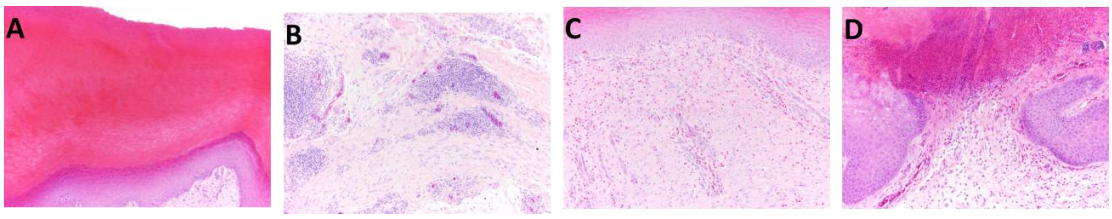


Abbildung 6: Histopathologische Schnitte: A=geringgradige Hyperkeratose, B=perivaskuläre Pododermatitis; C=Granulationsgewebe; D= Ulzeration

Ergebnisse:

Die Korrelation zwischen der Größe der Läsion und dem visuellen Score war hoch (rsp = 0,9), wobei kein linearer Anstieg der Schweregrade der jeweiligen histologischen Parameter mit ansteigender Scorestufe der makroskopischen Befundung festgestellt wurde. Hyperkeratose wurde in allen Scorestufen nachgewiesen (Score 0-3: zu 100%; Score 4: 90%), wobei hochgradige Hyperkeratosen vorwiegend in Scorestufe 2 auftraten. Erosionen wurden nicht nachgewiesen. Ulzerationen traten bereits ab der Scorestufe 1 auf. Hier waren bereits 55% der Füße von einer Ulzeration betroffen, in den Scorestufen 2-4 wiesen alle Füße Ulzerationen auf, wobei diese vorwiegend als hochgradig eingeschätzt wurden. Narbengewebe trat in allen Scorestufen auf (0/1: 20-35%, 2/3/4: 60-65%). Auch die multifokalen, lymphozytären, perivaskulären Pododermatitis wurde in allen Scorestufen nachgewiesen (0/1: 20-43,4%, 2/3/4: 62,5-73,8%).

II.1.1.6 Standardisierung

Im Rahmen des Projektes wurde eine Handlungsempfehlung erstellt, welches insbesondere auf die Relevanz der zu treffenden Einstellungen (Licht/Säuberung/Kallibierung), notwendiger Voruntersuchungen (Erfassung des geeigneteren Fußes, Vorgehen bei Beobachterabgleichen) und Maßnahmen zur Qualitätskontrolle (stichprobenartige Kontrolle der Erfassung) eingeht (Anlage 3).

Zusammenfassung der erzielten Ergebnisse und Gegenüberstellung zu den vorgegebenen Projektzielen AP.2.2:

Die Ziele des Arbeitspaketes wurden in der Projektlaufzeit erreicht. Aus den vorgestellten Ergebnissen ergibt sich weiterer Optimierungsbedarf hinsichtlich der Erkennung des Referenzbereiches (Abgrenzung Mittelfußballen) als auch zur Festlegung von geeigneten (standardisierten) Schwellenwerten.

II.1.2 Arbeitspaket 2.3 Brusthautveränderungen bei der Pute

II.1.2.1 Beobachterabgleich

Für die Bewertung der Brusthautveränderungen wurde ein 4-stufiger Score genutzt, angelehnt an das, an dem untersuchten Standort, übliche Protokoll (s. Tabelle 5). Anders als im ursprünglichen

Protokoll vorgesehen wurden Breast Buttons (nekrotische Veränderungen der Brusthaut) separat erfasst, hier wurde eine subjektive Einteilung in normal, leicht vergrößert und hochgradig vergrößert vorgenommen. Buttons konnten zusätzlich zu einer Brustblase auftreten.

Tabelle 5: Bewertungssystem Brusthautveränderungen (ausschließlich Brustblasen)

Score 0		**Intakte Brust**
Score 1	Ø >3cm	Kleine geringgradige Umfangsvermehrungen
Score 2	Ø <3 cm; Tiefe >1 cm < 3cm	Mittelgradige Umfangsvermehrungen mit geringen Farbabweichungen und geschlossener Haut
Score 3	Ø > 5 cm; Tiefe > 3cm	Hochgradige Umfangsvermehrungen mit Farbabweichungen und teils offener Haut

Der Beobachterabgleich wurde mit drei Personen durchgeführt. Hierfür wurden Tiere am Schlachtband markiert und bewertet.

Ergebnisse:

Für den Beobachterabgleich wurden 98 Tiere begutachtet. Es ergab sich ein PABAK von 0,73, der als gut gewertet werden kann.

II.1.2.2 Validierung: Erhebung der Prävalenz und Abgleich der automatischen Erkennung mit einer visuelle Beurteilung (am Band)

Insgesamt wurden 28 Herden (Hennen und Hähne) am Band bewertet. Ziel war es jeweils 200 Tiere verteilt auf Anfang, Mitte und Ende der Schlachtpartie zu beurteilen. Aufgrund der Geschwindigkeit des Schlachtbandes und der unterschiedlichen Herdengrößen, konnte diese Zielstellung allerdings nicht eingehalten werden. Insgesamt gehen daher 5095 Tiere in die Auswertung ein. Für die Bewertung wurde oben genannter Score verwendet. Zeitgleich wurden alle Tiere der entsprechenden Herden von einem automatisierten System erfasst. Hierbei handelt es sich um ein bildgebendes Kamerasystem das sowohl Farb- als auch 3-D-Erkennung kombiniert und bereits im Schlachthof der Firma Heidemark GmbH installiert war.

Die jeweiligen Durchschnittswerte (auf Herdenbasis) wurden miteinander verglichen.

Ergebnisse:

Es konnte eine Prävalenz von Brusthautveränderungen von insgesamt 15,4% festgestellt werden (Score 1: 5,0 %, Score 2: 1,6 %, Score 3: 0,9 %, Breast Buttons 7,9 %). Hähne waren im Vergleich zu Hennen deutlich häufiger betroffen (Hahn vs. Henne: Score 1: 5,4 % vs. 4,3 %; Score 2: 1,7 % vs. 1,3 %; Score 3: 1,2 % vs. 0,2 %; Button: 10,0 % vs. 3,4 %). Beim Vergleich der visuellen Bonitur mit dem automatischen System zeigte sich, dass die Kamera die Veränderungen deutlich positiver einstufte als ein visueller Beobachter.

II.1.2.3 Optimierung: Systematische Bewertung der Fehler und Optimierung des Algorithmus

Nach einer ersten systematischen Fehlersuche stellte sich heraus, dass die Bildqualität nicht den Anforderungen entsprach. Dies lag vermutlich an einer fehlenden Ausleuchtung des Systems. Nach Optimierung der Ausleuchtung wurde erneut die Qualität der automatisierten Klassifizierung bewertet. Hier stellte sich heraus, dass insbesondere die Brustblasen schlecht erkannt wurden, Breast Buttons wiesen eine etwas bessere Sensitivität auf (Brustblasen: Score 0: 0,74; Score 1: 0; Score 2: 0,07; Score 3: 0,08; Button: 0,37). Nach einer Optimierung des Algorithmus wurde erneut eine systematische Qualitätskontrolle vorgenommen. Hierfür wurden insgesamt 9426 Bilder der automatisierten Erfassung visuell bewertet, Abweichungen zur automatischen Erfassung wurden vermerkt. Auch nach der Optimierung des Algorithmus wurde keine deutliche Verbesserung der automatischen Erkennung von Brustblasen erreicht (s. Tabelle. 6).

Tabelle 6: Übereinstimmung der automatischen und der visuellen Klassifizierung von Brusthautveränderungen (n.b.=nicht bewertet).

	Visuell							
Automatisch	Score 0	Score 1	Score 2	Score 3		Button	n.b.	Gesamt
Score 0	5939	22	8	2		512	5	6488
Score 1	0	0	0	0		0	0	0
Score 2	1	0	3	0		0	0	4
Score 3	0	0	1	4		0	0	5
Button	464	7	4	2		464	0	941
n.b.	1660	24	28	14		262	0	1988
Gesamt	8064	53	44	22		1238	5	**9426**

Von 9426 bewerteten Brüsten wurden 67,9 % der Brüste übereinstimmend bewertet (perfekte Übereinstimmung), wobei die fehlerhafte Bewertung durch das automatisierte System nicht als Hauptproblem hervortrat. Größer wog die Problematik, dass ein Großteil der betroffenen Tiere aus der Bewertung ausgeschlossen wurde (nicht bewertet (n.b.) Brustblasen: 55,5 %, Breast Buttons: 21,2 %). Dies liegt unter anderem an Einstellungen des Algorithmus, die eine Detektion von falsch positiven Befunden minimieren soll. Bei den niedrigen Prävalenzen, die beim Auftreten von Brusthautveränderungen vorlagen, führt dieses Vorgehen allerdings zu fehlerhaften Ergebnissen. Zudem gestaltet sich insbesondere die Detektion von Tiefenveränderungen durch die 3-D Kamera als schwierig. Weitere Optimierungen zur Erkennung von Brustblasen mit dem derzeitigen System erscheinen daher als nicht zielführend.

Um die Erkennung von Breast Buttons (Erkennung vorwiegend basierend auf Detektion von Farbveränderungen) zu optimieren entwickelte CLK einen zweiten Algorithmus. Die erstellten Bilder durchlaufen hier zwei Zyklen; werden sie in einem ersten Bewertungszyklus als "nicht zu bewerten" aussortiert, durchlaufen sie einen zweiten, verfeinerten Durchgang. Ein erster Testdurchlauf zeigte,

dass bei einem zweiten Durchlaufen des Algorithmus von zehn zunächst als nicht bewertbar klassifizierten Bildern, bei sechs Bildern das Auftreten eines Brustknopfes richtig erkannt wurde.

II.1.2.4 Alternative Erkennung Brustblasen Probelauf NIR Kamerasystem

Eine Alternative zur 3-D Kamera könnte die Aufnahme der Veränderungen mit einer Nah-Infra-Rot Kamera darstellen. Diese misst Veränderungen von Wellenlängen, abhängig von der stofflichen Zusammensetzung verschiedener Strukturen.

Für einen ersten Probelauf wurde ein entsprechender Aufbau (NIR-Kamerasystem, CLK) am Schlachthof Heidemark aufgebaut. Durch einen erfahrenen Tierarzt (Dr. Günther, Heidemark) wurde eine Auswahl an Putenkörpern (n= 30) getroffen, um eine möglichst große Bandbreite an Brusthautveränderungen zu untersuchen. Die Puten wurden händisch im Aufbau eingehängt und entsprechende Bilder mit der NIR-Kamera erstellt. Zusätzlich fand eine visuelle Bonitur statt, wobei nach oben beschriebenem Scoring vorgegangen wurde. Außerdem wurde jede Brusthaut angeschnitten und der Inhalt durch einen Tierarzt (Dr. Günther) pathologisch beurteilt.
Diese pathologische Beurteilung von unterschiedlichen Zuständen (z.B. seröse Flüssigkeit, Blut, Fibrin) wurden in Zusammenarbeit mit einem Mitarbeiter der CLK mit den unterschiedlichen, von der NIR-Kamera erkannten, Wellenlängen verknüpft, um hier erste Ansätze zu entwickeln unterschiedliche Zustände voneinander zu unterscheiden und von einer gesunden Brust abzugrenzen (Abbildung 7).

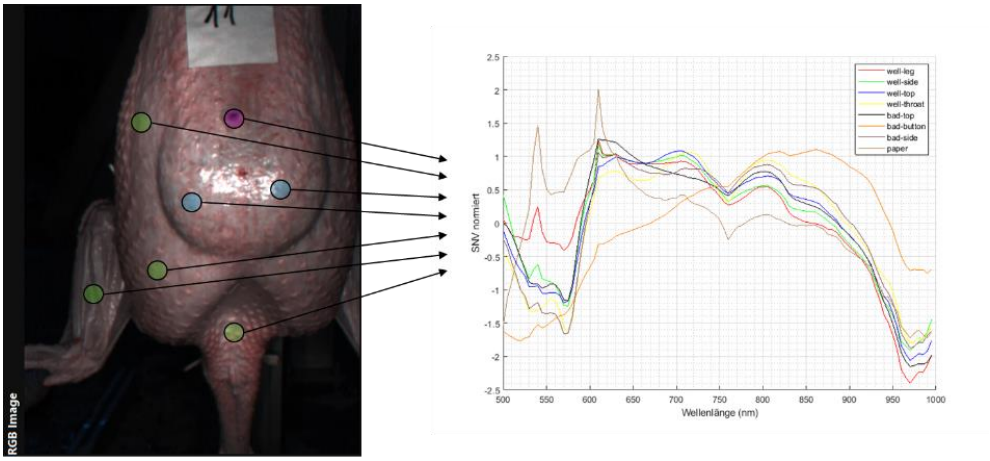


Abbildung 7: Ergebnisse der Probeaufnahmen (NIR-Kamera) ©CLK GmbH

Um eine Anwendung in der Praxis zu ermöglichen, wäre die Entwicklung eines geeigneten Algorithmus nötig, um diese Wellenlängen unterschiedlichen Scorestufen zuordnen zu können.

Zusammenfassung der erzielten Ergebnisse und Gegenüberstellung zu den vorgegebenen Projektzielen AP.2.3:

Das Ziel ein automatisiertes System zur Erkennung von Brusthautveränderungen zu validieren, zu standardisieren und für die Verwendung in der Praxis weiterzuentwickeln. wurde nicht vollumfänglich erreicht, da das untersuchte System nach Validierung und Standardisierung keine verwertbaren, reliablen Ergebnisse zur Erkennung von Brustblasen lieferte. Für die Erkennung von Brustknöpfen wurden erste Modifizierungen vorgenommen, hier besteht allerdings weiterer Optimierungsbedarf, um eine valide Anwendung in der Praxis zu gewährleisten. Im Rahmen des Projektes wurden zudem erste Alternativmethoden entwickelt, diese befinden sich zum jetzigen Zeitpunkt noch in der weiteren Entwicklung.

II.1.3 AP 2.4 Kratzer am Schlachthof bei Puten

Nach Absprache mit der Firma CLK und der Heidemark GmbH wurde sich darauf geeinigt in einem ersten Schritt zunächst Kratzverletzungen am Rücken zu erfassen. Dies ist der am besten zugängliche Bereich in der Schlachtkette und erscheint somit für einen Kameraaufbau am geeignetsten.

II.1.3.1 Beobachterabgleich

Für die Beurteilung der Kratzer wurde ein Score aus dem KTBL-Leitfaden modifiziert und ein Beobachterabgleich mit zwei Beobachtern durchgeführt. Es ergab sich ein PABAK von 0,69, der als gut bewertet werden kann.

Das verwendete Scoringsystem erwies sich allerdings nach der Erstellung erster Probebilder als nicht geeignet, um daraus einen geeigneten Algorithmus für die automatische Detektion zu entwickeln. Daher wurde der Beurteilungsrahmen in Zusammenarbeit mit der Firma CLK weitergehend modifiziert (Tabelle 7).

Tabelle 7: Bewertungssystem Kratzer

	Altes Bewertungssystem	**Modifiziertes Bewertungssystem**
Score 0	Intakter Rücken	Intakter Rücken
Score 1	Kratzer < 2cm	oberflächliche Kratzer die nicht in die Muskulatur gehen (hell/blass/blutig) < 8cm Länge
Score 2	Kratzer 2-8 cm	oberflächliche Kratzer die nicht in die Muskulatur gehen >8cm oder Verletzungen die in die Muskulatur reichen mit einer Länge <2cm und Breite <1cm
Score 3	Kratzer >8 cm oder flächige Verletzungen Ø< 5cm	in die Muskulatur reichende Verletzungen >2cm und/oder Plattenbildung (flächige Verletzungen) mit einem Durchmesser <5cm gewertet
Score 4	Flächige Verletzungen Ø > 5cm	Plattenbildungen >5cm

Ein erneuter Beobachterabgleich wurde in Zusammenarbeit mit der Universität Kassel durchgeführt. Insgesamt wurde die Übereinstimmung zwischen drei Beobachtern untersucht. Es wurde ein PABAK von 0,60-0,64 berechnet.

II.1.3.2 Prävalenzerhebung

Die Kratzverletzungen wurden am Band bewertet. Auch hier war es das Ziel jeweils 200 Tiere verteilt auf Anfang, Mitte und Ende der Schlachtpartie zu beurteilen. Aufgrund der Geschwindigkeit des Schlachtbandes und der unterschiedlichen Herdengrößen, konnte diese Zielstellung allerdings nicht eingehalten werden. Insgesamt wurden 15 Herden (2653 Tiere; Hahn und Henne) ausgewertet.

Ergebnisse:

56,5% der Tiere wiesen keine Kratzverletzungen auf. 22,8% der Tiere wiesen Kratzverletzungen der Stufe 1, 15,8% der Tiere Kratzverletzungen der Stufe 2 und 4,2% der Tiere Kratzverletzungen der Stufe 3 auf. Die Stufe 4 wurde bei 0,6% der Tiere gefunden. Es konnten keine Geschlechtsunterschiede festgestellt werden.

II.1.3.3 Kamerabilder

In Zusammenarbeit mit der CLK wurde ein Testaufbau im Schlachthaus installiert. Als automatisches System wurde ein bildverarbeitendes Kamerasystem genutzt, welches bisher Anwendung in Entenschlachthöfen findet. Es wurde eine Reihe von Testbildern erstellt, und diese anschließend von der CLK bearbeitet. Eine Auswahl von 5070 Bildern wurde an die TiHo versandt und dort visuell bewertet. Die hier klassifizierten Tiere sollten als Grundlage für die Entwicklung eines geeigneten Algorithmus für die automatische Klassifizierung dienen. Die Entwicklung des Algorithmus zur Klassifizierung der zu detektierenden Hautschäden befindet sich derzeit in Arbeit. Erste Ansätze sind bereits vorhanden, allerdings gestaltet sich insbesondere die Differenzierung von Hautverfärbungen (z.B. Hämatomen) und von maschinell entstandenen Schäden noch als schwierig.

Zusammenfassung der erzielten Ergebnisse und Gegenüberstellung zu den vorgegebenen Projektzielen AP.2.4:

Das Ziel ein automatisiertes System zur Erfassung von Kratzern und Pickverletzungen am Körper von Puten neu entwickeln, wurde nur teilweise erreicht, da sich die Entwicklung eines geeigneten Algorithmus als deutlich komplexer erwies, als ursprünglich gedacht. Insbesondere die Unterscheidung von tatsächlichen Verletzungen von durch maschinell entstandene Schäden am

Tierkörper, aber auch die Differenzierung von tatsächlichen Verletzungen von Hautverfärbungen und Ähnlichem erwies sich als schwierig und wurde bis zum Projektende nicht gelöst.

II.2 Wichtigste Positionen des zahlenmäßigen Nachweises

Die wichtigsten Positionen des zahlenmäßigen Nachweises umfassten die Personalkosten für die Projektmitarbeiterin sowie die wissenschaftliche Hilfskraft (Pathologie), Kosten für die Dienstreisen im Inland und Ausland (für die Datenerfassung sowie diverse Kongressteilnahmen), Kosten für die histopathologische Auswertung sowie Kosten für die durchgeführten Workshops (Einzelheiten siehe Zahlungsnachweis).

II.3 Notwendigkeit und Angemessenheit der geleisteten Arbeit

Die geleisteten Arbeiten waren zum Erreichen der beantragten Arbeitsziele erforderlich und in dem Umfang unter Berücksichtigung der Grundsätze des sparsamen Haushaltens notwendig.

II.4 Voraussichtlicher Nutzen, insbesondere der Verwertbarkeit des Ergebnisses im Sinne des fortgeschriebenen Verwertungsplans

Die Ergebnisse tragen dazu bei, zukünftig valide und zuverlässig repräsentative Daten zu erheben. Insbesondere im Hinblick auf aktuelle Tierwohldebatten können die Ergebnisse aus AP 2.2 genutzt werden, um die automatisierte Erfassung von FPD bei Puten zu standardisieren und die Qualität der Erfassung weiter zu optimieren. Insbesondere die Ergebnisse aus den histopathologischen Untersuchungen bieten sich als Diskussionsgrundlage an, um geeignete Grenzwerte festzulegen. Dies wäre sowohl für die behördliche Überwachung, als auch für ein mögliches nationales Monitoring von Vorteil. Nicht zuletzt hat die Geflügelwirtschaft selbst ein großes Interesse, im Sinne der betrieblichen Eigenkontrolle diese Entwicklungen für Verbesserungen in der Geflügelhaltung, mehr Transparenz bezüglich des Tierwohls und eine Verbesserung der Wettbewerbssituation der deutschen Geflügelwirtschaft im internationalen Vergleich zu nutzen.

Die im Rahmen von Arbeitspaket 2.2 erstellten Handlungsempfehlungen, ermöglichen es den beteiligten Unternehmen (Schlachthof, QS), Informationen zum System und zur Qualitätskontrolle abzurufen.

Auch wenn das finale Ziel in AP 2.3 und 2.4 innerhalb der ursprünglichen Projektlaufzeit nicht verwirklicht wurde, so ergaben sich dennoch erste wertvolle Ergebnisse, an denen angeknüpft werden kann:

- Die Erfassung von Brustknöpfen wurde optimiert. Weitere Optimierungsmaßnahmen und eine entsprechende Validierung sind notwendig. Nichtsdestotrotz erscheint die Erfassung der Brustknöpfe mit dem derzeitigen Prototyp (3-D/RGB Kamera) als möglich und vielversprechend.
- Für die Erfassung von Brustblasen ist der derzeitige Prototyp nicht geeignet, allerdings erscheint ein weiterer Ausbau der NIR-Technik als vielversprechend.
- Die ermittelten Prävalenzen für Brusthautveränderungen können als Diskussionsgrundlage für die Notwendigkeit einer automatischen vs. einer manuellen Erfassung herangezogen werden.
- Die ermittelten Prävalenzen für Kratzverletzungen weisen darauf hin, dass Kratzer als wichtige Tierschutzindikatoren herangezogen werden sollten (ggfs. schon während der Haltung).

Denkbar ist, noch ausstehende Fragestellungen in nachfolgenden Projekten weiter zu bearbeiten.
Alle Ergebnisse können zudem in der Lehre, für nationale und internationale Publikationen, Vorträge (s. Veröffentlichungen) und für weitergehende Untersuchungen zur Tierhaltung und zum Tierschutz genutzt werden.

II.4 Darstellung des während der Durchführung des Vorhabens dem ZE bekannt gewordenen Fortschritts auf dem Gebiet des Vorhabens bei anderen Stellen

Bisher sind keine Ergebnisse von Dritten bekannt. In Belgien soll dieses Jahr (2020) ein Projekt zur Überwachung von Tierwohl-Indikatoren beginnen, dieses wird auch eine automatische Detektion von Tierwohlindikatoren beinhalten.

II.6 Darstellung der erfolgten oder geplanten Veröffentlichungen des Ergebnisses nach Nr. 6

Im Rahmen des Projektes wurden zwei Workshops durchgeführt. Hier waren jeweils Teilnehmer aus unterschiedlichen Bereichen (Wirtschaft, Wissenschaft, Verbraucher) geladen (33 Teilnehmer Workshop 1 und 42 Teilnehmer Workshop 2).

Die Ergebnisse wurden außerdem auf Fachtagungen vorgestellt und in wissenschaftlichen Zeitschriften veröffentlicht. Die bereits erfolgten Veröffentlichungen der Stiftung Tierärztliche Hochschule sind in Anhang 5 chronologisch aufgeführt.

III. Erfolgskontrollbericht

III.1 Beitrag des Ergebnisses zu den förderpolitischen Zielen

Das geplante Projekt nimmt Bezug auf die Ziele des Programms zur Innovationsförderung entsprechend der Richtlinie zur Bewertung der Tiergerechtheit und des Tierwohls in der landwirtschaftlichen Nutztierhaltung unter Einsatz geeigneter Indikatoren, insbesondere auf die Fördergegenstände

- Validierung bestehender Indikatoren und ihrer Anwendung
- Erarbeitung von Methodenbeschreibungen und Hilfsmitteln zum praktischen Einsatz von Indikatoren
- Entwicklung verbesserter und neuer Messsysteme zur automatisierten Erhebung von Tierwohlindikatoren

III.2 Wissenschaftlich-technische Ergebnisse, erreichte Nebenergebnisse, gesammelte wesentliche Erfahrungen

Im Arbeitspaket 1.2 (Fussballenerkennung von Puten am Schlachthof) konnte ein bereits bestehendes automatisiertes Bonitursysteme zur Erfassung von Fußballenveränderungen (FPD) validiert, standardisiert und weiterentwickelt werden. Über zusätzliche histologische Untersuchungen konnten Rückschlüsse auf die histopathologischen Grundlagen der untersuchten Fußballenveränderungen getroffen werden. Diese können genutzt werden, um die automatisierte Erfassung von FPD bei Puten weiter zu standardisieren und die Qualität der Erfassung zu optimieren. Die Ergebnisse können unter anderem als Diskussionsgrundlage genutzt werden, um geeignete Grenzwerte festzulegen. Aus den Ergebnissen ergaben sich außerdem weiterführende Frage- und Aufgabenstellungen; insbesondere ist weiterer Optimierungsbedarf hinsichtlich der Erkennung des Referenzbereiches und zur Festlegung von geeigneten (standardisierten) Schwellenwerten vorhanden. Außerdem wurden erste Untersuchungen zu Zehenveränderungen durchgeführt – diese sind derzeit nicht Bestandteil der praxisüblichen Erfassung von FPD, könnten aber zu einer besseren Einordnung dieses Tierwohlindikators beitragen. Eine in diesem Arbeitspaket erarbeitete Handlungsempfehlung steht online zur Verfügung und kann genutzt werden um zusätzliche Informationen zum System und zur Qualitätskontrolle abzurufen. Diese sind, zusammen mit den in den anderen Arbeitspaketen erarbeiteten Handlungsempfehlungen und Merkblättern, langfristig in digitaler Form verfügbar (als pdf-Dateien zum Download auf der Website der Universität Kassel: http://www.uni-kassel.de/go/autowohl).

Im Arbeitspaket 1.3 (Automatische Erkennung von Brusthautveränderungen von Puten am Schlachthof) sollte ein bereits bestehendes automatisiertes System zur Erkennung von Brusthautveränderungen validiert, standardisiert und für die Verwendung in der Praxis weiterentwickelt werden. Zunächst wurden die Prävalenzen für das Auftreten von Brusthautveränderungen am untersuchten Betriebsstandort ermittelt. Nach einer Validierung und

Optimierung des derzeit existierenden Prototyps wurde festgestellt, dass dieser nicht ausreichend geeignet ist, um in der Praxis valide Ergebnisse für die Erkennung von Brustblasen zu liefern. Für die Erkennung von Brustknöpfen konnten erste Optimierungen vorgenommen werden, hier besteht allerdings weiterer Optimierungsbedarf, um eine valide Anwendung in der Praxis zu gewährleisten. Im Rahmen des Projektes wurde zudem eine Alternativmethode zur Detektion von Brustblasen entwickelt (Anwendung einer Nah-Infra-Rot-Kamera), diese befinden sich zum jetzigen Zeitpunkt allerdings noch in der weiteren Entwicklung.

Im Arbeitspaket 1.4 (Automatische Erkennung von Kratzverletzungen bei Puten am Schlachthof) wurde ein geeignetes Bonitursystem zur Detektion von Kratzverletzungen am Rücken von Puten zum Zeitpunkt der Schlachtung entwickelt und über einen Beobachterabgleich validiert. Die Entwicklung eines Algorithmus für eine automatische bildgestützte Detektion wurde angestoßen und befindet sich derzeit in der Weiterentwicklung.

III.3 Fortschreibung des Verwertungsplans

FKZ:	2817903515	
Projekttitel:	Automatisierte Erfassung von Tierwohlindikatoren bei Geflügel (Autowohl); Arbeitspakete 2.2-2.4	
1. Angestrebte Erfindungen/Schutzrechtsanmeldungen und bereits erteilte Schutzrechte		
Lfd. Nr.	Konkrete Verwertung	Zeithorizont
1.1	Es bestehen keine Schutzrechtsanmeldungen und erteilten Schutzrechte	Projektende
2. Wirtschaftliche Erfolgsaussichten nach Projektende		
Lfd. Nr.	Konkrete Verwertung	Zeithorizont
2.1	Es bestehen keine wirtschaftlichen Verwendungsmöglichkeiten für den Projektpartner ITTN. Hierfür verantwortlicher Projektpartner: CLK	
3. Wissenschaftliche und/oder technische Erfolgsaussichten nach Projektende		
Lfd. Nr.	Konkrete Verwertung	Zeithorizont
3.1	Durch den hohen Praxisbezug als auch durch die durchgeführten Workshops steigt die Sichtbarkeit des ITTN in der landwirtschaftlichen Praxis, der Geflügelwirtschaft und in entsprechenden Arbeitsgruppen.	unbefristet
3.2	Es entstand ein reger wissenschaftlicher Austausch mit anderen Arbeitsgruppen auf nationaler und internationaler Ebene, sowohl innerhalb des Konsortiums aber auch außerhalb der beteiligten Institutionen.	-Projektende
3.3	Die Ergebnisse wurden auf nationalen und internationalen Fachtagungen vorgestellt und diskutiert (ausführliche Darstellung in Anhang 5)	-Projektende
3.4	Die Ergebnisse der Untersuchungen werden intensiv im Rahmen der universitären Ausbildung und der Fortbildung von Tierärzten und Tierhaltern verwendet.	unbefristet
4. Wissenschaftliche und wirtschaftliche Anschlussfähigkeit		
Lfd. Nr.	Nächste Phase / nächste Schritte	Zeithorizont

4.1	Die Ergebnisse des Projektes (insbesondere in Arbeitspaket 2.3 und 2.4) machen deutlich, dass weitere Anpassungen nötig sind, um eine Marktfähigkeit zu realisieren. Auf der Basis der gewonnenen Projektergebnisse sollen weitere Projektideen entwickelt und realisiert werden.	

III.4 Arbeiten, die zu keiner Lösung geführt haben

Der überwiegende Teil der Arbeiten konnte wie geplant durchgeführt werden. In Arbeitspaket 2.3 konnte das Projektziel nicht vollumfänglich erreicht werden. Nach einer aufwendigen Validierung und Optimierung des derzeit bestehenden Prototyps kamen die beteiligten Projektpartner zu dem Schluss, dass das derzeitige System keine verwertbaren und reliablen Ergebnisse zur Detektion von Brustblasen liefern kann. Es erscheint aber als vielversprechend zur Detektion von Brustknöpfen, hier sind die Arbeiten zur Optimierung des Algorithmus allerdings noch nicht final abgeschlossen. Erste Ansätze mit einer alternativen Technik konnten bereits umgesetzt werden, eine Weiterentwicklung zum Einsatz in der Praxis, erwies sich allerdings als wesentlich komplexer als ursprünglich angenommen.

Auch für das Arbeitspaket 2.4 konnte das Ziel eines marktfähigen automatischen Bewertungssystems bisher nicht erreicht werden. Die zu detektierenden Verletzungen sind sehr variabel (sowohl in der Größe als auch in der Farbigkeit), die Entwicklung eines geeigneten Algorithmus auf der Basis einer Farberkennung ist auch hier wesentlich zeitaufwendiger, als ursprünglich angenommen. Hinzu kommt die Schwierigkeit der Erfassung im laufenden Schlachtprozess. Erste Ansätze wurden im Projekt entwickelt, ein funktionierender Algorithmus befindet sich allerdings noch in Entwicklung, eine entsprechende Validierung konnte bis zum Projektende nicht vorgenommen werden.

III.5 Präsentationsmöglichkeiten für mögliche Nutzer

Die modifizierten und/oder in Entwicklung befindlichen Systeme sollen auf der Eurotier 2020 präsentiert werden.

III.6 Einhaltung der Kosten- und Zeitplanung

Die ursprünglichen Kosten- und Zeitplanungen haben sich teilweise verschoben. Hier wird auf die bereits im Schlussbericht dargelegten Erläuterung verwiesen.

IV. Kurzfassung

Ein für die Erkennung von Fußballenveränderungen (FPD) bei Puten bereits bestehendes automatisiertes System wurde im Rahmen des Arbeitspaketes (AP) 2.2 zunächst auf Reliabilität untersucht. Anschließend wurde der Algorithmus des automatischen Systems modifiziert und somit für eine valide Erkennung der Veränderungen optimiert. Zudem wurde über histopathologische Untersuchungen der FPD untersucht, inwiefern die auf der Größe der Veränderungen basierende Klassifizierung auch Aufschluss geben kann über tatsächliche pathologische Befunde. Aus den erreichten Ergebnissen ergibt sich weiterer Optimierungsbedarf hinsichtlich der Abgrenzung des Mittelfußballens, als auch zur Festlegung von geeigneten Schwellenwerten der unterschiedlichen Bewertungsstufen.

Im Rahmen des AP 2.3 wurde ein bereits bestehender Prototyp zur automatischen Erkennung von Brusthautveränderungen untersucht. Nach wiederholter Optimierung des Algorithmus wurde festgestellt, dass keine verwertbaren Ergebnisse zur Erkennung von Brustblasen erzielt wurden. Erste Ansätze für Alternativmethoden wurden entwickelt. Für die Erkennung von Brustknöpfen wurden Modifizierungen vorgenommen, hier besteht weiterer Optimierungsbedarf, um eine valide Anwendung in der Praxis zu gewährleisten.

Im Rahmen des AP 2.4 sollte ein neues System zur Erkennung von Kratzverletzungen entwickelt werden. Es wurde ein geeignetes Bonitursystem entwickelt, Testbilder mit einem Prototyp wurden erstellt. Die Entwicklung eines für die Praxis geeigneten Algorithmus befindet sich in der Entwicklung.

Verwendete Literatur:

Allain, V.; Huonnic, D.; Rouina, M.; Michel, V. (2013). Prevalence of skin lesions in turkeys at slaughter. British Poultry Science 54: 33-41.

De Jong, I.; Reimert, H.; Gerritzen, M. (2008). Monitoring broiler welfare parameters at slaughter using video imaging. Final report WP 2.2, Task 2.2.14, Welfare Quality® Consortium, Lelystad, the Netherlands.

Hocking, P. M.; Mayne, R. K.; Else, R. W.; French, N. A.; Gatcliffe, J. (2008). Standard European footpad dermatitis scoring system for use in turkey processing plants. *World's Poultry Science Journal*, *64*(3), 323-328.

KTBL: Knierim, U.; Andersson, R.; Keppler, C.; Petermann, S.; Rauch, E.; Spindler, B.; Zapf, R. (2016): Tierschutzindikatoren: Leitfaden für die Praxis - Geflügel. *Kuratorium für Technik und Bauwesen in der Landwirtschaft (KTBL). Darmstadt*.

Lund, V.P.; Nielsen, L.R.; Oliveira, A.R.S.; Christensen, J.P. (2017). Evaluation of the Danish footpad lesion surveillance in conventional and organic broilers: Misclassification of scoring*. Poultry Science* 96, 2018–2028.

QS, Qualitätssicherung (2020): Leitfaden Befunddaten in der Geflügelschlachtung. file:///C:/Users/Pc/Downloads/lf_befund_g_frei_01012020_d.pdf.

Vanderhasselt, R. F.; Sprenger, M.; Duchateau, L.; Tuyttens, F. A. M. (2013). Automated assessment of footpad dermatitis in broiler chickens at the slaughter-line: Evaluation and correspondence with human expert scores. *Poultry science*, *92*(1), 12-18.

Van Harn, J. and de Jong, I. (2017). Validation of Meyn Footpad Inspection System. Wageningen *Livestock Research Report* 1044 B.

Verband Deutscher Putenerzeuger (VDP) (2013). Bundeseinheitliche Eckwerte für eine freiwillige Vereinbarung zur Haltung von Mastputen.

Welfare Quality®2009 Welfare Quality® Assessment Protocol for Poultry (Broilers, Laying Hens). Welfare Quality® Consortium: Lelystad, The Netherlands

Hannover, den

Anhang 1: Merkblatt zum Beurteilerabgleich

Autowohl: Merkblatt „Verfahrenshinweise“ zum Beurteilerabgleich

Das Bonitieren von Tieren ist eine wichtige Methode, um Tierschutzindikatoren zu erheben. Bonituren bergen allerdings auch immer die Gefahr, eine subjektive Einschätzung der Situation wiederzugeben - persönliche Erfahrungen, Erwartungen und Wahrnehmungsfähigkeiten können die Ergebnisse teils stark beeinflussen.

Neben klaren Messanweisungen und einem gutem Training der Beurteiler/Assessors stellt daher die Überprüfung der Reliabilität (Zuverlässigkeit) einen wesentlichen Bestandteil solch wissenschaftlicher Untersuchungen dar, um einerseits Messfehler so gering wie möglich zu halten und andererseits zu garantieren, dass die Untersuchungen auch von anderen Personen reproduzierbar wären.

Für die Erfassung der Reliabilität gibt es unterschiedlichste Methoden. Um die Ergebnisse im Projekt „Autowohl“ vergleichen zu können, werden in diesem Merkblatt die wichtigsten Rahmenbedingungen zur Erfassung der Beurteilerabgleiche festgehalten:

1. Mindestens 10 unabhängige[1] Samples/Score-Kategorie sind anzustreben, und wenigstens 30 Samples insgesamt sind in der Regel einzubeziehen
2. Gleichverteilung der Stichprobengröße in den einzelnen Score-Kategorien ist anzustreben
    a. Kommt eine Kategorie gar nicht oder bei einem der Beurteiler nicht vor, wird diese zwar in die Berechnung der Reliabilität mit einbezogen (da diese Kategorie ja theoretisch bei der Bewertung selbst eine Option gewesen wäre), im abschließenden Bericht (Veröffentlichung der Ergebnisse) wird dieses dann aber separat vermerkt. Sollte diese Kategorie später in der Praxis doch vorkommen, ist möglichst ein weiterer Abgleich unter Einbezug dieser Kategorie (evtl. auch nur anhand von Fotos) durchzuführen.
3. Die Reliabilität wird paarweise berechnet (jeweils zwei Beurteiler werden miteinander verglichen)
4. Als statistisches Maß wird der PABAK berechnet, verwendet und in den Ergebnisberichten angegeben (für die Berechnung s. Anlage 1).
    a. Der PABAK (*Prevalence-Adjusted Bias-Adjusted Kappa*) entspricht dem Cohen's Kappa, solange alle Kategorien gleich verteilt im Sample vorliegen. Wenn das nicht der Fall ist, unterstellt der PABAK eine gleichmäßige Verteilung der einzelnen Kategorien (wenn eine Befundkategorie nur selten auftritt, führt er daher zu höheren Werten als bspw. Cohens Kappa oder andere Koeffizienten, daher s. Punkt 2).
5. Für eine Bewertung der Reliabilität wird die empfohlene Klassifizierung von Landis und Koch (1977) (Anlage 2) verwendet:
    - <0.00 = poor
    - 0.00-0.20 = slight
    - 0.21-0.40 = fair
    - 0.41-0.60 = moderate
    - 0.61-0.8 = substantial
    - 0.81-1.00 =almost perfect

[1] Also z.B. noch von demselben Tier

Anhang 2: Rahmenpapier allgemeine Vorgehensweise für Veröffentlichungen

Autowohl: Allgemeine Vorgehensweise für Veröffentlichungen

Die Vorgehensweise bei Veröffentlichungen im Projekt Autowohl richtet sich nach den Vorgaben der DFG zur Sicherung guter wissenschaftlicher Praxis (Deutsche Forschungsgemeinschaft, 1998: Sicherung guter wissenschaftlicher Praxis, Wiley-VCH: Weinheim.), hier wird statuiert:

„Autorinnen und Autoren wissenschaftlicher Veröffentlichungen tragen die Verantwortung für deren Inhalt stets gemeinsam. Autorin oder Autor ist nur, wer einen wesentlichen Beitrag zu einer wissenschaftlichen Veröffentlichung geleistet hat."

Aus den Vorgaben der DFG für gute wissenschaftliche Praxis ergeben sich folgende Kriterien für eine Autorenschaft auf einer Veröffentlichung:
(Alle drei Punkte müssen für eine berechtigte Autorenschaft erfüllt sein.)

- Substantieller Beitrag bei Planung/Konzeption, Erhebung/Analyse/Interpretation der Daten
- Substantieller Beitrag bei dem Entwurf und der kritischen Überarbeitung einer Publikation
- Abschließende Genehmigung der Publikation in der endgültigen, zur Veröffentlichung freigegebenen Version

Folgende wesentlichen Beiträge reichen laut DFG in alleiniger Form nicht für eine Autorenschaft aus (**müssen aber in den Acknowledgments angemessen anerkannt werden**)

- Organisatorische Verantwortung für die Einwerbung von Fördermitteln
- Beistellung von Standard-Untersuchungsmaterialien,
- Unterweisung von Mitarbeiterinnen und Mitarbeitern in Standard-Methoden,
- (Technische) Mitwirkung bei der Datenerhebung,
- Beistellung von Geräten oder Versuchstieren,
- Bereitstellung von bloßen Datensätzen
- alleiniges Lesen des Manuskripts ohne substanzielle Mitgestaltung des Inhalts
- Leitung einer Institution oder Organisationseinheit, in der die Publikation entstanden ist.

Präzisierung für die Beobachterabgleiche im Projekt Autowohl:
Es ist wünschenswert, dass die gemeinsame Arbeit im Projekt durch gemeinsame Publikationen belegt wird. Voraussetzung dafür ist bei:

- Die erste Person, die beim Beobachterabgleich mitwirkt (aus der Einrichtung, die im Projektantrag als Partner aufgeführt ist):
  - Beitrag zu Planung und Erhebung der Daten ist gegeben
  - Mitarbeit an der Publikation (mindestens kritische Überarbeitung aller Manuskriptversionen in Form von Änderungs- bzw. Ergänzungsvorschlägen)
  - Genehmigung der Veröffentlichung
- Die zweite Person, die beim Beobachterabgleich mitwirkt (aus der Einrichtung, die im Projektantrag als Partner aufgeführt ist):
  - Beitrag zu Planung/Erhebung der Daten ist gegeben
  - verstärkte/substantielle Mitarbeit an der Publikation (z.B. Schreiben eines Teils der Veröffentlichung oder wesentliche und umfangreiche Überarbeitungen)
  - Genehmigung der Veröffentlichung

Anhang 3: Handlungsempfehlung „Automatisierte Erfassung Fußballenveränderung“

Automatisierte Erfassung von Tierwohlindikatoren bei Geflügel

➢ **Automatische Erfassung von Fußballenveränderungen bei der Pute am Schlachthof**

Einleitung und Hintergrund

Fußballenveränderungen (Food pad dermatitis; FPD) werden in der Putenhaltung als ein wichtiger Tierschutzindikator herangezogen. Eine Erhebung am Schlachthof kann genutzt werden, um retrospektiv Rückschlüsse auf die Haltung und das Management, unter anderem auf das Einstreumanagement in der Haltungsperiode, ziehen zu können. Im QS Leitfaden „Befunddaten in der Geflügelschlachtung“ (QS Fachgesellschaft Geflügel GmbH, Version 01.01.2019) wird ab einer gewissen Schlachtkapazität (mehr als 500 Tiere/Stunde) eine automatische Bewertung von Fußballenveränderungen gefordert. In den meisten großen deutschen Putenschlachthöfen hat sich ein automatisches Kamerasystem (Firma CLK GmbH) etabliert. Dieses Merkblatt geht insbesondere auf die Relevanz der zu treffenden Einstellungen, notwendiger Voruntersuchungen und Maßnahmen zur Qualitätskontrolle ein und soll somit dazu dienen die automatische Erhebung von FPD zu optimieren und zu standardisieren.

Automatisches System

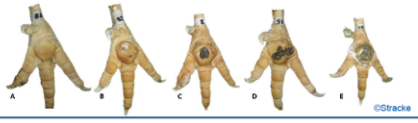
Für mehr Informationen siehe hier:

PDF Turkey Check V1.0 Handbuch

Hardware	Software
Objektiv: CLK 1/2", 6mm, FJN DF6HA-1B/1,2	TurkeyCheck (V1.0) (CLK GmbH)
Industriecomputer: CLK IPC 19" High Performance	Halcon Runtime Lizenz; MAC Adresse 00190FOA7909 (MVTec)
Beleuchtung: CLK Spot CoolWhite LED VA IP67	CLK Library Rev Nr. 5157 (CLK GmbH)
Kamera: CLK Kamera GigE 1MP SE	WinXP SP3 Embedded (Microsoft
Gehäuse: CLK Kameragehäuse VA IP67	

Mittelfußballen grün; durch FPD verfärbter Bereich in rot, © CLK

Bonitursystem



Fünf-stufiger Score nach HOCKING et al. (2008):

- Kategorie 0 = intakter Fußballen
- Kategorie 1 = Veränderungen betreffen <10 % des Fußballens
- Kategorie 2 = Veränderungen betreffen <25 % des Fußballens
- Kategorie 3 = Veränderungen betreffen <50 % des Fußballens
- Kategorie 4 = Veränderungen betreffen >50 % des Fußballens

Beobachter-abgleich

Für mehr Informationen siehe hier:

PDF Merkblatt_ Beobachterabgleich

Sollte in regelmäßigen Abständen durchgeführt werden zwischen:

- allen Personen, die eine händische Bonitur durchführen (anhand von Fotos, entnommenen Füßen oder am Band an vormarkierten Einzelfüßen)
- zwischen Kamerasystem und visueller Bonitur der entsprechenden Kamerabilder

➢ **Mindestens aber zur Einarbeitung neuer MitarbeiterInnen und nach Änderungen von Systemeinstellungen**

Unterschiedlich statistische Maßzahlen können berechnet werden:

- Prevalence-Adjusted Bias-Adjusted Kappa (PABAK)
- Cohens Kappa
- Krippendorff Alpha

Kritische Werte:

- 0,41-0,6 = akzeptabel
- 0,61-0,8 = gut
- 0,81-1,0 = sehr gut

Stichproben-größe

Für mehr Informationen siehe hier:

PDF Merkblatt Stichprobengröße

Empfohlene Stichprobengröße nach Hocking et al., 2012:

Herdengröße	Prävalenz 10 %	Prävalenz 20 %	Prävalenz 40 %
1000	122	198	270
2000	130	219	312
4000	139	246	369
8000	139	246	369
Empfohlener Standard	150	250	400

Die benötigte Stichprobengröße ist abhängig von verschiedenen Faktoren

- Prävalenz
- Herdengröße
- Gütekriterien der Methodik

➢ steigende Herdengröße = höherer Stichprobenumfang

Gütekriterien

Für mehr Informationen siehe hier:

PDF Merkblatt Datenqualität

	Scorenote	0	1	2	3	4
Mindest werte	Sensitivität	0,7	0,7	0,6	0,7	0,3
	Spezifität	0,9	0,7	0,8	0,9	0,9
	Treffergenauigkeit	0,9	0,7	0,6	0,8	0,9
	Fehlklassifikationsrate	0,1	0,2	0,3	0,1	0,1

- Für jede Score Kategorie separat zu erheben
- Standard:visuelle Bonitur; System Turkey Check V 1.0 (2019)

Die Qualität der Datenerfassung sollte in regelmäßigen Abständen kontrolliert werden (mindestens aber nach Systemänderungen)

- **Anforderungen sind betriebsintern im Vorfeld festzulegen**
  - Sensitivität / Spezifität
  - Treffergenauigkeit
  - Fehlklassifikationsrate
- **Für Dritte nachvollziehbare Dokumentation**

Qualitäts-sicherung

- Regelmäßige Kontrollen der Gütekriterien
- Regelmäßige Dokumentation aller Werte (inkl. der Ausfallraten / nicht detektierte Füße/Schlachtpartie)
- Regelmäßiger Check-Up der Einstellungen am System (Licht / Schärfe / Kontrast)
- Anpassung der Schwellenwerte an aktuelle Erkenntnisse aus der Literatur
- Sorgfältige betriebsspezifische Voreinstellungen des Systems (z.B. Auswahl des geeigneten Kamerastandortes, Auswahl des zu beurteilenden Fußes

Beeinflussende Faktoren

- Standort der Kamera
- Verschmutzung
- Federreste
- Licht
- Fehler / Defekte in der Verkabelung

➢ Informationen zur Reinigung, Wartung und Validierung der Kameraeinstellungen sind im Handbuch Turkey Check V1.0 hinterlegt

Stärken	Schwächen
• Erfassung der Fläche nekrotischer Veränderungen am Mittelfußballen • Große Stichprobenzahl pro Partie (>90 % der Füße erfassbar) • Geringer personeller Aufwand • Objektive Erfassung	• Keine eindeutige Erkennung des Mittelfußballens (in der Regel zu groß) • Tiefe nicht erfassbar • Leichte Farbabweichungen (Kategorie 1) teilweise nicht detektierbar
Chancen	**Risiken/Entwicklungsmöglichkeiten**
• Lückenlose Erfassung aller Füße, Einzeltiere/-Befunde können erfasst werden • Unabhängige, standardisierte Erfassung möglich	• Leichte Veränderungen werden vernachlässigt • Tendenz zu „besseren“ Fußballen/ Vergleichbarkeit zu visuellem Beobachter bisher nicht hinreichend gegeben • Schwellenwerte zwischen den Scorestufen frei wählbar – Ermittlung valider Schwellenwerte zur Vermeidung von z.B. Ulzerationen steht noch aus

Anhang 5: Veröffentlichungen der Stiftung Tierärztliche Hochschule im Rahmen des Projektes

Veröffentlichungen in Fachzeitschriften (peer review)

Stracke, J.; Klotz, D.; Wohlsein, P.; Döhring, S.; Volkmann, N.; Kemper, N.; Spindler, B. (2020): Scratch the surface: Histopathology of footpad dermatitis in turkeys (Meleagris gallopavo). In: Animal Welfare; accepted

Louton, H.; Piller, A.; Bergmann, S.; Erhard, M.; Stracke, J.; Spindler, B.; Kemper, N.; Schmidt, P.; Schade, B.; Boehm, B.; Kappe, E.; Bachmeier, J.; Schwarzer, A. (2020): Histologically validated scoring system for the assessment of hock burn in broilers. In: Avian Pathology, DOI: 10.1080/03079457.2020.1712328

Piller, A.; Bergmann, S.; Schwarzer, A.; Erhard, M.; Stracke, J.; Spindler, B.; Kemper, N.; Schmidt, P.; Bachmeier, J.; Schade, B.; Boehm, B.; Kappe, E.; Louton, H. (2020):Validation of histological and visual scoring systems for foot pad dermatitis in broiler chicken. In: Animal Welfare 29(2), 185-196

Beiträge auf wissenschaftlichen Tagungen

Spindler, B., Stracke, J., Kemper, N., Louton, H., Piller, A., Erhard, M., Döhring, S., Andersson, R., Gonzalez, J., Jung, L., Knierim, U. (2020): Möglichkeiten und Erfahrungen mit der automatisierten Erfassung von Tierwohlindikatoren bei Geflügel - Ein Beitrag zum Projekt AutoWohl. In: Tierschutzsymposium Oldenburg

Stracke, J.; Klotz, D.; Wohlsein, P.; Döhring, S.; Günther, R.; Spindler, B.; Kemper, N. (2019): Scratch the surface: histopathology of foot pad dermatitis in turkeys. In: 9th European Conference on Precision Livestock Farming Cork, Ireland, 2019, pp 421-427

Stracke, J.; Klotz, D.; Wohlsein, P.; Döhring, S.; Günther, R.; Kemper, N.; Spindler, B. (2019): Scratch the surface. Histopathology of foot pad dermatitis in turkeys. In: UFAW International Animal Welfare Science Symposium, 2019, pp. 24

Louton, H.; Piller, A.; Bergmann, S.; Schwarzer, A.; Stracke, J.; Spindler, B.; Kemper, N.; Schmidt, P.; Bachmeier, J.; Schade, B.; Boehm, B.; Kappe, E.; Erhard, M. (2019): Histological validation of a macroscopic visual scheme for the assessment of contact dermatitis of the hock (hock burn) of broilers. In: UFAW International Animal Welfare Science Symposium, 2019, pp.76

Stracke, J.; Klotz, D.; Wohlsein, P.; Döhring, S.; Günther, R.; Kemper, N.; Spindler, B. (2019): Scratch the surface: Histologische Untersuchungen der Fußballendermatitis (Foot Pad Dermatitis) bei Puten. In: 25. Internationale DVG-Fachtagung zum Thema Tierschutz, München 2019, pp. 333-337

Louton, H.; Piller, A.; Bergmann, S.; Schwarzer, A.; Stracke, J.; Spindler, B.; Kemper, N.; Schmidt, P.; Bachmeier, J.; Schade, B.; Boehm, B.; Kappe, E.; Erhard, M. (2019): What you see is what you get? – Histologische Validierung eines visuellen Bewertungssystems der Kontaktdermatitis am Fersenhöcker (Hock Burn) von Masthühnern. In: 25. Internationale DVG-Fachtagung zum Thema Tierschutz, München 2019, pp. 255-265

Piller, A.; Bergmann, S.; Schwarzer, A.; Erhard, M.; Spindler, B.; Stracke, J.; Kemper, N.; Bachmeier, J.; Schmidt, P.; Schade, B.; Boehm, B.; Kappe, E.; Louton, H. (2019): Validierung eines makroskopischen Bewertungssytems durch histologische Untersuchung der Pododermatitis (Foot Pad Dermatitis) bei Masthühnern. . In: 25. Internationale DVG-Fachtagung zum Thema Tierschutz, München 2019, pp. 389-393

Stracke, J.; Spindler, B.; Kemper, N. (2018): Walking on the tiptoe: Is footpad dermatitis in turkeys a problem of the foot pad only or is it going further? In: Book of Abstracts of the 69th Annual Meeting of the European Federation of Animal Science, 27-31 August, Dubrovnik, Kroatien, p. 559

Stracke, J.; Spindler, B.; Kemper, N. (2018): Walking on the tiptoe: Description of alterations of the digital pads as one indicator of foot pad dermatitis. In: The XVth European Poultry Conference, Conference Information and Proceedings 17-21 September, Dubrovnik, Kroatien, p.25